On the relaxation to quantum-statistical equilibrium of the Wigner–Weisskopf atom in a one-dimensional radiation field. VII. Emission in a finite system in the presence of an extra photon

Russell Davidson

1450 Chestnut Street, Vancouver, British Columbia V6J 3K3, Canada

John J. Kozak

Department of Chemistry and Radiation Laboratory, University of Notre Dame, Notre Dame, Indiana 46556*
(Received 20 January 1975)

In this paper, we present the exact solution to a problem previously unsolved in radiation theory: the emission of a two-level atom in a (one-dimensional) radiation field in the presence of an extra photon. The solution is obtained directly from the Schrödinger equation of the problem using techniques suggested by the work of Muskhelishvili on singular integral equations. The solution corresponding to a finite system as well as the one corresponding to a system infinite in extent are given, although our primary concern in this paper is the finite-system problem. For a particular choice of initial condition, the probability at time t that the two-level atom is in the excited state is found, and the effects of system size and choice of coupling function are studied numerically for a given coupling constant. Our results are compared with those obtained in an earlier paper of the series, wherein the spontaneous emission of a Wigner-Weisskopf atom in a (one-dimensional) field of radiation for a system finite in extent was studied. The physical effects calculated and the conclusions drawn from our comparative studies are all in accord with simple intuition regarding the problem. The paper concludes with some brief remarks on problems in radiation theory which are accessible to study given the methods laid down in this paper.

I. INTRODUCTION

The link between this paper and the earlier ones of the series[1-6] (hereafter referred to as I-VI, respectively) is that the same Hamiltonian is used as the basis of all the calculations. This Hamiltonian is thought of as providing a model for the interaction of a two-level atom, or more generally a two-level quantum system, with a one-dimensional radiation field. The field may be either finite or infinite in extent, but in this paper the main concern is with finite systems. Some indications will be given of results for infinite systems which correspond with those laid out here for finite systems, although a detailed study will be postponed to a future paper.

It should be said at the outset that the restriction to a one-dimensional field, which has been made throughout the series is not essential: It is made chiefly for ease of presentation and for simplicity of the numerical computations. The Hamiltonian is then

$$H=\epsilon_1\alpha\alpha^*+\epsilon_2\alpha^*\alpha+\sum_\lambda\left[\tfrac{1}{2}\hbar\omega_\lambda(a_\lambda^*a_\lambda+1)\right]+\sum_\lambda\left(h_\lambda^*\alpha^*a_\lambda+h_\lambda\alpha a_\lambda^*\right) \tag{1}$$

where ϵ_1 and ϵ_2 are the energies of the ground state $|1\rangle$ and excited state $|2\rangle$ of the two-level atom, and where the operators are defined by

$$\alpha=|1\rangle\langle 2|,\quad \alpha^*=|2\rangle\langle 1|,$$
$$\langle n_\lambda|a_\lambda|m_\lambda\rangle=[2(n_\lambda+1)]^{1/2}\delta^{\mathrm{Kr}}(m_\lambda-n_\lambda-1)=\langle m_\lambda|a_\lambda^*|n_\lambda\rangle.$$

Here the state $|n_\lambda\rangle$ is that which has $n_\lambda(=0,1,2,\cdots)$ photons in the λth mode of the radiation field, and $\delta^{\mathrm{Kr}}(\cdots)$ is the Kronecker delta. Further, $\hbar\omega_\lambda$ is the energy of a photon in the λth mode, $\hbar E=\epsilon_2-\epsilon_1$ is the energy separating the two levels of the atom, and the h_λ give the coupling between the atom and the radiation field. A basis for the Hilbert space of the system is given by the product states

$$|i;\{n_\lambda\}\rangle\equiv|i\rangle\prod_\lambda|n_\lambda\rangle \tag{2}$$

with $i=1,2$ and $n_\lambda=0,1,2,\cdots$.

An important property of the Hamiltonian of Eq. (1) is that in the basis (2) it becomes block diagonal. The blocks can be labelled by the number

$$N=i+\sum_\lambda n_\lambda.$$

These numbers are the eigenvalues of the operator

$$\alpha^*\alpha+\tfrac{1}{2}\sum_\lambda a_\lambda^*a_\lambda$$

which can easily be seen to commute with H. Previously in this series, only the "sector" where $N=2$ (that which describes the spontaneous emission of the atom) has been discussed. It is the aim here to provide an exact solution of the problem $N=3$ in the next section. It will be seen that methods similar to those used to do so will yield an exact solution in any sector.

The authors are not aware of any earlier work in which methods like those of this paper have been used to solve problems in the theory of radiation. These methods may therefore lead to several additions to the list of soluble models available for study. For example, a model has been constructed for a three-level system interacting with radiation, and its solution is being studied at present. The new aspect of these soluble models is that both boson and fermion operators are in-

 Copyright © 1975 American Institute of Physics

volved in the Hamiltonians [the α and a_λ in Eq. (1)] while at the same time an infinite number of modes of the boson field is considered. Since Glauber's coherent state representation for radiation fields was introduced,[7] it has been known that models involving only boson operators where the coupling is linear in these operators can be solved exactly. Problems of this kind have been extensively studied in the theory of lasers,[8-10] solid state,[11] and molecular physics.[12] On the other hand, in the numerous models in laser theory where fermion operators have been included in the Hamiltonian, it has previously been necessary to consider only a finite number (usually one) of modes of an interacting boson field if an exact solution was to be achieved. Probably the best-known example of such a model is the Dicke maser model[13] which was solved exactly by Tavis and Cummings[14] and has since received rather sophisticated treatment in connection with the phenomenon of superradiance.[15]

In Sec. II of this paper the Hamiltonian of Eq. (1) is restricted to the sector $N=3$, and an equation is derived the solution of which yields the resolvent of this restricted Hamiltonian. In Sec. III, this equation is solved and an expression found for the time dependence of a state evolving according to our Hamiltonian starting from a specified initial condition. A particular choice of this initial condition is made in Secs. IV and V, so that $\rho(t)$, the probability at time t that the two-level atom is in its excited state, can be found and computed numerically. The numerical results are then compared with results found in paper IV of this series for $\rho(t)$ in the $N=2$ sector. In Sec. VI there is some discussion of the arguments and results of the paper and of further work in progress.

II. THE RESOLVENT FOR THE $N=3$ SECTOR

Let us begin by listing those states of the set (2) for which $N=3$. If $i=2$, only one of the n_λ can be nonzero, and it must be 1. These states are then

$$|2\rangle\,|1_\lambda\rangle \prod_{\mu\neq\lambda} |0_\mu\rangle;$$

let us write these as $|2;\lambda\rangle$. If $i=1$, there are two possibilities. Either all but two of the n_λ are zero and these two are each one, or else there is one of the n_λ equal to 2 with all others zero. These states will be written, respectively, as $|1;\lambda_1,\lambda_2\rangle$ and $|1;2\lambda\rangle$. At this point it is useful to order the modes λ of the field. In our one-dimensional case this is simple—the modes correspond to the positive integers, as will be seen in Sec. V—and in general some ordering can always be found. It would seem from the notation that two distinct states $|1;\lambda_1,\lambda_2\rangle$ and $|1;\lambda_2,\lambda_1\rangle$ can exist, and this of course is not so. The mode ordering can thus be used so that meaning is given to $|1;\lambda_1,\lambda_2\rangle$ only if $\lambda_1<\lambda_2$ and in this way no over-counting of the states can occur.

It is now a simple matter to restrict the Hamiltonian, Eq. (1), to the sector spanned by the states $|2;\lambda\rangle$, $|1;2\lambda\rangle$ and $|1;\lambda_1,\lambda_2\rangle$. If the zero of energy is set at the energy of the state $|1;\{0_\lambda\}\rangle$, the result is

$$\begin{aligned}\frac{1}{\hbar}H_3=&\sum_\lambda (E+\omega_\lambda)|2;\lambda\rangle\langle 2;\lambda| \\ &+\sum_{\lambda_1<\lambda_2}\sum (\omega_{\lambda_1}+\omega_{\lambda_2})|1;\lambda_1,\lambda_2\rangle\langle 1;\lambda_1,\lambda_2| \\ &+\sum_\lambda 2\omega_\lambda|1;2\lambda\rangle\langle 1;2\lambda| \\ &+\sum_{\lambda_1<\lambda_2}\sum \frac{1}{\hbar}h_{\lambda_1}^*\sqrt{2}\,|1;\lambda_1,\lambda_2\rangle\langle 2;\lambda_2| \\ &+\sum_{\lambda_1>\lambda_2}\sum \frac{1}{\hbar}h_{\lambda_1}^*\sqrt{2}\,|1;\lambda_2,\lambda_1\rangle\langle 2;\lambda_2| \\ &+\sum_\lambda \frac{1}{\hbar}2h_\lambda^*|1;2\lambda\rangle\langle 2;\lambda| \\ &+\sum_{\lambda_1<\lambda_2}\sum \frac{\sqrt{2}}{\hbar}\{h_{\lambda_1}|2;\lambda_2\rangle\langle 1;\lambda_1,\lambda_2| + h_{\lambda_2}|2;\lambda_1\rangle\langle 1;\lambda_1,\lambda_2|\} \\ &+\sum_\lambda \frac{2}{\hbar}h_\lambda|2;\lambda\rangle\langle 1;2\lambda|.\end{aligned} \tag{3}$$

Now if the system is in the state $|\Psi(0)\rangle$ at time $t=0$, it evolves according to the equation

$$\begin{aligned}|\Psi(t)\rangle&=\exp(-iH_3t/\hbar)|\Psi(0)\rangle \\ &=\frac{1}{2\pi i}\int_C dz\exp(-izt)(H_3/\hbar-z)^{-1}|\Psi(0)\rangle\end{aligned} \tag{4}$$

where C is a Bromwich contour taken parallel to the positive direction of the real axis of the complex variable z and above all singularities of the integrand. The operator

$$(H_3/\hbar-z)^{-1}$$

is the resolvent of the operator of Eq. (3) and must now be found.

We are looking for the solution of the equation

$$(H_3/\hbar-z)|\Phi\rangle=|\Psi\rangle. \tag{5}$$

First, it can be seen that

$$\begin{aligned}&(H_3/\hbar-z)|2;\lambda\rangle \\ &\quad=(E+\omega_\lambda-z)|2;\lambda\rangle+\sum_{\mu<\lambda}\frac{1}{\hbar}h_\mu^*\sqrt{2}\,|1;\mu,\lambda\rangle \\ &\qquad+\sum_{\lambda<\mu}\frac{1}{\hbar}h_\mu^*\sqrt{2}\,|1;\lambda,\mu\rangle+\frac{2}{\hbar}h_\lambda^*|1;2\lambda\rangle; \\ &(H_3/\hbar-z)|1;\lambda_1,\lambda_2\rangle \\ &\quad=(\omega_{\lambda_1}+\omega_{\lambda_2}-z)|1;\lambda_1,\lambda_2\rangle+(\sqrt{2}/\hbar)(h_{\lambda_1}|2;\lambda_2\rangle \\ &\qquad+h_{\lambda_2}|2;\lambda_1\rangle); \\ &(H_3/\hbar-z)|1;2\lambda\rangle=(2\omega_\lambda-z)|1;2\lambda\rangle+(2/\hbar)h_\lambda|2;\lambda\rangle.\end{aligned} \tag{6}$$

The kets $|\Phi\rangle$ and $|\Psi\rangle$ can be expanded in terms of the basis states

$$\begin{aligned}|\Phi\rangle=&\sum_\lambda \phi_{2;\lambda}|2;\lambda\rangle \\ &+\sum_{\lambda_1<\lambda_2}\sum \phi_{1;\lambda_1,\lambda_2}|1;\lambda_1,\lambda_2\rangle+\sum_\lambda \phi_{1;2\lambda}|1;2\lambda\rangle\end{aligned}$$

with a similar expression for $|\Psi\rangle$. Putting these definitions into Eq. (5) and using the relations (6) leads to

$$\psi_{2;\lambda}=(E+\omega_\lambda-z)\phi_{2;\lambda}+\sum_{\mu<\lambda}\frac{1}{\hbar}h_\mu\sqrt{2}\,\phi_{1;\mu,\lambda}$$

$$+\sum_{\lambda<\mu}\frac{1}{\hbar}\,h_\mu\sqrt{2}\,\phi_{1;\lambda,\mu}+\frac{2}{\hbar}\,h_\lambda\,\phi_{1;2\lambda};$$

$$\psi_{1;\lambda_1,\lambda_2}=(\omega_{\lambda_1}+\omega_{\lambda_2}-z)\,\phi_{1;\lambda_1,\lambda_2}+\frac{1}{\hbar}\,h^*_{\lambda_1}\sqrt{2}\,\phi_{2;\lambda_2}$$
$$+\frac{1}{\hbar}\,h^*_{\lambda_2}\sqrt{2}\,\phi_{2;\lambda_1};$$

$$\psi_{1;2\lambda}=(2\omega_\lambda-z)\,\phi_{1;2\lambda}+\frac{2}{\hbar}\,h^*_\lambda\phi_{2;\lambda}.$$

The quantities $\phi_{1;2\lambda}$ and $\phi_{1;\lambda_1,\lambda_2}$ can be eliminated from these equations to yield an equation in the $\phi_{2;\lambda}$ only; the result is

$$\phi_{2;\lambda}\left(E+\omega_\lambda-z-\sum_\mu\frac{2|h_\mu|^2}{\hbar^2(\omega_\lambda+\omega_\mu-z)}\right)-\sum_\mu\frac{2h^*_\lambda h_\mu\phi_{2;\mu}}{\hbar^2(\omega_\lambda+\omega_\mu-z)}$$
$$=\psi_{2;\lambda}-\left(\sum_{\mu<\lambda}\frac{h_\mu\sqrt{2}\,\psi_{1;\mu,\lambda}}{\hbar(\omega_\lambda+\omega_\mu-z)}+\sum_{\lambda<\mu}\frac{h_\mu\sqrt{2}\,\psi_{1;\lambda,\mu}}{\hbar(\omega_\lambda+\omega_\mu-z)}\right.$$
$$\left.+\frac{2h_\lambda}{\hbar}\,\frac{\psi_{1;2\lambda}}{2\omega_\lambda-z}\right). \tag{7}$$

This expression is, in fact, an infinite set of simultaneous linear equations in the $\phi_{2;\lambda}$ to be solved in terms of the $\psi_{2;\lambda}$, $\psi_{1;\lambda_1,\lambda_2}$, and $\psi_{1;2\lambda}$. It is not immediately clear how to solve this set of equations, but it is helpful to see what happens in the limit when the system becomes of infinite extent and the mode spectrum becomes continuous. The discrete variables λ and μ are replaced by continuous variables, and sums are replaced by integrals. Without going into details, one can see that an integral equation of the following form results:

$$\phi(\lambda)\left(E+\lambda-z-\int_0^\infty\frac{|h(\mu)|^2\,d\mu}{\lambda+\mu-z}\right)-h^*(\lambda)\int_0^\infty\frac{h(\mu)\,\phi(\mu)\,d\mu}{\lambda+\mu-z}$$
$$=\psi(\mu,z). \tag{8}$$

This is an integral equation, still linear, for a function $\phi(\lambda)$ of the positive real variable λ. The form of the kernel suggests that the equation may be soluble by the methods developed by Muskhelishvili[16] for singular integral equations, and this is indeed so. The solution of Eq. (8) will not be given in this paper, but the set of equations (7) will be solved in the next section by a method suggested by that which can be used to solve Eq. (8).

It is perhaps worth remarking here that Eq. (8) can be obtained explicitly either by setting up a procedure to take the infinite system limit of all the quantities in Eq. (7), or else directly by seeking the resolvent of a suitable Hamiltonian which has a continuous spectrum from the beginning, in the manner of Friedrichs.[17] The infinite system limit of the $N=2$ sector can be obtained in the same way, and it is not hard to cast many of the results of Friedrichs' model into the form of the results of Papers V and VI in this series.

III. SOLUTION OF THE EQUATION FOR $\phi_{2,\lambda}$

The set of equations (7) can be conveniently rewritten with the following substitutions:

$$c_\lambda\equiv h_\lambda\phi_{2;\lambda},$$

$$g_\lambda(z)\equiv h_\lambda\left[\psi_{2;\lambda}-\left(\sum_{\mu<\lambda}\frac{h_\mu\sqrt{2}\,\psi_{1;\mu,\lambda}}{\hbar(\omega_\lambda+\omega_\mu-z)}\right.\right.$$
$$\left.\left.+\sum_{\lambda<\mu}\frac{h_\mu\sqrt{2}\,\psi_{1;\lambda,\mu}}{\hbar(\omega_\lambda+\omega_\mu-z)}+\frac{2h_\lambda}{\hbar}\,\frac{\psi_{1;2\lambda}}{2\omega_\lambda-z}\right)\right]. \tag{9}$$

The result is

$$\left(E-z+\omega_\lambda-\sum_\mu\frac{2|h_\mu|^2}{\hbar^2(\omega_\lambda+\omega_\mu-z)}\right)c_\lambda$$
$$-\frac{2|h_\lambda|^2}{\hbar^2}\sum_\mu\frac{c_\mu}{\omega_\lambda+\omega_\mu-z}=g_\lambda(z). \tag{10}$$

Now we introduce two functions of the complex variable ξ as follows:

$$X(\xi)\equiv\frac{1}{2\pi i}\sum_\mu\frac{c_\mu}{\omega_\mu-\xi};$$

and

$$H(\xi)\equiv\frac{1}{2\pi i}\left(E-\xi-\sum_\mu\frac{2|h_\mu|^2}{\hbar^2(\omega_\mu-\xi)}\right). \tag{11}$$

The use of the functions $X(\xi)$ and $H(\xi)$ to solve Eq. (10) will correspond to the use of Cauchy integrals in Muskhelishvili's techniques for singular integral equations. One notices that

$$c_\mu=-2\pi i\,\mathrm{Res}_{\omega_\mu}X,$$
$$\frac{2|h_\mu|^2}{\hbar^2}=2\pi i\,\mathrm{Res}_{\omega_\mu}H, \tag{12}$$

where $\mathrm{Res}_{\omega_\mu}X$ means the residue at the point $\xi=\omega_\mu$ of the meromorphic function X of ξ. Equation (10) can thus be written

$$H(z-\omega_\lambda)\,\mathrm{Res}_{\omega_\lambda}X+X(z-\omega_\lambda)\,\mathrm{Res}_{\omega_\lambda}H=-g_\lambda(z)/(2\pi i)^2. \tag{13}$$

Here H is to be regarded as a known function, and the equation is to be solved for X; z appears as a nonreal parameter. Let us now consider the function

$$F(\xi)\equiv H(z-\xi)X(\xi)+H(\xi)X(z-\xi). \tag{14}$$

This function has poles where $\xi=\omega_\lambda$ and $\xi=z-\omega_\lambda$. It is seen at once that

$$\mathrm{Res}_{\omega_\lambda}F=\text{left-hand side of Eq. (13)}$$
$$=-g_\lambda(z)/(2\pi i)^2,$$
$$\mathrm{Res}_{z-\omega_\lambda}F=-[\text{left-hand side of Eq. (13)}]$$
$$=g_\lambda(z)/(2\pi i)^2.$$

The function F has no other singularities than these poles, and it is finite (possibly zero) at infinity. This latter statement can be seen to follow from the definitions (11) provided $\phi_{2;\lambda}$ is an admissible solution to the problem in the sense that

$$\sum_\lambda|\phi_{2;\lambda}|^2$$

is finite. The couplings h_λ must also satisfy

$$\sum_\lambda|h_\lambda|^2\ \text{finite},$$

in general, although under certain circumstances this restriction can be replaced without affecting the remainder of the argument of this section. By the use of the Mittag–Leffler theorem,[18] it now follows that

$$F(\xi)=\frac{1}{(2\pi i)^2}\left(\sum_\lambda\frac{g_\lambda(z)}{\omega_\lambda-\xi}+\sum_\lambda\frac{g_\lambda(z)}{\omega_\lambda-(z-\xi)}\right)+c(z) \tag{15}$$

where $c(z)$, depending only on the parameter z, gives the value of F at infinity; the quantity $c(z)$ will be determined in a later paragraph. If we define

$$G(\xi) \equiv \frac{1}{(2\pi i)^2} \sum_\lambda \frac{g_\lambda(z)}{\omega_\lambda - \xi}, \tag{16}$$

then Eq. (15) becomes:

$$F(\xi) = G(\xi) + G(z-\xi) + c(z).$$

From the definition of F, Eq. (14), it is seen that

$$X(z-\xi) = [1/H(\xi)][G(\xi) + G(z-\xi) + c(z) - H(z-\xi)X(\xi)]. \tag{17}$$

Now the function $H(\xi)$ has a set of simple zeros at the points $\xi = \xi_\mu$, say, which interlace its poles at $\xi = \omega_\lambda$. That this is so has been shown in detail in Paper IV [see especially Eqs. (8) and (9) and Fig. 1 of this paper]. The function $(H(\xi))^{-1}$ therefore has simple poles at the points $\xi = \xi_\mu$ where the residues are $(H'(\xi_\mu))^{-1}$ (the prime here denotes differentiation). Thus the terms on the right-hand side of Eq. (17) severally have poles at $\xi = \xi_\mu$, and the total residue of the right-hand side at each such pole is

$$\frac{G(\xi_\mu)}{H'(\xi_\mu)} + \frac{G(z-\xi_\mu)}{H'(\xi_\mu)} + \frac{c(z)}{H'(\xi_\mu)} - H(z-\xi_\mu)\,\mathrm{Res}_{\xi_\mu}\left(\frac{X}{H}\right).$$

But this expression is equal to zero, since the left-hand side of Eq. (17), that is, $X(z-\xi)$, is not singular at $\xi = \xi_\mu$. (Note that z is *not* real.) This result enables us to write down an expression for

$$X(\xi)/H(\xi),$$

since this meromorphic function, equal to zero at infinity, has poles only at $\xi = \xi_\mu$. Where $\xi = \omega_\lambda$, both X and H have poles; but an easy calculation shows that the ratio is regular. Thus another use of the Mittag–Leffler theorem gives

$$X(\xi) = -H(\xi) \sum_\mu \frac{1}{\xi_\mu - \xi} \frac{G(\xi_\mu) + G(z-\xi_\mu) + c(z)}{H'(\xi_\mu) H(z-\xi_\mu)}. \tag{18}$$

The quantity $c(z)$ is not yet determined, but apart from that the Equation (10) is now solved.

From Eq. (12), we have

$$c_\lambda = -2\pi i\,\mathrm{Res}_{\omega_\lambda} X$$

$$= \frac{2|h_\lambda|^2}{\hbar^2} \sum_\mu \frac{G(\xi_\mu) + G(z-\xi_\mu) + c(z)}{(\xi_\mu - \omega_\lambda) H'(\xi_\mu) H(z-\xi_\mu)}. \tag{19}$$

[It may seem from Eq. (18) that X has more singularities than we want: at ω_λ, from the factor $H(\xi)$, and at ξ_μ from the denominator in the sum. But, in fact, $H(\xi_\mu) = 0$ by definition, and it is easy to see that X is regular at $\xi = \xi_\mu$.] Now, by definition [Eq. (11)],

$$X(\xi) = \frac{1}{2\pi i} \sum_\lambda \frac{c_\lambda}{\omega_\lambda - \xi}$$

$$= \frac{1}{2\pi i} \sum_\mu \frac{G(\xi_\mu) + G(z-\xi_\mu) + c(z)}{H'(\xi_\mu) H(z-\xi_\mu)} \times \sum_\lambda \frac{2|h_\lambda|^2}{\hbar^2(\omega_\lambda - \xi)(\xi_\mu - \omega_\lambda)} \tag{20}$$

using Eq. (19). Since [Eq. (11)]

$$\sum_\lambda \frac{2|h_\lambda|^2}{\hbar^2(\omega_\lambda - \xi)} = E - \xi - 2\pi i H(\xi)$$

and since $H(\xi_\mu) = 0$, it follows that

$$\sum_\lambda \frac{2|h_\lambda|^2}{\hbar^2(\omega_\lambda - \xi)(\xi_\mu - \omega_\lambda)} = 1 - \frac{2\pi i H(\xi)}{\xi_\mu - \xi}. \tag{21}$$

The expressions (18) and (20) for $X(\xi)$ can now be equated. The result is that $c(z)$ is determined, as follows:

$$c(z) = -\frac{1}{H_1(z)} \sum_\mu \frac{G(\xi_\mu) + G(z-\xi_\mu)}{H'(\xi_\mu) H(z-\xi_\mu)} \tag{22}$$

where the new function $H_1(z)$ is defined by

$$H_1(z) = \sum_\mu \frac{1}{H'(\xi_\mu) H(z-\xi_\mu)}. \tag{23}$$

The final answer can thus be written [using Eqs. (18), (22) and (9)]

$$\phi_{2;\lambda} = \frac{2h_\lambda^*}{\hbar^2} \frac{1}{H_1(z)} \sum_\mu \frac{1}{(\xi_\mu - \omega_\lambda) H'(\xi_\mu) H(z-\xi_\mu)} \times \sum_\kappa \frac{G(\xi_\mu) - G(\xi_\kappa) + G(z-\xi_\mu) - G(z-\xi_\kappa)}{H'(\xi_\kappa) H(z-\xi_\kappa)}.$$

The time-dependent form of this expression is obtained easily from Eq. (4). If the right-hand side $|\Psi\rangle$ of Eq. (5) is identified with the initial condition $|\Psi(0)\rangle$ in Eq. (4), then

$$\langle 2;\lambda | \Psi(t)\rangle = \frac{1}{2\pi i} \int_c dz \exp(-izt)\,\phi_{2;\lambda}$$

$$= \frac{1}{2\pi i} \frac{2h_\lambda^*}{\hbar^2} \int_c dz \exp(-izt) \frac{1}{H_1(z)} \times \sum_\mu \frac{1}{(\xi_\mu - \omega_\lambda) H'(\xi_\mu) H(z-\xi_\mu)} \times \sum_\kappa \frac{G(z,\xi_\mu) - G(z,\xi_\kappa) + G(z, z-\xi_\mu) - G(z, z-\xi_\kappa)}{H'(\xi_\kappa) H(z-\xi_\kappa)} \tag{24}$$

where the dependence of G on z coming from $g_\lambda(z)$ [see Eq. (16)] has been explicitly included.

IV. THE CASE OF INDUCED EMISSION

In this section and the next, we shall consider only the situation where, at time $t = 0$, the atom is in its excited state. This means that in the initial state vector $|\Psi\rangle$

$$\psi_{1;\lambda_1,\lambda_2} = \psi_{1;2\lambda} = 0.$$

Further, by normalization of $|\Psi\rangle$, we have

$$\sum_\lambda |\psi_{2;\lambda}|^2 = 1.$$

Then it is seen from Eq. (9) that

$$g_\lambda(z) = h_\lambda \psi_{2;\lambda};$$

the z-dependence has disappeared and we may now write merely g_λ. The direct dependence of the function $G(z,\xi)$ in Eq. (24) similarly disappears, and thus we obtain

$$\phi_\lambda(t) \equiv \langle 2;\lambda | \Psi(t)\rangle$$

$$= \frac{1}{2\pi i}\frac{2h_\lambda^*}{\hbar^2}\int_c dz \exp(-izt)\frac{1}{H_1(z)}$$
$$\times\sum_\mu \frac{1}{(\xi_\mu-\omega_\lambda)H'(\xi_\mu)H(z-\xi_\mu)}$$
$$\times\sum_\kappa \frac{G(\xi_\mu)-G(\xi_\kappa)+G(z-\xi_\mu)-G(z-\xi_\kappa)}{H'(\xi_\kappa)H(z-\xi_\kappa)}. \quad (25)$$

To evaluate the integral in this expression, it is necessary to locate the singularities of the integrand as a function of z. There is no difficulty about doing this term by term in the two summations, since both, being Mittag—Leffler expansions of meromorphic functions well-behaved at infinity, are uniformly convergent along the contour C. The function $H(z-\xi_\mu)$ has simple zeros where $z=\xi_\mu+\xi_\nu$ for any ξ_ν. But the function $H_1(z)$ has simple poles at these points, since from the definition (23) of H_1,

$$H_1(\xi_\mu+\xi_\nu)=\sum_\kappa \frac{1}{H'(\xi_\kappa)H(\xi_\mu+\xi_\nu-\xi_\kappa)}$$

and the denominators of the terms $\kappa=\mu,\nu$ in this expression vanish. There are therefore no poles of the integrand at the points $z=\xi_\mu+\xi_\nu$ or $z=\xi_\kappa+\xi_\nu$ for those terms in Eq. (25) where $\mu\neq\kappa$. But for any term where $\mu=\kappa$, the numerator

$$G(\xi_\mu)-G(\xi_\kappa)+G(z-\xi_\mu)-G(z-\xi_\kappa)$$

vanishes, and so these terms have no poles at the points under discussion either. The only remaining points where a singularity may occur are those where $H_1(z)$ is zero. It is useful at this stage to note that the function $(H(z))^{-1}$ is meromorphic, zero at infinity, and with simple poles where $z=\xi_\mu$. It follows then that

$$\frac{1}{H(z)}=\sum_\mu \frac{1}{H'(\xi_\mu)(z-\xi_\mu)}. \quad (26)$$

Thus, from Eq. (23),

$$H_1(z)=\sum_\mu\sum_\kappa \frac{1}{H'(\xi_\mu)H'(\xi_\kappa)(z-\xi_\mu-\xi_\kappa)}. \quad (27)$$

This equation is itself a Mittag—Leffler expansion for $H_1(z)$, which, as we saw just above, has poles where $z=\xi_\mu+\xi_\kappa$ for any μ,κ. Now the same reasoning as was used in Sec. III to establish that $H(z)$ had simple zeros at the points $z=\xi_\mu$ interlacing the poles at $z=\omega_\lambda$ can be used again to show that $H_1(z)$ has simple zeros at points $z=\zeta_\lambda$, say, interlacing the poles at $z=\xi_\mu+\xi_\kappa$. It is necessary for the reasoning to be valid that all the coefficients

$$\frac{1}{H'(\xi_\mu)H'(\xi_\kappa)}$$

have the same sign, but this is clearly true from the expression

$$H'(z)=-\frac{1}{2\pi i}\left(-1-\sum_\mu \frac{2|h_\mu|^2}{\hbar^2(\omega_\mu-z)^2}\right)$$

which follows from the second of Eqs. (11). The poles being simple, the residues of $(H_1)^{-1}$ at $z=\zeta_\lambda$ are

$$1/H_1'(\zeta_\lambda).$$

The results of the preceding paragraph can now be used to perform the z-integration in Eq. (25). The result is

$$\phi_\lambda(t)=\frac{2h_\lambda^*}{\hbar^2}\sum_\nu \frac{\exp(-i\zeta_\nu t)}{H_1'(\zeta_\nu)}\sum_\mu \frac{1}{(\xi_\mu-\omega_\lambda)H'(\xi_\mu)H(\zeta_\nu-\xi_\mu)}$$
$$\times\sum_\kappa \frac{G(\xi_\kappa)+G(\zeta_\nu-\xi_\kappa)}{H'(\xi_\kappa)H(\zeta_\nu-\xi_\kappa)}. \quad (28)$$

The terms in $G(\xi_\mu)$ and $G(z-\xi_\mu)$ have dropped out since

$$\sum_\kappa \frac{1}{H'(\xi_\kappa)H(\zeta_\nu-\xi_\kappa)}=H_1(\zeta_\nu)=0$$

by definition. A further simplication can be made at this point. If one considers the integral

$$\frac{1}{2\pi i}\int_S \frac{G(z')}{H(z')H(z-z')}\,dz'$$

where the contour S is a large circle $|z'|=R$, in the limit where $R\to\infty$, then firstly it is clear that the integral vanishes, since

$$\lim_{z'\to\infty}\frac{z'G(z')}{H(z')H(z-z')}=0.$$

But then inside the contour the integrand has poles where $z'=\xi_\kappa$ and $z'=z-\xi_\kappa$. The residues at these poles can be added and the sum set equal to zero, with the result

$$\sum_\kappa \frac{G(\xi_\kappa)}{H'(\xi_\kappa)H(z-\xi_\kappa)}-\sum_\kappa \frac{G(z-\xi_\kappa)}{H(z-\xi_\kappa)H'(\xi_\kappa)}=0.$$

Consequently, from Eq. (28),

$$\phi_\lambda(t)=\frac{4h_\lambda^*}{\hbar^2}\sum_\nu \frac{\exp(-i\zeta_\nu t)}{H_1'(\zeta_\nu)}$$
$$\times\sum_\mu \frac{1}{(\xi_\mu-\omega_\lambda)H'(\xi_\mu)H(\zeta_\nu-\xi_\mu)}$$
$$\times\sum_\kappa \frac{G(\xi_\kappa)}{H'(\xi_\kappa)H(\zeta_\nu-\xi_\kappa)}. \quad (29)$$

For the purposes of the numerical calculations described in the next section, the particular initial condition has been chosen where only one of the $\psi_{2;\lambda}$ is allowed to be nonzero. Thus, we have

$$\psi_{2;\lambda}=\delta_{\lambda\rho}, \quad (30)$$

say, where $\delta_{\lambda\rho}$ is the Kronecker delta. This means simply that the photon present in the initial state is in an energy eigenstate. For the present, this special choice allows us to make some more simplifications in Eq. (29). But, more importantly, we can also make a check on this equation by verifying that if Eq. (30) holds, then the equation gives

$$\phi_\lambda(0)=\delta_{\lambda\rho}.$$

When this has been seen, it will follow by linearity that Eq. (29) gives

$$\phi_\lambda(0)=\psi_{2;\lambda}$$

for any $\psi_{2;\lambda}$. Thus, if Eq. (30) is put into Eq. (29) at $t=0$,

$$\phi_\lambda(0)=\frac{4h_\lambda^*h_\rho}{(2\pi i)^2\hbar^2}\sum_\nu \frac{1}{H_1'(\zeta_\nu)}$$

$$\times\sum_{\mu}\frac{1}{(\xi_\mu-\omega_\lambda)H'(\xi_\mu)H(\xi_\nu-\xi_\mu)}$$

$$\times\sum_{\kappa}\frac{1}{(\omega_\rho-\xi_\kappa)H'(\xi_\kappa)H(\xi_\nu-\xi_\kappa)}. \tag{31}$$

The ν sum in this expression is

$$\sum_{\nu}\frac{1}{H_1'(\xi_\nu)H(\xi_\nu-\xi_\mu)H(\xi_\nu-\xi_\kappa)}.$$

The sum can be performed by considering the integral

$$\frac{1}{2\pi i}\int_C\frac{d\xi}{H_1(\xi)H(\xi-\xi_\mu)H(\xi-\xi_\kappa)}$$

for the case $\mu\neq\kappa$. It has already been seen that the functions

$$H_1(\xi)H(\xi-\xi_\mu)$$

are regular at $\xi=\xi_\mu+\xi_{\mu'}$ for any $\xi_{\mu'}$, and so, apart from the poles at $\xi=\xi_\nu$, the only other singularity of the integrand is at $\xi=\xi_\mu+\xi_\kappa$. Use of the expressions (26) and (27) shows that the residue here is 1/2. Again, these expressions yield the result

$$\lim_{\xi\to\infty}\frac{\xi}{H_1(\xi)H(\xi-\xi_\mu)H(\xi-\xi_\kappa)}=1.$$

To obtain this result, it was necessary to note that

$$\sum_{\mu}\frac{1}{H'(\xi_\mu)}=\frac{1}{2\pi i}\int_S\frac{d\xi}{H(\xi)}$$

$$=\lim_{\xi\to\infty}\frac{\xi}{H(\xi)}=-2\pi i$$

by Eq. (11). It therefore follows that

$$\sum_{\nu}\frac{1}{H_1'(\xi_\nu)H(\xi_\nu-\xi_\mu)H(\xi_\nu-\xi_\kappa)}=\frac{1}{2}\quad(\mu\neq\kappa).$$

A slightly more delicate calculation of the same kind extends the result for $\mu=\kappa$ and in general it turns out that

$$\sum_{\nu}\frac{1}{H_1'(\xi_\nu)H(\xi_\nu-\xi_\mu)H(\xi_\nu-\xi_\kappa)}=\tfrac{1}{2}+\delta_{\kappa\mu}\pi i H'(\xi_\mu). \tag{32}$$

Thus, Eq. (31) becomes

$$\phi_\lambda(0)=\frac{4h_\lambda^*h_\rho}{(2\pi i)^2\hbar^2}$$

$$\times\sum_{\mu}\sum_{\kappa}\frac{1}{H'(\xi_\mu)H'(\xi_\kappa)(\omega_\rho-\xi_\kappa)(\xi_\mu-\omega_\lambda)}[\tfrac{1}{2}+\delta_{\kappa\mu}\pi i H'(\xi_\mu)].$$

But

$$\sum_{\mu}\frac{1}{H'(\xi_\mu)(\omega_\lambda-\xi_\mu)}=\frac{1}{H(\omega_\lambda)}=0 \tag{33}$$

and so

$$\phi_\lambda(0)=\frac{2h_\lambda^*h_\rho}{(2\pi i)\hbar^2}\sum_{\mu}\frac{1}{H'(\xi_\mu)(\omega_\rho-\xi_\mu)(\xi_\mu-\omega_\lambda)}.$$

Plainly, if $\rho\neq\lambda$, this expression vanishes as in Eq. (33). If $\rho=\lambda$, the sum

$$\sum_{\mu}\frac{1}{H'(\xi_\mu)(\omega_\rho-\xi_\mu)^2}$$

can be evaluated by the same technique as gave Eq. (32), and the value of the sum is

$$-(2\pi i)\hbar^2/2|h_\rho|^2.$$

Consequently,

$$\phi_\lambda(0)=\delta_{\lambda\rho},$$

as we wished to prove.

Finally, in this section, a further expression will be derived for the purposes of numerical calculation. The probability at time t that the atom is in its excited state, if $\psi_{2;\lambda}=\delta_{\lambda\rho}$, is

$$\rho(t)\equiv\sum_{\lambda}|\phi_\lambda(t)|^2$$

$$=\frac{8|h_\rho|^2}{\hbar^2}\frac{1}{(2\pi i)^3}\sum_{\mu}\sum_{\nu}\frac{\exp(-i\xi_\nu t)}{H_1'(\xi_\nu)H(\xi_\nu-\xi_\mu)H'(\xi_\mu)}$$

$$\times\sum_{\nu'}\frac{\exp(i\xi_{\nu'}t)}{\overline{H_1'(\xi_{\nu'})H(\xi_{\nu'}-\xi_\mu)}}$$

$$\times\sum_{\kappa}\frac{1}{(\omega_\rho-\xi_\kappa)H'(\xi_\kappa)H(\xi_\nu-\xi_\kappa)}$$

$$\times\sum_{\kappa'}\frac{1}{(\omega_\rho-\xi_{\kappa'})\overline{H'(\xi_{\kappa'})H(\xi_{\nu'}-\xi_{\kappa'})}}, \tag{34}$$

where Eq. (29) has been used along with the result given in Eq. (21). The bars on the functions H and H_1 denote complex conjugates.

V. NUMERICAL CALCULATIONS

In all the previous papers of this series, numerical results were expressed in terms of dimensionless variables, which had the advantage of allowing both weak-coupling and infinite-system limits to be taken easily. Although neither limit is of concern in this paper, both will be in future work, and for this reason as well as to facilitate comparisons between the results of this section and those of earlier papers, the usual dimensionless variables will now be introduced.

The one-dimensional system of our problem is of length L. It has been seen in paper IV [see Eq. (8) of that paper] that a general form for the coupling function $|h_\lambda|^2$ is as follows:

$$|h_\lambda|^2=\frac{\alpha\hbar^2Ec}{L}\,f\left(\frac{c|k_\lambda|}{E}\right), \tag{35}$$

where α is a dimensionless coupling constant (analogous in one dimension to the fine-structure constant in quantum electrodynamics), c is the speed of light, and f is a dimensionless function whose argument involves k_λ, the wave number of the λth mode of the radiation field. In the usual way,

$$\omega_\lambda=c|k_\lambda|\quad\text{and}\quad k_\lambda=2\pi n/L$$

for some nonzero integer n. By a suitable choice of α it is always possible to impose $f(1)=1$, so that $|h_\lambda|^2$ has a certain value at resonance ($\omega_\lambda=E$) irrespective of f. Next, the time t and the various frequencies, ω_λ, ξ_μ, ξ_ν of the problem are made dimensionless by the following definitions:

$$\tau=\alpha Et,\quad\beta_\lambda=\omega_\lambda/\alpha E,\quad\gamma_\mu=\xi_\mu/\alpha E,\quad\delta_\nu=\xi_\nu/\alpha E.$$

A parameter which gives the length of the system in dimensionless form is

$$\sigma^2=\alpha EL/c$$

and it follows that

$$|h_\lambda|^2=[(\alpha E)^2\hbar^2/\sigma^2]\,f(\alpha\beta_\lambda).$$

It is convenient to define dimensionless functions corresponding to H and H_1:

$$H(\alpha E\xi)\equiv(\alpha E/2\pi i)\,\hat{H}(\xi)$$

and

$$H_1(\alpha E\xi)\equiv[(2\pi i)^2/(\alpha E)]\,\hat{H}_1(\xi),$$

whence

$$H'(\alpha E\xi)=(1/2\pi i)\,\hat{H}'(\xi)$$

and

$$H_1'(\alpha E\xi)=[(2\pi i)^2/(\alpha E)^2]\,\hat{H}_1'(\xi).$$

From the definitions of H and H_1 [Eqs. (11) and (23), respectively], it is seen that

$$\hat{H}(\xi)=\frac{1}{\alpha}-\xi-\frac{2}{\sigma^2}\sum_\mu\frac{f(\alpha\beta_\mu)}{\beta_\mu-\xi},\tag{36}$$

$$\hat{H}_1(\xi)=\sum_\kappa\frac{1}{\hat{H}'(\gamma_\kappa)\hat{H}(\xi-\gamma_\kappa)}.$$

Finally, the γ_κ and δ_ν defined above are the zeros of the functions $\hat{H}$ and $\hat{H}_1$, respectively. With all these definitions, Eq. (34) becomes

$$\begin{aligned}\hat{\rho}(\tau)&\equiv\rho(\alpha E t)\\&=-\frac{8f(\alpha\beta_\rho)}{\sigma^2}\sum_\mu\sum_\nu\frac{\exp(-i\delta_\nu\tau)}{\hat{H}_1'(\delta_\nu)\hat{H}(\delta_\nu-\gamma_\kappa)\hat{H}'(\gamma_\mu)}\\&\quad\times\sum_{\nu'}\frac{\exp(i\delta_{\nu'}\tau)}{\hat{H}_1'(\delta_{\nu'})\hat{H}(\delta_{\nu'}-\gamma_\mu)}\\&\quad\times\sum_\kappa\frac{1}{(\beta_\rho-\gamma_\kappa)\hat{H}'(\gamma_\kappa)\hat{H}(\delta_\nu-\gamma_\kappa)}\\&\quad\times\sum_{\kappa'}\frac{1}{(\beta_\rho-\gamma_{\kappa'})\hat{H}'(\gamma_{\kappa'})\hat{H}(\delta_{\nu'}-\gamma_{\kappa'})}.\end{aligned}\tag{37}$$

For the explicit evaluation of this expression, the β_μ, in accord with the relation

$$\omega_\mu=|2\pi nc/L|$$

above, are given by

$$\beta_\mu=|2\pi n/\sigma^2|$$

and thus, for example, from Eq. (11), we have

$$\hat{H}(\xi)=\frac{1}{\alpha}-\xi+\frac{4}{\sigma^2}\sum_{n=1}^{\infty}\frac{f(2\pi\alpha n/\sigma^2)}{\xi-(2\pi n/\sigma^2)}.$$

The expression for $\hat{\rho}(\tau)$ of Eq. (37) has been computed numerically for a moderately weak coupling, viz., $\alpha=0.1$, two values of the parameter σ^2, namely 1.0 and 10.0, and for two choices of the coupling function f: $f(x)=x^{-1/2}$ and $f(x)=x^{-1/4}$. In all cases, the mode ρ was chosen so that the frequency β_ρ was the closest frequency to the resonance frequency $-1/\alpha$ in dimensionless units. The choices of f were made, not because they have any intrinsic merit, but because they were the choices used in Papers IV and V of this series, and consequently comparisons can properly be made with the results of those papers.

To preserve the continuity of our subsequent discussion, we present here some remarks concerning the numerical work. Various checks can be performed to assess the reliability of the results calculated. For example, using the definitions of $\hat{H}$ and $\hat{H}'$, it can be shown that the following identity must be satisfied:

$$\frac{1}{\hat{H}(\delta_\lambda-\gamma_\mu)}=\sum_\kappa\frac{1}{(\delta_\lambda-\gamma_\mu-\gamma_\kappa)\hat{H}'(\gamma_\kappa)}.$$

The separate calculation of the left- and right-hand sides of this expression gives an estimate of the accuracy achieved in determining the γ's and δ's. This check was performed, with satisfactory results, although the details will not be reported here since a separate number would have to be given for each pairs (γ,δ) considered. A somewhat more convenient check is provided by examining the identity

$$\sum_\lambda\frac{1}{\delta_\lambda^2\hat{H}_1'(\delta_\kappa)}=1+\frac{\hat{H}_1'(0)}{[\hat{H}_1(0)]^2},\tag{38}$$

where

$$\hat{H}_1(0)=-\sum_\mu\sum_\kappa\frac{1}{(\gamma_\mu+\gamma_\kappa)\hat{H}'(\gamma_\mu)\hat{H}'(\gamma_\kappa)},$$

$$\hat{H}_1'(0)=-\sum_\mu\sum_\kappa\frac{1}{(\gamma_\mu+\gamma_\kappa)^2\hat{H}'(\gamma_\mu)\hat{H}'(\gamma_\kappa)}.$$

For $\alpha=0.1$, $\sigma^2=1.0$, and $f(x)=x^{-1/2}$ (with 12 γ's and 41 δ's) the left-hand side of Eq. (38) was found to be -0.0955, while calculation of the right-hand side of this expression yielded the result -0.1013. For $\alpha=0.1$, $\sigma^2=1.0$, and $f(x)=x^{-1/4}$ (with 13 γ's and 41 δ's), the left-hand side gave -0.1057, while the calculation of the right-hand side gave -0.1162. Finally, for $\alpha=0.1$, $\sigma^2=10.0$, and $f(x)=x^{-1/2}$ (with 50 γ's and 300 δ's), the left-hand side was -1.8270 whereas the right-hand side was -1.8889. Hence the greatest discrepancy found in these calculations was for the case $\alpha=0.1$, $\sigma^2=1.0$, and $f(x)=x^{-1/4}$ (roughly 8.5%), with the calculations for the coupling function $f(x)=x^{-1/2}$ for both choices of σ^2 in somewhat better agreement.

To assess the reliability of our calculations of $\hat{\rho}(\tau)$, one can calculate $\hat{\rho}(0)$, i.e., the probability that the atom is in its excited state at $\tau=0$; from the discussion of the preceding section, we know that this probability must be unity at $\tau=0$, given our choice of initial condition, Eq. (30). For $\alpha=0.1$, $\sigma^2=1.0$, and $f(x)=x^{-1/2}$, using 12 γ's and 41 δ's, the value $\hat{\rho}(0)=0.9946$ was obtained. This result may be compared with the one calculated using Eq. (IV-19) (the spontaneous-emission problem) for the same choice of α, σ^2, and $f(x)$; there we found that for 11 γ's, the value calculated for $\hat{\rho}(0)$ was 0.9939. For $\alpha=0.1$, $\sigma^2=1.0$, and $f(x)=x^{-1/4}$, using 13 γ's and 41 δ's, we found $\hat{\rho}(0)=0.9897$, a result which may be compared with the corresponding calculation in IV, $\hat{\rho}(0)=0.9888$. We see that the calculation of $\hat{\rho}(0)$ via Eq. (37) for $\alpha=0.1$, $\sigma^2=1.0$ and either choice of coupling function compares favorably with values calculated using the simpler Eq. (IV-19). A similar calculation was performed for the choice of parameters $\alpha=0.1$, $\sigma^2=10.0$, and $f(x)=x^{-1/2}$. Here, using 50 γ's and 300 δ's, a value $\hat{\rho}(0)=0.9353$ was obtained using Eq. (37), a result which is not as good as the one calculated using Eq. (IV-19) with 100 γ's, namely $\hat{\rho}(0)$

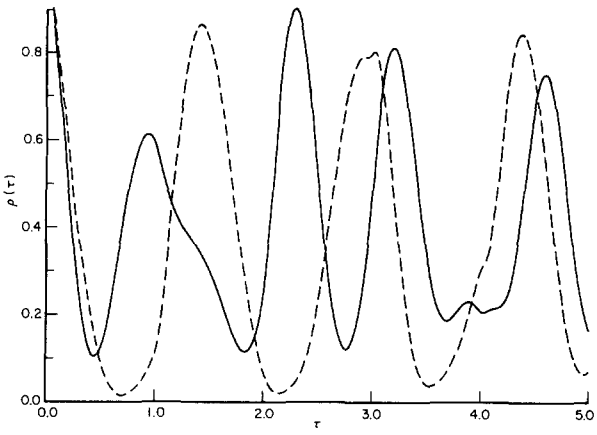


FIG. 1. Plots of $\hat{\rho}(\tau)$ versus τ for the choice of parameters: $\alpha = 0.1$, $\sigma^2 = 1.0$, and $f(x) = x^{-p}$ with $p = 1/2$. The solid line describes the time evolution of $\hat{\rho}(\tau)$ as determined from Eq. (37), the result corresponding to the $N = 3$ sector, while the dashed line displays the time evolution of $\hat{\rho}(\tau)$ as determined from Eq. (IV-19), the result corresponding to the $N = 2$ sector.

$= 0.9932$. The poorer estimate of $\hat{\rho}(0)$ for $\sigma^2 = 10.0$ as calculated from Eq. (37), as compared with the calculation based on Eq. (IV-19), can be attributed to the fact that only half as many γ's were used in the $N = 3$ sector calculation, this because of the excessive computer time required to perform the various sums in Eq. (37) for greater than 50 γ's (and a correspondingly greater number of δ's).

In Fig. 1, $\hat{\rho}(\tau)$ is displayed for $\alpha = 0.1$, $\sigma^2 = 1.0$, and $f(x) = x^{-1/2}$ along with the result obtained in IV for $\hat{\rho}(\tau)$—the probability at time τ of the atom's being excited in the spontaneous-emission sector $N = 2$ if $\hat{\rho}(0) = 1$—for the same values of all the parameters. Figure 2 is similar with $f(x) = x^{-1/4}$, and in Fig. 3, $\alpha = 0.1$, $\sigma^2 = 10.0$, and $f(x) = x^{-1/2}$. It is evident that in the very small systems where $\sigma^2 = 1.0$, the effect of the presence of the extra photon is quite marked. The greater intensity of radiation leads to a somewhat more rapid initial

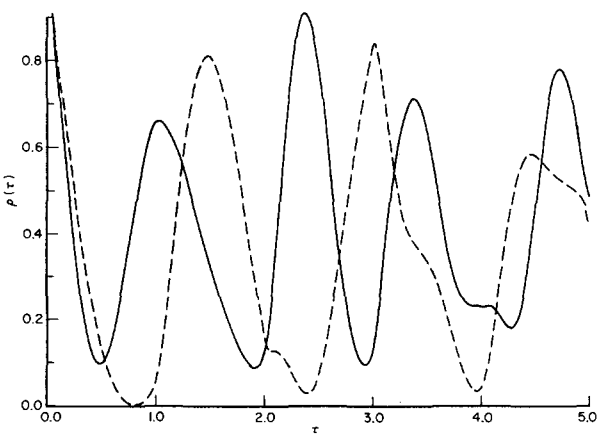


FIG. 2. Plots of $\hat{\rho}(\tau)$ versus τ for the choice of parameters: $\alpha = 0.1$, $\sigma^2 = 1.0$, and $f(x) = x^{-p}$ with $p = 1/4$. The solid line and dashed line refer to the $N = 3$ sector and the $N = 2$ sector, respectively (see the caption to Fig. 1).

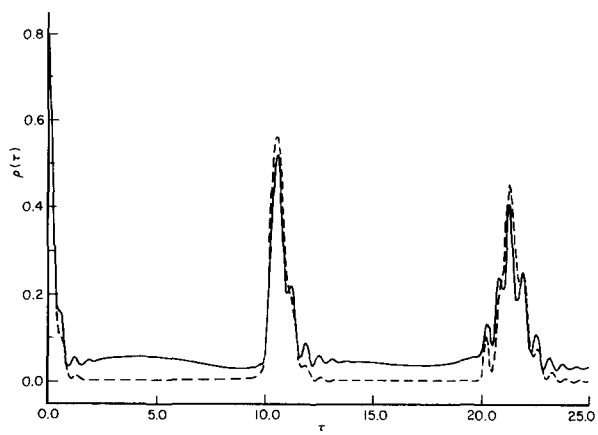


FIG. 3. Plots of $\hat{\rho}(\tau)$ versus τ for the choice of parameters: $\alpha = 0.1$, $\sigma^2 = 10.0$, and $f(x) = x^{-p}$ with $p = 1/2$. The solid line and the dashed line refer to the $N = 3$ sector and the $N = 2$ sector, respectively (see the caption to Fig. 1).

decay of the excited state, more frequent recurrences of high values of the probability of excitation and a raising of the values of this probability at its minima—all effects reasonably to be expected from elementary physical considerations. When $\sigma^2 = 10.0$ on the other hand, the similarities between the results for the two sectors are much more striking than the differences. The overall differences cited above are all still visible, but are plainly small effects. Rather the extent to which maxima and minima occur at the same times, along with the quite faithful reproduction of small details from one time-evolution curve to the other, is remarkable. A further calculation with the same parameters (i.e., $\alpha = 0.1$, $\sigma^2 = 10.0$, and $f(x) = x^{-1/2}$) was performed for much larger values of τ. In particular, the behavior of $\hat{\rho}(\tau)$ in the range $105 \leq \tau \leq 113.0$ was examined—an interval of interest because of the marked structure of the time-evolution profile found for emission in the $N = 2$ sector [see Fig. (IV-7.)] A comparison of the $N = 2$ and $N = 3$ time-evolution profiles for this range of τ is presented in Fig. 4; it can be seen that the similarities per-

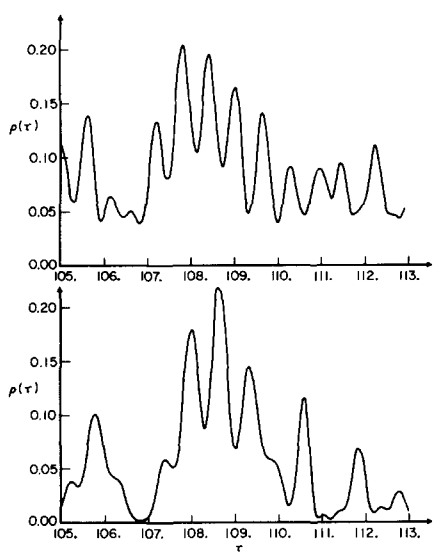


FIG. 4. An examination of the down-range behavior of $\hat{\rho}(\tau)$ versus τ for the choice of parameters: $\alpha = 0.1$, $\sigma^2 = 10.0$, and $f(x) = x^{-p}$ with $p = 1/2$. The upper curve corresponds to the result obtained in the $N = 3$ sector, while the lower curve corresponds to the result obtained in the $N = 2$ sector.

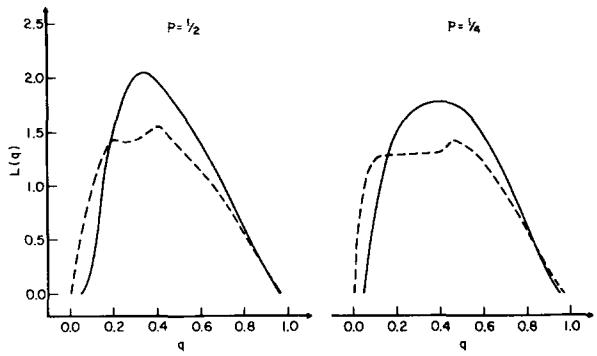


FIG. 5. Plots of $L(q)$ versus q for the choice of parameters: $\alpha = 0.1$, $\sigma^2 = 1.0$. The two figures on the left-hand side of the diagram correspond to the choice $f(x) = x^{-p}$ with $p = 1/2$, whereas the two figures on the right-hand side correspond to the same choice of $f(x)$ but with $p = 1/4$. In each pair of figures, the solid line corresponds to the result obtained in the $N = 3$ sector, while the dashed line corresponds to the result obtained in the $N = 2$ sector.

sist, although not so strikingly as at the beginning of the evolution. Of course, numerical errors are much more important for large τ (arising from the rounding off the values of γ_μ and δ_ν), and the displacement of the positions of the maxima and minima between the two curves displayed in Fig. 4 is most probably explicable merely by this.

A rather clearer indication for the case $\sigma^2 = 1.0$ of the differences between the two sectors is provided in Fig. 5. Here a function $L(q)$ is plotted which measures the number of times the value q is attained by $\hat{\rho}(\tau)$ over a range of τ. To determine $L(q)$, the same range of τ and resolution of $\hat{\rho}(\tau)$ were used as in IV. Specifically, $\hat{\rho}(\tau)$ was computed at intervals of 0.02 over the range $0.0 \leq \tau \leq 150.0$; for this range of τ, sensitive values of q occur about 300 times, and estimates of $L(q)$ were obtained via analysis (direct counting) of the plotted results. The interesting double humps in $L(q)$ for $N = 2$, which were previously reported in Paper IV and are reproduced in Fig. 5, are seen to disappear entirely in the higher sector, and the effects when $N = 3$ of a higher average value of $\hat{\rho}(\tau)$ and of more frequent recurrences are plainly visible. The double humps have most likely vanished because the greater degree of excitation of the radiation field means more frequent excitation and de-excitation of the atom, with consequently less likelihood of the system's being in a state in which there is either a very high or very low probability of the atom's being excited. It was, in any event, seen in Paper IV [see Fig. (IV-4)] that $L(q)$ was, for $\sigma^2 = 1.0$, rather sensitive to the choice of the coupling function f, and thus it is not especially plausible that the disappearance of the double humps has any profound significance.

Finally, we remark that a plot of $L(q)$ versus q was not constructed for the case $\alpha = 0.1$, $\sigma^2 = 10.0$, and $f(x) = x^{-1/2}$, since an excessive amount of computer time would have been required for this determination. To be specific, Fig. 3, which was constructed by calculating $\hat{\rho}(\tau)$ in units of τ of 0.1 over the range $0.0 \leq \tau \leq 25.0$, required 4.8 minutes per point on the Notre Dame IBM 370/158; to obtain a plot of $L(q)$ versus q with an accuracy comparable to that achieved in similar calculations in IV, a range of τ from 100.0 to 400.0 in units of 0.1 would have to be considered, a calculation which is obviously out of the question. In any case, the $L(q)$ versus q profile in the $N = 2$ sector was not characterized by humps [see Fig. (IV-5)], so that the calculation of $L(q)$ for this choice of parameters seems not as urgent.

VI. DISCUSSION

In this paper attention has been directed to the exact solution of our model in the $N = 3$ sector with certain choices of the initial condition and with σ^2 finite. The limit of an infinite system is certainly of interest, and will be considered in a later paper. It may be useful to say something here about why the weak-coupling $\alpha = 0$ limit, which was of central importance in earlier papers in the series, has not been treated. A main point about the limit is that with the scaled dimensionless variables of Sec. V—the choice of which was imposed if the limit was to exist at all—quantities like $\hat{H}(\xi)$ and $\hat{\rho}(\tau)$ are not analytic in α at $\alpha = 0$. This remains true for $N = 3$. Now, it was shown in Papers IV and V that $\rho(\tau)$ in the weak-coupling limit was very similar to the exact $\rho(\tau)$ calculated with small but nonzero α ($\alpha = 0.1$). Since the numerical computation of Eq. (37) is not substantially longer for $\alpha = 0.1$ than for $\alpha = 0$, and since there are no new qualitative features (such as nonanalyticity, ergodicity, etc., as discussed in Paper V) in the $N = 3$ sector over and above those already discussed fully for $N = 2$, it seemed unnecessary for us to perform any separate computations for $\alpha = 0$ here. There is, of course, no particular difficulty in doing them. For an infinite system ($\sigma^2 = \infty$) the weak-coupling solution has some interesting properties which will be studied in later work.

In Papers V and VI some emphasis was put on the choice of the coupling function f [Eq. (35)], and it was noted that difficulties arose, especially with σ^2 infinite, if f was lacking certain properties. The choices of f made in this paper have allowed us to compare results with earlier ones, but they are by no means the most useful or interesting choices. This is not likely to be of much importance for present purposes, as many of the problems associated with f arise only for σ^2 infinite. Further, if σ^2 is finite, unless it is very small, the behavior of $\hat{\rho}(\tau)$ is not much affected by the choice of f [see especially Figs. 1 and 2 of this paper, and Figs. 2 and 3 of Paper IV). However, when the problem with $\sigma^2 = \infty$ comes to be looked at, the choice of f will be important again, and will merit some discussion.

The results of this paper seem to us to be interesting chiefly because they give the solution to a problem in radiation theory which has not previously been amenable to exact treatment. The physical effects calculated are all in accord with simple intuition regarding the problem. But the methods used are capable of a good deal of extension. It has already been noted that a model

of a three-level atom in a radiation field can be tackled in the same way as the model of this paper. This model is able to give a description of many more effects than the two-level model, for example, of competing decay modes for an excited state, phosphorescence and fluorescence in molecules, and so forth. Sectors of the two-level model with higher values of N than 3 can be treated by the methods of this paper, the only limit to computation being the complexity of the expressions to be evaluated. There is every reason to hope that a genuine thermodynamic system can be discussed in which σ^2 becomes infinite along with N in such a way that a finite density of radiation is maintained over all space. Looking farther ahead, one can hope that perhaps even the Dicke–Maser model, involving ultimately an infinite number of two-level atoms in the thermodynamic limit, will prove to be exactly soluable with a full spectrum of modes of the radiation field rather than the one mode that is all that has been manageable up to now.[14,15] In summary, the results presented here are interesting in that they confirm simple intuition about the problem treated and in that they open the way to the exact treatment of many more realistic models.

*The Radiation Laboratory of the University of Notre Dame is operated under contract with the United States Atomic Energy Commission. This is AEC Document No. COO-38-986.

[1]R. Davidson and J.J. Kozak, J. Math. Phys. **11**, 189 (1970).
[2]R. Davidson and J.J. Kozak, J. Math. Phys. **11**, 1420 (1970).
[3]R. Davidson and J.J. Kozak, J. Math. Phys. **12**, 903 (1971).
[4]R. Davidson and J.J. Kozak, J. Math. Phys. **14**, 414 (1973).
[5]R. Davidson and J.J. Kozak, J. Math. Phys. **14**, 423 (1973).
[6]J.J.-W. Yang, R. Davidson, and J.J. Kozak, J. Math. Phys. **15**, 491 (1974).
[7]R.J. Glauber, Phys. Rev. **131**, 2766 (1963).
[8]*Quantum Optics, Proceedings of the International School of Physics, "Enrico Fermi" Course XLII*, edited by R.J. Glauber (Academic, New York, 1969).
[9]I.R. Senitzky, Phys. Rev. A **6**, 1171, 1175 (1972).
[10]M. Lax, Phys. Rev. **145**, 110 (1966).
[11]A. Nitzan and J. Jortner, Mole. Phys. **25**, 713 (1973).
[12]D.F. Walls and R. Barakat, Phys. Rev. A **1**, 446 (1970).
[13]R.H. Dicke, Phys. Rev. **93**, 99 (1954).
[14]M. Tavis and F.W. Cummings, Phys. Rev. **170**, 379 (1968).
[15]K. Hepp and E.H. Lieb, Ann. Phys. (N.Y.) **76**, 360 (1973); see also, Y.K. Wang and F.T. Hioe, Phys. Rev. C **7**, 831 (1973).
[16]N.I. Muskhelishvili, *Singular Integral Equations* (Noordhoff, Groningen, 1953).
[17]K.O. Friedrichs, *Spectral Theory of Operators in Hilbert Space* (Springer, New York, 1973).
[18]P. Dennery and A. Krzywicki, *Mathematics for Physicists* (Harper and Row, New York, 1967), p. 84 *et seq.*

On the curvature dynamics for metric gravitational theories

M. Camenzind

Institut für Theoretische Physik, Universität Bern, Sidlerstrasse 5, 3012 Bern, Switzerland
(Received 27 September 1974)

Any "metric gravitation theory" (including general relativity) is shown to determine transport equations for the connection and curvature of the Lorentz frame bundle P_4 defined by the metric g. Observers are generally defined as curves in P_4 which project down to timelike trajectories in space-time. The transport of curvature along an observer trajectory is then given by a Lorentz Lie algebra-valued current composed of an internal and external part. Einstein's equations are shown to define one part of the self-dual limit of the usual Yang–Mills gauge equations, here called a particular form of curvature dynamics. As a consequence, the Yang–Mills-like energy–momentum tensor, introduced for the Lorentz connection, vanishes identically under Einstein's vacuum conditions.

1. INTRODUCTION

The state of motion of observer systems (inertial frames of reference) in special relativity is not affected by all the interaction occurring on the basis of Minkowski space–time; this means that the observer systems of special relativity are considered to be in an absolute state with respect to interacting systems.

On the other side, any theory of gravitation has to take into account the fact of the observer's participation in the gravitational interaction between macroscopic matter. For this reason, a gravitation theory reveals new aspects, for any gravitation theory must be considered as an attempt to build an interaction theory including observers on the same footing as the description of matter. This fact justifies the use of geometrical methods; curving up space–time seems to have the origin for the purpose of making the state of observers dependent on the gravitational background, e.g., in the sense that in Einstein's theory there exists a frame-dragging effect.[1]

The following discussion is based on the two assumptions:

(I) Space–time (together with a differentiable structure) is endowed with a Lorentz metric g[2] (at least C^2, but in the following C^∞);

(II) space–time curvature is generated by matter of any kind (including fields); therefore, a flat space–time only admits of test fields.

Geometrically, assumption (I) enables us to introduce local Lorentz frames of reference as a model for the description of observer systems on space–time. We do not consider g to be directly related to the gravitational field itself. g defines a subbundle, the Lorentz frame bundle, of the linear frame bundle on space–time. A geometrical quantity—a tensor field or a tensorial form with values in the Lie algebra of a structure group of a principal fibre bundle[3]—will be called *"observable"*, if it is intrinsically related to an observer system or a field of observers on space–time. Connection and curvature of space–time are observables since the linear connection (as connection in the linear frame bundle) is reducible to a Lorentz connection (defined as connection in the Lorentz frame bundle P_4) under assumption (I) (together with a metric condition, see Sec. 2).

A connection in the Lorentz frame bundle, together with curvature and covariant derivatives of curvature, is called, in the following, *external geometry* of a space–time, contrasting with internal geometries defined as connections of other principal fibre bundles on the same base space.[3] In general relativity, the metric tensor serves as an object representing at the same time the state of an observer (it determines standard clocks and measuring rods) as well as the pure gravitational field. The aim of this paper is to find the role of every element in a connection theory for the Lorentz frame bundle: first, curves in the bundle itself which project down to timelike trajectories in space–time represent general *observers*, e.g., the freely falling systems are related to the horizontal curves defined by a connection in P_4, or the "Fermi–Walker curves" correspond to nonrotating observers on space–time (Sec. 3 and 4). Secondly, we shall develop (in Sec. 5) the dynamics for the Lorentz connection and curvature independently of any restriction on the elements (besides the connection has to be torsionless); this dynamical problem consists in giving the evolution for the connection and for all its associated elements along any observer trajectory in P_4.

The structure equations of the connection do not form however, a complete set of transport equations; the transport of the timelike part of the curvature is not yet prescribed. Einstein's equations, considered as a selection principle for curvature elements, imply a Yang–Mills type transport for these elements (Sec. 6). Furthermore, in Sec. 7, we show that the Yang–Mills-like energy–momentum complex for the external geometry vanishes identically under the restrictions implied by Einstein's vacuum equations. This fact suggests enlarging the set of "vacuum space–times" in the sense that it should contain also space–times with nonvanishing energy density for the external geometry; Yang–Mills-like equations for the external geometry surely satisfy this requirement.

2. GEOMETRICAL CONSEQUENCES OF ASSUMPTIONS (I) AND (II)

The description of the external geometry is usually based on the properties of the linear frame bundle

 Copyright © 1975 American Institute of Physics

$L(V, GL(4), \pi)$ over a space–time manifold V. We understand connection and curvature always as given in the sense of Refs. 3, 4.

Let L be provided with a linear connection Γ. Furthermore, we assume space–time to be totally and time-oriented (in the sense of Lichnerowicz[5,6]), then there exists a reduced subbundle $P_4(V, G_4, \pi)$ determined by the structure group G_4, the proper orthochronous Lorentz group; this means the elements $u \in P_4$ consist of orthonormal vector fields, in the sense of the given Lorentz metric g, the timelike element X_0 being future-oriented. Then locally a coordinate system U of V will induce a coordinatization of P_4, a point u in P_4 taking the the coordinates

$$u = (x^\mu, X_a^\mu), \quad \pi(u) = x. \tag{1}$$

Though the linear frame bundle is reducible in this case to the Lorentz frame bundle, a connection Γ in L has in general not to be reducible to a connection $\tilde{\Gamma}$ in the Lorentz frame bundle. Let us first define therefore the exact definition of a reducible linear connection:

Definition: The linear connection Γ in L determined by the connection form ω with values in the Lie algebra of $GL(4)$ is reducible to a connection $\tilde{\Gamma}$ in P_4, given by the connection form $\tilde{\omega}$ with values in the Lie algebra of the Lorentz group G_4, if the following diagram is commutative:

$$\begin{array}{ccc} TL & \xrightarrow{\omega} & \underline{GL_e} \\ {\scriptstyle i_*}\downarrow & & \downarrow{\scriptstyle \tilde{i}_*} \\ TP_4 & \xrightarrow{\tilde{\omega}} & \underline{G_e} \end{array}$$

with i: reduction map $L \to P_4$, determined by the metric g, $\tilde{i}$: reduction map $GL(4) \to G_4$.

This definition is equivalent to the following:

The linear connection is reducible to a Lorentz connection, if, for every $u \in L$, the horizontal subspace of $T_u L$ is tangent to P_4.

The reducibility of the linear connection means for our purposes that, instead of working with all the Christoffel symbols as expressions of ω with respect to a natural linear frame $u = (x, \partial_\mu|_x)$, we shall not lose information on the connection by using a reduced set of connection coefficients as expressions of $\tilde{\omega}$ with respect to a special Lorentz frame, or in other words, with respect to a special observer system.

Lemma 1[3]: Let a Lorentzian metric g be given on V so that g defines the reduction of the linear frame bundle to the Lorentz frame bundle P_4. A linear connection in L is reducible to a Lorentz connection $\tilde{\Gamma}$ in P_4 iff the linear connection is metric.

From this point of view, gravitational theories which are not based on a metric connection (in which the parallel displacement does not preserve the structure of the Lorentz frame bundle) cannot be described within our observer–observable formalism. For, the Lorentz connection defined with respect to an observer system, does not contain all the information about the external geometry. The metric condition on the linear connection is therefore one of the most basic assumptions in an observer-oriented construction of a gravitation theory.

Whether we allow the existence of a torsion for the connection or not, is completely on a different footing; torsion can be explained as part of a curvature of an enlarged Lorentz frame bundle,[4,7] the Poincaré-affine frame bundle on a space–time. This opens new possibilities of constructing internal geometries on space–time by enlarging the Lorentz frame bundle (see, e.g., Refs. 7, 8) in the sense that the Lorentz group appears as a subgroup of a bigger structure group of a corresponding principal fibre bundle and the Lorentz connection as a subconnection of this new connection. As a consequence, the first structure equation and the second one contain additional terms (such as, e.g., the torsion term) which act at the same time as source terms for the Lorentz connection. One example of this kind of a generalized Lorentz connection theory is found in the general relativistic formulation of the Yang–Mills theory by Lichnerowicz and Mme. Kerbrat.[9] In these theories the coupling between the external geometry and the geometry representing internal degrees of freedom is more direct than in Einstein's theory where the coupling is indirect over the energy–momentum tensor of these "fields."

3. GENERAL CURVES IN P_4

Let a connection $\tilde{\Gamma}$ in P_4 be given; then an arbitrary curve (smooth) $u(\lambda)$ in P_4 can be characterized by the corresponding tangent vector field $\mathring{u}(\lambda)$:

$$\mathring{u}(\lambda) = V^a B_{ua} + \tilde{A}^a{}_b E^{*b}_{ua} \tag{2}$$

decomposed with respect to the standard horizontal basis B_a, $a = 0, \ldots, 3$, and the vertical basis $E^{*a}{}_b$ induced by the right action of the structure group G_4 on P_4. $u(\lambda)$ is called *horizontal*, if the tangent vector field $\mathring{u}(\lambda)$ is horizontal, i.e., $\tilde{\omega}(\mathring{u}(\lambda)) = 0$, or

> the curve $u(\lambda)$ is horizontal iff $\tilde{A}^a{}_b = 0$, $\forall\, \lambda$, in the decomposition (2) of the tangent vector field.

The lifting lemma[3] guarantees that there exists, for a given curve x_λ, $\lambda \in [0,1]$, in V and a given frame u_0 for $\lambda = 0$, a uniquely defined horizontal curve $u(\lambda)$ satisfying the initial condition $u(0) = u_0$, with $\pi(u(\lambda)) = x_\lambda$.

A. Horizontal curves

The horizontal curves of a linear connection which are integral curves for standard horizontal vector fields $B(\xi)$ specify the geodesics in V:

Lemma 2[3]: Let $u(\lambda)$ be an integral curve of a standard horizontal vector field $B(\xi)$ on L. Then the projection $\pi(u(\lambda)) = x_\lambda$ determines a geodesic in V, and any geodesic can be obtained in this way. This is even valid for standard horizontal vector fields to be tangent to P_4 if the linear connection is reducible.

The second statement means that for a given geodesic x_λ in V we can find a Lorentz frame $X_a(\lambda)$, parallely propagated along x_λ, and a set of 4 numbers ξ^a such that

$$\dot{x}_\lambda = \xi^a X_a(\lambda), \quad \xi^a = \text{const}\ \forall\ \lambda,$$

which defines the projection of a standard horizontal vector field $B(\xi)$ under π_*.

B. Vertical curves

As a matter of construction, purely vertical curves can be generated by the transformations of the stability group of a group of isometry transformations of the space–time. We do not care about this special type of curves in P_4.

C. General curves in P_4

With respect to a local coordinate system, a general curve in P_4 is given by

$$u(\lambda)=(x^\mu_\lambda, X^\mu_a(\lambda)), \tag{3}$$

and the corresponding tangent vector field can be decomposed

$$V^\mu=\frac{dx^\mu}{d\lambda}, \tag{4}$$

$$A^\mu_a=\frac{dX^\mu_a}{d\lambda}, \tag{5}$$

or

$$u(\lambda)=V^a B_a+\tilde{A}^a{}_b\overset{*}{E}{}^b{}_a. \tag{2'}$$

Lemma 3: V^a: frame components of the tangent vector of the space–time part of the curve $u(\lambda)$; $\tilde{A}^a{}_b$: frame components of the covariant derivatives of the frame vectors $X_a(\lambda)$ along x_λ, or, in other words, $\tilde{A}^a{}_b$ determine how the frame $\{X_a\}$ is transported in the direction of $\dot{x}_\lambda$ with respect to a horizontal curve $\bar{u}(\lambda)=(x_\lambda, \bar{X}_a(\lambda))$.

Proof: 1. $\tilde{\theta}^a_{u^\cdot}(\dot{u})=V^a=V^\mu Y^a_\mu$ since $\tilde{\theta}^a_u(B_{ub})=\delta^a_b$, and $\tilde{\theta}^a$ takes the local expressions with respect to the frame $u(\lambda)$:

$$\tilde{\theta}^a_u=Y^a_\mu(\lambda)\,dx^\mu. \tag{6}$$

2. Look at

$$\tilde{\omega}^a_{ub}(\dot{u})=\tilde{A}^a{}_b, \text{ since } \tilde{\omega}^a{}_b(\overset{*}{E}{}^c{}_d)=\delta^a{}_d\delta^c{}_b.$$

According to the lifting lemma,[3] ∃ a unique horizontal curve $\bar{u}(\lambda)$ with $u(0)=\bar{u}(0)$ and

$$\pi(u(\lambda))=x_\lambda=\pi(\bar{u}(\lambda)),$$

and the corresponding connection forms are related to one another by the transformation law of the connection form under cross-section transformations

$$\tilde{\omega}(\dot{u}(\lambda))=\mathrm{ad}g^{-1}(\lambda)\tilde{\omega}(\dot{\bar{u}}(\lambda))+g^{-1}(\lambda)\dot{g}(\lambda) \tag{7}$$

with

$$u(\lambda)=\bar{u}(\lambda)g(\lambda). \tag{8}$$

$g(\lambda)$ determines a curve in the Lorentz group satisfying the initial condition $g(0)=e$. Since $\bar{u}(\lambda)$ is constructed to be horizontal $\tilde{\omega}(\dot{\bar{u}})=0$, and therefore

$$\tilde{\omega}(\dot{u}(\lambda))=g^{-1}(\lambda)\dot{g}(\lambda), \tag{9}$$

which shows that $\tilde{\omega}(\dot{u})$ determines how the frame $u(\lambda)$ is transported with respect to the corresponding horizontal frame.

Let us express the connection form with respect to the local cross section $u(\lambda)=(x_\lambda, X_a(\lambda))$:

$$\tilde{\omega}^a_{ub}=Y^a_\mu(\lambda)\{dX^\mu_b+\Gamma^\mu_{\rho\beta}(\lambda)X^\beta_b(\lambda)\,dx^\rho\} \tag{10}$$

where the $\Gamma^\mu_{\rho\beta}$ stand for the local expression of the corresponding linear connection form with respect to a natural frame. Apply this decomposition to the interior product:

$$\tilde{\omega}^a_b(\dot{u}(\lambda))=Y^a_\mu\{A^\mu_b+V^\rho\Gamma^\mu_{\rho\beta}X^\beta_b\}$$

$$=Y^a_\mu\left\{\frac{dX^\mu_b}{d\lambda}+V^\rho\Gamma^\mu_{\rho\beta}X^\beta_b\right\}=Y^a_\mu\{V^\rho\nabla_\rho X^\mu_\rho\}=V^\rho\tilde{\Gamma}^a_{\rho b},$$

where $\tilde{\Gamma}^a_{\rho b}$ is the local expression for the reduced Lorentz connection form in our local cross section. This shows that

$$\tilde{\omega}^a_{ub}(\dot{u}(\lambda))=V^\rho\tilde{\Gamma}^a_{\rho b}. \tag{11}$$

If x_λ is a timelike curve in V, the expression (11) contains the information on how our frame is transported along the timelike curve. If x_λ is a geodesic, then we can always choose $\dot{x}_\lambda\sim X_0$; therefore,

$$V^\rho\tilde{\Gamma}^a_{\rho b}=0. \tag{12}$$

Corollary[3]: If x_t represents the integral curve to a Killing vector field ξ on V and $u(t)$ the transformed frame of some u_0, then $\xi^\mu\tilde{\Gamma}^a_{\mu b}$ is constant and determines a one-parameter subgroup in the Lorentz group which gives the transformation between the horizontal frame along ξ and the frame $u(t)=\tilde{\phi}_t(u_0)$, where ϕ_t is generated by ξ.

Proof: Choose a fixed point on x_t, say x_0, and at x_0 a fixed Lorentz frame u_0. Then the Killing vector field ξ generates a one-parameter group of space–time transformations ϕ_t, $x_t=\phi_t(x_0)$, and a corresponding transformation in P_4, say $\tilde{\phi}_t$, with

$$u(t)=\tilde{\phi}_t(u_0).$$

Since ϕ_t is a group of isometry transformations, the Lorentz connnection is invariant under the bundle transformation group which means

$$\tilde{\omega}_{u(t)}(\dot{u}(t))=\tilde{\omega}_{u_0}(\dot{u}(0))$$

and according to (11), this means

$$\xi^\mu\tilde{\Gamma}^a_{\mu b}=\text{const} \quad \text{along the trajectory.}$$

Moreover, the lifting lemma guarantees the existence of a horizontal curve $\bar{u}(t)$ with $\bar{u}(0)=u_0$, and

$$u(t)=\bar{u}(t)g(t), \quad g\in G_4, \ g(0)=e,$$

$$\tilde{\omega}(\dot{u}(t))=g^{-1}(t)\dot{g}(t)=\text{const}=A\in\underset{e}{G_4},$$

which shows that A determines a one-parameter subgroup in the Lorentz group.

4. THE LORENTZ FRAME BUNDLE AS OBSERVER MODEL FOR RELATIVISTIC GRAVITATION THEORIES

The linear frame bundle of a four-dimensional manifold does not as such carry importance for physics, only when specifying the manifold to bear a Lorentz structure, given by a Lorentzian metric on the manifold, or a Galilei structure, given by a positive semi-definite contravariant tensor field γ of rank 3 and a never vanishing 1-form in the kernel of γ,[10] do we impose the first "physical structure" onto the manifold, namely the observer model. The Lorentz structure defines the Lorentz or relativistic observer model; the Galilei structure will define the Galilei or Newtonian

observer model. In the following discussion we restrict ourselves to the relativistic observer model in order to find the relations between our connection-oriented description of gravitation and the usual metric-oriented general relativity.

Definition: Every curve $u(\lambda)$ in the restricted Lorentz frame bundle P_4, such that $\pi(u(\lambda))=x_\lambda$ is a future-oriented timelike curve in the underlying space–time V, defines a *general observer* on space–time.

The state of the observer $u(\lambda)$ is given by the tangent vector $\dot{u}(\lambda)$ on P_4, since, as we have seen above, the space–time part of $\dot{u}$, $V^a=\theta^a_u(\dot{u})$, determines the components of the velocity vector in his own system and $\tilde{\omega}^a_{ub}(\dot{u})=V^\mu\tilde{\Gamma}^a_{\mu b}=V^c\tilde{\Gamma}^a_{cb}$ determines how this observer system is transported along the space–time trajectory with respect to a horizontal curve $\bar{u}(\lambda)$, $\pi(\bar{u}(\lambda))=x_\lambda$.

The horizontal curves $\bar{u}(\lambda)$ in P_4, as integral curves to the standard horizontal vector fields on P_4, represent a special type of these observers; they correspond to the freely falling systems along timelike geodesics, as a consequence of Lemma 2. But, on account of the lifting lemma we find along any arbitrary timelike curve in V a "freely transported system," with respect to which the components of the velocity vector field are, however, no longer constant. This seems to be a little bit in contradiction with the usual definition of observers on space–time; usually, observers appear to be attached to a congruence of timelike geodesics,[11] or a little bit more general in cosmological considerations[12]: There, the observer tetrad has to satisfy the conditions $V^a=\delta^a_0$, i.e., the observers are considered as the integral curves to the frame vector field X_0, which is future-pointing and normalized to unity by the given Lorentz structure. Anyhow, our general definition of observers reduces to these special observers, which were usually considered to be the only "physical observers." Nevertheless, Synge[1] proposed an interpretation of the three spacelike tetrad vectors for adapted observer frames: If we draw connecting vectors in the directions of the spacelike members of a Lorentz frame $\{x_\lambda, X_a(\lambda)\}$ with X_0 tangent to x_λ, then this defines a rigid system of four neighboring observer trajectories. Such an array of four observers serves to define the chronometric meaning of standard lengths.[1] These tetrads along a timelike trajectory determine a special curve in P_4. Moreover, Fermi-transported frames seem to offer a useful physical interpretation; they are the relativistic generalization of the classical "nonrotating frames."[1] Whether the state of an observer determines a Fermi-transported observer can be tested, e.g., by the "bouncing photon experiment" proposed also by Synge.[1]

Lemma 4: 1. For any timelike curve x_λ there exists the horizontal observer $\bar{u}(\lambda)$, defined by the horizontal condition $\tilde{\omega}(\dot{\bar{u}}(\lambda))=0$.

2. For any timelike curve x_λ we find an observer $\tilde{u}(\lambda)$ uniquely determined by $\tilde{u}(0)=u_0$ and the acceleration A for the trajectory x_λ. We call $\tilde{u}(\lambda)$ the general "Fermi–Walker observer."

3. The general Fermi–Walker observer adapted to x_λ, i.e., $\dot{x}^\mu_\lambda=V^\mu=X^\mu_0(\lambda)$, $\forall\lambda$, (λ proper time along the trajectory), corresponds to the usual Fermi–Walker transported tetrad system along x_λ with the initial condition $\tilde{u}(0)=u_0$.

Proof: $\tilde{\omega}_u(\dot{\tilde{u}}(\lambda))$ determines in general how the system $\{X_a\}$ is transported along x_λ.

1. The existence of the horizontal observer is guaranteed by the lifting lemma for the Lorentz connection; we always find a horizontal curve $\bar{u}(\lambda)$ in P_4 satisfying the initial conditions $\bar{u}(0)=u_0$ and $\pi(\bar{u}(0))=x_0$, $\pi(\bar{u}(\lambda))=x_\lambda$, $\forall\lambda$. If x_λ is a geodesic, then $\bar{X}_0$ can be chosen parallely to the velocity vector field V^μ (Lemma 3). If x_λ is not a geodesic, the scalar product $g(V,\bar{X}_0)$ is not constant along x_λ.

2. Fermi–Walker transport for vector fields[13] is that kind of transport that is only determined by the velocity vector field V^μ and the corresponding acceleration A^μ. Especially, for tetrad vector fields we obtain

$$V^\alpha\nabla_\alpha X^\mu_a=\frac{\nabla}{d\lambda}X^\mu_a=X_{a\rho}(V^\mu A^\rho-V^\rho A^\mu), \tag{13}$$

$$A^\mu:=V^\alpha\nabla_\alpha V^\mu,$$

where ∇ is the covariant derivative associated with the linear connection. On the other hand,

$$V^\alpha\nabla_\alpha X^\mu_a=\epsilon^b{}_a X^\mu_b, \quad \epsilon^b{}_a=V^\mu\tilde{\Gamma}^b_{\mu a}, \tag{14}$$

and

$$\epsilon^b{}_a=Y^b_\mu(V^\alpha\nabla_\alpha X^\mu_a)$$

$$=Y^b_\mu X_{a\rho}(V^\mu A^\rho-V^\rho A^\mu),$$

i.e.,

$$\epsilon^b{}_a=V^bA_a-V_aA^b. \tag{15}$$

This shows that $\epsilon\in\underline{G}^4_e$, $\epsilon^0{}_i=\epsilon^i{}_0$, $\epsilon^i{}_k=-\epsilon^k{}_i$.

$\epsilon^b{}_a$ is therefore completely determined by the frame components of the velocity and the acceleration vector field. Given a timelike curve x_λ in space–time, we find a curve in P_4, $\tilde{u}(\lambda)$ with $\tilde{u}(0)=u_0$, determined by the tangent vector field $\tilde{\omega}^a_{ub}(\dot{\tilde{u}}(\lambda))=\epsilon^a{}_b$, $\epsilon^a{}_b$ given by (15). For the same space–time curve the horizontal lift $\bar{u}(\lambda)$ with $\bar{u}(0)=u_0$ is uniquely defined; we consider the transformation between the two curves

$$\tilde{u}(\lambda)=\bar{u}(\lambda)g(\lambda), \quad \forall\lambda, \quad \text{with } g(0)=e. \tag{16}$$

The transformation property (7) for the Lorentz connection in this case is

$$\tilde{\omega}(\dot{\tilde{u}}(\lambda))=\theta_{g(\lambda)}(\dot{g}(\lambda))$$

(θ is the canonical 1-form for the Lorentz group G_4); therefore, the Fermi–Walker observer $\tilde{u}(\lambda)$ is a special curve in the Lorentz frame bundle given by a special curve in the Lorentz group satisfying the differential equation (7) and the initial condition $g(0)=e$.

3. Usually, we define Fermi–Walker transport for tetrads with $X^\mu_0=V^\mu$; for these adapted frames (15) reduces to

$$\epsilon^b{}_a=\delta^b{}_0A_a-\eta_{a0}A^b, \tag{17}$$

which shows that $\epsilon^b{}_a$ is a pure infinitesimal Lorentz transformation, $\epsilon^i{}_k=0$, with $\epsilon^0{}_i=A_i=\epsilon^i{}_0$. Besides the Fermi–Walker observers there are other purely "physical" curves in the Lorentz frame bundle. The counter-

part of the freely falling observers in space–time are the freely rotating observers attached to a small rotating mass. As the physical observer we would take e.g., the rest system for such a rotating body with negligible space extension, X_0 equal to the 4-momentum, which is no longer parallel to the 4-velocity, one of the X_i in the direction of the angular momentum of this rotating observer. Then gravitation, i.e., the external curvature structure of space–time, will determine completely the trajectory of this system in P_4,[14] the state of the observer $\tilde{\omega}(\mathring{u})$ depends on the angular momentum of the rotating mass and on the external curvature of space–time. This example of a freely rotating observer in space–time is self-determining since the state of the rest observer is determined by geometry itself. Even more complicated systems find an interpretation as curves in the Lorentz frame bundle such as the dynamics for a rigid body falling in space–time.[15] Other curves and cross sections of P_4 will not get in general a proper physical interpretation, but they often serve to simplify calculations, say in adapting the observer frame to expected or required symmetries of space–time, just as the usual metric calculations often use coordinate systems responding to symmetry properties of space–time. This is the meaning of the observer-covariance principle for gravitational theories; purely physical observers, as given by special curves in P_4, as well as purely mathematical observers, given by general curves or cross sections in P_4, are able to describe geometry; the only thing we need is the knowledge of the state of the observer, i.e., $\tilde{\theta}^a_u(\mathring{u})$ and $\tilde{\omega}_u(\mathring{u})$. This state will enter all the dynamical equations for the geometry, developed in the next section.

5. GENERAL CURVATURE DYNAMICS FOR THE EXTERNAL GEOMETRY

Every intrinsic geometrical quantity will be expressed in the following, with respect to an observer system thought to be given by a curve in the Lorentz frame bundle or a cross section. We shall develop the dynamics for the connection and the curvature of the Lorentz frame bundle P_4 as long as possibly independent of a special external restriction such as Einstein's equations or a similar restriction on the Lorentz curvature. The aim is to extract as much information as possible only out of the geometry of the Lorentz frame bundle. This procedure has a counterpart for an experimental point of view; in my opinion, it leads to a distinction between four types of experiments concerning the gravitational interaction:

- —locally dynamical experiments which fix the state of the observer system with respect to the gravitational surrounding;
- —experiments for the external geometry measuring some components of the Lorentz curvature, or higher order derivatives;
- —matter experiments giving information on the state of the matter which acts at the same time as source for the gravitational field;
- —experiments for internal geometries measuring the values of internal curvatures associated with some internal degrees of freedom of the matter.

The dynamical problem consists in solving the evolution for the connection and all its associated elements—especially for the Lorentz curvature—with respect to a given observer system. Thereby, each of the following structure equations of the Lorentz frame bundle contributes in a specific way:

- —the first structure equation (typical for linear connections);
- —the second structure equation;
- —the first and second Bianchi identities;
- —higher order structure equations for the curvature.

All these relations do not really restrict the whole set of Lorentzian geometries, nor yield a complete set of differential equations for all possible geometrical observables; the evolution of the timelike curvature does not show up at all in these equations, as in the dynamical picture of a Yang–Mills connection.[9]

A. First structure equation of a Lorentz connection

The first structure equation of a Lorentz connection determines the state of an observer system, i.e., it determines how an observer system is propagated along a curve in space–time. We restrict ourselves to the consideration of only Lorentz connections without torsion,

$$d\tilde{\theta}^a = -\tilde{\omega}^a{}_b \wedge \tilde{\theta}^b, \tag{18}$$

where the $\tilde{\omega}^a{}_b$ are the components of the Lorentz connection form in the decomposition

$$\tilde{\omega} = \tilde{\omega}^a{}_b E^b{}_a, \quad (E^a{}_b)^c{}_d = \eta_{ac}\delta_{bd} - \delta_{ad}\eta_{bc}. \tag{19}$$

Let $u(\lambda)$ be an observer in P_4, $\tilde{V}$ its state, B_c a standard horizontal vector field; then

$$\begin{aligned} &d\tilde{\theta}^a(\tilde{V}, B_c) = L_{\tilde{V}}\tilde{\theta}^a(B_c) - B_c \cdot \tilde{\theta}^a(\tilde{V}), \\ &L_{\tilde{V}}\tilde{\theta}^a(B_c) = B_c \cdot \tilde{\theta}^a(\tilde{V}) - \tilde{\omega}^a{}_c(\tilde{V}) \end{aligned} \tag{20}$$

since

$$\tilde{\omega}(B_c) = 0, \quad \tilde{\theta}^a(B_c) = \delta^a{}_c, \tag{21}$$

which shows how the canonical basis is transported along the observer trajectory. Now, we calculate (18) in local coordinates for the Lorentz frame bundle

$$i_{\tilde{V}}\, d\tilde{\theta}^a = -i_{\tilde{V}}(\tilde{\omega}^a{}_b \wedge \tilde{\theta}^b).$$

The tangent vector field to the observer curve in P_4 is

$$\tilde{V} = V^\mu \partial_\mu + A^\mu_a \frac{\partial}{\partial X^\mu_a} \tag{22}$$

and

$$\tilde{\theta}^a = Y^a_\mu\, dx^\mu, \quad \text{if } u(\lambda) = (x^\mu_\lambda, X^\mu_a(\lambda)). \tag{3'}$$

Therefore,

$$\begin{aligned} i_{\tilde{V}}\, d\tilde{\theta}^a &= -(Y^a_\alpha A^\alpha_b Y^b_\mu)\, dx^\mu - dY^a_\rho V^\rho \\ &= -(Y^a_\alpha A^\alpha_b Y^b_\mu)\, dx^\mu + Y^a_\mu (dX^\mu_b) Y^b_\rho V^\rho; \end{aligned}$$

on the other hand,

$$\begin{aligned} -i_{\tilde{V}}(\tilde{\omega}^a{}_b \wedge \tilde{\theta}^b) = &-\tilde{\omega}^a{}_b(\tilde{V}) Y^b_\mu\, dx^\mu \\ &+ Y^a_\rho (dX^\rho_b + \Gamma^\rho_{\mu\beta} X^\beta_b\, dx^\mu) V^b. \end{aligned}$$

The first structure equation gives therefore an equation for the components A_a^μ for the tangent vector field $\tilde{V}$ in the local decomposition (22)

$$A_c^\beta = \frac{dX_c^\beta}{d\lambda} = X_a^\beta \tilde{\omega}^a{}_c(\tilde{V}) - X_c^\mu \Gamma^\beta_{\mu\rho} V^\rho, \tag{23}$$

or in other words

$$\frac{dX_c^\beta}{d\lambda} + X_c^\mu \Gamma^\beta_{\mu\rho} V^\rho = X_a^\beta \tilde{\omega}^a{}_c(\tilde{V}). \tag{24}$$

The timelike connection coefficients $\tilde{\omega}^a{}_b(\tilde{V})$ determine therefore the transport of the observer frame along the vector field V^ρ if torsion vanishes.

Remark 1: We obtain the same Eq. (23) by transforming the components $\tilde{A}^a{}_b$ of $\tilde{V}$ in the decomposition

$$\tilde{V} = V^a B_{ua} + \tilde{A}^a{}_b \overset{*}{E}{}^b_{ua} \tag{2''}$$

to the local decomposition (22), since the standard horizontal vector fields are locally given by

$$B_{ua} = X_a^\mu \left(\partial_\mu - \Gamma^\alpha_{\mu\beta} X_b^\beta \frac{\partial}{\partial X_b^\alpha} \right), \tag{25}$$

$$\overset{*}{E}{}^a_{ub} = X_b^\mu \frac{\partial}{\partial X_a^\mu}, \tag{26}$$

$$V^a = V^\mu Y^a_\mu. \tag{27}$$

Remark 2: If the Lorentz connection is torsionless and $\tilde{V}$ is horizontal, i.e., $\tilde{\omega}(\tilde{V}) = 0$, then the frames are parallely propagated, as a consequence of (24). From Lemma 2 we know furthermore that, if $V^a = \text{constants}$, then x_λ will be a timelike geodesic and $\tilde{V}$ a standard horizontal vector field on P_4. At the same time we recognize the role of torsion in a Lorentz connection: Torsion will act in the sense of an exterior force term on the state of an observer field, since all Eqs. (18), (20), (23), and (24) have an exterior force term on the right-hand side given by the torsion coefficients.

B. Second structure equation of a Lorentz connection

The contraction of the second structure equation of the Lorentz connnection,

$$d\tilde{\omega}(\tilde{V}, B_c) = \tfrac{1}{2} [\tilde{\omega}(\tilde{V}), \tilde{\omega}(B_c)] + \Omega(\tilde{V}, B_c), \tag{28}$$

gives

$$L_{\tilde{V}} \tilde{\omega}(B_c) = B_c \cdot \tilde{\omega}(\tilde{V}) + \tilde{R}_{\mu c} V^\mu. \tag{29}$$

The Lie transport of the Lorentz connection in the direction of $\tilde{V}$ is determined by the state of the observer and the Lorentz curvature. For a given observer field σ as a local cross section of P_4, the second structure equation determines the curvature in terms of the Lorentz connection

$$\tilde{R}_{\mu\nu} = \partial_\mu \tilde{\Gamma}_\nu - \partial_\nu \tilde{\Gamma}_\mu + [\tilde{\Gamma}_\mu, \tilde{\Gamma}_\nu] \tag{30}$$

or, if we contract with V^μ and X_c, $\sigma = (x^\mu, X_c^\mu)$,

$$\tilde{R}_{\mu\nu} V^\mu X_c^\nu = V^\mu \partial_\mu (X_c^\nu \tilde{\Gamma}_\nu) - X_c^\nu \partial_\nu (V^\mu \tilde{\Gamma}_\mu) + [\tilde{\Gamma}_\mu V_\mu, X_c^\nu \tilde{\Gamma}_\nu] - (V^\mu \partial_\mu X_c^\nu) \tilde{\Gamma}_\nu + (X_c^\nu \partial_\nu V^\mu) \tilde{\Gamma}_\mu,$$

or

$$\frac{d}{d\lambda} \tilde{\Gamma}_c = X_c \cdot (V^\mu \tilde{\Gamma}_\mu) - [V^\mu \tilde{\Gamma}_\mu, \tilde{\Gamma}_c] + \tilde{R}_{\mu c} V^\mu + b^\mu{}_c \tilde{\Gamma}_\mu \tag{31}$$

with

$$[V, X_c] = b_c^\mu \partial_\mu = b^a{}_c X_a. \tag{32}$$

$b^a{}_c$ is determined by the first structure equation. Let σ be a horizontal field of observers, $V^\mu \tilde{\Gamma}_\mu = 0$,

$$\frac{d}{d\lambda} \tilde{\Gamma}_c = \tilde{R}_{\mu c} V^\mu + b_c^\mu \tilde{\Gamma}_\mu. \tag{33}$$

V^μ can be chosen parallel to X_0, since the V^a are constants,

$$\frac{d}{d\lambda} \tilde{\Gamma}_i = \tilde{R}_{0i} + b^k{}_c \tilde{\Gamma}_k. \tag{34}$$

Equations (31) and (24) imply the following interpretation: The timelike connection coefficient $V^\mu \tilde{\Gamma}_\mu$ determines the state of the observer in as much as it gives the vertical part of the tangent vector field to the observer trajectory in P_4; the spacelike part of the connection, $\tilde{\Gamma}_i$, contains the effective information about the Lorentz connection; its evolution along the observer trajectory is determined by the state of the observer and the timelike part of the Lorentz curvature $\tilde{R}_{\mu i} V^\mu$.

C. Bianchi identities for the Lorentz connection

The Bianchi identities for the Lorentz connection take the following form

$$\nabla_\mu \tilde{R}_{\alpha\beta} + \nabla_\beta \tilde{R}_{\mu\alpha} + \nabla_\alpha \tilde{R}_{\beta\mu} + [\tilde{\Gamma}_\mu, \tilde{R}_{\alpha\beta}] + [\tilde{\Gamma}_\beta, \tilde{R}_{\mu\alpha}] + [\tilde{\Gamma}_\alpha, \tilde{R}_{\beta\mu}] = 0. \tag{35}$$

∇ is the linear covariant derivative given by the corresponding linear connection which reduces to the Lorentz connection.

Contracting with $V^\mu X_a^\alpha X_b^\beta$, we have

$$V^\mu X_a^\alpha X_b^\beta \{ \nabla_{[\mu}, \tilde{R}_{\alpha\beta]} + (\tilde{\Gamma}_{[\mu}, \tilde{R}_{\alpha\beta]}) \} = 0.$$

For the following we need some identities:

(1) $$V^\mu X_a^\alpha X_b^\beta \nabla_\mu \tilde{R}_{\alpha\beta} = V^\mu \nabla_\mu (X_a^\alpha X_b^\beta \tilde{R}_{\alpha\beta}) - V^\mu (\nabla_\mu X_a^\alpha) X_b^\beta \tilde{R}_{\alpha\beta} - V^\mu X_a^\alpha (\nabla_\mu X_b^\beta) \tilde{R}_{\alpha\beta}$$

with

$$Y_\alpha^b (V^\mu \nabla_\mu X_a^\alpha) = Y_\alpha^b (V^\mu \partial_\mu X_a^\alpha + V^\mu \Gamma^\alpha_{\mu\rho} X_a^\rho) = V^\mu \tilde{\Gamma}^b_{\mu a},$$

$$V^\mu (\nabla_\mu X_a^\alpha) X_b^\beta \tilde{R}_{\alpha\beta} = (V^\mu \nabla_\mu X_a^\alpha) Y_\alpha^c \tilde{R}_{cb} = V^\mu \tilde{\Gamma}^c_{\mu a} \tilde{R}_{cb};$$

therefore,

$$V^\mu X_a^\alpha X_b^\beta \nabla_\mu \tilde{R}_{\alpha\beta} = V^\mu \partial_\mu \tilde{R}_{ab} - V^\mu \tilde{\Gamma}^c_{\mu a} \tilde{R}_{cb} - V^\mu \tilde{\Gamma}^c_{\mu b} \tilde{R}_{ac}. \tag{36}$$

(2) $$V^\mu X_a^\alpha X_b^\beta \nabla_\beta \tilde{R}_{\mu\alpha} = X_b^\beta \nabla_\beta (V^\mu X_a^\alpha \tilde{R}_{\mu\alpha}) - (X_b^\beta \nabla_\beta V^\mu) X_a^\alpha \tilde{R}_{\mu\alpha} - (X_b^\beta \nabla_\beta X_a^\alpha) V^\mu \tilde{R}_{\mu\alpha},$$

$$(X_b^\beta \nabla_\beta X_a^\alpha) Y_\alpha^c \tilde{R}_{\mu c} V^\mu = X_b^\beta \tilde{\Gamma}^c_{\beta a} \tilde{R}_{\mu c} V^\mu,$$

$$Y_\mu^e (X_b^\beta \nabla_\rho V^\mu) = Y_\mu^e X_b^\beta \nabla_\beta (X_d^\mu V^d) = (X_b \cdot V^e) + X_b^\beta (\nabla_\beta X_d^\mu) V^d Y_\mu^e = X_b \cdot V^e + X_b^\beta \tilde{\Gamma}^e_{\beta d} V^d. \tag{37}$$

Therefore, the contracted Bianchi identities read now

$$\frac{d}{d\lambda} \tilde{R}_{ab} + [V^\mu \tilde{\Gamma}_\mu, \tilde{R}_{ab}] = V^\mu \tilde{\Gamma}^e_{\mu a} \tilde{R}_{cb} - V^\mu \tilde{\Gamma}^c_{\mu b} \tilde{R}_{ca} - \tilde{\nabla}_b (\tilde{R}_{\mu a} V^\mu) - \tilde{\nabla}^a (\tilde{R}_{b\mu} V^\mu) + X_b^\beta \tilde{\Gamma}^c_{\beta a} \tilde{R}_{\mu c} V^\mu + X_a^\alpha \tilde{\Gamma}^c_{\alpha b} \tilde{R}_{c\mu} V^\mu + (X_b \cdot V^c) \tilde{R}_{ca}$$

$$+(X_a\cdot V^c)\tilde{R}_{bc}+X_b^\beta\tilde{\Gamma}_{\beta d}^c V^d\tilde{R}_{ca}+X_a^\alpha\tilde{\Gamma}_{\alpha d}^c V^d\tilde{R}_{bc}, \tag{38}$$

$$\tilde{\nabla}_b(\tilde{R}_{\mu a}V^\mu):=X_b\cdot(\tilde{R}_{\mu a}V^\mu)+[\tilde{\Gamma}_b,\tilde{R}_{\mu a}V^\mu],$$

where the "covariant derivative" stands for the Lorentz connection.

Let us restrict (38) to a horizontal curve with $V^\mu=X_0^\mu$, $a=i$, $b=k$,

$$\frac{d}{d\lambda}\tilde{R}_{ik}=-\tilde{\nabla}_k\tilde{R}_{0i}-\tilde{\nabla}_i\tilde{R}_{k0}+\tilde{\Gamma}_{ki}^m\tilde{R}_{0m}+\tilde{\Gamma}_{ik}^m\tilde{R}_{m0}$$
$$+\tilde{\Gamma}_{k0}^m\tilde{R}_{mi}+\tilde{\Gamma}_{i0}^m\tilde{R}_{km}. \tag{39}$$

Bianchi equations do not provide us with an equation for the timelike elements of the Lorentz curvature $\tilde{R}_{0i}$. In the decomposition of

$$\tilde{\Gamma}_i=f_i^lK_l+\Delta_i^lJ_l, \tag{40}$$

$$\tilde{R}_{ab}=f_{ab}^lK_l+\Delta_{ab}^lJ_l \tag{41}$$

(for $\underline{G}_e^4=\underline{so}(3)+\underline{m}$, J_l the rotation basis for $l=1,2,3$, K_l is the corresponding boost basis) together with the bracket relation

$$[\tilde{\Gamma}_i,\tilde{R}_{ab}]=\{\Delta_i^n\Delta_{ab}^m-f_i^nf_{ab}^m\}\epsilon_{nm}{}^lJ_l$$
$$+\{\Delta_i^nf_{ab}^m+f_i^n\Delta_{ab}^m\}\epsilon_{nm}{}^lK_l \tag{42}$$

we get the J terms,

$$\frac{d}{d\lambda}\Delta_{ik}^l=-X_k\cdot\Delta_{0i}^l-X_i\cdot\Delta_{k0}^l$$
$$-\{\Delta_k^n\Delta_{0i}^m\epsilon_{nm}{}^l-f_k^nf_{0i}^m\epsilon_{nm}{}^l\}$$
$$-\{\Delta_i^n\Delta_{k0}^m\epsilon_{nm}{}^l-f_i^nf_{k0}^m\epsilon_{nm}{}^l\}$$
$$+(\tilde{\Gamma}_{ki}^m-\tilde{\Gamma}_{ik}^m)\Delta_{0m}^l+\tilde{\Gamma}_{k0}^m\Delta_{mi}^l+\tilde{\Gamma}_{i0}^m\Delta_{km}^l, \tag{43}$$

the K terms,

$$\frac{d}{d\lambda}f_{ik}^l=-X_k\cdot f_{0i}^l-X_i\cdot f_{k0}^l$$
$$-\{\Delta_k^nf_{0i}^m\epsilon_{nm}{}^l+f_k^n\Delta_{0i}^m\epsilon_{nm}{}^l\}$$
$$-\{\Delta_i^nf_{k0}^m\epsilon_{nm}{}^l+f_i^n\Delta_{k0}^m\epsilon_{nm}{}^l\}$$
$$+(\tilde{\Gamma}_{ki}^m-\tilde{\Gamma}_{ik}^m)f_{0m}^l+\tilde{\Gamma}_{k0}^mf_{mi}^l+\tilde{\Gamma}_{i0}^mf_{km}^l. \tag{44}$$

Therefore, we obtain differential equations for Δ_{ik}^l and f_{ik}^l, $i,k,l=1,2,3$, i.e., six matrix equations.

D. Differential equations for the timelike curvature part

The evolution equation for the timelike part of the curvature is not implied by the structure of the Lorentz frame bundle; in general, this evolution will be governed by an internal current, built up by components of the connection and curvature itself, and an external current, which is a Lorentz Lie algebra-valued object of type $\mathrm{ad}G_4$—a fact which is suggested by the curvature structure of the gravitation theory. Since Einstein's equations, which are not of the type of a curvature evolution equation in the original form, give rise to a Yang–Mills equation for the Lorentz connection

$$\nabla_\mu\tilde{R}^{\mu\nu}+[\tilde{\Gamma}_\mu,\tilde{R}^{\mu\nu}]=\kappa\tilde{J}_{ex}^\nu, \tag{45}$$

as we shall see in the next section, we complete the general Lorentz curvature dynamics by this type of evolution equation for the timelike part of the curvature; we obtain the Einstein class of solutions by restricting the Lorentz curvature to have some linear dependencies between different components.

Let us transform Eq. (45) with various identities

$$\tilde{R}^{\mu\nu}=-\tfrac{1}{2}\eta^{\mu\nu\alpha\beta}{}^*\tilde{R}_{\alpha\beta},$$
$$\nabla_\mu\tilde{R}^{\mu\nu}=-\tfrac{1}{2}\eta^{\mu\nu\alpha\beta}\nabla_\mu{}^*\tilde{R}_{\alpha\beta},$$
$$[\tilde{\Gamma}_\mu,\tilde{R}^{\mu\nu}]=-\tfrac{1}{2}\eta^{\mu\nu\alpha\beta}[\tilde{\Gamma}_\mu,{}^*\tilde{R}_{\alpha\beta}]$$
$$\tilde{J}^\nu=-\tfrac{1}{2}\eta^{\mu\nu\alpha\beta}{}^*\tilde{J}_{\mu\alpha\beta}\ \text{(as a definition of } {}^*\tilde{J}_{\mu\alpha\beta}),$$
$$\nabla_{[\mu}{}^*\tilde{R}_{\alpha\beta]}+[\tilde{\Gamma}_{[\mu},{}^*\tilde{R}_{\alpha\beta]}]=\kappa{}^*\tilde{J}_{\mu\alpha\beta}. \tag{46}$$

General vacuum solutions will be characterized by ${}^*\tilde{J}_{\mu\alpha\beta}=0$. (46) represents the "dual Bianchi equations"; we apply the same procedure to them as in Sec. 5c and obtain the dual evolution equations for the elements ${}^*\tilde{R}_{ab}$. With respect to a horizontal observer the timelike curvature satisfies the evolution equation

$$\frac{d}{d\lambda}{}^*\tilde{R}_{ik}=-\tilde{\nabla}_k{}^*\tilde{R}_{0i}-\tilde{\nabla}_i{}^*\tilde{R}_{k0}$$
$$+(\tilde{\Gamma}_{ki}^m-\tilde{\Gamma}_{ik}^m){}^*\tilde{R}_{0m}+\tilde{\Gamma}_{k0}^m{}^*\tilde{R}_{mi}+\tilde{\Gamma}_{i0}^m{}^*\tilde{R}_{km}. \tag{47}$$

6. EINSTEIN'S EQUATIONS AS A DEGENERATE CURVATURE DYNAMICS FOR THE LORENTZ CONNECTION

Einstein's vacuum equations are not in the form of an evolution equation for the curvature with respect to a given observer, since they work with the general linear (torsionless) connection (which is also supposed to be metric) over space–time. Reduced to the Lorentz connection of the Lorentz frame bundle, Einstein's vacuum equations are equivalent to special linear dependencies between the components of the Lorentz curvature

$${}^*\tilde{R}_{ab}=\tilde{R}^*{}_{ab}\ \text{ are equivalent to }\ R_{\alpha\beta}=\lambda g_{\alpha\beta}, \tag{48}$$

where $\tilde{R}^*{}_{ab}$ denotes the algebra adjoint for the Lorentz Lie algebra, namely if

$$\tilde{R}_{ab}=f_{ab}^lK_l+\Delta_{ab}^lJ_l, \tag{41}$$

then

$$\tilde{R}^*{}_{ab}=-\Delta_{ab}^lK_l+f_{ab}^lJ_l. \tag{49}$$

Therefore, Einstein's vacuum equations (48) imply the following identities:

$$f_{ik}^l=\epsilon_{ik}{}^m\Delta_{0m}^l,\quad \Delta_{0m}^l=\tfrac{1}{2}\epsilon_m{}^{ik}f_{ik}^l, \tag{50}$$

$$\Delta_{ik}^l=-\epsilon_{ik}{}^mf_{0m}^l,\quad f_{0m}^l=-\tfrac{1}{2}\epsilon_m{}^{ik}\Delta_{ik}^l. \tag{51}$$

i.e., e.g., the relations

$$\Delta_{03}^l=f_{12}^l,\quad f_{03}^l=-\Delta_{12}^l. \tag{52}$$

Bianchi identities contain the evolution equations for the spacelike curvature components, i.e., for the f_{ik}^l's, and Δ_{ik}^l's, so that now Einstein's equations determine automatically the evolution equations for the timelike components, at least in the "Einstein" vacuum case.

Lemma 5: Einstein's vacuum equations together with the evolution equations for the spacelike part of the Lorentz curvature, given by Bianchi identities, imply a Yang–Mills type equation for the timelike Lorentz curvature.

Proof: We assume the differential equations for the

timelike curvature are given by an equation of type (45). All the equations are Lorentz covariant, and so we restrict ourselves to horizontal observers. We have therefore to show that Eqs. (47) together with Einstein's conditions on the Lorentz curvature reduce to Bianchi equations (43) and (44). We pick out, e.g., the components for $i=1$ and $k=2$.

1. Take (43) for $i=1$, $k=2$:

$$\frac{d}{d\lambda}\Delta^l_{12} = -X_2\cdot\Delta^l_{01} - X_1\cdot\Delta^l_{20} - \{\Delta^n_2\Delta^m_{01} - f^n_2 f^m_{01}\}\epsilon_{nm}{}^l - \{\Delta^n_1\Delta^m_{20} - f^n_1 f^m_{20}\}\epsilon_{nm}{}^l + (\tilde\Gamma^m_{21} - \tilde\Gamma^m_{12})\Delta^l_{0m} + \tilde\Gamma^m_{20}\Delta^l_{m1} + \tilde\Gamma^m_{10}\Delta^l_{2m}$$

Together with Einstein's conditions (52), we obtain

$$\frac{d}{d\lambda}\Delta^l_{12} = -X_2\cdot f^l_{23} - X_1\cdot f^l_{13} - \{\Delta^n_2 f^m_{23} + f^n_2\Delta^m_{23}\}\epsilon_{nm}{}^l - \{\Delta^n_1 f^m_{13} + f^n_1\Delta^m_{13}\}\epsilon_{nm}{}^l - \tilde\Gamma^1_{12}f^l_{23} - \tilde\Gamma^2_{21}f^l_{13} + (\tilde\Gamma^3_{21} - \tilde\Gamma^3_{12})f^l_{12} + \tilde\Gamma^m_{20}\Delta^l_{m1} + \tilde\Gamma^m_{10}\Delta^l_{2m}. \tag{53}$$

On the other hand, Eq. (47) for the $*\tilde R_{12}$ gives

$$\frac{d}{d\lambda}*f^l_{12} = -\frac{d}{d\lambda}f^l_{03} = -X_2\cdot f^l_{23} - X_1\cdot f^l_{13} - \{\Delta^n_2 f^m_{23} + f^n_2\Delta^m_{23}\}\epsilon_{nm}{}^l - \{\Delta^n_1 f^m_{13} + f^n_1\Delta^m_{13}\}\epsilon_{nm}{}^l - \tilde\Gamma^1_{12}f^l_{23} - \tilde\Gamma^2_{21}f^l_{13} + (\tilde\Gamma^3_{12} - \tilde\Gamma^3_{21})f^l_{12} + \tilde\Gamma^2_{20}f^l_{03} - \tilde\Gamma^3_{20}f^l_{02} + \tilde\Gamma^1_{10}f^l_{03} - \tilde\Gamma^3_{10}f^l_{01},$$

which reduces to (53) by introducing Einstein's conditions once again.

2. Take (44) for $i=1$, $k=2$:

$$\frac{d}{d\lambda}f^l_{12} = -X_2\cdot f^l_{01} - X_1\cdot f^l_{20} - \{\Delta^n_2 f^m_{01} + f^n_2\Delta^m_{01}\}\epsilon_{nm}{}^l - \{\Delta^n_1 f^m_{20} + f^n_1\Delta^m_{20}\}\epsilon_{nm}{}^l + (\tilde\Gamma^m_{21} - \tilde\Gamma^m_{12})f^l_{0m} + \tilde\Gamma^m_{20}f^l_{m1} + \tilde\Gamma^m_{10}f^l_{2m}$$

together with Einstein's conditions (52)

$$\frac{d}{d\lambda}f^l_{12} = X_2\cdot\Delta^l_{23} + X_1\cdot\Delta^l_{13} + \{\Delta^n_2\Delta^m_{23} - f^n_2 f^m_{23}\}\epsilon_{nm}{}^l + \{\Delta^n_1\Delta^m_{13} - f^n_1 f^m_{13}\}\epsilon_{nm}{}^l + \tilde\Gamma^1_{12}\Delta^l_{23} + \tilde\Gamma^2_{21}\Delta^l_{13} - (\tilde\Gamma^3_{21} - \tilde\Gamma^3_{12})\Delta^l_{12} + \tilde\Gamma^m_{20}f^l_{m1} + \tilde\Gamma^m_{10}f^l_{2m}. \tag{54}$$

On the other hand, Eq. (47) for the $*\tilde R_{12}$ gives

$$\frac{d}{d\lambda}*\Delta^l_{12} = -\frac{d}{d\lambda}\Delta^l_{03} = -X_2\cdot\Delta^l_{23} - X_1\cdot\Delta^l_{13} - \{\Delta^n_2\Delta^m_{23} - f^n_2 f^m_{23}\}\epsilon_{nm}{}^l - \{\Delta^n_1\Delta^m_{13} - f^n_1 f^m_{13}\}\epsilon_{nm}{}^l - \tilde\Gamma^1_{12}\Delta^l_{23} - \tilde\Gamma^2_{21}\Delta^l_{13} + (\tilde\Gamma^3_{21} - \tilde\Gamma^3_{12})\Delta^l_{12} + \tilde\Gamma^2_{20}\Delta^l_{03} - \tilde\Gamma^3_{20}\Delta^l_{02} + \tilde\Gamma^1_{10}\Delta^l_{03} - \tilde\Gamma^3_{10}\Delta^l_{01}.$$

Introducing Einstein's equations once again reduces this equation to (54).

These Yang—Mills type equations for the Lorentz curvature follow also over the corresponding equations for the linear connection (metric and torsionless)

$$\nabla_\rho R^\alpha{}_\beta{}^{\mu\rho} = 0. \tag{55}$$

The linear curvature is related to the corresponding Lorentz curvature over

$$R^\alpha{}_\beta{}^{\mu\rho} = X^\alpha_c Y^b_\beta \tilde R^c{}_b{}^{\mu\rho}.$$

Then

$$Y^e_\alpha X^\beta_f(\nabla_\rho(X^\alpha_c Y^b_\beta \tilde R^c{}_b{}^{\mu\rho})) = \nabla_\rho \tilde R^e{}_f{}^{\mu\rho} + \tilde\Gamma^e_{\rho c}\tilde R^c{}_f{}^{\mu\rho} - \tilde\Gamma^b_{\rho f}\tilde R^e{}_b{}^{\mu\rho} = 0,$$

i.e., in matrix notation

$$\nabla_\rho \tilde R^{\mu\rho} + [\tilde\Gamma_\rho, \tilde R_{\mu\rho}] = 0. \tag{45'}$$

7. ON THE YANG-MILLS CONNECTION ENERGY-MOMENTUM TENSOR OF THE LORENTZ STRUCTURE

The two algebra invariants I_1 and I_2 for the Lorentz connection, defined by

$$I_1 = \mathrm{Tr}(\tilde R_{ab}\tilde R^{ab}), \tag{56}$$

$$I_2 = \mathrm{Tr}(\tilde R_{ab}{}^*\tilde R^{ab}), \tag{57}$$

assume particularly simple values under Einstein's vacuum conditions

$$I_1 = 8\sum_{i=1}^{3}\{[(f^i_{12})^2 - (\Delta^i_{12})^2] + [(f^i_{23})^2 - (\Delta^i_{23})^2] + [(f^i_{31})^2 - (\Delta^i_{31})^2]\}, \tag{58}$$

$$I_2 = 8\sum_{i=1}^{3}(f^i_{12}\Delta^i_{12} + f^i_{23}\Delta^i_{23} + f^i_{31}\Delta^i_{31}). \tag{59}$$

In a general Yang—Mills connection theory, there is defined a gauge-invariant connection energy—momentum tensor T_{ab}; the analogous object for the Lorentz connection

$$\kappa T_{ab} = (-g)^{1/2}\,\mathrm{Tr}(\tilde R_a{}^c\tilde R_{cb} + \tfrac{1}{4}\eta_{ab}\tilde R_{cd}\tilde R^{cd}) \tag{60}$$

resembles Bel's super energy tensor[16] in as much as T_{ab} is also quadratic in the curvature. $\kappa = 8\pi G/c^2$ is the Einsteinian constant. The decomposition of the Lorentz Lie algebra into $\underline{G}_e = so(3) + \underline{m}$ gives a remarkable identity for the Lie algebra-valued curvature elements and their corresponding adjoint components, locally expressed in (41), (49),

$$\mathrm{Tr}(\tilde R_{ab}\tilde R_{ab}) = \mathrm{Tr}(\tilde R^*_{ab}\tilde R^*_{ab}) \tag{61}$$

(without summation over a and b).

The tensor T_{ab} satisfies the following condition:

Lemma 6: The connection energy—momentum tensor density T_{ab}, defined in (60), vanishes identically under Einstein's vacuum equations (48).

This property of T_{ab} follows directly by introducing (48) into T_{ab} and using the identity (61).

In general T_{00} will not be positive definite because of the noncompactness of the structure group (read "gauge group"); however, we hope that the physically important solutions of Eqs. (35) and (45) would imply a positivity condition for the connection energy density T_{00}. T_{ab} itself will not longer vanish, e.g., for the vacuum Einstein—Maxwell equations, even if the corresponding Lorentz connection satisfies the more general "vacuum"

equations (45) with vanishing current $\tilde{J}^a$, such as Lovelock's spherically symmetric connection.[17] The existence of further internal degrees of freedom for space–time manifolds modifies, in general, the Lorentz connection in a way that T_{ab} is nonvanishing.

8. CONCLUSIONS

We have seen that any metric gravitation theory[2] (without torsion) leads immediately to three different levels in the hierarchy of the dynamics: observer dynamics, curvature dynamics for the observer geometry, and matter dynamics. Besides Einstein's theory, for which all the curvature evolution equations follow from Bianchi identities, any metrical theory (e.g., the Brans–Dicke–Jordan theory[18] or Nordtvedt's theories[19]) can be brought to the form of our curvature dynamics; a particular theory is specified by the current $\tilde{J}_\mu$ in Eq. (45) and the particular values for the invariants $\tilde{I}_1$ and $\tilde{I}_2$ in (58) and (59) as well as for the Lorentz energy–momentum tensor (60). Our model is more general than the framework of metrical theories developed by Thorne and Will (see, e.g., Ref. 2) in the sense that we neither restrict ourselves to asymptotically flat space–times nor neglect internal degrees of freedom for the geometric structures. In our framework, general relativity is characterized among other metrical theories by the fact that the curvature-energy–momentum tensor (60) vanishes identically. However, the more general equations (45) include vacuum space–times with nonvanishing curvature-energy density. The meaning of such solutions, at least in the static spherically symmetric case, will be investigated in a forthcoming paper. What we are expecting is the existence of a richer class of space–times containing the Schwarzschild family as a subclass. Such enlarged sets of "vacuum" space–times are useful for different applications even in general relativity, since Einsteinian vacuum geometries appear then in a subset.

[1] J. L. Synge, *Relativity: The General Theory* (North-Holland, Amsterdam, 1971).
[2] So-called metric theories in the wider sense, see, e.g., K. S. Thorne and C. M. Will, Astrophys. J. **163**, 595 (1971).
[3] S. Kobayashi, and K. Nomizu, *Foundations of Differential Geometry, Vol. I* (Interscience, New York, 1963).
[4] A. Lichnerowicz, *Théorie des connexions et des groupes d'holonomie* (Cremonese, Rome, 1955).
[5] A. Lichnerowicz, *Battelle Rencontres* (Benjamin, New York, 1968), p. 107
[6] A. Lichnerowicz, General Relativity Gravitation **1**, 235 (1971).
[7] M. Camenzind, (to be published).
[8] R. Kerner, Ann. Inst. H. Poincaré **9**, 143 (1968).
[9] Mme H. Kerbrat, Ann. Inst. H. Poincaré **13**, 295 (1970).
[10] H. P. Künzle, "Galilei and Lorentz Structures on Space–Time, Comparison of the Corresponding Geometry and Physics," preprint, Department of Mathematics, University of Alberta, 1972.
[11] A. Trautman, *Brandeis Lectures* 1964, *Vol. I* (Prentice-Hall, Englewood Cliffs, N.J., 1965).
[12] R. Sachs, Lectures at Banff Summer School on Relativity, Astrophysics and Cosmology, 1972.
[13] C. Møller, *The Theory of Relativity* (Clarendon, Oxford, 1972), 2nd ed.
[14] H. P. Künzle, J. Math. Phys. **13**, 739 (1972).
[15] H. P. Künzle, Comm. Math. Phys. **27**, 23 (1972).
[16] L. Bel, *Les théories relativistes de la gravitation*, Colloques internationaux du C.N.R.S. (C.N.R.S., Paris, 1959), p. 119.
[17] D. Lovelock, Comm. Math. Phys. **5**, 205, 257 (1967).
[18] C. Brans and R.H. Dicke, Phys. Rev. **124**, 925 (1961).
[19] K. Nordtvedt, Astrophys. J. **161**, 1059 (1970).

Conformal invariance and Hamilton Jacobi theory for dissipative systems

R. M. Kiehn

Department of Physics, University of Houston, Houston, Texas 77004
(Received 25 October 1974)

For certain dissipative systems, a comparison can be made between the Hamilton–Jacobi theory and the conformal invariance of action theory. The two concepts are not identical, but the conformal action theory covers the Hamilton–Jacobi theory.

Recently, in this journal, two articles have appeared, both of which deal with dissipative systems. The first[1] considers the Hamilton—Jacobi theory approach, while the second[2] incorporates dissipation into Hamiltonian theory by considering the conformal invariance (rather than the absolute invariance) of the closed integral of action, $\mathcal{A} = p_\mu dq^\mu - H dt$. It is of interest to compare these two theories, and herein it will be demonstrated that the first theory is a special case of the second.

Classical Hamilton—Jacobi theory admits dissipation in the sense that $\partial H/\partial t$ need not be zero, but physical systems represented by Hamilton's equations of motion must be adiabatic, because the closed integral of work always vanishes when the force is a spatial gradient, $\partial H/\partial q^\mu$. The conformal theory yields a modified set of Hamilton equations of motion which admit nonadiabatic phenomena, as well as systems for which $\partial H/\partial t \neq 0$. These concepts as well as the explicit equations for v^μ and f_μ in the conformal case, are given in Ref. 2.

The key idea, which permits the ready comparison of the two developments, is based upon the observation that if a form $\mathcal{A}$ admits an integrating factor β, such that $\beta\mathcal{A}$ is an absolute invariant with respect to the parameterized vector field γV, then the form $\mathcal{A}$ is a conformal invariant of the vector field γV. That is, if

$$\mathcal{L}_{\gamma V}(\beta\mathcal{A}) = 0 \quad \text{(absolute invariance)} \tag{1}$$

then

$$\mathcal{L}_{\gamma V}\mathcal{A} = \Gamma\mathcal{A} \quad \text{(conformal invariance)}, \tag{2}$$

where the conformality factor Γ is given by the expression

$$-\Gamma/\gamma = \partial(\ln\beta)/\partial t + v^\mu \partial(\ln\beta)/\partial q^\mu + f_\mu \partial(\ln\beta)/\partial p_\mu. \tag{3}$$

For the one-dimensional viscous medium, the Lagrangian $L = \frac{1}{2}\exp(\nu t)\dot{q}^2$ generates a momentum $p = \exp(\nu t)\dot{q}$ and a Hamiltonian $H = \frac{1}{2}p^2\exp(-\nu t)$, such that the action $\mathcal{A} = p_\mu dq^\mu - H dt$ is a conformal invariant of the flow $\gamma V = \gamma(v^\mu, f_\mu, 1)$, with conformality factor $\Gamma/\gamma = -\nu$. The integrating factor for the action is $\beta = \exp(-\nu t)$, such that $\beta\mathcal{A} = \dot{q}^\mu dq^\mu - (\dot{q}^2/2)\,dt$ is an absolute invariant with respect to γV.

Similarly, for the linear damped particle with constant force $L = \exp(\nu t)(\frac{1}{2}\dot{q}^2 - gq)$, $p = \exp(\nu t)\dot{q}$, $H = \frac{1}{2}p^2\exp(-\nu t) + gq\exp(\nu t)$, such that the conformality factor is $\Gamma/\gamma = -\nu$, and the integrating factor is $\beta = \exp(-\nu t)$.

The linearly damped harmonic oscillator yields $L = \frac{1}{2}\exp(2\nu t)(\dot{q}^2 - \omega_0^2 q^2)$, $p = \dot{q}\exp(2\nu t)$, $H = \frac{1}{2}[p^2\exp(-2\nu t) + \omega_0^2 q^2\exp(2\nu t)]$, with a conformality factor $\Gamma/\gamma = -2\nu$, and an integrating factor $\beta = \exp(-2\nu t)$.

The Hamilton—Jacobi theory for these and other examples has been worked out by Denman[1] and others. It should be noted that each of the above examples is a special case of the conformal theory; the constraints of the special case are that the dissipation function Γ is computed to be a scalar constant, and the integrating factor is a function of time, alone. The conformal theory of dissipation includes a much wider class of dissipative systems for which Γ is an arbitrary function of q, p, and t. As an example of the more general case, consider the problem for which the equation of motion involves both linear and quadratic damping and an external force to be described below; i.e., the equation of motion is assumed to be,

$$\ddot{q} + \tfrac{1}{2}k\dot{q}^2 + \nu\dot{q} = -\frac{\partial V}{\partial q} - kV. \tag{4}$$

The Lagrangian generating this equation of motion is given by

$$L = e^\phi\{(\dot{q}^2/2) - V(q)\}, \quad \phi = kq - \nu t. \tag{5}$$

The canonical momentum is $p = \dot{q}e^\phi$, and the Hamiltonian is given by the expression

$$H = (p^2/2)\exp(-\phi) + V\exp(+\phi), \tag{6}$$

such that the action $\mathcal{A} = pdq - H dt$ is a conformal invariant with a conformality factor

$$\Gamma/\gamma = +\nu - k\dot{q} = \nu - kp\exp[-(kq - \nu t)]. \tag{7}$$

The integrating factor for the action is $\beta = \exp(-\phi)$.

The Hamilton—Jacobi theory may be viewed in light of a gauge transformation which reduces the rank of the action-generated volume element from $2N+1$ to $2N$ dimensions; i.e., consider the gauge function ψ, such that the equation

$$\mathcal{A} - d\psi = \left(p_\mu - \frac{\partial\psi}{\partial q^\mu}\right)dq^\mu - \left(\frac{\partial\psi}{\partial p_\mu}\right)dp_\mu - \left(H + \frac{\partial\psi}{\partial t}\right)dt \tag{8}$$

reduces to $\mathcal{A} - d\psi = -(\partial\psi/\partial p_\mu)\,dp_\mu$. Those actions that admit such a gauge reduction must satisfy $\Omega = \mathcal{A}\wedge d\mathcal{A}\wedge \cdots \wedge d\mathcal{A} = 0$; i.e., the $2N+1$ volume Ω is empty.

The constraints so imposed by the gauge reduction on (8) are the usual Hamilton—Jacobi conditions:

$$p_\mu = \partial\psi/\partial q^\mu \tag{9}$$

Copyright © 1975 American Institute of Physics

and

$$\frac{\partial \psi}{\partial t} + H(p, q, t) = 0. \tag{10}$$

The concept that an action which satisfies a conformal invariance equation (which implies dissipation) admit a global integrating factor β can be equilibrated with the idea that the action be reducible by means of a gauge transformation to a set of constraints that yield the Hamilton–Jacobi equations (9) and (10). The general 1-form of action does not satisfy $\mathcal{A} \wedge d\mathcal{A} \wedge \cdots \wedge d\mathcal{A} = 0$, and cannot be represented in the Hamilton–Jacobi format. Such systems are included, however, in the conformal theory of dissipation.

ACKNOWLEDGMENT

This work was partially supported by a NASA–AMES Research Center Grant NCA2-OR-295-502.

[1]H. H. Denman and L. H. Bach, J. Math. Phys. **14**, 326 (1973).
[2]R. M. Kiehn, J. Math. Phys. **15**, 9 (1974).

Weyl quantization of anharmonic oscillators

T. T. Truong*

Department of Physics, Columbia University, New York, New York 10027
(Received 8 August 1974)

It is shown that polynomial self-interactions appearing repeatedly in the mathematical physics and chemical physics literature for the case of one degree of freedom can be treated in the formalism suggested by Weyl long ago, as was the case for the harmonic oscillator. New properties of the Schrödinger wavefunctions are derived, and appropriate schemes of approximation for the eigenvalue problem arise naturally in a nonperturbative way.

I. INTRODUCTION

As far back as 1928, Weyl[1] had proposed to look at the Heisenberg canonical commutation relations (CCR) in their "exponential" form to describe the quantum kinematics of a system. The proposal did not generate much interest until 1949 when Moyal used it to show that quantum mechanics can be viewed as a nondeterministic statistical theory.[2] But it was only in the early 1960's that Weyl's suggestion received renewed interest in the formulation of quantum dynamics of a field by Araki[3] and Segal.[4] This was subsequently followed by the discovery of a realization of the "exponential" (or Weyl) CCR in the space of analytic functions of a complex variable by Bargmann.[5] It turns out that Bargmann's construction is a particular case of the theory of continous representation of Hilbert spaces formulated by Klauder.[6] Indeed with the discovery of coherent light at that time, the Weyl CCR enjoyed a great deal of popularity due to the work of Glauber, Sudarshan, Klauder and others.[7] However, despite recent attempts,[8] this formulation does not seem to be very successful in any other case than the harmonic oscillator problem.

The objective of this article is to demonstrate that nontrivial self-interactions of the polynomial type, Q^{2m} with $m = 2, 3,$ can be treated by the Weyl formulation. This seems to be a worthwhile enterprise since it is an alternative to the well-known Schrödinger approach, which leads too often to involved singular perturbation series. Moreover, since it is so customary to view the single degree of freedom with nontrivial interaction as the prototype of a nonlinearly coupled field with no space dimension, it is hoped that the present study would shed some light on perennial problems of quantum field theory.

In the rest of this section, we shall lay the foundations of our discussion by introducing the functional representation of the Hilbert space of states as a continous representation theory generated by the Weyl CCR and by recalling the essential functional properties that are derived from this approach. In Sec. II the dynamical problem is brought in, and it will be shown that an integral of the motion arises which leads to a partial solution of the problem. Sections IV and V are devoted to the study of the quartic and sextic anharmonic oscillators; it will be shown that new properties of the Schrödinger energy eigenfunctions can be derived. Finally, new connections with old problems are pointed out and their impact on physical applications emphasized.

A. Functional formulation of the Weyl's prescription

Let $\mathcal{H}$ be the abstract Hilbert space of states for a single degree of freedom system. An element of $\mathcal{H}$ will be denoted by a Dirac ket, i.e., $|\psi\rangle \in \mathcal{H}$. Quantum kinematics in $\mathcal{H}$ is described by two self-adjoint unbounded operators on $\mathcal{H}$ and their canonical commutation relation (CCR):

$$[Q, P] = i\mathrm{I} \tag{I.1}$$

where I represents the identity in $\mathcal{H}$ and the Planck constant is set equal to one for simplicity.

Following Weyl, consider the unitary operator

$$U(p, q) = \exp[i(pQ - qP)]$$

where p and q are two real numbers. The group multiplication rule for $U(p, q)$ can be easily derived from (I.1):

$$U(p, q)\,U(p', q') = \exp[i(pq' - p'q)\,U(p + p', q + q')].$$

A special state $|\Omega\rangle$ in $\mathcal{H}$ is called a cyclic state with respect to the $U(p, q)$ if the totality of the states generated by $U(p, q)$ as

$$|p, q\rangle = U(p, q)\,|\Omega\rangle$$

for all p and q, spans $\mathcal{H}$ and form a total set in $\mathcal{H}$.[6] Similarly, one would define

$$\langle p, q| = \langle\Omega|\,U^{\dagger}(p, q).$$

The set of states $|p, q\rangle$ are said to form an overcomplete family of states (OFS) in $\mathcal{H}$.

Once an OFS is known or made available, any state can be completely described by the amplitude

$$|\psi\rangle \in \mathcal{H}, \quad \psi(p, q) = \langle p, q|\psi\rangle.$$

The physical interpretation of the amplitude $\psi(p, q)$ has already been given in Ref. 9 and shall not be discussed here. It can be shown[6] that:

(a) $\psi(p, q)$ is a continuous, bounded function of p and q, the upper bound being the norm of the state $|\psi\rangle$;

(b) $\psi(p, q)$ is a square integrable function with respect to $dp\,dq$; moreover, we have

$$\frac{1}{2\pi}\int |\psi(p, q)|^2\,dp\,dq \leq \langle\psi|\psi\rangle.$$

Let $L^2(p, q)$ be the Hilbert space of the functions $\psi(p, q)$; then $L^2(p, q)$ is obviously a realization of $\mathcal{H}$ in the sense of a continuous representation of a Hilbert space.[6]

 Copyright © 1975 American Institute of Physics

B. Representation of operators in L^2 (p,q)

Assuming the most general choice of $|\Omega\rangle \in \mathcal{H}$, it is possible to construct the partial differential operators representing Q and P from the elementary relations

$$\frac{\partial}{\partial p} U^*(p,q) = U^*(p,q)\, i(\tfrac{1}{2}q - Q),$$

$$\frac{\partial}{\partial q} U^*(p,q) = U^*(p,q)\, i(\tfrac{1}{2}p + P),$$

It follows then that

$$\langle p,q|Q|\psi\rangle = \left(i\frac{\partial}{\partial p} + \tfrac{1}{2}q\right)\psi(p,q),$$
$$\langle p,q|P|\psi\rangle = \left(-i\frac{\partial}{\partial q} + \tfrac{1}{2}p\right)\psi(p,q)$$

for all real p and q. Consequently, the correspondence

$$Q \to Q = \left(i\frac{\partial}{\partial p} + \tfrac{1}{2}q\right), \qquad (I.2)$$
$$P \to P = \left(-i\frac{\partial}{\partial q} + \tfrac{1}{2}p\right)$$

defines a representation of the CCR in $L^2(p,q)$.

However, in most realistic cases, the functions $\psi(p,q)$ will have continuous derivatives of all orders and will exhibit a strong decrease at infinity with respect to p and q. Then the Fourier transform

$$\psi(p,q) = 1/2\pi \int_{-\infty}^{\infty} \psi(k,q)\exp(-ikp)\,dp$$

establishes an isometry between $L^2(p,q)$ and $L^2(k,q)$, the Hilbert space of the functions $\psi(k,q)$. In $L^2(k,q)$ the representation of the CCR is now defined by the correspondence

$$Q \to Q = (\tfrac{1}{2}q + k), \qquad (I.2')$$
$$P \to P = -i\left(\tfrac{1}{2}\frac{\partial}{\partial k} + \frac{\partial}{\partial q}\right).$$

Consider the linear transformation in the plane (k,q),

$$x = (\tfrac{1}{2}q + k), \quad y = (-\tfrac{1}{2}q + k), \qquad (I.3)$$

which induces the initary mapping $L^2(k,q) \to L^2(x,y)$ through

$$\psi(k,q) \to \psi(x,y) = \psi[\tfrac{1}{2}(x+y),(x-y)]. \qquad (I.4)$$

In $L^2(x,y)$, Eqs. (I.2′) can be rewritten in the variables x and y and yield the usual form of the Schrödinger representation of the CCR in the x variable:

$$Q \to Q = x, \quad P \to P = -i\frac{\partial}{\partial x}. \qquad (I.5)$$

Note that y remains a variable so that $L^2(x,y)$ is not the usual space of the Schrödinger representation.

What is presented here is, for an arbitrary choice of cyclic vector state $|\Omega\rangle$, three different ways of constructing two-variable representations of the CCR which are mutually unitary equivalent to each other, exhibiting, in particular, the main feature of von Neumann's theorem[6] on the decomposition of any arbitrary representation of the CCR into a direct sum of copies of the Schrödinger's representations as described by Eq. (I.5).

Finally, if the kinematics offer such a great flexibility in its description, the quantum dynamics in turn are more specific and lead generally to a definite irreducible representation of the CCR, i.e., a more concrete and specific space of wavefunctions where the dynamics is fully described and developed.

II. FORMULATION OF THE DYNAMICS

A. General assumptions

Dynamics of the system are defined by a Hamiltonian operator H, generator of the time translation group

$$\{t \in R,\ U(t) = \exp(-itH)\}.$$

The action of this group is defined by

$$U^{-1}(t)\,Q\,U(t) = Q(t),$$
$$U^{-1}(t)\,P\,U(t) = P(t). \qquad (II.1)$$

The Hamiltonians considered here will be assumed to have pure point spectrum bounded from below, as well as the "canonical" form, defined by the Heisenberg equations of motion

$$[iH, Q] = P,$$
$$[iH, P] = -\frac{dV(Q)}{dQ}, \qquad (II.2)$$

where $V(Q)$ is the given interaction potential. H will consequently appear as a construct of Q and P.

We shall assume, moreover, that in the context of the Weyl picture described in the previous section we deal with an irreducible representation space of functions $L^2(p,q)$, $L^2(k,q)$, or $L^2(x)\times L^2(y)$, where H is a self-adjoint operator represented by the differential operators

$$H\left(-i\frac{\partial}{\partial q} + \tfrac{1}{2}p,\ i\frac{\partial}{\partial p} + \tfrac{1}{2}q\right) \text{ in } L^2(p,q),$$

$$H\left(-i\tfrac{1}{2}\frac{\partial}{\partial k} - i\frac{\partial}{\partial q},\ k + \tfrac{1}{2}q\right) \text{ in } L^2(k,q),$$

$$H\left(-i\frac{\partial}{\partial x},\ x\right) \text{ in } L^2(x)\times L^2(y). \qquad (II.3)$$

The point of considering an irreducible representation is that it allows us to choose any state as a cyclic state[6] and to relate the amplitude $\psi(p,q)$ to the Schrödinger wavefunctions and thereby to study their properties.

B. Existence of an integral of the motion

The task of solving for the spectrum of the differential operators (II.3) turns out to be more tractable in the Weyl formulation because one can explicitly construct an integral of the motion, i.e., another differential operator H' which commutes with H.

To derive the expression for H' consider the OFS generated by a choosen eigenstate $|\epsilon_n\rangle$ of the Hamiltonian

$$H|\epsilon_n\rangle = \epsilon_n|\epsilon_n\rangle,$$

namely,

$$\{(p,q) \in \mathbb{R}\times\mathbb{R},\quad |pq,n\rangle = U(p,q)|\epsilon_n\rangle\}$$

and form the matrix element of H between $|pq,n\rangle$ and an arbitrary state $|\psi\rangle$. From Stone's theorem[6] one can write

$$\langle pq,n|H|\psi\rangle = i\frac{d}{dt}\langle pq,n|\exp(-itH)|\psi\rangle|_{t=0} \qquad \text{(II. 4)}$$

The matrix element of $\exp(-itH)$ can now be transformed into

$$\langle pq,n|\exp(-itH)|\psi\rangle = \langle\epsilon_n|\exp(-i\epsilon_n t)U^+(p,q;t)|\psi\rangle, \qquad \text{(II. 5)}$$

where

$$U^+(p,q,t) = \exp[-ipQ(t)+iqP(t)],$$

thanks to the well-known Baker—Hausdorff formula[6]

$$\exp(-A)B\,e^A = B + \frac{1}{1!}[A,B] + \frac{1}{2!}[A,[A,B]] + \cdots$$

and Eq. (II. 1). It remains to substitute (II. 5) into (II. 4) and take the time derivative according to the formula[10]

$$i\frac{d}{dt}\exp[-iA(t)]\Big|_{t=0} = \exp[-iA(0)]\sum_{m=0}^{\infty}\frac{i^m}{(m+1)!} \times [A(0),\underbrace{[\cdots[A(0),B]\cdots]}_{m\text{ times}}]$$

where

$$B = \text{s-}\lim_{t\to 0}\frac{A(t)-A(0)}{t-0}.$$

Again because of (II. 7) the value of the matrix element (II. 4) is

$$\langle pq,n|H|\psi\rangle = \langle pq,n|\left(\epsilon_n + pP - \tfrac{1}{2}p^2 - \sum_{m=0}^{\infty}\frac{(-1)^{m+1}}{(m+1)!}\,q^{m+1} \times \frac{d^{m+1}}{dQ^{m+1}}V(Q)\right)|\psi\rangle.$$

But, formally one has

$$V(Q) = V(Q-q) = -\sum_{m=0}^{\infty}\frac{(-q)^{m+1}}{(m+1)!}\left(\frac{d}{dQ}\right)^{m+1}V(Q)$$

and

$$pP - \tfrac{1}{2}p^2 = \tfrac{1}{2}[P^2 - (P-p)^2].$$

Consequently, one winds up with the expression

$$\langle pq,n|H|\psi\rangle = \langle pq,n|\epsilon_n + [H(Q,P) - H(Q-q,P-p)]|\psi\rangle, \qquad \text{(II. 6)}$$

where $iH(Q,P)$ is the canonical form of the Hamiltonian of Eq. (II. 2):

$$H(Q,P) = \tfrac{1}{2}P^2 + V(Q).$$

The integral of the motion is the difference between the Hamiltonian and its translated form

$$H'(Q,P) = H(Q,P) - H(Q-q,P-p).$$

From (II. 6) it is clear that H' has the following expressions as differential operator:

$$H' = H\left(i\frac{\partial}{\partial p} + \tfrac{1}{2}q,\ -i\frac{\partial}{\partial q} + \tfrac{1}{2}p\right) - H\left(i\frac{\partial}{\partial p} - \tfrac{1}{2}q,\ -i\frac{\partial}{\partial q} - \tfrac{1}{2}p\right),$$

$$H' = H\left(\tfrac{1}{2}q + k,\ -i\tfrac{1}{2}\frac{\partial}{\partial k} - i\frac{\partial}{\partial q}\right) - H\left(k - \tfrac{1}{2}q,\ i\tfrac{1}{2}\frac{\partial}{\partial k} - i\frac{\partial}{\partial q}\right),$$

$$H' = H\left(x,\ -i\frac{\partial}{\partial x}\right) - H\left(y,\ i\frac{\partial}{\partial y}\right).$$

From the last expression, it can be seen that H' is essentially the difference between two Schrödinger Hamiltonians of the same type in two independent variables x and y, hence one verifies that

$$[H,H'] = 0.$$

C. Applications to the study of energy eigenstates

If one chooses ψ to be $|\epsilon_m\rangle$, another energy eigenstate, then $\psi(p,q)$ takes the form

$$E_{mn}(p,q) = \langle\epsilon_n|U^+(p,q)|\epsilon_m\rangle \qquad \text{(II. 7a)}$$

The great advantage in dealing with an irreducible representation is that it allows us to express $E_{mn}(p,q)$ in terms of the Schrödinger wavefunctions[6]:

$$E_{mn}(p,q) = 1/2\pi\int_{-\infty}^{\infty}\psi_n(k-q)\exp(ikp)\,\psi_m(k+\tfrac{1}{2}q)\,dk;$$

but we also have

$$E_{mn}(p,q) = 1/2\pi\int_{-\infty}^{\infty}\tilde{E}_{mn}(k,q)\exp(ikp)\,dk. \qquad \text{(II. 7b)}$$

This would simply imply that

$$\tilde{E}_{mn}(k,q) = \psi_n(k-\tfrac{1}{2}q)\,\psi_m(k+\tfrac{1}{2}q), \qquad \text{(II. 8)}$$

or, alternatively, thanks to (I. 3),

$$E_{mn}(x,y) = \psi_n^*(x)\,\psi_m(y). \qquad \text{(II. 9)}$$

Now without spelling out the variables it can be verified from the expression (II. 4) that the amplitude E_{mn} satisfies the following system of equations:

$$H'E_{mn} = (\epsilon_m - \epsilon_n)E_{mn},$$
$$HE_{mn} = {}_mE_{mn}. \qquad \text{(II. 10)}$$

These equations are greatly simplified when $m=n$, i.e.,

$$H'E_{nn} = 0,$$
$$HE_{nn} = \epsilon_n E_{nn}. \qquad \text{(II. 11)}$$

Hence in order to study an energy eigenstate $|\epsilon_n\rangle$ in the Weyl representation one studies the amplitude E_{nn} which is nothing else than the average value of $U(p,q)$ in the state $|\epsilon_n\rangle$. The interesting point is that for the interactions we are considering it is easy to find the solution to the first equation of (II. 11) since the knowledge of the spectrum of H' is not required and since H' appears as a separable partial differential operator. Finally, one can use the normalization condition on $\psi_n(x)$ to normalize properly E_{nn}.

III. THE HARMONIC OSCILLATOR

The harmonic oscillator is a well-known testing ground which is widely discussed in the literature, in particular by Moyal and Bartlett.[11] Nevertheless, we would like to show in this section how the ideas put forward in the previous section can be applied to this simple and soluble problem. Moyal's results will be rederived with the help of Eqs. (II. 10), quite in contrast

to the use of the Schrödinger energy eigenfunctions in his approach.

The Hamiltonian

$$H = \tfrac{1}{2}[P^2 + \omega^2 Q^2]$$

is realized in $L^2(p, q)$ by the second order partial differential operator

$$H(p, q) = \tfrac{1}{2}\left(-i\frac{\partial}{\partial q} + \tfrac{1}{2}p\right)^2 + \frac{\omega^2}{2}\left(i\frac{\partial}{\partial p} + \tfrac{1}{2}q\right)^2 \qquad \text{(III. 1)}$$

which admits the integral of the motion

$$H'(p, q) = i\left(\omega^2 q\frac{\partial}{\partial p} - p\frac{\partial}{\partial q}\right). \qquad \text{(III. 2)}$$

Geometrically $H'(p, q)$ is nothing else than the generator of the rotation around the origin of the space of the variables p, q.

According to Eqs. (II. 10) the amplitude

$$E_{mn}(p, q) = \langle \epsilon_n | U^+(p, q) | \epsilon_m \rangle$$

are the eigenfunctions of $H(p, q)$ and $H'(p, q)$ with eigenvalues ϵ_m and $(\epsilon_n - \epsilon_m)$, respectively.

The single valued eigenfunctions of $H'(p, q)$ are of the form

$$E_{mn}(p, q) = f_{mn}(p^2 + \omega^2 q^2)\exp[-i(\epsilon_n - \epsilon_m)]\omega^{-1}\theta, \qquad \text{(III. 3)}$$

where

$$\epsilon_n - \epsilon_m = r\omega \quad \text{with} \quad r = 0, \pm 1, \pm 2, \cdots,$$

$$\tan\theta = p/\omega q.$$

The unknown functions $f_{mn}(x)$ can be found, if one requires that E_{mn} be eigenstate of the Hamiltonian. It can be verified that

$$\left(x\frac{d^2}{dx^2} + \frac{d}{dx} + \frac{\epsilon_n - \epsilon_m}{\omega} - x - \frac{r^2}{4x}\right) f_{mn} = 0,$$

where 4 $\omega x = (p^2 + \omega^2 q^2)$. With the change of variable $\rho = 2x$, we define a new function $g_{mn}(\rho)$ through the relation

$$f_{mn}(\rho) = \exp(-\rho)\sqrt{\tfrac{1}{2}\rho}\; g_{mn}(\rho).$$

Then the $g_{mn}(\rho)$ will satisfy

$$\left(\rho\frac{d^2}{d\rho^2} + (r + 1 - \rho)\frac{d}{d\rho} + (\epsilon_m + \epsilon_n)\omega^{-1} - (r + 1)\right) g_{mn}(\rho) = 0.$$

One can attempt to solve this equation by power series in ρ. However, in order to prevent its growth at infinity from overcoming the factor $\exp(-\rho)\sqrt{\tfrac{1}{2}\rho}$, one has to require that the series terminate: $g_{mn}(\rho)$ assumes then the form of an associated Laguerre polynomial[12] $L_n^{n-m}(\rho)$ if the following condition is satisfied:

$$(\epsilon_n + \epsilon_m) - \omega(r + 1) = 2m\omega. \qquad \text{(III. 4)}$$

Taking into account (III. 3) one finally obtains the energy levels

$$\epsilon_m = \omega(m + \tfrac{1}{2})$$

and the expression of $E_{mn}(p, q)$ up to a normalization factor N_{mn}:

$$E_{mn}(p, q) = N_{mn}\exp[-\tfrac{1}{2}\rho - i(n - m)\theta]\rho^{-(n-m)/2}L_n^{n-m}(\rho).$$

When $n = m$, $E_{nn}(p, q)$ is just a Laguerre function $N_{nn}\exp(-\rho/2)L_n(\rho)$. When $n = 0$ the ground state being chosen as the cyclic state of the representation we obtain the familiar coherent state representation if one sets $z = \sqrt{\tfrac{1}{2}\rho}\;\exp(i\theta)$:

$$E_{m0} = \langle z | \epsilon_m \rangle.$$

We shall not discuss the properties of the Laguerre polynomials further, since there exists ample literature on the subject.[12] However, it is interesting to notice that they are introduced here in a natural way as the solutions of differential operators arising from the representation of the CCR.

Finally, we should mention that Moyal in Ref. 2 also obtained the first Eq. (II. 11) with the time dependence, starting from a different approach.

IV. THE QUARTIC ANHARMONIC OSCILLATOR

A. The main result of the Weyl formulation

The quartic anharmonic oscillator has been quite a popular topic in recent years.[13] In this section we would like to present the derivation of some results reported earlier.[14] The model contemplated here is described by the Hamiltonian

$$H = \tfrac{1}{2}(P^2 + \omega^2 Q^2 + 2\lambda Q^4). \qquad \text{(IV. 1)}$$

Had we used a normal ordered term $:Q^4:$ we would have to introduce a correction to the frequency ω and a shift in the energy levels. However, these corrections do not add any new physical content to the problem and shall not be discussed.

The central problem here is the study of the eigenstate $|\epsilon_n\rangle$ of H in (IV. 1). From Sec. II, it is shown that the realization

$$H(p, q) = \tfrac{1}{2}\left(-i\frac{\partial}{\partial q} + \tfrac{1}{2}p\right)^2 + \frac{\omega^2}{2}\left(i\frac{\partial}{\partial p} + \tfrac{1}{2}q\right)^2 + \lambda\left(i\frac{\partial}{\partial p} + \tfrac{1}{2}q\right)^4$$

of H admits the integral of motion $H'(p, q)$,

$$H'(p, q) = i\left(\omega^2 q\frac{\partial}{\partial p} - p\frac{\partial}{\partial q}\right) - \lambda\left(4q\frac{\partial^3}{\partial p^3} - q^3\frac{\partial}{\partial p}\right),$$

which unfortunately does not possess a simple geometrical meaning.

Since $|\epsilon_n\rangle$ is the state to be studied, its properties can be described by the amplitude $E_{nn}(p, q) = \langle \epsilon_n | U^+(p, q) | \epsilon_m \rangle$ or by its Fourier transform $E_{nn}(k, q)$ given by Eq. (II. 7). It just happens that H' turns out to be a separable partial differential operator in $L^2(k, q)$ through the combinations

$$u = (\tfrac{1}{4}q^2 + k^2) \quad \text{and} \quad v = (\tfrac{1}{4}q^2 - k^2); \qquad \text{(IV. 2a)}$$

namely,

$$H'(k, q) = qk\left[\frac{\partial^2}{\partial u^2} - \frac{\partial^2}{\partial v^2} - 4\lambda u - \omega^2\right].$$

Hence Eq. (II. 11) for the $|\epsilon_n\rangle$ eigenstate of the quartic anharmonic oscillator will admit the general solution

$$E_{nn}(u, v) = \int_{-\infty}^{\infty} \phi(z) \begin{matrix}\cosh zv \\ \sinh zv\end{matrix} \mathrm{Ai}\left(\gamma u + \frac{\omega^2 + z^2}{\gamma^2}\right) dz,$$

where $4\lambda=\gamma^3$ and $\mathrm{Ai}(x)$ is the Airy function decreasing for $x\to\infty$, i.e.,

$$\mathrm{Ai}(x)=\int_{-\infty}^{\infty}\exp[i(t^3/3+xt)]\,dt.$$

Finally, $\phi_n(x)$ is an unknown function.

Noting that, because of Eqs. (II. 2), (II. 3) we have

$$u=\tfrac{1}{2}(x^2+y^2)\quad\text{and}\quad v=xy \tag{IV. 2b}$$

and that E_{nn} is related to the Schrödinger amplitudes by Eqs. (II. 9), we obtain the expansion

$$\psi_n(x)\,\psi_n(y)=\int_{-\infty}^{\infty}\phi_n(z)\begin{matrix}\cosh xyz\\ \sinh xyz\end{matrix}\mathrm{Ai}\left(\gamma\,\frac{x^2+y^2}{2}+\frac{\omega^2+z^2}{\gamma^2}\right)dz. \tag{IV. 3}$$

The choice of cosh or sinh in this equation depends obviously on the parity of $\psi_n(x)$.

To determine $\phi_n(z)$, we let the Schrödinger Hamiltonian

$$H(x)=-\tfrac{1}{2}\frac{d^2}{dx^2}+\tfrac{1}{2}\omega^2x^2+\lambda x^4$$

act on both sides of (IV. 3) and using partial integration reexpress this action as a differential operator on $\phi_n(z)$. To simplify the notation let us define

$$T=\left(\gamma\,\frac{x^2+y^2}{2}+\frac{\omega^2+z^2}{\gamma^2}\right).$$

Then

$$H(x)\,\psi_n(x)\,\psi_n(y)$$
$$=\int_{-\infty}^{\infty}\phi_n(z)[\tfrac{1}{2}z^2(x^2+y^2)-\tfrac{1}{4}\gamma^3x^2y^2]\begin{matrix}\cosh xyz\\ \sinh xyz\end{matrix}\mathrm{Ai}(T)\,dz$$
$$+\int_{-\infty}^{\infty}\phi_n(z)\left[-xyz\begin{matrix}\sinh xyz\\ \cosh xyz\end{matrix}-\begin{matrix}\cosh xyz\\ \sinh xyz\end{matrix}\right]\mathrm{Ai}'(T)\,dz.$$

Every expression containing x and y can be rewritten as a differential operator in the variable z, i.e.,

$$-\tfrac{1}{2}z^2(x^2+y^2)\,\mathrm{Ai}(T)$$
$$=\left(\frac{\gamma^3}{4z}\,\frac{d}{dz}-\frac{1}{4}\gamma^3\frac{d^2}{dz^2}+\frac{1}{\gamma^3}z^2(z^2+\omega^2)\right)\mathrm{Ai}(T),$$

$$x^2y^2\begin{pmatrix}\cosh xyz\\ \sinh xyz\end{pmatrix}=\frac{d^2}{dz^2}\begin{pmatrix}\cosh xyz\\ \sinh xyz\end{pmatrix}.$$

Hence

$$H(x)\,\psi_n(x)\psi_n(y)=\int_{-\infty}^{\infty}\phi_n(z)\left(-\tfrac{1}{4}\gamma^3\frac{d^2}{dz^2}+\gamma^{-3}z^2(z^2+\omega^2)\right)\times\begin{matrix}\cosh xyz\\ \sinh xyz\end{matrix}\mathrm{Ai}(T)\,dz.$$

From this last equation it can be seen that the original Schrödinger eigenvalue problem is now shifted to an eigenvalue problem on $\phi_n(z)$ which satisfies the equation

$$\left(-\frac{1}{4}\gamma^3\frac{d^2}{dz^2}+\gamma^{-3}z^2(z^2+\omega^2)\right)\phi_n(z)=\epsilon_n\,\phi_n(z)$$

with the new boundary conditions

$$\left(\phi_n(z)\,\mathrm{Ai}(T)\begin{matrix}\cosh xyz\\ \sinh xyz\end{matrix}\right)_{-\infty}^{\infty}=0,$$

$$\left[\phi_n(z)\,\frac{d}{dz}\left(\mathrm{Ai}(T)\begin{matrix}\cosh xyz\\ \sinh xyz\end{matrix}\right)\right]_{-\infty}^{\infty}=0,$$

which force $\phi_n(x)$ to be of the same type as $\psi_n(x)$.

Let us perform the following scale transformations on the equation satisfied by $\phi_n(z)$:

$$z=\sigma t,\quad \omega=\sigma\omega',\quad \epsilon_n=\sigma\epsilon_n', \tag{IV. 4}$$

where $\gamma^3=2\sigma^3$. The transformed equation is now

$$\left(-\frac{1}{2}\frac{d^2}{dt^2}+\tfrac{1}{2}\omega'^2t^2+\tfrac{1}{2}t^4\right)\phi_n(t)=\epsilon_n'\,\phi_n(t). \tag{IV. 5}$$

The previous transformations reveal the correct dependence of the energy levels on the cubic root of the coupling constant λ.[13] Since it is well known that there is no loss in generality as far as the properties of the Schrödinger wavefunctions are concerned in assuming $\lambda=\frac{1}{2}$, we shall stick to this particular value of λ in the rest of the discussion, and set

$$\phi_n(z)=\mu_n\psi_n(z)$$

where μ_n is a constant of proportionality.

The expansion (IV. 3) assumes now the well-known form of an integral equation for zonal spherical functions[15] (e.g., Legendre, Gegenbauer polynomials, Bessel functions ⋯)

$$\psi_n(x)\,\psi_n(y)=\mu_n\int_{-\infty}^{\infty}\psi_n(z)\begin{matrix}\cosh xyz\\ \sinh xyz\end{matrix}\mathrm{Ai}\left(\frac{x^2+y^2+z^2+\omega^2}{\gamma^2}\right)dz \tag{IV. 6}$$

where $\gamma^3=2$.

We observe that if $\psi_n(x)$ is a solution of (IV. 6) corresponding to μ_n then $-\psi_n(x)$ is also a solution corresponding to the value $-\mu_n$; since both wavefunctions describe the same physical probability distribution we can assume that $\mu_n>0$ for all n.

B. Homogenous integral transformation for $\psi_n(n)$

A linear homogenous integral equation can be derived from equation (IV. 6) if one multiplies by $\psi_n(y)$ on both sides of it and replaces the product $\psi_n(y)\,\psi_n(z)$ on the right-hand side by its expression given by (IV. 6). The normalization condition on $\psi_n(y)$ yields

$$\psi_n(x)=\mu_n^2\int_{-\infty}^{\infty}\psi_n(y)K_\pm(x,y)\,dy, \tag{IV. 7}$$

where the kernels $K_\pm(x,y)$ are defined by

$$K_\pm(x,y)=\int\int\begin{matrix}\cosh xuv\,\cosh yuv\\ \sinh xuv\,\sinh yuv\end{matrix}\mathrm{Ai}\left(\frac{x^2+u^2+v^2+\omega^2}{\gamma^2}\right)\times\mathrm{Ai}\left(\frac{y^2+u^2+v^2+\omega^2}{\gamma^2}\right)du\,dv.$$

The integral transforms $T_\pm$ defined by the symmetric kernels $K_\pm(x,y)$ commute with the Schrödinger Hamiltonian $H(x)$. It is interesting to notice that T is analogous to the Fourier transform in the problem of the harmonic oscillator. Equation (IV. 7) is of the Fredholm type. The conventional approach to Fredholm integral equations is well known,[16] but we shall seek an alternative approach to the solutions of (IV. 7).

We shall restrict our discussion to the case of even parity wavefunctions, and the results can be trivially extended to odd parity wavefunctions.

Consider the function of x defined by the integral

$$F_0(x)=\int_{-\infty}^{\infty}\cosh xu\ \mathrm{Ai}\left(\gamma u^2+\frac{\omega^2+x^2}{\gamma^2}\right)du.$$

$F_0(x)$ is obviously a square integrable function in x since it is bounded and integrable over the real line. $F_0(x)$ does not have any zero for $\omega^2>\gamma r$ where $r=-2.338\cdots$ is the first root of the Airy function $\mathrm{Ai}(x)$. However, as ω^2 goes increasingly negative a first zero will appear at the origin when ω^2 satisfies the condition

$$\int_{-\infty}^{\infty}\mathrm{Ai}\left(\gamma u^2+\frac{\omega^2}{\gamma^2}\right)du=0.$$

In fact, this equation has an infinite number of roots ω_j^2, $j=1,2,\cdots$, for which when $\omega^2<\omega_j^2$ there exists at least j zeros of $F_0(x)$ on $x>0$ and for which the sign of $F_0(0)$ is $(-1)^j$.

Equation (IV. 6) shows that

$$\mu_{2k}^{-1}=(\psi_{2k},F_0),\quad k=0,1,2,\cdots, \tag{IV. 8}$$

and using the Sturm—Liouville theorem for $H(x)$ on $L^2(x)$, we have the expression

$$F_0(x)=\sum_{k=0}^{\infty}\mu_{2k}^{-1}\psi_{2k}(x). \tag{IV. 9}$$

More generally, if one defines the functions

$$F_m(x)=T_+^mF_0(x),$$

then they would have the expansion

$$F_m(x)=\sum_{k=0}^{\infty}(\mu_{2k})^{-2m-1}\ \psi_{2k}(x) \tag{IV. 10}$$

and T^m are trace class operators: $\mathrm{Tr}\ T_+^m=(F_{m-1},F_0)$ for $m=1,2,3,\cdots$, where $(\ ,\ )$ is the inner product in $L_+^2(x)$, the Hilbert space of square integrable functions which are even on R. Moreover, using (IV. 6), (IV. 9), (IV. 10) the kernel of T_+^m can be either calculated explicitly from $F_m(x)$ or be expanded in terms of the Schrödinger wavefunctions as follows:

$$K_+^m(x,y)=\int_{-\infty}^{\infty}F_{m-1}(z)\cosh xyz\ \mathrm{Ai}\left(\frac{x^2+y^2+z^2+\omega^2}{\gamma^2}\right)dz,$$

$$K_+^m(x,y)=\sum_{k=0}(\mu_{2k})^{-2m}\psi_{2k}(x)\psi_{2k}(y). \tag{IV. 11}$$

Finally, from (IV. 9), (IV. 10), (IV. 11) we can derive the numerical equations

$$(F_{n-1},F_0)=\sum_{k=0}^{\infty}(\mu_{2k}^{-2})^n, \tag{IV. 12}$$

$$(F_{n-1}HF_0)=\sum_{k=0}^{\infty}(\mu_{2k}^{-2})^n\epsilon_{2k} \tag{IV. 13}$$

for $n=1,2,\cdots$.

Equation (IV. 8) shows that μ_{2k}^{-1} is a positive real number for an appropriate choice of phase for $\psi_{2k}(x)$ and consequently that T_+ is a positive transformation. For a fixed value of ω^2, the overlap amplitude (ψ_{2k},F_0) will be maximal when $\psi_{2k}(x)$ looks close to $F_0(x)$, in particular when both $\psi_{2k}(x)$ and $F_0(x)$ have the same number of zeros on the real line. From Schwartz inequality it follows that $\mu_{2k}^{-2}\leqslant(F_0,F_0)$.

There are, however, values of ω^2 for which the overlap amplitude (ψ_{2k},F_0) is the same for a finite number of k (since for k sufficiently large, μ_{2k}^{-2} will tend to 0); then μ_{2k}^{-2} is a degenerate eigenvalue of T_+.

Let μ_{2i}^{-2} be the largest eigenvalue of T_+ appearing d_i times in (IV. 12); then independently of d_i

$$\mu_{2i}^{-2}=\lim_{n\to\infty}\frac{(F_n,F_0)}{(F_{n-1},F_0)}\ .$$

Again from Eq. (IV. 12) one can see that $(F_{n-1},F_0)^{1/n}$ will converge to μ_{2i}^{-2} if and only if $d_i=1$, otherwise d_i would be the number for which

$$\lim_{n\to\infty}\left(\frac{(F_{n-1},F_0)}{d_i}\right)^{1/n}=\mu_{2i}^{-2}.$$

Then if $d_i=1$, Eqs. (IV. 10) and (IV. 12) show that one can determine the wavefunction as

$$\psi_{2i}(x)=\lim_{m\to\infty}\left[(\mu_{2i})^{2m+1}F_m(x)\right]$$

and the energy levels

$$\epsilon_{2i}=\lim_{n\to\infty}\mu_{2i}^{2n}(F_{n-1},HF_0).$$

So long as there is no degeneracy, this process can be repeated to determine all subsequent μ_{2k}^{-2}, ϵ_{2k}, and $\psi_{2k}(x)$ in a recursive way.

Unfortunately, the determination of $\psi_{2k}(x)$ and ϵ_{2k} in the case of $d_i>2$ remains an unsolved problem. The previous procedure of taking limits would yield only a mixture of equal weights of eigenstates and consequently of energy levels. It seems that one would need an additional condition to lift this degeneracy.

But the values of ω^2 for which transformation T_+ has degenerate eigenvalues (in finite multiplicity) form only a discrete set of numbers which annihilate the Van der Monde determinant[17] of the system of linear equations in ϵ_{2k} (IV. 12). We note that the crossing of some of the μ_{2k}^{-2} as functions of ω^2 may correspond to some of the physical features of the quartic anharmonic oscillator unseen before in the Schrödinger representation. In particular, this would happen most likely when $\omega^2<\gamma^2v$ in the usual situation of the double well potential.

Finally, we would like to mention that numerical calculations have been performed in [14] as an illustration of the approximation scheme which is apparent from this approach.

C. Adjunction of a radial potential

When a centrifugal potential, arising from a separation of variables in polar coordinates, for example, the term

$$L(L+1)/2x^2,\quad x>0,$$

is added to the Schrödinger Hamiltonian, the operator H' again remains soluble in the variable u and v as defined by Eq. (IV. 2):

$$H'(u,v)=\left(\frac{\partial^2}{\partial u^2}-\frac{\partial^2}{\partial v^2}-2u-\omega^2+\frac{L(L+1)}{v^2}\right).$$

Let $\psi_{mL}(x)$ be the eigenfunction of the Schrödinger Hamiltonian

$$H(x)=\left(-\frac{1}{2}\frac{d^2}{dx^2}+\tfrac{1}{2}\omega^2x^2+\tfrac{1}{2}x^4+\frac{L(L+1)}{2x^2}\right);$$

then we obtain, following the same procedure, the integral equation

$$\psi_{mL}(x)\,\psi_{mL}(y)=\mu_{mL}\int_0^\infty \psi_{mL}(z)\sqrt{xyz}\;I_{L+1/2}(xyz)\,\mathrm{Ai}(T)\,dz \tag{IV. 14}$$

where

$$T=(\omega^2+x^2+y^2+z^2)/\gamma^2$$

and

$$\left(\frac{d^2}{dv^2}-\frac{L(L+1)}{v^2}-z^2\right)\sqrt{zv}\;I_{L+1/2}(zv)=0,$$

$I_{L+1/2}$ being a modified Bessel function regular at the origin. For $L\to 0$ we recover the odd solutions of the previous case. It is interesting to note that the "zonal" character is not destroyed by the centrifugal potential and that $\psi_{mL}(x)$ again satisfies a homogeneous integral equation

$$\psi_{mL}(x)=\mu_{mL}\int_0^\infty K(x,y)\,\psi_{mL}(y)\,dy$$

with kernel

$$K(x,y)=\int_0^\infty F_{0L}(z)\sqrt{xyz}\;I_{L+1/2}(xyz)\,\mathrm{Ai}(T)\,dz,$$

where

$$F_{0L}(z)=\int_0^\infty u\sqrt{z}\;I_{L+1/2}(zu^2)\,\mathrm{Ai}\left(\gamma u^2+\frac{\omega^2+z^2}{\gamma^2}\right)du.$$

V. THE SEXTIC ANHARMONIC OSCILLATOR

A. Main results of Weyl's formulation

The Hamiltonian

$$H=\tfrac{1}{2}P^2+\tfrac{1}{2}\omega^2Q^2+\tfrac{1}{2}\Omega^2Q^6 \tag{V. 1}$$

has been studied in the past in the Schrödinger representation through singular perturbations[18] and through the Hill determinant method.[19] Moreover, one can show that the Schrödinger eigenfunctions of the Hamiltonian (V. 1) are expressible in terms of a confluent form of the Heun functions.[20] In this section we shall demonstrate that Weyl's formulation of the Hamiltonian (V. 1) leads to the expansion of the Schrödinger eigenfunctions in terms of parabolic cylinder functions and the eigenvalue problem can be reduced to the finding of the eigenvalues of a very simple symmetric matrix.

We first remark that the operator H' of the Hamiltonian H is again separable in the variables u and v defined by Eq. (IV. 2):

$$H'(u,v)=\left(-\frac{1}{2}\frac{\partial^2}{\partial u^2}+\frac{4\Omega^2}{2}u^2\right)-\left(-\frac{1}{2}\frac{\partial^2}{\partial v^2}+\frac{\Omega^2}{2}v^2\right)+\frac{\omega^2}{2}. \tag{V. 2}$$

Hence $E_{nn}(k,q)$ or $\tilde{E}_{nn}(u,v)$, Fourier transform of

$$E_{nn}(p,q)=\langle\epsilon_n|\,U^*(p,q)\,|\epsilon_n\rangle$$

where $|\epsilon_n\rangle$ is the nth eigenstate of H, has the expansion with coefficient $c_\mu(n)$:

$$E_{nn}(u,v)=\sum_\mu c_\mu(n)D_\nu(2\sqrt{\Omega}\,u)D_\mu(\sqrt{2\Omega}v).$$

In the expansion $D_\nu(x)$ is the well-known parabolic cylinder function solution of the equation

$$\left(\frac{d^2}{dx^2}-\tfrac{1}{2}x^2+(2\nu+1)\right)D_\nu(x)=0$$

which decreases at infinity ($x\to+\infty$). The indices μ and ν of the parabolic cylinder functions according to Eq. (II. 11) are related by

$$\mu=2\nu+\beta$$

where β is a naturally scale invariant parameter: $\frac{1}{2}(1+\omega^2/\Omega)$

Equations (II. 8, II. 9) help to obtain the expansion of the Schrödinger eigenfunctions product $\psi_n(x)\psi_n(y)$ with real numerical coefficients $c_\mu(n)$:

$$\psi_n(x)\,\psi_n(y)=\sum_\mu c_\mu(n)\,D_\nu(\sqrt{\Omega}\,(x^2+y^2))\,D_\mu(\sqrt{2\,\Omega}\,xy). \tag{V. 3}$$

To ensure that the right-hand side of this expansion shows the appropriate behavior under the space inversion $x\to -x$ (or $y\to -y$), we are led to choose μ even (odd) positive integers whenever n is even (odd). The parabolic cylinder functions $D_\mu(x)$ are just in this case the harmonic oscillator Schrödinger eigenfunctions $H_\mu(x)$ of frequency Ω.

To determine $c_\mu(n)$ we let the Schrödinger Hamiltonian act on the expansion (IV. 3):

$$\begin{aligned}&[\psi_n''-(\omega^2x^2+\Omega^2x^6)\psi_n]\,\psi_n(y)\\&\quad=\sqrt{\Omega}\sum_\mu c_\mu(n)\,2D_\nu'(\sqrt{\Omega}\,(x^2+y^2))\,[H_\mu(\sqrt{2\Omega}\,xy)\\&\qquad+2\sqrt{2\Omega}\,xy\,H_\mu(\sqrt{2\Omega}\,xy)]\\&\qquad+\sqrt{\Omega}\sum_\mu c_\mu(n)\sqrt{\Omega}\,(x^2+y^2)D_\nu\,(\sqrt{\Omega}\,(x^2+y^2))\\&\qquad\times[2\Omega x^2y^2-(2\mu+1)]H_\mu(\sqrt{2\Omega}\,xy).\end{aligned}$$

From the known recursion relations for $D_\nu(x)$ (Ref. 12) we have with the use of (IV. 2)

$$2D_\nu'(2\sqrt{\Omega}\,u)=\nu D_{\nu-1}(2\sqrt{\Omega}\,u)-D_{\nu+1}(2\sqrt{\Omega}\,u),$$

$$2\sqrt{\Omega}\,uD_\nu(2\sqrt{\Omega}\,u)=\nu D_{\nu-1}(2\sqrt{\Omega}\,u)-D_{\nu+1}(2\sqrt{\Omega}\,u),$$

and similarly from the recursion relations of the harmonic oscillator eigenfunctions:

$$\begin{aligned}H_\mu(\sqrt{2\Omega}\,v)+2vH_\mu'(\sqrt{2\Omega}\,v)&=\sqrt{\mu(\mu-1)}\;H_{\mu-2}(\sqrt{2\Omega}\,v)\\&\quad-\sqrt{(\mu+1)(\mu+2)}\;H_{\mu+2}(\sqrt{2\Omega}\,v),\\[2\Omega v^2-(2\mu+1)]H_\mu(\sqrt{2\Omega}\,v)&=\sqrt{\mu(\mu-1)}\,H_{\mu-2}(\sqrt{2\Omega}\,v)\\&\quad+\sqrt{(\mu+1)(\mu+2)}\,H_{\mu+2}(\sqrt{2\Omega}\,v).\end{aligned}$$

Hence

$$\begin{aligned}&[\psi''-(\omega^2x^2+\Omega^2x^6)\psi_n]\psi_n(y)\\&\quad=2\sqrt{\Omega}\sum_\mu[\sqrt{\mu(\mu-1)}\;c_{\mu-2}(n)\\&\qquad+\nu\sqrt{(\mu+1)(\mu+2)}\;c_{\mu+2}(n)]D_\nu(\sqrt{\Omega}\,(x^2+y^2)H_\mu(\sqrt{2\Omega}\,xy).\end{aligned}$$

The eigenvalue problem in the Schrödinger representa-

TABLE. I. Convergence of the energy levels ϵ_k as a function of N, the dimension of the symmetric matrix approximating A.

$N=$	40	80	160	320
ϵ_0	0.7647	0.7501	0.7402	0.7336
ϵ_1	1.689	1.639	1.609	1.588
ϵ_2	5.812	5.537	5.360	5.248
ϵ_3	7.607	7.214	6.962	6.795

tion is now shifted to the difference equation

$$\sqrt{\mu(\mu+1)}\; c_{\mu-2}(n) + \nu\sqrt{(\mu+1)(\mu+2)}\; c_{\mu+2}(n) + (\epsilon_n/\sqrt{\Omega})c_\mu(n) = 0. \tag{V.4}$$

This is a numerical system of linear homogeneous equations in $c_\mu(n)$, with coefficients depending on two dimensionless quantities β and the scaled energy value

$$\epsilon' = \epsilon/\sqrt{\Omega}.$$

The existence of a nontrivial solution requires that the determinant should be zero.

Since μ takes values in the set of positive even or odd integers, we shall restrict ourselves to the case where $\mu = 2p$ ($p = 0, 1, 2, \cdots$) corresponding to the Schrödinger wavefunction $\psi_{2p}(x)$. The extension to odd wavefunctions $\psi_{2p+1}(x)$ is trivial. For $\mu = 2p$, we have $\nu(p) = p - \beta/2$ with $p = 0, 1, 2, \cdots$. Hence if $\beta/2 < 0$, then all $\nu > 0$; and if $p < \beta/2 < p+1$, ν will take $(p+1)$ negative values. In particular, when $\beta/2 = p$ then $\nu = 0$, and the system (V. 4) is split into two groups of equations. The first group of p linear homogeneous equations in $c_0, \ldots, c_{2(p-1)}$ has, in general, the trivial solution $c_0 = \cdots = c_{2(p-1)} = 0$ because the vanishing of its determinant leads to imaginary values for ϵ':

$$\begin{vmatrix} \epsilon' & \nu(1)\sqrt{1.2} & 0 & \cdots & 0 \\ \sqrt{1.2} & \epsilon' & \nu(2)\sqrt{3.4} & \cdots & 0 \\ 0 & \sqrt{3.4} & \epsilon' & \cdots & 0 \\ & & & & \vdots \\ \vdots & \vdots & \vdots & & \nu(p-1)\sqrt{(2p)(2p-1)} \\ 0 & 0 \cdots & \sqrt{(2p)(2p-1)} & & \epsilon' \end{vmatrix} = 0.$$

The second group of equations for $c_{2p}, c_{2p+2}, \cdots$ is of the usual type which shall be discussed. Consequently, for those special values of $\beta/2$ the expansion (V. 3) would have started with $H_{2p}(x)$ rather than $H_0(x)$.

When $p < \beta/2 < p+1$, since all c_{2k} and ϵ_{2k} for $k < p$ are continuous in $\beta/2$, the reality of the expansion (V. 3) requires that $c_{2k} = 0$ for $k < p$. The only system of equations to be considered is (V. 4) with $\mu > 2p$ its involves only one negative value of ν, i. e., $\nu(p) < 0$ and $\nu(k) > 0$ for $k > p$. Then in this latter situation we have

$$\epsilon' c_{2p} + \nu(p+1)\sqrt{(2p+1)(2p+2)}\, c_{2p+2} = 0,$$

$$\sqrt{(2p+1)(2p+2)}\, c_{2p} + \epsilon' c_{2p+2} + \nu(p+2)\sqrt{(2p+3)(2p+4)}\, c_{2p+4} = 0,$$

$$\sqrt{(2p+3)(2p+4)}\, c_{2p+2} + \epsilon' c_{2p+4} + \nu(p+3)\sqrt{(2p+5)(2p+6)}\, c_{2p+6} = 0,$$

$$\cdots$$

Since

$$\nu(p+1) - 1 = \nu(p) = 1 + \nu(p-1),$$

one can define new coefficients d_{2k} by the relations

$$c_{2p} = \left(\frac{\nu(p+1)}{\Gamma[\nu(p+1)]}\right)^{1/2} d_{2p},$$

$$c_{2k} = \frac{1}{\sqrt{\Gamma[\nu(k)]}}\; d_{2k} \quad \text{for} \quad k > p$$

and obtain a completely symmetric system of linear homogeneous equations in d_{2k}, $k > p$. The vanishing of the determinant now can be viewed as the characteristic equation of a real symmetric matrix of the type

$$A = \begin{pmatrix} 0 & a_p & 0 & 0 & & 0 \cdots \\ a_p & 0 & a_{p+1} & 0 & & \\ 0 & a_{p+1} & 0 & a_{p+2} & & \\ & \ddots & \ddots & \ddots & \ddots & \\ 0 & & & a_{p+k-1} & 0 & a_{p+k} \\ \vdots & & & & \ddots & \ddots \end{pmatrix}$$

with

$$a_{p+k} = \sqrt{\nu(p+k)}\; \sqrt{(2p+2k+1)(2p+2k+2)}.$$

As expected, the eigenvalues of A obviously real and the coefficients $c_\mu(n)$ can be computed theoretically up to a factor which in turn will be determined by the normalization condition on $\psi_n(x)$.

Finally, we note that for $\beta/2 < 0$ and large, the case of the double well potential, there is no problem in symmetrizing the matrix for there is no negative ν. The expansion (V. 3) starts only with a large value of ν.

B. Numerical results

A straightforward approximation scheme can be set by restricting ourselves to the $N \times N$ upper submatrix of A. As an illustration we have choosen $\omega = \Omega = 1$ and proceed with various values of N. The results for low-lying levels are reported in Table I and show the slow convergence of the approximation steps. We have made no attempt to improve this numerical convergence but have just used a standard routine to diagonalize a real symmetric matrix. Graphical interpolation yields the following values:

$$\epsilon_0 = 0.718, \quad \epsilon_1 = 1.571,$$

$$\epsilon_2 = 5.08, \quad \epsilon_3 = 6.64.$$

These numbers can be compared to those of Ref. 19.

C. Adjunction of other interactions

The Schrödinger Hamiltonian

$$H_J(x) = -\frac{1}{2}\frac{d^2}{dx^2} + \tfrac{1}{2}\omega^2 x^2 + \tfrac{1}{2}\Omega^2 x^6 + \frac{J^2 - 1/4}{2x^2}$$

admits in Weyl's formulation the following integral of the motion:

$$H'_J = \left(-\frac{1}{2}\frac{\partial^2}{\partial u^2} + \frac{4\Omega^2}{2}u^2\right) - \left(-\frac{1}{2}\frac{\partial^2}{\partial v^2} + \frac{\Omega^2}{2}v^2 + \frac{J^2-1/4}{2v^2}\right) + \frac{\omega^2}{2},$$

where u and v have the definition given in Sec. V.A. Then $\psi_{nJ}(x)$, the Schrödinger eigenfunction of $H(x)$, verifies

$$\psi_{nJ}(x)\,\psi_{nJ}(y) = \sum_q c_q(n,J)\,D_\nu(\sqrt{\Omega}\,(x^2+y^2))\,G_{qJ}(\sqrt{\Omega}\,xy), \quad \text{(V. 5)}$$

where $G_{qJ}(v)$ is the solution of the differential equation

$$\left(-\frac{1}{2}\frac{d^2}{dv^2} + \frac{1}{2}\omega^2 v^2 + \frac{J^2-1/4}{2v^2}\right)G_{qJ} = (2q+J+1)\omega\,G_{qJ}$$

vanishing at the origin and normalized as the radial harmonic oscillator eigenfunction.

The indices ν and q are related by

$$2q + (J+\tfrac{1}{2}) = 2\nu + \beta.$$

Again, to determine $c_q(n,J)$ we calculate the action of $H(x)$ on (V. 5):

$$\begin{aligned} &H_J(x)\,\psi_{nJ}(x)\,\psi_{nJ}(y) \\ &= \sqrt{\Omega}\sum_0^\infty c_q(n,J)\,2D'_\nu(\sqrt{\Omega}\,(x^2+y^2))\,[G_{qJ}(\sqrt{\Omega}\,xy) \\ &\quad + 2\sqrt{\Omega}\,xy\,G'_{qJ}(\sqrt{\Omega}\,xy)] \\ &\quad + 2\sqrt{\Omega}\sum_0^\infty c_q(n,J)\,\sqrt{\Omega}\,(x^2+y^2)\,D_\nu(\sqrt{\Omega}\,(x^2+y^2)) \\ &\quad \times[\Omega\,x^2y^2 - (2q+J+1)]G_{qJ}(\sqrt{\Omega}\,xy). \end{aligned} \quad \text{(V. 6)}$$

From the recursion relations of the Laguerre polynomials[12]

$$[\Omega x^2y^2 - (2q+J+1)]G_{qJ} = -\sqrt{(q+1)(q+J+1)}\;G_{(q+1)J} - \sqrt{(q+J)q}\;G_{(q-1)J},$$

$$G_{qJ}(\sqrt{\Omega}\,xy) + 2\sqrt{\Omega}\,xyG'_{qJ}(\sqrt{\Omega}\,xy) = 2\sqrt{(q+1)(q+1+J)}\,G_{(q+1)J} - 2\sqrt{(q+J)q}\;G_{(q-1)J}.$$

This implies in (V. 6) that

$$\nu\sqrt{(q+1)(q+1+J)}\;c_{q+1}(n,J) + \sqrt{q(q+J)}\;c_{q-1} = \epsilon'\,c_q, \quad \text{(V. 7)}$$

where the ϵ' the scaled energy value given by

$$\epsilon' = \epsilon/2\sqrt{\Omega}$$

appears to be the eigenvalue of an infinite real symmetrizable matrix analogous to A, along the lines discussed before in the preceding paragraphs.

Finally, one can consider the Hamiltonian

$$H_{J\lambda}(x) = \left[\left(-\frac{1}{2}\frac{d^2}{dx^2} + \tfrac{1}{2}\omega^2x^2 + \tfrac{1}{2}\lambda x^4 + \frac{1}{2}\frac{J^2-1/4}{x^2}\right) + \tfrac{1}{2}\Omega^2x^6\right].$$

Its eigenfunction $\psi_{nJ}(x)$ satisfy the expansion

$$\psi_{nJ}(x)\;\psi_{nJ}(y) = \sum_0^\infty c_q(n,J,\lambda)\,D_\nu(\sqrt{\Omega}\,(x^2+y^2+\lambda/2\Omega^2))\,G_{qJ}(\sqrt{\Omega}\,xy),$$

where

$$2q + J + \tfrac{1}{2} = 2\nu + \beta'$$

with

$$\beta' = (\beta - \lambda^2/4\Omega^3)$$

and c_q satisfying a recursion relation similar to (V. 7).

VI. CONCLUSION

In this article we have shown that Weyl's formulation of the CCR can be used to treat nontrivial interactions in the quantum mechanics of a single degree of freedom system. This formalism is based on the existence of an integral of the motion. It leads to the natural separation of odd and even states. It does predict the correct scaling of the energy levels in terms of the coupling constant. It does accomodate the addition of a radial centrifugal potential. And most crucial of all, it does set the eigenvalue problem in a novel context suitable for a natural approximation scheme. Moreover, the expansions (IV. 6) and (V. 3) exhibit more explicitly the analytic properties of the Schrödinger eigenfunctions.

However, we did not try to deal with the interaction Q^8 since the integral of the motion is unfortunately not separable in the variables u and v.

The present study has by no means exhausted the new revelations brought into light by Weyl's formulation. Equation (IV. 6) proves that some group theoretical background is still to be discovered[21] and it is quite obvious that from the structure of its integral of the motion H', the sextic anharmonic oscillator is related to the two-dimensional anisotropic oscillator in Minskowski space and therefore to its group of dynamical symmetry. We hope to treat these problems in a forthcoming publication.

Finally, needless to say that in many physical situations the radial potential arises from the separation of variables in spherical coordinates of some higher dimensional problems of statistical mechanics.[22] It may be there that the degeneracy of some eigenvalues of the transformation $T_\pm$ found in the quartic anharmonic oscillator may be of some unsuspected physical relevance.

ACKNOWLEDGMENTS

Part of this work was done when the author was a resident visitor at Bell Laboratories at Murray Hill, New Jersey. He is grateful to Dr. John R. Klauder and Professor H. Ezawa for having many instructive discussions on the subject at all stages. Professor Henry M. Foley has made available the computing facilities of Columbia University and Dr. Huynh B. Hanh has given his advice on numerical calculations.

*Present address: Institut für Theoritische Physik, Freie Universität Berlin, 1-Berlin 33, W. Germany.

[1]H. Weyl, *Group Theory and Quantum Mechanics* (Dover, New York, 1928).

[2]J. E. Moyal, Proc. Camb. Phil. Soc. **45**, 99 (1949).
[3]H. Araki, J. Math. Phys. **1**, 492 (1960).
[4]I. E. Segal, *Mathematical Problems of Relativistic Physics* (American Mathematical Society, Providence, R. I., 1963).
[5]V. Bargmann, Commun. Pure Appl. Math. **14**, 187 (1961).
[6]J. R. Klauder, J. Math. Phys. **5**, 878 (1964).
[7]R. Glauber in *Quantum Optics and Electronics*, edited by C. de Witt (Gordon and Breach, New York, 1964); E. G. C. Sudarshan and J. R. Klauder, *Quantum Optics* (Benjamin, New York, 1968).
[8]J. Grodnick, Phys. Rev., D **3**, 2955 (1971).
[9]J. R. Klauder, "Coherence, Correlation and Quantum Field Theory" in 1967 *Brandeis Lectures*, edited by M. Chretien and S. Schweber (Gordon and Breach, New York, 1972).
[10]T. T. Truong, Columbia University Thesis, unpublished, 1973.
[11]M. S. Bartlett and J. E. Moyal, Proc. Camb. Phil. Soc. **45**, 545 (1949).
[12]W. Magnus, F. Oberhettinger, and R. Soni, *Formula and Theorems for the Special Functions of Mathematical Physics* (Springer, New York, 1966).
[13]T. T. Wu and C. M. Bender, Phys. Rev. **184**, 1231 (1969); Phys. Rev. D **7**, 1620 (1973); B. Simon, Ann. Phys. (N.Y.) **58**, 76 (1970); *Cargese Lectures in Physics* 1970, edited by O. Bessis (Gordon and Breach, New York, 1972),
[14]T. T. Truong, Nuovo Cimento Lett. **14**, 533 (1974).
[15]N. J. Vilenkin, "Group Representations," Amer. Math. Soc. Trans. **22**, (1968).
[16]H. Widom, *Integral Equations* (Van Nostrad, New York, 1969).
[17]W. Greub, *Linear Alegebra* (Springer, New York, 1963).
[18]C. M. Bender, J. Math. Phys. **11**, 796 (1970).
[19]S. N. Biswas *et al.*, J. Math. Phys. **14**, 1190 (1973).
[20]A. Lemieux and A. K. Bose, Ann. Inst. H. Poincaré **10**, 259 (1969).
[21]F. P. Berezin and I. M. Gelfand, Amer. Math. Soc. Trans. (2), **21**, 193 (1962).
[22]R. Balian and G. Toulouse, Ann. Phys. (N.Y.) **83**, 28 (1974).

An integral equation for the Gel'fand–Levitan kernel in terms of the scattering potential in the one-dimensional case

H. E. Moses

Air Force Cambridge Laboratories, Hanscom Air Force Base, Massachusetts 01731
(Received 6 November 1974; revised manuscript received 10 January 1975)

An integral equation for the Gel'fand–Levitan kernel is given in terms of the scattering potential. This integral equation may be regarded as complementary to the Gel'fand–Levitan equation which is an integral equation for the kernel in terms of the Fourier transform of the reflection coefficient.

1. REVIEW OF THE GEL'FAND-LEVITAN ALGORITHM FOR ONE DIMENSION

We take as our space variable x with $-\infty < x < +\infty$. We shall assume the scattering potential $V(x)$ vanishes with sufficient rapidity as $x \to \pm\infty$ for the operator $-d^2/dx^2 + V(x)$ to be Hermitian with the usual complex inner product and have a spectrum and eigenfunctions of the usual character described below.

The operator $-d^2/dx^2 + V(x) = H$ has a continuous spectrum which is the entire positive real axis and in addition may have negative point eigenvalues. Denoting an eigenvalue of the continuous spectrum by $E = p^2$ with $-\infty < p < +\infty$, the corresponding eigenfunctions (there are two, one for each sign of p) are denoted by $\psi(x\,|\,p)$. The eigenvalue equation $H\psi(x\,|\,p) = p^2\psi(x\,|\,p)$ is a second order differential equation which requires boundary conditions to specify ψ uniquely. One such boundary condition gives rise to $\psi_s(x\,|\,p)$ which are of interest in scattering theory:

$$\lim_{x\to-\infty} \psi_s(x\,|\,p) = \exp(ipx) + b(p)\exp(-ipx),$$

$$\lim_{x\to+\infty} \psi_s(x\,|\,p) = t(p)\exp(ipx). \tag{1}$$

The quantities $b(p)$ and $t(p)$ are called the reflection and transmission coefficients respectively. Their properties are discussed in Ref. 1, for example. The inverse problem which we are about to present is discussed in Refs. 1–3 and a very extensive bibliography of inverse scattering theory is given in Ref. 4.

It is readily shown that the eigenfunctions $\psi_s(x\,|\,p)$ satisfy the integral equation

$$\psi_s(x\,|\,p) = \exp(ipx) - \frac{i}{2p}\int_{-\infty}^{+\infty} \exp(ip\,|\,x - x'\,|) \times V(x')\psi_s(x'\,|\,p)\,dx'. \tag{2}$$

From this integral equation one can obtain $b(p)$ and $t(p)$ in terms of ψ_s.

Another set of eigenfunctions $\psi_a(x\,|\,p)$ of the continuous spectrum is defined by the boundary conditions

$$\lim_{x\to-\infty} \psi_a(x\,|\,p) = \exp(ipx). \tag{3}$$

These eigenfunctions have simpler analytic properties as functions of p than $\psi_s(x\,|\,p)$. The two sets of eigenfunctions are not independent but are related. Clearly

$$\psi_s(x\,|\,p) = \psi_a(x\,|\,p) + b(p)\psi_a(x\,|\,-p). \tag{4}$$

The eigenfunctions ψ_a satisfy the following integral equation:

$$\psi_a(x\,|\,p) = \exp(ipx) + \frac{1}{p}\int_{-\infty}^{x} \sin p(x - x')V(x')\psi_a(x'\,|\,p)\,dx'. \tag{5}$$

As mentioned above, there may also be point eigenvalues $E_i < 0$. The corresponding eigenfunctions $\psi_i(x)$ may always be chosen real and their normalization is arbitrary:

$$\int_{-\infty}^{+\infty} [\psi_i(x)]^2\,dx = A_i > 0.$$

The functions $\psi_a(x\,|\,p)$, $\psi_i(x)$ satisfy the following completeness relationship, from which an expansion theorem can be derived:

$$(2\pi)^{-1}\int_{-\infty}^{+\infty} \psi_a(x\,|\,p)\psi_a(x'\,|\,-p)\,dp + (2\pi)^{-1}\int_{-\infty}^{+\infty} \psi_a(x\,|\,p) \times b(-p)\psi_a(x'\,|\,p)\,dp + \sum_i \frac{\psi_i(x)\psi_i(x')}{A_i} = \delta(x - x'). \tag{7}$$

We shall now give the Gel'fand—Levitan algorithm: Define $B(x)$ by

$$B(x) = (2\pi)^{-1}\int_{-\infty}^{+\infty} b(p)\exp(-ipx)\,dp + \sum_i \frac{1}{A_i}\exp[(-E_i)^{1/2}x]. \tag{8}$$

Then let the Gel'fand—Levitan kernel $K(x, y)$ defined for $x \geq y$ satisfy the integral equation

$$K(x, y) = -B(x + y) - \int_{-\infty}^{x} K(x, z)B(z + y)\,dz. \tag{9}$$

It follows that

$$V(x) = 2\frac{d}{dx}K(x, x), \tag{10}$$

$$\psi_a(x\,|\,p) = \exp(ipx) + \int_{-\infty}^{x} K(x, y)\exp(ipy)\,dy, \tag{11}$$

$$\psi_i(x) = \exp[(-E_i)^{1/2}x] + \int_{-\infty}^{x} K(x, y)\exp[(-E_i)^{1/2}y]\,dy. \tag{12}$$

2. THE INTEGRAL EQUATION FOR THE KERNEL IN TERMS OF THE POTENTIAL

One can also obtain an integral equation for the kernel $K(x, y)$ in terms of the potential $V(x)$. The equation is the principal result of the present paper. We have

 Copyright © 1975 American Institute of Physics

$$K(x,y)=\tfrac{1}{2}\int_{-\infty}^{(x+y)/2}V(x')\,dx'+\tfrac{1}{2}\int_{-\infty}^{(x+y)/2}V(x')\times\int_{y-x+x'}^{x'}K(x',z)\,dz\,dx'+\tfrac{1}{2}\int_{(x+y)/2}^{x}V(x')\times\int_{y-x+x'}^{y+x-x'}K(x',z)\,dz\,dx'. \tag{13}$$

We shall now derive Eq. (13). It is convenient to introduce the Heaviside function $\eta(x)$ defined by

$$\eta(x)=1 \text{ for } x>0,\quad =0 \text{ for } x<0. \tag{14}$$

We use Eq. (11) on the right-hand side of Eq. (5) and then equate the resulting expression for ψ_a with the expression for ψ_a given by Eq. (11). We find

$$\int_{-\infty}^{+\infty}K(x,y)\eta(x-y)\exp(ipy)\,dy=\frac{1}{p}\int_{-\infty}^{+\infty}\sin p(x-y)\eta(x-y)\times V(y)\exp(ipy)\,dy+\frac{1}{p}\int_{-\infty}^{+\infty}\eta(x-x')\sin p(x-x')V(x')\times\int_{-\infty}^{+\infty}\eta(x'-y)K(x',y)\exp(ipy)\,dy\,dx'. \tag{15}$$

We now multiply Eq. (15) by $(2\pi)^{-1}\exp(-ipz)$ and integrate with respect to p. We use the fact that

$$(2\pi)^{-1}\int_{-\infty}^{+\infty}\sin kx\exp(iky)k^{-1}\,dk=\tfrac{1}{2}[\eta(y+x)-\eta(y-x)]. \tag{16}$$

Then

$$K(x,z)\eta(x-z)\equiv K(x,z)=\tfrac{1}{2}\int_{-\infty}^{+\infty}\eta(x-y)[\eta(x-z)-\eta(2y-z-x)]\times V(y)\,dy+\tfrac{1}{2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\eta(x-x')\eta(x'-y)\times[\eta(y-z+x-x')-\eta(y-z-x+x')]\times V(x')K(x',y)\,dx'\,dy. \tag{17}$$

But as we shall show

$$\eta(x-y)[\eta(x-z)-\eta(2y-z-x)]=\eta(x-z)\eta((x+z)/2-y), \tag{18}$$

$$\eta(x-x')\eta(x'-y)[\eta(y-z+x-x')-\eta(y-z-x+x')]=\eta(x-z)\eta((x+z)/2-x')[\eta(y-z+x-x')-\eta(y-x')]+\eta(x-z)[\eta(x'-(x+z)/2)-\eta(x'-x)]\times[\eta(y-z+x-x')-\eta(y-z-x+x')]. \tag{19}$$

Substitution of Eqs. (18) and (19) into Eq. (17) leads to Eq. (13).

The proofs of Eq. (18) and (19) are not altogether trivial. Hence we shall sketch them.

To prove Eq. (18), we shall first prove that $\eta(x-y)\times\eta(2y-z-x)=0$ if $x<z$. If $x<y$, this result follows immediately. Hence take $x>y$. Then, if $x<z$, we have $y<x<z$ and so $2y<x+z$. Thus the factor $\eta(2y-z-x)=0$ and we may write

$$\eta(x-y)\eta(2y-z-x)\equiv\eta(x-z)\eta(x-y)\eta(2y-z-x). \tag{20}$$

The left-hand side of Eq. (18) is then

$$\eta(x-z)\eta(x-y)[1-\eta(2y-z-x)]=\eta(x-z)\eta(x-y)\eta(x+z-2y), \tag{21}$$

where we have used the relation $1-\eta(x)=\eta(-x)$. On also using the identity $\eta(2x)=\eta(x)$, the last two factors on the right-hand side of Eq. (21) become $\eta(x-y)\times\eta((x+z)/2-y)$. For a nonzero result $y<(x+z)/2<x$. Hence

$$\eta(x-y)\eta((x+z)/2-y)=\eta((x+z)/2-y), \tag{22}$$

and Eq. (18) results.

To prove Eq. (19), we denote the left-hand side by A and define y_1 and y_2 by

$$y_1=z-x+x',\quad y_2=z+x-x'. \tag{23}$$

Then

$$A=\eta(x-x')\eta(x'-y)[\eta(y-y_1)-\eta(y-y_2)]. \tag{24}$$

First let us prove that $A=0$ if $z>x$. We may take $x>x'$, for the result is trivial otherwise. Then

$$y_1<y_2. \tag{25}$$

Furthermore,

$$x'<y_1. \tag{26}$$

By considering A as a function of y, it is clear that $A=0$, since one of the two y-dependent factors is always zero. Hence we may insert the factor $\eta(x-z)$ on the right-hand side of Eq. (24) without a change of meaning.

Now let us take $x>z$ and $x>x'$ in evaluating the expression

$$B=\eta(x'-y)[\eta(y-y_1)-\eta(y-y_2)]. \tag{27}$$

[We note

$$A=\eta(x-z)\eta(x-x')B.] \tag{28}$$

Equation (25) holds. Moreover, $x'>y_1$.

We can now consider two cases: (a) in which $x'<y_2$ which is equivalent to $x'<(x+z)/2$ and (b) in which $x'>y_2$ which is equivalent to $x'>(x+z)/2$.

Let us first consider case (a). Considering B as a function of y, we have

$$B=\eta(y-y_1)-\eta(y-x'). \tag{29}$$

In case (b)

$$B=\eta(y-y_1)-\eta(y-y_2). \tag{30}$$

Thus combining the two cases and using Eq. (28), we have

$$A=\eta(x-z)\eta(x-x')\eta((x+z)/2-x')[\eta(y-y_1)-\eta(y-x')]+\eta(x-z)\eta(x-x')\eta(x'-(x+z)/2)[\eta(y-y_1)-\eta(y-y_2)]. \tag{31}$$

But on using $x>z$ with the consequence $x>(x+z)/2$,

$$\eta(x-x')\eta((x+z)/2-x')=\eta((x+z)/2-x'),$$

$$\eta(x-x')\eta(x'-(x+z)/2)=\eta(x'-(x+z)/2)-\eta(x-x'), \tag{32}$$

considering the left-hand sides as functions of x'. Then the substitution of Eq. (23) and (32) into Eq. (31) yields Eq. (19).

We shall now discuss the integral equation (13) for the Gel'fand—Levitan kernel in terms of the scattering potential briefly. First of all, we note that this integral equation is one for the *two* variables in $K(x, y)$ in contrast to the Gel'fand—Levitan equation (9) which involves the variable y only and in which x is fixed. Secondly, it is readily seen that the kernel satisfies the hyperbolic equation

$$K_{xx} - K_{yy} = V(x)K \tag{33}$$

with the boundary condition

$$K(x, x) = \tfrac{1}{2}\int_{-\infty}^{x} V(x')\,dx'. \tag{34}$$

Indeed, Eqs. (33) and (34) are used in deriving the Gel'fand—Levitan equation (see Refs. 1 and 5 for the radial equation). The integral equation (13) can be considered an alternate approach and can in fact be derived from Eq. (33) with boundary conditions (34) through the use of the theory of characteristics.

It should also be mentioned that an analog of Eq. (13) arising from the three-dimensional problem of inverse scattering at fixed energy but various angles is given in Ref. 6.

Note added in proof: The integral equation (13) can also be obtained from the integral equation for the kernel $B_2(x, y)$ of L. D. Fadeev [AMS Transl. Ser. 2 **65**, 142 (1967)] by a suitable change of variables. The author is grateful to Professor R. G. Newton for pointing out this reference to him.

ACKNOWLEDGMENT

The author expresses his thanks to Professor R. G. Newton for bringing to his attention Ref. 6 and for conversations which led to the author's finding an error in the integral equation (13).

[1] I. Kay, "The Inverse Scattering Problem," Research Report EM-74, Institute of Mathematical Sciences of New York University, Division of Electromagnetic Research (1955).
[2] I. Kay and H. E. Moses, Nuovo Cimento **X**, 276 (1956).
[3] H. E. Moses and C. M. deRidder, "Properties of Dielectrics from Reflection Coefficients in One Dimension" Lincoln Laboratory Report No. 322 Astia No. AD-422-444 (1963).
[4] L. Colin, Ed., "Mathematics of Profile Inversion. Proceedings of a Workshop held at Ames Research Center, July 12-16, 1971," NASA Technical Memorandum TMX-62-150, August 1972.
[5] N. Levinson, Phys. Rev. **89**, 755 (1953).
[6] P. J. Kelemen, J. Math. Phys. **12**, 1142 (1971).

Variable dimensionality in the group-theoretic prediction of configuration mixings for doubly-excited helium

David R. Herrick

Bell Laboratories, Murray Hill, New Jersey 07974
(Received 4 November 1974)

Variable dimensionality ($= D$) is used to interpret recent group-theoretic predictions of configuration mixings in doubly-excited helium. Calculated $2s^2{:}2p^2\ {}^1S$ mixings agree with the group theory over the range $1 \leq D \leq \infty$. General results for $D = 2$ predicted mixings are given and energies of states within the $N = 2$ atomic shell confirm predicted level orderings. The $D = 1$ model atom is described exactly by the group theory, with quantum numbers and a selection rule characterizing the stability of Coulomb matrix elements as $D \to 1$. The exact $D = 1$ results have a physical interpretation in approximate autoionization and energy selection rules for Rydberg series at $D = 3$.

I. INTRODUCTION

Recent investigations showed that mixing ratios of doubly-excited, degenerate, two-electron hydrogenic configurations (for example, s, $2s^2{:}2p^2\,{}^1S$) due to the Coulomb interaction $1/R_{12}$ are predicted quite well by a new group-theoretical technique.[1] Diagonalization of an operator $\hat{B}^2$ equal to the square of the distance between two one-electron Runge—Lenz vectors in the hydrogenic configuration basis yielded a "doubly-excited symmetry basis" (DESB) with configuration mixing ratios extremely close to the directly calculated ratios. Qualitatively, the ability of the DESB states to describe correctly the near-degeneracy two-electron correlation is related to the fact that average values of the Runge—Lenz vector $\mathbf{A}$ are proportional to average values of the position vector $\mathbf{R}$ in the one-electron hydrogenic basis. It is expected then that two-electron states in which $\hat{B}^2 = (\mathbf{A}_1 - \mathbf{A}_2)^2$ is largest will have the largest average values of $(\mathbf{R}_1 - \mathbf{R}_2)^2$, resulting in the smallest average values of $1/R_{12}$. This assertion was extremely well verified, and simple empirical formulas relating $\hat{B}^2$ and the total angular momentum to exact doubly excited energies were obtained. A natural consequence of the diagonalization of $\hat{B}^2$ was the appearance of two new quantum numbers K and T which labeled each configuration-mixed state. Extensive configuration interaction calculations of the doubly-excited helium spectrum demonstrated conclusively that K and T labeled uniquely each Rydberg series in helium. In addition an analysis of the relationship of K and T to the angular and radial correlation properties of each zeroth order DESB state gave approximate "selection rules" for predicting correctly the relative weakness of autoionization and UV absorption processes for the actual helium states.

Although the DESB classification of the near-degeneracy correlation is good the fact remains that the method does not exactly diagonalize the energy, thus complicating attempts to determine both the precise role of $\hat{B}^2$ in the group-theoretical description of $1/R_{12}$ and a group-theoretical derivation of the autoionization selection rules. Some insight to $\hat{B}^2$ can be gained by investigation of configuration mixings of the asymptotic region of the wavefunction, for which the "dipole approximation"[2,3] is quite good. As we have noted elsewhere,[4] a more accurate description of the long-range mixings was given by Macek[5] in a diagonalization of asymptotically degenerate configuration interactions to order r^{-2} in the hyperspherical coordinate $r = (r_1^2 + r_2^2)^{1/2}$. One readily sees that the major portion of the mixings is due to $\hat{B}^2$.[6] As we emphasized in Ref. 1, there is sufficient empirical evidence to suggest that $\hat{B}^2$ is fundamentally more important than these long-range results indicate.

In order to elucidate the role of $\hat{B}^2$ in determining configuration mixings, we investigate in the present paper the model problem of a helium-like atom of variable dimensionality ($\equiv D$), with physical results corresponding to the case $D = 3$. Calculations of the actual $2s^2{:}\,2p^2\,{}^1S$ mixing are made, and verify the group-theoretic predicted mixings over the range $1 \leq D \leq \infty$. Details of the group theory and the calculation are described in Sec. II. Further investigation of the operator $\hat{B}^2$ for the case $D = 2$ appears in Sec. III. While our use of variable D in this problem represents in itself an interesting mathematical problem, it is our goal to use the results for general D to better interpret the $D = 3$ results. To this end we show in Sec. IV that in the limit $D = 1$ the DESB mixings predicted by $\hat{B}^2$ are in *exact* agreement with those given by calculation. This correspondence leads directly to a new interpretation of the $D = 3$ quantum numbers and autoionization "selection rules."

Variable dimensionality has been used previously by Herrick and Stillinger[7] to describe the ground state binding energy of the helium isoelectronic sequence, and resulted in the discovery that the doubly-excited $2p^2\ {}^3P^e$ ($D = 3$) state is identical (within a simple angular factor) to the $1s^2\,{}^1S^e$ ($D = 5$) ground state. Similar degeneracies exist for excited states of the same symmetries, and in Sec. V we show that the group-theoretical DESB does in fact predict identical configuration mixings for these $D = 3$ and $D = 5$ states. Interpretation and discussion of the variable D classification of states appears in Sec. VI, and we show that the $D = 1$ results are particularly useful for interpreting the $D = 3$ classification.

II. THE $2s^2{:}2p^2\ {}^1S$ CONFIGURATION INTERACTION IN D DIMENSIONS

A. Group theoretic prediction of mixings

The degeneracy group of the bound hydrogen atom

 Copyright © 1975 American Institute of Physics

TABLE I. Calculated and predicted $2s^2, 2p^2\ ^1S$ mixing coefficients for variable dimensionality ($=D$). Calculated values are for the lowest energy state from a 2×2 diagonalization of the $1/R_{12}$ interaction in a hydrogenic basis. Predicted values are from Eq. 6.

	Calculated coefficients		Predicted coefficients	
D	$(2s^2,S)$	$(2p^2,S)$	$(2s^2,S)$	$(2p^2,S)$
1	0.7071	0.7071	0.7071	0.7071
2	0.8250	0.5651	0.8165	0.5774
3	0.8796	0.4756	0.8660	0.5000
4	0.9094	0.4160	0.8944	0.4472
5	0.9277	0.3734	0.9129	0.4082
8	0.9552	0.2958	0.9428	0.3333
10	0.9644	0.2645	0.9535	0.3162
100	0.9965	0.0834	0.9950	0.0995
∞	1.0000	0.0000	1.0000	0.0000

states in D dimensions is $SO(D+1)$,[8-10] in analogy to the well-known results of Fock[11] for $D=3$. The group algebra is generated by the $D(D-1)/2$ angular momentum operators

$$\hat{L}_{jk}=-i\left(X_j\frac{\partial}{\partial X_k}-X_k\frac{\partial}{\partial X_j}\right),\quad (j<k\leq D) \tag{1}$$

and the generalized Runge–Lenz vector operator[9,10]

$$\hat{A}_k=-\left\{\left[\left(\mathbf{R}\cdot\frac{\partial}{\partial\mathbf{R}}\right)+\left(\frac{D-1}{2}\right)\right]\frac{\partial}{\partial X_k}-X_k[2H+R^{-1}]\right\}[-2H]^{-1/2}\quad (k\leq D). \tag{2}$$

Here $\mathbf{R}$ is the position vector $(X_1, X_2, \cdots, X_D)$ and H is the one-electron energy $H=-\Delta/2-1/R$. Both $\hat{L}_{jk}$ and $\hat{A}_k$ commute with H. The one-electron energy spectrum is $E_n=-2(2n+D-3)^{-2}$, with the principal quantum number $n=1, 2, \cdots$. The bound state wavefunctions are $\Psi^{(D)}=R^D_{n\Lambda}\cdot Y^D_\Lambda$, where Y^D_Λ is a D-dimensional spherical harmonic of total angular momentum $\Lambda(\Lambda+D-2)$, $\Lambda=0, 1, \cdots, n-1$. $R^D_{n\Lambda}$ is the radial function[12]

$$R^D_{n\Lambda}=N^D_{n\Lambda}\exp(-kR)R^\Lambda F(-n+\Lambda+1;\ 2\Lambda+D-1;\ 2kR), \tag{3a}$$

$$N^D_{n\Lambda}=\frac{(-1)^\Lambda}{2\Gamma(2\Lambda+D-1)}\left[\frac{\Gamma(n+\Lambda+D-2)(2k)^{2\Lambda+D+1}}{\Gamma(n-\Lambda)}\right]^{1/2}, \tag{3b}$$

with $k=\sqrt{-2E_n}$. We label states with $\Lambda=0, 1, 2, \cdots$, respectively, ns, np, $nd, \cdots$. $\Gamma(X)$ is Euler's gamma function.

For $n=2$ there are $D+1$ degenerate states. There is one spherical ($\Lambda=0$) $|2s)$ state, and a D-component $|2\mathbf{p})$ vector state ($\Lambda=1$). For all dimensionalities we find that

$$\hat{A}_k|2s)=|2p_k), \tag{4a}$$

$$A_k|2p_j)=|2s)\delta_{jk}, \tag{4b}$$

where δ_{jk} is the Kronecker delta function. The corresponding two-electron configurations with zero total angular momentum are

$$|2s^2,S)\equiv|2s)_1|2s)_2, \tag{5a}$$

$$|2p^2,S)\equiv-[(2\mathbf{p})_1\cdot(2\mathbf{p})_2]/\sqrt{D}. \tag{5b}$$

Each configuration is normalized to unity, and subscripts label electron 1 and 2. The phase in (5b) is chosen to be consistent with the Condon–Shortley[13] convention for $D=3$.

As in Ref. 1, we can define generators of an $SO(D+1)_1\times SO(D+1)_2$ algebra with the two-electron angular momentum and the antisymmetric operator $\hat{\mathbf{B}}=\hat{\mathbf{A}}_2-\hat{\mathbf{A}}_1$. We expect that diagonalization of $\hat{B}^2$ in the $|2s^2)$, $|2p^2)$ basis will yield correlated states which approximate the exact mixings caused by $1/R_{12}$. The resulting DESB states are

$$\chi_+=\left[\frac{D}{D+1}\right]^{1/2}|2s^2,S)+\left[\frac{1}{D+1}\right]^{1/2}|2p^2,S), \tag{6}$$

for which $[B^2-2(D+1)]\chi_+=0$ and

$$\chi_-=-\left[\frac{1}{D+1}\right]^{1/2}|2s^2,S)+\left[\frac{D}{D+1}\right]^{1/2}|2p^2,S) \tag{7}$$

for which $\hat{B}^2\chi_-=0$. χ_+ is expected to give a lower $1/R_{12}$ repulsion energy since it is characterized by the larger $\hat{B}^2$ eigenvalue.

Note that in the infinite-D limit χ_+ and χ_- approach pure $|2s^2,S)$ and $|2p^2,S)$ states, respectively, while at $D=1$ the mixing is predicted to be strongest.

B. Calculated $2s^2:2p^2$ mixings

Each configuration has zero total angular momentum, and transformation to the coordinates R_1, R_2, R_{12} allows all Coulomb integrals to be evaluated as a continuous function of D.[7] Defining $E^D_{ss}=(2s^2|R_{12}^{-1}|2s^2)$, $E^D_{pp}=(2p^2|R_{12}^{-1}|2p^2)$, and $E^D_{sp}=(2s^2|R_{12}^{-1}|2p^2)$, we obtain

$$E^D_{ss}=\alpha[D(16D^3+56D^2+17D+91)N^{-1}], \tag{8}$$

$$E^D_{pp}=\alpha[(2D+1)(2D+3)(4D^2+15D-7)N^{-1}], \tag{9}$$

$$E^D_{sp}=-\alpha[15(D+3)(2D+1)D^{1/2}N^{-1}], \tag{10}$$

with[14]

$$N=16D(D+1)(D+2)(D+4), \tag{11}$$

$$\alpha=\frac{2\Gamma[(D-1)/2]\Gamma(D+\frac{1}{2})}{\Gamma(D+2)(\Gamma(D/2)}. \tag{12}$$

Diagonalization of the $2s^2$, $2p^2\ ^1S$ interaction matrix yields two configuration-mixed states which can be compared with χ_+ and χ_-. The mixings are determined uniquely by the parameter

$$\eta=(E_{ss}-E_{pp})/2E_{sp} \tag{13a}$$

$$=\frac{(D-1)(12D^2+41D+7)}{10(D+3)(2D+1)\sqrt{D}}, \tag{13b}$$

with the ratio of the $2s^2{:}2p^2$ mixing ($\equiv C_{sp}$) in the lower energy state given by

$$C^{\text{calc}}_{sp}=\eta+\sqrt{1+\eta^2}. \tag{14}$$

Values of this calculated ratio appear in Table I for comparison with the predicted value $C^{\text{pred}}_{sp}=\sqrt{D}$ obtained from χ_+ in Eq. (6). Excellent agreement between the predicted and calculated ratios is found, and at $D=1$ they are identical. For increasing D the predicted ratio is smaller than the calculated ratio and at $D=\infty$, $C^{\text{calc}}_{sp}/C^{\text{pred}}_{sp}=6/5$. However, the actual state at $D=\infty$ is pure $|2s^2,S)$ in agreement with the group theory. It is clear then (at least to first order in a Z^{-1} expansion of the doubly-excited energy) that the contribution of the near-degeneracy $2s^2{:}2p^2$ correlation to the total energy diminishes with increasing D.

TABLE II. Group theoretical DESB states and Coulomb repulsion energies for hydrogenic configurations with both orbitals in the $N=2$ shell.

(K,M,spin)	$\hat{B}^2$	$\langle 1/R_{12}\rangle$[a]	Label	Wavefunction
(1,0,0)	6	197	^1S	$\sqrt{\frac{2}{3}}(2s2s)+\sqrt{\frac{1}{3}}(2p2p,S)$[b]
(1,1,1)	5	210	^3P	$-(2s2p_{+1}-2p_{+1}2s)/\sqrt{2}$
(0,0,1)	2	280	^3S	$(2p_{-1}2p_{+1}-2p_{+1}2p_{-1})/\sqrt{2}$
(1,2,0)	2	367.5	^1D	$-(2p_{+1}2p_{+1})$
(0,1,0)	1	460	^1P	$(2s2p_{+1}+2p_{+1}2s)/\sqrt{2}$
(−1,0,0)	0	576	^1S	$-\sqrt{\frac{1}{3}}(2s2s)+\sqrt{\frac{2}{3}}(2p2p,S)$[b]

[a] In units $\pi/3072$; $\langle(1,0,0)|1/R_{12}|(-1,0,0)\rangle=4\sqrt{2}$
[b] Defined in Eq. (24).

III. DESB MIXINGS FOR $D=2$

The results of Sec. II show that the DESB approximation to the actual $2s^2{:}2p^2\,{}^1S$ mixing is better in two dimensions than in three dimensions. This suggests that group-theoretical investigations of the Helium doubly-excited states can be made without a loss of generality in $D=2$ rather than $D=3$. The results obtained for this model problem should allow similar results to be formulated in $D=3$. There are two features of the $D=2$ problem which make it particularly appealing. The first is that the angular portion of the wavefunction is simpler than in $D=3$. This places a greater emphasis in the calculation of Coulomb matrix elements upon the radial portion of the wavefunctions, and facilitates their evaluation. A comparison of the matrix elements for two and three dimensions appears in the Appendix. Secondly, the underlying group theory of the DESB is simpler in $D=2$ than for $D=3$, although the fundamental role of $\hat{B}^2$ in both dimensions is unchanged. For these reasons we describe in detail the DESB states for $D=2$.

A. One-electron states

Defining $\rho=2kR$, we use the radial states in Eq. (3) to construct the $D=2$ hydrogenic wavefunctions

$$\Psi_{nm}(R)=M_{nm}\rho^{|m|}\exp(-\rho/2)F(|m|-j;\,2|m|+1;\rho)\frac{\exp(im\phi)}{\sqrt{2\pi}}, \tag{15a}$$

$$M_{nm}=\frac{(-1)^m 4}{(2|m|)!}\left(\frac{(j+|m|)!}{(j-|m|)!\,(2j+1)^3}\right)^{1/2}, \tag{15b}$$

where $j\equiv n-1$ and $m=0,\pm1,\ldots,\pm j$. The angular momentum operator $\hat{L}=-i\partial/\partial\phi$ and the Runge-Lenz vector in Eq. (2) are used to define

$$J_0=-i\frac{\partial}{\partial\phi}, \tag{16}$$

$$J_\pm=\hat{A}_x\pm i\hat{A}_y, \tag{17}$$

$$=\exp(\pm i\phi)\left(\left(\tfrac{1}{2}\pm J_0\right)\frac{\partial}{\partial R}-\left(J_0^2\pm\tfrac{1}{2}J_0\right)\frac{1}{R}+1\right)[-2H]^{-1/2} \tag{18}$$

$$=\exp(\pm i\phi)\left((1\pm 2J_0)\frac{\partial}{\partial\rho}-(2J_0^2\pm J_0)\frac{1}{\rho}+(j+\tfrac{1}{2})\right). \tag{19}$$

J_0, $J_\pm$ are generators of the hydrogenic $SO(3)$ degeneracy algebra, with commutation relations

$$[J_0,J_\pm]=\pm J_\pm, \tag{20a}$$

$$[J_+,J_-]=2J_0 \tag{20b}$$

and

$$J_0\Psi_{nm}=m\Psi_{nm}, \tag{21a}$$

$$J_\pm\Psi_{nm}=[(j\mp m)(j\pm m+1)]^{1/2}\Psi_{nm\pm1}. \tag{21b}$$

$J^2=J_0(J_0-1)+J_+J_-$ is diagonal in the Ψ_{nm} basis with eigenvalue $j(j+1)$.

We use the convention that states with $|m|=0,1,2,\cdots$ are labeled s, p_m, $d_m,\cdots$. For instance, when $n=3$ there are a total of five degenerate states corresponding to the $j=2$ irreducible representation of $SO(3)$: $3s$, $3p_{\pm1}$, and $3d_{\pm2}$.

B. Two-electron DESB states

The DESB states obtained by diagonalizing $\hat{B}^2=(\hat{\mathbf{A}}_1-\hat{\mathbf{A}}_2)^2$ subject to conservation of total angular momentum ($\equiv M$) and parity are

$$\phi(Nn,KM)=\sum_m(-1)^{m+n}\sqrt{2J+1}\times\begin{pmatrix}N-1 & n-1 & J\\ m & M-m & -M\end{pmatrix}\Psi_{Nm}(1)\Psi_{nM-m}(2), \tag{22}$$

where $(:::)$ is a $3-j$ symbol. For convenience we assume $n\geq N$. J [the usual $SO(3)$ quantum number] is restricted to integer values $|j_1-j_2|\leq J\leq j_1+j_2$, but the more important DESB quantum number is $K\equiv J-n+1$, in analogy to our definition of K for the $D=3$ states.[1] Note that the total parity for the $D=2$ states is $(-1)^M$, and need not be specified.[15] The phase factor $(-1)^m$ arises because $\hat{B}^2$ depends upon the difference of the operators $\mathbf{A}_1$ and $\mathbf{A}_2$ rather than their sum. The states $\phi(Nn,KM)$ are bases for irreducible representations of the $SO(3)$ algebra generated by $J_0=J_0(1)+J_0(2)$ and $J_\pm=J_\pm(2)-J_\pm(1)$. The eigenvalues of $\hat{B}^2$ are $J(J+1)-M^2$. When $n=N$ the states have exchange symmetry $(-1)^{J+M}$, but when $n\neq N$ exchange-symmetrized states must be constructed. Multiplication by appropriate spin functions then gives the antisymmetrized states required by the Pauli principle.

The K quantum number is more useful than J since it labels an entire Rydberg series of states $n=N$, $N+1,\cdots$ for each value of M. For instance the DESB S states ($M=0$) with $N=2$, $K=+1$ have the mixing

$$\varphi(2n,+10)=\left(\frac{n}{2n-1}\right)^{1/2}|2sns,S)+\left(\frac{n-1}{2n-1}\right)^{1/2}|2pnp,S), \tag{23}$$

where

$$|2pnp,S)\equiv-[|2p_{-1}np_{+1})+|2p_{+1}np_{-1})]/\sqrt{2}. \tag{24}$$

C. Example: $N=2$ states

The DESB states in which both electrons are in orbitals with principal quantum number $N=2$ illustrate quite effectively the relationship of the $\hat{B}^2$ eigenvalues to the average Coulomb repulsion energies. The wavefunctions, quantum numbers, and energies appear in Table II. As expected, the states with larger values of $\hat{B}^2$ have a smaller Coulomb repulsion energy. This supports our assertion that the fundamental role of $\hat{B}^2$ may be independent of dimensionality. Each state is clearly the analogue of a $D=3$ state, as indicated by our

TABLE III. Group-theoretic DESB configuration-mixed states for $D=1$ doubly-excited states. Each wavefunction for $n>N$ must be exchange symmetrized to form a singlet ($S=0$) and a triplet ($S=1$) state. The labels ^{2S+1}L are used to indicate correspondence to states in $D=3$. The other quantum numbers are described in the text.

K	$\hat{B}^2$	$\hat{\pi}$	$\hat{C}_1\hat{C}_2$	Label	Wavefunction ($n\geq N\geq 2$)
$+(N-1)$	$(N+n-2)^2$	$+1$	-1	$^{1,3}S$	$(Ns_1ns_2 - Np_1np_2)/\sqrt{2}$
$+(N-1)$	$(N+n-2)^2$	-1	-1	$^{3,1}P$	$(Ns_1np_2 - Np_1ns_2)/\sqrt{2}$
$-(N-1)$	$(n-N)^2$	-1	$+1$	$^{1,3}P$	$(Ns_1np_2 + Np_1ns_2)/\sqrt{2}$
$-(N-1)$	$(n-N)^2$	$+1$	$+1$	$^{1,3}S$	$(Ns_1ns_2 + Np_1np_2)/\sqrt{2}$

notation. The ^3S state ($D=2$) corresponds to the $^3P^e$ state at $D=3$.

IV. CONFIGURATION MIXINGS IN ONE DIMENSION

We show in this section that Coulomb mixings of hydrogenic configurations in the limit $D=1$ are exactly those predicted group-theoretically by diagonalization of $\hat{B}^2$. We noted in Sec. II, for example, that the predicted DESB mixing of $2s^2:2p^2\,^1S$ converges to the calculated mixing ratio at $D=1$, even though the Coulomb matrix elements diverge in that limit. In general, problems of convergence can be avoided by assuming that the limit $D=1$ is not taken until after the radial integrals and matrix diagonalization are performed. The same convention is used in the determination of commutation relations (for instance, of the Runge–Lenz operator and the one-electron Hamiltonian) on the hydrogenic basis. A second problem is the conceptual difficulty of interpreting the physical nature of the $D=1$ hydrogenic states, since there is no "$SO(1)$" angular momentum. The one-electron states are described in part A, and are classified by an $SO(2)$ hydrogenic degeneracy group. The two-electron classification appears in part B.

A. One-electron states

While each radial function $R^D_{n\Lambda}$ in Eq. (3) has a definite limit at $D=1$, only states with zero total angular momentum [$=\Lambda(\Lambda+D-2)$] have a physical interpretation. Clearly, these are the states with $\Lambda=0$ and $\Lambda=1$. The corresponding "angular" functions in each case are $s\equiv 1/\sqrt{2}$ for $\Lambda=0$, and $p\equiv \mathrm{sgn}(X)/\sqrt{2}$ for $\Lambda=1$. Note that s and p have even and odd parity, respectively. The full hydrogenic wavefunctions are then

$$ns = R^D_{n0}\cdot s, \tag{25a}$$

$$np = R^D_{n1}\cdot p \tag{25b}$$

in the limit $D=1$. The $1s$ ground state lies at $E=-\infty$, and has been shown[7] to represent the single bound state of an attractive Dirac delta-function potential when length and energy are suitably scaled. All excited bound states ($n\geq 2$) have energy $E_n=-\frac{1}{2}(n-1)^{-2}$.

The most important property of the hydrogenic states is that *the normalized ns and np radial functions are identical at* $D=1$, corresponding to a collapse into the origin of the first radial node in each ns wavefunction as $D\to 1$. This is easily shown,

$$\lim_{D\to 1} R^D_{n0} = [n-1]^{-3/2}\lim_{D\to 1}(D-1)F(1-n;\,D-1;2kR)\exp(-kR) = R^1_{n1}, \tag{26}$$

using the hypergeometric identity

$$\lim_{b\to 0} bF(a;b;z) = azF(a+1;2;z). \tag{27}$$

More fundamentally, the equality of the ns and np radial functions is a result of the Fock-type $SO(2)$ degeneracy of the state, with the Runge–Lenz operator in the $D=1$ limit being simply

$$\hat{A} = (n-1)\,\mathrm{sgn}(X) \tag{28}$$

when acting on ns and np. The states

$$U_{\pm m} = (ns \pm np)/\sqrt{2} \tag{29}$$

with $m\equiv n-1$ satisfy

$$\hat{A}U_{\pm m} = \pm m U_{\pm m}, \tag{30}$$

and define the irreducible representations of the $SO(2)$ degeneracy algebra.[16,17]

B. Two-electron states

The DESB for $D=1$ is obtained by diagonalizing $\hat{B}^2 = (\hat{A}_2-\hat{A}_1)^2$ on the $(Ns_1,Np_1)\times(ns_2,np_2)$ product space subject to conservation of parity. These wavefunctions and the $\hat{B}^2$ eigenvalues for the entire doubly excited spectrum $n\geq N$ appear in Table III. In each case $\hat{B}^2$ is an integer ($=M^2$), and the DESB quantum number $K\equiv |M|-n+1$ labels each Rydberg series $n=N, N+1, \cdots$. Because the ns and np radial functions are equal the DESB wavefunctions factor into a product of a two-electron radial function times one of the four "angular" functions $(s_1s_2\pm p_1p_2)$, $(s_1p_2\pm p_1s_2)$. Subsequent symmetrization of the functions in Table III with respect to electron exchange therefore affects only the radial portion of the wavefunctions.

The one-electron Runge–Lenz operator $\hat{A}=\mathrm{sgn}(X)/\sqrt{-2E}$ is energy-normalized, and therefore commutes with R_{12}^{-1} only in subspaces of constant n_1 and n_2. Thus

$$\hat{C}\equiv\hat{A}\sqrt{-2E} \tag{31}$$

is clearly a more useful one-electron operator because $\hat{C}_1\hat{C}_2$ commutes with both $1/R_{12}$ and the partity operator for all $n_1,n_2\geq 2$. On the DESB $\hat{C}_1\hat{C}_2$ affects only the "angular" portion of each wavefunction, and the corresponding eigenvalues (given in Table III) are

$$\hat{C}_1\hat{C}_2 = -\,\mathrm{sgn}(K). \tag{32}$$

We therefore have the result that matrix elements of $1/R_{12}$ between DESB states having different parity ($\hat{\pi}$), exchange symmetry ($\hat{P}_{12}$), or $\hat{C}_1\hat{C}_2$ tend to vanish as $D\to 1$, *even when the states represent nondegerate hydrogenic configurations*. Note that $\hat{C}_1\hat{C}_2$ is a function only of the $SO(2)$ algebra generators, and does not depend upon the corresponding noncompact $SO(2,1)$ description.[17]

For the $2s^2:2p^2\,^1S$ mixing described in Sec. III the off-diagonal Coulomb matrix element in the hydrogenic DESB is

$$\langle\chi_+|R_{12}^{-1}|\chi_-\rangle = \frac{6(D-1)^2(D+4)D^{1/2}\alpha}{(D+1)N}, \tag{33}$$

where N and α were defined in Eqs. (11), (12). Since α diverges as $(D-1)^{-1}$ at $D=1$, the matrix element (33) vanishes linearly in $(D-1)$.

C. Diagonal DESB Coulomb energies

A discussion of $1/R_{12}$ matrix elements for arbitrary D appears in the Appendix. Assuming as we have through throughout that the $D=1$ matrix elements can be obtained by continuous scaling of D from higher dimensionalities, these matrix elements have the form

$$\langle R_{12}^{-1}\rangle = A(D-1)^{-1} + B + C(D-1) + \cdots \tag{34}$$

in a $(D-1)$ expansion. The properties of this series, and in particular the value of A are important in understanding the relative magnitudes of the Coulomb matrix elements for DESB states at $D=2,3$. In this section we shall consider only this leading order singularity at $D=1$. Evaluation of A is facilitated by the fact that $(D-1)/R_{12}$ may be replaced in the integrations by the delta-function $2\delta(X_1 - X_2)$ in the limit $D \to 1$. In the present paper this result is obtained immediately from Eq. (A6) using the Fourier representation of the δ-function,

$$\delta(x) = (2\pi)^{-1} \int_{-\infty}^{\infty} dk \exp(ikx).$$

Within DESB subspace of constant N and n defined in Table III, the Coulomb matrix elements can be written uniquely in terms of the exchange operator $\hat{P}_{12}$, $\hat{C}_1\hat{C}_2$, and the matrix elements

$$F^0_{Nn} = \langle Ns_1 ns_2 | R_{12}^{-1} | Ns_1 ns_2\rangle (D-1), \tag{35a}$$

$$G^0_{Nn} = \langle Ns_1 ns_2 | R_{12}^{-1} | ns_1 Ns_2\rangle (D-1), \tag{35b}$$

$$F^1_{Nn} = \langle Ns_1 ns_2 | R_{12}^{-1} | Np_1 np_2\rangle (D-1), \tag{35c}$$

$$G^1_{Nn} = \langle Ns_1 ns_2 | R_{12}^{-1} | np_1 Np_2\rangle (D-1). \tag{35d}$$

The form of the diagonal DESB energies is such that the Coulomb operator may be represented by

$$(D-1)R_{12}^{-1} \to F^0_{Nn} + G^0_{Nn}\hat{P}_{12} + (F^1_{Nn} + G^1_{Nn}\hat{P}_{12})\hat{C}_1\hat{C}_2 \tag{36}$$

for $n > N$. When $n=N$ (36) must be halved because of a normalization factor in the wavefunction. Further reduction of the energy expression follows from application of the δ-function behavior of $(D-1)/R_{12}$ to (35), and we find that $F^0_{Nn} = G^0_{Nn} = F^1_{Nn} = G^1_{Nn}$ at $D=1$. Values of the coefficient A of the leading order Coulomb singularity may then be calculated using

$$(D-1)R_{12}^{-1} \to F^0_{Nn}(1+\hat{P}_{12})(1+\hat{C}_1\hat{C}_2) \tag{37}$$

within the DESB subspaces. In particular, note that there is no singularity for DESB states having either antisymmetric spatial exchange or a -1 eigenvalue (i.e., $K>0$) of $\hat{C}_1\hat{C}_2$. Only states with $K<0$ can have the $(D-1)^{-1}$ Coulomb singularity. For fixed N and n the nonzero values of A are the same for each state since the magnitude depends only upon F^0_{Nn}.

The preceding group theoretical classification of the Coulomb singularity at $D=1$ can be used to interpret states of higher dimensionality only if they connect to "physical" states at $D=1$ upon continuous scaling of D. It is possible that similar singularities exist in the energies of higher dimensional states which have no physical interpretation at lower dimensionalities.

V. $(3p^2:3d^2)^3P^e$ MIXINGS FOR $D=3$

We have shown elsewhere[7] that every exact nonrelativistic two-electron wavefunction of total orbital angular momentum $L=0$ in $D+2$ dimensions is identical (within an angular factor) to a state in D dimensions having $L=1$ when $D>1$. This equivalence was effected in the zeroth-order hydrogenic radial functions by the transformation[18]

$$(\Lambda_1, \Lambda_2, n_1, n_2)_{D+2} \to (\Lambda_1+1, \Lambda_2+1, n_1+1, n_2+1)_D \tag{38}$$

where Λ_k and n_k are the one-electron orbital angular momentum and principal quantum numbers, respectively. A consequence of the equivalence of the states is that the mixing ratio of $(3p^2:3d^2)^3P^e$ for $D=3$ is identical to the $(2s^2:2p^2)^1S$ mixing at $D=5$. It remains to be seen whether the group-theoretical DESB mixings for these two states are also equal.

The $D=3$ DESB states from Ref. 1 are

$$\phi_+ = \left(\frac{5}{6}\right)^{1/2}(3p^2, {}^3P) + \left(\frac{1}{6}\right)^{1/2}(3d^2, {}^3P) \tag{39a}$$

with $(\hat{B}^2 - 14)\phi_+ = 0$, and

$$\phi_- = -\left(\frac{1}{6}\right)^{1/2}(3p^2, {}^3P) + \left(\frac{5}{6}\right)^{1/2}(3d^2, {}^3P) \tag{39b}$$

with $(\hat{B}^2 - 2)\phi_- = 0$. Comparison with Eqs. (6) and (7) shows that the ${}^1S$ $(D=5)$ states χ_+ and χ_- are in fact transformed by Eq. (38) into the ϕ_+ and ϕ_- ${}^3P^e$ $(D=3)$ states, respectively.

A similar relationship exists between the $(2s^2:2p^2)^1S$ $(D=4)$ state and the $(3p^2:3d^2)^3S$ $(D=2)$ state. Using the $D=2$ coordinates $X = R\cos\theta$, $Y = R\sin\theta$, this latter state (for $K=+1$) is obtained from Eq. (22) as

$$\phi(33, +10) = \left[\left(\frac{4}{5}\right)^{1/2}\left(\frac{-3p_{-1}3p_{+1} + 3p_{+1}3p_{-1}}{\sqrt{2}}\right) + \left(\frac{1}{5}\right)^{1/2}\left(\frac{3d_{-2}3d_{+2} - 3d_{+2}3d_{-2}}{\sqrt{2}}\right)\right]_{D=2} \tag{40a}$$

$$= Q\left[\left(\frac{4}{5}\right)^{1/2} R_{3p}R_{3p} - \left(\frac{1}{5}\right)^{1/2} R_{3d}R_{3d}\cos\theta_{12}\right]_{D=2}, \tag{40b}$$

where $Q = (X_1Y_2 - Y_1X_2)/\pi i\sqrt{2}$ and θ_{12} is the angle between electrons 1 and 2. The second factor in Eq. (40b) is identical to the $D=4$ state χ_+ defined in Eq. (6) when $(\cos\theta_{12})_{D=2}$ is replaced by $(\cos\theta_{12})_{D=4}$. Both states have the same DESB mixing coefficients.

The agreement of the DESB mixing coefficients with the exact invariance of the mixing ratios for these pairs of states in D and $D+2$ dimensions is interesting. It is likely that the mixing invariance in the DESB results from the explicit use of hydrogenic radial functions. The results of this section are further evidence that variable dimensionality does not affect significantly the role of $\hat{B}^2$ in classifying the two-electron states.

VI. INTERPRETATION AND DISCUSSION OF RESULTS

By using variable dimensionality in the calculation and group theoretic prediction of configuration mixings we have obtained several results which elucidate the previous DESB classification of the doubly-excited helium spectrum for the physical case $D=3$. Foremost of these results is that simultaneous diagonalization of $\hat{B}^2$, parity, angular momentum, and exchange yields configuration-mixed states nearly identical to the states obtained from a direct diagonalization of $1/R_{12}$, for all

dimensionalities. Future investigations of the operator may therefore be performed in a two-dimensional vector space, for which both the group theory and integral evaluation are considerably simpler than in three dimensions.

More important, however, is our application of the DESB classification to the limiting case $D=1$. That the calculated mixing ratios are predicted exactly by the group theory for this model problem can be attributed directly to the equivalence of the *ns* and *np* hydrogenic radial functions. All the features of the $D=1$ classification can be used to interpret the $D=3$ results. For instance, Eq. (37) shows that states with DESB quantum number $K>0$ have a lower repulsion energy than states with $K<0$. This is due to the presence of a node in the "angular" portion of the wavefunction at the position $X_1 = X_2$. Thus of the two ^1P Rydberg series (see Table III) $\phi_{K=+1} = (2s_1np_2 - 2p_1ns)$, and $\phi_{K=-1} = (2s_1np_2 + 2p_1ns_2)$; the $K=+1$ states lie lower in energy. In addition, the $K=+1$ series is stable against autoionization to the $(1s\,kp)$ continuum, whereas the $K=-1$ series can decay due to its nonzero amplitude when $X_1=X_2$. These exact results for $D=1$ are identical to the qualitative interpretation given by Cooper, Fano, and Prats[19] in their initial classification of the experimental helium absorption spectrum. This example also illustrates how autoionization "selection rules" arise naturally from the DESB classification. Since $\hat{C}_1\hat{C}_2$ exactly commutes with $1/R_{12}$, interactions between discrete and continuum channels are allowed only when $\hat{C}_1\hat{C}_2$ is conserved. Although no analog of $\hat{C}_1\hat{C}_2$ has yet been found for the $D=3$ states, we have noted in Ref. (1) approximate autoionization selection rules which agree with experiment. These rules are in fact exact at $D=1$.

Conservation of $\hat{C}_1\hat{C}_2$ within a given Rydberg series at $D=1$ also has a physical interpretation at $D=3$. For instance, a configuration interaction calculation at $D=1$ including the configurations $(2sns)$, $(2pnp)^1S$ for $n=2,3,\cdots$ would give a block-diagonalized energy matrix in which all matrix elements connecting $K=+1$ DESB states with $K=-1$ DESB states are zero. The extensive configuration calculations in Ref. 1 showed that the $D=3$ energy eigenvectors are very strongly (although not exactly) classified by the K quantum number over the entire helium isoelectronic sequence. Use of K as the "natural" quantum number for the $D=3$ Rydberg series is thus well justified by its exact conservation in the $D=1$ limit.

That $\hat{C}$ rather than $\hat{A}$ is the operator of physical importance for Rydberg series at $D=1$ suggests that it may also be useful in three dimensions for describing the mixings of configurations (Nl,nl') with $n>N$. The DESB approximation to these mixings does not have the correct energy dependence at large n. Investigation of this possibility is in progress.

In conclusion, the use of a nonphysical variable dimensionality has given new insight to the classification of the doubly-excited helium spectrum and the K quantum number of the DESB, primarily because exact results are obtained at $D=1$ which have physical interpretation at $D=3$.

APPENDIX: EVALUATION OF TWO-ELECTRON COULOMB MATRIX ELEMENTS

For the wavefunctions $A_{lm}=a(R)\,Y_{lm}(\theta,\phi)$ in three dimensions, Coulomb matrix elements are usually evaluated using the Legendre expansion

$$R_{12}^{-1}=\sum_{k=0}^{\infty}P_k(\cos\theta_{12})f_k(R_<,R_>), \tag{A1}$$

$$f_k(R_<,R_>)=\frac{R_<^k}{R_>^{k+1}} \tag{A2}$$

where $R_>$ and $R_<$ are, respectively, the larger and smaller of R_1 and R_2. The resulting integrals are

$$\begin{aligned}&\langle l_am_a,l_bm_b|R_{12}^{-1}|l_cm_c,l_dm_d\rangle\\&=\sum_{k=0}^{\infty}C^k(l_am_a,l_cm_c)C^k(l_dm_d,l_bm_b)\\&\quad\times R^k(ab,cd)\delta(m_a+m_b,m_c+m_d),\end{aligned} \tag{A3}$$

where

$$C^k(lm,l'm')=[(2l+1)(2l'+1)]^{1/2}\times\begin{pmatrix}l & l' & k\\0 & 0 & 0\end{pmatrix}\begin{pmatrix}l & l' & k\\-m & m' & m-m'\end{pmatrix}, \tag{A4}$$

and

$$R^k(ab,cd)=\int_0^\infty dR_1\int_0^\infty dR_2(R_1^2a_1c_1)f_k(R_<,R_>)(R_2^2b_2d_2). \tag{A5}$$

The expansion contains a finite number of terms due to symmetry restrictions on the $3-j$ symbols. A similar expansion for $D=2$ (or for any *even* dimensionality) is not convenient, since in general an infinite number of R^k integrals must be determined. In order to evaluate the Coulomb integrals we therefore use a Fourier Convolution in D dimensions:

$$\begin{aligned}&\int d\mathbf{R}_1^{(D)}\int d\mathbf{R}_2^{(D)}\Psi(\mathbf{R}_1)R_{12}^{-1}\chi(\mathbf{R}_2)\\&=\frac{\Gamma[(D-1)/2]}{2\Gamma(\frac{1}{2})^{D+1}}\int d\mathbf{t}^{(D)}F_\Psi(-\mathbf{t})t^{1-D}F_\chi(\mathbf{t}),\end{aligned} \tag{A6}$$

with

$$F_u=\int d\mathbf{R}u(\mathbf{R})\exp(-i\mathbf{t}\cdot\mathbf{R}) \tag{A7}$$

for $u=\Psi,\chi$. This illustrates the $(D-1)^{-1}$ divergence at $D=1$ discussed in Sec. V.

For $D=3$ Eq. (A6) leads directly to Eq. (A3), except that now

$$R^k(ab,cd)=\int_0^\infty dtM_{ac}^k(t)M_{bd}^k(t), \tag{A8}$$

with

$$M_{ac}^k(t)=\left(\frac{2k+1}{t}\right)^{1/2}\int_0^\infty dRa(R)c(R)J_{k+1/2}(tR)R^{3/2}. \tag{A9}$$

$J_\nu(Z)$ is the Bessel function. Equation (A8) is particularly convenient because it reduces the two-electron radial integrals to an average (over t) of a product of one-electron expectation values of $J_{k+1/2}(tR)/\sqrt{Rt}$ in the radial basis. In this sense, the $M_{ac}^k(t)$ are similar to the $c^k(lm,l'm')$ coefficients, which are equivalent to averages of $P_k(\cos\theta)$ between the one-electron angular functions. The difference is of course that averages of products of two functions are effected in Eq. (A3) by a summation over k, while the average over products of M_{ac}^k coefficients involve integration.

For $D=2$ the one-electron wavefunctions are $A_m = a(R)\exp(im\phi)/\sqrt{2\pi}$. Equation (A6) leads quite simply to the result

$$\langle m_a m_b | R_{12}^{-1} | m_c m_d \rangle = \delta(m_a + m_b, m_c + m_d)(2k+1)^{-1} S^k(ab, cd). \quad \text{(A10)}$$

Here $k = |m_c - m_a| = |m_d - m_b|$, and S^k is the radial integral

$$S^k(ab, cd) = \int_0^\infty dt G^k_{ac}(t) G^k_{bd}(t), \quad \text{(A11)}$$

$$G^k_{ac}(t) = (2k+1)^{1/2} \int_0^\infty dR a(R) c(R) J_k(tR) R. \quad \text{(A12)}$$

Note again that use of the $R_<, R_>$ coordinates is avoided. We can return to these coordinates by substituting (A12) into (A11) and integrating first over t to get

$$S^k(ab, cd) = \int_0^\infty dR_1 \int_0^\infty dR_2 (a_1 c_1 R_1) g_k(R_<, R_>)(b_2 d_2 R_2), \quad \text{(A13)}$$

with

$$g_k(R_<, R_>) = \frac{\Gamma(k+\frac{3}{2})}{\Gamma(k+1)\Gamma(\frac{3}{2})} \cdot f_k(R_<, R_>) \times F(k+\tfrac{1}{2}, \tfrac{1}{2}; k+1; (R_</R_>)^2). \quad \text{(A14)}$$

The special case $g_0(R_<, R_>)$, representing a circular average ($m_a = m_b = 0$) of $1/R_{12}$, was given in Ref. 7. Note that the leading term in (A14) is $R_<^k / R_>^{k+1}$. Comparison with the Legendre expansion in (A1) gives the selection rule

$$\int_0^{2\pi} d\phi_1 \int_0^{2\pi} d\phi_2 P_{k'}(\cos\phi_{12}) \exp(\pm ik\phi_{12}) = 0 \quad \text{(A15)}$$

for integer $k > k' \geq 0$.

We illustrate the difference between the $D=2$ and $D=3$ Coulomb integrals by comparing the quantities E^D_{ss}, E^D_{pp}, and E^D_{sp} defined in Eqs. (8)–(10). They are for $D=3$:

$$E^3_{ss} = R^0(2s2s, 2s2s),$$
$$E^3_{pp} = R^0(2p2p, 2p2p) + \tfrac{2}{5}R^2(2p2p, 2p2p), \quad \text{(A16)}$$
$$E^3_{sp} = -\sqrt{3}R^1(2s2p, 2s2p);$$

and for $D=2$:

$$E^2_{ss} = S^0(2s2s, 2s2s),$$
$$E^2_{pp} = S^0(2p2p, 2p2p) + \tfrac{1}{5}S^2(2p2p, 2p2p), \quad \text{(A17)}$$
$$E^2_{sp} = -\sqrt{2}S^1(2s2p, 2p2s).$$

[1] (a) D.R. Herrick and O. Sinanoglu, Phys. Rev. A 11, 97 (1975). (b) O. Sinanoglu and D.R. Herrick, J. Chem. Phys. 62, 886 (1975); (c) D.R. Herrick, Ph.D. thesis, Yale University, 1973.

[2] M.J. Seaton, Proc. Phys. Soc. 77, 173 (1961).

[3] M. Gailitis and R. Damburg, Proc. Phys. Soc. 82, 192 (1963).

[4] D.R. Herrick and O. Sinanoglu, Rev. Mex. Fis. 22, 1 (1973).

[5] J. Macek, J. Phys. B 1, 831 (1968).

[6] A recent study of the configuration interaction has resulted in a classification of mixings similar to the DESB. See C. Wulfman, Chem. Phys. Lett. 23, 370 (1973). Our classification and quantum numbers (Ref. 1) are valid for entire Rydberg series, and are thus more general than Wulfman's.

[7] D.R. Herrick and F.H. Stillinger, Phys. Rev. A 11, 42 (1972).

[8] S.P. Alliluev, Zh. Eksp. Teor. Fis. 33, 200 (1957) [Sov. Phys. JETP 6, 156 (1958)].

[9] E.C.G. Sudarshan, N. Mukunda, and L. O'Raifeartaigh, Phys. Lett. 19, 322 (1965).

[10] See. Ref. 4. This paper also describes the noncompact extension $SO(D+1,2)$.

[11] V. Fock, Z. Phys. 98, 145 (1935).

[12] With normalization $\int_0^\infty (R^D_{n\Lambda})^2 R^{D-1} dR = 1$.

[13] E.U. Condon and G.H. Shortley, *The Theory of Atomic Spectra* (Cambridge U.P., London, 1953).

[14] $\alpha = 4\epsilon^1_D/(D^2-1)$, where ϵ^1_D is the first order $1s^2$ ground state energy defined in Eq. (45) of Ref. 7.

[15] This represents the main differences between the $D=2$ and $D=3$ states. For $D=3$ there are two classes of states, having parity $(-1)^L$ or $(-1)^{L+1}$, where L is the total angular momentum number. For $D=2$ the parity is a redundant quantum number. No analog to the $D=3$ quantum number T exists for $D=2$.

[16] In momentum ($=P$) space the U_+ are equal (within normalization to $\cos^2(\phi/2)\exp(\mp im\phi)$, with $\phi = 2\arctan(P/\sqrt{-2E_n})$ being the hyperspherical angle of the Fock space.

[17] The corresponding noncompact algebra which classifies the entire spectrum $\pm m = 0, 1, \cdots$ is $SO(2,1)$. See, for example, Ref. 4.

[18] D.R. Herrick, J. Math. Phys. 16, 281 (1975).

[19] J.W. Cooper, U. Fano, and F. Prats, Phys. Rev. Lett. 10, 518 (1963).

The inverse scattering transform: Semi-infinite interval

M. J. Ablowitz and H. Segur

Mathematics Department, Clarkson College of Technology, Potsdam, New York 13676
(Received 12 September 1974; revised manuscript received 1 November 1974)

Certain nonlinear evolution equations can be solved on the semi-infinite interval by the method of inverse scattering. These equations are a subset of those which can be solved on the full interval. The equations have even dispersion relations when linearized, and are subject to appropriate homogeneous boundary conditions at the origin.

1. INTRODUCTION

Many nonlinear evolution equations can be solved exactly as initial value problems on the infinite interval $-\infty < x < \infty$ by the method of inverse scattering. The form of each of these equations is characterized by the dispersion relation of its linearized form and an integro-differential operator. As shown in Ref. 1, the problems that can be solved by use of the generalized Zakharov–Shabat eigenvalue problem have the form

$$\frac{\partial}{\partial t}\begin{pmatrix} r \\ -q \end{pmatrix} + 2A_0(L^+)\begin{pmatrix} r \\ q \end{pmatrix} = 0, \tag{1.1}$$

where L^+ is the integro–differential operator

$$L^+ = \frac{1}{2i}\begin{bmatrix} \dfrac{\partial}{\partial x} - 2r\displaystyle\int_{-\infty}^{x} dy\, q & 2r\displaystyle\int_{-\infty}^{x} dy\, r \\ -2q\displaystyle\int_{-\infty}^{x} dy\, q & -\dfrac{\partial}{\partial x} + 2q\displaystyle\int_{-\infty}^{x} dy\, r \end{bmatrix} \tag{1.2}$$

and $A_0(\zeta)$ is proportional to the dispersion relation of the linearized problem. As a simple example, if

$$A_0(\zeta) = -2i\zeta^2, \quad r = \pm q^*, \tag{1.3}$$

then (1.1) reduces to the nonlinear Schrödinger equation[2]

$$q_t - iq_{xx} \pm 2i|q|^2 q = 0. \tag{1.4}$$

It was also shown in Ref. 1 that the method of inverse scattering, as applied to nonlinear equations, is analogous to the method of Fourier transforms for linear equations.

In this paper, we consider the question of whether any of these equations can be solved on the semi-infinite interval, $0 < x < \infty$. The question has a simple answer, which reemphasizes the close connection between inverse scattering and Fourier transforms (which we use here in the narrow sense, excluding other eigenfunction expansions). We find that those evolution equations of the form (1.1) which can be solved on the semi-infinite interval are exactly those which, when linearized, can be solved by either a Fourier sine or cosine transform.

For a linear evolution equation,

$$u_t = Lu, \quad x, t > 0, \tag{1.5}$$

where L is a linear differential (in x) operator with constant coefficients, Fourier transforms are useful if the (homogeneous) boundary conditions at $x = 0$ allow $u(x, t)$ to be extended to $-\infty < x < \infty$ as either (a) an odd or (b) an even function of x. Specifically, the solution of (1.1) may then be written as

$$u(x, t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} r_0(k)\exp\{i[kx + \omega(k)t]\}\,dk, \tag{1.6}$$

where $r_0(k) = \int_{-\infty}^{\infty} u(x, 0)\exp(-ikx)\,dx$, and $\omega(k)$ is the dispersion relation of the equation. Since $u(x, 0)$ is now (a) odd or (b) even, so is $r_0(k)$. Then if and only if $\omega(k)$ is even, $u(x, t)$ remains (a) odd or (b) even for all $t > 0$.

For a nonlinear problem that can be solved by inverse scattering (on $-\infty < x < \infty$), similar conclusions hold. We will show that if the linearized dispersion relation is even and appropriate homogeneous boundary conditions at $x = 0$ are imposed, the solution of the nonlinear equation can be extended as an odd (or even) function. Moreover, if the linearized dispersion relation is odd, a counterexample is given which suggests no extension will provide the solution by inverse scattering.

2. ANALYSIS

The generalized Zakharov–Shabat eigenvalue problem is ($-\infty < x < \infty$, $t > 0$):

$$v_{1x} + i\zeta v_1 = qv_2, \quad v_{2x} - i\zeta v_2 = rv_1, \tag{2.1}$$

with the associated time-dependence

$$v_{1t} = Av_1 + Bv_2, \quad v_{2t} = Cv_1 - Av_2. \tag{2.2}$$

A, B, C are given functions of $q(x, t)$, $r(x, t)$, and ζ. Cross differentiation of (2.1) and (2.2) yields, as a compatibility condition, the evolution equations for q, r.[3] Suppose that, at time $t = 0$, (a) $q(x)$ and $r(x)$ are odd functions or (b) $q(x)$ and $r(x)$ are even functions. Then there are symmetry relations for (2.1) and (2.2):

(a) $v_1(x, \zeta) = v_1(-x, -\zeta)$,
$v_2(x, \zeta) = v_2(-x, -\zeta)$,
$A(x, \zeta) = A(-x, -\zeta)$,
$B(x, \zeta) = B(-x, -\zeta)$,
$C(x, \zeta) = C(-x, -\zeta)$;

(b) $v_1(x, \zeta) = v_1(-x, -\zeta)$,
$v_2(x, \zeta) = -v_2(-x, -\zeta)$,
$A(x, \zeta) = A(-x, -\zeta)$,

 Copyright © 1975 American Institute of Physics

$$B(x,\zeta)=B(-x,-\zeta),$$
$$C(x,\zeta)=C(-x,-\zeta), \tag{2.3}$$

In the case where $r, q \to 0$ sufficiently rapidly as $|x| \to \infty$, define particular solutions of (2.1) at $t=0$ by

$$\begin{aligned}
&\phi \to \begin{pmatrix}1\\0\end{pmatrix}\exp(-i\zeta x) \\
&\qquad\qquad , \quad x \to -\infty, \\
&\bar{\phi} \to \begin{pmatrix}0\\-1\end{pmatrix}\exp(i\zeta x) \\
&\psi \to \begin{pmatrix}0\\1\end{pmatrix}\exp(i\zeta x) \\
&\qquad\qquad , \quad x \to +\infty. \\
&\bar{\psi} \to \begin{pmatrix}1\\0\end{pmatrix}\exp(-i\zeta x)
\end{aligned} \tag{2.4}$$

Here ϕ and $\bar{\phi}$ (or ψ and $\bar{\psi}$) are two linearly independent solutions of (2.1), rather than complex conjugates. It follows from (2.3) that:

$$\begin{aligned}
&\text{(a)}\ \ \phi(x,\zeta)=\bar{\psi}(-x,-\zeta), \\
&\qquad \bar{\phi}(x,\zeta)=-\psi(-x,-\zeta); \\
&\text{(b)}\ \ \phi_1(x,\zeta)=\bar{\psi}_1(-x,-\zeta), \\
&\qquad \phi_2(x,\zeta)=-\bar{\psi}_2(-x,-\zeta), \\
&\qquad \bar{\phi}_1(x,\zeta)=\psi_1(-x,-\zeta), \\
&\qquad \bar{\phi}_2(x,\zeta)=-\psi_2(-x,-\zeta).
\end{aligned} \tag{2.5}$$

The Wronskians of these solutions, which comprise the scattering data (see, for example, Ref. 1), satisfy:

$$\begin{aligned}
&\text{(a)}\ \ a(\zeta)=\bar{a}(-\zeta), \qquad \text{(b)}\ \ a(\zeta)=\bar{a}(-\zeta), \\
&\qquad b(\zeta)=-b(-\zeta), \qquad\quad b(\zeta)=b(-\zeta), \\
&\qquad \bar{b}(\zeta)=-\bar{b}(-\zeta); \qquad\quad \bar{b}(\zeta)=\bar{b}(-\zeta).
\end{aligned} \tag{2.6}$$

For $t>0$, the scattering data evolves as

$$\begin{aligned}
&a(\zeta,t)=a(\zeta), \quad b(\zeta,t)=b(\zeta)\exp[-2A_0(\zeta)t], \\
&\bar{a}(\zeta,t)=\bar{a}(\zeta), \quad \bar{b}(\zeta,t)=\bar{b}(\zeta)\exp[2A_0(\zeta)t],
\end{aligned} \tag{2.7}$$

where

$$A_0(\zeta)=\lim_{|x|\to\infty} A(x,\zeta;t),$$

and is directly related to the dispersion relation of the linearized problem. Thus (2.6), which follows from the assumption of oddness or evenness at $t=0$, remains valid for $t>0$ if and only if $A_0(\zeta)$ is an even function of ζ. This is analogous to the results in the linear problem.

Conversely, we show that if the eigenfunctions satisfy (2.5) [so that the scattering data satisfy (2.6)], then the corresponding potentials must be (a) odd or (b) even functions of x. These eigenfunctions have the representations

$$\begin{aligned}
&\phi(x,\zeta) \\
&\quad=\begin{pmatrix}1\\0\end{pmatrix}\exp(-i\zeta x)-\int_{-\infty}^{x} L(x,s)\exp(-i\zeta s)\,ds, \\
&\bar{\phi}(x,\zeta) \\
&\quad=-\begin{pmatrix}0\\1\end{pmatrix}\exp(i\zeta x)-\int_{-\infty}^{x} \bar{L}(x,s)\exp(i\zeta s)\,ds,
\end{aligned} \tag{2.8}$$

$$\begin{aligned}
&\psi(x,\zeta) \\
&\quad=\begin{pmatrix}0\\1\end{pmatrix}\exp(i\zeta x)+\int_{x}^{\infty} K(x,s)\exp(i\zeta s)\,ds, \\
&\bar{\psi}(x,\zeta) \\
&\quad=\begin{pmatrix}1\\0\end{pmatrix}\exp(-i\zeta s)+\int_{x}^{\infty} \bar{K}(x,s)\exp(-i\zeta s)\,ds,
\end{aligned}$$

where $L=\binom{L_1}{L_2}$, etc. It can be shown that

$$\begin{aligned}
&K_1(x,x)=-\bar{L}_1(x,x)=-\tfrac{1}{2}q(x), \\
&L_2(x,x)=\bar{K}_2(x,x)=-\tfrac{1}{2}r(x).
\end{aligned} \tag{2.9}$$

Substituting (2.8) into (2.5) yields ($s \geq x$):

$$\begin{aligned}
&\text{(a)}\ \ K(x,s)=\bar{L}(-x,-s), \qquad \text{(b)}\ \ K_1(x,s)=-\bar{L}_1(-x,-s), \\
&\qquad \bar{K}(x,s)=-L(-x,-s); \qquad\quad K_2(x,s)=\bar{L}_2(-x,-s), \\
&\qquad\qquad\qquad\qquad\quad \bar{K}_1(x,s)=-L_1(-x,-s), \\
&\qquad\qquad\qquad\qquad\quad \bar{K}_2(x,s)=L_2(-x,-s).
\end{aligned} \tag{2.10}$$

The relations (2.9) and (2.10) ensure that the potentials q, r remain (a) odd or (b) even for all time.

The solution procedure is to solve either of the sets of equations,

$$\begin{aligned}
&\bar{K}(x,y)+\begin{pmatrix}0\\1\end{pmatrix}F(x+y)+\int_{x}^{\infty} K(x,s)\,F(s+y)\,ds=0 \quad (y>x), \\
&K(x,y)-\begin{pmatrix}1\\0\end{pmatrix}\bar{F}(x+y)-\int_{x}^{\infty} \bar{K}(x,s)\,\bar{F}(s+y)\,ds=0 \quad (y>x),
\end{aligned} \tag{2.11}$$

or

$$\begin{aligned}
&\bar{L}(x,y)+\begin{pmatrix}1\\0\end{pmatrix}G(x+y)-\int_{-\infty}^{x} L(x,s)\,G(s+y)\,ds=0 \quad (x>y), \\
&L(x,y)+\begin{pmatrix}0\\1\end{pmatrix}\bar{G}(x+y)+\int_{-\infty}^{x} \bar{L}(x,s)\,\bar{G}(s+y)\,ds=0 \quad (x>y),
\end{aligned} \tag{2.12}$$

for $K, \bar{K}$ or $L, \bar{L}$. Here

$$\begin{aligned}
&F(z)\equiv\frac{1}{2\pi}\oint\frac{b(\zeta)}{a(\zeta)}\exp[i\zeta z-2A_0(\zeta)t]\,d\zeta, \\
&\bar{F}(z)\equiv\frac{1}{2\pi}\oint\frac{\bar{b}(\zeta)}{\bar{a}(\zeta)}\exp[-i\zeta z+2A_0(\zeta)t]\,d\zeta,
\end{aligned} \tag{2.13}$$

$$\begin{aligned}
&G(z)\equiv\frac{1}{2\pi}\oint\frac{\bar{b}(\zeta)}{a(\zeta)}\exp[-i\zeta z+2A_0(\zeta)t]\,d\zeta, \\
&\bar{G}(z)\equiv\frac{1}{2\pi}\oint\frac{b(\zeta)}{\bar{a}(\zeta)}\exp[i\zeta z-2A_0(\zeta)t]\,d\zeta.
\end{aligned} \tag{2.14}$$

In (2.13)–(2.14) the contours C ($\bar{C}$) refer to integration above (below) all zeros of $a(\zeta)$ ($\bar{a}(\zeta)$), assuming $q(x,0)$, $r(x,0)$ are on compact support. $F, \bar{F}, G, \bar{G}$ satisfy the linearized equations and obey the symmetry relations,

$$\begin{aligned}
&\text{(a)}\ \ F(z)=-\bar{G}(-z), \quad \text{(b)}\ \ F(z)=\bar{G}(-z), \\
&\qquad \bar{F}(z)=-G(-z), \qquad\quad \bar{F}(z)=G(-z),
\end{aligned} \tag{2.15}$$

which ensure that (2.10) are consistent with (2.11)–(2.12). Note that in the special case $r(x,t)=0$ [the evolution equation for $q(x,t)$ is rendered linear] then (2.13)–(2.14) properly reduce to (a) sine transforms or (b)

cosine transforms. Thus, the nonlinear evolution equation which can be solved by the inverse scattering method on $0 < x < \infty$ are exactly those which, when linearized, can be solved by a Fourier sine or cosine transform.

It is a consequence of the method of inverse scattering that with the time dependence (2.7) an infinite set of globally conserved quantities on $|x| < \infty$ exist. If $g(x)$, $r(x)$ are (a) odd or (b) even, half of these vanish identically. The other half yield conserved quantities of the form $\int_0^\infty (\cdots)\,dx$.

A much different situation arises when the dispersion relation is odd. Extensions to the full interval cannot be expected to yield valid solutions. Indeed we know this to be true of linear problems as well. For example, consider

$$q_t + q_{xxx} = 0 \tag{2.16}$$

on the interval $-\infty < x < \infty$ [$r(x,t) = 0$, $A_0 = -4i\xi^3$ in our scattering formalism]. The solution has the form

$$q(x,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} r_0(k)\exp[i(kx + k^3t)]\,dk. \tag{2.17}$$

If (2.16) holds on the semi-infinite interval $0 < x < \infty$, subject to the boundary condition $q(x=0) = 0$, then $q(x,t) \equiv 0$ is the only solution of the form (2.17) which vanishes at the origin for all time. Thus, the method of Fourier transforms (or inverse scattering) yields only the trivial solution. There are nontrivial solutions of (2.16) of the form,

$$u(x,t) = \int_0^\infty r_0(k)\exp(-kx/2)\sin(\tfrac{1}{2}\sqrt{3}\,kx)\exp(-k^3t)\,dk, \tag{2.18}$$

but they are not recovered by this method.

Finally, we note that, in solving nonlinear evolution equations, use of the full interval $-\infty < x < \infty$ is important. If one attempts to use inverse scattering procedures on $0 < x < \infty$, then the time dependence of the eigenfunctions at $x = 0$ is required information. The solution process then depends on the solution of (2.2) at the origin. Unfortunately, $A(x,t)$, $B(x,t)$, $C(x,t)$ introduce further unknowns, not given by the boundary conditions, into the equations.

The conclusion is that the solution procedure for nonlinear evolution equations by inverse scattering on the interval $0 < x < \infty$ is in close analogy to that of Fourier analysis. The method is valid when the linearized dispersion relation is even and appropriate homogeneous boundary conditions are given at the origin.

ACKNOWLEDGMENT

This work was supported by N. S. F. Grants GA-27727A1 and GP-32839X.

[1]M.J. Ablowitz, D.J. Kaup, A.C. Newell and H. Segur, Stud. Appl. Math. **53**, 255 (1974).
[2]V.E. Zakharov and A.B. Shabat, Zh. Eksp. Teor. Fiz. **61**, 118 (1971) [Sov. Phys. JETP **34**, 62 (1972)].
[3]M.J. Ablowitz, D.J. Kaup, A.C. Newell, and H. Segur, Phys. Rev. Lett. **31**, 125 (1973).

On the transformation from random Schrödinger Hamiltonians to random Hamiltonian matrices

J. F. McDonald*

Department of Mathematics, University of Windsor, Windsor, Ontario, Canada

L. D. Favro

Department of Physics, Wayne State University, Detroit, Michigan
(Received 5 August 1974)

In this article we consider the effects on ensembles of random Hamiltonian matrices when certain restrictions are imposed on the potentials of the corresponding Schrödinger Hamiltonians. In particular, we investigate the validity of the assumption which is usually made when ensembles of random matrices are used to predict statistical properties of energy levels of complex systems, namely that the Hamiltonian matrix elements are statistically independent.

1. INTRODUCTION

Recently there has been considerable interest in studying the restrictions imposed on random Hamiltonian matrices by the requirement that they represent physically realistic forces. In particular, extensive numerical computations have been undertaken to study the implications of the limitation of the forces to two-body interactions and to study the effects of the Pauli principle.[1] These computations start with assumptions about the effect of the physical conditions on matrix element distributions and proceed by computer diagonalization of the resulting matrices.

The purpose of this article is to provide a systematic method for determining the matrix element distribution which corresponds to any physical set of potentials. The method utilizes a particular representation (harmonic oscillator states are used as a basis) to calculate analytic expressions for the transformation from a random Schrödinger equation. The results are essentially identical (except for the dimensionality of the matrices involved) for a single particle in an arbitrarily complicated external potential or for a system of particles with arbitrarily complicated interactions. Explicit transformation matrices are given for both cases, and it is shown that, for sufficiently general Schrödinger equations, all of the matrix elements H_{ij}, $j \geq i$, are independent. Since relatively simple analytic expressions can be obtained for both the transformation matrix and its inverse one can transform from Schrödinger equation to Hamiltonian matrix and vice versa at will.

As an example of the effect of the imposition of physical restrictions on the Hamiltonian matrices we then consider an ensemble of velocity-independent Hamiltonians. This imposes so severe a restriction on the matrix that now instead of having all of the matrix elements independent we find that only the diagonal elements are independent. The off-diagonal elements are linear combinations of the diagonal ones. This causes restrictions on the resulting eigenvalue distribution, but does not destroy the functional independence of the eigenvalues.

We then turn to the question of the effects of the statistics of the particles on the matrix element distribution. This is done by reformulating the problem in second-quantized form and again calculating the transformation matrices explicitly. For bosons the algebra is unchanged from the previous version of the transformation. For fermions the algebra is different but the results are effectively the same. In both cases one finds that, provided the interactions are complicated enough, the matrix elements are all functionally independent. However, this requires the use of many-body forces up to and including N-body forces, where N is the number of particles present. Thus, in practical cases where the order of the interaction is quite low the matrix elements will not be independent. Nonetheless, as one increases the order of the admissible interactions, the number of restrictions on the matrix elements decreases rapidly so that one would expect "approximate" independence to prevail.[7]

2. DEFINITIONS AND CONVENTIONS

A basic assumption in the statistical theory of spectra is that the Hamiltonian matrix of a sufficiently complex system will be a typical sample of the same multivariate distribution in almost any representation.[8] Thus, we are essentially free in our choice of a representation for describing a complex Hamiltonian.[9]

We will choose as our basis functions, the eigenfunctions of an isotropic harmonic oscillator. (In the beginning, we shall not consider spin dependent ensembles.) Such basis functions are commonly described by defining so-called "ladder operators" for each degree of freedom.[10] These operators may be defined in terms of the coordinate $\mathbf{x}$ and momentum $\mathbf{p}$ associated with a given degree of freedom as

$$\beta^* \equiv (\mathbf{x} - i\mathbf{p})/\sqrt{2}, \tag{2.1}$$

$$\beta \equiv (\mathbf{x} + i\mathbf{p})/\sqrt{2}, \tag{2.2}$$

where the system of units has been chosen such that $m = k = \hbar = 1$ (m is the mass of the oscillator, k is the spring constant, and $\hbar$ is Planck's constant divided by 2π).

It follows from the canonical commutation relation $[\mathbf{p}, \mathbf{x}] = -i$ that

 Copyright © 1975 American Institute of Physics

$$[\beta^m, (\beta^+)^n] = \sum_{j=1}^{m} \binom{m}{j}\binom{n}{j} j!(\beta^+)^{n-j}\beta^{m-j}, \quad m,n = 0,1,2,\cdots. \tag{2.3}$$

Using this result it can then be verified that

$$\mathbf{x}^n = 2^{-n/2}(\beta^+ + \beta)^n = 2^{-n/2}\sum_{j=0}^{n} \eta_j^n \sum_{k=0}^{j}\binom{j}{k}(\beta^+)^{j-k}\beta^k \tag{2.4}$$

and

$$\mathbf{p}^n = i^n 2^{-n/2}(\beta^+ - \beta)^n = i^n 2^{-n/2}\sum_{j=0}^{n}(-1)^j \eta_j^n \sum_{k=0}^{j} (-1)^k\binom{j}{k}(\beta^+)^{j-k}\beta^k, \tag{2.5}$$

where n is the nonnegative integer, and where

$$\eta_j^n \equiv \left(\frac{1+(-1)^{n+j}}{2}\right)\frac{n!}{j![(n-j)/2]!\,2^{(n-j)/2}}. \tag{2.6}$$

Finally, we choose the normalization and phase of the oscillator eigenfunctions such that

$$\langle n|m\rangle = \delta_{nm} \tag{2.7}$$

and

$$\beta^m|n\rangle = \left(\frac{n!}{(n-m)!}\right)^{1/2}|n-m\rangle, \tag{2.8}$$

where we have used Dirac notation

3. ENSEMBLES OF NONSINGULAR HAMILTONIANS

First we consider the most general ensemble of Hamiltonians corresponding to a one-dimensional particle in a nonsingular potential (i. e., we consider Hamiltonians which are expandable in nonnegative integer powers of **x** and **p**). The ensembles can be specified mathematically by writing the Hamiltonian, **H**, in the form[11]

$$\mathbf{H} = \sum_{n=0}^{N}\sum_{k=0}^{n} b_{nk}(\beta^+)^{n-k}\beta^k, \tag{3.1}$$

where the b_{nk} are parameters, which have an associated multivariate distribution that determines the ensemble. The largest power involved in **H** has been chosen arbitrarily to be N. Later, we will have reason to choose a particular value of N, which will be determined by considerations of functional independence of matrix elements.

Since **H** is Hermitian, the parameters b_{nk} must satisfy

$$b_{nk}^* = b_{n\,n-k}. \tag{3.2}$$

Also, if we restrict H to be time reversal invariant, then

$$b_{nk}^* = b_{nk} \tag{3.3}$$

(i. e., the b_{nk} are real).

We wish now to calculate the matrix elements corresponding to this ensemble. As is the usual procedure in the statistical theory of energy levels we shall treat the Hamiltonian as a matrix of arbitrary but finite dimension, say $M+1$. (Then, of course, one must take a limit as $M \to \infty$.)

From Eqs. (2.8) and (3.1) it follows easily that

$$H_{ij} \equiv \langle i|H|j\rangle = \frac{1}{2}\sum_{n=0}^{N}\frac{[1+(-1)^{n+j-i}](i!j!)^{1/2}}{[(i+j-n)/2]!}\, b_{n,\,(n+j-i)/2}, \tag{3.4}$$

where $i,j = 0,1,\ldots,M$, and where $b_{nl} \equiv 0$, $l > n$. Note that only odd or even values of n contribute to the sum on the right if $j-i$ is odd or even, respectively. This is because our states have definite parity, and each term in the Hamiltonian (but not the Hamiltonian itself) has definite parity. It should be pointed out that if $N > 2M$, those b_{nl} with $n > 2M$ do not contribute to any matrix element. Thus, terms in Eq. (3.1) involving those b_{nl} are superfluous. Note also that due to properties (3.2) and (3.3), it follows that **H** is real and symmetric. Thus, from this point we need consider only the matrix elements H_{ij}, $j \geq i$.

Equation (3.4) can be rewritten as

$$v_i^{(s)} = \sum_{k=0}^{[\mu_s]}\frac{1}{(i-k)!}\,u_k^{(s)}, \quad s = 0,1,\ldots,M,$$
$$i = 0,1,2,\ldots,M-s, \tag{3.5}$$

where $\mu_s \equiv (N-s)/2$, $[m/2]$ is the largest integer less than or equal to $m/2$,

$$v_i^{(s)} = H_{i\,i+s}/(i!\,(i+s)!)^{1/2}, \tag{3.6}$$

and

$$u_k^{(s)} \equiv b_{s+2k,\,s+k}. \tag{3.7}$$

Note that if $N < 2M$ there will be a smaller number of $u_k^{(s)}$ than there will be $v_i^{(s)}$. This means that the $v_i^{(s)}$, and thus the H_{ij}, $j \geq i$, are not functionally independent. Thus the assumption usually made in the statistical theory of energy levels that the H_{ij}, $i \leq j$, are statistically independent would certainly be false if the ensemble consisted of only Hamiltonians with $N < 2M$. That is, if the ensemble does not contain a large enough class of Hamiltonians, then the assumption of statistical independence of the matrix elements is untrue.

If $N \geq 2M$, then Eq. (3.5) can be rewritten as

$$v_i^{(s)} = \sum_{k=0}^{i}\frac{1}{(i-k)!}\,u_k^{(s)}. \tag{3.8}$$

Clearly, there are exactly the same number of $u_k^{(s)}$ involved in (3.8) as there are $v_i^{(s)}$. Further, since the Jacobian of the transformation does not vanish, i. e.,

$$\left|J\begin{pmatrix} v_i^{(s)} \\ u_k^{(s)}\end{pmatrix}\right| = 1, \tag{3.9}$$

the $v_i^{(s)}$ are all functionally independent provided the $u_k^{(s)}$ are functionally independent.

At this point it is clear that if the multivariate distribution for the b_{nk} is known, the multivariate distribution for the H_{ij}, $j \geq i$, can be obtained by direct calculation by use of Eqs. (3.6)–(3.9). Further, if $N \geq 2M$ the inverse of the transformation given by (3.8) can be obtained explicitly. In particular, it can be shown that

$$u_i^{(s)} = \sum_{k=0}^{i}\frac{(-1)^{i+k}}{(i-k)!}\,v_k^{(s)}, \tag{3.10}$$

where $s = 0,1,\ldots,M$, $i = 0,1,\ldots,M-s$.

Note that not all of the b_{nl} can be obtained from Eqs.

(3.10) and (3.7). However, all other b_{nl} (excluding the superfluous ones) can be obtained from relations (3.2) and (3.3) (i.e., they are not independent). Furthermore, if N is exactly equal to $2M$, through (3.8) and (3.10) there is a *known* one-to-one transformation between *all* of the independent b_{nl} and the independent H_{ij}. Thus, given any ensemble of one-dimensional Schrödinger Hamiltonians [defined through (3.1)] the corresponding matrix element distribution can be obtained explicitly by using (3.8). Conversely, if the distribution of matrix elements is known, the corresponding ensemble of Schrödinger Hamiltonians can be obtained through (3.10).

From the above discussion it is clear that provided the ensemble of matrices is general enough, (i.e., $N \geq 2M$) almost any joint distribution of the H_{ij}, $i \leq j$, is possible. That is, even though we are considering a fairly special case of Hamiltonians (i.e., one-dimensional particle with potential expandable in powers of $\mathbf{x}$ and $\mathbf{p}$) there seems to be no constraint on the matrix element distribution.

If we add the condition that the Hamiltonian is invariant under space inversion, the matrix elements between states of different parity will vanish. Thus, we can choose our basis such that all Hamiltonian matrices in the ensemble consist of two blocks along the diagonal, one for each parity. Since all the matrix elements are still independent functions of the remaining b_{nl}, the transformations being exactly the same as above, there is essentially no change in the conclusions. However, we now would, of course, consider each block separately.

It is now a simple matter to generalize to an ensemble of Hamiltonians in a *multi*dimensional space. Again we will use Hamiltonians which are expandable in positive integer powers of the canonical variables, $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_K$ and $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_K$. Such a Hamiltonian corresponds to an n particle system in an m-dimensional space, such that $nm = K$.

The most general Hamiltonian of this form can be written as

$$\mathbf{H} = \sum_{n_1=0}^{N_1} \sum_{k_1=0}^{n_1} \cdots \sum_{n_K=0}^{N_K} \sum_{k_K=0}^{n_K} b_{n_1 \cdots n_K; k_1 \cdots k_K} \times (\beta_1^*)^{n_1-k_1} \beta_1^{k_1} \cdots (\beta_K^*)^{n_K-k_K} \beta_K^{k_K}. \tag{3.11}$$

A calculation analogous to that for the one-dimensional case above yields

$$H_{i_1 \cdots i_K; j_1 \cdots j_K} = \sum_{n_1=0}^{N_1} \cdots \sum_{n_K=0}^{N_K} A_{\mathbf{ijn}} \times b_{n_1 \cdots n_K; (n_1+j_1-i_1)/2 \cdots (n_K+j_K-i_K)/2}, \tag{3.12}$$

$$i_1, j_1 = 0, 1, \ldots, M_1, \cdots, \quad i_K, j_K = 0, 1, \ldots, M_K,$$

where

$$A_{\mathbf{ijn}} \equiv \frac{1}{2^K} \prod_{k=1}^{K} \frac{[1 + (-1)^{n_k+j_k-i_k}]}{[(i_k+j_k-n_k)/2]!} \sqrt{i_k! j_k!}, \tag{3.13}$$

and

$$|i_1 i_2 \cdots i_k\rangle \equiv |i_1\rangle |i_2\rangle \cdots |i_k\rangle. \tag{3.14}$$

If we now define

$$v_{\mathbf{i}}^{(s)} \equiv H_{i_1 \cdots i_K; i_1+s_1 \cdots i_K+s_K} \Big/ \prod_{k=1}^{K} (i_k! (i_k+s_k)!)^{1/2} \tag{3.15}$$

and

$$u_{\mathbf{k}}^{(s)} \equiv b_{s_1+2k_1 \cdots s_K+2k_K; s_1+k_1 \cdots s_K+k_K}, \tag{3.16}$$

Eq. (3.13) can be rewritten as

$$v_{\mathbf{i}}^{(s)} = \sum_{k_1=0}^{[\mu_1]} \cdots \sum_{k_K=0}^{[\mu_K]} T_{\mathbf{ik}} u_{\mathbf{k}}^{(s)}, \tag{3.17}$$

where

$$T_{\mathbf{ik}} \equiv T_{i_1 \cdots i_K; k_1 \cdots k_K} = \prod_{m=1}^{K} \frac{1}{(i_m - k_m)!} \tag{3.18}$$

and where $\mu_k \equiv (N_k - s_k)/2$.

It should be clear at this point all calculations and conclusions are essentially unchanged from the one-dimensional case. In particular, the Hamiltonian matrix is real and symmetric (assuming time reversal invariance), and the matrix elements can be functionally independent only if $N_m \geq 2M_m$, $m = 1, \ldots, K$. If these conditions are satisfied, then (3.17) can be rewritten as

$$v_{\mathbf{i}}^{(s)} = \sum_{k_1=0}^{i_1} \cdots \sum_{k_K=0}^{i_K} T_{\mathbf{ik}} u_{\mathbf{k}}^{(s)}. \tag{3.19}$$

The $H_{i_1 \cdots i_K; j_1 \cdots j_K}$, $j_1 \geq i_1, \ldots, j_K \geq i_K$, will be functionally independent provided the $u_{\mathbf{k}}^{(s)}$ are independent. (In fact, the absolute value of the Jacobian will again be unity.) Further, it is easily verified that

$$(\mathbf{T}^{-1})_{i_1 \cdots i_K; j_1 \cdots j_K} = \prod_{k=1}^{K} \frac{(-1)^{i_k+j_k}}{(i_k - j_k)!}. \tag{3.20}$$

If $N_m = 2M_m$, $m = 1, 2, \ldots, K$, we have a *known* one-to-one transformation between the independent $b_{\mathbf{nj}}$ and the independent matrix elements. Thus, given an ensemble of Hamiltonians of the form (3.11) the corresponding matrix element distribution can be obtained using (3.17). Conversely, if the matrix element distribution is known, the ensemble of Schrödinger Hamiltonians can be obtained using inverse transformation given by (3.20).

It should be noted that using the same techniques one can treat the problem of an ensemble of Hamiltonians of the form (3.11) corresponding to a single particle in three dimensions with the *additional* restriction that the Hamiltonians be spherical symmetric. Since the conclusions are basically the same as those above the explicit calculations will not be given here.

4. AN ENSEMBLE CORRESPONDING TO HAMILTONIAN WITH NONSINGULAR VELOCITY INDEPENDENT POTENTIALS

As an example of an ensemble for which the matrix elements are not independent we consider one which is restricted to velocity independent potentials. In particular, let us consider an ensemble of one-dimensional Hamiltonians of the form

$$\mathbf{H} = 2A\mathbf{p}^2 + \sum_{k=0}^{N} a_k (\sqrt{2}\mathbf{x})^k, \tag{4.1}$$

where A is a real positive constant, and the a_k are parameters which have some distribution associated with them. Clearly, the a_k are real since H is Hermi-

tian. Further, since each H is real and spin independent it is also time reversal invariant.

Inserting Eqs. (2.4) and (2.5) into (4.1), we obtain

$$H = -A[(\beta^+)^2 - 2\beta^+\beta + \beta^2] + \sum_{n=0}^{N}\sum_{k=0}^{n}\binom{n}{k} b_n (\beta^+)^{n-k}\beta^k, \tag{4.2}$$

where

$$b_{2n} = \sum_{k=n}^{[N/2]} \eta_{2n}^{2k} a_{2k} + A\delta_{n0} \tag{4.3}$$

and

$$b_{2n+1} \equiv \sum_{k=n}^{[(N-1)/2]} \eta_{2n+1}^{2k+1} a_{2k+1}, \tag{4.4}$$

with η_m^l defined by Eq. (2.6).

It should be noted that

$$\left| J\binom{b_{2n}}{a_{2k}} \right| = \left| J\binom{b_{2n+1}}{a_{2k+1}} \right| = 1. \tag{4.5}$$

At this point we shall restrict our discussion to Hamiltonians which are invariant with respect to space inversion (i.e., all $a_{2j+1} \equiv 0$). Equation (4.2) can then be rewritten as

$$\mathrm{H} = -A[(\beta^+)^2 - 2\beta^+\beta + \beta^2] + \sum_{n=0}^{[N/2]}\sum_{k=0}^{2n}\binom{2n}{k} b_{2n}(\beta^+)^{2n-k}\beta^k. \tag{4.6}$$

If now we use Eq. (2.8) to evaluate the corresponding matrix elements, we find that

$$H_{j,\,j+2k+1} = 0, \quad j = 0, 1, 2, \ldots, M, \quad k = 0, 1, \ldots, [(M-j-1)/2], \tag{4.7}$$

and

$$H_{j,\,j+2k} = -A[\sqrt{(j+1)(j+2)}\,\delta_{1k} - 2j\delta_{0k}] + \sum_{m=0}^{j+k}\binom{2m}{m+k}\frac{[j!\,(j+2k)!\,]^{1/2}}{[j+k-m]!}\, b_{2m}, \tag{4.8}$$

$$j = 0, 1, \ldots, M, \quad k = 0, 1, \ldots, [(M-j)/2],$$

where $b_{2m} = 0$ if $2m > N$. Note that since the matrix is real and symmetric, only the diagonal and superdiagonal matrix elements have been given explicitly.

As above, if $N > 2M$ one has superfluous b_{2n}'s which do not contribute to any matrix element. Further, if $N = 2M$ there are only $M+1$ of the b_{2n}'s. Thus, it is impossible for the H_{ij}, $j \geq i$, to be functionally independent. However, as is shown below, it is possible to choose a subset of $M+1$ of the H_{ij}'s which are independent. In particular, the diagonal elements

$$H_{jj} = 2jA + \sum_{m=0}^{j}\binom{2m}{m}\frac{j!}{(j-m)!}\, b_{2m}, \tag{4.9}$$

form a functionally independent set.

Note that if we define

$$v_j \equiv (H_{jj} - 2jA)/j! \tag{4.10}$$

and

$$u_m \equiv \binom{2m}{m} b_{2m}, \tag{4.11}$$

then Eq. (4.9) becomes

$$v_j = \sum_{m=0}^{j}\frac{1}{(j-m)!}\, u_m. \tag{4.12}$$

This is of course exactly the same transformation as that given by Eq. (3.8).

Thus from Eq. (3.10) we have

$$\binom{2m}{m} b_{2m} = \sum_{j=0}^{m}\frac{(-1)^{m-j}}{(m-j)!\,j!}\,(H_{jj} - 2jA). \tag{4.13}$$

If now we insert this into Eq. (4.8), we obtain

$$H_{j\,j+2k} = -A\sqrt{(j+1)(j+2)}\,\delta_{1k} + \sum_{m=0}^{j+k} B_{jkm}(H_{mm} - 2mA), \quad k > 0, \tag{4.14}$$

where

$$B_{jkm} \equiv \sqrt{j!\,(j+2k)!}\sum_{l=m}^{j+k}\frac{(-1)^{l-m}(l!)^2}{(l-k)!\,(l+k)!\,(l-m)!\,m!}. \tag{4.15}$$

Thus, the off-diagonal elements can all be written explicitly as linear combinations of the diagonal elements.[12] Further, it can be shown that this is true in almost any representation, not just for the particular basis we have chosen.[13]

Clearly, if $N \geq 2M$ and the a_{2n}'s are independent, the matrix element distribution can be expressed as a joint distribution containing only the diagonal elements. An obvious question which arises is whether the various statistical properties of the eigenvalues are affected by the fact that the multivariate distribution does not involve all of the matrix elements (i.e., how does the functional dependence of the matrix elements manifest itself in the statistical properties?)

To study the statistical properties of the eigenvalues one must first obtain the joint distribution of the eigenvalues. Immediately, there is the question of whether or not the eigenvalues are functionally independent. Appendix I contains a proof of that independence. However, it has not proven feasible to obtain the joint eigenvalue distribution explicitly when the dimensionality of the matrices is arbitrary. An indication of the nature of the problem will be given below by obtaining the solution for a low-dimensional case.

In keeping with the usual assumptions, we shall consider only ensembles of matrices which are representationally invariant. Thus, the ensemble will be of the form

$$P(H_{jj})d\mathrm{H} = f(\mathrm{Tr}\mathrm{H}^k)[\,|\det g|\,]^{1/2} d\mathrm{H}, \tag{4.16}$$

where $d\mathrm{H} \equiv \prod_{j=1}^{m} dH_{jj}$, where all traces are to be expressed in terms of the diagonal elements only, and where the symmetric matrix g is determined as follows. Define length in the matrix space as[14]

$$ds^2 \equiv \mathrm{Tr}(d\mathrm{H})^2 = \sum_{i,\,j=0}^{M}(dH_{ij})^2. \tag{4.17}$$

If now Eqs. (4.7) and (4.8) are used to express the dH_{ij}, $i \neq j$, in terms of the dH_{ii}, Eq. (4.17) can be put in the form

$$ds^2 = \sum_{i,j=0}^{M} g_{ij} dH_{ii} dH_{jj}, \tag{4.18}$$

where $g_{ij} = g_{ji}$. Equation (4.18) then defines the matrix g.

For definiteness, we shall restrict our discussion from this point to the special case $M = 2$. The Hamiltonian matrices are 3×3, and can be put in the form

$$\mathbf{H} = \begin{pmatrix} H_{11} & 0 & 0 \\ 0 & H_{00} & H_{02} \\ 0 & H_{02} & H_{22} \end{pmatrix}, \tag{4.19}$$

where

$$H_{02} = -2\sqrt{2}A + (\sqrt{2}/2)(H_{11} - H_{00}). \tag{4.20}$$

Note that the matrix has been blocked according to the parity of the states. Note, however, that the blocks corresponding to different parity are not statistically independent since H_{02} is a function of H_{11}. Thus, both blocks must be considered together.

It is easily shown that $|\det g| = 3$. Further, for definiteness, we shall assume that

$$f(\mathrm{Tr}\mathbf{H}^k) = (C/\sqrt{3}) \exp(-\beta \mathrm{Tr}\mathbf{H}^2), \tag{4.21}$$

where $-\infty < H_{ii} < +\infty$, $i = 0, 1, 2$, and where

$$C \equiv \sqrt{3}(\pi/\beta)^{3/2} \exp(16A^2\beta/3). \tag{4.22}$$

Thus,

$$P(H_{jj})\, d\mathbf{H} = C \exp(-\beta \mathrm{Tr}\mathbf{H}^2)\, dH_{00}\, dH_{11}\, dH_{22}, \tag{4.23}$$

where $\mathrm{Tr}H^2$ is to be expressed in terms of only the H_{jj}, using Eq. (4.20).

To obtain the joint eigenvalue distribution we must now make a change of variables to the three eigenvalues. It is easily shown that the eigenvalues are given by

$$E_0 = H_{11}, \tag{4.24}$$

$$E_\pm = \tfrac{1}{2}[(H_{00} + H_{22}) \pm \{(H_{00} - H_{22})^2 + 4H_{02}^2\}^{1/2}]. \tag{4.25}$$

Note that E_0 corresponds to an odd parity state, while the $E_\pm$ correspond to even parity states. The joint eigenvalue distribution is given by

$$p(E_0, E_+, E_-)\, dE_0\, dE_+ dE_- = CF(\mathbf{E}) \exp[-\beta(E_0^2 + E_+^2 + E_-^2)]\, dE_0\, dE_+ dE_-, \tag{4.26}$$

where

$$F(\mathbf{E}) \equiv \int_{-\infty}^{+\infty} dH_{00} \int_{-\infty}^{+\infty} dH_{11} \int_{-\infty}^{+\infty} dH_{22}\, \delta(E_0 - H_{11}) \times \delta(E_+ - E_+(H_{jj}))\, \delta(E_- - E_-(H_{ii})). \tag{4.27}$$

It can be shown by a straightforward calculation that

$$F(E) = \frac{\sqrt{2}(E_+ - E_-)}{\sqrt{(E_0' - E_-)(E_+ - E_0')}} H(E_+ - E_-) H(E_+ - E_0') H(E_0' - E_-), \tag{4.28}$$

where

$$E_0' \equiv E_0 - 4A \tag{4.29}$$

and

$$H(u) \equiv \begin{cases} 1, & u > 0, \\ 0, & u < 0. \end{cases} \tag{4.30}$$

Note that even though the three eigenvalues are functionally independent their *ranges* are not independent. In particular, note that if E_- and E_+ are fixed then for a nonzero probability, E_0 must be in the interval

$$E_- + 4A < E_0 < E_+ + 4A.$$

This of course is quite different from the results one obtains for ensembles of matrices containing no functional dependence among the matrix elements.

Since the various spacing distributions play a crucial role in the statistical theory of energy levels, let us next consider one of the spacing distributions for this ensemble. The spacing distribution for the two levels with even parity (i. e., E_+ and E_-) is defined by

$$P(S)\, dS = \left[\frac{1}{2} \int_{-\infty}^{+\infty} dE_0 \int_{-\infty}^{+\infty} dT\, p\left(E_0, \frac{T+S}{2}, \frac{T-S}{2}\right)\right] dS, \tag{4.31}$$

where

$$S \equiv E_+ - E_- \geq 0 \tag{4.32}$$

and

$$T \equiv E_+ + E_-. \tag{4.33}$$

If now Eqs. (4.26) and (4.28) are inserted into (4.31) it follows after some straightforward calculations that

$$P(S) = \{2\pi/(3\beta)\}^{1/2} CS \exp(-\beta S^2/2) G(S), \tag{4.34}$$

where

$$G(S) \equiv \int_{-S/2}^{S/2} [(u + S/2)(S/2 - u)]^{-1/2} \times \exp[-\tfrac{2}{3}\beta(u + 4A)^2]\, du. \tag{4.35}$$

Since $\lim_{s \to 0+} G(S) = \pi \exp(-32\beta A^2/3)$, for small spacings

$$P(S) \approx C'S \tag{4.36}$$

where C' is a constant. Thus, the repulsion of energy levels of like parity is still present.[15]

It seems likely that for larger dimensions the effects noted here will be present. In particular, even though the eigenvalue s are functionally independent their ranges will not be independent. Also, there is no reason not to expect the repulsion effect to remain.

5. THE EFFECT OF PARTICLE STATISTICS

We now consider a system of N identical particles. A general Hamiltonian for such a system can be written in the second quantized form as

$$\mathbf{H} = \sum_{l=0}^{L} \sum_{n_1=0}^{I} \cdots \sum_{m_K=0}^{I} C_{\mathbf{nm}}^{(l)} (\beta_1^\dagger)^{n_1} \cdots (\beta_K^\dagger)^{n_K} \times \beta_K^{m_K} \cdots \beta_1^{m_1} \delta_{\hat{n}l} \delta_{\hat{m}l}, \tag{5.1}$$

where $\mathbf{n} \equiv (n_1, n_2, \ldots, n_K)$ and $\hat{n} \equiv \sum n_i$ and where $\beta_i^\dagger$ and β_i are, respectively, the particle creation and annihilation operators for the ith single particle state. Note it has been assumed that there are only a finite number K of single particle states involved, and that the number of particles N is finite. (We shall assume that $K \geq 2N$.) The effect of these assumptions is to make the dimensionality of the Hamiltonian matrix finite. Note that the index l indicates the number of particles interacting and

also note that the maximum number L of interacting particles must be less than or equal to N. The commutation relations for the operators are

$$\beta_i\beta_j^+ \pm \beta_j^+\beta_i = \delta_{ij} \tag{5.2}$$

and

$$\beta_i\beta_j \pm \beta_j\beta_i = \beta_i^+\beta_j^+ \pm \beta_j^+\beta_i^+ = 0, \tag{5.3}$$

where the upper sign is to be used for fermions and the lower sign is to be used for bosons.

We shall assume that the parameters $C_{nm}^{(l)}$ are real.[16] Since **H** is Hermitian, it then follows that $C_{n,m}^{(l)} = C_{m,n}^{(l)}$ (i.e., for each l, $C_{nm}^{(l)}$ is symmetric). It follows that an ensemble of Hamiltonians will be completely specified if the joint distribution for the diagonal elements and the superdiagonal elements of the $C_{nm}^{(l)}$ are given.

The different commutation relations of the β's for fermions and bosons cause the matrix elements of **H** to be quite different. Thus, we will consider the two cases separately.

Since the commutation relations of the boson creation and annihilation operators are identical to those of the ladder operators given above, the calculation of matrix elements in this case is identical to the one given above for obtaining (3.12). The result is

$$H_{pq} = \sum_{l=0}^{L} \sum_{n_1=0}^{l} \cdots \sum_{n_K=0}^{l} C_{\mathbf{n};\mathbf{n}+\mathbf{q}-\mathbf{p}}^{(l)} \prod_{i=1}^{K} \frac{\sqrt{p_i!q_i!}}{(p_i-n_i)!}, \quad \delta_{l\hat{n}}\delta_{N\hat{p}}\delta_{N\hat{q}}, \tag{5.4}$$

where p_i and q_i are the number of particles in the ith state of the bra and ket vectors, respectively, and where

$$C_{\mathbf{n};\mathbf{n}+\mathbf{q}-\mathbf{p}}^{(l)} \equiv 0 \quad \text{if any } n_i + q_i - p_i < 0.$$

If the number of particles is larger than the maximum order of the interaction (i.e., $N > L$) there will be matrix elements which are identically zero. For example, any element with $p_i = N$ and $q_j = N$, $i \neq j$, is easily seen to vanish. Thus, if $N > L$ the matrix elements (diagonal and superdiagonal) certainly can not be treated as independent variables. To illustrate the effects on the matrix elements if $N > L$ let us consider the case when $L = 2$ (i.e., two-body interactions at most).

From Eq. (5.4) it follows directly that the only nonzero matrix elements are

$$H_{\mathbf{p};\mathbf{p}} = C_{0;0}^{(0)} + \sum_{k=1}^{K} C_{1_k;1_k}^{(1)} p_k + \sum_{k=1}^{k} C_{2_k;2_k}^{(2)} p_k(p_k-1) + \sum_{k>l} C_{1_k+1_l;1_k+1_l}^{(2)} p_k p_l, \tag{5.5}$$

$$H_{\mathbf{p};\mathbf{p}+1_k-1_l} = C_{1_l;1_k}^{(1)}\sqrt{p_l(p_k+1)} + C_{2_l;1_l+1_k}^{(2)}\sqrt{p_l(p_l-1)^2(p_k+1)} + C_{1_l+1_k;2_k}^{(2)}\sqrt{p_l p_k^2(p_k+1)} + \sum_{\substack{m\neq l \\ m\neq k}} C_{1_l+1_m;1_k+1_m}^{(2)}\sqrt{p_l(p_k+1)p_m^2}, \quad k \neq l, \tag{5.6}$$

$$H_{\mathbf{p};\mathbf{p}+2_k-2_l} = C_{2_l;2_k}^{(2)}\sqrt{p_l(p_l-1)(p_k+2)(p_k+1)}, \quad k \neq l, \tag{5.7}$$

$$H_{\mathbf{p};\mathbf{p}+2_k-1_l-1_m} = H_{\mathbf{p}+2_k-1_l-1_m;\mathbf{p}} + C_{1_l+1_m;2_k}^{(2)}\sqrt{p_l p_m(p_k+2)(p_k+1)}, \tag{5.8}$$

$$H_{\mathbf{p};\mathbf{p}+1_k+1_l-1_m-1_n} = C_{1_m+1_n;1_k+1_l}^{(2)}\sqrt{p_m p_n(p_k+1)(p_l+1)}, \tag{5.9}$$

where

$$\mathbf{p}+1_k = (p_1, p_2, \ldots, p_k+1, p_{k+1}, \ldots, p_K), \tag{5.10}$$

etc.

Note that each type of nonzero element depends on a different set of C's. Thus, any given matrix element is functionally independent of matrix elements of another type. However, for matrix elements of one type there are more matrix elements than C's. For example, the diagonal elements are functions of $(K+1)(K+2)/2$ distinct C's. Thus, all diagonal elements can be written in terms of a particular set of $(K+1)(K+2)/2$ diagonal elements.

The results for larger values of L will be similar. That is, some matrix elements will be identically zero and those which are not will not be functionally independent. As L increases, however, the number of zero matrix elements decreases and the number of functional relationships between the matrix elements decreases.

On the other hand, if $L = N$ the diagonal and superdiagonal elements can be used as independent variables provided the corresponding $C_{\mathbf{k},\mathbf{q}}^{(N)}$ are independent. This follows easily from (5.4). (i.e., each $C_{\mathbf{k},\mathbf{q}}^{(N)}$ occurs in one and only one of the elements of **H**.)

Consider now the case when the particles are fermions. The calculation of the matrix elements is analogous to the case of the bosons. We find that

$$H_{pq} = \sum_{l=0}^{L} \sum_{n_1=0}^{p_1} \cdots \sum_{n_K=0}^{p_K} {p_1}^{n_1} \cdots {p_k}^{n_K} {q_1}^{n_1+q_1-p_1} \cdots {q_K}^{n_K+q_K-p_K} \times C_{\mathbf{n};\mathbf{n}+\mathbf{q}-\mathbf{p}}^{(l)} \delta_{l\hat{n}}\delta_{N\hat{p}}\delta_{N\hat{q}} \tag{5.11}$$

where $p_i = 0, 1$, $q_i = 0, 1$, and $C_{\mathbf{n};\mathbf{n}+\mathbf{q}-\mathbf{p}}^{(l)} \equiv 0$, if any $n_i + q_i - p_i < 0$. Note that since we are dealing with fermions the matrix elements are zero if any p_i or q_i is larger than unity. Thus, from this point we consider only matrix elements with $p_i = 0, 1$ and $q_i = 0, 1$, $i = 1, 2, \ldots, K$. Note that again if $N > L$ some of the matrix elements are identically zero. For example, if $p_i \neq q_i$, $i = i_1, i_2, \ldots, i_{2N}$, then $p_i = q_i = 0$ for all other i, and it follows easily that $H_{pq} = 0$. Therefore, the matrix elements cannot be used as independent variable. Again to illustrate the effect on the matrix elements we consider the case $L = 2$.

When $L = 2$ the only nonzero matrix elements are

$$H_{pp} = C_{0;0}^{(0)} + \sum_{k} (C_{1_k;1_k}^{(1)} + C_{2_k;2_k}^{(2)})\delta_{1p_k} + \sum_{k>l} C_{1_k+1_l;1_k+1_l}^{(2)}\delta_{1p_k}\delta_{1p_l}, \tag{5.12}$$

$$H_{pp+1_k-1_l} = \delta_{0p_k}\delta_{1p_l}\left(C_{1_l;1_k}^{(1)} + \sum_{\substack{m\neq k \\ m\neq l}} C_{1_l+1_m;1_k+1_m}^{(2)}\delta_{1p_m}\right), \tag{5.13}$$

and

$$H_{pp+1_k+1_l-1_m-1_n} = \delta_{0p_r}\delta_{0p_l}\delta_{1p_m}\delta_{1p_n} C_{1_m+1_n;1_k+1_l}^{(2)}. \tag{5.14}$$

The three classes of nonzero matrix elements are independent of each other. However, in each class only a proper subset of matrix elements can be taken as independent variables.

The results for $L > 2$ are similar. However, as L in-

creases the number of types of nonzero matrix elements increases and the size of the proper subset of independent variables for each type increases.

As was true for bosons, if $L=N$ the diagonal and superdiagonal elements can be used as independent variables provided the $C_{pq}^{(N)}$ are independent. That is each $C_{pq}^{(N)}$ occurs only in the element H_{pq}.

Finally, it should be noted that if the system is a mixture of bosons and fermions we could choose as our basis products of the antisymmetrized state functions for the fermion and symmetrized state function for the bosons. The details of the calculation are somewhat different than those above but the conclusions are basically the same.

APPENDIX: PROOF THAT THE EIGENVALUES CORRESPONDING TO A RANDOM HAMILTONIAN MATRIX, WITH A FUNCTIONAL DEPENDENCE AMONG MATRIX ELEMENTS GIVEN BY (4.14), CAN BE USED AS INDEPENDENT VARIABLES

The procedure will be to show that there exists a region in the matrix element space where

$$J\begin{pmatrix}E_i\\H_{jj}\end{pmatrix}\neq 0. \tag{A1}$$

The region is selected such that nondegenerate perturbation theory is applicable.

In particular, let

$$\mathbf{H}=\mathbf{H}^0+\mathbf{V} \tag{A2}$$

where

$$H^0_{mm}=2mA, \quad m=0,1,\ldots,4P-1, \tag{A3}$$

and where

$$H^0_{jj+2}=H^0_{j+2j}=-A\sqrt{(j+1)(j+2)}, \tag{A4}$$

are the only nonzero off diagonal elements. Note that the matrix elements of $\mathbf{H}_0$ satisfy (4.14). Thus, the matrix elements of $\mathbf{V}$ (which will be treated below as a perturbation) must also satisfy (4.14). Note also, that we have chosen $M=4P-1$ for simplicity (i.e., $\mathbf{H}$ is $4P\times 4P$).

Clearly, $\mathbf{H}^0$ can be written in the form

$$\mathbf{H}=\begin{pmatrix}H^0_0 & & & & & & \\ & H^0_1 & & & & & \\ & & H^0_4 & & & & \\ & & & H^0_5 & & & \\ & & & & \cdot & & \\ & & & & & \cdot & \\ & & & & & & H^0_{4P-3}\end{pmatrix} \tag{A5}$$

where

$$\mathbf{H}^0_r=\begin{pmatrix}H^0_{rr} & H^0_{r\,r+2}\\ H^0_{r\,r+2} & H^0_{r+2\,r+2}\end{pmatrix}. \tag{A6}$$

If we let $\mathbf{A}^0$ be defined by

$$\mathbf{E}^0=\tilde{\mathbf{A}}^0\mathbf{H}^0\mathbf{A}^0, \tag{A7}$$

where $\mathbf{E}^0$ is the diagonal matrix with the eigenvalues of $\mathbf{H}^0$ on the diagonal, it is clear that

$$\mathbf{A}^0_0=\begin{pmatrix}\mathbf{T}^0_0 & & & & & & \\ & \mathbf{T}^0_1 & & & & & \\ & & \mathbf{T}^0_4 & & & & \\ & & & \cdot & & & \\ & & & & \cdot & & \\ & & & & & \cdot & \\ & & & & & & \mathbf{T}^0_{4P-3}\end{pmatrix}, \tag{A8}$$

where

$$\begin{pmatrix}E^0_r & 0\\ 0 & E^0_{r+2}\end{pmatrix}=\tilde{\mathbf{T}}^0_r\mathbf{H}^0_r\mathbf{T}^0_r, \quad r=0,1,4,5,\ldots,4P-4,4P-3. \tag{A9}$$

If now we let

$$\mathbf{T}^0_r=\begin{pmatrix}a_r & -b_r\\ b_r & a_r\end{pmatrix}, \tag{A10}$$

it follows easily that

$$a_r=(1+S_r^{-1})/\sqrt{2}, \tag{A11}$$

$$b_r=(1-S_r^{-1})/\sqrt{2}, \tag{A12}$$

$$E^0_r=2A[(r+1)-S_r], \tag{A13}$$

and

$$E^0_{r+2}=2A[(r+1)+S_r], \tag{A14}$$

where

$$S_r\equiv[(r+1)(r+2)+4]^{1/2}/2. \tag{A15}$$

It is easily verified that none of the $4P$ eigenvalues are equal (i.e., there are no degeneracies). Thus, we can use nondegenerate perturbation theory to approximate the eigenvalues of $\mathbf{H}$ (provided of course $\mathbf{V}=\mathbf{H}-\mathbf{H}^0$ can be treated as a perturbation).

If we let

$$V_{mm}=\epsilon_m, \tag{A16}$$

(where the ϵ_m are as small in absolute value, as need be for perturbation theory to hold) it follows from (4.14) that the only nonzero superdiagonal element of V are

$$V_{jj+2k}=\sum_{m=0}^{j+k}B_{jkm}\epsilon_m. \tag{A17}$$

From nondegenerate perturbation theory it follows that

$$E_m=E^0_m+u_m+O(\epsilon^2), \quad m=0,1,\ldots,4P-1, \tag{A18}$$

where the E_m are the eigenvalues of $\mathbf{H}$, and the u_m are the diagonal elements of $\mathbf{V}$ in the representation where $\mathbf{H}^0$ is diagonal, i.e.,

$$u_m=\sum_{k,n}A^0_{km}A^0_{nm}V_{kn}. \tag{A19}$$

Clearly, since $\mathbf{A}^0$ and the E^0_m are constant,

$$J\begin{pmatrix}E_i\\H_{jj}\end{pmatrix}=J\begin{pmatrix}u_i\\\epsilon_j\end{pmatrix}+O(\epsilon), \tag{A20}$$

and we need only show that

$$J\begin{pmatrix}u_i\\\epsilon_j\end{pmatrix}\neq 0.$$

It can be shown directly using (A8), (A10), (A17), and (A19) that

$$u_{4r} = (1 - \delta_{0r})v_{4r} + \alpha_{4r}\epsilon_{4r} + \gamma_{4r}\mu_{4r,4r+1}\epsilon_{4r+1} + b_{4r}^2\epsilon_{4r+2}, \quad \text{(A21)}$$

$$u_{4r+2} = -(1 - \delta_{0r})v_{4r} + \beta_{4r}\epsilon_{4r} - \gamma_{4r}\mu_{4r,4r+1}\epsilon_{4r+1} + a_{4r}^2\epsilon_{4r+2}, \quad \text{(A22)}$$

$$u_{4r+1} = (1 - \delta_{0r})v_{4r+1} + \gamma_{4r+1}\mu_{4r+1,4r}\epsilon_{4r} + \alpha_{4r+1}\epsilon_{4r+1} + \gamma_{4r+1}\mu_{4r+1,4r+2}\epsilon_{4r+2} + b_{4r+1}^2\epsilon_{4r+3}, \quad \text{(A23)}$$

$$u_{4r+3} = -(1 - \delta_{0r})v_{4r+1} - \gamma_{4r+1}u_{4r+1,4r}\epsilon_{4r} + \beta_{4r+1}\epsilon_{4r+1} - \gamma_{4r+1}\mu_{4r+1,4r+2}\epsilon_{4r+2} + a_{4r+1}^2\epsilon_{4r+3}, \quad \text{(A24)}$$

where $r = 0, 1, \ldots, P-1$, and where

$$\mu_{jm} \equiv B_{j1m}, \quad \text{(A25)}$$

$$\gamma_s \equiv 2a_s b_s, \quad \text{(A26)}$$

$$\alpha_s \equiv a_s^2 + \gamma_s \mu_{ss}, \quad \text{(A27)}$$

$$\beta_s \equiv b_s^2 - \gamma_s \mu_{ss}, \quad \text{(A28)}$$

and

$$v_s \equiv \gamma_s \sum_{m=0}^{4r-1} \mu_{sm}\epsilon_m. \quad \text{(A29)}$$

It is easily seen from (A21)—(A24) and (A29) that

$$\left| J\begin{pmatrix} u_i \\ \epsilon_j \end{pmatrix} \right| = \prod_{r=0}^{P-1} |J_r| \quad \text{(A30)}$$

where

$$J_r \equiv J\begin{pmatrix} u_{4r} & u_{4r+1} & u_{4r+2} & u_{4r+3} \\ \epsilon_{4r} & \epsilon_{4r+1} & \epsilon_{4r+2} & \epsilon_{4r+3} \end{pmatrix}. \quad \text{(A31)}$$

It can be shown by direct calculation that

$$J_r = [2 - (R+1)(R^3 + 5R^2 + 4R - 6)]/D_r \quad \text{(A32)}$$

where $R = 4r$, and

$$D_r = 2[\{(R+1)(R+2)+4\}\{(R+2)(R+3)+4\}]^{1/2}. \quad \text{(A33)}$$

It is easily verified that $J_0 < 0$, while $J_r > 0$, $r > 0$. Thus, clearly

$$J\begin{pmatrix} E_i \\ H_{jj} \end{pmatrix} \not\equiv 0$$

in the region we have considered, and the eigenvalues are functionally independent.

*Supported in part by NRC Grant A4625.

[1]See for example, Ref. 2—6 below.
[2]J.B. French and S.S.M. Wong, Phys. Lett. B **33**, 449 (1970).
[3]O. Bohigas and J. Flores, Phys. Lett. B **34**, 261 (1971).
[4]J.B. French and S.S.M. Wong, Phys. Lett. B **35**, 5 (1971).
[5]O. Bohigas and J. Flores, Phys. Lett. B **35**, 383 (1971).
[6]S.S.M. Wong and J.B. French, Much. Phys. A **198**, 188 (1972).
[7]This is consistent with the numerical computations of French and Wong, Ref. 4 above.
[8]See for example, C.E. Porter and N. Rosenzweig, Ann. Acad. Sci. Fennicae: Ser. AVI, No. 44 (1960).
[9]Clearly the representation chosen should not be one where the Hamiltonian matrix of the particular system of interest is diagonal (or almost diagonal).
[10]See, for example, R.H. Dicke and J.P. Wittke, *Introduction to Quantum Mechanics* (Addison-Wesley, Reading, Mass. 1960).
[11]If the Hamiltonian is given in the form $H = \Sigma_{n,k} a_{nk} x^n p^k$, it can always be put in the form (3.1) by using (2.4—(2.6).
[12]We might add that if we do not impose the condition of space inversion invariance there will be no change in Eq. (4.15), but there will be a similar set of equations which give the matrix elements $H_{i,i+2k+1}$ as linear combinations of the $H_{i,i+1}$.
[13]It is easily shown that the Jacobian of the transformation from the diagonal elements in one representation to the diagonal elements in another representation vanishes only for a set of points of measure zero in the space of rotations. Thus, the diagonal elements are "good" variables in almost any representation.
[14]This approach is due to Porter. See the introductory review of C.E. Porter, *Statistical Theories of Spectra Fluctations* (Academic, New York, 1965).
[15]See, for example, Ref. 8 above.
[16]Since every matrix in our ensemble is then real (as shown below), we are in general considering only time-reversal invariant systems. In addition, if the total angular momentum of the system is half-integer, we are considering only rotationally invariant systems. See introductory review of Ref. 14.

Lattice approximation of the $(\lambda\phi^4 - \mu\phi)_3$ field theory in a finite volume

Yong Moon Park*

Mathematical Sciences Department, IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598
(Received 26 July 1974)

For space–time regions with finite volumes, the lattice cutoff unnormalized Schwinger functions converge to the (unnormalized) Schwinger functions as the lattice cutoff is removed.

1. INTRODUCTION

We study the lattice approximation of the $(\lambda\phi^4 - \mu\phi)_3$ field theory in a finite volume. The main progress on the ϕ_3^4 model has been the proof of the existence[1] and semiboundedness[2] of the spatially cutoff Hamiltonian, and the proof of the convergence[3] of the ultraviolet cutoff Schwinger functions as the ultraviolet cutoff is removed. Recently, Guerra, Rosen and Simon[4] have shown how the $p(\phi)_2$ field theory can be approximated by a general Ising model, which they called the "lattice approximation." As a consequence of this approximation, one has been able to apply results of the classical lattice theory to the $p(\phi)_2$ field theory.[4,5,6] The main purpose of this work is to establish the convergence of the lattice cutoff Schwinger functions of the ϕ_3^4 model. Then one may develop the ϕ_3^4 field theory parallel to the $p(\phi)_2$ model.

We are concerned with the Euclidean approach to the $(\lambda\phi^4 - \mu\phi)_3$ theory for $\lambda \in R^+$, $\mu \in R$. Let $\phi_\delta(f)$, $f \in \mathcal{S}(R^3)$, be the Euclidean free field with the lattice cutoff δ and let dq_0 be the free measure.[4] Then the lattice cutoff partition function Z_δ and the lattice cutoff unnormalized Schwinger functions S_δ^{un} are given by

$$Z_\delta(g) = \int \exp[-V(\delta, \lambda g) + \mu\phi_\delta(g)]\, dq_0,$$

$$S_\delta^{un}(g; f_1, \ldots, f_n) = \int \phi_\delta(f_1), \ldots, \phi_\delta(f_n) \times \exp[-V(\delta, \lambda g) + \mu\phi_\delta]\, dq_0,$$

where $\exp[-V(\delta, \lambda g) + \mu\phi_\delta(g)]\, dq_0$ is the interaction measure of the $(\lambda\phi^4 - \mu\phi)_3$ model, $\lambda \in R^+$, $\mu \in R$, with the lattice cutoff δ and the spatial cutoff g. See Sec. 2 for the detailed definition of the interaction measure. Throughout this paper we assume that the spatial cutoff g is the product of a positive function in $C_0^\infty(R^3)$ with the characteristic function of a union of unit lattice cubes. We also assume that $\lambda \in R^+$ and $\mu \in R$.

Theorem 1.1: (a) Let $f_i \in \mathcal{S}(R^3)$. There are constants $K_1(\lambda, \mu, f_i\text{'s})$ and $K_2(\lambda, \mu)$ independent of δ and g such that

$$|S_\delta^{un}(g; f_1, \ldots, f_n)| \leq K_1 \exp[K_2 A(g)],$$

where $A(g)$ is the volume of the set of points within a distance unity from the support of g.

(b) As $\delta \to 0$, $Z_\delta(g) \to Z(g)$ and

$$S_\delta^{un}(g, f_1, \ldots, f_n) \to S^{un}(g; f_1, \ldots, f_n),$$

where $Z(g)$ and $S^{un}(g; f_1, \ldots, f_n)$ are the partition function and the unnormalized Schwinger functions.

(c) Let $\mu \in \mathbb{C}$. Then the results of (a) and (b) are uniform with respect to μ on a compact subset of $\mathbb{C}$.

Remark 1.1: Feldman[3] has proven the existence of $Z(g)$ and $S^{un}(g; f_1, \ldots, f_n)$ for the case of $\mu = 0$. For an arbitrary $\mu \in R$ the existence follows immediately from Ref. 3, Theorem 3, by setting $f = \mu g$.

Remark 1.2: For small coupling constant λ (depending on g) the lattice cutoff normalized Schwinger functions $S_\delta(g; g_1, \ldots, g_n) = Z_\delta(g)^{-1} S_\delta^{un}(g; f_1, \ldots, f_n)$ converge to the normalized Schwinger functions $S(g; f_1, \ldots, f_n) = Z(g)^{-1} S^{un}(g; f_1, \ldots, f_n)$ as $\delta \to 0$. The above assertion follows from Theorem 1.1 and from the fact that $Z(g) \neq 0$ for small coupling constant λ.

Instead of proving Theorem 1.1 directly, we will introduce the ultraviolet cutoff κ in Z_δ and S_δ^{un} to get the triple cutoff partition function $Z_{\delta,\kappa}$ and Schwinger functions $S_{\delta,\kappa}^{un}$. For a given lattice cutoff δ we will show the uniform convergence of $Z_{\delta,\kappa}$ and $S_{\delta,\kappa}^{un}$ as the ultraviolet cutoff is removed (Theorem 3.1). The basic idea of our proof for the uniform convergence is essentially the same as that of Glimm and Jaffe[2] or Feldman.[3] The theorem will follow as a consequence of Theorem 3.1.

In Sec. 2 we introduce the definition of the lattice cutoff interaction measure for the interaction $(\lambda\phi^4 - \mu\phi)_3$, $\lambda > 0$, $\mu \in R$. We also establish technical lemmas which we need in Sec. 3. Section 3 is devoted to proving Theorem 1.1 by showing the uniform convergence of $Z_{\delta,\kappa}$ and $S_{\delta,\kappa}$. In the Appendix we give estimates of small graphs.

Readers are referred to Refs. 2, 3, and 4 for further background material, notation, and references.

We conclude this section by proposing two problems. The applications of the lattice approximation to the $P(\phi)_2$ field theory are frequently dependent on boundary conditions (for example, the Dirichlet boundary condition).[4,6,7] Notice that the results of this paper hold for the free boundary condition. The first problem is the convergence of the lattice cutoff normalized Schwinger functions with the free boundary condition and the second is the extension to the Dirichlet boundary condition. That is, we wish to prove:

(i) For the free boundary condition the lattice cutoff normalized Schwinger functions $S_\delta(g; f_1, \ldots, f_n)$ converge to the normalized Schwinger functions $S(g; f_1, \ldots, f_n)$ as $\delta \to 0$.

(ii) Let $Z_\delta^D(g)$ and $S_\delta^D(g; f_1, \ldots, f_n)$ be the corresponding partition function and Schwinger functions with the Dirichlet boundary condition. Then $Z_\delta^D(g) \to Z^D(g)$ and $S_\delta^D(g; f_1, \ldots, f_n) \to S^D(g; f_1, \ldots, f_n)$ as $\delta \to 0$.

If one is able to show that the partition function $Z(g)$ is

 Copyright © 1975 American Institute of Physics

not vanishing for $\mu = 0$ and $0 \leq \lambda \leq \lambda_0$, where λ_0 is a constant independent of g, it then follows that $Z(g) \neq 0$ for all $\mu \in \mathbb{C}$ and $0 \leq \lambda \leq \lambda_0$ by the Lee–Yang theorem for the lattice cutoff partition function together with Hurwitz's theorem.[6] Hence the first problem will be solved by Theorem 1.1. In this case, we conclude that the Lee–Yang theorem and any correlation inequalities, which hold for $\delta \neq 0$, hold for ϕ_3^4 field theory with the free boundary condition in a finite volume.

2. NOTATION, DEFINITIONS, AND BASIC ESTIMATES

We introduce the lattice cutoff interaction measure. The free Euclidean field theory is given on the path space $L^2(\mathcal{S}'_R(R^3), dq_0)$ where dq_0 is the Gaussian measure with mean zero and covariance $(-\Delta + m^2)^{-1}$. The Euclidean fields are the linear coordinate functions on $\mathcal{S}'_R(R^3)$: $\phi(f)(q) = \langle q, f\rangle$ for all $q \in \mathcal{S}'_R(R^3)$ and $f \in \mathcal{S}(R^3)$. Following Ref. 4, in Sec. IV we introduce the lattice approximation of the free fields. Let $\delta > 0$ be the spacing parameter for the lattice $L_\delta = \{n\delta : n = (n^{(0)}, n^{(1)}, n^{(2)}) \in Z^3\}$ in R^3. We approximate $-\Delta$ by

$$(-\Delta_\delta f)(n\delta) = \delta^{-2}\Big[6 f(n\delta) - \sum_{|n'-n|=1} f(n'\delta)\Big],$$

where $|n| = \sum_{i=0}^{2} |n^{(i)}|$. We consider the Fourier transformation from $l^2(\delta Z^3, \delta^3)$ to $L_2([-\pi/\delta, \pi/\delta]^3, d^3k)$ defined by

$$\tilde{h}_\delta(k) = \frac{\delta^3}{(2\pi)^{3/2}} \sum_{n \in Z^3} h(n\delta) \exp(-ik\cdot n\delta). \tag{2.1}$$

Then the image of $(-\Delta_\delta + m^2)$ on L^2 is mutliplication by

$$\mu_\delta(k)^2 = \delta^{-2}\Big(6 - 2\sum_{i=0}^{2} \cos(\delta k^{(i)})\Big) + m^2. \tag{2.2}$$

We define the lattice cutoff free fields and Wick product of the free fields by

$$\phi_\delta(n) = (2\pi)^{-3/2} \int_\delta \exp(-ik\cdot n\delta)\,[a(k)^* + a(-k)]\,\mu_\delta(k)^{-1} d^3k,$$

$$:\phi_\delta^m(n): = (2\pi)^{-3m/2} \int_\delta \exp(-i\delta n\cdot \Sigma k_j) \sum_{j=1}^{m} \binom{m}{j} a(k_1)^* \cdots a(k_j)^*$$

$$\times a(-k_{j+1}) \cdots a(-k_m) \prod_{i}^{m} \mu_\delta(k_i)^{-1} d^3k_i, \tag{2.3}$$

where $\int_\delta$ means that the range of integration is $|k^{(i)}| \leq \pi/\delta$. For $g \in L_2(R^3)$ we write

$$\phi_\delta(g) = \sum_{n \in Z^3} \delta^3 \phi_\delta(n)\, g(n\delta),$$

$$:\phi_\delta^m:(g) = \sum_{n \in Z^3} \delta^3 :\phi_\delta^m(n): g(n\delta). \tag{2.4}$$

The partition function Z_δ and unnormalized Schwinger functions S_δ^{un} of the lattice cutoff interaction theory are the mass and moments of the unnormalized lattice cutoff interaction measure $dq(\delta, \lambda g)$:

$$Z_\delta(g) = \langle 1\rangle_{dq(\delta,g)},$$
$$S_\delta^{un}(g; f_1, \ldots, f_n) = \langle \phi_\delta(f_1)\cdots\phi_\delta(f_n)\rangle_{dq(\delta,g)} \tag{2.5}$$

where $\langle M\rangle_{dq} = \int M\, dq$. The measure is given by

$$dq(\delta, g) = \exp[-V(\delta, \lambda g) + \mu\phi_\delta(g)]\, dq,$$
$$V(\delta, \lambda g) = V_I(\delta, \lambda g) + V_C(\delta, \lambda g),$$

$$V_I(\delta, \lambda g) = \lambda :\phi_\delta^4:(g),$$
$$V_C(\delta, \lambda g) = \tfrac{1}{2}\langle V_I(\delta, \lambda g)^2\rangle_{dq_0} - \tfrac{1}{6}\langle V_I(\delta, \lambda g)^3\rangle_{dq_0}$$
$$- \tfrac{1}{2}\lambda^2\, \delta m_\delta^2 :\phi_\delta^2:(g^2), \tag{2.6}$$
$$\delta m_\delta^2 = -4^2 \times 6 \times (2\pi)^{-9}$$
$$\sum_{\substack{n \in N^3 \\ n^{(i)} = 0, \pm 1}} \int_{Z_n(\delta)} \prod_{i=2}^{4} \mu_\delta\Big(k_i - \frac{1}{3}\zeta + \frac{2n\pi}{3\delta}\Big)^{-2} d_\zeta k,$$

where

$$\zeta = k_2 + k_3 + k_4, \quad Z_n(\delta) = \Big\{k_2, k_3, k_4 \Big|\Big| k_i^{(l)} - \frac{1}{3}\zeta + \frac{2n\pi}{3\delta}\Big| \leq \pi/\delta,$$
$$2 \leq i \leq 4,\ 0 \leq l \leq 2\Big\},$$

and the measure $d_\zeta k$ is proportional to Euclidean measure on the space $\zeta = 0$,

$$d\zeta\, d_\zeta k = d^3k_2\, d^3k_3\, d^3k_4.$$

The mass renormalization counter term is logarithmically divergent. We will discuss the counter term V_c in more detail in Appendix A3–A4.

We now establish technical lemmas which we will use in the next section. Let

$$\mu(k) = (k^2 + m^2)^{1/2}$$

where $k^2 = |k^{(0)}|^2 + |k^{(1)}|^2 + |k^{(2)}|^2$.

Lemma 2.1 (Adapted from Lemma IV.2 of Ref. 4):

(a) For each $k \in R^3$, $\mu_\delta(k) \to \mu(k)$ as $\delta \to 0$;

(b) if $|k^{(i)}| \leq \pi/\delta$, $0 \leq i \leq 2$, $\mu_\delta(k)^{-1} \leq \frac{1}{2}\pi\,\mu(k)^{-1}$.

We remark that Guerra, Rosen, and Simon[4] have proved the lemma for $k \in R^2$. The proof was based on the inequality

$$1 - \cos(y) \geq 2\pi^{-2}y^2 \quad \text{if } y \in [-\pi, \pi].$$

Hence the extension to the case of $k \in R^3$ is trivial.

Lemma 2.2: Let $f \in \mathcal{S}(R^3)$. Then for $|k^{(i)}| \leq \pi/\delta$

$$|\tilde{f}_\delta(k)| \leq O(1)\,\mu(k)^{-2}.$$

Proof: We note that by the definitions

$$|\mu_\delta(k)^2 \tilde{f}_\delta(k)| \leq 0(1)\,\delta \sum_{n \in Z^3} \Big|6f(n\delta) - \sum_{|n'-n|=1} f(n'\delta)\Big|.$$

Since

$$\Big|2f(n^{(0)}\delta, n^{(1)}\delta, n^{(2)}\delta) - f((n^{(0)}+1)\delta, n^{(1)}\delta, n^{(2)}\delta)$$
$$- f((n^{(0)}-1)\delta, n^{(2)}\delta, n^{(3)}\delta)\Big|$$
$$\leq \Big| \int_{(n^{(0)}-1/2)\delta}^{(n^{(0)}+1/2)\delta} \Big[\int_{y-\delta/2}^{y+\delta/2} \Big(\frac{\partial^2}{\partial x^2} f(x, n^{(1)}\delta, n^{(2)}\delta)\Big) dx\Big] dy \Big|$$
$$\leq \delta^2 \sup_{x \in [(n^{(0)}-1)\delta, (n^{(0)}+1)\delta]} \Big|\frac{\partial^2}{\partial x^2} f(x, n^{(1)}\delta, n^{(2)}\delta)\Big|$$
$$\leq \delta^2 \sup_{x^{(i)} \in [(n^{(i)}-1)\delta, (n^{(i)}+1)\delta]} \Big|\frac{\partial}{\partial x^{(0)}}^2 f(x^{(0)}, x^{(1)}, x^{(2)})\Big|,$$

we obtain

$$|\mu_\delta(k)^2\tilde{f}_\delta(k)| \le O(1)\delta^3 \sum_{n\in Z^3}\sum_{i=0}^{2} \sup_{x^{(i)}\in[(n^{(i)}-1)\delta,(n^{(i)}+1)\delta]} \left|\left(\frac{\partial}{\partial x^{(i)}}\right)^2 f(x)\right|.$$

Since $f\in\mathcal{S}(R^3)$, $|f''(x)| \le O(1)\prod_{i=0}^{2}[1+|x^{(i)}|]^{-m}$ for any $m>0$. Hence the above is bounded by

$$O(1)\delta^3 \sum_{n\in Z}\prod_{i=0}^{2}[1+|n^{(i)}\delta|]^{-m} \le O(1)$$

for $m\ge 2$. The lemma follows by Lemma 2.1(b) and the above bound. ∎

If one applies a method similar to that used above, one may show the stronger result: For any $m>0$

$$|\tilde{f}_\delta(k)| \le O(1)\,\mu(k)^{-m} \text{ for } |k^i| \le \pi/\delta.$$

In applications we only use the case $m=2$.

In the next section we will divide a bounded region into a union of cubes. To prevent double counts of lattice points on the boundary of cubes we consider half open cubes of the form

$$\Delta = \{x : x^i \in [a^{(i)}, b^{(i)}),\ |a^{(i)} - b^{(i)}| = d,\quad 0\le i\le 2\}. \tag{2.7}$$

We write

$$(h\chi_\Delta)\tilde{}_\delta(k) = \frac{\delta^3}{(2\pi)^{3/2}} \sum_{\substack{n\\ n\delta\in\Delta}} h(n\delta)\exp(-ik\cdot n\delta),$$

$$F_{\delta,\Delta}(k) = \prod_{l=0}^{2}[\,|\Delta|^{1/3}\mu_\delta(k^{(l)})+1]^{-1},\quad \mu_\delta(k^{(l)}) = \delta^{-2}[2-2\cos(\delta k^{(l)})]+m^2,$$

$$D_j^m = \prod_{l=0}^{2}\left(\frac{\partial}{\partial k_j^{(l)}}\right)^{m^{(l)}},\quad \sum_l m^{(l)} = |m|, \tag{2.8}$$

$$F_\Delta(k) = \prod_{l=0}^{2}(1+|\Delta|^{1/3}|k^{(l)}|)^{-1},$$

where $|\Delta|$ is the volume of Δ. Notice that $F_\Delta(k)$ is the comparison function introduced in (5.2.10) of Ref. 2.

Lemma 2.3: (a) For $|k^{(i)}| \le \pi/\delta$

$$|D^m\mu_\delta(k)^{-2}| \le O(1)\,\mu(k)^{-2-|m|};$$

(b) let $g\in\mathcal{S}(R^3)$ and let $k=\sum^P k_j$, $p\le 4$. We assume that $|\Delta|\le 1$ and the center of Δ is at the origin:

$$|D^m(g\chi_\Delta)\tilde{}_\delta(k)| \le O(1)\,|\Delta|^{1+|m|/3}F_{\delta,\Delta}(k);$$

(c) for $|k^{(i)}|\le\pi/\delta$ and $|k'^{(i)}|\le\pi/\delta$, and for $0\le\epsilon\le 1/3$,

$$|\mu_\delta(k)^{-2}-\mu_\delta(k')^{-2}| \le O(1)\,\mu(k-k')^{3\epsilon}\times[\mu_\delta(k)^{-2-\epsilon}+\mu_\delta(k')^{-2-\epsilon}].$$

Proof: (a) To show the method of proof we first consider the case where $D^m=(\partial/\partial k^{(0)})^2$. For this case we have

$$\left(\frac{\partial}{\partial k^{(0)}}\right)^2\mu_\delta(k)^{-2} = -2^2\,\mu_\delta(k)^{-6}\,[2\delta^{-1}\sin(\delta k^{(0)})]^2 + 4\,\mu_\delta(k)^{-4}\cos(\delta k^{(0)}).$$

The lemma follows from Lemma 2.1(b) and from the following inequalities:

$$|\delta^{-1}\sin(\delta k^{(i)})| \le O(1)|k^{(i)}|,$$

$$|\cos(\delta k^{(i)})| \le 1 \quad\text{and}\quad |k^{(i)}| \le \pi/\delta. \tag{2.9}$$

The lemma for the general case follows by a method similar to that used above. ∎

(b) Let

$$a_\delta(k) = 2\delta^{-1}\sin(\tfrac{1}{2}k\delta) = i\delta^{-1}[\exp(ik\delta/2)-\exp(ik\delta/2)].$$

It follows that $a_\delta(k)^2 = \delta^{-2}[2-2\cos(\delta k)]$ and so

$$|\mu_\delta(k^{(l)})| \le O(1)|a_\delta(k^{(l)})|. \tag{2.10}$$

By definition we have

$$\left|\prod_{i=0}^{2} a_\delta(k^{(i)})\,D^m(g\chi_\Delta)\tilde{}_\delta(k)\right| \le O(1)\left|\sum_{\substack{n\\ n\delta\in\Delta}} g(n\delta)\exp(-ik\cdot n\delta)\prod_{i=0}^{2}(n^{(i)}\delta)^{m^{(i)}} \times[\exp(i\tfrac{1}{2}k^{(i)}\delta)-\exp(-i\tfrac{1}{2}k^{(i)}\delta)]\right|. \tag{2.11}$$

We only consider the case of $g=1$ on the cube Δ. The simple modification of the method for the above case gives the lemma. Let $\partial\Delta$ be the set of points within a distance $\frac{1}{2}\delta$ from the boundary of Δ and let $\Delta_{\text{int}}=\Delta\sim\partial\Delta$. By rearranging the sum in (2.11) we bound (2.11) by

$$O(1)\prod_{i=0}^{2}\Bigg(\sum_{\substack{n^{(i)}\\ n\delta\in\partial\Delta}}|(n^{(i)}\delta)^{m^{(i)}}| + \sum_{\substack{n^{(i)}\\ n\delta\in\Delta}}|[(n^{(i)}+\tfrac{1}{2})\delta]^{m^{(i)}}-[(n^{(i)}-\tfrac{1}{2})\delta]^{m^{(i)}}|\Bigg)$$

$$\le O(1)\prod_{i=0}^{2}\left(|\Delta|^{m^{(i)}/3}+\left|\int_{n^{(i)}-\delta/2}^{n^{(i)}+\delta/2}\frac{\partial}{\partial x^i}(x^{m^{(i)}})\,dx^i\right|\right) \le O(1)|\Delta|^{|m|/3}.$$

On the other hand, one may check that $|D^m(gx_\Delta)\tilde{}_\delta(k)| \le O(1)|\Delta|^{1+|m|/3}$. From (2.10) and the above result we have proved the lemma for $g=1$ on Δ. By using a summation version of the integration by parts and using the above argument we proved the lemma. We leave the detailed proof to the reader. ∎

(c) We write

$$\mu_\delta(k)^{-2}-\mu_\delta(k')^{-2} = (\mu_\delta(k)\mu_\delta(k'))^{-2}(\mu_\delta(k')^2-\mu_\delta(k)^2),$$

$$\mu_\delta(k)^2-\mu_\delta(k')^2 = 2\delta^{-2}\sum_{i=0}^{2}(\cos(k'^{(i)}\delta)-\cos(k^{(i)}\delta)).$$

If $k'^{(i)}$ and $k^{(i)}$ have the same direction, then we have

$$\delta^{-2}|\cos(k'^{(i)}\delta)-\cos(k^{(i)}\delta)| \le O(1)\left|\int_{k'^{(i)}}^{k^{(i)}}\delta^{-1}\sin(\delta k)\,dk\right| \le O(1)|k'^{(i)}-k^{(i)}| \tag{2.12}$$

by (2.9). Since $\cos(\delta k)=\cos(-\delta k)$, the above bound holds for the case where $k'^{(i)}$ and $k^{(i)}$ have opposite directions. The lemma follows from (2.12).

3. CONVERGENCE OF THE LATTICE APPROXIMATION

In this section we prove Theorem 1.1. We first in-

troduce the ultraviolet cutoff in the interaction measure $dq(\delta, g)$. We define

$$dq(\delta, \kappa, g) = \exp[-V(\delta, \kappa, \lambda g) + \mu\phi_\delta(g)]\, dq_0,$$

$$dq(\delta, \kappa, \lambda g) = \exp[-V(\delta, \kappa, \lambda g)]\, dq_0, \tag{3.1}$$

where $V(\delta, \kappa, g)$ is defined by replacing $\mu_\delta(k)^{-1}$ by $\kappa(k)\mu_\delta(k)^{-1}$ in the definition of $V(\delta, g)$ in (2.6). We assume that the momentum cutoff $\kappa(k)$ has the form[2]

$$\kappa(k) = \prod_{i=0}^{2} \left[\eta\left(\frac{k^{(i)}}{\beta^{(i)}}\right) - \eta\left(\frac{k^{(i)}}{\alpha^{(i)}}\right)\right], \tag{3.2}$$

$$\alpha^{(i)} < \beta^{(i)},$$

$$\alpha^{(i)}, \beta^{(i)} \in \{M_0 = 0,\ M_j = M_1^{(1+\nu)^{j-1}} \text{ if } j \geq 1\},$$

where $M_1 > 1$ and $\nu > 0$ are constants given in Ref. 4 and η is a fixed $C_0^\infty(R^1)$ function satisfying

$$\eta(x) = 1 \text{ for } |x| \leq 1/2, \quad 0 < \eta(x) < 1 \text{ for } 1/2 < |x| < 2,$$

$$\eta(x) = 0 \text{ for } |x| \geq 2 \text{ and } \eta(x) = \eta(-x),$$

and by convention $\eta(k/0) = 0$. The triple-cutoff partition function and unnormalized Schwinger function are defined by

$$Z_{\delta,\kappa}(g) = \langle 1 \rangle_{dq(\delta,\kappa,g)},$$
$$S^{\mathrm{un}}_{\delta,\kappa}(g; f_1, \ldots, f_n) = \langle \phi_\delta(f_1) \cdots \phi_\delta(f_n) \rangle_{dq(\delta,\kappa,g)}. \tag{3.3}$$

Theorem 1.1 is a consequence of the following result:

Theorem 3.1: (a) Let $f_i \in \mathcal{S}(R^3)$. There are constants $k_1(\lambda, \mu, f_i, s)$ and $k_2(\lambda, \mu)$ independent of δ, κ, and g such that for $\beta^{(i)} \geq 2\pi/\delta$ and $\alpha^{(i)} = 0$, $i = 0, 1, 2$,

$$|S^{\mathrm{un}}_{\delta,\kappa}(g; f_1, \ldots, f_n)| \leq k_3 \exp[k_2 A(g)].$$

(b) As $\kappa \to 1$,

$$Z_{\delta,\kappa}(g) \to Z_\delta(g) \text{ uniformly in } \delta,$$

$$S^{\mathrm{un}}_{\delta,\kappa}(g; f_1, \ldots, f_n) \to S^{\mathrm{un}}_\delta(g; f_1, \ldots, f_n) \text{ uniformly in } \delta.$$

(c) Let $Z \in \mathbb{C}$. Then the results of (a) and (b) are uniform with respect to μ on a compact subset of $\mathbb{C}$.

Proof of Theorem 1.1: (a) This follows from the boundedness of Theorem 3.1(a) and the fact that $S^{\mathrm{un}}_{\delta,\kappa} = S^{\mathrm{un}}_\delta$ for $\beta^{(i)} \geq 2\pi/\delta$ and $\alpha^{(i)} = 0$, $i = 0, 1, 2$. (b) We note that

$$|S^{\mathrm{un}}_\delta - S^{\mathrm{un}}| \leq |S^{\mathrm{un}}_\delta - S^{\mathrm{un}}_{\delta,\kappa}| + |S^{\mathrm{un}}_{\delta,\kappa} - S^{\mathrm{un}}_\kappa| + |S^{\mathrm{un}}_\kappa - S^{\mathrm{un}}| \tag{3.4}$$

where $S^{\mathrm{un}}_\kappa = S^{\mathrm{un}}_{\delta=0,\kappa}$. By virtue of the momentum cutoff κ we have that for given κ and ϵ there exists a constant $a(\kappa, \epsilon)$ depending on κ and ϵ such that

$$|S^{\mathrm{un}}_{\delta,\kappa} - S^{\mathrm{un}}_\kappa| \leq \epsilon \text{ for } \delta \leq a(\kappa, \epsilon). \tag{3.5}$$

We use Theorem 3.1(b) and the result of Ref. 3 (see also remark in Introduction). The theorem follows from the 3ϵ argument. (c) This follows from Theorem 3.1(c) and a method similar to that used in the proof of Theorem 1.1(b). ∎

The rest of this paper is devoted to showing Theorem 3.1. The proof is based on the method developed in Ref. 2 (and Ref. 3). We organize the proof in the order of Ref. 3. Apparently, we have to show the necessary results corresponding to those of Ref. 3. We will adapt freely the notation and the results in Refs. 2, 3. Feldman[3] has introduced a norm on the kernel of Graph G: For $\gamma > 2\alpha > 0$,

$$\|G\|_{1,\gamma,\alpha} = \sup_{\rho^e_\alpha} \sup_{C} \| \rho^e_\alpha\, C\, M^\gamma |G| \|_{\mathrm{H.S.}}$$

ρ^e_α, C, M^γ, and $|\cdot|$ are "operations" that modify the Graph G and its kernel. See Ref. 3 for the detailed definitions. For convenience we include the characteristic function on the range of the integration into the kernel of the Graph G. From the definition we have

$$\|\phi_\delta(f)\|_{1,\gamma,\alpha} = (2\pi)^{-3/2} K_1 \|\mu^\gamma \mu_\delta^{-1} \tilde{f}_\delta \chi_\delta\|_2. \tag{3.6}$$
$$\equiv |f_\delta|_\gamma,$$

where $\chi_\delta(k)$ is the characteristic function on $\{k : k^{(i)} \in [-\pi/\delta, \pi/\delta],\ 0 \leq i \leq 2\}$ and K_1 is the constant given in Ref. 3. If $f \in \mathcal{S}(R^3)$, then it follows that $|f_\delta|_\gamma \leq$ const uniformly in δ for small γ by Lemmas 2.1, 2.2. Consequently, we obtain that for small γ and $f_i \in \mathcal{S}(R^3)$

$$\|\prod^n \phi_\delta(f_i)\|_{1,\gamma,\alpha} \leq \text{const uniformly in } \delta. \tag{3.7}$$

Here we have used the fact that $\|G\|_{\mathrm{H.S.}} \leq \Pi_j \|g_j\|_{\mathrm{H.S.}}$ if G is a union of subgraphs g_j.

We reduce the problem by a method similar to that used in Ref. 3. Theorem 3.1 follows immediately from Corollary 3.3 and Theorem 3.4 listed below, and from the result (3.7).

Theorem 3.2: Assume G is a graph having N external legs. Then there is a constant $K_4(\lambda, \mu, \gamma, \alpha)$ independent of δ and κ such that

$$|\langle G \exp[-V(\delta, \kappa, \lambda g)\rangle_{dq_0}| \leq N^N \|G\|_{1,\gamma,\alpha} \exp[K_4 A(g)].$$

Corollary 3.3 Under the assumption on κ as Theorem 3.1(a), there is $K_5(\lambda, \mu)$ such that

$$|\langle G \exp[\mu\phi_\delta(g)] \exp[-V(\delta, \kappa, \lambda g)\rangle_{dq_0} \leq N^N \|G\|_{1,\gamma,\alpha} \exp[K_5 A(g)].$$

The constant K_5 is uniformly bounded with respect to μ in a compact subset of $\mathbb{C}$.

Theorem 3.4: Let $\|G\|_{1,\gamma,\alpha} < \infty$ for some $0 < \gamma < \gamma_0$. Then as $\kappa \to 1$

$$\langle G \exp[\mu\phi_\delta(g)] \exp[-V(\delta, \kappa, \lambda g)]\rangle_{dq_0} \to \langle G\rangle_{dq(\delta,g)}$$

uniformly in δ,

and $\langle G\rangle_{dq(\delta,g)}$ obeys the bound of Corollary 3.3. The convergence is uniform in μ on a compact subset of C.

The above are the results corresponding to Ref. 3. Theorem 2, Corollary 2.1, and Theorem 3, respectively, for the case of $G_1 = G$, $G_2 = 1$, and $f = \mu g$.

Before proving the results we describe briefly the reason why the method used in Refs. 2, 3 can apply to our case. Roughly speaking, the authors of the cited papers have proved the results corresponding to Theorem 3.1—Theorem 3.4 by expanding $\langle G \exp[-V(\kappa, \lambda g)]\rangle_{dq_0}$ by a so-called "inductive construction."[2] Each inductive step consists of three main parts:

(a) a high momentum (P − C) expansion,

(b) a low momentum (Wick) expansion,

(c) combinatoric estimates.

Then the problem was reduced to estimates of the elementary integration labeled by Feynman graphs. The combinatoric estimates have been used to bound the number of terms in a sum of Feynman graphs. To estimate each term the majorizing functions for the kernels of P, C, and W vertices (except the mass renormalization cancellation) have been used. For the details we refer the reader to Refs. 2 and 3. By replacing $\exp[-V(\kappa,\lambda g)]\,dq_0$ and $\exp[\phi(f)]$ in Ref. 3 by $\exp[-V(\delta,\kappa,\lambda g)]\,dq_0$ and $\exp[\mu\phi_\delta(g)]$, we perform exactly the same inductive construction used in Ref. 3 (and Ref. 2). We also assign the same combinatoric factors for the combinatoric estimates. We then use Lemma 2.1—Lemma 2.3 to obtain majorizing functions for the kernels of P, C, and W vertices similar to those of Ref. 2. In the Appendix we estimate small graphs to obtain the same bounds as those of Ref. 2. Hence the upper bound of the elementary integration for each Feynman graph will turn out to be same as that in Refs. 2, 3.

We begin to show Theorem 3.2 with a few technical assumptions. We assume that the lattice spacing parameter has the form

$$\delta = 2^{-j}, \quad j \in Z^+.$$

We also assume that the allowed cubes have the form

$$\Delta = 2^{-j}\Delta_0 + 2^{-j}n\delta, \quad k \in Z^+, \quad n \in Z^3,$$

where $\Delta_0 = \{x : x^{(i)} \in [-\tfrac{1}{2}, \tfrac{1}{2})\}$. Theorem 3.2 for an arbitrary δ will follow from the method for $\delta = 2^{-j}$ with a modification of the allowed cubes. Following the main steps used in Refs. 2, 3 we summarize our proof.

Sketch of the Proof of Theorem 3.2

The main steps are as follows:

Step 1. *The Inductive Expansions*: We expand $\langle G\exp[-V(\delta,\kappa,\lambda g)]\rangle_{dq_0}$ by the inductive construction given in Ref. 2 Secs. 2, 3 (and in the proof of Ref. 3, Theorem 2). First we write

$$\langle G'\rangle_{dq(\delta,\kappa,\lambda g)} = N^N \|G'\|_{1,\gamma,\alpha} \langle G''\rangle_{dq(\delta,\kappa,\lambda g)} \tag{3.8}$$

where now $\|G''\|_{1,\gamma,\alpha} = N^{-N}$. We will show that for this new G''

$$|\langle G''\rangle_{dq(\delta,\kappa,\lambda g)}| \le \exp[KA(g)]. \tag{3.9}$$

When the inductive expansion is applied to $\langle G''\exp[-V(\delta,\kappa,\lambda g)\rangle_{dq_0}$ the G'' vertex remains completely passive (initiates no action) with one exception. The exception is the low-momentum contraction operation. Since

$$V(\delta,\kappa,\lambda g) = V(\delta,\kappa',\lambda g) \quad \text{for } \beta,\ \beta' \ge 2\pi/\delta, \tag{3.10}$$

where β and β' are the upper momentum cutoffs for κ and κ', respectively, we may assume that the upper momentum cutoff κ is smaller than $2\pi/\delta$. Undoing the Wick ordering it is easy to check that on the cube Δ

$$V(\delta,\kappa,\lambda g\chi_\Delta) \ge -O(1)\,|\Delta|M^2_{\mu(\Delta)},$$

where $M_{\mu(\Delta)}$ is the maximum upper momentum cutoff of κ. Hence, in the Wick construction, the exponent is completely removed when Δ is below the minimum size, i.e., when $|\Delta| \le (\delta/2\pi)^2 < M^{-2}_{\mu(\Delta)}$. From Proposition A1.2 and Lemma A3.1 in the Appendix, and from the method used in Ref. 2, Sec. 3.2, one may check that a lower bound of δV (the change of interation during the Wick construction) similar to that of Ref. 2 (3.2.5) holds for our case. Replacing ϕ_κ by $\phi_{\delta,\kappa}$ in Ref. 2, we apply the Wick construction same as Ref. 2, Sec. 3.2. We perform the low-momentum contraction operation on the G'' vertex according to the process given in the proof of Ref. 3, Theorem 2 (set $G_1 = G''$, $G_2 = 1$). After completing the inductive expansion, we obtain

$$|\langle G\rangle_{dq(\delta,\kappa,\lambda g)}| \le |\sum_G I(G)|, \tag{3.11}$$

where $I(G)$ is the elementary integration labeled by the Feynman Graph G. We note that $I(G)$'s can be obtained from those of Ref. 3 (and Ref. 2) by replacing

$$\left.\begin{matrix}\mu(k)^{-1}\\ (g\chi_\Delta)\tilde{}\,(k)\\ \int dk\\ \delta m^2_\kappa\end{matrix}\right\} \text{by} \left\{\begin{matrix}\mu_\delta(k)^{-1}\\ (g\chi\ )\tilde{}_\delta(k)\\ \int_\delta dk\\ \delta m^2_{\delta,\kappa}\end{matrix}\right. \quad \text{in } P,\ C, \text{ and } W \text{ vertices,} \tag{3.12}$$

where $\delta m^2_{\delta,\kappa}$, is defined by replacing μ_δ^{-1} by $\kappa(k)\mu_\delta^{-1}$ in the definition of δm^2_δ. $I(G)$ was written $T(\sigma)$ in Ref. 3, Lemma 5.1. For our case we set $G_1 = G''$ and $G_2 = 1$ in Ref. 3. The estimate leading to Theorem 3.2 is completed by first using the method of combinatoric factors to bound the number of terms in the sum of Graph G and then bounding the size of each term.

Step 2. *The Combinatoric Estimates*: We assert that the combinatoric bounds given in Ref. 2, Theorem 4.1 apply equally well to our case provided we include in addition a factor of

$$K'N\lambda^{8\epsilon} \quad \text{for each (external) } G'' \text{ leg.}$$

The above assertion follows from the argument used in the proof of Ref. 3, Lemma 4.1 on setting $G_1 = G''$ and $G_2 = 1$.

Step 3. *Localization Factors*: Following the method used in Ref. 2, Sec. 5.2, we isolate the distance factors from $I(G)$'s. We first divide cubes in P, C, and W vertices such that the conditions in Ref. 2, Sec. 5.2, are satisfied, and we translate each P, C, and W vertex to the origin. Then one obtains

$$I(G) = \sum_\alpha I(G_\alpha),$$

$$I(G_\alpha) = \int \int_{\delta \atop P,C,W \text{ vertices}} w(k) \prod_{\substack{\text{lines } l \\ \text{in } P,C, \text{ and } W \text{ vertices}}} \times \exp(ik_l \cdot n_{l\alpha}\delta)\,dk,$$

where $n_{l\alpha}\delta$ is the vector displacement of the lth line between the center of the cubes it connected. For more details we refer to Ref. 2, Sec. 5.2. We first note that for any momentum cutoff κ_1 and κ_2

$$\mu_\delta(k)^{-2}\kappa_1(k)\kappa_2(k)\exp(-ik\cdot n\delta)\Big|_{-\pi/\delta}^{\pi/\delta} = 0 \quad \text{for any } n \in Z^3.$$

(3.13)

We use (3.13) and integration by parts to get

$$\int\int_{\delta} w(k) \prod_{\text{lines } l} (\Delta_l^2)^{m_l} \exp(ik_l \cdot n_{l\alpha}\delta)\, dk$$

$$= \int\int_{\delta} \prod_{\text{lines } l} (\Delta_l^2)^{m_l} w(k) \exp(ik_l \cdot n_{l\alpha}\delta)\, dk, \tag{3.14}$$

where Δ_l is the Laplacian $\sum_{i=0}^{2}(\partial/\partial k_l^{(i)})^2$. If we use Lemma 2.1(b) and Lemma 2.3(b) on the range of integration, and if we use (3.14) and the method of Ref. 2, Sec. 5.2 and Sec. 6.2, then we may obtain

$$|I(G)| \leq \left(\prod_{\substack{\text{lines } l \\ \text{connected to } P,C,W \text{ vertices}}} (\alpha_1 d_l^{-m}) \right) \bar{I}(G), \tag{3.15}$$

$$I(G) = \int \bar{w}(k)\, dk,$$

where $\bar{w}$ majorizes w. Here $\bar{w}(k)$ is a product of the following factors: (i) the absolute value of the kernel of the G'' vertix in $w(k)$; (ii) a $\mu(k)^{-2}$ for each factor $\mu_\delta(k)^{-2}$ in P, C, and W vertices in $w(k)$; (iii) $|\Delta| F_{\delta,\Delta}(q)$ bounding each $(g\chi_\Delta)^{\sim}_\delta(q)$ in P and C vertices in $w(k)$; (iv) the bound in (A3.2) in Appendix for $|\nu(k)|$ from the W vertix in $w(k)$; (v) the characteristic function on the support $w(k)$ with respect to lines connected to P, C, and W vertices; (vi) any constants not absorbed in combinatoric factors. We note that $\bar{I}(G)$ is dependent on the lattice spacing parameter δ. The following step gives the uniform estimates of $\bar{I}(G)$.

Step 4. *Estimate of* $\bar{I}(G)$: We assert that $|\bar{I}(G)|$ is bounded above by a product of factors given by those of Ref. 2, Sec. 5.3 and

$\|G''\|_{1,\gamma,\alpha}$ per G'' graph,

$K''\lambda^{-\gamma/4}$ per G'' leg,

where $K'' = K_{11}K_1^{-1}$, γ and α are given in Ref. 3, Lemma 5.1. The bound of (3.19) follows from (3.15), the assertion and the method in the proof of Ref. 3, Theorem 3. We now prove our assertion. We write

$$|\bar{I}(G)| \leq \int \prod_{\text{lines } l} d^3k_l \prod_{\substack{G'' \\ \text{graphs } p}} |G''((k_p))\kappa_p((k))_p| \prod_{\substack{P,C,W \\ \text{vertices}}} \cdots, \tag{3.16}$$

where the $\kappa(k)$ are momentum cutoffs that were introduced into the G'' graph during the low-momentum operation.

We now use the method of decomposing big graphs to estimate $\bar{I}(G)$. We adapt the decomposing process of Refs. 2, 3. $\bar{I}(G)$ is the union of the disjoint subgraphs that consist of

(1) a single G'' graph,

(2) a single W vertex,

(3) a P vertex and the C vertices it generated.

The estimates of the P, C, and W graphs are given in Propositions A.1.1—5 in the Appendix, which are the same as those of Ref. 2, Sec. 5. We note that the expression (3.16) is of the form in the proof of Ref. 3, Lemma 5.1, for the case of $G_1 = G''$ and $G_2 = 1$. Hence the proof of our assertion follows from a simple modification of that of Ref. 3, Lemma 5.1, and from Propositions A1.1—5 [and also the method used in proving Proposition A1.5(b)].

Proof of Corollary 3.3: We first assert that there is a constant $K_5(\lambda, \mu)$ independent of δ, κ and g such that

$$|\langle G \exp[-V(\delta,\kappa,\lambda g) + \mu\phi_{\delta,\kappa}(g)]\rangle_{dq_0}| \leq N^N \|G\|_{1,\gamma,\alpha} \exp[K_5 A(g)].$$

Then Corollary 3.3 follows from our assertion and from the fact that $\mu\phi_{\delta,\kappa}(g) = \mu\phi_\delta(g)$ under the assumption in Corollary 3.3.

The proof of our assertion follows from an easy modification of the proof of Theorem 3.2. That is, in the proof of Theorem 3.2 we replace $V(\delta,\kappa,\lambda g)$ by $V(\delta,\kappa,\lambda g) - \mu\phi_{\delta,\kappa}(g)$ and follow each step. For complex μ we use the fact that $|\exp(-\mu\phi_{\delta,\kappa}(g))| \leq \exp[-(\mathrm{Re}\,\mu)\phi_{\delta,\kappa}(g)]$. We note that there is essentially no change in the combinatoric bounds. Let w be the kernel of $\phi_{\delta,\kappa}(g)$ vertex. Then it follows that

$$\|\mu^\epsilon w\|_2 \leq O(1)|\Delta|^{1/4-\epsilon/2}, \quad \epsilon < 1/4. \tag{3.17}$$

We now use the method of decomposing big graphs into small graphs, (3.17) and the results in the Appendix. The proof then follows a straight modification of Step 4 in the proof of Theorem 3.2. Obviously the constant K_5 in the corollary is uniformly bounded in μ on a compact subset of $\mathbb{C}$. ∎

Proof of Theorem 3.4. We employ the method used in the proof of Ref. 3, Theorem 3. We assume that $\kappa_b \geq \kappa_a$ and that neither κ_a nor κ_b has a lower cutoff. Let $\kappa_a = \kappa_0 < \kappa_1 \cdots < \kappa_M = \kappa_b$. Let $\kappa_i(s) = s\kappa_{i+1} + (1-s)\kappa_i$. If we follow the method used in the proof of Ref. 3, Theorem 3, then we obtain

$$\begin{aligned}
&|\langle G\rangle_{dq(\delta,\kappa_b,g)} - \langle G\rangle_{dq(\delta,\kappa_a,g)}| \\
&\leq \sum_{i=0}^{M-1} \left| \int_0^1 ds \left\langle G \exp[\mu\phi_\delta(g)] \left\{ \frac{d}{ds} V(\delta,\kappa_i(s),\lambda g) \right\} \right.\right. \\
&\left.\left. \times \exp[-V(\delta,\kappa_i(s),\lambda g)] \right\rangle_{dq_0} \right| \\
&\leq \sum_{i=0}^{M-1} \sum_{\sigma\in\Gamma} \int_0^1 ds \left| \langle G_{\sigma,i} \exp[\mu\phi_\delta(g)] \right. \\
&\left. \exp[-V(\delta,\kappa_i(s),\lambda g)] \rangle_{dq_0} \right|,
\end{aligned} \tag{3.18}$$

where $(d/ds)V(\delta,\kappa_i(s),\lambda g)$ is the sum of a finite P vertices, Γ is a finite index set, and each Graph $G_{\sigma,i}$ contains one G graph, one P vertex, at most 16 C vertices and at most 12 $\mu\phi_\delta(g)$ vertices. We have used a single P step (C step) of the P-C expansion to obtain the first (second) inequality of (3.18). Following the process used in Ref. 3 and using Corollary 3.3, we bound (3.18) by

$$\sup_{i,\sigma} K' \log\lambda_i \|G_{\sigma,i}\|_{1,\gamma,\alpha}, \tag{3.19}$$

where the constant K' depends on almost everything except the κ's and δ, and λ_i is the maximum low cutoff of the P vertex in $G_{\sigma,i}$. Since the topological

Graph $G_{\sigma,i}$ is the same as that of Ref. 3, the upper bound of the kernel norm $\|G_{\sigma,i}\|_{1,\gamma,\alpha}$ may now be estimated by the method of Ref. 3.

We divide $P^e_\alpha \subset M^\nu |G_{\sigma,i}|$ into two subgraphs $P^e_{\alpha,1} G$ and $P^e_{\alpha,2} R$ where $P^e_{\alpha,1}$ includes those P^e vertices that contract to G only and $P^e_{\alpha,2}$ includes the rest (we suppress the μ^ν's, σ's, etc.). R is the graph containing the P, C, and $\mu\phi_\delta(g)$ vertices. Following the process used in the proof of Ref. 3, Theorem 3, the problem is essentially reduced to the estimate of $\|P^2_{\alpha,2} R\|_{\text{H.S.}}$. We again use Propositions A1.1–5, Remark A3, and also (3.17) to handle the $\mu\phi_\delta(g)$. Following the process used in the proof of Ref. 3, Theorem 3, to obtain the same upper bounds of $\|P^e_{\alpha,2} R\|_{\text{H.S.}}$ as those of Ref. 3. In this way we bound (3.19) by

$$\sup_{i,\sigma} K' \log\lambda_i \|G_{\sigma,1}\|_{1,\gamma,\alpha} \leq O(1)\lambda_0^{-\epsilon}, \quad \epsilon > 0,$$

for sufficiently small γ and $\alpha < \gamma/2$, where λ_0 is the upper momentum cutoff κ_a. Since $\lambda_0 \to \infty \kappa_a \to 1$, we have proven the theorem. The convergence is uniform in μ on a compact subset of C by Corollary 3.3 and by the uniform boundedness of $|\mu f_\delta|_\gamma$. We leave any omitted details of the proof to the readers. ∎

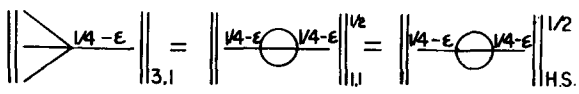
APPENDIX: ESTIMATES FOR SMALL GRAPHS

1. The results

In this Appendix we establish the basic estimates for small graphs and obtain upper bounds like those of Refs. 2 and 3. We keep the notation of the former. Throughout this section we assume that the upper momentum cutoffs are smaller than $2\delta^{-1}$ for the reason stated in Step 1 of the proof for Theorem 3.2. We first collect the main results in this section. We let $M(1), \ldots, M(4)$ denote the upper cutoffs in the legs of a P or C vertex, numbered so that $M(1) \leq M(2) \leq M(3) \leq M(4)$.

Propositions A1.1: Let w be the kernel of a $\phi^4 P$ or C vertex in Δ.

(a) For sufficiently small $\epsilon \geq 0$ and $\epsilon_1 \leq 1/4$,



$$\leq O(1)|\Delta|\epsilon_1 (\log M(2))^{1/2}.$$

(b)

$$\leq O(1)\ |\Delta|^\epsilon \lambda^{-\epsilon}, \quad \epsilon < 1/16.$$

Proposition A1.2: The uncancelled mass-renormalization diagram

$$M(\Delta, \Delta') = -P \equiv C -$$

and the mass counterterm

$$\delta M = \delta M(\Delta, \Delta')\delta_{\Delta,\Delta'}$$

($\delta_{\Delta,\Delta'} = 0$ for $\Delta \neq \Delta'$ and 1 for $\Delta = \Delta'$) have Hilbert–Schmidt norms that are $O(1)(|\Delta|\ |\Delta'|)^{1/4} \log M(2)^{1/2}$. Their sum with kernel w satisfies

$$\|w\| \leq O(1)|\Delta'|^\epsilon |\Delta|^\epsilon \lambda^{-\epsilon} d^{-\epsilon} d^{-n}$$

for some $\epsilon > 0$. Here d is the scaled distance from Δ to Δ'.

Proposition A1.3: Let w be the kernel of a W vertex with each momentum leg restricted to an interval $[M_j, M_{j+1}]$. Let $\epsilon < 1/4$ and let ν of (3.2) be sufficiently small. Then

$$\|w\| \leq O(|\Delta|^\epsilon) \prod_{\text{legs}} b^{-3(1+\nu)/2}.$$

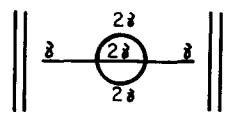
Proposition A1.4: The two-vertex diagram $P \equiv W$ for contraction between a P vertex in Δ and a W vertex in Δ' is bounded by

$$O(1)(|\Delta|\ |\Delta'|)^\epsilon \lambda^{-\epsilon} \prod_{\text{legs}} b^{-3(1+\nu)/2}$$

for small $\epsilon > 0$ provided ϵ and ν are sufficiently small.

Proposition A1.5:

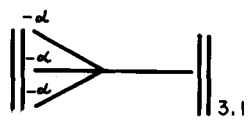
(a) Let u be the smallest upper cutoff of any of the three contraction lines. Then



$$\leq O(u^{6\gamma}),$$

when $|\Delta| = |\Delta'| = 1$.

(b) Let w be a ϕ^4 P vertex and let $|\Delta| = 1$. Then



$$\leq O(1)\lambda^{-\alpha}.$$

We note that propositions A1.1–4 and Proposition A1.5 are the results corresponding, respectively, to Ref. 2 Propositions 5.3.3–6, and Ref. 3, Theorem 6. It will be clear that the other estimates of small graphs corresponding to those used in the proof of Ref. 3, Theorem 3, follow from the method used in proving Propositions A1.1–5 and its slight modification. The method of proof for our results is essentially the same as that of Ref. 2, Secs. 6.1–6.3 (and Ref. 3, Appendix). We will skip any detailed discussions given in Ref. 2, and exhibit only the necessary modifications.

2. Convolution estimates

We start with convolution estimates for $F_{\delta,\Delta}(k)$ as defined in (2.8). Since $|\mu_\delta(k^{(l)})| \geq O(1)\delta^{-1}$ for $k^{(l)} \in [\pi/\delta,\ 3\pi/2\delta]$, it follows that

$$\mu(k) \leq O(1)\mu_\delta(k) \quad \text{and} \quad F_{\delta,\Delta}(k) \leq O(1) F_\Delta(k) \qquad \text{(A2.1)}$$

for $|k^{(l)}| \leq 3\pi/2\delta$, by Lemma 2.1 (b). Since $2 - 2\cos(y) \leq 2y^2$ for all $y \in R$, we have

$$\mu_\delta(k) \leq O(1)\mu(k) \quad \text{and} \quad F_\Delta(k) \leq O(1) F_{\delta,\Delta}(k), \qquad \text{(A2.2)}$$

for all $k \in R^3$. From now on, we assume $|k^{(l)}| \leq 4\pi/\delta$, $l = 0, 1, 2$.

Lemma A2.1:

(a) For $\gamma + \lambda > 1$, $\nu = \min\{\gamma, \lambda, \gamma + \lambda - 1\}$, $\bar{\nu} = \max(\gamma, \lambda)$, $\epsilon > 0$,

$$|\Delta| F_{\delta,\Delta}^{\gamma} * F_{\delta,\Delta}^{\lambda} \leq O(1) \begin{cases} F_{\delta,\Delta}^{\nu} & \text{if } \bar{\nu} \neq 1 \\ F_{\delta,\Delta}^{\nu-\epsilon} & \text{if } \bar{\nu} = 1. \end{cases}$$

(b) Let $\lambda \geq 0$; then part (a) still holds, if we replace $F_{\delta,\Delta}^{\lambda}$ by F_{Δ}^{λ}.

Corollary A2.2:

(a) For $0 \leq \gamma \neq 1$, $0 \leq \lambda \neq 1$, $\gamma + \lambda > 1$, $\nu = \min(\gamma, \lambda, \gamma + \lambda - 1)$,

$$|\Delta| \leq |\Delta'|,$$

$$|\Delta|^{\gamma} |\Delta'|^{1-\gamma} F_{\delta,\Delta}^{\gamma} * F_{\delta,\Delta'}^{\lambda} \leq O(1) F_{\delta,\Delta'}^{\nu} \quad \text{if } \gamma < 1,$$

$$|\Delta| F_{\delta,\Delta}^{\gamma} * F_{\delta,\Delta'}^{\lambda} \leq O(1) F_{\delta,\Delta'}^{\lambda} \quad \text{if } \lambda < 1 < \gamma,$$

$$|\Delta|^{\nu} |\Delta'|^{1-\nu} F_{\delta,\Delta}^{\gamma} * F_{\delta,\Delta'}^{\lambda} \leq O(1) F_{\delta,\Delta'}^{\nu} \quad \text{if } 1 < \lambda, \ 1 < \gamma.$$

(b) Let $0 < \epsilon < 1/2$, $0 \leq \alpha_i < 1/4$, $\alpha = \alpha_1 + \alpha_2$, $|\Delta| \leq |\Delta'| \leq |\Delta''| \leq 1$. Then

$$\int F_{\delta,\Delta}(k+p) F_{\delta,\Delta''}(p)^{\alpha} F_{\delta,\Delta'}(k'-p)\, dp$$

is bounded by $O(1)$ times any of the following:

$$|\Delta|^{-3/4-\epsilon/2} |\Delta'|^{-1/4+\epsilon/2}$$

$$F_{\delta,\Delta'}(k+k)^{(1+\epsilon)/2} F_{\delta,\Delta''}(k)^{\alpha 1} F_{\delta,\Delta''}(k')^{\alpha_2},$$

$$|\Delta|^{-3/4-\epsilon/2} |\Delta'|^{-3/4} F_{\delta}(k+k')^{(1+\epsilon)/2} F_{\delta,\Delta''}(k)^{\alpha 1} F_{\delta,\Delta''}(k')^{\alpha_2},$$

$$|\Delta|^{-3/4-\epsilon/2} |\Delta'|^{-3/4} |\Delta''|^{-\alpha} F_{\delta}(k-k')^{(1+\epsilon)/2} F_{\delta}(k)^{\alpha 1} F_{\delta}(k')^{\alpha_2},$$

where $F_{\delta} = F_{\delta,\Delta}$ for $|\Delta| = 1$.

Lemma A2.3: (Ref. 2, Proposition 6.1.5 and Corollary 6.1.6). If we replace every $F_{\delta,\Delta}$ by F_{Δ}, then Lemma A2.1 and Corollary A2.2 still hold.

Proof of Corollary A2.2:

(a) This follows by a direct application of Lemma A2.1 (a).

(b) This follows from Lemma A2.1 (a) and from the Schwartz inequality. ∎

Proof of Lemma A2.1: Since part (b) follows from part (a) and (A2.2), we only need to show part (a). We note that

$$\int_{4\delta} F_{\delta,\Delta}(k-k')^{\gamma} F_{\delta,\Delta}(k')^{\lambda} dk'$$

$$\leq (4)^3 \int_{\delta} F_{\delta,\Delta}(k-k')^{\gamma} F_{\delta,\Delta}(k')^{\lambda} dk'$$

$$= (4)^3 \int_{\delta} F_{\delta,\Delta}\left(k \pm \frac{2n\pi}{\delta} - k'\right)^{\gamma} F_{\delta,\Delta}(k')^{\lambda} dk', \tag{A2.3}$$

since, by the periodic property of the cosine function,

$$F_{\delta,\Delta}(k) = F_{\delta,\Delta}(k \pm 2n\pi/\delta), \ n \in \mathbb{N}^3. \tag{A2.4}$$

Consequently, it is sufficient to prove the lemma for $|k^{(1)}| \leq \pi/\delta$ and $|k'^{(1)}| \leq \pi/\delta$. For simplicity we write that, for $k \in R$ and $|k| \leq \pi/\delta$,

$$I_{\delta,\Delta}(k) = \int_{\delta} F_{\delta,\Delta}(k-k')^{\gamma} F_{\delta,\Delta}(k')^{\lambda} dk'.$$

We split the region of integration into

I: $0 \leq k' \leq \pi/\delta$, II: $-\pi/\delta \leq k' \leq 0$.

In region I: If $-\pi/2\delta \leq k \leq \pi/\delta$, $|k-k'| \leq 3\pi/2\delta$. We use (A2.1), (A2.2) and Lemma A2.3 to obtain

$$|I_{\delta,\Delta}(k)| \leq O(1) \int F_{\Delta}(k-k')^{\gamma} F_{\Delta}(k')\, dk$$

$$\leq O(1) |\Delta|^{1/3} \begin{cases} F_{\Delta}(k)^{\nu} \\ F_{\Delta}(k)^{\nu-\epsilon} \end{cases}$$

$$\leq O(1) |\Delta|^{1/3} \begin{cases} F_{\delta,\Delta}(k)^{\nu} & \text{if } \bar{\nu} \neq 1 \\ F_{\delta,\Delta}(k)^{\nu-\epsilon} & \text{if } \nu = 1. \end{cases} \tag{A2.5}$$

If $-\pi/\delta \leq k \leq -\pi/2\delta$, we subdivide region I into $I_1 = \{k' | 0 \leq k' \leq \pi/2\delta\}$ and $I_2 = \{k' | \pi/2\delta \leq k' \leq \pi/\delta\}$. In region I_1 we have $|k-k'| \leq 3\pi/2\delta$ and hence the bound in (A2.6) still holds. In region I_2 we also have

$$\int_{\pi/2\delta}^{\pi/\delta} F_{\delta,\Delta}(k-k')^{\gamma} F_{\delta,\Delta}(k')^{\lambda} dk''$$

$$= \int_{-\pi/\delta}^{-\pi/2\delta} F_{\delta,\Delta}(k-k'')^{\gamma} F_{\delta,\Delta}(k'')^{\lambda} dk''$$

by (A2.4). Since $|k''-k| \leq \pi/\delta$, the argument used in (A2.6) gives the bound in (A2.5).

In Region II: The argument used in region I gives the bound in (A2.4). Collecting the above results we have proved the lemma. ∎

3. Estimates for small graphs

In this section we prove Proposition A1.1 and Propositions A1.3—5 by using the results in Sec. 2, Sec. A2, and Ref. 2, Propositions 6.1.1—4. We will show the mass-renormalization cancellation (Proposition A1.2) in the following section. We adopt the main process and the results given in Ref. 2, Secs. 6.2, 6.3, and Ref. 3, Appendix. Roughly speaking, if we use Lemma A2.1 and Corollary A2.2 in the places where Lemma A2.3 has been used in Ref. 2, then the method used in Ref. 2, Secs. 6.2, 6.3 gives us Proposition A1.1, Propositions A1.3, 4, and Proposition A1.5 (b). To show the remainders we need to split the region of integration into several parts and use Lemma A2.3 in each region. We note that the kernel of an individual vertex has the form (from Step 3 in the proof of Theorem 3.2)

$$w(k_1, \ldots, k_l) = (\mu_1, \ldots, \mu_l)^{-1} v_{\delta}(k_1, \ldots, k_l) \tag{A3.1}$$

on the range of integration, where v_{δ} denotes the various momentum cutoffs and localization factors. We keep in mind that the ranges of integration are restricted to $[-\pi/\delta, \pi/\delta]^{3l}$. We first show Proposition A1.3.

Proof of Proposition A1.3: The proof is similar to that of Ref. 2, Sec. 6.2. Replacing F_{Δ} by $F_{\delta,\Delta}$ in Ref. 2 (6.2.16), it is easy to check that

$$|v_{\delta}(k)| \leq \sum_{J_1''} O(|\Delta|^{(1/4)+(|J_1''|/3)}) \left(\prod_{i \in I_{j_i}} \mu_i\right) \int_0^1 d\lambda_1 \cdots \int_0^1 \times d\lambda_{\alpha} F_{\delta,\lambda}, \tag{A3.2}$$

where $F_{\delta,\lambda}$ is defined through replacing F_{Δ} by $F_{\delta,\Delta}$ in Ref. 2 (6.2.15). For the details we cite Ref. 2, Sec. 6.2. The bound corresponding to the last expression in Ref. 2, p. 370, follows from (A2.1), (A2.2). Hence, by the same reasoning as that in Ref. 2, p. 371, the following estimate is sufficient to prove Proposition A1.3:

$$\sup_{\xi} \int_{J_1'' \cup I_2}^{\delta} dk \quad \frac{F_{\delta,\Delta}(\xi+P)^2}{\Pi_{i \subset J_1'' \cup I_2} \mu_i^2} \leq O(|\Delta|^{(1-|I_2|-|J_1''|)-\epsilon)/2}). \tag{A3.3}$$

We now show (A3.3) and complete the proof.

For $|J''| + |I_2| = 1$, by using the lower cutoff $|\Delta|^{-1/2}$ on the I_2 leg and also Corollary A2.2 (a), we have that

$$\sup_{\xi} \int_{|p|>|\Delta|^{-1/2}} dP \, \frac{F_{\delta,\Delta}(\xi+P)^2}{\mu(p)^2} \leq O(|\Delta|^{-\epsilon}). \tag{A3.4}$$

If $|J_1''| + |I_2| = 2$, Corollary A2.2 and the method above yield $O(|\Delta|^{-1/2-\epsilon})$. If $|J_1''| + |I_2| = 3$, then Proposition 6.1.4 of Ref. 2 yields

$$\sup_{\xi} \int_{3\delta} dP \, F_{\delta,\Delta}(\xi+P)^2$$

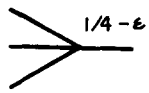
$$\int dk_2 \, dk_3 [\mu(P-k_2-k_3)^2 \mu(k_2)^2 \mu(k_3)^2]^{-1}$$

$$\leq O(1) \log M_\mu(\Delta) \sup_{\xi} \int_{3\delta} dP \, F_{\delta,\Delta}(\xi+P)^2$$

$$\leq O(1) |\Delta|^{-1-\epsilon}.$$

Here we have used Lemma A2.1 and the fact that $M_\mu(\Delta) \leq O(1)(\log|\Delta|^{-1})$. By the method above (also Ref. 2, p. 371]), one may also prove (A 3.3) for $|J_1''| + |I_2| = 4$. This completes the proof. ∎

Proof of Proposition A1.1: We again follow the method in Ref. 2, Sec. 6.3. We only need to check the final estimates, which differ from our results (i.e., estimates for $F_{\delta,\Delta}$ instead of F_Δ).

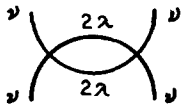
(a) We set $A =$  We use Lemma A2.1 (a) and (b), and the method in Ref. 2, p. 372, to bound $\|A*A\|$ by

$$O(1)(\log^2 M(2))|\Delta|^2 \int_{\delta} \frac{F_{\delta,\Delta}(k_1+k_5)^{2-2\gamma}}{(\mu_1\mu_5)^{(3+\epsilon)/2}} \, dk_1 \, dk_5$$

$$\leq O(1)(\log^2 M(2)) |\Delta|^2 \int_{\delta} \mu_1^{-(3+\epsilon)/2} F_{\delta,\Delta}(k_1)^{(3+\epsilon)/6} \, dk_1$$

$$\leq O(1)(\log^2 M(2)) |\Delta|.$$

This proves the proposition.

(b) Let v be the kernel corresponding to  $\lambda + \nu < 1/8$. By using Corollary A2.2 (b) and by following the method in Ref. 2, p. 342, we bound $|v(k_1, k_5)|$ by

$$O(|\Delta|)(\mu_1\mu_2\mu_3\mu_4)^{-1+\nu} F_{\delta,\Delta}(K_1+K_2)^{(1+\gamma)/2}$$

$$\times F_\delta(K_1)^{(1-4\lambda)/6} F_\delta(K_2)[(1-4\lambda)/6],$$

where $K_1 = k_1 + k_2$, $K_2 = k_3 + k_4$. We again use Lemma A2.1 (a) and (b), and the method of Ref. 2, p. 373, to bound $\|v\|_2^2$ by $O(|\Delta|)$. By obtaining $\lambda^{-\epsilon}$ from $\mu^{-\lambda}$ or $\mu^{-\nu}$ from the kernel of the diagram in Proposition (b), we have proved our proposition. ∎

Proof of Proposition A1.4: The follows by the argument used in the proof of Ref. 2, Proposition 5.3.6, and by the method in the proof of Proposition A1.3 (for $|\Delta'| \leq |\Delta|$). We do not exhibit the detailed proof, and refer instead to Ref. 2. ∎

Proof of Proposition A1.5. (a): Let $w(k_1, k_5)$ be the kernel by the corresponding diagram. We note that by Lemma 2.1 (b)

$$w(k_1, k_5) \leq O(1) \int_{\delta} dk_2 \, dk_3 \, dk_4 \, \frac{F_\delta(k_1+k_2+k_4) F_\delta(-k_2-k_3-k_4-k_5)}{\mu_1^{1-\gamma} \mu_5^{1-\gamma} (\mu_2\mu_3\mu_4)^{2-2\gamma}}. \tag{A3.5}$$

We assume that $k_2^{(0)}$ has the upper momentum cutoff u. If we use k_2, k_3 and $P = k_2 + k_3 + k_4$ as integration variables, then

$$|w(k_1, k_5)| = O(1)(\mu_1\mu_5)^{-1+\gamma} \int_{|p^{(l)}| \leq 3\pi/\delta} dp \int_{\delta} dk_2 \mu_2^{-2+2\gamma} \times \mu(p-k_2)^{-1+4\gamma} \times F_\delta(-p+k_5) F_\delta(p+k_1). \tag{A3.6}$$

We split the P integration into the regions

$$I_n = \{(P^{(0)}, P^{(1)}, P^{(2)}) \mid |P^{(l)} - 2n\pi/\delta| \leq \pi/\delta, \; l = 0, 1, 2\}. \tag{A3.7}$$

In the region I_0, $n = (0,0,0)$: If both $|k_1^{(l)}| \leq \pi/2\delta$ and $|k_5^{(l)}| \leq \pi/2\delta$ for all $l = 0, 1, 2$, then $|p^{(l)} - k_1^{(l)}| \leq 3\pi/2\delta$ and $|p^{(l)}| \leq 3\pi/2\delta$ for all $l = 0, 1, 2$. Then the proposition follows from (A2.2) and the method used in Ref. 3, Appendix. Otherwise, there exists an $i(=1, 5)$ and an $l(=0, 1, 2)$ such that $|k_i^{(l)}| \geq \pi/2\delta$, for example, $i = 1$ and $l = 0$. We multiply $(u^{-1}\mu_2^{(0)})^{-8\gamma}$ in the integrand of (A3.6) and integrate with respect to k_2 to obtain

$$|w(k_1 k_5) \leq O(1) u^{8\gamma} \int_{I_0} dp \, F_\delta(k_1+p) F_\delta(-p-k_5) \cdot (\mu_1\mu_5)^{-1+\gamma}$$

$$\leq O(1) u^{8\gamma} F_\delta(k_1+k_5)^{1-\epsilon} (\mu_1\mu_5)^{-1+\gamma} \quad \text{if } \gamma < 1/4,$$

where we have used Lemma A2.1. Hence

$$\|w\|_2 \leq O(1) u^{8\gamma} (\delta)^{-1/8-2\gamma}$$

$$\leq O(1) u^{6\gamma} \qquad \text{if } \gamma < 1/4.$$

Here we have used the fact that $u \leq O(1)\delta^{-1}$.

In region I_n, $n \neq (0,0,0)$: There exists $l(=0, 1 \text{ or } 2)$ such that $n^{(l)} \neq 0$, for example, $l = 1$. We split the $k_2^{(l)}$ integration into two parts:

A: $|k_2^{(l)}| \leq \pi/2\delta$, B: $|k_2^{(l)}| \geq \pi/2\delta$.

In region A: Then $|p^{(l)} - k_2^{(l)}| \geq \pi/2\delta$. We multiply $(\delta\mu_2)^{-2\gamma}(\delta u(p - k_2^{(l)}))^{1-4\gamma}$ [$\geq O(1)$ on the support of the integrand] and integrate to obtain

$$|w(k_1, k_5)| \leq O(1) \delta^{1-6\gamma} \int \mu_2^{-2} \, dk_2 \, F_\delta(k_1+k_5)^{1-\epsilon} (\mu_1\mu_5)^{-1+\gamma}$$

$$\leq O(1) \delta^{1-6\gamma} (\pi/\delta)^{2/3} u^{1/3} F_\delta(k_1+k_5)^{1-\epsilon} (\mu_1\mu_5)^{-1+\gamma}$$

$$\leq O(1) u^{6\gamma} F_\delta(k_1+k_5)^{1-\epsilon} (\mu_1\mu_5)^{-1+\gamma} \quad \text{if } \gamma < 1/6$$

Here we have used Lemma A2.1. The proposition follows from the above result.

In region B: Since $|p^{(l)}| \geq \pi/2\delta$, the method used in the region A gives the corresponding assertion. Summarizing these results, we have proved the proposition.

(b) This follows as a corollary of Proposition A1.1 (a).

∎

Remark A3: We have used estimates on one- or two-

vertex graphs in the proof of Theorem 3.4. These are essentially propositions A1.1—5 and simple modifications of those results.

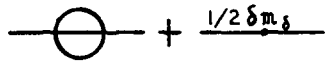
We conclude this section with the following result:

Lemma A3.1: The changes, $\delta E_2(\Delta)$ and δE_3, of the second- and third-order vacuum counterterms arising in the Wick construction in δV are bounded by $O(|\Delta|^{1-\gamma}M(1)\log M(2)) \leq O(1)$.

Proof: This follows by the method in the proof of Ref. 1, Proposition 6.3.1 and Lemma A 2.1, and its corollary. ∎

4. Mass renormalization cancellations

Finally, we consider the cancellation between the uncancelled mass renormalization diagram and the mass counterterm. The kernels of their sum



has the form

$$w_{\Delta,\Delta'}(k_1,k_5) = \text{const}(\mu_\delta(k_1)\mu_\delta(k_5))^{-1}\Big(\int \alpha\beta\gamma\tau\, dk_2\, dk_3\, dk_4 - \sum_{\substack{n\in N^3 \\ n^{(1)}=0,\pm1}} \int_{|P^{(1)}|\leq\pi/\delta} \bar\alpha_n\bar\beta_n\bar\gamma_n\bar\tau_n\, dk_2\, dk_3\, dk_4\Big)$$

$$= \hat w_{\Delta,\Delta'} + \bar w_{\Delta,\Delta'}, \tag{A4.1}$$

where

$$\alpha = \left(\frac{d}{ds}\kappa_s\right)(k_1,k_2,k_3,k_4)\,\kappa_s(k_2,k_3,k_4,k_5), \tag{A4.2}$$

$$\beta = \bar\beta_n = (g\chi_\Delta)\tilde{}_\delta(P+k_1)(g\chi_{\Delta'})\tilde{}_\delta(-P-k_5),$$

$$\gamma = \prod_{i=2}^4 \mu_\delta(k_i)^{-2}, \quad \tau = \prod_{i=2}^4 \chi\left(\frac{\pi k_i}{\delta}\right),$$

$$\bar\alpha_n = \frac{1}{2}\frac{d}{ds}\kappa_s\left(P_2+\frac{2n\pi}{3\delta},\ P_3+\frac{2n\pi}{3\delta},\ P_4+\frac{2n\pi}{3\delta}\right)^2 \kappa_1(k_1)\chi_1(k_5),$$

$$\bar\gamma_n = \prod_{i=2}^4 \mu_\delta\left(p_i+\frac{2n\pi}{3\delta}\right),\quad \bar\tau_n = \prod_{i=2}^4 \chi\left(\frac{\pi}{\delta}\left(p_i+\frac{2n\pi}{3\delta}\right)\right).$$

Here $\chi(k)$ is the characteristic function of $[-1,1]^3$, $P = k_2+k_3+k_4$, $P_i = k_i - \frac{1}{3}P$. For $\Delta \neq \Delta'$, we note that $\bar w_{\Delta,\Delta'} = 0$ by the fact that $g^2\chi_\Delta\chi_{\Delta'} = 0$.

To compare $\hat w_{\Delta,\Delta'}$ with $\bar w_{\Delta,\Delta'}$ we rewrite $\hat w_{\Delta,\Delta'}$. We split the P integration into the regions I_n defined in (A3.7). We translate the center of I_n to the origin by redefining k_i by $k_i - 2n\pi/3\delta$ and use the fact that $(g\chi_\Delta)\tilde{}_\delta(k) = (g\chi_\Delta)\tilde{}_\delta(k \pm 2n\pi/\delta)$ to obtain

$$\hat w_{\Delta,\Delta'}(k_1,k_5) = \sum_{\substack{n\in N^3 \\ n^{(1)}=0,\pm1}} \hat w^{(n)}_{\Delta,\Delta'} \tag{A4.3}$$

$$\hat w^{(n)}_{\Delta,\Delta'} = \text{const}(\mu_\delta(k_1)\mu_\delta(k_5))^{-1}\int_{|P^{(1)}|\leq\pi/\delta} \alpha_n\beta_n\gamma_n\tau_n\, dk_2\, dk_3\, dk_4,$$

where $\alpha_n,\ldots,\tau_n$ are defined through replacing k_i by $k_i + 2n\pi/3\delta$ in the definitions of $\alpha,\ldots,\tau$. We also define $\hat w^{(n)}_{\Delta,\Delta}$ in a similar manner. We rewrite (A4.1) as

$$w_{\Delta,\Delta'} = \sum_{\substack{n\in N^3 \\ n^{(1)}=0,\pm1}} (\hat w^{(n)}_{\Delta,\Delta'} + \bar w^{(n)}_{\Delta,\Delta'}). \tag{A4.4}$$

We now complete the proof by estimating each $\hat w^{(n)}_{\Delta,\Delta'} + \bar w^{(n)}_{\Delta,\Delta'}$.

Proof of Proposition A1.2: We consider the first part of the proposition. We use Lemma 2.1 (b) and Lemma 2.3 (b) to estimate

$$|\bar w^{(n)}_{\Delta,\Delta'}(k_1,k_5)| \leq O(|\Delta|\,|\Delta'|)(\mu_1\mu_5)^{-1}\cdot$$

$$\times \int |\alpha_n| \prod_{i=2}^4 \mu(k_i+2n\pi/3\delta)$$

$$F_{\delta,\Delta}(P+k_1)F_{\delta,\Delta'}(-P-k_5)\,dk_5\,dk_3\,dk_4.$$

We assume that the upper cutoff $M(2)$ for k_2 is the smallest upper cutoff among k_2, k_3, k_4. We then have Ref. 2, (6.3.1),

$$\int dk_2\,dk_3\,[\mu_2^2\mu_3^2\mu(P-k_2-k_3)^2]^{-1} \leq O(\log M(2)). \tag{A4.5}$$

We let $k_2+2n\pi/3\delta$, $k_3+2\pi n/3\delta$, and P be independent variables and use (A4.5) and Corollary A2.2 (b) ($\gamma = 3/4$) to obtain

$$|\hat w^{(n)}_{\Delta,\Delta'}| \leq O(1)(|\Delta|\,|\Delta'|)^{1/4}\log M(2)(\mu_1\mu_5)^{-1}\,|\Delta'|^{3/4} \times F_{\delta,\Delta'}(k_1+k_4)^{3/4-\epsilon}. \tag{A4.6}$$

Here we have assumed that $|\Delta| \leq |\Delta'|$. The first part of the proposition for $\hat w^{(n)}_{\Delta,\Delta'}$ follows from (A4.6) and Lemma A2.1(a). A method similar to that used above gives us the bound for $\bar w^{(n)}_{\Delta,\Delta'}$. This proves the first part of Proposition A1.2.

We now prove the renormalization cancellation. Since we use only the comparison function $F_{\delta,\Delta}$, the distance factor d^{-n} can be obtained by Lemma 2.3(b) and the method used in Step 3 of the proof for Theorem 3.2. If $|k_i^{(l)}| \geq \pi/2\delta$ for some i ($=1$, or 5) and l, then one may obtain the convergence factor $\lambda^{-\epsilon} < O(1)\delta^\epsilon$ from the $k_i^{(l)}$ integration by using the bound in (A4.6) for $|\hat w^{(n)}_{\Delta,\Delta'}|$ and $|\bar w^{(n)}_{\Delta,\Delta'}|$, respectively. Hence we only need to consider the following region:

$$|k_1^{(l)}| \leq \pi/2\delta, \quad |k_5^{(l)}| \leq \pi/2\delta \quad \text{and all } l. \tag{A4.7}$$

On the support of the integration we estimate

$$|\hat w^{(n)}_{\Delta,\Delta'} + \bar w^{(n)}_{\Delta,\Delta'}| \leq O(1)(\mu_1\mu_5)^{-1}\int_{|P^{(l)}|\leq\pi/\delta} |(\alpha_n\gamma_n\tau_n - \bar\alpha_n\bar\gamma_n\bar\tau_n) \times \beta|\,dk_2\,dk_3\,dk_4. \tag{A4.8}$$

We expand

$$\alpha_n\beta_n\gamma_n\tau_n - \bar\alpha_n\bar\beta_n\bar\gamma_n\bar\tau_n = \tau_n\bar\tau_n(\alpha_n\gamma_n - \bar\alpha_n\bar\beta_n)\beta + \tau_n(1-\bar\tau_n)\alpha_n\gamma_n\beta + \bar\tau_n(\tau_n-1)\bar\alpha_n\bar\gamma_n\beta). \tag{A4.9}$$

We consider the first term in (A4.9). Lemma 2.3(c) yields

$$|\gamma_n - \bar\gamma_n| \leq O(1)\mu(P)^{3\epsilon}\left\{\left[\mu\left(k_2+\frac{2n\pi}{3\delta}\right)^{-2+\epsilon} + \mu\left(P_2+\frac{2n\pi}{3\delta}\right)^{-2+\epsilon}\right] \times \mu\left(k_3+\frac{2n\pi}{3\delta}\right)^{-2}\ \mu\left(k_4+\frac{2n\pi}{3\delta}\right)^{-2} + \cdots\right\} \tag{A4.10}$$

on the support of the integrand. By the method used in Ref. 2, p. 374, we also have

$$|\alpha_n - \bar\alpha_n| \leq O(1)\,|\Delta'|^{-\epsilon/2}F_\Delta(P)^\epsilon\lambda^{-\epsilon/2}M(2)^{-\epsilon/2}. \tag{A4.11}$$

We again assume that $|\Delta| \leq |\Delta'|$. We write

$$\tau_n\bar{\tau}_n(\alpha_n\gamma_n - \bar{\alpha}_n\bar{\gamma}_n)\beta = \tau_n\bar{\tau}_n(\alpha_n - \bar{\alpha}_n)\gamma_n\beta + \tau_n\bar{\tau}_n(\gamma_n - \bar{\gamma}_n)\bar{\alpha}_n\beta. \tag{A4.12}$$

We estimate the first term in (A4.12). If we use (A4.11), Lemma 2.1(b), Lemma 2.3(b), (A2.2), and (A4.7), we bound the kernel corresponding to the first term of (A4.12) by

$$O(1)(|\Delta|\ |\Delta'|\lambda^{-\epsilon/2}M(2)^{-\epsilon/4}(\mu_1\mu_5)^{-1} \times \int_{|P^{(1)}|\leq\pi/\delta} dP\,dp_1\,dp_2\,|\bar{\tau}_n| \times \frac{F_\Delta(k_1+P)\,F_{\Delta'}(P)^{-\epsilon}\,|\Delta'|^{-\epsilon/3}F_{\Delta'}(-P-k_5)}{\Pi_{i=2}^4\,\mu(P_i+2\pi n/3\delta)^2} \tag{A4.13}$$

We first use (A4.5) for integrations with respect to $P_2 + 2n\pi/3\delta$ and $P_3 + 2n\pi/3\delta$, and we next use Lemma A2.3 for the P integration to bound the above expression by

$$O(|\Delta|^{1/4-\epsilon'/3})\,\lambda^{-\epsilon'/2} \times F(k_1+k_5)^{(1+\epsilon)/2}\,F(k_5)^{-\epsilon'}(\mu_1\mu_5)^{-1}. \tag{A4.14}$$

The L_2 norm of (A4.13) is dominated by $O(1)(|\Delta|\,|\Delta'|)^\epsilon \times\lambda^{-\epsilon}$ for some $\epsilon > 0$. This proves the proposition for the first term in (A4.12). Using (A4.10), the second term in (A4.12) yields a similar bound.

We next consider the second term in (A4.9). For convenience we denote $k_i + 2n\pi/3\delta$ by k_i for $i = 2, 3, 4$. We note that P_i's are not changed. The kernel corresponding to the second term in (A4.9) is bounded by

$$O(1)(|\Delta|\ |\Delta'|)(\mu_1\mu_5)^{-1}\int_{|P^{(1)}-2n^{(1)}\pi/\delta|\leq\pi/\delta} |\alpha\beta\gamma\tau| \times(1-\bar{\tau}_n)\,F_\Delta\left(P-\frac{2n\pi}{\delta}+k_1\right)\,F_{\Delta'}\left(-k-P+\frac{2n\pi}{\delta}\right) \times dk_2\,dk_3\,dk_4. \tag{A4.15}$$

We split the region of P integration into two parts:

I: $|P^{(1)} - 2n\pi/\delta| \leq (\pi/\delta)^{1/2}$;

II: $|P^{(1)} - 2n\pi/\delta| \geq (\pi/\delta)^{1/2}$;

In region I, on the support of the integrations,

$$\pi(\delta^{-1} - \delta^{-1/2}) \leq |k_i^{(1)}| \leq \pi\delta^{-1}$$

for some i and l, for example, $i = 3$ and $l = 0$. Also,

$$\int_{\pi(\delta^{-1}-\delta^{-1/2})\leq|k_3{}^{(0)}|\leq\pi\delta^{-1}} |\tau|\,\mu_3{}^{-2/3}\,dk_3{}^{(0)} \leq O(1)\,\delta^{1/6}. \tag{A4.16}$$

We use k_1, k_3, and P as the variables of integration and use (A4.16) and Lemma A2.3 to bound (A4.15) by

$$O(1)(|\Delta|\ |\Delta'|^{1/4}\delta^{1/6-\epsilon}(\mu_1\mu_5)^{-1}|\Delta'|^{3/4}F_{\Delta'}(k_1+k_5)^{3/4-\epsilon'}$$

for any $\epsilon > 0$ and $\epsilon' > 0$. Hence the L_2-norm for the kernel corresponding to the second term in (A4.9) yields the desired bound.

In region II, we multiply $[\delta F(p - 2n\pi/\delta)]^{-\epsilon} (\geq O(1))$ to the integrand of (A4.14) and apply the method used in bounding the L_2 norm of (A4.14). We then dominate the L_2 norm of (A4.15) by $O(1)(|\Delta|\,|\Delta'|)^\epsilon\lambda^{-\epsilon}$.

Collecting these results yields the proposition in the case of $|\Delta| \leq |\Delta'|$. Since our estimates are symmetric, this proves the proposition. ∎

ACKNOWLEDGMENT

In would like to thank Dr. Joel Feldman for finding errors in the original manuscript and suggesting the modifications.

Correction to the manuscript: The ultraviolet cutoff function $\kappa(k)$ should be replaced by $\kappa(k_\delta)$, where $k_\delta{}^{(i)} = 2\delta^{-1}\sin(\delta k^{(i)}/2)$, $i = 0, 1, 2$. It then follows that for $|k^{(i)}| \leq \pi/\delta$, we have:

(a) $|k^{(i)}/k_\delta^{(i)}|^{\pm 1} \leq O(1)$.

(b) Let u and l be the upper and lower momentum cutoffs of $\kappa(k)$ and let u_δ and l_δ be the corresponding momentum cutoffs of $\kappa(k_\delta)$. Then there exists a positive constant c $(1 \leq c \leq 2)$ such that $u \leq u_\delta \leq cu$ and $cl \geq cl_\delta \geq l$ uniformly in δ.

(c) $(D^m\chi)(k = \pm\pi/\delta) = 0$ for odd m.

Then the relation (3.14) follows by (c). It is also easy to check that the replacement causes no harm as a consequence of (b).

*Present address: Department of Mathematics, Yonsei University, Seoul, Korea.

[1]J. Glimm, Commun. Math. Phys. **10**, 1 (1968).
[2]J. Glimm and A. Jaffe, Fortschr. Phys. **21**, 327 (1973).
[3]J. Feldman, Commun. Math. Phys. **37**, 93 (1974).
[4]F. Guerra, L. Rosen, and B. Simon, Ann. Math. (to be published).
[5]J. Glimm, A. Jaffe, and T. Spencer, "The particle structure of the weakly coupled $p(\phi)_2$ model and other applications of high temperature expansion, I, II," in *Constructive Quantum Field Theory*, edited by G. Velo and A. Wightman (Springer-Verlag, Berlin, 1973).
[6]R. Griffiths and B. Simons, Commun. Math. Phys. **33**, 145 (1973).
[7]E. Nelson, "Probability theory and Euclidean field theory," in *Constructive Quantum Field Theory*, edited by G. Velo and A. Wightman (Springer-Verlag, Berlin, 1973).
[8]J. Glimm, A. Jaffe, and T. Spencer, Ann. Math. (to be published).

Linked cluster expansions in the density fluctuation formulation of the many-body problem*

C. E. Campbell

School of Physics and Astronomy, University of Minnesota, Minneapolis, Minnesota 55455
(Received 12 December 1974)

A perturbation theory is developed for the logarithm of the normalization integral (or partition function) of an N-body system which is either a Bose liquid in its ground state or a classical fluid in the canonical ensemble. The perturbation in the former is an n-body factor in the ground state wavefunction and in the latter is an n-body potential. The normalization function serves as a generating function for the cumulants (or static correlation functions) of the density fluctuation operator $\rho_{\mathbf{k}}$. The expansions of the perturbed partition function and correlation functions are shown to be linked cluster expansions involving the correlation functions in the unperturbed system; each term in these expansions is manifestly of the porper order in N. Several approximations involving truncations of the cumulants and/or resummation of part of the terms in the linked cluster expansions are discussed.

I. INTRODUCTION

The analogy between the Boltzmann factor for a classical fluid and the square of the ground state wavefunction for a quantum mechanical system of bosons has been exploited extensively to obtain approximations for n-body distribution functions for the quantum system.[1-6] For example, the ground state properties of liquid He^4, most notably the binding energy and the liquid structure factor, can be obtained approximately from a trial Jastrow function

$$\psi_J = \exp \frac{1}{2} \sum_{i<j}^{N} u_2(r_{ij}) \tag{1}$$

where u_2 is a real function chosen to minimize the expectation value of the Hamiltonian of N He^4 atoms.[4-7] For a given $u_2(r)$ it is only necessary to know the radial distribution function $g(r)$ to obtain the energy E and liquid structure function $S(k)$[8]:

$$E_J = \langle \psi_J | H | \psi_J \rangle / \langle \psi_J | \psi_J \rangle$$

$$= \tfrac{1}{2} N\rho \int d\mathbf{r}\, g_J(r) \left(V(r) - \frac{\hbar^2}{4m} \nabla^2 u_2(r) \right),$$

$$S_J(k) = N^{-1} \langle \psi_J | \rho_{\mathbf{k}} \rho_{-\mathbf{k}} | \psi_J \rangle / \langle \psi_J | \psi_J \rangle = 1 + \rho \int d\mathbf{r} (g_J(r) - 1) \times \exp(i\mathbf{k} \cdot \mathbf{r})$$

where $V(r)$ is the potential energy between two helium atoms, ρ is the number density ($\rho = N/\Omega$ where Ω is the volume), m is the mass of the helium atom, and $\rho_{\mathbf{k}}$ is the density fluctuation function

$$\rho_{\mathbf{k}} = \sum_{i=1}^{N} \exp(i\mathbf{k} \cdot \mathbf{r}_i). \tag{2}$$

The radial distribution function is defined in terms of u_2 by

$$g_J(|\mathbf{r}_1 - \mathbf{r}_2|) = \frac{N(N-1)}{\rho^2} \frac{\int d\mathbf{r}_3 \cdots d\mathbf{r}_N \psi_J^2(\mathbf{r}_1 \cdots \mathbf{r}_N)}{\int d\mathbf{r}_1 \cdots d\mathbf{r}_N \psi_J^2(\mathbf{r}_1 \cdots \mathbf{r}_N)}. \tag{3}$$

(The subscript J is attached to the energy, liquid structure function, and radial distribution function in order to emphasize that these are obtained from a trial Jastrow function.)

The evaluation of Eq. (3) is equivalent to the determination of the radial distribution function of a classical fluid in the canonical ensemble at temperature T interacting via the two body potential $\phi_2(r)$ defined by[1-3]

$$-\phi_2(r)/k_B T = u_2(r). \tag{4}$$

Then it is possible to use the various approximations developed in the theory of classical fluids (BBGKY–KSA, PY, HNC) to replace Eq. (3) by an approximate relation between $g_J(r)$ and $u_2(r)$. Generally, these approximations relate the classical radial distribution function $g_{cl}(r)$ to the two-body potential $\phi_2(r)$ through some functional equation of the form

$$g_{cl}(r) = F[\phi_2/k_B T,\ g_{cl}; r]$$

which is then solved for g_{cl}.[8] Then the analogous approximation for Jastrow functions is

$$g_J(r) = F[-u_2, g_J; r].$$

While Jastrow functions give reasonably good account of some of the properties of liquid He^4, recent interest has been focused upon enlarging the class of trial functions to obtain improved agreement with experiment.[9-11] This is accomplished by defining an extended Jastrow function as

$$\psi(\mathbf{r}_1 \cdots \mathbf{r}_N) = \exp \tfrac{1}{2} \chi(\mathbf{r}_1 \cdots \mathbf{r}_N) \tag{5a}$$

where χ is a real function given by

$$\chi(\mathbf{r}_1 \cdots \mathbf{r}_N) = \sum_{n=1}^{N} \frac{1}{n!} \sum_{\substack{i_1 \cdots i_n \\ i_s \neq i_t}} u_n(\mathbf{r}_{i_1} \cdots \mathbf{r}_{i_n}) \tag{5b}$$

where u_n is symmetric under interchange of its arguments. Indeed, the exact ground state wavefunction of a system of bosons can be chosen in the form of Eq. (5). For this equation to be unique, the u_n functions must be short-ranged under separation of subsets of the n particles.[10] This cluster property is shared by an n-body potential energy $\phi_n(\mathbf{r}_1 \cdots \mathbf{r}_n)$. Thus the analogy of the square of the wavefunction ψ with the Boltzmann factor for a classical system of particles interacting with many-body potentials is retained through the generalization of Eq. (4):

Copyright © 1975 American Institute of Physics

$$-\phi_n(\mathbf{r}_1 \cdots \mathbf{r}_n)/k_B T = u_n(\mathbf{r}_1 \cdots \mathbf{r}_n).$$

Consequently, an approximation which relates the distribution functions to the logarithm of the N-body distribution function will be useful for both the classical and the quantum fluid.

In this paper we develop a perturbation theory for the distribution functions defined in terms of the N-body probability density under the assumption that some zeroth order problem has been solved; i.e., the zeroth order problem is given by the N-body probability

$$W_N^{(0)}(\mathbf{r}_1 \cdots \mathbf{r}_N) = [\exp \chi_0(\mathbf{r}_1 \cdots \mathbf{r}_N)]/Q_N^{(0)} \tag{6}$$

where

$$Q_N^{(0)} = \int \exp\chi_0(\mathbf{r}_1 \cdots \mathbf{r}_N)\, d\mathbf{r}_1 \cdots d\mathbf{r}_N. \tag{7}$$

We assume that the necessary n-body distribution functions $g_n^{(0)}$ are known. For a general N-body probability W_N the distribution functions are defined by

$$g_n(\mathbf{r}_1 \cdots \mathbf{r}_n) = \frac{N!}{(N-n)!}\, \rho^{-n} \int W_N\, d\mathbf{r}_{n+1} \cdots d\mathbf{r}_N. \tag{8}$$

Then supposing that $\ln W_N$ is given by

$$\chi(\mathbf{r}_1 \cdots \mathbf{r}_N) - \ln Q_N = \chi_0(\mathbf{r}_1 \cdots \mathbf{r}_N) - \ln Q_N^0 + \chi_1(\mathbf{r}_1 \cdots \mathbf{r}_N), \tag{9}$$

we wish to find an expression for Q_N in terms of $g_n^{(0)}$ and χ_1.

The problem is formulated in Sec. II in terms of density fluctuation functions $\rho_{\mathbf{k}}$, where it is shown that the perturbation χ_1 has a simple series expansion in terms of these functions. It is convenient to replace the distribution functions g_n by the cumulants $U_p(\mathbf{k}_1 \cdots \mathbf{k}_p)$ of the $\rho_{\mathbf{k}}$, which are defined in Sec. II following Wu's development.[13] In Sec. III it is shown that the partition function for the perturbed classical system is a generating function for the cumulants U_p, and that there are linked cluster expansions for the perturbed partition function and cumulant functions in terms of the unperturbed cumulants $U_p^{(0)}$. The main result of this paper is Eq. (39) of this section. A reasonably simple diagrammatic notation is developed to express these results. Finally, in Sec. IV we discuss several approximations which involve truncations and resummations of these linked cluster series, and review the previous usage of some of these approximations. New applications of these results to specific physical models will be given elsewhere.[14]

II. DENSITY FLUCTUATION FORMULATION

In this paper, attention is restricted to functions $\chi_1(\mathbf{r}_1 \cdots \mathbf{r}_N)$ which are Fourier transformable. As in Eq. (5b), χ_1 is decomposed into a sum of n-body short-ranged functions, $n \leq N$:

$$\chi_1(\mathbf{r}_1 \cdots \mathbf{r}_N) = \sum_{n=1}^{N} \frac{1}{n!} \sum_{\substack{i_1 \cdots i_n \\ i_s \neq i_t}} u_n^{(1)}(\mathbf{r}_{i_1} \cdots \mathbf{r}_{i_n}). \tag{10}$$

The Fourier transforms of u_n are defined by

$$u_n^{(1)}(\mathbf{r}_1 \cdots \mathbf{r}_n) = N^{(1-n)} \sum_{\substack{\mathbf{k}_1 \cdots \mathbf{k}_n \\ \mathbf{k}_i \neq 0}} \left(\prod_{i=1}^{n} \exp(i\mathbf{k}_i \cdot \mathbf{r}_i) \right) C_n(\mathbf{k}_1 \cdots \mathbf{k}_n). \tag{11}$$

The restriction to nonzero k values in the sum in (11) is required by the cluster condition. The terms with one or more k equal to zero are included in $u_m^{(1)}$ with $m < n$. With this definition, Eq. (10) becomes

$$\chi_1(\mathbf{r}_1 \cdots \mathbf{r}_N) = \sum_{n=1}^{N} \frac{N^{(1-n)}}{n!} \sum_{\substack{\mathbf{k}_1 \cdots \mathbf{k}_n \\ \mathbf{k}_i \neq 0}} C_n(\mathbf{k}_1 \cdots \mathbf{k}_n) \rho^{(n)}(\mathbf{k}_1 \cdots \mathbf{k}_n) \tag{12}$$

where $\rho^{(n)}$ is the symmetrized plane wave[10]:

$$\rho^{(n)}(\mathbf{k}_1 \cdots \mathbf{k}_n) = \sum_{\substack{i_1 \cdots i_n \\ i_s \neq i_t}}^{N} \prod_{p=1}^{n} \exp(i\mathbf{k}_p \cdot \mathbf{r}_{i_p}). \tag{13}$$

The usefulness of this expansion of χ_1 in plane waves lies in the connection of these plane waves to products of density fluctuation functions $\rho_{\mathbf{k}}$ [defined in Eq. (2)] which have simple cluster properties. Thus we note that

$$\rho_{\mathbf{k}_1} = \rho^{(1)}(\mathbf{k}_1), \tag{14}$$

$$\rho_{\mathbf{k}_1}\rho_{\mathbf{k}_2} = \rho^{(2)}(\mathbf{k}_1, \mathbf{k}_2) + \rho^{(1)}(\mathbf{k}_1 + \mathbf{k}_2), \tag{15}$$

$$\begin{aligned}\rho_{\mathbf{k}_1}\rho_{\mathbf{k}_2}\rho_{\mathbf{k}_3} = {} & \rho^{(3)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + \rho^{(2)}(\mathbf{k}_1 + \mathbf{k}_2, \mathbf{k}_3) \\ & + \rho^{(2)}(\mathbf{k}_1 + \mathbf{k}_3, \mathbf{k}_2) + \rho^{(2)}(\mathbf{k}_2 + \mathbf{k}_3, \mathbf{k}_1) \\ & + \rho^{(1)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3),\end{aligned} \tag{16}$$

etc., and

$$\prod_{i=1}^{n} \rho_{\mathbf{k}_i} = \sum_{l=1}^{n} \rho^{(l)} \quad \text{(all distinct partitions of the k's)}. \tag{17}$$

The inverse expression gives $\rho^{(n)}(\mathbf{k}_1, \ldots, \mathbf{k}_n)$ as an nth order polynomial in $\rho_{\mathbf{k}}$ (see Appendix A), which can be used in (12) to obtain

$$\chi_1(\mathbf{r}_1 \cdots \mathbf{r}_n) = \sum_{n=1}^{N} \sum_{\substack{(\mathbf{k}_1 \cdots \mathbf{k}_n) \\ \mathbf{k}_i \neq 0}}{}' y_n(\mathbf{k}_1 \cdots \mathbf{k}_n) \prod_{i=1}^{n} \rho_{\mathbf{k}_i} \tag{18}$$

where the functions y_n are linear combinations of $C_n(\mathbf{k}_1 \cdots \mathbf{k}_n)$. The prime on the second summation restricts the sum to distinct sets $(\mathbf{k}_1 \cdots \mathbf{k}_n)$. The procedure for generating the y_n's from a given χ_1 is described in Appendix A.

With this expression for χ_1, the normalization integral (partition function) for W_N is [using Eq. (9)]

$$Q_N = \int \exp\chi(\mathbf{r}_1 \cdots \mathbf{r}_N)\, d\mathbf{r}_1 \cdots d\mathbf{r}_N = Q_N^{(0)}\, J_N[\chi_1], \tag{19}$$

where

$$J_N[\chi_1] = \left\langle \exp\left(\sum_{n=1}^{N} \sum_{(\mathbf{k}_1 \cdots \mathbf{k}_n)}{}' y_n(\mathbf{k}_1 \cdots \mathbf{k}_n) \prod_{i=1}^{n} \rho_{\mathbf{k}_i} \right) \right\rangle_0 \tag{20}$$

where $\langle \cdots \rangle_0$ means the average taken with respect to the normalized N-body probability $W_N^{(0)}(r_1 \cdots r_N)$, defined in Eq. (6).

To evaluate J_N in terms of the y_n, it is convenient to introduce the cumulants of the $\rho_{\mathbf{k}}$ functions. The average value of a product of n $\rho_{\mathbf{k}}$'s is defined as

$$I_n(\mathbf{k}_1 \cdots \mathbf{k}_n) = \left\langle \prod_{i=1}^{n} \rho_{\mathbf{k}_i} \right\rangle = \int \prod_{i=1}^{n} \rho_{\mathbf{k}_i}\, W_N(\mathbf{r}_1 \cdots \mathbf{r}_N)\, d\mathbf{r}_1 \cdots d\mathbf{r}_N \tag{21}$$

where

$$W_N(\mathbf{r}_1 \cdots \mathbf{r}_N) = \exp[\chi(\mathbf{r}_1 \cdots \mathbf{r}_N)]/Q_N. \tag{22}$$

A subscript or superscript "0" will be afixed to indicate averaging with respect to $W_N^{(0)}$. Then the nth cumulant $U_n(\mathbf{k}_1 \cdots \mathbf{k}_n)$ is defined by the first n equations of the following sequence[13]:

$$I_1(\mathbf{k}_1) = U_1(\mathbf{k}_1)$$

$$I_2(\mathbf{k}_1, \mathbf{k}_2) = U_1(\mathbf{k}_1)U_1(\mathbf{k}_2) + U_2(\mathbf{k}_1, \mathbf{k}_2)$$

$$I_3(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = U_1(\mathbf{k}_1)U_1(\mathbf{k}_2)U_1(\mathbf{k}_3) + U_1(\mathbf{k}_1)U_2(\mathbf{k}_2, \mathbf{k}_3) + U_1(\mathbf{k}_2)U_2(\mathbf{k}_1\mathbf{k}_3) + U_1(\mathbf{k}_3)U_2(\mathbf{k}_1\mathbf{k}_2) + U_3(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3).$$

$$\vdots$$

$$I_n(\mathbf{k}_1 \cdots \mathbf{k}_n) = \sum_{0 \le n_1 \cdots \le n_n \atop \sum n_i = n} \sum_{\{P_{1,1} \cdots P_{n,n_\mu}\}} \prod_{i=1}^{n} U_{n_i}(\mathbf{k}_{P_{i,1}} \cdots \mathbf{k}_{P_{i,n_i}}) \tag{23}$$

where $\{\ \}$ means the sum over all permutations of $(1 \cdots n)$ which produce a distinct term in the sum when $\mathbf{k}_i$ are treated as independent variables. Wu[13] has shown that if the cumulants $f_n(r_1 \cdots r_n)$ of the n-body distribution functions $g_n(r_1 \cdots r_n)$ satisfy the cluster condition, then the U_n are of order N. This result allows us to explicitly keep track of the order of magnitude in N of the expressions of interest. It was also shown in Ref. 13 that the Fourier transforms of g_n are linear combinations of the $I_m(\mathbf{k}_1 \cdots \mathbf{k}_m)$ with $m \le n$; thus g_n can be expressed in terms of U_m with $m \le n$. With that in mind, approximations for U_m are developed in Sec. IV, which in turn amount to approximations for g_n. Some of these approximations take the form of integral equations for the $U_n(\mathbf{k}_1 \cdots \mathbf{k}_n)$ in terms of $U_p^{(0)}(\mathbf{h}_1 \cdots \mathbf{h}_p)$ and $y_q(\mathbf{l}_1 \cdots \mathbf{l}_q)$.

From a physical point of view these cumulants U_n are probably better described as static correlation functions. For example, in a uniform liquid, $U_2(\mathbf{k}, -\mathbf{k})$ is essentially the liquid structure function $S(k)$ measured by x-ray scattering and integrated neutron scattering:

$$U_2(\mathbf{k}, -\mathbf{k}) = NS(k).$$

Both U_2 and U_4 are related to the dynamic correlation function which is detected in second-order Raman scattering:

$$\mathcal{J}(\omega, \boldsymbol{\epsilon}) = \sum_{\mathbf{k},\mathbf{l}} t(\mathbf{k})t(\mathbf{l})S_4(\mathbf{k}, -\mathbf{k}-\boldsymbol{\epsilon}, \mathbf{l}, -\mathbf{l}+\boldsymbol{\epsilon}, \omega),$$

where

$$\int d\omega\, S_4(\mathbf{k}, -\mathbf{k}-\boldsymbol{\epsilon}, \mathbf{l}, -\mathbf{l}+\boldsymbol{\epsilon}, \omega) = (\delta_{\mathbf{k},\mathbf{l}} + \delta_{\mathbf{k},-\mathbf{l}+\boldsymbol{\epsilon}})U_2(\mathbf{k}, -\mathbf{k})U_2(\mathbf{k}+\boldsymbol{\epsilon}, -\mathbf{k}-\boldsymbol{\epsilon}) + U_4(\mathbf{k}, -\mathbf{k}-\boldsymbol{\epsilon}, \mathbf{l}, -\mathbf{l}+\boldsymbol{\epsilon})$$

where $\mathcal{J}(\omega, \boldsymbol{\epsilon})$ is the light intensity measured at frequency shift ω and momentum shift $\boldsymbol{\epsilon}$, and $t(\mathbf{k})$ is related to the coupling between the light and the liquid.[12]

III. LINKED CLUSTER THEOREMS

The functional $J_N[\chi_1]$ is a generator for the functions I_n. The simplest expression for I_n in terms of $J_N[\chi_1]$ is

$$I_n(\mathbf{k}_1 \cdots \mathbf{k}_n) = \frac{\partial}{\partial y_n(\mathbf{k}_1 \cdots \mathbf{k}_n)} \ln J_N[\chi_1]. \tag{24}$$

A more useful result, however, is that $J_N[\chi_1]$ is a generating function for the cumulants U_n. To show this, it is convenient to introduce a new notation for the cumulant expansion; Eq. (23) becomes

$$I_n(\mathbf{k}_1 \cdots \mathbf{k}_n) = \sum_{0 \le n_1 \cdots \le n_n} \left[\prod_{i=1}^{n} U_{n_i}(-)\right]\{\mathbf{k}_1 \cdots \mathbf{k}_n\} \tag{25}$$

where the notation

$$\left[\prod_{i=1}^{n} U_{n_i}(-)\{\mathbf{k}_1 \cdots \mathbf{k}_n\}\right]$$

refers to the sum of all functionally distinct terms obtained by distributing the vectors $\mathbf{k}_i$ in the arguments of U_{n_j} [denoted by $(-)$]. This can be rewritten as

$$U_n(\mathbf{k}_1 \cdots \mathbf{k}_n) = I_n(\mathbf{k}_1 \cdots \mathbf{k}_n) - \sum_{p=1}^{n-1} [U_{n-p}(\mathbf{k}_1, -)I_p(-)]\{\mathbf{k}_2 \cdots \mathbf{k}_n\}. \tag{26}$$

(The choice of $\mathbf{k}_1$ in U_{n-p} is arbitrary.) Another equation for U_n is obtained by differentiating (26) with respect $y_1(\mathbf{k})$. Then the right-hand side involves $\partial I_p/\partial y_1$, which can be obtained by differentiating (21), using (22), (18), and (19):

$$\frac{\partial}{\partial y_1(\mathbf{k})} I_p(\mathbf{h}_1 \cdots \mathbf{h}_p) = I_{p+1}(\mathbf{h}_1 \cdots \mathbf{h}_p, \mathbf{k}) - U_1(\mathbf{k})I_p(\mathbf{h}_1 \cdots \mathbf{h}_p) \tag{27}$$

where the second term comes from the derivative of the denominator in (22). The expression obtained by using (27) for $\partial I_p/\partial y_1$ in the derivative of (26) can be simplified by using (26) for I_{n+1} to obtain

$$\frac{\partial}{\partial y_1(\mathbf{k}_{n+1})} U_n(\mathbf{k}_1 \cdots \mathbf{k}_n) = U_{n+1}(\mathbf{k}_1 \cdots \mathbf{k}_{n+1}) + \sum_{p=1}^{n-1} \left[I_p(-)\left(U_{n-p+1}(\mathbf{k}_1, \mathbf{k}_{n+1}, -) - \frac{\partial}{\partial y_1(\mathbf{k}_{n+1})} U_{n-p}(\mathbf{k}_1, -)\right)\right] \times \{\mathbf{k}_2 \cdots \mathbf{k}_n\}. \tag{28}$$

It is readily verified by induction that Eq. (28) gives

$$\frac{\partial}{\partial y_1(\mathbf{k}_{n+1})} U_n(\mathbf{k}_1 \cdots \mathbf{k}_n) = U_{n+1}(\mathbf{k}_1 \cdots \mathbf{k}_n, \mathbf{k}_{n+1}) \tag{29}$$

from which it follows by using (24) and (25) for $n = 1$ and then using (29) $n-1$ times that $\ln J_N$ is a generating function for U_n:

$$\left(\prod_{i=1}^{n} \frac{\partial}{\partial y_1(\mathbf{k}_i)}\right) \ln J_N = U_n(\mathbf{k}_1 \cdots \mathbf{k}_n). \tag{30}$$

The first of a set of linked cluster theorems may be obtained immediately from this result. We set

$$y_n = 0, \quad n > 1, \tag{31}$$

so that J_N is a function of only $y_1(\mathbf{k})$. Equation (30) is a set of differential equations for J_N, with the boundary condition

$$\ln J_N = y_0 \quad \text{if} \quad y_1(\mathbf{k}) = 0, \quad \text{all } \mathbf{k}.$$

The most straightforward way of solving (30) is to note that it gives all of the coefficients of a Taylor series in y_1 for $\ln J_N$ when $y_1(\mathbf{k})$ is set equal to zero. That is, ignoring questions of convergence,

$$\ln J_N = y_0 + \sum_{p=1}^{\infty} \frac{1}{p!} \sum_{\mathbf{k}_1 \cdots \mathbf{k}_p} \left(\prod_{i=1}^{p} y_1(\mathbf{k}_i) \right) U_p^{(0)}(\mathbf{k}_1 \cdots \mathbf{k}_p) \tag{32}$$

where $U_n^{(0)}$ is defined by (21) and (23), with $W_N^{(0)}$ in (21).

The sense in which (32) is a linked cluster series will be clear when the generalization of (32) [lifting the restriction (31)] is exhibited below. It is easy to see that (32) shares an important property with linked cluster expansions, which is that corrections to the thermodynamic potentials are manifestly extensive. For example, $\ln J_N$ is the correction to the thermodynamic potential for a classical system (when J_N is the partition function) due to the addition of an external single body potential $\phi_1(\mathbf{r})$ with Fourier transform $-k_B T y_1(\mathbf{k})$. To see that $\ln J_N = O(N)$ consider a uniform zeroth order system; then $U_n^{(0)}$ vanishes unless its momenta sum to zero, in which case it is of order N. Thus each nonzero term in (32) has a factor N implicit in $U_p^{(0)}$. Furthermore, since $\phi_1(\mathbf{r})$ is of order unity the sums over k in (32) do not increase the order in N of $\ln J_N$ [see Eq. (11) and Appendix A], nor does the sum over p. The easiest example to consider is an external periodic potential for which $y_1(\mathbf{k})$ will only be nonzero at reciprocal lattice vectors, and generally goes to zero rapidly with increasing reciprocal lattice vector. Then clearly $\ln J_N$ is proportional to N.

An application of this result to a quantum mechanical problem is the determination of the effect of a periodic substrate on absorbed liquid He^4. Details of that application are given elsewhere. [14]

To obtain a general linked cluster expansion, we remove restriction (31) so that J_N is a function of all the variables y_n. Using (24) in the left-hand side of (27), interchanging the order of differentiation and using (30) gives

$$\frac{\partial}{\partial y_n(\mathbf{k}_1 \cdots \mathbf{k}_n)} U_1(\mathbf{h}) = I_{n+1}(\mathbf{h}, \mathbf{k}_1 \cdots \mathbf{k}_n) - U_1(\mathbf{h}) I_n(\mathbf{k}_1 \cdots \mathbf{k}_n). \tag{33}$$

We use (26) to eliminate I_{n+1} from this expression. The second term on the right-hand side of (33) cancels the last term in the sum on the right-hand side of (26) giving

$$\frac{\partial}{\partial y_n(\mathbf{k}_1 \cdots \mathbf{k}_n)} U_1(\mathbf{h}) = \sum_{j=0}^{n-1} [U_{n+1-j}(\mathbf{h}, -) I_j(-)] \{\mathbf{k}_1 \cdots \mathbf{k}_n\}. \tag{34}$$

Then using (25) for I_j, the right-hand side of (34) is simply the sum over all products of U_p's containing distinct partitions of the set of variables $\mathbf{h}, \mathbf{k}_1, \ldots, \mathbf{k}_n$ and connected by the set of variables $(\mathbf{k}_1 \cdots \mathbf{k}_n)$. It is convenient to introduce a new notation to represent a connected product of U's:

$$[\prod_j U_{n_j}]_c \{(\mathbf{h}_1 \cdots \mathbf{h}_{p_1})(\mathbf{k}_1 \cdots \mathbf{k}_{p_2}) \cdots (\mathbf{l}_1 \cdots \mathbf{l}_{p_n})\} \tag{35}$$

means the sum over all distinct partitions of the momenta in braces in the arguments of the U_{n_j} factors in such a way that the product over j is connected by the subsets of momenta in parentheses. To illustrate:

$$[U_1 U_2 U_2]_c \{(\mathbf{h}), (\mathbf{k}_1, \mathbf{k}_2), (\mathbf{l}_1, \mathbf{l}_2)\} = \sum_{\alpha \neq \beta} \sum_{i \neq j} U_1(\overbrace{\mathbf{l}_\alpha}) U_2(\overbrace{\mathbf{l}_\beta, \mathbf{k}_i}) U_2(\mathbf{h}, \overbrace{\mathbf{k}_j}) + \sum_{\alpha \neq \beta} \sum_{i \neq j} U_1(\overbrace{\mathbf{k}_i}) U_2(\overbrace{\mathbf{k}_j, \mathbf{l}_\alpha}) U_2(\mathbf{h}, \overbrace{\mathbf{l}_\beta})$$

where $\overbrace{\quad}$ aid the eye in observing that all terms are connected. A term which is disconnected and therefore does not appear in the sum is

$$U_1(\mathbf{h}) U_2(\overbrace{\mathbf{l}_\alpha, \mathbf{k}_i}) U_2(\overbrace{\mathbf{l}_\beta, \mathbf{k}_j}),$$

since the first factor is not connected to the last two by any momentum set. Note that a single momentum in parenthesis within the braces in (35) does not provide any connections between the U factors. Thus (34) may be rewritten as

$$\frac{\partial U_1(\mathbf{h}_1)}{\partial y_n(\mathbf{k}_1 \cdots \mathbf{k}_n)} = \sum_{0 \leq n_1 \cdots \leq n_{n+1}} \left[\prod_{i=1}^{n+1} U_{n_i} \right]_c \{(\mathbf{h}_1), (\mathbf{k}_1 \cdots \mathbf{k}_n)\}. \tag{36}$$

Using (29), differentiating (36) $p-1$ times with respect to y_1 and exchanging the orders of differentiation gives

$$\frac{\partial}{\partial y_n(\mathbf{k}_1 \cdots \mathbf{k}_n)} U_p(\mathbf{h}_1 \cdots \mathbf{h}_p) = \sum_{\substack{\text{distinct partitions} \\ \text{of } n+p}} \left[\prod_{i=1}^{n+p} U_{n_i} \right]_c \times \{(\mathbf{h}_1) \cdots (\mathbf{h}_p)(\mathbf{k}_1 \cdots \mathbf{k}_n)\}. \tag{37}$$

The right-hand side of (37) is the sum over all functionally distinct partitions of $n+p$ and momenta $\mathbf{h}_i$ and $\mathbf{k}_j$ which are connected by the set $(\mathbf{k}_1 \cdots \mathbf{k}_n)$. Thus differentiation of a cumulant by $y_n(\mathbf{k}_1 \cdots \mathbf{k}_n)$ introduces all combinations of cumulants which are connected only by the new momentum set. It follows by using (24), (25), and (37) t times that

$$\prod_{i=1}^{t} \frac{\partial}{\partial y_{m_i}(\mathbf{k}_1^{(i)} \cdots \mathbf{k}_{m_i}^{(i)})} \ln J_N = \sum_{\substack{n_1 \leq n_2 \cdots \leq n_l \\ \Sigma n_j = \Sigma m_i}} [\prod_j U_{n_j}]_c \{(\mathbf{k}_1^{(1)} \cdots \mathbf{k}_{m_1}^{(1)}) \cdots (\mathbf{k}_1^{(t)} \cdots \mathbf{k}_{m_t}^{(t)})\}. \tag{38}$$

Note that Eq. (30) is a special case of (38) since in (30) $m_i = 1$ for all i, and thus the connecting momentum sets each contain but one momentum, $(\mathbf{k}_i)$, and cannot provide any connection. In that case the product of U's in (38) can contain but one factor, as seen on the right-hand side of (30).

As in the development of (32), we can use (38) to obtain all the Taylor series coefficients for the $\ln J_N$ in terms of all the variables y_n by setting $y_n = 0$ for all n in (38). The only consequence is that all U_n becomes $U_n^{(0)}$ on the right-hand side of (38). Then the series for $\ln J_N$ is

$$\ln J_N = y_0 + \prod_{i=1}^{N} \left[\sum_{n_i} \frac{1}{n_i!} \prod_{j=1}^{n_i} \left(\sum_{(\mathbf{k}_{(ij)}^{(1)} \cdots \mathbf{k}_{(ij)}^{(i)})}{}' y_i(\mathbf{k}_{(ij)}^{(1)} \cdots \mathbf{k}_{(ij)}^{(i)}) \right) \right] \otimes \sum_{\substack{m_1 \leq m_2 \leq \cdots \leq m_p \leq \cdots \\ \Sigma m_l = \Sigma n_i}} [\prod_l U_{m_l}^{(0)}]_c \times \{(\mathbf{k}_{(1,1)}^{(1)}), (\mathbf{k}_{(1,2)}^{(1)}), \cdots (\mathbf{k}_{(1,n_1)}^{(1)}) \cdots (\mathbf{k}_{(\alpha\beta)}^{(\alpha)} \cdots \mathbf{k}_{(\alpha\beta)}^{(\alpha)}) \cdots\}. \tag{39}$$

The prime on the summation sign means that the sum

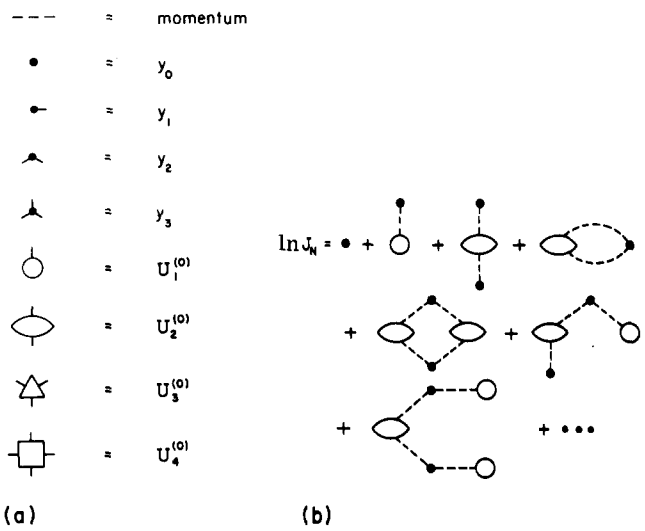


FIG. 1. (a) The elements of a diagrammatic expansion. y_n is an n-pronged vertex and U_n is an n-sided polygon. (b) A schematic expansion of the logarithm of the normalization integral. Momentum must be assigned to the dashed lines, ----, and summed over. Appropriate combinatorial factors must be assigned to each term. (See the discussion in the text.)

is to be restricted to run over distinct sets of momenta $(\mathbf{k}_{(ij)}^{(1)} \cdots \mathbf{k}_{(ij)}^{(i)})$.

This complicated expression is simplified by introducing a diagrammatic representation, which is defined in Fig. 1: A dashed line represents a momentum, a vertex composed of n prongs connected at a dot represents y_n and an m-sided polygon represents $U_m^{(0)}$. The momentum lines connect dots to polygons by connecting a prong to a side of the polygon. Then a connected term is a diagram which is topologically connected. Then $\ln J_N$ is a sum over all connected diagrams with no free momentum lines. The counting factor for each diagram is chosen so that any distribution of the momentum variables which is not equivalent to the first by the permutation of the $U_m^{(0)}$'s or the arguments of the $U_m^{(0)}$ or a combination of both must be considered a new term. Furthermore, there is a factor of $1/n_i!$ associated with every diagram containing n_i vertices y_i. It might seem that the $n_i!$ permutations of the y_i would cancel these factors, but some of these permutations are equivalent to permutations of the $U_m^{(0)}$ and/or their arguments. This counting argument is sufficiently tedious that it is best to consider special cases, as is done in the next section.

The diagrammatic expansion for $U_n(\mathbf{h}_1 \cdots \mathbf{h}_n)$ is obtained by differentiating $\ln J_N$ with respect to $y_1(\mathbf{h}_i)$, $i = 1 \to n$ (Eq. 30). This is equivalent to removing n single prong vertices in each term of $\ln J_N$ in Fig. 1 in all possible ways, and labelling the dashed lines from which these have been removed by $\mathbf{h}_1 \cdots \mathbf{h}_n$. The diagram rules for U_n are identical to those for $\ln J_N$ as can be seen by noting that (37) can be used to generate the Taylor series for $U_n(\mathbf{h}_1 \cdots \mathbf{h}_n)$, which has the same form as (39) but with $(\mathbf{h}_1), (\mathbf{h}_2), \ldots, (\mathbf{h}_n)$ included in the braces.

The sense in which these expressions for $\ln J_N$ and U_n are linked cluster expansions is clear from the diagrams as well as the notation in Eq. (39). It is not quite so obvious that each of the linked diagrams produces a term of order N. Nevertheless, $\ln J_N$ must be of order N since it is the shift in thermodynamic potential in the classical system; the U_n are also proportional to N.[13] It is one of the main advantages of linked cluster theorems that the contribution of each diagram is the correct order of magnitude in N, so that further approximations do not need to be concerned with keeping track of powers of N. To see that each diagram contributes $O(N)$ to the expansion for $\ln J_N$, consider a particular linked diagram and the corresponding momentum sums in Eq. (39). If we focus on a single $y_n(\mathbf{k}_1 \cdots \mathbf{k}_n)$ and the associated sums over $\mathbf{k}_1 \cdots \mathbf{k}_n$, we can use the fact that $u_n(\mathbf{r}_1 \cdots \mathbf{r}_n)$ is of order unity (which is to say that it does not change as N increases) to conclude that y_n and the n k sums contribute order unity to the diagram if all other constraints on the momentum sums (which come from the $U_m^{(0)}$ factors) are ignored. Now, if we consider the case where the zeroth order problem is translationally invariant, so that $U_m^{(0)}$ is zero unless the sum of the momenta in its arguments vanish, then each $U^{(0)}$ introduces a single constraint on the momentum sums, reducing the overall N dependence by one order of magnitude. But $U_m^{(0)}$ is of order N, cancelling out the effect of this constraint. For a linked factor, however, one (and only one) of these momentum constraints is equivalent to overall momentum conservation, which is already guaranteed by the y_n's within this linked cluster (supposing that the perturbation is also translationally invariant), thereby leaving one uncancelled factor of N. Thus each linked factor contributes $O(N)$, and it is the fact that there is just one linked factor in each term in the above expansion which shows that each diagram is $O(N)$. A similar argument can be made for nontranslationally invariant systems, except that the role of momentum conservation must be replaced by restrictions on the possible values of the total momentum in $U_m^{(0)}$ or y_n. For example, in a periodic system these totals must always be reciprocal lattice vectors. The fact that these possibilities do not increase with increasing N plays the same role as the momentum conservation argument above.

IV. APPROXIMATIONS

There are numerous possible approximation schemes involving truncations and/or resummations of (39). In this section we discuss several of the approximations in differing detail, deferring the application of these approximations to physical problems of interest until later.

A. Finite series truncation

The simplest approximation is to truncate (39) after a finite number of terms in each of the variables y_n. For example, if only single-body terms are retained (i.e., $y_n = 0$, $n \neq 1$) and the series is truncated beyond second order, then

$$\ln J_N \approx y_0 + \sum_{\mathbf{k}} U_1^{(0)}(\mathbf{k}) y_1(\mathbf{k}) + \tfrac{1}{2} \sum_{\mathbf{k},\mathbf{l}} U_2^{(0)}(\mathbf{k},\mathbf{l})\, y_1(\mathbf{k})\, y_1(\mathbf{l}). \qquad (40)$$

This is the approximation used elsewhere to obtain the effects of a periodic substrate potential on adsorbed helium.[14]

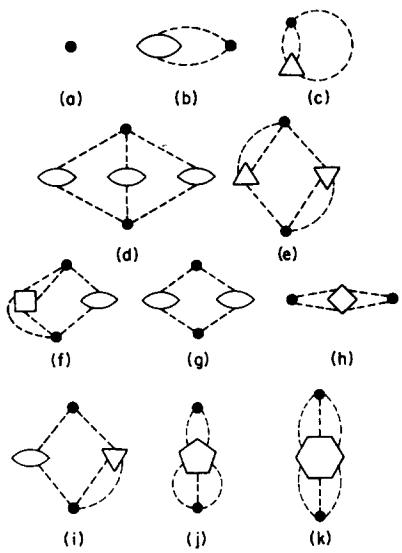


FIG. 2. The diagrams for $\ln J_N$ which are linear and bilinear in y_2 and y_3. Expressions for each term are given in Eqs. (B4a)—(B4k), Appendix B.

As another example of finite series truncation, consider the inclusion of a p-body potential in the classical problem or a p-body term in an extended Jastrow wavefunction, so that

$$\chi_1 = \frac{1}{p!} \sum_{\substack{i_1 \cdots i_p \\ i\alpha \neq i\beta}} u_p^{(1)}(\mathbf{r}_{i_1} \cdots \mathbf{r}_{i_p}). \tag{41}$$

Introducing the Fourier transform of u_p gives Eq. (12) with only the $n=p$ term in the first summation. To obtain the expression of χ_1 in terms of density fluctuation operators, $\rho^{(p)}(\mathbf{k}_1 \cdots \mathbf{k}_p)$ must be expressed as a pth order polynomial in $\rho_{\mathbf{k}}$ (the general result for $\rho^{(p)}$ is given in Appendix A). Then Eq. (18) will have $y_n \neq 0$, $n \leq p$. Expressions for y_n for the cases $p=2$ and $p=3$ are also given in Appendix A. Consider the case $p=3$ and suppose u_3 is translationally invariant. Then the expression for χ_1 in (A4)—(A8) involves only terms with total momentum zero:

$$\chi_1 = y_0 + \sum_{\mathbf{k}, k_x > 0} y_2(\mathbf{k}, -\mathbf{k})\rho_{\mathbf{k}}\rho_{-\mathbf{k}} + \sum_{\substack{(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3) \\ \mathbf{k}_i \neq 0}}' y_3(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3)\rho_{\mathbf{k}_1}\rho_{\mathbf{k}_2}\rho_{\mathbf{k}_3}, \tag{42}$$

y_0, y_2 and y_3 are all linear in C_3. Thus truncation of the expansion for $\ln J_N$ at second order in C_3 requires retention of terms which are second order or lower in y_3 and y_2. Diagrams for the linear and second order terms in C_3 are shown in Fig. 2. Expressions for each diagram are given in Appendix B. Note that a complete evaluation of these terms requires $U_2^{(0)}$, $U_3^{(0)}$, $U_4^{(0)}$, $U_5^{(0)}$, and $U_6^{(0)}$ since each of the terms are $O(N)$. Nothing is known about $U_4^{(0)}$, $U_5^{(0)}$, or $U_6^{(0)}$ for any interacting system.

B. Truncation of higher correlations

Since $U_n^{(0)}$ is unknown beyond some value of n (usually two), it is necessary to approximate the $U_n^{(0)}$ in terms of other known functions. There are several approximations which have been used for this truncation. The simplest is to set $U_n^{(0)}$ equal to zero beyond some value of n. For example,

$$U_n^{(0)} = 0, \quad n > 2. \tag{43}$$

If only y_1 is nonzero, this approximation is equivalent to Eq. (40) for $\ln J_N$. The two procedures give different expressions for U_n, however.

A more interesting example is the case when χ_1 contains both one- and two-body potentials or Jastrow factors, so that both $y_1(\mathbf{k})$ and $y_2(\mathbf{k}, \mathbf{l})$ are nonzero. In that case (39) still has an infinite number of terms even when approximation (43) is used. With only $U_1^{(0)}$, $U_2^{(0)}$, y_1, and y_2 present, the topology of the diagrams and the rules for the diagrams simplify considerably. The four types of diagrams, shown in Fig. 3, are those which contain all U_2's and y_2's, those which contain one y_1 and one U_1, those which contain two y_1's and those which contain two U_1's. The corresponding equation for $\ln J_N$ is

$$\begin{aligned}
\ln J_N = {} & \sum_{p=1}^{\infty} \frac{1}{p!} \prod_{i=1}^{p} \left(\sum_{(\mathbf{h}_i \mathbf{q}_i)}' y_2(\mathbf{h}_i \mathbf{q}_i) \right) [(U_2^{(0)})^p]_c \{(\mathbf{h}_1\mathbf{q}_1) \cdots (\mathbf{h}_p\mathbf{q}_p)\} \\
& + \sum_{\mathbf{k}} y_1(\mathbf{k}) \sum_{p=0}^{\infty} \frac{1}{p!} \prod_{i=1}^{p} \left(\sum_{(\mathbf{h}_i \mathbf{q}_i)}' y_2(\mathbf{h}_i \mathbf{q}_i) \right) [U_1^{(0)}(U_2^{(0)})^p]_c \\
& \times \{(\mathbf{k})(\mathbf{h}_1\mathbf{q}_1) \cdots (\mathbf{h}_p\mathbf{q}_p)\} \\
& + \tfrac{1}{2} \sum_{\mathbf{k}_1\mathbf{k}_2} y_1(\mathbf{k}_1) y_1(\mathbf{k}_2) \sum_{p=0}^{\infty} \frac{1}{p!} \prod_{i=1}^{p} \left(\sum_{(\mathbf{h}_i \mathbf{q}_i)}' y_2(\mathbf{h}_i \mathbf{q}_i) \right) \\
& \times [(U_2^{(0)})^{p+1}]_c \{(\mathbf{k}_1)(\mathbf{k}_2)(\mathbf{h}_1\mathbf{q}_1) \cdots (\mathbf{h}_p\mathbf{q}_p)\} \\
& + \sum_{p=1}^{\infty} \frac{1}{p!} \prod_{i=1}^{p} \left(\sum_{(\mathbf{h}_i \mathbf{q}_i)}' y_2(\mathbf{h}_i \mathbf{q}_i) \right) \\
& \times [(U_1^{(0)})^2 (U_2^{(0)})^{p-1}]_c \{(\mathbf{h}_1\mathbf{q}_1) \cdots (\mathbf{h}_p\mathbf{q}_p)\}.
\end{aligned} \tag{44}$$

Succeeding lines correspond to diagrams (a), (b), (c), and (d), respectively, in Fig. 3. To simplify the expressions, assign the momenta to specific dashed lines in the figure. The prime on the sums can be removed in the first line by noting that each pair $(\mathbf{h}_i, \mathbf{q}_i)$ gives rise to two equivalent terms except the first, which must have a factor $\frac{1}{2}$ to remove the prime. Similarly, the $1/p!$ becomes $1/p$ from permutations of the y_2's which

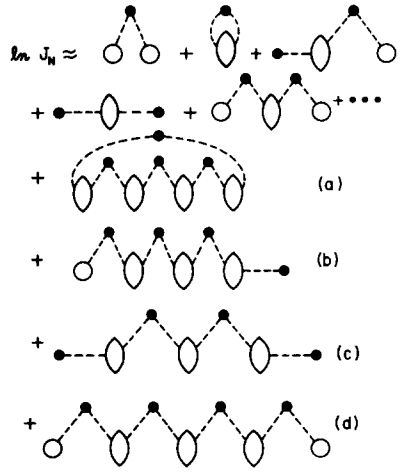


FIG. 3. The diagrams for $\ln J_N$ containing y_1, y_2, U_1, and U_2. The four basic types of diagrams are given in (a)—(d).

reproduce the term only after a change of variables. The $1/p$ comes from cyclic permutations which leaves the original term invariant. Then

$$(\ln J_N)_a = \tfrac{1}{2}\sum_{p=1}^{\infty}\frac{1}{p!}\sum_{h_1q_1\cdots h_pq_p} y_2(h_1,q_1)U_2^{(0)}(q_1h_2)y_2(h_2q_2) \times U_2^{(0)}(q_2h_3)\cdots y_2(h_pq_p)U_2^{(0)}(q_ph_1). \tag{45a}$$

Similar arguments eliminate the $p!$ in (b), (c) and (d) as well as all of the primes on the sums in (b) and (c) and all but one in (d); removing the last prime in (d) requires an additional factor of $\frac{1}{2}$. Then

$$(\ln J_N)_b = \sum_{\mathbf{k}}\sum_{p=0}^{\infty}\sum_{h_1q_1\cdots h_pq_p} U_1^{(0)}(h_1)\,y_2(h_1q_1)\,U_2^{(0)}(q_1h_2)\cdots y_2(h_pq_p) \times U_2^{(0)}(q_p\mathbf{k})\,y_1(\mathbf{k}), \tag{45b}$$

$$(\ln J_N)_c = \tfrac{1}{2}\sum_{\mathbf{k}_1\mathbf{k}_2}\sum_{p=0}^{\infty}\sum_{h_1q_1\cdots h_pq_p} y_1(\mathbf{k}_1)\,U_2^{(0)}(\mathbf{k}_1h_1)\,y_2(h_1q_1) \times U_2^{(0)}(q_1h_2)\cdots y_2(h_pq_p)U_2^{(0)}(q_p\mathbf{k}_2)\,y_1(\mathbf{k}_2), \tag{45c}$$

$$(\ln J_N)_d = \tfrac{1}{2}\sum_{p=1}^{\infty}\sum_{h_1q_1\cdots h_pq_p} U_1^{(0)}(h_1)\,y_2(h_1q_1)\,U_2^{(0)}(q_1h_2)\cdots y_2(h_pq_p) \times U_1^{(0)}(q_p). \tag{45d}$$

The cumulant $U_1(\mathbf{k}_1)$ is obtained by differentiating $\ln J_N$ with respect to $y_1(\mathbf{k}_1)$ and $U_2(\mathbf{k}_1,\mathbf{k}_2)$ is obtained from $U_1(\mathbf{k}_1)$ by differentiating with $y_1(\mathbf{k}_2)$ [Eq. (29)]. Note then that U_1 comes entirely from (45b) and (45c), and U_2 comes entirely from (45c). Diagrammatic expansions are given in Fig. 4. The expressions are

$$U_1(\mathbf{k}_1) = U_1^{(0)}(\mathbf{k}_1) + \sum_{p=1}^{\infty}\sum_{h_1q_1\cdots h_pq_p} U_1^{(0)}(h_1)\,y_2(h_1q_1) \times U_2^{(0)}(q_1h_2)\cdots y_2(h_pq_p)U_2^{(0)}(q_p\mathbf{k}_1) + \sum_{\mathbf{k}_2}\sum_{p=0}^{\infty}\sum_{h_1q_1\cdots h_pq_p} U_2^{(0)}(\mathbf{k}_1h_1)\,y_2(h_1q_1) \times U_2^{(0)}(q_1h_2)\cdots y_2(h_pq_p)U_2^{(0)}(q_p\mathbf{k}_2)\,y_1(\mathbf{k}_2) \tag{46}$$

and

$$U_2(\mathbf{k}_1,\mathbf{k}_2) = \sum_{p=0}^{\infty}\sum_{h_1q_1\cdots h_pq_p} U_2^{(0)}(\mathbf{k}_1h_1)\,y_2(h_1q_1) \times U_2^{(0)}(q_1h_2)\cdots y_2(h_pq_p)U_2^{(0)}(q_p\mathbf{k}_2). \tag{47}$$

Note that U_2 is independent of y_1 in this approximation. These two expressions can be summed formally to produce Dyson-like equations

$$U_1(\mathbf{k}_1) = U_1^{(0)}(\mathbf{k}_1) + \sum_{hq} U_1^{(0)}(h)\,y_2(hq)\,U_2(q\mathbf{k}_1) + \sum_{\mathbf{k}} y_1(\mathbf{k})\,U_2(\mathbf{k},\mathbf{k}_1), \tag{48}$$

$$U_2(\mathbf{k}_1,\mathbf{k}_2) = U_2^{(0)}(\mathbf{k}_1\mathbf{k}_2) + \sum_{hq} U_2^{(0)}(\mathbf{k}_1h)\,y_2(hq)\,U_2(q\mathbf{k}_2). \tag{49}$$

The consistency of these results can be checked by noting that $I_2(\mathbf{k}_1,\mathbf{k}_2)$ can be obtained from $\ln J_N$ by differentiation with respect to $y_2(\mathbf{k}_1,\mathbf{k}_2)$ [Eq. (24)], while from Eq. (23)

$$I_2(\mathbf{k}_1,\mathbf{k}_2) = U_1(\mathbf{k}_1)\,U_1(\mathbf{k}_2) + U_2(\mathbf{k}_1\mathbf{k}_2). \tag{50}$$

It is shown in Appendix C that (50) is indeed obtained from differentiation of (45) when (46) and (47) are used for U_1 and U_2 in (50).

Special cases of the results (45) and (49) have been obtained for classical systems by using for the zeroth order system the noninteracting gas[15] and the hard sphere gas.[16] These results have also been used to include phonon zero point motion in the ground state wavefunction of liquid He^4 (Ref. 17) and to optimize the Jastrow trial function for He^4.[7,18] In each of these cases the zeroth order system is translationally invariant so that

$$U_1^{(0)}(\mathbf{k}) = N\delta_{\mathbf{k},0}, \tag{51}$$

$$U_2^{(0)}(\mathbf{k}_1\mathbf{k}_2) = \delta_{\mathbf{k}_1+\mathbf{k}_2,0}NS_0(k_1), \tag{52}$$

where $S_0(k)$ is the liquid structure function for the zeroth-order system. Furthermore, the "perturbation" χ_1 is a sum of translationally invariant two-body functions

$$\chi_1(\mathbf{r}_1\cdots\mathbf{r}_N) = \tfrac{1}{2}\sum_{i<j}^{N} u_2^{(1)}(|\mathbf{r}_i-\mathbf{r}_j|), \tag{53}$$

so that

$$y_1(\mathbf{k}) = 0, \tag{54}$$

$$y_2(\mathbf{k}_1,\mathbf{k}_2) = \delta_{\mathbf{k}_1+\mathbf{k}_2,0}\frac{1}{N}C_2(k_1) \tag{55}$$

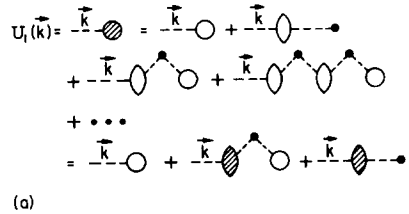
where C_2 is of order unity. Then (49) can be solved algebraically for U_2, giving the new liquid structure function

$$S(k) = \frac{1}{N}U_2(\mathbf{k},-\mathbf{k}) = S_0(k)/[1-C_2(k)S_0(k)]. \tag{56}$$

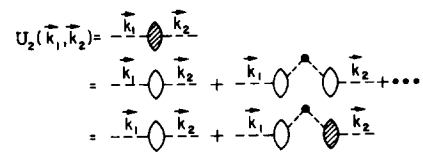
Furthermore, since y_1 is zero, $\ln J_N$ is given entirely by (45a), which, because of the Kronecker delta functions in (52) and (55), becomes

$$(\ln J_N)_a = \sum_{p=1}^{\infty}\frac{1}{p}\sum_{\mathbf{k},\mathbf{k}_\nu>0}[C_2(k)S_0(k)]^p = -\sum_{\mathbf{k},\mathbf{k}_\nu>0}\ln[1-C_2(k)S_0(k)]. \tag{57}$$

The result (56) first appears in the theory of classical



(a)



(b)

FIG. 4. Diagrams for (a) U_1 and (b) U_2 containing only y_1, y_2, U_1, and U_2.

fluids where the zeroth order system is composed of noninteracting particles and the perturbation is a Fourier transformable potential.[15] Then

$$S_0(k)=1,$$
$$C_2(k)=-\phi_2(k)/k_BT, \tag{58}$$

and thus

$$S(k)=\frac{1}{1+\beta\phi_2(k)}. \tag{59}$$

More recently, Eq. (56) has been used to include the long range correlations in the ground state wavefunction of liquid He^4 which come about because of the zero point motion of the long wavelength phonons.[17] In that case χ_0 is the logarithm of the square of a wavefunction chosen to give a good description of the short-range correlations of liquid He^4. The long range correlations are found by quantizing the phonon field. The phonon operator (the collective coordinate) is $\rho_{\mathbf{k}}$, and the ground state of the phonon field is a Gaussian in this collective coordinate. Then

$$\psi_{lrc}=\exp(\chi_1/2)=\prod_{\mathbf{k}}{}' \exp\left(-\tfrac{1}{2}\frac{2mc}{N\hbar k}\ |\rho_{\mathbf{k}}|^2\right) \tag{60}$$

where c is the sound velocity in He^4 and m is the He^4 mass. The prime on the product restricts the k values in the product to long wavelengths. Then $S_0(k)$ is the liquid structure function calculated from the short-ranged wavefunction

$$\psi_0=\exp(\chi_0/2) \tag{61}$$

and is very nearly the experimental liquid structure function except for small k, where it approaches a positive constant as $k\to 0$. Sum rule arguments show that the proper long wavelength limit in the ground state is[8]

$$S(k)\underset{k\to 0}{\to}\frac{\hbar k}{2mc}. \tag{62}$$

This behavior is obtained by multiplying (60) by (61):

$$\psi=\psi_0\psi_{lrc}.$$

In that case

$$v_2(k)=-2mc/N\hbar k$$

and

$$S(k)=\frac{S_0(k)}{1+(2mc/\hbar k)S_0(k)}. \tag{63}$$

Note that (62) is indeed the limit of (63) as k goes to zero.

The approximation given in (56) has also been used to obtain the Jastrow function u_2 [Eq. (1)] which minimizes the energy for liquid He^4 from the class of all Jastrow functions with $u_2(r)$ real.[7] The procedure, known as the *paired phonon analysis* since only second order cumulants are retained,[18] begins with a zeroth order Jastrow function which is obtained from some previous minimization procedure involving a small number of parameters[19]:

$$\chi_0=\sum_{i<j}^{N} u_2^{(0)}(r_{ij}).$$

χ_1 is chosen in the same form:

$$\chi_1=\sum_{i<j}^{N} u_2^{(1)}(r_{ij}),$$

so that minimizing the energy with respect to the function $u_2^{(1)}$ is equivalent to minimizing the energy with respect to u_2 in Eq. (1). The liquid structure function for ψ_J is given by (56), where $C_2(k)$ is now the Fourier transform of $u_2(r)$.

The error introduced by the neglect of $U_n^{(0)}$ for n larger than two is reduced by iterating the paired phonon analysis, $u_2=u_2^{(0)}+u_2^{(1)}$ becoming the new $u_2^{(0)}$.[7] With an appropriate choice for the initial $u_2^{(0)}$ these iterations converge rapidly to the optimum u_2.

The results (45) and (49) for uniform systems have also been obtained previously by Coopersmith and Brout to include the effects of a weak attractive potential when added to a hard core repulsion in a classical gas.[16] Then $S_0(k)$ is the hard core liquid structure function which is obtained approximately from an independent calculation.

In problems where the error introduced by setting $U_n^{(0)}=0$ for $n>2$ cannot be reduced by some procedure such as the iterations of the paired phonon analysis, a better truncation approximation may be necessary. The approximation (43) is not even valid for the noninteracting system. One approximation which may be an improvement but hasn't to our knowledge been used is to use the noninteracting evaluation of $U_n^{(0)}$ for $n>2$. The noninteracting value of $U_n^{(0)}$ is obtained from Ref. 13 where may be found the expression for U_n in terms of the Fourier transforms $F_p(\mathbf{k}_1\cdots\mathbf{k}_p)$ of the cumulants $f_p(\mathbf{r}_1\cdots\mathbf{r}_p)$ of the p-body distribution $g_p(\mathbf{r}_1\cdots\mathbf{r}_p)$ [see Eq. (8)], which vanish for the noninteracting system.[20] The result is that $U_n^{(0)}=N$ except for those sets of $(\mathbf{k}_1\cdots\mathbf{k}_n)$ for which the sum of a subset of the k's vanishes. That is also the limiting behavior of $U_n^{(0)}$ in the interacting system as $k_i\to\infty$, all i.

Neither of the above approximations give the long wavelength properties of $U_n^{(0)}$ even qualitatively correctly. An approximation which gives some long wavelength properties correctly is the convolution approximation,[21] which approximates U_n in terms of $U_2(\mathbf{k},-\mathbf{k})=NS(k)$. For example, the convolution approximation for U_3 and U_4 are, respectively,[21]

$$U_3^c(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3)=NS(k_1)S(k_2)S(k_3),$$
$$U_4^c(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{k}_4)=NS(k_1)S(k_2)S(k_3)S(k_4)[-2+S(\mathbf{k}_1+\mathbf{k}_2)+S(\mathbf{k}_1+\mathbf{k}_3)+S(\mathbf{k}_1+\mathbf{k}_4)].$$

The general expression for U_n^c is given in Ref. 13. The long range properties are expected to be reasonable because the convolution approximation is chosen as the simplest form which satisfies the sequential relation required by Eq. (8):

$$\rho\int g_{n+1}^c(\mathbf{r}_1\cdots\mathbf{r}_{n+1})\,d\mathbf{r}_{n+1}=(N-n)g_n^c(\mathbf{r}_1\cdots\mathbf{r}_n), \tag{64}$$

where g_n^c is the approximation for g_n deduced from U_n^c. U_n^c also has the proper high momentum limit. It should be pointed out, however, that there is an infinitude of other such approximations which satisfy (64).[22]

Still another truncation is the generalized Kirkwood superposition approximation,[18] which truncates the U_n by approximating the g_n in terms of g_2. For a translationally invariant system this is

$$g_p^K(\mathbf{r}_1 \cdots \mathbf{r}_p) = \prod_{i<j}^{p} g(r_{ij}) \tag{65}$$

where

$$g(r_{ij}) = g_2(\mathbf{r}_i, \mathbf{r}_j)$$

is the radial distribution function. Furthermore,

$$U_2(\mathbf{k}_1 \mathbf{k}_2) = N \delta_{\mathbf{k}_1+\mathbf{k}_2,0}\, S(k).$$

where

$$S(k) = 1 + \rho \int (g(r) - 1) \exp(i\mathbf{k} \cdot \mathbf{r}). \tag{66}$$

The general relationship between the set of functions U_n and the set g_p is given in Ref. 13. The consequence of (65) and (66) is that the resultant approximation for U_n is a sum of integrals involving only U_2. As in the convolution approximation, U_n^K is a functional of U_2, but the emphasis of the Kirkwood approximation is designed primarily to treat the short range behavior correctly. Specifically, when the particles under consideration have a strong, short range inter-particle repulsion, then (65) gives the correct qualitative result that g_p vanishes as any two of its arguments approach one another.

We have not applied the generalized convolution approximation or the generalized Kirkwood superposition approximation to the linked cluster expansion nor can we comment on whether the infinite series can be summed in a simple manner. It would be an interesting exercise to combine Wu's diagrammatic formulation of the convolution approximation in terms of Cayley trees[13] with the diagrammatic expansions derived here to see if there are any useful resummations for $\ln J_N$ or U_n.

ACKNOWLEDGMENTS

I would like to thank C.-C. Chang, J.W. Halley, P. Kleban, and T. Paulick for helpful comments on this manuscript.

APPENDIX A

A procedure for obtaining the expression of χ_1 in terms of powers of $\rho_{\mathbf{k}}$ [Eq. (18)] from the expression for χ_1 in terms of symmetrized plane waves [Eq. (12)] is given in this section. We first consider the specific cases for a single u_n in χ_1 with $n=2$ and $n=3$. The general procedure is outlined thereafter.

The first step is to obtain the expression for $\rho^{(n)}(\mathbf{k}_1 \cdots \mathbf{k}_n)$ as an nth order polynomial in $\rho_{\mathbf{k}}$. For $n=2$ and $n=3$, this means inverting equations (15) and (16), respectively, giving

$$\rho^{(1)}(\mathbf{k}_1) = \rho_{\mathbf{k}_1}, \tag{A1}$$

$$\rho^{(2)}(\mathbf{k}_1\mathbf{k}_2) = \rho_{\mathbf{k}_1}\rho_{\mathbf{k}_2} - \rho_{\mathbf{k}_1+\mathbf{k}_2}, \tag{A2}$$

$$\rho^{(3)}(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3) = \rho_{\mathbf{k}_1}\rho_{\mathbf{k}_2}\rho_{\mathbf{k}_3} - \rho_{\mathbf{k}_1+\mathbf{k}_2}\rho_{\mathbf{k}_3} - \rho_{\mathbf{k}_1+\mathbf{k}_3}\rho_{\mathbf{k}_2} - \rho_{\mathbf{k}_2+\mathbf{k}_3}\rho_{\mathbf{k}_1} + 2\rho_{\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3}. \tag{A3}$$

The general expression has alternating sign according to the order of the terms, and the coefficient is generally different than 1 [note the last term in (A3)].

Then, for $n=2$,

$$\chi_1 = \frac{N^{-1}}{2!} \sum_{\substack{\mathbf{k}_1\mathbf{k}_2 \\ \mathbf{k}_i \neq 0}} C_2(\mathbf{k}_1\mathbf{k}_2)\left[\rho_{\mathbf{k}_1}\rho_{\mathbf{k}_2} - \rho_{\mathbf{k}_1+\mathbf{k}_2}\right]$$

can be written as

$$\chi_1 = y_0 + \sum_{\mathbf{l}\neq 0} y_1(\mathbf{l})\rho_{\mathbf{l}} + \sum_{(\mathbf{k}_1,\mathbf{k}_2)}{}' y_2(\mathbf{k}_1\mathbf{k}_2)\rho_{\mathbf{k}_1}\rho_{\mathbf{k}_2}$$

where

$$y_0 = -\tfrac{1}{2}\sum_{\mathbf{k}\neq 0} C_2(\mathbf{k}, -\mathbf{k}),$$

$$y_1(\mathbf{l}) = \tfrac{1}{2}N^{-1}\sum_{\mathbf{h}\neq 0,\mathbf{l}} C_2(\mathbf{l}-\mathbf{h}, \mathbf{h}),$$

$$y_2(\mathbf{k}_1\mathbf{k}_2) = N^{-1}C_2(\mathbf{k}_1\mathbf{k}_2)/(1+\delta_{\mathbf{k}_1,\mathbf{k}_2}).$$

The y_0 comes from the term $\mathbf{k}_1+\mathbf{k}_2=0$, where $\rho_{\mathbf{k}_1+\mathbf{k}_2}=N$.

In a similar fashion, for $n=3$,

$$\chi_1 = \frac{N^{-2}}{3!} \sum_{\substack{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3 \\ \mathbf{k}_i\neq 0}} C_3(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3)\left[\rho_{\mathbf{k}_1}\rho_{\mathbf{k}_2}\rho_{\mathbf{k}_3} - 3\rho_{\mathbf{k}_1+\mathbf{k}_2}\rho_{\mathbf{k}_3} + 2\rho_{\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3}\right].$$

Then χ_1 becomes

$$\chi_1 = y_0 + \sum_{\mathbf{l}\neq 0} y_1(\mathbf{l})\rho_{\mathbf{l}} + \sum_{\substack{(\mathbf{h}_1\mathbf{h}_2)' \\ \mathbf{h}_i\neq 0}} y_2(\mathbf{h}_1\mathbf{h}_2)\rho_{\mathbf{h}_1}\rho_{\mathbf{h}_2} + \sum_{\substack{(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3)' \\ \mathbf{k}_i\neq 0}} y_3(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3)\rho_{\mathbf{k}_1}\rho_{\mathbf{k}_2}\rho_{\mathbf{k}_3} \tag{A4}$$

where

$$\begin{aligned} y_3(\mathbf{k}_1\mathbf{k}_1\mathbf{k}_1) &= N^{-2}C_3(\mathbf{k}_1\mathbf{k}_1\mathbf{k}_1)/3!, \\ y_3(\mathbf{k}_1\mathbf{k}_1\mathbf{k}_2) &= N^{-2}C_3(\mathbf{k}_1\mathbf{k}_1\mathbf{k}_2)/2, \quad \mathbf{k}_1 \neq \mathbf{k}_2, \\ y_3(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3) &= N^{-2}C_3(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3), \quad \mathbf{k}_i \neq \mathbf{k}_j \end{aligned} \tag{A5}$$

$$y_2(\mathbf{h}_1\mathbf{h}_2) = -N^{-2}\left(\sum_{\mathbf{k}\neq 0,\mathbf{h}_1} C_3(\mathbf{h}_1 - \mathbf{k}, \mathbf{k}, \mathbf{h}_2) + \sum_{\mathbf{k}\neq 0,\mathbf{h}_2} C_3(\mathbf{h}_2 - \mathbf{k}, \mathbf{k}, \mathbf{h}_1)\right)\Big/(1+\delta_{\mathbf{h}_1,\mathbf{h}_2}), \tag{A6}$$

$$y_1(\mathbf{l}) = N^{-2}\sum_{\substack{\mathbf{k}_1\mathbf{k}_2 \\ (\mathbf{k}_i\neq 0,\ \mathbf{k}_1+\mathbf{k}_2\neq \mathbf{l})}} C_3(\mathbf{k}_1, \mathbf{k}_2, \mathbf{l}-\mathbf{k}_1-\mathbf{k}_2)/3 - N^{-1}\sum_{\mathbf{k}\neq 0} C_3(-\mathbf{k}, \mathbf{l}, \mathbf{k})/2, \tag{A7}$$

$$y_0 = N^{-2}\sum_{\substack{\mathbf{k}_1\mathbf{k}_2 \\ (\mathbf{k}_i\neq 0, -\mathbf{k}_j)}} C_3(\mathbf{k}_1, \mathbf{k}_2, -\mathbf{k}_1-\mathbf{k}_2). \tag{A8}$$

Although these results are adequate for most purposes, a general prescription may be of some use. The general expression for $\rho^{(n)}$ as an nth order polynomial in $\rho_{\mathbf{k}}$ is obtained by first noting that $\rho^{(n)}$ satisfies a recursion relation which may be deduced by multiplying $\rho_{\mathbf{k}_{n+1}}$ [Eq. (2)] into Eq. (13) for $\rho^{(n)}$. Then

$$\rho_{\mathbf{k}_{n+1}}\rho^{(n)}(\mathbf{k}_1 \cdots \mathbf{k}_n) = \rho^{(n+1)}(\mathbf{k}_1 \cdots \mathbf{k}_{n+1}) + \sum_{i=1}^{n} \rho^{(n)}(\mathbf{k}_1 \cdots, \mathbf{k}_i + \mathbf{k}_{n+1}, \cdots, \mathbf{k}_n). \tag{A9}$$

The second term in (A9) comes from honoring the

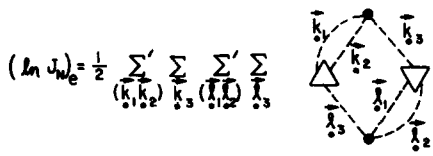


FIG. 5. Fig. 2(e) properly labelled and summed.

restriction $i_s \neq i_{n+1}$ in (13). Then (A9) may be solved for $\rho^{(n+1)}$, giving an expression in terms of $\rho_{\mathbf{k}}$ and $\rho^{(p)}$, $p < n+1$. Changing $n+1$ to n, this is

$$\rho^{(n)}(\mathbf{k}_1 \cdots \mathbf{k}_n) = \rho_{\mathbf{k}_n} \rho^{(n-1)}(\mathbf{k}_1 \cdots \mathbf{k}_{n-1}) - \sum_{i=1}^{n-1} \rho^{(n-1)}(\mathbf{k}_1 \cdots, \mathbf{k}_i + \mathbf{k}_n, \cdots, \mathbf{k}_{n-1}). \quad \text{(A10)}$$

The expression for $\rho^{(n)}$ which this generates has a general term of the form

$$(-1)^{n-q}\gamma \left[\prod_{i=1}^{q} R_{p_i}(-)\right] \{\mathbf{k}_1 \cdots \mathbf{k}_n\}$$

where the notation is that of Eq. (25), the function R_p is defined by

$$R_p(\mathbf{k}_1 \cdots \mathbf{k}_p) = \rho_{\mathbf{k}_1+\mathbf{k}_2+\cdots+\mathbf{k}_p}$$

and γ is a combinatoric factor which may be deduced in the following manner. Consider a term in $\rho^{(n)}$ which includes as a factor $R_p(\mathbf{k}_n, \mathbf{k}_1, \ldots, \mathbf{k}_{p-1})$. Since $\mathbf{k}_n$ appears as an argument, this term can only be obtained from the second term (the sum) in (A10), and there are $p-1$ such members of this second term. Then consider a particular one of these, say $\rho^{(n-1)}(\mathbf{k}_n + \mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_{n-1})$. Then the only terms which give an R_p are those obtained from the second term of (A10) with $n \to n-1$ and $\mathbf{k}_n \to \mathbf{k}_n + \mathbf{k}_1$, and there will be $(p-2)$ of these terms. Continuing this procedure we find that the factor $R_p(\mathbf{k}_n, \mathbf{k}_1, \ldots, \mathbf{k}_{p-1})$ must have an associated factor $(p-1)!$ Because of the symmetry of the problem this is independent of the choice of arguments of R_p. Then we may define a new function

$$\tilde{R}_p(\mathbf{k}_1 \cdots \mathbf{k}_p) = -(p-1)!\ \rho_{\mathbf{k}_1+\cdots+\mathbf{k}_p}$$

in which case the expansion for $\rho^{(n)}$ in terms of $\rho_{\mathbf{k}}$ can be written

$$\rho^{(n)}(\mathbf{k}_1 \cdots \mathbf{k}_n) = (-1)^n \sum_{\substack{1 \le p_1 \cdots \le p_q \\ \sum p_i = n}} \left[\prod_i \tilde{R}_{p_i}(-)\right] \{\mathbf{k}_1 \cdots \mathbf{k}_n\}. \quad \text{(A11)}$$

Now this expression must be put back into Eq. (12) and the overall sum rearranged according to the number of $\rho_{\mathbf{k}}$ factors. Each R factor contributes a $\rho_{\mathbf{k}}$ factor unless its total momentum is zero, in which case it contributes N. The remainder of the development of the expression for y_p involves relabelling the momentum sums, but is too tedious to include here.

APPENDIX B

As an example of the calculation of the contribution of a diagram to $\ln J_N$, consider diagram e of Fig. 2. From Eq. (39), its contribution is

$$(\ln J_N)_e = \frac{1}{2!} \sum_{(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3)}{}' \sum_{(\mathbf{l}_1\mathbf{l}_2\mathbf{l}_3)}{}' y_3(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3)\, y_3(\mathbf{l}_1\mathbf{l}_2\mathbf{l}_3) \times [U_3^{(0)} U_3^{(0)}]_c \{(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3)(\mathbf{l}_1\mathbf{l}_2\mathbf{l}_3)\}. \quad \text{(B1)}$$

The unlabelled diagram refers to

$$[U_3^{(0)} U_3^{(0)}]_c \{(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3)(\mathbf{l}_1\mathbf{l}_2\mathbf{l}_3)\}, \quad \text{(B2)}$$

which in the most general case is nine terms, coming from the three different ways of assigning the momenta $(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ to dashed lines connected to one vertex and the three different ways of assigning $(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3)$ to the other vertex. There are only three ways in each case since the two lines from a vertex to the same triangle are equivalent. Similarly, interchanging the two vertices produces an equivalent term. These nine diagrams can be collapsed to one by letting the odd momentum range independently over all momenta. Thus

$$(\ln J_N)_e = \tfrac{1}{2} \sum_{(\mathbf{k}_1\mathbf{k}_2)}{}' \sum_{\mathbf{k}_3} \sum_{(\mathbf{l}_1\mathbf{l}_2)}{}' \sum_{\mathbf{l}_3} y_3(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3)\, U_3^{(0)}(\mathbf{k}_1\mathbf{k}_2\mathbf{l}_3)\, y_3(\mathbf{l}_1\mathbf{l}_2\mathbf{l}_3) \times U_3^{(0)}(\mathbf{l}_1\mathbf{l}_2\mathbf{k}_3) \quad \text{(B3)}$$

which is shown as a labelled diagram in Fig. 5. Note that sums over equivalent bonds from a particular vertex must still be restricted to distinct [whence the primes in (B3)].

Momentum conservation in $U_n^{(0)}$ and y_n simplifies this sum, giving

$$(\ln J_N)_e = \tfrac{1}{2} \sum_{(\mathbf{k}_1\mathbf{k}_2)}{}' \sum_{(\mathbf{l}_1\mathbf{l}_2)}{}' y_3(\mathbf{k}_1, \mathbf{k}_2, -\mathbf{k}_1 - \mathbf{k}_2) \times U_3^{(0)}(\mathbf{k}_1, \mathbf{k}_2, -\mathbf{k}_1 - \mathbf{k}_2) \otimes y_3(\mathbf{l}_1, \mathbf{l}_2, -\mathbf{l}_1 - \mathbf{l}_2)\, U_3^{(0)}(\mathbf{l}_1, \mathbf{l}_2, -\mathbf{l}_1 - \mathbf{l}_2). \quad \text{(B4e)}$$

The remaining diagrams in the second order expansion for $\ln J_N$ shown in Fig. 2 are calculated similarly. Then

$$(\ln J_N)_2 = \sum_{\alpha=a\cdots k} (\ln J_N)_\alpha$$

where, in the translationally invariant case,

$$(\ln J_N)_a = y_0, \quad \text{(B4a)}$$

$$(\ln J_N)_b = \sum_{\mathbf{k}, k_x>0} y_2(\mathbf{k}, -\mathbf{k}) U_2^{(0)}(\mathbf{k}, -\mathbf{k}), \quad \text{(B4b)}$$

$$(\ln J_N)_c = \sum_{(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3)}{}' y_3(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3)\, U_3^{(0)}(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3) \quad \text{(B4c)}$$

$$(\ln J_N)_d = \tfrac{1}{2} \sum_{(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3)}{}' y_3(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3)\, U_2^{(0)}(\mathbf{k}_1 - \mathbf{k}_1)\, U_2^{(0)}(\mathbf{k}_2 - \mathbf{k}_2) \times U_2^{(0)}(\mathbf{k}_3 - \mathbf{k}_3)\, y_3(-\mathbf{k}_1 - \mathbf{k}_2 - \mathbf{k}_3), \quad \text{(B4d)}$$

$$(\ln J_N)_f = \tfrac{1}{2} \sum_{(\mathbf{k}_1\mathbf{k}_2)}{}' \sum_{(\mathbf{l}_1\mathbf{l}_2)}{}' y_3(\mathbf{k}_1, \mathbf{k}_2, -\mathbf{k}_1 - \mathbf{k}_2)\, U_4^{(0)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{l}_1, \mathbf{l}_2) \times y_3(\mathbf{l}_1, \mathbf{l}_2, -\mathbf{l}_1 - \mathbf{l}_2) \otimes U_2^{(0)}(-\mathbf{k}_1, -\mathbf{k}_2, -\mathbf{l}_1 - \mathbf{l}_2)\, \delta_{\mathbf{k}_1+\mathbf{k}_2, -\mathbf{l}_1-\mathbf{l}_2}, \quad \text{(B4f)}$$

$$(\ln J_N)_g = \tfrac{1}{2} \sum_{\mathbf{k}, k_x>0} U_2^{(0)}(\mathbf{k}, -\mathbf{k})^2\, y_2(\mathbf{k}, -\mathbf{k})^2, \quad \text{(B4g)}$$

$$(\ln J_N)_h = \tfrac{1}{2} \sum_{\mathbf{k}, k_x>0} \sum_{\mathbf{l}, l_x>0} U_4^{(0)}(\mathbf{k}, -\mathbf{k}, \mathbf{l}, -\mathbf{l})\, y_2(\mathbf{k}, -\mathbf{k})\, y_2(\mathbf{l}, -\mathbf{l}), \quad \text{(B4h)}$$

$$(\ln J_N)_i = \sum_{(\mathbf{k}_1\mathbf{k}_2)}{}' U_3^{(0)}(\mathbf{k}_1, \mathbf{k}_2, -\mathbf{k}_1 - \mathbf{k}_2)\, U_2^{(0)}(-\mathbf{k}_1 - \mathbf{k}_2, \mathbf{k}_1 + \mathbf{k}_2)$$

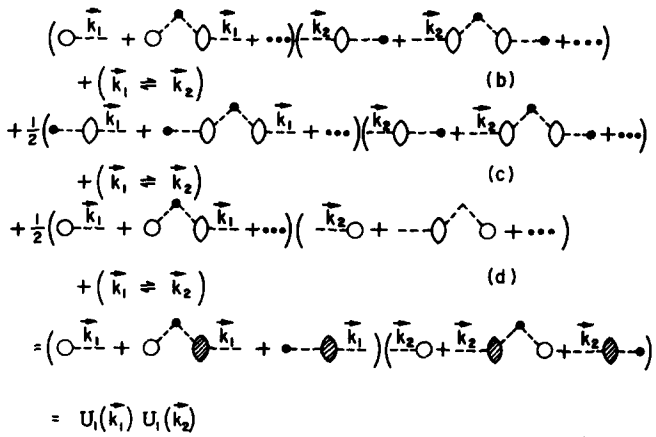


FIG. 6. The disconnected terms in $I_2(\mathbf{k}_1, \mathbf{k}_2)$ as they arise from the derivative of $\ln J_N$ with respect to $y_2(\mathbf{k}_1, \mathbf{k}_2)$. Contributions (b), (c), and (d) come from Fig. 3(b), 3(c), and 3(d), respectively.

$$\otimes y_2(\mathbf{k}_1+\mathbf{k}_2, -\mathbf{k}_1-\mathbf{k}_2)\, y_3(-\mathbf{k}_1-\mathbf{k}_2, \mathbf{k}_1, \mathbf{k}_2), \qquad \text{(B4i)}$$

$$(\ln J_N)_j = \sum_{(\mathbf{k}_1\mathbf{k}_2)}{}' \sum_{(\mathbf{l}_1\mathbf{l}_2\mathbf{l}_3)}{}' U_5^{(0)}(\mathbf{k}_1\mathbf{k}_2\mathbf{l}_1\mathbf{l}_2\mathbf{l}_3)\, y_2(\mathbf{k}_1\mathbf{k}_2)\, y_3(\mathbf{l}_1\mathbf{l}_2\mathbf{l}_3), \qquad \text{(B4j)}$$

$$(\ln J_N)_k = \tfrac{1}{2} \sum_{(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3)}{}' \sum_{(\mathbf{l}_1\mathbf{l}_2\mathbf{l}_3)}{}' U_6^{(0)}(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{l}_1\mathbf{l}_2\mathbf{l}_3)\, y_3(\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3)\, y_3(\mathbf{l}_1\mathbf{l}_2\mathbf{l}_3). \qquad \text{(B4k)}$$

The three-body modification u_3 to the Jastrow wavefunction has been calculated for liquid He^4 using terms (a), (b), (c), and (d) above.[9] It can be argued that terms $f-j$ do not contribute by considering instead the trial wavefunction for which $y_3 \neq 0$ but $y_2 = 0$ in the modification χ_1. Since $f-j$ are proportional to y_2, they will not contribute. We have omitted term (e), however, which may be an important term, particularly at high density. We hope to examine the contribution of e to the ground state properties of He^4 in the future. Term k requires $U_6^{(0)}$, which is difficult to approximate.

APPENDIX C

We wish to show that (50) is obtained from differentiation of (45) by y_2 when (46) and (47) are used for U_1 and U_2 in (50). To see this we note that differentiation with respect to $y_2(\mathbf{k}_1, \mathbf{k}_2)$ is the same as removing one two-prong vertex from each diagram in all possible ways, then assigning the momenta $\mathbf{k}_1$ and $\mathbf{k}_2$ to the remaining two external dashed lines in all possible ways. When this is done to the pth ring diagram from Fig. 3(a) it produces $2p$ equivalent diagrams of the form of the pth diagram for U_2, Fig. 4(b). Thus

$$\frac{\partial}{\partial y_2(\mathbf{k}_1\mathbf{k}_2)} (\ln J_N)_a = U_2(\mathbf{k}_1, \mathbf{k}_2). \qquad \text{(C1)}$$

Removing a two-prong vertex from any of the diagrams (b), (c), or (d) in Fig. 3 gives the product of two connected diagrams, one containing $\mathbf{k}_1$ and the other containing $\mathbf{k}_2$. This is shown in Fig. 6 with properly labelled diagrams. These products of pairs of connected graphs add to give $U_1(\mathbf{k}_1)U_1(\mathbf{k}_2)$:

$$\frac{\partial}{\partial y_2(\mathbf{k}_1\mathbf{k}_2)} [(\ln J_N)_b + (\ln J_N)_c + (\ln J_N)_d] = U_1(\mathbf{k}_1)U_1(\mathbf{k}_2). \qquad \text{(C2)}$$

Thus (50) is obtained by adding (C1) and (C2), which is the desired result.

*Work supported in part by the National Science Foundation (Grant No. NSF-GH-43836) and by Research Corporation.

[1] R. Jastrow, Phys. Rev. **98**, 1479 (1955).
[2] R. Abe, Prog. Theor. Phys. **19**, 57, 407 (1958).
[3] F. Y. Wu and E. Feenberg, Phys. Rev. **122**, 739 (1961).
[4] W. E. Massey, Phys. Rev. **151**, A153 (1966).
[5] W. L. McMillan, Phys. Rev. **164**, 256 (1967); D. Schiff and L. Verlet, Phys. Rev. **160**, 208 (1967).
[6] C. E. Campbell and M. Schick, Phys. Rev. A **3**, 691 (1971); M. Miller, C.-W. Woo, and C. E. Campbell, Phys. Rev. A **6**, 1942 (1972).
[7] C. E. Campbell and E. Feenberg, Phys. Rev. **188**, 396 (1969).
[8] E. Feenberg, *Theory of Quantum Fluids* (Academic, New York, 1969).
[9] C.-W. Woo, Phys. Rev. Lett. **28**, 1442 (1972); Phys. Rev. A**6**, 2312 (1972).
[10] C. E. Campbell, Phys. Lett. **44A**, 471 (1973); the details of this paper are given in Sec. 12 of E. Feenberg, Proceedings of 1974 Latin American School of Physics.
[11] E. Feenberg, Ann. Phys. (N.Y.) **84**, 128 (1974). An earlier analysis in which similar results were obtained is given in E. P. Gross, Ann. Phys. (N.Y.) **20**, 44 (1962).
[12] P. Kleban and J. W. Halley, Phys. Rev. B, 1 May 1975; A. D. B. Woods and R. A. Cowley, Rep. Prog. Phys. **36**, 1135 (1973).
[13] F. Y. Wu, J. Math. Phys. **12**, 1923 (1971).
[14] A. D. Novaco and C. E. Campbell, Phys. Rev. B, 1 April 1975; C. E. Campbell (to be published).
[15] J. K. Percus and G. J. Yevick, Nuovo Cimento **5**, 1057 (1957).
[16] R. Brout, *Phase Transitions* (Benjamin, New York, 1965), and references cited therein. For more recent work using the hard core model as the starting point, see H. C. Anderson, J. D. Weeks, and D. Chandler, Phys. Rev. A **4**, 1597 (1971).
[17] G. V. Chester and L. Reatto, Phys. Lett. **22**, 276 (1966); Chap. 3 of Ref. 8.
[18] H. W. Jackson and E. Feenberg, Ann. Phys. (N.Y.) **15**, 266 (1961).
[19] Strictly speaking, it is not necessary for ψ_0 to be a Jastrow function for the *paired phonon analysis* to be used. What is actually being found is the Jastrow function ψ, which minimizes the energy when multiplied by ψ_0.
[20] T. Paulick, private communication.
[21] H. W. Jackson and E. Feenberg, Phys. Rev. **128**, 943 (1962); D. K. Lee, Phys. Rev. **162**, 134 (1967); F. Y. Wu and M. K. Chien, J. Math. Phys. **11**, 1912 (1970).
[22] E. Feenberg, preprint; K. Y. Chung, Washington University Dissertation, 1973 (unpublished); D. K. Lee, Phys. Rev. **187**, 326 (1969); P. Berdahl (to be published).

A note on coherent state representations of Lie groups

Enrico Onofri

Istituto di Fisica dell'Università, Parma, Italy

Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Milano, Italy
(Received 25 January 1974)

The analyticity properties of coherent states for a semisimple Lie group are discussed. It is shown that they lead naturally to a classical "phase space realization" of the group.

In a recent article Peremolov[1] introduced the concept of "coherent states" for unitary irreducible representations (UIR) of any Lie group G. The idea to consider the translates $T_g|0\rangle$ of a fixed cyclic vector $|0\rangle$ under the group action is as old as the celebrated Gel'fand–Raikov theorem on locally bicompact groups.[2] The contribution of Ref. 1 was then to show how the concept of coherent state, first introduced in the case of Heisenberg–Weyl group, fitted in the general theory of group representations. Let us note that no mention to analyticity question is made in the very general approach given in Ref. 1. Yet all the examples [Heisenberg–Weyl group, $SU(2)$, $SU(1,1)$] have in common the following properties: (a) the homogeneous space G/H has a complex homogeneous structure, i.e., G acts on G/H by means of holomorphic transformations [in the above examples G/H is given by the complex plane, by $S^2 \approx$ complex projective space $P_1(\mathbb{C})$ and by the unit disc $D_1 \subset \mathbb{C}$, respectively]; (b) the Hilbert space of the UIR is identified, in the coherent state basis, with a space of holomorphic functions on G/H [namely, $\exp(\frac{1}{2}|z|^2)$-bounded entire functions, polynomials of degree $2l$ in z and $(1-|z|^2)^{-l}$-bounded analytic functions in D_1, respectively].

Purpose of the present note is to show that

(1) *a homogeneous complex structure is actually present quite generally*, and

(2) that on the basis of the homogeneous complex structure *the manifolds G/H are just the classical phase spaces on which G acts through canonical transformations*.

From this point of view, coherent states appear just as probability wave packets over the classical phase space, a well-known result for the harmonic oscillator coherent states. The converse problem, i.e., the construction of a UIR of G starting from a phase space realization of G. (*quantization of a dynamical group*[3]) was considered in Ref. 4 and found a definite mathematical setting in Refs. 5, 6; here, however, we shall adhere to Ref. 1 scheme and deduce the phase space structure from the UIR.

First of all we restrict our attention to *compact semisimple Lie groups*. Let us fix the notation: $\mathbf{g}$ is the real Lie algebra of the group G, $\mathbf{g}_c = \mathbf{g} \oplus i\mathbf{g}$ its complexification, H a Cartan subgroup of G, $\mathbf{g}_c = \mathbf{h}_c \oplus \sum_{\alpha\in\Delta} \mathbf{g}_\alpha$ the corresponding Cartan decomposition, Δ_+ the set of positive roots (chosen once for all), $\rho = \frac{1}{2}\sum_{\alpha\in\Delta_+}\alpha$. Let $g \in G \to U(g)$ be a UIR of G in a (finite-dimensional) Hilbert space $\mathcal{H}$. Let λ be the highest weight and let us suppose that it is nonsingular. Then there exists a vector $|0\rangle \in \mathcal{H}$ such that

$$U(h)|0\rangle = e^{\lambda}(h)|0\rangle, \quad h \in H$$
$$\{e^{\lambda}\exp(X) = \exp[\lambda(X)], \quad X \in \mathbf{g}, \quad \lambda \in i\mathbf{g}^*\}, \tag{1}$$
$$X|0\rangle = 0, \quad X \in \mathbf{n}_+ = \sum_{\alpha\in\Delta_+} \mathbf{g}_\alpha,$$

or, alternatively,

$$\langle 0|U(h) = e^{\lambda}(h)\langle 0|, \quad h \in H,$$
$$\langle 0|X = 0, \quad X \in \mathbf{n}_- = \sum_{\alpha\in\Delta_+} \mathbf{g}_{-\alpha}. \tag{2}$$

Given $U(g)$ there is no problem in extending it to a holomorphic representation $T(g)$ of G_c in $\mathcal{H}$ (it is sufficient to exponentiate the finite-dimensional matrices which represent the elements of $\mathbf{g}_c$; for noncompact groups the situation is different). Equation (2) shows that the stability subgroup of $\langle 0|$ under $T(g)$ is just the *Borel subgroup* B, which is the complex Lie subgroup of G_c with Lie algebra $\mathbf{b} = \mathbf{h}_c \oplus \mathbf{n}_-$. Let us call $\pi(b)$ the *holomorphic character* of B defined by

$$\langle 0|T(b) = \pi(b)\langle 0|. \tag{3}$$

It is wellknown that the quotient manifolds G/H and G_c/B coincide[7]; then G/H inherits a complex homogeneous structure from G_c, being B a complex subgroup. Let us now consider the family of coherent states $T(g^{-1})^\dagger|0\rangle$; the representatives

$$\Psi(g) = \langle 0|T(g^{-1})|\Psi\rangle \tag{4}$$

are holomorphic on G_c; moreover,

$$\Psi(gb) = \langle 0|T(b^{-1})\,T(g^{-1})|\Psi\rangle$$
$$= \pi(b^{-1})\Psi(g), \quad g \in G_c, \ b \in B. \tag{5}$$

This means that $\Psi(g)$ defines a *holomorphic section of the homogeneous line bundle $E_\pi(G/H, \mathbb{C})$ associated to the principal fibre bundle $B \to G_c \to G_c/B$ by the holomorphic character π* (see Ref. 8 for the definition of "associated" fibre bundle). The local trivializations of the line bundle E_π associates to every such Ψ a holomorphic function on an open domain in G/H. For instance, let $\mathcal{U} = \{g \in G_c | \langle 0|T(g^{-1})|0\rangle \neq 0\}$; then the function

$$\Psi(g) = \frac{\langle 0|T(g^{-1})|\Psi\rangle}{\langle 0|T(g^{-1})|0\rangle} \tag{6}$$

is actually a function only of the projection $g \xrightarrow{p} g\{H\} \in G/H$ and is holomorphic in the domain $p(\mathcal{U}) \subset G/H$. Let $z = (z_1, \ldots, z_n)$ be a local chart $p(\mathcal{U}) \to \mathbb{C}^n$ such that $0 = (0, \ldots, 0)$ correspond to $\{H\}$. We can, therefore, introduce a new set of states in a one-to-one correspondence with the points in $p(\mathcal{U})$:

Copyright © 1975 American Institute of Physics

$$|g\cdot 0\rangle=\frac{T(g^{-1})^\dagger|0\rangle}{\langle 0|T(g^{-1})^\dagger|0\rangle},\quad g\in G_c, \tag{7}$$

or, restricting to G,

$$|g\cdot 0\rangle=\frac{U(g)|0\rangle}{\langle 0|U(g)|0\rangle},\quad g\in G. \tag{8}$$

Owing to the orthogonality relations

$$\int_G \overline{\langle\Psi_1|U(g)|\Psi_2\rangle}\langle\Psi_3|U(g)|\Psi_4\rangle dg=d_\lambda^{-1}\langle\Psi_2|\Psi_4\rangle\langle\Psi_3|\Psi_1\rangle$$

(being $d_\lambda=\Pi_{\alpha\in\Delta_+}(\langle\lambda+\rho,\alpha\rangle/\langle\rho,\alpha\rangle)=$ dimension of U), the scalar product $\langle\Psi_1|\Psi_2\rangle$ can be written as follows (completeness relation for the coherent states $|z\rangle=|g\cdot 0\rangle$):

$$\begin{aligned}\langle\Psi_1|\Psi_2\rangle&=d_\lambda\int_G\overline{\langle\Psi_2|U(g)|0\rangle}\langle\Psi_1|U(g)|0\rangle dg\\&=d_\lambda\int_G\overline{\langle 0|U(g)^\dagger|\Psi_1\rangle}\langle 0|U(g)^\dagger|\Psi_2\rangle dg\\&=d_\lambda\int_G\overline{\langle g\cdot 0|\Psi_1\rangle}\langle g\cdot 0|\Psi_2\rangle|\langle 0|U(g)|0\rangle|^2 dg\\&=d_\lambda\int_{G/H}\overline{\psi_1(z)}\psi_2(z)\exp[-f(z,\bar z)]\dot z,\end{aligned} \tag{9}$$

with

(i) $z=g\cdot 0$,

(ii) $f(z,\bar z)=\log|\langle 0|U(g)|0\rangle|^{-2}\geq 0\quad(f=0\Rightarrow z=0)$,

(iii) $\dot z=$ invariant volume density on G/H induced by dg.

Thus H is isomorphic to a Hilbert space of holomorphic functions bounded by $\exp[\frac{1}{2}f(z,\bar z)]$. In this derivation we tacitly assumed that $\{g\in G|\langle 0|U(g)|0\rangle=0\}$ is of measure zero in G so that integrals over G/H can be restricted to integrals over $p(\mathcal{U})$.

The representation of G in the basis of the coherent states takes the form of a "multiplier" representation:

$$\begin{aligned}[U(g')\psi](z)&=\frac{\langle 0|U(g)^\dagger U(g')|\Psi\rangle}{\langle 0|U(g)^\dagger|0\rangle}\\&=\frac{\langle 0|U(g'^{-1}g)^\dagger|0\rangle}{\langle 0|U(g)^\dagger|0\rangle}\psi(g'^{-1}z)\\&=\mu(g',z)\psi(g'^{-1}z).\end{aligned} \tag{10}$$

Let us now construct the symplectic (Kaehler) structure on G/H. Let

$$\omega=i\partial\bar\partial f(z,\bar z),\quad z\in p(\mathcal{U}), \tag{11}$$

with ∂ and $\bar\partial$ the exterior differentiation operators with respect to z and $\bar z$, respectively $(\partial\bar\partial+\bar\partial\partial=\partial^2=\bar\partial^2=0)$;

explicitly, $\omega=i\sum_{j,k}\frac{\partial^2 f}{\partial z^j\partial\bar z^k}dz^j\wedge d\bar z^k$.

It can be shown that $\omega(X,Y)$ coincides at $z=0$ with the bilinear functional $-i\lambda([X,Y])$, with $X,Y\in\mathbf{g}$ and λ identified with its image in $i\mathbf{g}^*$. Then ω is nonsingular at $z=0$. Moreover, since ω is G-invariant, it is nonsingular everywhere. Let us prove the G-invariance of ω:

$$\begin{aligned}g'_*\omega&=i\partial\bar\partial f(g'^{-1}z,\overline{g'^{-1}z})\\&=i\partial\bar\partial\log|\langle 0|U(g'^{-1}g)|0\rangle|^{-2}\\&=i\partial\bar\partial f(z,\bar z)-i\partial\bar\partial\log|\mu(g',z)|^2\end{aligned} \tag{12}$$

and the invariance follows, being $\mu(g',z)$ holomorphic in z.

Up to now we have made only local considerations. Actually, it can be easily shown that a Kaehler form ω_α can be given in a whole covering $\mathcal{O}_\alpha$ of G/H and $\omega_\alpha=\omega_\beta$ on the overlapping $\mathcal{O}_\alpha\cap\mathcal{O}_\beta$. Then we have a homogeneous Kaehler manifold G/H associated to the UIR, i.e., a *G-homogeneous Hamiltonian dynamical system*. Given $X\in\mathbf{g}$, let $\hat X$ be its self-adjoint representation in H, $\mathcal{X}$ the holomorphic vectorfield on G/H, $H(z,\bar z)$ the "classical generating function," defined by $i\chi_{X*}\bar\chi\omega=dH(z,\bar z)$. For every ψ in the domain of $\hat X$ it holds that

$$[\hat X\psi](z)=[H(z,\bar z)+i\chi(f)]\psi(z)-i(\mathcal{X}\psi)(z). \tag{13}$$

From this it follows that

$$H(z,\bar z)=\langle z|\hat X|z\rangle/\langle z|z\rangle; \tag{14}$$

i.e., *the value of the classical generating function at a point z coincides with the expectation value of the corresponding self-adjoint operator on the coherent state $|z\rangle$*. In general, for a normalized $|\Psi\rangle$ it holds that

$$\begin{aligned}\langle\Psi|\hat X|\Psi\rangle&=\gamma\int_{G/H}\exp[-f(z,\bar z)]|\psi(z)|^2H(z,\bar z)\dot z\\&=\int_{G/H}\rho(z,\bar z)H(z,\bar z)\dot z,\end{aligned} \tag{15}$$

with γ a factor depending on the UIR (but not on X!); i.e., *the expectation values in the* UIR *coincide with a phase space average of the classical generating function with a probability distribution*

$$\begin{aligned}\rho(z,\bar z)&=\gamma|\psi(z)|^2\exp[-f(z,\bar z)]=\\&=\gamma\frac{\langle z|\hat\rho|z\rangle}{\langle z|z\rangle},\quad\hat\rho\equiv|\Psi\rangle\langle\Psi|.\end{aligned} \tag{16}$$

$\rho(z,\bar z)$ is then proportional to the diagonal matrix element of the density matrix in the coherent states representation. Equations (15), (16) hold for a generic density matrix $\hat\rho$. Let us note that $\rho(z,\bar z)$ is positive definite, unlike other correspondential statistical distributions, such as Wigner's function.

A simple relation exists between $\mu(g,z)$ and the classical generating function of the (finite) canonical transformations in G. Let us define $\vartheta=-i\partial f$, then

$$\omega=\mathrm{d}\vartheta\quad(\mathrm{d}=\partial+\bar\partial\Rightarrow\mathrm{d}\partial=\bar\partial\partial=-\partial\bar\partial)$$

and

$$S(g,z)=\int_0^z(\vartheta-g_*\vartheta)+S(g,0)$$

is the (holomorphic) generating function, analogous to the usual $S=\int(pdq-p'dq')$. Then we have

$$\mu(g,z)=\exp[iS(g,z)] \tag{17}$$

and we recover an expression first introduced by Van Hove[9] for a Euclidean phase space:

$$[U(g)\psi](z)=\exp[iS(g,z)]\psi(g^{-1}z). \tag{18}$$

Let us now consider the case of a *singular* λ. In this case the stability subgroup of the highest weight vector $\langle 0|$ is bigger than H; by complexifying the parameters, one finds a complex subgroup $P\subset G_c$ (called a *parabolic* subgroup) which includes B. Yet the same statements hold as in the nonsingular case; we simply get Kaehler manifolds with dimension smaller than $\dim(G/H)$, i.e.,

the so-called singular realizations.[10] It is actually known that every G-homogeneous symplectic manifold is Kaehler,[11] a result which does not generalize to noncompact groups.

We give now some hints about the question of the extension of the above results to (a) noncompact semisimple Lie groups and (b) to a wider class of Lie groups. In Ref. 6 we have reviewed the relevant results for the noncompact semisimple case. Obviously we must restrict ourselves to discrete series, in order to maintain orthogonality relations; yet the simple results of the compact case hold only in the "Hermitian symmetric" case, while in general we must consider Hilbert spaces of *harmonic differential forms* rather than holomorphic functions. Details can be found in Ref. 6.

As for point (b), a similar treatment as presented here can be given for nilpotent and, more generally, *solvable* Lie groups, on the basis of the UIR theory developed by Kostant, Auslander, Moore and others.[12] For Lie groups having an intermediate structure (between solvable and semisimple) results are yet incomplete. Let us conclude by remarking that the theory as it stands can already deal with groups such as $SU(n,m)$, $SO(2n,1)$, and $SO(n,2)$, which have attracted physicists' attention as good candidates for a fundamental dynamical group.

ACKNOWLEDGMENTS

The ideas presented in this paper arose during a collaboration with Professor Massimo Pauri. I am indebted to Professor Fiorenzo Duimio for reading the manuscript.

[1] A. M. Peremolov, Commun. Math. Phys. **26**, 222 (1972); see also in this connection the article by A. O. Barut and L. Girardello cited by Peremolov.
[2] I. M. Gel'fand and D. Raikov, Mat. Sb. (N.S.) **13(55)**, 301 (1942) [Am. Math. Soc. Transl. **2(36)**, 1 (1964)]. 1 (1964)].
[3] J. M. Souriau, *Structure des systèmes dynamiques* (Dunod, Paris, 1970).
[4] E. Onofri and M. Pauri, J. Math. Phys. **13**, 533 (1972).
[5] E. Onofri and M. Pauri, Lett. Nuovo Cimento **2**(3), 35 (1972).
[6] E. Onofri, "Quantization theory for homogeneous Kaehler manifolds," Università di Parma, Preprint IFPR-T-038 (1974).
[7] R. Bott, Ann. Math. **66**, 203 (1957); P. Griffiths and W. Schmid, Acta Math. **123**, 253 (1969).
[8] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* (Interscience, New York, 1963), Vol. 1.
[9] L. Van Hove, Acad. Roy. Belg. Bull. Classe Sci. Mem. **37**, 610 (1951).
[10] M. Pauri and G. M. Prosperi, J. Math. Phys. **9**, 1146 (1968).
[11] J. P. Serre, "Représentation linéaires et espaces homogènes Kaehleriens des groupes de Lie semisimples compacts," Sém. Bourbaki, Exp. 100 (1954).
[12] L. Auslander and C. C. Moore, Amer. Math. Soc. Mem. **62**, (1966); R. F. Streater, Commun. Math. Phys. **4**, 217 (1967); L. Auslander and B. Kostant, Bull. Amer. Math. Soc. **73**, 692 (1967); J. Brezin, Amer. Math. Soc. Mem. **79**, (1968); B. Kostant, in *Lecture Notes in Mathematics*, edited by C. C. Taam (Springer, New York, 1970).

The de Donder coordinate condition and minimal class 1 space-time

Jorge Krause

Instituto de Física, Universidad Católica de Chile, Casilla 114-D, Santiago, Chile
(Received 11 December 1974)

By means of immersion techniques a set of "adapted coordinates" are introduced as preferred coordinates for class 1 space-time. It is proved that the necessary and sufficient condition for the adapted coordinates to be harmonic coordinates is that class 1 space-time be a minimal variety. Some interesting features of the embedding approach to curved space-time are also shown in terms of these adapted coordinates.

1. INTRODUCTION

Some reasons on behalf of the embedding approach to general relativity have been presented by the author in a previous paper.[1] In the present note we prove that if class 1 space–time[2] is a minimal variety[3] there exist harmonic coordinates, which represent a kind of preferred coordinates according to a well defined geometric feature. This property will be proved and briefly discussed by means of immersion techniques.

Einstein's theory of gravitation is commited to an enormous gauge freedom, notwithstanding the fact that the principle of general covariance, by itself, is devoid of physical content, for, indeed, every physical theory may be written in a general covariant manner.[4,5] On the other hand, it is clear that the geometry of some generic space–times may admit a class of preferred coordinates and that the group of general covariance becomes, thereby, unnecessarily broad for handling the dynamics of such particular geometries. This point of view has been stressed by Fock[6] in connection with the existence of harmonic coordinates as the preferred ones for some kind of gravitational problems.

The most favored approach to the issue of preferred coordinates in general relativity consists in fixing the gauge to some extent by imposing a set of *ad hoc* coordinate conditions directly upon the metric tensor.[7] The question thus arises, quite naturally, whether there exist some variational principle leading us to the desired coordinate conditions,[8] i.e., allowing us to obtain a set of "optimal" coordinates defined by a stationary integral property. The study of some extremal behavior of the coordinates, however, is barren as far as we lack the geometric meaning of the variational principle involved.[8] In this paper we show that such a principle exists for fitting harmonic coordinates in curved space–time, at least when space–time belongs to the simplest embedding class. As is well known, the concept of minimal variety is arrived at by generalizing the definition of minimal surface. Thus (according to the result to be shown in this note), for curved space–time embeddable in five dimensions, the variational principle leading to the harmonic coordinate condition[9] states that the four-dimensional volume integral must be stationary; i.e.,

$$\delta \int d^4x\sqrt{-g(x)} = 0. \tag{1.1}$$

It is interesting to observe that if we handle this problem directly, the corresponding Euler–Lagrange equation collapses to a useless identity; namely, we get

$$(w - 1/2)\left[(-g)^w\right]_{,\mu} \equiv 0, \tag{1.2}$$

for $w = 1/2$ in this case.[10] Therefore, a different approach must be followed in order to relate (1.1) with the de Donder coordinate condition.

The present note dwells only on space–time geometry. The dynamical content of class 1 space–time will be discussed elsewhere. In Sec. 2 we prepare the five-dimensional immersion scaffolding. Covariant derivatives afforded by the embedding formalism are briefly presented in Sec. 3, while in Sec. 4 we analyze the second fundamental form of the embedded space–time. Finally, in Sec. 5 the connection between minimal class 1 space–time and the existance of harmonic coordinates is proved.

2. THE EMBEDDING SCAFFOLDING AND THE ADAPTED COORDINATES

Let us consider the class 1 relativistic embeddings from a synthetic point of view; namely, we consider the curved space–time manifold as a given four-dimensional (normal hyperbolic) hypersurface $E_{(4)}$, already embedded in a five-dimensional flat space $M_{(5)}$. In $E_{(4)}$ we adopt signature (-2); therefore, according to the local isometric embedding theorem,[11] the pseudo-Euclidean metric tensor belonging to the host space is given by $\eta_{AB} = \mathrm{diag}(+ - - - \pm)$, in terms of a system $\{X^A\}$ of Cartesian rectangular coordinates.[12] We explicitly decompose the fundamental form of $M_{(5)}$ in the following way:

$$ds^2 = \eta_{AB}\, dX^A\, dX^B = \eta_{\mu\nu}\, dX^\mu\, dX^\nu + \eta (dX^4)^2, \tag{2.1}$$

where, clearly, $\eta_{\mu\nu} = \mathrm{diag}(+ - - -)$ denotes the usual four-dimensional Minkowski metric, and $\eta = \eta_{44} = \pm 1$. We now assume that the embedded $E_{(4)}$ is defined by introducing a coordinate relation in $M_{(5)}$ of the form

$$E(X^A) = E(X, X^4) = e, \tag{2.2}$$

say, where e is a constant. Since we want $E_{(4)}$ to be a space–time-like hypersurface, somehow leaning smoothly on the (X^0, X^1, X^2, X^3) hyperplane (i.e., Minkowski space–time $M_{(4)}$), we require

$$\eta\eta^{AB} E_{,A}(X) E_{,B}(X)\big|_{X \in E_{(4)}} > 0. \tag{2.3}$$

Copyright © 1975 American Institute of Physics

Hence the unit 5-vector N_A normal to $E_{(4)}$, at points on $E_{(4)}$, in terms of the $\{X^A\}$ coordinates, obtains

$$N_A(X)=(\eta\eta^{BC}E_{,B}E_{,C})^{-1/2}E_{,A}(X)\big|_{X\in E_{(4)}}. \tag{2.4}$$

Next we introduce new curvilinear coordinates $\{x^A\}$ in the embedding space. These we choose as the following *adapted coordinates*[13]

$$\begin{aligned} x^\mu &= X^\mu, \\ x^4 &= E(X^\nu, X^4), \end{aligned} \tag{2.5}$$

so that in terms of these coordinates $E_{(4)}$ is simply given by the equation $x^4=e$ (hence the name). The Jacobian of this transformation is equal to $\partial E/\partial X^4 = E_{,4}$, which we obviously assume to be different from zero. Then, for the inverse transformation of (2.5) we write, say,

$$\begin{aligned} X^\mu &= x^\mu, \\ X^4 &= F(x^\nu, x^4). \end{aligned} \tag{2.6}$$

If we now define the function $\phi(x)=F(x^\nu, e)$, we observe, from (2.6), that the following parametric equations hold as a definition of $E_{(4)}$:

$$\begin{aligned} X^\mu &= x^\mu, \\ X^4 &= \phi(x). \end{aligned} \tag{2.7}$$

Hence, for a local isometric embedding we have, as usual,[14]

$$g_{\mu\nu}(x)=X^A{}_{,\mu}X^B{}_{,\nu}\eta_{AB}, \tag{2.8}$$

and thus we get

$$g_{\mu\nu}(x)=\eta_{\mu\nu}+\eta\phi_{,\mu}(x)\phi_{,\nu}(x). \tag{2.9}$$

Henceforth we use $\{x^\mu\}$, i.e., the first four adapted coordinates, as a set of internal curvilinear coordinates for $E_{(4)}$, while retaining the old Cartesian coordinates $\{X^A\}$ in $M_{(5)}$.

The decomposition (2.9) of the metric $g_{\mu\nu}$ holds locally, at least, over that coordinate patch on which representation (2.6), with $x^4=e$, is valid. Furthermore, the stated decomposition of the curved metric has general tensorial character.[15] The adapted coordinates $\{x^\mu\}$, however, are preferred coordinates, as are the Galilean coordinates in sperial relativity, for they bring the *flat part* of the curved metric $g_{\mu\nu}$ to the canonical Minkowskian form $\eta_{\mu\nu}$. In paper I we have shown, for the general local embedding scheme, that the introduction of this special kind of preferred coordinates reduces the general covariance of the theory, investing Einstein's theory of gravitation with a new restricted covariance under a group of transformations which represents an enlargement of the Poincaré group.[15] For class 1 space–time this group corresponds to five-dimensional rotations and translations in $M_{(5)}$.

3. COVARIANT DERIVATIVES IN THE EMBEDDING FORMALISM

In this section we present some useful formulas which will be needed in the forthcoming discussion. First we observe that (2.3) means that the determinant $g(x)$ of the space–time metric tensor $g_{\mu\nu}(x)$ has the property

$$-g(x)=1+\eta\eta^{\mu\nu}\phi_{,\mu}(x)\phi_{,\nu}(x)>0, \tag{3.1}$$

everywhere on the embedded patch. The expression for the contravariant metric tensor in $E_{(4)}$, in terms of the adapted coordinates, immediately obtains as

$$g^{\mu\nu}(x)=\eta^{\mu\nu}-\frac{\eta\phi^{,\mu}\phi^{,\nu}}{1+\eta\phi_{,\lambda}\phi^{,\lambda}}, \tag{3.2}$$

where

$$\phi^{,\mu}=\eta^{\mu\nu}\phi_{,\nu}=\frac{\phi;\mu}{1-\eta\phi_{;\lambda}\phi^{;\lambda}}, \tag{3.3}$$

and also

$$\phi^{;\mu}=g^{\mu\nu}\phi_{,\nu}=\frac{\phi^{,\mu}}{1+\eta\phi_{,\lambda}\phi^{,\lambda}}. \tag{3.4}$$

These relations hold indeed, for the "fundamental potential" $\phi(x)$ behaves as a scalar field on both (curved and flat) space–times.[16] Furthermore, it is interesting to observe that the fields $g_{\mu\nu}, \eta_{\mu\nu}, \phi_{,\mu}, \phi^{;\mu}$, have tensorial character in both space–times.[17]

Finally, for the Christoffel symbols, in terms of the adapted set $\{x^\mu\}$, we get the expressions

$$\left\{ {\lambda \atop \mu\nu} \right\}=\frac{\eta\phi_{,\mu\nu}\,\phi^{,\lambda}}{1+\eta\phi_{,\sigma}\phi^{,\sigma}}=\frac{\eta\phi_{;\mu\,;\nu}\phi^{;\lambda}}{1-\eta\phi_{;\sigma}\phi^{;\sigma}} \overset{\text{def}}{=} \Gamma^\lambda_{\mu\nu}. \tag{3.5}$$

These expressions manifestly define a tensor field $[\Gamma^\lambda_{\mu\nu}(x)$, say] belonging both in $E_{(4)}$ and $M_{(4)}$. Indeed, they represent a space–time tensor whose components, once a set of adapted coordinates is introduced, become identical with the components of the affine connection. Of course, this result is a "virtue" of the adapted coordinates only, for, clearly, we have to transform these quantities differently (as a tensor and as the affine connection) while going to a general set of space–time coordinates. In effect, (3.5) shows that $\Gamma^\lambda_{\mu\nu}$ is that *part* of $\left\{ {\lambda \atop \mu\nu} \right\}$ which always transforms as a tensor while using general coordinates; i.e., this decomposition of the affine connection preserves its geometric character under general transformation of coordinates, since the transformation law obeyed by the Christoffel symbols will not mix up the two parts of the affine connection. This result strongly suggests, as does the concommitant decomposition of the metric, the very special character of the adapted coordinates.

4. THE SECOND FUNDAMENTAL FORM OF CLASS 1 SPACE-TIME

The components of the unit normal to $E_{(4)}$, at points on $E_{(4)}$, in terms of the $\{X^A\}$ coordinates, can be written explicitly as functions of the internal (adapted) coordinates $\{x^\mu\}$. One finds that

$$N_A(X)=\begin{cases} N_\mu=\dfrac{\sigma\phi_{,\mu}}{(1+\eta\phi_{,\nu}\phi^{,\nu})^{1/2}} \\ N_4=\dfrac{\sigma}{(1+\eta\phi_{,\nu}\phi^{,\nu})^{1/2}} \end{cases}, \tag{4.1}$$

where we define

$$\sigma=E_{,4}(x,\phi(x))/\left|E_{,4}(x,\phi(x))\right|=\pm 1. \tag{4.2}$$

It is well known from the Riemannian geometry of sub-

spaces that the $X^A{}_{,\mu}$ [cf. Eq. (2.7)] are the components of the unit vectors tangent to the x^μ-parametric lines, in terms of the $\{X^A\}$ coordinates. These are vectors in $M_{(5)}$ tangent to $E_{(4)}$. As vectors in $E_{(4)}$, their covariant derivatives respect to g-differentiation are

$$X^A{}_{,\mu;\nu} = \begin{cases} X^\lambda{}_{,\mu;\nu} = -\dfrac{\eta\phi_{,\mu\nu}\phi^{,\lambda}}{1+\eta\phi_{,\sigma}\phi^{,\sigma}} \\ X^4{}_{,\mu;\nu} = \dfrac{\phi_{,\mu\nu}}{1+\eta\phi_{,\sigma}\phi^{,\sigma}} \end{cases} \tag{4.3}$$

We now calculate the expression for the components of the tensor which gives us the second fundamental form of the embedded space–time (we call it the *Gauss tensor*), namely,

$$\psi = \Omega_{\mu\nu}\,dx^\mu\,dx^\nu. \tag{4.4}$$

It is well known that, since the $\{X^A\}$ is a Cartesian set, the Gauss tensor is given by[18]

$$\Omega_{\mu\nu} = N_A X^A{}_{,\mu;\nu}. \tag{4.5}$$

Therefore, using (4.1) and (4.3), we get

$$\Omega_{\mu\nu} = \frac{\sigma\phi_{,\mu\nu}}{\sqrt{1+\eta\phi_{,\lambda}\phi^{,\lambda}}} = \frac{\sigma\phi_{;\mu;\nu}}{\sqrt{1-\eta\phi_{;\lambda}\phi^{;\lambda}}}. \tag{4.6}$$

The trace of the space–time Gauss tensor is, thus,

$$\Omega = g^{\mu\nu}\Omega_{\mu\nu} = \frac{\sigma\phi^{;\mu}{}_{;\mu}}{\sqrt{1-\eta\phi_{;\lambda}\phi^{;\lambda}}}. \tag{4.7}$$

5. CONCLUSION: THE ADAPTED COORDINATES AS HARMONIC COORDINATES

We are now in position to prove the connection, stated in the Introduction, between our adapted coordinates and the harmonic coordinates. Indeed, the de Donder condition for harmonic coordinates is

$$(\sqrt{-g}\,g^{\mu\nu})_{,\nu} \equiv -\sqrt{-g}\,g^{\nu\lambda}\left\{{\mu \atop \nu\lambda}\right\} = 0. \tag{5.1}$$

Therefore, according to our previous results, cf. Eq. (3.2), (3.3), (3.5), and (4.7), we get

$$(\sqrt{-g}\,g^{\mu\nu})_{,\nu} = -\sigma\eta\phi^{,\mu}\,\Omega. \tag{5.2}$$

Let us recall that the necessary and sufficient condition for a curved space–time (immersed in a five-dimensional space) to be a minimal variety is that the Gauss tensor be traceless. Thus we conclude: For class 1 space–time, the necessary and sufficient condition for the adapted coordinates to be harmonic coordinates is that space–time be a minimal variety. (Clearly so, since $\phi_{,\mu} = 0$ corresponds to Minkowski space–time and affords a trivial model.) Incidentally, this fact tells us that for a minimal class 1 space–time the "fundamental potential" has to satisfy the equation $\phi^{;\mu}{}_{;\mu} = 0$; namely, the general covariant homogeneous wave equation in the curved space–time manifold generated by the potential itself! We wish to remark this fact here, although this paper does not explicitly touch on dynamical questions.

Since the de Donder coordinate condition forms a convenient mathematical tool for treating some problems of general relativity, it is certainly interesting to have a variational principle, with a clear geometric meaning, related with the existance of harmonic coordinates. We have shown, for those space–time metrics which admit a five-dimensional embedding, that the de Donder condition is essentially equivalent with the requirement that curved space–time be a minimal variety. In other words, this means that the underlying variational principle related with harmonic coordinates is that (class 1) space–time must be a solution of the corresponding four-dimensional Plateau's problem.[19] How far can we push this principle into physics, we do not know. Let us remark, however, that a hopeful analogy between the soap film minimal surface and the geometry involved in Einstein's field equations was suggested by Wheeler some years ago.[19]

To end up, we wish to mention here that the detailed study of class 1 space–time deserves some interest by itself for, as is well known, many cosmological solutions to the Einstein field equations belong to this class.[20]

ACKNOWLEDGMENT

This work was strongly influenced by several ideas stated by J. A. Wheeler in a private conversation.

[1] J. Krause, Nuovo Cimento B18, 302 (1973); hereafter referred to as Paper I.

[2] Let us remind that, in general, an n-dimensional Riemannian metric $g_{\mu\nu}$ if of class p if the minimal flat embedding space of $g_{\mu\nu}$ is of dimensions $N = n+p$, with $n \le N \le (n/2)(n+1)$. This definition of p-embeddability alludes to local embeddings. See A. Friedman, J. Math. Mech. 10, 625 (1961).

[3] L. P. Eisenhart, *Riemannian Geometry* (Princeton U.P., Princeton, N.J., 1926), p. 176.

[4] E. Kretchmann, Ann. der Phys. 53, 575 (1917).

[5] A. Einstein, Ann. der Phys. 55, 241 (1918).

[6] V. Fock, Rev. Mod. Phys. 29, 325 (1957); also V. Fock, *The Theory of Space, Time and Gravitation* (Pergamon, New York, 1964), 2nd ed.

[7] See, for instance, P.G. Bergmann and A. Komar, Int. J. Theor. Phys. 5, 15 (1972).

[8] J. Plebanski and J. Rytten, J. Math. Phys. 2, 677 (1961).

[9] T. de Donder, *La gravifique Eisteinienne* (Cauthier-Villars, Paris, 1921).

[10] See J. Plebanski and J. Rytten, Ref. 8, Eqs. (2.4) and (2.6) of that paper.

[11] A. Friedman, Ref. 2.

[12] Capital italic suffixes denote tensor indices in $M_{(5)}$; thus $A, B, C, \cdots = 0, 1, 2, 3, 4$. Greek suffixes $\mu, \nu, \lambda, \cdots = 0, 1, 2, 3$, have their usual meaning as space–time tensor indices.

[13] See Paper I.

[14] L. P. Eisenhart, Ref. 3, p. 143.

[15] See Paper I for a more comprehensive discussion of the issue.

[16] See Paper I.

[17] The possible role of the flat metric field in curved space–time has been discussed by several authors; cf. W.R. Davis, Nuovo Cimento B43, 200 (1966). Also, N. Rosen, Ann. Phys. (N.Y.) 84, 455 (1974).

[18] See L. P. Eisenhart, Ref. 3.

[19] J. A. Wheeler, *Geometrodynamics* (Academic, New York, 1962).

[20] J. Rosen, Rev. Mod. Phys. 37, 204 (1965).

Bifurcate nondiverging null hypersurfaces and trapped surfaces

Paul N. Demmie

University of Pittsburgh at Johnstown, Johnstown, Pennsylvania 15904
(Received 24 June 1974)

An empty spacetime containing a bifurcate nondiverging null hypersurface is investigated, and conditions are given that are necessary and sufficient for the existence of trapped surfaces near this hypersurface to its future. These conditions involve a topological requirement that the two-surface of bifurcation be compact and an inequality that must be satisfied by the characteristic data for this spacetime—the metric of the two-surface of bifurcation and an arbitrary function given on this surface that is shown to be related to angular momentum. The existence of a bifurcate Killing horizon in this spacetime is established. Finally, a Kerr spacetime containing bifurcate Killing horizons is examined, and results pertaining to the existence of trapped surfaces near these horizons to their futures are obtained. These results involve the parameters representing mass and angular momentum per unit mass.

1. INTRODUCTION

In the gravitational collapse of an object to sufficiently small dimensions in an asymptotically flat spacetime, it is conjectured that a domain of trapped surfaces is formed and bounded by a null hypersurface that eventually becomes the absolute event horizon.[1]

A trapped surface is a compact, spacelike two-surface having the property that all null geodesics meeting it orthogonally converge locally to the future.[2] This concept provides for a precise characterization of gravitational collapse that has proceeded beyond the point of no return. Physically, its existence indicates that the gravitational field is so strong that even light rays emitted outward from it cannot escape. In addition, under rather general conditions, its existence implies that the spacetime containing it is singular.[3]

In an asymptotically flat spacetime with a well-defined future null infinity, the absolute event horizon is the boundary of the union of all timelike and null curves that terminate at this future null infinity.[1] This boundary, being a null hypersurface, acts as a one way membrane, permitting light, matter, and radiation to enter the region not extending to infinity but forbidding them to escape. Under rather general conditions, the absolute event horizon must have a normal whose divergence is positive or zero.[4]

A bifuricate nondiverging null hypersurface is the union of two intersecting null hypersurfaces each having a normal whose divergence is zero. The spacelike two-surface on which these nondiverging null hypersurfaces intersect is called the two-surface of bifurcation. It follows from the nature of such a hypersurface that it represents one possible model for an absolute event horizon. Therefore, with the role conjectured for trapped surfaces and the absolute event horizon in gravitational collapse and with the possibility that this horizon is a bifurcate nondiverging null hypersurface, an important problem in the theory of gravitational collapse is the determination of the conditions under which trapped surfaces exist near a bifurcate nondiverging null hypersurface to its future.

Important steps toward solving this problem in the case of an empty spacetime containing a nondiverging null hypersurface were taken by Pajerski and Newman[5] and Demmie and Janis (DJ).[6] Exploiting the property of the Schwarzschild spacetime[7] that the boundary of the region containing trapped surfaces is a nondiverging null hypersurface,[8] they generalized this spacetime to a class of spacetimes each containing a nondiverging null hypersurface. Restrictions on the characteristic data for these spacetimes that are sufficient for the development of trapped surfaces were then determined. In the present work, the case of an empty spacetime containing a bifurcate nondiverging null hypersurface is investigated.

In Sec. 2 a formalism appropriate for the investigation will be presented. This formalism is particularly useful, since it is well suited for problems involving null hypersurfaces and provides for a convenient characterization of trapped surfaces. In Sec. 3 it will be used to determine all empty spacetimes containing a bifurcate nondiverging null hypersurface and their characteristic data. The restrictions on these data and spacetimes that constitute necessary and sufficient conditions for the existence of trapped surfaces near the bifurcate nondiverging null hypersurface to its future will be given in Sec. 4. Also in this section the characteristic data will be discussed. In particular, evidence will be presented there suggesting that a certain piece of these data is related to angular momentum. A bifurcate Killing horizon is the union of two intersecting null hypersurfaces each having a normal that can be normalized to coincide with a Killing vector field.[9] In Sec. 5 the existence of such a horizon in an empty spacetime containing a bifurcate nondiverging null hypersurface will be established. This result implies that a Kerr spacetime[10] whose angular momentum per unit mass does not exceed its mass in magnitude is an example of the type of spacetime considered here. Therefore, in Sec. 6 the Kerr spacetime will be examined. Finally, in Sec. 7 the results of this investigation will be summarized and discussed.

2. THE FORMALISM

The spin-coefficient formalism of Newman and Penrose (NP)[11] is particularly useful for the investigation of an empty spacetime containing a bifurcate non-

Copyright © 1975 American Institute of Physics

diverging null hypersurface. This formalism requires introducing into the tangent space at each point of the spacetime a null tetrad system,[12]

$$\{D = l^\mu \partial/\partial x^\mu,\ \Delta = n^\mu \partial/\partial x^\mu,\ \delta = m^\mu \partial/\partial x^\mu,\ \bar{\delta} = \bar{m}^\mu \partial/\partial x^\mu\}, \tag{2.1}$$

that consists of two real null vectors, D and Δ, and a pair of complex null vectors,

$$\delta = (s_1 + is_2)/\sqrt{2} \quad \text{and} \quad \bar{\delta} = (s_1 - is_2)/\sqrt{2},$$

where s_1 and s_2 are real, orthonormal, spacelike vectors, and that satisfies the orthonormality conditions

$$\begin{aligned} &l_\mu n^\mu = -m_\mu \bar{m}^\mu = 1, \\ &l_\mu l^\mu = n_\mu n^\mu = m_\mu m^\mu = \bar{m}_\mu \bar{m}^\mu = 0, \\ &l_\mu m^\mu = l_\mu \bar{m}^\mu = n_\mu m^\mu = n_\mu \bar{m}^\mu = 0. \end{aligned} \tag{2.2}$$

The components $g^{\mu\nu}$ of the contravariant metric are found from (2.2) to be[13]

$$g^{\mu\nu} = 2l^{(\mu} n^{\nu)} - 2m^{(\mu} \bar{m}^{\nu)}. \tag{2.3}$$

The formalism then provides for a set of partial differential equations equivalent to the Einstein field equations for the determination of the $g^{\mu\nu}$. For an empty spacetime these equations are given in terms of the five independent physical components[14] of the Weyl tensor $C_{\mu\nu\rho\sigma}$,

$$\begin{aligned} \psi_0 &= -C_{\mu\nu\rho\sigma} l^\mu m^\nu l^\rho m^\sigma, \\ \psi_1 &= -C_{\mu\nu\rho\sigma} l^\mu n^\nu l^\rho m^\sigma, \\ \psi_2 &= -C_{\mu\nu\rho\sigma} \bar{m}^\mu n^\nu l^\rho m^\sigma, \\ \psi_3 &= -C_{\mu\nu\rho\sigma} \bar{m}^\mu n^\nu l^\rho n^\sigma, \\ \psi_4 &= -C_{\mu\nu\rho\sigma} \bar{m}^\mu n^\nu \bar{m}^\rho n^\sigma, \end{aligned} \tag{2.4}$$

and the twelve spin coefficients,[15]

$$\begin{aligned} &\kappa = l_{\mu;\nu} m^\mu l^\nu, \quad \nu = -n_{\mu;\nu} \bar{m}^\mu n^\nu, \\ &\rho = l_{\mu;\nu} m^\mu \bar{m}^\nu, \quad \mu = -n_{\mu;\nu} \bar{m}^\mu m^\nu, \\ &\sigma = l_{\mu;\nu} m^\mu m^\nu, \quad \lambda = -n_{\mu;\nu} \bar{m}^\mu \bar{m}^\nu, \\ &\tau = l_{\mu;\nu} m^\mu n^\nu, \quad \pi = -n_{\mu;\nu} \bar{m}^\mu l^\nu, \\ &\alpha = \tfrac{1}{2}(l_{\mu;\nu} n^\mu \bar{m}^\nu - m_{\mu;\nu} \bar{m}^\mu \bar{m}^\nu), \\ &\beta = \tfrac{1}{2}(l_{\mu;\nu} n^\mu m^\nu - m_{\mu;\nu} \bar{m}^\mu m^\nu), \\ &\gamma = \tfrac{1}{2}(l_{\mu;\nu} n^\mu n^\nu - m_{\mu;\nu} \bar{m}^\mu n^\nu), \\ &\epsilon = \tfrac{1}{2}(l_{\mu;\nu} n^\mu l^\nu - m_{\mu;\nu} \bar{m}^\mu l^\nu). \end{aligned} \tag{2.5}$$

Before exhibiting the NP equations, a class of null tetrad systems appropriate for this investigation will be given. This class of null tetrad systems, which consists of those systems associated in a particular way with a class of null coordinate systems, simplifies these equations somewhat. In DJ it was shown that:

In a spacetime there exists a class of null coordinate systems such that any one of these coordinate systems,[16]

$$\{u, r, x^m\}, \tag{2.6a}$$

satisfies the coordinate conditions[17]

$$g^{0\mu} = \delta_1{}^\mu, \tag{2.6b}$$

and has associated with it a particular null tetrad system

$$\{D,\ \Delta,\ \delta,\ \bar{\delta}\}, \tag{2.6c}$$

where

$$\begin{aligned} &D = \partial/\partial r, \\ &\Delta = \partial/\partial u + U\,\partial/\partial r + X^m \partial/\partial x^m, \\ &\delta = \xi^m\, \partial/\partial x^m, \end{aligned}$$

which satisfies the orthonormality conditions (2.2), is unique up to spatial rotations

$$\tilde{D} = D,\ \tilde{\Delta} = \Delta,\ \tilde{\delta} = \exp(iC)\delta \tag{2.6d}$$

with $C = C(u, x^m) = \bar{C}$, and has spin coefficients satisfying

$$\kappa = \epsilon = 0, \quad \rho = \bar{\rho}, \quad \tau = \bar{\alpha} + \beta. \tag{2.6e}$$

The coordinate u in (2.6a) labels null hypersurfaces, $u = \text{const}$, while the coordinate r is an affine parameter along null geodesics in $u = \text{const}$, each having the null vector D as their tangent and each labeled by the x^m coordinates.

From (2.1), (2.2), and (2.6c) the components $g^{\mu\nu}$ of the contravariant metric are

$$\begin{aligned} &g^{0\mu} = \delta_1{}^\mu, \quad g^{11} = 2U, \quad g^{1m} = X^m, \\ &g^{mn} = -(\xi^m \bar{\xi}^n + \bar{\xi}^m \xi^n). \end{aligned} \tag{2.7}$$

With the null tetrad system (2.6c) chosen, the NP equations for an empty spacetime will now be exhibited in three classes[18]: the commutator equations applied to the coordinates, the spin-coefficient equations, and the spin-coefficient form of the Bianchi identities. The commutator equations applied to the coordinates imply that the spin coefficients τ, π, and μ satisfy

$$\tau = \bar{\pi}, \quad \mu = \bar{\mu}, \tag{2.8}$$

and the metric variables U, X^m, and ξ^m satisfy

$$D\xi^m = \rho\xi^m + \sigma\bar{\xi}^m, \tag{2.9a}$$

$$DX^m = 2(\bar{\tau}\xi^m + \tau\bar{\xi}^m), \tag{2.9b}$$

$$DU = -(\gamma + \bar{\gamma}), \tag{2.9c}$$

$$\delta\bar{\xi}^m - \bar{\delta}\xi^m = (\bar{\alpha} - \beta)\bar{\xi}^m + (\bar{\beta} - \alpha)\xi^m, \tag{2.9d}$$

$$\Delta\xi^m - \delta X^m = -(\mu + \bar{\gamma} - \gamma)\xi^m - \bar{\lambda}\bar{\xi}^m, \tag{2.9e}$$

$$\bar{\delta}U = -\nu. \tag{2.9f}$$

With (2.6e) and (2.8) satisfied, the spin-coefficient equations for an empty spacetime are

$$D\rho = \rho^2 + \sigma\bar{\sigma}, \tag{2.10a}$$

$$D\sigma = 2\rho\sigma + \psi_0, \tag{2.10b}$$

$$D\tau = 2\rho\tau + 2\sigma\bar{\tau} + \psi_1, \tag{2.10c}$$

$$D\alpha = (\alpha + \bar{\tau})\rho + \beta\bar{\sigma}, \tag{2.10d}$$

$$D\beta = \rho\beta + (\alpha + \bar{\tau})\sigma + \psi_1, \tag{2.10e}$$

$$D\gamma = 2\tau\alpha + 2\bar{\tau}\beta + \tau\bar{\tau} + \psi_2, \tag{2.10f}$$

$$D\lambda - \bar{\delta}\bar{\tau} = \rho\lambda + \bar{\sigma}\mu + \bar{\tau}^2 + (\alpha - \bar{\beta})\bar{\tau}, \tag{2.10g}$$

$$D\mu - \delta\bar{\tau} = \rho\mu + \sigma\lambda + \tau\bar{\tau} - (\bar{\alpha} - \beta)\bar{\tau} + \psi_2, \tag{2.10h}$$

$$\delta\rho - \bar{\delta}\sigma = \rho\tau - (3\alpha - \bar{\beta})\sigma - \psi_1, \tag{2.10i}$$

$$\delta\alpha - \bar{\delta}\beta = \rho\mu - \sigma\lambda + \alpha\bar{\alpha} - \beta\bar{\beta} - 2\alpha\beta - \psi_2, \tag{2.10j}$$

$$\delta\lambda - \bar{\delta}\mu = \mu\bar{\tau} + (\bar{\alpha} - 3\beta)\lambda - \psi_3, \tag{2.10k}$$

$$\Delta\tau - D\bar{\nu} = -2\mu\tau - 2\bar{\tau}\bar{\lambda} + (\gamma - \bar{\gamma})\tau - \bar{\psi}_3, \tag{2.10l}$$

$$\Delta\lambda - \bar{\delta}\nu = (\bar{\gamma} - 3\gamma - 2\mu)\lambda + (3\alpha + \bar{\beta})\nu - \psi_4, \tag{2.10m}$$

$$\Delta\mu - \delta\nu = -\mu^2 - \lambda\bar{\lambda} - (\gamma + \bar{\gamma})\mu + 2\beta\nu + \nu\bar{\tau}, \tag{2.10n}$$

$$\Delta\beta - \delta\gamma = -\mu\tau + \sigma\nu + (\gamma - \bar{\gamma} - \mu)\beta - \alpha\bar{\lambda}, \tag{2.10o}$$

$$\Delta\sigma - \delta\tau = -\mu\sigma - \rho\bar{\lambda} - 2\beta\tau + (3\gamma - \bar{\gamma})\sigma, \tag{2.10p}$$

$$\Delta\rho - \bar{\delta}\tau = -\mu\rho - \sigma\lambda - 2\alpha\tau + (\gamma + \bar{\gamma})\rho - \psi_2, \tag{2.10q}$$

$$\Delta\alpha - \bar{\delta}\gamma = \rho\nu - (\tau + \beta)\lambda + (\bar{\gamma} - \gamma - \mu)\alpha - \psi_3, \tag{2.10r}$$

and the spin-coefficient form of the Bianchi identities for an empty spacetime are

$$D\psi_1 - \bar{\delta}\psi_0 = 4\rho\psi_1 - (4\alpha - \bar{\tau})\psi_0, \tag{2.11a}$$

$$\Delta\psi_0 - \delta\psi_1 = (4\gamma - \mu)\psi_0 - 2(2\tau + \beta)\psi_1 + 3\sigma\psi_2, \tag{2.11b}$$

$$D\psi_2 - \bar{\delta}\psi_1 = 3\rho\psi_2 + 2\bar{\beta}\psi_1 - \lambda\psi_0, \tag{2.11c}$$

$$\Delta\psi_1 - \delta\psi_2 = \nu\psi_0 + 2(\gamma - \mu)\psi_1 - 3\tau\psi_2 + 2\sigma\psi_3, \tag{2.11d}$$

$$D\psi_3 - \bar{\delta}\psi_2 = 2\rho\psi_3 + 3\bar{\tau}\psi_2 - 2\lambda\psi_1, \tag{2.11e}$$

$$\Delta\psi_2 - \delta\psi_3 = 2\nu\psi_1 - 3\mu\psi_2 - 2\bar{\alpha}\psi_3 + \sigma\psi_4, \tag{2.11f}$$

$$D\psi_4 - \bar{\delta}\psi_3 = \rho\psi_4 + 2(\alpha + 2\bar{\tau})\psi_3 - 3\lambda\psi_2, \tag{2.11g}$$

$$\Delta\psi_3 - \delta\psi_4 = 3\nu\psi_2 - 2(\gamma + 2\mu)\psi_3 + (4\beta - \tau)\psi_4. \tag{2.11h}$$

The formalism presented here provides not only for the determination of the metric variables, spin coefficients, and physical Weyl tensor components of an empty spacetime, but also for a convenient characterization of trapped surfaces. This characterization was discovered by considering the spacelike two-surface

$$S(u, r) = \{(u, r, x^m) : u \text{ and } r \text{ are constant}\}.$$

In order for $S(u, r)$ to be a trapped surface, it must be compact and have the property that all null geodesics meeting it orthogonally converge locally to the future. In DJ it was shown that:

The spacelike two-surface

$$S(u, r) = \{(u, r, x^m) : u \text{ and } r \text{ are constant}\}$$

is a trapped surface if and only if it is compact and everywhere on it the spin coefficients ρ and μ satisfy

$$\rho > 0 \quad \text{and} \quad \mu < 0. \tag{2.12}$$

The divergence referred to in the definition of a trapped surface and a nondiverging null hypersurface is the optical scalar[19] $\frac{1}{2}k^{\mu}{}_{;\mu}$, where the k^{μ} are the components of the vector $k = k^{\mu}\partial/\partial x^{\mu}$. The spin-coefficient form of this quantity and the remaining optical scalars for k, the rotation and shear, are[20]

$$d(k) = \tfrac{1}{2}\eta^{ab}(k_{a,b} - \gamma_{acb}k^c), \tag{2.13a}$$

$$r(k) = [\tfrac{1}{2}(k_{[a,b]} - \gamma_{[a|c|b]}k^c)(k^{a,b} - \gamma^{adb}k_d)]^{1/2}, \tag{2.13b}$$

$$s(k) = [\tfrac{1}{2}(k_{(a,b)} - \gamma_{(a|c|b)}k^c)(k^{a,b} - \gamma^{adb}k_d) - d^2(k)]^{1/2}, \tag{2.13c}$$

respectively, where k_a and $k_{a,b}$ are the physical components of k_{μ} and $(k_a)_{;\mu}$, respectively,

$$(\eta_{ab}) = (\eta^{ab}) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

is the null form of the Minkowski metric, and the $\gamma_{abc} = z_{a\mu;\nu}z_b{}^{\mu}z_c{}^{\nu}$ with $(z_a{}^{\mu}) = (l^{\mu}, n^{\mu}, m^{\mu}, \bar{m}^{\mu})$ are the Ricci rotation coefficients.

The spin-coefficient form of Killing's equations is also particularly useful for this investigation. A Killing vector field is a vector field along whose trajectories the metric does not change.[21] The components of a Killing vector $K = K^{\mu}\partial/\partial x^{\mu}$ satisfy Killing's equations

$$K_{(\mu;\nu)} = 0.$$

These equations are equivalent to

$$K_{(a,b)} = \gamma_{(a|c|b)}K^c,$$

where

$$K = K_1 D + K_0\Delta + K_2\bar{\delta} + \bar{K}_2\delta,$$

the $K_{a,b}$ are the physical components of $(K_a)_{;\mu}$, and the γ_{abc} are the Ricci rotation coefficients. The independent equations obtained from these are in any null tetrad system (2.1)[22]:

$$DK_0 = (\epsilon + \bar{\epsilon})K_0 - \bar{\kappa}K_2 - \kappa\bar{K}_2, \tag{2.14a}$$

$$\Delta K_1 = -(\gamma + \bar{\gamma})K_1 - \nu K_2 - \bar{\nu}\bar{K}_2, \tag{2.14b}$$

$$\delta K_2 = \bar{\lambda}K_0 - \sigma K_1 - (\bar{\alpha} - \beta)K_2, \tag{2.14c}$$

$$\Delta K_0 + DK_1 = (\gamma + \bar{\gamma})K_0 - (\epsilon + \bar{\epsilon})K_1 + (\pi - \bar{\tau})K_2 + (\bar{\pi} - \tau)\bar{K}_2 \tag{2.14d}$$

$$DK_2 + \delta K_0 = (\bar{\alpha} + \beta + \bar{\pi})K_0 - \kappa K_1 + (\epsilon - \bar{\epsilon} - \bar{\rho})K_2 - \sigma\bar{K}_2, \tag{2.14e}$$

$$\Delta K_2 + \delta K_1 = \bar{\nu}K_0 - (\bar{\alpha} + \beta + \tau)K_1 + (\mu + \gamma - \bar{\gamma})K_2 + \bar{\lambda}\bar{K}_2, \tag{2.14f}$$

$$\bar{\delta}K_2 + \delta\bar{K}_2 = (\mu + \bar{\mu})K_0 - (\rho + \bar{\rho})K_1 + (\alpha - \bar{\beta})K_2 + (\alpha - \beta)\bar{K}_2. \tag{2.14g}$$

Equations (2.14) are the spin-coefficient form of Killing's equations. If K is also hypersurface orthogonal, then it must satisfy[23]

$$K_{[\mu;\nu}K_{\rho]} = 0.$$

These equations are equivalent to

$$K_{(a,b}K_{c)} = \gamma_{(a|d|b}K_{c)}K^d.$$

The only equation obtained from this that cannot be derived from (2.14) is in any null tetrad system (2.1):

$$\begin{aligned} &K_2\Delta K_0 + K_1 DK_2 + K_0\delta K_1 \\ &\quad = [-(\bar{\alpha} + \beta)K_1 + \bar{\lambda}\bar{K}_2 + \mu K_2]K_0 + [\bar{\pi}K_0 - (\epsilon + \bar{\epsilon})K_1 \\ &\qquad + (\epsilon - \bar{\epsilon})K_2]K_1 + [(\gamma + \bar{\gamma})K_0 - \tau\bar{K}_2 - \bar{\tau}k_2]K_2. \end{aligned} \tag{2.15}$$

Considerable simplifications result in subsequent expressions by utilization of the differential operators edth, $\eth$, and edth conjugate, $\bar{\eth}$.[24] To define these operators, the notion of a spin-weighted function on a two-surface must first be introduced. Consider a space-

like two-surface S_0 which has in the region $\{(x^\mu)\}$ the metric

$$ds_0^2 = -P^{-2}(z,\bar{z})\,dz\otimes d\bar{z},$$

where $P(z,\bar{z})$ is real and $z=(x^2-ix^3)/\sqrt{2}$. The tangent space at each point of S_0 has a basis

$$\delta_0=(1/\sqrt{2})P\,\partial/\partial z,\quad \bar{\delta}_0=(1/\sqrt{2})P\,\partial/\partial\bar{z}.$$

A function $\eta(z,\bar{z})$ is said to have spin weight s if and only if under the rotation of this basis,

$$\tilde{\delta}_0=(\exp i\psi)\delta_0,$$

where ψ is real, η becomes

$$\tilde{\eta}=(\exp is\psi)\eta.$$

The differential operators $\eth$ and $\bar{\eth}$ are then defined as

$$\eth\eta=P^{1-s}\frac{\partial}{\partial z}(P^s\eta)\quad\text{and}\quad\bar{\eth}\eta=P^{1+s}\frac{\partial}{\partial\bar{z}}(P^{-s}\eta),\quad(2.16)$$

where η is a function with spin weight s.

3. EMPTY SPACETIMES CONTAINING A BIFURCATE NONDIVERGING NULL HYPERSURFACE

The problem of determining all empty spacetimes containing a bifurcate nondiverging null hypersurface can be solved using the formalism presented in Sec. 2. For any one of these spacetimes with a particular null coordinate system (2.6a) and associated null tetrad system (2.6c) introduced in it, the main conditions adopted are that in the region $\{(u,r,x^m)\}$ the bifurcate nondiverging null hypersurface is given by $ur=0$,[25] and the metric variables, spin coefficients, and physical Weyl tensor components are analytic functions of (u,r,x^m).

Since the implications of these quantities being analytic functions of (u,r,x^m) and the $u=0$ null hypersurface being nondiverging were determined in DJ, the results given there that are pertinent to the present investigation will now be summarized.[26] Subject to these conditions, the NP equations involving D yielded the metric variables, spin coefficients, and physical Weyl tensor components in terms of a set of functions given on the spacelike two-surface

$$S_0=\{(u,r,x^m):u=0=r\}.$$

Some of these functions had conditions placed on them in order that the coordinate and tetrad systems chosen initially be specified up to scale transformations,

$$\begin{aligned}&\tilde{u}=Au,\quad \tilde{r}=A^{-1}r,\quad \tilde{x}^m=x^m,\\ &\tilde{D}=A^{-1}D,\quad \tilde{\Delta}=A\Delta,\quad \tilde{\delta}=\delta,\end{aligned}\qquad(3.1)$$

where A is constant, whereas others were determined by the remaining NP equations. The remaining functions constitute the characteristic data for this spacetime. It was shown that[27]:

The null coordinate system (2.6a) and associated null tetrad system (2.6c) can be specified up to scale transformations (3.1) by imposing the conditions

$$[\xi^m(0,0,x^n)]=(1/\sqrt{2})(P,iP),\qquad(3.2a)$$

where P is real, is an analytic function of $(z,\bar{z})$ with $z=(x^2-ix^3)/\sqrt{2}$, and has spin weight zero,

$$\mu(0,0,x^m)=0\ \text{ and }\ \mu(0,r,x^m)\neq 0\ \text{ for }\ r\neq 0,\qquad(3.2b)$$

$$X^m(u,0,x^n)=0,\qquad(3.2c)$$

$$\tau_0(x^m)=\tau(0,0,x^m)=i\eth g,\qquad(3.2d)$$

where g is real, is an analytic function of $(z,\bar{z})$, and has spin weight zero,

$$U(u,0,x^m)=0,\qquad(3.2e)$$

$$\gamma(u,0,x^m)=0.\qquad(3.2f)$$

These conditions will be adopted for subsequent considerations. In DJ it was shown that

In the region $\{(u,r,x^m)\}$ of an empty spacetime containing a nondiverging null hypersurface $u=0$, the metric variables, spin coefficients, and physical Weyl tensor components for this spacetime are determined from Eqs. (2.9), (2.10), and (2.11) by specifying the arbitrary functions

$$\begin{aligned}&P(x^m),\quad \tau_0(x^m)=\tau(0,0,x^m),\\ &\lambda_0(x^m)=\lambda(0,0,x^m),\quad \psi_4(u,0,x^m).\end{aligned}\qquad(3.3)$$

The arbitrary functions (3.3) constitute the characteristic data for an empty spacetime containing a nondiverging null hypersurface.

The implications of the $r=0$ null hypersurface also being nondiverging can be determined by considering its normal. The general $r=$ const hypersurface has a normal k, where

$$k=k^\mu\partial/\partial x^\mu=g^{\mu\nu}r,_\nu\partial/\partial x^\mu=\partial/\partial u.$$

From this, the $r=$ const hypersurface is spacelike, null, or timelike according to the sign of

$$g_{\mu\nu}k^\mu k^\nu=-2U$$

being positive, zero, or negative, respectively, on $r=$ const. That the $r=0$ hypersurface is null follows from (3.2e). Furthermore, from (3.2c) and (3.2e)

$$\Delta=\partial/\partial u$$

on $r=0$. Therefore, the $r=0$ hypersurface is a null hypersurface whose normal is Δ. From (2.12) it follows that the normal to this hypersurface has divergence $\mu(u,0,x^m)$, zero rotation, and shear $\lambda(u,0,x^m)$. Equations (2.10m) and (2.10n) and conditions (2.8) and (3.2) imply that

$$\dot{\mu}(u,0,x^m)=-\mu^2(u,0,x^m)-\lambda(u,0,x^m)\bar{\lambda}(u,0,x^m),\qquad(3.4a)$$

$$\mu(0,0,x^m)=0,\qquad(3.4b)$$

$$\dot{\lambda}(u,0,x^m)=-2\mu(u,0,x^m)\lambda(u,0,x^m)-\psi_4(u,0,x^m),\qquad(3.4c)$$

$$\lambda(0,0,x^m)=\lambda_0(x^m).\qquad(3.4d)$$

Therefore, through these, the characteristic data $\lambda_0(x^m)$ and $\psi_4(u,0,x^m)$ determine $\mu(u,0,x^m)$ and $\lambda(u,0,x^m)$ and hence the divergence and shear of the normal to the $r=0$ null hypersurface. Since this hypersurface is nondiverging,

$$\mu(u,0,x^m)=0.$$

This and Eq. (3.4a) imply that

$$\lambda(u,0,x^m)=0.$$

Therefore,

$$\lambda_0(x^m)=\lambda(0,0,x^m)=0.$$

The result $\lambda(u,0,x^m)=0$ and Eq. (3.4c) imply that

$$\psi_4(u,0,x^m)=0.$$

With $\lambda_0(x^m)$ and $\psi_4(u,0,x^m)$ zero as a result of (3.4) and $r=0$ being nondiverging, it has been established that:

In the region $\{(u,r,x^m)\}$ of an empty spacetime containing a bifurcate nondiverging null hypersurface $ur=0$, the metric variables, spin coefficients, and physical Weyl tensor components for this spacetime are determined from Eqs. (2.9), (2.10), and (2.11) by specifying the arbitrary functions

$$P(x^m) \text{ and } \tau_0(x^m) \tag{3.5}$$

on the two-surface of bifurcation

$$S_0=\{(u,r,x^m): u=0=r\}.$$

The arbitrary functions (3.5) constitute the characteristic data for an empty spacetime containing a bifurcate nondiverging null hypersurface.

The metric variables, spin coefficients, and physical Weyl tensor components of this spacetime can be found by solving Eqs. (2.9), (2.10), and (2.11) subject to (3.2). By doing this it can be established that:

In the region $\{(u,r,x^m)\}$ of an empty spacetime containing a bifurcate nondiverging null hypersurface $ur=0$, this spacetime has:

metric variables:

$$\begin{aligned}U=&(\tfrac{1}{2}K-3\tau_0\bar{\tau}_0)r^2+[\tfrac{1}{12}\bar{\eth}\eth K+(3/4\sqrt{2})(\tau_0\bar{\eth}K+\bar{\tau}_0\eth K)\\&+\tfrac{1}{4}K^2-6\tau_0\bar{\tau}_0K+\tfrac{1}{2}(\bar{\eth}\tau_0)^2+\tfrac{5}{6}(\tau_0\bar{\eth}^2\tau_0-\bar{\tau}_0\eth\bar{\eth}\tau_0)\\&-(2/\sqrt{2})(\tau_0{}^2\bar{\eth}\bar{\tau}_0+\bar{\tau}_0{}^2\eth\tau_0)+8(\tau_0\bar{\tau}_0)^2]r^3u+\cdots\\&+U_i(x^m)r^{i+2}u^i+\cdots,\end{aligned} \tag{3.6a}$$

$$\begin{aligned}X^m=&2(\bar{\tau}_0\xi_0{}^m+\tau_0\bar{\xi}_0{}^m)r+\{[-(1/2\sqrt{2})\bar{\eth}K+3\bar{\tau}_0K-6\tau_0\bar{\tau}_0{}^2\\&-\tfrac{1}{2}\bar{\eth}^2\tau_0+(3/\sqrt{2})\bar{\tau}_0\bar{\eth}\tau_0+(3/\sqrt{2})\tau_0\bar{\eth}\tau_0]\xi_0{}^m\\&+[(-1/2\sqrt{2})\eth K+3\tau_0K-6\tau_0{}^2\bar{\tau}_0+\tfrac{1}{2}\eth\bar{\eth}\tau_0-(3/\sqrt{2})\bar{\tau}_0\bar{\eth}\tau_0\\&+(3/\sqrt{2})\bar{\tau}_0\eth\tau_0]\bar{\xi}_0{}^m\}r^2u+\cdots+X_i{}^m(x^n)r^{i+1}u^i+\cdots,\end{aligned} \tag{3.6b}$$

$$\begin{aligned}\xi^m=&\xi_0{}^m+\{(\tfrac{1}{2}K-\tau_0\bar{\tau}_0)\xi_0{}^m+[(1/\sqrt{2})\eth\tau_0-\tau_0{}^2]\bar{\xi}_0{}^m\}ru\\&+\cdots+\xi_i{}^m(x^n)r^iu^i+\cdots;\end{aligned} \tag{3.6c}$$

spin coefficients:

$$\begin{aligned}\rho=&(\tfrac{1}{2}K-\tau_0\bar{\tau}_0)u+[\tfrac{1}{4}K^2-\tau_0\bar{\tau}_0K+2(\tau_0\bar{\tau}_0)^2\\&+\tfrac{1}{2}(\eth\tau_0)(\bar{\eth}\bar{\tau}_0)-(1/\sqrt{2})\tau_0{}^2\bar{\eth}\bar{\tau}_0-(1/\sqrt{2})\bar{\tau}_0{}^2\eth\tau_0]ru^2+\cdots\\&+\rho_i(x^m)r^iu^{i+1}+\cdots,\end{aligned} \tag{3.7a}$$

$$\begin{aligned}\sigma=&[(1/\sqrt{2})\eth\tau_0-\tau_0{}^2]u+[-\tfrac{1}{8}\eth^2K+(2/\sqrt{2})\tau_0\eth K-4\tau_0{}^2K\\&+(1/4\sqrt{2})\eth^2\bar{\eth}\tau_0-2\tau_0\eth\bar{\eth}\tau_0+(6/\sqrt{2})\tau_0{}^2\bar{\eth}\tau_0-(2/\sqrt{2})\tau_0\bar{\tau}_0\eth\tau_0\\&+4\tau_0{}^3\bar{\tau}_0+(1/\sqrt{2})K\eth\tau_0]ru^2+\cdots+\sigma_i(x^m)r^iu^{i+1}+\cdots,\end{aligned} \tag{3.7b}$$

$$\begin{aligned}\mu=&-(\tfrac{1}{2}K-\tau_0\bar{\tau}_0)r+[-\tfrac{1}{4}\bar{\eth}\eth K-\tfrac{1}{2}K^2+(1/2\sqrt{2})(\tau_0\bar{\eth}K\\&+\bar{\tau}_0\eth K)+6\tau_0\bar{\tau}_0K-5(\tau_0\bar{\tau}_0)^2-\tfrac{3}{2}(\bar{\eth}\tau_0)^2+\tfrac{3}{2}(\bar{\tau}_0\eth\bar{\eth}\tau_0\\&-\tau_0\bar{\eth}^2\tau_0)+(\eth\tau_0)(\bar{\eth}\bar{\tau}_0)]r^2u+\cdots+\mu_i(x^m)r^{i+1}u^i\\&+\cdots,\end{aligned} \tag{3.7c}$$

$$\begin{aligned}\lambda=&[(1/\sqrt{2})\bar{\eth}\bar{\tau}+\bar{\tau}_0^2]r+[-\tfrac{1}{8}\bar{\eth}^2K+(1/\sqrt{2})\bar{\tau}_0\bar{\eth}K+\tfrac{5}{2}\bar{\tau}_0{}^2K\\&+(3/2\sqrt{2})K\bar{\eth}\bar{\tau}_0-(1/4\sqrt{2})\bar{\eth}^3\tau_0+\tfrac{1}{2}\bar{\tau}_0\bar{\eth}^2\tau_0-(3/\sqrt{2})\tau_0\bar{\tau}_0\bar{\eth}\bar{\tau}_0\\&-5\tau_0\bar{\tau}_0{}^3+\tfrac{1}{2}\tau_0\bar{\eth}^2\bar{\tau}_0+(\bar{\eth}\tau_0)(\bar{\eth}\bar{\tau}_0)]r^2u+\cdots\\&+\lambda_i(x^m)r^{i+1}u^i+\cdots,\end{aligned} \tag{3.7d}$$

$$\begin{aligned}\tau=&\tau_0+[-(1/2\sqrt{2})\eth K+\tfrac{5}{2}\tau_0K+\tfrac{1}{2}\eth\bar{\eth}\tau_0-(3/\sqrt{2})\tau_0\bar{\eth}\tau_0\\&+(2/\sqrt{2})\bar{\tau}_0\eth\tau_0-4\tau_0{}^2\bar{\tau}_0]ru+\cdots+\tau_i(x^m)r^iu^i+\cdots,\end{aligned} \tag{3.7e}$$

$$\nu=-\bar{\eth}U, \tag{3.7f}$$

$$\begin{aligned}\alpha=&(\alpha_0+\tfrac{1}{2}\bar{\tau}_0)+\{\tfrac{3}{4}\bar{\tau}_0K-2\tau_0\bar{\tau}_0{}^2+(1/2\sqrt{2})\tau_0\bar{\eth}\bar{\tau}_0\\&+(\tfrac{1}{2}K-\tau_0\bar{\tau}_0)\alpha_0+[-(1/\sqrt{2})\bar{\eth}\bar{\tau}_0+\bar{\tau}_0{}^2]\bar{\alpha}_0\}ru+\cdots\\&+\alpha_i(x^m)r^iu^i+\cdots,\end{aligned} \tag{3.7g}$$

$$\begin{aligned}\beta=&(-\bar{\alpha}_0+\tfrac{1}{2}\tau_0)+\{-(1/2\sqrt{2})\eth K+\tfrac{7}{4}\tau_0K\\&+(3/2\sqrt{2})\bar{\tau}_0\eth\tau_0+\tfrac{1}{2}\eth\bar{\eth}\tau_0\\&-(3/\sqrt{2})\tau_0\bar{\eth}\tau_0-2\tau_0{}^2\bar{\tau}_0+[(1/\sqrt{2})\eth\tau_0-\tau_0{}^2]\alpha_0+(-\tfrac{1}{2}K\\&+\tau_0\bar{\tau}_0)\bar{\alpha}_0\}ru+\cdots+\beta_i(x^m)r^iu^i+\cdots,\end{aligned} \tag{3.7h}$$

$$\begin{aligned}\gamma=&[-\tfrac{1}{2}K+3\tau_0\bar{\tau}_0+(1/\sqrt{2})\bar{\eth}\bar{\tau}_0+2\alpha_0\tau_0-2\bar{\alpha}_0\bar{\tau}_0]r\\&+\{-\tfrac{1}{8}\bar{\eth}\eth K+(1/4\sqrt{2})\tau_0\bar{\eth}K-(5/2\sqrt{2})\bar{\tau}_0\eth K-\tfrac{3}{8}K^2\\&+(3/2\sqrt{2})K\bar{\eth}\tau_0\\&+9\tau_0\bar{\tau}_0K+(1/4\sqrt{2})\eth\bar{\eth}^2\tau_0-\tfrac{3}{4}(\bar{\eth}\tau_0)^2+\tfrac{5}{4}\bar{\tau}_0\eth\bar{\eth}\tau_0-\tfrac{5}{4}\tau_0\bar{\eth}^2\tau_0\\&-(6/\sqrt{2})\tau_0\bar{\tau}_0\eth\tau_0+(5/2\sqrt{2})\tau_0{}^2\bar{\eth}\bar{\tau}_0+(7/2\sqrt{2})\bar{\tau}_0{}^2\eth\tau_0\\&-12(\tau_0\bar{\tau}_0)^2\\&+[(-1/2\sqrt{2})\eth K+3\tau_0K+\tfrac{1}{2}\eth\bar{\eth}\tau_0\\&-(3/\sqrt{2})\tau_0\bar{\eth}\tau_0+(3/\sqrt{2})\bar{\tau}_0\eth\tau_0\\&-6\tau_0{}^2\bar{\tau}_0]\alpha_0+[(1/2\sqrt{2})\bar{\eth}K-3\bar{\tau}_0K+\tfrac{1}{2}\bar{\eth}^2\tau_0-(3/\sqrt{2})\bar{\tau}_0\bar{\eth}\tau_0\\&-(3/\sqrt{2})\tau_0\bar{\eth}\bar{\tau}_0+6\bar{\tau}_0{}^2\tau_0]\bar{\alpha}_0\}r^2u+\cdots+\gamma_i(x^m)r^{i+1}u^i\\&+\cdots;\end{aligned} \tag{3.7i}$$

physical Weyl tensor components:

$$\begin{aligned}\psi_0=&[-\tfrac{1}{8}\eth^2K+(2/\sqrt{2})\tau_0\eth K-3\tau_0{}^2K+(1/4\sqrt{2})\eth^2\bar{\eth}\tau_0-2\tau_0\eth\bar{\eth}\tau_0\\&+(6/\sqrt{2})\tau_0{}^2\bar{\eth}\tau_0]u^2+\cdots+\psi_0{}^i(x^m)r^iu^{i+2}+\cdots,\end{aligned} \tag{3.8a}$$

$$\begin{aligned}\psi_1=&[(-1/2\sqrt{2})\eth K+\tfrac{3}{2}\tau_0K+\tfrac{1}{2}\eth\bar{\eth}\tau_0-(3/\sqrt{2})\tau_0\bar{\eth}\tau_0]u+\cdots\\&+\psi_1{}^i(x^m)r^iu^{i+1}+\cdots,\end{aligned} \tag{3.8b}$$

$$\psi_2=[-\tfrac{1}{2}K+(1/\sqrt{2})\bar{\eth}\tau_0]+\cdots+\psi_2{}^i(x^m)r^iu^i+\cdots, \tag{3.8c}$$

$$\begin{aligned}\psi_3=&[(-1/2\sqrt{2})\bar{\eth}K-\tfrac{3}{2}\bar{\tau}_0K+\tfrac{1}{2}\bar{\eth}^2\tau_0+(3/\sqrt{2})\bar{\tau}_0\bar{\eth}\tau_0]r+\cdots\\&+\psi_3{}^i(x^m)r^{i+1}u^i+\cdots,\end{aligned} \tag{3.8d}$$

$$\begin{aligned}\psi_4=&[-\tfrac{1}{8}\bar{\eth}^2K-(2/\sqrt{2})\bar{\tau}_0\bar{\eth}K-3K\bar{\tau}_0{}^2+(1/4\sqrt{2})\bar{\eth}^3\tau_0+2\bar{\tau}_0\bar{\eth}^2\tau_0\\&+(6/\sqrt{2})\bar{\tau}_0{}^2\bar{\eth}\tau_0]r^2+\cdots+\psi_4{}^i(x^m)r^{i+2}u^i+\cdots;\end{aligned} \tag{3.8e}$$

where $K=\bar{\eth}\eth\ln P$, $\alpha_0=(1/2\sqrt{2})\bar{\eth}\ln P$, and $i=0,1,2,\cdots$.

The general form of these expressions in powers of u and r can be established by mathematical induction[28] and is an important consequence of the existence of a bifurcate nondiverging null hypersurface.

4. NECESSARY AND SUFFICIENT CONDITIONS FOR THE EXISTENCE OF TRAPPED SURFACES

An empty spacetime containing a bifurcate nondiverging null hypersurface will be considered, and necessary and sufficient conditions will be given for the existence of trapped surfaces near this hypersurface to its future.

Consider a region $\{(u, r, x^m)\}$ in this spacetime containing a bifurcate nondiverging null hypersurface $ur=0$. The future or $ur=0$ is the set

$$\{(u, r, x^m) : u>0 \text{ and } r>0\}.$$

The two-surface of bifurcation, S_0, separates $ur=0$ into four components: the $r>0$, $r<0$, $u>0$, and $u<0$ branches given by

$$\{(u, r, x^m) ; u=0 \text{ and } r>0\}, \quad \{(u, r, x^m) : u=0 \text{ and } r<0\},$$

$$\{(u, r, x^m) : u>0 \text{ and } r=0\}, \quad \{(u, r, x^m) : u<0 \text{ and } r=0\},$$

respectively.

Since both the $u=0$ and $r=0$ null hypersurfaces are nondiverging, it follows from the sufficient conditions for the existence of trapped surfaces given in DJ that:

In the region $\{(u, r, x^m)\}$ of an empty spacetime containing a bifurcate nondiverging null hypersurface $ur=0$, trapped surfaces exist to the future of this hypersurface if

$$S_0 \text{ is compact} \tag{4.1a}$$

and the characteristic data P and τ_0 satisfy

$$\tfrac{1}{2}K - \tau_0\bar{\tau}_0 > 0 \quad \text{everywhere on } S_0 \tag{4.1b}$$

where $K=\bar{\eth}\eth \ln P$ is the Gaussian curvature of S_0.[29]

It is possible to extend this result to one for which (4.1a) and (4.1b) are both necessary and sufficient conditions for the existence of trapped surfaces by specifying more precisely their location.

Consider the two-surface $S(u_0, r_0)$ in the region $\{(u, r, x^m)\}$. An ϵ-neighborhood of $S(u_0, r_0)$ is the set

$$N_\epsilon(u_0, r_0) = \{(u, r, x^m) : |u-u_0| < \epsilon, \ |r-r_0| < \epsilon\}$$

A neighborhood of $S(u_0, r_0)$ is a set N such that

(1) $S(u_0, r_0)$ is contained in N

(2) For every point $(u', r', x^{m'})$ in N there exists an $\epsilon>0$ such that $N_\epsilon(u', r')$ is contained in N.

Clearly an ϵ-neighborhood of $S(u_0, r_0)$ is a neighborhood of $S(u_0, r_0)$.

With these definitions given, it will now be shown that

In the region $\{(u, r, x^m)\}$ of an empty spacetime containing a bifurcate nondiverging null hypersurface $ur=0$, trapped surfaces exist to the future of $ur=0$ in every neighborhood of S_0 and all two-surfaces $S(u, r)$ contained in the $r>0$ and $u>0$ branches of this hypersurface if and only if

$$S_0 \text{ is compact} \tag{4.2}$$

and the characteristic data P and τ_0 satisfy

$$\tfrac{1}{2}K - \tau_0\bar{\tau}_0 > 0 \quad \text{everywhere on } S_0,$$

where $K=\bar{\eth}\eth \ln P$ is the Gaussian curvature of S_0.

Proof: Consider a region $\{(u, r, x^m)\}$ of an empty spacetime containing a bifurcate nondiverging null hypersurface $ur=0$.

Suppose that there exist trapped surfaces to the future of $ur=0$ in every neighborhood of S_0 and all two-surfaces $S(u, r)$ contained in the $r>0$ and $u>0$ branches of $ur=0$. Let N be some neighborhood of S_0. Then there exists a trapped surface $S(u, r)$ in N. Since $S(u, r)$ is compact and for fixed u and r the function

$$f_{(u,r)} : S_0 \to S(u, r), \quad \text{where } f_{(u,r)}(0, 0, x^m) = (u, r, x^m),$$

is a homeomorphism of S_0 onto $S(u, r)$, S_0 is also compact. Now, consider any two-surface $S(0, r_0)$ in the $r>0$ branch of $ur=0$. Let $\{N_{1/n}(0, r_0)\}$, $n=1, 2, \cdots$, be a sequence of ϵ neighborhoods of $S(0, r_0)$. Then for each n there exists a trapped surface $S(u_n, r_n)$ in $N_{1/n}(0, r_0)$. Since the spin coefficients ρ and μ are continuous functions of (u, r, x^m) and are given by (3.7a) and (3.7c), respectively, it follows that

$$\lim_{n\to\infty} \rho(u_n, r_n, x^m) = \rho(0, r_0, x^m) = 0$$

and

$$\lim_{n\to\infty} \mu(u_n, r_n, x^m) = \mu(0, r_0, x^m) = -(\tfrac{1}{2}K - \tau_0\bar{\tau}_0)r_0.$$

From these, since the $S(u_n, r_n)$ are trapped surfaces and hence by (2.12) $\rho>0$ and $\mu<0$ everywhere on $S(u_n, r_n)$, it follows that $\frac{1}{2}K - \tau_0\bar{\tau}_0 \geq 0$ everywhere on S_0. The possibility that $\frac{1}{2}K - \tau_0\bar{\tau}_0 = 0$ everywhere on S_0 is inconsistent with (3.2b). Therefore, $\frac{1}{2}K - \tau_0\bar{\tau}_0 > 0$ everywhere on S_0.

Conversely, suppose that S_0 is compact and everywhere on it $\frac{1}{2}K - \tau_0\bar{\tau}_0 > 0$. Let $S(0, r_0)$ be any two-surface contained in the $r>0$ branch of $ur=0$ and let N be any neighborhood of $S(0, r_0)$. Then there exists an $\epsilon>0$ such that $N_\epsilon(0, r_0)$ is contained in N. From (3.7a) and (3.7c), ϵ can be chosen sufficiently small that $\rho>0$ and $\mu<0$ in the intersection of $N_\epsilon(0, r_0)$ and the future of $ur=0$, since $\frac{1}{2}K - \tau_0\bar{\tau}_0 > 0$ everywhere on S_0. Let $S(u, r)$ be any two-surface in the intersection of $N_\epsilon(0, r_0)$ and the future of $ur=0$. Then $\rho>0$ and $\mu<0$ everywhere on $S(u, r)$. Furthermore, since S_0 is compact and is homeomorphic to $S(u, r)$ as previously shown, $S(u, r)$ is also compact. Hence by (2.12), $S(u, r)$ is a trapped surface. Therefore, trapped surfaces exist to the future of $ur=0$ in every neighborhood of all two-surfaces contained in the $r>0$ branch of $ur=0$. Similarly, it can be shown that trapped surfaces exist to the future of $ur=0$ in every neighborhood of S_0 and all two-surfaces contained in the $u>0$ branch of $ur=0$. QED

Consider the characteristic data (3.5). The function P is the most important of (3.5) for the existence of trapped surfaces, since unless the spacelike two-surface S_0 with induced covariant metric, $ds_0^2 = -P^{-2}\, dz \otimes d\bar{z}$, has strictly positive Gaussian curvature, $K=\bar{\eth}\eth \ln P$, there is no possibility of satisfying (4.1b). In the case of a spherically symmetric spacetime it is known that

S_0 is a two-sphere with $K>0$.[27] Although in the case of an arbitrary spacetime S_0 may be chosen to be a two-sphere, there do exist other compact two-surfaces with strictly positive Gaussian curvature,[30] and hence this choice is not imperative.

Although a compact S_0 with $K>0$ is essential for the existence of trapped surfaces according to (4.2), this is not the only consequence of these conditions. It follows from these conditions and the Gauss–Bonnet theorem that S_0 is topologically a two-sphere.[31] Furthermore, under these conditions, the $u=0$ and $r=0$ nondiverging null hypersurfaces must each be topologically the product of a two-sphere and the real line.

The function τ_0 in (3.5) is also very important for the existence of trapped surfaces, since even if $K>0$, the magnitude of τ_0 could be sufficiently large that (4.1b) is violated. This possibility suggests that there may exist a relationship between τ_0 and angular momentum. That such a relationship exists was shown in DJ for the linearized Kerr spacetime and will later be shown for the Kerr spacetime.

Additional evidence that τ_0 is related to angular momentum can be given by considering the propagation of the null tetrad system (2.6c) along the generators of $ur=0$. A tetrad system is normally said to be propagated without rotation along a timelike curve if and only if it is Fermi propagated along this curve, which in the case of a timelike geodesic is equivalent to being parallelly propagated.[32] If this notion is extended to null geodesics, then it can be said that the null tetrad system (2.6c) is propagated without rotation along the generators of $ur=0$ if and only if it is parallelly propagated along them. From (2.5), subject to the conditions (2.6e), (2.8), and (3.2), it follows that

$$l^{\mu}{}_{;\nu}l^{\nu}=0, \quad m^{\mu}{}_{;\nu}l^{\nu}=\tau l^{\mu},$$

and

$$n^{\mu}{}_{;\nu}l^{\nu}=\bar{\tau}m^{\mu}+\tau\bar{m}^{\mu}$$

on $u=0$ and

$$n^{\mu}{}_{;\nu}n^{\nu}=0, \quad m^{\mu}{}_{;\nu}n^{\nu}=-\bar{\tau}n^{\mu},$$

and

$$l^{\mu}{}_{;\nu}n^{\nu}=-\bar{\tau}m^{\mu}-\tau\bar{m}^{\mu}$$

on $r=0$. Therefore, since (2.7e) implies that $\tau=\tau_0$ everywhere on $ur=0$, the null tetrad system (2.5c) is parallelly propagated along the generators of $ur=0$ if and only if $\tau_0=0$.

5. EXISTENCE OF A BIFURCATE KILLING HORIZON

The existence of a bifurcate Killing horizon in an empty spacetime containing a bifurcate nondiverging null hypersurface will now be established. More precisely, it will be shown that:

A pair of intersecting null hypersurfaces in an empty spacetime is a bifurcate Killing horizon if and only if it is a bifurcate nondiverging null hypersurface. (5.1)

Proof: Suppose M is a spacetime containing a bifurcate Killing horizon $H=H_1\cup H_2$ for the Killing vector field K. In M introduce a null coordinate system (2.6a) and associated null tetrad system (2.6c) such that the conditions (2.6b), (2.6e), and (3.2) are satisfied for (2.6a) and (2.6c), H_1 is given by $u=0$, and H_2 is given $r=0$. Since H_1 and H_2 are Killing horizons for K,

$$K=K_1D+K_0\Delta+K_2\bar{\delta}+\bar{K}_2\delta$$

must satisfy

$$K_0=K_2=0 \text{ and } K_1\neq 0 \text{ everywhere on } H_1 \qquad (5.2a)$$

and

$$K_1=K_2=0 \text{ and } K_0\neq 0 \text{ everywhere on } H_2. \qquad (5.2b)$$

From (2.13a) and (2.6e) the divergence of D is $-\rho$ and from (2.13a), (2.8), and (3.2f) the divergence of Δ on H_2 is μ evaluated on H_2. Equation (2.14g) and (5.2a) imply that $\rho=0$ on H_1. Similarly, Eq. (2.14g) and (5.2b) imply that $\mu=0$ on H_2. Therefore, H_1 and H_2 are both nondiverging null hypersurfaces, and hence H is a bifurcate nondiverging null hypersurface.

Conversely, suppose that M is an empty spacetime containing a bifurcate nondiverging null hypersurface H. In M introduce a null coordinate system (2.6a) and associated null tetrad system (2.6c) such that the conditions (2.6b), (2.6e), and (3.2) are satisfied for (2.6a) and (2.6c) and H is given by $ur=0$. Consider the vector field

$$K=K_1D+K_0\Delta+K_2\bar{\delta}+\bar{K}_2\delta, \qquad (5.3)$$

where

$$K_0=au, \quad K_1=-a(r+Uu),$$

and

$$K_2=(\xi^2X^3-\xi^3X^2)au/(\xi^2\bar{\xi}^3-\bar{\xi}^2\xi^3),$$

with a being a nonzero real number. Also, consider the spin-coefficient form of Killing's equations, Eqs. (2.14), expressed in the null coordinate and tetrad system introduced into M. Since

$$DK_0=0,$$

K satisfies Eq. (2.14a). By Eq. (2.9c),

$$\Delta K_0+DK_1=a-a(1+uDU)=au(\gamma+\bar{\gamma})=(\gamma+\bar{\gamma})K_0.$$

Therefore, K satisfies Eq. (2.14d). Similarly after a slightly more tedious calculation using Eqs. (2.9a) and (2.9b), it can be shown that K satisfies Eq. (2.14e). Thus far the bifurcate nondiverging null hypersurface property of H was not needed to establish that K satisfies Eqs. (2.14a), (2.14d), and (2.14e). This property, however, is essential for K to satisfy the remaining equations of (2.14). After some very tedious calculations employing Eqs. (2.9), it can be shown that K satisfies Eqs. (2.14b), (2.14c), (2.14f), and (2.14g) if and only if K satisfies the equations

$$(r\partial/\partial r-u\partial/\partial u-2)U=0,$$

$$\xi^2(u\partial/\partial u-r\partial/\partial r)\xi^3=\xi^3(u\partial/\partial u-r\partial/\partial r)\xi^2,$$

$$\xi^2(u\partial/\partial u-r\partial/\partial r+1)X^3=\xi^3(u\partial/\partial u-r\partial/\partial r+1)X^2$$

$$\xi^2(u\partial/\partial u-r\partial/\partial r)\bar{\xi}^3-\bar{\xi}^2(u\partial/\partial u-r\partial/\partial r)\xi^3$$

$$=\xi^3(u\partial/\partial u-r\partial/\partial r)\bar{\xi}^2-\bar{\xi}^3(u\partial/\partial u-r\partial/\partial r)\xi^2,$$

respectively. From the form of U, X^m, and ξ^m in powers of u and r that is given by (2.6) and that, as indicated previously, is a consequence of $ur=0$ being a bifurcate nondiverging null hypersurface, it follows immediately that these equations are satisfied by U, X^m, and ξ^m. Thus K satisfies Eqs. (2.14) and is therefore a Killing vector field. Furthermore, from these expressions for K_0, K_1, and K_2 and from (3.6)

$$K=-arD \quad \text{on } u=0$$

and

$$K=au\Delta \quad \text{on } r=0.$$

Therefore, $u=0$ and $r=0$ are Killing horizons for K and hence $ur=0$ is a bifurcate Killing horizon for K. QED

From the proof of (5.1) it follows that:

For an empty spacetime containing a bifurcate nondiverging null hypersurface H there exists a Killing vector field K given in the region $\{(u, r, x^m)\}$ by

$$K=a(u\partial/\partial u - r\partial/\partial r + 2uX^m\partial/\partial x^m), \tag{5.4}$$

where a is a real constant, having H, which in $\{(u, r, x^m)\}$ is given by $ur=0$, for a bifurcate Killing horizon.

The condition (3.2c) and the expression for K given in (5.4) imply that K vanishes on S_0 and therefore S_0 is a fixed point surface for K. On $u=0$ or $r=0$, S_0 separates each nondiverging, nonrotating, and shear-free null geodesic generator of one branch from its continuation on the other branch.

Near S_0 the branches of $ur=0$, divide the spacetime into four disjoint open parts such that the trajectories of K are timelike in two of these regions and spacelike in the others. An examination of the square of the norm of K, $g(K,K)=(K_1K_0-K_2\bar{K}_2)$, reveals that

$$g(K,K)=-2a^2ur+\cdots+(\quad)u^ir^i+\cdots, \quad i=1,2,\cdots.$$

Therefore, at least for small $|u|$ or $|r|$, K is timelike in the regions where $ur<0$ and spacelike in the regions where $ur>0$. Hence an empty spacetime containing a bifurcate nondiverging null hypersurface is locally stationary.

If K is a hypersurface orthogonal Killing vector field, then K_0, K_1, and K_2 must satisfy Eq. (2.15) in addition to Eqs. (2.14). By substituting these quantities in Eq. (2.15) and examining this equation in powers of u, it can be shown that τ_0 must satisfy

$$\eth\bar{\eth}\tau_0=0. \tag{5.5}$$

The tedious nature of the calculations required to find possible further conditions resulting from K satisfying Eq. (2.15) inhibits the search for them. Thus conditions other than (5.5) have not been determined.

6. BIFURCATE NONDIVERGING NULL HYPERSURFACES AND THE KERR SPACETIME

The Kerr spacetime[10] will now be examined as an example of an empty spacetime that contains a bifurcate nondiverging null hypersurface. This spacetime has a metric that depends on two parameters, m and a, representing mass and angular momentum per unit mass, respectively. That it is an appropriate example follows from its containing, when $a^2<m^2$, the bifurcate Killing horizons $H_\pm$ for the Killing vector fields $K_\pm$,[33] respectively, where in the region $\{(\tilde{u}, \tilde{r}, \tilde{\theta}, \tilde{\phi})\}$

$$H_\pm=\{(\tilde{u}, \tilde{r}, \tilde{\theta}, \tilde{\phi}): \tilde{r}=r_\pm=m\pm(m^2-a^2)^{1/2}\}$$

and

$$K_\pm=\partial/\partial\tilde{u}+[a/(r_\pm^2+a^2)]\partial/\partial\tilde{\phi}.$$

The characteristic data that determine the Kerr spacetime in the regions containing $H_\pm$ will now be obtained. To accomplish this, the contravariant metric tensor, whose components in $\{(\tilde{u}, \tilde{r}, \tilde{\theta}, \phi)\}$ are[34]

$$\frac{1}{(\tilde{r}^2+a^2\cos^2\tilde{\theta})} \times \begin{bmatrix} -a^2\sin^2\tilde{\theta} & \tilde{r}^2+a^2 & 0 & -a \\ \tilde{r}^2+a^2 & -(\tilde{r}^2-2m\tilde{r}+a^2) & 0 & a \\ 0 & 0 & -1 & 0 \\ -a & a & 0 & -\csc^2\tilde{\theta} \end{bmatrix},$$

must first be expressed in a null coordinate system $\{u, r, x^m\}$ such that the conditions (2.6) and (3.2) are satisfied by $\{u, r, x^m\}$ and its associated null tetrad system and $H_\pm$ is given by $ur=0$. Once this is done, the characteristic data $(\xi_0{}^m)$ and τ_0 can be determined from (2.7), (2.9b), and (3.6). By considering coordinate transformations of the form

$$u=A_0+A_1(\tilde{r}-r_\pm)+A_2(\tilde{r}-r_\pm)^2+\cdots,$$
$$r=B_1(\tilde{r}-r_\pm)+B_2(\tilde{r}-r_\pm)^2+\cdots,$$
$$x^k=Y_0{}^k+Y_1{}^k(\tilde{r}-r_\pm)+Y_2{}^k(\tilde{r}-r_\pm)^2+\cdots,$$

where the A_i, B_{i+1}, and $Y_i{}^k$, $i=0, 1, \cdots$, are functions of $(\tilde{u}, \tilde{\theta}, \tilde{\phi})$, and requiring as implied by the conditions (2.6) and (3.2) that the resulting $(g^{\mu\nu})$ satisfy

$$g^{0\mu}=\delta_1^\mu, \quad g^{1k}(u, 0, x^m)=0,$$
$$g^{11}(u, 0, x^m)=0, \quad \partial g^{11}(u, 0, x^m)/\partial r=0,$$

it can be shown after some tedious calculations that, by taking

$$A_0=-\frac{2(r_\pm^2+a^2\cos^2\tilde{\theta})}{(r_\pm-r_\mp)}\exp\left(\frac{-(r_\pm-r_\mp)}{2(r_\pm^2+a^2)}\tilde{u}\right),$$

$$A_1=(a^2\sin^2\tilde{\theta}A_{0,u}^2+A_{0,\tilde{\theta}}^2)/(r_\pm^2+a^2)A_{0,u},$$

$$B_1=(r_\pm^2+a^2\cos^2\tilde{\theta})^{1/2}\exp\left(\frac{(r_\pm-r_\mp)}{2(r_\pm^2+a^2)}\tilde{u}\right),$$

$$Y_0{}^2=\tilde{\theta}, \quad Y_0{}^3=\tilde{\phi}-a\tilde{u}/(r_\pm^2+a^2),$$

$$Y_1{}^2=2a^2\sin\tilde{\theta}\cos\tilde{\theta}/(r_\pm-r_\mp)(r_\pm^2+a^2\cos^2\tilde{\theta}),$$

and

$$Y_1{}^3=a(r_\pm^2+a^2\cos^2\tilde{\theta})/(r_\pm^2+a^2)^2,$$

enough terms in these coordinate transformations are obtained to enable $(\xi_0{}^m)$ and τ_0 to be calculated. After calculating g^{mn} and g^{1m}, it then follows from

$$g^{mn}(0, r, x^l) = -(\xi_0{}^m \bar{\xi}_0{}^n + \bar{\xi}_0{}^m \xi_0{}^n) = g^{mn}(u, 0, x^l)$$

and

$$DX^m(0, r, x^l) = (\bar{\tau}_0 \xi_0{}^m + \tau_0 \bar{\xi}_0{}^m) = g^{1m}(0, r, x^l)/r$$
$$= \lim_{r \to 0} g^{1m}(u, r, x^l)/r$$

that the characteristic data $(\xi_0{}^m)$ and τ_0 are

$$(\xi_0{}^m) = \frac{1}{\sqrt{2}} \left[\frac{1}{(r_\pm^2 + a^2 \cos^2\theta)^{1/2}}, \quad \frac{i(r_\pm^2 + a^2 \cos^2\theta)^{1/2}}{(r_\pm^2 + a^2)\sin\theta} \right]$$

and (6.1a)

$$\tau_0 = \frac{-ia \sin\theta}{2\sqrt{2}\,(r_\pm^2 + a^2)(r_\pm^2 + a^2 \cos^2\theta)^{3/2}}$$

$$\times [2r_\pm(r_\pm^2 + a^2) + (r_\pm - r_\mp)(r_\pm^2 + a^2 \cos^2\theta)]. \quad (6.1b)$$

With the characteristic data (6.1) determined it is possible to address the question of the existence of trapped surfaces in the Kerr spacetime near the bifurcate Killing horizons $H_\pm$. Since S_0 is compact, trapped surfaces exist as described in (4.2) if and only if

$$\tfrac{1}{2}K - \tau_0 \bar{\tau}_0 > 0 \quad \text{everywhere on } S_0, \quad (6.2)$$

where K is the Gaussian curvature of S_0.

To examine this inequality, K must first be calculated. The metric of S_0,

$$-ds_0{}^2 = (r_\pm^2 + a^2 \cos^2\theta)d\theta^2 + \frac{(r_\pm^2 + a^2)^2 \sin^2\theta}{(r_\pm^2 + a^2 \cos^2\theta)}\, d\phi^2,$$

can be expressed as

$$-ds_0{}^2 = \theta_1 \otimes \theta_1 + \theta_2 \otimes \theta_2,$$

where

$$\theta_1 = (r_\pm^2 + a^2 \cos^2\theta)^{1/2}\, d\theta$$

and

$$\theta_2 = [(r_\pm^2 + a^2)\sin\theta/(r_\pm^2 + a^2 \cos^2\theta)^{1/2}]\, d\phi.$$

The first structural equations[28]

$$d\theta_1 = \omega_{12} \wedge \theta_2 \quad \text{and} \quad d\theta_2 = \omega_{21} \wedge \theta_1,$$

where d is the exterior derivative and $\wedge$ indicates exterior multiplication, yield the connection form

$$\omega_{12} = [(r_\pm^2 + a^2)^2 \cos\theta/(r_\pm^2 + a^2 \cos^2\theta)^2]\, d\phi$$

The second structural equation[28]

$$d\omega_{12} = -K\theta_1 \wedge \theta_2$$

then yields the Gaussian curvature

$$K = (r_\pm^2 + a^2)(r_\pm^2 - 3a^2 \cos^2\theta)/(r_\pm^2 + a^2 \cos^2\theta)^3, \quad (6.3)$$

in agreement with the work of Smarr.[35]

In order that there is any possibility of satisfying (6.2), K must be strictly positive on S_0. From (6.3), $K > 0$ on S_0 if and only if

$$r_\pm^2 - 3a^2 > 0,$$

where the "+" or "−" sign is used when S_0 is the fixed point two-surface of the bifurcate Killing borizon H_+ or H_-, respectively. In terms of $\xi = |a|/m$ these conditions become

$$[1 \pm (1 - \xi^2)^{1/2}]^2 - 3\xi^2 > 0.$$

An examination of the functions

$$f_\pm(\xi) = [1 \pm (1 - \xi^2)^{1/2}]^2 - 3\xi^2$$

shows that $f_\pm(\xi) > 0$ for $0 \leq \xi < \sqrt{3}/2$ and $f_-(\xi) < 0$ for all ξ. Therefore, $K > 0$ for S_0 on H_+ if and only if $0 \leq \xi < \sqrt{3}/2$ and $K \leq 0$ for S_0 on H_-. With these results and (4.2) it has been established that:

In the region $\{(u, r, \theta, \phi)\}$ of the Kerr spacetime containing either the bifurcate nondiverging null hypersurface H_+ or H_-,

(1) if $a^2/m^2 \geq 3/4$, then for S_0 or any two-surface $S(u, r)$ contained in the $r > 0$ or $u > 0$ branches of H_+ there exists a neighborhood of S_0 or $S(u, r)$, respectively, that contains no trapped surfaces in the region common to it and the future of H_+, and

(6.4)

(2) for S_0 or any two-surface $S(u, r)$ contained in the $r > 0$ or $u > 0$ branches of H_- there exists a neighborhood of S_0 or $S(u, r)$, respectively, that contains no trapped surfaces in the region common to it and the future of H_-.

To obtain a positive result concerning the existence of trapped surfaces to the future of H_+, the entire inequality (6.2) must be considered. From (6.1b) and (6.3), this inequality becomes

$$\frac{1}{8(r_+^2 + a^2)^2(r_+^2 + a^2 \cos^2\theta)^3} \{4(r_+{}^2 + a^2)^3(r_+{}^2 - 3a^2 \cos^2\theta)$$

$$- a^2[2r_+(r_+{}^2 + a^2) + (r_+ - r_-)(r_+{}^2 + a^2 \cos^2\theta)]^2 \sin^2\theta\} > 0.$$

Clearly, this holds if and only if the quantity in braces is strictly positive. In terms of ξ and after some simplification this quantity can be rewritten as

$$16m^2 F(\xi, \theta),$$

where

$$F(\xi, \theta) = 2[1 + (1 - \xi^2)^{1/2}]^3 \{2[1 + (1 - \xi^2)^{1/2}] - \xi^2(1 + 3\cos^2\theta)\}$$
$$- \xi^2[(3 - 2\xi^2) + (1 - \xi^2)^{1/2}(3 - \tfrac{1}{2}\xi^2 \sin^2\theta)]^2 \sin^2\theta.$$

Thus the problem of determining the criterion for the existence of trapped surfaces to the future of H_+ as described by (4.2) becomes a problem of determining the value of ξ, ξ_0, such that $F(\xi, \theta) > 0$ for all ξ, $0 \leq \xi < \xi_0$, and for all θ, $0 \leq \theta \leq \pi$. That ξ_0 falls between 0 and $\sqrt{3}/2$ follows from $F(0, \theta) > 0$ for all θ, $0 \leq \theta \leq \pi$, and $F(\sqrt{3}/2, 0) = 0$.

Even though $F(\xi, \theta)$ is rather cumbersome, it is possible to make an exact determination of ξ_0. A calculation of $\partial F(\xi, \theta)/\partial\theta$ shows that $\partial F(\xi, \theta)/\partial\theta$ is zero when θ is 0, $\pi/2$, π. When $\partial^2 F(\xi, \theta)/\partial\theta^2$ is evaluated at $\theta = 0$, $\pi/2$, π, it is found that

$$\partial^2 F(\xi, 0)/\partial\theta^2 = \partial^2 F(\xi, \pi)/\partial\theta^2 > 0$$

and

$$\partial^2 F(\xi, \pi/2)/\partial\theta^2 < 0$$

for all ξ, $0 < \xi < 1$. Therefore, for fixed ξ, $0 < \xi < 1$, the minimum value of $F(\xi, \theta)$ on the interval $0 \leq \theta \leq \pi$ occurs at $\theta = 0$, since $F(\xi, \theta)$ is continuous on $0 \leq \theta \leq \pi$

and $F(\xi, 0) = F(\xi, \pi)$. This result implies that $F(\xi, \theta) > 0$ for all θ, $0 \leq \theta \leq \pi$, if and only if $F(\xi, 0) > 0$. The expression for $F(\xi, \theta)$ implies that $F(\xi, 0) > 0$ if an only if $f_+(\xi) > 0$. Therefore, since $f_+(\xi) > 0$ for all ξ, $0 \leq \xi < \sqrt{3}/2$, ξ_0 must be $\sqrt{3}/2$.

With ξ_0 determined, it follows from (4.2) and (6.4) that:

In the region $\{u, r, \theta, \phi, \}$ of the Kerr spacetime containing the bifurcate nondiverging null hypersurface H_+, trapped surfaces exist to the future of H_+ in every neighborhood of S_0 and all two-surfaces $S(u, r)$ contained in the $r > 0$ and $u > 0$ branches of H_+ if and only if $0 \leq a^2/m^2 < 3/4$. (6.5)

7. SUMMARY AND DISCUSSION

In this work an investigation of an empty spacetime containing a bifurcate nondiverging null hypersurface was undertaken. Its principle objective was to determine conditions that are necessary and sufficient for the existence of trapped surfaces near this hypersurface to its future.

A formalism appropriate for this investigation was presented in Sec. 2.

This formalism was then used in Sec. 3 to determine all empty spacetimes containing a bifurcate nondiverging null hypersurface and their characteristic data. There it was shown that the metric variables, spin coefficients, and physical Weyl tensor components of one such spacetime are determined in the region $\{(u, r, x^m)\}$ containing the bifurcate nondiverging null hypersurface $ur = 0$ by specifying the arbitrary functions (3.5)—$P(x^m)$, which determines the induced metric $ds_0^2 = -P^{-2}\, dz \otimes d\bar{z}$ of the two-surface of bifurcation $S_0 = \{(u, r, x^m) : u = 0 = r\}$, and $\tau_0(x^m)$. Furthermore, the first few terms in the power series expansion of these quantities as well as the general form of these series in powers of u and r are displayed there in (3.6), (3.7), and (3.8).

In Sec. 4 conditions that are necessary and sufficient for the existence of trapped surfaces in this spacetime near the bifurcate nondiverging null hypersurface to its future were given. The main result of this section is (4.2), which states that, in the region $\{(u, r, x^m)\}$ of an empty spacetime containing a bifurcate nondiverging null hypersurface $ur = 0$, trapped surfaces exist to the future of $ur = 0$ in every neighborhood of S_0 and all two-surfaces $S(u, r)$ contained in the $r > 0$ and $u > 0$ branches of $ur = 0$ if and only if S_0 is compact and the characteristic data P and τ_0 satisfy (4.1b), $\frac{1}{2}K - \tau_0\bar{\tau}_0 > 0$ everywhere on S_0, where $K = \bar{\eth}\eth \ln P$ is the Gaussian curvature of S_0. Also in this section the characteristic data (3.5) were discussed. Of particular importance in this discussion is the evidence presented in support of the interpretation of τ_0 in terms of angular momentum.

The role of K in the criteria (4.1) for the existence of trapped surfaces suggests that the collapse of an object maintaining a cylindrical or torroidal shape during collapse will not result in the formation of trapped surfaces, since a cylinder has zero Gaussian curvature[36] and a torus has regions of negative Gaussian curvature.[36] Indeed, it has been shown that at least one cylindrical collapse model results in complete collapse to a singularity without the formation of trapped surfaces.[36] Furthermore, the role of τ_0 in these criteria and its relationship to angular momentum emphasizes the importance of angular momentum for determining whether or not trapped surfaces are formed during gravitational collapse.

In Sec. 5 it was shown that a pair of intersecting null hypersurfaces in an empty spacetime is a bifurcate nondiverging null hypersurface if and only if it is a bifurcate Killing horizon. This result establishes the existence of a bifurcate Killing horizon in an empty spacetime containing a bifurcate nondiverging null hypersurface. Moreover, it was shown in this section that, in the region $\{(u, r, x^m\}$ of such a spacetime with this hypersurface given by $ur = 0$, the Killing vector field having $ur = 0$ as its bifurcate Killing horizon is $K = a(u\partial/\partial u - r\partial/\partial r + 2u\mathcal{X}^m\partial/\partial x^m)$, where a is a real constant.

The previously known result that a Killing horizon is a nondiverging null hypersurface[37] formed the basis for the widely held belief that trapped surfaces can always be found in the region interior to a Killing horizon.[38] However, it follows from (4.2) that if (4.1b) is not satisfied, then trapped surfaces do not exist near a bifurcate Killing horizon to its future. Thus, this belief is not supported by the results of this investigation.

Since the Kerr spacetime, when $a^2 < m^2$, contains the bifurcate Killing horizons $H_\pm$, which by (5.1) are bifurcate nondiverging null hypersurfaces, this spacetime was examined in Sec. 6. After determining the characteristic data (6.1), two results pertaining to the existence of trapped surfaces were obtained. An examination of the Gaussian curvature (6.3) of the two-surfaces of bifurcation for $H_\pm$ yielded the first of these (6.4) which asserts that trapped surfaces do not exist near H_- to its future and trapped surfaces do not exist near H_+ to its future if $a^2/m^2 \geq 3/4$. A positive result pertaining to the existence of trapped surfaces was obtained by considering the inequality (4.1b). This result is (6.5) which asserts that trapped surfaces exist near H_+ to its future if and only if $0 \leq a^2/m^2 < 3/4$.

ACKNOWLEDGMENTS

It is a pleasure to thank the relativity group at the University of Pittsburgh, especially Dr. Alan Janis, Dr. Ezra Newman, Dr. John Porter, Dr. Alan Thompson, and Dr. Jeffrey Winicour, for their encouragement and helpful comments in this work.

[1] R. Penrose, Revista Nuovo Cimento **1**, 252 (1969).
[2] R. Penrose, Phys. Rev. Lett. **14**, 57 (1965).
[3] See, for example, S.W. Hawking and R. Penrose, Proc. Roy. Soc. (London) A **314**, 529 (1970)
[4] S.W. Hawking, Commun. Math. Phys. **25**, 152 (1972).
[5] D.W. Pajerski and E.T. Newman, J. Math. Phys. **12**, 1929 (1971).
[6] P.N. Demmie and A.I. Janis, J. Math. Phys. **14**, 793 (1973).

[7]K. Schwarzschild, Sitz. Preuss. Akad. Wiss. **189** (1916).
[8]Actually it is a bifurcate nondiverging null hypersurface. See, for example, Ref. 5 or Ref. 27.
[9]R.H. Boyer, Proc. Roy. Soc. A **311**, 245 (1969).
[10]R.P. Kerr, Phys.Rev. Lett. **11**, 237 (1963).
[11]E.T. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).
[12]Greek indices and lower-case Latin indices a through d range and sum over the values (0, 1, 2, 3) and the remaining lower-case Latin indices sum over the values (2, 3) unless stated otherwise.
[13]Parentheses denote symmetrization: $A_{(\mu\nu)} = (1/2)(A_{\mu\nu}+A_{\nu\mu})$; brackets denote antisymmetrization: $A_{[\mu\nu]} = (1/2)(A_{\mu\nu}-A_{\nu\mu})$.
[14]The physical components $T_{\hat{a}_1\cdots\hat{a}_s\hat{b}_1\cdots\hat{b}_t}$ relative to a tetrad system $\{z^\mu_{\hat{a}}\partial/\partial x^\mu\}$ of a tensor with components $T_{\mu_1\cdots\mu_s\nu_1\cdots\nu_t}$ are given by $T_{\hat{a}_1\cdots\hat{a}_s\hat{b}_1\cdots\hat{b}_t} = T_{\mu_1\cdots\mu_s\nu_1\cdots\nu_t} z_{\hat{a}_1}{}^{\mu_1}\cdots z_{\hat{a}_s}{}^{\mu_s} z_{\hat{\nu}_1}{}^{\nu_1}\cdots z_{\hat{\nu}_t}{}^{\nu_t}$.
[15]A comma denotes partial differentiation, a semicolon denotes covariant differentiation, and Å denotes A, u.
[16]$\{x^\mu\}$ is the coordinate system, $\{(x^\mu)\}$ is the region consisting of all points of the manifold covered by $\{x^\mu\}$, and (x^μ) is a point in the region $\{(x^\mu)\}$.
[17]$\delta^\mu_\nu = 1$ if $\mu = \nu$ and $\delta^\nu_\mu = 0$ if $\mu \neq \nu$.
[18]These equations are obtained from the corresponding equations in Ref. 6 by specializing to the case of an empty spacetime.
[19]R. Sachs, Proc. Roy. Soc. (London) A **264**, 309 (1961).
[20]The subscript notations $(a \mid c \mid b)$ and $[a \mid c \mid b]$ denote symmetrization and antisymmetrization, respectively, with respect to a and b.
[21]See, for example, A. Trautman, *Lectures on General Relativity*, 1964 Brandeis Summer Institute of Theoretical Physics, edited by S. Deser and K. Ford (Prentice-Hall, Englewood Cliffs, N.J., 1965), p. 97
[22]These can also be found in C.D. Collison and D.G. French, J. Math. Phys. **8**, 701 (1967).
[23]Parentheses denote symmetrization: $A_{(\mu\nu\rho)} = (A_{\mu\nu\rho}+A_{\mu\rho\nu}+A_{\nu\mu\rho}+A_{\nu\rho\mu}+A_{\rho\mu\nu}+A_{\rho\nu\mu})/6$; brackets denote antisymmetrization: $A_{[\mu\nu\rho]} = (A_{\mu\nu\rho}-A_{\mu\rho\nu}+A_{\nu\rho\mu}-A_{\nu\mu\rho}+A_{\rho\mu\nu}-A_{\rho\nu\mu})/6$; and the subscript notation $(a \mid d \mid bc)$ denotes symmetrization with respect to a, b, and c.
[24]E.T. Newman and R. Penrose, J. Math. Phys. **7**, 863 (1966).
[25]That is, the bifurcate nondiverging null hypersurface $ur = 0$ is the set of all points (u, r, x^m) where $ur = 0$.
[26]The appropriate results in Ref. 6 are, of course, specialized to the case of an empty spacetime.
[27]P.N. Demmie, Ph.D. Thesis (University of Pittsburgh, 1971).
[28]See, for example, G. Birkhoff and S. MacLane, *A Survey of Modern Algebra* (Macmillan, New York, 1953), p. 11.
[29]See, for example, B. O'Neill, *Elementary Differential Geometry* (Academic, New York, 1966), p. 310.
[30]Compact two-surfaces with positive Gaussian curvature are called ovaloids. Such surfaces are discussed in D. Langwitz, *Differential and Riemannian Geometry* (Academic, New York, 1965), p. 205.
[31]B. O'Neill, Ref. 29, p. 385.
[32]See, for example, J.L. Synge, *Relativity: The General Theory* (North-Holland, Amsterdam, 1960), Chap. II, Sec. 8.
[33]R.H. Boyer and R.W. Lindquist, J. Math. Phys. **8**, 265 (1967).
[34]This metric follows from that given in E. Newman *et al.*, J. Math. Phys. **6**, 918 (1965), by setting $e = 0$.
[35]L. Smarr, Phys. Rev. D **7**, 289 (1973).
[36]B. O'Neill, Ref. 29, p. 231.
[37]B. Carter, J. Math. Phys. **10**, 70 (1969).
[38]K.S. Thorne, "Nonspherical Gravitational Collapse—A Short Review," (preprint, 1971).

Variational principles, variational identities, and supervariational principles for wavefunctions

E. Gerjuoy*

Physics Department, University of Pittsburgh, Pittsburgh, Pennsylvania 15213

A. R. P. Rau

Department of Physics and Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803

L. Rosenberg† and Larry Spruch‡

Physics Department, New York University, New York, New York 10003
(Received 18 September 1974)

We develop variational principles and variational identities for bound state and continuum wavefunctions in a general context, paying particular attention to the proper choice of defining equations and boundary conditions which will lead to unique and unambiguous wavefunctions even when these functions are complex. Any functional, such as a matrix element, calculated with such a variationally determined wavefunction, will also be accurate to second order in the error of the starting choice. This provides, therefore, an alternative procedure for getting variational estimates of matrix elements to the one that already exists in the literature and we establish the equivalence of the two. Of even more interest is the possibility which now seems open of going beyond the variational principle and generating "supervariational" estimates of wavefunctions and matrix elements which are good to better than second order. We also give a simple prescription for the construction of variational identities for wavefunctions, that is, identities which lead readily to variational principles and, more significantly, might well serve as a starting point for the development of variational bounds.

1. INTRODUCTION

In a series of papers,[1–5] we have addressed ourselves to the general construction of variational estimates of functionals $F(\phi)$, where the ϕ are unknown functions whose defining equations are too complicated to be solved for exactly. Typically, in the quantum-mechanical problem,[6] $F(\phi)$ may be a matrix element of some operator between wavefunctions ϕ which are defined through the appropriate Schrödinger equation and boundary conditions. We showed in Ref. 2 that there is a simple and direct method for writing down a variational expression $F_v \equiv \langle F(\phi_t)\rangle_v$ in terms of trial solutions ϕ_t, the expression being guaranteed not to have any errors of first order in $\phi_t - \phi (\equiv \delta\phi)$. The same procedure has been noted in a limited way in many places[7] in the literature, but usually with unnecessarily restrictive conditions and without sufficient regard to possible difficulties associated with uniquely specifying the functions involved. In Ref. 1 we gave a general prescription for formulating the corresponding "variational identities," which should be useful since they might well serve as the starting point for the development of variational bounds.

These general procedures for constructing variational principles and identities turn out, however, to be of practical significance only when a prescription is given for obtaining trial approximations to the various functions that are involved in the variational expression. A basic shortcoming, which has held up wider applications of such general variational principles, has been that no such procedure was known which would work in general and would be ambiguity-free and singularity-free. This has now been overcome, at least formally, for bound states in Ref. 3 and for continuum states in Refs. 4 and 5, where extremum principles that provide such a procedure were given; some preliminary calculations strongly indicate that the approach is a practical one. In the light of this development, it seems appropriate to take a fresh look at certain problems; some formal developments of quite some time ago seem to have been aborted without having seen practical application because of the difficulty of actually using them in any meaningful way for nontrivial problems. One such problem, which is the subject of this paper, is the development of variational principles for wavefunctions[8–10] and the subsequent use of these variationally good wavefunctions in the calculations of various properties of the system of interest.

In the general formulation of a variational principle (VP) for a functional $F(\phi)$, the feature to be emphasized is that the variational expression is specific to the functional of interest, depending both on the wavefunctions involved and on the operators of interest. When, for example, the desired functional is the ground-state energy of a discrete spectrum, the variational expression reduces to the familiar Rayleigh–Ritz principle; if some other property of the system is of interest, the expression is correspondingly different (and usually more complicated). An alternative procedure for getting a variational estimate of $F(\phi)$ would be to develop a "variationally good" wavefunction ϕ_v, which, *as a wavefunction*, would have no first order error $\delta\phi$ in the starting ϕ_t. Having done this once and for all, it is evident that any $F(\phi_v)$ evaluated with such a function will automatically be a variational estimate of $F(\phi)$, regardless of the specific nature of the functional. This then

 Copyright © 1975 American Institute of Physics

is the motivation of the present paper. The basic procedure for the construction of a VP for ϕ is no different from the general method we have followed for obtaining a VP for $F(\phi)$ and needs only the identification of the functional with the wavefunction itself, that is, $F(\phi)=\phi$, since there is no restriction in the general formulation[2] on the nature of the functional. In fact, in a previous paper,[1] we derived a VP for $g^\dagger\phi$ with the dagger representing some appropriate inner product with a known function g when ϕ is defined by $M\phi=\omega$, where M is some known differential operator and ω a given function. One might, therefore, expect that with the identification $g=\delta(\mathbf{r}-\mathbf{r}')$ for which $g^\dagger\phi=\int d\mathbf{r}'\,\delta(\mathbf{r}-\mathbf{r}')\phi(\mathbf{r}')=\phi(\mathbf{r})$ such a VP would follow immediately.[11] This is true, but there are complications associated with defining uniquely the functions ϕ when one considers bound state and continuum functions which are not real; there are still further complications when one looks at systems that are not invariant under both time reversal and rotation. These complications, which did not arise in our earlier work,[12] and were not considered in the existing literature[8–10] on VP's for wavefunctions, have never been spelled out adequately and have, therefore, remained unappreciated.[13] A main emphasis of this paper is a thorough consideration of these problems and a unified presentation of VP's for wavefunctions of all kinds, where, furthermore, the possibility of any near singularities in the Lagrange multipliers that appear in the VP's is to be unambigously avoided.

The arrangement of the paper is as follows. In Sec. 2, we first develop a VP for a bound state wavefunction when it is purely real. This is the simplest case and is free of troubles arising from the specification of the phase of complex wave functions. At the end of Sec. 2, we indicate how, even with complex ϕ, hardly any change is necessary so long as the system is invariant under rotation and time reversal. In Sec. 3, we address ourselves to the general problem at its most complicated level when we do not assume such invariance and show how the problem of the phase is to be handled in setting up the VP. This is not at all an academic consideration since such complications have to be faced for systems which have external electric and/or magnetic fields present. In Sec. 4, we briefly record, for the sake of completeness, a VP for continuum wavefunctions which was essentially contained in a previous paper.[1] Section 5 establishes contact between the two procedures for generating a variational estimate of $F(\phi)$, that is, the direct method of previous papers, where we construct the variational expression $F_v=\langle F(\phi_t)\rangle_v$, and the evaluation of $F(\phi_v)$ with the variational wavefunctions derived in the earlier sections of the present paper. We demonstrate that, as expected, they can only differ in second-order quantities. Throughout, in each section, we present along with the VP's the corresponding "variational identities" which are expressions which give the *exact* $F(\phi)$ (not merely a variational estimate) in terms of ϕ_t, the original unknown functions ϕ themselves, and certain Lagrange multipliers. The merit of such identities as compared to the corresponding VP's is that they exhibit explicitly the second order "error" terms and might serve as an excellent starting point for establishing variational bounds on the quantities of interest.[14–16] We had presented earlier[1] a simple, general procedure for writing down such identities, and we point out here an even simpler one which consists essentially of realizing that the "half-way point" in the general procedure for constructing VP's *is* the identity. In Sec. 6, we indicate briefly the development of "supervariational principles," made possible by the results of this paper for ϕ_v; use of these variationally good wavefunctions as trial functions in the expression $\langle F(\phi_t)\rangle_v$, that is, the use of $\langle F(\phi_v)\rangle_v$, leads to estimates which are good not just to second order but to third or fourth order—as opposed to VP's, they will contain no second order errors. [Coupled with an iterative procedure of starting from a ϕ_t and deriving an improved estimate ϕ_v from it the way may be open for generating super-VP's for wavefunctions, and, therefore, for any $F(\phi)$, which are good to still higher orders; there is no point elaborating on this, however, until the practicability of super-VP's good to third or fourth order has been tested.] The entire development of these sections hinges, as regards its practical utility, on being able to obtain trial approximations to auxiliary functions that are introduced during the construction of the VP's. In the final section, Sec. 7, we discuss briefly how, using methods introduced in Refs. 3–5, possible difficulties associated with near singularities in the defining equations for these functions can be removed. Subsidiary minimum principles are then available to aid in the construction of these trial functions.

2. VARIATIONAL PRINCIPLE FOR A REAL BOUND STATE WAVEFUNCTION

The construction of a VP for ϕ, where ϕ is a real bound state wavefunction of energy eigenvalue E associated with a real Hamiltonian H follows immediately from the general procedure of our earlier papers.[1,2] The quantity ϕ is uniquely defined by

$$(H-E)\phi=0, \tag{2.1a}$$

$$\phi^\dagger\phi=1. \tag{2.1b}$$

In our notation, taken from Ref. 2, the dagger is understood to imply integration over continuum variables and, where appropriate, summation over discrete variables. For simplicity only, we assume that the state is nondegenerate. We write then, in terms of some real trial approximation ϕ_t, with $\mathbf{r}$ or $\mathbf{r}'$ the totality of coordinates,

$$\phi_v(\mathbf{r})=\phi_t(\mathbf{r})+\Lambda_{at}^\dagger(\mathbf{r},\mathbf{r}')\{[H(\mathbf{r}')-E_t]\phi_t(\mathbf{r}')\}+L_{nt}(\mathbf{r})(\phi_t^\dagger\phi_t-1), \tag{2.2}$$

where we have incorporated the defining equations as constraints through the use of trial Lagrange multipliers Λ_{at} and L_{nt} and where E_t, a trial estimate of the eigenvalue, would most usually be $\phi_t^\dagger H\phi_t$, the natural choice for it; E_t would then be a variational estimate of E. We use the subscript n on L_{nt} since L_{nt} arose in connection with the normalization condition, while the subscript a on Λ_{at} facilitates comparison with a more general case in Sec. 3. The form of the multipliers Λ and L is immediately evident from the nature of the constraints and the fact that we are looking for a VP for the function ϕ at any arbitrary point $\mathbf{r}$. Thus Eq. (2.1a) is multiplied by a two-variable function Λ_{at} with integration over a

dummy variable $\mathbf{r}'$ whereas Eq. (2.1b) is multiplied by a one-variable function $L_n(\mathbf{r})$. Λ_{at} and L_{at} in Eq. (2.2) are trial approximations to the functions Λ_a and L_n which are defined by demanding that Eq. (2.2) contain no terms linear in the first-order errors:

$$\delta\phi \equiv \phi_t - \phi, \quad \delta\Lambda_a \equiv \Lambda_{at} - \Lambda_a, \quad \delta L_n \equiv L_{nt} - L_n. \tag{2.3}$$

There are no terms linear in $\delta\Lambda_a$ and δL_n because of Eqs. (2.1) and we only have to examine the term in $\delta\phi$. We find

$$\delta\phi(\mathbf{r}) + \Lambda_a^\dagger(\mathbf{r},\mathbf{r}')\{[H(\mathbf{r}') - E]\delta\phi(\mathbf{r}')\} + L_n(\mathbf{r})(\phi^\dagger\delta\phi + \delta\phi^\dagger\phi) = 0. \tag{2.4}$$

Before extracting the defining equations for Λ_a and L_n from this, we point out that Eq. (2.4) is the variational identity associated with the VP, Eq. (2.2). Thus, using Eqs. (2.1) and (2.3), we can rewrite Eq. (2.4) as

$$\phi(\mathbf{r}) = \phi_t(\mathbf{r}) + \Lambda_a^\dagger(\mathbf{r},\mathbf{r}')\{[H(\mathbf{r}') - E]\phi_t(\mathbf{r}')\} + L_n(\mathbf{r})(\phi^\dagger\phi_t + \phi_t^\dagger\phi - 2). \tag{2.5}$$

[The passage to the VP can be seen as follows. $(H-E)\phi_t$ is of first-order and $(\phi^\dagger\phi_t + \phi_t^\dagger\phi - 2)$ differs from $(\phi_t^\dagger\phi_t - 1)$ by a term of second order, $(\delta\phi^\dagger\delta\phi)$. Therefore, the right-hand sides of Eqs. (2.2) and (2.5) differ only in terms of second order. We will return to this identity shortly, but we note now that the result that the identity coincides with the equation obtained by setting the terms in $\delta\phi$ equal to zero is valid not only for the case under consideration but, generally, working with linear and nonlinear equations alike. This is a completely general prescription for obtaining identities and is superior to the one we gave in an earlier paper.[1] That Eq. (2.5) is indeed an identity follows by construction. Writing $\phi_t = \phi + \delta\phi$, the ϕ terms vanish and the $\delta\phi$ terms are as in Eq. (2.4) and will vanish because Λ_a and L_n will be chosen to make them vanish. An explicit check that Eq. (2.5) is an identity will be presented shortly.]

To extract the defining equations for the Lagrange multipliers, it is necessary[2] to rewrite Eq. (2.4) in the form $K(\mathbf{r},\mathbf{r}')^\dagger\delta\phi(\mathbf{r}') = 0$, where K does not contain differential operations in $\mathbf{r}'$, so that $K(\mathbf{r},\mathbf{r}') = 0$ follows. We begin by writing

$$\Lambda_a^\dagger(\mathbf{r},\mathbf{r}')\{[H(\mathbf{r}') - E]\delta\phi(\mathbf{r}')\} = \{[H(\mathbf{r}') - E]\Lambda_a(\mathbf{r},\mathbf{r}')\}^\dagger\delta\phi(\mathbf{r}'), \tag{2.6}$$

which follows from the hermiticity of H and the decay property of $\delta\phi$. Further, since we are for the moment restricting ourselves to real functions, we have

$$\delta\phi^\dagger\phi = \phi^\dagger\delta\phi. \tag{2.7}$$

Writing the first term in Eq. (2.4) as $[\delta(\mathbf{r}-\mathbf{r}')]^\dagger\delta\phi(\mathbf{r}')$, we can set the coefficient of $\delta\phi(\mathbf{r}')$ equal to zero and obtain (we write down the Hermitian adjoint of the coefficient)

$$\delta(\mathbf{r}-\mathbf{r}') + [H(\mathbf{r}') - E]\Lambda_a(\mathbf{r},\mathbf{r}') + 2\phi(\mathbf{r}')L_n^\dagger(\mathbf{r}) = 0. \tag{2.8}$$

Multiplication from the left by $\phi^\dagger(\mathbf{r}')$ and use of Eqs. (2.1) determines L_n to be

$$L_n(\mathbf{r}) = -\tfrac{1}{2}\phi(\mathbf{r}). \tag{2.9}$$

Insertion of Eq. (2.9) in Eq. (2.8) gives the defining equation for the function Λ_a

$$[H(\mathbf{r}') - E]\Lambda_a(\mathbf{r},\mathbf{r}') = -\delta(\mathbf{r}-\mathbf{r}') + \phi(\mathbf{r}')\phi^\dagger(\mathbf{r}) \equiv -Q, \tag{2.10}$$

where $Q \equiv 1 - P$ and where $P \equiv \phi\phi^\dagger$ is the projection operator for the eigenstate of interest, Λ_a is, therefore, a "Green's function in the generalized sense."[17] Since $\phi^\dagger Q = 0$, the function Λ_a is perfectly well defined when appropriate boundary conditions are imposed even though E is an eigenvalue of H. In the usual bilinear representation, in terms of the complete set of eigenfunctions $|n\rangle$ of H, this means that $\Lambda_a = \sum_n'|n\rangle\langle n|(E-E_n)^{-1}$, where the prime denotes the omission of the state with eigenvalue E from the summation. We also note that the full complexity of the Green's function is not involved in the variational expression, Eq. (2.2), since only its projection on $(H-E_t)\phi_t$ is of interest.

Equations (2.9) and (2.10) fully define the Lagrange functions. Trial solutions of these equations when fed into Eq. (2.2) give the desired variational estimate of ϕ; the procedure for obtaining trial solutions is the subject of Sec. 7.

We now return to Eq. (25) to show that it really is an identity. One can readily check that the right-hand side of Eq. (2.5) is annihilated by $(H-E)$, but this does not check the normalization. Using Eq. (2.9), we have

$$\phi(\mathbf{r})(\phi^\dagger\phi_t) = \phi_t(\mathbf{r}) + \Lambda_a^\dagger(\mathbf{r},\mathbf{r}')\{[H(\mathbf{r}') - E]\phi_t(\mathbf{r}')\}. \tag{2.11}$$

This form can be recognized from Eq. (2.10) to be merely the obvious identity $P\phi_t = (1-Q)\phi_t$. Such an identity has been noted previously.[18]

This derivation of the VP and the variational identity proceeded without a hitch; a crucial step that made this possible was Eq. (2.7) which used the reality of the functions involved. Were the functions not real, ϕ and $\phi^\dagger$ would have to be treated as linearly independent quantities as would, therefore, $\delta\phi$ and $\delta\phi^\dagger$ and it would not have been possible to consider Eq. (2.4) as the Hermitian inner product of $\delta\phi$ with some $K(\mathbf{r},\mathbf{r}')$. $\delta\phi$ and $\delta\phi^\dagger$ would have had to be treated independently, and it would have been necessary to incorporate not just Eq. (2.1a) as a constraint but also its Hermitian adjoint as an independent constraint which would have made the analysis more complicated. We will face these complexities of the completely general VP in Sec. 3, but note here that even when ϕ is complex, the above simpler derivation remains essentially valid so long as the system has invariance under rotation and time reversal. This is because of a powerful theorem due to Wigner[19] for such a situation which states that $\phi(\mathbf{r})$ can then be expanded in terms of a known orthonormal basis set $\Omega_\alpha^{JM}(\hat{r})$ of complex functions which contain all the spin and angular momentum dependence, and whose phases can be so chosen that the coefficients of the expansion, $R_\alpha(r)$, which contain the radial behavior, can be chosen to be real. Thus, we have

$$\phi(\mathbf{r}) = \sum_\alpha R_\alpha(r)\Omega_\alpha^{JM}(\hat{r}), \tag{2.12}$$

as the unique structure of the solutions of Eq. (2.1). If we make an analogous choice for ϕ_t in terms of the same Ω's, with real $R_{\alpha t}(r)$, we have

$$\phi_t(\mathbf{r}) = \sum_\alpha R_{\alpha t}(r)\Omega_\alpha^{JM}(\hat{r}), \tag{2.13}$$

and Eq. (2.7) remains valid since

$$\phi^\dagger\delta\phi=\sum_\alpha R_\alpha^\dagger(r)\delta R_\alpha(r)=\sum_\alpha\delta R_\alpha^\dagger(r)R_\alpha(r)=\delta\phi^\dagger\phi. \quad (2.14)$$

Thus, the fact that the defining equations for ϕ involve, through the normalization condition, the linearly independent $\phi^\dagger$ causes no difficulty since this coupling in an equation such as (2.4) unravels because of Eq. (2.7). We need not, therefore, incorporate in the VP the constraint given by the adjoint of Eq. (2.1a). [Were we to incorporate this through a Λ_b as in what follows in Sec. 3, we would find that $(H-E)\Lambda_b=0$ and, therefore, consistently (to second order terms) we drop the term in Λ_b in the VP].

The representation (2.12) of ϕ, with a specific choice of the phases of the Ω_α^{JM}, represents a specific and standard[19] choice of the phases of ϕ. In the following section, we introduce a specification of the phase of ϕ which is applicable under any and all circumstances, a specification which via the introduction of an arbitrary function χ, introduces a constraint in a form which requires an additional Lagrange function. The procedure above, which is applicable for almost all cases of interest, effectively builds in the constraint in the choice of ϕ_t; no additional Lagrange function need be introduced. It may be useful to compare the above constraint with the normalization constraint. Either we can choose ϕ_t to be normalized or we can simply choose ϕ_t to be close to ϕ and use a Lagrange multiplier to build in the normalization condition; the latter is often algebraically advantageous because the variation of parameters in ϕ_t, for ϕ_t chosen to be normalized, can be cumbersome. With regard to the phase condition, however, we suspect that the above approach, when applicable, will be simpler to apply, since it merely calls for a standard form of ϕ_t in terms of a basis set such as that of spherical harmonics; the basis set is fixed, and as far as the phase condition is concerned, any parameters contained in the real functions $R_{\alpha t}(r)$ can be freely varied. It will, of course, be necessary to apply both these approaches to incorporating the phase condition before any definite conclusions concerning their relative merits can be drawn.

3. VARIATIONAL PRINCIPLE FOR A GENERAL BOUND STATE WAVEFUNCTION

In this section, we will develop a VP for ϕ, the wavefunction of a nondegenerate bound state of energy E of a Hermitian Hamiltonian, without making any further assumptions about the system; in particular, the system need not be invariant with respect to rotation or time reversal or the parity transformation. The Schrödinger equation and the normalization condition in Eqs. (2.1) do not then serve to define uniquely the function ϕ but leave the phase ambiguous. To have a well-defined problem, therefore, the phase of ϕ must be fixed; one possible way of doing this is to fix the phase of ϕ with respect to some known function χ, say by demanding that $\phi^\dagger\chi$ be pure real or pure imaginary (a more general choice is possible but unnecessary), that is,

$$\phi^\dagger\chi\mp\chi^\dagger\phi=0. \quad (3.1)$$

χ may be an arbitrary function apart from the requirement that $\phi^\dagger\chi$ should not be trivially fixed at some value because of some symmetry of the system, since we could not use such a χ to fix the phase. Thus, for instance, choices for χ of functions proportional to ϕ or orthogonal to ϕ in spin or some other quantum number would not be satisfactory since normalization or orthogonality, respectively, will trivially fix the value of $\phi^\dagger\chi$ in such cases, independent of the phase of ϕ. With this in mind, Eq. (3.1), together with Eq. (2.1) and their Hermitian adjoints, define completely ϕ and its linearly independent adjoint $\phi^\dagger$. In seeking a VP for ϕ we start again by incorporating all these constraints through Lagrange multipliers and write[2]

$$\begin{aligned}\phi_v(\mathbf{r})&=\phi_t(\mathbf{r})+\Lambda_{at}^\dagger(\mathbf{r},\mathbf{r}')\{[H(\mathbf{r}')-E_t]\phi_t(\mathbf{r}')\}\\&+\{[H(\mathbf{r}')-E_t)\phi_t(\mathbf{r}')\}^\dagger\Lambda_{bt}(\mathbf{r},\mathbf{r}')\\&+L_{nt}(\mathbf{r})(\phi_t^\dagger\phi_t-1)+L_{pt}(\mathbf{r})\phi_t^\dagger\chi\mp\chi^\dagger\phi_t).\end{aligned} \quad (3.2)$$

The constraint (2.1a) necessitates a two-variable Lagrange multiplier, while the normalization and phase constraints require, respectively, L_{nt} and L_{pt}, single variable functions; we use the subscript p on L_{pt} since L_{pt} arose in connection with the phase condition. $\mathbf{r}$ collectively represents continuum and discrete (such as spin) coordinates, and when the latter are present so that ϕ is a column vector in the space of these coordinates (with elements that are functions of the space coordinates), $\phi^\dagger$ will correspondingly be a row vector and L_{nt} and L_{pt} column vectors. Λ_{at} will be a square matrix in the same space whereas, as can already be anticipated, Λ_{bt} will have the rather unusual structure of two column vectors side by side (no multiplication involved). All this will follow quite naturally from the formalism and we are merely anticipating some of the structure at this time simply by inspection of Eq. (3.2). Varying Eq. (3.2) and using a number of relations of the form of Eq. (2.3), we obtain

$$\begin{aligned}&\delta\phi(\mathbf{r})\\&+\Lambda_a^\dagger(\mathbf{r},\mathbf{r}')\{[H(\mathbf{r}')-E]\delta\phi(\mathbf{r}')\}\\&+\{[H(\mathbf{r}')-E]\delta\phi(\mathbf{r}')\}^\dagger\Lambda_b(\mathbf{r},\mathbf{r}')\\&+L_n(\mathbf{r})(\phi^\dagger\delta\phi+\delta\phi^\dagger\phi)+L_p(\mathbf{r})[\delta\phi^\dagger\chi\mp\chi^\dagger\delta\phi)=0.\end{aligned} \quad (3.3)$$

Using

$$\begin{aligned}&\Lambda_a^\dagger[(H-E)\delta\phi]=[(H-E)\Lambda_a]^\dagger\delta\phi,\\&[(H-E)\delta\phi]^\dagger\Lambda_b=\delta\phi^\dagger[(H-E)\Lambda_b],\end{aligned} \quad (3.4)$$

and setting the (adjoint of the) coefficient of $\delta\phi(\mathbf{r}')$ equal to zero gives

$$\delta(\mathbf{r}-\mathbf{r}')+[H(\mathbf{r}')-E]\Lambda_a(\mathbf{r},\mathbf{r}')+\phi(\mathbf{r}')L_n^\dagger(\mathbf{r})\mp\chi(\mathbf{r}')L_p^\dagger(\mathbf{r})=0, \quad (3.5)$$

whereas the coefficient of $\delta\phi^\dagger(\mathbf{r}')$ gives

$$[H(\mathbf{r}')-E]\Lambda_b(\mathbf{r},\mathbf{r}')+L_n(\mathbf{r})\phi(\mathbf{r}')+L_p(\mathbf{r})\chi(\mathbf{r}')=0. \quad (3.6)$$

L_n and L_p are determined by multiplying the above equations by $\phi^\dagger(\mathbf{r}')$ from the left and carrying out the integrations over $\mathbf{r}'$. We have, using Eq. (3.1),

$$\phi(\mathbf{r})+L_n(\mathbf{r})-(\phi^\dagger\chi)L_p(\mathbf{r})=0, \quad (3.7a)$$

$$L_n(\mathbf{r})+(\phi^\dagger\chi)L_p(\mathbf{r})=0. \quad (3.7b)$$

Adding gives

$$L_n=-\tfrac{1}{2}\phi, \quad (3.8a)$$

$$L_p = \tfrac{1}{2}\phi/(\phi^\dagger\chi). \tag{3.8b}$$

With these solutions we return to examine Eqs. (3.5) and (3.6) for Λ_a and Λ_b. Taking the latter first, we have

$$[H(\mathbf{r}') - E]\Lambda_b(\mathbf{r},\mathbf{r}') - \tfrac{1}{2}\phi(\mathbf{r})\phi(\mathbf{r}') + \tfrac{1}{2}\phi(\mathbf{r})\chi(\mathbf{r}')/(\phi^\dagger\chi) = 0. \tag{3.9}$$

The solution of Eq. (3.9) is clearly given by

$$\Lambda_b(\mathbf{r},\mathbf{r}') = \tfrac{1}{2}\phi(\mathbf{r})L_b(\mathbf{r}'), \tag{3.10}$$

where

$$(H - E)L_b = \phi - \chi/(\phi^\dagger\chi). \tag{3.11}$$

Note that, as anticipated, Λ_b is the product of two column vectors with elements which are functions.

Equation (3.5) for Λ_a becomes, by using Eq. (3.1),

$$[H(\mathbf{r}') - E]\Lambda_a(\mathbf{r},\mathbf{r}') = -\delta(\mathbf{r}-\mathbf{r}') + \tfrac{1}{2}\phi(\mathbf{r}')\phi^\dagger(\mathbf{r}) + \tfrac{1}{2}\chi(\mathbf{r}')\phi^\dagger(\mathbf{r})/(\phi^\dagger\chi). \tag{3.12}$$

Contrast this with Eq. (2.10) which defined the similar Green's function in the pure real case. [The solution of Eq. (3.12) can be expressed as the sum of solutions,

$$\Lambda_a = \Lambda_{a1} + \Lambda_{a2},$$

where

$$(H - E)\Lambda_{a1} = -1 + \phi\phi^\dagger \equiv -Q, \tag{3.13a}$$

$$(H - E)\Lambda_{a2} = -\tfrac{1}{2}\phi\phi^\dagger + \tfrac{1}{2}\chi\phi^\dagger/(\phi^\dagger\chi), \tag{3.13b}$$

provided both Λ_{a1} and Λ_{a2} can be made to satisfy the usual boundary conditions. Equation (3.13a) is exactly similar to Eq. (2.10) while Eq. (3.13b) is of the form of Eq. (3.9), and allows the factorization of Λ_{a2}, one of the factors being $\phi^\dagger(r)$.]

L_b and Λ_a are solutions of inhomogeneous equations, (3.11) and (3.12), respectively, and are uniquely defined only to within a multiple of the solution ϕ of the associated homogeneous equation. We can choose to make L_b and Λ_a unique by demanding that

$$L_b^\dagger\phi = 0, \quad \Lambda_a^\dagger\phi = 0.$$

(In particular, the phases of L_b and of Λ_a are now uniquely determined by the source terms in their defining equations.) It is then natural to choose L_{bt} and Λ_{at} such that

$$L_{bt}{}^\dagger\phi_t = 0, \quad \Lambda_{at}{}^\dagger\phi_t = 0,$$

so that we can have $L_{bt} \to L_b$ and $\Lambda_{at} \to \Lambda_a$ as $\phi_t \to \phi$. Note that for the choice $E_t = \phi_t^\dagger H\phi_t$, any $\phi_t(r')$ components of L_{bt} or of Λ_{at} would have made no contribution in the VP of Eq. (3.2). The VP in Eq. (3.2) is now complete with all the functions properly defined. Trial solutions to them (see Sec. 7) yield through Eq. (3.2) the desired ϕ_v.

The variational identity has also, just as in Sec. 2, been derived along the way—it is just Eq. (3.3), arrived at by setting the first-order terms equal to zero. Rewriting it in terms of ϕ_t and ϕ and using Eqs. (2.1) and (3.1), we have

$$\phi = \phi_t + \Lambda_a^\dagger[(H - E)\phi_t] + [(H - E)\phi_t]^\dagger\Lambda_b + L_n(\phi^\dagger\phi_t + \phi_t^\dagger\phi - 2) + L_p(\phi_t^\dagger\chi \mp \chi^\dagger\phi_t). \tag{3.14}$$

The connection to the corresponding VP in Eq. (3.2) is immediate and exactly analogous to the discussion in Sec. 2. That Eq. (3.14) is an identity follows, as in Sec. 2, by construction.

4. VARIATIONAL PRINCIPLE FOR CONTINUUM WAVEFUNCTIONS

The construction of a VP for continuum wavefunctions is simpler than for a bound-state function for two reasons. The first is that there is usually no complication of an ambiguous phase, the phase of the continuum ϕ being specified by the nature of the asymptotic boundary condition. The second reason is that the defining equations for ϕ do not contain $\phi^\dagger$ (as in the normalization condition for the bound state case) and, therefore, the equations obeyed by $\phi^\dagger$ need not be incorporated as constraints. A VP for continuum ϕ has already been given by us in a previous paper,[1] where we derived a VP for $g^\dagger\phi$ with g arbitrary. The choice $g = \delta(\mathbf{r} - \mathbf{r}')$ in Sec. V of that paper immediately yields ϕ_v. Merely for the sake of completeness we record here (in the present context) the results which were derived in that paper. For simplicity, we consider the scattering of a spinless particle by a spherically-symmetric potential $V(r)$.

If we look at the (real) component $\phi_l(r)$ of the lth partial wave which is defined by

$$(H_l - E)\phi_l(r) = 0, \tag{4.1a}$$

$$H_l \equiv \frac{\hbar^2}{2\mu}\left(-\frac{d^2}{dr^2} + \frac{l(l+1)}{r^2}\right) + V(r), \tag{4.1b}$$

$$\phi_l(0) = 0, \quad \phi_l(r) \sim \sin(kr - \tfrac{1}{2}l\pi) + \tan\eta_l \cos(kr - \tfrac{1}{2}l\pi), \quad r \sim \infty, \tag{4.1c}$$

where η_l is the phase shift and we pick a trial ϕ_{lt} which satisfies the boundary conditions in (4.1c) with some trial phase shift η_{lt}, we have

$$\phi_{lv} = \phi_{lt} + \Lambda_t^\dagger[(H_l - E)\phi_{lt}]. \tag{4.2}$$

The requirement that $\delta\phi_l = 0$ gives the condition that the Green's function Λ is defined by

$$(H_l - E)\Lambda = -1, \tag{4.3a}$$

$$\Lambda(r,0) = 0, \quad \Lambda(r,r') \sim f(r)\cos(kr' - \tfrac{1}{2}l\pi), \quad r' \sim \infty, \tag{4.3b}$$

where the boundary conditions in Eq. (4.3b) are necessary if, now that we are dealing with continuum functions, the surface terms that arise in transferring the operator H_l from right to left are to vanish.

If we look at the full wave function $\phi(\mathbf{r})$ instead of a specific partial wave, the defining equations are

$$(H - E)\phi = 0, \tag{4.4a}$$

$$\phi(\mathbf{r}) \sim \exp(i\mathbf{k}\cdot\mathbf{r}) + f(\theta)\exp(ikr)/r, \tag{4.4b}$$

where $f(\theta)$ is the scattering amplitude. Once again the VP for ϕ is straightforward. We have

$$\phi_v = \phi_t + \Lambda_t^\dagger[(H - E)\phi_t], \tag{4.5}$$

where Λ now obeys

$$(H - E)\Lambda = -1, \ \Lambda(\mathbf{r},\mathbf{r}')_{\text{outgoing}}. \tag{4.6}$$

The identities are also immediate and follow from the

simple general procedure. In this case, since Eqs. (4.2) and (4.5) are linear in Λ_t and ϕ_t, it is clear that use of the exact Λ in place of Λ_t in Eqs. (4.2) or (4.5) yields the corresponding identity. The choices $\phi_{lt}=kr\,j_l(kr)$ or $\phi_t(\mathbf{r})=\exp(i\mathbf{k}\cdot\mathbf{r})$ in these identities give the usual integral equations for the wavefunctions.

There would be no difficulty in extending the above derivation to the development of a VP in the case of scattering of one system by another, taking into account the Pauli principle, elastic and inelastic scattering, etc. We will not consider the extension to nontime reversible systems. There is no matter of principle involved. It is simply that the extension could be very complicated. Independent of the question of VP's (and independent of the question of time reversibility) the specification of the boundary conditions associated with scattering in the presence of an external magnetic field can be rather difficult.

5. CONTACT BETWEEN THE TWO METHODS OF VARIATIONALLY ESTIMATING MATRIX ELEMENTS

As a remarked upon in the Introduction, when a VP is desired for $F(\phi)$ where $F(\phi)$ is a diagonal or off-diagonal matrix element (we consider bound state wavefunctions in this section), one can directly construct a VP, $\langle F(\phi_t)\rangle_v$, by the methods of our earlier papers[1,2] or, alternatively, one can evaluate $F(\phi_v)$ with variationally accurate ϕ_v from Secs. 2 and 3 of this paper. We will now establish or show how to establish the equivalence of these two procedures for a number of cases. There is no real need to prove the equivalence to second-order of two different VP's except as a check on their validity.

A. Bound state diagonal matrix elements

In Sec. III of Ref. 1, we derived a VP for $\phi^\dagger W\phi$, where W is a linear, self-adjoint operator (the extension to W not self-adjoint is trivial), which coincides with a VP given originally for this case by Schwartz.[20] The VP takes the form

$$\langle W\rangle_v=\phi_t^\dagger W\phi_t+L_t^\dagger[(H-E_t)\phi_t]+[(H-E_t)\phi_t]^\dagger L_t-\lambda_t(\phi_t^\dagger\phi_t-1), \tag{5.1}$$

where the Lagrange function L is defined by

$$(H-E)L=-W\phi+\lambda\phi, \tag{5.2a}$$

$$\lambda=\phi^\dagger W\phi,\quad L^\dagger\phi=0. \tag{5.2b}$$

[In Ref. 1, f was used in place of our (now) standard use of L for such Lagrange multipliers and the last term involving λ_t was not explicitly shown since ϕ_t was chosen to be normalized. For a more general choice such a Lagrange multiplier is present because of the normalization condition and the value of λ is as in Eq. (5.2b) as can be seen by premultiplying Eq. (5.2a) by $\phi^\dagger$.] In that derivation, we made no reference to the specification of the phase of ϕ, such a specification being irrelevant for the diagonal element $\phi^\dagger W\phi$. We will now demonstrate that the evaluation of the matrix element with ϕ_v from Eq. (3.2) makes complete contact with Eq. (5.1) and, as expected, the terms specifying the phase drop out of the picture. Writing Eq. (3.2) as

$$\phi_v=\phi_t+\phi_t^{(1)}, \tag{5.3}$$

with $\phi_t^{(1)}$ denoting collectively the four first-order terms in Eq. (3.2), we have

$$\phi_v^\dagger W\phi_v=\phi_t^\dagger W\phi_t+(\phi_t^{(1)})^\dagger W\phi_t+\phi_t^\dagger W(\phi_t^{(1)}), \tag{5.4}$$

where we have dropped terms of second order. Using the explicit form of $\phi_t^{(1)}$ from Eq. (3.2) and using Eqs. (3.8) and (3.10), we get after considerable algebra

$$\phi_v^\dagger W\phi_v=\phi_t^\dagger W\phi_t[1-(\phi_t^\dagger\phi_t-1)]+[(H-E_t)\phi_t]^\dagger L_{ct}+L_{ct}^\dagger[(H-E_t)\phi_t], \tag{5.5}$$

where we have defined

$$L_{ct}=\tfrac{1}{2}(\phi_t^\dagger W\phi_t)L_{bt}+\Lambda_{at}W\phi_t, \tag{5.6}$$

and dropped a second-order term proportional to $(\phi_t^\dagger\chi\mp\chi^\dagger\phi_t)^2$ that arises from the term in L_{pt} in Eq. (3.2). That the term in χ drops out is in agreement with our expectation that the specification of the phase should be irrelevant for this case of a diagonal matrix element. Introducing L_c defined by

$$L_c=\tfrac{1}{2}(\phi^\dagger W\phi)L_b+\Lambda_a W\phi, \tag{5.7}$$

so that L_{ct} is a trial approximation to L_c, and using Eqs. (3.11) and (3.12) which define, respectively, L_b and Λ_a, we again find that the term (in χ) specifying the phase disappears, and we have

$$(H-E)L_c=-W\phi+(\phi^\dagger W\phi)\phi. \tag{5.8}$$

L_c is, therefore, the function L in Eq. (5.2a) and the two variational expressions in Eqs. (5.1) and (5.5) coincide. This demonstrates the formal equivalence to second order of the two approaches. We emphasize that the philosophy and mechanics of the application of the two procedures could be different since the choice of trial functions in the two methods need not be related at all through expressions such as Eqs. (5.6) and (5.8). In any numerical calculation, different values will normally be obtained for the two variational estimates because they do differ in second order; the exact numerical difference between them will depend on the choice and accuracy of the trial approximations. What is ensured is that as the trial functions get more accurate, the two expressions will rapidly (quadratically) converge to the same value.

In the application of Eq. (5.4), one may ask how Λ_{at} and L_{bt} in $\phi_t^{(1)}$ are to be chosen. This need not necessarily be with a view to the operator W of interest, that is, guided by Eqs. (5.7) and (5.8). (Whereas L_{ct} or L_t is tailored to the W of interest, Λ_{at} and L_{bt} need not be.) After all, in writing down ϕ_v, no specific W need be in mind and, therefore, one may work once and for all to get a very good Λ_{at} and L_{bt} and, thereby, a very good ϕ_v so that one can get accurate estimates for any $\langle W\rangle$ that one wishes. Practically, however, it is clear that this will be a harder task since getting a good trial Green's function, Λ_{at}, means getting a good description of the entire spectrum (not just the state of interest) of the operator H whereas getting a good L_{ct} (or L_t) may require much less information depending on the nature of W (essentially, only those states of the spectrum that are strongly connected to ϕ by W need be described well). It would seem, therefore, that as a practical pro-

cedure of getting an estimate $\langle W\rangle$ for some W the direct method of Eq. (5.1) will be preferable though the more global approach of getting a good ϕ_v in Eq. (5.3) once and for all so that one may evaluate any functional (and not even necessarily only bilinear diagonal matrix elements) has its own appeal, such a use for many problems compensating for the greater initial effort in obtaining the ϕ_v.

B. Bound state off-diagonal matrix elements

In Ref. 2, we derived a VP for $\phi_1^\dagger W\phi_2$ where ϕ_1 and ϕ_2 are two eigenstates of H. As distinct from the diagonal case, the specification of the relative phase of the functions is now important though the absolute phases are not and in Ref. 2, we chose to consider $\phi_1^\dagger W\phi_2$ to be pure real or pure imaginary:

$$\phi_1^\dagger W\phi_2 \mp \phi_2^\dagger W\phi_1 = 0. \tag{5.9}$$

This element makes it a little more difficult to establish the equivalence of $(\phi_1^\dagger W\phi_2)_v$ that we derived there with $\phi_{1v}^\dagger W\phi_{2v}$ as evaluated with the ϕ_v's obtained in Sec. 3 of this paper since the phase specifications made now are different, namely, each of the phases is fixed absolutely by the conditions

$$\phi_1^\dagger\chi_1 \mp \chi_1^\dagger\phi_1 = 0, \quad \phi_2^\dagger\chi_2 \mp \chi_2^\dagger\phi_2 = 0, \tag{5.10}$$

where the χ's are known. To connect Eqs. (5.9) and (5.10) would require $\chi_1 = W\phi_2$ and $\chi_2 = W\phi_1$, but ϕ_1 and ϕ_2 are unknown and this would contradict our assumption that the χ's are known. This slight complication can be overcome in a number of ways. We could rederive our earlier results for $(\phi_1^\dagger W\phi_2)_v$ with Eq. (5.10) specifying the phase in place of Eq. (5.9). Alternatively, we could consider the derivation of ϕ_{1v} in Sec. 3 of this paper with the choice $\chi_1 = W\phi_{2t}$ and of ϕ_{2v} with $\chi_2 = W\phi_{1t}$. The desired identification follows using either of the two approaches, and we will not bother with the details.

6. SUPERVARIATIONAL PRINCIPLES

We make a few remarks about going beyond the VP to what one might call a super-VP which gives an estimate that is accurate to higher than second order. These comments are suggested by our development of a variationally accurate ϕ_v which is itself good to second order, so that, starting from a ϕ_t with first-order errors, we have

$$\phi_v - \phi = \mathcal{S}, \tag{6.1}$$

where $\mathcal{S}$ contains second order errors like $(\delta\phi)^2$, $\delta L_b\delta\phi$, and $\delta\Lambda_a\delta\phi$. As elaborated upon in Sec. 5, were we to use this ϕ_v and evaluate $\phi_v^\dagger W\phi_v$, we would obtain as in Eq. (5.5) a VP for $\langle W\rangle$ which would have similar second-order errors. Direct use of ϕ_t in $\langle W\rangle_v$ in Eq. (5.1) would give an equivalent VP for $\langle W\rangle$ with second-order errors of the form $(\delta\phi)^2$ and $\delta L\delta\phi$. However, were we to use ϕ_v as the trial function in Eq. (5.1), we would get something better than either of these results, something better than a VP, namely, a super-VP for $F(\phi)$ with errors of the form $(\delta L)\mathcal{S}$ and $\mathcal{S}^2$. These are errors of third or fourth order.

Though the matter should be pursued only after the utility of the super-VP for $F(\phi)$ noted above has been verified, we would at least like to note the possibility of using our VP for ϕ in an iterative fashion to produce super-VP's for ϕ (and, as a result, for any functional of ϕ) of increasingly (arbitrarily) higher order. This has also been noted in Ref. 9. Starting from an initial ϕ_t and solving for Λ_{at} and L_{bt} as in Sec. 7 would give a ϕ_v good to second order. With this as the trial function, we could repeat the process, recalculating a new Λ_{at} and L_{bt} and obtain a new ϕ_v, which we might call $\phi_v^{(2)}$, which would be good to third or fourth order, and iterate this process. It is also conceivable that Λ_{at} and L_{bt} need not be evaluated anew at each stage of the iteration but left unchanged from their values in the first step. Since the errors $\delta\Lambda_a$ and δL_b appear only in the form of a product with $\delta\phi^{(1)}$, successive improvements of $\delta\phi^{(1)}$ alone, with Λ_{at} and L_{bt} held fixed, may suffice. These are remarks that have to await for their confirmation an actual practical calculation.

7. PROCEDURE FOR CHOOSING TRIAL AUXILIARY FUNCTIONS

The validity of the VP requires, of course, that the error in each trial function is of first or higher order. As we have discussed previously,[3] some care must be exercised in choosing approximate solutions of equations of the type Eq. (3.11) and Eq. (3.12). Such solutions involve, essentially, the inverse of the operator $(H-E)$ modified by the removal of the singularity corresponding to the solution ϕ of the homogeneous equation. An approximate solution may introduce a new singularity which contributes an unacceptable error of *zeroth* order. The procedure suggested in Ref. 3 for avoiding this difficulty may be taken over directly to determine the trial auxiliary functions encountered here. We shall briefly summarize this procedure, confining ourselves for simplicity to the case where ϕ is the ground-state function.

Consider the equation

$$(H-E)L = q(\phi), \tag{7.1a}$$

where q is a square-integrable function and

$$\phi^\dagger q(\phi) = 0. \tag{7.1b}$$

We complete the definition of L by requiring that $\phi^\dagger L = 0$. In place of Eq. (7.1a) we consider

$$(H^{(1)}_{\mathrm{mod},t} - E_t)L_t = q(\phi_t), \tag{7.2}$$

with

$$H^{(1)}_{\mathrm{mod},t} \equiv H - HP_tH/E_t\,; \tag{7.3a}$$

we have defined

$$P_t \equiv \phi_t\phi_t^\dagger, \tag{7.3b}$$

and have chosen $E_t = \phi_t^\dagger H\phi_t$. We note that $\phi_t^\dagger H^{(1)}_{\mathrm{mod},t} = 0$. We impose the condition $\phi_t^\dagger q(\phi_t) = 0$, as suggested by Eq. (7.1b). We also note that $H^{(1)}_{\mathrm{mod},t}$ approaches $H - E\phi\phi^\dagger$ as $\phi_t \to \phi$. By projecting Eq. (7.2) onto $\phi_t^\dagger$ we easily see that $\phi_t^\dagger L_t = 0$. Since the equation and the boundary condition that define L_t approach those that define L as $\phi_t \to \phi$, it follows that $L_t \to L$ as $\phi_t \to \phi$. Thus, for ϕ_t a good approximation to ϕ, Eq. (7.2) may be used as a replacement for Eq. (7.1) for the purpose

of defining a solution L_t which is a good approximation to L. The advantage of this reformulation lies in the fact that for ϕ_t sufficiently accurate, a condition that can be stated precisely,[3] $H^{(1)}_{\mathrm{mod},t}$ will have its lowest eigenvalue above E_t.[3] Thus, no singularity difficulties appear in this this reformulation. A practical procedure for determining the variational parameters to be used in a trial function L_{tt}, an approximation to L_t, is to minimize

$$M(L_{tt}) = L_{tt}^{\dagger}[H^{(1)}_{\mathrm{mod},t} - E_t]L_{tt} - L_{tt}^{\dagger}q(\phi_t) - q^{\dagger}(\phi_t)L_{tt} \qquad (7.4)$$

with respect to these parameters.

The above discussion can be applied directly to Eq. (3.11) with $q = \phi - \chi/(\phi^{\dagger}\chi)$. In a similar way, Eq. (3.12) may be replaced by

$$(H^{(1)}_{\mathrm{mod},t} - E_t)\Lambda_{at} = -1 + \tfrac{1}{2}\phi_t\phi_t^{\dagger} + \tfrac{1}{2}\chi\phi_t^{\dagger}/(\phi_t^{\dagger}\chi). \qquad (7.5)$$

In this case the inhomogeneous term is not square-integrable so that a minimum principle of the type described in connection with Eq. (7.4) cannot be used directly to determine Λ_{at}. We can, however, consider the function

$$L_t(\mathbf{r}') = \int \Lambda_{at}(\mathbf{r},\mathbf{r}')F(\mathbf{r})\,d\mathbf{r} \equiv \Lambda_{at}^{\dagger}(\mathbf{r}',\mathbf{r})F(\mathbf{r}),$$

where in the second step we used $\Lambda_{at}^{\dagger}(\mathbf{r}',\mathbf{r}) = \Lambda_{at}(\mathbf{r},\mathbf{r}')$ and the notation $A^{\dagger}B$ implies integration, and where F is an arbitrary square-integrable function. Then L_t satisfies Eq. (7.2) with

$$q(\phi_t) = -F + \tfrac{1}{2}\phi_t(\phi_t^{\dagger}F) + \tfrac{1}{2}\chi(\phi_t^{\dagger}F)/(\phi_t^{\dagger}\chi) \qquad (7.6)$$

and the minimum principle can be used to determine the variational parameters in Λ_{att}, the approximation to the trial Green's function Λ_{at}. [We note that the choice $F = \chi$ in Eq. (7.6) gives $q(\phi_t) = -(1/2)Q_t\chi$]. This procedure would be acceptable in practice if the optimum values of the variational parameters change only slightly as the arbitrary function F is changed. We also note that a natural choice for F will be $(H - E_t)\phi_t$ because the projection of Λ_{at} on this is what appears in the variational expression and, typically, ϕ_t will be some given trial function (say, from a Rayleigh–Ritz calculation) that we are trying to improve. For any χ and for $F = (H - E_t)\phi_t$, we have $q(\phi_t) = -(H - E_t)\phi_t$. Numerical tests of the subsidiary minimum principle discussed above are now in progress.

We note, finally, that the choice of modified Hamiltonian which appears in the above analysis is not unique. As an alternative to the choice made in Eq. (7.3a) we may consider a modified Hamiltonian that has been studied previously[18,21] in a slightly different context,

$$\hat{H}^{(1)}_{\mathrm{mod},t} \equiv P_tHP_t + Q_tHQ_t, \qquad (7.7)$$

where $Q_t \equiv 1 - P_t$ with P_t given by Eq. (7.3b). The lowest eigenvalue of $\hat{H}^{(1)}_{\mathrm{mod},t}$ may be shown[22] to lie above E_t for "sufficiently accurate" ϕ_t.

*Work supported in part by the Advanced Research Project Agency under Contract No. DA-31-124-ARO-D-440, and by the National Aeronautics and Space Administration under Contract No. NGL-39-011-035.
†Supported by the National Science Foundation.
‡Work supported by the U.S. Army Research Office, Durham, under Grant No. DA-ARO-D-31-124-72-G92 and by the Office of Naval Research, under Contract No. N00014-67-A-0467-0007.

[1]E. Gerjuoy, A.R.P. Rau, and L. Spruch, J. Math. Phys. **13**, 1797 (1972).
[2]E. Gerjuoy, A.R.P. Rau, and L. Spruch, Phys. Rev. A **8**, 662 (1973).
[3]E. Gerjuoy, A.R.P. Rau, L. Rosenberg, and L. Spruch, Phys. Rev. A **9**, 108 (1974).
[4]L. Rosenberg, Phys. Rev. D **9**, 1789 (1974).
[5]L. Rosenberg and L. Spruch, Phys. Rev. A **10**, 2002 (1974).
[6]As far as the considerations of this paper are concerned, ϕ could be anything and, in particular, it could be a classical quantity such as the amplitude of vibration of a string or the electrostatic potential between the charged plates of a capacitor. Our use of quantum-mechanical language and terming ϕ a wavefunction is done only on grounds of familiarity and convenience.
[7]Some of the relevant references were included in our previous papers. One which we should have, but were unaware of at that time, is G.C. Pomraning, J. Math. Phys. **8**, 149 (1967).
[8]L.C. Biedenharn and J.M. Blatt, Phys. Rev. **93**, 230 (1954).
[9]T. Kikuta, Progr. Theor. Phys. (Japan) **12**, 10 (1954); **14**, 457 (1955). See also the reference quoted in Footnote 7.
[10]S. Altshuler, Phys. Rev. **109**, 1830 (1958).
[11]A similar procedure was carried out in the paper quoted in Footnote 7, but the form of g selected was more complicated.
[12]We considered in Ref. 1 either the case $\omega \neq 0$ which fixes the phase of ϕ or functionals F which did not depend on the phase of ϕ.
[13]For an alternative method of constructing variational identities and the related VP's for wavefunctions, see Ref. 4.
[14]L. Spruch and L. Rosenberg, Phys. Rev. **116**, 1034 (1959).
[15]L. Rosenberg, L. Spruch, and T.F. O'Malley, Phys. Rev. **118**, 184 (1960).
[16]L. Spruch, *Lectures in Theoretical Physics, Boulder*, 1961, edited by W.E. Brittin, W.B. Downs, and J. Downs (Wiley, New York, 1962), Vol. 4, p. 161.
[17]R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Wiley, New York, 1965).
[18]S. Aranoff and J. Percus, Phys. Rev. **166**, 1255 (1968).
[19]E.P. Wigner, Nachr. Ges. Wiss. Goett. Math.-Phys. Kl. **32**, 35 (1932).
[20]C. Schwartz, Ann. Phys. (N.Y.) **2**, 156, 170 (1959). See also A. Dalgarno and A.L. Stewart, Proc. R. Soc. Lond. A **257**, 534 (1960), and L.M. Delves, Proc. Phys. Soc. Lond. **92**, 55 (1967).
[21]S. Borowitz and E. Gerjuoy, Univ. of Colorado JILA Report No. 36, 1965 (unpublished).
[22]E. Gerjuoy, L. Rosenberg, and L. Spruch, J. Math. Phys. **16**, 455 (1975).

Pyramidal composition rules for Wronskians upon Wronskians

Robert C. Brunet

Département de Mathématiques, Université de Montreal, Montréal, Québec, Canada

Institut de Physique Nucléaire de Lyon, Université Claude Bernard, 69621 Villeurbanne, France*
(Received 27 March 1974)

We give composition rules for Wronskians which have Wronskians as arguments. These pyramids of Wronskians are shown to reduce to products of Wronskians of different order. Many symmetric and antisymmetric regroupings of functions are then possible. The arithmetic of Wronskians can efficiently be reduced to the application of the $\overleftrightarrow{\partial}$ of the LSZ formalism.

I. INTRODUCTION

The role played by Wronskians in differential equations as they relate solutions to each other, and to the coefficients is well known,[1,2] as is their convenience in describing Green's functions.[3] What does not appear known is that the Wronskians on a set of linearly independent differentiable functions satisfy a number of interesting relations quite apart from any reference to differential equations. We shall refer to them as pyramidal composition rules as they relate Wronskians taken upon Wronskians to products of Wronskians taken upon the original set of functions, but of lower or higher order. These algorithms where Wronskians are nested within Wronskians are due to the peculiar combination of the determinental and differential features of the Wronskian which is a completely antisymmetric multilinear differential form. The simplest realization lies in the repeated use of the antisymmetric operator $\overleftrightarrow{\partial}$, used by Lehmann, Symanzyk and Zimmermann,[4] which is equivalent to forming a Wronskian of order 2. Consider the differential form

$$\delta(\varphi_1, \varphi_2, \varphi_3) = (\varphi_1 \overleftrightarrow{\partial} \varphi_2) \overleftrightarrow{\partial} (\varphi_1 \overleftrightarrow{\partial} \varphi_3)$$

$$= W_2[W_2(\varphi_1, \varphi_2), W_2(\varphi_1, \varphi_3)]. \tag{1.1}$$

This form is manifestly antisymmetric under the exchange $\varphi_2 \leftrightarrow \varphi_3$; however, it is a simple matter to verify that

$$\varphi_1^{-1}\delta(\varphi_1, \varphi_2, \varphi_3) = W_3(\varphi_1, \varphi_2, \varphi_3). \tag{1.2}$$

We see that the privileged stance of the repeated function φ_1 in Eq. (1.1) can be removed easily thru appropriate "normalization" to obtain $W_3(\varphi_1, \varphi_2, \varphi_3)$ which is completely antisymmetric. This illustrates also the previous statement that Wronskians of Wronskians are related to Wronskians on the original set, but of a different order.

In this article we give general composition rules for Wronskians upon Wronskians that reduce to Wronskians on the original functions. These yield many combinations where not only antisymmetric regroupings are possible, but symmetric ones as well. The pyramidal algorithms can also be used to simplify computations involving Wronskians. We present the results in Sec. III as three simple theorems, and in Sec. IV we give graphical rules which exhibit the pyramidal structure.

II. PRELIMINARIES

Given a set of linearly independent functions of one variable $\Phi = \{\varphi_1(x), \varphi_2(x), \cdots\}$ sufficiently differentiable, we define the Wronskian on a subset of these functions $\Phi_n \subset \Phi$ as

$$W_n(\varphi_1, \varphi_2, \cdots, \varphi_n) \equiv \begin{vmatrix} \varphi_1 & \varphi_2 & \cdots & \varphi_n \\ \varphi_1' & \varphi_2' & \cdots & \varphi_n' \\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ \varphi_1^{(n-1)} & \varphi_2^{(n-1)} & \cdots & \varphi_n^{(n-1)} \end{vmatrix}, \tag{2.1}$$

where $\varphi_i', \varphi_i'', \cdots, \varphi_i^{(n-1)}$ indicate the first, second $\cdots$, $(n-1)$th derivative of a function $\varphi_i(x)$. It is also convenient to define $W_0 \equiv 1$ where W_0 is the Wronskian on the empty subset. (If one is dealing with functions of several variables, this definition is easily extended to the total or partial derivatives.)

Wronskians have the following useful features:

(i) The derivative of a Wronskian differs from it by the last row only:

$$\frac{d}{dx} W_n(\varphi_1, \cdots, \varphi_n) = \begin{vmatrix} \varphi_1 & \cdots & \varphi_n \\ \varphi_1' & \cdots & \varphi_n' \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \varphi_1^{(n-2)} & \cdots & \varphi_n^{(n-2)} \\ \varphi_1^{(n)} & \cdots & \varphi_n^{(n)} \end{vmatrix}; \tag{2.2}$$

(ii) $$W_2[\varphi_1\varphi_2, \varphi_3\varphi_4] = \varphi_1\varphi_3 W_2[\varphi_2, \varphi_4] + \varphi_2\varphi_4 W_2[\varphi_1, \varphi_3]$$
$$= \varphi_1\varphi_4 W_2[\varphi_2, \varphi_3] + \varphi_2\varphi_3 W_2[\varphi_1, \varphi_4];$$

(2.3)

(iii) from which for a differentiable function $f(x)$, we

 Copyright © 1975 American Institute of Physics

get

$$W_2\,[f\varphi_1, f\varphi_2] = f^2 W_2[\varphi_1, \varphi_2], \tag{2.4}$$

and by systematic expansion of the Wronskians of ever higher order by their last row we obtain

$$W_n[f\varphi_1, f\varphi_2, \cdots, f\varphi_n] = f^n W_n[\varphi_1, \cdots, \varphi_n]. \tag{2.5}$$

It is also useful to recall Laplace's development of determinants by minors (see, for example, Smirnov's book[5]). From a determinant Δ of order n, define a minor of order $l < n$:

$$A\begin{pmatrix} p_1, \cdots, p_l \\ q_1, \cdots, q_l \end{pmatrix} = \begin{vmatrix} \delta_{p_1 q_1} \cdots \delta_{p_1 q_l} \\ \cdot \quad\quad \cdot \\ \cdot \quad\quad \cdot \\ \cdot \quad\quad \cdot \\ \delta_{p_l q_1} \cdots \delta_{p_l q_l} \end{vmatrix}, \tag{2.6}$$

obtained by deleting $(n-l)$ rows and columns from Δ. The indices $p_1 \cdots p_l$ of the selected rows and $q_1 \cdots q_l$ of the selected columns are arranged in increasing order. With the deleted rows: $r_1, \cdots, r_{n-l}$ and deleted columns: $s_1, \cdots, s_{n-l}$ arranged in increasing order we define a complementary minor to A called

$$\mathcal{A}\begin{pmatrix} p_1, \cdots, p_l \\ q_1, \cdots, q_l \end{pmatrix} = \begin{vmatrix} \delta_{r_1 s_1} & \delta_{r_1 s_{n-l}} \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \delta_{r_{n-l} s_1} & \cdots \delta_{r_{n-l} s_{n-l}} \end{vmatrix}. \tag{2.7}$$

Laplace's theorem says that for a fixed set of row indices $p_1, \cdots, p_l$ we can write

$$\Delta = (-1)^{p_1+\cdots+p_l} \left(\sum_{q_1<q_2<\cdots<q_l} (-1)^{q_1+\cdots+q_l} \times A\begin{pmatrix} p_1 \cdots p_l \\ q_1 \cdots q_l \end{pmatrix} \mathcal{A}\begin{pmatrix} p_1 \cdots p_l \\ q_1 \cdots q_l \end{pmatrix} \right) \tag{2.8}$$

where the sum is over the $n!/l!(n-l)!$ subsequences $q_1 < q_2 \cdots < q_l$ of the sequence $1, 2, \cdots, n$. A similar expansion exists if one starts with a fixed set of columns, and sums over ordered subsequences of rows.

III. WRONSKIANS UPON WRONSKIANS

Definition 1: Given a Wronskian $W_n(\varphi_1, \cdots, \varphi_n)$ on the subset $\Phi_n \subset \Phi$, we define a stepped-up Wronskian $W_{n+1}(\varphi_1, \cdots, \varphi_n; \varphi_p)$ as the one built on the subset Φ_n enlarged by the new function $\varphi_p \notin \Phi_n$ (the " ;" is only for convenience in bookeeping).

Definition 2: Similarly given a Wronskian $W_n(\varphi_1, \cdots \varphi_n)$ on the subset $\Phi_n \subset \Phi$, we define a stepped-down Wronskian $W_{n-1}(\varphi_1, \cdots, \not{\varphi}_j, \cdots, \varphi_n)$ as the one built on the subset Φ_n diminished by one of the functions $\varphi_j \in \Phi_n$.

Theorem 1: If the Wronskian $W_n(\varphi_1, \cdots, \varphi_n)$ is stepped up alternatively with each of two new functions φ_{n+1} and φ_{n+2}, then the Wronskian of order 2 built on these two stepped-up Wronskians is equal to the product of the basic W_n with the twice stepped-up Wronskian $W_{n+2}(\varphi_1, \cdots, \varphi_n; \varphi_{n+1}, \varphi_{n+2})$. That is,

$$W_2\,[W_{n+1}(\varphi_1, \cdots, \varphi_n; \varphi_{n+1}),\ W_{n+1}(\varphi_1, \cdots, \varphi_n; \varphi_{n+2})] = W_n(\varphi_1, \cdots, \varphi_n) \cdot W_{n+2}(\varphi_1, \cdots, \varphi_n; \varphi_{n+1}, \varphi_{n+2}). \tag{3.1}$$

This result is obtained by imbedding the above W_n and W_{n+2} in a determinant Δ of order $2n+2$. Consider

$$\Delta = \left|\begin{array}{ccc|ccccc} \varphi_1 & \cdots & \varphi_n & \varphi_1 & \cdots & \varphi_n & \varphi_{n+1} & \varphi_{n+2} \\ \cdot & & \cdot & \cdot & & \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot & & \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot & & \cdot & \cdot & \cdot \\ \varphi_1^{(n-1)} & \cdots & \varphi_n^{(n-1)} & \varphi_1^{(n-1)} & \cdots & \varphi_n^{(n-1)} & \varphi_{n+1}^{(n-1)} & \varphi_{n+2}^{(n-1)} \\ \hline 0 & \cdots & 0 & \varphi_1 & \cdots & \varphi_n & \varphi_{n+1} & \varphi_{n+2} \\ 0 & \cdots & 0 & \cdot & & \cdot & \cdot & \cdot \\ 0 & \cdots & 0 & \cdot & & \cdot & \cdot & \cdot \\ 0 & \cdots & 0 & \cdot & & \cdot & \cdot & \cdot \\ 0 & \cdots & 0 & \varphi_1^{(n-1)} & \cdots & \varphi_n^{(n-1)} & \varphi_{n+1}^{(n-1)} & \varphi_{n+2}^{(n-1)} \\ \varphi_1^{(n)} & \cdots & \varphi_n^{(n)} & \varphi_1^{(n)} & \cdots & \varphi_n^{(n)} & \varphi_{n+1}^{(n)} & \varphi_{n+2}^{(n)} \\ 0 & \cdots & 0 & \varphi_1^{(n+1)} & \cdots & \varphi_n^{(n+1)} & \varphi_{n+1}^{(n+1)} & \varphi_{n+2}^{(n+1)} \end{array}\right|. \tag{3.2}$$

Choosing a Laplace development for Δ with the first n columns fixed, it is clear that Δ is equal to the product $W_n(\varphi_1, \cdots, \varphi_n)$ times $W_{n+2}(\varphi_1, \cdots, \varphi_n; \varphi_{n+1}, \varphi_{n+2})$ since only the diagonal blocks contribute, for any minor $(n \times n)$ formed with the second to the last row has a complementary minor with a repeated row. We then permute the second to the last row upstairs n times and get

$$\Delta = (-1)^n \left|\begin{array}{ccc|ccccc} \varphi_1 & \cdots & \varphi_n & \varphi_1 & \cdots & \varphi_n & \varphi_{n+1} & \varphi_{n+2} \\ \cdot & & \cdot & \cdot & & \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot & & \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot & & \cdot & \cdot & \cdot \\ \varphi_1^{(n-1)} & \cdots & \varphi_n^{(n-1)} & \varphi_1^{(n-1)} & \cdots & \varphi_n^{(n-1)} & \varphi_{n+1}^{(n-1)} & \varphi_{n+2}^{(n-1)} \\ \varphi_1^{(n)} & \cdots & \varphi_n^{(n)} & \varphi_1^{(n)} & \cdots & \varphi_n^{(n)} & \varphi_{n+1}^{(n)} & \varphi_{n+2}^{(n)} \\ \hline & & & \varphi_1 & \cdots & \varphi_n & \varphi_{n+1} & \varphi_{n+2} \\ & 0 & & \cdot & & \cdot & \cdot & \cdot \\ & & & \cdot & & \cdot & \cdot & \cdot \\ & & & \cdot & & \cdot & \cdot & \cdot \\ & & & \varphi_1^{(n-1)} & \cdots & \varphi_n^{(n-1)} & \varphi_{n+1}^{(n-1)} & \varphi_{n+2}^{(n-1)} \\ & & & \varphi_1^{(n+1)} & \cdots & \varphi_n^{(n+1)} & \varphi_{n+1}^{(n+1)} & \varphi_{n+2}^{(n+1)} \end{array}\right|. \tag{3.3}$$

Choosing this time a Laplace development with the top $(n+1)$ rows fixed, the minors and complementary minors are both of order $(n+1)$. The only minors which neither have repeated rows, nor have a complementary minor with zero columns, are the two minors formed with the first n columns and either of the last two col-

umns. Thus,

$$\Delta = (-1)^n(-1)^{1+\cdots+n+(n+1)}$$

$$\times\left[\begin{vmatrix} & & \\ & & \\ & & \\ & & \\ \varphi_1^{(n-1)} & \cdots & \varphi_n^{(n-1)} \\ \varphi_1^{(n)} & \cdots & \varphi_n^{(n)} \end{vmatrix}\begin{vmatrix} \varphi_1 & \cdots & \varphi_n & \varphi_{n+1} \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \varphi_1^{(n-1)} & \cdots & \varphi_n^{(n-1)} & \varphi_{n+1}^{(n-1)} \\ \varphi_1^{(n)} & \cdots & \varphi_n^{(n)} & \varphi_{n+1}^{(n)} \end{vmatrix}\right.$$

$$\times\begin{vmatrix} \varphi_1 & \cdots & \varphi_n & \varphi_{n+2} \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \varphi_1^{(n-1)} & \cdots & \varphi_n^{(n-1)} & \varphi_{n+2}^{(n-1)} \\ \varphi_1^{(n+1)} & \cdots & \varphi_n^{(n+1)} & \varphi_{n+2}^{(n+1)} \end{vmatrix} + (-1)^{1+\cdots+n+(2n+2)}$$

$$\left.\times\begin{vmatrix} \varphi_1 & \cdots & \varphi_n & \varphi_{n+2} \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \varphi_1^{(n-1)} & \cdots & \varphi_n^{(n-1)} & \varphi_{n+2}^{(n-1)} \\ \varphi_1^{(n)} & \cdots & \varphi_n^{(n)} & \varphi_{n+2}^{(n)} \end{vmatrix}\times\begin{vmatrix} \varphi_1 & \cdots & \varphi_n & \varphi_{n+1} \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \varphi_1^{(n-1)} & \cdots & \varphi_n^{(n-1)} & \varphi_{n+1}^{(n-1)} \\ \varphi_1^{(n+1)} & \cdots & \varphi_n^{(n+1)} & \varphi_{n+1}^{(n+1)} \end{vmatrix}\right]. \tag{3.4}$$

Noting that the complementary minors are derivatives of Wronskians of order $(n+1)$, we have the desired result:

$$\Delta = W_{n+1}(\varphi_1, \cdots, \varphi_n; \varphi_{n+1}) \cdot W'_{n+1}(\varphi_1, \cdots, \varphi_n; \varphi_{n+2}) - W_{n+1}(\varphi_1, \cdots, \varphi_n; \varphi_{n+2}) \cdot W'_{n+1}(\varphi_1, \cdots, \varphi_n; \varphi_{n+1}). \tag{3.5}$$

Corollary 1: Since $W_n(\varphi_1, \cdots, \varphi_n)$ stepped up once alternatively with φ_{n+1} and φ_{n+2} is equivalent to $W_{n+2}(\varphi_1, \cdots, \varphi_n; \varphi_{n+1}, \varphi_{n+2})$ stepped down once alternatively with φ_{n+2} and φ_{n+1}, putting $n \to n-2$ in (3.1), we write for future reference

$$W_2\,[W_{n-1}(\varphi_1, \cdots; \varphi_{n-1}, \phi_n),\ W_{n-1}(\varphi_1, \cdots; \phi_{n-1}, \varphi_n)] = W_n(\varphi_1, \cdots, \varphi_n) \cdot W_{n-2}(\varphi_1, \cdots, \varphi_{n-2}; \phi_{n-1}, \phi_n), \tag{3.6}$$

where $W_{n-2}(\varphi_1, \cdots, \varphi_{n-2}; \phi_{n-1}, \phi_n)$ is the twice stepped-down Wronskian, and $n \geq 2$.

Theorem 2: If the Wronskian $W_n(\varphi_1, \cdots, \varphi_n)$ is stepped up once, alternatively with k new functions, say $\varphi_{n+1}, \cdots, \varphi_{n+k}$, then the Wronskian of order k built upon these stepped up Wronskians is equal to the product of the basic Wronskian W_n to the power $(k-1)$ with the Wronskian stepped up k times. That is,

$$W_k\{W_{n+1}(\varphi_1, \cdots, \varphi_n; \varphi_{n+1}),\ W_{n+1}(\varphi_1, \cdots, \varphi_n; \varphi_{n+2}), \cdots, W_{n+1}(\varphi_1, \cdots, \varphi_n; \varphi_{n+k})\} = W_n^{k-1}(\varphi_1, \cdots, \varphi_n) \cdot W_{n+k}(\varphi_1, \cdots, \varphi_n; \varphi_{n+1}, \varphi_{n+2}, \cdots, \varphi_{n+k}). \tag{3.7}$$

Proof: Since in Theorem 1 we have shown this formula to be valid for $k=2$ and any value of n, a proof by induction will suffice ($k=1$ is an identity). Let us therefore assume it true for k and show it true for $k+1$. Taking (3.7) to be true for k with $n=1$ on a set $\Psi = \{\psi_1, \psi_2, \cdots, \psi_{k+1}\}$ we have

$$\psi_1^{k-1} \cdot W_{k+1}(\psi_1; \psi_2, \cdots, \psi_{k+1}) = W_k\{W_2(\psi_1; \psi_2), W_2(\psi_1; \psi_3), \cdots, W_2(\psi_1; \psi_{k+1})\}. \tag{3.8}$$

We can identify the functions ψ_i with the following $k+1$ Wronskians of order $(n+1)$ such that: $\psi_1 \equiv W_{n+1}(\varphi_1, \cdots, \varphi_n; \varphi_{n+1}), \cdots, \psi_{k+1} \equiv W_{n+1}(\varphi_1, \cdots, \varphi_n; \varphi_{n+k+1})$.

Each of the $W_2(\psi_1; \psi_j)$, for $j = 2, \cdots, k+1$, on the rhs of (3.8) can be written using Theorem 1:

$$W_2(\psi_1; \psi_j) = W_2\{W_{n+1}(\varphi_1, \cdots, \varphi_n; \varphi_{n+1}), W_{n+1}(\varphi_1, \cdots, \varphi_n; \varphi_{n+j})\} = W_n(\varphi_1, \cdots, \varphi_n) \cdot W_{n+2}(\varphi_1, \cdots, \varphi_{n+1}; \varphi_{n+j}). \tag{3.9}$$

Substituting on both sides of (3.8) and noting that each $W_2(\psi_1; \psi_j)$ gives rise to a common function $W_n(\varphi_1, \cdots, \varphi_n)$ which can be factored out according to (2.5), one gets

$$W_{n+1}^{k-1}(\varphi_1, \cdots, \varphi_n; \varphi_{n+1}) \cdot W_{k+1}\{W_{n+1}(\varphi_1, \cdots, \varphi_n; \varphi_{n+1}), \cdots, W_{n+1}(\varphi_1, \cdots, \varphi_n; \varphi_{n+k+1})\} = W_n^k(\varphi_1, \cdots, \varphi_n) \cdot W_k\{W_{n+2}(\varphi_1, \cdots, \varphi_{n+1}; \varphi_{n+2}), \cdots, W_{n+2}(\varphi_1, \cdots, \varphi_{n+1}; \varphi_{n+k+1})\}. \tag{3.10}$$

Having assumed (3.7) true for k and any value of n, we can use it, with $n \to n+1$, to write the W_k on the rhs of (3.10) as equal to

$$W_{n+1}^{k-1}(\varphi_1, \cdots, \varphi_{n+1}) \cdot W_{n+1+k}(\varphi_1, \cdots, \varphi_{n+1}; \varphi_{n+2}, \cdots, \varphi_{n+k+1}). \tag{3.11}$$

After simplifying through the factor $W_{n+1}^{k-1}(\varphi_1, \cdots, \varphi_{n+1})$, (3.10) with (3.11) gives

$$W_{k+1}\{W_{n+1}(\varphi_1, \cdots, \varphi_n; \varphi_{n+1}), \cdots, W_{n+1}(\varphi_1, \cdots, \varphi_n; \varphi_{n+k+1})\} = W_n^k(\varphi_1, \cdots, \varphi_n) \cdot W_{n+k+1}(\varphi_1, \cdots, \varphi_n; \varphi_{n+1}, \cdots, \varphi_{n+k+1}). \tag{3.12}$$

QED

Theorem 3: If the Wronskian $W_n(\varphi_1, \cdots, \varphi_n)$ is stepped down once, alternatively with k ($\leq n$) of the functions of the subset Φ_n, then the Wronskian of order k built upon these stepped-down Wronskians is equal to the product of the basic Wronskian W_n to the power $(k-1)$ with the Wronskian stepped down k times. That is,

$$W_k\{W_{n-1}(\varphi_1, \cdots, \phi_n), \cdots, W_{n-1}(\varphi_1, \cdots, \phi_{n-k+1}, \cdots, \varphi_n)\} = W_n^{k-1}(\varphi_1, \cdots, \varphi_n) \cdot W_{n-k}(\varphi_1, \cdots, \varphi_{n-k}; \phi_{n-k+1}, \cdots, \phi_n), \tag{3.13}$$

where $W_{n-k}(\varphi_1, \cdots, \varphi_{n-k}; \phi_{n-k+1}, \cdots, \phi_n)$ is the original W_n stepped down k times by the removal of the functions $\{\varphi_{n-k+1}, \cdots, \varphi_n\}$. For $k=n$ we have $W_0 = 1$.

Proof: By Corollary 1 we know (3.13) to be true for $k=2$ (any value of $n \geq 2$), and also that (3.13) is identically true for $k=1$. We will again proceed by induction, this time assuming the formula true for k and $k-1$ and showing it true for $k+1$.

Consider (3.1) with its $n = k-1$ on a set $\{\psi_1, \psi_2, \cdots,$

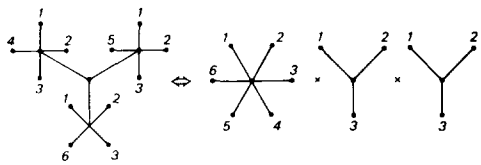


FIG. 1. The stepping up algorithm for Wronskians on three core functions enlarged by three new ones.

$\psi_{k+1}\}$; we have

$$W_{k-1}(\psi_1, \cdots, \psi_{k-1}) \cdot W_{k+1}(\psi_1, \cdots, \psi_{k-1}; \psi_k, \psi_{k+1}) = W_2\{W_k(\psi_1, \cdots, \psi_{k-1}; \psi_k), W_k(\psi_1, \cdots, \psi_{k-1}; \psi_{k+1})\}. \tag{3.14}$$

We identify $\psi_1 = W_{n-1}(\varphi_1, \cdots, \not{\varphi}_n), \cdots, \psi_{k-1} = W_{n-1}(\varphi_1, \cdots, \not{\varphi}_{n-k+2}, \cdots, \varphi_n), \psi_k = W_{n-1}(\varphi_1, \cdots, \not{\varphi}_{n-k+1}, \cdots, \varphi_n)$, and $\psi_{k+1} = W_{n-1}(\varphi_1, \cdots, \not{\varphi}_{n-k}, \cdots, \varphi_n)$.

This substitution applied to the terms of (3.14) yields the following intermediate results when (3.13) is assumed true for $k-1$ and k:

$$W_{k-1}(\psi_1, \cdots, \psi_{k-1}) = W_n^{k-2}(\varphi_1, \cdots, \varphi_n) \cdot W_{n-k+1}(\varphi_1, \cdots, \varphi_{n-k+1}), \tag{3.15}$$

$$W_k(\psi_1, \cdots, \psi_{k-1}; \psi_k) = W_n^{k-1}(\varphi_1, \cdots, \varphi_n) \cdot W_{n-k}(\varphi_1, \cdots, \varphi_{n-k-1}; \varphi_{n-k}), \tag{3.16}$$

$$W_k(\psi_1, \cdots, \psi_{k-1}; \psi_{k+1}) = W_n^{k-1}(\varphi_1, \cdots, \varphi_n) \cdot W_{n-k}(\varphi_1, \cdots, \varphi_{n-k-1}; \varphi_{n-k+1}). \tag{3.17}$$

The latter is obtained after some reordering of the sequences to make (3.13) applicable. Now when the Wronskian of order 2 is built on (3.16) and (3.17) we can factor out the common function $W_n^{k-1}(\varphi_1, \cdots, \varphi_n)$ with (2.4), and using Theorem 1 we have for the rhs of (3.14)

$$W_2\{W_k(\psi_1, \cdots, \psi_{k-1}; \psi_k), W_k(\psi_1, \cdots, \psi_{k-1}; \psi_{k+1})\} = W_n^{2k-2}(\varphi_1, \cdots, \varphi_n) \cdot W_{n-k+1}(\varphi_1, \cdots, \varphi_{n-k+1}) \cdot W_{n-k-1}(\varphi_1, \cdots, \varphi_{n-k-1}). \tag{3.18}$$

Simplification on each side of (3.14) by the explicit factors provided by (3.15) gives finally

$$W_{k+1}\{W_{n-1}(\varphi_1, \cdots, \not{\varphi}_n), \cdots, W_{n-1}(\varphi_1, \cdots, \not{\varphi}_{n-k}, \cdots, \varphi_n)\} = W_n^k(\varphi_1, \cdots, \varphi_n) \cdot W_{n-k-1}(\varphi_1, \cdots, \varphi_{n-k-1}). \tag{3.19}$$

QED

IV. GRAPHICAL REPRESENTATION

It is often convenient for bookkeeping purposes to have graphical illustrations of the above composition rules, especially when several "layers" of Wronskians are used as building blocks to other Wronskians (i.e., if the φ_i above are themselves Wronskians, etc.).

(1) We describe a Wronskian of order j by the vertex of a pyramid with j lines conveying to it its arguments: Wronskians or functions (the latter are but Wronskians of order 1).

(2) Each line stems from a Wronskian of order l.

(3) The labels of the arguments are arranged in increasing order. In the case where the arguments are made of a sequence of once stepped-up (or down) Wronskians, the ordering of these Wronskians is determined by the sum of the labels of their own respective argument functions.

(4) Take k cores of n functions. Enlarge each core by one of k new functions (or remove from each core one of k original functions). Build k Wronskians of order $n+1$ (or $n-1$) on these enlarged (or diminished) cores. Finally, use these k Wronskians as arguments to form a Wronskian of order k. This pyramid is equal to:

(5) The product of (i) the Wronskian of order n, built on the core of n functions, raised to the power $k-1$; (ii) the Wronskian of order $n+k$ (or $n-k$) built on the core of n functions enlarged (or diminished) by k new (or original) functions.

(6) The limitation $k \leq n$ applies only to the case of a diminished core.

Note: As expected the frequency of occurence of each of the core functions involved is the same on both sides of the equation. They are simply rearranged with respect to the Wronskian "operations."

In Fig. 1, we give an example of stepping up a basic core of three functions $\{\varphi_1, \varphi_2, \varphi_3\}$ by three new functions $\{\varphi_4, \varphi_5, \varphi_6\}$. Using Theorem 2 we have

$$W_3\{W_4(\varphi_1, \varphi_2, \varphi_3; \varphi_4), W_4(\varphi_1, \varphi_2, \varphi_3; \varphi_5), W_4(\varphi_1, \varphi_2, \varphi_3; \varphi_6)\} = W_3^2(\varphi_1, \varphi_2, \varphi_3) \cdot W_6(\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5, \varphi_6). \tag{4.1}$$

In Fig. 2, we give an example of stepping down a basic core of four functions $\{\varphi_1, \varphi_2, \varphi_3, \varphi_4\}$ by three of them, say $\{\varphi_2, \varphi_3, \varphi_4\}$. Using Theorem 3 we have

$$W_3\{W_3(\varphi_1; \varphi_2, \varphi_3), W_3(\varphi_1; \varphi_2, \varphi_4), W_3(\varphi_1; \varphi_3, \varphi_4)\} = W_4^2(\varphi_1, \varphi_2, \varphi_3, \varphi_4) \cdot \varphi_1. \tag{4.2}$$

V. CONCLUDING REMARKS

In the introduction, using the antisymmetric pairs $(\varphi_1 \overleftrightarrow{\partial} \varphi_2)$ and $(\varphi_1 \overleftrightarrow{\partial} \varphi_3)$, we used essentially our stepping

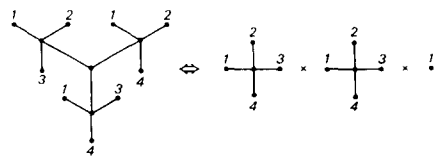


FIG. 2. The stepping down algorithm for Wronskians on four core functions dimenished by three of them.

up procedure on $W_1(\varphi_1) \equiv \varphi_1$ to obtain the antisymmetric $W_3(\varphi_1, \varphi_2, \varphi_3)$. However, with the other possible antisymmetric pair $(\varphi_2 \overleftrightarrow{\partial} \varphi_3)$, we can consider the three antisymmetric pairs as stepped-down versions of $W_3(\varphi_1, \varphi_2, \varphi_3)$ and Eq. (3.13) gives instead a totally symmetrized form:

$$W_3\{(\varphi_1 \overleftrightarrow{\partial} \varphi_2), (\varphi_1 \overleftrightarrow{\partial} \varphi_3), (\varphi_2 \overleftrightarrow{\partial} \varphi_3)\} = W_3{}^2(\varphi_1, \varphi_2, \varphi_3). \tag{5.1}$$

Theorems 1, 2, and 3 and the related graphical rules suggest countless ways of combining functions in order to obtain desired symmetries without leaving the family of Wronskians considered either as "operations" or as argument functions.

From a computational point of view these pyramidal algorithms can result in considerable simplification and time saving when one needs to calculate several Wronskians as with Green's functions. In fact, the whole family of Wronskians to any order can be generated with the sole definition of the operation $\overleftrightarrow{\partial} \equiv W_2$, and repeated use of (3.1). Even systems which are not of the categories described above are often conveniently reduced; thus,

$$\begin{aligned} &W_3\{(\varphi_1 \overleftrightarrow{\partial} \varphi_2), (\varphi_2 \overleftrightarrow{\partial} \varphi_3), (\varphi_3 \overleftrightarrow{\partial} \varphi_4)\} \\ &\quad = W_2^{-1}(\varphi_2, \varphi_3) \cdot W_2\{W_2[W_2(\varphi_1, \varphi_2), W_2(\varphi_2, \varphi_3)], \\ &\qquad W_2[W_2(\varphi_2, \varphi_3), W_2(\varphi_3, \varphi_4)]\} \\ &\quad = W_2^{-1}(\varphi_2, \varphi_3) \cdot W_2\{[\varphi_2 \cdot W_3(\varphi_1, \varphi_2, \varphi_3)], \\ &\qquad [\varphi_3 \cdot W_3(\varphi_2, \varphi_3, \varphi_4)]\} \\ &\quad = W_2^{-1}(\varphi_2, \varphi_3) \cdot \{\varphi_2 \cdot \varphi_3 \cdot W_2[W_3(\varphi_1, \varphi_2, \varphi_3), W_3(\varphi_2, \varphi_3, \varphi_4)] \\ &\qquad + W_2(\varphi_2, \varphi_3) \cdot W_3(\varphi_1, \varphi_2, \varphi_3) \cdot W_3(\varphi_2, \varphi_3, \varphi_4)\} \\ &\quad = \varphi_2 \cdot \varphi_3 \cdot W_4(\varphi_1, \varphi_2, \varphi_3, \varphi_4) + W_3(\varphi_1, \varphi_2, \varphi_3) \\ &\qquad \cdot W_3(\varphi_2, \varphi_3, \varphi_4). \end{aligned} \tag{5.2}$$

*Present address for the academic year 1973—74.

[1]E. Kamke, *Differentialgleichungen* (Academische Verlagsgesellschaft, Leipzig, 1967).
[2]K. Yosida, *Lectures on Differential and Integral Equations* (Interscience, New York, 1960).
[3]F.W. Byron and R.W. Fuller, *Mathematics of Classical and Quantum Physics* (Addison-Wesley, Reading, Massachusetts, 1970), Vol. II.
[4]H. Lehmann, K. Symanzik, and W. Zimmermann, Nuovo Cimento **1**, 205 (1955).
[5]V.I. Smirnov, *Linear Algebra and Group Theory* (McGraw-Hill, New York, 1961).

Extension of the Case formulas to L_p. Application to half and full space problems

E. W. Larsen

Courant Institute of Mathematical Sciences, New York University, New York, New York 10011

S. Sancaktar and P. F. Zweifel

Department of Physics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061

(Received 7 October 1974)

The singular eigenfunction expansions originally applied by Case to solutions of the transport equation are extended from the space of Hölder-continuous functions to the function spaces $X_p = \{f \mid \mu f(\mu) \in L_p\}$, where the expansions are now to be interpreted in the X_p norm. The spectral family for the transport operator is then obtained explicitly, and is used to solve transport problems with X_p sources and incident distributions.

I. INTRODUCTION

In 1960, Case[1] introduced a method of solving one-speed, one-dimensional neutron transport problems by expanding the solution in terms of "normal modes." The expansion coefficients were obtained from certain singular integral equations determined by the boundary conditions of the problem. Although the completeness theorems proved by Case were not rigorous from a mathematical point of view, his method achieved great popularity because of its close analogy with the classical method of solving boundary value problems by eigenfunction expansions.

Recently, a rigorous derivation of the Case solutions has been obtained[2,3] by considering the spectral resolution of an operator K, which, for isotropic scattering, is defined by

$$Kf(x,\mu)=\mu f(x,\mu)+\frac{c}{2(1-c)}\int_{-1}^{1} sf(x,s)\,ds. \tag{1}$$

(The independendent variables x and μ represent, respectively, the neutron position and the cosine of the angle between the neutron velocity vector and the x axis, and c is a positive constant $\neq 1$.) In terms of K, which acts only on μ, the transport equation can be written as

$$\frac{\partial}{\partial x}\psi(x,\mu)+K^{-1}\psi(x,\mu)=h(x,\mu), \tag{2}$$

where $h=\mu^{-1}g$ and g represents the neutron source.

A different approach to the rigorous determination of the Case formalism has been given by Hangelbroek,[4] who has shown for $c<1$ that the operator K ($=A^{-1}[\mu]$ in Hangelbroek notation) is topologically equivalent to a self-adjoint operator on the Hilbert space $L_2(-1,1)$. By well-known spectral theorems,[5] this guarantees the existence of a spectral family for K. (This family is abstractly obtained by application of the Gel'fand–Naimark theorem after a commutative C^*-algebra is generated from K and the identity I.)

In all of these works, i.e., Refs. 1,2,4, final attention is restricted to Hölder-continuous (or piecewise Hölder-continuous) functions since the explicit formulas involve principle value integrals and boundary values of Cauchy integrals. The purpose of the present note is to show how these results can be extended to a much larger class of functions, namely, the spaces $X_p = \{f \mid \mu f \in L_p,\ p>1\}$. We do this by extending the results of Ref. 1 to these spaces. We also show that the expansion formulas of that reference can be used to construct the spectral family for the operator K. Finally, we indicate how this spectral family may be used to solve typical transport problems and suggest possible applications in other areas.

II. EXTENSION OF THE CASE FORMULAS

Our first step is to quote a theorem which will guarantee that an integral operator A: $L_p \to L_p$ of the form

$$Af(x)=\int_{-\infty}^{\infty}\frac{f(t)}{t-x}\,dt \equiv g(x) \tag{3}$$

is a bijection. This theorem is crucial for all the subsequent analysis.

Theorem 0[6]: Let $f(x)\in L_p(-\infty,\infty)$, $p>1$. Then the formula

$$g(x)=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{f(t)}{t-x}\,dt \tag{4}$$

defines almost everywhere a function $g(x)$ also belonging to $L_p(-\infty,\infty)$. The reciprocal formula

$$f(x)=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{g(t)}{t-x}\,dt \tag{5}$$

also holds almost everywhere and

$$\int_{-\infty}^{\infty}|g(x)|^p\,dx \le (M_p)^p\int_{-\infty}^{\infty}|f(x)|^p\,dx, \tag{6}$$

where M_p depends on p only. If $p=2$, then

$$\int_{-\infty}^{\infty}|g(x)|^2\,dx=\int_{-\infty}^{\infty}|f(x)|^2\,dx.$$

Now let us consider the formulas[1,2]

$$f'(\mu)=\int_{-1}^{1}A(\nu)\phi(\nu,\mu)\,d\nu+A(\nu_0)\phi(\nu_0,\mu)+A(-\nu_0)\phi(-\nu_0,\mu), \tag{7}$$

$$A(\nu)=\frac{1}{N(\nu)}\int_{-1}^{1}\mu f'(\mu)\phi(\nu,\mu)\,d\mu, \tag{8}$$

which hold for Hölder-continuous f' and which we shall refer to as Case transforms. We can simplify notation by defining

Copyright © 1975 American Institute of Physics

$$f(\mu)=f'(\mu)-A(\nu_0)\phi(\nu_0,\mu)-A(-\nu_0)\phi(-\nu_0,\mu); \tag{9}$$

then Eqs. (7) and (8) become

$$f(\mu)=\int_{-1}^{1}A(\nu)\phi(\nu,\mu)\,d\nu, \quad -1<\mu<1, \tag{10}$$

$$A(\nu)=\frac{1}{N(\nu)}\int_{-1}^{1}\mu f(\mu)\phi(\nu,\mu)\,d\mu, \quad -1<\nu<1. \tag{11}$$

We wish to show that Eqs. (7) and (8) are valid for $\mu f'(\mu)\in L_p(-1,1)$, $p>1$. Since the discrete parts $A(\pm\nu_0)\phi(\pm\nu_0,\mu)$ are in L_p, then it suffices to show that (10) and (11) hold for functions f (from which the contribution of the discrete modes has been subtracted out).

Lemma 1: For each

$$f\in X_p, X_p=\{f\mid \|f\|_p\equiv(\int_{-1}^{1}|\mu f(\mu)|^p\,d\mu)^{1/p}<\infty\}, \quad p>1, \tag{12}$$

there is a corresponding $A(\nu)\in X_p$ defined by Eq. (11), and $A(\nu)$ depends continuously on f.

Proof: Using the definition

$$\phi_\nu(\nu,\mu)\equiv\frac{c\nu}{2}\frac{1}{\nu-\mu}+\delta(\nu-\mu)\lambda(\nu) \tag{13}$$

and the Case transform, we obtain

$$\nu A(\nu)=\frac{\nu}{N(\nu)}\left(\frac{c\nu}{2}\int_{-1}^{1}\frac{\mu f(\mu)}{\nu-\mu}d\mu+\int_{-1}^{1}\mu f(\mu)\lambda(\nu)\delta(\nu-\mu)\,d\mu\right), \tag{14}$$

which we consider as a formal abbreviation for

$$\nu A(\nu)=\frac{\lambda(\nu)}{\Lambda^+(\nu)\Lambda^-(\nu)}\nu f(\nu)+\frac{c}{2}\frac{\nu}{\Lambda^+(\nu)\Lambda^-(\nu)}\int_{-1}^{1}\frac{\mu f(\mu)}{\nu-\mu}\,d\mu. \tag{15}$$

In these equations, the expression

$$N=\nu\Lambda^+(\nu)\Lambda^-(\nu) \tag{16}$$

has been utilized.

The functions $\lambda(\nu)/\Lambda^+(\nu)\Lambda^-(\nu)$ and $\nu/\Lambda^+(\nu)\Lambda^-(\nu)$ are continuous and bounded on $[-1,1]$.[1] Thus the first term in Eq. (15) is in $L_p(-1,1)$ and the second term is in $L_p(-1,1)$ if

$$g(\nu)=\frac{1}{\pi}\int_{-1}^{1}\frac{\mu f(\mu)}{\nu-\mu}d\mu \tag{17}$$

is in $L_p(-1,1)$. But by Theorem 0,

$$\int_{-1}^{1}|g(\nu)|^p\,d\nu\le\int_{-\infty}^{\infty}|g(\nu)|^p\,d\nu\le(M_p)^p\int_{-1}^{1}|\mu f(\mu)|^p\,d\mu. \tag{18}$$

Thus from Eq. (15), it is clear that

$$\int_{-1}^{1}|\nu A(\nu)|^p\,d\nu\le N_p^p\int_{-1}^{1}|\mu f(\mu)|^p\,d\mu, \tag{19}$$

so that

$$\|A\|_p\le N_p\|f\|_p. \tag{20}$$

Let us define the map T: $X_p\to X_p$ by

$$A(\nu)=(Tf)(\mu); \tag{21}$$

then

$$\|T\|_p\le N_p. \tag{22}$$ ∎

Note: Using Eq. (15), one can show that if one multiplies by $\lambda(\nu)$, then

$$\int_{-1}^{1}|\nu\lambda(\nu)A(\nu)|^p\,d\nu\le\hat{N}_p^p\int_{-1}^{1}|\mu f(\mu)|^p\,d\mu, \tag{23}$$

where $\hat{N}_p$ is a constant depending only on p.

Thus, if

$$\lambda(\nu)A(\nu)=(\hat{T}f)(\mu), \tag{24}$$

then

$$\|\hat{T}\|_p\le\hat{N}_p. \tag{25}$$

Lemma 2: For each $A(\nu)$ such that $A(\nu)$ and $\lambda(\nu)A(\nu)\in X_p$, there exists an $f\in X_p$ defined by Eq. (10).

Proof: By definition

$$\int_{-1}^{1}|\nu A(\nu)|^p\,d\nu<\infty \tag{26}$$

and

$$\int_{-1}^{1}|\nu\lambda(\nu)A(\nu)|^p\,d\nu<\infty. \tag{27}$$

By Eq. (10) we define the function f by

$$f(\mu)\equiv\lambda(\nu)A(\nu)+\frac{c}{2}\int_{-1}^{1}\frac{\nu A(\nu)}{\nu-\mu}\,d\nu. \tag{28}$$

Clearly the first term is in X_p, and so is the second by application of Theorem 0. ∎

Also, if we have a sequence $\{A_n\}$ such that $\|A_n-A\|_p\to 0$ and $\|\lambda(\nu)A_n(\nu)-\lambda(\nu)A(\nu)\|_p\to 0$, then it is obvious that $\|f_n-f\|_p\to 0$.

It is known[2,7] that for $\mu f(\mu)$ Hölder-continuous [and f of the form of Eq. (9)], Eqs. (10) and (11) hold simultaneously. Since the Hölder-continuous functions are dense in L_p, let us choose a sequence $\{f_n\}$ such that $\mu f_n(\mu)$ is Hölder-continuous and $f_n\to f\in X_p$. Then $A_n\to A$ and $\lambda(\nu)A_n\to\lambda(\nu)A$ by Eqs. (22)—(25). Thus by the above paragraph, Eq. (10) holds in the limit.

The above results can be summarized by a lemma:

Lemma 3: Eqs. (10) and (11) hold for any $f\in X_p$ which is of the form of Eq. (9). We can combine Lemmas 1, 2, and 3 as the following theorem:

Theorem I: The domain of the reduced transport operator K may be extended to the spaces X_p, $p>1$, and the Case transforms (7), (8) hold for each f' such that $\mu f'(\mu)\in L_p(-1,1)$.

3. RESOLUTION OF IDENTITY OF K

For $-1<\omega<1$, we define the operator $E(\omega)$ as

$$E(\omega)f(\mu)=\int_{-1}^{\omega}A(\nu)\phi(\nu,\mu)\,d\nu \tag{29}$$

$$=\begin{cases}A(\mu)\lambda(\mu)+\dfrac{c}{2}\displaystyle\int_{-1}^{\omega}\frac{\nu A(\nu)}{\nu-\mu}d\nu, & -1<\mu\le\omega,\\[2ex] \dfrac{c}{2}\displaystyle\int_{-1}^{\omega}\frac{\nu A(\nu)}{\nu-\mu}d\nu, & \omega<\mu<1.\end{cases} \tag{30}$$

From the above analysis, it is clear that the terms in Eq. (30) represent bounded operators on X_p, acting on f. Thus for $-1<\omega<1$, $E(\omega)$ is a bounded operator. In this section we shall show that the family $E(\omega)$ forms part of the spectral family of K.

First, for $\epsilon>0$ we have

$$[E(\omega+\epsilon)-E(\omega)]f(\mu)=\int_{\omega}^{\omega+\epsilon}A(\nu)\phi(\nu,\mu)\,d\nu \tag{31}$$

$$= \left\{ \begin{array}{ll} \lambda(\mu)A(\mu), & \omega \le \mu \le \omega+\epsilon, \\ 0, & \text{otherwise}, \end{array} \right\} + \frac{c}{2}\int_{\omega}^{\omega+\epsilon} \frac{\nu A(\nu)}{\nu-\mu}\,d\nu. \quad (32)$$

The norm of the first term is just

$$\left[\int_{\omega}^{\omega+\epsilon} |\mu\lambda(\mu)A(\mu)|^p\,d\mu\right]^{1/p}$$

which tends to zero for a fixed A (i.e., fixed f) as $\epsilon \to 0$. Call

$$g(\mu,\epsilon) = \frac{1}{\pi}\int_{\omega}^{\omega+\epsilon} \frac{\nu A(\nu)}{\nu-\mu}\,d\nu. \quad (33)$$

Then by Theorem 0,

$$\|g(\mu,\epsilon)\|^p = \int_{-1}^{1} |\mu g(\mu,\epsilon)|^p\,d\mu \le \int_{-\infty}^{\infty} |g(\mu,\epsilon)|^p\,d\mu$$
$$\le M_p^p \int_{\omega}^{\omega+\epsilon} |\nu A(\nu)|^p\,d\nu, \quad (34)$$

and this term also approaches 0 for a fixed f as $\epsilon \to 0$. Thus we can state the result as a lemma.

Lemma 4: For each $f \in X_p$ and $-1 \le \omega < 1$,

$$\lim_{\substack{\epsilon\to 0 \\ \epsilon>0}} \|E(\omega+\epsilon) - E(\omega)f\|_p = 0. \quad (35)$$

That is, $E(\omega)$ is a continuous function of ω (in the strong operator topology). Now, we will verify the following lemma.

Lemma 5:

$$E(\omega_1)E(\omega_2) = E(\omega_2)E(\omega_1) = E(\omega) \quad (36)$$

where $\omega = \min\{\omega_1, \omega_2\}$.

Proof: For definiteness, we consider

$$\omega_1 \le \omega_2,$$
$$E(\omega_2)f(\mu) = \int_{-1}^{\omega_2} A(\nu)\phi(\nu,\mu)\,d\nu. \quad (37)$$

Then the expansion coefficients of $E(\omega_2)f(\mu)$ are

$$B(\nu) = \left\{ \begin{array}{ll} A(\nu), & -1 \le \nu \le \omega_2, \\ 0, & \omega_2 < \nu \le 1. \end{array} \right. \quad (38)$$

Thus,

$$E(\omega_1)E(\omega_2)f(\mu) = \int_{-1}^{\omega_1} B(\nu)\phi(\nu,\mu)\,d\nu$$
$$= \int_{-1}^{\omega_1} A(\nu)\phi(\nu,\mu)\,d\nu$$
$$= E(\omega_1)f(\mu). \quad (39)$$

In a similar way, we obtain $E(\omega_2)E(\omega_1)f = E(\omega_1)f$. ∎

Before proving Lemma 7, we will prove the following lemma which is essential in proving Lemma 7.

Lemma 6:

$$Kf = \int_{-1}^{1} zA(z)\phi(z,\mu)\,dz. \quad (40)$$

Proof: Applying the same method as in Ref. 2, we get for Γ a simple closed curve containing the line segment $[-1,1]$.

$$Kf(\mu) = \frac{1}{2\pi i}\oint_{\Gamma} K(z-K)^{-1}f(\mu)\,dz$$
$$= \frac{1}{2\pi i}\oint_{\Gamma} (K-z+z)(z-K)^{-1}f(\mu)\,dz$$
$$= \frac{1}{2\pi i}\oint_{\Gamma} (z(z-K)^{-1} - I)f(\mu)\,dz$$
$$= \frac{1}{2\pi i}\oint_{\Gamma} z(z-K)^{-1}f(\mu)\,d\mu. \quad (41)$$

We now compute the contour integral exactly as in Ref. 2 to obtain

$$Kf(\mu) = \int_{-1}^{1} zA(z)\phi(z,\mu)\,dz. \quad \blacksquare \quad (42)$$

Lemma 7: For each $f \in X_p$ of the form (9), $Kf = \int_{-1}^{1} \omega\,dE(\omega)f$.

Proof: Let

$$U(\omega) = \int_{-1}^{1} (E(\omega)f(\mu))g(\mu)\,d\mu, \quad (43)$$

where $f \in X_p$ and $g \in X_q$ (the dual space of X_p: $1/p + 1/q = 1$). Then

$$U(\omega) = \int_{-1}^{\omega} A(\mu)\lambda(\mu)g(\mu)\,d\mu + \int_{-1}^{1} d\mu \int_{-1}^{\omega} d\nu \tfrac{1}{2}c\nu \frac{A(\nu)}{\nu-\mu}g(\mu)$$
$$= \int_{-1}^{\omega} A(\mu)\lambda(\mu)g(\mu)\,d\mu + \int_{-1}^{\omega} d\nu \tfrac{1}{2}c\nu A(\nu)\int_{-1}^{1} \frac{g(\mu)}{\nu-\mu}\,d\mu$$
$$\equiv \int_{-1}^{\omega} A(\mu)\lambda(\mu)g(\mu)\,d\mu + \int_{-1}^{\omega} \tfrac{1}{2}c\mu A(\mu)Lg(\mu)\,d\mu$$
$$= \int_{-1}^{\omega} A(\mu)[\lambda(\mu)g(\mu) + \tfrac{1}{2}c\mu Lg(\mu)]\,d\mu. \quad (44)$$

Thus $U(\omega)$ is differentiable a.e. and

$$U'(\omega) = A(\omega)[\lambda(\omega)g(\omega) + \tfrac{1}{2}c\omega Lg(\omega)]. \quad (45)$$

Now we get

$$\int_{-1}^{1} \omega\,dU(\omega) = \int_{-1}^{1} \omega U'(\omega)\,d\omega$$
$$= \int_{-1}^{1} \omega A(\omega)[\lambda(\omega)g(\omega) + \tfrac{1}{2}c\omega Lg(\omega)]\,d\omega$$
$$= \int_{-1}^{1} \omega A(\omega)\lambda(\omega)g(\omega)$$
$$+ \int_{-1}^{1} \omega A(\omega)\frac{c\omega}{2}\int_{-1}^{1} \frac{g(\mu)}{\omega-\mu}\,d\mu\,d\omega$$
$$= \int_{-1}^{1}\left[\omega A(\omega)\lambda(\omega) + \int_{-1}^{1} \nu A(\nu)\frac{c\nu}{2}\frac{d\nu}{\nu-\mu}\right]g(\omega)\,d\omega. \quad (46)$$

By the previous lemma,

$$\int_{-1}^{1} \omega\,dU(\omega) = \int_{-1}^{1} Kf(\omega)g(\omega)\,d\omega.$$

Thus

$$\int_{-1}^{1} Kf(\mu)g(\mu)\,d\mu = \int_{-1}^{1} \omega\,dU(\omega) \quad (47)$$
$$= \int_{\omega=-1}^{1} \omega\,d\int_{\mu=-1}^{1} [E(\omega)f](\mu)g(\mu)\,d\mu$$
$$= \int_{-1}^{1} d\mu g(\mu)\int_{\omega=-1}^{1} \omega\,d[E(\omega)f](\mu), \quad (48)$$

where the interchange of limits is justified because ω and $E(\omega)f$ are continuous in ω. The above equation implies

$$Kf(\mu)=\int_{-1}^{1}\omega\, d[E(\omega)f](\mu), \tag{49}$$

or

$$K=\int_{-1}^{1}\omega\, dE(\omega),$$

where the integral is defined in the weak sense. ∎

Note:

$$\int_{-1}^{1}\omega\, dE(\omega)f=\int_{-1}^{1}\omega\frac{dE(\omega)}{d\omega}\,d\omega f$$

$$=\int_{-1}^{1}\omega A(\omega)\phi(\omega,\mu)\,d\omega=Kf \tag{50}$$

gives the result directly, but only formally.

Lemma 8: $KE(\omega)=E(\omega)K.$ (51)

Proof: By Lemma 6,

$$Kf=\int_{-1}^{1}zA(z)\phi(z,\mu)\,dz. \tag{52}$$

Thus by Eq. (29),

$$E(\omega)Kf(\mu)=\int_{-1}^{\omega}\nu A(\nu)\phi(\nu,\mu)\,d\nu \tag{53}$$

In a similar way we obtain

$$K(Ef)=\int_{-1}^{\omega}zA(z)\phi(z,\mu)\,dz. \tag{54}$$

Equality of Eqs. (53) and (54) proves the lemma. ∎

Also, the following identities hold by definition:

$$E(-1)=0, \tag{55}$$

$$E(1)=1. \tag{56}$$

Lemmas 4, 5, 7, and 8, together with Eqs. (55) and (56) complete the proof of the following theorem[5]:

Theorem II: For $\mu f'(\mu)\in L_p(-1,1)$, we define

$$E(\pm\nu_0)f'(\mu)=\frac{1}{N(\pm\nu_0)}\int_{-1}^{1}sf'(s)\phi(\pm\nu_0,s)\,ds\,\phi(\pm\nu_0,\mu).$$

For $\nu\in[-1,1]$, we let $E(\nu)$ be defined by Eq. (29). Then

$$K^nf'(\mu)=\nu_0^nE(\nu_0)f'(\mu)+(-\nu_0)^nE(-\nu_0)f'(\mu)$$

$$+\int_{-1}^{1}\nu^n\,dE(\nu)f(\mu),$$

and $E(\nu)$ is the spectral family of projection operators for the operator K.

IV. HALF-RANGE THEORY

Let $\mu\psi_0(\mu)$ be defined and Hölder-continuous on $0\leq\mu\leq1$. Define[3]

$$\psi_e(\mu)=\begin{cases}\psi_0(\mu), & 0<\mu\leq1,\\ \int_0^1 J(\mu,s)\psi_0(s)\,ds, & -1\leq\mu<0,\end{cases} \tag{57}$$

where

$$J(\mu,s)=\frac{1}{X(\mu)X(-s)}\frac{cs}{2}\frac{1}{s-\mu}, \tag{58}$$

and

$$X(z)=\Lambda^{1/2}(\infty)(z-\nu_0)\exp\left(\frac{1}{2\pi i}\int_0^1\ln\frac{\Lambda^+(s)}{\Lambda^-(s)}\frac{ds}{s-z}\right), \tag{59}$$

$$\Lambda(z)=X(z)X(-z). \tag{60}$$

Then the full range coefficients $A(\nu)$ of $\psi_e(\mu)$ are zero for $\nu<0$;

$$\psi_e(\mu)=\int_0^1A(\nu)\phi(\nu,\mu)\,d\nu+A(\nu_0)\phi(\nu_0,\mu). \tag{61}$$

Extending $\mu\psi_0$ to $L_p(0,1)$, we obtain by continuity the extension ψ_e of ψ_0, and $\psi_e\in X_p$, also.

By definition of $E(\omega)$, we have for $-1\leq\omega\leq1$,

$$E(\omega)\psi_e(\mu)=\int_{-1}^{\omega}A(\nu)\phi(\nu,\mu)\,d\nu$$

$$=\begin{cases}\int_0^{\omega}A(\nu)\phi(\nu,\mu)\,d\nu, & 0<\omega\leq1,\\ 0, & -1\leq\omega\leq0.\end{cases} \tag{62}$$

Thus

$$\psi_e=\int_0^1dE(\omega)\psi_e(\mu)+E(\nu_0)\phi(\nu_0,\mu). \tag{63}$$

For $0<\mu<1$, this reduces to

$$\psi_0=\int_0^1dE(\omega)\psi_e(\mu)+E(\nu_0)\phi(\nu_0,\mu), \tag{64}$$

which is just a statement of the half-range completeness theorem.

V. SOLUTIONS OF TRANSPORT PROBLEMS

Consider the problem

$$\frac{\partial\psi}{\partial x}+K^{-1}\psi=0, \qquad x>0,$$
$$\psi(0,\mu)=\psi_0(\mu), \quad \mu\psi_0\in L_p(0,1), \quad 0<\mu\leq1. \tag{65}$$

Letting $\psi_e(\mu)$ be the extension of ψ_0 described in Eq. (57), we claim that the solution of this problem is

$$\psi(x,\mu)=\int_0^1\exp(-x/\nu)\,d[E(\nu)\psi_e(\mu)]. \tag{66}$$

This function satisfies the boundary conditions, and it also satisfies the transport equation, as can be seen by inspection.

Consider next the problem

$$\frac{\partial\psi}{\partial x}+K^{-1}\psi=q_0(x,\mu), \qquad x_0<x<x_1,$$
$$q_0(x,\mu)=q(x,\mu)/\mu\in X_p, \quad p>1. \tag{67}$$

We will look only for a particular solution. Boundary conditions can be met by using solutions of the homogeneous equation.

We look for a particular solution of the form

$$\psi(x,\mu)=\int_{-1}^{1}d[E(\nu)\psi(x,\mu,\nu)]. \tag{68}$$

We have the identity

$$q_0(x,\mu)=\int_{-1}^{1}d[E(\nu)q_0(x,\mu)]. \tag{69}$$

Inserting Eqs. (67) and (69) in Eq. (70), we obtain

$$\int_{-1}^{1}d[E(\nu)q_0(x,\mu)]=\int_{-1}^{1}dE(\nu)\left[\frac{\partial\psi}{\partial x}(x,\mu,\nu)+\frac{1}{\nu}\psi(x,\mu,\nu)\right]. \tag{70}$$

This is solved by taking

$$q_0(x,\mu)=\frac{\partial\psi}{\partial x}(x,\mu,\nu)+\frac{1}{\nu}\psi(x,\mu,\nu), \tag{71}$$

or

$$\frac{d}{dx}\exp(x/\nu)\psi(x,\mu,\nu)=\exp(x/\nu)q_0(x,\mu). \tag{72}$$

For $\nu > 0$, we integrate from x_0 to x to get

$$\exp(x/\nu)\psi(x,\mu,\nu) - \exp(x_0/\nu)\psi(x_0,\mu,\nu) = \int_{x_0}^{x} \exp(s/\nu) q_0(s,\mu)\, ds, \tag{73}$$

so that

$$\psi(x,\mu,\nu) = \exp[(x_0 - x)/\nu]\psi(x_0,\mu,\nu) + \int_{x_0}^{x} \exp[(s-x)/\nu] q_0(s,\mu)\, ds. \tag{74}$$

For $\nu < 0$, we integrate from x_1 to x to get a similar equation:

$$\psi(x,\mu,\nu) = \exp[(x_1 - x)/\nu]\psi(x_1,\mu,\nu) + \int_{x_1}^{x} \exp[(s-x)/\nu] q_0(s,\mu)\, ds \tag{75}$$

The general solution of Eq. (67) which is bounded is then given by Eq. (68), where ψ is defined in Eqs. (74) and (75), and $\psi(x_i,\mu,\nu)$ is arbitrary. A particular solution is obtained by setting $\psi(x_i,\mu,\nu) = 0$.

VI. DISCUSSION

The Case transforms, Eqs. (7) and (8) were originally derived for Hölder-continuous functions f', and we have shown that they can be extended to functions $f' \in X_p$. Furthermore, we have constructed the spectral family for K in each of the X_p spaces, $p > 1$, and we have shown how to use this spectral family to solve typical problems. Several aspects of these results seem worthy of further comment.

First, the conditions that $f'(\mu)$ be in L_p means that the Case formulas hold for functions f' which can be highly singular at $\mu = 0$. However, this feature was shown in Sec. V to be essential in solving problems with sources, since the modified source $q_0 = \mu^{-1} q$ had to be written as a "full-range" expansion. Second, the removal of the unphysical Hölder condition on f' constitutes an obvious generalization. We emphasize that for Hölder-continuous f', Eqs. (7) and (8) hold pointwise for each μ, while, for $f' \in X_p$, Eqs. (7) and (8) hold only in the (integral) norm of X_p.

The existence of a spectral family for K was established by Hangelbroek[4] for L_2 and $c < 1$ by showing that K is topologically equivalent to a self-adjoint operator. Our results show that this family exists not only for L_2 and $c < 1$, but for the much larger spaces X_p, $p > 1$, and for any value of $c > 0$.

This physically natural space for the transport operator is X_1, but in this space interesting mathematical difficulties occur. For $p = 1$, Theorem 0 is no longer true; in fact the principal value operator is definitely unbounded.[6] Also the projection operators $E(\omega)$ for $-1 < \omega < 1$ are unbounded, so formulas such as Eq. (42) cannot hold in the usual Stieltjes sense.

We shall not consider here the problem of generalizing the Case formulas to X_1. Instead, we refer the reader to Ref. 7, in which this problem has been solved by methods different from those introduced here.

Since the solutions of the transport equations obtained in Ref. 7 include those obtained in Sec. V of the present paper as a special case, we should comment on the applicability of the present results. First, convenient orthogonality relations such as Eq. (11) and a corresponding formula for the half-range expansion coefficient are available in the Case approach, but not in the method used in Ref. 7. Second, the completeness and spectral properties of the Case eigenfunctions or closely allied techniques have been applied to a number of problems independent of simply solving the transport equation. For example, the solution of the inverse problem,[8] the rigorous derivation of the equation of invariant imbedding,[9] the justification of the multiple scattering expansion,[10] the solution of the transport equation in cylindrical or spherical geometries,[11] as well as a whole host of problems involving the solution of equations which have arisen in a number of astronomical, physical or engineering applications.[11] It is hoped that generalizing the class of functions which may be expanded in Case eigenfunctions may increase the number of possible applications such as those enumerated above.

Finally, we point out that it is of some mathematical interest to demonstrate the spectral family for an operator (not normal) for which a general spectral theorem has not been proved. Other such operators may be treated by methods similar to those used here.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation Grant No. GK-35903 at Virginia Polytechnic Institute, and by the Air Force Office of Scientific Research under Grant No. AFOSR-71-2013 at Courant Institute of Mathematical Sciences, New York University.

[1] K. M. Case, Ann. Phys. (N.Y.) 9, 1 (1960).
[2] E. W. Larsen and G. J. Habetler, Commun. Pure Appl. Math. 26, 525 (1973).
[3] E. W. Larsen, Commun. Pure Appl. Math. 27, 523 (1974).
[4] R. Hangelbroek, "A Functional Analytic Approach to the Linear Transport Equation," thesis, University of Groningen, Netherlands, 1973.
[5] G. Bachman and L. Narici, *Functional Analysis* (Academic, New York, 1960).
[6] E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals* (Oxford U. P., London, 1950).
[7] E. W. Larsen, "Solutions of Neutron Transport Problems in L_1," to appear in J. Math. Anal. Appl.
[8] K. M. Case, Phys. Fluids 16, 1607 (1973); E. W. Larsen, "The Inverse Source Problem in Radiative Transfer," to appear in J. Quant. Spectrosc. Radiat. Transfer.
[9] S. Pahor and P. F. Zweifel, J. Math. Phys. 10, 581 (1969).
[10] Madhoo Kanal, J. Math. Phys. 14, 888 (1973).
[11] K. M. Case and P. F. Zweifel, *Linear Transport Theory* (Addison-Wesley, Reading, Mass., 1967).

Global existence of solutions to the Cauchy problem for time-dependent Hartree equations

J. M. Chadam* and R. T. Glassey*

Department of Mathematics, Indiana University, Bloomington, Indiana 47401
(Received 9 September 1974)

The existence of global solutions to the Cauchy problem for time-dependent Hartree equations for N electrons is established. The solution is shown to have a uniformly bounded $H^1(\mathbf{R}^3)$ norm and to satisfy an estimate of the form $\|\psi(t)\|_{H^1} \leq c \exp(kt)$. It is shown that "negative energy" solutions do not converge uniformly to zero as $t \to \infty$.

1. INTRODUCTION

We consider the existence of global solutions to the Cauchy problem for the equations

$$i\frac{\partial\psi}{\partial t} = \frac{1}{2}\Delta\psi + \frac{2}{|x|}\psi - v(\psi)\psi \tag{1.1}$$

and

$$i\frac{\partial\psi_j}{\partial t} = \frac{1}{2}\Delta\psi_j + \frac{Z}{|x|}\psi_j - V_{0p}\psi_j \quad (j=1,2,\ldots,N). \tag{1.2}$$

(1.1) is the time-dependent version of the Hartree equation for the helium atom, and (1.2) is the corresponding equation for an N-electron system. In (1.1), $v=v(\psi)$ is the Newtonian potential of the charge density $\bar{\psi}\psi = |\psi|^2$:

$$v = \int_{\mathbb{R}^3} \frac{|\psi(y,t)|^2\,dy}{|x-y|}. \tag{1.3}$$

The notation in V_{0p} in (1.2) denotes an operator whose action on the function $\psi_j(x,t)$ is

$$v_{0p}\psi_j(x,t) = \sum_{k=1}^{N}\left(\psi_j(x,t)\int_{\mathbb{R}^3}\frac{|\psi_k(y,t)|^2}{|x-y|}\,dy - \psi_k(x,t)\int_{\mathbb{R}^2}\frac{\bar{\psi}_k(y,t)\psi_j(y,t)\,dy}{|x-y|}\right). \tag{1.4}$$

The "interchange" operation is to be noted here; that is, in the second term of (1.4) ψ_j has been "brought under the integral sign." Also, in (1.2), Z is a positive integer, the atomic number.

Stationary solutions of these equations are of great interest. Here one asks for solutions of the form

$$\psi(x,t) = \exp(i\lambda t)\Psi(x) \quad \text{for (1.1)} \quad (\lambda\in\mathbb{R})$$

or

$$\psi_j(x,t) = \exp(i\lambda_j t)\,\Psi_j(x) \quad \text{for} \quad (1.2) \quad (\lambda_j\in\mathbb{R}).$$

The resulting nonlinear elliptic equation for Ψ has been studied by Reeken[1] and Sather and Gustafson[2]; for a physical derivation see Ref. 3. Wolkowisky[4] considers stationary solutions of a generalized form of Eq. (1.1), and Lieb and Simon[5] have treated the corresponding problem for (1.2).

Our interest in these equations has arisen from a problem in relativistic quantum mechanics. Specifically, we have shown in Ref. 6 that the coupled Maxwell–Dirac equations, in the case of zero magnetic field ($\mathbf{H}\equiv\operatorname{curl}\mathbf{A}=0$) can be reduced to a single partial differential equation

$$\frac{\partial\psi}{\partial t} = \sum_{k=1}^{3}\gamma^k\gamma^0\frac{\partial\psi}{\partial x_k} - iM\gamma^0\psi + \frac{ig^2}{4\pi}v(\psi)\psi$$

for the spinor function ψ, where $v(\psi)$ has precisely the form (1.3) and the γ^k's denote the Dirac matrices. In Ref. 6 we have shown that the above equation has (an appropriately defined) global solution in two space dimensions (in which case $-v=\log r * |\psi|^2$), but the three-dimensional problem remains unsolved, thus motivating the present work.

In Sec. 2 we shall treat the local existence problem for (1.1), (1.2). The nonlinear term $v(\psi)\psi$ (or $v_{0p}\psi_j$) is Lipschitz in an appropriate manner, but the presence of the singular coefficient $1/|x|$ complicates the situation.

In Sec. 3 we prove that these local solutions exist globally, by obtaining uniform bounds on the norms

$$\|\nabla\psi_j(t)\|_{L^2(\mathbb{R}^3)}.$$

This is accomplished by finding, after a very tedious calculation, an invariant, the "energy." In addition, a bound of the form $\|\psi(t)\|_{H^2}$ and $\exp(kt)$ is established. The situation arising from applying a periodic driving term $f(t)\psi$ [$f(t)\psi_j$] to Eq. (1.1) [(1.2) resp.] is discussed at the end of Sec. 3.

We conclude in Sec. 4 by showing that "negative energy" solutions do not decay uniformly to zero as $t\to\infty$.

We shall employ the usual notation for L^p norms

$$\|u\|_p = \left[\int |u(x)|^p\,dx\right]^{1/p}$$

and for Sobolev norms

$$\|u\|_{H^m} = \left(\sum_{|k|\leq m}\int |D^k u(x)|^2\,dx\right)^{1/2}$$

where H^m is the space of functions whose derivatives up through order m lie in $L^2(\mathbb{R}^3)$. Furthermore, $f * g(x)$ will denote the spatial convolution of two functions defined on $\mathbb{R}^3$.

2. LOCAL EXISTENCE

In this paper, when we speak of a solution to the Cauchy problem for Eqs. (1.1) and (1.2), we shall mean a solution of the t-ingegrated form of these equations. Specifically, if we denote the Coulomb Hamiltonian $-\Delta/2 - 2/|x|$ by H_c, then a finite-energy solution of Eq. (1.1) over the time interval $[t_0, T)$ is a continuous map $t\to\psi(t)$: $[t_0, T)\to H^1$ which satisfies

 Copyright © 1975 American Institute of Physics

$$\psi(t)=\psi^{c}(t)+i\int_{t_0}^{t}\exp[iH_c(t-s)]v(\psi(s))\psi(s)\,ds, \quad (1.1')$$

where $\psi^{c}(t)=\exp(iH_c t)\psi_0$ is the solution of the Coulomb equation with the same Cauchy data and the integral is to be interpreted in the strong Riemann sense in H^1 [similarly for Eq. (1.2)]. The treatment of integral equations such as (1.1'), modelled after the Picard theory for ordinary differential equations, is well established (cf. Segal[7]). In particular then we shall prove in this section that solutions of (1.1') and (1.2') exist locally {i.e., for a small enough interval $[t_0, T)$} by showing that the nonlinearity is locally Lipschitzian in $H^1(\mathbb{R}^3)$. This coupled with the proof in Sec. 3 that the H^1 norm of the solution remains finite for each t shows that T can be taken to be $+\infty$. As mentioned in the Introduction, the Coulomb term ($2\psi/|x|$ or $z\psi_j/|x|$) is not treated as part of the perturbation because of the complications which arise when one tries to apply the standard nonlinear techniques to it. For this reason it is included in the "free" part of the equation and handled with the well-known linear theory.[8] We begin with a technical result which effectively permits the Coulomb Hamiltonian to be treated using the Sobolev space H^1.

Lemma 2.1: For each z there is a γ_z such that $-\Delta/2 - z/|x| + \gamma_z$ is a positive self-adjoint operator with domain $= D(-\Delta)$. Thus $(-\Delta/2 - z/|x| + \gamma_z)^{1/2}$ exists as a positive self-adjoint operator with domain $D((-\Delta)^{1/2})$. The norm $\|(-\Delta/2 - z/|x| + \gamma_z)^{1/2}\cdot\|$ is equivalent to the H^1 norm.

Proof: If γ_z is chosen $\geq (z/2)^2/\frac{1}{2} = z^2/2$ (= – the lowest eigenvalue of $-\Delta/2 - z/|x|$), the first two statements are standard results (Ref. 8, pp. 410 and 281), which are included for completeness. That $\|(-\Delta/2 - z/|x| + \gamma_z)^{1/2}\|$ is a norm is a direct consequence of the second statement. Finally if $\psi \in D(-\Delta) = D(-\Delta/2 - z/|x| + \gamma_z)$, then

$$\begin{aligned}
&\left(\frac{-\Delta}{2} - \frac{z}{|x|} + \gamma_z\right)\psi, \psi\Big) \\
&= \left(\left(\frac{-\Delta}{4} - \frac{z}{|x|} + \gamma_z\right)\psi, \psi\right) + \left(\frac{-\Delta\psi}{4}, \psi\right) \\
&\geq \left(\left(-\frac{(z/2)^2}{\frac{1}{4}} + \gamma_z\right)\psi, \psi\right) + \left(\frac{-\Delta\psi}{4}, \psi\right) \\
&\geq (\tfrac{1}{4}\psi, \psi) + \left(\frac{-\Delta\psi}{4}, \psi\right) = (\tfrac{1}{4}(1-\Delta)\psi, \psi)
\end{aligned}$$

if γ_z is chosen $\geq \frac{1}{4} + 4(z/2)^2 > z^2/2$. Thus

$$\begin{aligned}
(\tfrac{1}{4}(1-\Delta)\psi, \psi) &\leq \left(\left(\frac{-\Delta}{2} - \frac{z}{|x|} + \gamma_z\right)\psi, \psi\right) \\
&\leq (\tfrac{1}{2}(2\gamma_z - \Delta)\psi, \psi)
\end{aligned}$$

the last inequality following from $-(z/|x|\,\psi, \psi) \leq 0$. Since all operators are positive, the above can be rewritten as

$$\begin{aligned}
\tfrac{1}{4}\|(1-\Delta)^{1/2}\psi\|^2 &\leq \left\|\left(\frac{-\Delta}{2} - \frac{z}{|x|} + \gamma_z\right)^{1/2}\psi\right\|^2 \\
&\leq \tfrac{1}{2}\|(2\gamma_z - \Delta)^{1/2}\psi\|^2.
\end{aligned}$$

Since $D(-\Delta)$ is a core for all of the above operators, the inequality can be extended to $\psi \in D((-\Delta)^{1/2})$, thus completing the proof since the first and last norms are equivalent to the H^1 norm by the spectral theorem.

Corollary 2.2: The Coulomb propagator $\exp(iH_c t)$ is a uniformly bounded one-parameter group of operators in $H^1(\mathbb{R}^3)$.

Returning to the question of existence locally of solutions to Eqs. (1.1') and (1.2'), all that requires checking in Segal's fundamental result (Ref. 7, p. 343, Theorem 1), in view of the above corollary is the locally Lipschitz character of the nonlinear term in H^1.

Lemma 2.3: The nonlinear term $v(\psi)\psi$ in (1.1) is locally Lipschitzian on $H^1(\mathbb{R}^3)$; that is, there exists a constant $c = c(\|\psi\|_{H^1}, \|\tilde{\psi}\|_{H^1})$, depending on $\|\psi\|_{H^1}$ and $\|\psi\|_{H^1}$, such that

$$\|v(\psi)\psi - v(\tilde{\psi})\tilde{\psi}\|_{H^1} \leq c(\|\psi\|_{H^1}, \|\tilde{\psi}\|_{H^1})\,\|\psi - \tilde{\psi}\|_{H^1}.$$

Proof: From (1.3) we have

$$v = v(\psi) = \int \frac{|\psi|^2}{|x-y|}\,dy.$$

Hence, using the basic inequality (cf. Ref. 9, p. 446)

$$\int (|\psi|^2/|x|^2)\,dx \leq 4\|\nabla\psi\|_2^2,$$

we find

$$\|v\|_\infty \leq \|\psi\|_2\,\|\psi/|x-y|\,\|_2 \leq 2\|\psi\|_2\,\|\nabla\psi\|_2$$

and $\|\nabla v\|_\infty \leq \text{const}\,\|\nabla\psi\|_2^2$.

Then, writing $\tilde{v}$ for $v(\tilde{\psi})$ and v for $v(\psi)$, we find

$$\begin{aligned}
&\|v\psi - \tilde{v}\tilde{\psi}\|_{H^1} = \|v(\psi - \tilde{\psi}) - \tilde{\psi}(\tilde{v} - v)\|_{H^1} \\
&\leq \text{const}(\|v(\psi - \tilde{\psi})\|_2 + \|\tilde{\psi}(\tilde{v} - v)\|_2 + \|\nabla v(\psi - \tilde{\psi})\|_2 \\
&+ \|v(\nabla\psi - \nabla\tilde{\psi})\|_2 + \|\nabla\tilde{\psi}(\tilde{v} - v)\|_2 + \|\tilde{\psi}(\nabla v - \nabla\tilde{v})\|_2) \\
&\equiv \text{const}(I_1 + I_2 + \cdots + I_6).
\end{aligned}$$

We now estimate the I_k's using the above inequalities. We have

$$\begin{aligned}
I_1 = \|v(\psi - \tilde{\psi})\|_2 &\leq \|v\|_\infty\|\psi - \tilde{\psi}\|_2 \leq 2\|\psi\|_2\,\|\nabla\psi\|_2\|\psi - \tilde{\psi}\|_2 \\
&\leq \|\psi\|_{H^1}^2\|\psi - \tilde{\psi}\|_{H^1}.
\end{aligned}$$

Using the Sobolev inequality $\|\varphi\|_6 \leq \text{const}\,\|\nabla\varphi\|_2$, we obtain

$$I_2 = \|\tilde{\psi}(\tilde{v} - v)\|_2 \leq \|\tilde{\psi}\|_6\|\tilde{v} - v\|_3 \leq \text{const}\,\|\tilde{\psi}\|_{H^1}\|\tilde{v} - v\|_3.$$

Now from Ref. 10, p. 220, part c, there follows

$$\begin{aligned}
\|\tilde{v} - v\|_3 &\leq \text{const}\,\|\,|\tilde{\psi}|^2 - |\psi|^2\|_1 \\
&\leq \text{const}\,\|\psi + \tilde{\psi}\|_2\,\|\psi - \tilde{\psi}\|_2.
\end{aligned}$$

Thus

$$I_2 \leq \text{const}\,\|\tilde{\psi}\|_{H^1}(\|\psi\|_{H^1} + \|\tilde{\psi}\|_{H^1})\,\|\psi - \tilde{\psi}\|_{H^1}.$$

Similarly we find

$$\begin{aligned}
I_3 = \|\nabla v(\psi - \tilde{\psi})\|_2 &\leq \|\nabla v\|_\infty\|\psi - \tilde{\psi}\|_2 \\
&\leq \text{const}\,\|\psi\|_{H^1}^2\|\psi - \tilde{\psi}\|_{H^1}
\end{aligned}$$

and

$$I_4 = \|v(\nabla\psi - \nabla\tilde{\psi})\|_2 \leq \|v\|_\infty\|\nabla\psi - \nabla\tilde{\psi}\|_2 \leq 2\|\psi\|_{H^1}^2\|\psi - \tilde{\psi}\|_{H^1}.$$

For I_5 we write

$$I_5 = \|\nabla\tilde{\mathcal{V}}(\tilde{v}-v)\|_2 \le \|\nabla\tilde{\mathcal{V}}\|_2 \|\tilde{v}-v\|_\infty$$
$$\le \|\nabla\tilde{\mathcal{V}}\|_2 \|\psi-\tilde{\psi}\|_2 \|(\psi+\tilde{\psi})/|x-y|\|_2$$
$$\le 2\|\nabla\mathcal{V}\|_2 (\|\nabla\psi\|_2 + \|\nabla\tilde{\mathcal{V}}\|_2) \|\psi-\tilde{\mathcal{V}}\|_{H^1}$$

and finally I_6 is estimated as

$$I_6 = \|\mathcal{V}(\nabla v - \nabla\tilde{v})\|_2 \le \|\tilde{\mathcal{V}}\|_2 \|\nabla v - \nabla\tilde{v}\|_2$$
$$\le \text{const}\|\tilde{\mathcal{V}}\|_2 \left\| \int \big| |\psi|^2 - |\tilde{\psi}|^2 \big| \, dy/|x-y|^2 \right\|_\infty$$
$$\le \text{const}\|\tilde{\mathcal{V}}\|_{H^1} \|(\psi+\tilde{\psi})/|x-y|\|_2 \|(\psi-\tilde{\psi})/|x-y|\|_2$$
$$\le \text{const}\|\tilde{\mathcal{V}}\|_{H^1}(\|\nabla\psi\|_2 + \|\nabla\tilde{\mathcal{V}}\|_2) \|\nabla\psi - \nabla\tilde{\mathcal{V}}\|_2$$
$$\le \text{const}\|\tilde{\mathcal{V}}\|_{H^1}(\|\psi\|_{H^1} + \|\tilde{\mathcal{V}}\|_{H^1}) \|\psi - \tilde{\mathcal{V}}\|_{H^1}.$$

This proves the lemma.

We can prove a similar result for Eq. (1.2); complications only arise in algebraic manipulations.

Lemma 2.4: The nonlinear term $v_{0\rho}\psi_j$ in (1.2) is locally Lipschitzian on $H^1(\mathbb{R}^3)$.

Proof: If we set

$$v_{jk} = (1/r) * \psi_j\psi_k$$

and $v_k = v_{kk}$, we must obtain the stated estimate on the quantity

$$\sum_{k=1}^{N} [(\psi_j v_k - \psi_k v_{jk}) - (\tilde{\mathcal{V}}_j\tilde{v}_k - \tilde{\mathcal{V}}_k\tilde{v}_{jk})],$$

where $\tilde{v}_k = (1/r) * |\tilde{\psi}_k|^2$, $\tilde{v}_{jk} = (1/r) * \tilde{\psi}_j\tilde{\psi}_k$. We rewrite the above expression in the form

$$\sum_{k=1}^{N} [\psi_j(v_k - \tilde{v}_k) - \tilde{v}_k(\tilde{\psi}_j - \psi_j) - \psi_k(v_{jk} - \tilde{v}_{jk}) + \tilde{v}_{jk}(\tilde{\mathcal{V}}_k - \psi_k)].$$

We choose the norm $|\Psi| = \max_{1\le i\le N} |\psi_i|$ for a vector $\Psi = (\psi_1, \psi_2, \ldots, \psi_N)$. Then all the estimates are essentially contained in the preceding lemma. Clearly, the first two terms above,

$$\psi_j(v_k - \tilde{v}_k) - \tilde{v}_k(\tilde{\psi}_j - \psi_j),$$

are handled exactly as before. The remaining terms are estimated, for example, as

$$\|\nabla\tilde{v}_{jk}(\tilde{\mathcal{V}}_k - \psi_k)\|_2 \le \|\nabla\tilde{v}_{jk}\|_\infty \|\tilde{\mathcal{V}}_k - \psi_k\|_2$$
$$\le \text{const}\|\tilde{\Psi} - \Psi\|_2 \int \tilde{\mathcal{V}}_j\bar{\tilde{\mathcal{V}}}_k \, dy/|x-y| \, \|_\infty$$
$$\le \text{const}\|\tilde{\Psi} - \Psi\|_{H^1} \|\tilde{\mathcal{V}}_j\|_2 \|\nabla\tilde{\mathcal{V}}_k\|_2$$

and as

$$\|\nabla\psi_k(v_{jk} - \tilde{v}_{jk})\|_2$$
$$\le \|v_{jk} - \tilde{v}_{jk}\|_\infty \|\nabla\psi_k\|_2$$
$$\le \|\Psi\|_{H^1} \int \frac{|\psi_j\bar{\psi}_k - \tilde{\psi}_j\bar{\tilde{\psi}}_k| \, dy}{|x-y|}$$
$$= \|\Psi\|_{H^1} \int \frac{|\psi_j(\bar{\psi}_k - \bar{\tilde{\psi}}_k) - \bar{\tilde{\psi}}_k(\tilde{\psi}_j - \psi_j)| \, dy}{|x-y|}$$
$$\le \|\Psi\|_{H^1} \left\| \frac{\psi_j}{|x-y|} \right\|_2 \|\psi_k - \tilde{\psi}_k\|_2$$
$$+ \left\| \frac{\tilde{\psi}_k}{|x-y|} \right\|_2 \|\tilde{\psi}_j - \psi_j\|_2$$
$$\le \text{const}\|\Psi\|_{H^1}[\|\Psi\|_{H^1} \|\Psi - \tilde{\Psi}\|_2$$
$$+ \|\tilde{\Psi}\|_{H^1} \|\Psi - \tilde{\Psi}\|_{H^1}], \quad \text{etc.}$$

3. GLOBAL EXISTENCE

We now show that the local solutions found above exist globally. In view of Segal's theorem,[7] we must show that the H^1 norms remain finite at each $t \ge 0$. Indeed we will prove that, for a solution ψ of (1.1) or (1.2)

$$\|\psi(t)\|_{H^1}$$

is a uniformly bounded function of t. In addition it can be shown that $\|\psi(t)\|_{H^2}$ grows no faster than exponentially for large t. We will proceed formally with the calculations and will supply justification at the end of this section. We begin with the following basic result.

Lemma 3.1: (a) Let ψ be a solution of (1.1). Then $\|\psi(t)\|_2 = \|\psi(0)\|_2$.

(b) Let ψ_j $(j = 1, 2, \ldots, N)$ be solutions of (1.2). Then $\|\psi_j(t)\|_2 = \|\psi_j(0)\|_2$ for all $j = 1, 2, \ldots, N$.

Proof: (a) We multiply (1.1) by $\bar{\psi}$ to get

$$\bar{\psi}\psi_t = (-i/2)\bar{\psi}\,\Delta\psi - (2i/|x|)\,|\psi|^2 + iv|\psi|^2.$$

We conjugate this and add the result to the above expression. This gives

$$\frac{\partial}{\partial t}|\psi|^2 = \frac{i}{2}(\psi\,\Delta\bar{\psi} - \bar{\psi}\,\Delta\psi)$$
$$= \frac{i}{2}\nabla\cdot(\psi\nabla\bar{\psi} - \bar{\psi}\nabla\psi).$$

An integration establishes the result. To prove (b), let us write (1.2) in the form

$$i\frac{\partial\psi_j}{\partial t} = \frac{1}{2}\Delta\psi_j + \frac{z}{|x|}\psi_j - \sum_{k=1}^{N}(\psi_j v_k - \psi_k v_{jk}), \tag{3.1}$$

where

$$v_{jk} = \int \frac{\psi_j(y,t)\bar{\psi}_k(y,t)}{|x-y|}\,dy \tag{3.1'}$$

and $v_k = v_{kk}$ is real. A multiplication of (3.1) by $\bar{\psi}_j$ gives

$$i\bar{\psi}_j\frac{\partial\psi}{\partial t} = \frac{1}{2}\bar{\psi}_j\,\Delta\psi_j + \frac{z}{|x|}|\psi_j|^2$$
$$- \sum_{k=1}^{N}(|\psi_j|^2 v_k - \bar{\psi}_j\psi_k v_{jk}).$$

We now take imaginary parts of this equation. The left-hand side becomes

$$\text{Im}\left(i\bar{\psi}_j\frac{\partial\psi_j}{\partial t}\right) = \text{Re}\left(\bar{\psi}_j\frac{\partial\psi_j}{\partial t}\right) = \frac{1}{2}\frac{\partial}{\partial t}|\psi_j|^2,$$

and the right-hand side equals

$$\tfrac{1}{2}\text{Im}[\nabla\cdot(\bar{\psi}_j\nabla\psi_j)] + \text{Im}\sum_{k=1}^{N}\bar{\psi}_j\psi_k v_{jk}.$$

Now from (3.2) we have

$$\Delta v_{jk} = -4\pi\psi_j\bar{\psi}_k$$

so that

$$\operatorname{Im}\sum_{k=1}^{N}\bar{\psi}_j\psi_k v_{jk}=\frac{-1}{4\pi}\operatorname{Im}\sum_{k=1}^{N}v_{jk}\Delta\bar{v}_{jk}$$

$$=\frac{-1}{4\pi}\operatorname{Im}\sum_{k=1}^{N}[\nabla\cdot(v_{jk}\nabla\bar{v}_{jk})-|\nabla v_{jk}|^2].$$

Hence

$$\frac{1}{2}\frac{\partial}{\partial t}|\psi|^2=\operatorname{Im}\nabla\cdot\left(\frac{1}{2}(\bar{\psi}_j\nabla\psi_j)-\frac{1}{4\pi}\sum_{k=1}^{N}v_{jk}\nabla\bar{v}_{jk}\right),$$

and an integration over $\mathbb{R}^3$ completes the proof.

Let us now confine our attention to solutions of Eq. (1.1) and obtain the desired estimates.

Lemma 3.2: Let ψ be a solution of (1.1) with $\psi(\cdot,0)\in H^1(\mathbb{R}^3)$. Then we have

$$\int[|\nabla\psi(x,t)|^2+v(x,t)|\psi(x,t)|^2-(4/|x|)|\psi(x,t)|^2]\,dx$$
$$=\text{const}=C_0. \tag{3.2}$$

Hence $\|\nabla\psi(t)\|_2^2$ is uniformly bounded by some constant depending only on $\|\psi(0)\|_{H^1}$.

Proof: We multiply (1.1) by $\bar{\psi}_t$ to get

$$|\psi_t|^2=(-i/2)\bar{\psi}_t\Delta\psi-(2i/|x|)\bar{\psi}_t\psi+iv\bar{\psi}_t\psi.$$

We conjugate this and subtract the result from the above equation. This gives

$$0=\frac{-i}{2}(\bar{\psi}_t\Delta\psi+\psi_t\Delta\bar{\psi})-\frac{2i}{|x|}\frac{\partial}{\partial t}|\psi|^2+iv\frac{\partial}{\partial t}|\psi|^2$$

$$=\frac{-i}{2}\nabla\cdot(\bar{\psi}_t\nabla\psi+\psi_t\nabla\bar{\psi})+\frac{i}{2}\frac{\partial}{\partial t}|\nabla\psi|^2$$

$$-\frac{\partial}{\partial t}\left(\frac{2i}{|x|}|\psi|^2\right)+iv\frac{\partial}{\partial t}|\psi|^2.$$

Now from (1.3) we have

$$\Delta v=-4\pi|\psi|^2.$$

Hence we may rewrite the last term above as

$$v\frac{\partial}{\partial t}|\psi|^2=\frac{-v}{4\pi}\Delta v_t=\frac{-1}{4\pi}\nabla\cdot(v\nabla v_t)+\frac{1}{4\pi}\nabla v\cdot\nabla v_t$$

$$=\frac{-1}{4\pi}\nabla\cdot(v\nabla v_t)+\frac{1}{8\pi}\frac{\partial}{\partial t}|\nabla v|^2$$

$$=\frac{-1}{4\pi}\nabla\cdot(v\nabla v_t)+\frac{1}{2}\frac{\partial}{\partial t}(v|\psi|^2).$$

By integrating over $\mathbb{R}^3$ we obtain the desired result

$$\frac{d}{dt}\int\left(|\nabla\psi|^2+v|\psi|^2-\frac{4|\psi|^2}{|x|}\right)dx=0.$$

To prove the second assertion, we find from (3.2)

$$\|\nabla\psi(t)\|_2^2\le C_0+4\int\frac{|\psi(x,t)|^2}{|x|}\,dx.$$

Using the fundamental inequality

$$\|\psi/|x|\|_2\le 2\|\nabla\psi\|_2$$

(cf. Ref. 9, p. 446), we obtain by the Schwarz inequality the estimate

$$\|\nabla\psi(t)\|_2^2\le C_0+4\|\psi(t)\|_2\|\psi(t)/|x|\|_2$$
$$\le C_0+8\|\psi(0)\|_2\|\nabla\psi(t)\|_2,$$

which proves the result.

Before establishing the H^2 bounds, we remark that this invariant can also be obtained through a multiplication of (1.1) by $\Delta\bar{\psi}$.

Lemma 3.3: Let ψ be a solution of (1.1) with $\psi(\cdot,0)\in H^2(\mathbb{R}^3)$. Then there exist positive constants c, k, depending only on $\|\psi(0)\|_{H^2}$, such that

$$\|\psi(t)\|_{H^2}\le c\exp(kt)$$

for all $t\ge 0$.

Proof: From (1.1) we have

$$\psi_{tt}=(-i/2)\Delta\psi_t-(2i/|x|)\psi_t+iv_t\psi+iv\psi_t.$$

Hence

$$\bar{\psi}_t\psi_{tt}=(-i/2)\bar{\psi}_t\Delta\psi_t-(2i/|x|)|\psi_t|^2+iv_t\bar{\psi}_t\psi+iv|\psi_t|^2$$

and

$$\psi_t\bar{\psi}_{tt}=(i/2)\psi_t\Delta\bar{\psi}_t+(2i/|x|)|\psi_t|^2-iv_t\psi_t\bar{\psi}-iv|\psi_t|^2.$$

We add these equations and integrate over $\mathbb{R}^3$ to get

$$\frac{d}{dt}\|\psi_t(t)\|_2^2=2\operatorname{Im}\int v_t\psi_t\bar{\psi}\,dx.$$

Now

$$\|v_t(t)\|_\infty\le 2\|\int|\psi||\psi_t|\,dy/|x-y|\,\|_\infty$$
$$\le 2\|\psi_t(t)\|_2\|\psi(t)/|x-y|\,\|_2$$
$$\le 4\|\psi_t(t)\|_2\|\nabla\psi(t)\|_2$$
$$\le\text{const}\,\|\psi_t(t)\|_2.$$

Therefore,

$$\frac{d}{dt}\|\psi_t(t)\|_2^2\le\text{const}\|v_t(t)\|_\infty\|\psi_t(t)\|_2\|\psi(t)\|_2$$
$$\le\text{const}\,\|\psi_t(t)\|_2^2.$$

Hence by Gronwall's inequality, $\|\psi_t(t)\|_2$ is an exponentially bounded function of t. Then from (1.1) we find

$$\|\Delta\psi(t)\|_2\le 2\|\psi_t(t)\|_2+4\left\|\frac{\psi(t)}{|x|}\right\|_2+2\|v(t)\psi(t)\|_2$$

$$\le 2\|\psi_t(t)\|_2+8\|\nabla\psi(t)\|_2+2\|v(t)\|_\infty\|\psi(0)\|_2$$

$$\le 2\|\psi_t(t)\|_2+8\|\nabla\psi(t)\|_2+4\|\psi(0)\|_2^2\|\nabla\psi(t)\|_2.$$

Thus $\|\Delta\psi(t)\|_2\le\text{const}(1+\|\psi_t(t)\|_2)$ so that $\|\Delta\psi(t)\|_2$, and also $\|\psi(t)\|_{H^2}$, are exponentially bounded.

We have now proved the following result.

Theorem 3.1: The Cauchy problem for the (integrated form of the) time-dependent Hartree equation (1.1) has a unique global solution in $H^1(\mathbb{R}^3)$, provided $\psi(\cdot,0)\in H^1(\mathbb{R}^3)$. This solution satisfies the bounds

$$\|\psi(t)\|_{H^1}\le M,\quad t\ge 0,$$

$$\|\psi(t)\|_{H^2}\le c\exp(kt)\quad\text{if}\quad\psi(\cdot,0)\in H^2,\quad t\ge 0,$$

where the constants M and c, k depend only on $\|\psi(0)\|_{H^1}$ and $\|\psi(0)\|_{H^2}$, respectively.

We shall now develop the corresponding theorem for

Eq. (1.2). Recall that it can be written as

$$i\frac{\partial\psi_j}{\partial t}=\frac{1}{2}\Delta\psi_j+\frac{z}{|x|}\psi_j-\sum_{k=1}^{N}(\psi_j v_k-\psi_k v_{jk})\quad(1\leq j\leq N)\tag{1.2'}$$

where $v_{jk}=\int[\psi_j(y,t)\bar{\psi}_k(y,t)/|x-y|]dy$ and $v_k=v_{kk}$. We thus have $\bar{v}_{jk}=v_{kj}$ and

$$\Delta v_{jk}=-4\pi\psi_j\bar{\psi}_k.\tag{3.3}$$

We shall consistently use the notation

$$\sum_{j=1}^{N}\sum_{k=1}^{N}{}' a_j b_k,$$

which will mean the term for which $k=j$ is to be omitted from the summation. Notice that the term corresponding to $k=j$ in the sum on the right side of (1.2) vanishes. We now obtain the analog of Lemma 3.2.

Lemma 3.4: Let ψ_j, $1\leq j\leq N$, be solutions of (1.2) with $\psi_j(\cdot,0)\in H^1(\mathbb{R}^3)$ for all j. Then we have

$$\sum_{j=1}^{N}\int_{\mathbb{R}^3}\left(\frac{1}{4}|\nabla\psi_j(x,t)|^2-\frac{z}{2}\frac{|\psi_j(x,t)|^2}{|x|}\right.$$

$$+\frac{1}{4}\sum_{k=1}^{N}{}' v_k(x,t)|\psi_j(x,t)|^2$$

$$\left.-\frac{1}{16\pi}\sum_{k=1}^{N}{}'|\nabla v_{jk}(x,t)|^2\right)dx=\text{const}=C_0.$$

Moreover, $\sum_{j=1}^{N}\|\nabla\psi_j(t)\|_2^2$ is uniformly bounded by a constant depending only on the norms $\|\psi_j(0)\|_{H^1}$, $1\leq j\leq N$.

Proof: The proof is an exceptionally laborious calculation. We follow the method of Lemma 3.2; that is, we multiply (1.2) by $\partial\bar{\psi}_j/\partial t$, sum over j, and then take real parts. As before, the left-hand side is

$$\text{Re}\left(i\sum_{j=1}^{N}\left|\frac{\partial\psi_j}{\partial t}\right|^2\right)\equiv 0.$$

The first two terms on the right-hand side give

$$\text{Re}\sum_j\frac{\overline{\partial\psi_j}}{\partial t}\left(\frac{1}{2}\Delta\psi_j+\frac{z}{|x|}\psi_j\right)$$

$$=\text{Re}\sum_j\left[\frac{1}{2}\nabla\cdot\left(\frac{\overline{\partial\psi_j}}{\partial t}\nabla\psi_j\right)-\frac{1}{2}\nabla\psi_j\cdot\nabla\left(\frac{\overline{\partial\psi_j}}{\partial t}\right)\right.$$

$$\left.+\frac{z}{|x|}\psi_j\frac{\overline{\partial\psi_j}}{\partial t}\right]$$

$$=\nabla\cdot\left(\frac{1}{2}\text{Re}\sum_j\frac{\overline{\partial\psi_j}}{\partial t}\nabla\psi_j\right)-\frac{\partial}{\partial t}\left[\sum_j\left(\frac{1}{4}|\nabla\psi_j|^2-\frac{z}{2|x|}|\psi_j|^2\right)\right].$$

Hence

$$\frac{d}{dt}\sum_j\int\left(\frac{1}{4}|\nabla\psi_j|^2-\frac{z}{2|x|}|\psi_j|^2\right)dx$$

$$=-\text{Re}\int\sum_j\sum_k{}'\left(\psi_j\frac{\overline{\partial\psi_j}}{\partial t}v_k-\psi_k\frac{\overline{\partial\psi_j}}{\partial t}v_{jk}\right)dx.\tag{3.4}$$

It remains for us to show that the right side of (3.4) can be written as the time derivative of the appropriate quantity. For this purpose we shall integrate by parts at will without explicit mention. All double sums below $\sum_j\sum_k'$ are taken over the indices $j=1,2,\ldots,N$, $k=1,2,\ldots,N$, $k\neq j$. First we calculate

$$\frac{d}{dt}\int\sum_j\sum_k{}' v_k|\psi_j|^2\,dx$$

$$=\int\sum_j\sum_k{}'\left(\frac{\partial v_k}{\partial t}|\psi_j|^2+v_k\frac{\partial}{\partial t}|\psi_j|^2\right)dx$$

$$=\int\sum_j\sum_k{}'\left(-\frac{1}{4\pi}\frac{\partial v_k}{\partial t}\Delta v_j+2\,\text{Re}\,v_k\psi_j\frac{\partial\bar{\psi}_j}{\partial t}\right)dx$$

$$=\int\sum_j\sum_k{}'\left[-\frac{1}{8\pi}\frac{\partial v_k}{\partial t}\Delta v_j-\frac{1}{8\pi}v_j\Delta\left(\frac{\partial v_k}{\partial t}\right)\right.$$

$$\left.+2\text{Re}\,v_k\psi_j\frac{\partial\bar{\psi}_j}{\partial t}\right]dx$$

$$=-\frac{1}{8\pi}\int\sum_j\sum_k{}'\frac{\partial v_k}{\partial t}\Delta v_j\,dx-\frac{1}{8\pi}\frac{\partial}{\partial t}\int\sum_j\sum_k{}' v_j\Delta v_k\,dx$$

$$+\frac{1}{8\pi}\int\sum_j\sum_k{}'\frac{\partial v_j}{\partial t}\Delta v_k\,dx+2\text{Re}\int\sum_j\sum_k{}' v_k\psi_j\frac{\overline{\partial\psi_j}}{\partial t}\,dx.$$

Hence

$$\frac{d}{dt}\int\sum_j\sum_k{}'(v_k|\psi_j|^2-\tfrac{1}{2}v_j|\psi_k|^2)\,dx$$

$$=\frac{1}{8\pi}\int\sum_j\sum_k{}'\left(\frac{\partial v_j}{\partial t}\Delta v_k-\frac{\partial v_k}{\partial t}\Delta v_j\right)dx$$

$$+2\,\text{Re}\int\sum_j\sum_k{}' v_k\psi_j\frac{\partial\bar{\psi}_j}{\partial t}\,dx.\tag{3.5}$$

The first term on the right above vanishes on account of

$$\sum_j\sum_k{}'\left(\frac{\partial v_j}{\partial t}\Delta v_k-\frac{\partial v_k}{\partial t}\Delta v_j\right)$$

$$=\sum_j\sum_k\left(\frac{\partial v_j}{\partial t}\Delta v_k-\frac{\partial v_k}{\partial t}\Delta v_j\right)$$

$$-\sum_j\frac{\partial v_j}{\partial t}\Delta v_j+\sum_j\frac{\partial v_j}{\partial t}\Delta v_j\equiv 0.$$

The integrand on the left-hand side of (3.5) can be written

$$\sum_j\sum_k{}'(v_k|\psi_j|^2-\tfrac{1}{2}v_j|\psi_k|^2)$$

$$=\sum_j\sum_k v_k|\psi_j|^2-\sum_j v_j|\psi_j|^2$$

$$-\tfrac{1}{2}\sum_j\sum_k v_j|\psi_k|^2+\tfrac{1}{2}\sum_j v_j|\psi_j|^2$$

$$=\tfrac{1}{2}\sum_j\sum_k v_k|\psi_j|^2-\tfrac{1}{2}\sum_j v_j|\psi_j|^2$$

$$=\tfrac{1}{2}\sum_j\sum_k{}' v_k|\psi_j|^2.$$

Thus from (3.5) we obtain

$$\mathrm{Re}\int \sum_j \sum_k{}' v_k \psi_j \frac{\partial \bar{\psi}_j}{\partial t}\,dx = \frac{1}{4}\frac{d}{dt}\int \sum_j \sum_k{}' v_k |\psi_j|^2\,dx.$$

The use of this result in (3.4) gives

$$\frac{d}{dt}\sum_j \int \left[\frac{1}{4}|\nabla\psi_j|^2 - \frac{z}{2|x|}|\psi_j|^2 + \frac{1}{4}\sum_k{}' v_k|\psi_j|^2\right]dx$$

$$= \mathrm{Re}\int \sum_j \sum_k{}' \psi_k \frac{\partial \bar{\psi}_j}{\partial t} v_{jk}\,dx. \tag{3.6}$$

To evaluate the last term above, we calculate

$$\frac{1}{4\pi}\frac{d}{dt}\int \sum_j \sum_k{}' |\nabla v_{jk}|^2\,dx$$

$$= \frac{1}{4\pi}\frac{d}{dt}\int \sum_j \sum_k{}' [\nabla\cdot(\bar{v}_{jk}\nabla v_{jk}) - \bar{v}_{jk}\Delta v_{jk}]\,dx$$

$$= -\frac{1}{4\pi}\int\left[\sum_j \sum_k{}' \frac{\partial \bar{v}_{jk}}{\partial t}\Delta v_{jk} + \bar{v}_{jk}\Delta\left(\frac{\partial v_{jk}}{\partial t}\right)\right]dx$$

$$= -\frac{1}{4\pi}\int\left[\sum_j \sum_k{}' \Delta\left(\frac{\partial \bar{v}_{jk}}{\partial t}\right)v_{jk} + \bar{v}_{jk}\Delta\left(\frac{\partial v_{jk}}{\partial t}\right)\right]dx$$

$$= \int \sum_j \sum_k{}' v_{jk}\frac{\partial}{\partial t}(\bar{\psi}_j\psi_k) + \bar{v}_{jk}\frac{\partial}{\partial t}(\psi_j\bar{\psi}_k)\,dx$$

$$= 2\mathrm{Re}\int \sum_j \sum_k{}' v_{jk}\frac{\partial}{\partial t}(\bar{\psi}_j\psi_k)\,dx$$

$$= 2\,\mathrm{Re}\int \sum_j \sum_k{}' v_{jk}\left(\bar{\psi}_j\frac{\partial\psi_k}{\partial t} + \psi_k\frac{\partial\bar{\psi}_j}{\partial t}\right)dx$$

$$= 2\,\mathrm{Re}\int \sum_j \sum_k v_{jk}\left(\bar{\psi}_j\frac{\partial\psi_k}{\partial t} + \psi\frac{\partial\bar{\psi}_j}{\partial t}\right)dx$$

$$- 2\,\mathrm{Re}\int \sum_j v_{jj}\left(\bar{\psi}_j\frac{\partial\psi_j}{\partial t} + \psi_j\frac{\partial\bar{\psi}_j}{\partial t}\right)dx$$

$$= 2\,\mathrm{Re}\int \sum_j \sum_k v_{jk}\psi_k\frac{\overline{\partial\psi_j}}{\partial t}\,dx$$

$$\overline{+\ 2\,\mathrm{Re}\int \sum_j \sum_k \bar{v}_{jk}\psi_j\frac{\partial\bar{\psi}_k}{\partial t}\,dx}$$

$$- 2\,\mathrm{Re}\int \sum_j v_{jj}\left(\bar{\psi}_j\frac{\partial\psi_j}{\partial t} + \psi_j\frac{\partial\bar{\psi}_j}{\partial t}\right)dx$$

$$= 2\,\mathrm{Re}\int \sum_j \sum_k v_{jk}\psi_k\frac{\partial\bar{\psi}_j}{\partial t}\,dx$$

$$\overline{+\ 2\,\mathrm{Re}\int \sum_j \sum_k v_{kj}\psi_j\frac{\partial\bar{\psi}_k}{\partial t}\,dx}$$

$$- 2\,\mathrm{Re}\int \sum_j v_{jj}\left(\bar{\psi}_j\frac{\partial\psi_j}{\partial t} + \psi_j\frac{\partial\bar{\psi}_j}{\partial t}\right)dx.$$

Here we have used the property $\bar{v}_{jk} = v_{kj}$, which now allows us to combine the first two sums in the line above. Hence

$$\frac{1}{4\pi}\frac{d}{dt}\int \sum_j \sum_k{}' |\nabla v_{jk}|^2\,dx$$

$$= 4\,\mathrm{Re}\int \sum_j \sum_k v_{jk}\psi_k\frac{\partial\bar{\psi}_j}{\partial t}\,dx - 4\,\mathrm{Re}\int \sum_j v_{jj}\psi_j\frac{\partial\bar{\psi}_j}{\partial t}\,dx$$

$$= 4\,\mathrm{Re}\int \sum_j \sum_k{}' v_{jk}\psi_k\frac{\partial\bar{\psi}_j}{\partial t}\,dx$$

because v_{jj} is real. Thus we have calculated the right-hand side of (3.6). Substitution of this result there yields

$$\frac{d}{dt}\sum_j \int \left[\frac{1}{4}|\nabla\psi_j|^2 - \frac{z}{2|x|}|\psi_j|^2 + \frac{1}{4}\sum_k{}' v_k|\psi_j|^2 - \frac{1}{16\pi}\sum_k{}' |\nabla v_{jk}|^2\right]dx = 0,$$

which is equivalent to the first part of the statement of the lemma. We show finally that the expressions

$$I_j \equiv \int \sum_k{}'\left(\frac{1}{4}v_k|\psi_j|^2 - \frac{1}{16\pi}|\nabla v_{jk}|^2\right)dx \tag{3.7}$$

are nonnegative for each $j = 1, 2, \ldots, N$. Then we will have

$$\sum_{j=1}^N \|\nabla\psi_j(t)\|_2^2 \le 4C_0 + 2Z\sum_{j=1}^N \int \frac{|\psi_j|^2}{|x|}\,dx$$

and this inequality provides the uniform bound on $\sum_j \|\nabla\psi_j(t)\|_2^2$, as in Lemma 3.2. To establish (3.7), we note that

$$\frac{1}{4\pi}\int |\nabla v_{jk}|^2\,dx = \int \bar{v}_{jk}\psi_j\bar{\psi}_k\,dx$$

follows from (3.3). Hence

$$4I_j \ge \int \sum_k{}' (v_k|\psi_j|^2 - |\bar{v}_{jk}\psi_j\bar{\psi}_k|)\,dx.$$

Now

$$|\bar{v}_{jk}| = \left|\int \frac{\bar{\psi}_j\psi_k\,dy}{|x-y|}\right| = \left|\int \frac{\bar{\psi}_j}{|x-y|^{1/2}}\,\frac{\psi_k}{|x-y|^{1/2}}\,dy\right|$$

$$\le \left(\int \frac{|\psi_j|^2}{|x-y|}dy\right)^{1/2}\left(\int \frac{|\psi_k|^2}{|x-y|}\,dy\right)^{1/2} = v_j^{1/2}v_k^{1/2}$$

so that

$$|\bar{v}_{jk}\psi_j\bar{\psi}_k| \le (v_k^{1/2}|\psi_j|)(v_j^{1/2}|\psi_k|)$$

$$\le \tfrac{1}{2}v_k|\psi_j|^2 + \tfrac{1}{2}v_j|\psi_k|^2.$$

Therefore,

$$4I_j \ge \int \sum_k{}' (\tfrac{1}{2}v_k|\psi_j|^2 - \tfrac{1}{2}v_j|\psi_k|^2)\,dx$$

$$= \frac{1}{8\pi}\int \sum_k{}' (v_j\Delta v_k - v_k\Delta v_j)\,dx = 0. \qquad \text{QED}$$

As in Lemma 3.3, we can also derive H^2 bounds of the form

$$\|\psi_j(t)\|_{H^2} \le c\exp(kt) \quad (j = 1, 2, \ldots, N)$$

for solutions ψ_j of (1.2). The method is exactly that of Lemma 3.3; that is, we take $\partial/\partial t$ of (1.2), multiply the expressions obtained by $\partial\bar{\psi}_j/\partial t$, etc., and then integrate over $\mathbb{R}^3$ and sum the results over $1 \le j \le N$. Obviously the proof will go through as before provided we can estimate each of the norms

$$\left\| \psi_j \frac{\partial \bar{\psi}_j}{\partial t} \frac{\partial v_k}{\partial t} \right\|_1, \quad \left\| v_{jk} \frac{\partial \bar{\psi}_j}{\partial t} \frac{\partial \psi_k}{\partial t} \right\|_1, \quad \left\| \psi_k \frac{\partial \bar{\psi}_j}{\partial t} \frac{\partial v_{jk}}{\partial t} \right\|_1$$

above by

$$\text{const} \sum_{j=1}^{N} \left\| \frac{\partial \psi_j}{\partial t}(t) \right\|_2^2.$$

However, this is straightforward using the result of Lemma 3.4:

$$\left\| \psi_j \frac{\partial \bar{\psi}_j}{\partial t} \frac{\partial v_k}{\partial t} \right\|_1$$

$$\leq \left\| \frac{\partial v_k}{\partial t} \right\|_\infty \|\psi_j\|_2 \left\| \frac{\partial \bar{\psi}_j}{\partial t} \right\|_2$$

$$\leq \text{const} \left\| \frac{\psi_k}{|x-y|} \right\|_2 \left\| \frac{\partial \psi_k}{\partial t} \right\|_2 \|\psi_j\|_2 \left\| \frac{\partial \bar{\psi}_j}{\partial t} \right\|_2$$

$$\leq \text{const} \sum_j \left\| \frac{\partial \psi_j}{\partial t} \right\|_2^2,$$

$$\left\| \psi_k \frac{\partial \bar{\psi}_j}{\partial t} \frac{\partial v_{jk}}{\partial t} \right\|_1 \leq \left\| \frac{\partial v_{jk}}{\partial t} \right\|_\infty \|\psi_k\|_2 \left\| \frac{\partial \bar{\psi}_j}{\partial t} \right\|_2$$

$$\leq \text{const} \left\| \frac{\psi_j}{|x-y|} \right\|_2 \left\| \frac{\partial \psi_k}{\partial t} \right\|_2 \|\psi_k\|_2 \left\| \frac{\partial \bar{\psi}_j}{\partial t} \right\|_2$$

$$\leq \text{const} \sum_j \left\| \frac{\partial \psi_j}{\partial t} \right\|_2^2, \quad \text{etc.}$$

Thus finally we can state the main result:

Theorem 3.2: The Cauchy problem for the (integrated form of the) time-dependent Hartree system (1.2) has a unique global solution in $H^1(\mathbb{R}^3)$, provided $\psi_j(\cdot, 0) \in H^1(\mathbb{R}^3)$ for all $j = 1, 2, \ldots, N$. This solution obeys the bounds

$$\sum_{j=1}^{N} \|\psi_j(t)\|_{H^1}^2 \leq M$$

and

$$\sum_{j=1}^{N} \|\psi_j(t)\|_{H^2}^2 \leq c \exp(kt) \quad \text{if} \quad \psi(\cdot, 0) \in H^2,$$

where the constants M, c, k, depend only on N, Z, and the H^2 norms of the Cauchy data.

We now turn our attention to a justification of the above formal calculations. Clearly it would suffice to give a regularity theorem proving that our H^1 solution of the integral equation satisfies the differential equation in the usual sense. Such smoothness might be too much to ask for especially in view of the fact that the required estimates are of a much weaker variety. For this reason we approach the problem from a Galerkin viewpoint (Ref. 7, p. 347); i.e., we obtain a differential equation in a subspace of H^1, obtain estimates using the computational ideas of the previous lemmas, and obtain the final result by passing to the limit. We restrict our attention to the two-particle case, pointing out that, as usual, the N-particle situation follows similarly but with more calculating.

Let P_n be the spectral projection for $H_c = -\Delta/2 - 2/|x|$ on whose range the restriction of H_c is bounded by n. Multiplying equation (1.1′) by $P_n = P_n^2$, one obtains

$$P_n\psi(t) = P_n\psi^\circ(t) + i \int_{t_0}^{t} P_n \exp[iH_c(t-s)] P_n v(\psi(s))\psi(s)\, ds.$$

Now strong differentiation in t of this equation is permitted because of the boundedness of $P_n H_c$ (Ref. 7, p. 353, Theorem 3). This gives rise to the following differential equation for $p_n\psi = \psi_n$:

$$\frac{d}{dt}\psi_n = iH_c\psi_n + iP_n v(\psi)\psi. \tag{3.8}$$

This can be used just as effectively as Eq. (1.1) to establish the conservation of charge. Forming the inner product in $L^2(\mathbb{R}^3)$ with $\psi_n(t)$ and taking real parts, one obtains

$$\frac{d}{dt}\|\psi_n(t)\|^2 = 2\text{Im}(P_n V(\psi(t))\psi(t), \quad \psi_n(t)).$$

Integrating from t_0 to $t \in (t_0, T)$, we obtain

$$\|\psi_n(t)\|^2 - \|\psi_n(t_0)\|^2 = 2\int_{t_0}^{t} \text{Im}(v(\psi(s))\psi(s), \ P_n\psi(s))\, ds.$$

Letting $n \to \infty$, we obtain

$$\|\psi(t)\|^2 = \|\psi(t_0)\|^2 + 2\int_{t_0}^{t} \text{Im}(v(\psi(s))\psi(s), \ \psi(s))\, ds$$
$$= \|\psi(t_0)\|^2,$$

thus justifying the calculations of Lemma 3.1.

In order to justify the conservation law, we again begin with Eq. (3.8). As in the formal proof, we form the $L^2(\mathbb{R}^3)$ inner product with $(d/dt)\ \psi_n$ and take imaginary parts obtaining

$$0 = \frac{1}{2}\frac{d}{dt}\|\nabla\psi_n\|_2^2 - \frac{d}{dt}\left(\frac{2}{|x|}\psi_n, \psi_n\right) + (P_n v\psi, \psi_{nt}) + (\psi_{nt}, P_n v\psi).$$

By using the fact that $P_n\psi_{nt} = \psi_{nt}$ [directly from Eq. (3.8)], the last two terms can be written as

$$(v(\psi_n)\psi_n, \ \psi_{nt}) + (\psi_{nt}, \ v(\psi_n)\psi_n)$$
$$+ (v(\psi)\psi - v(\psi_n)\psi_n, \ \psi_{nt}) + (\psi_{nt}, \ v(\psi)\psi - v(\psi_n)\psi_n).$$

The first two terms are precisely $\int v(\psi_n)(\partial/\partial t)|\psi_n|^2 d^3x$ for which the formal calculations in Lemma 3.2 are now rigorous because $\psi_n \in D(H_c) = D(-\Delta)$. Thus

$$\tfrac{1}{2}\|\nabla\psi_n(t)\|_2^2 - ((2/|x|)\psi_n(t), \ \psi_n(t)) + \tfrac{1}{2}(v(\psi_n(t))\psi_n(t), \ \psi_n(t))$$
$$= \text{const} - 2\,\text{Re}\int_{t_0}^{t} (v(\psi)\psi - v(\psi_n)\psi \ , \psi_{nt})\, ds.$$

All that remains is to take the limit:

$$\|\nabla(\psi_n(t) - \psi(t))\|_2 \leq \|(H_c + \gamma_2)^{1/2}(P_n\psi(t) - \psi(t))\|_2$$
$$\leq \|(P_n - I)(H_c + \gamma_2)^{1/2}\psi(t)\|_2.$$

The first inequality follows from Lemma 2.1. The last term tends to zero because of the spectral theorem and the fact that for each $t \in [t_0, T)$, $\psi(t) \in H^1(\mathbb{R}^3) = D((H_c + \gamma_2)^{1/2})$. The second term can be treated by writing

$$|((2/|x|)\psi_n(t), \ \psi_n(t)) - ((2/|x|)\psi(t), \ \psi(t))|$$
$$\leq |((2/|x|)(\psi_n - \psi), \ \psi_n)| + |((2/|x|)\psi, \ \psi_n - \psi)|$$

$$\leq 2\{\|[(\psi_n-\psi)/|x|]\|_2\ \|\psi_n\|_2+\|\psi/|x|\|_2\ \|\psi_n-\psi\|_2\}$$
$$\leq 4[\|\nabla(\psi_n-\psi)\|_2\ \|\psi_n\|_2+\|\nabla\psi\|_2\ \|\psi_n-\psi\|_2].$$

The proof now follows from the previous argument. For the third term write

$$|(v_n\psi_n,\ \psi_n)-(v\psi,\ \psi)|$$
$$\leq |(v_n\psi_n-v\psi,\ \psi_n)|+|(v\psi,\ \psi_n-\psi)|$$
$$\leq \|v_n\psi_n-v\psi\|_2\ \|\psi_n\|_2+\|v\psi\|_2\ \|\psi_n-\psi\|_2,$$

and use the calculations presented in Lemma 2.3 for terms I_1 and I_2. Finally the integrand can be written as

$$i[(v\psi-v_n\psi_n,\ (H_c+\gamma_2)\psi_n)+(v\psi-v\psi_n, P_nv\psi)-\gamma_2(v\psi-v_n\psi_n,\ \psi_n)].$$

The last term was treated above, and the middle one can be handled similarly since $\|v\psi\|_2 \leq \|v\|_\infty\ \|\psi\|_2$ $\leq 2\|\nabla\psi\|_2\|\psi\|_2^2$. The first term is written as

$$((H_c+\gamma_2)^{1/2}(v\psi-v_n\psi_n),\ (H_c+\gamma_2)^{1/2}\psi_n)$$

and hence bounded by $\|v\psi-v_n\psi_n\|_{H^1}\|\psi_n\|_{H^1}$. The calculations of Lemma 2.3 now apply. Notice finally that the integrand is uniformly (in n) bounded by $2\ \text{const}[\|v\psi\|_{H^1}\|\psi\|_{H^1}+\|v\psi\|_2^2]$, which is integrable on $[t_0, T]$. Thus the integral tends to zero by the dominated convergence theorem and the conservation law is rigorously justified.

The H^2 estimate proceeds in the same manner requiring in addition the justification via the regularity theorem of Segal (Ref. 7, p. 353, Theorem 3) that Eq. (3.8) can be differentiated with respect to t.

Finally we remark that Eqs. (1.1) and (1.2) with periodic driving terms of the form $f(t)\psi$ (a situation of recent interest in the physical chemistry literature) can be treated with the same methods. The locally Lipschitz nature of this term is obvious. In the proof of the conservation of charge it behaves (i.e., disappears) as the other terms on the right-hand side of Eqs. (1.1) and (1.2). The conservation law for the energy now becomes

$$\int |\nabla\psi(x,t)|^2+v(x,t)|\psi(x,t)|^2-(4/|x|)\,|\psi(x,t)|^2 + \left(\int_{t_0}^{t} f'(s)|\psi(x,s)|^2\,ds\right)dx$$
$$= C_0+f(t)\|\psi(t)\|_2^2-f(t_0)\|\psi(t_0)\|_2^2$$
$$= C_0+[f(t)-f(t_0)]\,\|\psi(t_0)\|_2^2.$$

Thus the "energy" of Lemma 3.2 satisfies the inequality

$$\int [|\nabla\psi(x,t)|^2+v(x,t)|\psi(x,t)|^2 - (4/|x|)\,|\psi(x,t)|^2]\,dx$$
$$\leq C_0+\left(\int_{t_0}^{t}|f'(s)|\,ds+f(t)-f(t_0)\right)\|\psi(t_0)\|_2^2.$$

Since the right hand side is a locally integrable function the Gronwall-type argument used in Lemma 3.2 to prove the finiteness of $\|\nabla\psi(t)\|_2^2$ obtains. Suitable modifications can also be made to the H^2 estimate in order that it apply to this case.

4. ASYMPTOTIC BEHAVIOR

Now that we have established the existence of global solutions to (1.1), (1.2), it is natural to ask if they converge in some manner as $t\to\infty$ to a solution of some time-independent problem. We state and prove one result in this direction.

Theorem 4.1: Let $\psi(x,t)$ be a solution of (1.1). Denote by C_0 the (constant) energy of ψ:

$$C_0=\int_{\mathbb{R}^3}[|\nabla\psi(x,t)|^2+v(x,t)|\psi(x,t)|^2 - (4/|x|)\,|\psi(x,t)|^2]\,dx.$$

If ψ is a solution with negative energy, then $\|\psi(t)\|_\infty \not\to 0$ as $t\to\infty$.

Proof: Let $C_0<0$ and suppose ψ is a solution such that $\|\psi(t)\|_\infty\to 0$ as $t\to\infty$. Then

$$\int v|\psi|^2(x,t)\,dx=\int v|\psi|^{2/3}\,|\psi|^{4/3}(x,t)\,dx$$
$$\leq \|\psi(t)\|_\infty^{2/3}\,\|v(t)\,\psi^{4/3}(t)\|_1$$
$$\leq \|\psi(t)\|_\infty^{2/3}\,\|v(t)\|_3\ \|\psi^{4/3}(t)\|_{3/2}.$$

Now from Ref. 10, p. 220, estimate (c) we have

$$\|v(t)\|_3 \leq \text{const}\,\|\psi(t)\|_2^2=\text{const}\,\|\psi(0)\|_2^2.$$

Hence

$$\int v|\psi|^2(x,t)\,dx \leq \text{const}\,\|\psi(t)\|_\infty^{2/3}\ \|\psi(0)\|_2^{10/3} \to 0 \quad \text{as } t\to\infty.$$

Next we write

$$\int \frac{|\psi(x,t)|^2}{|x|}\,dx = \int_{|x|\leq 1}\frac{|\psi(x,t)|^2\,dx}{|x|}+\int_{|x|\geq 1}\frac{|\psi(x,t)|^2\,dx}{|x|}\equiv I_1+I_2.$$

By the Schwarz inequality we have

$$I_1 \leq \|\psi(t)\|_\infty\int_{|x|\leq 1}\frac{|\psi(x,t)|}{|x|}\,dx$$
$$\leq \|\psi(t)\|_\infty\left(\int_{|x|\leq 1}|\psi(x,t)|^2\,dx\right)^{1/2} \times\left(\int_{|x|\leq 1}\frac{dx}{|x|^2}\right)^{1/2} \leq \text{const}\,\|\psi(0)\|_2\ \|\psi(t)\|_\infty\to 0 \quad \text{as } t\to\infty.$$

Also we have

$$I_2=\int_{|x|\geq 1}\frac{|\psi(x,t)|^{1/2}\,|\psi(x,t)|^{3/2}\,dx}{|x|}$$
$$\leq \|\psi(t)\|_\infty^{1/2}\ \|\psi^{3/2}(t)\|_{4/3}\left(\int_{|x|\geq 1}\frac{dx}{|x|^4}\right)^{1/4}$$
$$\leq \text{const}\,\|\psi(0)\|_2^{3/2}\ \|\psi(t)\|_\infty^{1/2}\to 0 \quad \text{as } t\to\infty.$$

Thus since C_0 is a constant,

$$\lim_{t\to\infty}\|\nabla\psi(t)\|_2^2$$

exists and

$$\lim_{t\to\infty}\|\nabla\psi(t)\|_2^2=C_0<0,$$

a contradiction. This proves the theorem. [An analogous result holds for solutions of Eq. (1.2); it need not be written down explicitly.]

We conclude with a simple example showing that the condition $C_0 < 0$ is easily realized. Consider the spherically symmetric solution $\psi = \psi(r, t)$ with Cauchy data $\psi(r, 0) = \psi_0(r) = \exp(-ar/2)$, $a > 0$. From Ref. 1 or 4 we have

$$v(r, t) = 4\pi\left(\frac{1}{r}\int_0^r s^2 |\psi(s, t)|^2 \, ds + \int_0^\infty s |\psi(s, t)|^2 \, ds\right)$$

$$\leq \frac{4\pi}{r}\int_0^\infty s^2 |\psi(s, t)|^2 \, ds = \frac{1}{r} \|\psi_0\|_2^2.$$

Thus for $C_0 = C_0(a)$ we find the estimate

$$C_0(a)/4\pi$$

$$\leq \int_0^\infty r^2 [\psi_0'(r)]^2 \, dr + (\|\psi_0\|_2^2 - 4)\int_0^\infty r \, \psi_0^2(r) \, dr$$

$$\leq \frac{1}{4} a^2 \int_0^\infty r^2 \exp(-ar) \, dr + \left(\int_0^\infty r^2 \exp(-ar) \, dr - 4\right)$$

$$\times \int_0^\infty r \exp(-ar) \, dr$$

$$= 1/2a + (2/a^3 - 4)/a^2 .$$

Clearly, if a lies in a neighborhood of 3, then $C_0(a) < 0$, and the condition obtains.

We remark finally that the decay of positive energy solutions remains an open question.

ACKNOWLEDGMENT

The authors wish to thank the referee for bringing to their attention the related results of Ref. 11. There Bove, *et al.* consider an abstract form of Eq. (1.2) with $N = 2$ and with the Coulomb potential $1/r$ replaced by a bounded two-body potential. Using fixed point theorems for local contractions in Banach spaces, they establish an existence and uniqueness theorem for operator-valued solutions of the time-dependent Hartree equation on an interval $0 \leq t \leq \tau$.

*Supported by NSF Grant GP 37630.

[1] M. Reeken, J. Math. Phys. **11**, 2505 (1970).
[2] D. Sather and K. Gustafson, Rend. Mat. **4**, 723 (1971).
[3] H. Bethe and E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Academic, New York, 1957).
[4] J. H. Wolkowisky, Ind. Univ. Math. J. **22**, 551 (1972).
[5] E. H. Lieb and B. Simon, J. Chem. Phys. **61**, 735 (1974).
[6] J. M. Chadam and R. T. Glassey, "On the Maxwell-Dirac Equations with Zero Magnetic Field and Their Solution in Two Space Dimensions," to appear.
[7] I. Segal, Ann. Math. **78**, 339 (1963).
[8] T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1966), pp. 348, 410.
[9] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1966), Vol. I.
[10] L. Bers, F. John, and M. Schecter, *Partial Differential Equations* (Interscience, New York, 1964).
[11] A. Bove, G. Da Prato, and G. Fano, Comm. Math. Phys. **37**, 183 (1974).

The classical nonlinear oscillator and the coherent state

K. Bhaumik and B. Dutta-Roy

Saha Institute of Nuclear Physics, Calcutta, India
(Received 9 April 1974)

The coherent state constructed out of quantum oscillator states is employed to develop a method for solving the classical nonlinear oscillator problem. The perturbation solution of the Duffing oscillator is used to illustrate the method and to obtain the result of the classical procedure.

I. INTRODUCTION

Many of the problems of applied mathematics possess some essential features which preclude exact analytic solutions. One of these features is nonlinearity.[1] To obtain information about solutions of such equations we are forced to resort to approximations and numerical solutions. Foremost among approximation methods are the perturbation methods,[2] whereby the solution is represented by the first few terms of an asympototic expansion. These expansions may be in terms of a parameter which appears naturally in the problem (or which may be artificially introduced) or these may be expansions in terms of a coordinate. It is the rule, rather than the exception, the straightforward expansion in powers of a parameter have limited regions of validity and break down in certain regions called regions of nonuniformity. To render these expansions valid (even if in the sense of asymptotic expansions) a number of techniques have been developed in the past, for example, among many others, the method of strained coordinates (where uniformity is achieved by expanding the dependent as well as the independent variables in terms of new independent parameters), the method of multiple scale, and the method of averaging.[2] Confining our attention to canonical nonlinear oscillators, let us consider, for the sake of illustration, the Duffing nonlinear oscillator which follows the equation of motion

$$\ddot{x} + a^2 x + \beta x^3 = 0, \tag{1}$$

with boundary conditions given by, say,

$$x(t=0) = A, \quad \dot{x}(t=0) = 0. \tag{1'}$$

The straightforward perturbation approach consists in expanding the dependent variable (the coordinate) in a series in the parameter β, namely,

$$x = \sum_{n=0}^{\infty} \beta^n x_n(t), \tag{2}$$

substituting in the equation of motion [Eq. (1)], equating like powers of β to zero, and obtaining

$$\ddot{x}_0 + \omega^2 x_0 = 0, \tag{3a}$$

$$\ddot{x}_1 + \omega^2 x_1 = -x_0^3, \tag{3b}$$

wherein putting $x_0 = A\cos\omega t$ and solving, we get

$$x = A\cos\omega t + \beta A^3 \left(-\frac{3t}{8\omega}\sin\omega t + \frac{1}{32\omega^2}(\cos 3\omega t - \cos\omega t)\right) + O(\beta^2). \tag{4}$$

This two-term expansion cannot approximate the solution. As $t \to \infty$, $x_1/x_0 \to \infty$ due to the term $t\sin\omega t$, which is called for this reason the secular term. The true solution for Eq. (1) for positive β must be bounded, since multiplying Eq. (1) by $\dot{x}$, integrating, and imposing the boundary conditions we get

$$\dot{x}^2 + \omega^2 x^2 + \tfrac{1}{2}\beta x^4 = \omega^2 A^2 + \tfrac{1}{2}\beta A^4 \tag{5}$$

which describes bounded motion for $\beta > 0$, and hence the secular term expresses the nonuniformity of the straightforward perturbation expansion. Various techniques have evolved to render such approximate solutions uniformly valid, for example, the method of strained coordinates due to Lindstedt[3] and Poincaré,[4] the method of averaging used by Krylov and Bogoliubov,[5] the generalized method of averaging due to Krylov–Bogoliubov–Mitropolski[6] or by Struble,[7] by the method of averaging and variation of parameters using canonical variables by Von Zeipel,[8] and other methods.[2]

In the present work we shall use the coherent state to obtain a perturbative solution to the quantum nonlinear oscillator problem and obtain the classical solution in the appropriate limit. The use of the coherent state is motivated by the fact that it is the minimum uncertainty state of the oscillator for which the expectation value of x possesses the classical time dependence. In the sense of expectation values the coherent state is indeed a "classical state" and the difference lies in the appearance of quantum correlations which tend to zero in the classical limit ($\hbar \to 0$). An important advatange in the present method stems from the fact that in quantum mechanics the time appears in the form $\exp(iEt)$ which is a bounded function and thus the problem of secular terms does not arise. In the next section the well-known properties, which we shall need, of the coherent states are briefly reviewed. In Sec. III the method is developed and employed to solve the Duffing equation as an example. Section IV contains our concluding remarks.

II. THE COHERENT STATE

The eigenstates of the Hamiltonian

$$H_0 = \frac{p^2}{2m} + \tfrac{1}{2}m\omega^2 x^2, \tag{6}$$

for the harmonic oscillator, may easily be obtained[9] using the "annihilation" and "creation" operators

$$a = \frac{p - im\omega x}{\sqrt{2m\omega\hbar}}, \quad a^\dagger = \frac{p + im\omega x}{\sqrt{2m\omega\hbar}}, \tag{7}$$

with $[a, a^\dagger] = 1$, $[a, a] = 0 = [a^\dagger, a^\dagger]$, in terms of which the Hamiltonian may be written as

$$H_0 = \hbar\omega(a^\dagger a + \tfrac{1}{2}), \tag{8}$$

and the state of "n quanta," namely, $|n\rangle$, is the eigenstate of the Hamiltonian belonging to the eigenvalue

 Copyright © 1975 American Institute of Physics

$\hbar\omega(n+\frac{1}{2})$, and may be generated from the ground or "vacuum" state $|0\rangle$ by repeated application of the operator $a^\dagger$, namely,

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}\,|0\rangle. \tag{9}$$

A set of normalized states, known as *coherent states*,[10] may now be obtained by superposing these states, to wit,

$$\begin{aligned}|\alpha\rangle &= \exp(-|\alpha|^2/2)\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\,|n\rangle \\ &= \exp(-\alpha^* a + \alpha a^\dagger)\,|0\rangle,\end{aligned} \tag{10}$$

where α is a complex number. The last step of Eq. (10) follows from the fact that $\exp(A)\exp(B) = \exp(A+B+\frac{1}{2}[A,B])$ if A and B are operators whose commutator is a c-number. These coherent states, eigenstates of the annihilation operator a (belonging to the eigenvalue α), forming an over-complete set, are not orthogonal. In fact,

$$|\langle\alpha|\beta\rangle|^2 = \exp(-|\alpha-\beta|^2); \tag{11}$$

however, this overlap goes to zero as $|\alpha-\beta|\to\infty$. Writing

$$\alpha = -\frac{i\lambda}{2}\exp(i\omega t), \tag{12}$$

it is easily verified that

$$\langle\alpha|x|\alpha\rangle = \lambda\sqrt{\hbar/2m\omega}\cos\omega t. \tag{13}$$

Thus λ is the "amplitude" in units of the "zero-point amplitude" and the classical limit is achieved when

$$\hbar\to 0,\ \lambda\to\infty:\ \lambda\sqrt{\hbar/2m\omega}\to A \quad \text{(the corresponding classical amplitude)}. \tag{14}$$

III. THE CLASSICAL NONLINEAR OSCILLATOR AND THE COHERENT STATE

The Ehrenfest theorem[11] assures us that the expectation value of a dynamical variable satisfies the classical equation. Thus, for example, for the Duffing oscillator [see Eq. (1)]

$$\frac{d^2}{dt^2}\langle x\rangle + \omega^2\langle x\rangle + \beta\langle x^3\rangle = 0. \tag{15}$$

Since we are interested in the solution of the classical nonlinear oscillator problem, we shall consider these expectation values in the coherent state [see Eq. (10)] and take the classical limit [see Eq. (13)] and observe that for expectation values in the coherent state

$$\langle x^3\rangle = \langle x\rangle^3 + \frac{3\hbar}{2m\omega}\langle x\rangle \xrightarrow[\hbar\to 0]{} \langle x\rangle^3, \tag{16}$$

and thus $\langle x\rangle$ should yield the classical solution.

The Hamiltonian corresponding to the Duffing equation is given by

$$H = \frac{p^2}{2m} + \tfrac{1}{2}m\omega^2x^2 + \tfrac{1}{4}\beta m x^4. \tag{17}$$

We shall next adapt the coherent state $|\alpha\rangle$ to the perturbed Hamiltonian H to obtain states $|\alpha\rangle'$ in terms of the modified quantum states $|n\rangle'$ which are eigenstates of H and the energies $\hbar\omega(n+\frac{1}{2})$ of the unperturbed problem are replaced by the corresponding energies E_n'. Thus

$$|\alpha\rangle' = N_\lambda\sum_{n=0}^{\infty}\frac{(i\lambda)^n}{\sqrt{n!}}\exp(iE_n't/\hbar)\,|n\rangle', \tag{18}$$

where N_λ is a normalization constant. In view of the Ehrenfest theorem [providing Eq. (15)] and the vanishing quantum correlations in the classical limit [shown in Eq. (16)], the solution of the classical problem may be obtained from the classical limit of the quantum solution. Thus

$$x = \lim_{\substack{\hbar\to 0\\ \lambda\to\infty\\ \lambda\sqrt{\hbar/2m\omega}\to A}} {}'\langle\alpha|x|\alpha\rangle'. \tag{19}$$

Thus, using perturbation theory (retaining terms to first order), we obtain

$$E_n' = n\omega\hbar + \tfrac{3}{4}\beta\left(\frac{\hbar^2}{4m\omega^2}\right)(2n^2+2n+1) + O(\beta^2) \tag{20}$$

and

$$\begin{aligned}|n\rangle' = |n\rangle + \frac{m\beta}{4\hbar\omega}\left(\frac{\hbar}{2m\omega}\right)^2 [&-\tfrac{1}{4}\sqrt{(n+1)(n+2)(n+3)(n+4)} \\ &\times |n+4\rangle + (2n+3)\sqrt{(n+1)(n+2)}\,|n+2\rangle \\ &-(2n-1)\sqrt{n(n-1)}\,|n-2\rangle \\ &+\tfrac{1}{4}\sqrt{n(n-1)(n-2)(n-3)}\,|n-4\rangle] + \mathcal{O}(\beta^2).\end{aligned} \tag{21}$$

Thus to order β we obtain, using Eq. (19),

$$\begin{aligned}x = A\cos\left[\omega\left(1+\tfrac{3}{8}\beta\frac{A^2}{\omega^2}t\right)\right] + \frac{A^3\beta}{32\omega^2}(\cos 3\omega t - 6\cos\omega t) \\ + O(\beta^2),\end{aligned} \tag{22}$$

which agrees with the classical solution (see, for example, Von Ziepel's[8] method).

IV. CONCLUSION

We have obtained a method for the solution of the classical nonlinear oscillator problem provided the system is canonical and have shown that the perturbation method yields the same solution as that obtained by classical methods. An important advantage of the method stems from the fact that the independent variable (the time t) appears in the form $\exp(iEt)$ and is thus bounded and the problem of secular terms does not arise. It may also be interesting to use the nonperturbative solutions to the quantum nonlinear oscillator[12] to obtain interesting classical solutions.

ACKNOWLEDGMENTS

We would like to thank B. K. Shome and Anutosh Chatterjee for helpful discussions.

[1]N. Minorsky, *Nonlinear Oscillations* (Van Nostrand, Princeton, New Jersey, 1962).
[2]A.H. Nayfeh, *Perturbation Methods* (Wiley, New York, 1973).
[3]A. Lindstedt, Astron. Nach. **103**, 211 (1882); see also Ref. 2.
[4]H. Poincaré, *New Methods of Celestial Mechanics*, English translation (NASA, TTF-450, 1967), Vols. I—III.
[5]N. Krylov and N.N. Bogoliubov, *Introduction to Nonlinear Mechanics* (Princeton U.P., Princeton, New Jersey, 1947).
[6]N.N. Bogoliubov and Y.A. Mitropolski, *Asymptotic Methods in the Theory of Nonlinear Oscillation* (Gordon and Breach, New York, 1961).
[7]R.A. Struble, *Nonlinear Differential Equations* (McGraw-Hill, New York, 1962).
[8]H. Von Zeipel, Ark. Mat. Astron. Fysik, Stockholm, **11** (7), 1; see also Ref. 2.
[9]P.A.M. Dirac *The Principles of Quantum Mechanics* (Oxford U.P., New York, 1967).
[10]R.J. Glauber, Phys. Rev. **131**, 2766 (1963).
[11]L.I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1968).
[12]S.N. Biswas, K. Datta, R.P. Saxena, P.K. Srivastava, and V.S. Verma, Phys. Rev. D **4**, 3617 (1971).

An iterative method for solving a two-point boundary-value problem

C. V. Pao

Department of Mathematics, North Carolina State University, Raleigh, North Carolina 27607
(Received 28 June 1974)

The purpose of this paper is to study the existence, uniqueness, and method of construction of a nonnegative solution as well as the question of criticality for a two-point boundary-value problem arising in the transport process of n different types of particles in a rod of finite length subjecting incident fluxes and internal source. A recursion formula is derived for the calculation of the maximal and the minimal solution which are the respective limits of a monotonically nonincreasing sequence and a monotonically nondecreasing sequence. The behavior of these sequences leads to a characterization for the criticality question of the transport problem. It is shown under some physically reasonable conditions that the minimal and maximal solutions coincide so that it leads to a uniqueness theorem.

1. INTRODUCTION

In the transport process of n different types of particles in a finite rod of length $(b-a)$ the equation governing the particle's density is given by the following system of equations:

$$\left.\begin{aligned} \frac{du}{dx} + A_0(x)u = A_1(x)u + A_2(x)v + p(x) \\ -\frac{dv}{dx} + B_0(x)v = B_1(x)u + B_2(x)v + q(x) \end{aligned}\right\} \quad (a \le x \le b), \qquad (1.1)$$

where A_i, B_i $(i=0, 1, 2)$ are $n\times n$ matrices and u, v, p, q are n-vectors. The components $u_1, \ldots, u_n$ of the vector u represent the n distinct type of particles moving in the forward direction along the rod while the components $v_1, \ldots, v_n$ of v are the ones moving in the backward direction. The derivation of (1.1) is based on the balance relation for each of the n particles. The ith component of A_0u is the loss of ith type particles due to collision while the corresponding component of $(A_1u + A_2v)$ is the gain caused by the interactions. The components of p represent internal sources independent of the interactions. Similar physical interpretation applies to the second equation in (1.1) for the backward moving particles. [For a derivation of (1.1) and the related problems see Refs. 1, 2, 3 and for the corresponding time-dependent problem see Refs. 4, 5.] When the ends of the rod are subjected to incident fluxes, the boundary condition becomes

$$u(a) = u_a, \quad v(b) = v_b, \qquad (1.2)$$

where the vectors u_a, v_b are given. By physical reasons, we assume that A_0, B_0 are diagonal matrices and all the elements in the matrices A_i, B_i $(i=0, 1, 2)$ and the vectors p, q, u_a, v_b are nonnegative functions on $[a, b]$. We also assume that these elements are only piecewise continuous. The requirement of piecewise continuity rather than continuity is the most physically interesting case (cf. Refs. 3, 4).

The purpose of this paper is to characterize when the problem (1.1), (1.2) has a nonnegative solution (or no nonnegative solution) and when the solution is unique if it exists. We also present a recursion formula for the calculation of the maximal and the minimal solution (see Definition 2.2). It will be shown that the existence or nonexistence of a solution can be characterized by the convergence or nonconvergence, respectively, of the minimal sequence which is obtained from the recursion formula with zero initial iteration (see Definition 2.3). In fact, the boundedness or unboundedness of the minimal sequence is sufficient to determine whether the problem (1.1), (1.2) has a nonnegative solution. This fact leads to a characterization of the criticality problem for the system (1.1), (1.2). In this paper by a solution of (1.1), (1.2) we mean a pair of continuous vector functions (u, v) which are piecewise continuously differentiable and satisfy (1.1), (1.2) at every point x at which A_l, B_l, p, q are continuous.

The problem (1.1), (1.2) has some other physical applications. Consider, for example, a beam under static loading conditions. If the beam has nonuniform stiffness, such as a sudden change in cross section, or the beam is made of different materials, then the equations governing the deflection $w(x)$ of the beam is given by

$$(EI(x)w_{xx})_{xx} - \beta(x)w_{xx} = f(x), \qquad (1.3)$$

where $E(x)$ is the Young's modulus, $I(x)$ is the moment of inertia, $\beta(x)$ is the axial force, and $f(x)$ is the distributed forces along the beam. All of these functions can be piecewise continuous as is in most problems of interests. Now by letting $u_1 = w$, $u_2 = w_x$, $v_1 = EIw_{xx}$, $v_2 = -(EIw_{xx})_x$, Eq. (1.3) reduces to the form of (1.1) with

$$A_0 = (0), \quad A_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 \\ (EI)^{-1} & 0 \end{pmatrix}, \quad p = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$
$$B_0 = B_1 = (0), \quad B_2 = \begin{pmatrix} 0 & 1 \\ (EI)^{-1}\beta & 0 \end{pmatrix}, \quad q = \begin{pmatrix} 0 \\ f \end{pmatrix}. \qquad (1.4)$$

Hence our recursion formula can be used for the calculation of the deflection of a beam with irregular cross section.

It is to be pointed out that the corresponding time-dependent transport problem has recently been discussed by Bellman[4] for the case of constant matrices and by Pao[5] for the case of nonconstant matrices and with internal sources. The technique used in Ref. 5 is

 Copyright © 1975 American Institute of Physics

also by successive approximations, but the convergence proof is based on the contraction mapping theorem which does not seem suitable in the present situation. Here, we use the classical monotone argument which has frequently been used for the treatment of elliptic-type boundary-value problems (e.g., see Refs. 6—10). This approach also leads to a recursion formula for the calculation of a solution, but, in general, the construction of a solution depends on the choice of initial iteration so that different initial iterations may lead to distinct solutions. However, by imposing some additional conditions on the matrices A_l, B_l, we can obtain a uniqueness theorem.

2. MINIMAL AND MAXIMAL SOLUTIONS AND CRITICALITY PROBLEM

In this section we discuss the existence and construction of minimal and maximal solutions for the problem (1.1), (1.2), by a monotone iteration scheme. We also characterize the criticality problem in terms of the minimal sequence obtained from the iterations. Our process of iteration is as follows: Starting from a suitable pair of continuous functions $(u^{(0)}, v^{(0)})$ we can construct a sequence $\{u^{(k)}, v^{(k)}\}$ from the following uncoupled but inter-related initial-value problems

$$\left.\begin{aligned} &\frac{du^{(k)}}{dx} + A_0 u^{(k)} = A_1 u^{(k-1)} + A_2 v^{(k-1)} + p \\ &u^{(k)}(a) = u_a \\ &-\frac{dv^{(k)}}{dx} + B_0 v^{(k)} = B_1 u^{(k-1)} + B_2 v^{(k-1)} + q \\ &v^{(k)}(b) = v_b \end{aligned}\right\}, \quad k = 1, 2, \cdots. \tag{2.1}$$

The construction of the sequence is clear since A_0, B_0 are diagonal matrices and the right-side of both equations in (2.1) are known functions. In fact, if we write

$$A_0(x) = \operatorname{diag}(a_1^{(0)}(x), \ldots, a_n^{(0)}(x)),$$

$$B_0(x) = \operatorname{diag}(b_1^{(0)}(x), \ldots, b_n^{(0)}(x))$$

and define

$$\left.\begin{aligned} \alpha_i(x_1, x_2) &= \exp\left(-\int_{x_1}^{x_2} a_i^{(0)}(\eta)\, d\eta\right) \\ \beta_i(x_1, x_2) &= \exp\left(-\int_{x_1}^{x_2} b_i^{(0)}(\eta)\, d\eta\right) \end{aligned}\right\}, \quad i = 1, \ldots, n, \tag{2.2}$$

then an integration of the first equation in (2.1) from a to x and the second equation from x to b lead to the following recursion formula:

$$\begin{aligned} u_i^{(k)}(x) &= \alpha_i(a, x)(u_a)_i + \int_a^x \alpha_i(\xi, x)\Big(\sum_{j=1}^{n} [a_{ij}^{(1)}(\xi) u_j^{(k-1)}(\xi) \\ &\quad + a_{ij}^{(2)}(\xi) v_j^{(k-1)}(\xi)] + p_i(\xi)\Big)\, d\xi, \\ v_i^{(k)}(x) &= \beta_i(x, b)(v_b)_i + \int_x^b \beta_i(x, \xi)\Big(\sum_{j=1}^{n} [b_{ij}^{(1)}(\xi) u_j^{(k-1)}(\xi) \\ &\quad + b_{ij}^{(2)}(\xi) v_j^{(k-1)}(\xi)] + q_i(\xi)\Big)\, d\xi, \quad i = 1, \ldots, n. \end{aligned} \tag{2.3}$$

In (2.3), the functions $a_{ij}^{(l)}$, $b_{ij}^{(l)}$, u_i, v_i, p_i, q_i ($l = 1, 2$, $i, j = 1, \ldots, n$) are the respective elements of A_l, B_l, u, v, p, and q. Notice that, for each $k = 1, 2, \ldots$, $u_i^{(k)}$, $v_i^{(k)}$ are continuous on $[a, b]$ and are continuously differentiable at each point x where the prescribed functions are continuous. Set

$$D_\alpha = \operatorname{diag}(\alpha_1, \ldots, \alpha_n), \quad D_\beta = \operatorname{diag}(\beta_1, \ldots, \beta_n).$$

Then (2.3) may be written in the vector form

$$\begin{aligned} u^{(k)}(x) &= D_\alpha(a, x) u_a + \int_a^x D_\alpha(\xi, x)[A_1(\xi) u^{(k-1)}(\xi) \\ &\quad + A_2(\xi) v^{(k-1)}(\xi) + p(\xi)]\, d\xi \\ v^{(k)}(x) &= D_\beta(x, b) v_b + \int_x^b D_\beta(x, \xi)[B_1(\xi) u^{(k-1)}(\xi) \\ &\quad + B_2(\xi) v^{(k-1)}(\xi) + q(\xi)]\, d\xi \end{aligned} \tag{2.4}$$

Hence our crucial problem is to determine when the sequence $\{u^{(k)}, v^{(k)}\}$ converges to a solution of (1.2), (1.2). As we shall see in the subsequent discussions, the convergence of this sequence depends on the existence of an upper solution and the choice of the initial iteration $(u^{(0)}, v^{(0)})$. The definition of an upper solution is given in the following and it is similar to that used in Refs. 8—10 for the treatment of certain elliptic-type boundary-value problems. We recall that for any two vectors $u = (u_1, \ldots, u_n)$, $v = (v_1, \ldots, v_n)$ the inequality $u \leq v$ means that $u_i \leq v_i$ for each $i = 1, \ldots, n$.

Definition 2.1: A pair of nonnegative functions $(\tilde{u}, \tilde{v})$ is called an upper solution of (1.1), (1.2) if it is differentiable at every point where A_l, B_l, p, q, ($l = 0, 1, 2$) are continuous and satisfies the following conditions:

$$\text{(i)} \quad \frac{d\tilde{u}}{dx} + A_0 \tilde{u} \geq A_1 \tilde{u} + A_2 \tilde{v} + p, \quad \tilde{u}(a) \geq u_a; \tag{2.5}$$

$$\text{(ii)} \quad -\frac{d\tilde{v}}{dx} + B_0 \tilde{v} \geq B_1 \tilde{u} + B_2 \tilde{v} + q, \quad \tilde{v}(b) \geq v_b. \tag{2.6}$$

The above definition implies that every nonnegative solution of (1.1), (1.2) is also an upper solution.

Definition 2.2: Let $(\bar{u}, \bar{v})$, $(\underline{u}, \underline{v})$ be any nonnegative solutions of (1.1), (1.2) with $\underline{u} \leq \bar{u} \leq \tilde{u}$, $\underline{v} \leq \bar{v} \leq \tilde{v}$. Then $(\bar{u}, \bar{v})$ and $(\underline{u}, \underline{v})$ are called maximal and minimal solutions, respectively, if every other nonnegative solution (u, v) with $u \leq \tilde{u}$, $v \leq \tilde{v}$ satisfies the relations

$$\underline{u}(x) \leq u(x) \leq \bar{u}(x), \quad \underline{v}(x) \leq v(x) \leq \bar{v}(x) \quad \text{for all } x \in [a, b].$$

In the following lemmas we show some monotone properties of the sequence determined from (2.4).

Lemma 2.1: Let $(\tilde{u}, \tilde{v})$ be an upper solution of (1.1), (1.2). Then the sequence $\{\bar{u}^{(k)}, \bar{v}^{(k)}\}$ given by the formula (2.4) with $u^{(0)} = \tilde{u}$, $v^{(0)} = \tilde{v}$ is monotonically nonincreasing, that is, $\bar{u}^{(k+1)} \leq \bar{u}^{(k)}$, $\bar{v}^{(k+1)} \leq \bar{v}^{(k)}$ for every $k = 0, 1, 2, \cdots$.

Proof: It suffices to show that the sequence determined from (2.1) is monotonically nonincreasing. We first show that $\bar{u}^{(1)} \leq \bar{u}^{(0)}$, where $\bar{u}^{(0)} \equiv u^{(0)} \equiv \tilde{u}$. Let $w = \bar{u}^{(1)} - \bar{u}^{(0)}$. Then, by (2.1) and (2.5),

$$\frac{dw}{dx} + A_0 w = (A_1 \bar{u}^{(0)} + A_2 \bar{v}^{(0)} + p) - \left(\frac{d\bar{u}^{(0)}}{dx} + A_0 \bar{u}^{(0)}\right) \leq 0, \tag{2.7}$$

$$w(a) = \bar{u}^{(1)}(a) - \bar{u}^{(0)}(a) = u_a - \bar{u}^{(0)}(a) \leq 0. \tag{2.8}$$

By a transformation $w \to e^{\lambda x}w$ for some $\lambda > 0$, if necessary, we may assume that the elements $a_i^{(0)}$ of A_0 is strictly positive. Suppose, by contradiction, that $w(x)$ has a component $w_i(x)$ which assumes a positive maximum at some point $x_0 \in [a,b]$. Then, clearly, $x_0 \neq a$ and $a_i^{(0)}(x_0)w_i(x_0) > 0$. Since $[dw_i(x_0)/dx] \geq 0$ for $a < x_0 \leq b$ we see that the left side of (2.7) is positive, which leads to a contradiction. Hence $w(x) \leq 0$, that is, $\overline{u}^{(1)} \leq \overline{u}^{(0)}$. The same argument shows that $\overline{v}^{(1)} \leq \overline{v}^{(0)}$. We complete the proof by induction. Assume $\overline{u}^{(k)} \leq \overline{u}^{(k-1)}$, $\overline{v}^{(k)} \leq \overline{v}^{(k-1)}$. Then the function $w(x) = \overline{u}^{(k+1)}(x) - \overline{u}^{(k)}(x)$ satisfies the equation

$$\frac{dw}{dx} + A_0 w = A_1(\overline{u}^{(k)} - \overline{u}^{(k-1)}) + A_2(\overline{v}^{(k)} - \overline{v}^{(k-1)}) \tag{2.9}$$

and the initial condition $w(a) = 0$. Since the elements of A_1, A_2 are nonnegative, we see that the right side of (2.9) is nonpositive and thus

$$\frac{dw}{dx} + A_0 w \leq 0, \quad w(a) = 0. \tag{2.10}$$

It follows from the same proof as for $(\overline{u}^{(1)} - \overline{u}^{(0)})$ that $w(x)$ must be nonpositive, that is, $\overline{u}^{(k+1)} \leq \overline{u}^{(k)}$. By a similar argument, we have $\overline{v}^{(k+1)} \leq \overline{v}^{(k)}$. This completes the proof of the lemma.

Lemma 2.2: The sequence $\{\underline{u}^{(k)}, \underline{v}^{(k)}\}$ given by (2.4) with $u^{(0)} = v^{(0)} = 0$ is monotonically nondecreasing, that is, $\underline{u}^{(k+1)} \geq \underline{u}^{(k)}$, $\underline{v}^{(k+1)} \geq \underline{v}^{(k)}$ for every $k = 0, 1, 2, \cdots$.

Proof: The proof is similar to that given in Lemma 2.1 and we sketch it as follows: For $k = 1$, we have

$$\frac{d\underline{u}^{(1)}}{dx} + A_0 \underline{u}^{(1)} = p \geq 0, \quad \underline{u}^{(1)}(a) = u_a \geq 0, \tag{2.11}$$

$$-\frac{d\underline{v}^{(1)}}{dx} + B_0 \underline{v}^{(1)} = q \geq 0, \quad \underline{v}^{(1)}(b) = v_b \geq 0. \tag{2.12}$$

Let $w = -\underline{u}^{(1)}$. Then (2.11) is reduced to

$$\frac{dw}{dx} + A_0 w \leq 0, \quad w(a) \leq 0, \tag{2.13}$$

which is of the same form as in (2.7), (2.8). Hence from the proof of Lemma 2.1 we have $w(x) \leq 0$, that is, $\underline{u}^{(1)}(x) \geq 0$. The same treatment for $\underline{v}^{(1)}$ in (2.12) leads to $\underline{v}^{(1)} \geq 0$. By an induction argument as in the proof of Lemma 2.1 we conclude that $\underline{u}^{(k+1)} \geq \underline{u}^{(k)}$, $\underline{v}^{(k+1)} \geq \underline{v}^{(k)}$ for every k.

Definition 2.3: The sequences $\{\overline{u}^{(k)}, \overline{v}^{(k)}\}$ and $\{\underline{u}^{(k)}, \underline{v}^{(k)}\}$ given in Lemmas 2.1 and 2.2 are called maximal and minimal sequences, respectively.

Thus a maximal sequence is nonincreasing while the minimal sequence is nondecreasing. In the following theorem we show the convergence of these sequences.

Theorem 2.1: Let there exist an upper solution $(\tilde{u}, \tilde{v})$. Then the maximal sequence $\{\overline{u}^{(k)}, \overline{v}^{(k)}\}$ converges uniformly from above to a maximal solution $(\overline{u}, \overline{v})$ of (1.1), (1.2) while the minimal sequence $\{\underline{u}^{(k)}, \underline{v}^{(k)}\}$ converges uniformly from below to a minimal solution $(\underline{u}, \underline{v})$. Furthermore,

$$\begin{aligned} &0 \leq \underline{u}^{(1)} \leq \underline{u}^{(2)} \leq \cdots \leq \underline{u} \leq \overline{u} \leq \cdots \leq \overline{u}^{(2)} \leq \overline{u}^{(1)} \leq \tilde{u}, \\ &0 \leq \underline{v}^{(1)} \leq \underline{v}^{(2)} \leq \cdots \leq \underline{v} \leq \overline{v} \leq \cdots \leq \overline{v}^{(2)} \leq \overline{v}^{(1)} \leq \tilde{v}. \end{aligned} \tag{2.14}$$

Proof: We first show that $\underline{u}^{(k)} \leq \overline{u}^{(k)}$ and $\underline{v}^{(k)} \leq \overline{v}^{(k)}$ for every $k = 0, 1, 2, \cdots$. The proof is by induction. Since $\overline{u}^{(0)} = \tilde{u}$, $\overline{v}^{(0)} = \tilde{v}$ and $\underline{u}^{(0)} = \underline{v}^{(0)} = 0$, it is clear that $\underline{u}^{(0)} \leq \overline{u}^{(0)}$ and $\underline{v}^{(0)} \leq \overline{v}^{(0)}$. Suppose that $\underline{u}^{(k-1)} \leq \overline{u}^{(k-1)}$ and $\underline{v}^{(k-1)} \leq \overline{v}^{(k-1)}$. Then from (2.1) and the nonnegative properties of the matrices A_l, B_l, $l = 0, 1, 2$, we obtain

$$\begin{aligned} &\frac{d}{dx}(\overline{u}^{(k)} - \underline{u}^{(k)}) + A_0(\overline{u}^{(k)} - \underline{u}^{(k)}) \\ &= A_1(\overline{u}^{(k-1)} - \underline{u}^{(k-1)}) + A_2(\overline{v}^{(k-1)} - \underline{v}^{(k-1)}) \geq 0, \\ &\overline{u}^{(k)}(a) - \underline{u}^{(k)}(a) = 0, \end{aligned} \tag{2.15}$$

$$\begin{aligned} &-\frac{d}{dx}(\overline{v}^{(k)} - \underline{v}^{(k)}) + B_0(\overline{v}^{(k)} - \underline{v}^{(k)}) \\ &= B_1(\overline{u}^{(k-1)} - \underline{u}^{(k-1)}) + B_2(\overline{v}^{(k-1)} - \underline{v}^{(k-1)}) \geq 0, \\ &\overline{v}^{(k)}(b) - \underline{v}^{(k)}(b) = 0. \end{aligned} \tag{2.16}$$

By the same argument as in the proof of Lemma 2.1 with $w(x) = \overline{u}^{(k)} - \underline{u}^{(k)}$ [or $w(x) = \overline{v}^{(k)} - \underline{v}^{(k)}$] we conclude that $\overline{u}^{(k)} \geq \underline{u}^{(k)}$, $\overline{v}^{(k)} \geq \underline{v}^{(k)}$. This fact together with the results in Lemmas 2.1 and 2.2 imply that the sequence $\{\overline{u}^{(k)}, \overline{v}^{(k)}\}$ is monotonically nonincreasing and is bounded from below by $(0,0)$ and the sequence $\{\underline{u}^{(k)}, \underline{v}^{(k)}\}$ is monotonically nondecreasing and is bounded from above by $(\tilde{u}, \tilde{v})$. Therefore, the pointwise limits

$$\begin{aligned} &\lim_{k\to\infty} \overline{u}^{(k)}(x) = \overline{u}(x), \quad \lim_{k\to\infty} \overline{v}^{(k)}(x) = \overline{v}(x), \\ &\lim_{k\to\infty} \underline{u}^{(k)}(x) = \underline{u}(x), \quad \lim_{k\to\infty} \underline{v}^{(k)}(x) = \underline{v}(x) \end{aligned} \tag{2.17}$$

exist and the inequalities in (2.14) hold. To show that $(\overline{u}, \overline{v})$ and $(\underline{u}, \underline{v})$ are solutions of the problem (1.1), (1.2), we recall that the sequences $\{\overline{u}^{(k)}, \overline{v}^{(k)}\}$ and $\{\underline{u}^{(k)}, \underline{v}^{(k)}\}$ both satisfy the recursion formula (2.4) and, in view of (2.14), are uniformly bounded. This implies that both sequences are equicontinuous and thus, by an elementary argument, their limits $(\overline{u}, \overline{v})$, $(\underline{u}, \underline{v})$ are continuous. By the Dini's theorem, the convergence of these sequences are uniform. It follows by letting $k \to \infty$ in (2.4) with $\{u^{(k)}, v^{(k)}\}$ representing either $\{\overline{u}^{(k)}, \overline{v}^{(k)}\}$ or $\{\underline{u}^{(k)}, \underline{v}^{(k)}\}$ that both $(\overline{u}, \overline{v})$ and $(\underline{u}, \underline{v})$ are solutions of the integral equation

$$\begin{aligned} u(x) &= D_\alpha(a,x)u_a \\ &\quad + \int_a^x D_\alpha(\xi,x)[A_1(\xi)u(\xi) + A_2(\xi)v(\xi) + p(\xi)]\,d\xi, \\ v(x) &= D_\beta(x,b)v_b \\ &\quad + \int_x^b D_\beta(x,\xi)[B_1(\xi)u(\xi) + B_2(\xi)v(\xi) + q(\xi)]\,d\xi. \end{aligned} \tag{2.18}$$

Now since $(\overline{u}, \overline{v})$ and $(\underline{u}, \underline{v})$ are continuous, the integrals in (2.18) are continuously differentiable at any point x where A_l, B_l $(l = 0, 1, 2)$ and p, q are continuous. This implies, by direct differentiation of (2.18), that $(\overline{u}, \overline{v})$ and $(\underline{u}, \underline{v})$ are classical solutions of the boundary-value problem (1.1), (1.2). Finally, if (u, v) is any nonnegative solution of (1.1), (1.2) with $u \leq \tilde{u}$, $v \leq \tilde{v}$, then by the same induction argument it is easily seen that $u \leq \overline{u}^{(k)}$, $v \leq \overline{v}^{(k)}$ and $u \geq \underline{u}^{(k)}$, $v \geq \underline{v}^{(k)}$ for every $k = 0, 1, 2, \cdots$. Hence we have $\underline{u} \leq u \leq \overline{u}$ and $\underline{v} \leq v \leq \overline{v}$. This shows that

$(\bar{u},\bar{v})$ is a maximal solution and $(\underline{u},\underline{v})$ is a minimal solution. The proof of the theorem is completed.

It is seen from Theorem 2.1 that the convergence of the maximal and minimal sequences depends on the existence of an upper solution. Since, by definition, any nonnegative solution of the problem (1.1), (1.2) is an upper solution, we see that the minimal sequence $\{\underline{u}^{(k)},\underline{v}^{(k)}\}$ must converge to the minimal solution $(\underline{u},\underline{v})$ unless the problem (1.1), (1.2) has no solution. This observation leads to the following conclusion:

Theorem 2.2: A necessary and sufficient condition for the problem (1.1), (1.2) to have a nonnegative solution is that the minimal sequence $\{\underline{u}^{(k)},\underline{v}^{(k)}\}$ converges.

The above theorem can be used to characterize the criticality problem in terms of the minimal sequence. We recall that the problem (1.1), (1.2) is subcritical or critical according to whether it has a unique or more than one nonnegative solution.

Theorem 2.3: The problem (1.1), (1.2) is subcritical if the minimal sequence $\{\underline{u}^{(k)},\underline{v}^{(k)}\}$ converges to a unique nonnegative solution, and is critical, if it converges but the solution is not unique.

Proof: The proof follows directly from definition and the results in Theorem 2.2.

Since the minimal sequence is monotonically nondecreasing, a necessary and sufficient condition for the pointwise convergence of $\{\underline{u}^{(k)},\underline{v}^{(k)}\}$ is that it be bounded on $[a,b]$. It is shown in the proof of Theorem 2.1 that the pointwise convergence of $\{\underline{u}^{(k)},\underline{v}^{(k)}\}$ implies its uniform convergence so that the limit $(\underline{u},\underline{v})$ is a classical solution of (1.1), (1.2). We therefore have the following:

Corollary: The problem (1.1), (1.2) is critical or subcritical if and only if the minimal sequence $\{\underline{u}^{(k)},\underline{v}^{(k)}\}$ is bounded.

3. UNIQUENESS PROBLEM

In the preceding section it is shown that if there exists an upper solution $(\tilde{u},\tilde{v})$, then the problem (1.1), (1.2) has maximal and minimal solutions which can be constructed by successive integrations of the recursion formula (2.4). However, these solutions are not necessarily the same so that there is no uniqueness theorem. Consider, for example, the simple problem:

$$\begin{aligned} &\frac{du}{dx}+u=v, \quad u(0)=0,\\ &-\frac{dv}{dx}+v=(1+\beta^2)u, \quad v(b)=0, \end{aligned} \tag{3.1}$$

where $\beta>0$ is a given constant and u, v are scalar functions. If the length b of the rod is taken to be $\beta^{-1}(\pi - \arctan\beta)$, then the pair $(u,v)=(\sin\beta x,\ \sin\beta x+\beta\cos\beta x)$ is a nontrivial solution of (3.1). Therefore, in order to insure the uniqueness of a solution, it is necessary to impose some additional conditions on either the matrices A_l, B_l or the length of the interval $[a,b]$. In this section we impose some additional conditions on A_l, B_l. Specifically, we have the following:

Theorem 3.1: Assume that

$$\left.\begin{aligned} &\sum_{i=1}^{n}[a_{ij}^{(1)}(x)+b_{ij}^{(1)}(x)]<a_j^{(0)}(x)\\ &\sum_{i=1}^{n}[a_{ij}^{(2)}(x)+b_{ij}^{(2)}(x)]<b_j^{(0)}(x) \end{aligned}\right\},[x\in(a,b)] \quad j=1,\ldots,n. \tag{3.2}$$

Then the maximal solution $(\bar{u},\bar{v})$ and the minimal solution $(\underline{u},\underline{v})$ obtained in Theorem 2.1 coincide on $[a,b]$, that is, $\bar{u}(x)=\underline{u}(x)$, $\bar{v}(x)=\underline{v}(x)$ for $x\in[a,b]$.

Proof: Let $\lambda>0$ be an arbitrary constant and let $w=e^{-\lambda x}(\bar{u}-u)$, $w^*=e^{-\lambda x}(\bar{v}-\underline{v})$. Then from (2.14), $w(x)\geq 0$, $w^*(x)\geq 0$ for each $x\in[a,b]$. Since both $(\bar{u},\bar{v})$, $(\underline{u},\underline{v})$ are solutions of the problem (1.1), (1.2), we have

$$\frac{dw}{dx}=-\lambda w+(A_1-A_0)w+A_2w^*, \quad w(a)=0, \tag{3.3}$$

$$-\frac{dw^*}{dx}=\lambda w^*+B_1w+(B_2-B_0)w^*, \quad w(b)=0. \tag{3.4}$$

Addition of Eqs. (3.3), (3.4) leads to

$$\begin{aligned} &\frac{d}{dx}(w-w^*)\\ &\quad=-\lambda(w-w^*)+(A_1+B_1-A_0)w+(A_2+B_2-B_0)w^*. \end{aligned} \tag{3.5}$$

Let $C=A_1+B_1-A_0$, $D=A_2+B_2-B_0$ and let the elements of C and D be c_{ij}, d_{ij}, respectively. Then (3.5) may be written as

$$\begin{aligned} &\frac{d}{dx}(w_i-w_i^*)=-\lambda(w_i-w_i^*)+\sum_{j=1}^{n}(c_{ij}w_j+d_{ij}w_j^*),\\ &i=1,\ldots,n, \end{aligned} \tag{3.6}$$

where w_i, w_i^* are the corresponding components of w and w^*. By adding the equations in (3.6) and using the condition (3.2) we obtain

$$\begin{aligned} &\frac{d}{dx}\left(\sum_{i=1}^{n}(w_i-w_i^*)\right)\\ &\quad=-\lambda\sum_{i=1}^{n}(w_i-w_i^*)+\sum_{j=1}^{n}\left[\left(\sum_{i=1}^{n}c_{ij}\right)w_j+\left(\sum_{i=1}^{n}d_{ij}\right)w_j^*\right]\\ &\quad\leq-\lambda\sum_{i=1}^{n}(w_i-w_i^*) \quad [x\in(a,b)] \end{aligned} \tag{3.7}$$

The above inequality shows that the scalar function

$$f(x)\equiv\sum_{i=1}^{n}[w_i(x)-w_i^*(x)] \quad (x\in[a,b])$$

is strictly decreasing in (a,b) so long as it is different from zero. This implies that, for any fixed $x\in[a,b]$,

$$f(x)\leq f(a)=\sum_{i=1}^{n}[0-w_i^*(a)]\leq 0$$

and

$$f(x)\geq f(b)=\sum_{i=1}^{n}[w_i(b)-0]\geq 0.$$

Therefore, we must have $f(x)=0$ for all $x\in[a,b]$. In view of the equality relation in (3.7) we obtain

$$\sum_{j=1}^{n}\left[\left(\sum_{i=1}^{n} c_{ij}(x)\ w_j(x)\right)+\left(\sum_{i=1}^{n} d_{ij}(x)\right) w_j^*(x)\right]=0 \quad [x\in(a,b)]. \tag{3.8}$$

Since $w_j(x)$, $w_j^*(x)$ are nonnegative and, by (3.2), $\sum_{i=1}^{n} c_{ij}(x)$ and $\sum_{i=1}^{n} d_{ij}(x)$ are nonpositive for $x\in(a,b)$, $j=1,\ldots,n$, we have

$$\sum_{i=1}^{n}[c_{ij}(x)]w_j(x)=\left(\sum_{i=1}^{n} d_{ij}(x)\right) w_j^*(x)=0 \quad [x\in(a,b)]. \tag{3.9}$$

It follows from (3.2) and the continuity of w_j, w_j^* that $w_j(x)$ and $w_j^*(x)=0$ for each $x\in[a,b]$, $j=1,\ldots,n$. This proves $\bar{u}=\underline{u}$, $\bar{v}=\underline{v}$ and thus the theorem.

Since every nonnegative solution (u,v) of (1.1), (1.2) is also an upper solution, the above theorem leads to the following conclusion.

Therem 3.2: Let there exist an upper solution and let the condition (3.2) hold. Then the problem (1.1), (1.2) has a unique nonnegative solution which can be constructed from (2.4) with $u^{(0)}=v^{(0)}=0$.

Proof: Let (u,v) be any nonnegative solution. Then the pair $\tilde{u}=u$, $\tilde{v}=v$ is also an upper solution and thus the process of construction given in (2.4) insures that $\bar{u}=u$, $\bar{v}=v$. By Theorem 3.1, we have $u=\underline{u}$, $v=\underline{v}$. Since this is true for any nonnegative solution $(u,\bar{v})$, we conclude that the problem (1.1), (1.2) has only one such solution. The construction of the solution follows from Theorem 2.1.

Remark: From the example given in (3.1) we see that for any $\epsilon>0$, no matter how small it is, the condition

$$\sum_{i=1}^{n} c_{ij}(x)\leq 0, \quad \sum_{i=1}^{n} d_{ij}(x)\leq\epsilon, \quad j=1,\ldots,n, \tag{3.10}$$

is not sufficient to insure the uniqueness problem. Thus, except the case of pure scattering (that is, with $\epsilon=0$), the conditions in (3.2) cannot be improved significantly without some restrictions on the interval $[a,b]$. (See Ref. 11 for the case of pure scattering.) On the other hand, if the solutions of the problem (1.1), (1.2) are not unique, that is, if the maximal and minimal solutions do not coincide, then the physical system is critical and the corresponding value $(b-a)$ is the critical length of the rod. The question of determining the critical length of the rod in terms of the matrices A_l, B_l needs further exploration.

[1]G.M. Wing, *An Introduction to Transport Theory* (Wiley, New York, 1962).
[2]G.H. Meyer, *Initial Value Methods for Boundary Value Problems* (Academic, New York, 1973).
[3]R.C. Allen, Jr. and G.M. Wing, J. Math. Anal. Appl. 29, 141 (1970).
[4]R. Bellman, J. Math. Anal. Appl. 43, 232 (1973).
[5]C.V. Pao, J. Math. Anal. Appl., (to appear).
[6]A. McNabb, J. Math. Anal. Appl. 3, 133 (1961).
[7]D.S. Cohen, J. Math. Mech. 17, 209 (1967).
[8]H.B. Keller, Arch. Ratl. Mech. Anal. 35, 363 (1969).
[9]H. Amman, Indiana Univ. Math. J. 21, 125 (1971).
[10]D.H. Sattinger, Indiana Univ. Math. J. 21, 979 (1972).
[11]R. Bellman, *Matrix Analysis* (McGraw-Hill, New York, 1970), 2nd ed.

$O(4,2)$ symmetry and the classical Kepler problem

K. C. Tripathy, Rakesh Gupta, and J. D. Anand

Department of Physics, University of Delhi, Delhi-110007, India
(Received 8 July 1974)

The spectrum generating algebra associated with the negative energy motions of a classical dynamical system, namely the Kepler problem, has been systematically studied with the aid of the Poisson brackets. The canonical map between our realization and that of an earlier analysis by Barut and Bornzin has been established.

I. INTRODUCTION

Since the days of Laplace,[1] the Kepler problem has been extensively studied as this system furnishes an excellent prototype for building up dynamical symmetry models. In a previous communication,[2] we systematically analysed the possible dynamical symmetries associated with the Kepler problem and the three-dimensional harmonic oscillator with the aid of Poisson brackets. It was shown that in the Kepler problem (i) there are two conserved vectors (the well-known Lenz vector **A** and the vector **L**×**A**) and three second rank, traceless tensors which together with the angular-momentum vector **L** separately close to the Lie algebra of $O(4)$ and $SU(3)$, respectively (for $E<0$); (ii) because of the existence of the preferential points (perihelion/apehelion) on the orbit, it was shown why $O(4)$ but not $SU(3)$ is a dynamical group for the underlying classical dynamical system; (iii) since the vectors and tensors involve arbitrary functions of energy, there is an ambiguity in the relation of the Hamiltonian and the Casimir operators of the dynamical group.

Here, following Dothan,[3] we address ourselves to a systematic investigation of the "spectrum generating algebras" for the Kepler problem. We make use of the following assumptions.

1. Let M be the state space of the classical dynamical system. It admits a maximal dynamical symmetry group K (correspondingly, $\mathbb{K}$ is the Lie algebra of infinetismal canonical transformations on M) actingly transitively on each energy surface $M_E \approx K/K_0$, where K_0 is the stability subgroup of some point on M_E.[4] This implies that all the orbits of the dynamical system are diffeomorphic to one another and that the Hamiltonian is a certain function of the canonical invariants of K.

2. The vector field X_H on M (H, the Hamiltonian of the system) generates a global action in Ω (Ω, the symplectic manifold or the phase space of the dynamical system) of $O(2)$ $[O(1,1)]$ for compact (noncompact) orbits. This defines the "Hamiltonian group" G_H $[O(2)$ or $O(1,1)]$ and the Hamiltonian appears as a function of *single element* of the Lie algebra.

3. We now define the spectrum generating algebra $\mathbb{G}$ with the following properties: (a) It contains $\mathbb{K}\oplus\mathbb{G}_H$ ($\mathbb{G}_H$ is the corresponding Lie algebra for the Hamiltonian group) in such a way that $\mathbb{G}_H$ commutes with all the elements in $\mathbb{K}$. (b) The compact and noncompact orbits may be realised corresponding to different open intervals of the energy and correspondingly there exists analytic continuation within the above submanifolds. It is only in this sense that the group G admits a global canonical action in the entire phase space Ω. (c) The elements of $\mathbb{G}$ satisfy the classical equation of motion,

$$\frac{\partial \mathbb{G}}{\partial t} - [\mathbb{G}, H] = 0.$$

This implies that the elements of the algebras $\mathbb{K}$ and $\mathbb{G}_H$ are independent of time and the rest of the elements of $\mathbb{G}$ have explicit time dependence.

Our material is arranged as follows. In Sec. II, we briefly recapitulate the ideas underlying the spectrum generating algebra for the Kepler problem. Then, we outline the general method of constructing the elements of this algebra with the aid of equal-time Poisson brackets. The "symmetry algebra" for $E<0$ is $O(4)$ which possesses a local canonical action on the energy surface $M_E \approx O(4)/O(2)$; i.e., each energy surface is a homogeneous space with the stability subgroup $O(2)$. We now demand that the "spectrum generating algebra" $\mathbb{G}$ be such that the commutant of $O(2)$ is precisely $O(4)$ which leads to the structure of $\mathbb{G}=O(4,2)$. The representations of $O(4,2)$ possess the following properties:

(i) A unitary irreducible representation of $O(4,2)$ on restriction to $O(4,1)$ remains irreducible. We recall that the group $O(4,1)$ is identified as the "group of quantum numbers" in the quantum counterpart of the Kepler problem by the Boulder School led by Barut. Dothan[3] calls it the spectrum generating group since it describes the multiplicity structure of the energy levels (later, on inclusion of the Hamiltonian as a generator of the group, he adopts $O(4,2)$ as the spectrum generating group of the Kepler problem).

(ii) When restricted to $O(4)$, it splits up into the direct sum of $\oplus D(n, k_0=0)$, each $D(n, k_0=0)$ having multiplicity one.[5] We note here that the Casimir operator (the quadratic one) of $O(4)$ commutes with the Hamiltonian group generated by L_{56}. We obtain various realisations of $\mathbb{G}$ and in the special case, rederive the well-known results of Dothan.[3]

In Sec. III, we examine the correspondence of our realisations with that of an earlier analysis by Barut and Bornzin.[6] Starting with the usual nilpotent Heisenberg algebra and with the given spectrum of the Hamiltonian with the $1/q$ potential, these authors generate the Lie algebra of $O(2,1)$ and subsequently taking $O(2,1)$ as the building block construct the Lie algebra of $O(4,2)$. In order to recover our present realisations from that of Barut and Bornzin, we first make a canonical transformation to the elements of $\mathbb{G}$ given in Ref. 6:

$$L'^{(F)}_{ab} = e^{F(q,p)} L_{ab}$$

 Copyright © 1975 American Institute of Physics

where, $F(\mathbf{q},\mathbf{p})$ is an arbitrary differential function of the coordinates q and p. We note that under the canonical transformation, the elements $L_{ab}^{(F)}$ preserve the poisson bracket relations. Now, taking a suitable linear combination of these transformed elements $L_{ab}^{(F)}$, we rederive our realisations for the Lie algebra $\mathbb{G}$.

In Sec. IV, we briefly comment on our choice of $O(4,2)$ as the "spectrum generating algebra." Some light is also shed on the analytic continuation when the Hamiltonian is identified with a noncompact generator L_{46}. The bearing of our analysis in studying the symmetry algebras for a perturbed Hamiltonian dynamical system in celestial mechanics is also briefly commented upon.

II. SPECTRUM GENERATING ALGEBRAS

A. Introduction

Let M = the state space of a classical dynamical system. In general, it could be of infinite dimension; and $F(M)$ = the C^∞ real-valued functions defined on M. Let $F(M)=\{f_1, f_2, \ldots, f_n\}$, and Ω = the symplectic closed form on M such that $d\Omega=0$ (i.e., the exterior derivative is zero). Ω is called the phase space for the underlying dynamical system and in the canonical coordinate system is given by[7]

$$\Omega=\sum_i dp_i\wedge dq_i - dH\wedge dt. \tag{2.1}$$

For each $f_a\in F(M)$, the differential df_a is a covariant vector field and the contravariant vector field X_{f_a} on M is defined as

$$df_a = X_{f_a}\ \lrcorner\ \Omega=\sum_i\left(\frac{\partial f_a}{\partial p_i}\,dp_i+\frac{\partial f_a}{\partial q_i}\,dq_i\right)+\frac{\partial f_a}{\partial t}\,dt \tag{2.2}$$

and

$$X_{f_a}=\frac{\partial}{\partial t}+\sum_i\left(\frac{\partial f_a}{\partial p_i}\frac{\partial}{\partial q_i}-\frac{\partial f_a}{\partial q_i}\frac{\partial}{\partial p_i}\right) \tag{2.3}$$

where $X_{f_a}\lrcorner\Omega$ defines the contraction of Ω by X_{f_a}.[8] Let $\{X_{f_a}\}=V(M)$. For every f_a, $f_b\in F(M)$, we write

$$\begin{aligned}X_{f_a}(f_b)&=\frac{\partial f_b}{\partial t}+\sum_i\left(\frac{\partial f_a}{\partial p_i}\frac{\partial f_b}{\partial q_i}-\frac{\partial f_a}{\partial q_i}\frac{\partial f_b}{\partial p_i}\right)\\&=\frac{\partial f_b}{\partial t}+[f_a, f_b]\end{aligned} \tag{2.4}$$

Under the Poisson bracket (2.4), the real vector space $F(M)$ becomes a Lie algebra. The map $f_a\to X_{f_a}$ is a Lie algebra homomorphism of $F(M)$ into $V(M)$ on M. Thus,

$$\begin{aligned}&X_{\alpha f_a+\beta f_b}=\alpha X_{f_a}+\beta X_{f_b},\\&X_{[f_a,f_b]}=X_{f_a}X_{f_b}-X_{f_b}X_{f_a},\quad f_a,\ f_b\in F(M).\end{aligned} \tag{2.5}$$

Constants of the motion

A function $H\in F(M)$ is called a Hamiltonian function if it has no critical points (i.e., $dH=0$). The triplet (M,Ω,H) is called a Hamiltonian dynamical system. Thus,

$$H\to X_H=\frac{\partial}{\partial t}+\sum_i\left(\frac{\partial H}{\partial p_i}\frac{\partial}{\partial q_i}-\frac{\partial H}{\partial q_i}\frac{\partial}{\partial p_i}\right) \tag{2.6}$$

and

$$dH=\sum_i\left(\frac{\partial H}{\partial p_i}\,dp_i+\frac{\partial H}{\partial q_i}\,dq_i\right)+\frac{\partial H}{\partial t}\,dt. \tag{2.7}$$

If $H=H(p_i,q_i)$, then $\partial H/\partial t=0$. Now,

$$X_H(f_a)=\frac{\partial f_a}{\partial t}+[H,f_a],\quad [\text{using (2.4)}]. \tag{2.8}$$

Using the classical equation of motion, we have

$$\frac{df_a}{dt}=\frac{\partial f_a}{\partial t}+[H,f_a]=0, \tag{2.9}$$

i.e., f_a, $a=1,2,3,\ldots,n$, are constants of the motion and at equal times generate the Lie algebra

$$\mathbb{G}=\left\{f_a,\ 1\le a\le n\ \middle|\ \frac{df_a}{dt}=\frac{\partial f_a}{\partial t}+[H,f_a]=0\right\}.$$

We call $\mathbb{G}$ the *spectrum generating algebra* for the given Hamiltonian H. Further, if $[H,f_a]=0$, then $\partial f_a/\partial t=0$; i.e., f_a's are *independent of time*. Let

$$\mathbb{K}=\left\{f_a,\ 1\le a\le m<n\ \middle|\ \frac{\partial f_a}{\partial t}=0,\ [H,f_a]=0\right\}$$

be the subalgebra of $\mathbb{G}$, which we call the *symmetry algebra* for H. We note here that the elements of $\mathbb{G}$ so constructed will have some elements independent of time and the rest have explicit time dependence.

B. Construction of the Lie algebra $\mathbb{K}$

We will briefly summarise the construction of the symmetry algebra $\mathbb{K}\subset\mathbb{G}$. For a detailed discussion, the reader is referred to our earlier work.[2]

The Hamiltonian H for the Kepler problem is given by

$$H=(p^2/2)+V(q),\quad V(q)=V(q^2)=-1/q, \tag{2.10}$$

where we have used the reduced mass $\mu=1$ and the coupling parameter $\lambda=1$ for sake of convenience. The constants of the motion which have vanishing Poisson bracket with H are given by[2]

$$\mathbf{L}=\mathbf{q}\times\mathbf{p}, \tag{2.11}$$

$$\begin{aligned}\mathbf{f}&=a_1(q^2,H,l^2)\mathbf{q}+a_2(q^2,H,l^2)\mathbf{p}.\\&=(\alpha_0'/l^2)\,\mathbf{L}\times\mathbf{A}-\alpha_2'\,\mathbf{A},\end{aligned} \tag{2.12}$$

where

$$\mathbf{A}=\mathbf{p}\times\mathbf{L}-\mathbf{q}/q\quad(\text{the conventional Lenz-vector}), \tag{2.13}$$

$$\mathbf{L}\times\mathbf{A}=L^2[1-q/l^2]\mathbf{p}+\left(\frac{u}{ql^2}\right)\mathbf{q},\quad u=\mathbf{p}\cdot\mathbf{q}, \tag{2.14}$$

$$a_1=\alpha_0'\left(\frac{u}{l^2q}\right)-\alpha_2'(2H+1/q). \tag{2.15}$$

$$a_2=\alpha_0'(1-q/l^2)+\alpha_2'u. \tag{2.16}$$

α_0', α_2' are two arbitrary constants depending on H and l^2. Further, the coefficients a_1 and a_2 satisfy the following equations:

$$\begin{aligned}&\frac{u}{q}\frac{\partial a_2}{\partial q}+a_1=0,\\&\frac{a_2}{q^3}-\frac{u}{q}\frac{\partial a_1}{\partial q}=0\quad[\because\ [H,f_i]=0].\end{aligned} \tag{2.17}$$

The vectors $\mathbf{A}$ and $\mathbf{L}\times\mathbf{A}$ lie on the plane of the orbit.

So f is an arbitrary vector in the plane of the orbit which points in an arbitrary direction. For negative energy motion ($E < 0$),

$$\mathbf{f}\cdot\mathbf{f} = a_1^2 q^2 + 2u\, a_1 a_2 + a_2^2 p^2 = -\sigma l^2 + \sigma_0(H),$$

where

$$\sigma_0(H) > l^2 > -1/2H \quad \text{and} \quad \sigma = +1. \tag{2.18}$$

For $E > 0$ and $E = 0$, σ takes values -1 and 0, respectively. Thus, the symmetry algebra $\mathbb{K}$ is spanned by $(\mathbf{L}, \mathbf{f})$ whose Poisson bracket relations satisfy the Lie algebra isomorphic to that of $O(4)$.

C. Construction of the spectrum generating algebra $\mathbb{G}$

In order to get the multiplicity structure of the energy levels, we consider now a bigger algebra $\mathbb{G}$ which contains $\mathbb{K} = \{f_a = \mathbf{L}, \mathbf{f} \mid [H, f_a] = 0,\ \partial f_a/\partial t = 0\}$ as a subalgebra. As discussed earlier, the additional generators of $\mathbb{G}$ will have time dependence through Eq. (2.9). Let us consider now that the elements of $\mathbb{G} = \{f_a = \mathbf{L}, \mathbf{f}, \mathbf{M}, \Gamma, T, \Gamma_4, \Gamma_0 \mid \partial f_a/\partial t + [H, f_a] = 0\}$ satisfy the following equal-time Poisson bracket relations of $O(4, 2)$:

$$[L_{ab}, L_{cd}]_t = g_{ac} L_{bd} + g_{bd} L_{ac} - g_{ad} L_{bc} - g_{bc} L_{ad} \tag{2.19}$$

where the subscript "t" denotes the Poisson bracket relation at equal time t, $g_{ii} = g_{44} = -g_{55} = -g_{66} = 1$, $i = 1, 2, 3$; and we identify $L_i = \frac{1}{2}\epsilon_{ijk} L_{jk}$, $f_i = L_{i4}$, $M_i = L_{i5}$, $\Gamma_i = L_{i6}$, $T = L_{45}$, $\Gamma_4 = L_{46}$, and $\Gamma_0 = L_{56}$.

Construction of B(Γ and M), S, (Γ₄ and T), and Γ₀

Leave

$$\mathbf{B} = A_1 \mathbf{q} + A_2 \mathbf{p} \tag{2.20}$$

to be determined.

Now,

$$[\mathbf{f}, \mathbf{B}]_t = a_1 A_2 - a_2 A_1 + \text{terms which involve } \mathbf{p}\times\mathbf{p},\ \mathbf{q}\times\mathbf{p},\ \mathbf{p}\times\mathbf{q} \text{ and } \mathbf{q}\times\mathbf{q} \tag{2.21}$$

[using (2.12) and (2.20)].

From (2.19), we have

$$[\mathbf{f}, \mathbf{B}]_t = S.$$

Thus, identifying $S = a_1 A_2 - a_2 A_1$, we find that the coefficients which occur with $\mathbf{p}\times\mathbf{p}$, $\mathbf{q}\times\mathbf{p}$, etc. in (2.21) identically vanish. Again, using (2.12), (2.19), and (2.20), we have

$$[S, \mathbf{f}]_t = \mathbf{B} = A_1 \mathbf{q} + A_2 \mathbf{p}. \tag{2.22}$$

Simplifying the Poisson bracket relation of S and f at equal time and matching the coefficients of q and p on either side of (2.22), we have

$$A_1 = \left(\frac{u}{q}\frac{\partial a_1}{\partial H} + \frac{a_2}{q}\right)\frac{\partial S}{\partial q} + \frac{1}{l}(a_1 u + a_2 p^2)\frac{\partial S}{\partial l}$$

and

$$A_2 = \frac{u}{q}\frac{\partial S}{\partial q}\frac{\partial a_2}{\partial H} - \frac{1}{l}(a_1 q^2 + a_2 u)\frac{\partial S}{\partial l}. \tag{2.23}$$

Thus,

$$S = a_1 A_2 - a_2 A_1 = \left[\frac{u}{q}\left(a_1\frac{\partial a_2}{\partial H} - a_2\frac{\partial a_1}{\partial H}\right) - \frac{a_2^2}{q}\right]\frac{\partial S}{\partial q} + (l + 1/2Hl)\frac{\partial S}{\partial l}. \tag{2.24}$$

[using (2.23)].

Now, substituting for a_1 and a_2 from (2.15) and changing the variables from (l, q) to (u, q), we could in principle solve for S in (2.24). However, since a_1 and a_2 involve arbitrary functions α_0' and α_2', the general structure of S does not appear transparent. So, henceforth, we will restrict our analysis to the following cases:

1. $\alpha_0' = 0$, $\alpha_2' = (1/\sqrt{-2H})$, $\mathbf{f} = -\alpha_2' \mathbf{A}$; (2.25a)

2. $\alpha_2' = 0$, $(\alpha_0'/l) = (1/\sqrt{-2H})$, $\mathbf{f} = \dfrac{1}{l\sqrt{-2H}}\,\mathbf{L}\times\mathbf{A}$. (2.25b)

Case 1: The structure of S is obtained as follows (see Appendix A for the detailed derivation of S):

$$S = \frac{X(H)}{\sqrt{-2H}}\left(-\frac{1+2Hq}{\sqrt{-2H}}\cos\beta - u\sin\beta\right) + \frac{Y(H)}{\sqrt{-2H}}\left(-\frac{1+2Hq}{\sqrt{-2H}}\sin\beta + u\cos\beta\right),$$

$$\beta = (-2H)^{1/2}(u - 2Ht). \tag{2.26}$$

We note that S is a linear combination of two rotational scalars Γ_4 and T (say) involving the arbitrary coefficients $X(H)$ and $Y(H)$.

Let

$$\Gamma_4 = -\frac{1+2Hq}{\sqrt{-2H}}\cos\beta - u\sin\beta$$

and

$$T = -\frac{1+2Hq}{\sqrt{-2H}}\sin\beta + u\cos\beta. \tag{2.27}$$

In (2.27) we have chosen the arbitrary coefficients $X(H)$ and $Y(H)$ as

$$[X(H)/\sqrt{-2H}] = 1, \quad [Y(H)/\sqrt{-2H}] = 1. \tag{2.28}$$

To compute the vectors B, we substitute (2.27) in (2.23) and using (2.25a), we finally obtain after a little algebraic manipulation

$$\mathbf{M} = q\mathbf{p}\cos\beta - (1/\sqrt{-2H})(u\mathbf{p} - \mathbf{q}/q)\sin\beta$$

and

$$\mathbf{\Gamma} = q\mathbf{p}\sin\beta + (1/\sqrt{-2H})(u\mathbf{p} - \mathbf{q}/q)\cos\beta. \tag{2.29}$$

Now, using the equal-time Poisson bracket relation between M and Γ, we find

$$\Gamma_0 = (-2H)^{-1/2}. \tag{2.30}$$

We note here that Γ_0 is independent of time as it should be. It satisfies all the properties mentioned in the in-

troductory chapter to obtain the structure of the "spectrum generating Lie algebra" for the negative energy motion.

Case 2: We obtain the following expressions for M, Γ, Γ_4, T, and Γ_0:

$$\mathbf{M}=\frac{1}{l}\left[(2+2Hq)\mathbf{q}-uq\mathbf{p}\right]\cos\beta+\frac{1}{l\sqrt{-2H}}\left(-\frac{u}{q}(1+2Hq)\mathbf{q}+(u^2-q)\mathbf{p}\right)\sin\beta$$

$$\mathbf{\Gamma}=\frac{1}{l}\left[(2+2Hq)\mathbf{q}-uq\mathbf{p}\right]\sin\beta-\frac{1}{l\sqrt{-2H}}\left(-\frac{u}{q}(1+2Hq)\mathbf{q}+(u^2-q)\mathbf{p}\right)\cos\beta. \tag{2.31}$$

The structure of Γ_4, T, and Γ_0 are the same as in the Case 1.

III. CONNECTION WITH $O(4,2)$ REALIZATIONS OF BARUT AND BORNZIN

The canonical realisations of the Lie algebra of $O(4,2)$ as computed by Barut and Bornzin[6] are as follows:

$$\begin{aligned}
&\mathbf{L}'=\mathbf{q}\times\mathbf{p},\\
&\mathbf{f}'=\tfrac{1}{2}p^2\mathbf{q}-u\mathbf{p}-\tfrac{1}{2}\mathbf{q},\\
&\mathbf{M}'=\tfrac{1}{2}p^2\mathbf{q}-u\mathbf{p}+\tfrac{1}{2}\mathbf{q},\\
&\mathbf{\Gamma}'=q\mathbf{p}.\\
&\Gamma_0'=\tfrac{1}{2}(qp^2+q),\\
&\Gamma_4'=\tfrac{1}{2}(qp^2-q),\\
&T'=u.
\end{aligned} \tag{3.1}$$

To establish the canonical map between the realisations of Barut and Bornzin and those of our present analysis, we first define the canonical transformations of the elements in (3.1)[9]:

$$L_{ab}^{(F)}=e^{F}L_{ab}'=L_{ab}'+[F,L_{ab}']+\frac{1}{2!}[F,[F,L_{ab}']]+\cdots. \tag{3.2}$$

Note that under the canonical transformation (3.2), the Poisson bracket relations satisfied by (3.1) remain unaltered. Let us assume $F=a\,\mathbf{p}\cdot\mathbf{q}=au$ and "a" is an arbitrary function of energy E which has to be fixed later. Now, using $e^{au}u=u$, $e^{au}\mathbf{q}=e^{a}\mathbf{q}$ and $e^{au}\mathbf{p}=e^{-a}\mathbf{p}$, we obtain

$$\begin{aligned}
&\mathbf{L}^{(F)}=\mathbf{L}'=\mathbf{L},\\
&\mathbf{f}^{(F)}=\tfrac{1}{2}e^{-a}\mathbf{q}\,p^2-e^{-a}u\mathbf{p}-\tfrac{1}{2}e^{a}\mathbf{q},\\
&\mathbf{M}^{(F)}=\tfrac{1}{2}e^{-a}\mathbf{q}\,p^2-e^{-a}u\mathbf{p}+\tfrac{1}{2}e^{a}\mathbf{q},\\
&\mathbf{\Gamma}^{(F)}=e^{-a}q\mathbf{p},\\
&\Gamma_0^{(F)}=\tfrac{1}{2}(e^{-a}qp^2+qe^{a}),\\
&\Gamma_4^{(F)}=\tfrac{1}{2}(e^{-a}qp^2-qe^{a}),\\
&T^{(F)}=u.
\end{aligned} \tag{3.3}$$

Let $\exp(a)=(-2E)^{1/2}$.[10] Using $p^2=2H+2/q$, we simplify (3.3) to

$$\begin{aligned}
&\mathbf{M}^{(F)}=(-2E)^{-1/2}\left[-u\mathbf{p}+\mathbf{q}/q\right]+(-2E)^{-1/2}(H-E)\mathbf{q},\\
&\mathbf{\Gamma}^{(F)}=(-2E)^{-1/2}q\mathbf{p},\\
&\Gamma_0^{(F)}=(-2E)^{-1/2}+(-2E)^{-1/2}(H-E)q,\\
&\Gamma_4^{(F)}=(-2E)^{-1/2}+\left[1+q(H+E)\right],\\
&T^{(F)}=u.
\end{aligned} \tag{3.4}$$

Thus, for $H-E=0$ and taking linear combination of the elements $(\mathbf{M}^{(F)},\ \mathbf{\Gamma}^{(F)})$, $(\Gamma_4^{(F)},\ T^{(F)})$, we finally rederive the following structure for L_{ab}:

$$\begin{aligned}
&\mathbf{L}=\mathbf{q}\times\mathbf{p},\\
&\mathbf{f}=\frac{1}{\sqrt{-2H}}\left(\frac{1+2Hq}{q}\,\mathbf{q}-u\mathbf{p}\right),\\
&\mathbf{M}=\sqrt{-2H}\,\cos\beta\,\mathbf{\Gamma}^{(F)}-\sin\beta\,\mathbf{M}^{(F)},\\
&\mathbf{\Gamma}=\sqrt{-2H}\,\sin\beta\,\mathbf{\Gamma}^{(F)}+\cos\beta\,\mathbf{M}^{(F)},\\
&\Gamma_0=(-2H)^{-1/2},\\
&\Gamma_4=-\cos\beta\,\Gamma_4^{(F)}-\sin\beta\,T^{(F)},\\
&T=-\sin\beta\,\Gamma_4^{(F)}+\cos\beta\,T^{(F)}.
\end{aligned} \tag{3.5}$$

IV. CONCLUSION

The mathematical structure of the "spectrum generating algebras" was studied by Dothan,[3] Hermann,[11] and Barut and Bornzin.[6] Hermann addresses himself to the time independent case and squarely confines discussion to the spectrum generating algebra of $SO(5,R)$ and only outlines the Cartan's construction of obtaining the Lie algebra of the noncompact $SO(4,1)$. As stressed in Sec. II, with the appropriate choice of $X(H)$ and $Y(H)$, we reproduce Dothan's calculation (see our discussion for Case 1). Although, the spectrum generating algebra $O(4,1)$ precisely describes the multiplicity structure of energy levels, Dothan later adopts the Hamiltonian as a generator and extends the notion of spectrum generating algebra to $O(4,2)$. To us, the choice of $O(4,2)$ as the "spectrum generating algebra" has more intrinsic values. As stressed in the introductory chapter, we demand that G must contain the symmetry algebra $\mathbb{K}$ and $\mathbb{G}_H$ as subalgebras in such a way that $\mathbb{G}_H$ commutes with $\mathbb{K}$. By identifying L_{56} (the generator of $\mathbb{G}_H$) as the Hamiltonian flow, we have laid down the basis for the Hamiltonian quantization.

As noted earlier, Barut and Bornzin[6] construct the Lie algebra of $O(4,2)$ taking $O(2,1)$ (T, Γ_4, Γ_0) as the prototype building block. The elements of $O(4,2)$ so constructed do not depend upon time. Since $O(4,2)$ contains $O(4,1)$ as a subalgebra (which takes care of explaining the multiplicity structure of energy levels) and current elements (L_{i6}, L_{56}) which account for the electromagnetic transitions (also, the strong and weak decay transitions in elementary particles),[12–14] it is referred to as the dynamical group by the Boulder school while $O(2,1)$ which solves the spectrum of the Hamiltonian is known as the spectrum generating algebra.[6] Although, such a nomenclature looks more reasonable, we, however, take the attitude that, after all, "the dynamics of yesterday is known as kinematics of today" and have adopted the philosophy of Hermann[11] and Dothan.[3]

Now, coming to the case of noncompact orbits, we

essentially deal with the energy in the positive interval. The maximal symmetry (or, simply the symmetry) algebra $\mathbb{K}$ is $O(3,1)$, the Hamiltonian flow has to be identified with L_{46} and $\mathbb{G}$ which contains $\mathbb{K}$ and L_{46} as the elements of the subalgebras is still $O(4,2)$.

We want to recall here that the differential operators and the unitary irreducible representations of $O(4,2)$ are realised in the Hilbert space L^2 $(\Omega_{2n+1},\ n=3)$ equipped with the inner product[15]

$$(\phi_{nlm}^{(p,q,t)},\ \phi_{n'l'm'}^{(p,q,t)}) = \delta_{nn'}\,\delta_{ll'}\,\delta_{mm'},$$

where $\{\phi_{nlm}\}$ are the eigenvalues of the quadratic Casimir operator C_1 of $O(4)_{1,2,3,4}$ or Γ_0 (L_{56}), $\mathbf{L}^2$ and L_z.

Additional remark: The spectrum generating algebras are also useful in studying the problems in celestial mechanics where the Hamiltonian has a small perturbation; e.g., one could analyse the motion of the Earth and the Sun described by the unperturbed Hamiltonian, the solvable two-body problem and the perturbation could arise due to say Venus. By chosing the phase space Ω of this dynamical system as

$$\Omega = \sum_i (dp_i \wedge dq_i - dH_0 \wedge dt) - dH' \wedge dt$$
$$= \sum_i dp'_i \wedge dq'_i - dH' \wedge dt,$$

where H' is a function of p'_i and q'_i, $H = H_0 + H'$, one could obtain a detailed Lie algebraic structure.

ACKNOWLEDGMENTS

One of the authors (K.C.T.) is extremely thankful to Professor A.O. Barut for many instructive communications without which the work could not have been brought to the present form.

APPENDIX: COMPUTATION OF S (Γ_4 AND T)

We consider the case when $\mathbf{f} = [-1/(-2H)^{1/2}]\mathbf{A}$. Let $\alpha'_0 = 0$, $\alpha'_2 = 1/(-2H)^{1/2}$. Then

$$a_1 = -1/(-2H)^{1/2}(2H + 1/q), \quad a_2 = u/(-2H)^{1/2}$$

and

$$\mathbf{f}\cdot\mathbf{f} = -1/2H. \tag{A1}$$

Substituting (A1) in (2.24), we have

$$S = a_1A_2 - a_2A_1 = \frac{1}{q}\left(\frac{l^2-q}{2H}\right)\frac{\partial S}{\partial q} + \frac{1}{l}\left(l^2 + \frac{1}{2H}\right)\frac{\partial S}{\partial l}. \tag{A2}$$

Changing the variables (l, q) to (u, q), we rewrite (A2) as

$$S = \frac{1}{q}\left(\frac{\partial S}{\partial q} + \frac{1+2Hq}{u}\frac{\partial S}{\partial u}\right)\left(\frac{l^2-q}{2H}\right) - \frac{1}{u}\left(l^2 + \frac{1}{2H}\right)\frac{\partial S}{\partial u}. \tag{A3}$$

Let $S = \sum_i S_i q^i$, where $S_0 \neq 0$, S_i's are functions of H, and u only a solution of (A3). Substituting for S in (A3), we have the following recursion relations:

$$2H\,S_0 = S_1 - u\frac{\partial S_1}{\partial u},$$
$$u^2 S_1 + u\frac{\partial S_0}{\partial u} = 0, \tag{A4}$$

and S_2, S_3, ... identically vanish. Eliminating S_0 in (A4), we have

$$\frac{\partial^2 S_1}{\partial u^2} + (-2H/u)S_1 = 0 \tag{A5}$$

and

$$S = S_0 + qS_1 = \left(\frac{1+2Hq}{2H}\right)S_1 - (u/2H)\frac{\partial S_1}{\partial u}. \tag{A6}$$

Let $S_1 = C_0(H,t)\cos(u\sqrt{-2H}) + C_1(H,t)\sin(u\sqrt{-2H})$ be a solution of (A5). Then,

$$S = \left(\frac{1+2Hq}{2H}\right)[C_0\cos(u\sqrt{-2H}) + C_1\sin(u\sqrt{-2H})]$$
$$+ \frac{u}{\sqrt{-2H}}[C_0\sin(u\sqrt{-2H}) - C_1\cos(u\sqrt{-2H})] \tag{A7}$$

[substituting for S_1 in (A6)].

To estimate C_0 and C_1, we make use of the equation of motion

$$\frac{\partial S}{\partial t} + [H, S] = 0$$

or,

$$\frac{u}{q}\frac{\partial S}{\partial q} + \frac{\partial S}{\partial t} = 0, \tag{A8}$$

Using (A8) in (A7), we have

$$\frac{\partial C_0}{\partial t} + (-8H^3)^{1/2}C_1 = 0,$$
$$\frac{\partial C_1}{\partial t} - (-8H^3)^{1/2}C_0 = 0. \tag{A9}$$

Thus solving for C_0, C_1 in (A9), we find

$$C_0 = X(H)\cos(-8H^3)^{1/2}t + Y(H)\sin(-8H^3)^{1/2}t, \tag{A10}$$
$$C_1 = -\frac{1}{(-8H^2)^{1/2}}\frac{\partial C_0}{\partial t} = X(H)\sin(-8H^3)^{1/2}t - Y(H)$$
$$\times\cos(-8H^3)^{1/2}t.$$

Substituting (A10) in (A7), we finally obtain

$$S = \frac{X(H)}{(-2H)^{1/2}}\left(-\frac{1+2Hq}{\sqrt{-2H}}\cos[\sqrt{-2H}\,(u-2Ht)]\right.$$
$$\left. - u\sin[\sqrt{-2H}\,(u-2Ht)]\right)$$
$$+ \frac{Y(H)}{(-2H)^{1/2}}\left(-\frac{1+2Hq}{\sqrt{-2H}}\sin[\sqrt{-2H}\,(u-2Ht)]\right.$$
$$\left. + u\cos[\sqrt{-2H}\,(u-2Ht)]\right). \tag{2.26}$$

QED

[1]P.S. Laplace, *A Treatise of Celestial Mechanics* (Dublin, 1827).
[2]K.C. Tripathy and J.D. Anand, Nuovo Cimento 17B, 71 (1973).
[3]Y. Dothan, Phys. Rev. D 2, 2944 (1970).
[4]The energy surfaces M_E are submanifolds of M. The topological structure of M_E is

$$M_E \approx O(4)/O(2) = V_{1,2},$$

where $V_{1,2}$ is known in the literature as the Stiefel manifold.

[5]The two Casimir operators of $O(4)_{1,2,3,4}$ have the following values:

$$C_1 = \tfrac{1}{2} L_{\alpha\beta} L^{\alpha\beta} = k_0^2 + (|k_0| + n)^2 = -1/2H$$

$$C_2 = \tfrac{1}{8} \epsilon_{\alpha\beta\gamma\delta} L^{\alpha\beta} L^{\gamma\delta} = k_0(|k_0| + n) = 0.$$

This implies that $k_0 = 0$ and $C_1 = n^2 = -1/2H$.

[6]A. O. Barut and G. L. Bornzin, J. Math. Phys. **12**, 841 (1971) and Trieste preprint IC/72/142 (unpublished).

[7]R. Hermann, *Differential Geometry and the Calculus of Variations* (Academic, New York, 1968).

[8]Ω is used here as a metric tensor to construct a covariant vector field on M from a contravariant one.

[9]P. Chand, C. L. Meheta, N. Mukunda, and E. C. G. Sudarshan, J. Math. Phys. **8**, 2048 (1967).

[10]To fix $\exp(a) = (-2E)^{1/2}$, we consider the following operator equation:

$$\Gamma_0^{(F)} = e^{-a} q(H + 1/q) + \tfrac{1}{2} q\, e^{a} \quad \text{[from (3.3)]},$$

or

$$[e^{a}\Gamma_0^{(F)} - 1]\varphi = q(H + \tfrac{1}{2} e^{2a})\varphi.$$

Using $H\varphi = E\varphi$ and by demanding

$$q(H + \tfrac{1}{2} e^{2a})\varphi = 0,$$

we obtain

$$\exp(2a) = -2E, \quad \text{or} \quad \exp(a) = (-2E)^{1/2}.$$

Under this case, $\Gamma_0^{(F)} = \sqrt{-2H}$ which exactly coincides with our realisation for Γ_0.

[11]R. Hermann, J. Math. Phys. **13**, 833, 838 (1972).

[12]A. O. Barut and K. C. Tripathy, Phys. Rev. Lett. **19**, 918, 1081 (1967) and Phys. Rev. **175**, 2278 (1969). A. O. Barut, D. Corrigan, and H. Kleinert, Phys. Rev. **167**, 1527 (1968). A. O. Barut and A. Baiquni, Phys. Rev. D **9**, 1084 (1974).

[13]Milan Noga, Phys. Rev. D **3**, 3047 (1971).

[14]F. Mansouri, Phys. Rev. D **8**, 11, 59 (1973).

[15]In our present realizations of $O(4,2)$, the Casimir operators take the following values:

$$Q_2 = \tfrac{1}{2} L_{ab} L^{ab} = 0, \quad a, b = 1, 2, 3, 4, 5, 6;$$

$$Q_3 = \tfrac{1}{48} \epsilon^{abcdef} L_{ab} L_{cd} L_{ef} = 0;$$

$$Q_4 = \tfrac{1}{4}(L_{ab} L^{bc} L_{cd} L^{da} - Q_2^2 - 8Q_2) = 0.$$

The Casimir operators of $O(4)_{1,2,3,4}$ as state in Footnote 5 are given by

$$C_1 = -1/2H = (\Gamma_0)^2,$$

$$C_2 = 0.$$

A relativistic approach to quadrupole effects on planetary orbits

Daniel Yaakobi

Department of Environmental Sciences, Tel-Aviv University, Ramat Aviv, Israel
(Received 2 August 1974)

The geodesic curves for a static axially symmetric metric and their first integrals of motion are calculated. The conformal mapping found by Erez and Rosen leading from the axially symmetric metric to the Schwarzschild spherically symmetric metric is applied and if a quadrupole moment is taken into account initially, a correction to the motion of a planet moving in the field of the sun in the equatorial plane is found. The results are compatible with Schwarzschild solution and motion in a classical quadrupole field. The main result is a justification of the usual procedure of simply adding relativistic and classical perturbations on planetary orbits.

The stationary axially symmetric metric is according to Weyl and Levi—Civita in polar coordinates[1]

$$ds^2 = \exp(2\psi)\,dt^2 - \exp(2\gamma - 2\psi)(dr^2 + r^2\,d\theta^2) - \exp(-2\psi r^2 \sin^2\theta\, d\phi^2, \tag{1}$$

where r, θ, ϕ are the polar coordinates and t is time.

The functions $\psi(r, \theta)$ and $\gamma(r, \theta)$ satisfy the equations (a suffix denotes differentiation with respect to the coordinate)

$$\psi_{rr} + \frac{2}{r}\psi_r + \frac{1}{r^2}\psi_{\theta\theta} + \frac{1}{r^2}\cot\theta\psi_\theta = 0, \tag{2a}$$

$$\gamma_r = \left(\psi_r^2 - \frac{1}{r^2}\psi_\theta^2\right) r\sin^2\theta + 2\psi_r\psi_\theta \sin\theta\cos\theta, \tag{2b}$$

$$\gamma_\theta = -\left(\psi_r^2 - \frac{1}{r^2}\psi_\theta^2\right) r^2\sin\theta\cos\theta + 2\psi_r\psi_\theta\, r\sin^2\theta. \tag{2c}$$

The solutions are[1]

$$\psi(r, \theta) = \sum_{n=-\infty}^{\infty} A_n r^n P_n(\cos\theta) \tag{3a}$$

$$\gamma(r, \theta) = \sum_{m,n=-\infty}^{\infty}{}' A_m A_n \frac{mn}{m+n} r^{m+n}(P_m P_n - P_{m-1}P_{n-1}) + 2\sum_{m=1}^{\infty} A_m A_{-m}\int_0^\theta (P_m'^2 + m^2P_m^2)\sin\theta\,d\theta + C. \tag{3b}$$

$P_m(\cos\theta)$ are the Legendre polynomials and $\sum'$ is a summation on m and n excluding the case where $m = -n$, P_m' denotes $dP_m/d\theta$.

The geodesic equations are

$$\frac{d^2r}{ds^2} - (\psi_r - \gamma_r)\left(\frac{dr}{ds}\right)^2 - 2(\psi_\theta - \gamma_\theta)\frac{dr}{ds}\frac{d\theta}{ds} - r(1 - r\psi_r + r\gamma_r)\left(\frac{d\theta}{ds}\right)^2 - \exp(-2\gamma)\, r\sin^2\theta\,(-2\psi_r)\left(\frac{d\psi}{ds}\right)^2 + \exp(4\psi - 2\gamma)\,\psi_r\left(\frac{dt}{ds}\right)^2 = 0, \tag{4a}$$

$$\frac{d^2\theta}{ds^2} + \frac{1}{r^2}(\psi - \gamma)\left(\frac{dr}{ds}\right)^2 + 2\left(\frac{1}{r} - \psi_r + \gamma_r\right)\frac{dr}{ds}\frac{d\theta}{ds} - (\psi_\theta - \gamma_\theta)\left(\frac{d\theta}{ds}\right)^2 - \exp(-2\gamma)\sin^2\theta(\cot\theta - \psi_\theta)\left(\frac{d\psi}{ds}\right)^2 + \exp(4\psi - 2\gamma)\left(\frac{dt}{ds}\right)^2 = 0, \tag{4b}$$

$$\frac{d^2\phi}{ds^2} + 2\left(\frac{1}{r} - \psi_r\right)\frac{dr}{ds}\frac{d\phi}{ds} + 2(\cot\theta - \psi_\theta)\frac{d\theta}{ds}\frac{d\psi}{ds} = 0, \tag{4c}$$

$$\frac{d^2t}{ds^2} + 2\psi_r\frac{dr}{ds}\frac{dt}{ds} + 2\psi_\theta\frac{d\theta}{ds}\frac{dt}{ds} = 0. \tag{4d}$$

The last two equations can be integrated:

$$r^2\sin^2\theta\frac{d\phi}{ds} = he^{2\psi}, \tag{5a}$$

$$\frac{dt}{ds} = ke^{-2\psi}. \tag{5b}$$

Equation (5a) is the angular momentum integral while (5b) is the energy integral. Substituting (5a, b) in (1), we get

$$\left(\frac{dr}{ds}\right)^2 + r^2\left(\frac{d\theta}{ds}\right)^2 = \left(k^2\exp(-2\psi) - 1 - \frac{h^2}{r^2\sin^2\theta}\times\exp(2\psi)\right)\exp(2\psi - 2\gamma). \tag{6}$$

Equation (6) can be substituted for (4a).

A possible solution for θ is

$$\theta = \text{const} = \pi/2. \tag{7}$$

This is the case of equatorial motion of a particle in the Sun's field. This is possible when all the Sun's mass is distributed symmetrically with respect to the equatorial plane and it excludes from Eq. (3) for ψ all the odd Legendre polynomials and we are left only with even n. Applying (7) to (5a) and (6), we obtain after division

$$\left(\frac{dr}{ds}\Big/\frac{d\phi}{ds}\right)^2 = \left(\frac{dr}{d\phi}\right)^2 = \exp(-2\gamma)\left(\left[k^2\exp(-2\psi) - 1\right]\frac{r^4\exp(-2\psi)}{h^2} - r^2\right). \tag{8}$$

Choosing a new variable $u \equiv 1/r$, we find

$$\left(\frac{du}{d\phi}\right)^2 \equiv u'^2 = \exp(-2\gamma)\left(\left[k^2\exp(-2\psi) - 1\right]\frac{\exp(-2\psi)}{h^2} - u^2\right). \tag{9}$$

A solution for $\psi(r, \theta)$ which includes only a point mass m and a quadrupole moment q is

$$\psi(r, \theta) = \frac{-m}{r} + \frac{1}{2}\frac{q}{r^3}(3\cos^2\theta - 1) \tag{10}$$

and the corresponding value of γ is

$$\gamma(r, \theta) = \frac{-\sin^2\theta}{2}\left(\frac{m^2}{r^2} - \frac{3}{2}\frac{mq}{r^4}(5\cos^2\theta - 1)\right.$$

Copyright © 1975 American Institute of Physics

$$+\frac{3}{2}\frac{q^2}{r^6}(25\cos^4\theta-14\cos^2\theta-1)\Big)\ . \tag{11}$$

Taking $\theta=\pi/2$ these expressions reduce to

$$\psi(u)=-mu-\tfrac{1}{2}qu^3, \tag{12a}$$

$$\gamma(u)=-\tfrac{1}{2}(m^2u^2+\tfrac{3}{2}mqu^4+\tfrac{3}{2}q^2u^6). \tag{12b}$$

Differentiating (9) with respect to ψ and denoting $u''\equiv d^2u/d\phi^2$, $\psi'\equiv d\psi/du$, $\gamma'\equiv d\gamma/du$, we obtain a second order differential equation for u:

$$u''=\exp(-2\gamma)\left[-\gamma'\left(\frac{1}{h^2}[k^2\exp(-4\psi)-\exp(-2\psi)]-u^2\right)+\frac{1}{h^2}\psi'[-2k^2\exp(-4\psi)+\exp(-2\psi]-u\right]. \tag{13}$$

Expanding the exponentials up to the second order of the mass m and quadrupole moment q, we get

$$\exp(-2\gamma)=1-2\gamma, \tag{14a}$$

$$\exp(-2\psi)=1-2\psi+2\psi^2, \tag{14b}$$

$$\exp(-4\psi)=1-4\psi+8\psi^2. \tag{14c}$$

The derivatives of (12) are

$$\psi'=-m-\tfrac{3}{2}qu^2, \tag{15a}$$

$$\gamma'=-m^2u-3mqu^3-\tfrac{9}{2}q^2u^5. \tag{15b}$$

Inserting these values into (13), we obtain

$$\begin{aligned}u''+u=&\frac{1}{h^2}(2k^2-1)(m+\tfrac{3}{2}qu^2)+\frac{3m^2}{h^2}u(3k^2-1)\\&+\frac{mq}{h^2}u^3(19k^2-7)+\frac{3}{2}\frac{q^2}{h^2}u^5(7k^2-4)\\&-u(2m^2u^2+\tfrac{9}{2}mqu^4+6q^2u^6)\\&+\text{terms of higher order in }(m,q).\end{aligned} \tag{16}$$

In the case of slow motion of a test particle, k is very close to unity and so Eq. (16) becomes

$$u''+u=\frac{m}{h^2}+\frac{3}{2}\frac{q}{h^2}u^2+\frac{6m^2}{h^2}u+\frac{12mqu^3}{h^2}+\frac{9}{2}\frac{q^2}{h^2}u^5+\text{terms of higher order in }(mu,qu^3). \tag{17}$$

The conformal transform performed by Erez and Rosen[2] which takes a rod of length $2m$ and transforms it to the Schwarzschild spherical solution of a mass point m is

$$r^2=\bar r^2-2m\bar r+m^2\cos^2\bar\theta, \tag{18a}$$

$$\cos\theta=\frac{\cos\bar\theta(\bar r-m)}{(\bar r^2-2m\bar r+m^2\cos\bar\theta)^{1/2}} \tag{18b}$$

$$\phi=\bar\phi. \tag{18c}$$

In the equatorial plane $\theta=\bar\theta=\pi/2$ this gives

$$u=\bar u/(1-2m\bar u)^{1/2} \tag{19}$$

where $u=1/r$, $\bar u=1/\bar r$.

The derivatives of u with respect to $\phi=\bar\phi$ are

$$u'=\bar u'[1-m\bar u)/(1-2m\bar u)^{3/2}], \tag{20a}$$

$$u''=\bar u''\frac{(1-m\bar u)}{(1-2m\bar u)^{3/2}}+\frac{m\bar u'^2}{(1-2m\bar u)^{5/2}}(2-m\bar u). \tag{20b}$$

Expanding (19) and (20) up to linear terms in $m\bar u$, we get

$$u=\bar u(1+m\bar u), \tag{21a}$$

$$u'=\bar u'(1+2m\bar u), \tag{21b}$$

$$u''=\bar u''(1+2m\bar u)+2m\bar u'^2, \tag{21c}$$

$$u'^2=\bar u'^2(1+4m\bar u). \tag{21d}$$

We have found before, Eq. (9), that

$$u'^2=\exp(-2\gamma)\left([\exp(-2\psi)-1]\frac{\exp(-2\psi)}{h^2}-u^2\right) \tag{22}$$

and after using the expressions for ψ and γ, Eq. (12), we obtain

$$u'^2=\frac{1}{h^2}(2mu+qu^3)-u^2=\bar u'^2(1+4m\bar u). \tag{23}$$

The value for $\bar u'^2$ is

$$\begin{aligned}\bar u'^2&=u'^2(1-4m\bar u)\\&=(1-4m\bar u)\left(\frac{1}{h^2}[2m\bar u(1+m\bar u)+q\bar u^3(1+3m\bar u)]-\bar u^2(1+2m\bar u)\right)\\&=\frac{1}{h^2}[2m\bar u(1-3m\bar u)+q\bar u^3(1-m\bar u)]-\bar u^2(1-2m\bar u).\end{aligned} \tag{24}$$

From Eqs. (21) we get for u'' (after omitting terms higher than $m^2\bar u/h^2$, $mq\bar u^3/h^2$):

$$u''=\bar u''(1+2m\bar u)+\frac{4m}{h^2}\left(m\bar u+\frac{q}{2}\bar u^3\right)-2m\bar u^2. \tag{25}$$

From Eq. (17) we find

$$\begin{aligned}u''+u=&\bar u''(1+2m\bar u)+\frac{4m^2}{h^2}\bar u+\frac{2q}{h^2}m\bar u^3\\&-2m\bar u^2+\bar u(1+m\bar u)\\=&\frac{1}{h^2}m+\tfrac{3}{2}q\bar u^2(1+2m\bar u)+\frac{6m^2}{h^2}\bar u(1+m\bar u)\\&+\frac{12mq}{h^2}\bar u^3(1+3m\bar u)+\frac{9}{2}\frac{q^2}{h^2}\bar u^5(1+5m\bar u).\end{aligned} \tag{26}$$

Dividing Eq. (26) by $(1+2m\bar u)$ and expanding, the result is

$$\begin{aligned}\bar u''+&\frac{4m^2}{h^2}\bar u(1-2m\bar u)+\frac{2mq}{h^2}\bar u^3(1-2m\bar u)-m\bar u^2(1-2m\bar u)\\&+\bar u(1-2m\bar u)\\=&\frac{m}{h^2}(1-2m\bar u)+\tfrac{3}{2}q\bar u^2+\frac{6m^2}{h^2}\bar u(1-m\bar u)\\&+\frac{12mq}{h^2}\bar u^3(1+m\bar u)+\frac{9}{2}\frac{q^2}{h^2}\bar u^5(1+3m\bar u).\end{aligned} \tag{27}$$

The final result is (after omitting the bar over the u):

$$u''+u=\frac{m}{h^2}+3mu^2+\frac{3}{2}\frac{q}{h^2}u^2+10\frac{mq}{h^2}u^3+\frac{9}{2}\frac{q^2}{h^2}u^5. \tag{28}$$

This equation describes the motion of a test particle in the equatorial plane in the field of a gravitational mass m having a quadrupole distribution q. The first two terms at right give the Schwarzschild solution for the precession of the perihelion of a planet around the Sun, the third term gives the precession due to a classical quadrupole effect (see Ref. 3) while the last two terms are a general relativistic correction to the

Schwarzschild solution due to the quadrupole moment of the Sun.

Now applying Eq. (28) to the motion of a planet (e.g., Mercury) around the Sun we can drop the last two terms and we are left with

$$u''+u=\frac{m}{h^2}+3\left(m+\frac{1}{2}\frac{q}{h^2}\right)u^2=\frac{m}{h^2}+3m\left(1+\frac{1}{2}\frac{q}{mh^2}\right)u^2$$

$$=\frac{m}{h^2}+3m(1+C)u^2. \tag{29}$$

As Dicke and Goldberg[4] have pointed out, the Sun has an oblateness of

$$(r_{eq}-r_{pole})/r=(5\pm 0.7)\times 10^{-5} \tag{30}$$

which would contribute 8% of the general relativistic effect which means that C which is the ratio between the classical (quadrupole precession of a planet and the Schwarzschild precession, is 0.08. Let us calculate it according to Adler, Bazin, and Schiffer.[3]

Comparing the potential $\psi(r)$ given by (12a) with the potential of Ref. 3

$$f(r)=-\frac{m}{r}-\frac{B}{r^3}=-mu-\tfrac{1}{2}qu^3, \tag{31}$$

we find in natural units that

$$B=q/2. \tag{32}$$

The classical precession is (according to Ref. 3)

$$\delta\phi_c=3\pi\frac{mq}{h^2}, \tag{33a}$$

while the Schwarzschild precession is

$$\delta\phi_s=2\pi(3m^2/h^2). \tag{33b}$$

Their ratio is exactly C as defined in (29). The conclusion is that Mercury's precession around the Sun can be described by the superposition of the general relativistic effect and the Sun's quadrupole moment. This paper rigorously justifies the intuitively obvious procedure of adding relativistic and classical perturbations on planetary orbits around the Sun.

ACKNOWLEDGMENTS

My thanks are due to Professor G. E. Tauber for the many valuable discussions of this subject.

[1]N. Rosen and D. Yaakobi, Bull. Res. Counc. of Israel 10F, 49 (1961).
[2]G. Erez and N. Rosen, Bull. Res. Counc. of Israel 8F, 47 (1959).
[3]Adler, Bazin, and Schiffer, *Introduction to General Relativity* (McGraw-Hill, New York, 1965).
[4]R. H. Dicke and H. Mark Goldberg, Phys. Rev. Lett. 18, 313 (1967).

An explicit model for the renormalization of field equations

P. de Mottoni* and A. Tesei*

G.N.A.F.A., C.N.R., Roma, Italy
(Received 15 July 1974)

We provide a simple model where a precedure for renormalizing field equations yields nontrivial distribution solutions.

1. INTRODUCTION

The renormalization of field equations has been recently considered[1,2] in a frame which can be summarized as follows:

(a) Write down a sequence of approximated "renormalized" equations

$$\begin{cases} \dfrac{d}{dl} U_n(t) = LU_n(t) + b(n)f(U_n), \\ U_n(0) = U_{n0}, \end{cases} \tag{1}$$

where $n \in \mathbb{N}$, the unknown $U_n(t)$ being an m-tuple of operator-valued functions (of the space variables), L a $m \times m$ matrix of linear operators, f an m-tuple of nonlinear functions, U_{n0} an m-tuple of given operator-valued functions, and $b(n)$ a sequence of positive numbers. The operator L should give rise to the free dynamics, the U_n's should represent the approximated fields, and U_{n0} the corresponding initial data.

(b) Suppose we are able to establish existence and uniqueness for the solution of $(1)_n$, $n \in \mathbb{N}$. Suppose further that, for $n \to \infty$, U_{n0} converges to an m-tuple of operator-valued distributions (the $t=0$ fields), and $b(n)$ converges to zero.

(c) We may ask (i) whether the corresponding solutions converge for $n \to \infty$, or, at least, whether they belong to a compact set in the space of distributions; If the answer is affirmative, we may ask (ii) whether the so-obtained limiting points are trivial, that is, whether they satisfy an equation where no interaction is present. Of course, both answers should involve conditions on $b(n)$ and U_{n0}, to be interpreted as renormalization conditions. If we succeed in finding nontrivial limiting points, these are good candidates for describing fields which are at any time tempered distributions in the space variables, and satisfy in some weak sense a nonlinear equation of motion, which is perfectly defined in the approximated form (1).

The question (i) has been answered in some cases, such as the massive Thirring and Federbush models, in Ref. 2. The aim of the present note is to provide a very simple model in which, first, by the same general method developed in Ref. 2, we establish a condition for the existence of limiting points, and, secondly, we perform an explicit calculation of the solutions, yielding a unique nontrivial limiting point under conditions which are compatible with those required by the general method.

2. THE MODEL

Cur model derives from that proposed by Dell'Antonio,[3] and is of type (1) with $m=1$. The approximated operator-valued function u_n satisfies the equation

$$\begin{aligned} &\partial_t u_n = \partial_x u_n - n^{-b} u_n^k, \\ &u_n(0) = u_{n0}, \end{aligned} \tag{2}$$

where k is a positive integer. We first deal with (2) for fixed n. Let V be a Hilbert space [scalar product $(.,.)$], $H(V)$ the space of linear continuous self-adjoint operators on V (norm $\|\cdot\|$); let $X \equiv C_0(\mathbb{R}; H(V))$ be the Banach space of the continuous functions on $\mathbb{R}$, with values in $H(V)$, vanishing at $\pm\infty$, with the norm $u \to |u| \equiv \sup_{x \in \mathbb{R}} \|u(x)\|$; let $H^+(V)$ be the closed convex cone of the positive operators of $H(V)$, and $X^+ \equiv C_0(\mathbb{R}; H^+(V))$ the closed convex cone of the positive operator valued functions of X. It is easily seen[4-6] that for any $u_{n0} \in X^+$ there is a unique continuous function from $\mathbb{R}^+$ to X^+, $t \to u_n(t)$, satisfying

$$(u_n(t))(x) = u_{n0}(x+t) - n^{-b} \int_0^t (u_n^k(s))(x+t-s)\, ds. \tag{3}$$

In addition, the following estimate holds:

$$|u_n(t)| \leq |u_{n0}| \quad \forall\ t \in \mathbb{R}^+.$$

The next step is going to the limit $n \to \infty$. Let us first define $\mathcal{S}^1 \equiv \mathcal{S}^1(\mathbb{R}; H(V))$ as the space of linear continuous mappings from $\mathcal{S}(\mathbb{R})$ to $H(V)$: $T: f \to \langle T, f\rangle \in H(V)$ for $f \in \mathcal{S}(\mathbb{R})$. Any $g \in X$ can be identified with a $T_g \in \mathcal{S}'$ by putting $\langle T_g, f\rangle = \int g(x) f(x)\, dx$. We shall endow $\mathcal{S}'$ with the weak*—weak operator topology, whose seminorms are

$$T \to p_{f,A}(T) = \max_{v_i, v_j \in A} |(v_i, \langle T, f\rangle v_j)|$$

for $f \in \mathcal{S}(\mathbb{R})$, and A a finite subset of V. We shall also consider another (stronger) topology, called weak*-norm operator topology, with seminorms given by

$$T \to q_f(T) = \|\langle T, f\rangle\| \quad \text{for } f \in \mathcal{S}(\mathbb{R}).$$

We further define $\mathcal{S}'^+$, the space of positive operator-valued distributions, namely the subset of the T's of $\mathcal{S}'$ such that $\langle T, f\rangle \in H^+(V)$ whenever $f \in \mathcal{S}(\mathbb{R})$, $f > 0$. Eventually, for $g \in \mathcal{S}(\mathbb{R})$, we put $g_t(x) = g(x-t)$ (for any $t \in \mathbb{R}$). After these preliminaries, we may apply (3), thought of as an equation in $\mathcal{S}'^+$, to an arbitrary positive function f of $\mathcal{S}(\mathbb{R})$. We get

$$\langle u_n(t), f\rangle = \langle u_{n0}, f_t\rangle - n^{-b} \int_0^t \langle u_n^k(s), f_{t-s}\rangle\, ds,$$

whence

$$\|\langle u_n(t), f\rangle\| \leq \|\langle u_{n0}, f_t\rangle\|,\ \exp(n^{-b} |u_{n0}|^{k-1} t),$$

where use was made of Gronwall's lemma and of the inequality

$$\left\| \int h(x) k(x)\, dx \right\| \leq \sup_x \|h(x)\| \cdot \left\| \int k(x)\, dx \right\|$$

which holds for any two continuous integrable $H^+(V)$-valued functions h, k on $\mathbb{R}$, such that $h(x)k(x) = k(x)h(x)$,

 Copyright © 1975 American Institute of Physics

for any $x \in \mathbf{R}$.[5] Observing that $t \to f_t$ is a continuous (and hence locally equicontinuous[7,8]) group on $\mathcal{S}(\mathbf{R})$, and using uniform boundedness arguments,[9] we see that $\|\langle u_n(t), f\rangle\|$ is bounded in n for f in a bounded set of $\mathcal{S}(\mathbf{R})$ and t in a bounded interval when u_{n0} converges in the weak*-weak operator topology, provided

$$\{|u_{n0}|^{k-1} n^{-b}\} \tag{4}$$

is a bounded sequence. But, by a straightforward generalization of the Banach—Alaoglu theorem,[5,7] this implies that $\{u_n(t)\}$ belongs to a compact set in the weak*-weak operator topology of $\mathcal{S}'$. Hence we may conclude: *If the initial data u_{n0} converge in the weak*-weak operator topology, and the "renormalization condition" (4) is fulfilled, the (mild) solutions $u_n(t)$ of (2) do have weak*-weak operator limiting points in $\mathcal{S}'^{+}$, uniformly in t on the compact sets of $\mathbf{R}^+$.* Three remarks are in order:

(1) The same procedure applies and the same result holds when replacing ∂_x in (2) by any differential operator L, provided it generates a strongly continuous semigroup on X, preserving positivity.[5]

(2) The procedure we followed is essentially the same as in Ref. 2; it is quite general and does not take into account the explicit form of the solutions of (2).

(3) The renormalization condition (4) above amounts to a relationship between the damping factor n^{-b} and the growth of the C_0-norms of the initial data.

The above procedure gives no guarantee that the renormalization condition (4) is too strong, in the sense that it may force the limiting points to coincide with the free solution, namely, with the distribution $U(t,.)$ satisfying

$$\frac{\partial}{\partial t} U(t) = \frac{\partial}{\partial x} U(t), \quad U(0) = u_0$$

(where u_0 is the limit of the u_{n0}'s). We shall now show, in a simple example, that this is not the case. The explicit calculation of $u_n(t,x)$ can be performed, and yields

$$u_n(t,x) = u_{n0}(x+t) J(n^{-b}(k-1)^{-1} t, u_0^{k-1}(x+t))^{1/(k-1)}$$

where $J(z,A)$ denotes $(zA+I)^{-1}$. If we take $u_{n0} = \delta_{c,n} \otimes B$, with $B \in H^+(V)$, $\delta_{c,n} = n^c (2\pi)^{-1/2} \exp(-n^{2c} \cdot x^2/2)$, $c \in \mathbf{N}$, then u_{n0} converges to $\delta \otimes B$ in the weak*-norm operator topology of $\mathcal{S}'$. The C_0-norm of the initial data can be easily computed: $|u_{n0}| = \|B\| n^c$. With $f \in \mathcal{S}(\mathbf{R})$, $f \geq 0$, we get

$$\begin{aligned}
&\langle u_n(t,), f\rangle \\
&= n^c B \int \exp(-n^{2c} x^2/2) \cdot [(k-1)^{-1} n^{-b} t n^{(k-1)c} \\
&\quad \times \exp(-n^{2c}(k-1)x^2/2) B^{k-1} + 1]^{-1/(k-1)} f(x-t)\, dx \\
&= B \int \exp(-x^2/2) [(k-1)^{-1} n^{-b+(k-1)c} \\
&\quad \times \exp(-(k-1)x^2/2) B^{k-1} + 1]^{-1/(k-1)} f(x n^{-c} - t)\, dx.
\end{aligned}$$

A straightforward calculation then yields

(a) if $(k-1)c < b$, u_n converges to the free solution;

(b) if $(k-1)c = b$, u_n converges to a nontrivial limit;

(c) if $(k-1)c > b$, u_n converges to zero.

Here, convergence is meant in the weak*-norm operator topology (which implies convergence in the weak*-weak operator topology). The nontrivial limit of $\langle u_n(t), f\rangle$ is

$$B f(-t) \int dx \exp(-x^2/2) J(t(k-1)^{-1} \exp[-(k-1)x^2/2, B^{k-1})^{1/k-1}.$$

Cases (a) and (b) correspond to the condition (4); namely, (a) to $|u_{n0}|^{k-1} n^{-b} \to 0$ and (b) to $|u_{n0}|^{k-1} n^{-b} \to$ const. Case (b), therefore, gives the nontrivial solution compatible with the condition (4).

The general characterization of nontrivial limiting points—independent of the explicit knowledge of the solutions—is still an open problem, which deserves further investigation.

*Postal address: A. Tesei, Via S. Giovanni in Lanterano, 84 I 00184, Rome, Italy.

[1] A. Tesei, J. Math. Phys. **16**, 140 (1975).
[2] P. de Mottoni and A. Tesei, "On the Renormalization of Field Equations," preprint (1974).
[3] G. F. Dell'Antonio, "A Model Field Theory: the Thirring Model," in *Proceedings of the Schladming Winter School*, 1973.
[4] E. Salusti and A. Tesei, Nuovo Cimento A **2**, 122 (1971).
[5] P. de Mottoni and A. Tesei, "Soluzioni distribuzioni per equazioni di evoluzione non lineari," *Seminario di Analisi non lineare*, Istituto Matematico dell'Università di Roma (1974).
[6] M. Iannelli, Boll. U. M. I. **6**, 1015 (1970).
[7] K. Yosida, *Functional Analysis* (Springer, Berlin, 1968).
[8] T. Komura, J. Funct. Anal. **2**, 258 (1968).
[9] M. Reed and B. Simon, *Methods of Mathematical Physics I: Functional Analysis* (Academic, New York, 1972).

On the equivalence of the Ornstein–Zernike relation and Baxter's relations for a one-dimensional simple fluid

M. Chen

Department of Mathematics, Vanier College, Snowdon Campus, Montreal, Quebec, Canada

Department of Mathematics, Sir George Williams University, Montreal, Quebec, Canada

(Received 18 June 1974)

If the direct correlation function $c(x)$ for a one-dimensional simple fluid vanishes for $|x|>l$, where l is finite, it can be proved that the Ornstein–Zernike relation and Baxter's relations are equivalent provided that $\tilde{h}(K)$, the Fourier transform of the total correlation function $h(x)$, is bounded for real K.

I. INTRODUCTION

For simplicity our discussion in this note will be restricted to the one-dimensional simple fluid.[1] The Ornstein—Zernike relation[2] can be written as

$$h(x)=c(x)+\rho\int_{-\infty}^{\infty}h(x-x')c(x')\,dx' \tag{1}$$

where $h(x)$ is the total correlation function and $c(x)$ is the direct correlation function.

The Fourier transform of (1) yields

$$\tilde{h}(K)=\tilde{c}(K)+\rho\tilde{h}(K)\tilde{c}(K) \tag{2}$$

where

$$\tilde{h}(K)=\int_{-\infty}^{\infty}h(x)\exp(iKx)\,dx, \tag{3}$$

$$\tilde{c}(K)=\int_{-\infty}^{\infty}c(x)\exp(iKx)\,dx. \tag{4}$$

Let

$$\tilde{A}(K)=1-\rho\tilde{c}(K). \tag{5}$$

We then obtain from (2) and (5) the following relation:

$$1+\rho\tilde{h}(K)=[\tilde{A}(K)]^{-1}. \tag{6}$$

If $c(x)$ vanishes beyond a finite interval, i.e., $c(x)=0$ for $|x|>l$, then

$$\tilde{c}(K)=\int_{-l}^{l}c(x)\exp(iKx)\,dx=2\int_{0}^{l}c(x)\cos Kx\,dx \tag{7}$$

where we have assumed $c(x)=c(-x)$.

For a disordered fluid $\tilde{h}(K)$ is bounded for real K. Thus $\tilde{h}(K)$ is finite and, from (6), $\tilde{A}(K)$ has no zeros on the real K axis. Following the same proof given by Baxter for the three-dimensional case,[3] we can obtain

Proposition 1: If $c(x)=0$ for $|x|>l$ and $\tilde{h}(K)$ is bounded for real K, the Ornstein—Zernike relation (1) can be transformed into the following relations:

$$c(x)=Q(x)-\rho\int_{x}^{l}Q(x')Q(x'-x)\,dx', \quad 0<x<l, \tag{8}$$

$$h(x)=Q(x)+\rho\int_{0}^{l}h(x-x')Q(x')\,dx', \quad x>0, \tag{9}$$

where $Q(x)$ is a bounded piece-wise continuous real function for $0<x<l$ and $Q(x)=0$ for $x\geq l$.[4]

II. EQUIVALENCE OF BAXTER'S RELATIONS AND THE ORNSTEIN-ZERNIKE RELATION

In this section we first prove that the boundedness of $\tilde{h}(K)$ for real K is not only a necessary condition but also a sufficient condition of Baxter's relations (8) and (9). We then show the equivalence of Baxter's relations and the Ornstein—Zernike relation.

Since $Q(x)=0$ for $x<0$ and $x\geq l$, we define

$$\tilde{Q}(K)=1-\rho\int_{0}^{l}Q(x)\exp(iKx)\,dx \quad (K \text{ real}); \tag{10}$$

then

$$\tilde{Q}^{*}(K)=\tilde{Q}(-K)=1-\rho\int_{0}^{l}Q(x)\exp(-iKx)\,dx \quad (K \text{ real}). \tag{11}$$

From (7), (8), (10), and (11) we can obtain

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{Q}(K)\tilde{Q}(-K)\exp(-iKx)\,dK=\delta(x)-\rho c(x). \tag{12}$$

Hence

$$\tilde{Q}(K)\tilde{Q}(-K)=1-\rho\tilde{c}(K)=\tilde{A}(K). \tag{13}$$

From (3), (9), (10), and (11) we can obtain

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{Q}(K)\tilde{Q}(-K)[1+\rho\tilde{h}(K)]\exp(-iKx)\,dK=\delta(x). \tag{14}$$

Hence

$$\tilde{Q}(-K)\tilde{Q}(K)[1+\rho\tilde{h}(K)]=\tilde{A}(K)[1+\rho\tilde{h}(K)]=1 \tag{15}$$

or

$$\tilde{Q}(K)[1+\rho\tilde{h}(K)]=[\tilde{Q}(-K)]^{-1} \tag{16}$$

But from (3), (9), and (10) we can obtain

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{Q}(K)[1+\rho\tilde{h}(K)]\exp(-iKx)\,dK=0; \tag{17}$$

therefore,

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}[\tilde{Q}(-K)]^{-1}\exp(-iKx)\,dK=0. \tag{18}$$

Although $\tilde{Q}(K)$ and $\tilde{Q}(-K)$ are defined for real K, we can analytically continue $\tilde{Q}(K)$ and $\tilde{Q}(-K)$ into the upper half-plane (uhp) and the lower half-plane (lhp), respectively. $\tilde{Q}(K)$ is then regular in the uhp and $\tilde{Q}(-K)$ is regular in the lhp. From (11) it can be shown that $\tilde{Q}(-K)\to 1$ as $|K|\to\infty$ in the lhp. We can close the contour of integration of (18) in the lhp. It then follows that $\tilde{Q}(-K)$ has no zeros on the real K axis and in the lhp. Since $\tilde{Q}^{*}(K)=\tilde{Q}(-K)=\tilde{Q}(K)$ for real K, $\tilde{Q}(K)$ has no zeros for real K. From (15) we then conclude that $\tilde{h}(K)$ is bounded for real K.

 Copyright © 1975 American Institute of Physics

Proposition 2: Suppose $c(x)=0$ for $|x|>l$. The Ornstein–Zernike relation (1) can be transformed into Baxter's relations (8) and (9) if and only if $\tilde{h}(K)$ is bounded for real K.

Since (15) is identical with (6), we can obtain (2) from (8) and (9). The inverse Fourier transform of (2) then gives (1). Thus from Baxter's relations we can obtain the Ornstein–Zernike relation.

Proposition 3: If $c(x)=0$ for $|x|>l$ and $\tilde{h}(K)$ is bounded for real K, the Ornstein–Zernike relation can be transformed into Baxter's relations. On the other hand, Baxter's relations can be transformed back into the Ornstein–Zernike relation. The Ornstein–Zernike relation and Baxter's relations are equivalent.[5]

III. PERCUS-YEVICK (P.Y.) APPROXIMATION FOR HARD RODS

Wertheim[6] and Thiel[7] have obtained analytic solution of P.Y. integral equation[8] for hard spheres. However, Baxter has been able to obtain the same result from Baxter's relations in a much simpler method.[3] For hard rods with diameter l_0, P.Y. approximation assumes $h(x)=-1$ for $|x|<l_0$. From (10) and (11) we immediately obtain

$$Q(x)=-(1-\rho l_0)^{-1}, \tag{19}$$

$$c(x)=-(1-\rho x)(1-\rho l_0)^{-2}. \tag{20}$$

In terms of Wertheim's notation $Q^2=(1-\rho l_0)^{-2}$, (19) and (20) become $Q(x)=-Q$ and $-c(x)=Q^2(1-\rho x)$ which is identical with Wertheim's solution for the one-dimensional case.

For hard rods with a finite tail potential $v_T(x)$ where $v_T(x)\neq 0$ for $l_0<|x|<l_0+a$ and $v_T(x)=0$ for $|x|<l_0$, $|x|>l_0+a$, a finite, Wertheim[6] had derived the following expression:

$$-c(x)=\xi(-x)+\xi(x)+\rho\int_{-\infty}^{\infty}\xi(x')\xi(x'-x)\,dx' \tag{21}$$

where

$$\xi(x)=Q\theta(x,l_0)-\rho^{-1}n(-x)-m(x-l_0),$$

$$\theta(x,l_0)=\begin{cases}1 & \text{for } 0<x<l_0\\ 0 & \text{otherwise}\end{cases} \tag{22}$$

and $n(x)$, $m(x)$ are involved with a closed set of coupled quadratic type integral equations.

It is interesting to observe the striking similarity between (8) and (21). Particularly for hard rods, we find $\xi(x)=-Q(x)$ and (21) is then identical with (8). However, it should be remarked that Baxter's relations (8) and (9) are valid for any finite l whereas Wertheim's method is valid only for $0<a<l_0$. This indicates that Baxter's relations have a wider range of applicability. It will be of great interest to supplement Baxter's relations with the P.Y. approximation and check if Wertheim's results can be modified for any finite l. This is currently under more careful investigation by the author.

[1]For the three-dimensional case the problem is much more complicated. However, the method discussed in this note can be applied to Propositions 2 and 3.

[2]L.S. Ornstein, and F. Zernike, Proc. Sec. Sci. K. ned. Akad. Wet. **17**, 793 (1914).

[3]Equations (8) and (9) can be obtained following Baxter's method for the three-dimensional case. R.J. Baxter, Aust. J. Phys. **21**, 563 (1968).

[4]If we wish $c(x)$ and $h(x)$ to be bounded and piece-wise continuous for $0<x<l$, then $Q(x)$ must be bounded and piece-wise continuous for $0<x<l$ and vice-versa.

[5]Proposition 3 is still valid if we consider l as a parameter and finally take the limit $l\to\infty$.

[6]M.S. Wertheim, J. Math. Phys. **5**, 643 (1964).

[7]E. Thiel, J. Chem. Phys. **39**, 474 (1963).

[8]J.K. Percus and G. Yevick, Phys. Rev. **110**, 1 (1958).

Geodesics and Newtonian trajectories

Jack Cohn

Department of Physics and Astronomy, University of Oklahoma, Norman, Oklahoma 73069
(Received 25 February 1974; final revised manuscript received 31 July 1974)

In the framework of Riemannian geometry, the problem is considered whether there is a closest geodesic, for all possible space–time metrics, to a given Newtonian trajectory. The meaning of closeness is given precision and the question is answered—in the negative.

I. INTRODUCTION

In a previous paper by the author[1] it was shown that the path taken by objects moving according to Newtonian mechanics (i.e., Newtonian trajectories), which are of the form $d^2x^i/dt^2 = \varphi_{,i}$, will not be geodesics in any Riemannian space–time, if the metric is time independent, and Galilean at infinity.

The question naturally arises then, if, in some sense, there is a closest geodesic to a Newtonian trajectory. The present work will be concerned with the clarification of this question and the formulation of its answer—which turns out to be in the negative.

II. FORMULATION OF PROBLEM

In this section we shall engage in a qualitative discussion and development of the problem, so that it can be stated concisely as a theorem to be proven in the following section.

Consider a Newtonian trajectory (hereafter designated as N-T) in space–time described by the relation

$$\frac{d^2x^i}{dt^2} = \varphi_{,i}, \quad i = 1, 2, 3, \tag{II. 1}$$

in some coordinate frame $\{x^\alpha\}$, and let Latin indices go from 1–3 and Greek indices go from 1–4. Further, let P denote some (space–time) event on this path. P will be taken as the "origin" of the trajectory. Also, we assume that space–time is characterized by some metric $g_{\mu\nu}(x^\sigma)$ (where $x^4 = t$, and $c = 1$), and a corresponding world interval $d\tau^2 = g_{\alpha\beta}\,dx^\alpha\,dx^\beta$, which, for the moment, is arbitrary. Consider the unique tangent geodesic (in this metric) that goes through P and that has the same 4-velocity, $V^\mu = dx^\mu/d\tau$, as does the N-T there. As the space–time metric is changed in the vicinity of P the N-T will be unaffected but the tangent geodesic will generally change. Also, as the magnitude and/or direction of the velocity of the N-T at P changes the tangent geodesic will change. In a given coordinate frame, for some choice of metric, the tangent geodesic may remain closer to the N-T, for any N-T 3-velocity at P, than for any other metric. That is, there might exist a metric for which all the tangent geodesics remain closer to their respective Newtonian trajectories, for all N-T's passing through P, than for any other metric. The existence of such a metric is the problem of concern in this paper.

To make this problem more meaningful, we must sharpen the notion of closeness referred to above. To do this, we appeal to a result of Synge[2] when he considers two near-neighboring space–time paths, one of which is a geodesic. At any world distance τ along the N-T we construct a spacelike geodesic orthogonal to the trajectory and terminating at the tangent geodesic. The world length of this space–like geodesic depends on τ and is denoted as $\phi(\tau)$, where $\tau = 0$ corresponds to event P. If ϕ is expanded about $\tau = 0$, Synge shows that one obtains

$$\phi(\tau) = \tfrac{1}{2}\lambda^\alpha DA_\alpha\,\tau^2 + \tfrac{1}{6}(3D\lambda^\beta DA_\beta + \lambda^\beta D^2A_\beta)\tau^3 + \cdots \tag{II. 2}$$

where all quantities (other than τ) are evaluated at $\tau = 0$ (i.e., event P) and where, furthermore, λ^α denotes a unit 4-vector in the direction of the orthogonal spacelike geodesic; D signifies the absolute derivative with respect to τ, and A^α is the tangent 4-vector to the N-T.

The closeness of a tangent geodesic to its N-T is then characterized by the associated $\phi(\tau)$. We are now interested in that metric which makes $\phi(\tau)$ a minimum, for all τ and all $\mathbf{v}$ (3-velocity of the N-T), at event P—subject to the requirement that, for any given τ, τ is not to change as the range of possible metrics [relative to which $\phi(\tau)$ is to be a minimum] is considered. (So, as the metric is changed—for any given τ—we are actually considering *different* orthogonal spacelike geodesics, i.e., geodesics beginning at different events on the N-T).

The relevant theorem will now be formulated.

III. STATEMENT OF THEOREM

Consider all N-T's issuing through the event P in space–time, each with its associated tangent geodesic (i.e., tangent at P). Let $\mathbf{a}$ denote the 3-vector d^2x^i/dt^2 at P. With each N-T and its tangent geodesic is associated a closeness function $\phi(\tau)$ (as defined above), which also depends on the metric in the vicinity of P and the velocity $\mathbf{v}$ of the N-T at P. Then there is no metric which makes $\phi(\tau)$ a minimum, for all τ (in the sense defined above) and $\mathbf{v}$, if $\mathbf{a} \neq 0$.

IV. PROOF OF THEOREM

It is sufficient to prove the theorem for sufficiently small τ. We assume then that τ is sufficiently small that ϕ can be approximated as

$$\phi(\tau) = \tfrac{1}{2}\lambda^\alpha DA_\alpha\,\tau^2. \tag{IV. 1}$$

To describe the variations we shall need, let $D_{\mu\nu} \equiv \delta g_{\mu\nu}(\delta/\delta g_{\mu\nu})$ denote a variation where only $g_{\mu\nu}$ or $g_{\nu\mu}$ changes by the amount $\delta g_{\mu\nu}$.

Then we shall seek that metric such that

$$D_{\mu\nu}(\lambda^\alpha DA_\alpha) = 0 \tag{IV. 2}$$

for all different pairs (μ, ν), and for all $\mathbf{v}$ at event P; where we also recall that $D_{\mu\nu}\tau = 0$.

 Copyright © 1975 American Institute of Physics

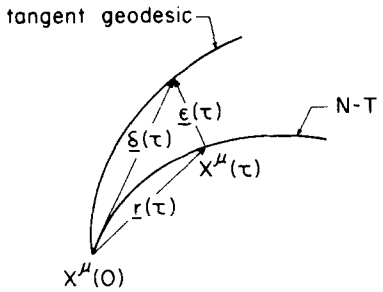


FIG. 1. Synge's construction.

Before we evaluate the above variations, we note a simple but important point. λ^α is a limiting unit spacelike vector at P directed from a N-T to its tangent geodesic there. It must be then that at P, $\lambda^\alpha \| DA^\alpha$ for this reason: Introduce a locally geodesic coordinate system $\{\bar{x}^\mu\}$ at P. Let $\bar{x}^\sigma(\tau)$ denote the geodesic coordinates of an event on the N-T which is world distance τ from P, where τ is very small. Then, expanding about $\tau = 0$ gives

$$\bar{x}^\mu(\tau) - \bar{x}^\mu(0) = \bar{x}^{\mu\prime}(0)\tau + \tfrac{1}{2}\bar{x}^{\mu\prime\prime}(0)\tau^2 + \cdots \tag{IV. 3}$$

where primes signify derivatives.

Now, $\bar{x}^{\mu\prime}(0)$ is just the tangent vector $\bar{T}^\mu$ to the N-T at P. Similarly, $\bar{x}^{\mu\prime\prime}(0)$ is proportional to the principle normal $\bar{N}^\mu$ to the same N-T at P. So, we have this relation in *any* frame,

$$x^\mu(\tau) - x^\mu(0) = T^\mu\tau + \gamma N^\mu\tau^2 + \cdots, \tag{IV. 4}$$

where γ is essentially the proportionality constant mentioned.

We write this in a more suggestive vector notation as

$$\mathbf{r}(\tau) = \mathbf{T}\tau + \gamma\mathbf{N}\tau^2 + \cdots. \tag{IV. 5}$$

Referring to Fig. 1, we also have, in this notation, that

$$\mathbf{r}(\tau) + \boldsymbol{\epsilon}(\tau) = \boldsymbol{\delta}(\tau) \tag{IV. 6}$$

where $\delta(\tau)$ is the 4-vector displacement along the tangent geodesic and $\epsilon(\tau)$ is the spacelike 4-vector, orthogonal to the N-T at τ, going to the tangent geodesic.

Combining these equations yields

$$\boldsymbol{\epsilon}(\tau) = \boldsymbol{\delta}(\tau) - \mathbf{T}\tau - \gamma\mathbf{N}\tau^2 + \cdots. \tag{IV. 7}$$

Now, $\delta(\tau) \| \mathbf{T}$, so from this relation we conclude that, for sufficiently small τ, $\epsilon(\tau)$ is in the plane of **N** and **T**. Thus, the limiting position of $\epsilon(\tau)$, as $\tau \to 0$, must be in the osculating plane at P. Since the limiting direction of $\epsilon(\tau)$ is parallel to λ^α at P, we have then that $\lambda^\alpha \| DA^\alpha$ at P.

Therefore, we have that

$$\lambda^\alpha = -\frac{DA^\alpha}{(DA^\sigma DA_\sigma)^{1/2}} \tag{IV. 8}$$

where the minus sign is introduced since λ^α is actually antiparallel to DA^α, and the signature of the metric is taken as $+2$.

Now, we have the relation

$$\Lambda \equiv \lambda^\alpha DA_\alpha = -(DA^\sigma DA_\sigma)^{1/2} \tag{IV. 9}$$

and we shall be concerned with variations of Λ, i.e., with $D_{\mu\nu}\Lambda$.

Some relevant relations are

$$D_{\mu\nu}(DA^\alpha DA_\alpha) = 2DA_\sigma D_{\mu\nu}(DA^\sigma) + (D_{\mu\nu}g_{\alpha\sigma})DA^\alpha DA^\sigma \tag{IV. 10}$$

and

$$D_{\mu\nu}g_{\alpha\sigma} = \begin{cases} \delta g_{\alpha\sigma} & \text{if } \mu=\alpha,\ \nu=\sigma \text{ or } \mu=\sigma,\ \nu=\alpha \\ \delta g_{\alpha\sigma} & \text{if } \mu=\nu=\alpha=\sigma \\ 0 & \text{otherwise} \end{cases} \tag{IV. 11}$$

Therefore,

$$(D_{\mu\nu}g_{\alpha\sigma})DA^\alpha DA^\sigma = (2-\delta_{\mu\nu})DA^\mu DA^\nu \text{ (no sum on } \mu \text{ and } \nu). \tag{IV. 12}$$

Combining this result with Eqs. (IV. 9) and (IV. 10) then gives

$$D_{\mu\nu}\Lambda = -4(DA^\sigma DA_\sigma)^{-1/2}\{2DA_\alpha D_{\mu\nu}(DA^\alpha) + (2-\delta_{\mu\nu})DA^\mu DA^\nu \delta g_{\mu\nu}\} \tag{IV. 13}$$

where there is no sum on μ and ν on the right-hand side.

Our basic requirement on the metric, Eq. (IV. 2), then becomes

$$2DA_\sigma D_{\mu\nu}(DA^\sigma) = (\delta_{\mu\nu} - 2)DA^\mu DA^\nu \delta g_{\mu\nu} \tag{IV. 14}$$

for all μ and ν, and where μ and ν are not summed. Further, this relation is to hold for all v at P.

We see from this expression that we need to evaluate DA^σ and $D_{\mu\nu}(DA^\sigma)$, which we now proceed to do.

Concerning DA^σ we have the relations

$$DA^\sigma = \frac{d^2x^\sigma}{d\tau^2} + \Gamma^\sigma_{\alpha\beta}\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} \tag{IV. 15}$$

and

$$\frac{dx^\sigma}{d\tau} = \left(\mathbf{v}^i\frac{dt}{d\tau}, \frac{dt}{d\tau}\right), \tag{IV. 16a}$$

$$\frac{d^2x^\sigma}{d\tau^2} = \left(\mathbf{a}^i\left[\frac{dt}{d\tau}\right]^2 + \mathbf{v}^i\frac{d^2t}{d\tau^2}, \frac{d^2t}{d\tau^2}\right), \tag{IV. 16b}$$

$$\frac{dt}{d\tau} = (g_{ij}\mathbf{v}^i\mathbf{v}^j + 2g_{i4}\mathbf{v}^i + g_{44})^{-1/2}, \tag{IV. 16c}$$

$$\frac{d^2t}{d\tau^2} = -\frac{1}{2}\left(\frac{dt}{d\tau}\right)^4(g_{\lambda\alpha,\sigma}\mathbf{v}^\lambda\mathbf{v}^\alpha\mathbf{v}^\sigma + 2g_{i\alpha}\mathbf{a}^i\mathbf{v}^\alpha), \tag{IV. 16d}$$

where $\mathbf{v}^\alpha \equiv dx^\alpha/dt$, $\mathbf{a}^\sigma \equiv d^2x^\sigma/dt^2$, and commas denote ordinary differentiation.

Next, we must find how $D_{\mu\nu}$ affects the various parts of DA^σ, and, finally, DA^σ itself.

By straightforward application of the above relations, we obtain

$$D_{ij}\frac{dt}{d\tau} = -\frac{1}{2}\left(\frac{dt}{d\tau}\right)^3 (2-\delta_{ij})\,\mathbf{v}^i\,\mathbf{v}^j\,\delta g_{ij} \quad \text{(no sum on } i \text{ and } j\text{)}, \tag{IV.17a}$$

$$D_{i4}\frac{dt}{d\tau} = -\frac{1}{2}\left(\frac{dt}{d\tau}\right)^3 2\mathbf{v}^i\,\delta g_{i4} \quad \text{(no sum on } i\text{)}, \tag{IV.17b}$$

$$D_{44}\frac{dt}{d\tau} = -\frac{1}{2}\left(\frac{dt}{d\tau}\right)^3 \delta g_{44}. \tag{IV.17c}$$

Further, from the relation $g^{\sigma\lambda}g_{\sigma\epsilon}=\delta^\lambda_\epsilon$, we obtain the relation

$$\delta\Gamma^\sigma_{\alpha\beta} = \tfrac{1}{2}g^{\sigma\lambda}\{(\delta g_{\alpha\lambda})_{,\beta} + (\delta g_{\lambda\beta})_{,\alpha} - (\delta g_{\alpha\beta})_{,\lambda}\} - \Gamma^\rho_{\alpha\beta}g^{\epsilon\sigma}\delta g_{\rho\epsilon} \tag{IV.18}$$

where δ signifies a variation where *all* the $g_{\mu\nu}$ can vary, and we use the fact that $\delta[(g_{\alpha\beta})_{,\sigma}]=(\delta g_{\alpha\beta})_{,\sigma}$.

From Eq. (IV.18) we then obtain

$$\begin{aligned}(D_{\mu\nu}\Gamma^\sigma_{\alpha\beta})A^\alpha A^\beta = {} & A^\beta(g^{\sigma\nu}A^\mu + g^{\sigma\mu}A^\nu)(\delta g_{\mu\nu})_{,\beta} \\ & - g^{\sigma\lambda}A^\mu A^\nu(\delta g_{\mu\nu})_{,\lambda} \\ & - (\Gamma^\mu_{\alpha\beta}g^{\nu\sigma}A^\alpha A^\beta + \Gamma^\nu_{\alpha\beta}g^{\mu\sigma}A^\alpha A^\beta)\delta g_{\mu\nu}\end{aligned}$$

$(\mu\neq\nu)$ (no sum on μ and ν) (IV.19a)

and

$$\begin{aligned}(D_{\mu\mu}\Gamma^\sigma_{\alpha\beta})A^\alpha A^\beta = {} & \tfrac{1}{2}A^\beta g^{\sigma\mu}A^\mu(\delta g_{\mu\mu})_{,\beta} \\ & - \tfrac{1}{2}g^{\sigma\lambda}A^\mu A^\mu(\delta g_{\mu\mu})_{,\lambda} - \Gamma^\mu_{\alpha\beta}g^{\mu\sigma}A^\alpha A^\beta\,\delta g_{\mu\mu} \\ & \text{(no sum on } \mu\text{)}.\end{aligned} \tag{IV.19b}$$

Having developed the general expressions that will be needed, we now consider several special cases that prove useful.

Case i: All $\mathbf{v}^i=0$.

We now specialize considerations to the case where all $\mathbf{v}^i=0$, at event P. This will enable us to obtain relations which remain valid in the general case where the $\mathbf{v}^i$ may not be zero.

Expressions (IV.17) and (IV.19) now have simpler forms, and, further, we now have the relation

$$D_{i4}\frac{d^2x^\sigma}{d\tau^2} = (0,0,0,-g^{-2}_{44}\mathbf{a}^i\,\delta g_{i4}) \quad \text{(no sum on } i\text{)}. \tag{IV.19c}$$

We can now evaluate our basic requirement on the metric, Eq. (IV.14). As will be explained later we only need explicitly consider the case where $D_{\mu\nu}=D_{i4}$. The left-hand side of the relation is evaluated by expressing it as

$$\begin{aligned}g_{\lambda\sigma}DA^\lambda D_{i4}(DA^\sigma) = {} & g_{lm}DA^l D_{i4}(DA^m) + g_{l4}DA^l D_{i4}(DA^4) \\ & + g_{4l}DA^4 D_{i4}(DA^l) + g_{44}DA^4 D_{i4}(DA^4).\end{aligned} \tag{IV.20}$$

Utilizing Eqs. (IV.17) and (IV.19) the terms here on the right-hand side can all be evaluated. The resulting expression will not be written out here as it is of unwieldy length, however.

Continuing on, the right-hand side of Eq. (IV.14) can be expressed as

$$DA^iDA^4 = g^{-2}_{44}(\mathbf{a}^i + \Gamma^i_{44})[-\tfrac{1}{2}g^{-1}_{44}(g_{44,4} + 2g_{k4}\mathbf{a}^k) + \Gamma^1_{44}]. \tag{IV.21}$$

Inserting Eqs. (IV.20) and (IV.21) into Eq. (IV.14) then yields a very lengthy relation containing the variations δg_{i4} and $(\delta g_{i4})_{,4}$. Since these quantities may be chosen independently at P, their coefficients in the above relation must *separately* vanish. Vanishing of the coefficient of $(\delta g_{i4})_{,4}$ yields the relation

$$\mathbf{a}^i = -\Gamma^i_{44}, \quad i=1,2,3, \tag{IV.22}$$

and then the vanishing of the coefficient of δg_{i4} can be shown to merely yield the identity $0=0$.

Further, it can be shown that making similar considerations on Eq. (IV.14) involving all the other $D_{\mu\nu}\neq D_{i4}$ yields no relations in addition to Eq. (IV.22). We either get relations which are direct consequences of Eq. (IV.22) (such as $\mathbf{a}^i\mathbf{a}^j=\Gamma^i_{44}\Gamma^j_{44}$) or relations of the form $0=0$.

Thus, minimizing Λ with respect to the metric, when all $\mathbf{v}^i=0$ at P, only yields the relations

$$\varphi_{,i} = \mathbf{a}^i = -\Gamma^i_{44}, \quad i=1,2,3. \tag{IV.23}$$

We note that this is the equation of a geodesic if all $\mathbf{v}^i=0$. Thus, the requirement of a closest tangent geodesic to the N-T, when all $\mathbf{v}^i=0$, implies that the N-T is a geodesic *at* P. Thus, when all $\mathbf{v}^i=0$, the closest geodesic *coincides* with (and is not merely tangent to) the N-T at P.

Case ii: Not all $\mathbf{v}^i=0$.

Here, we consider the more general case where not all $\mathbf{v}^i=0$. It will be sufficient for our purposes to assume, however, that only $\mathbf{v}^1\neq 0$ (of course, $\mathbf{v}^4\neq 0$).

Straightforward application of Eqs. (IV.16b), (IV.17a), (IV.17b), (IV.19a), and (IV.19b) yields the following relations:

$$D_{ij}\frac{d^2x^\sigma}{d\tau^2} = (0,0,0,0) \quad (i,j\neq 1), \tag{IV.24a}$$

$$D_{i4}\frac{d^2x^\sigma}{d\tau^2} = \left(-\mathbf{v}^k\left[\frac{dt}{d\tau}\right]^4\mathbf{a}^i\,\delta g_{i4}, \ -\left[\frac{dt}{d\tau}\right]^4\mathbf{a}^i\,\delta g_{i4}\right) \tag{IV.24b}$$

$(i\neq 1)$ (no sum on i),

$$D_{ij}(\Gamma^\sigma_{\alpha\beta}A^\alpha A^\beta) = \left(\frac{dt}{d\tau}\right)^2\left\{-\mathbf{v}^1\mathbf{v}^1 T^{ij\sigma}_{11}\delta g_{ij} - 2\mathbf{v}^1 T^{ij\sigma}_{14}\delta g_{ij} - T^{ij\sigma}_{44}\delta g_{ij}\right\} \quad (i,j\neq 1)\ (i\neq j), \tag{IV.24c}$$

$$D_{i4}(\Gamma^\sigma_{\alpha\beta}A^\alpha A^\beta) = \left(\frac{dt}{d\tau}\right)^2\{-\mathbf{v}^1\mathbf{v}^1 T^{i4\sigma}_{11}\delta g_{i4} + \mathbf{v}^1[g^{\sigma i}(\delta g_{i4})_{,1} - 2T^{i4\sigma}_{14}\delta g_{i4}] + g^{\sigma i}(\delta g_{i4})_{,4} - T^{i4\sigma}_{44}\delta g_{i4}\} \quad (i\neq 1), \tag{IV.24d}$$

where there is no summation on i and j, and where

$$T^{i\lambda\sigma}_{\alpha\beta} \equiv \Gamma^{i}_{\alpha\beta}\, g^{\lambda\sigma} + \Gamma^{\lambda}_{\alpha\beta}\, g^{i\sigma}. \tag{IV.24e}$$

Now, returning to Eq. (IV.14) for the case $D_{\mu\nu} = D_{ij}$ $(i, j \neq 1,\ i \neq j)$, and utilizing the above relations, we obtain the expression

$$S^{ij}_{11}\, \mathbf{v}^1\mathbf{v}^1 + S^{ij}_{14}\, 2\mathbf{v}^1 + S^{ij}_{44} = \left(\frac{dt}{d\tau}\right)^2 (\mathbf{a}^i + \Gamma^{i}_{\alpha\beta}\, \mathbf{v}^\alpha \mathbf{v}^\beta)(\mathbf{a}^j + \Gamma^{j}_{\lambda\sigma}\mathbf{v}^\lambda \mathbf{v}^\sigma) \tag{IV.25}$$

where there is no sum on i and j, and where

$$S^{ij}_{\alpha\beta} \equiv DA^i \Gamma^{j}_{\alpha\beta} + DA^j \Gamma^{i}_{\alpha\beta} \tag{IV.26}$$

Inserting Eq. (IV.16c) in for $dt/d\tau$ in Eq. (IV.25) and then equating coefficients of $(\mathbf{v}^1)^4$ terms yields the relations

$$\Gamma^{k}_{11} = 0 \quad \text{all } k \neq 1. \tag{IV.27}$$

Equating coefficients of $(\mathbf{v}^1)^3$ terms then gives $0 = 0$. Equating coefficients of $(\mathbf{v}^1)^2$ terms then yields the relations

$$\Gamma^{k}_{14} = 0 \ \text{ all } k \neq 1. \tag{IV.28}$$

And, finally, equating coefficients of $\mathbf{v}^1$ terms then yields $0 = 0$.

Since the x^1 direction is not preferred, and the $\Gamma^{\alpha}_{\beta\sigma}$ do not depend on the $\mathbf{v}$, we can generalize the above relations to

$$\underset{\text{(no sum)}}{\Gamma^{k}_{ii}} = 0 \quad \text{all } k \neq i \text{ and } \Gamma^{k}_{i4} = 0, \quad \text{all } k \neq i. \tag{IV.29}$$

We finally consider Eq. (IV.14) once more, where we take $D_{\mu\nu} = D_{i4}$, with $i \neq 1$. Utilizing Eqs. (IV.24b) and (IV.24d) and Eqs. (IV.29), the very lengthy expression reduces to the relation

$$\{-\mathbf{v}^1\mathbf{v}^1\, \Gamma^{4}_{11} DA^i\, \delta g_{i4} + \mathbf{v}^1 [DA^i(\delta g_{i4})_{,1} + \Gamma^{4}_{14} DA^i]\, \delta g_{i4}$$
$$+ DA^i(\delta g_{i4})_{,4} - S^{i4}_{44}\, \delta g_{i4}] - (g_{\lambda 1} DA^\lambda \mathbf{v}^1 \mathbf{a}^i + g_{\lambda 4} DA^\lambda \mathbf{a}^i)\delta g_{i4}$$
$$= (\mathbf{a}^i + \Gamma^{i}_{\alpha\beta}\mathbf{v}^\alpha \mathbf{v}^\beta)\left\{-\frac{1}{2}\left(\frac{dt}{d\tau}\right)^4 \quad g_{\lambda\alpha,\sigma}\mathbf{v}^\lambda \mathbf{v}^\alpha \mathbf{v}^\sigma + 2g_{i\alpha}\, \mathbf{a}^i\, \mathbf{v}^\alpha \right.$$
$$\left. + \Gamma^{4}_{\alpha\beta}\, \mathbf{v}^\alpha \mathbf{v}^\beta \left(\frac{dt}{d\tau}\right)^2 \right\} \delta g_{i4} \tag{IV.30}$$

where there is no sum on i, and $S^{i\sigma}_{\alpha\beta}$ is defined as in Eq. (IV.26).

The entire right-hand side of this equation is zero since the quantity in the first parenthesis is zero by the relations (IV.23) and (IV.29). Considering the left-hand side of the equation, and setting the coefficients of the variations $(\delta g_{i4})_{,4}$, δg_{i4}, and $(\delta g_{i4})_{,k}$ separately to zero yields, first,

$$DA^i = 0, \quad i \neq 1, \tag{IV.31}$$

a relation slightly more general than Eq. (IV.23), and, second,

$$\Gamma^{i}_{44}(g_{11}\, \mathbf{v}^1\mathbf{v}^1 + 2g_{14}\mathbf{v}^1 + g_{44}) = -(g_{41}\mathbf{v}^1\mathbf{a}^i + g_{44}\mathbf{a}^i). \tag{IV.32}$$

Equating coefficients of $(\mathbf{v}^1)^0$, $\mathbf{v}^1$, $(\mathbf{v}^1)^2$, respectively, here, yields the relations

$$\Gamma^{i}_{44} = -\mathbf{a}^i, \quad i \neq 1, \tag{IV.32a}$$

$$2\Gamma^{i}_{44} = -\mathbf{a}^i, \ i \neq 1, \tag{IV.32b}$$

$$\Gamma^{i}_{44} \text{ and/or } g_{11} = 0, \quad i \neq 1. \tag{IV.32c}$$

And these equations can only be compatible if $\Gamma^{i}_{44} = 0 = \mathbf{a}^i$ $(i \neq 1)$, which implies that $\mathbf{a} = 0$, by suitable selection of coordinate axes—thus concluding the theorem.

[1] J. Cohn, Phys. Rev. **149**, 1040 (1966).
[2] J. L. Synge, *Relativity: The General Theory* (North-Holland, Amsterdam, 1960), p. 134.

Long-wavelength normal mode vibrations of infinite, ionic crystal lattices

John A. Davies and Carol L. Mainville

Department of Physics, Clark University, Worcester, Massachusetts 01610
(Received 29 August 1974)

This paper is an extension of an earlier article [J. Math. Phys. **13**, 1207 (1972)] in which a study is made of the long-wavelength normal mode vibrations of infinite, ionic crystal lattices. The work is applicable to lattices of all symmetries but is restricted to the rigid ion model without retardation. This article completes the mathematical formalism introduced in the earlier article. In addition, two theorems are proved. The first states that all branches of the dispersion relations for an ionic lattice will approach definite frequencies as the propagation vector approaches zero if the point group of the space group of the lattice belongs to the regular system. The second states that, for all other lattices, at least two branches of the dispersion relations will approach frequencies which depend upon the direction of the propagation vector as the propagation vector approaches zero. A number of useful techniques are introduced for determining both the qualitative and quantitative behavior of the long-wavelength normal modes. Finally it is shown that the work in this paper is easily extended to include some but not all lattice models with polarizable and deformable ions.

I. INTRODUCTION

This article is an extension of an earlier paper[1] in which we presented a mathematical study of the properties of the long-wavelength normal mode vibrations of crystal lattices in which Coulomb interactions are present.

The dynamical matrix for an infinite lattice with Coulomb interactions does not approach a definite value as the propagation vector k for normal mode vibrations approaches zero. This singular behavior arises from the fact that the expansion of the dynamical matrix about $k=0$ contains a nonanalytic zero order term which depends upon the direction of k.[2-4] The presence of this nonanalytic term has the following consequences:

1. The normal mode frequencies and the eigenvectors of the dynamical matrix are undefined at $k=0$ for infinite ionic lattices. Any attempts to define them lead to contradictions.[5]

2. The normal mode frequencies and eigenvectors are well defined in the long-wavelength limit in the following sense. Given any nonzero k in a small neighborhood of $k=0$, the normal mode frequencies and eigenvectors are well defined at that k.[6] On the other hand, the usual group theoretical methods appropriate at $k=0$ in the absence of Coulomb forces cannot be used without modification to study the long-wavelength normal modes of crystals with Coulomb interactions.[7,8]

3. In the case of some but not all lattices with Coulomb interactions, branches of the dispersion relations occur whose frequencies do not approach definite values as k approaches zero. The value approached depends upon the direction from which k approaches zero.[9,10]

Reference 1 contains a mathematical formalism designed to enable one to extend group theoretical arguments to the analysis of long-wavelength normal modes in crystals with Coulomb interactions. The purpose of this paper is to complete this theory.

The work in this paper and in Ref. 1 is applicable to crystals of all symmetries. Our rigorous analysis is, however, subject to the following restrictions: The lattices dealt with are infinite, retardation is not included, and the rigid ion model is used. Later in this paper arguments based upon the phenomenological theory of ionic lattices[11] are used to show that the treatment in this paper can be extended with only a trivial modification to include a restricted class of lattice models in which polarizable atoms are present.

It is noted above that branches of the dispersion relations which do not approach definite frequencies as k approaches zero exist for some lattices but not for all. Probably the chief deficiency in Ref. 1 is that no simple rule is derived for determining from symmetry alone whether or not a lattice will possess such branches. In this paper we derive such a rule, which is presented in the form of two theorems. These two theorems are stated in detail in Secs. VIII and X, respectively. Briefly, Theorem I states that all branches of the dispersion relations for a lattice will approach definite frequencies if the point group of the space group for the lattice belongs to the regular system. Theorem II states that some branches of the dispersion relations for all other lattices will not approach definite frequencies.

The following is an outline of the work included in this article. In Sec. II, we review the mathematical theory introduced in Ref. 1. Section III is a discussion of the physical significance of some of the operators and linear vector spaces introduced mathematically in Ref. 1. In Secs. IV–VII, we introduce some additional operations and definitions and prove a number of lemmas required for the proofs of Theorems I and II. These sections actually contain more material than is required for the proofs since they are also intended to complete the mathematical theory introduced in Ref. 1. Section VIII is a proof of Theorem I. This section also contains results which are useful in constructing the long-wavelength dispersion relations and normal mode eigenvectors for cubic, ionic lattices. An alternative proof of Theorem I is given in Sec. IX. A by-product of the alternative proof is a technique by which one can obtain the long-wavelength normal mode eigenvectors for cubic

 Copyright © 1975 American Institute of Physics

lattices once they are obtained for a single direction of k. Theorem II is proved in Sec. X. In addition, we obtain results which are useful for analyzing the long-wavelength normal modes for noncubic lattices. Finally, in Sec. XI, we show that the results obtained in this paper can be extended to include a limited class of lattices with polarizable ions.

II. BACKGROUND

We begin by summarizing those results, conventions of notation, and equations from Ref. 1 which are essential for the work developed in this paper. Many of the results of Ref. 1 are restated as lemmas in order that they may be easily referred to in later sections.

Consider an infinite crystal lattice whose particles interact through both Coulomb and short range forces. The particles are assumed to be point ions. There are f particles per primitive cell and the κth particle in the lth primitive cell has the equilibrium position $\eta(l,\kappa)$, mass μ_κ, and charge Z_κ. These quantities are dimensionless since throughout this paper all lengths, masses, and charges are divided by a typical cell dimension a, a typical mass m, or the electronic charge e, respectively. Since each primitive cell is electrically neutral.

$$\sum_\kappa Z_\kappa = 0. \tag{1}$$

A Cartesian coordinate system is imbedded in the lattice. In a normal mode vibration, the ith Cartesian component of the displacement (divided by a) of the κth particle in the lth primitive cell is given by

$$u_i(l,\kappa) = \mu_\kappa^{-1/2}\Psi_{\kappa i}(\phi)\exp[2\pi i\phi\cdot\eta(l,\kappa) - i\omega(\phi)t], \tag{2}$$

where $\phi = ka$ is the dimensionless propagation vector, t is the time, and $\omega(\phi)$ is the frequency. The amplitude factors $\Psi_{\kappa i}(\phi)$ form the elements of a $3f$-component column matrix $\Psi(\phi)$. We shall refer to $\Psi(\phi)$ as a normal mode vector.

In Ref. 1 it is shown that the normal mode vibrations are not defined at $\phi = 0$ for infinite lattices with Coulomb interactions. However, as ϕ approaches zero, the behavior of the normal modes is governed by the eigenvector equation

$$\begin{aligned}\mathbf{C}^0(\hat\phi)\Psi^0(\hat\phi) &= [\mathbf{A} + (4\pi a^3/v_a)\mathbf{N}(\hat\phi)]\Psi^0(\hat\phi)\\ &= \lambda^0(\hat\phi)\Psi^0(\hat\phi).\end{aligned} \tag{3}$$

The $3f\times 3f$ matrix $\mathbf{C}^0(\hat\phi)$ is the long-wavelength limit of the dynamical matrix and is real and symmetric. It is the sum of two real, symmetric matrices, the $\hat\phi$-independent matrix A, and the $\hat\phi$-dependent matrix $(4\pi a^3/v_a)\mathbf{N}(\hat\phi)$, where v_a is the volume of a primitive cell. The column matrix $\Psi^0(\hat\phi)$ gives the behavior of the long-wavelength normal mode eigenvectors, and the eigenvalue $\lambda^0(\hat\phi)$ is related to the long-wavelength dispersion relations by

$$\lambda^0(\hat\phi) = a^3 m\omega^2(\hat\phi)/e^2. \tag{4}$$

The normal mode eigenvectors span a $3f$-dimensional linear vector space denoted by S_{3f}(total).

As we explain in Ref. 1, most $3f\times 3f$ matrices K used in this work consist of three by three submatrices $\mathbf{K}_{\kappa\nu}$, where κ and ν $(=1,2,\ldots,f)$ refer to particle names in the primitive cell. The elements of $\mathbf{K}_{\kappa\nu}$ are $K_{\kappa i,\nu j}$, where i and j $(=1,2,3)$ refer to Cartesian coordinates. Similarly most $3f$-component column vectors V consist of three-component subvectors $\mathbf{V}_\kappa$ with elements $V_{\kappa i}$. Throughout this paper Roman letters will be used to refer to Cartesian components and Greek letters to refer to components designating submatrices.

The submatrices of $\mathbf{N}(\hat\phi)$ are defined by

$$\mathbf{N}_{\kappa\nu}(\hat\phi) = [Z_\kappa Z_\nu/(\mu_\kappa\mu_\nu)^{1/2}]\mathbf{L}(\hat\phi), \tag{5}$$

where

$$\mathbf{L}(\hat\phi) = \phi^{-2}\begin{bmatrix}\phi_1^2 & \phi_1\phi_2 & \phi_1\phi_3\\ \phi_1\phi_2 & \phi_2^2 & \phi_2\phi_3\\ \phi_1\phi_3 & \phi_2\phi_3 & \phi_3^2\end{bmatrix}. \tag{6}$$

The matrix $\mathbf{L}(\hat\phi)$ has the properties that

$$\mathbf{L}(\hat\phi)\hat\phi = \hat\phi, \tag{7}$$

and

$$\mathbf{L}(\hat\phi)\hat\eta(\hat\phi) = 0, \tag{8}$$

for any $\hat\eta(\hat\phi)$ obeying

$$\hat\eta(\hat\phi)\cdot\hat\phi = 0.^{12} \tag{9}$$

An important property of $\mathbf{L}(\hat\phi)$ is the following. Let R be any element of $O(3)$. Then,

$$\mathbf{L}(R\hat\phi) = R\mathbf{L}(\hat\phi)R^\dagger. \tag{10}$$

In Ref. 1, it is shown that the eigenvalue zero of $\mathbf{N}(\hat\phi)$ is $(3f-1)$-fold degenerate. The corresponding eigenvectors span a $(3f-1)$-dimensional subspace of S_{3f}(total) denoted by $S_{3f-1}(\lambda_N = 0;\hat\phi)$. The remaining independent eigenvector of $\mathbf{N}(\hat\phi)$ is provided by the normalized vector $\Psi^{ln}(\hat\phi)$, orthogonal to $S_{3f-1}(\lambda_N = 0;\hat\phi)$, defined by

$$\Psi_\kappa^{ln}(\hat\phi) = \left(\sum_\nu \mu_\nu^{-1}Z_\nu^2\right)^{-1/2} Z_\kappa\mu_\kappa^{-1/2}\hat\phi. \tag{11}$$

The corresponding eigenvalue is $\sum_\nu Z_\nu^2/\mu_\nu$. As $\hat\phi$ varies the subspace $S_{3f-1}(\lambda_N = 0;\hat\phi)$ varies. The intersection for all $\hat\phi$ of the $S_{3f-1}(\lambda_N = 0;\hat\phi)$ is a $(3f-3)$-dimensional subspace of S_{3f}(total) denoted by S_{3f-3}(zero). Similarly, as $\hat\phi$ varies, $\Psi^{ln}(\hat\phi)$ varies and traces out a three-dimensional subspace S_3(normal). S_3(normal) consists of all vectors orthogonal to S_{3f-3}(zero).

The matrix A in Eq. (3) has the usual symmetry properties of a dynamical matrix at $\phi = 0$ for lattices with short range forces only. In particular the unitary matrices of the representation $\{\mathbf{T}(0,\mathrm{R})\}$ of the point group G of the space group of the lattice[13] commute with A. That is,

$$\mathbf{T}(0,\mathrm{R})\mathbf{A}\mathbf{T}^\dagger(0,\mathbf{R}) = \mathbf{A}, \tag{12}$$

where

$$\mathbf{T}_{\kappa\nu}(0,\mathbf{R}) = \mathbf{R}\delta(\kappa, F_0(\nu,\mathbf{R})). \tag{13}$$

In Eq. (13), R is an orthogonal matrix representing an element of the point group G for the lattice, and $F_0(\nu,\mathbf{R}) = \nu'$, where particles of type ν are brought into coincidence with identical particles of type ν' under an operation of the space group involving the point group operation R. Since ν and $F_0(\nu,\mathbf{R})$ are identical particles,

$$Z_{F_0(\nu,\mathbf{R})} = Z_\nu, \tag{14}$$

and

$$\mu_{F_0(\nu,\mathbf{R})} = \mu_\nu. \tag{15}$$

Properties of the eigenvalues and eigenvectors of **A** are then found, using the usual group theoretical methods, by determining the irreducible representations of the point group contained in the representation $\{\mathbf{T}(0, \mathbf{R})\}$. (For details see Refs. 1 and 13.) The matrix **A** also obeys the relation

$$\sum_\kappa \mu_\kappa^{1/2}\mathbf{A}_{\kappa\nu} = \sum_\kappa \mathbf{A}_{\nu\kappa}\mu_\kappa^{1/2} = 0. \tag{16}$$

Equation (16) insures the existence of three independent acoustic modes for **A**. The matrices $\mathbf{N}(\hat{\phi})$ and $\mathbf{C}^0(\hat{\phi})$ also obey equations of the form of Eq. (16).

A great deal of information concerning the eigenvalues and eigenvectors of $\mathbf{C}^0(\hat{\phi})$ can be determined once a group theoretical analysis of the eigenvalues and eigenvectors of A is carried out. In Ref. 1, a number of rules (called Cases a, b, and c) for carrying out such an analysis are derived. These cases are rewritten below as lemmas:

Lemma I: If an eigenvector of **A** corresponding to the eigenvalue λ^a lies in S_{3f-3}(zero), then it is also an eigenvector of $\mathbf{C}^0(\hat{\phi})$ corresponding to the same eigenvalue λ^a. Suppose A has a q-fold degenerate eigenvalue λ^a with corresponding eigenvectors spanning a q-dimensional subspace of S_{3f}(total) contained in S_{3f-3}(zero). If $\{\Psi^{an}\}$, $n = 1, 2, \ldots, q$, is any set of q linearly independent eigenvectors of **A** corresponding to the eigenvalue λ^a, then the long-wavelength eigenvectors for q branches of the dispersion relations can be constructed from linear combinations[14] of the Ψ^{an}. These branches of the dispersion relations will approach the definite frequency corresponding to λ^a as ϕ approaches zero.

Lemma II: Suppose **A** has a q-fold degenerate eigenvalue λ^a with corresponding eigenvectors spanning a q-dimensional subspace S_q of S_{3f}(total), where S_q is not contained in S_{3f-3}(zero). Let $\{\Psi^{an}\}$ be a set of q linearly independent eigenvectors of A corresponding to the eigenvalue λ^a. For any given $\hat{\phi}$, the intersection of S_q and $S_{3f-1}(\lambda_N = 0; \hat{\phi})$ is $(q-1)$-dimensional.[15] Thus, for any given $\hat{\phi}$, we can construct $(q-1)$ independent eigenvectors of $\mathbf{C}^0(\hat{\phi})$ which are linear combinations of the Ψ^{an} and lie in $S_{3f-1}(\lambda_N = 0; \hat{\phi})$. As $\hat{\phi}$ varies, the eigenvectors so constructed vary and, if properly chosen,[16] trace out the long-wavelength normal mode vectors for $(q-1)$ branches of the dispersion relations. Each of the above $(q-1)$ branches approaches the definite frequency corresponding to λ^a as $\hat{\phi}$ approaches zero.

Lemma III: Suppose **A** has a q-fold degenerate eigenvalue λ^a with corresponding eigenvectors spanning a q-dimensional subspace S_q of S_{3f}(total). Also suppose that S_q contains S_3(normal). Let $\{\Psi^{an}\}$ be a set of q linearly independent eigenvectors of **A** corresponding to the eigenvalue λ^a. Then for any $\hat{\phi}$, we can construct a linear combination of the Ψ^{an} which is parallel to $\Psi^{ln}(\hat{\phi})$. As $\hat{\phi}$ varies, $\Psi^{ln}(\hat{\phi})$ varies and provides the long-wavelength normal mode vectors for one branch of the dispersion relations which approaches the definite frequency corresponding to $\lambda^a + (4\pi a^3/v_a)\sum_\kappa Z_\kappa^2/\mu_\kappa$ as ϕ approaches zero. Further, for each $\hat{\phi}$, we can construct $(q-1)$ linear combinations of the Ψ^{an} which are mutually orthogonal and orthogonal to $\Psi^{ln}(\hat{\phi})$. The $(q-1)$ linear combinations lie in $S_{3f-1}(\lambda_N = 0; \hat{\phi})$ and are, thus, eigenvectors of $\mathbf{C}^0(\hat{\phi})$. As $\hat{\phi}$ varies these vectors vary and, if properly chosen, trace out the long-wavelength normal mode vectors for $(q-1)$ branches of the dispersion relations. Each such branch approaches the definite frequency corresponding to λ^a as $\hat{\phi}$ approaches zero.

When Coulomb interactions are present in an infinite lattice, some of the branches of its dispersion relations may not approach definite frequencies in the long-wavelength limit. If the eigenvalue $\lambda^0(\hat{\phi})$ in Eq. (3) actually is $\hat{\phi}$-dependent, then the frequency approached will depend upon the direction from which ϕ approaches zero. In Ref. 1, a necessary and sufficient condition that a single given branch approaches a definite frequency is derived. This condition is rewritten below as Lemma IV.

Lemma IV: A necessary and sufficient condition that a single given branch of the dispersion relations approaches a definite frequency as ϕ approaches zero is that the $\Psi^0(\hat{\phi})$ for the branch lie either in $S_{3f-1}(\lambda_N = 0; \hat{\phi})$ or in $S_{3f-2}(\lambda_M = 0; \hat{\phi})$ for all $\hat{\phi}$.

The subspace $S_{3f-2}(\lambda_M = 0; \hat{\phi})$ is defined as follows. First define the $3f \times 3f$ matrix $\mathbf{M}(\hat{\phi})$ by

$$\mathbf{M}_{\kappa\nu}(\hat{\phi}) = [Z_\kappa Z_\nu/(\mu_\kappa\mu_\nu)^{1/2}]\mathbf{T}(\hat{\phi}), \tag{17}$$

where

$$\mathbf{T}(\hat{\phi}) = \phi^{-2}\begin{bmatrix} \phi_2^2+\phi_3^2 & -\phi_1\phi_2 & -\phi_1\phi_3 \\ -\phi_1\phi_2 & \phi_1^2+\phi_3^2 & -\phi_2\phi_3 \\ -\phi_1\phi_3 & -\phi_2\phi_3 & \phi_1^2+\phi_2^2 \end{bmatrix}. \tag{18}$$

The $(3f-2)$-dimensional subspace $S_{3f-2}(\lambda_M = 0; \hat{\phi})$ is the set of all vectors satisfying the equation

$$\mathbf{M}(\hat{\phi})\Psi = 0. \tag{19}$$

Thus, it is the subspace occupied by the eigenvectors of $\mathbf{M}(\hat{\phi})$ corresponding to the $(3f-2)$-fold degenerate eigenvalue zero. In Ref. 1 it is noted that $S_{3f-1}(\lambda_N = 0; \hat{\phi}) \cap S_{3f-2}(\lambda_M = 0; \hat{\phi}) = S_{3f-3}$(zero) for any $\hat{\phi}$.

The matrix $\mathbf{M}(\hat{\phi})$ also has the twofold degenerate eigenvalue $\sum_\kappa Z_\kappa^2/\mu_\kappa$ with corresponding eigenvectors Ψ of the form

$$\Psi_\kappa = \left(\sum_\nu \mu_\nu^{-1}Z_\nu^2\right)^{-1/2}\mu_\kappa^{-1/2}Z_\kappa\hat{\eta}(\hat{\phi}), \tag{20}$$

where

$$\hat{\phi}\cdot\hat{\eta}(\hat{\phi}) = 0. \tag{21}$$

The two-dimensional subspace of S_{3f}(total) consisting of these eigenvectors is denoted by $S_2(\lambda_M = \sum_\kappa Z_\kappa^2/\mu_\kappa; \hat{\phi})$. The union for all $\hat{\phi}$ of the $S_2(\lambda_M = \sum_\kappa Z_\kappa^2/\mu_\kappa; \hat{\phi})$ is S_3(normal). Since the latter result is not proved in Ref. 1, we prove it in Sec. VI.

The 3×3 matrix $\mathbf{T}(\hat{\phi})$ has the properties that

$$\mathbf{T}(\hat{\phi})\hat{\phi} = 0 \tag{22}$$

and

$$\mathbf{T}(\hat{\phi})\hat{\eta}(\hat{\phi}) = \hat{\eta}(\hat{\phi}). \tag{23}$$

In Ref. 1, we considered but did not solve the problem of deriving a necessary and sufficient condition, based on symmetry, that all branches of the dispersion relations for a lattice approach definite frequencies (independent of the direction of $\hat{\phi}$) in the long-wavelength limit. In Ref. 1 the following lemma is stated:

Lemma V: A necessary and sufficient condition that all branches of the dispersion relations for a lattice approach definite frequencies as ϕ approaches zero is that there exists a unitary matrix $\mathbf{Q}(\mathbf{R})$ such that for all $\mathbf{R} \in O(3)$ and some $\hat{\phi}$.

$$\mathbf{C}^0(\mathbf{R}\hat{\phi}) = \mathbf{Q}^\dagger(\mathbf{R})\mathbf{C}^0(\hat{\phi})\mathbf{Q}(\mathbf{R}). \tag{24}$$

The above condition is too general to provide any direct relation between the lattice symmetry and the property that all branches approach definite frequencies. Bakr[17] attempted, with only limited success, to prove the two theories quoted in Sec. I by directly considering the relation between the lattice symmetry and the existence or nonexistence of $\mathbf{Q}(\mathbf{R})$. In this paper, we prove the theorems by using an alternative approach and thus obtain the corollary that $\mathbf{Q}(\mathbf{R})$ exists for all lattices whose point groups belong to the regular system and does not exist for other lattices (barring accidental degeneracies in the long-wavelength dispersion relations). In Sec. IX of this paper, we show how $\mathbf{Q}(\mathbf{R})$ can be constructed when it exists.

III. SOME REMARKS ON THE PHYSICS OF THE FORMALISM

As is well known, the term $(4\pi a^3/v_a)\mathbf{N}(\hat{\phi})$ in Eq. (3) is interpreted as the contribution of a macroscopic electric dipole field of amplitude $\mathbf{E}$ to the zero order dynamical matrix $\mathbf{C}^0(\hat{\phi})$. In most treatments of infinite lattices, including this paper, this macroscopic field arises solely from electric dipole moments which appear at the lattice sites due to lattice vibrations and is not in part due to a field imposed from outside (at infinity). Referring either to Eq. (31.14) of Ref. 3 or to Eq. (6.2.27) of Ref. 4 and taking into account the dimensionless nature of the quantities used in this paper, we see that

$$\left[4\pi\frac{a^3}{v_a}\mathbf{N}(\hat{\phi})\Psi\right]_\kappa = -\frac{Z_\kappa}{\mu_\kappa^{1/2}}\mathbf{E}, \tag{25}$$

where $\mathbf{E}$ is a dimensionless macroscopic field amplitude related to $\mathbf{E}$ by

$$\mathbf{E} = (a^2/e)\mathbf{E}, \tag{26}$$

and $\mathbf{E}$ is the macroscopic electric field amplitude arising from a long-wavelength normal mode vibration propagating in the $\hat{\phi}$ direction and characterized by the normal mode vector Ψ. We shall thus sometimes refer to $\mathbf{N}(\hat{\phi})$ as the macroscopic electric field operator.

Next consider the operator $\mathbf{N}(\hat{\phi}) + \mathbf{M}(\hat{\phi})$. Using Eqs. (5), (6), (17), and (18), we see that

$$\mathbf{N}_{\kappa\nu}(\hat{\phi}) + \mathbf{M}_{\kappa\nu}(\hat{\phi}) = \frac{Z_\kappa Z_\nu}{(\mu_\kappa\mu_\nu)^{1/2}}\mathbf{I}, \tag{27}$$

where $\mathbf{I}$ is the 3×3 identity matrix. Notice that $\mathbf{M}(\hat{\phi}) + \mathbf{N}(\hat{\phi})$ is independent of $\hat{\phi}$. From Eq. (27), one easily obtains the result

$$\left[4\pi\frac{a^3}{v_a}[\mathbf{N}(\hat{\phi}) + \mathbf{M}(\hat{\phi})]\Psi\right]_\kappa = 4\pi\frac{Z_\kappa}{\mu_\kappa^{1/2}}\mathbf{P}, \tag{28}$$

where $\mathbf{P}$ is the amplitude of the dimensionless macroscopic polarization vector, related to the amplitude of the macroscopic polarization vector $\mathbf{P}$ by

$$\mathbf{P} = (a^2/e)\mathbf{P}, \tag{29}$$

where

$$\mathbf{P} = \frac{1}{v_a}\sum_\kappa (eZ_\kappa)a\frac{\Psi_\kappa}{\mu_\kappa^{1/2}}. \tag{30}$$

We sometimes refer to $\mathbf{N}(\hat{\phi}) + \mathbf{M}(\hat{\phi})$ as the macroscopic polarization operator.

Finally, from Eqs. (25), (26), (28), and (29), we obtain the result

$$\left[4\pi\frac{a^3}{v_a}\mathbf{M}(\hat{\phi})\Psi\right]_\kappa = \frac{Z_\kappa}{\mu_\kappa^{1/2}}(\mathbf{E} + 4\pi\mathbf{P}) = \frac{Z_\kappa}{\mu_\kappa^{1/2}}\mathbf{D}, \tag{31}$$

where

$$\mathbf{D} = (a^2/e)\mathbf{D} \tag{32}$$

and, clearly, $\mathbf{D}$ is the amplitude of the macroscopic electric displacement vector. We shall, therefore, refer to $\mathbf{M}(\hat{\phi})$ as the macroscopic electric displacement operator.

In the long-wavelength limit the electrostatic equations $\nabla\times\mathbf{E}(\mathbf{r}) = 0$ and $\nabla\cdot\mathbf{D}(\mathbf{r}) = 0$ reduce to $\hat{\phi}\times\mathbf{E} = 0$ and $\hat{\phi}\cdot\mathbf{D} = 0$, respectively. Using Eqs. (5), (8), (17), (22), (25), and (31), one sees that the latter two equations are automatically obeyed by $\mathbf{E}$ and $\mathbf{D}$.

Clearly the subspace $S_{3f-1}(\lambda_N = 0; \hat{\phi})$ consists of all normal mode eigenvectors Ψ for which the corresponding long-wavelength vibrations with propagation vectors in the $\hat{\phi}$ direction produce no macroscopic electric field. These are vibrations whose longitudinal components produce zero dipole moment per primitive cell (that is, $\hat{\phi}\cdot\mathbf{P} = 0$). Similarly, the subspace $S_{3f-2}(\lambda_M = 0; \hat{\phi})$ consists of all normal mode vectors Ψ for which the corresponding long-wavelength vibrations with propagation vectors in the $\hat{\phi}$ direction produce no macroscopic electric displacement vector. These are vibrations whose transverse components produce zero dipole moment per primitive cell (that is, $\hat{\eta}(\hat{\phi})\cdot\mathbf{P} = 0$ for all $\hat{\eta}(\hat{\phi}) \perp \hat{\phi}$). The subspace S_{3f-3}(zero) consists of all Ψ whose corresponding long-wavelength normal mode vibrations (for any $\hat{\phi}$) produce no polarization vector. The latter subspace is the intersection of $S_{3f-1}(\lambda_N = 0; \hat{\phi})$ and $S_{3f-2}(\lambda_M = 0; \hat{\phi})$ for any $\hat{\phi}$. It follows that whenever $\mathbf{P} = 0$ both $\mathbf{E}$ and $\mathbf{D}$ vanish. This result must, of course, be true since (as we state above) the macroscopic electric field is assumed to arise solely from dipole moments which arise from particle displacements and is not in part due to an applied outside field.

A long-wavelength normal mode vibration propagating in the $\hat{\phi}$ direction and having the normal mode vector $k\Psi^{ln}(\hat{\phi})$, where k is a constant, is, of course, purely longitudinal. Clearly $\mathbf{M}(\hat{\phi})\Psi^{ln}(\hat{\phi}) = 0$. From Eqs. (11), (25), (28), and (31), it follows that the above vibration produces $\mathbf{E} = -4\pi\mathbf{P} = -k(4\pi a^3/v_a)(\sum_\omega Z_\omega^2/\mu_\omega)^{1/2}\hat{\phi}$ and $\mathbf{D} = 0$. Similarly, consider a long-wavelength normal mode vibration propagating in the $\hat{\phi}$ direction with normal

mode vector $k\Psi$, where Ψ is given by Eq. (20). Such a vibration is purely transverse and produces $\mathbf{D} = 4\pi\mathbf{P} = k(4\pi a^3/v_a)(\sum_\omega Z_\omega^2/\mu_\omega)^{1/2}\hat{\eta}(\hat{\phi})$ and $\mathbf{E} = 0$.

Lemma IV is stated in mathematical terms. Clearly it can be restated as follows: *A necessary and sufficient condition that a single given branch of the dispersion relations approaches a definite frequency is that for all long-wavelength normal mode vibrations belonging to the branch either* $\mathbf{E} = 0$ *or* $\mathbf{D} = 0$.

IV. AN ALTERNATIVE FORM FOR $C^0(\hat{\phi})$. THE MATRIX A'

Usually the matrix $\mathbf{C}^0(\hat{\phi})$ is written in the form given by Eq. (3) where the nonanalytic $\hat{\phi}$-dependent term is proportional to the macroscopic electric field operator $\mathbf{N}(\hat{\phi})$. An alternative method of writing $\mathbf{C}^0(\hat{\phi})$ is obtained as follows.

Define the matrix $\mathbf{A}'$ by

$$\mathbf{A}' = \mathbf{A} + (4\pi a^3/v_a)[\mathbf{N}(\hat{\phi}) + \mathbf{M}(\hat{\phi})]. \tag{33}$$

According to Eq. (27), the polarization operator $\mathbf{N}(\hat{\phi}) + \mathbf{M}(\hat{\phi})$ is independent of $\hat{\phi}$ and is real and symmetric. Thus, the matrix $\mathbf{A}'$ is constant, real, and symmetric. In terms of $\mathbf{A}'$, we may express $\mathbf{C}^0(\hat{\phi})$ in the form

$$\mathbf{C}^0(\hat{\phi}) = \mathbf{A}' - (4\pi a^3/v_a)\mathbf{M}(\hat{\phi}). \tag{34}$$

In this alternative form, the nonanalytic part of $\mathbf{C}^0(\hat{\phi})$ is proportional to the electric displacement operator $\mathbf{M}(\hat{\phi})$.

Consider the representation $\{\mathbf{T}(0, \mathbf{R})\}$ of the point group G for a lattice as defined by Eq. (13). The matrices of this representation, of course, commute with $\mathbf{A}$. Using Eqs. (13), (14), (15), and (27), we see that all of the $\mathbf{T}(0, \mathbf{R})$ also commute with $\mathbf{N}(\hat{\phi}) + \mathbf{M}(\hat{\phi})$. It then follows from Eq. (33) that $\{\mathbf{T}(0, \mathbf{R})\}$ provides a representation of the point group all of whose matrices commute with $\mathbf{A}'$. We conclude that any group theoretical analysis of the properties of the eigenvectors and eigenvalues of $\mathbf{A}$ based upon the point group apply equally well to the eigenvalues and eigenvectors of $\mathbf{A}'$.

From Eqs. (16) and (33), the remarks immediately following Eq. (16), and the equation $\sum_\kappa \mu_\kappa^{1/2}\mathbf{M}_{\kappa\nu}(\hat{\phi}) = \sum_\kappa \mathbf{M}_{\nu\kappa}(\hat{\phi})\mu_\kappa^{1/2} = 0$ [which follows immediately from Eq. (1)], we see that $\mathbf{A}'$ also obeys an equation of the form of Eq. (16). The latter result ensures the existence of three independent acoustic modes for $\mathbf{A}'$. Both $\mathbf{A}$ and $\mathbf{A}'$ are real and symmetric. Since both are real, any general conclusions based upon time reversal invariance apply equally to $\mathbf{A}$ and $\mathbf{A}'$.

We finally conclude that any properties of the eigenvectors and eigenvalues of $\mathbf{A}$ derived in this paper by group theoretical methods can be applied with equal validity to the eigenvectors and eigenvalues of $\mathbf{A}'$.

V. ACCIDENTAL AND NONACCIDENTAL DEGENERACIES

For purposes of this paper a degeneracy in an eigenvalue of $\mathbf{A}$ or of $\mathbf{A}'$ will be considered as not accidental under any of the following circumstances:

1. If it follows from translational invariances; that is, if the degeneracy follows from Eq. (16) or from the corresponding equation for $\mathbf{A}'$.

2. If it follows from an analysis of the space group of the lattice; that is, if it follows from the irreducible representations of the point group G contained in $\{\mathbf{T}(0, \mathbf{R})\}$.

3. If it follows from time reversal invariance.

A degeneracy in an eigenvalue of $\mathbf{C}^0(\hat{\phi})$ is considered to be nonaccidental under any of the following circumstances:

1. If it follows from translational invariance; that is, if it follows from the equation involving $\mathbf{C}^0(\hat{\phi})$ which corresponds to Eq. (16).

2. If it follows from time reversal invariance.

3. If by using methods described in this paper and Ref. 1, we can show that the degeneracy is a consequence of nonaccidental degeneracies in eigenvalues of $\mathbf{A}$ and $\mathbf{A}'$ (that is, if the degeneracy follows directly or indirectly from Lemmas I, II, or III).

A final type of degeneracy in an eigenvalue of $\mathbf{C}^0(\hat{\phi})$ which will be considered as nonaccidental arises as follows. Consider the subgroup of the point group G for the lattice consisting of all $\mathbf{R} \in G$ having the property that $\mathbf{R}\hat{\phi} = \hat{\phi}$. We denote this subgroup by $G(\hat{\phi})$ and refer to it as the group of the direction of $\hat{\phi}$. A representation of $G(\hat{\phi})$, all of whose elements commute with $\mathbf{C}^0(\hat{\phi})$ is the set of all $\mathbf{T}(0, \mathbf{R})$ for which $\mathbf{R} \in G(\hat{\phi})$. A degeneracy in an eigenvalue of $\mathbf{C}^0(\hat{\phi})$ will be considered as nonaccidental if it follows from an analysis of the irreducible representations of $G(\hat{\phi})$ contained in $\{\mathbf{T}(0, \mathbf{R})\}$, $\mathbf{R} \in G(\hat{\phi})$.

All other degeneracies are regarded as accidental. In the interests of simplicity, we assume throughout this paper that such accidental degeneracies are not present.

VI. PROJECTION OPERATORS AND SOME ASSOCIATED THEOREMS

The work to follow in this paper depends heavily upon the use of projection operators to the various subspaces of S_{3f}(total) defined in Sec. II. Most of the projection operators are derived in Ref. 1 and the remainder can be derived by standard methods. Thus, we merely list expressions for them below [where $\mathbf{P}(S_q)$ is the projection operator to the subspace S_q, $\mathbf{I}$ is the 3×3 identity matrix, and $\mathbf{I}$ is the $3f\times3f$ identity matrix]:

$$\mathbf{P}(S_{3f-1}(\lambda_N = 0; \hat{\phi})) = \mathbf{I} - \left(\sum_\omega Z_\omega^2/\mu_\omega\right)^{-1}\mathbf{N}(\hat{\phi}) \tag{35}$$

$$\mathbf{P}(S_{3f-2}(\lambda_M = 0; \hat{\phi})) = \mathbf{I} - \left(\sum_\omega Z_\omega^2/\mu_\omega\right)^{-1}\mathbf{M}(\hat{\phi}), \tag{36}$$

$$\mathbf{P}_{\kappa\nu}(S_3(\text{acoustic})) = \left(\sum_\omega \mu_\omega\right)^{-1}(\mu_\kappa\mu_\nu)^{1/2}\mathbf{I}, \tag{37}$$

$$\mathbf{P}(\Psi^{In}(\phi)) = \left(\sum_\omega Z_\omega^2/\mu_\omega\right)^{-1}\mathbf{N}(\hat{\phi}), \tag{38}$$

$$\mathbf{P}_{\kappa\nu}(S_3(\text{normal})) = \left(\sum_\omega \mu_\omega^{-1}Z_\omega^2\right)^{-1}\frac{Z_\kappa Z_\nu}{(\mu_\kappa\mu_\nu)^{1/2}}\mathbf{I}, \tag{39}$$

$$\mathbf{P}(S_{3f-3}(\text{zero})) = \mathbf{I} - \mathbf{P}(S_3(\text{normal})). \tag{40}$$

and

$$\mathbf{P}\left[S_2\left(\lambda_M=\sum_\kappa Z_\kappa^2/\mu_\kappa;\hat{\boldsymbol{\phi}}\right)\right]=\mathbf{I}-\mathbf{P}(S_{3f-2}(\lambda_M=0;\hat{\boldsymbol{\phi}})). \tag{41}$$

In Eq. (37), S_3(acoustic) is the subspace of acoustic mode eigenvectors for **A**, **A**′, and $\mathbf{C}^0(\hat{\boldsymbol{\phi}})$.

The projection operator to the subspace of S_{3f}(total) spanned by eigenvectors of **A** belonging to the ith row of the μth irreducible representation of the point group G for the lattice is denoted by $\mathbf{P}^{\mu,i}$ which is defined by

$$\mathbf{P}^{\mu,i}=\frac{n_\mu}{g}\sum_{\mathbf{R}\in G}D_{ii}^{(\mu)*}(\mathbf{R})\mathbf{T}(0,\mathbf{R}), \tag{42}$$

where g is the order of the point group, n_μ is the dimension of the μth irreducible representation, and $\mathbf{D}^{(\mu)}(\mathbf{R})$ is the matrix representing the point group element **R** in the μth irreducible representation.[18] Summing both sides of Eq. (37) over i, we obtain the projection operator $\mathbf{P}^\mu$ to the subspace of S_{3f}(total) spanned by eigenvectors of **A** belonging to the μth irreducible representation. We get

$$\mathbf{P}^\mu=\frac{n_\mu}{g}\sum_{\mathbf{R}}\chi^{(\mu)*}(\mathbf{R})\mathbf{T}(0,\mathbf{R}), \tag{43}$$

where $\chi^{(\mu)}(\mathbf{R})$ is the character (trace) of $D^{(\mu)}(\mathbf{R})$.

Suppose that $\mathbf{D}^{(\mu)}$ is an n_μ-dimensional irreducible representation of the point group for the lattice and $\Psi^{\mu,i}$ is a vector belonging to the ith row of $\mathbf{D}^{(\mu)}$. Then a set of n_μ mutually orthogonal vectors $\Psi^{\mu,j}$ which span an n_μ-dimensional subspace of S_{3f}(total) which is invariant under the operations of the $\mathbf{T}(0,\mathbf{R})$ is obtained by applying the formula

$$\Psi^{\mu,j}=\frac{n_\mu}{g}\sum_{\mathbf{R}}D_{ji}^{(\mu)*}(\mathbf{R})\mathbf{T}(0,\mathbf{R})\Psi^{\mu,i}. \tag{44}$$

The vector $\Psi^{\mu,j}$ transforms according to the jth row of $\mathbf{D}^{(\mu)}$. [For a derivation of Eq. (44) see Ref. 18.]

We now derive a number of formulas and theorems concerning the above projection operators and their corresponding subspaces of S_{3f}(total).

The subspace S_{3f-3}(zero) is defined as the intersection of the subspaces $S_{3f-1}(\lambda_N=0;\hat{\boldsymbol{\phi}})$ for all $\hat{\boldsymbol{\phi}}$. The intersection of the subspaces $S_{3f-2}(\lambda_M=0;\boldsymbol{\phi})$ for all $\boldsymbol{\phi}$ is the set of all Ψ obeying the equation $\mathbf{M}(\hat{\boldsymbol{\phi}})\Psi=0$ for all $\hat{\boldsymbol{\phi}}$. In terms of submatrices the latter equation becomes $\mathbf{T}(\hat{\boldsymbol{\phi}})\sum_\lambda Z_\kappa Z_\lambda(\mu_\kappa\mu_\lambda)^{-1/2}\Psi_\lambda=0$. Using Eq. (18), one can easily show that this equation holds true for all $\hat{\boldsymbol{\phi}}$ if and only if $\sum_\lambda Z_\kappa Z_\lambda(\mu_\kappa\mu_\lambda)^{-1/2}\Psi_\lambda=0$. That is, Ψ lies in S_{3f-3}(zero). We, therefore, add the following lemma:

Lemma VI: The intersection of the subspaces $S_{3f-2}(\lambda_M=0;\hat{\boldsymbol{\phi}})$ for all $\hat{\boldsymbol{\phi}}$ is S_{3f-3}(zero).

The subspace $S_2(\lambda_M=\sum_\kappa Z_\kappa^2/\mu_\kappa;\hat{\boldsymbol{\phi}})$ consists of all vectors normal to $S_{3f-2}(\lambda_M=0;\hat{\boldsymbol{\phi}})$. Using Eq. (20), one can easily prove that the union of the $S_2(\lambda_M=\sum_\kappa Z_\kappa^2/\mu_\kappa;\boldsymbol{\phi})$ for all $\hat{\boldsymbol{\phi}}$ is at least three-dimensional. The subspace S_3(normal) consists of all vectors normal to S_{3f-3}(zero). It follows from Lemma VI that all Ψ contained in the union of the $S_2(\lambda_M=\sum_\kappa Z_\kappa^2/\mu_\kappa;\hat{\boldsymbol{\phi}})$ for all $\hat{\boldsymbol{\phi}}$ are contained in S_3(normal). Thus, we add the following lemma:

Lemma VII: The union of the $S_2(\lambda_M=\sum_\kappa Z_\kappa^2/\mu_\kappa;\hat{\boldsymbol{\phi}})$ for all $\hat{\boldsymbol{\phi}}$ is S_3(normal).

Using Eqs. (1), (37), (39), and (40), we obtain the equation

$$\begin{aligned}&\mathbf{P}(S_{3f-3}(\text{zero}))\mathbf{P}(S_3(\text{acoustic}))\\&\quad=\mathbf{P}(S_3(\text{acoustic}))\mathbf{P}(S_{3f-3}(\text{zero}))\\&\quad=\mathbf{P}(S_3(\text{acoustic})).\end{aligned} \tag{45}$$

An immediate consequence is the following lemma:

Lemma VIII: S_3(acoustic)$\subseteq S_{3f-3}$(zero). Lemma VIII is a statement of the obvious physical fact that acoustic modes are nonpolar.

Using Eqs. (13), (14), (15), (39), and (40), we find that $\mathbf{T}(0,\mathbf{R})$ and $\mathbf{P}(S_{3f-3}(\text{zero}))$ commute for all **R** of the point group. That is,

$$[\mathbf{T}(0,\mathbf{R}),\mathbf{P}(S_{3f-3}(\text{zero}))]=0. \tag{46}$$

It follows that Ψ is contained in S_{3f-3}(zero) if and only if $\mathbf{T}(0,\mathbf{R})\Psi$ is contained in S_{3f-3}(zero) for all **R** of the point group. [For, using Eq. (46) and the fact that $\{\mathbf{T}(0,\mathbf{R})\}$ is a group, we see that $\mathbf{P}(S_{3f-3}(\text{zero}))\Psi=\Psi$ implies $\mathbf{P}(S_{3f-3}(\text{zero}))\mathbf{T}(0,\mathbf{R})\Psi=\mathbf{T}(0,\mathbf{R})\Psi$. Conversely, multiplying the latter equation by $\mathbf{T}(0,\mathbf{R}^{-1})=\mathbf{T}^{-1}(0,\mathbf{R})$, we regain the former equation.] Further, suppose we are given a set of vectors which form a basis for an irreducible representation of the point group G and then form any nonvanishing linear combination of these vectors. Successive applications of the $\mathbf{T}(0,\mathbf{R})$ to this linear combination must result in a set of vectors spanning the subspace spanned by the original basis vectors. Consequently, we obtain the following lemma:

Lemma IX: If, under the transformations of $\{\mathbf{T}(0,\mathbf{R})\}$, a set of eigenvectors of **A** form a basis for an irreducible representation of the point group G, then the subspace spanned by these eigenvectors is either contained in S_{3f-3}(zero) or has no intersection with S_{3f-3}(zero).

The matrix **A** is real. Consequently the operation of time reversal is simply complex conjugation. From Eq. (40), we see that $\mathbf{P}(S_{3f-3}(\text{zero}))$ is a real operator. Consequently, the operations of time reversal and projection to S_{3f-3}(zero) commute, and by reasoning similar to that above we obtain the following lemma:

Lemma X: Consider a subspace of eigenvectors of **A** (corresponding to the same eigenvalue) which is invariant under time reversal. The subspace is either contained in S_{3f-3}(zero) or has no intersection with S_{3f-3}(zero).

Combining Lemmas IX and X, and recalling that the consideration of accidental degeneracies is excluded from our discussion, we finally arrive at the following:

Lemma XI: Any set of eigenvectors of **A** corresponding to the same eigenvalue occupy a subspace of S_{3f}(total) which is either contained in S_{3f-3}(zero) or has zero intersection with S_{3f-3}(zero).

Consider the matrix $\mathbf{C}^0(\hat{\boldsymbol{\phi}})$ for any given $\hat{\boldsymbol{\phi}}$. Using arguments similar to those above, one can easily show that eigenvectors of $\mathbf{C}^0(\hat{\boldsymbol{\phi}})$ which transform among themselves under the $\{\mathbf{T}(0,\mathbf{R})\}$ with $\mathbf{R}\in G(\hat{\boldsymbol{\phi}})$ or under the

operation of time reversal must either occupy subspaces which are contained in S_{3f-3}(zero) or have no intersection with S_{3f-3}(zero). Subspaces of eigenvectors of $\mathbf{C}^0(\hat{\phi})$ corresponding to the same eigenvalue can also be determined using Lemmas I–III. However, such subspaces are always contained within invariant subspaces of eigenvectors of **A**. Thus, we obtain a Lemma for $\mathbf{C}^0(\hat{\phi})$ which corresponds to Lemma XI for **A**:

Lemma XII: Any set of eigenvectors of $\mathbf{C}^0(\hat{\phi})$ (for any given $\hat{\phi}$) corresponding to the same eigenvalue occupy a subspace of S_{3f}(total) which is either contained in S_{3f-3}(zero) or has no intersection with S_{3f-3}(zero).[19]

Equations analogous to Eq. (46) and theorems analogous to Lemmas IX–XII also hold for the subspaces S_3(acoustic) and S_3(normal).

Choose an orthonormal set of eigenvectors of **A** which span S_{3f}(total). From Lemma XI, we see that a unique number of these eigenvectors will be contained in S_{3f-3}(zero) independent of our particular choice. The space S_{3f}(total) is then the sum of two mutually orthogonal subspaces denoted by S(nonpolar) and S(polar). The subspace S(nonpolar) is that subspace spanned by eigenvectors of **A** lying in S_{3f-3}(zero) and the subspace S(polar) is the space spanned by the eigenvectors of **A** lying outside of S_{3f-3}(zero).

Suppose that S(polar) is q-dimensional and S(nonpolar) is $(3f-q)$-dimensional. By Lemma I, the $\Psi^0(\hat{\phi})$ for $(3f-q)$ branches of the dispersion relations lie in S(nonpolar) for all $\hat{\phi}$. Consider the $\Psi^0(\hat{\phi})$ for any one of the remaining branches. For any $\hat{\phi}$, $\Psi^0(\hat{\phi})$ for the branch is orthogonal to S(nonpolar) and, thus, lies in S(polar). Further for no $\hat{\phi}=\hat{\phi}_0$ does $\Psi^0(\hat{\phi})$ for the branch lie in S_{3f-3}(zero). For then, $\Psi^0(\hat{\phi}_0)$ would be an eigenvector of **A** lying in S_{3f-3}(zero) and would thus lie in S(nonpolar), leading to a contradiction. Therefore, we obtain the following lemma:

Lemma XIII: The branches of the dispersion relations fall into two sets: those branches whose $\Psi^0(\hat{\phi})$ lie in S(polar) for all $\hat{\phi}$ and those branches whose $\Psi^0(\hat{\phi})$ lie in S(nonpolar) for all $\hat{\phi}$. Further, if the $\Psi^0(\hat{\phi})$ for a branch lie in S(polar), then for no $\hat{\phi}$ does a $\Psi^0(\hat{\phi})$ of that branch lie in S_{3f-3}(zero).

Consider a branch of the dispersion relations whose long-wavelength eigenvectors are contained in S(polar). Also assume that the eigenvectors $\Psi^0(\hat{\phi})$ for the branch are contained in $S_{3f-2}(\lambda_M=0;\hat{\phi})$. We now show that the $\Psi^0(\hat{\phi})$ span a three-dimensional subspace of S_{3f}(total). The condition that the $\Psi^0(\hat{\phi})$ lie in $S_{3f-2}(\lambda_M=0;\hat{\phi})$ for all $\hat{\phi}$ is given by the equation

$$\mathbf{M}(\hat{\phi})\Psi^0(\hat{\phi})=0, \tag{47}$$

for all $\hat{\phi}$. Using Eq. (17) and the fact that not all Z_ν vanish for ionic lattices, we see that Eq. (47) is equivalent to

$$\mathbf{T}(\hat{\phi})\sum_\kappa Z_\kappa\mu_\kappa^{-1/2}\Psi_\kappa^0(\hat{\phi})=0. \tag{48}$$

The sum in Eq. (48) cannot vanish, for, referring to Eq. (40), we see that then $\Psi^0(\hat{\phi})$ would lie in S_{3f-3}(zero), contrary to assumption. It then follows from Eqs. (22) and (23) that $\Psi^0(\hat{\phi})$ obeys Eq. (47) if and only if

$$\sum_\kappa Z_\kappa\mu_\kappa^{-1/2}\Psi_\kappa^0(\hat{\phi})=k(\hat{\phi})\hat{\phi}, \tag{49}$$

for all $\hat{\phi}$, where $k(\hat{\phi})$ is a number not equal to zero. Let $\hat{\phi}^{(1)}$, $\hat{\phi}^{(2)}$, and $\hat{\phi}^{(3)}$ be three mutually orthogonal vectors in ϕ-space. From Eq. (49) we obtain the relation

$$\sum_{\kappa,\nu}\frac{Z_\kappa Z_\nu}{(\mu_\kappa\mu_\nu)^{1/2}}\Psi_\kappa^{0\dagger}(\hat{\phi}^{(i)})\Psi_\nu^0(\hat{\phi}^{(j)}) = k^*(\hat{\phi}^{(i)})k(\hat{\phi}^{(j)})\delta_{ij}, \tag{50}$$

where $k^*(\hat{\phi}^{(i)})k(\hat{\phi}^{(j)})\neq 0$. Equation (50) not only indicates that the three-dimensional vectors $\hat{\phi}^{(i)}$ are mutually orthogonal in the usual sense, but it also represents a mutual orthogonality relation among the $3f$-dimensional vectors $\Psi^0(\hat{\phi}^{(i)})$ under the inner product

$$\{\Omega,\Phi\}=\sum_{\kappa,\nu}\frac{Z_\kappa Z_\nu}{(\mu_\kappa\mu_\nu)^{1/2}}\Omega_\kappa^\dagger\Phi_\nu. \tag{51}$$

If follows that $\Psi^0(\hat{\phi}^{(1)})$, $\Psi^0(\hat{\phi}^{(2)})$, and $\Psi^0(\hat{\phi}^{(3)})$ are linearly independent. Therefore, the $\Psi^0(\hat{\phi})$ obeying Eq. (47) must span a space of at least three dimensions. To show that the eigenvectors span a space of no more than three dimensions, note that the $\Psi^0(\hat{\phi})$ for the branch under consideration are all eigenvectors of $\mathbf{A}'$ corresponding to the same eigenvalue $\lambda'=\lambda^0$. [See Eq. (34).] None of the irreducible representations for the point groups G exceed three dimensions. Referring to Sec. IV and remembering that accidental degeneracies are excluded from our consideration, we see that the $\Psi^0(\hat{\phi})$ cannot span a subspace of more that three dimensions. Therefore, we obtain the following lemma:

Lemma XIV: Consider a branch of the dispersion relations all of whose long-wavelength eigenvectors $\Psi^0(\hat{\phi})$ lie in S(polar). If the $\Psi^0(\hat{\phi})$ are contained in $S_{3f-2}(\lambda_M=0;\hat{\phi})$ for each $\hat{\phi}$, then the $\Psi^0(\hat{\phi})$ span a subspace of S_{3f}(total) which is exactly three-dimensional.

By a proof similar to that of Lemma XIV (see Appendix A), we obtain the following lemma:

Lemma XV: Consider a branch of the dispersion relations all of whose long-wavelength eigenvectors $\Psi^0(\hat{\phi})$ lie in S(polar). If the $\Psi^0(\hat{\phi})$ are contained in $S_{3f-1}(\lambda_N=0;\hat{\phi})$ for each $\hat{\phi}$, then the $\Psi^0(\hat{\phi})$ span a subspace of S_{3f}(total) which is either two-dimensional or three-dimensional.

We conclude this section by a further refinement of the statement of Lemma IV. Consider a branch of the dispersion relations which approaches a definite frequency as ϕ approaches zero and whose normal mode eigenvectors $\Psi^0(\hat{\phi})$ lie in S(polar). As stated, Lemma IV allows for the possibility that the $\Psi^0(\hat{\phi})$ might lie in $S_{3f-1}(\lambda_N=0;\hat{\phi})$ for some (but not all) values of $\hat{\phi}$ and in $S_{3f-2}(\lambda_M=0;\hat{\phi})$ for the remaining values of $\hat{\phi}$. A second possibility which is not discounted is that the $\Psi^0(\hat{\phi})$ for a branch in S(polar) lie in both $S_{3f-1}(\lambda_N=0;\hat{\phi})$ and $S_{3f-2}(\lambda_M=0;\hat{\phi})$ for some value of $\hat{\phi}$. We now show that neither of these possibilities can occur. The second possibility is immediately discounted by Lemma XIII since the intersection of $S_{3f-1}(\lambda_N=0;\hat{\phi})$ and $S_{3f-2}(\lambda_M=0;\hat{\phi})$ is S_{3f-3}(zero). Our proof that the first possibility cannot occur is somewhat lengthy and is given in Appendix B. We conclude the following:

Lemma XVI: Consider a branch of the dispersion relations all of whose eigenvectors $\Psi^0(\hat{\phi})$ lie in S(polar). A necessary and sufficient condition that the branch ap-

proach a definite frequency in the long-wavelength limit is that either the $\Psi^0(\hat{\phi})$ lie in $S_{3f-1}(\lambda_N = 0; \hat{\phi})$ for all $\hat{\phi}$ or that the $\Psi^0(\hat{\phi})$ lie in $S_{3f-2}(\lambda_M = 0; \hat{\phi})$ for all $\hat{\phi}$. Further, for no $\hat{\phi}$ does $\Psi^0(\hat{\phi})$ lie in both $S_{3f-1}(\lambda_N = 0; \hat{\phi})$ and $S_{3f-2}(\lambda_M = 0; \hat{\phi})$.

In physical terms the above lemma can be stated as follows. *Suppose a branch of the dispersion relations has long-wavelength vibrations which produce a nonzero polarzation. The branch will approach a definite frequency as $\hat{\phi}$ approaches zero if and only if either all of its long-wavelength vibrations produce no macroscopic electric field or alternatively all of its long-wavelength vibrations produce no macroscopic electric displacement field.*

VII. ORTHOGONALITY RELATIONS AND RELATED LEMMAS

Let $\mathbf{D}^{(\mu)}$ and $\mathbf{D}^{(\nu)}$ be (unitary) irreducible representations of the point group G for the lattice. One of the basic orthogonality relations of the theory of group representations[20] gives us the equation

$$\sum_{\mathbf{R}\in G} D_{ij}^{(\mu)*}(\mathbf{R}) D_{lm}^{(\nu)}(\mathbf{R}) = (g/n_\mu)\delta_{\mu\nu}\delta_{il}\delta_{jm}, \tag{52}$$

where g is the order of G and n_μ is the dimension of $D^{(\mu)}$.

First consider any point group G belonging to the regular system. Referring to any character table for the 32 crystal point groups, we see that all of the point groups belonging to the regular system have as one of their irreducible representations a representation consisting of 3×3, orthogonal rotation matrices R. We shall use the symbol Γ to refer to this representation. Using Eq. (52), we obtain the orthogonality relation

$$\sum_{\mathbf{R}} D_{ij}^{(\mu)*}(\mathbf{R}) R_{lm} = \frac{g}{3}\delta_{\mu\Gamma}\,\delta_{il}\delta_{jm}, \tag{53}$$

where μ is any irreducible representation of the point group G belonging to the regular system. From Eq. (53), we immediately obtain the relation

$$\sum_{\mathbf{R}} \chi^{(\mu)*}(\mathbf{R}) R_{lm} = \sum_{\mathbf{R}} \chi^{(\mu)}(\mathbf{R}) R_{lm} = \tfrac{1}{3} g \delta_{\mu\Gamma}\delta_{lm}, \tag{54}$$

or equivalently

$$\sum_{\mathbf{R}} \chi^{(\mu)*}(\mathbf{R})\mathbf{R} = \sum_{\mathbf{R}} \chi^{(\mu)}(\mathbf{R})\mathbf{R} = \tfrac{1}{3} g \delta_{\mu\Gamma} I, \tag{55}$$

where $\chi^{(\mu)}(\mathbf{R})$ is the character of $\mathbf{D}^{(\mu)}(\mathbf{R})$. [Since R_{lm} and the left-hand side of Eq. (53) are real, Eq. (53) remains true if the complex conjugate sign * is removed from the $D_{ij}^{(\mu)*}(\mathbf{R})$.]

Next consider crystal point groups which do not belong to the regular system. Choose a system of Cartesian coordinates such that the z axis lies along the principal axis of rotation for the point group. Referring to any set of character tables for the crystal point groups, we see that the function z belongs to a one-dimensional, real, irreducible representation of each point group not belonging to the regular system. We now divide such point groups into three classes (referred to as Types I, II, and III) in accordance with the transformation properties of the functions x and y.

Type I crystal point groups are those for which x and y belong to one-dimensional, real, irreducible representations. Point groups belonging to Type I are C_1, C_i, C_2, C_s, C_{2h}, C_{2v}, D_2, and D_{2h}. All point groups of this type have only one-dimensional, real irreducible representations.

Type II crystal point groups are those for which x and y belong to the same two-dimensional, real irreducible representation. Point groups of this type are C_{3v}, D_3, C_{4v}, D_4, D_{2d}, D_6, C_{6v}, D_{3h}, D_{3d}, D_{4h}, and D_{6h}.

Type III crystal point groups are those for which $x + iy$ and $x - iy$ belong to different one-dimensional, complex, irreducible representations which are complex conjugates. Because of time reversal invariance, eigenvectors of A belonging to either of these representations correspond to twofold degenerate eigenvalues. The matrix A is real. Thus, if Ψ is an eigenvector of A transforming according to one of these irreducible representations, then Ψ^* is an eigenvector of A corresponding to the same eigenvalue and belonging to the other irreducible representation. Point groups belonging to Type III are C_4, S_4, C_3, C_6, C_{3h}, C_{4h}, C_{6h}, and S_6.

Each of the crystal point groups not belonging to the regular system has a reducible representation $\mathbf{D}^{(\Gamma)}$ consisting of 3×3, real, orthogonal rotation matrices R. For Type I point groups, $\mathbf{D}^{(\Gamma)}$ is automatically in the reduced form

$$\mathbf{D}^{(\Gamma)} = \mathbf{D}^{(x)} \oplus \mathbf{D}^{(y)} \oplus \mathbf{D}^{(z)}, \tag{56}$$

where $\mathbf{D}^{(x)}$, $\mathbf{D}^{(y)}$, and $\mathbf{D}^{(z)}$ are the irreducible representations to which x, y, and z belong respectively. ($\mathbf{D}^{(x)}$, $\mathbf{D}^{(y)}$, and $\mathbf{D}^{(z)}$ are not always distinct.) Using Eq. (52) and the fact that all $\mathbf{R}\in\mathbf{D}^{(\Gamma)}$ are real, we obtain the following relation for Type I point groups:

$$\sum_{\mathbf{R}} \chi^{(\mu)*}(\mathbf{R})\mathbf{R} = \sum_{\mathbf{R}} \chi^{(\mu)}(\mathbf{R})\mathbf{R} = g\begin{bmatrix} \delta_{\mu x} & 0 & 0 \\ 0 & \delta_{\mu y} & 0 \\ 0 & 0 & \delta_{\mu z} \end{bmatrix}, \tag{57}$$

where μ is any irreducible representation of the point group.

For Type II point groups, $\mathbf{D}^{(\Gamma)}$ is automatically in the reduced form,

$$\mathbf{D}^{(\Gamma)} = \mathbf{D}^{(xy)} \oplus \mathbf{D}^{(z)}, \tag{58}$$

where $\mathbf{D}^{(xy)}$ is the two-dimensional, irreducible representation to which x and y belong and $\mathbf{D}^{(z)}$ is the one-dimensional irreducible representation to which z belongs. Again using Eq. (52) and the reality of the R matrices, we obtain the following relation for the Type II point groups:

$$\sum_{\mathbf{R}} D_{ij}^{(\mu)*}(\mathbf{R})\mathbf{R} = \sum_{\mathbf{R}} D_{ij}^{(\mu)}(\mathbf{R})\mathbf{R} = \frac{g}{2}\begin{bmatrix} & & 0 \\ \mathbf{H}\delta_{\mu(xy)} & & 0 \\ 0 & 0 & 2\delta_{\mu z} \end{bmatrix}, \tag{59}$$

where

$$\mathbf{H} = \begin{bmatrix} \delta_{i1}\delta_{j1} & \delta_{i1}\delta_{j2} \\ \delta_{i2}\delta_{j1} & \delta_{i2}\delta_{j2} \end{bmatrix}, \tag{60}$$

and μ is any irreducible representation of the point group. We also obtain the relation

$$\sum_{\mathbf{R}}\chi^{(\mu)*}(\mathbf{R})\mathbf{R}=\sum_{\mathbf{R}}\chi^{(\mu)}(\mathbf{R})\mathbf{R}$$
$$=g\begin{bmatrix}\frac{1}{2}\delta_{\mu(xy)} & 0 & 0\\ 0 & \frac{1}{2}\delta_{\mu(xy)} & 0\\ 0 & 0 & \delta_{\mu z}\end{bmatrix}, \tag{61}$$

where μ is any irreducible representation of the point group.

In the case of Type III point groups, $\mathbf{D}^{(\Gamma)}$ is not automatically in reduced form. It is reduced by carrying out a similarity transformation with the unitary matrix U defined by

$$\mathbf{U}=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & i & 0\\ 1 & -i & 0\\ 0 & 0 & \sqrt{2}\end{bmatrix}. \tag{62}$$

Then $\mathbf{D}^{(\Gamma)}$ reduces according to

$$\mathbf{D}^{(\Gamma)}=\mathbf{D}^{(+)}\oplus\mathbf{D}^{(-)}\oplus\mathbf{D}^{(z)}, \tag{63}$$

where +, −, and z are the one-dimensional irreducible representations to which $x+iy$, $x-iy$, and z belong, respectively. ($\mathbf{D}^{(+)}$, $\mathbf{D}^{(-)}$, and $\mathbf{D}^{(z)}$ are distinct.) Using Eq. (52) and the identity $\mathbf{U}^{\dagger}\mathbf{U}\mathbf{R}\mathbf{U}^{\dagger}\mathbf{U}=\mathbf{R}$, we obtain the following equation for Type III point groups:

$$\sum_{\mathbf{R}}\chi^{(\mu)*}(\mathbf{R})\mathbf{R}$$
$$=g\begin{bmatrix}\frac{1}{2}(\delta_{\mu+}+\delta_{\mu-}) & \frac{1}{2}i(\delta_{\mu+}-\delta_{\mu-}) & 0\\ -\frac{1}{2}i(\delta_{\mu+}-\delta_{\mu-}) & \frac{1}{2}(\delta_{\mu+}+\delta_{\mu-}) & 0\\ 0 & 0 & \delta_{\mu z}\end{bmatrix}. \tag{64}$$

Let $\mathbf{P}^{\mu}$ be the projection operator to the subspace of eigenvectors belonging to the μth irreducible representation of the point group G for the lattice. Using Eq. (40), we see that

$$\mathbf{P}^{\mu}\mathbf{P}(S_{3f-3}(\text{zero}))=\mathbf{P}^{\mu}-\mathbf{P}^{\mu}\mathbf{P}(S_3(\text{normal})). \tag{65}$$

With the aid of Eqs. (14), (15), (39), and (43), one finds

$$[\mathbf{P}^{\mu}\mathbf{P}(S_3)\text{normal}))]_{\kappa\nu}=\frac{n_{\mu}}{g}(\textstyle\sum_{\omega}Z_{\omega}^2/\mu_{\omega})^{-1}Z_{\kappa}Z_{\nu}(\mu_{\kappa}\mu_{\nu})^{-1/2}$$
$$\times\sum_{\mathbf{R}}\chi^{\mu}*(\mathbf{R})\mathbf{R}. \tag{66}$$

First consider the case where none of the functions x, y, or z belong to $D^{(\mu)}$. Using Eqs. (55), (57), (61), and (64), we see that for any crystal point group $\sum_{\mathbf{R}}\chi^{(\mu)*}(\mathbf{R})\mathbf{R}=0$.

Thus

$$\mathbf{P}^{\mu}\mathbf{P}(S_{3f-3}(\text{zero}))=\mathbf{P}^{\mu}. \tag{67}$$

On the other hand, if $\mathbf{D}^{(\mu)}$ is one of the irreducible representations to which x, y, or z does belong, then

$$\mathbf{P}^{\mu}\mathbf{P}(S_3(\text{normal}))\neq 0, \tag{68}$$

for ionic lattices. The implications of Eqs. (67) and (68) are summarized in the following lemma:

Lemma XVII: Let Ψ be an eigenvector of A which belongs to the irreducible representation $\mathbf{D}^{(\mu)}$ of the point group for the lattice. If none of x, y, nor z belong to $\mathbf{D}^{(\mu)}$, then Ψ lies in S_{3f-3}(zero). The subspace S(polar) is spanned by eigenvectors of A belonging to the irreducible representations to which x, y, and z belong. Further, given any $\mathbf{D}^{(\mu)}$ to which x, y, or z belongs, at least one eigenvector of A belonging to $\mathbf{D}^{(\mu)}$ will be included among those spanning S(polar).

Using a proof similar to the above, we obtain an analogous lemma pertaining to S_3(acoustic):

Lemma XVIII: The subspace S_3(acoustic) is spanned by eigenvectors of A [and $\mathbf{C}^0(\hat{\phi})$] belonging to irreducible representations of the point group to which x, y, or z belong. Further, given any irreducible representation to which x, y, or z belongs, S_3(acoustic) contains a set of eigenvectors of A transforming among themselves according to that irreducible representation.

We conclude this section by introducing a few items of mathematical notation to be used in the following sections. The various particle types κ can be divided into equivalence classes such that κ and ν are in the same equivalence class if and only if $\nu=F_0(\kappa,\mathbf{R})$ for some R in the point group G for the lattice, where $F_0(\nu,\mathbf{R})$ is defined in Eq. (13). We indicate that κ and ν are in the same equivalence class by writing $\kappa\approx\nu$. According to Eqs. (14) and (15), $\kappa\approx\nu$ implies that $Z_{\kappa}=Z_{\nu}$ and that $\mu_{\kappa}=\mu_{\nu}$. Each equivalence class will be denoted by a typical element of the class τ.

In what follows we shall often sum a quantity $f(\kappa)$ over all κ in the equivalence class characterized by τ. We indicate such a sum by writing

$$\sum_{\kappa\approx\tau}f(\kappa). \tag{69}$$

A useful property of this sum is that

$$\sum_{\kappa\approx\tau}f(F_0(\kappa,\mathbf{R}))=\sum_{\kappa\approx\tau}f(F_0^{-1}(\kappa,\mathbf{R}))=\sum_{\kappa\approx\tau}f(\kappa), \tag{70}$$

where R is any element of the point group G.

Another type of sum which will be employed is the following. Suppose that $g(\kappa)$ is a quantity which is the same for all κ's within any given equivalence class. A sum of $g(\kappa)$ over all equivalence classes [with just one $g(\kappa)$ for each class] is indicated by

$$\sum_{\tau}{}'g(\tau). \tag{71}$$

It follows that

$$\sum_{\kappa}f(\kappa)=\sum_{\tau}{}'\sum_{\kappa\approx\tau}f(\kappa). \tag{72}$$

VIII. LATTICES WHOSE POINT GROUPS BELONG TO THE REGULAR SYSTEM

We now use the formulas and lemmas developed in the preceding sections to prove the following theorem:

Theorem I: Consider an ionic lattice in the point-ion approximation. If the point group of the space group for the lattice belongs to the regular system, then all branches of the phonon dispersion relations for the lattice approach definite frequencies (independent of $\hat{\phi}$) in the long-wavelength limit.

In the remainder of this section, we will prove the above theorem. Assume that A has m independent eigenvectors which belong to irreducible representations of the point group to which x, y, and z do not belong. By Lemmas I and XVII it follows that m corresponding branches of the dispersion relations approach definite frequencies.

Thus, we need consider further only those eigenvectors of A belonging to the irreducible representation Γ. Consider a three-dimensional subspace $S_3(\Gamma)$ of S_{3f}(total) spanned by a set of eigenvectors of A which transform among themselves according to Γ. The intersection $S_3(\Gamma) \cap S_{3f-1}(\lambda_N = 0; \hat{\phi})$ is at least two-dimensional and $S_3(\Gamma) \cap S_{3f-2}(\lambda_M = 0; \hat{\phi})$ is at least one-dimensional for each $\hat{\phi}$. (See Footnote 14.) If $S_3(\Gamma) \cap S_{3f-1}(\lambda_N = 0; \hat{\phi})$ is three-dimensional, then $S_3(\Gamma)$ has a nonzero-dimensional intersection with $S_{3f-1}(\lambda_N = 0; \hat{\phi}) \cap S_{3f-2}(\lambda_M = 0; \hat{\phi}) = S_{3f-3}$(zero). It follows from Lemma IX that $S_3(\Gamma) \subset S_{3f-3}$(zero) and then from Lemma I that three corresponding branches of the dispersions approach a definite frequency. (From Lemmas VIII and XVIII, we see that the acoustic modes provide three such branches.)

Finally we must consider those eigenvectors of $\mathbf{C}^0(\hat{\phi})$ which lie in S(polar). From the preceding paragraph, we see that these eigenvectors of $\mathbf{C}^0(\hat{\phi})$ lie in a subspace of S_{3f}(total) spanned by sets of eigenvectors of A belonging to Γ for which $S_3(\Gamma) \cap S_{3f-1}(\lambda_N = 0; \hat{\phi})$ is exactly two-dimensional and $S_3(\Gamma) \cap S_{3f-2}(\lambda_M = 0; \hat{\phi})$ is exactly one-dimensional. Assume there are n such sets with corresponding three-dimensional subspaces $S_3^\mu(\Gamma)$, where $\mu = 1, 2, \ldots, n$. In the following paragraphs we will show that any vector in $S_3^\mu(\Gamma) \cap S_{3f-1}(\lambda_N = 0; \hat{\phi})$ is orthogonal to any vector in the one-dimensional subspace $S_3^\mu(\Gamma) \cap S_{3f-2}(\lambda_M = 0; \hat{\phi})$. Once the preceding statement is established, the remainder of the proof of Theorem I is as follows. Since we assume that no accidental degeneracies are present, the subspaces $S_3^\mu(\Gamma)$ and $S_3^\nu(\Gamma)$ must be orthogonal for $\mu \neq \nu$. Within each $S_3^\mu(\Gamma)$, the eigenvectors of A lying in $S_3^\mu(\Gamma) \cap S_{3f-1}(\lambda_N = 0; \hat{\phi})$ are also eigenvectors of $\mathbf{C}^0(\hat{\phi})$. Lemma II is applicable and $2n$ additional branches of the dispersion relations [two from each $S_3^\mu(\Gamma)$] approach definite frequencies in the long-wavelength limit. The only eigenvectors of $\mathbf{C}^0(\hat{\phi})$ remaining for consideration lie in an n-dimensional subspace of S_{3f}(total) which for each $\hat{\phi}$ is orthogonal to all of the eigenvectors of $\mathbf{C}^0(\hat{\phi})$ so far considered. Clearly this is the n-dimensional subspace spanned by one eigenvector of A from each of the one-dimensional subspaces $S_3^\mu(\Gamma) \cap S_{3f-2}(\lambda_M = 0; \hat{\phi})$. But any vector in this n-dimensional subspace is, for each $\hat{\phi}$, contained in $S_{3f-2}(\lambda_M = 0; \hat{\phi})$. Thus, by either Lemma IV or XVI the remaining n branches of the dispersion relations also approach definite frequencies in the long-wavelength limit. Referring to Eq. (34), we see that the eigenvectors for these n branches are also eigenvectors of $\mathbf{A}'$.

As a result of the above argument, Theorem I will be established once we show that, for each of the subspaces $S_3^\mu(\Gamma)$, any vector in $S_3^\mu(\Gamma) \cap S_{3f-1}(\lambda_N = 0; \hat{\phi})$ is orthogonal to $S_3^\mu(\Gamma) \cap S_{3f-2}(\lambda_M = 0; \hat{\phi})$. The subspace $S_3^\mu(\Gamma)$ is spanned by three vectors $\Psi^{\Gamma\mu i}$ (with $i = 1$, 2, or 3) which transform according to the first, second, or third rows of Γ, respectively. These vectors are mutually orthogonal since they transform according to different rows of the same irreducible representation of the point group. From Eq. (44) and the fact that $\mathbf{C}_3$ is an element of every point group belonging to the regular system, we find that

$$\Psi^{\Gamma\mu 1} = \mathbf{T}(0, \mathbf{C}_3)\Psi^{\Gamma\mu 3}, \quad \text{and} \quad \Psi^{\Gamma\mu 2} = \mathbf{T}(0, \mathbf{C}_3^2)\Psi^{\Gamma\mu 3}. \tag{73}$$

With the aid of Eqs. (13), (42), (53), (69), (70), and (73), we obtain the result[21]

$$\sum_{\nu \approx \tau} \Psi_\nu^{\Gamma\mu i} = c_\tau^\Gamma(\mu) \begin{bmatrix} \delta_{i1} \\ \delta_{i2} \\ \delta_{i3} \end{bmatrix}, \tag{74}$$

where

$$c_\tau^\Gamma(\mu) = \sum_{\nu \approx \tau} \Psi_{\nu 3}^{\Gamma\mu 3}. \tag{75}$$

A general vector Ψ in $S_3^\mu(\Gamma)$ has the form

$$\Psi = \sum_i f_i \Psi^{\Gamma\mu i}. \tag{76}$$

In order to determine the general form of a vector in $S_3^\mu(\Gamma) \cap S_{3f-1}(\lambda_N = 0; \hat{\phi})$, we impose the condition $\mathbf{N}(\hat{\phi})\Psi = 0$. With the aid of Eqs. (5) and (74) and the fact that not all Z_ν vanish for ionic lattices, we find that our condition is equivalent to

$$\left[\sum_\tau{}' Z_\tau \mu_\tau^{-1/2} c_\tau^\Gamma(\mu)\right] \mathbf{L}(\hat{\phi}) \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = 0, \tag{77}$$

where the sum is defined by Eq. (71). Using Eas. (39), (40), (74), and (75), we see that the vanishing of the sum in Eq. (77) implies that the $\Psi^{\Gamma\mu i}$ are contained in S_{3f-3}(zero). But this possibility is contrary to our assumption. Thus, using Eq. (6) and the fact that not all components of ϕ vanish, we see that Eq. (77) is equivalent to the equation

$$\sum_i f_i \phi_i = 0. \tag{78}$$

That is, the f_i form the components of a three-dimensional vector transverse to $\hat{\phi}$. It follows that two mutually orthogonal vectors spanning $S_3^\mu(\Gamma) \cap S_{3f-1}(\lambda_N = 0; \hat{\phi})$ are given by

$$\Phi^{\mu i}(\hat{\phi}) = \sum_j \eta_j^i(\hat{\phi}) \Psi^{\Gamma\mu j}, \tag{79}$$

where $i = 1$ or 2 and the $\eta^i(\hat{\phi})$ are two mutually orthogonal vectors both orthogonal to $\hat{\phi}$.

A vector in $S_3^\mu(\Gamma) \cap S_{3f-2}(\lambda_M = 0; \hat{\phi})$ is obtained by setting $\mathbf{M}(\hat{\phi})\Psi = 0$ in Eq. (76). Using arguments similar to those in the preceding paragraph, we find that now the f_i must form the components of a vector parallel to $\hat{\phi}$. Thus, a general vector in $S_3^\mu(\Gamma) \cap S_{3f-2}(\lambda_M = 0; \hat{\phi})$ is of the form

$$\Omega^\mu(\hat{\phi}) = k \sum_i \phi_i \Psi^{\Gamma\mu i}, \tag{80}$$

where k is some number not equal to zero. For convenience assume that the $\Psi^{\Gamma\mu i}$ are normalized. The scalar product of $\Omega^\mu(\hat{\phi})$ with either $\Phi^{\mu i}(\hat{\phi})$ is easily seen to be

$$\Phi^{\mu i\dagger}(\hat{\phi})\Omega^{\mu}(\hat{\phi}) = k\hat{\eta}^{i}\cdot\hat{\phi} = 0.$$

Thus, we have proved that any vector in $S_3^{\mu}(\Gamma) \cap S_{3f-1}(\lambda_N = 0; \hat{\phi})$ is orthogonal to $S_3^{\mu}(\Gamma) \cap S_{3f-2}(\lambda_M = 0; \hat{\phi})$, and Theorem I is established.

As was stated in Sec. V, the general development in this paper is carried out under the assumption that accidental degeneracies (as defined in Sec. V) are not present. However, it is clear that Theorem I remains true even if such accidental degeneracies occur. All of the branches considered above approach definite frequencies [that is, the eigenvalues of $\mathbf{C}^0(\hat{\phi})$ to which the $\Psi^0(\hat{\phi})$ correspond are independent of $\hat{\phi}$]. If additional degeneracies appear between certain branches, we must replace these branches by new branches whose long-wavelength eigenvectors are linear combinations of those of the original branches. However, the corresponding eigenvalue of $\mathbf{C}^0(\hat{\phi})$ remains $\hat{\phi}$-independent and all of the new branches approach the same definite, $\hat{\phi}$-independent frequency.

We summarize the work in this section as follows. For the lattices considered in this section, all of the values of λ^0 approached by the various branches of the dispersion relations will be either eigenvalues of **A** or of **A′**. If $3f-q$ branches have eigenvectors lying in S(nonpolar), then the corresponding eigenvalues λ^0 of $\mathbf{C}^0(\hat{\phi})$ will be common eigenvalues of **A** and **A′**. The corresponding long-wavelength lattice vibrations will produce $\mathbf{E} = \mathbf{D} = \mathbf{P} = 0$. The remaining branches will have long-wavelength eigenvectors lying in S(polar). Of these, $\frac{2}{3}q$ branches will approach values of λ^0 which are eigenvalues of **A** for eigenvectors of **A** lying in S(polar). The corresponding long-wavelength lattice vibrations produce $\mathbf{E} = 0$. The remaining $\frac{1}{3}q$ branches approach eigenvalues of **A′** for eigenvectors of **A′** lying in S(polar). The corresponding long-wavelength lattice vibrations produce $\mathbf{D} = 0$.

IX. ALTERNATIVE PROOF OF THEOREM I. THE MATRIX $Q(R)$

In Ref. 1 and in Lemma V, it is stated that a necessary and sufficient condition that all branches of the dispersion relations for a lattice approach definite frequencies as ϕ approaches zero is the existence of a unitary matrix $\mathbf{Q}(\mathbf{R})$ satisfying Eq. (24), for all $\mathbf{R} \in O(3)$. Theorem I implies that such a matrix exists for all lattices whose point groups belong to the regular system. Once the matrix $\mathbf{Q}(\mathbf{R})$ is determined, the eigenvectors of $\mathbf{C}^0(\hat{\phi})$ can be determined for all $\hat{\phi}$ by determining the eigenvectors of $\mathbf{C}^0(\hat{\phi}_0)$ for one convenient $\hat{\phi}_0$ only. For if the $\Psi^{0n}(\hat{\phi}_0)$, with $n = 1, 2, \ldots, 3f$, are a set of mutually orthogonal eigenvectors of $\mathbf{C}^0(\hat{\phi}_0)$, then the $\mathbf{Q}^{\dagger}(\mathbf{R})\Psi^{0n}(\hat{\phi}_0)$ provide a mutually orthogonal set of eigenvectors of $\mathbf{C}^0(\mathbf{R}\hat{\phi}_0)$ for all $\mathbf{R} \in O(3)$.

In the remainder of this section, we show how the matrix $\mathbf{Q}(\mathbf{R})$ can be constructed once the eigenvectors of **A** are determined. Our construction of course constitutes an alternative proof of Theorem I. Denote the members of a complete, orthonormal set of eigenvectors of **A** by $\Psi^{\mu\sigma i}$. Here μ and i are integers which indicate that $\Psi^{\mu\sigma i}$ belongs to a set of eigenvectors of **A** which transform among themselves under the $\mathbf{T}(0, \mathbf{R})$ according to the μth irreducible representation of the point group and that $\Psi^{\mu\sigma i}$ belongs to the ith row of that irreducible representation. There may be several such sets of eigenvectors and σ distinguishes among them. We order the $\Psi^{\mu\sigma i}$ in lexigraphical order and assign the largest value of μ to the Γ representation. Assume that the Γ representation occurs n_Γ times in the reducible representation $\{\mathbf{T}(0, \mathbf{R})\}$.

Transform $\mathbf{C}^0(\hat{\phi}) = \mathbf{A} + (4\pi a^3/v_a)\mathbf{N}(\hat{\phi})$ to the $\Psi^{\mu\sigma i}$ basis. If we denote the transformed $\mathbf{C}^0(\hat{\phi})$ by $\overline{\mathbf{C}}^0(\hat{\phi})$, then $\overline{\mathbf{C}}^0(\hat{\phi}) = \mathbf{U}^{\dagger}\mathbf{C}^0(\hat{\phi})\mathbf{U} = \overline{\mathbf{A}} + (4\pi a^3/v_a)\overline{\mathbf{N}}(\hat{\phi})$, where **U** is the unitary matrix each of whose columns consists of the elements of one of the $\Psi^{\mu\sigma i}$ (with the $\mu\sigma i$ arranged in lexigraphical order beginning with the first column). First consider the form of $\overline{\mathbf{A}} = \mathbf{U}^{\dagger}\mathbf{A}\mathbf{U}$. It is, of course, diagonal. More explicitly, it has the following form:

$$\overline{\mathbf{A}} = \left[\begin{array}{cccc|ccc} \alpha_1\mathbf{I}_1 & & & & & & \\ & \alpha_2\mathbf{I}_2 & & & & 0 & \\ & & \ddots & & & & \\ & 0 & & \alpha_m\mathbf{I}_m & & & \\ \hline & & & & \beta_1\mathbf{I} & & 0 \\ & & & & & \beta_2\mathbf{I} & \\ & 0 & & & & & \ddots \\ & & & & 0 & & \beta_{n_\Gamma}\mathbf{I} \end{array}\right]. \tag{81}$$

In the above matrix $\alpha_1, \alpha_2, \ldots, \alpha_m$ and $\beta_1, \beta_2, \ldots, \beta_{n_\Gamma}$ are the eigenvalues of **A**. The quantities $\mathbf{I}_1, \mathbf{I}_2, \ldots, \mathbf{I}_m$ are unit submatrices which are either one-, two-, or three-dimensional. The unit matrices I occurring in the lower, right-hand, $3n_\Gamma \times 3n_\Gamma$ block are three-dimensional since Γ is a three-dimensional irreducible representation.

Next consider the form of $\overline{\mathbf{N}}(\hat{\phi}) = \mathbf{U}^{\dagger}\mathbf{N}(\hat{\phi})\mathbf{U}$. Its elements are given by

$$\overline{\mathbf{N}}_{\nu\sigma i,\,\omega\kappa j}(\hat{\phi}) = \Psi^{\nu\sigma i\dagger}\mathbf{N}(\hat{\phi})\Psi^{\omega\kappa j}. \tag{82}$$

It follows from Lemma XVII that these elements vanish unless both ν and ω are the irreducible representation Γ. Thus, the matrix $\overline{\mathbf{N}}(\hat{\phi})$ has nonzero elements only in the lower, right-hand $3n_\Gamma \times 3n_\Gamma$ block. We now calculate the values of the elements in this block. Using Eqs. (5) and (72), we obtain

$$\begin{aligned}\Psi^{\Gamma\sigma i\dagger}&\mathbf{N}(\hat{\phi})\Psi^{\Gamma\kappa j} \\ &= \sum_{\lambda,\nu} Z_\lambda Z_\nu (\mu_\lambda\mu_\nu)^{-1/2}\Psi_\lambda^{\Gamma\sigma i\dagger}\mathbf{L}(\hat{\phi})\Psi_\nu^{\Gamma\kappa j} \\ &= \sum_\tau{}' \sum_\omega{}' Z_\tau Z_\omega (\mu_\tau\mu_\omega)^{-1/2} \sum_{\lambda\approx\tau} \Psi_\lambda^{\Gamma\sigma i\dagger}\mathbf{L}(\hat{\phi}) \sum_{\nu\approx\omega} \Psi_\nu^{\Gamma\kappa j}.\end{aligned} \tag{83}$$

Then applying Eq. (74), we find that

$$\begin{aligned}\overline{N}&_{\Gamma\sigma i,\,\Gamma\kappa j}(\hat{\phi}) \\ &= \sum_\tau{}' \sum_\omega{}' Z_\tau Z_\omega (\mu_\tau\mu_\omega)^{-1/2} c_\tau^{\Gamma *}(\sigma) c_\omega^{\Gamma}(\kappa) L_{ij}(\hat{\phi}).\end{aligned} \tag{84}$$

Let **R** be any element of $O(3)$. Using Eq. (10), we find that

$$\begin{aligned}\overline{N}&_{\Gamma\sigma i,\,\Gamma\kappa j}(\mathbf{R}\hat{\phi}) \\ &= \sum_{l,m} R_{il}\left(\sum_\tau{}' \sum_\omega{}' Z_\tau Z_\omega (\mu_\tau\mu_\omega)^{-1/2} c_\tau^{\Gamma *}(\sigma) c_\omega^{\Gamma}(\kappa) L_{lm}(\hat{\phi})\right)(\mathbf{R}^{\dagger})_{mj}\end{aligned}$$

$$= \sum_{i,m} R_{il} \bar{N}_{\Gamma\sigma l, \Gamma\kappa m}(\hat{\phi})(\mathbf{R}^\dagger)_{mj}. \tag{85}$$

Construct the unitary matrix $\bar{\mathbf{Q}}(\mathbf{R})$ defined as follows:

$$\bar{\mathbf{Q}}(\mathbf{R}) = \begin{bmatrix} \mathbf{I} & & & 0 \\ \hline & \mathbf{R} & & \\ & & \mathbf{R} & 0 \\ & & & \ddots \\ 0 & & 0 & \ddots \\ & & & & \mathbf{R} \end{bmatrix}, \tag{86}$$

where I is the $(3f-3n_\Gamma)\times(3f-3n_\Gamma)$ unit matrix. It follows from Eq. (81) that

$$\bar{\mathbf{Q}}^\dagger(\mathbf{R})\bar{\mathbf{A}}\bar{\mathbf{Q}}(\mathbf{R}) = \bar{\mathbf{A}}. \tag{87}$$

Further, using Eq. (85) and the discussion following Eq. (82), we see that

$$\bar{\mathbf{Q}}(\mathbf{R})\bar{\mathbf{N}}(\hat{\phi})\bar{\mathbf{Q}}(\mathbf{R}) = \bar{\mathbf{N}}(\mathbf{R}\hat{\phi}). \tag{88}$$

Thus, for all $\mathbf{R} \in O(3)$,

$$\bar{\mathbf{C}}^0(\mathbf{R}\hat{\phi}) = \bar{\mathbf{Q}}^\dagger(\mathbf{R})\bar{\mathbf{C}}^0(\hat{\phi})\bar{\mathbf{Q}}(\mathbf{R}). \tag{89}$$

Finally, we obtain the equation

$$\mathbf{C}^0(\mathbf{R}\hat{\phi}) = \mathbf{Q}^\dagger(\mathbf{R})\mathbf{C}^0(\hat{\phi})\mathbf{Q}(\mathbf{R}), \tag{90}$$

where

$$\mathbf{Q}(\mathbf{R}) = \mathbf{U}\bar{\mathbf{Q}}(\mathbf{R})\mathbf{U}^\dagger. \tag{91}$$

We summarize these results. To construct the matrix $\mathbf{Q}(\mathbf{R})$ in Eq. (24), first construct the matrix $\mathbf{U}$ whose columns are the eigenvectors of $\mathbf{A}$ ordered as described earlier in this section. Then write down the matrix $\bar{\mathbf{Q}}(\mathbf{R})$ defined by Eq. (86). The unitary matrix $\mathbf{Q}(\mathbf{R})$ is then obtained from Eq. (91). If the eigenvectors $\Psi^{0n}(\hat{\phi}_0)$ of $\mathbf{C}^0(\hat{\phi}_0)$ are determined, the eigenvectors of $\mathbf{C}^0(\hat{\phi})$ for any $\hat{\phi} = \mathbf{R}\hat{\phi}_0$ may be obtained from the equation

$$\Psi^{0n}(\mathbf{R}\hat{\phi}_0) = \mathbf{Q}^\dagger(\mathbf{R})\Psi^{0n}(\hat{\phi}_0). \tag{92}$$

X. LATTICES WHOSE POINT GROUPS DO NOT BELONG TO THE REGULAR SYSTEM

Recall that in Sec. VII the nonregular point groups are divided into Types I, II, and III. The following theorem is proved in the first part of this section:

Theorem II: Consider an ionic lattice in the point ion approximation. Assume that the point group G for the lattice does not belong to the regular system and that no accidental degeneracies exist in the eigenvalues of $\mathbf{A}$, $\mathbf{A}'$, or $\mathbf{C}^0(\hat{\phi})$.

a. If the point group is of Type I, then the number of branches of the dispersion relations which do not approach definite frequencies as ϕ approaches zero is equal to the dimension of the subspace S(polar).

b. If the point group is of Type II [Type III], then the following is true. Let r and m be the number of independent eigenvectors of $\mathbf{A}$ lying in S(polar) and transforming according to $\mathbf{D}^{(z)}$ and $\mathbf{D}^{(xy)}$ [$\mathbf{D}^{(\pm)}$], respectively. Then the number of branches of the dispersion relations which do not approach definite frequencies as ϕ approaches zero is equal to $r + m/2$ where $r + m/2 \geq 2$.

To prove the theorem, let q be the dimension of S(polar) (where clearly $q \geq 3$). Then, by Lemma I, $(3f-q)$ branches of the dispersion relations approach definite frequencies in the long-wavelength limit. Thus, we need consider further only those q branches whose eigenvectors are linear combinations of eigenvectors of $\mathbf{A}$ spanning S(polar). By Lemma XVII these eigenvectors of $\mathbf{A}$ will belong to irreducible representations of the point group to which x, y, or z belong. If any such branch is to approach a definite frequency it must satisfy the conditions of Lemma XVI. That is, either its long-wavelength normal mode vectors $\Psi^0(\hat{\phi})$ lie in $S_{3f-1}(\lambda_N = 0; \hat{\phi})$ for all $\hat{\phi}$ or its normal mode eigenvectors $\Psi^0(\hat{\phi})$ lie in $S_{3f-2}(\lambda_M = 0; \hat{\phi})$ for all $\hat{\phi}$.

The irreducible representations of the point groups not belonging to the regular system are at most two-dimensional even when time reversal invariance is considered. It follows from the preceding paragraph and from Lemma XIV that none of the branches of the dispersion relations with eigenvectors in S(polar) have eigenvectors $\Psi^0(\hat{\phi})$ which lie in $S_{3f-2}(\lambda_M = 0; \hat{\phi})$ for all $\hat{\phi}$. Thus, in the following, we need only consider the possibility of the $\Psi^0(\hat{\phi})$ lying in $S_{3f-1}(\lambda_N = 0; \hat{\phi})$ for all $\hat{\phi}$.

First consider lattices with point groups of Type I. Such point groups have only one-dimensional, real, irreducible representations. It follows from Lemma XV and the above discussion that for such lattices none of the dispersion relations whose eigenvectors lie in S(polar) approach definite frequencies in the long-wavelength limit. Thus, part a of Theorem II is established.

Next consider lattices whose point groups are of Type II [Type III]. By Lemmas XV and XVII the only invariant subspaces of eigenvectors of $\mathbf{A}$ lying in S(polar) which can contain eigenvectors $\Psi^0(\hat{\phi})$ of $\mathbf{C}^0(\hat{\phi})$ lying in $S_{3f-1}(\lambda_N = 0; \hat{\phi})$ for all $\hat{\phi}$ are the two-dimensional subspaces spanned by eigenvectors transforming according to $\mathbf{D}^{(xy)}$ [$\mathbf{D}^{(\pm)}$]. Each such two-dimensional subspace will contain at least one set of such $\Psi^0(\hat{\phi})$ since its intersection with $S_{3f-1}(\lambda_N = 0; \hat{\phi})$ is at least one-dimensional for each $\hat{\phi}$. However, such a two-dimensional invariant subspace cannot contain two independent sets of eigenvectors of $\mathbf{C}^0(\hat{\phi})$ lying in $S_{3f-1}(\lambda_N = 0; \hat{\phi})$ for all $\hat{\phi}$. For, if it did, the invariant subspace would lie in $S_{3f-1}(\lambda_N = 0; \hat{\phi})$ for all $\hat{\phi}$ and thus be contained in S_{3f-3}(zero), contrary to our assumption. Thus, if r independent eigenvectors of $\mathbf{A}$ transforming according to $D^{(z)}$ and m independent eigenvectors of $\mathbf{A}$ transforming according to $D^{(xy)}$ [$D^{(\pm)}$] lie in S(polar), then $m/2$ corresponding branches of the dispersion relations approach definite frequencies in the long-wavelength limit and $r + m/2$ do not. By Lemma XVII, $r = 1, 2, 3, \cdots$ and $m = 2, 4, 6, \cdots$. Since $r + m \geq 3$, we see that $r + m/2 \geq 2$.

The above completes the formal proof of Theorem II. In the following we analyze the long-wavelength normal modes for crystals with point groups of Types II and III in greater detail.

We first consider crystals whose point groups are of Type II and determine the form of $\Psi^0(\hat{\phi})$ for a branch of the dispersion relations which lies outside of S_{3f-3}(zero) but which approaches a definite frequency. Let $\Psi^{(xy)i}$ ($i = 1$ or 2) be eigenvectors of A which transform among

themselves according to the first and second rows of $\mathbf{D}^{(xy)}$, respectively, and which are related by Eq. (44). A general vector in the invariant subspace spanned by the $\Psi^{(xy)i}$ is of the form

$$\Psi^{(xy)} = a_1\Psi^{(xy)1} + a_2\Psi^{(xy)2}. \tag{93}$$

From Eq. (35) we see that the condition that $\Psi^{(xy)}$ lies in $S_{3f-1}(\lambda_N = 0; \hat{\phi})$ is that $\mathbf{N}(\hat{\phi})\Psi^{(xy)} = 0$. Applying this condition together with Eqs. (5), (14), (15), and (72), we obtain the equation.

$$\mathbf{L}(\hat{\phi})\sum_\tau{}' \frac{Z_\nu Z_\tau}{(\mu_\nu\mu_\tau)^{1/2}}\left(a_1\sum_{\kappa\approx\tau}\Psi_\kappa^{(xy)1} + a_2\sum_{\kappa\approx\tau}\Psi_\kappa^{(xy)2}\right) = 0. \tag{94}$$

Using Eqs. (42), (44), (59), and (70), and a procedure similar to that used in deriving Eq. (74), we find that

$$\sum_{\kappa\approx\tau}\Psi_\kappa^{(xy)i} = c_\tau^{(xy)}\begin{bmatrix}\delta_{i1}\\ \delta_{i2}\\ 0\end{bmatrix}, \tag{95}$$

where

$$c_\tau^{(xy)} = \sum_{\kappa\approx\tau}\Psi_{\kappa 1}^{(xy)1}. \tag{96}$$

It follows from Eq. (94) and (95) and the fact that not all Z_ν vanish that

$$\left(\sum_\tau{}' Z_\tau\mu_\tau^{-1/2}c_\tau^{(xy)}\right)\begin{bmatrix}a_1\phi_1{}^2 + a_2\phi_1\phi_2\\ a_1\phi_1\phi_2 + a_2\phi_2{}^2\\ a_1\phi_1\phi_3 + a_2\phi_2\phi_3\end{bmatrix} = 0. \tag{97}$$

The vanishing of the first factor on the left of Eq. (97) would imply that $\Psi^{(xy)} \in S_{3f-3}(\text{zero})$, contrary to our assumption. Thus, the second factor must vanish. In order that it vanish for all $\hat{\phi}$, we must set $a_1 = k\phi_2$ and $a_2 = -k\phi_1$ where k is a proportionality factor. Thus, the normalized eigenvectors of $\mathbf{C}^0(\hat{\phi})$ which lie in $S_{3f-1}(\lambda_N = 0; \hat{\phi})$ for all $\hat{\phi}$ have the form

$$\Psi^{0a}(\hat{\phi}) = (\phi_1{}^2 + \phi_2{}^2)^{-1/2}(\phi_2\Psi^{(xy)1} - \phi_1\Psi^{(xy)2}), \tag{98}$$

where $\Psi^{(xy)1}$ and $\Psi^{(xy)2}$ are normalized. For lattices whose point groups are of Type II, Eq. (98) gives the general form of the normal mode eigenvectors for those branches of the dispersion relations which approach definite frequencies and whose normal mode eigenvectors lie in S(polar).

We note that for the particular values $\hat{\phi} = (0, 0, \pm 1)$, the two-dimensional invariant subspace is contained in $S_{3f-1}(\lambda_N = 0; \hat{\phi})$. For these values of $\hat{\phi}$, the branch of the dispersion relations becomes degenerate with another branch (one not approaching a definite frequency), and, as a result, the eigenvector $\Psi^{0a}(\hat{\phi})$ given by Eq. (98) becomes indeterminant for $\hat{\phi} = (0, 0, \pm 1)$. This degeneracy is required by symmetry from an analysis of $G(\hat{\phi})$, the group of the direction of $\hat{\phi}$, when $\hat{\phi} = (0, 0, \pm 1)$.

The normalized vector in the two-dimensional invariant subspace orthogonal to $\Psi^{0a}(\hat{\phi})$ is given by

$$\Psi^{0b}(\hat{\phi}) = (\phi_1{}^2 + \phi_2{}^2)^{-1/2}(\phi_1\Psi^{(xy)1} + \phi_2\Psi^{(xy)2}). \tag{99}$$

The eigenvectors for branches not approaching definite frequencies can be constructed as linear combinations of the $\Psi^{0b}(\hat{\phi})$ (one from each two-dimensional invariant subspace) and of eigenvectors of $\mathbf{A}$ belonging to $\mathbf{D}^{(z)}$ and lying within S(polar).

Next consider crystals whose point groups are of Type III. Let $\Psi^{(+)}$ and $\Psi^{(-)}$ be eigenvectors of $\mathbf{A}$, lying within S(polar), corresponding to the same eigenvalue, and transforming according to $\mathbf{D}^{(+)}$ and $\mathbf{D}^{(-)}$, respectively. These vectors span a two-dimensional subspace invariant under the operations of $\{\mathbf{T}(0, \mathbf{R})\}$ and time reversal. Since $\mathbf{A}$ is real, we choose

$$\Psi^{(-)} = \Psi^{(+)*}. \tag{100}$$

A general vector in the two-dimensional subspace has the form

$$\Psi = a_+\Psi^{(+)} + a_-\Psi^{(-)}. \tag{101}$$

To determine the form of the $\Psi^0(\hat{\phi})$ for a branch approaching a definite frequency, set $\mathbf{N}(\hat{\phi})\Psi = 0$. Applying Eqs. (5), (14), (15), and (72), we find that the latter condition reduces to

$$\mathbf{L}(\hat{\phi})\sum_\tau{}' \frac{Z_\nu Z_\tau}{(\mu_\nu\mu_\tau)^{1/2}}\left(a_+\sum_{\kappa\approx\tau}\Psi_\kappa^{(+)} + a_-\sum_{\kappa\approx\tau}\Psi_\kappa^{(-)}\right) = 0. \tag{102}$$

Using Eqs. (43), (64), (70), and (100), we obtain the result

$$\sum_{\kappa\approx\tau}\Psi_\kappa^{(+)} = \sum_{\kappa\approx\tau}\Psi_\kappa^{(-)*} = c(\tau)\begin{bmatrix}1\\ -i\\ 0\end{bmatrix}, \tag{103}$$

where

$$c(\tau) = \tfrac{1}{2}\sum_{\lambda\approx\tau}(\Psi_{\lambda 1}^{(+)} + i\Psi_{\lambda 2}^{(+)}). \tag{104}$$

Define the quantity

$$C = \sum_\tau{}' Z_\tau\mu_\tau^{-1/2}c(\tau). \tag{105}$$

Then, using the fact that all of the Z_ν are not zero, we reduce Eq. (102) to the equation

$$\mathbf{L}(\hat{\phi})\left\{Ca_+\begin{bmatrix}1\\ -i\\ 0\end{bmatrix} + C^*a_-\begin{bmatrix}1\\ i\\ 0\end{bmatrix}\right\} = 0. \tag{106}$$

One way to satisfy Eq. (106) is to set $Ca_+ + C^*a_- = 0$ and $Ca_+ - C^*a_- = 0$. Since both a_+ and a_- cannot vanish, the latter two equations hold only if $C = 0$. But it is easily shown that $C = 0$ implies that $\Psi^{(\pm)}$ lie in $S_{3f-3}(\text{zero})$, contrary to our assumption. To find solutions outside of $S_{3f-3}(\text{zero})$, we use Eq. (6) to reduce Eq. (106) to the single equation

$$Ca_+(\phi_1 - i\phi_2) + C^*a_-(\phi_1 + \phi_2) = 0. \tag{107}$$

Let $(1, \theta, \varphi)$ be the spherical coordinates of $\hat{\phi}$ in ϕ-space. Then Eq. (107) reduces to

$$\sin\theta[Ca_+\exp(-i\varphi) + C^*a_-\exp(i\varphi)] = 0. \tag{108}$$

The latter equation is satisfied for all $\hat{\phi}$ if and only if

$$a_+ = k\exp(i\varphi)C^*$$

and

$$a_- = -k\exp(-i\varphi)C, \tag{109}$$

where k is a factor of proportionality. Thus, if a branch of the dispersion relations approaches a definite frequency and its normal mode eigenvectors $\Psi^0(\hat{\phi})$ lie in S(polar), then the eigenvectors must be of the normalized form

$$\Psi^{0a}(\hat{\phi}) = \frac{1}{\sqrt{2}\,|C|}\,[C^* \exp(i\varphi)\Psi^{(+)} - C\exp(-i\varphi)\Psi^{(+)*}], \tag{110}$$

where $\Psi^{(+)}$ is assumed normalized.

From Eq. (108), we see that when $\theta = 0$ or π [that is, $\hat{\phi} = (0, 0, \pm 1)$] the two-dimensional invariant subspace is contained in $S_{3f-1}(\lambda_N = 0; \hat{\phi})$. The branch under consideration then becomes degenerate with another branch (one not approaching a definite frequency), and as a result $\Psi^{0a}(\hat{\phi})$ as given by Eq. (110) becomes indeterminant. Again this degeneracy is required by symmetry from an analysis of $G(\hat{\phi})$.

The normalized vector in the two-dimensional invariant subspace orthogonal to $\Psi^{0a}(\hat{\phi})$ is

$$\Psi^{0b}(\hat{\phi}) = \frac{1}{\sqrt{2}\,|C|}\,[C^* \exp(i\varphi)\Psi^{(+)} + C\exp(-i\varphi)\Psi^{(+)*}]. \tag{111}$$

The normal mode eigenvectors for branches not approaching definite frequencies are linear combinations of the $\Psi^{0b}(\hat{\phi})$ (one from each two-dimensional invariant subspace) and eigenvectors of A belonging to $\mathbf{D}^{(\varepsilon)}$ and lying in S(polar).

To summarize we state the following. For lattices whose point groups are of Type I, only those branches of the dispersion relations whose corresponding normal mode vibrations produce $\mathbf{P} = 0$ (and thus $\mathbf{E} = \mathbf{D} = 0$) approach definite frequencies in the long-wavelength limit. In the case of lattices whose point groups are of Types II or III, those branches of the dispersion relations which produce $\mathbf{P} = 0$ approach definite frequencies. In addition, definite frequencies will be approached by branches producing $\mathbf{E} = 0$ and $\mathbf{P} \neq 0$. For lattices considered in this section, there will be no branches with $\mathbf{P} \neq 0$ and $\mathbf{E} \neq 0$ which approach definite frequencies.

XI. CONCLUDING REMARKS. POLARIZABLE ATOMS

This paper completes the mathematical theory introduced in Ref. 1. Our analysis has been limited to the rigid-ion model. We have done little work with crystals with deformable or polarizable atoms. However, by referring to Sec. VI.5 of Maradudin, Montroll, Weiss, and Ipatova (Ref. 4), we can immediately obtain some idea of the additional complications imposed by the existence of such atoms.

In the above reference, the long-wavelength normal mode eigenvector equations appropriate for lattices with polarizable and deformable atoms are derived using a phenomenological approach. The eigenvector equations are given by Eqs. (6. 5. 34), (6. 5. 37), and (6. 5. 38) of Ref. 4. The long-wavelength dynamical matrix appearing in these equations depends upon the effective transverse electric charges $f_{\mu\nu}(\kappa)$ and upon the susceptibilities $\chi^{\infty}_{\mu\nu}$ which relate the macroscopic electric field to its contribution to the polarization.[22]

Using Ref. 4, we immediately find one special case where the presence of deformable and polarizable atoms affects nothing of substance in this paper. This is the case where the deformabilities and polarizabilities of the atoms are isotropic, that is, where,

$$f_{\mu\nu}(\kappa) = eZ_\kappa \delta_{\mu\nu}, \tag{112}$$

and

$$\chi^{\infty}_{\mu\nu} = \chi^{\infty}\delta_{\mu\nu}. \tag{113}$$

(Note that we have never restricted the Z_κ to integral values.) In this case, we need merely replace $\mathbf{C}^0(\hat{\phi})$ as given by Eq. (3) with

$$\mathbf{C}^0(\hat{\phi}) = \mathbf{A} + 4\pi \frac{a^3}{v_a}\frac{1}{\epsilon^{\infty}_L}\,\mathbf{N}(\hat{\phi}), \tag{114}$$

where ϵ^{∞}_L (the longitudinal, optical frequency dielectric constant) is independent of $\hat{\phi}$ and is given by

$$\epsilon^{\infty}_L = 1 + 4\pi\chi^{\infty}. \tag{115}$$

Thus, we need merely replace the constant $4\pi a^3/v_a$ with the new constant $4\pi a^3/(v_a\epsilon^{\infty}_L)$ everywhere in this paper. Since the particular value of $4\pi a^3/v_a$ enters into none of our arguments, nothing of substance is changed.

However, it is clear that a more general formalism than that presented in this paper and Ref. 1 is required to deal with lattices for which either of Eqs. (112) or (113) fails to hold true. If Eq. (112) is true and Eq. (113) is false, then $\mathbf{C}^0(\hat{\phi})$ is still given by Eq. (114). Now, however, ϵ^{∞}_L is $\hat{\phi}$-dependent. As a consequence much of the formalism developed in this paper breaks down. For example, Lemmas I and II remain true. However, Lemma III must be modified, for now the branch of the dispersion relations whose long-wavelength normal mode eigenvectors are the $\Psi^{In}(\hat{\phi})$ approaches the $\hat{\phi}$-dependent frequency corresponding to $\lambda^a + 4\pi a^3/[v_a\epsilon^{\infty}_L(\hat{\phi})]$. Similarly, Lemma IV must be modified. We find that the condition that $\mathbf{N}(\hat{\phi})\Psi^0(\hat{\phi}) = 0$, for all $\hat{\phi}$, remains a sufficient condition that a branch of the dispersion relations approaches a definite frequency, but the condition $\mathbf{M}(\hat{\phi})\Psi^0(\hat{\phi}) = 0$, for all $\hat{\phi}$, is no longer sufficient. Clearly much of the remaining work must be changed. If the condition of Eq. (112) is relaxed, then Eq. (114) is no longer true and the mathematical theory of this paper and Ref. 1 is not applicable. Finally we point out that Eq. (113) is required by symmetry to hold for all lattices whose point groups belong to the regular system. A sufficient condition that Eq. (112) hold for such lattices is that $\delta(\kappa, F_0(\nu, \mathbf{R})) = \delta_{\kappa\nu}$ for all κ, ν, and R of the point group.

To summarize, we state the following. The work in this paper remains applicable to lattices with deformable and polarizable atoms provided the deformabilities and polarizabilities are isotropic. If they are not isotropic, then a more general mathematical formalism is required.

APPENDIX A

The proof of Lemma XV is the following. The condition that $\Psi^0(\hat{\phi})$ lies in $S_{3f-1}(\lambda_N = 0; \hat{\phi})$ is $\mathbf{N}(\hat{\phi})\Psi^0(\hat{\phi}) = 0$. Using Eqs. (5), (7), (8), and (9), and the fact that not all Z_ν vanish, we see that the latter condition is equivalent to

$$\sum_\kappa Z_\kappa \mu_\kappa^{-1/2}\Psi^0_\kappa(\hat{\phi}) = q(\hat{\phi})\hat{\eta}(\hat{\phi}), \tag{A1}$$

where $q(\hat{\phi})$ is some number not equal to zero and

$\hat{\eta}^\dagger(\hat{\phi})\hat{\phi}=0$. Note that here $\hat{\eta}(\hat{\phi})$ may be complex. Its most general form is $\hat{\eta}(\hat{\phi})=\alpha_1\hat{\eta}^{(1)}(\hat{\phi})+\alpha_2\hat{\eta}^{(2)}(\hat{\phi})$ where $\hat{\eta}^{(1)}(\hat{\phi})$ and $\hat{\eta}^{(2)}(\hat{\phi})$ are independent, real vectors transverse to $\hat{\phi}$, and α_1 and α_2 are complex numbers.[23] [The caret in $\hat{\eta}(\hat{\phi})$ indicates that $\hat{\eta}^\dagger(\hat{\phi})\hat{\eta}(\hat{\phi})=1$.] We refer to the three-dimensional space containing the $\hat{\eta}(\hat{\phi})$ as complex ϕ-space. *However, the symbol $\hat{\phi}$ will continue to stand for a real unit vector.* Any three, linearly independent (real) $\hat{\phi}^{(i)}$ ($i=1$, 2, or 3) span complex ϕ-space. Consider the corresponding set of three vectors $\{\hat{\eta}(\hat{\phi}^{(i)})\}$. From this set, it must be possible to select two vectors which are linearly independent of each other. Otherwise any member of the set would be orthogonal to each of the $\hat{\phi}^{(i)}$, contrary to the fact that complex ϕ-space is three-dimensional.

It follows that, given any $\hat{\phi}^{(a)}$, there exists a $\hat{\phi}^{(b)}$ such that $\hat{\eta}(\hat{\phi}^{(a)})$ is linearly independent of $\hat{\eta}(\hat{\phi}^{(b)})$. It then follows from Eq. (A1) that $\sum_\kappa Z_\kappa\mu_\kappa^{-1/2}\Psi_\kappa^0(\hat{\phi}^{(a)})$ and $\sum_\kappa Z_\kappa\mu_\kappa^{-1/2}\Psi_\kappa^0(\hat{\phi}^{(b)})$ are linearly independent vectors in complex ϕ-space, and, therefore, $\Psi^0(\hat{\phi}^{(a)})$ and $\Psi^0(\hat{\phi}^{(b)})$ are linearly independent in S_{3f}(total). Thus, we have shown that the $\Psi^0(\hat{\phi})$ for the branch span a subspace of at least two dimensions. The $\Psi^0(\hat{\phi})$ are all eigenvectors of **A** corresponding to a common eigenvalue $\lambda=\lambda^0$. Since we are excluding accidental degeneracies from consideration and none of the irreducible representations of the point groups G are greater than three-dimensional, the space spanned by the $\Psi^0(\hat{\phi})$ for the branch is either two- or three-dimensional.

APPENDIX B

Consider a branch of the dispersion relations whose $\Psi^0(\hat{\phi})$ are contained in S(polar) and which approaches a definite frequency in the long-wavelength limit. Here we show that it is not possible that the $\Psi^0(\hat{\phi})$ are elements of $S_{3f-1}(\lambda_N=0;\hat{\phi})$ for some (but not all) $\hat{\phi}$ and are elements of $S_{3f-2}(\lambda_M=0;\hat{\phi})$ for the remaining $\hat{\phi}$.

Suppose that such a branch were to exist. Let $\Psi^0(\hat{\phi}^{(a)})$ and $\Psi^0(\hat{\phi}^{(b)})$ be eigenvectors of the branch where $\Psi^0(\hat{\phi}^{(a)})\in S_{3f-1}(\lambda_N=0;\hat{\phi}^{(a)})$ and $\Psi^0(\hat{\phi}^{(b)})\in S_{3f-2}(\lambda_M=0;\hat{\phi}^{(b)})$. Let λ^0 be the $\hat{\phi}$-independent eigenvalue of $\mathbf{C}^0(\hat{\phi})$ for the branch. Since $\mathbf{N}(\hat{\phi}^{(a)})\Psi^0(\hat{\phi}^{(a)})=0$ and $\mathbf{M}(\hat{\phi}^{(b)})\Psi^0(\hat{\phi}^{(b)})=0$, we see from Eqs. (3) and (34) that

$$\mathbf{A}\Psi^0(\hat{\phi}^{(a)})=\lambda^0\Psi^0(\hat{\phi}^{(a)}), \tag{B1}$$

and

$$\{\mathbf{A}+(4\pi a^3/v_a)[\mathbf{N}(\hat{\phi})+\mathbf{M}(\hat{\phi})]\}\Psi^0(\hat{\phi}^{(b)})=\lambda^0\Psi^0(\hat{\phi}^{(b)}). \tag{B2}$$

A simple calculation shows that

$$\Psi^{0\dagger}(\hat{\phi}^{(b)})[\mathbf{M}(\hat{\phi})+\mathbf{N}(\hat{\phi})]\Psi^0(\hat{\phi}^{(a)})=0, \tag{B3}$$

and from Eqs. (27) and (51) we immediately obtain the result

$$\{\Psi^0(\hat{\phi}^{(b)}),\Psi^0(\hat{\phi}^{(a)})\}=\sum_\kappa\frac{Z_\kappa}{\mu_\kappa^{1/2}}\Psi_\kappa^{0\dagger}(\hat{\phi}^{(b)})\sum_\nu\frac{Z_\nu}{\mu_\nu^{1/2}}\Psi_\nu^0(\hat{\phi}^{(a)})=0. \tag{B4}$$

According to Eqs. (49) and (A1),

$$\sum_\nu\frac{Z_\nu}{\mu_\nu^{1/2}}\Psi_\nu^0(\hat{\phi}^{(a)})=q(\hat{\phi}^{(a)})\hat{\eta}(\hat{\phi}^{(a)}) \tag{B5}$$

and

$$\sum_\kappa\frac{Z_\kappa}{\mu_\kappa^{1/2}}\Psi_\kappa^0(\hat{\phi}^{(b)})=k(\hat{\phi}^{(b)})\hat{\phi}^{(b)}, \tag{B6}$$

where $q(\hat{\phi}^{(a)})$ and $k(\hat{\phi}^{(b)})$ are nonvanishing numerical factors. It follows from Eq. (B4) that

$$\hat{\phi}^{(b)\dagger}\hat{\eta}(\hat{\phi}^{(a)})=0. \tag{B7}$$

Thus, $\hat{\phi}^{(b)}$ and $\hat{\eta}(\hat{\phi}^{(a)})$ must be orthogonal.

In Appendix A we defined the term complex ϕ-space. If the type of branch under consideration is to exist, then the $\hat{\phi}$ for which $\Psi^0(\hat{\phi})\in S_{3f-2}(\lambda_M=0;\hat{\phi})$ must either span complex ϕ-space or span a two-dimensional or a one-dimensional subspace of complex ϕ-space. We consider each of these three cases in turn.

Suppose that there exist $\Psi^0(\hat{\phi})\in S_{3f-2}(\lambda_M=0;\hat{\phi})$ for three linearly independent vectors $\hat{\phi}^{(1)}$, $\hat{\phi}^{(2)}$, and $\hat{\phi}^{(3)}$ in complex ϕ-space. Then for no $\hat{\phi}$ can $\hat{\eta}(\hat{\phi})$ satisfy Eq. (B7) for every $\hat{\phi}^{(i)}$ ($i=1$, 2, or 3), and $\Psi^0(\hat{\phi})\in S_{3f-2}(\lambda_M=0;\hat{\phi})$ for all $\hat{\phi}$.

Next assume that $\Psi^0(\hat{\phi})\in S_{3f-2}(\lambda_M=0;\hat{\phi})$ for just two linearly independent values of $\hat{\phi}$, say $\hat{\phi}^{(1)}$ and $\hat{\phi}^{(2)}$. Then the $\hat{\eta}(\hat{\phi})$ for those $\Psi^0(\hat{\phi})\in S_{3f-1}(\lambda_N=0;\hat{\phi})$ must be orthogonal to the $\hat{\phi}^{(1)}-\hat{\phi}^{(2)}$ plane. Consider $\hat{\phi}^{(3)}$ orthogonal to the $\hat{\phi}^{(1)}-\hat{\phi}^{(2)}$ plane. Clearly $\Psi^0(\hat{\phi}^{(3)})\in S_{3f-1}(\lambda_N=0;\hat{\phi}^{(3)})$ since $\hat{\phi}^{(3)}$ is linearly independent of $\hat{\phi}^{(1)}$ and $\hat{\phi}^{(2)}$. But $\hat{\eta}(\hat{\phi}^{(3)})$ lies in the $\hat{\phi}^{(1)}-\hat{\phi}^{(2)}$ plane. We are led to a contradiction and conclude that this case cannot occur.

Finally assume that $\Psi^0(\hat{\phi})\in S_{3f-2}(\lambda_M=0;\hat{\phi})$ for $\hat{\phi}=\pm\hat{\phi}^{(3)}$ (for one particular $\hat{\phi}^{(3)}$) and for no other $\hat{\phi}$. Then $\Psi^0(\hat{\phi})\in S_{3f-1}(\lambda_N=0;\hat{\phi})$ for all $\hat{\phi}\neq\pm\hat{\phi}^{(3)}$ and the corresponding $\hat{\eta}(\hat{\phi})$ in Eq. (B5) lie in the plane in complex ϕ-space orthogonal to $\hat{\phi}^{(3)}$. In the following paragraphs we prove that the preceding sentence implies that $\Psi^0(\hat{\phi})\in S_{3f-1}(\lambda_N=0;\hat{\phi})$ for all $\hat{\phi}$, and we obtain a contradiction. Thus, this final case cannot occur. Therefore, we have completed the proof of Lemma XVI.

To prove the above statement [that $\Psi^0(\hat{\phi})\in S_{3f-1}(\lambda_N=0;\hat{\phi})$ for all ϕ] consider the following hypothetical situation. We are given that a branch of the dispersion relations approaches a definite eigenvalue λ^0 of $\mathbf{C}^0(\hat{\phi})$. Further, the $\Psi^0(\hat{\phi})$ for the branch lie in S(polar) and are elements of $S_{3f-1}(\lambda_N=0;\hat{\phi})$ except possibly for $\hat{\phi}=\pm\hat{\phi}^{(3)}$ (for one particular $\hat{\phi}^{(3)}$). Finally, we are given that, for each $\hat{\phi}\neq\pm\hat{\phi}^{(3)}$, $\hat{\eta}(\hat{\phi})$ in Eq. (B5) lies in the plane in complex ϕ-space orthogonal to $\hat{\phi}^{(3)}$. Notice that, for $\hat{\phi}=\pm\hat{\phi}^{(3)}$, $\hat{\eta}(\hat{\phi})$ is well defined to within a phase factor since it is orthogonal to both $\hat{\phi}$ and $\hat{\phi}^{(3)}$. In the remainder of this section we prove that $\Psi^0(\pm\hat{\phi}^{(3)})\in S_{3f-1}(\lambda_N=0;\pm\hat{\phi}^{(3)})$.

Choose $\hat{\phi}^{(1)}$ and $\hat{\phi}^{(2)}$ such that neither equals $\pm\hat{\phi}^{(3)}$ and $\hat{\eta}(\hat{\phi}^{(1)})$ and $\hat{\eta}(\hat{\phi}^{(2)})$ are linearly independent. (Our discussion in Appendix A shows that the above choice is always possible.) Then $\Psi^0(\hat{\phi}^{(1)})$ and $\Psi^0(\hat{\phi}^{(2)})$ are linearly independent vectors in S_{3f}(total). First we show that for $\hat{\phi}\neq\pm\hat{\phi}^{(3)}$ any $\Psi^0(\hat{\phi})$ of the branch can be constructed as a linear combination of $\Psi^0(\hat{\phi}^{(1)})$ and $\Psi^0(\hat{\phi}^{(2)})$. For all $\hat{\phi}$, the $\Psi^0(\hat{\phi})$ are eigenvectors of **A** corresponding to the eigenvalue λ^0. Thus, a general $\Psi^0(\hat{\phi})$ for the branch ($\hat{\phi}\neq\pm\hat{\phi}^{(3)}$) is of the form

$$\Psi^0(\hat{\phi}) = f_1(\hat{\phi})\Psi^0(\hat{\phi}^{(1)}) + f_2(\hat{\phi})\Psi^0(\hat{\phi}^{(2)}) + \Omega(\hat{\phi}), \tag{B8}$$

where the $f_i(\hat{\phi})$ are numbers, and either $\Omega(\hat{\phi}) = 0$ or $\Omega(\hat{\phi})$ is linearly independent of $\Psi^0(\hat{\phi}^{(1)})$ and $\Psi^0(\hat{\phi}^{(2)})$ and is an eigenvector of **A** corresponding to the eigenvalue λ^0. We now show that $\Omega(\hat{\phi}) = 0$. Applying Eq. (A1) to $\Psi^0(\hat{\phi})$, $\Psi^0(\hat{\phi}^{(1)})$, and $\Psi^0(\hat{\phi}^{(2)})$ in Eq. (B8), we obtain the result

$$f_1(\hat{\phi})q(\hat{\phi}^{(1)})\hat{\eta}(\hat{\phi}^{(1)}) + f_2(\hat{\phi})q(\hat{\phi}^{(2)})\hat{\eta}(\hat{\phi}^{(2)}) + \sum_\kappa \frac{Z_\kappa}{\mu_\kappa^{1/2}}\,\Omega_\kappa(\hat{\phi}) = q(\hat{\phi})\hat{\eta}(\hat{\phi}) \tag{B9}$$

Since $\hat{\eta}(\hat{\phi})$ lies in the $\hat{\eta}(\hat{\phi}^{(1)})$–$\hat{\eta}(\hat{\phi}^{(2)})$ plane, $\sum_\kappa Z_\kappa \mu_\kappa^{-1/2} \times \Omega_\kappa(\hat{\phi})$ must be a linear combination of $\hat{\eta}(\hat{\phi}^{(1)})$ and $\hat{\eta}(\hat{\phi}^{(2)})$. Thus, applying Eq. (A1), we find that

$$\sum_\kappa (Z_\kappa/\mu_\kappa^{1/2})[\Omega_\kappa(\hat{\phi}) - c_1(\hat{\phi})\Psi_\kappa^0(\hat{\phi}^{(1)}) - c_2(\hat{\phi})\Psi_\kappa^0(\hat{\phi}^{(2)})] = 0, \tag{B10}$$

where $c_1(\hat{\phi})$ and $c_2(\hat{\phi})$ are numbers. Let $\Theta(\hat{\phi}) = \Omega(\hat{\phi}) - c_1(\hat{\phi})\Psi^0(\hat{\phi}^{(1)}) - c_2(\hat{\phi})\Psi^0(\hat{\phi}^{(2)})$. Equation (B10) implies that $\Theta(\hat{\phi})$ lies in S_{3f-3}(zero). But either $\Theta(\hat{\phi}) = 0$ or it is an eigenvector of **A** corresponding to the eigenvalue λ^0. The latter possibility violates Lemma XI and we conclude that $\Theta(\hat{\phi}) = 0$. But then $\Omega(\hat{\phi}) = c_1(\hat{\phi})\Psi^0(\hat{\phi}^{(1)}) + c_2(\hat{\phi})\Psi^0(\hat{\phi}^{(2)})$. We have stated that either $\Omega(\hat{\phi})$ is linearly independent of $\Psi^0(\hat{\phi}^{(1)})$ and $\Psi^0(\hat{\phi}^{(2)})$ or $\Omega(\hat{\phi}) = 0$. Clearly, $\Omega(\hat{\phi}) = 0$.

Thus, for any $\hat{\phi} \neq \pm\hat{\phi}^{(3)}$,

$$\Psi^0(\hat{\phi}) = f_1(\hat{\phi})\Psi^0(\hat{\phi}^{(1)}) + f_2(\hat{\phi})\Psi^0(\hat{\phi}^{(2)}). \tag{B11}$$

Again applying Eq. (A1) to $\Psi^0(\hat{\phi})$, $\Psi^0(\hat{\phi}^{(1)})$, and $\Psi^0(\hat{\phi}^{(2)})$, we obtain

$$\Psi^0(\hat{\phi}) = q(\hat{\phi})(1 - |\hat{\eta}^\dagger(\hat{\phi}^{(1)})\hat{\eta}(\hat{\phi}^{(2)})|^2)^{-1} \times\left(\frac{g_1(\hat{\phi})}{q(\hat{\phi}^{(1)})}\,\Psi^0(\hat{\phi}^{(1)}) + \frac{g_2(\hat{\phi})}{q(\hat{\phi}^{(2)})}\,\Psi^0(\hat{\phi}^{(2)})\right), \tag{B12}$$

where

$$g_1(\hat{\phi}) = \hat{\eta}^\dagger(\hat{\phi}^{(1)})\hat{\eta}(\hat{\phi}) - [\hat{\eta}^\dagger(\hat{\phi}^{(1)})\hat{\eta}(\hat{\phi}^{(2)})][\hat{\eta}^\dagger(\hat{\phi}^{(2)})\hat{\eta}(\hat{\phi})] \tag{B13}$$

and

$$g_2(\hat{\phi}) = \hat{\eta}^\dagger(\hat{\phi}^{(2)})\hat{\eta}(\hat{\phi}) - [\hat{\eta}^\dagger(\hat{\phi}^{(2)})\hat{\eta}(\hat{\phi}^{(1)})][\hat{\eta}^\dagger(\hat{\phi}^{(1)})\hat{\eta}(\hat{\phi})]. \tag{B14}$$

In Eq. (B12), the term $1 - |\hat{\eta}^\dagger(\hat{\phi}^{(1)})\hat{\eta}(\hat{\phi}^{(2)})|^2$ is nonzero by the Schwartz inequality. If $\Psi^0(\hat{\phi}^{(1)})$ and $\Psi^0(\hat{\phi}^{(2)})$ are given, Eq. (B12) determines $\Psi^0(\hat{\phi})$ for any $\hat{\phi} \neq \pm\hat{\phi}^{(3)}$ to within a numerical factor $q(\hat{\phi})$. [The undetermined phase factor in $\hat{\eta}(\hat{\phi})$ can be absorbed into $q(\hat{\phi})$.]

The behavior of the branch at $\hat{\phi} = \pm\hat{\phi}^{(3)}$ is determined by extending Eq. (B12) to include these points. The vectors $\Psi^0(\pm\hat{\phi}^{(3)})$ are not well defined since the $\hat{\eta}(\pm\hat{\phi}^{(3)})$ are only restricted to lie in the plane in complex ϕ-space orthogonal to $\hat{\phi}^{(3)}$. As $\hat{\eta}(\hat{\phi}^{(3)})$ circles about the $\hat{\phi}^{(3)}$ direction, $\Psi^0(\hat{\phi}^{(3)})$ as given by Eq. (B12) sweeps out a two-dimensional subspace S_2 of S_{3f}(total). This fact presents no difficulties. It merely shows that at $\hat{\phi} = \pm\hat{\phi}^{(3)}$ the branch under consideration becomes degenerate with another branch of the dispersion relations. At $\hat{\phi} = \pm\hat{\phi}^{(3)}$ the eigenvectors for the branches meeting at $\hat{\phi} = \pm\hat{\phi}^{(3)}$ are vectors lying in S_2. But all vectors in S_2 are contained in $S_{3f-1}(\lambda_N = 0; \pm\hat{\phi}^{(3)})$. Thus, all eigenvectors $\Psi^0(\hat{\phi})$ for the branch under consideration lie in $S_{3f-1}(\lambda_N = 0; \hat{\phi})$.

We conclude by noting that the hypothetical situation considered above does occur in the case of the branches of the dispersion relations analyzed in the latter part of Sec. X.

[1] J. A. Davies, J. Math. Phys. 13, 1207 (1972).

[2] In Ref. 1 we obtain the dynamical matrix for the Coulomb case by beginning with an r^{-p} potential between the particles with $p > 1$ and then taking the limit of the dynamical matrix as p approaches one. We find that, even for a given direction of **k**, there exists no unique double limit of the dynamical matrix as p approaches one and **k** approaches zero. On the other hand, a well-defined limit of the dynamical matrix as p approaches one exists for all nonzero **k** in any small neighborhood of **k** = 0.

[3] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford, London, 1954), pp. 248—77.

[4] A. A. Maradudin, E. W. Montroll, G. H. Weiss, and I. P. Ipatova, *Theory of Lattice Dynamics in the Harmonic Approximation* (Academic, New York and London, 1971), 2nd ed., Sec. VI.

[5] H. B. Rosenstock, Phys. Rev. 121, 416 (1961).

[6] Reference 1, Sec. IV.

[7] J. L. Warren, Rev. Mod. Phys. 40, 38 (1968).

[8] Reference 1, Secs. III—VI.

[9] Reference 3, pp. 269—70.

[10] Reference 1, Secs. VIII—IX.

[11] Reference 3, pp. 265—70; Ref. 4, Sec. VI. 5.

[12] In all sections of this article except Appendices A and B, the symbol $\hat{\eta}(\hat{\phi})$ represents a real vector orthogonal to $\hat{\phi}$. Complex $\hat{\eta}(\hat{\phi})$ are used in the Appendices. The symbol $\hat{\phi}$, of course, represents a real vector throughout this article including Appendices A and B.

[13] A. A. Maradudin and S. H. Vosko, Rev. Mod. Phys. 40, 1 (1968).

[14] If $q > 1$, the correct linear combinations are determined by expanding $\mathbf{C}(\phi)$ to higher orders in ϕ and applying the usual techniques of degenerate perturbation theory.

[15] In a Footnote to Ref. 1, the following statement is proved: Let S_m and S_n' be subspaces of S_{3f}(total) spanned by m and n linearly independent vectors, respectively, where $m + n \geq 3f$. Then the intersection of S_m and S_n' must be at least $(m + n - 3f)$-dimensional.

[16] See Footnote 14 with q replaced by $q - 1$.

[17] M. Bakr, M.A. thesis, Clark University (1973) (unpublished).

[18] Reference 14, Sec. 4.

[19] Note that the proof of Lemma XII does not depend upon the second paragraph of Lemma I. When proving this part of Lemma I in Ref. 1, we did not consider the possibility that $\mathbf{C}^0(\hat{\phi})$ might be required by symmetry to have a degenerate eigenvalue with corresponding eigenvectors lying both inside and outside of S_{3f-3}(zero). Lemma XII shows that the above possibility does not occur. If two eigenvectors of $\mathbf{C}^0(\hat{\phi})$, one lying in S_{3f-3}(zero) and one lying outside S_{3f-3}(zero), correspond to the same degenerate eigenvalue, then the degeneracy is accidental.

[20] See, for example, M. Hamermesh, *Group Theory* (Addison-Wesley, Reading, Mass., 1962), Chap. 3.

[21] The first step in the derivation of Eq. (74) is to write down the identity

$$\sum_{\nu\simeq\tau} \Psi_\nu^{\Gamma\mu 3} = \sum_{\nu\simeq\tau}\sum_\lambda P_{\nu\lambda}^{\Gamma 3}\,\Psi_\lambda^{\Gamma\mu 3}.$$

[22] The subscripts μ and ν in $f_{\mu\nu}(\kappa)$ and $\chi_{\mu\nu}^\infty$ refer to Cartesian coordinates. Here we use Greek instead of Roman letters in order to conform with the notation used in Ref. 4.

[23] See Footnote 12.

Exact results in quantum many-body systems of interacting particles in many dimensions with $\overline{SU(1,1)}$ as the dynamical group

Pascal J. Gambardella

Institute for Theoretical Physics, State University of New York at Stony Brook, Stony Brook, New York 11790
(Received 1 October 1974)

We consider a class of system of N interacting particles in any dimension—the potential includes a quadratic pair potential and an arbitrary translation-invariant position-dependent potential that is homogeneous of degree -2. The group $\overline{SU(1,1)}$ is the dynamical group for the Hamiltonian. We illustrate the significance of the Casimir operator in relation to the separation of variables method; obtain a series of eigenfunctions that transform under the unitary irreducible representations of $\overline{SU(1,1)}$ labeled by the ground state energy; indicate the structure of arbitrary eigenfunctions; and specify when the complete energy spectrum is linear. We treat N-body examples which include two- and three-body forces. For N identical particles in one dimension interacting with a quadratic pair potential and an inverse square pair potential, we exhibit a series of eigenfunctions characterized by four quantum numbers. These eigenfunctions reduce to the complete set of eigenfunctions for five particles. We indicate how a complete set of eigenfunctions for N particles are obtained.

I. INTRODUCTION

The study of exactly solvable nontrivial quantum systems of N interacting particles proceeds with the hope that the systems have some general characteristics in common with real systems. Also, exact results are crucial in estimating the accuracy of some of the approximation techniques used, for example, in statistical mechanics or nuclear physics.

Consider the system of N interacting identical particles characterized by the Hamiltonian

$$H=\frac{1}{2m}\sum_{i=1}^{N}\mathbf{p}_i^2+\frac{m\omega^2}{4}\sum_{i>j=1}^{N}\mathbf{r}_{ij}^2+V_\lambda(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N) \tag{1.1}$$

where $\mathbf{r}_i$ is the s-dimensional coordinate of the ith particle, and $\mathbf{r}_{ij}=\mathbf{r}_i-\mathbf{r}_j$; $\mathbf{p}_i$ is the s-dimensional momentum of the ith particle; m is the particle mass; and ω is the frequency. The potential V_λ is position-dependent, translation-invariant, and homogeneous of degree -2. There is no restriction on the dimension s, unless specified. Some examples of the potential V_λ are

$$\text{(i)}\quad V_{\lambda_1}=\lambda_1\sum_{i>j=1}^{N}\frac{1}{\mathbf{r}_{ij}^2}\qquad\text{(two body)}, \tag{1.2}$$

$$\text{(ii)}\quad V_{\lambda_2}=\lambda_2\sum_{k=1}^{N}\sum_{i>j=1}^{N}\frac{\mathbf{r}_{ki}\cdot\mathbf{r}_{kj}}{\mathbf{r}_{ki}^2\mathbf{r}_{kj}^2}\qquad\text{(three body)}, \tag{1.3}$$

$$\text{(iii)}\quad V_{\lambda_3}=\lambda_3\sum_{i>k=1}^{N}\sum_{\substack{j>e=1\\ i\neq j\neq k\\ i\neq e\neq k}}^{N}\frac{\mathbf{r}_{ik}\cdot\mathbf{r}_{je}}{\mathbf{r}_{ik}^2\mathbf{r}_{je}^2}\qquad\text{(four body)}, \tag{1.4}$$

$$\text{(iv)}\quad V_{\lambda_4}=\lambda_4/r^2,\quad r^2=\frac{1}{N}\sum_{i>j=1}^{N}\mathbf{r}_{ij}^2 \tag{1.5}$$

where λ_i $(1\leq i\leq 4)$ are constants.

The universal covering group of $SU(1,1)$, i.e., $\overline{SU(1,1)}$, forms the noninvariance or dynamical group of the class of systems governed by Hamiltonian (1.1). The eigenvalue of its only Casimir operator completely determines the energy spectrum. The Hamiltonian is a generator of its Lie algebra[1]; thus, the Lie algebra is referred to as a spectrum generating algebra.

One purpose of this paper is to give a concrete example of the role of a noncompact group in solving a many-body Schrödinger equation in many dimensions. This provides a splendid opportunity for those unfamiliar with noncompact groups to gain insight into the power and elegance these groups display in physical applications. So in Sec. II, we discuss the spinor group $SU(1,1)$ and its covering $\overline{SU(1,1)}$ with some of the necessary background and perspective. In Sec. III, we prove a theorem on the significance of the Casimir operator in relation to the separation of variables method, justify the restrictions on the potential V_λ, and give some examples.

Calogero and Marchioro[2] considered a system of N interacting particles in three dimensions consisting of distinguishable particles or bosons. The Hamiltonian they treated is (1.1) with a specific linear combination of the potentials V_{λ_1} (1.2) and V_{λ_2} (1.3) for V_λ. They exhibited a subset of completely symmetric eigenfunctions including the ground state and the corresponding energy eigenvalues. In Sec. IV, we obtain these results in any dimension as a consequence of the group theoretical approach.

We prove that the complete energy spectrum of the Hamiltonian (1.1) is linear if V_λ does not contain the potential V_{λ_4} (1.5) as a term. In addition, we indicate the structure of arbitrary eigenfunctions.

The N-body system in one dimension with the Hamiltonian H(1.1) and potential V_{λ_1} (1.2) has been solved completely by Calogero[3] for the case of two particles ($N=2$) and the case of three particles ($N=3$). For the case of N particles, he[4,5] exhibited the eigenfunction of the ground state, the complete energy spectrum, and a series of eigenfunctions for the excited states. The energy spectrum is independent of the statistics (Bose, Fermi) which the particles satisfy. In addition, it coincides except for a constant shift of all energy levels to the energy spectrum of the corresponding problem with just the harmonic forces, i.e., an $(N-1)$-dimensional harmonic oscillator. This important correspondence provided one reason for Calogero[5] to

 Copyright © 1975 American Institute of Physics

suggest the possibility of an underlying group structure in the solution of this N-body problem. Perelomov[6] showed that the Hamiltonian is a generator of the Lie[1] algebra of $SO(2,1)$. He exhibited a series of eigenfunctions of the N-body case characterized by three quantum numbers. These eigenfunctions reduce to the complete set of eigenfunctions for four particles. The eigenfunctions of the multidimensional harmonic oscillator can be expressed in terms of energy raising operators acting on the ground state; so, Perelomov expressed the eigenfunctions of this one-dimensional Calogero problem in terms of energy raising operators acting on the ground state. A necessary condition for the existence of these eigenfunctions is that the raising operators mutually commute. Perelomov verified this explicitly for each pair of raising operators. We generalize the method of Perelomov and exhibit a series of eigenfunctions for the N body case characterized by four quantum numbers. These eigenfunctions reduce to the complete set of eigenfunctions for five particles. We also express the eigenfunctions in terms of raising operators acting on the ground state; but, we provide a more general proof that the raising operators mutually commute.

II. THE SPINOR GROUP $SU(1,1)$ AND ITS UNIVERSAL COVERING GROUP $\overline{SU(1,1)}$ [7]

A representation of a dynamical group (in general a noncompact group) of the Hamiltonian determines the energy spectrum and energy states including the degeneracy of the levels; whereas, a representation of a symmetry group usually determines just the energy states for a given energy. The noncompact group $\overline{SU(1,1)}$ is the dynamical group for the Hamiltonians considered in this paper.

A Lie group is noncompact if and only if at least one of its parameters varies over an infinite range. The noncompact Lie group $SU(1,1)$ is the group of all two-dimensional complex matrices $g(\alpha,\beta)$ of the form

$$g(\alpha,\beta)=\begin{pmatrix}\alpha & \beta\\ \beta^* & \alpha^*\end{pmatrix} \tag{2.1a}$$

where α and β are complex parameters, and

$$|\alpha|^2-|\beta|^2=1. \tag{2.1b}$$

The matrices $g(\alpha,\beta)$ correspond to conformal transformations of the unit circle's interior into itself, given by

$$z'=\frac{\alpha^* z+\beta^*}{\beta z+\alpha} \tag{2.2}$$

where z' and z are complex variables. In addition, $SU(1,1)$ is the group of linear transformations on two complex variables z_1 and z_1 leaving the quadratic form $|z_1|^2-|z_2|^2$ invariant, i.e.,

$$g(\alpha,\beta)\begin{pmatrix}z_1\\ z_2\end{pmatrix}=\begin{pmatrix}z_1'\\ z_2'\end{pmatrix} \quad \text{with } |z_1|^2-|z_2|^2=|z_1'|^2-|z_2'|^2. \tag{2.3}$$

All local properties of a Lie group can be described in terms of its Lie algebra. We can characterize the group element $g(\alpha,\beta)$ by three independent real parameters since the complex parameters α and β are restricted by (2.1b). Thus, the Lie algebra has three generators; the generators I_1, I_2, and I_3 satisfy the commutation relations

$$[I_1,I_2]=iI_3, \tag{2.4a}$$

$$[I_3,I_1]=-iI_2, \tag{2.4b}$$

$$[I_2,I_3]=iI_1. \tag{2.4c}$$

In our application, the generator I_2 is proportional to the Hamiltonian; thus, the eigenvalue ϵ is proportional to the energy.

We discuss the class of unitary irreducible representations (UIR's) of $SU(1,1)$ called the positive discrete series $D^+(\Phi)$. The Casimir operator Q, defined by

$$Q\equiv -I_3^2-I_1^2+I_2^2, \tag{2.5}$$

commutes with all the elements of the Lie algebra. The Casimir operator Q and the generator I_2 are diagonal in this representation with eigenvalues

$$q=\Phi(\Phi+1) \tag{2.6}$$

and

$$\epsilon=a-\Phi, \tag{2.7a}$$

respectively, where a is a nonnegative integer, and Φ is real with $\Phi<0$. Each UIR is labeled by constant Φ. The eigenvalue ϵ can be rewritten in terms of q, i.e.,

$$\epsilon_\pm=a+\tfrac{1}{2}\pm(q+\tfrac{1}{4})^{1/2} \tag{2.7b}$$

where ϵ_- has the restriction $(q+\frac{1}{4})^{1/2}<\frac{1}{2}$ since $\Phi<0$.

The group manifold of $SU(1,1)$ is the product of a circle and a plane (picture three-dimensional space with a line deleted) indicating that the manifold is not simply connected and also that the group possesses a $U(1)$ subgroup. The universal covering group $\overline{G}$ of a Lie group G is essentially defined as the Lie group that is locally isomorphic to G and has a simply connected group manifold. The group manifold of $\overline{SU(1,1)}$ is three-dimensional Euclidean space. The Lie algebra of $SU(1,1)$ and the Lie algebra of $\overline{SU(1,1)}$ are locally isomorphic, i.e., each has three generators satisfying (2.4). Any UIR of $SU(1,1)$ can be extended to a UIR of $\overline{SU(1,1)}$. If the latter representation is not a representation of $SU(1,1)$, it is called a projective representation of $SU(1,1)$.

For the group $SU(1,1)$ there is a discreteness condition on ϵ (and, thus on Φ) as a consequence of the $U(1)$ subgroup, i.e., ϵ is an integer or half-integer. This discreteness condition is eliminated when the representation $D^+(\Phi)$ is extended to a representation of $\overline{SU(1,1)}$ since the group manifold of $SU(1,1)$ is not restricted by a $U(1)$ subgroup. In this projective representation the eigenvalue ϵ can be an arbitrary real number. We are mainly concerned, in this paper, with projective representations of $SU(1,1)$ since the energy eigenvalue will not take just integer or half integer values.

The raising and lowering operators of I_2 are

$$A_2^\pm=(I_3\pm iI_1), \quad A_2^-\equiv A_2 \tag{2.8}$$

with the commutation relations

$$[I_2,A_2^+]=A_2^+, \tag{2.9a}$$

$$[I_2,A_2]=-A_2, \tag{2.9b}$$

$$[A_2^+,A_2]=-2I_2 \tag{2.9c}$$

and matrix elements

$$A_2^+|\Phi,a\rangle=[(a-2\Phi)(a+1)]^{1/2}|\Phi,a+1\rangle, \tag{2.10a}$$

$$A_2|\Phi,a\rangle=[a(a-2\Phi-1)]^{1/2}|\Phi,a-1\rangle \tag{2.10b}$$

where

$$\langle\Phi,a|\Phi,a'\rangle=\delta_{aa'}. \tag{2.10c}$$

In a projective representation of $SU(1,1)$, relations such as (2.10) remain the same since the Lie algebra of $SU(1,1)$ is isomorphic to the Lie algebra of $\overline{SU(1,1)}$.

III. THE HAMILTONIAN AND CASIMIR OPERATOR

The group $\overline{SU(1,1)}$ is a dynamical group for the Hamiltonian (1.1); Perelomov[6] demonstrated this for one dimension; Burdet and Perrin[8] indicated this for a linear combination of the potentials V_{λ_1} (1.2), V_{λ_2} (1.3), and V_{λ_3} (1.4). We show that an important consequence of this group structure is the separation of the Schrödinger equation into an equation in the radial variable and an equation in the "angular variables"; the Casimir operator eigenvalue equation is, apart from a constant, the equation in the angular variables. This is more than just the conventional separation of variables method since the eigenvalue of Casimir operator determines the complete energy spectrum. In addition, we are provided with a lowering operator which gives an equation for the ground state and with a raising operator which generates a subset of eigenfunctions once the ground state is known. We assume the potential V_λ is translation invariant; we will show that the additional restrictions on the potential V_λ are consistent with an induced group structure.

We eliminate the center of mass motion from the Hamiltonian (1.1) to obtain it in the form[9]

$$H=H_{N,s}+\alpha^2 C_{N,s} \tag{3.1}$$

where

$$\alpha=(N/2)^{1/2}\omega, \tag{3.2}$$

$$H_{N,s}=\frac{1}{2mN}\sum_{i>k=1}^{N}\mathbf{p}_{ik}^2+V_\lambda, \tag{3.3}$$

$$C_{N,s}=\frac{m}{2N}\sum_{i>k=1}^{N}\mathbf{r}_{ik}^2, \tag{3.4}$$

and $\mathbf{p}_{ik}=\mathbf{p}_i-\mathbf{p}_k$. The subscripts N and s refer to the total number of particles and dimension, respectively. Equation (3.1) is obtained from (1.1) by applying the identity

$$\frac{1}{2m}\sum_{i=1}^{N}\mathbf{p}_i^2=\frac{1}{2mN}\left(\sum_{i=1}^{N}\mathbf{p}_i\right)^2+\frac{1}{2mN}\sum_{i>k=1}^{N}\mathbf{p}_{ik}^2, \tag{3.5}$$

where the first term on the rhs of (3.5) is the kinetic energy of the center of mass. After defining

$$D_{N,s}\equiv\frac{1}{2N}\sum_{i>k=1}^{N}(\mathbf{r}_{ik}\cdot\mathbf{p}_{ik}+\mathbf{p}_{ik}\cdot\mathbf{r}_{ik}), \tag{3.6}$$

we transform the operators $D_{N,s}$, $C_{N,s}$, and $H_{N,s}$ to the Jacobi position coordinates $\mathbf{X}_j$ and momentum coordinates $\mathbf{P}_j$ defined by

$$\mathbf{X}_j=\left(\frac{1}{j(j+1)}\right)^{1/2}\left(\sum_{k=1}^{j}\mathbf{r}_k-j\mathbf{r}_{j+1}\right), \tag{3.7a}$$

$$\mathbf{P}_j=\left(\frac{1}{j(j+1)}\right)^{1/2}\left(\sum_{k=1}^{j}\mathbf{p}_k-j\mathbf{p}_{j+1}\right), \tag{3.7b}$$

where

$$[\mathbf{X}_j,\mathbf{P}_k]=i\hbar s\,\delta_{ij}. \tag{3.8}$$

The transformed operators are given by

$$D_{N,s}=\frac{1}{2}\sum_{i=1}^{N-1}(\mathbf{X}_i\cdot\mathbf{P}_i+\mathbf{P}_i\cdot\mathbf{X}_i), \tag{3.9a}$$

$$H_{N,s}=\frac{1}{2m}\sum_{i=1}^{N-1}\mathbf{P}^2+V_\lambda, \tag{3.9b}$$

$$C_{N,s}=\frac{m}{2}\sum_{i=1}^{N-1}\mathbf{X}_i^2. \tag{3.9c}$$

The operators $C_{N,s}$, $D_{N,s}$, and $H_{N,s}$ form a closed Lie algebra with commutation relations

$$[C_{N,s},D_{N,s}]=2i\hbar C_{N,s}, \tag{3.10a}$$

$$[C_{N,s},H_{N,s}]=i\hbar D_{N,s}, \tag{3.10b}$$

$$[H_{N,s},D_{N,s}]=-2i\hbar H_{N,s} \tag{3.10c}$$

if and only if the potential V_λ satisfied the commutation relations

$$[V_\lambda,D_{N,s}]=-2i\hbar V_\lambda \tag{3.11a}$$

$$[V_\lambda,C_{N,s}]=0. \tag{3.11b}$$

The commutation relations (3.10) transform into the commutation relations of $\overline{SU(1,1)}$ (2.4) by defining the generators I_1, I_2, and I_3 as follows:

$$I_1=(-1/2\hbar)D_{N,s}, \tag{3.12a}$$

$$I_2=(1/2\hbar\alpha)H, \tag{3.12b}$$

$$I_3=(1/2\hbar\alpha)(-H_{N,s}+\alpha^2 C_{N,s}). \tag{3.12c}$$

The summation in each of the equations (3.9) contains $(N-1)s\equiv t$ terms. We relabel each term with the subscript k and sum from $k=1$ to $k=t$ to obtain

$$D_{N,s}=\frac{1}{2}\sum_{k=1}^{t}\{X_k,P_k\}_+, \tag{3.13a}$$

$$H_{N,s}=\frac{1}{2m}\sum_{k=1}^{t}P_k^2+V_\lambda, \tag{3.13b}$$

$$C_{N,s}=\frac{m}{2}\sum_{k=1}^{t}X_k^2, \tag{3.13c}$$

where $\{\ ,\ \}_+$ is the anticommutator bracket, and

$$[X_i,P_k]=i\hbar\delta_{ik} \tag{3.14}$$

with

$$P_k=\frac{\hbar}{i}\frac{\partial}{\partial X_k}. \tag{3.15}$$

We use (3.14) and (3.15) to rewrite $D_{N,s}$ (3.13a) in the form

$$D_{N,s}=\frac{\hbar}{i}\sum_{k=1}^{t}X_k\frac{\partial}{\partial X_k}-\frac{i\hbar t}{2}. \tag{3.16}$$

Finally, substituting (3.16) into (3.11a), we obtain a restriction on the potential V_λ in a more obvious form, i.e.,

$$\sum_{k=1}^{t}X_k\frac{\partial}{\partial X_k}V_\lambda=-2V_\lambda. \tag{3.17}$$

Thus, the translation-invariant potential V_λ is homogeneous of degree -2. Also, the assumption that V_λ is just a function of coordinates is consistent with (3.11b). The Casimir operator (2.5), after some manipulation (see Appendix A), is given by

$$Q_{N,s}=\frac{m}{2\hbar^2}\left(\frac{t(t-4)\hbar^2}{8m}+r^2V_\lambda+\frac{\hbar^2}{2m}L^2_{N,s}\right),\quad t=s(N-1), \tag{3.18}$$

where

$$r^2=\sum_{k=1}^{t}X_k^2 \tag{3.19}$$

and

$$\hbar^2L^2_{N,s}=\begin{cases}\sum_{i>k=1}^{t}(X_iP_k-X_kP_i)^2, & t>1, \qquad (3.20a)\\ 0, & t=1\ (N=2,s=1). \qquad (3.20b)\end{cases}$$

We are now ready to discuss the significance of the Casimir operator. Arbitrary eigenfunctions of the Schrödinger equation have the form

$$\Psi_{ak}(r,\Omega)=R_{ak}(r)W_k(\Omega) \tag{3.21}$$

where Ω represents the angular variables. The Schrödinger equation separates into an eigenvalue equation for $R_{ak}(r)$ and an eigenvalue equation for $W_k(\Omega)$. The eigenvalue equation for $W_k(\Omega)$ is

$$H_\Omega W_k(\Omega)=\delta_k W_k(\Omega) \tag{3.22}$$

where

$$H_\Omega=\frac{\hbar^2}{2m}L^2_{N,s}+r^2V_{\lambda_1}. \tag{3.23}$$

We write the potential

$$V_\lambda=V_{\lambda_1}+\lambda_2/r^2, \tag{3.24}$$

where λ_2 is a constant, and V_{λ_1} does not contain a term of the form $1/r^2$. The eigenvalue equation for R_{ak} is

$$\left(H_r+\frac{\delta_k}{r^2}\right)R_k(r)=E_{ak}R_{ak}(r) \tag{3.25}$$

where

$$H_r\equiv\frac{-\hbar^2}{2m}r^{1-t}\frac{\partial}{\partial r}r^{t-1}\frac{\partial}{\partial r}+\frac{m\alpha^2}{2}r^2+\frac{\lambda_2}{r^2}. \tag{3.26}$$

Thus, the Schrödinger equation is written

$$H\Psi_{ak}=E_{ak}\Psi_{ak} \tag{3.27}$$

with

$$H=H_r+\frac{1}{r^2}H_\Omega. \tag{3.28}$$

Note that the Casimir operator $Q_{N,s}$ (3.18), apart from a constant, is the operator H_Ω (3.23). Recall that $SU(1,1)$ has only one Casimir operator with

$$[H,Q_{N,s}]=0. \tag{3.29}$$

Finally, we insert (3.23) and (3.28) into (3.29), to obtain

$$[H_r,H_\Omega]=0. \tag{3.30}$$

Therefore, the group property (3.29) naturally implies the "separation of variables" indicated by (3.30).

The energy eigenvalue is

$$E_{aq}=\hbar\alpha\{2a+1+(1+4q)^{1/2}\}, \tag{3.31}$$

where a is a nonnegative integer, and q is the eigenvalue of the Casimir operator. Equation (3.31) is obtained from ϵ_+ (2.7b) and (3.12b). In the next section, we indicate how the energy ϵ_- (2.7b) leads to physically unacceptable eigenfunctions.

We consider, as an example, a system of N bosons with Hamiltonian (3.28) and potential

$$V_\lambda=\lambda/r^2. \tag{3.32}$$

The Casimir operator is now, apart from a constant, the "grand angular momentum" operator $L^2_{N,s}$. One can construct an orthonormal set of eigenfunctions for $L^2_{N,s}$ in terms of the harmonic polynomials[11,12] $H^t_\Lambda(\rho_1,\rho_2,\ldots,\rho_t)$, with $\rho_i=X_i/r$. These polynomials are homogeneous of degree Λ in the $t\equiv s(N-1)$ coordinates ρ_i and satisfy the eigenvalue equation

$$L^2_{N,s}H^t_\Lambda(\rho_1,\rho_2,\ldots,\rho_t)=\Lambda(\Lambda+t-2)H^t_\Lambda(\rho_1,\rho_2,\ldots,\rho_t) \tag{3.33a}$$

and the Laplace equation

$$\sum_{i=1}^{t}\frac{\partial^2}{\partial X_i^2}[r^\Lambda H^t_\Lambda(\rho_1,\rho_2,\ldots,\rho_t)]=0. \tag{3.33b}$$

For each Λ, there are

$$\frac{(2\Lambda+t-2)(\Lambda+t-3)}{\Lambda!(t-2)!} \tag{3.34}$$

linearly independent polynomials. Inserting (3.33a) into (3.18), we obtain

$$Q_{N,s}H^t_\Lambda(\rho_1,\rho_2,\ldots,\rho_t)=qH^t_\Lambda(\rho_1,\rho_2,\ldots,\rho_t) \tag{3.35}$$

with

$$q=\frac{\Lambda(\Lambda+t-2)}{4}+\frac{m\lambda}{2\hbar^2}+\frac{t(t-4)}{16}. \tag{3.36}$$

Inserting (3.36) into (3.31), we obtain the complete energy spectrum

$$E_{a\Lambda}=\hbar\alpha\{2a+1+[(\Lambda+t/2-1)^2+2m\lambda/\hbar^2]^{1/2}\} \tag{3.37a}$$

where $t\equiv s(N-1)>1$, and Λ and a are nonnegative integers. For $t=1$, i.e., $(N=2,\ s=1)$, we have (recall 3.20b)

$$E_a=\hbar\alpha\{2a+1+[\tfrac{1}{4}+2m\lambda/\hbar^2]^{1/2}\}. \tag{3.37b}$$

To insure a real spectrum, i.e., the system does not collapse to the origin,[10] the restriction on λ, from (3.37), is

$$\lambda>-(\hbar^2/8m)(sN-s-2)^2. \tag{3.38}$$

The energy spectrum of the particles interacting with just harmonic forces,[12] i.e., $\lambda\equiv0$, is given by

$$E_{a\Lambda_{\lambda=0}}=\hbar\alpha(2a+\Lambda+t/2),\quad t>1, \tag{3.39a}$$

$$E_{a_{\lambda=0}}=\hbar\alpha(2a+\tfrac{3}{2}),\quad t=1. \tag{3.39b}$$

Comparing (3.37a) and (3.39a), we note that the pres-

ence of the potential λ/r^2 drastically breaks the degeneracy of the system with just harmonic forces.

IV. RAISING AND LOWERING OPERATORS AND EIGENFUNCTIONS

We scale the (relabeled) Jacobi coordinates X_i according to

$$y_i=(m\alpha/\hbar)^{1/2}X_i, \quad 1\leq i\leq t\equiv s(N-1), \tag{4.1}$$

and use energy units $\hbar\alpha\equiv 1$; thus,

$$\rho^2\equiv\sum_{i=1}^{t} y_i^2=\frac{m}{\hbar^2}\,r^2. \tag{4.2}$$

Under this transformation, the Hamiltonian (3.28) becomes

$$H=-\frac{1}{2}\sum_{i=1}^{t}\frac{\partial^2}{\partial y_i^2}+\tfrac{1}{2}\rho^2+V_\lambda(y_1,y_2,\ldots,y_t). \tag{4.3}$$

The ground state of the Hamiltonian (4.3) has the form

$$\Psi_0(y_1,y_2,\ldots,y_t)=T_0K(y_1,y_2,\ldots,y_t)\exp(-\tfrac{1}{2}\rho^2) \tag{4.4}$$

where K is a translation-invariant function that is homogeneous of degree k, $k>0$, and T_0 is the normalization constant. In this section, we obtain a series of eigenfunctions for the Hamiltonian (4.3), derive a differential equation for the function K in terms of the potential V_λ, and give some examples. The methods Perelomov[6] used to obtain these results in one dimension are essentially the same for many dimensions. In addition, we construct the form for an arbitrary eigenfunction and show that the energy spectrum is linear if the potential V_λ does not contain γ/ρ^2 (γ a constant) as a term.

Inserting (3.12a) and (3.12c) into (2.8), we have the raising and lowering operators

$$A_2^{\pm}=\frac{1}{4}\sum_{i=1}\frac{\partial^2}{\partial y_i^2}-\frac{V_\lambda}{2}+\frac{\rho^2}{4}\pm\frac{1}{2}\left(-\sum_{i=1}^{t}y_i\frac{\partial}{\partial y_i}-\frac{t}{2}\right). \tag{4.5}$$

The Hamiltonian (4.3) and the raising and lowering operators (4.5) can be rewritten compactly in the conventional form

$$H=\frac{1}{2}\sum_{i=1}^{t}a_i^{+}a_i+V_\lambda+\frac{t}{2}, \tag{4.6}$$

$$A_2^{+}=\frac{1}{4}\sum_{i=1}^{t}a_i^{+2}-\frac{V_\lambda}{2}, \tag{4.7}$$

$$A_2=\frac{1}{4}\sum_{i=1}^{t}a_i^{2}-\frac{V_\lambda}{2}, \tag{4.8}$$

where

$$a_i^{+}=\left(y_i-\frac{\partial}{\partial y_i}\right), \quad a_i=\left(y_i+\frac{\partial}{\partial y_i}\right) \tag{4.9}$$

and

$$[a_i^{+},a_j^{+}]=[a_i,a_j]=0, \quad [a_i,a_j^{+}]=2\delta_{ij}. \tag{4.10}$$

We can generate a series of eigenfunctions by repeated application of the raising operator A_2^{+} (4.5) on the ground state ψ_0 (4.4); thus, with (2.10a) we obtain

$$\psi_{a\Phi}=\left(\frac{\Gamma(-2\Phi)}{a!\,\Gamma(a-2\Phi)}\right)^{1/2}(A_2^{+})^a\psi_0 \tag{4.11}$$

where

$$E_{a\Phi}=2a-2\Phi \tag{4.12}$$

is the energy eigenvalue, a is a nonnegative integer, $q=\Phi(\Phi+1)$ is the eigenvalue of the Casimir operator, and $\Gamma(1+z)\equiv z!$. The quantity (-2Φ) is not necessarily the ground state, i.e., there may be additional quantum numbers. The eigenfunctions $\psi_{a\Phi}$ (4.11) which transform under the UIR's of $\overline{SU(1,1)}$, labeled by constant Φ, are normalized if ψ_0 is normalized.

The equation

$$A_2\psi_0\equiv 0 \tag{4.13}$$

defines the ground state ψ_0. Inserting A_2 (4.5) into (4.13), we rewrite (4.13) as

$$[\exp(\rho^2/2)A_2\exp(-\rho^2/2)]K=0. \tag{4.14}$$

Inserting A_2 (4.5) into (4.14) and using the identity

$$\exp(\gamma\rho^2/2)\frac{\partial^n}{\partial y_i^n}\exp(-\gamma\rho^2/2)=\left(\frac{\partial}{\partial y_i}-\gamma y_i\right)^n \tag{4.15}$$

with $n=2$ and $\gamma=1$, we obtain an equation for K in terms of V_λ, i.e.,

$$V_\lambda=\frac{1}{2K}\sum_{i=1}^{t}\frac{\partial^2}{\partial y_i^2}K. \tag{4.16}$$

Equation (4.16) is the starting point for the generation of many solvable models, provided the eigenfunctions are normalizable.

We now obtain the series of eigenfunctions of the Hamiltonian that transform under UIR's of $\overline{SU(1,1)}$ labeled by the ground state, i.e., Φ is a constant [see (4.12)]. After operating on the ground state ψ_0 (4.4) with the Hamiltonian H (4.3) and using (4.16), we obtain the ground state energy,

$$-2\Phi=E_0=t/2+k. \tag{4.17}$$

Although E_0 and ψ_0 determine the series of eigenfunctions (4.11) with Φ a constant, we show that these eigenfunctions have the form

$$\psi_a=\psi_0T_a(\rho) \tag{4.18}$$

where $T_a(\rho)$ is a function of ρ alone; the eigenvalue equation is [see (4.12)]

$$H\psi_0=(2a+E_0)\psi_a. \tag{4.19}$$

We define the transformation $H\to\overline{H}$ by

$$\overline{H}=(\psi_0)^{-1}H\psi_0; \tag{4.20}$$

using the identity (4.15) and the result (4.16), we transform H to

$$\overline{H}=\sum_{i=1}^{t}\left(-\frac{1}{2}\frac{\partial^2}{\partial y_i^2}-\frac{\partial K}{\partial y_i}\frac{\partial}{\partial y_i}+y_i\frac{\partial}{\partial y_i}\right)+\frac{t}{2}+k. \tag{4.21}$$

If the operator $\overline{H}$ acts on functions of ρ alone it becomes effectively

$$\overline{H}_\rho=-\frac{1}{2}\frac{\partial^2}{\partial z^2}-\left(\frac{(\delta+\frac{1}{2})}{z}-z\right)\frac{\partial}{\partial z}+\delta+1 \tag{4.22}$$

where $\rho=z^{1/2}$ and

$$\delta=\frac{t}{2}+k-1=E_0-1. \tag{4.23}$$

We recognize the eigenvalue equation for $\overline{H}$ (4.22) as

the equation for the generalized Laguerre polynomial[13] $L_a^{E_0-1}(\rho^2)$ with

$$\bar{H}L_a^{E_0-1}(\rho^2)=\bar{H}_\rho L_a^{E_0-1}(\rho^2)=(2a+E_0)L_a^{E_0-1}(\rho^2), \tag{4.24}$$

where a is a nonnegative integer. Using (4.20), we rewrite (4.24) as

$$H[\psi_0 L_a^{E_0-1}(\rho^2)]=(2a+E_0)[\psi_0 L_a^{E_0-1}(\rho^2)]. \tag{4.25}$$

Hence, after comparing (4.25) with (4.19), we conclude that the series of eigenfunctions (4.18) are

$$\psi_a=N_{a,E_0}K(y_1,y_2,\ldots,y_t)\exp(-\rho^2/2)\,L_a^{E_0-1}(\rho^2) \tag{4.26}$$

where N_{a,E_0} is a normalization constant that depends on a and E_0 [see (4.11)].

We return to the example of N particles with $V_\lambda=\lambda/r^2$. Rewriting this potential in scaled Jacobi coordinates with (4.2), we have

$$V_\lambda=\gamma/\rho^2 \tag{4.27}$$

where $\gamma=m\lambda/\hbar^2$. Substituting (4.27) into (4.16), we find that

$$K=\rho^k \tag{4.28}$$

is the translation-invariant solution, where

$$k_\pm=-\frac{(t-2)}{2}\pm\left[\left(\frac{t-2}{2}\right)^2+2\gamma\right]^{1/2}. \tag{4.29}$$

The solution k_- does not lead to physically acceptable eigenfunctions. To see this, we examine the behavior of the radial eigenfunction $R(\rho)$ near the origin, with

$$R(\rho)\sim A\rho^{k_+}+B\rho^{k_-} \tag{4.30}$$

where A, B are constants; and γ is adjusted so that $k_+<0$ (the worse condition). We construct a sphere of radius ρ_0 about the origin. The radial eigenfunction $R(\rho)$ (4.30) and $dR(\rho)/d\rho$ must be continuous across the boundary of this sphere for $R(\rho)$ to be a physically acceptable eigenfunction. One can show[10] that these conditions on $R(\rho)$ lead to

$$\lim_{\rho_0\to 0}\frac{B}{A}\propto\lim_{\rho_0\to 0}\rho_0^{(k_+-k_-)}\to 0.$$

Inserting k_+ of (4.29) into (4.17), we obtain the ground state energy

$$E_0=1+\left[\left(\frac{t-2}{2}\right)^2+\frac{2m\lambda}{\hbar^2}\right]^{1/2}. \tag{4.31}$$

This is the ground state we obtained previously [see (3.37a)] from the eigenvalue of the Casimir operator. Thus, choosing k_+ instead of k_- is equivalent to choosing ϵ_+ over ϵ_- in (2.7b).

Inserting (4.28) into (4.26), we have the orthonormal (see Appendix B) radial eigenfunctions for this example

$$R_a(\rho)=\left(\frac{2\Gamma(E_0)}{a!\,\Gamma(a-E_0)}\right)^{1/2}\rho^{E_0-t/2}\exp(-\rho^2/2)\,L_a^{E_0-1}(\rho^2) \tag{4.32}$$

where E_0 is given by (4.31).

We now assume the potential V does not contain a term γ/ρ^2. An arbitrary eigenfunction of the Hamiltonian (4.3) is given by

$$\psi_{a\mu}=\frac{N}{\rho^\beta}K(y_1,y_2,\ldots,y_t)\,P_{\mu,s}(y_1,y_2\cdots y_t)\,\phi_{a\mu}(\rho^2) \tag{4.33}$$

where

$$\phi_{a\mu}(\rho^2)=\left(\frac{2\Gamma(-2\Phi)}{a!\Gamma(a+2\Phi)}\right)^{1/2}\rho^{-2\Phi-t/2}\exp(-\rho^2/2)\,L_a^{-2\Phi-1}(\rho^2) \tag{4.34}$$

are the orthonormal radial eigenfunctions with

$$-2\Phi=t/2+k+\mu, \tag{4.35}$$

$P_{\mu,s}$ is a homogeneous translation-invariant function of degree μ for the s-dimensional case, N is a normalization constant, and $\beta=k+\mu$. The Casimir operator eigenvalue equation is written

$$Q_{N,s}\left(\frac{KP_{\mu,s}}{\rho^\beta}\right)=q\left(\frac{KP_{\mu,s}}{\rho^\beta}\right). \tag{4.36}$$

In Appendix C we show (4.36) leads to an equation for $P_{\mu,s}$, i.e.,

$$\sum_{i=1}^{t}\left(\frac{\partial^2}{\partial y_i^2}P_{\mu,s}+\frac{2}{K}\frac{\partial K}{\partial y_i}\frac{\partial}{\partial y_i}P_{\mu,s}\right)=0 \tag{4.37}$$

and eigenvalue

$$q=\frac{t(t-4)}{16}+\frac{\beta^2}{4}+\frac{\beta(t-2)}{4},\quad \beta=k+\mu. \tag{4.38}$$

The function $P_{\mu,s}$ is a physically acceptable eigenfunction if it is nonsingular, i.e., $\mu\geq 0$.

The physical argument for choosing ϵ_+ over ϵ_- [see (2.7b)] when calculating the energy is essentially unchanged for this case since the radial eigenvalue equation (3.25) has the same structure for both cases. Inserting (4.38) into (3.31), we obtain the linear energy spectrum

$$E_{a\mu}=2a+\mu+E_0, \tag{4.39}$$

with $E_0=t/2+k$; thus, -2Φ (4.35) is equal to the ground state energy E_0 only if $\mu=0$.

Although the independent Jacobi coordinates have made calculations less complicated thus far, they tend to obscure the symmetry of the potential in the examples we now treat. So, following Perelomov,[6] we transform to more "democratic" coordinates. The transformation between the original coordinates $\mathbf{r}_i=(\mu_{i_1},\mu_{i_2},\ldots,\mu_{i_s})$, $1\leq i\leq N$, of Sec. I and the dependent coordinates $\boldsymbol{\Gamma}_i=(\xi_{i_1},\xi_{i_2},\ldots,\xi_{i_s})$, $1\leq i\leq N$, is given by

$$\xi_{i_\omega}=\left(\frac{m\alpha}{\hbar}\right)^{1/2}\mu_{i_\omega}-U_\omega,\quad \frac{\partial}{\partial\xi_{i_\omega}}=\left(\frac{\hbar}{m\alpha}\right)^{1/2}\frac{\partial}{\partial\mu_{i_\omega}}-\frac{\partial}{\partial U_\omega} \tag{4.40}$$

in the dimension ω, where the center of mass coordinates are

$$U_\omega=\left(\frac{m\alpha}{\hbar}\right)^{1/2}\frac{1}{N}\sum_{j=1}^{N}\mu_{j_\omega},\quad \frac{\partial}{\partial U_\omega}=\left(\frac{\hbar}{m\alpha}\right)^{1/2}\frac{1}{N}\sum_{j=1}^{N}\frac{\partial}{\partial\mu_{j_\omega}}. \tag{4.41}$$

In addition, we have

$$\sum_{i=1}^{N}\xi_{i_\omega}=\sum_{i=1}^{N}\frac{\partial}{\partial\xi_{i_\omega}}=0, \tag{4.42a}$$

$$\frac{\partial \xi_{i_\omega}}{\partial \xi_{j_\omega}} = \delta_{ij} - \frac{1}{N}, \tag{4.42b}$$

and

$$\Gamma^2 = \sum_{i=1}^{N} \Gamma_i^2. \tag{4.43}$$

Under this transformation, the Hamiltonian (1.1) becomes

$$H = -\frac{N}{2}\sum_{k=1}^{s}\frac{\partial^2}{\partial U_k^2} - \frac{1}{2}\sum_{i=1}^{N}\Delta_i + \tfrac{1}{2}\Gamma^2 + V_\lambda(\xi_1, \xi_2, \ldots, \xi_N) \tag{4.44}$$

where

$$\sum_{i=1}^{N}\Delta_i = \sum_{i=1}^{N}\sum_{k=1}^{s}\frac{\partial^2}{\partial \xi_{i_k}^2}. \tag{4.45}$$

Since we are only concerned with translation-invariant eigenfunctions, we rewrite (4.44) as

$$H = -\frac{1}{2}\sum_{i=1}^{N}\Delta_i + \tfrac{1}{2}\Gamma^2 + V_\lambda(\xi_1, \xi_2, \ldots, \xi_N). \tag{4.46}$$

Suppose

$$K = \prod_{i>j=1}^{N} F_{ij} \tag{4.47}$$

where F_{ij} is only a function of $\Gamma_{ij} = |\boldsymbol{\Gamma}_i - \boldsymbol{\Gamma}_j|$; from (4.16) [note that $\sum_{i=1}^{t}\partial^2/\partial y_i^2$ in Jacobi coordinates equals (4.45)] and (4.47), we have

$$V_\lambda = \frac{1}{2}\sum_{m=1}^{N}\sum_{i>j=1}^{N}\frac{\Delta_m F_{ij}}{F_{ij}} + \sum_{m=1}^{N}\sum_{k=1}^{N}\sum_{i>j=1}^{N}\frac{\boldsymbol{\nabla}_m F_{ki}\cdot\boldsymbol{\nabla}_{kj}F_{kj}}{F_{ki}F_{kj}} \tag{4.48}$$

Since F_{ij} is only a function of Γ_{ij} alone, we rewrite (4.48) as

$$V_\lambda = \sum_{i>j=1}^{N}\left(\frac{(s-1)F_{ij}'}{\Gamma_{ij}F_{ij}} + \frac{F_{ij}''}{F_{ij}}\right) + \sum_{k=1}^{N}\sum_{i>j=1}^{N}\frac{(\boldsymbol{\Gamma}_{ki}\cdot\boldsymbol{\Gamma}_{kj})}{\Gamma_{ki}\Gamma_{kj}}\frac{F_{ki}'}{F_{ki}}\frac{F_{kj}'}{F_{kj}} \tag{4.49}$$

where the prime indicates differentiation with respect to the argument and s is the dimension. Equations (4.48) and (4.49) represent a direct link between the ground state

$$\psi_0 = K\exp(-\Gamma^2/2) \tag{4.50}$$

and the potential V_λ. We now turn to some examples.

Calogero and Marchioro[2] presented the result (4.49) in three dimensions and used it to obtain a series of eigenfunctions, corresponding to (4.33) with $P_{\mu,3} \equiv 1$, i.e., $\mu = 0$, for some many body problems. We have shown that their method has a firm group theoretical foundation. Suppose

$$F_{ki} = \Gamma_{ki}^{\Theta_s} \tag{4.51}$$

in (4.47), then (4.49) becomes

$$V_\lambda = \lambda\sum_{i>j=1}^{N}\frac{1}{\Gamma_{ij}^2} + G\sum_{k=1}^{N}\sum_{i>j=1}^{N}\frac{(\boldsymbol{\Gamma}_{ki}\cdot\boldsymbol{\Gamma}_{kj})}{\Gamma_{ki}^2\Gamma_{kj}^2} \tag{4.52}$$

where

$$\Theta_s = \tfrac{1}{2}[(s-2)^2 + 4\lambda]^{1/2} - s/2 + 1, \tag{4.53}$$

$$G = \Theta_s^2, \tag{4.54}$$

and

$$\lambda > -\tfrac{1}{4}(s-2)^2. \tag{4.55}$$

The function K, for this example, is homogeneous of degree

$$k = [N(N-1)/2]\Theta_s; \tag{4.56}$$

hence, a series of eigenfunctions of H (4.46) with V_λ (4.52) is given by (4.33) in the $\boldsymbol{\Gamma}_i$ coordinates (with $\rho^2 = \Gamma^2$) with ground state energy

$$E_0 = \frac{s(N-1)}{2} + \frac{N(N-1)}{2}\Theta_s. \tag{4.57}$$

It is remarkable that in one and three dimensions the potential can be attractive, but in two dimensions $\lambda > 0$ [see (4.55)], i.e., the potential is only repulsive. We can understand this result intuitively as follows: For attractive potentials, i.e., $\lambda < 0$, the restriction (4.55) prevents two body collapse by providing a nonzero real quantum zero point energy [see (4.57)]. In one dimension each particle has only two nearest neighbors. However, in two and higher dimensions there can be as many as $(N-1)$ nearest neighbors. In two dimensions, we have the lowest zero point energy per particle which must be shared with a maximum of $(N-1)$ particles to prevent two body collapse. In three dimensions, the space per particle and the zero point energy per particle are greater; hence, the possibility of an attractive potential with no collapse as we have verified.

The second term in the potential (4.52) is a three-body potential; the term corresponding to a given three particles vanishes when the particles coincide.[2] Thus, in one dimension, i.e., $s = 1$, (4.52), becomes

$$V_\lambda = \lambda\sum_{i>j=1}^{N}\frac{1}{\xi_{ij}^2}. \tag{4.58}$$

Perelomov[6] first displayed the group theoretical structure associated with the one-dimensional potential (4.58). This potential preserves the particle order[5]; therefore, we can solve the Schrödinger equation in the sector $\xi_1 \leq \xi_2 \leq \cdots \leq \xi_N$ to obtain the eigenfunction $\psi(\xi_1, \xi_2, \ldots, \xi_N)$. We assume the particles are spinless. The total wavefunction is given by

$$\Psi(\xi_1, \xi_2, \ldots, \xi_N) = \sum_{\text{all permutations}} \eta_\rho P\psi(\xi_1, \xi_2, \ldots, \xi_N)$$

where P is the permutation operator, and

$$\eta_\rho = 1 \quad \text{for bosons,}$$

$$\eta_\rho = \begin{cases} 1 & \text{even permutations} \\ -1 & \text{odd permutations} \end{cases} \quad \text{for fermions.}$$

Thus, the energy spectrum for fermions or bosons is the same.

The eigenfunctions (4.33), for the potential (4.52) in higher dimensions, are physically realizable in terms of a system of N interacting bosons if the functions $P_{\mu,s}$ are completely symmetric.

The analog of equation (4.37) for $P_{\mu,s}$ in the coordinates (4.40) is

$$\sum_{i=1}^{N}\Delta_i P_{\mu,s} + \frac{2}{K}\sum_{\omega=1}^{s}\sum_{i=1}^{N}\frac{\partial K}{\partial \xi_{i_\omega}}\frac{\partial}{\partial \xi_{i_\omega}}P_{\mu,s} = 0. \tag{4.59}$$

For the potential (4.52) with (4.47) and (4.51), Eq. (4.59) becomes

$$\sum_{i=1}^{N} \Delta_i P_{\mu,s} + 2\Theta_s \sum_{\omega=1}^{s} \sum_{i>j=1}^{N} \frac{\xi_{i_\omega j_\omega}}{\Gamma_{ij}^2} \left(\frac{\partial}{\partial \xi_{i_\omega}} - \frac{\partial}{\partial \xi_{j_\omega}} \right) P_{\mu,s} = 0 \tag{4.60}$$

where

$$\Gamma_{ij}^2 = \sum_{\omega=1}^{s} \xi_{i_\omega j_\omega}^2. \tag{4.61}$$

To obtain (4.60), we have used the following two equations:

$$\frac{1}{K} \frac{\partial K}{\partial \xi_{i_\omega}} = \Theta_s \sum_{i=1}^{N} \frac{\xi_{i_\omega j_\omega}}{\Gamma_{ij}^2} \tag{4.62}$$

and

$$\sum_{\substack{i=1 \\ i\neq j}}^{N} \sum_{j=1}^{N} \frac{\xi_{i_\omega j_\omega}}{\Gamma_{ij}^2} \frac{\partial}{\partial \xi_{i_\omega}} = \sum_{i>j=1}^{N} \frac{\xi_{i_\omega j_\omega}}{\Gamma_{ij}^2} \left(\frac{\partial}{\partial \xi_{i_\omega}} - \frac{\partial}{\partial \xi_{j_\omega}} \right). \tag{4.63}$$

We reserve an analysis of (4.60) for $s>1$ to another paper.

We discuss the one-dimensional potential (4.58) and $P_{\mu,1}$. This discussion is essentially a summary of the observations of Calogero[5] and Perelomov[6] concerning the functions $P_{\mu,1}$. In one dimension, (4.60) becomes

$$\sum_{i=1}^{N} \frac{\partial^2}{\partial \xi_i^2} P_{\mu,1} + 2\Theta_1 \sum_{i>j=1}^{N} T_{ij} P_{\mu,1} = 0, \quad N \geq 3, \tag{4.64}$$

where

$$T_{ij} = \frac{1}{\xi_{ij}} \left(\frac{\partial}{\partial \xi_i} - \frac{\partial}{\partial \xi_j} \right). \tag{4.65}$$

Calogero has shown that the translation-invariant functions $P_{\mu,1}(\xi_1, \xi_2, \ldots, \xi_N)$ are completely symmetric under exchange of any two coordinates ξ_i and ξ_j; thus, we can construct[6] the functions $P_{\mu,1}$ from the elementary functions

$$h_l = \sum_{k=1}^{N} \xi_k^l \tag{4.66}$$

where at least $l \geq 0$ since $\mu \geq 0$. Operating on (4.66) with T_{ij} (4.65), we have

$$T_{ij} h_l = \frac{l}{\xi_{ij}} \left(\xi_i^{l-1} - \xi_j^{l-1} \right); \tag{4.67}$$

but, (4.67) (and consequently $P_{\mu,1}$) is nonsingular only if l is an integer greater than or equal to two. The solution of (4.64) with the lowest μ is $P_3 = h_3$, and thus μ is an integer greater than or equal to three. The number of different functions $P_{\mu,1}$ for constant μ is given by the solutions to the equation[5,6]

$$\mu = \sum_{\omega=3}^{N} \omega a_\omega \tag{4.68}$$

where a_ω are nonnegative integers. Inserting (4.68) into (4.39) and using (4.56) and (4.53) with $s=1$, we obtain the complete energy spectrum

$$E = \sum_{\omega=2}^{N} \omega a_\omega + \frac{(N-1)}{2} + \frac{N(N-1)}{4} \left[1 + (1+4\lambda)^{1/2} \right]. \tag{4.69}$$

The solution of (4.64) for general $P_{\mu,1}$ is not known. Instead of finding the complete set of eigenfunctions (4.33) ($s=1$), we will find a linearly independent set of eigenfunctions which can be put into a one-to-one correspondence with the eigenfunctions (4.33); hence, these latter eigenfunctions will form a complete set. We do not need to explicitly solve (4.64) for $P_{\mu,1}$.

V. THE EIGENFUNCTIONS OF THE ONE-DIMENSIONAL PROBLEM WITH AN INVERSE SQUARE PAIR POTENTIAL

We obtain eigenfunctions characterized by four quantum numbers for the Hamiltonian

$$H = \frac{1}{2} \sum_{i=1}^{N} \left\{ -\frac{\partial^2}{\partial \xi_i^2} + \xi_i^2 \right\} + V_\lambda \tag{5.1}$$

where V_λ is given by (4.58). These eigenfunctions for any N reduce to the complete set of eigenfunctions for five particles. The raising and lowering operators

$$A_2^{\pm} = \frac{1}{4} \sum_{i=1}^{N} \left(\frac{\partial^2}{\partial \xi_i^2} + \xi_i^2 \right) - \frac{V_\lambda}{2} \pm \frac{1}{2} \left(-\frac{(N-1)}{2} - S \right) \tag{5.2a}$$

satisfy

$$[H, A_2^{\pm}] = \pm 2A_2^{\pm}, \quad A_2^{-} \equiv A_2, \tag{5.2b}$$

where

$$S \equiv \sum_{i=1}^{N} \xi_i \frac{\partial}{\partial \xi_i}. \tag{5.3}$$

If $F(\xi_1, \xi_2, \ldots, \xi_N)$ is a homogeneous function of degree f, then the operator S (5.3) has the property

$$[S, F(\xi_1, \xi_2, \ldots, \xi_N)] = fF(\xi_1, \xi_2, \ldots, \xi_N). \tag{5.4}$$

We first define the transformation $T \to \tilde{T}$ by

$$\tilde{T} = \exp(-\Gamma^2/2)\, T \exp(\Gamma^2/2) \tag{5.5}$$

where $\Gamma^2 = \sum_{i=1}^{N} \xi_i^2$; then, we apply it to H and A_2^{+}, respectively, obtaining

$$\tilde{H} = -\frac{1}{2} \sum_{i=1}^{N} \frac{\partial^2}{\partial \xi_i^2} - S - \frac{(N-1)}{2} + V_\lambda, \tag{5.6}$$

and

$$\tilde{A}_2^{+} = \frac{1}{4} \sum_{i=1}^{N} \frac{\partial^2}{\partial \xi_i^2} - \frac{V_\lambda}{2}, \tag{5.7}$$

where

$$[\tilde{H}, \tilde{A}_2^{+}] = 2\tilde{A}_2^{+}. \tag{5.8}$$

We discuss the eigenfunctions for the Hamiltonian (5.1) with $\lambda = 0$. Define the mutually commuting operators

$$\tilde{C}_p^{+} \equiv \sum_{i=1}^{N} \frac{\partial^p}{\partial \xi_i^p}, \quad N \geq p \geq 2; \tag{5.9}$$

from (5.6) (with $V_\lambda = 0$) and (5.4), we have

$$[(\tilde{H})_{\lambda=0}, \tilde{C}_p^{+}] = p\tilde{C}_p^{+}. \tag{5.10}$$

The eigenfunctions of $(H)_{\lambda=0}$ are

$$\phi_{a_2 a_3 \cdots a_N} = \prod_{\omega=2}^{N} (C_\omega^{+})^{a_\omega} \phi_0 \tag{5.11}$$

where ϕ_0 is the ground state; the a_ω ($2 \leq \omega \leq N$) are nonnegative integers and the C_ω^{+} are obtained from $\tilde{C}_\omega^{+}$ through (5.5). The operators C_p^{+} ($2 \leq p \leq N$) are deter-

mined within a multiplicative constant which determines the normalization and phase of the eigenfunctions. The energy spectrum is

$$E=\sum_{\omega=2}^{N}\omega a_\omega+\epsilon_0 \tag{5.12}$$

where ϵ_0 is the ground state energy. The eigenfunctions (5.11) are different within a given energy, i.e., if

$$\sum_{\omega=2}^{N}\omega a'_\omega=\sum_{\omega=2}^{N}\omega a_\omega \quad \text{and} \quad a_\omega\neq a'_\omega$$

for at least two values of ω, then

$$\psi_{a_2a_3\cdots a_N}\neq\psi_{a'_2a'_3\cdots a'_N}.$$

Also, the eigenfunctions are linearly independent.

The energy spectrum (5.12) is the same as the energy spectrum (4.69) except for the ground state. This observation led Perelomov[6] to search for a set of mutually commuting operators A_p^+ such that

$$[\tilde{H},A_p^+]=p\tilde{A}_p^+, \quad 2\leq p\leq N, \tag{5.13}$$

and

$$(\tilde{A}_p^+)_{\lambda=0}=\tilde{C}_p^+. \tag{5.14}$$

The eigenfunctions of H would then be given by

$$\psi_{a_2a_3\cdots a_N}=\prod_{\omega=2}^{N}(A_\omega^+)^{a_\omega}\psi_0 \tag{5.15}$$

with

$$\psi_0=\left(\prod_{i>j=1}^{N}\xi_{ij}\right)^{\Theta_1}\exp(-\Gamma^2/2), \tag{5.16}$$

where Θ_1 is given by (4.53) with $s=1$, the a_ω $(2\leq\omega\leq N)$ are nonnegative integers, and the A_ω^+ are obtained from $\tilde{A}_\omega^+$ through (5.5). Property (5.14) guarantees that these eigenfunctions are different within a given energy and linearly independent. Also, there is a one-to-one correspondence between the eigenfunctions (5.15) and the linearly independent eigenfunctions (4.33); thus, by finding the operators A_ω^+ we are indirectly solving Eq. (4.64) for $P_{\mu,1}$. Substituting A_2^+ into H, and the resulting expression for H into (5.13), we have

$$[\tilde{H},\tilde{A}_p^+]=-[\tilde{A}_2^+,\tilde{A}_p^+]-[S,\tilde{A}_p^+]=p\tilde{A}_2^+; \tag{5.17}$$

thus, *sufficient* conditions for $\tilde{A}_p$ to satisfy (5.17) are

$$\text{(i)}\quad\left[\frac{1}{2}\sum_{i=1}^{N}\frac{\partial^2}{\partial\xi_i^2}-V_\lambda,\tilde{A}_p^+\right]=0. \tag{5.18a}$$

and

$$\text{(ii)}\quad\tilde{A}_p^+ \text{ is homogeneous of degree } -p. \tag{5.18b}$$

Condition (ii) is crucial in proving the necessary condition

$$\text{(iii)}\quad[\tilde{A}_p^+,\tilde{A}_{p'}^+]=0 \tag{5.18c}$$

for all p' and p such that $2\leq p,\ p'\leq N$.

We define the transformation $T\to\tilde{\tilde{T}}$ by

$$\tilde{\tilde{T}}=\exp(\Gamma^2/2)\,T\exp(-\Gamma^2/2), \tag{5.19}$$

and then we apply it to H and A_2, respectively, obtaining

$$\tilde{\tilde{H}}=-\frac{1}{2}\sum_{i=1}^{N}\frac{\partial^2}{\partial\xi_i^2}+\frac{(N-1)}{2}+S+V_\lambda, \tag{5.20}$$

$$\tilde{\tilde{A}}_2=\frac{1}{4}\sum_{i=1}^{N}\frac{\partial^2}{\partial\xi_i^2}-\frac{V_\lambda}{2} \tag{5.21}$$

with

$$[\tilde{\tilde{H}},\tilde{\tilde{A}}_2]=-2\tilde{\tilde{A}}_2. \tag{5.22}$$

Thus in analogy to (5.17), we write

$$[\tilde{\tilde{H}},\tilde{\tilde{A}}_p]=-[\tilde{\tilde{A}}_2,\tilde{\tilde{A}}_p]+[S,\tilde{\tilde{A}}_p]=-p\tilde{\tilde{A}}_p, \quad 2\leq p\leq N,$$

leading to the sufficient conditions for the lowering operators $\tilde{\tilde{A}}_p$;

$$\text{(i')}\quad\left[\frac{1}{2}\sum_{i=1}^{N}\frac{\partial^2}{\partial\xi_i^2}-V_\lambda,\tilde{\tilde{A}}_p\right]=0, \tag{5.23a}$$

$$\text{(ii')}\quad\tilde{\tilde{A}}_p \text{ are homogeneous of degree } -p. \tag{5.23b}$$

Note that the form of $\tilde{\tilde{A}}_2$ (5.21) is exactly the same as $\tilde{A}_2^+$ (5.7), leading to exactly the same two sufficient conditions; thus, once we have determined A_p^+, we have also determined A_p. In fact, once we have $\tilde{A}_p^+$, we make the following prescription to obtain A_p^+ and A_p:

$$\text{terms } \frac{\partial^n}{\partial\xi_i^n} \text{ in } \tilde{A}_p^+\to\begin{cases}(-c_i^+)^n=\left(\dfrac{\partial}{\partial\xi_i}-\xi_i\right)^n \text{ in } A_p^+ & \text{(5.24a)}\\ c_i^n=\left(\dfrac{\partial}{\partial\xi_i}+\xi_i\right)^n \quad \text{in } A_p & \text{(5.24b)}\end{cases}$$

where

$$c_i^+=\left(\xi_i-\frac{\partial}{\partial\xi_i}\right), \quad c_i=\left(\xi_i+\frac{\partial}{\partial\xi_i}\right), \tag{5.24c}$$

$$[c_i,c_j]=[c_i^+,c_j^+]=0, \quad [c_i,c_j^+]=2(\delta_{ij}-1/N), \tag{5.24d}$$

and

$$H=\frac{1}{2}\sum_{i=1}^{N}c_i^+,c_j^++\frac{N-1}{2}+V_\lambda. \tag{5.24e}$$

Perelomov[6] used conditions (5.18) to find operators A_3^+ and A_4^+. He verified by direct calculation that $[A_3^+,A_4^+]=0$. Thus, he obtained the eigenfunctions of H (5.1) characterized by three quantum numbers, with[6]

$$A_3^+=-\sum_{i=1}^{N}(c_i^+)^3-\sum_{i=1}^{N}g_i^{(3)}c_i^+, \tag{5.25}$$

$$A_4^+=\sum_{i=1}^{N}(c_i^+)^4+\sum_{\substack{i,\omega\\ i\neq\omega}}^{N}g_{i\omega}^{(4)}c_i^+c_\omega^+-\sum_{i=1}^{N}g_{ii}^{(4)}(c_i^+)^2-\sum_{i=1}^{N}g_i^{(4)}c_i^++g_0^{(4)} \tag{5.26}$$

where

$$g_i^{(3)}=-3\lambda\sum_{\omega=1}^{N}\frac{1}{\xi_{i\omega}^2}, \tag{5.27}$$

$$g_{i\omega}^{(4)}=-2\lambda\frac{1}{\xi_{\omega i}^2}, \tag{5.28}$$

$$g_{ii}^{(4)}=g_{i^2}^{(4)}=-4\lambda\sum_{\omega=1}^{N}\frac{1}{\xi_{\omega i}^2}, \tag{5.29}$$

$$g_i^{(4)}=-4\lambda\sum_{\omega=1}^{N}\frac{1}{\xi_{\omega i}^3}, \tag{5.30}$$

$$g_0^{(4)} = -6\lambda \sum_{i,j}^{N} \frac{1}{\xi_{ij}^4} + \lambda^2 \left(2 \sum_{\substack{i,k,j \\ j\neq k}}^{N} \frac{1}{\xi_{ij}^2} \frac{1}{\xi_{ik}^2} - \sum_{i,j}^{N} \frac{1}{\xi_{ij}^4} \right). \tag{5.31}$$

In summations where terms like $1/(\xi_{i\omega})^n$, $n>0$, appear, we automatically assume $i \neq \omega$. The lowering operators are obtained by replacing $(-c_i^+)$ with c_i, in (5.25) and (5.26)

In what is to follow, we use the notation

$$\frac{\partial}{\partial \xi_i} \equiv \partial_i. \tag{5.32}$$

We will find an operator $\tilde{A}_5^+$ which satisfies the conditions (5.18); thus obtaining a series of eigenfunctions characterized by four quantum numbers.

We assume $\tilde{A}_5^+$ is homogeneous of degree -5 and has the following form:

$$\tilde{A}_5^+ = \sum_{i=1}^{N} \left(\partial_i^5 + g_{i^3}^{(5)} \partial_i^3 + 3 \sum_{\substack{m \\ i\neq m}}^{N} g_{i^2{}_m}^{(5)} \partial_i^2 \partial_m + g_{i^2}^{(5)} \partial_i^2 + g_i^{(5)} \partial_i \right) + g_0^{(5)}. \tag{5.33}$$

The problem is to find the functions $g_{i_1 i_2 i_3}^{(5)}$ in (5.33) such that (5.18a) is satisfied. We insert $\tilde{A}_5^+$ into (5.18a) and obtain equations for the functions $g_{i_1 i_2 i_3}$ by equating coefficients of like differential elements. We list these equations with the corresponding differential elements [we drop the superscript (5) on the functions $g_{i_1 i_2 i_3}^{(5)}$ for convenience]:

Iteration equations

$$\partial_i^4\colon \quad \partial_i g_{i^3} = -5\partial_i V_\lambda, \tag{5.34}$$

$$\partial_m \partial_i^3\colon \quad \partial_i g_{i^2{}_m} = -\tfrac{1}{3}\partial_m g_{i^3}, \tag{5.35}$$

$$\partial_i^3\colon \quad \partial_i g_{i^2} = -10\partial_i^2 V_\lambda - \frac{1}{2}\sum_{k=1}^{N} \partial_k^2 g_{i^3}, \tag{5.36}$$

$$\partial_i^2\colon \quad \partial_i g_i = -10\partial_i^3 V_\lambda - \frac{1}{2}\sum_{k=1}^{N} \partial_k^2 g_{i^2} - 3g_{i^3}\partial_i V_\lambda - 3\sum_{\substack{m=1 \\ m\neq i}}^{N} g_{i^2{}_m}\partial_m V_\lambda, \tag{5.37}$$

$$\partial_i\colon \quad \partial_i g_0 = -5\partial_i^4 V_\lambda - \frac{1}{2}\sum_{k=1}^{N} \partial_k^2 g_i - 2g_{i^2}\partial_i V_\lambda - 3g_{i^3}\partial_i^2 V_\lambda - 6\sum_{\substack{m=1 \\ m\neq i}}^{N} g_{i^2{}_m}\partial_i\partial_m V_\lambda - 3\sum_{\substack{m \\ m\neq i}}^{N} g_{m^2{}_i}\partial_m^2 V_\lambda; \tag{5.38}$$

Consistency equations

$$\partial_m^2 \partial_i^2\colon \quad \partial_m g_{i^2{}_m} + \partial_i g_{m^2{}_i} = 0, \quad i\neq m, \tag{5.39}$$

$$\partial_i^2 \partial_m\colon \quad \tfrac{3}{2}\partial_i^2 g_{i^2{}_m} + \tfrac{3}{2}\partial_m^2 g_{i^2{}_m} + \partial_m g_{i^2} = 0, \quad i\neq m, \tag{5.40}$$

$$\partial_i \partial_m\colon \quad \partial_m g_i + \partial_i g_m + 6g_{i^2{}_m}\partial_i V_\lambda + 6g_{m^2{}_i}\partial_m V_\lambda = 0, \quad i\neq m, \tag{5.41}$$

$$\sum_{i=1}^{N} g_i \partial_i V_\lambda + \sum_{i=1}^{N} g_{i^2}\partial_i^2 V_\lambda + 3\sum_{\substack{i,m \\ m\neq i}}^{N} g_{i^2{}_m}\partial_m \partial_i^2 V_\lambda + \sum_{i=1}^{N} g_{i^3}\partial_i^3 V_\lambda + \sum_{i=1}^{N} \partial_i^5 V_\lambda + \frac{1}{2}\sum_{k=1}^{N} \partial_k^2 g_0 = 0. \tag{5.42}$$

In deriving the above equations, we have assumed $\partial_k g_{i^2{}_m} = 0$ for $i\neq k\neq m$, $i\neq m$. This is justified in the subsequent analysis.

The five unknown functions $g_{i_1 i_2 i_3}^{(5)}$ are obtained by sequential solution of the first five equations (5.34)–(5.38), beginning with (5.34), i.e., the rhs of each equation is expressed in terms of known functions after the previous equations have been solved. We need only present a particular solution of Eqs. (5.34)–(5.38) which satisfy the four consistency equations (5.39)–(5.42); so in what is to follow we automatically choose a particular solution and then verify the consistency relations.

The solution of (5.34) is

$$g_{i^3} = -5\lambda \sum_{\omega=1}^{N} \frac{1}{\xi_{\omega i}^2}. \tag{5.43}$$

Substituting (5.43) into (5.35) and integrating, we have

$$g_{i^2{}_m} = -(5\lambda/3)(1/\xi_{mi}^2). \tag{5.44}$$

The function $g_{i^2{}_m}$ clearly satisfies the first consistency equation (5.39). Inserting (5.43) into (5.36) and integrating, we have

$$g_{i^2} = -10\lambda \sum_{\omega=1}^{N} \frac{1}{\xi_{\omega i}^3}. \tag{5.45}$$

It is also clear that g_{i^2} and $g_{i^2{}_m}$ satisfy the second consistency relation (5.40). The integration of (5.37) for g_i is described in Appendix D with the result

$$g_i = (-30+5\lambda)\lambda \sum_{\omega=1}^{N} \frac{1}{\xi_{\omega i}^4} - 20\lambda^2 \sum_{m=1}^{N} \frac{1}{\xi_{mi}} T_{m3(i)} + 15\lambda^2 \sum_{m=1}^{N} \frac{1}{\xi_{mi}^2} T_{m2(i)} \tag{5.46a}$$

where

$$T_{m\alpha(i)} \equiv \sum_{\substack{k=1 \\ k\neq i}}^{N} \frac{1}{(\xi_{km})^\alpha}. \tag{5.46b}$$

The verification of the consistency equation (5.41) is nontrivial and appears in Appendix E.

The functions g_i and g_0 have the general form

$$g_i = W_i\lambda + U_i\lambda^2, \tag{5.47}$$

$$g_0 = W_0\lambda + U_0\lambda^2, \tag{5.48}$$

where the functions W_0, W_i, U_0, and U_i are independent of λ; thus, (5.38) reduces to the two equations

$$\partial_i W_0 = \frac{-5}{\lambda}\partial_i^4 V_\lambda - \frac{1}{2}\sum_{k=1}^{N} \partial_k^2 W_i \tag{5.49}$$

and

$$\partial_i U_0 = -\frac{2}{\lambda^2} g_{i^2}\partial_i V_\lambda - \frac{3}{\lambda^2} g_{i^3}\partial_i^2 V_\lambda - \frac{6}{\lambda^2}\sum_{\substack{m=1 \\ m\neq i}}^{N} g_{i^2{}_m}\partial_i\partial_m V_\lambda - \frac{1}{2}\sum_{k=1}^{N} \partial_k^2 U_i - \frac{3}{\lambda^2}\sum_{\substack{m=1 \\ m\neq i}}^{N} g_{m^2{}_i}\partial_m^2 V_\lambda. \tag{5.50}$$

Inserting V_λ (4.58) and W_i from (5.46a) into (5.49), we have

$$\partial_i W_0 = 0; \tag{5.51}$$

but, g_0 is a homogeneous function of degree -5, hence

$$W_0 \equiv 0. \tag{5.52}$$

In Appendix F, we show

$$\partial_i U_0 = 0, \tag{5.53}$$

implying $U_0 \equiv 0$; therefore, we have

$$g_0 \equiv 0. \tag{5.54}$$

Finally, we verify the last consistency relation (5.42) in Appendix G.

We now discuss and prove the remaining condition, i.e.,

$$[A_p, A_{p'}] = [A_p^+, A_{p'}^+] = 0 \tag{5.55}$$

for $2 \leq p'$, $p' \leq 5$. It is sufficient to prove $[\tilde{A}_p^+, \tilde{A}_{p'}^+] = 0$. Perelomov proved (5.55) for $2 \leq p, p' \leq 4$ by explicitly calculating each commutator. We provide a proof of (5.55) based on general properties of the $g_{i_1 i_2 i_3}^{(p)}$ functions for $2 \leq p \leq 5$. This in turn gives us insight into the structure of the functions $g_{i_1 i_2 \cdots i_j}^{(p)}$ for $p > 5$

We assume

$$[\tilde{A}_p^+, \tilde{A}_{p'}^+] = K_{p+p'} \tag{5.56}$$

where $K_{p+p'}$ is a nonzero operator. The operator $K_{p+p'}$ is homogeneous of degree $-(p+p')$ since $\tilde{A}_p^+$ and $\tilde{A}_{p'}^+$ are homogeneous of degree $(-p)$ and $(-p)$, respectively. We also have

$$[\tilde{H}, K_{p+p'}] = (p+p') K_{p+p'}. \tag{5.57}$$

The commutation relation (5.57) follows from the Jacobi identity

$$[\tilde{H}, [\tilde{A}_p^+, \tilde{A}_{p'}^+]] + [\tilde{A}_p^+, [\tilde{A}_{p'}^+, \tilde{H}]] + [\tilde{A}_{p'}^+, [\tilde{H}, \tilde{A}_p^+]] = 0$$

and Eqs. (5.13) and (5.56). Inserting $\tilde{H}$ (5.6) into (5.57), we find that $K_{p+p'}$ must satisfy the relation

$$\left[\frac{1}{2} \sum_{j=1}^{N} \frac{\partial^2}{\partial \xi_j^2} - V_\lambda, K_{p+p'} \right] = 0. \tag{5.58}$$

We will prove the only solution to (5.58) is $K_{p+p'} \equiv 0$. We observe that some general properties of the functions $g_{i_1 i_2 i_3}^{(p)}$ $(2 \leq p \leq 5)$ are:

G_1: the subscripts of $g_{i_1 i_2 i_3}^{(p)}$ indicate the explicit functional dependence,

i.e., $g_{i_1 i_2 i_3}^{(p)} = g_{i_1 i_2 i_3}^{(p)}(\xi_{i_1}, \xi_{i_2}, \xi_{i_3})$;

G_2: the functions $g_{i_1 i_2 i_3}^{(p)}$ depend only on difference of the ξ_i variables, i.e., on $(\xi_i - \xi_j)$.

Property G_2 guarantees that we can treat the formal derivative $\partial_i \equiv \partial / \partial \xi_i$ as an ordinary derivative, i.e., if f is just a function of $\xi_{ij} \equiv \xi_i - \xi_j$, then

$$\partial_\omega f_{ij}(\xi_{ij}) = (\delta_{\omega i} - 1/N - \delta_{\omega j} + 1/N) \frac{\partial}{\partial \xi_{ij}} f_{ij}(\xi_{ij})$$

$$= (\delta_{\omega i} - \delta_{\omega j}) \frac{\partial}{\partial \xi_{ij}} f_{ij}(\xi_{ij}). \tag{5.59}$$

The terms $1/N$ from $\partial \xi_j / \partial \xi_i = \delta_{ij} - 1/N$ always cancel. Therefore, if $\partial_\omega f_{ij}(\xi_{ij}) = 0$, then $f_{ij}(\xi_{ij})$ is independent of ξ_ω. The functions $\tilde{A}_p^+ (2 \leq p \leq 5)$ have the general form

$$\tilde{A}_p^+ = \sum_{i=1}^{N} \partial_i^p + \sum_{j=1}^{p-2} \sum_{i_1 i_2 \cdots i_j}^{N} g_{i_1 i_2 \cdots i_j}^{(p)} \partial_{i_1 i_2 \cdots i_j} + g_0^{(p)}. \tag{5.60}$$

Inserting (5.60) into (5.56), we obtain

$$K_{p+p'} = \sum_{j=1}^{\tau} \sum_{i_1 i_2 \cdots i_j}^{N} h_{i_1 i_2 \cdots i_j} \partial_{i_1 i_2 \cdots i_j} + h_0, \quad \tau = p + p' - 3, \tag{5.61}$$

where $\tau \leq 6$ (the maximum value of τ occurs for $p=4$, $p'=5$); the functions $h_{i_1 i_2 \cdots i_j}$ are homogeneous of degree

$$(j - p - p') < 0 \text{ for all } j, \ 1 \leq j \leq \tau, \tag{5.62}$$

and the function h_0 is homogeneous of degree $-(p+p')$. The functions $h_{i_1 i_2 \cdots i_j}$ and h_0 are constructed from derivatives of the functions $g_{i_1 i_2 \cdots i_k}^{(p)}$ and $g_{i_1 i_2}^{(p')} \cdots i_l$; hence, these functions also possess properties G_1 and G_2 of the $g_{i_1 i_2 i_3}^{(p)}$ functions. Inserting (5.61) into (5.58), the term with the highest order differential leads to the following equation:

$$\sum_{i_1 i_2 \cdots i_\tau i_{\tau+1}}^{N} (\partial_{i_{\tau+1}} h_{i_1 i_2 \cdots i_\tau}) \partial_{i_1 i_2 \cdots i_\tau i_{\tau+1}} = 0. \tag{5.63}$$

The coefficient of each different differential must vanish for (5.58) to hold. We assume $h_{i_1 i_2 \cdots i_\tau}$ are nonzero. The maximum number of different subscripts in $h_{i_1 i_2 \cdots i_\tau}$ is four, thus $\tau < 4$ in (5.63). This arises from terms in $[\tilde{A}_p^+, \tilde{A}_{p'}^+]$ of the form

$$\sum_{i, \omega, m, s}^{N} [\, g_{i\omega}^{(4)} \partial_i \partial_\omega, g_{s^2 m}^{(5)} \partial_s^2 \partial_m]. \tag{5.64}$$

We discuss some necessary notation and definitions:
(i) Suppose that there are s i-subscripts and t m-subscripts $(i \neq m)$ in the function $h_{i_1 i_2 \cdots i_{s+t}}$, then we write

$$h_{i_1 i_2 \cdots i_{s+t}} \equiv h_{i^s m^t}. \tag{5.65}$$

The functions $h_{i_1 i_2 \cdots i_k}$ are coefficients of differentials; hence, the order of the indices is unimportant, i.e., $h_{i m^2} = h_{m i m} = h_{m^2 i}$, $(i \neq m)$ (although $h_{i m^2} \neq h_{i^2 m}$, in general).

(ii) The constant $C_{k_1, k_2 \cdots k_n}$ is defined by

$$C_{k_1, k_2, \ldots, k_n} = k! \Big/ \prod_{i=1}^{n} k_i! \tag{5.66}$$

where $k = \sum_{i=1}^{n} k_i$ and k_i are nonnegative integers.

(iii) The permutation operator $P_{i_1 i_2 \cdots i_k}$ is defined by

$$P_{i_1 i_2 \cdots i_k} h_{i_1 i_2 \cdots i_k, i_{k+1} \cdots i_s} = \sum h_{i_1 i_2 \cdots i_k, i_{k+1} \cdots i_s} \tag{5.67}$$

where the sum is over all cyclic permutations of $(i_1, i_2, \ldots, i_k)$. For example, $P_{mw}(\partial_m h_{i^\tau \omega}) = \partial_m h_{i^\tau \omega} + \partial_w h_{i^\tau m}$.

Equation (5.63) is equivalent to the following equations:

Iteration equations

$$\partial_i h_{i^\tau} = 0, \quad 1 \leq \tau \leq 4, \tag{5.68a}$$

$$C_{\tau-1,1} \partial_i h_{i^{\tau-1} m} = -\partial_m h_{i^\tau}, \quad 2 \leq \tau \leq 4, \quad i \neq m, \tag{5.68b}$$

$$C_{2,2} \partial_i h_{i^2 m^2} = -C_{3,1} \partial_m h_{i^3 m}, \quad \tau = 4, \quad i \neq m, \tag{5.68c}$$

$$C_{\tau-2,1,1} \partial_i h_{i^{\tau-2} m w} = -C_{\tau-1,1} P_{mw}(\partial_m h_{i^{\tau-1} w}), \ 3 \leq \tau \leq 4, \quad i \neq m \neq w, \ w \neq i, \tag{5.68d}$$

$$C_{1,1,1,1}\partial_i h_{imws} = -C_{2,1,1}P_{mws}(\partial_m h_{i^2sw}), \quad \tau=4, \quad i\neq m\neq w\neq s, \quad m\neq s\neq i\neq w; \tag{5.68e}$$

Consistency equations

$$\partial_i h_{im^2} + \partial_m h_{i^2m} = 0, \quad \tau=3, \quad i\neq m, \tag{5.69a}$$

$$C_{2,1,1}P_{im}(\partial_i h_{m^2wi}) + \partial_w h_{m^2i^2} = 0, \quad \tau=4, \quad i\neq m\neq w, \quad i\neq w, \tag{5.69b}$$

$$P_{i_1i_2\cdots i_{\tau+1}}(\partial_{i_{\tau+1}} h_{i_1i_2\cdots i_\tau}) = 0, \quad 1\leq\tau\leq 4, \tag{5.69c}$$

where all the subscripts $i_1, i_2\cdots i_{\tau+1}$ are different in (5.69c). Equation (5.68a) implies $h_{i\tau}$ is independent of ξ_i; but, since $h_{i\tau}$ is only an explicit function of ξ_i, it must be a constant. All the functions $h_{i_1i_2\cdots i_\tau}$ are homogeneous of a degree less than zero; hence, the constant function $h_{i\tau}$ is zero. Equation (5.68b) can now be rewritten

$$\partial_i h_{i\tau-1m} = 0, \quad 1\leq\tau-1\leq 3, \quad i\neq m.$$

We advance the same argument; thus $h_{i\tau-1}=0$. Note that the rhs of each of the Eqs. (5.68) is expressed in terms of functions $h_{i_1i_2\cdots i_\tau}$ from the preceding equations; thus with our argument, each succeeding function is zero if the previous ones are zero, i.e., the system of Eqs. (5.68) collapses to $h_{i_1i_2\cdots i_\tau}\equiv 0$, a contradiction. Thus, the coefficient of the highest order differential in $K_{p+p'}$ (5.61) is now $h_{i_1i_2\cdots i_{\tau-1}}$, with the analog of (5.63), i.e.,

$$\sum_{i_1i_2\cdots i_\tau}^{N} (\partial_\tau h_{i_1i_2\cdots i_{\tau-1}})\,\partial_{i_1i_2\cdots i_\tau} = 0.$$

Since $h_{i_1i_2\cdots i_{\tau-1}}$ is homogeneous of degree less than zero, the argument that we used to prove $h_{i_1i_2\cdots i_\tau}\equiv 0$ applies here; thus $h_{i_1i_2\cdots i_{\tau-1}}\equiv 0$. All the h-functions in the expansion of (5.61) are homogeneous of degree less than zero; so, we proceed in a similiar manner to show $h_{i_1i_2\cdots i_j}=0$, $1\leq j\leq\tau$. Equation (5.58) cannot be satisfied by the function h_0 alone, thus $h_0=0$. The consistency equations (5.69) are satisfied by this solution. Hence, we have a contradiction; therefore $K_{p+p'}=0$. The $\tilde{A}_p^+$ functions are nonzero and satisfy Eq. (5.58) because the coefficient of the highest order differential in $\tilde{A}_p^+$ is homogeneous of degree zero. We can relax a number of our assumptions. The structure of the Eqs. (5.68) remains the same independently of τ, i.e., if none of the $h_{i_1i_2\cdots i_j}$ functions are homogeneous of degree zero, then the equations collapse to $h_{i_1i_2\cdots i_j}=0$, $i\leq j\leq\tau$. If we assume the operators $\tilde{A}_p^+$, $2\leq p\leq N$, have the form (5.60) and the corresponding $g^{(p)}_{i_1i_2\cdots i_j}$ functions have properties G_1 and G_2, then we can show $[\tilde{A}_p^+, \tilde{A}_{p'}^+]=0$ for $1\leq p, p'\leq N$. We conjecture that $\tilde{A}_p^+$ with the postulated structure exists for arbitrary p. The equations for the $g^{(p)}_{i_1i_2\cdots i_j}$ functions for arbitrary p have the same structure as the Eqs. (5.34)–(5.38) for $p=5$, i.e., the rhs of each equation is expressed in terms of known functions after the previous equations have been solved. We reserve this analysis for another paper.

SUMMARY

In this paper, we have used algebraic techniques (group theory, raising and lowering operators) to obtain a series of N-body eigenfunctions for systems of N interacting particles in any dimension. We provide an equation [see (4.16)] that relates the ground state eigenfunction to the interaction potential; we obtain a series of N-body eigenfunctions in terms of an energy raising operator acting on the ground state. These eigenfunctions follow immediately from the $SU(1,1)$ group structure. We indicated when the complete energy spectrum is linear.

For the particular example of N identical particles in one dimension interacting with a quadratic pair potential and an inverse square pair potential, we derived a series of N-body eigenfunctions characterized by four quantum numbers. If we let $N=5$ in these eigenfunctions, they form the complete set of eigenfunctions for five particles. We have indicated how all the N-body eigenfunctions may be generated. This one-dimensional system has generated much interest in statistical mechanics,[14] especially with the Hamiltonian (5.1) without the harmonic term.[15,16]

One-dimensional systems with Hamiltonian (1.1) have also generated interest in scattering theory.[5,17]

ACKNOWLEDGMENT

The author would like to express his sincere thanks to Dr. Max Dresden for his encouragement and suggestions during the course of the present work and for critically reading this paper.

APPENDIX A: DERIVATION OF (3.18) FOR THE CASIMIR OPERATOR

Inserting the Eqs. (3.12) into the defining equation for the Casimir operator (2.5), yields

$$Q_{N,s} = \frac{1}{4\hbar^2}\{-D_{N,s}^2 + 2(H_{N,s}C_{N,s} + C_{N,s}H_{N,s})\}. \tag{A1}$$

Inserting the Eqs. (3.13) into (A1), we have

$$Q_{N,s} = \frac{1}{16\hbar^2}\left[-\left(\sum_{i=1}^{t}\{X_i,P_i\}_+^2 + \sum_{k>i=1}^{t}\{X_k,P_k\}_+\{X_i,P_i\}_+\right) + 2\left(\sum_{i=1}^{t}\{X_i^2,P_i^2\}_+ + \sum_{k>i=1}^{t}\{X_k^2,P_i^2\}_+\right) + 8mr^2V_\lambda\right]. \tag{A2}$$

Substitute the lhs of each of the indentities

$$\sum_{i=1}^{t}(2\{X_i^2,P_i^2\}_+ - \{X_i,P_i\}_+^2) = -3t\hbar^2, \tag{A3}$$

$$\sum_{k>i=1}^{t} 2\{X_i,P_i\}_+\{X_k,P_k\}_+ = 4\sum_{k>i=1}^{t}(X_iP_iP_kX_k + X_kP_kP_iX_i) - t(t-1)\hbar^2 \tag{A4}$$

into (A2) to obtain (3.18).

APPENDIX B: NORMALIZATION OF RADIAL EIGENFUNCTIONS

The eigenfunctions

$$R_a(\rho^2) = N_0\rho^{E_0-t/2}\exp(-\rho^2/2)\,L_a^{E_0-1}(\rho^2) \tag{B1}$$

are normalized over the volume $\Pi_{i=1}^t y_i \propto \rho^{t-1}\,d\rho\,d\Omega$ where Ω represents the angular variables. Thus, we have

$$1 = N_0^2\int_0^\infty (\rho^2)^{E_0-t/2}\exp(-\rho^2)[L_a^{E_0-1}(\rho^2)]^2(\rho^{t-1}\,d\rho). \tag{B2}$$

Inserting $X=\rho^2$ into (B2), we have

$$\frac{2}{N_0^2}=\int_0^\infty X^{E_0-1}\exp(-X)[L_a^{E_0-1}(X)]^2\,dX. \tag{B3}$$

The rhs of (B3) is equal[13] to $\Gamma(E_0+a)/a!$; therefore, $N_0=[2a!/\Gamma(E_0+a)]^{1/2}$. The eigenfunctions (B1) are also orthogonal, since

$$0=\int_0^\infty X^{E_0-1}\exp(-X)\,L_a^{E_0-1}(x)\,L_b^{E_0-1}(x)\,dX \quad \text{for } a\neq b.$$

APPENDIX C: ANGULAR EIGENFUNCTIONS

The angular equation [see (3.22)] is given by

$$\left(\frac{-\rho^2}{2}\sum_{i=1}^{t}\frac{\partial^2}{\partial y_i^2}+\rho^2V_\lambda\right)\frac{KP_{\mu,s}}{\rho^\beta}=\delta\,\frac{KP_{\mu,s}}{\rho^\beta} \tag{C1}$$

with [see (3.18)]

$$q=\frac{\delta}{2}+\frac{t(t-4)}{16}. \tag{C2}$$

In (C1), we have used $\rho^2\Sigma_{i=1}^t\partial^2/\partial y_i^2$ instead of $L_{N,s}$ [see (3.23)] since the angular eigenfunctions are independent of ρ. Expanding (C1), we have

$$\frac{1}{\rho^{\beta-2}}\left(-\sum_{i=1}^{t}\frac{\partial^2}{\partial y_i^2}+V_\lambda\right)KP_{\mu,s}=\frac{\rho^2}{2}\,\chi(y_1,y_2,\ldots,y_t)+\delta\,\frac{KP_{\mu,s}}{\rho^\beta} \tag{C3}$$

where

$$\chi(y_1,y_2,\ldots,y_t)=(KP_{\mu,s})\,\rho^{t-1}\frac{\partial}{\partial\rho}\rho^{t-1}\frac{\partial}{\partial\rho}\rho^{-\beta}+2\sum_{i=1}^{t}\frac{\partial}{\partial y_i}(KP_{\mu,s})\frac{\partial}{\partial y_i}\rho^{-\beta}. \tag{C4}$$

After substituting $\partial\rho^{-\beta}/\partial y_i=y_i/\rho^{\beta+2}$ into (C4) and expanding, we have

$$\chi(y_1,y_2,\ldots,y_t)=[-\beta(t-2)-\beta^2]\,\frac{KP_{\mu,s}}{\rho^{\beta+2}};$$

thus,

$$\delta=\tfrac{1}{2}[\beta(t-2)+\beta^2], \tag{C5}$$

and the rhs of (C3) is zero. After inserting (C5) into (C2), we obtain (4.38). Finally, expanding the lhs of (C3) and using (4.16), we obtain (4.37).

APPENDIX D: INTEGRATION OF (5.37) FOR $g_i^{(5)}$

Inserting (5.43)–(5.45) into (5.37), we have

$$\partial_i g_i=(-5+5\lambda/6)\,\partial_i^3V+10\lambda^2\sum_{m=1}^{N}\frac{1}{\xi_{mi}^2}\sum_{\substack{w=1\\ w\neq i\\ m\neq i}}^{N}\frac{1}{\xi_{wm}^3}+30\lambda^2\sum_{\substack{m,w\\ m\neq w}}^{N}\frac{1}{\xi_{wi}^2}\frac{1}{\xi_{mi}^3}. \tag{D1}$$

In this and the following appendices, we will use the identity

$$\frac{1}{\xi_{wi}^p\,\xi_{mi}^s}=(-1)^p\sum_{k=1}^{N}\binom{k+p-2}{k-1}\frac{1}{(\xi_{mi})^{s+1-k}(\xi_{mw})^{k+p-1}}+(-1)^s\sum_{k=1}^{p}\binom{k+s-2}{k-1}\frac{1}{(\xi_{wi})^{p+1-k}(\xi_{wm})^{k+s-1}},\quad w\neq m, \tag{D2}$$

where $\binom{s}{t}=s!/(s-t)!t!$. The identity (D2) is derived by induction or by repeated differentiation of

$$\frac{1}{\xi_{wi}\xi_{mi}}=-\frac{1}{\xi_{mi}\xi_{mw}}-\frac{1}{\xi_{wi}\xi_{wm}} \tag{D3}$$

with respect to ξ_w and ξ_m.

We first substitute (D2) with $p=2$ and $s=3$ into (D1), and then integrate (D1) to obtain (5.46).

APPENDIX E: VERIFICATION OF THE CONSISTENCY EQUATION (5.41)

Inserting (5.47) into (5.41), we obtain the two equations,

$$\partial_i W_m+\partial_m W_i=0,\quad i\neq m, \tag{E1}$$

$$\partial_i U_m+\partial_m U_i+\frac{6}{\lambda^2}\,g_{i^2m}\partial_i V_\lambda+\frac{6}{\lambda^2}\,g_{m^2i}\partial_m V_\lambda=0,\quad i\neq m. \tag{E2}$$

Equation (E1) is clearly satisfied by W_i from (5.46a). Differentiating U_i [from (5.46a)] with respect to ξ_m, we have

$$\partial_m U_i=-20\,\frac{1}{\xi_{mi}^5}+20\sum_{\substack{w=1\\ w\neq i}}^{N}\frac{1}{\xi_{mi}^2\xi_{wm}^3}-30\sum_{\substack{w=1\\ w\neq i\\ w\neq m\neq i}}^{N}T_{miw} \tag{E3}$$

where

$$T_{miw}=A_{miw}+2B_{miw}+2C_{miw}+D_{miw}+E_{miw} \tag{E4}$$

with

$$A_{miw}=\frac{1}{\xi_{mw}^3\xi_{mi}^2},\quad B_{miw}=\frac{1}{\xi_{iw}\xi_{mw}^4},\quad C_{miw}=\frac{1}{\xi_{mw}^4\xi_{mi}},$$

$$D_{miw}=\frac{1}{\xi_{mw}^2\xi_{mi}^3},\quad E_{miw}=\frac{1}{\xi_{iw}^2\xi_{mw}^3}.$$

In addition, from (4.58) and (5.44), we have

$$6g_{m^2i}\partial_m V_\lambda=20\lambda^2\left(\frac{1}{\xi_{mi}^5}-\sum_{\substack{w=1\\ w\neq i}}^{N}\frac{1}{\xi_{mi}^2\xi_{wm}^3}\right). \tag{E5}$$

Inserting (E3) and (E5) into (E2), we have

$$\sum_{\substack{w=1\\ w\neq i\\ w\neq m\neq i}}^{N}(T_{miw}+T_{imw})=0. \tag{E6}$$

In order to verify (E6), we must prove three indentities:

$$0=-(B_{miw}+B_{imw}+C_{miw}+C_{imw}+E_{miw}+E_{imw}), \tag{E7}$$

$$0=-A_{imw}+A_{miw}+3B_{miw}+3C_{miw}+2E_{miw}+E_{imw}, \tag{E8}$$

$$0=2A_{imw}+3B_{imw}+3C_{imw}+D_{imw}+D_{miw}+E_{imw}. \tag{E9}$$

If we add equations (E7), (E8), and (E9) together, we obtain $T_{miw}+T_{imw}=0$, thus verifying (E6). Equation (E7) is a special case of (D2) with $p=4$ and $s=1$; and, Eq. (E8) is a special case of (D2) with $p=3$ and $s=2$. Equation (E9) follows from (E8) with the following prescription:

$$(E8)\rightarrow(E9)\ \text{ as }\ (m,i,w)\rightarrow(w,m,i).$$

APPENDIX F: DERIVATION OF (5.53)

Differentiating (E2) with respect ξ_m, we have

$$\lambda^2\partial_m^2 U_i = -\lambda^2\partial_i\partial_m U_m - 6g_{i^2m}\partial_i\partial_m V_\lambda - 6g_{m^2i}\partial_m^2 V_\lambda - 6\partial_m g_{i^2m}\partial_i V_\lambda - 6\partial_m g_{m^2i}\partial_m V_\lambda; \tag{F1}$$

and thus, with $g_{i^2m} = g_{m^2i}$ [see (5.44) and (5.39)], we have

$$\lambda^2\partial_m^2 U_i = \lambda^2\partial_i^2 U_m - \lambda^2(\partial_i\partial_m U_m - \partial_m\partial_i U_i) - 6g_{m^2i}(\partial_m^2 V_\lambda - \partial_i^2 V_\lambda) - 12\partial_m g_{i^2m}(\partial_i V_\lambda + \partial_m V_\lambda). \tag{F2}$$

Summing (F2) from $m=1$ to $m=N$, $m\neq i$, and adding $\partial_i^2 U_i$ to both sides yields

$$\frac{\lambda^2}{3}\sum_{m=1}^{N}\partial_m^2 U_i = \frac{\lambda^2}{2}\partial_i^2\left(\sum_{m=1}^{N} U_m - \lambda^2\partial_i\right)\left(\sum_{m=1}^{N}\partial_m U_m\right) - 3\sum_{\substack{m=1\\m\neq i}}^{N} g_{m^2i}\partial_m^2 V_\lambda + g_{i^3}\partial_i^2 V_\lambda - 6\sum_{\substack{m=1\\m\neq i}}^{N}\partial_m g_{i^2m}\partial_m V_\lambda + 2g_{i^2}\partial_i V_\lambda, \tag{F3}$$

where we have used the relations [see (5.43)–(5.45)]

$$g_{i^3} = 3\sum_{\substack{m=1\\m\neq i}}^{N} g_{m^2i} \text{ and } -\tfrac{1}{3}g_{i^2} = \sum_{\substack{m=1\\m\neq i}}^{N}\partial_m g_{i^2m} \tag{F4}$$

and

$$\sum_{\substack{m=1\\m\neq i}}^{N}(-\partial_i\partial_m U_m + \partial_m\partial_i U_i) = -\partial_i\left(\sum_{m=1}^{N}\partial_m U_m - \partial_i U_i\right) + \sum_{\substack{m=1\\m\neq i}}^{N}\partial_m(\partial_i U_i) = -\partial_i\left(\sum_{m=1}^{N}\partial_m U_m\right) + \partial_i^2 U_i - \partial_i^2 U_i = -\partial_i\left(\sum_{m=1}^{N}\partial_m U_m\right),$$

since

$$\sum_{\substack{m=1\\m\neq i}}^{N}\partial_m = -\partial_i. \tag{F5}$$

Using U_i from (D1), we have

$$\frac{1}{\lambda^2}\sum_{i=1}^{N}\partial_i U_i = 20\sum_{i,w}^{N}\frac{1}{\xi_{wi}^5} + 40\sum_{\substack{i,m,w\\w\neq m}}^{N}\frac{1}{\xi_{wi}^2\xi_{mi}^3}. \tag{F6a}$$

The first term on the rhs of (F6a) is clearly zero, and Perelomov[6] proved that the second term on the rhs of (F6) is zero; thus

$$\partial_i\left(\sum_{m=1}^{N}\partial_m U_m\right) = 0. \tag{F6b}$$

We now calculate

$$\partial_i^2\left(\sum_{\substack{m=1\\m\neq i}}^{N} U_m\right).$$

Summing (E2) from $m=1$ to $m=N$, $m\neq i$, and using (F4) and (F5), yields

$$\frac{\lambda^2}{2}\partial_i\left(\sum_{m=1}^{N} U_m\right) = \lambda^2\partial_i U_i - g_{i^3}\partial_i V_\lambda - 3\sum_{m=1}^{N} g_{m^2i}\partial_m V_\lambda. \tag{F7}$$

Inserting $\partial_i U_i$ [from (5.37) into (F7)], we have

$$\frac{\lambda^2}{2}\partial_i\left(\sum_{m=1}^{N} U_m\right) = -4g_{i^3}\partial_i V_\lambda - 6\sum_{\substack{m=1\\m\neq i}}^{N} g_{m^2i}\partial_m V_\lambda. \tag{F8}$$

We differentiate (F8) with respect to ξ_i and use the equation [see (5.43) and (5.45)] $\partial_i g_{i^3} = g_{i^2}$ to obtain

$$\frac{\lambda^2}{2}\partial_i^2\left(\sum_{m=1}^{N} U_m\right) = -4g_{i^3}\partial_i^2 V_\lambda - 6\sum_{m=1}^{N} g_{m^2i}\partial_m\partial_i V_\lambda - 4g_{i^2}\partial_i V_\lambda + 6\sum_{\substack{m=1\\i\neq m}}^{N}\partial_m g_{m^2i}\partial_m V_\lambda \tag{F9}$$

where we have used (5.39). Inserting (F9) and (F6b) into (F3), and the resulting expression into (5.38), we obtain (5.53).

APPENDIX G: VERIFICATION OF THE CONSISTENCY EQUATION (5.42)

Equation (5.42) with $g_0 = 0$ and (5.47) reduces to the three equations:

$$A = \sum_{i=1}^{N}\partial_i^5 V_\lambda = 0, \tag{G1}$$

$$B = \lambda\sum_{i=1}^{N} W_i\partial_i V_\lambda + \sum_{i=1}^{N} g_{i^2}\partial_i^2 V_\lambda + 3\sum_{\substack{i,m\\m\neq i}}^{N} g_{i^2m}\partial_m\partial_i^2 V_\lambda + \sum_{i=1}^{N} g_{i^3}\partial_i^3 V_\lambda = 0, \tag{G2}$$

$$C = \sum_{i=1}^{N} U_i\partial_i V_\lambda = 0. \tag{G3}$$

The expression (G1) is verified easily, since

$$\sum_{i,m}^{N}\frac{1}{\xi_{mi}^7} = 0. \tag{G4}$$

Inserting (4.58), (5.43)–(5.45), and W_i [from (5.46a)] into (G2) and using (G4), we have

$$B = -5!\sum_{\substack{i,w,k\\w\neq k}}^{N}\left(\frac{1}{\xi_{wi}^4\xi_{ki}^3} + \frac{1}{\xi_{wi}^2\xi_{ki}^5}\right). \tag{G5}$$

Perelomov proved[6] the identity

$$T = \sum_{\substack{i,k,w\\w\neq k}}^{N}\frac{1}{\xi_{wi}^2\xi_{ki}^3} = 0. \tag{G6}$$

One can show that

$$B = -5\sum_{e=1}^{N}\partial_e^2 T; \tag{G7}$$

thus, from (G6), we have $B=0$, i.e., (G2) is verified.

We insert (4.58) and U_i [from (5.46a)] into (G3), expand, and use $B=0$ (G5) to obtain

$$C = C_1 + C_2 + C_3 + C_4, \tag{G8}$$

where

$$C_1 = 40\lambda\sum_{\substack{i,k,w,s\\k\neq s\neq w\neq i}}^{N}\frac{1}{\xi_{ki}\xi_{kw}^3\xi_{si}^2}, \tag{G9}$$

$$C_2 = -40\lambda \sum_{i,k,s}^{N} \frac{1}{\xi_{ki}\xi_{sk}^3\xi_{si}^3}, \tag{G10}$$

$$C_3 = 30\lambda \sum_{\substack{i,k,w,s \\ k\neq s\neq w\neq i}}^{N} \frac{1}{\xi_{ki}^2\xi_{kw}^2\xi_{si}^3}, \tag{G11}$$

$$C_4 = 30\lambda \sum_{i,k,s}^{N} \frac{1}{\xi_{ki}^2\xi_{sk}^2\xi_{si}^3}. \tag{G12}$$

The expression C_1 (G9) is zero since the dummy variable change $(k,i,w,s) \to (i,k,s,w)$ yields $C_1 = -C_1$.

We expand the summation in C_2 as follows:

$$\begin{aligned}\sum_{i,k,s}^{N} \frac{1}{\xi_{ki}\xi_{sk}^3\xi_{si}^3} &= \sum_{i,k,s}^{N} \frac{1}{\xi_{ki}\xi_{sk}^4\xi_{si}^2} - \sum_{i,k,s}^{N} \frac{1}{\xi_{sk}^4\xi_{si}^3} \\ &= \sum_{i,k,s}^{N} \frac{1}{\xi_{ki}\xi_{sk}^5\xi_{si}} - \sum_{\substack{i,k,s \\ i\neq k}}^{N} \left(\frac{1}{\xi_{sk}^5\xi_{si}^2} + \frac{1}{\xi_{sk}^4\xi_{si}^3}\right) \\ &= \sum_{\substack{i,s,k \\ i\neq s}}^{N} \frac{1}{\xi_{ki}\xi_{sk}^6} - \sum_{\substack{i,k,s \\ i\neq k}}^{N} \frac{1}{\xi_{si}\xi_{sk}^6} \\ &\quad - \sum_{\substack{i,k,s \\ i\neq k}}^{N} \left(\frac{1}{\xi_{sk}^5\xi_{si}^2} + \frac{1}{\xi_{sk}^4\xi_{si}^3}\right)\end{aligned} \tag{G13}$$

where we have repeatedly used the identity

$$\left(\frac{1}{\xi_{ki}} - \frac{1}{\xi_{si}}\right) = \frac{\xi_{sk}}{\xi_{si}\xi_{ki}}. \tag{G14}$$

The last summation on the rhs of (G13) is zero since $B=0$ (G5).

The dummy variable change $(i,k,s) \to (i,s,k)$ in the first term on the rhs of (G13) yields

$$\sum_{\substack{i,k,s \\ i\neq s}}^{N} \frac{1}{\xi_{ki}\xi_{sk}^6} = \sum_{\substack{i,k,s \\ i\neq k}}^{N} \frac{1}{\xi_{si}\xi_{sk}^6};$$

thus, $C_2 = 0$. Using the identity (G14), we expand (G12) as follows:

$$C_4 = 30\lambda \left(\sum_{i,k,s}^{N} \frac{1}{\xi_{ki}^2\xi_{sk}^3\xi_{si}^2} - \sum_{i,k,s}^{N} \frac{1}{\xi_{ki}\xi_{sk}^3\xi_{si}^3}\right). \tag{G15}$$

The second term on the rhs of (G15) is zero since $C_2 = 0$.

Finally, using (G14) once again in the expansion of the first term on the rhs of C_4, we have

$$C_4 = 30\lambda \sum_{i,k,s}^{N} \left(\frac{1}{\xi_{ki}^2\xi_{sk}^4\xi_{si}} - \sum_{i,k,s}^{N} \frac{1}{\xi_{ki}\xi_{sk}^4\xi_{si}^2}\right). \tag{G16}$$

The dummy variable change $(i,k,s) \to (i,s,k)$ in the first term on the rhs of (G16) yields

$$\sum_{i,k,s}^{N} \frac{1}{\xi_{ki}^2\xi_{sk}^4\xi_{si}} = \sum_{i,k,s}^{N} \frac{1}{\xi_{ki}\xi_{sk}^4\xi_{si}^2}; \tag{G17}$$

thus $C_4 = 0$.

We now prove $C_3 = 0$. We first rewrite C_3 (G11), as follows:

$$C_3 = (30\lambda) \sum_{w>i>s>k=1}^{N} (T_{wisk} + T_{wiks} + T_{iwsk} + T_{iwks} + T_{swik}) \tag{G18}$$

where

$$T_{wisk} = \frac{1}{\xi_{kw}^2\xi_{si}^3}\left(\frac{1}{\xi_{ki}^2} - \frac{1}{\xi_{ks}^2} + \frac{1}{\xi_{wi}^2} - \frac{1}{\xi_{ws}^2}\right). \tag{G19}$$

We define

$$\alpha \equiv \xi_{kw}, \quad \beta \equiv \xi_{sw}, \quad \gamma = \xi_{iw}, \tag{G20}$$

so that

$$\xi_{ki} = \alpha - \gamma, \quad \xi_{si} = \beta - \gamma, \quad \xi_{ks} = \alpha - \beta, \tag{G21}$$

and thus,

$$T_{wisk} = \frac{1}{M}\{(\beta + \gamma - 2\alpha)\gamma^2\beta^2 + (\alpha - \gamma)^2(\alpha - \beta)^2(\beta + \gamma)\}, \tag{G22}$$

$$T_{wiks} = \frac{1}{M}\{(\alpha + \gamma - 2\beta)\gamma^2\alpha^2 + (\beta - \gamma)^2(\alpha - \beta)^2(\alpha + \gamma)\}, \tag{G23}$$

$$T_{wski} = \frac{1}{M}\{(\alpha + \beta - 2\gamma)\beta^2\alpha^2 + (\beta - \gamma)^2(\alpha - \gamma)^2(\alpha + \beta)\}, \tag{G24}$$

$$T_{iwsk} = \frac{1}{M}\{(\beta - 2\alpha)\gamma^2(\beta - \gamma)^2 + \alpha^2(\alpha - \beta)^2(\beta - 2\gamma)\}, \tag{G25}$$

$$T_{iwks} = \frac{1}{M}\{(\alpha - 2\beta)\gamma^2(\alpha - \gamma)^2 + \beta^2(\alpha - \beta)^2(\alpha - 2\gamma)\}, \tag{G26}$$

$$T_{swik} = \frac{1}{M}\{(\gamma - 2\alpha)\beta^2(\beta - \gamma)^2 + (\alpha - \gamma)^2\alpha^2(\gamma - 2\beta)\}, \tag{G27}$$

where

$$M = \alpha^2\beta^2\gamma^2(\alpha - \gamma)^2(\beta - \gamma)^2(\alpha - \beta)^2. \tag{G28}$$

We expand each of the Eqs. (G22)–(G27), and add to obtain

$$(T_{wisk} + T_{wiks} + T_{wski} + T_{iwsk} + T_{iwks} + T_{swik}) = 0; \tag{G29}$$

and thus $C_3 = 0$. Therefore, C (G8) is zero.

[1] The Lie algebras of $SU(1,1)$, $\overline{SU(1,1)}$, $SO(2,1)$, $SP(2,R)$ and $O(2,1)$ are isomorphic, i.e., the generators of each Lie algebra satisfy the same commutation relations. We are concerned with representations of the Lie algebra $\overline{SU(1,1)}$ that in general are not representations of the Lie algebra of $SU(1,1)$ (see Sec. II).

[2] F. Calogero, C. Marchioro, J. Math. Phys. **14**, 182 (1973).

[3] F. Calogero, J. Math. Phys. **10**, 2191 (1969).

[4] F. Calogero, J. Math. Phys. **10**, 2197 (1969).

[5] F. Calogero, J. Math. Phys. **12**, 419 (1971).

[6] A.M. Perelomov, Teor. Mat. Fiz. **6**, 364 (1971).

[7] A nonexhaustive set of references on $SU(1,1)$ and $\overline{SU(1,1)}$ are: A. O. Barut, C. Fronsdal, Proc. Roy. Soc. (Lond.) A **287**, 532 (1965); H. Ui, Ann. Phys. **49**, 69 (1968); V. Bargmann, Ann. Math. **48**, 568 (1947); L.C. Biedenharn, J. Nuyts, and N. Straumann, Ann. Inst. Henri Poincaré, III **13** (1965).

[8] G. Burdet and M. Perrin, Nuovo Cimento Lett. **4**, 651 (1972). They were concerned with the group of all permissible transformations that leave invariant the free Schrödinger equation for a system of N interacting particles. This group is called the "conformal extended Galilei group" (or Schrödinger group or Hagen group) and contains the Galilei group and two additional generators which form with the Hamiltonian a group isomorphic to the group $SU(1,1)$.

[9] If we replace the harmonic potential in (1.1) by $(m\omega^2/4)\Sigma_{i=1}^{N}\mathbf{r}_i^2$

and eliminate the center of mass motion, then we still obtain (3.1), but with $\alpha = \omega/\sqrt{2}$.
[10]L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Pergamon, New York, 1965), Sec. 35.
[11]A. Erdelyi, W. Magnus, W. Oberhettinger, and F. G. Tricomi, *High Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. 2, Chap. XI.
[12]M. Grynberg and Z. Koba, Ann. Phys. (N.Y.) **26**, 418 (1964).
[13]I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals Series and Products* (Academic, New York, 1965), Vol. II.
[14]C. Marchioro and E. Presutti, Nuovo Cimento Lett. **4**, 488 (1970).
[15]B. Sutherland, J. Math. Phys. **12**, 246 (1971); J. Math. Phys. **12**, 251 (1971); Phys. Rev. A **4**, 2019 (1971); Phys. Rev. A **5**, 1372 (1972).
[16]The eigenfunctions of this latter Hamiltonian are obtained once the functions $P_{i,\mu}$ are known (see, Ref. 5). The functions $P_{i,\mu}$ can be obtained from the raising and lowering operators corresponding to Hamiltonian (5.1) (see, Ref. 6).
[17]J. Wolfes, Ann. Phys. **85**, 454 (1974).

Cylindrically symmetric stationary beam of electromagnetic radiation

A. Banerjee

Department of Physics, Jadavpur University, Calcutta-32, India
(Received 5 June 1974; revised manuscript received 9 September 1974)

An exact solution of Einstein–Maxwell equations corresponding to an infinite stationary beam of electromagnetic radiation with the energy flux along the axial direction of a cylindrically symmetric system is obtained. Explicit forms of the electromagnetic field tensors are given. Further, the timelike and null geodesics in the region of the beam are also discussed.

INTRODUCTION

In a recent communication Nackoney[1] obtained an exact solution of Einstein field equations corresponding to an infinite stationary light beam (or neutrino flux) of circular cross section where the energy flux is along the axial direction, that is, in the direction of Z axis of a cylindrically symmetric system. The metric was dependent on the radial coordinate alone. He also discussed the timelike and null trajectories in the interior as well as in the exterior regions of such a beam. The present work has a similar approach for a beam of electromagnetic radiation with the difference that the solutions obtained in this case satisfy both the Einstein and the Maxwell equations. The dual restrictions of Maxwell's equations, however, allow only a definite set of functions for the quantity σ, which is interpreted as a measure of the energy density of the beam in the chosen coordinate system, while it is arbitrary in the case considered in the geometrical limit by Nackoney. Again, Nackoney was able to present an exterior solution which continuously fits at the boundary of the beam with its interior metric, but in our case no such matching is possible because of the vanishing electromagnetic field tensors outside the system. The 4-velocity of the beam or the null propagation vector is shown to satisfy the necessary conditions that it is geodetic and shearfree, so that according to Robinson's[2] theorem there must be an associated electromagnetic field satisfying Maxwell's equations.

Explicit forms for the electromagnetic field tensors are given and the timelike and null trajectories in the region of the beam are also discussed.

SOLUTIONS OF EINSTEIN-MAXWELL EQUATIONS

We consider a cylindrically symmetric stationary beam with the energy flow along the axial direction, that is, the positive Z direction. Let this be the null direction, so that the propagation vector K^μ is given by

$$K^\mu = \delta^\mu_3. \tag{1}$$

Thus $g_{33}=0$ and the line element can be written as

$$ds^2 = f\,dt^2 + 2a\,dz\,dt - m\,dr^2 - r^2\,d\theta^2, \tag{2}$$

where f, a, m are functions of the radial coordinate r alone. Einstein's field equations can be written as

$$R_{\mu\nu} = -\sigma K_\mu K_\nu. \tag{3}$$

σ may be interpreted as a measure of the energy density in the chosen coordinate system because of the fact that in relativity the energy density may be given by $T_{\mu\nu}v^\mu v^\nu = (8\pi)^{-1}\sigma(K_\mu v^\mu)^2$, v^μ being a suitable time like vector and is thus coordinate dependent. Field equations (3) give rise to two classes.

Case 1. $\partial a/\partial r \neq 0$: One gets in such a case from the field equations $a = Br^{-4}$, B being the integration constant. This is a singular solution for $r=0$ and is thus rejected.

Case 2. $\partial a/\partial r = 0$: In this case one obtains $a=\text{const}$ and also $m=\text{const}$. Choosing both of them equal to 1 for simplicity, the line element reduces to a very simple form

$$ds^2 = f\,dt^2 + 2\,dz\,dt - dr^2 - r^2\,d\theta^2, \tag{4}$$

which is the same as has been used by Nackoney. t, r, θ, z, are represented by x^0, x^1, x^2, x^3, respectively. Now in view of (1) and (3), $K_0=1$ and

$$K^\mu{}_{;\mu} = 0, \tag{5}$$

$$K_{\mu;\nu}K^\nu = 0, \tag{6}$$

and

$$K_{(\mu;\nu)}K^{\mu;\nu} = 0. \tag{7}$$

From (4), (5), (6) one can conclude that the beam represents a plane wave beam with its rays being geodetic and shearfree. Moreover, since $K_{[\mu;\nu]}=0$, the rays may be said to be twistfree also and K_μ is orthogonal to a null hypersurface. Now since the conditions (6) and (7) are satisfied, the problem remains to find some suitable electromagnetic field tensors F_μ, constructed with the help of the null propagation vector K_μ in (1). $F_{\mu\nu}$, however, satisfies Maxwell's equations. The most general form of $F_{\mu\nu}$ subject to the change of amplitude and polarization is given by (cf. Raychaudhuri and Dutta[3])

$$F^{\mu\nu} = \sigma^{1/2}[K^\mu(a^\nu\cos\beta + b^\nu\sin\beta) - K^\nu(a^\mu\cos\beta + b^\mu\sin\beta)] \tag{8}$$

where a_μ, b_μ are unit spacelike vectors orthogonal to K_μ. a_μ and b_μ are themselves mutually orthogonal. $\sigma^{1/2}$, β are the amplitude and polarization factors, so that

$$a^\mu a_\mu = b^\mu b_\mu = -1, \quad a^\mu b_\mu = 0. \tag{9}$$

Since $K^3=1$, we may assume that the only nonvanishing components of a^μ and b^μ are a^1 and b^2, respectively, and thus in view of (9) the electromagnetic field tensors are

$$F^{31} = \sigma^{1/2}\cos\beta,$$

Copyright © 1975 American Institute of Physics

$$F^{32} = (\sigma^{1/2}, r)\sin\beta. \tag{10}$$

Since $F^{\mu\nu}$ satisfies Maxwell's equations and also because σ is a function of r alone,

$$\beta_{,1} = \beta_{,3} = 0$$

and

$$(\sigma^{1/2}r)_{,1}/\sigma^{1/2} = \beta_{,2} = A, \tag{11}$$

A being a constant quantity. From (11)

$$\beta = (A\theta + \psi(t)), \tag{12}$$

where $\psi(t)$ is an arbitrary function of time (cf. Banerjee[4]). In this the electromagnetic field tensor of the null field for the given metric is not unique (Witten[5]). Finally, one can write explicitly the $F^{\mu\nu}$ components

$$\begin{aligned} F^{31} &= \sigma^{1/2}\cos(A\theta + \psi(t)), \\ F^{32} &= (\sigma^{1/2}/r)\sin(A\theta + \psi(t)). \end{aligned} \tag{13}$$

The only restriction on σ is through the relation (11), that is

$$\frac{1}{\sigma^{1/2}}\frac{d}{dr}(\sigma^{1/2}r) = A. \tag{14}$$

Equation (14) gives the explicit form of σ as

$$\sigma = Dr^C \tag{15}$$

where D and C are constants. The constant C is related to A by $C = 2(A-1)$. It is evident that the above restrictions on the form of σ is solely due to Maxwell's equations and is independent of the gravitational field. It can be shown purely from classical considerations also that if in the absence of the gravitational field if one proceeds to solve the two-dimensional Laplace equation in the (r, θ) plane of the cylindrical coordinate system suppressing the time and axial dependencies with the electric field vector written as

$$E_r = \sigma^{1/2}(r)\sin\beta(r,\theta), \quad E_\theta = \sigma^{1/2}(r)\cos\beta(r,\theta),$$

one arrives at the same functional forms for σ and β except for the arbitrary function of time $\psi(t)$. In the special case by a suitable choice of the function $\psi(t)$ one can interpret the system as being due to a monochromatic beam.

Now in view of (3) the Einstein field equations may be written as

$$\tfrac{1}{2}f'' + \tfrac{1}{2}f'/r = \sigma. \tag{16}$$

Using (15) the solutions of the field equations (16) is then given by

$$f = (ar^b + 1). \tag{17}$$

One of the integration constants is chosen to be unity, so that at $r = 0$, $f = 1$. The constant a, b are related to C, D, of (15) by

$$a = 2D/(C+2)^2 \quad \text{and} \quad b = (C+2) \tag{18}$$

The constant b must be greater than or equal to 2. Otherwise, for $b < 2$, $c < 0$ and $\sigma \to \infty$ as $r \to 0$, which means a singularity along the axis of the beam.

Contrary to the case of Nackoney, one can not, however, find any suitable metric in the region where $\sigma = 0$, which satisfies all the continuity conditions at any possible boundary because there the electromagnetic field is vanishing. Thus the beam may be said to extend over all space.

HELICAL AND CIRCULAR TRAJECTORIES OF TEST PARTICLES AND RAYS IN THE REGION OF THE BEAM

It is of some interest to discuss the timelike and null trajectories in the interior of the beam under consideration. The treatment and the method of approach are similar to those of Nackoney. The results are, however, different to some extent. Using the same symbols as in the paper of Nackoney, viz., $v^0 = A$, $v^2 = h/r^2$, $v^1 = 0$, $v^3 = Q$ for a trajectory with rotation at a fixed radial distance from the axis of the beam, A, h, and Q being constants along a definite trajectory. The results for the trajectories in the r-z plane are, however, not different from those in Nackoney's paper, including the point that rays travelling parallel to the beam are neither deflected nor retarded (cf. Tolman, Ehresfest, and Podolsky[6]). We therefore omit such cases and discuss only the trajectories with rotation.

Timelike orbits

For a timelike geodesic

$$h^2/r^3 = A^2/2 \cdot f'$$

and

$$fA^2 - \tfrac{1}{2}A^2 rf' + 2QA = 1. \tag{19}$$

In view of the solution (17) the relation (19) can be written as

$$\tfrac{1}{2}ar^b(b-2) = [1 - 1/A^2 + 2Q/A]. \tag{19'}$$

The orientations of the timelike orbits depend on the magnitude of A.

Case I. $b > 2$: Two different situations occur according to whether $A \leq 1$ or $A > 1$. (i) When $A \leq 1$, $Q > 0$ from (19') and the test particle has a timelike trajectory spiraling in the forward direction. (ii) When $A > 1$, the orbits may be circular ($Q = 0$) or they may be helical with inclinations either in the forward or in the backward direction ($Q > 0$ or $Q < 0$).

Case II. $b = 2$: There are three distinct situations. (i) When $A < 1$, $Q > 0$ and the orbits are spiraling in the forward direction. (ii) When $A = 1$, $Q = 0$ and timelike orbits are circular. (iii) When $A > 1$, $Q < 0$ and the orbits are spiraling in the backward direction. So for $A < 1$ no circular timelike orbit can exist anywhere in the region of the beam. For circular timelike geodesics corresponding to $Q = 0$ the following two relations hold in view of (19).

$$A^2 = 2/(2f - rf')$$

and

$$h^2 = r^3 f'/(2f - rf'). \tag{20}$$

Null orbits

The conditions controlling the orientations of the null orbits are, however, independent of the value of A, that is, the value of v^0. Rather, they will depend on the magnitudes of a and b or in other words on the strength of the beam. For such a null trajectory of a test ray

$$fA^2 - \tfrac{1}{2}A^2 rf' + 2QA = 0. \tag{21}$$

Again there are two distinctly different situations occurring for nonzero Q. (i) When $ar^b(b-2) > 2$, $Q > 0$ and the rays while rotating about the axis proceed in the forward direction, that is in the positive Z direction. (ii) When $ar^b(b-2) < 2$, $Q < 0$ and the rays spiral in the opposite direction. This includes the case $b = 2$. The equation for null geodesics along circular orbits ($Q = 0$) is, however, explicitely given by

$$ar^b(b-2) = 2. \tag{22}$$

This relation in fact restricts the value of the constant b, so that one can obtain null circular orbits in the finite distance from the axis. For $b = 2$ one can not get such orbits in the finite regions.

Stability of circular orbits

Stability conditions for the circular timelike or null orbits are analyzed exactly in accordance with the method being used by Nackoney. We given here only the results relevant in our case without going into the details, in view of (17), (19) and (20)

$$\frac{dh^2}{dr} = -\frac{4r^3}{(2f - rf')^2}\frac{ab}{2} r^{(b-2)}[ar^b(b-2) - (b+2)]. \tag{23}$$

Also,

$$\frac{d^2\Delta r}{ds^2} = -\frac{\Delta r}{2r^3 f}(2f - rf')\frac{dh^2}{dr}.$$

We know that for timelike circular orbits to exist ($2f - rf'$) must be greater than zero or in other words

$$ar^b(b-2) < 2. \tag{24}$$

One can immediately conclude from (30) that $dh^2/dr > 0$ and in consequence $d^2\Delta r/ds^2 < 0$ for Δr positive. The real timelike circular orbits are therefore stable.

ACKNOWLEDGMENT

The author is extremely grateful to the referee for his very valuable comments on this paper. He would also like to thank Professor A.K. Raychaudhuri of Presidency College, Calcutta for useful discussions.

[1]R.W. Nackoney, J. Math. Phys. **14**, 1239 (1973).
[2]I. Robinson, *Reports on Royaumont Conference* (1959).
[3]A.K. Raychaudhuri and A.K. Dutta, J. Math. Phys. (to be published).
[4]A. Banerjee, J. Math. Phys. **11**, 51 (1970).
[5]L. Witten, in *Gravitation: An Introduction to Current Research*, edited by L. Witten (Wiley, New York, 1962).
[6]R.C. Tolman, P. Ehrenfest, and B. Podolsky, Phys. Rev. **37**, 602 (1931).

Models of Hamiltonian several-particle interactions

Richard Arens

Mathematics Department, University of California, Los Angeles, California
(Received 7 January 1975)

A method of Hamiltonian surfaces proposed recently as a semiclassical completely Hamiltonian way of defining the motions of N particles (with finitely many degress of freedom) is shown to be mathematically feasible for $N \geq 2$. The interactions are invariant under the Poincaré group. The possibility of describing each motion by N world lines, the same for all observers, is dispensed with. The usual commutation relations are retained.

1. INTRODUCTION

The zero-interaction theorem of Currie, Jordan, and Sudarshan[1] for Hamiltonian, Lorentz–Einstein–Minkowski relativistically invariant two-particle systems have shown that not all conditions expected of such systems can be retained. Thus the earlier concepts[2,3] in which the world line condition is not met, deserve reexamination.

Some have preferred to retain that condition, and abandon another.[4] In Ref. 5 we have suggested a way to replace the world line condition by a space–time construct (the "complete Hamiltonian surface") with the intent of preserving some of that connection with the idea of N particles which had been provided by the world line condition. Because of the novelty of the concept, we undertook to show in Ref. 6 that our definition did allow systems of nonzero interaction. It dealt with just two particles.

The present paper shows the same thing for N particles, and provides examples which are different from that of Ref. 6 for $N=2$.

The method of complete Hamiltonian surfaces can be used to provide examples of systems of the kind considered in Refs. 2,3. Indeed, it can provide examples not expected by Foldy[3] who conjectured that the role played by "internal" variables in the work of Ref. 2 would be the rule in general. (Incidentally, we show that there is a purely Lie algebra theoretic definition, not only for external, but also internal variables.)

We show that a weakened form of the axioms for Hamiltonian surfaces will still provide *local* or infinitesimal completely Hamiltonian systems[7] of the type considered by Refs. 2,3. On the other hand, we do not believe although we cannot prove, that each of the examples constructed in Refs. 2,3 can be obtained by a "local" Hamiltonian surface. As indicated already, our intent is to *restrict* by imposing what is hoped to be a reasonable restriction, namely being derivable from such a surface.

We present a contact-transformation method which transforms the complete Hamiltonian surface of a system into that of another, and show that in one case it changes a zero-interaction system into a nonzero interaction system. In some applications of this method, merely *local* systems are obtained, but as remarked above, such systems are also interesting.

We give only a few examples of our method, simply to illustrate various possibilities. None of these examples has an aesthetic appeal remotely approaching that of the inverse square law in Galilei–Newton relativity. However, we have taken care to make interparticle accelerations to die off in some sense as the particles separate, to keep the particles essentially interchangable, and to provide for the possibility of inserting a parameter which can be adjusted to give zero interaction.

2. AN N-PARTICLE SYSTEM

According to our definition,[7] to construct a model of a completely Hamiltonian N particle interaction we have to define a surface $\mathcal{S}_N$ of dimension $2.4N - N = 7N$ in the $8N$-dimensional cotangent bundle $T_1(\mathbb{R}^{4N})$ over the product space of N copies of four-dimensional space–time $\mathbb{R}^4$. This surface $\mathcal{S}_N$ has to be a component (see Ref. 5, H – 1; we discuss this later in this section) of the locus defined by N equations $F_1 = 0, \ldots, F_N = 0$ where these functions have mutually vanishing Poisson brackets: $\{F_\alpha, F_\beta\} = 0$. If we want invariance under the Poincaré group, then $\mathcal{S}_N$ must be invariant under the action of that group in $T_1(\mathbb{R}^{4N})$, which action is induced by the action of the group in (each) $\mathbb{R}^4$. This much is easily assured if the F_α commute and are themselves invariant under the group. But there is a further requirement (see Ref. 5, H – 2): that these equations $F_1 = \cdots = F_N = 0$ can be solved without singularities for the N time components of the N momentum "4"-vectors. This necessity of regular solvability presents the only analytic difficulty. We refer to the functions $H_1, \ldots, H_N$ representing the negatives of these time components in terms of the other $7N$ coordinates in $T_1(\mathbb{R}^{4N})$ as the *partial Hamiltonians*, since the Hamiltonian of the N particle system is obtained by taking their sum, and then setting the N time coordinates equal to zero.

We might remark that in a sense H_α is associated with the αth particle, but that F_α is not necessarily associated with one particle more than any other.

We introduce some notation. The momentum 4-vector of the αth particle will be denoted by p_α. Its time component we can denote by $-H_\alpha$. Its space component is a vector $\mathbf{p}_\alpha$ in 3-space. The sum $p_1 + \cdots + p_N$ shall be denoted by p, and the sum of the $\mathbf{p}_\alpha$ shall be denoted by $\mathbf{p}$. For the space–time coordinate of the αth particle we will use x_α (instead of the more classical q_α). For the temporal component of x_α we use t_α. For the space component, or position in 3-space, we use $\mathbf{x}_\alpha$. We denote $x_\alpha - x_\beta$ by $z_{\alpha\beta}$, $\mathbf{x}_\alpha - \mathbf{x}_\beta$ by $\mathbf{z}_{\alpha\beta}$, etc.

If a and b are 4-vectors, we use $a \cdot b$ to stand for $a^4 b^4$

 Copyright © 1975 American Institute of Physics

$-a^1b^1-a^2b^2-a^3b^3$. If **a** and **b** are 3-vectors, we use $\mathbf{a}\cdot\mathbf{b}$ for $a^1b^1+a^2b^2+a^3b^3$.

We make explicit something which was assumed implicitly in Refs. 5,6. *The surface $\mathcal{S}_N$ must lie in that part of $T_1(\mathbb{R}^{4N})$ on which all $z_{\alpha\beta}$ are spacelike*, i.e., $z_{\alpha\beta}\cdot z_{\alpha\beta}<0$. This is a very acceptable requirement because we are usually interested only in values of $x_1,\ldots,x_N$ for which the t_α are near to zero, or at any rate, close together.

In order to ensure Poincaré invariance, we define the F_α in terms of the Poincaré invariants of $p_1,\ldots,p_N$, $x_1,\ldots,x_N$. The linear combinations of the Lorentz invariants $p_\alpha\cdot p_\beta$, $p_\alpha\cdot x_\beta$, $x_\alpha\cdot x_\beta$ form a Lie algebra L_N and *this* Lie *algebra has a structure depending only on N* (i.e., independent of the dimension and the metric of space–time.) This can be seen from the following. Let $q_1,\ldots,q_4$ be any four from the set $p_1,\ldots,x_N$. Then $\{q_1\cdot q_2,\ q_3\cdot q_4\}$ is the sum of four terms $T_{13}+T_{23}+T_{14}+T_{24}$ where in T_{ij} the q_i and q_j are treated as constants. Thus to verify our statement about the nondependence on the space–time structure, we have only to verify it for $\{a\cdot p_\alpha,\ b\cdot x_\beta\}$ where a and b are constant. Here $a\cdot p_\alpha=$ = (the sum) $a^ig_{ij}p_\alpha^i=a^ip_{\alpha i}$ where g_{ij} is the metric. Thus $\{a\cdot p_\alpha,\ b\cdot x_\alpha\}=a^i\{p_{\alpha i},x_\beta^j\}g_{jk}b^k$ and $\{p_{\alpha i},x_\beta^j\}=\delta_{\alpha\beta}\delta_i^j$ essentially by definition. Thus the result is $a^ig_{jk}b^k=a\cdot b$, and this agrees with what one would get in the one-dimensional case. This shows the nondependence asserted.

We introduce the first $N-1$ of the N functions F_α associated with the Hamiltonian surface $\mathcal{S}_N$ for the interaction we are about to present. Consider the system of equations

$$p\cdot(p_\alpha-p_\beta)=0,\quad 1\le\alpha,\beta\le N.$$

This system is in fact equivalent to the system

$$F_\alpha\equiv p\cdot(Np_\alpha-p)=0,\quad 1\le\alpha\le N-1. \tag{2.1}$$

One certainly has $\{F_\alpha,F_\beta\}=0$ here because these F's contain only p's. This will not be true of F_N, to which we turn.

For $1\le\alpha,\ \beta\le N$ let

$$g(\alpha,\beta)=\begin{vmatrix} p\cdot p & p\cdot z_{\alpha\beta}\\ p\cdot z_{\alpha\beta} & z_{\alpha\beta}\cdot z_{\alpha\beta}\end{vmatrix},$$

and let

$$F_N=p_1\cdot p_1+\cdots+p_N\cdot p_N-\varphi \tag{2.2}$$

where

$$\frac{1}{\varphi}=1-\frac{(p\cdot p)^2}{\sum_{\alpha<\beta}g(\alpha,\beta)}. \tag{2.3}$$

Theorem: $\{F_\alpha,F_N\}=0$ *for* $\alpha=1,2,\ldots,N-1$.

Proof: It is obvious that in one-dimensional space–time we have $\{p\cdot p_\alpha,\ a\cdot z_{\beta\gamma}\}=p\cdot a\,(\delta_{\alpha\beta}-\delta_{\alpha\gamma})$, because p itself is constant as far as the translation-invariant $z_{\beta\gamma}$ is concerned. Armed with this result, we prove that $\{p\cdot p_\alpha,\ g(\beta,\gamma)\}=0$. For the moment, denote $\{p\cdot p_\alpha,\ f\}=f'$. Then

$$g(\beta,\gamma)'=\begin{vmatrix}(p\cdot p)' & p\cdot z_{\beta\gamma}\\ (p\cdot z_{\beta\gamma})' & z_{\beta\gamma}\cdot z_{\beta\gamma}\end{vmatrix}+\begin{vmatrix} p\cdot p & (p\cdot z_{\beta\gamma})'\\ p\cdot z_{\beta\gamma} & (z_{\beta\gamma}\cdot z_{\beta\gamma})'\end{vmatrix}.$$

Now $(p\cdot p)'=0$, $(p\cdot z_{\beta\gamma})'=\epsilon p\cdot p$, $(z_{\beta\gamma}\cdot z_{\beta\gamma})'=2\epsilon p\cdot z_{\beta\gamma}$ where $\epsilon=\delta_{\alpha\beta}-\delta_{\alpha\gamma}$, and this makes $g(\beta,\gamma)'=0$. The theorem follows easily from this. A shorter "proof" is to regard space–time as two-dimensional. Then the gram matrix is the square of the determinant D of p and $z_{\beta\gamma}$. Since $z'_{\beta\gamma}=\epsilon p$, we obtain

$$D'=\begin{vmatrix}p\\ z_{\beta\gamma}\end{vmatrix}'=\begin{vmatrix}p\\ z'_{\beta\gamma}\end{vmatrix}=0,\quad\text{whence } g'=0.$$

It remains to express this system of N equations: $F_1=\cdots=F_N=0$ in terms of the "unknowns" $H_1,\ldots,H_N$ and to verify that they have a global solution.

Theorem: These N equations do have a regular solution. The Hamiltonian $H=H_1+\cdots+H_N$ satisfies an equation of sixth degree with coefficients depending on the euclidean invariants of the p_α and $\mathbf{x}_\alpha$.

Proof: Recall the notation $p_\alpha=(-H_\alpha,\mathbf{p}_\alpha)$, to which we adjoin $p=(-H,\mathbf{p})$. When this notation is introduced into Eq. (2.1), they take the form

$$H_\alpha=\frac{H}{N}+\frac{\mathbf{p}\cdot(N\mathbf{p}_\alpha-\mathbf{p})}{NH} \tag{2.4}$$

for $\alpha=1,2,\ldots,N-1$. This equation holds also for $\alpha=N$. It appears from this that if a nonzero solution H can be found for $F_N=0$, then the "unknowns" $H_1,\ldots,H_N$ will become "known." We must show that such an H exists.

Inserting these expressions for the H_α into $F_N=0$ gives this the form

$$\sum_{\alpha=1}^{N}\left[\left(\frac{H}{N}+\frac{\mathbf{p}\cdot(N\mathbf{p}_\alpha-\mathbf{p})}{NH}\right)^2-\mathbf{p}_\alpha\cdot\mathbf{p}_\alpha\right]=\varphi, \tag{2.5}$$

because the expression between the braces is indeed $p_\alpha\cdot p_\alpha$.

The expression $g(H)$ on the left of (2.5) has the form

$$\frac{H^2}{N}+B+\frac{C}{H^2},$$

where C, at any rate, is nonnegative. To give an idea of the behavior of $g(H)$ as a function of H, we will establish that $g(H)\le 0$ at $H=b\equiv(\mathbf{p}\cdot\mathbf{p})^{1/2}$. Take first the case $\mathbf{p}\neq 0$. Then for this value of H, we have

$$g(b)=\sum\left[\left(\frac{\mathbf{p}\cdot(N\mathbf{p}_\alpha)}{Nb}\right)^2-\mathbf{p}_\alpha\cdot\mathbf{p}_\alpha\right]$$
$$=\sum\left(\frac{(\mathbf{p}_\alpha\cdot\mathbf{p}_\alpha)^2-(\mathbf{p}\cdot\mathbf{p})(\mathbf{p}_\alpha\cdot\mathbf{p}_\alpha)}{b^2}\right).$$

Each term here is at most 0 by the inequality of Cauchy and Schwarz. If $\mathbf{p}=0$ then

$$g(H)=\frac{H^2}{N}-\sum\mathbf{p}_\alpha\cdot\mathbf{p}_\alpha$$

and so $g(b)=g(0)\le 0$, as before.

For $b>0$, the graph of g is as in Fig. 1. If $b=0$ the graph is a parabola passing under or through the origin.

We want now to show that this curve $y=g(H)$ cuts the curve $y=\Phi(H)$ where $\Phi(H)$ is the result of using (2.4) to express φ entirely in terms of H. We want also to make sure that this point of intersection varies smoothly as the parameters t_α, $\mathbf{p}_\alpha$, and $\mathbf{x}_\alpha$ change. Certain inequali-

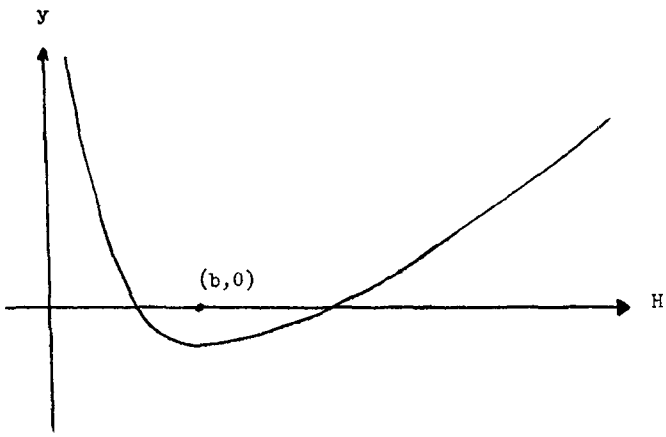


FIG. 1.

ties established in Ref. 6, (5.1) enable us to verify these properties, in the following manner.

Let $f(H)=f_{\alpha\beta}(H)$ be the result of expressing $g(\alpha,\beta)/(p\cdot p)$ in terms of H, using (2.4). Specifically,

$$f(H)=\frac{-ZH^2-2atH-a^2+b^2(Z-t^2)}{(H^2-b^2)^2} \tag{2.6}$$

where $b^2=\mathbf{p}\cdot\mathbf{p}$ as before, $Z=\mathbf{z}_{\alpha\beta}\cdot\mathbf{z}_{\alpha\beta}$, $t=t_\alpha-t_\beta$, and $a=\mathbf{p}\cdot\mathbf{z}_{\alpha\beta}$. We know two things from Ref. 6, (5.7). First, that $f(H)$ is negative for $b<H<\infty$ and increases monotonically from $-\infty$ to 0. Second, there is an inequality

$$\frac{Z}{(H+b)^2}\leq -f(H)\leq\frac{Z}{H^2-b^2}. \tag{2.6'}$$

The inequality won't concern us at this point, but the monotonic behavior shows that

$$\frac{-\sum_{\alpha<\beta}g(\alpha,\beta)}{(p\cdot p)^2}\equiv g(H)$$

decreases monotonically from $+\infty$ to 0 on $b<H<\infty$. Observing that

$$1/\Phi(H)=1+1/g(H),$$

we deduce that $\Phi(H)$ varies from $+\infty$ to 1 on the interval (b,∞). This guarantees us an intersection of these curves. The dependence is obviously smooth.

We define the surface S_N as the surface given parametrically as follows. Pick any values of the t_α, $\mathbf{p}_\alpha$, $\mathbf{x}_\alpha$. Calculate H from (2.5) and then the time components of the momenta from (2.4). This gives us a generic point of S_N in $T_1(\mathbb{R}^{4N})$. (Thus S_N is one connected component of the submanifold defined by $F_1=\cdots=F_N=0$.)

A third condition [Ref. 5, (H – 3)] has also to be verified in order that this example qualify as having a Hamiltonian surface S_N. This condition essentially postulates that solutions to the relevant Hamilton canonical equations exist with arbitrary initial conditions and for all time. With $\varphi=0$ this is surely true. Therefore, it will also be true for other φ close to 0 *in a suitable sense*, but rather than go into this now we will just assume that we have a φ which makes (H – 3) hold.

In a formal way, the example can be embellished with coupling constants. One can take (for $\alpha<N$)

$$F_\alpha=p\cdot\left(\frac{p_\alpha}{m_\alpha}-\frac{\mathbf{p}}{m}\right)$$

where $m_1,\ldots,m_N$ are nonzero numbers and $m=m_1+\cdots+m_N$ is also not zero. One can take

$$F_N=\sum_{\alpha,\beta}F_{\alpha\beta}(p_\alpha\cdot p_\beta)-\varphi \tag{2.7}$$

where $F_{\alpha\beta}$ is a positive-definite matrix of numbers.

Here one could define φ by

$$\frac{1}{\varphi}=K-\frac{(p\cdot p)^2}{\sum_{\alpha<\beta}k_{\alpha\beta}g(\alpha,\beta)} \tag{2.8}$$

where K and $k_{\alpha\beta}$ are positive constants; or by

$$\frac{1}{\varphi}K-(p\cdot p)^2\sum_{\alpha<\beta}\frac{k_{\alpha\beta}}{g(\alpha,\beta)} \tag{2.9}$$

where $K>0$, $k_{\alpha\beta}\geq 0$. [Keep in mind that $g(\alpha,\beta)<0$.]

Any two usable φ's can be combined as in

$$\varphi=\varphi_1+\varphi_2 \quad\text{or}\quad \frac{1}{\varphi}=1/\varphi_1+1/\varphi_2$$

to give another, since it is only the general behavior described in Fig. 1 that matters.

3. DISCUSSION OF THE HAMILTONIAN FOR THE EXAMPLES OF SEC. 2

The Hamiltonian of any system of the kind introduced in Ref. 5 is obtained from the sum of the partial Hamiltonians by equating to each other the N time coordinates (thereby yielding a Hamiltonian H independent of time.) An implicit equation for H can be obtained in this way from (2.5) where, however, we use the more general (2.8) instead of (2.3). Retaining the notation $b^2=\mathbf{p}\cdot\mathbf{p}$, we introduce $Z=\sum_{\alpha<\beta}k_{\alpha\beta}(\mathbf{z}_{\alpha\beta}\cdot\mathbf{z}_{\alpha\beta})$, and $a\geq 0$ by $a^2=\sum_{\alpha<\beta}k_{\alpha\beta}(\mathbf{p}\cdot\mathbf{z}_{\alpha\beta})^2$. Then

$$\frac{H^2}{N}-\sum_{\alpha=1}^{N}\mathbf{p}_\alpha\cdot\mathbf{p}_\alpha+\frac{1}{N^2H^2}\sum_{\alpha=1}^{N}[\mathbf{p}\cdot(N\mathbf{p}_\alpha-\mathbf{p})]^2$$

$$=\frac{1}{K+(H^2-b^2)^2/[Z(H^2-b^2)+a^2]} \tag{3.1}$$

The analysis of Sec. 2 assures us that there is a real root greater than b and no more than one such root.

In this equation H^2 satisfies a fourth-degree polynomial for every N.

If instead, we use φ as given by (2.9), then

$$\frac{H^2}{N}-\sum_{\alpha=1}^{N}\mathbf{p}_\alpha\cdot\mathbf{p}_\alpha+\frac{1}{N^2H^2}\sum_{\alpha=\beta}[\mathbf{p}\cdot(N\mathbf{p}_\alpha-\mathbf{p})]^2$$

$$=\frac{1}{K+(H^2-b^2)^2\sum_{\alpha<\beta}k_{\alpha\beta}/[Z_{\alpha\beta}(H^2-b^2)+a_{\alpha\beta}^2]} \tag{3.2}$$

where $Z_{\alpha\beta}=\mathbf{z}_{\alpha\beta}\cdot\mathbf{z}_{\alpha\beta}$ and $a_{\alpha\beta}^2=\mathbf{p}\cdot\mathbf{z}_{\alpha\beta}$.

This again has a unique root $H>b$, but the *polynomial* for H has a degree which grows with N.

In every case, the third term on the left side of Eq. (3.1) or (3.2) can be simplified to

$$\frac{1}{H^2}\sum_{\alpha}(\mathbf{p}\cdot\mathbf{p}_\alpha)^2-\frac{b^4}{N}\Big). \tag{3.3}$$

The ith *velocity component* of the αth particle is given by

$$H^{\alpha,i} \equiv \frac{\partial H}{\partial p_{\alpha,i}}$$

It is usually demanded of a Hamiltonian that these $3N$ components should be linearly independent. This may be called the requirement of *nondegeneracy*. We will now show:

Theorem: These Hamiltonians are nondegenerate.

We will give a proof of this only for the case of two-dimensional space–time, so that each $\mathbf{p}_\alpha$ has but one component P_α. Let P be the sum of these P_α. Let T be the sum of their squares. Then the left side of (3.1) and (3.2) is [observing (3.3)]

$$\frac{H^2}{N} - T + \frac{P^2}{H^2}\,T - \frac{P^2}{N}\ ,$$

which equals

$$[(H^2 - P^2)/NH^2](H^2 + P^2 - NT). \tag{3.4}$$

In this discussion, the position variables are not involved at all. Hence, looking at (3.1) and (3.2), we may suppose that the right side depends only on H and P. (The argument applies to any other φ which shares that property.) Looking again at (3.4) we see that H satisfies a relation (3.5)

$$H^2 + P^2 - NT = R(H, P). \tag{3.5}$$

This we differentiate with respect to P_α, obtaining

$$2HH^\alpha + 2P - 2NP_\alpha = R_1 H^\alpha + R_2$$

where $H^\alpha = \partial H/\partial P_\alpha$, and R_1, R_2 are the two partial derivatives of R. Now suppose there were constants $c_1, \ldots, c_N$, not all zero, such that $\sum c_\alpha H^\alpha = 0$ for all values of $P_1, \ldots, P_N$. Then we obtain another equation for H:

$$2P\sum c_\alpha - 2N\sum c_\alpha P_\alpha = R_2 \sum c_\alpha .$$

This forces H to depend only on P and $\sum c_\alpha P_\alpha$. We apply this finding to (3.5) and conclude that T depends only on P and $\sum c_\alpha P_\alpha$. This it obviously does not (assuming $N > 1$!), so the supposition of linear dependence must be rejected.

In case of the more general F_N in (2.7), the nondegeneracy of H would follow from the nondegeneracy of the matrix $(F_{\alpha\beta})$.

There is another property one could demand of a Hamiltonian when it is supposed to describe *interacting particles*. At the very least one could ask the following. Suppose the particles are separated into two constellations G_1, G_2 where $\alpha = 1, 2, \ldots, M$ for those in G_1 and $\alpha = M+1, \ldots, N$ for those in G_2. Then if these two constellations are moved far apart, the total momentum of each constellation ought to be practically constant on each motion. This requirement has to be more precisely formulated, but we can say now that our systems fulfill it, by virtue of the inequality (2.6′).

A valuable discussion of such requirements, under the general name of "separability" is provided by Foldy.[3] It is easy to formulate requirements that are stronger than the above (except that for $N = 2$ one could not ask for more.) However, the above is all that we can verify. We now enter into details.

Theorem: Let H be a Hamiltonian resulting from φ as in (2.9) or from a sum of such φ's. Separate the N particles into n subsets $G_1, \ldots, G_n$. Hold fixed all p_α. If α, β belong to the same G_i, hold $\mathbf{z}_{\alpha\beta}$ fixed. If α, β belong to different G_i, make $|\mathbf{z}_{\alpha\beta}| \to \infty$. Then H has a limit H_∞ commuting with the total momentum of each G_i.

Proof: By (2.6′), when $|\mathbf{z}_{\alpha\beta}| \to \infty$, then that particular f tends to ∞. The way in which these f's enter into (2.8) or (2.9) makes this tantamount to setting that particular $g(\alpha, \beta) = -\infty$. In general it wipes out the variable $\mathbf{z}_{\alpha\beta}$ in the resulting Hamiltonian H_∞. Therefore, H_∞ contains $\mathbf{z}_{\alpha\beta}$ only when α, β belong to the same G_i. Such a $\mathbf{z}_{\alpha\beta}$ commutes with the momentum $\sum_{\alpha \in G_i} \mathbf{p}_\alpha$ for that constellation, and of course it commutes with the momentum for the other constellations. The theorem is thus proved.

In the particular case of (2.8), an even more remarkable thing happens. Let G_1 be just one particle and let G_2 be all the others. The limiting Hamiltonian H_∞ in this case is the one obtained from (2.8) with all $k_{\alpha\beta} = \infty$. This H_∞ gives a zero interaction. Thus each particle is essential for binding the others together.

Return now to the statement of the theorem. Can one say more? Of course it is not true for general particle mechanics that if H commutes with the momentum of some subset (of more than one particle) then those particles are independent of the others. It does not seem likely that the special form of our Hamiltonian could help.

A third property of the interaction (2.3) is its *symmetry*, where this means that it is invariant under each permutation of the particles (or indices α). This property may be loosened up by requiring only that after a permutation, only some smooth readjustment of parameters is needed to reproduce the original interaction. Let us call such an interaction *symmetrizable*.

This property must be kept in mind in order to rule out interactions which are easy to construct and yield specious Hamiltonians. We give such an example. For F_α we will take $p_\alpha \cdot p_\alpha - 1$ when $\alpha < N$. For F_N we take $p_N \cdot p_N - \varphi$ where φ is any function of the $N-1$ gram determinants.

$$g_\alpha = \begin{vmatrix} p_\alpha \cdot p_\alpha & p_\alpha \cdot z_{\alpha N} \\ p_\alpha \cdot z_{\alpha N} & z_{\alpha N} \cdot z_{\alpha N} \end{vmatrix}, \quad \alpha < N.$$

Note that $\{F_\alpha, g_\beta\} = 0$ for $\alpha < N$. If we set $F_\alpha = 0$ then $H_\alpha = \sqrt{1 + \mathbf{p}_\alpha \cdot p_\alpha}$ for $\alpha < N$. These H_α can now be inserted into the equation $F_N = 0$, giving

$$H_N^2 - \mathbf{p}_N^2 = \Phi$$

where Φ is what results from replacing the time component of p_α by $-H_\alpha$. The result is

$$H = \sqrt{1 + \mathbf{p}_1 \cdot \mathbf{p}_1} + \cdots + \sqrt{1 + \mathbf{p}_{N-1} \cdot \mathbf{p}_{N-1}} + \sqrt{\Phi + \mathbf{p}_N \cdot \mathbf{p}_N}.$$

Any attempt to make this example symmetrizable by insertion of other Φ's encounters the danger of causing $\{F_\alpha, F_\beta\} \neq 0$. Thus it can describe only an interaction in which one particle is qualitatively different from the others.

Besides symmetrizability, there is another desideratum we have kept in mind: One should be able to insert a parameter which when set equal to 0, gives a satisfactory zero interaction.

4. INTERNAL AND EXTERNAL VARIABLES

To insure a Poincaré-invariant interaction, one is almost forced to make each F_α dependent only on the inner products in the space–time metric of the set of $N(N+1)/2$ vectors p_α and $z_{\beta\gamma}$. The requirement that $\{F_\alpha, F_\beta\}=0$ suggests a study of the commutation properties of the linear space E_N of linear combinations of these $N(N+1)(N^2+N+2)/8$ inner products.

Theorem: E_N is a Lie algebra under $\{\ ,\ \}$. Let R be the set of linear combinations of the $p\cdot p_\alpha$ and $p\cdot z_{\beta\gamma}$ where $p=\sum p_\alpha$. Then R is a subalgebra and indeed the radical of E_N. Let S be the set of linear combinations of the inner products of the set of all $p_\alpha - p_\beta$, $z_{\gamma\delta}$. Then S is a subalgebra, $R\cap S$ is the zero element of E_N, and $E_N = R + S$. The Lie algebra structure of E_N is independent of the the metric and dimension of space–time.

Proof: The nondependence on the particular space–time was shown earlier. It is clear that E_N is closed under $\{\ ,\ \}$. Moreover, if r is one of the generators of R, and e is an element of E_N then $\{r,e\}$ is sure to have a "factor" p and thus belongs to R. The bracket of two elements of R is always a multiple of $p\cdot p$ (which constitutes the center). Thus R is solvable (Ref. 8, p.24).

Clearly, $E_N = R + S$, and $R\cap S = (0)$. To show that R is the radical, we show first that S has no solvable ideal greater than (0).

In order to prove this, we consider a different subalgebra Q and prove that $E_N = R + Q$, $Q\cap R = (0)$, Q has no solvable ideals (the same will then follow for S). The algebra Q is the linear span of all $p_\alpha\cdot p_\beta$, $p_\alpha\cdot z_{\beta N}$, $z_{\alpha N}\cdot z_{\beta N}$ with α and β less than N. This algebra is isomorphic to the Lie algebra L_{N-1} of Lorentz invariants of $p_1, \ldots, p_{N-1}, x_1, \ldots, x_{N-1}$.

We have already noted that as far as the Lie algebraic structure is concerned, in studying $L_M (M = N-1)$ we might as well regard space–time as one-dimensional. Thus each f in L_M is a polynomial

$$f = \tfrac{1}{2}a_{ij}p_ip_j + b_{ij}p_ix_j + \tfrac{1}{2}c_{ij}x_ix_j$$

in $\mathbb{R}^{2M}$. By virtue of the symplectic structure $\omega = dp\wedge dx + \cdots + dp_M\wedge dx_M$, f generates an infinitesimal linear transformation V (or vector field) or $\mathbb{R}^{2M}$ which preserves that structure, and which has a matrix readily found to be the $2M\times 2M$ matrix

$$\begin{pmatrix} -b^t & a \\ -c & b \end{pmatrix}$$

where b^t is the transpose of b. Conversely, every such matrix determines a V that preserves ω, and each V is generated by an f. (Hence L_M is the Lie algebra of the real symplectic group.)

Making use of this matrix representation, one can discover that L_M has no *Abelian* ideal (except 0). Thus, by Ref. 8, p.25, it has no solvable ideals. Therefore, R is the *radical* of E_N. Thus the theorem is proved.

The radical R is distinguished by purely Lie algebra properties, but the complementary subalgebra is not unique. It can be made unique by observing that the symmetric group acts on $1,2,\ldots,N$ because they refer to *particles* (not *dimensions*.) Thus S is the only *symmetric* complement to the radical.

In the case of the Euclidean group in $\mathbb{R}^3$ instead of the Poincaré group in $\mathbb{R}^4$, the functions in the complementary algebra S have been called *internal* variables.[2,3] Since R is an algebraically distinguished class, its members also deserve a name: the *external* variables. In the case of $\mathbb{R}^4$, we will call them internal and external *space–time* variables.

In Ref. 3, Foldy attaches great importance to interactions in which the Hamiltonian H satisfies $H^2 - p^2 = \varphi$ where φ depends only on internal variables, and in fact regrets that he cannot prove that Hamiltonians always have that form. We wish to offer an example to settle this question (in the negative).

Let $p = p_1 + \cdots + p_N$ and let $\bar{p} = p/N$. Similarly, define $\bar{x} = (x_1 + \cdots + x_N)/N$. Choose a positive real number μ. Define $f_\alpha = p_\alpha - \bar{p} + \mu(x_\alpha - \bar{x})$, for $\alpha = 1,2,\ldots,N$. Notice that any set of $N-1$ of the N equations

$$p\cdot f_1 = 0, \ldots, p\cdot f_N = 0$$

implies the remaining one. We take $F_\alpha = p\cdot f_\alpha$ for $\alpha < N$.

Now select a real-valued function φ defined for real values on the half-line $[0,\infty)$ which is positive, continuous, and monotonely decreasing. Let

$$f = -f_1\cdot f_1 - \cdots - f_N\cdot f_N$$

and define $F_N = p\cdot p - \varphi(f)$.

Theorem: The system of equations

$$F_1 = F_2 = \cdots = F_N = 0$$

admits an analytic solution for $H_1, \ldots, H_N$ in terms of the remaining components of $p_1, \ldots, p_N, x_1, \ldots, x_N$. Moreover, $\{F_\alpha, F_\beta\} = 0$ for $1 \le \alpha < \beta \le N$.

Proof: Since $\{f_\alpha, f_\beta\} = 0$ for one-dimensional space–time, the stated commutation relations are true in general.

Let the time component of f_α be called k_α, and the space component be the 3-vector $\mathbf{L}_\alpha$. As before, $-H_\alpha$ is the time component of p_α, and H is the sum of these H_α. Now the equations $F_1 = \cdots = F_{N-1} = 0$ insure that

$$-Hk_\alpha - \mathbf{p}\cdot\mathbf{L}_\alpha = 0 \tag{4.1}$$

for $\alpha = 1,2,\ldots,N$. We solve these equations for $k_1, \ldots, k_N$ and insert that into the equation $F_N = 0$, which takes the form

$$H^2 - \mathbf{p}^2 - \varphi\left(\sum_\alpha L_\alpha^2 - \frac{\sum(\mathbf{p}\cdot\mathbf{L}_\alpha)^2}{H^2}\right) = 0 \tag{4.2}$$

because $f = -k_1^2 - \cdots - k_N^2 + \mathbf{L}_1^2 + \cdots + \mathbf{L}_N^2$. If we can solve (4.2) analytically for H, our problem is solved because $k_\alpha = \mathbf{p}\cdot\mathbf{L}_\alpha/H$ and $k_\alpha = H_\alpha - H/N + \mu[t_\alpha - (t_1 + \cdots + t_N)/N]$.

Denote the left side of (5.2) by $F(H)$, and consider its behavior for $H \ge b$. [Here b is as before $(\mathbf{p}\cdot\mathbf{p})^{1/2}$.] For $H = b$ the argument (the original f) in φ is not less than 0, and hence the value of φ is positive. Thus $F(b) < 0$.

Observe that $\varphi[\cdots]$ in its dependence on H is decreasing. This shows that $F(H)$ is increasing. It is also easy to verify that $F(H)\to\infty$ as $H\to\infty$, although it requires taking into account that $\varphi(0)\neq 0$. We conclude that (4.2) has exactly one solution H with $H>b$, and in fact that H is greater than b. This proves the theorem.

To obtain the final Hamiltonian H for the system defined by this Hamiltonian surface, we set all $t_\alpha=0$. Then H still satisfies (4.2). Let us show that in this particular case $H^2-\mathbf{p}^2\equiv\psi$ is *not* an internal variable (unless φ is constant.) For if ψ were internal, then (φ being monotonely increasing and L_α being internal) the expression

$$\frac{\sum(\mathbf{p}\cdot\mathbf{L}_\alpha)^2}{\mathbf{p}^2+\psi}$$

would be internal. It is easy to see that it can vary even though all internal variables are held fixed.

Therefore, this example does not fulfill Foldy's expectation, which thus becomes Foldy's *condition*. This condition is just as reasonable as it ever was, but now it is seen to be not automatically fulfilled.

On the other hand, the method of Hamiltonian surfaces is not *incompatible* with Foldy's condition either. Our next example shows this.

We treat only the case $N=2$. Let q stand for p_1-p_2. Let $H_1-H_2=D$ and $\mathbf{p}_1-\mathbf{p}_2=\mathbf{q}$. Let $x_1-x_2=z$, with components t and $\mathbf{z}$. Let ψ be any smooth function of one variable and define

$$f=(q\cdot z)^2+\psi(z^2),\quad g=q^2z^2-\psi(z^2).$$

[Here $\psi(z^2)$ means z^2 (or $t^2-\mathbf{z}^2$) substituted into ψ. Perhaps it would help to remind the reader that the H's are the negatives of the time component of the momentum 4-vectors.] Then

$$\{f,g\}=0.$$

Choose $\psi(x)$ as $x\omega(x)$ and define the surface S_2 by the equations

$$H_1+H_2=[\mathbf{p}^2+\varphi((tE+\mathbf{q}\cdot\mathbf{z})^2+z^2\omega(z^2))]^{1/2}$$

and

$$H_1-H_2=E$$

where $E=[\mathbf{q}^2+\omega(z^2)]^{1/2}$.

The second equation can be written $D=E$. Squaring, we get $D^2=\mathbf{q}^2+\omega(z^2)$ or $q^2-\omega(z^2)=0$ which (since we work entirely in the region $z^2<0$) is equivalent to $F_1\equiv q^2z^2-\psi(z^2)=0$. The first equation is $H=[\mathbf{p}^2+\varphi((z\cdot q)^2+z^2\omega(z^2))]^{1/2}$. Squaring, we get $F_2\equiv H^2-\mathbf{p}^2-\varphi(f)=0$. Thus S_2 is a component of the variety defined by $F_1=F_2=0$, and $\{F_1,F_2\}=0$.

The Hamiltonian for this system is

$$H=[\mathbf{p}^2+\varphi((\mathbf{q}\cdot\mathbf{z})^2-\mathbf{z}^2\omega(-\mathbf{z}^2))]^{1/2}.$$

It is easy to select ω and φ so that we have a limiting Hamiltonian

$$H_\infty=[\mathbf{p}^2+a]^{1/2}$$

where a is constant, for $\mathbf{z}^2\to\infty$. Thus we have a degree of "separability," but the limiting Hamiltonian H_∞ here is degenerate. (To that degree this example is inferior to that given in Secs. 3 and 4, but at least its Hamiltonian can be explicitly presented.)

5. INFINITESIMAL HAMILTONIAN SYSTEMS

The "Hamiltonian surface" method[5] illustrated in this paper gives rise to a "Hamiltonian space-time interaction" (Ref. 7, p.253). From here one may obtain an *infinitesimal* Hamiltonian interaction,[7] in which the interest centers on generating functions which represent the Lie algebra of infinitesimal space–time transformations.

This computation was made in Ref. 5, with the following result (but different notation). For an infinitesimal space–time Poincaré transformation

$$U=U^1\frac{\partial}{\partial x^1}+\cdots+U^4\frac{\partial}{\partial x^4} \tag{5.1}$$

in $\mathbb{R}^4$, the generating function g_U for the corresponding canonical transformation in $T_1(\mathbb{R}^N)$ is

$$\sum_{\alpha=1}^{N}\sum_{i=1}^{3}U^i_\alpha p_{\alpha i}-U^4_\alpha H^0_\alpha \tag{5.2}$$

where U^i_α is U^i with $(x^1,\ldots,x^4)$ replaced by $(x^1_\alpha,x^2_\alpha,x^3_\alpha,0)$, and the superscribed 0 on H_α indicates that $t_1,\ldots,t_N$ are all set equal to 0.

The concept of infinitesimal Hamiltonian interaction is the one underlying the work of Thomas, Bakamjian, Foldy, and others.[2,3] (We regret not having seen these papers until Ref. 7 had appeared. Several of the principal concepts had already been explicitly considered by Foldy.)

The class of infinitesimal Hamiltonian interactions is considerably larger than the class of those derived from (global) Hamiltonian space–time interactions, and the latter class is larger than those obtainable from Hamiltonian surfaces S_N. To put it informally, if all one wants is a set of generating functions g_U representing the Lie algebra of the Poincaré group, it shouldn't be necessary to worry about global solvability of the type expressed by $(H-2)$.

More precisely, retain the conditions $(H-1)$ and $(H-3)$ of Ref. 5 and replace $(H-2)$ by the weaker condition $(H-2$ local) which reads just like $(H-2)$ except that $a^4_1,\ldots,a^4_N$ are replaced by 0.

More loosely, we require that the equations $F_1=0,\ldots,F_N=0$ have some single-valued smooth solution for $H_1,\ldots,H_N$ defined on some neighborhood of the manifold characterized by $t_1=\cdots=t_N=0$.

Then the formula (6.2) *gives generating functions with the property that*

$$\{g_U,g_V\}=g_{[U,V]} \tag{5.3}$$

whenever U,V are infinitesimal Poincaré transformations in space–time, and $[U,V]$ is their commutator.

Proof: It is fairly clear that we obtain a *local* action of the Poincaré group and hence an action of the Lie group. This shows that if g_U,g_V generate Δ'_U and Δ'_V (respectively) in $T_1(\mathbb{R}^{3N})$, then $\{g_U,g_V\}$ generates $\Delta'_{[U,V]}$. This gives us $\{g_U,g_V\}=g_{[U,V]}+c_{U,V}$ where $c_{U,V}$

a constant. To show that $c_{U,V}$ is 0 we begin again, using direct substitution.

We must recognize that the Poisson bracket in (5.3) is the one appropriate to $\mathbb{R}^{3N}$. Therefore, let us denote it by $\{\ ,\ \}'$, to distinguish it from the Poisson bracket $\{\ ,\ \}$ in $\mathbb{R}^{4N}$. For functions such as these g_U, g_V one has $\{g_U, g_V\}' = \{g_U, g_V\}$ because they contain no $p_{\alpha 4}$. Now suppose

$$U = \sum_{i=1}^{4} U^i \frac{\partial}{\partial x^i}, \quad V = \sum_{j=1}^{4} U^j \frac{\partial}{\partial x^j}$$

are two vector fields in space–time $\mathbb{R}^4$, not necessarily related to the Poincaré group. Let

$$G_U = \sum_{\alpha} \sum_{i=1}^{4} U^{i\alpha} p_{\alpha i}$$

and G_V be formed analogously. It is easy to verify

$$\{G_U, G_V\} = G_{[U,V]}. \tag{5.4}$$

Here superscript α means that $x^1, \ldots, x^4$ are replaced by $x^1_\alpha, \ldots, x^4_\alpha$.

Now we postulate the invariance under these transformations of the surface S_N, which implies that

$$\{G_U, p_{4\gamma} + H_\gamma\} = 0 \tag{5.5}$$

for each γ, on S_N. Finally, we have

$$\{p_{4\alpha} + H_\alpha, p_{4\beta} + H_\beta\} = 0, \tag{5.6}$$

a consequence of $(H-1)$ and $(H-2$ local). It is obvious that $g_U = G_U + U^{4\alpha}(p_{\alpha 4} + H_\alpha)$, where we use the summation convention for α. We now calculate

$$\begin{aligned}\{g_U, g_V\}' = &\{G_U, G_V\} \\ &+ \{G_U, V^{4\beta}(p_{\beta 4} + H_\beta)\} - \{G_V, U^{4\alpha}(p_{\alpha 4} + H_\alpha)\} \\ &+ \{U^{4\alpha}(p_{\alpha 4} + H_\alpha),\ V^{4\beta}(p_{\beta 4} + H_\beta)\}.\end{aligned} \tag{5.7}$$

We propose to evaluate the right-hand side on S_N. (This does not change the value of $\{g_U, g_V\}'$ because there are no $p_{\alpha 4}$ to set equal to $-H_\alpha$). Consider a term like

$$\begin{aligned}&\{G_U, V^{4\beta}(p_{4\beta} + H_\beta)\} = \\ &= V^{4\beta}\{G_U, p_{4\beta} + H_\beta\} + (p_{4\beta} + H_\beta)\{G_U, V^{4\beta}\}.\end{aligned} \tag{5.8}$$

The second term on the right side of (5.8) is clearly 0 on S_N, but so is the first, by (5.5). The third term on the right of (5.7) is 0 by (5.6). We obtain $\{g_U, g_V\}' = G_{[U,V]}$, but on S_N this is the same as $g_{[U,V]}$. Thus ($H-2$ local) (and *a fortiori* $H-2$) implies (5.3), as asserted.

6. CANONICAL TRANSFORMATIONS IN MULTIPLE SPACE-TIME

Canonical transformations in $T_1(\mathbb{R}^{4N})$ can be used to transform one N-particle interaction into another. In particular, a zero-interaction can be transformed into a nonzero interaction if the canonical transformation is properly conditioned. Actually, our examples will produce only infinitesimal interactions (in the sense of Sec. 5) but for some of them, the Hamiltonian can be exhibited explicitly.

The construction of canonical transformations in $T_1(\mathbb{R}^{4N})$ is nothing new, since $T_1(\mathbb{R}^{4N})$ is also the phase space for a single particle moving in real $4N$-dimensional *space*. There is no need therefore to review the general theory. For the moment, let $y^1, \ldots, y^{4N}$ be the Cartesian coordinates, and $q_1, \ldots, q_{4N}$ the canonically dual momenta. Select a function U of $8N$ variables. Define

$$q_i = \frac{\partial U}{\partial y^i}(y^1, \ldots, y^{4N}, \bar{q}_1, \ldots, \bar{q}_{4N}), \tag{6.1}$$

$$\bar{y}^i = \frac{\partial U}{\partial \bar{q}^i}(y_1, \ldots, y^{4N}, \bar{q}_1, \ldots, \bar{q}_{4N}). \tag{6.2}$$

Solve these equations for the $\bar{q}$'s and $\bar{y}$'s in terms of the $\bar{q}$'s and $\bar{y}$'s. If this is possible, define a map $\mathcal{U}: T_1(\mathbb{R}^{4N}) \to T_1(\mathbb{R}^{4N})$ wherein

$$\bar{q}^i(\Phi) = q_i(T(\Phi)), \quad \bar{y}_i(\Phi) = y_i(T(\Phi)).$$

Then $\mathcal{U}$ is symplectic, i.e., $\{f\circ\mathcal{U},\ g\circ\mathcal{U}\} = \{f,g\}\circ\mathcal{U}$ because (as is easy to show) $\sum \bar{q}_i\, d\bar{y}^i - \sum q_i\, dy^i$ is exact.

Now let an interaction be given, with a Hamiltonian surface S_N. Then $\mathcal{U}(S_N)$ is also a Hamiltonian surface. If S_N is Poincaré invariant, and if $\mathcal{U}$ commutes with the action of the Poincaré group in $T_1(\mathbb{R}^{4N})$, then $\mathcal{U}(S_N)$ is also invariant. (We will call such $\mathcal{U}$ Poincaré *constructions*.) We will want to insure that ($H-2$ local) is safisfied for $\mathcal{U}(S_N)$. That will always be a matter for investigation, for which we set a pattern in our examples.

We now revert to our notation x^i_α and $p_{\alpha i}$ for $T_1(\mathbb{R}^{4N})$. Recall also that $p = p_1 + \cdots + p_N$. We will investigate the $\mathcal{U}$ "generated" by

$$U = \sum_{\alpha=1}^{N} x_\alpha \circ p_\alpha + g$$

where g depends on the various $\bar{p}\cdot z_{\alpha\beta}$ and $z_{\alpha\beta}\cdot z_{\gamma\delta}$. The U is not Poincaré invariant (but the g is). Nevertheless, $\mathcal{U}$ is a Poincaré construction, as will become apparent when we write down the formulas (6.1) and (6.2).

Let I be the set of indices $1, \ldots, 4N$ such that $\sum_{i\in I} q_i$ is one of the four components of p. From (6.1) we obtain $q_i = \bar{q}_i + \partial g/\partial y^i$, whence $\sum_{i\in I} q_i = \sum_{i\in I} \bar{q}_i + \sum_{i\in I}(\partial/\partial y^i)g$. But this last term is 0 because g is invariant under those translations in $\mathbb{R}^{4N}$ induced by translations in $\mathbb{R}^4$. Thus we see that the 4-momentum is preserved by $\mathcal{U}$:

$$\bar{p} = p. \tag{6.3}$$

From this it follows that $\bar{H} = H$, but this does not mean that the new interaction is zero if the old one is. What is crucial is how $\bar{H}$ depends on the $\bar{\mathbf{p}}_\alpha$ and $\bar{\mathbf{x}}_\alpha$, not how it depends on $\mathbf{p}_\alpha$ and $\mathbf{x}_\alpha$.

To see that $\mathcal{U}$ is a Poincaré construction we have to examine (6.1), and (6.2) in terms of the p's and z's. To this end denote $\bar{p}\cdot z_{\alpha\beta}$ by $\eta_{\alpha\beta}$ and $z_{\alpha\beta}\cdot z_{\gamma\delta}$ by $\zeta_{\alpha\beta\gamma\delta}$ using only enough of these variables to represent the variables without repetition (note: $\eta_{11} = 0$, $\eta_{12} = -\eta_{21}$, etc.) We now write down (6.1) and (6.2), and combine the formulas for $p_{\alpha 1}, \ldots, p_{\alpha 4}$ and $x^1_\alpha, \ldots, x^4_\alpha$ (respectively) into 4-vector formula. The result for (6.1) is

$$p_\epsilon = \bar{p}_\epsilon + \bar{p}A_\epsilon + \sum_{\alpha} B_{\alpha\epsilon} z^{\alpha\epsilon} \tag{6.4}$$

where

(a) A_ϵ and $B_{\alpha\epsilon}$ are linear combinations with constant coefficients of the partial derivatives $\partial g/\partial \eta_{\alpha\beta}$ and $\partial g/\partial \zeta_{\alpha\beta\gamma\delta}$, and

(b) $z^{\alpha\epsilon}$ is the covariant form of $z_{\alpha\epsilon}$, so that if the latter is (a,b,c,d), then the former is $(-a,\ -b,\ -c,\ d)$. Recalling that by (6.3) we can replace $\bar{p}$ by p and transposing, we see that $\bar{p}_\epsilon$ depends on the p's and x's in a Poincaré-invariant way, because the A's and B's are Poincaré invariants.

In the same way, we can see that $\bar{x}_\epsilon$ depends on the x's and p's in a Poincaré-invariant way. We omit writing down these latter formulas, but we must now assume that they, together with (6.4), can be inverted, at least locally. If this is possible, we may take the barred coordinates to be the original coordinates while p,x are the transformed ones.

Recall that (6.4) implied that

$$\bar{p}_\epsilon = p_\epsilon - pA_\epsilon - \sum_\alpha B_{\alpha\epsilon} z^{\alpha\epsilon} \tag{6.5}$$

where A,B depend only on p and the $z_{\alpha\beta}$.

We begin with the zero interaction defined by

$$\bar{p}_1^2 = m^2,\ \ldots,\bar{p}_N^2 = m_N^2$$

or more precisely $\bar{H}_\alpha = [m_\alpha^2 + \bar{p}_\alpha^2]^{1/2}$. This is obviously Poincaré invariant. Using (6.5), these give equations relating the p's and x's, namely

$$H_\epsilon - HA_\epsilon - \sum B_{\alpha\epsilon} z^{\alpha\epsilon} = [m_\epsilon^2 + (\mathbf{p}_\epsilon - \mathbf{p}A_\epsilon - \sum B_{\alpha\epsilon} z^{\alpha\epsilon})^2]^{1/2}. \tag{6.6}$$

Here the A,B depend on H, the $\mathbf{p}_\alpha$, and the $\mathbf{x}_\alpha$, as well as the $l_{\alpha\beta}$. These formula will give the desired solution for $H_1,\ \ldots,H_N$ provided we can first evaluate H. This we do by summing (6.6) with respect to ϵ. The left side, being $\bar{H}_\epsilon$ sums to $\bar{H}=H$, as already remarked. [Indeed, $\sum_\epsilon A_\epsilon = 0$ and $\sum_{\alpha,\epsilon} B_{\alpha\epsilon} z^{\alpha\epsilon} = 0$, reflecting the argument establishing (7.3).] Thus

$$H = \sum_{\epsilon=1}^{N} [m_\epsilon^2 + (\mathbf{p}_\epsilon - \mathbf{p}A_\epsilon - \sum B_{\alpha\epsilon} z^{\alpha\epsilon})^2]^{1/2}. \tag{6.7}$$

This has to be solved for H. *When* $t_1 = \cdots = t_N = 0$, *there is no* H *left on the right-hand side*, so H can be found. This is the *final* Hamiltonian. Appropriate conditions can be imposed on U to ensure that when $\mathbf{p}_\alpha$, $\mathbf{x}_\alpha$ are fixed fixed, then for any sufficiently small departure from $t_1 = \cdots = t_N = 0$, (6.7) can still be solved. This gives us (H – 2 local), and an infinitesimal interaction is established.

For the old system, the motions are characterized by *constant* $\bar{p}_\epsilon$, and $\bar{\mathbf{x}}_\epsilon$, a *linear* function of $\bar{t}_\epsilon$. Our enterprise would have been in vain if the same were true for p_ϵ and x_ϵ. To see that it need not, we will consider afresh the case $N=2$, in which there is but one $\eta = (\bar{p}_1 + \bar{p}_2)\cdot z$, $z = x_1 - x_2$, and one $\zeta = z\cdot z$. To further simplify we let $p_1 = p$, $p_2 = q$, $x_1 = x$, $x_2 = y$.

We select $g = \eta\varphi(\zeta)$. Then (6.4) takes the form

$$p = \bar{p} + (\bar{p} + \bar{q})\varphi(\zeta) + 2\eta\varphi'(\zeta)z^*$$

(where z^* is the covariant form of z) and

$$q = \bar{q} - (\bar{p} + \bar{q})\varphi(\zeta) - 2\eta\varphi'(\zeta)z^*.$$

From (6.2) we obtain

$$\bar{x} = x + z\varphi(\zeta),\quad \bar{y} = y + z\varphi(\zeta).$$

These equations are equivalent to

$$x = \bar{x} - \bar{z}\varphi(\bar{z}^2),\quad y = \bar{y} - \bar{z}\varphi(\bar{z}^2).$$

Therefore, this particular $/\!/$ is globally invertible.

Transposing the p and q equations and using $\bar{p} + \bar{q} = p + q$ gives

$$\bar{p} = p - (p+q)\varphi - 2(p+q)\cdot z\varphi' z^*,$$

$$\bar{q} = q + (p+q)\varphi + 2(p+q)\cdot z\varphi' z^*.$$

Thereupon (6.6) yields

$$H_1 - H\varphi - 2H(t_1 - t_2)^2\varphi'$$
$$= \{m_1^2 + [\mathbf{p} - (\mathbf{p}+\mathbf{q})\varphi - 2(\mathbf{p}+\mathbf{q})(t_1 - t_2)^2\varphi']^2\}^{1/2} \tag{6.8}$$

and a similar equation for H_2. The argument in φ, φ' is $\zeta = (t_1 - t_2)^2 - \mathbf{z}^2$. In this example, there is no H in the right side of (6.8) and by adding it to its mate we obtain H. When $t_1 = t_2$, we obtain the final Hamiltonian

$$H = \{m_1^2 + [\mathbf{p} - (\mathbf{p}+\mathbf{q})\varphi]^2\}^{1/2} + \{m_2^2 + [\mathbf{q} + (\mathbf{p}+\mathbf{q})\varphi]^2\}^{1/2}.$$

Here $\varphi = \varphi(-z^2)$. Hence for $\varphi(s) = (-s)^{-1/4}$, this approaches the free Hamiltonian when $|\mathbf{z}| \to \infty$.

We now make a canonical transformation in $T_1(\mathbb{R}^6)$ defined by

$$u = \bar{p}_1 x^1 + \bar{p}_2 x^2 + \bar{p}_3 x^3 + \bar{q}_1 x^1 + \bar{q}_2 x^2 + \bar{q}_3 x^3$$
$$+ [(\bar{p}_1 + \bar{q}_1)z^1 + (\bar{p}_2 + \bar{q}_2)z^2 + (\bar{p}_3 + \bar{q}_3)z^3]\varphi(-z\cdot z).$$

This gives a new Hamiltonian with no $\bar{z}$, and hence zero interaction. For its solutions, the $\bar{p}$ and $\bar{q}$ are constant, while $\bar{x}$, $\bar{y}$ are linear functions of time. Must the same be true of x and y? From

$$\bar{\mathbf{x}} = \mathbf{x} + \mathbf{z}\varphi(-\mathbf{z}^2),\quad \bar{\mathbf{y}} = \mathbf{y} + \mathbf{z}\varphi(-\mathbf{z}^2),$$

we get $\bar{\mathbf{z}} = \mathbf{z}$ and

$$\mathbf{x} = \bar{\mathbf{x}} - \bar{\mathbf{z}}\varphi(-\bar{\mathbf{z}}^2),\quad \mathbf{y} = \bar{\mathbf{y}} - \bar{\mathbf{z}}\varphi(-\bar{\mathbf{z}}^2).$$

Evidently, if φ is not constant, then with at least some initial conditions, $\mathbf{x}$ and $\mathbf{y}$ are not linear in time and hence the interaction is not zero.

If φ is chosen so that $\varphi(-\infty) = 0$, then whenever the particles drift apart, their motion will approach that for zero interaction.

*Preparation of this paper was supported in part by USF Grant GP 3369X.

[1]D.G. Currie, T.F. Jordan, and E.C.G. Sudarshan, Rev. Mod. Phys. **35**, 350 (1963).
[2]B. Bakamjian and L.H. Thomas, Phys. Rev. **92**, 1300 (1953).
[3]Leslie L. Foldy, Phys. Rev. **122**, 275 (1961).
[4]L. Bel, Ann. Inst. H. Poincaré **18**, 57 (1973).
[5]R. Arens, Nuovo Cimento B **21**, 395 (1974).
[6]R. Arens, Int. J. Theor. Phys. **13** (to be published).
[7]R. Arens and Donald G. Babbitt, Pacific J. Math. 28, 243 (1969).
[8]Nathan Jacobson, *Lie Algebras* (Wiley, New York, 1962).

Conserved vector densities and their curl expressions

Richard Pavelle

Innovation Industries, Woburn, Massachusetts 01801
(Received 13 September 1974)

It has been shown that all Lagrange densities which are concomitants of sets of independent geometrical objects possess conserved vectors of the Komar type. For the general conserved vector derived from these Lagrangians we give an explicit method for rewriting the vector as the curl of a two index object.

INTRODUCTION

In a recent paper[1] we discussed a new type of conserved vector density defined uniquely for each Lagrangian density on an n-dimensional Riemannian manifold. The vector density possesses an identically vanishing ordinary divergence (and covariant divergence since it is a density), and it therefore belongs to the class of strongly conserved quantities. For the Lagrangian of the Brans–Dicke theory[2] we derived a conserved vector density[3] similar to the well-known Komar conservation law of general relativity,[4] while for other Lagrangians we found new vector densities whose physical significance is not yet fully understood.[5]

A mathematical problem arises if we wish to obtain useful physical results from a conserved vector density. We must first rewrite the vector density as the curl of an antisymmetric object with two indices. This accomplished we may then integrate over a fixed volume of the manifold to obtain the physical conservation law. In Ref. 1 we gave a difficult and awkward method of accomplishing this for certain cases of interest. Now our purpose is to consider the general problem we encounter in this type of conservation law and give a simple method for finding the corresponding curl form.

I. CURL FORM FOR THE CONSERVED VECTOR

Consider a vector density V^i which is identically divergence free and which is defined in the following manner[1]:

$$V^i = \sum_{r=0}^{p} A_a^{ij_1j_2\cdots j_r}\,\xi^a{}_{,j_1j_2\cdots j_r}. \tag{1}$$

Here ξ^a is an arbitrary vector field, the quantities A_a^i, $A_a^{ij_1}, \ldots, A_a^{ij_1j_2\cdots j_r}$ are given functions of the coordinates and the comma denotes differentiation with respect to the coordinates. The fact that V^i is divergence free implies, from Poincare's lemma,[6] that V^i may be written locally in the form of a curl

$$V^i \equiv T^{ij}{}_{,j} \tag{2}$$

where T^{ij} is skew-symmetric. Our problem is to express (1) in this form. This is important for only then can we integrate T^{ij} and employ the divergence theorem to obtain the physical quantities which are conserved for particular choices of V^i and ξ^a.

In order to accomplish this it is clear that at one stage we must differentiate (1) and employ the arbitrary nature of ξ^a to obtain symmetry identities involving the A's. Unfortunately, the direct approach of forming the divergence of (1) becomes complicated by the fact that the symmetry identities thus obtained are coupled, whereas we shall find that the curl form is not. By employing the arbitrariness of ξ^a at the start, however, we can simplify the problem by showing that the vanishing of the divergence of (1) implies the vanishing of the divergence of each term in (1) separately. To see this let us choose the arbitrary vector ξ^a to have one component, x^1. With this choice the first term in the summation in (1) is the only survivor and it follows that this term can be written in the curl form as $A_a^i\xi^a = T^{ij}_{(1),j}$. Thus in general we may write (1) as

$$V^i - T^{ij}_{(1),j} = \sum_{r=1}^{p} A_a^{ij_1j_2\cdots j_r}\,\xi^a{}_{,j_1j_2\cdots j_r}. \tag{3}$$

Now choose $\xi^a = (x^1)^2$. With this choice only the first term on the right-hand side of (3) survives and we may write it in the form $A_a^{ij}\,\xi^a{}_{,j_1} = T^{ij}_{(2),j}$. By continuing this process we see that each term in (1) must be independently expressible as a curl and we thus need only consider the general term in (1).

Let us then consider the object

$$W^i_{(r)} \equiv A_a^{ij_1j_2\cdots j_r}\,\xi^a{}_{,j_1j_2\cdots j_r} \tag{4}$$

where $W^i{}_{,i} \equiv 0$. Since partial derivatives commute we may write (4) as

$$W^i \equiv A_a^{i(j_1j_2\cdots j_r)}\,\xi^a{}_{,j_1j_2\cdots j_r}, \tag{5}$$

where the round brackets surrounding enclosed indices denote the complete symmetrization of the enclosed indices.[7] If we consider the identical vanishing of the divergence of (5) for small values of r we find that the object defined by

$$S_a^{ij_1j_2\cdots j_r} \equiv A_a^{i(j_1j_2\cdots j_r)} - A_a^{j_1(ij_2\cdots j_r)} \tag{6}$$

is useful. This object has the following symmetries: It is skew-symmetric under the interchange of the index pair (i, j_1) but symmetric under the interchange of (j_s, j_t) for $s, t = 2, 3, \ldots, p$ and $s \neq t$.

Once again, by considering low values of r in (5) one sees that it is possible to write $W^i(r)$ in the form

$$W^i(r) = \{C_1 S_a^{ij_1j_2\cdots j_r}\,\xi^a{}_{,i_2j_3\cdots j_r} + C_2 S_a^{ij_1j_2\cdots j_r}{}_{,j_2}\,\xi^a{}_{,j_3j_4\cdots j_r} + \cdots + C_r S_a^{ij_1j_2\cdots j_r}{}_{,j_2j_3\cdots j_r}\,\xi^a\}_{,j_1} \tag{7}$$

where $C_1, C_2, \ldots, C_r$ are constants to be determined. Clearly, this is a curl form for $W^i(r)$. Since the set of partial derivatives of ξ^a are mutually independent we differentiate (7) and collect terms in corresponding orders of derivatives of ξ^a. Owing to the independence of the

 Copyright © 1975 American Institute of Physics

derivatives of the vector ξ^a, we find the identities

$$A_a^{ij_1j_2\cdots j_r}\xi^a{}_{,j_1j_2\cdots j_r} \equiv C_1 S_a^{ij_1j_2\cdots j_r}\xi^a{}_{,j_1j_2\cdots j_r}, \tag{8a}$$

$$(C_l S_a^{ij_1j_2\cdots j_r} + C_{l+1} S_a^{ij_{l+1}j_2j_3\cdots j_l j_1 j_{l+2}\cdots j_r})_{,j_1j_2\cdots j_l} \times \xi^a{}_{,j_{l+1}\cdots j_r} \equiv 0 \qquad (l=2,3,\ldots,r). \tag{8b}$$

We have other identities too. Let us differentiate (5) and set it to zero. We then find

$$A_a^{i(j_1j_2\cdots j_r)}{}_{,i} \equiv 0 \tag{9a}$$

and

$$A_a^{(ij_1j_2\cdots j_r)} \equiv 0. \tag{9b}$$

The sets (8) and (9) will enable us to determine the constants in (7).

It is a simple matter to establish the identity

$$A_a^{(ij_1j_2\cdots j_r)} \equiv \frac{1}{r}[A_a^{i(j_1j_2\cdots j_r)} + A_a^{j_1(ij_2\cdots j_r)} + \cdots + A_a^{j_r(ij_1j_2\cdots j_{r-1})}]. \tag{10}$$

Hence, from (9b) and (10) we find

$$A_a^{j_1(ij_2\cdots j_r)}\xi^a{}_{,j_1j_2\cdots j_r} \equiv -\frac{1}{r}A_a^{i(j_1j_2\cdots j_r)}\xi^a{}_{,j_1j_2\cdots j_r}. \tag{11}$$

Therefore, (8a) implies

$$C_1 = r/(r+1). \tag{12}$$

Now consider (8b). From (6) this is

$$(C_l A_a^{i(j_1\cdots j_r)} - C_l A^{j_1(ij_2\cdots j_r)} + C_{l+1}A_a^{i(j_1\cdots j_r)} - C_{l+1}A_a^{j_{l+1}(i\cdots j_r)})_{,j_1j_2\cdots j_l}\xi^a{}_{,j_{l+1}\cdots j_r}. \tag{13}$$

However, from (9a) and (10) it follows that

$$A_a^{j_{l+1}(i\cdots j_r)}{}_{,j_1j_2\cdots j_l}\xi^a{}_{,j_{l+1}\cdots j_r} \equiv -\frac{1}{r-l}A_a^{i(j_1\cdots j_r)}{}_{,j_1\cdots j_l}\xi^a{}_{,j_{l+1}}. \tag{14}$$

Hence, upon substitution of (14) into (13) and by employing (9a) to eliminate the second term in (13) we find that (8b) or (13) gives the recursion

$$C_{m+1} = -[(r-m)/(r-m+1)]C_m \tag{15}$$

This tells us that

$$C_{m+1} = (-1)^m[(r-m)/r]C_1 \tag{16}$$

and from (12) we then have

$$C_{m+1} = (-1)^m[(r-m)/(r+1)], \quad m=0,1,2,\ldots,r-1, \tag{17}$$

for all m. Hence from (1), (7), and (17) we have

$$V^i = \sum_{r=0}^{P} W^i(r), \tag{18}$$

where

$$W^i(r) = \left\{\frac{r}{r+1}S_a^{ij_1j_2\cdots j_r}\xi^a{}_{,j_2j_3\cdots j_r} - \frac{r-1}{r+1}S_a^{ij_1j_2\cdots j_r}{}_{,j_3} \times \xi^a{}_{,j_3j_4\cdots j_r} + \frac{r-2}{r+1}S_a^{ij_1j_2\cdots j_r}{}_{,j_2j_3}\xi^a{}_{,j_4\cdots j_r}+\cdots\right\}_{,j_1}$$

and $S_a^{ij_1j_2\cdots j_r}$ is defined by (6).

[1] R. Pavelle, J. Math. Phys. **16**, 696 (1975).
[2] C. Brans and R.H. Dicke, Phys. Rev. **124**, 925 (1961).
[3] R. Pavelle, Phys. Rev. D **8**, 2369 (1973).
[4] A. Komar, Phys. Rev. **113**, 934 (1959).
[5] A.J. Fennelly and R. Pavelle, American Physical Society Meeting, Washington, D.C., April 1974, Paper HJ5 (to be published).
[6] H. Flanders, *Differential Forms with Applications to the Physical Sciences* (Academic, New York, 1963), p. 27.
[7] For example, $A_a^{i(jk)} \equiv 1/2!\,(A_a^{ikj} + A_a^{ijk})$.

Erratum: On the existence of weakly retarded and advanced Green's functions [J. Math. Phys. 15, 2093 (1974)]

A. Z. Capri

Department of Physics, The University of Alberta, Edmonton, Alberta, Canada T6G 2E1
(Received 22 January 1975)

The transformations given by Eqs. (6) and (16), (17), (18) are correct only if the function f is independent of $\mathbf{x}$. Thus throughout the paper f should be considered as a function of t only. This does not change the conclusions of the paper.

CUMULATIVE AUTHOR INDEX

All authors published so far in the current volume are listed alphabetically with the issue and page numbers following the dash. A cumulative author and subject index covering issues 1 through 12 is published yearly. An (E) *after the page number indicates an Erratum.*