Quantum dynamical semigroups and multipole relaxation of a spin in isotropic surroundings[a)]

Maurizio Verri[b)]

Istituto di Matematica del Politecnico, Milano, Italy
and Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Italy

Vittorio Gorini

Istituto di Fisica dell'Università, Milano, Italy
and Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Italy
(Received 8 February 1978)

We derive and discuss three different parametrizations of the generator of a dynamical semigroup which describes the Markovian relaxation of a spin j under the influence of isotropic surroundings. The relevant parametrizations that we consider are the strengths of the polarities of the interaction, the relaxation rates of the different multipoles and the transition probabilities per unit time among the Zeeman sublevels. The results are model independent and allow us to derive a set of relations and inequalities for the transition probabilities and for the relaxation rates whose validity is not bound to any specific assumption concerning the mechanisms which govern the relaxation.

1. INTRODUCTION

Quantum dynamical semigroups[1-3] with various degrees of symmetry are widely used in the phenomenological description of optical pumping of atoms and molecules.[4,5] In this paper we discuss three physically meaningful parametrizations of the generator L of a dynamical semigroup of an N-level quantum system S, under the assumption of invariance of L under the irreducible representation D^j $[j=(N-1)/2]$ of the three-dimensional rotation group SO(3). From the physical point of view, the system S can be thought of as representing a single spin j which evolves irreversibly in a Markovian fashion under the influence of isotropic surroundings, according to the equation

$$\frac{d}{dt}\rho(t) = L\rho(t) \tag{1.1}$$

for the density matrix $\rho(t)$ which describes the state of the system. One of the best known examples of a situation of this kind is provided by the relaxation, in a weak external magnetic field, among the Zeeman sublevels of the electronic spin of the ground state (or of some excited state) of the atoms of an optically pumped atomic vapor, when the atoms have zero nuclear spin.[4]

It is generally argued that the external mechanisms which are responsible for the relaxation are such as to justify the Markovian approximation (1.1) to the generalized master equation which gives in principle an exact description of the subdynamics of the spin.[4,6-8] Then, the isotropy assumption is a good approximation whenever the temperature of the vapor is high enough and the external field is sufficiently weak that the energy difference among the different levels is much smaller than kT so that, at thermal equilibrium, the Zeeman levels are almost uniformly populated.

Even though our results are to a large extent known,[9-12] they have been derived so far only within the context of specific phenomenological models for the mechanisms governing the relaxation. On the other hand, our treatment, which is based on the theory of dynamical semigroups, has the advantage of providing a unified and simple proof of the results, which is generally valid, granted the assumptions of Markovicity and isotropy, independently of any particular model for the interaction.

In Sec. 2 we give the general form of the generator L of a dynamical semigroup for a spin j, under the assumption of spatial isotropy, as a function of the different strengths λ_J $(J=1,2,\ldots,2j)$ of the interaction polarities. From this, we derive a set of inequalities which have to be satisfied by the relaxation rates γ_J of the different multipoles. Furthermore, we show that in the special case under consideration positivity and complete positivity[13,14] are equivalent.

In Sec. 3 we consider the parametrization of L in terms of the transition probabilities per unit time $W_{mm'}$ among the Zeeman sublevels and we derive the constraints imposed upon the $W_{mm'}$ by their relations with the parametrizations γ_J and λ_J.

2. THE PARAMETRIZATIONS λ_J AND γ_J

We recall that a *dynamical semigroup* of an N-level system is a one-parameter continuous semigroup $t \to \Lambda_t = \exp(Lt)$, $t \geq 0$, of completely positive[13,14] trace preserving linear maps on the algebra $M(N)$ of the $N\times N$ complex matrices. Complete positivity, which is a stronger property than positivity, has been recognized to be a general feature of the dynamics of quantum open systems.[15,1-3]

It was shown in Ref. 1 (see also Ref. 2) that a linear operator $L: M(N) \to M(N)$ is the generator of a dynamical semigroup iff it has the form

$$L: \rho \to L\rho = -i[H,\rho] + \tfrac{1}{2}\sum_{i,j=1}^{N^2-1} c_{ij}\{[F_i, \rho F_j^*] + [F_i\rho, F_j^*]\}, \tag{2.1}$$
$$\rho \in M(N),$$

[a)]Supported in part by NATO Research Grant No. 1380 and by CNR Research Contract No. 77.01543.63.
[b)]A fellowship from the Italian National Research Council (CNR) is acknowledged.

 0022-2488/78/1909-1803$1.00

where

$$H = H^*, \quad \mathrm{tr}(H) = 0, \tag{2.1a}$$

$$\mathrm{tr}(F_i) = 0, \quad \mathrm{tr}(F_i^* F_j) = \delta_{ij} \quad (i, j = 1, 2, \ldots, N^2 - 1), \tag{2.1b}$$

$$\{c_{ij}\} \text{ is a complex positive matrix.} \tag{2.1c}$$

For a given L, H is uniquely determined by the condition $\mathrm{tr}(H) = 0$ and $\{c_{ij}\}$ is uniquely determined by the choice of the F_i's.

A straightforward consequence of Theorem 5 of Ref. 16 is that the linear operator $L : M(N) \to M(N)$ is the generator of a one-parameter continuous semigroup of positive trace preserving linear maps on $M(N)$ iff it is of the form (2.1), where H and $\{F_i\}_{i=1,2,\ldots,N^2-1}$ satisfy (2.1a) and (2.1b) above and where $\{c_{ij}\}$ is a self-adjoint matrix with the property that

$$\sum_{i,j=1}^{N^2-1} \langle x | F_i | y \rangle c_{ij} \overline{\langle x | F_j | y \rangle} \geq 0 \tag{2.1c'}$$

for all pairs of orthogonal vectors $|x\rangle$ and $|y\rangle$ in $\mathbb{C}^N$. Denote by $\{|jm\rangle\}$ $(j = (N-1)/2,\ m = j, j-1, \ldots, -j)$ the standard spin basis, i.e., $J_3 |jm\rangle = m |jm\rangle$, and by $D^j : R \to D^j(R)$, $R \in \mathrm{SO}(3)$, the standard form of the irreducible representation of SO(3) acting on $\mathbb{C}^{2j+1}$, namely $D^j(R) = \exp(-i\alpha J_3) \exp(-i\beta J_2) \exp(-i\gamma J_3)$, where α, β, γ are the Euler angles which define the rotation R.[17]

Let $\{T_{JM}\}$ $(J = 0, 1, \ldots, 2j;\ M = J, J-1, \ldots, -J)$ denote the orthonormal basis [w.r.t. the Hilbert–Schmidt product $(A | B) = \mathrm{tr}(A^* B)$] in $M(N)$ of the irreducible spherical tensors,

$$D^j(R) T_{JM} D^j(R)^* = \sum_{M'=-J}^{J} D^J_{M'M}(R) T_{JM'}, \tag{2.2}$$

$$\langle jm | T_{JM} | jm' \rangle = (-)^{j-m} \sqrt{2J+1} \begin{pmatrix} j & j & J \\ m & -m' & -M \end{pmatrix}, \tag{2.3}$$

where the phase conventions are those of Ref. 17, Appendix B.

We say that a dynamical semigroup on $M(2j+1)$, $\Lambda_t = \exp(Lt)$, is invariant under D^j if

$$L(D^j(R) \rho D^j(R)^*) = D^j(R)(L\rho) D^j(R)^*,$$

$$\forall R \in \mathrm{SO}(3) \text{ and } \forall \rho \in M(2j+1). \tag{2.4}$$

Proposition 2.1: The general form of the generator L of a dynamical semigroup Λ_t on $M(2j+1)$ which is invariant under D^j is

$$L : \rho \to L\rho = -\tfrac{1}{2} \sum_{J=1}^{2j} \lambda_J \sum_{M=-J}^{J} [T^*_{JM}, [T_{JM}, \rho]], \tag{2.5}$$

$$\rho \in M(2j+1),$$

where

$$\lambda_J \geq 0 \quad (J = 1, 2, \ldots, 2j). \tag{2.6}$$

Proof: Choosing for the set $\{F_i\}$ in Eq. (2.1) the traceless tensors T_{JM}, $J = 1, 2, \ldots, 2j$, $M = J, J-1, \ldots, -J$, we have

$$L : \rho \to L\rho = -i[H, \rho]$$

$$+ \tfrac{1}{2} \sum_{J,J'=1}^{2j} \sum_{M=-J}^{J} \sum_{M'=-J'}^{J'} c^{JJ'}_{MM'} \{[[T_{JM}, \rho T^*_{J'M'}]$$

$$+ [T_{JM} \rho, T^*_{J'M'}]\}. \tag{2.7}$$

From (2.4), (2.2) and the uniqueness of the decomposition (2.7) we get

$$[H, D^j(R)] = 0, \ \forall \ R \in \mathrm{SO}(3),$$

and

$$c^{JJ'}_{MM'} = \sum_{Q=-J}^{J} \sum_{Q'=-J'}^{J'} D^J_{MQ}(R)\, c^{JJ'}_{QQ'} \overline{D^{J'}_{M'Q'}(R)}, \ \forall \ R \in \mathrm{SO}(3).$$

Hence, from the irreducibility of D^j it follows that $H = 0$ and that $c^{JJ'}_{MM'} = \lambda_J \delta_{JJ'} \delta_{MM'}$, where, by (2.1c), $\lambda_J \geq 0$. Then (2.5) follows from the property $T^*_{JM} = (-)^M T_{J-M}$.

QED

Regarding the physical meaning of the parameters λ_J, if we think for instance of the semigroup Λ_t as obtained in the limit of white noise $\epsilon \to 0$, when the effect of the surroundings on the spin is represented by an isotropic stationary Gaussian fluctuating Hamitonian $\tilde{H}(T) = \sum_{JM} T_{JM} V_{JM}(t)$, $\langle V_{JM}(t) V_{J'M'}(s) \rangle = \lambda_J \delta_{JJ'} \delta_{MM'} (1/\epsilon\sqrt{\pi}) \exp[-(t-s)^2/\epsilon^2]$,[18] we see that the parameter λ_J characterizes the strength of the $(2)^J$-pole component of the interaction. An anologous conclusion would hold in a weak coupling model.[19,20]

If the semigroup Λ_t of Proposition 2.1 is required to be positive (instead of completely positive), then, by (2.1c'), its generator L still has the form (2.5) with the coefficients λ_J satisfying

$$\sum_{J=1}^{2j} \lambda_J \sum_{M=-J}^{J} |\langle x | T_{JM} | y \rangle|^2 \geq 0 \tag{2.8}$$

for all orthogonal vectors $|x\rangle$ and $|y\rangle$ in $\mathbb{C}^{2j+1}$. Choosing

$$|x\rangle = |jj\rangle$$

and

$$|y^{(K)}\rangle = \sum_{l=0}^{K} \alpha_l^{(K)} |j, -j+l\rangle, \quad K = 0, 1, \ldots, 2j-1,$$

where the coefficients $\alpha_l^{(K)}$ are defined by the recurrence formula

$$\alpha^{(K)}_{2j-J} = -\sum_{l=2j-J+1}^{K} \alpha_l^{(K)} \frac{\langle jj | T_{J,2j-l} | j, -j+l \rangle}{\langle jj | T_{JJ} | j, j-J \rangle},$$

$$\alpha_K^{(K)} \neq 0,$$

one checks that

$$\sum_{M=-J}^{J} |\langle x | T_{JM} | y^{(K)} \rangle|^2 = \text{const} \times \delta_{J,2j-K},$$

where the constant is nonzero. From this it follows that $\lambda_{2j-K} \geq 0$ for $K = 0, 1, \ldots, 2j-1$, which proves that a trace preserving positive semigroup on $M(2j+1)$ which is invariant under D^j is automatically completely positive. Note that this is by no means the general situation, since condition (2.1c') is a strictly weaker requirement than the positivity of the matrix $\{c_{ij}\}$. A particularly simple example is provided by the axially symmetric relaxation of a spin 1/2 in a strong external magnetic field. As shown in Ref. 1, in this case com-

plete positivity implies the inequality $T_{\parallel} \geq \frac{1}{2}T_{\perp}$ between the longitudinal ($T_{\parallel}$) and transverse ($T_{\perp}$) relaxation times, whereas positivity alone does not impose any restriction on the values of $T_{\parallel}$ and $T_{\perp}$. It is to be expected that complete positivity should imply stronger restrictions than positivity on observable parameters also in the general case of axially symmetric relaxation of an arbitrary spin j, as well as in the case of a dynamical semigroup which is invariant under a *reducible* representation of the rotation group, which corresponds for example to the description of the relaxation of two coupled spins (such as electronic and nuclear). These problems are currently under investigation.

From (2.2) and (2.4) and Schur's lemma we have

$$LT_{JM} = -\gamma_J T_{JM} \quad (J = 0, 1, \ldots, 2j;\ M = J, J-1, \ldots, -J), \tag{2.9}$$

where $\gamma_0 = 0$, as follows from the fact that Λ_t is trace preserving, and (see Appendix A)

$$\gamma_J = \sum_{K=1}^{2j} \Gamma_{JK} \lambda_K \quad (J = 1, 2, \ldots, 2j), \tag{2.10}$$

with

$$\Gamma_{JK} = (2K+1)\left(\frac{1}{2j+1} + (-)^{2j+1+J+K} \begin{Bmatrix} K & j & j \\ J & j & j \end{Bmatrix}\right) \tag{2.11}$$

$$(J, K = 1, 2, \ldots, 2j).$$

Equation (2.10) has been previously obtained by Happer[12] in the framework of a model of a weak fluctuating perturbation with exponentially decaying two-point time correlation functions.

In Appendix B we prove the following properties (Γ^{-1} denotes the inverse of the matrix $\Gamma = \{\Gamma_{JK}\}$):

$$\Gamma_{JK} > 0 \quad (J, K = 1, 2, \ldots, 2j), \tag{2.12}$$

$$\sum_{K=1}^{2j} \Gamma_{JK} = 2j+1 \quad (J = 1, 2, \ldots, 2j), \tag{2.13}$$

$$(\Gamma^{-1})_{JK} = \Gamma_{JK} - \frac{2K+1}{2j+1} = (-)^{2j+1+J+K} \begin{Bmatrix} K & j & j \\ J & j & j \end{Bmatrix}, \tag{2.14}$$

$$A^{(+)}\Gamma_{J+1,K} + A^{(0)}\Gamma_{JK} + A^{(-)}\Gamma_{J-1,K} = \frac{2K(K+1)(2K+1)}{2j+1}(2J+1), \tag{2.15}$$

where

$$A^{(0)} = (2J+1)[2K(K+1) + J(J+1) - 4j(j+1)], \tag{2.15a}$$

$$A^{(+)} = g(J+1), \quad A^{(-)} = g(J), \tag{2.15b}$$

$$g(x) = x(2j+1-x)(2j+1+x). \tag{2.15c}$$

The parameters γ_J represent the relaxation rates of the different multipoles T_{JM} and they are the quantities which are usually more directly accessible to experimental measurements.[4] It follows from (2.6), (2.10), and (2.12) that, except in the trivial case $L = 0$, all relaxation rates are different from zero. In particular, this implies that any initial state relaxes to the terminal unpolarized state $\rho(\infty) = [1/(2j+1)]\,\mathbb{1}$.

From (2.11), the first two rows of the matrix $\{\Gamma_{JK}\}$ are easily found to be

$$\Gamma_{1K} = \frac{K(K+1)(2K+1)}{j(2j+1)(2j+2)}, \tag{2.16}$$

$$\Gamma_{2K} = \frac{6K(K+1)(2K+1)}{(2j-1)2j\cdots(2j+3)}[(2j-1)(2j+3) - (K-1)(K+2)], \tag{2.17}$$

and the recurrence relation (2.15) allows us to compute the remaining matrix elements.

Equation (2.13) implies that

$$\Gamma\,|e\rangle = (2j+1)\,|e\rangle,$$

where $\langle jm\,|\,e\rangle = \mu$ ($\mu \in \mathbb{C}$) for all $m = j, j-1, \ldots, -j$ (by (2.12) there are no other linearly independent eigenvectors with positive components[21]). In other words, all relaxation rates are the same if and only if the strengths of the different multipole components of the interaction are all equal.

By (2.10) and (2.6), the relaxation rates γ_J must satisfy the set of inequalities

$$\sum_{K=1}^{2j} (\Gamma^{-1})_{JK}\gamma_K \geq 0 \quad (J = 1, 2, \ldots, 2j). \tag{2.18}$$

Equation (2.18) defines a closed hyperpyramid which, except for its vertex, is contained in the open positive hyperoctant. For example, in the lowest nontrivial cases $j = 1$ and $j = \frac{3}{2}$ we have

$$\tfrac{3}{5}\gamma_1 \leq \gamma_2 \leq 3\gamma_1 \quad \text{for } j = 1, \tag{2.19}$$

and

$$\left.\begin{aligned} &9\gamma_1 - 5\gamma_2 + \gamma_3 \geq 0 \\ &3\gamma_1 - 15\gamma_2 + 7\gamma_3 \leq 0 \\ &11\gamma_1 + 5\gamma_2 - 21\gamma_3 \leq 0 \end{aligned}\right\} \text{ for } j = \tfrac{3}{2}. \tag{2.20}$$

Relation (2.18) is equivalent to

$$\min_{1\leq K\leq 2j}\left(\frac{\Gamma_{JK}}{\Gamma_{J'K}}\right) \leq \frac{\gamma_J}{\gamma_{J'}} \leq \max_{1\leq K\leq 2j}\left(\frac{\Gamma_{JK}}{\Gamma_{J'K}}\right) \tag{2.21}$$

and, for a given value of K, $\Gamma_{JK}/\Gamma_{J'K}$ is the value of the ratio $\gamma_J/\gamma_{J'}$ for an interaction of pure polarity $(2)^K$ (i.e., $\lambda_J = \lambda\delta_{JK}$; compare Ref. 12, Fig. 1). In particular, using (2.16) and (2.17) and the recurrence formula (2.15), one finds

$$\frac{3}{2j+3} \leq \frac{\gamma_2}{\gamma_1} \leq 3 \tag{2.22}$$

and

$$\gamma_3 \leq 6\gamma_1. \tag{2.23}$$

There is no simple expression for the lower limit

$$\min_{1\leq K\leq 2j}(\Gamma_{3K}/\Gamma_{1K})$$

of γ_3/γ_1 as a function of j. Explicitly, we have, for example,

$$\min_{1\le K\le 2j}\frac{\Gamma_{3K}}{\Gamma_{1K}}=\begin{cases}\dfrac{\Gamma_{3,2j}}{\Gamma_{1,2j}}=1-\dfrac{5j}{(j+2)(2j+3)} & \text{for } j=\tfrac{3}{2},2\\[2ex] \dfrac{\Gamma_{3,2j-1}}{\Gamma_{1,2j-1}}=1-\dfrac{15j(2j-3)}{(2j-1)(j+2)(2j+3)} & \text{for } 2\le j\le\tfrac{13}{2}.\end{cases} \tag{2.24}$$

Since $\Gamma_{J1}/\Gamma_{11}=J(J+1)/2$, we have

$$\max_{1\le K\le 2j}\frac{\Gamma_{JK}}{\Gamma_{1K}}=\max\frac{\gamma_J}{\gamma_1}\ge\tfrac{1}{2}J(J+1). \tag{2.25}$$

We suspect that (2.25) holds with the equality sign (compare Ref. 12), but we have not been able to find a proof of this conjecture.

3. THE PARAMETRIZATION $W_{mm'}$

Another useful parametrization of the generator L is provided by the transition probabilities per unit time $W_{mm'}$ among the Zeeman sublevels, which appear in the Pauli equation

$$\frac{d}{dt}\rho_m(t)=\sum_{m'=-j}^{j}[W_{mm'}\rho_{m'}(t)-W_{m'm}\rho_m(t)] \tag{3.1}$$

for the relative level populations $\rho_m(t)\equiv\rho_{mm}(t)$ $=\mathrm{tr}(P_{mm}\rho(t))=\langle jm|\rho(t)|jm\rangle$, where we use the notation $P_{mm'}=|jm\rangle\langle jm'|$. We can write the master equation (1.1) in terms of the matrix elements of $\rho(t)$ in the standard basis $|jm\rangle$, $\rho_{mm'}(t)=\mathrm{tr}(P_{m'm}\rho(t))=\langle jm|\rho(t)|jm'\rangle$, as

$$\frac{d}{dt}\rho_{mm'}(t)=\sum_{mm'nn'}L_{mm'nn'}\rho_{nn'}(t), \tag{3.2}$$

where

$$\begin{aligned}L_{mm'nn'}&=(P_{mm'}|LP_{nn'})\\&=-\sum_{JM}(P_{mm'}|T_{JM})\gamma_J(T_{JM}|P_{nn'})\\&=(-)^{2j+1-m-n}\sum_{JM}(2J+1)\begin{pmatrix}j&j&J\\m&-m'&-M\end{pmatrix}\\&\quad\times\begin{pmatrix}j&j&J\\n&-n'&-M\end{pmatrix}\gamma_J.\end{aligned} \tag{3.3}$$

In particular, for the diagonal elements, we have

$$\begin{aligned}L_{mmm'm'}&=(P_{mm}|LP_{m'm'})\\&=(-)^{2j+1-m-m'}\sum_{J=1}^{2j}(2J+1)\begin{pmatrix}j&j&J\\m&-m&0\end{pmatrix}\\&\quad\times\begin{pmatrix}j&j&J\\m'&-m'&0\end{pmatrix}\gamma_J,\end{aligned} \tag{3.4}$$

and

$$\begin{aligned}\gamma_J&=-(T_{J0}|LT_{J0})\\&=-\sum_{mm'}(T_{J0}|P_{mm})(P_{mm}|LP_{m'm'})(P_{m'm'}|T_{J0})\\&=\sum_{mm'}(-)^{2j+1-m-m'}(2J+1)\begin{pmatrix}j&j&J\\m&-m&0\end{pmatrix}\\&\quad\times\begin{pmatrix}j&j&J\\m'&-m'&0\end{pmatrix}L_{mmm'm'}.\end{aligned} \tag{3.5}$$

Inserting (3.5) into (3.3) we find that the off-diagonal elements of the matrix $\{L_{mm'nn'}\}$ are completely determined by the diagonal ones. Furthermore, setting $m=m'$ in Eq. (3.2), we get (3.1) with

$$W_{mm'}=\delta_{mm'}+L_{mmm'm'}. \tag{3.6}$$

Therefore, under the assumption of space isotropy, coherence effects do not play any role in the evolution of the relative populations.

It follows from (3.4) and (3.5) that only $2j$ of the $(2j+1)^2$ transition probabilities $W_{mm'}$ are independent. Upon insertion of (3.5) into (3.4) and using (3.6), we obtain the relations satisfied by the $W_{mm'}$

$$\begin{aligned}W_{mm'}=\frac{1}{2j+1}+(-)^{2j-m-m'}\sum_{J=1}^{2j}\sum_{n,n'=-j}^{j}(-)^{2j-n-n'}\\\times(2J+1)^2\begin{pmatrix}j&j&J\\m&-m&0\end{pmatrix}\begin{pmatrix}j&j&J\\m'&-m'&0\end{pmatrix}\begin{pmatrix}j&j&J\\n&-n&0\end{pmatrix}\\\times\begin{pmatrix}j&j&J\\n'&-n'&0\end{pmatrix}W_{nn'}.\end{aligned} \tag{3.7}$$

Of the above relations, the following ones can also be read directly from (3.4)

$$W_{mm'}=W_{m'm}, \tag{3.8}$$

$$W_{mm'}=W_{-m-m'}, \tag{3.9}$$

$$\sum_m W_{mm'}=1, \tag{3.10}$$

whereas the remaining ones have to be computed explicitly from (3.7). For example, if $j=1$, (3.8)—(3.10) are the only independent relations implied by (3.7). Choosing W_{10} and W_{1-1} as the two independent transition probabilities, we have by (3.4)

$$\left.\begin{aligned}W_{10}&=\gamma_2/3,\\W_{1-1}&=(\gamma_1/2)-(\gamma_2/6).\end{aligned}\right\} \tag{3.11}$$

If $j=3/2$, using (3.8)—(3.10), one derives from (3.7) the following extra constraint

$$\begin{aligned}&4W_{(3/2)(1/2)}-4W_{(3/2)(-1/2)}+3W_{(3/2)(-3/2)}-3W_{(1/2)(-1/2)}\\&\quad=0\end{aligned} \tag{3.12}$$

and, by (3.4), we have

$$\begin{aligned}W_{(3/2)(1/2)}&=\tfrac{1}{20}(-3\gamma_1+5\gamma_2+3\gamma_3),\\W_{(3/2)(-1/2)}&=\tfrac{1}{20}(3\gamma_1+5\gamma_2-3\gamma_3),\\W_{(3/2)(-3/2)}&=\tfrac{1}{20}(9\gamma_1-5\gamma_2+\gamma_3),\\W_{(1/2)(-1/2)}&=\tfrac{1}{20}(\gamma_1-5\gamma_2+9\gamma_3).\end{aligned} \tag{3.13}$$

Equations (3.12) and (3.13) have been previously reported by Papp and Franz.[22]

Using (3.4), (3.6), and (2.10), one obtains the following formula which expresses the $W_{mm'}$ as functions of the parameters λ_J:

$$\begin{aligned}W_{mm'}=\delta_{mm'}\left[1-\frac{1}{2j+1}\sum_{J=1}^{2j}(2J+1)\lambda_J\right]\\+\sum_{J=1}^{2j}\sum_{M=-J}^{J}(2J+1)\begin{pmatrix}j&j&J\\m&-m'&-M\end{pmatrix}\begin{pmatrix}j&j&J\\m&-m'&-M\end{pmatrix}\lambda_J.\end{aligned} \tag{3.14}$$

From this equation it is easy to derive the usual selection rules on the transition probabilities, depending on the polarities of the interactions. For instance, if $\lambda_J = 0$ for $J > K$, we get the well known selection rule $|\Delta m| \leq K$.

Finally, imposing the positivity of the $W_{mm'}$, one obtains from (3.4) and (3.6) a set of inequalities for the relaxation rates γ_J (compare, e.g., Ref. 22). However, these inequalities are strictly weaker than those expressed by (2.18). This is a purely quantum mechanical effect, which follows from the fact that the expression $\mathrm{tr}[P_1 L(P_2)]$ must be nonnegative for an arbitrary pair of mutually orthogonal one dimensional self-adjoint projections P_1 and P_2 (compare theorem 5 of Ref. 16).

ACKNOWLEDGMENTS

We are grateful to A. Kossakowski who inspired our interest in the problem, and we acknowledge useful discussions with A. Frigerio, G. Parravicini, and A. Zecca.

APPENDIX A

Equations (2.9) and (2.5) give

$$\gamma_J = -\,\mathrm{tr}[T^*_{JM} L(T_{JM})] = \sum_{K=1}^{2j} \Gamma_{JK}\lambda_K \quad (J = 1, 2, \ldots, 2j),$$

with

$$\Gamma_{JK} = \tfrac{1}{2} \sum_{Q=-K}^{K} \mathrm{tr}(T^*_{JM}[T^*_{KQ}, [T_{KQ}, T_{JM}]]) \tag{A1}$$

or, equivalently, by the completeness of the irreducible spherical tensors,

$$\Gamma_{JK} = \tfrac{1}{2} \sum_{K'=1}^{2j} \sum_{Q=-K}^{K} \sum_{Q'=-K'}^{K'} \mathrm{tr}(T^*_{K'Q'}[T_{KQ}, T_{JM}]) \times \mathrm{tr}(T^*_{JM}[T^*_{KQ}, T_{K'Q'}]). \tag{A2}$$

Then, one obtains (2.11) by inserting into (A2) the explicit expression of $\mathrm{tr}(T^*_{J_1M_1}[T_{J_2M_2}, T_{J_3M_3}])$ in terms of products of 3-j symbols and using standard identities and sum rules for the 3-j and 6-j symbols.

APPENDIX B

We can rewrite (A1) as

$$\Gamma_{JK} = \tfrac{1}{2} \sum_{Q=-K}^{K} \|[T_{KQ}, T_{JM}]\|_2^2, \tag{B1}$$

where $\|A\|_2 = \sqrt{\mathrm{tr}(A^*A)}$ denotes the Hilbert–Schmidt norm. Then, since for example the matrix element $\langle jj|[T_{K1}, T_{J-1}]|jj\rangle$ is nonzero for all $J, K = 1, 2, \ldots, 2j$, it follows that the rhs of (B1) does not vanish, which proves (2.12).

From Eq. (2.11) we have

$$\sum_{K=1}^{2j} \Gamma_{JK} = \frac{(2j+1)^2 - 1}{2j+1} + (-)^{J+1} \times \sum_{K=1}^{2j} (-)^{2j+K}(2K+1) \begin{Bmatrix} K & j & j \\ J & j & j \end{Bmatrix}$$

and (2.13) follows from the identity

$$\sum_{K=1}^{2j} (-)^{2j+K}(2K+1) \begin{Bmatrix} K & j & j \\ J & j & j \end{Bmatrix} = \frac{(-)^{J+1}}{2j+1} \quad (J = 1, \ldots, 2j). \tag{B2}$$

In order to prove (2.14), we need the relation

$$(\Gamma^2)_{JK} = \delta_{JK} + (2K+1) \tag{B3}$$

which follows from (2.11), (B2), and the identity

$$\sum_{K'=1}^{2j} (2K'+1) \begin{Bmatrix} J & j & j \\ K' & j & j \end{Bmatrix} \begin{Bmatrix} K' & j & j \\ K & j & j \end{Bmatrix} = \frac{\delta_{JK}}{2K+1} + \frac{(-)^{J+K+1}}{(2j+1)^2}.$$

Then, by (B3) and (2.13), we have

$$\sum_{K'=1}^{2j} \Gamma_{JK'}\left(\Gamma_{K'K} - \frac{2K+1}{2j+1}\right) = (\Gamma^2)_{JK} - \frac{2K+1}{2j+1} \sum_{K'=1}^{2j} \Gamma_{JK'} = \delta_{JK}.$$

Finally, the recurrence relation (2.15) is obtained from (2.11) and the analogous recurrence relation satisfied by the 6-j symbols.[23]

Note added in proof: After this paper went into print, we discovered the important review work on optical pumping and relaxation by A. Omont, Progr. Quantum Electronics, 5, 69 (1977). In Appendix B of the paper the author derives a general form for the rotationally invariant generator in the reducible case, using the general phenomenological theory of relaxation. In the special case of reducibility he obtains formulas (2.10) and (2.11) and explicitly derives Eq. (2.22).

[1] V. Gorini, A. Kossakowski, and E.C.G. Sudarshan, J. Math. Phys. **17**, 821 (1976).
[2] G. Lindblad, Commun. Math. Phys. **48**, 119 (1976).
[3] E.B. Davies, *Quantum Theory of Open Systems* (Academic, New York, 1976).
[4] W. Happer, Rev. Mod. Phys. **44**, 169 (1972).
[5] M. Broyer, G. Gouedard, J.C. Lehmann, and J. Vigué, in *Advances in Atomic and Molecule Physics*, edited by D.R. Bates and B. Bederson (Academic, New York, 1976), Vol. 12, p. 165.
[6] F. Haake, in *Springer Tracts in Modern Physics* (Springer-Verlag, Berlin, 1973), Vol. 66, p. 98.
[7] A. Abragam, *The Principles of Nuclear Magnetism* (Clarendon, Oxford, 1961).
[8] P.N. Argyres and P.L. Kelley, Phys. Rev. **134**, A98 (1964).
[9] A. Omont, J. Phys. (Paris) **26**, 26 (1965).
[10] A.I. Okunevich and V.I. Perel', Sov. Phys.—JETP **31**, 356 (1970); see also M.I. D'Yakonov and V.I. Perel', Sov. Phys.—JETP **21**, 227 (1965).
[11] H. Gabriel, Phys. Rev. **181**, 506 (1969).
[12] W. Happer, Phys. Rev. B **1**, 2203 (1970).
[13] W.F. Stinespring, Proc. Am. Math. Soc. **6**, 211 (1955).
[14] M.D. Choi, Linear Alg. Appl. **10**, 285 (1975).
[15] K. Kraus, Ann. Phys. (N.Y.) **64**, 311 (1971).
[16] A. Kossakowski, Bull. Acad. Pol. Sci. Sér. Math. Astron. Phys. **20**, 1021 (1972).
[17] A. Messiah, *Mécanique Quantique* (Dunod, Paris, 1960), App.. C.
[18] V. Gorini and A. Kossakowski, J. Math. Phys. **17**, 1298 (1976); see also Ref. 12.
[19] E.B. Davies, Commun. Math. Phys. **39**, 91 (1974).
[20] V. Gorini, A. Frigerio, M. Verri, A. Kossakowski, and E.C.G. Sudarshan, Rep. Math. Phys. **12**, 359 (1977).
[21] F.R. Gantmacher, *The Theory of Matrices* (Chelsea, New York, 1964), Chap. XIII, Rem. 3.
[22] J.F. Papp and F.A. Franz, Phys. Rev. A **5**, 1763 (1972).
[23] A.P. Yutsis, I.B. Levinson, and V.V. Vanagas, *The Theory of Angular Momentum* (Israel Program for Scientific Translations, Jerusalem, 1962), Eq. (16.8).

Generation of stationary axisymmetric Einstein–Maxwell fields

Yukio Tanabe

Department of Physics, Nagoya University, Nagoya 464, Japan
(Received 3 February 1978)

A method is presented which allows one to generate new solutions from old ones. Kinnersley's functions u, v, w associated with the old solutions must be stationary, axially-symmetric, and satisfy a linear equation $a^*u + b^*v - c^*w = 0$. The method generalizes work of Bonner and Misra *et al.* who obtained a electric/magnetic dipole solution from the Kerr solution.

I. INTRODUCTION

Mainly because of the complicated nonlinear structure it is a very difficult problem to find new solutions of the Einstein–Maxwell equations. Several authors attacked this problem by seeking methods by which we can generate new solutions from the known ones.[1–7] In this paper we formulate a new generating method which generalizes work of Bonner[6] and Misra *et al.*[7] The method involves an unfamiliar symbol j with properties $j^2 = -1$, $j^* = j$. Its mathematical implication is discussed in the Appendix.

II. FIELD EQUATION

We consider stationary Einstein–Maxwell fields whose metric can be written as

$$ds^2 = f(dt + \omega_m\, dx^m)^2 - f^{-1}\gamma_{mn}\, dx^m\, dx^n, \tag{2.1}$$

where the indices m,n take values $1,2,3$ and the functions f, ω_m, γ_{mn} do not depend on t. Israel and Wilson[8] showed that the Einstein–Maxwell equations are equivalent to the following simultaneous equations for two complex functions $\mathcal{E}$, Φ and the metric γ_{mn}:

$$(\mathrm{Re}\,\mathcal{E} + |\Phi|^2)\nabla^2\mathcal{E} = \nabla\mathcal{E}\cdot(\nabla\mathcal{E} + 2\Phi^*\nabla\Phi), \tag{2.2}$$

$$(\mathrm{Re}\,\mathcal{E} + |\Phi|^2)\nabla^2\Phi = \nabla\Phi\cdot(\nabla\mathcal{E} + 2\Phi^*\nabla\Phi), \tag{2.3}$$

$$R_{mn}(\gamma) = (\mathrm{Re}\,\mathcal{E} + |\Phi|^2)^{-2}[(1/2)\mathcal{E}_{(m}\mathcal{E}^*_{n)} + \Phi\mathcal{E}_{(m}\Phi^*_{n)} + \Phi^*\mathcal{E}^*_{(m}\Phi_{n)} - (\mathcal{E} + \mathcal{E}^*)\Phi_{(m}\Phi^*_{n)}], \tag{2.4}$$

where $\nabla = (\nabla^1, \nabla^2, \nabla^3)$ is the covariant derivative with respect to γ_{mn}, $R_{mn}(\gamma)$ the Ricci tensor associated with γ_{mn}, and $\chi_{(m}\psi_{n)}$ the operation $\frac{1}{2}[(\partial_m\chi)(\partial_n\psi) + (\partial_n\chi)(\partial_m\psi)]$. From $\mathcal{E}$ and $\Phi(=A_0 + iA')$ the functions f, ω_m and the electromagnetic field $F_{\mu\nu}$ are obtained by the formulas

$$f = \mathrm{Re}\,\mathcal{E} + |\Phi|^2,$$

$$\partial_m\omega_n - \partial_m\omega_n = -\epsilon_{mnk}\gamma^{kl}f^{-2}\,\mathrm{Im}(\partial_l\mathcal{E} + 2\Phi^*\partial_l\Phi),$$

$$F_{0m} = \partial_m A_0,$$

$$F_{mn} = \omega_m\partial_n A_0 - \omega_n\partial_m A_0 + f^{-1}\epsilon_{mnk}\gamma^{kl}\partial_l A',$$

where ϵ_{mnk} is the totally antisymmetric tensor with $\epsilon_{123} = [\det(\gamma_{mn})]^{1/2}$ and γ^{mn} the inverse matrix of γ_{mn}.

Kinnersley[5] replaced $\mathcal{E}$ and Φ by three complex functions u,v,w:

$$\mathcal{E} = \frac{u-w}{u+w}, \quad \Phi = \frac{v}{u+w}. \tag{2.5}$$

The equations (2.2) and (2.3) become

$$wU - uW = 0, \tag{2.6}$$

$$-v(U+W) + (u+w)V = 0, \tag{2.7}$$

where

$$U = (|u|^2 + |v|^2 - |w|^2)\nabla^2 u - 2(u^*\nabla u + v^*\nabla v - w^*\nabla w)\cdot\nabla u,$$

$$V = (|u|^2 + |v|^2 - |w|^2)\nabla^2 v - 2(u^*\nabla u + v^*\nabla v - w^*\nabla w)\cdot\nabla v,$$

$$W = (|u|^2 + |v|^2 - |w|^2)\nabla^2 w - 2(u^*\nabla u + v^*\nabla v - w^*\nabla w)\cdot\nabla w.$$

Concerning the replacement (2.5) Kinnersley argued as follows: "There is redundancy in a description of this sort. In particular, we may choose w to obey any field equation we please, in order to obtain simple ones for u, v." His choice is $W = 0$. Then we have $U = V = 0$.

However it is more convenient to choose $W = Fw$ so that we have

$$U = Fu, \quad V = Fv, \quad W = Fw, \tag{2.8}$$

where F may be a function of x^m. The function F is closely related to the fact that only the ratio of u,v,w enter into Eq. (2.5). In fact if a set of functions (u,v,w) satisifes Eq. (2.8), a new set $(u',v',w') = (h^{-1}u, h^{-1}v, h^{-1}w)$ also satisfies Eq. (2.8) with the new function F' given by

$$F' = h^{-1}|h|^{-2}[hF - (|u|^2 + |v|^2 - |w|^2)\nabla^2 h + 2(u^*\nabla u + v^*\nabla v - w^*\nabla w)\cdot\nabla h],$$

where h is an arbitrary function of x^m.

In Sec. IV we shall formulate a generating method on the basis of Eq. (2.8). For later convenience we rewrite Eqs. (2.8) and (2.4) in the form

$$\langle \mathbf{u}^*, \mathbf{u}\rangle\nabla^2\mathbf{u} - 2u^*_\alpha\nabla u^\alpha\cdot\nabla\mathbf{u} = F\mathbf{u}, \tag{2.9}$$

$$R_{mn}(\gamma) = \frac{2}{\langle \mathbf{u}^*, \mathbf{u}\rangle^2}[\langle \mathbf{u}^*, \mathbf{u}_{(m}\rangle\langle \mathbf{u}^*_{n)}, \mathbf{u}\rangle - \langle \mathbf{u}^*, \mathbf{u}\rangle\langle \mathbf{u}^*_{(m}, \mathbf{u}_{n)}\rangle], \tag{2.10}$$

where

$$\mathbf{u} = \{u^\alpha\} = (u,v,w),$$

$$\langle \mathbf{u}^*, \mathbf{u}\rangle = |u|^2 + |v|^2 - |w|^2 = \eta_{\alpha\beta}u^{*\alpha}u^\beta,$$

$$u_\alpha = \eta_{\alpha\beta}u^\beta, \quad \|\eta_{\alpha\beta}\| = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}. \tag{2.11}$$

 0022-2488/78/1909-1808$1.00

III. OLD SOLUTIONS

Old solutions must fulfil the following two requirements:

(A) The metric γ_{mn} must be of such a form that Eq. (2.9) is practically independent of γ_{mn}. This implies that the metric γ_{mn} is determined after the solution $\mathbf{u}$ of Eq. (2.9) is obtained. In this paper we consider the axially-symmetric case in which the metric takes the form

$$ds^2 = f(dt - \omega d\phi)^2 - f^{-1}[\exp(2\gamma)(d\rho^2 + dz^2) + \rho^2 d\phi^2] \tag{3.1}$$

where f, ω, and γ are functions of ρ and z.

(B) Kinnersley's vector $\mathbf{u}$ associated with old solutions must satisfy

$$\langle \mathbf{a}^*, \mathbf{u} \rangle = a^* u + b^* v - c^* w = 0, \tag{3.2}$$

where $\mathbf{a} = (a, b, c)$ is a constant vector. The equation (3.2) implies that $\mathbf{u}$ lies in the plane orthogonal to $\mathbf{a}$. We introduce two linearly-independent constant vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ on this plane such that

$$\langle \mathbf{a}^*, \mathbf{e}_A \rangle = 0, \quad \langle \mathbf{e}_A{}^*, \mathbf{e}_B \rangle = \tilde{\eta}_{AB}, \quad A, B = 1, 2. \tag{3.3}$$

Then the vector $\mathbf{u}$ can be written in the form

$$\mathbf{u} = z^1 \mathbf{e}_1 + z^2 \mathbf{e}_2 . \tag{3.4}$$

Because $\mathbf{u}$ is a solution of Eq. (2.9), the functions z^A satisfy the following equation:

$$\langle \mathbf{z}^*, \mathbf{z} \rangle \nabla^2 \mathbf{z} - 2 z_A{}^* \nabla z^A \cdot \nabla \mathbf{z} = F \mathbf{z}, \tag{3.5}$$

where

$$\begin{aligned} &\mathbf{z} = (z^1, z^2), \ z_A = \tilde{\eta}_{AB} z^B, \\ &\langle \mathbf{z}^*, \mathbf{z} \rangle = \tilde{\eta}_{AB} z^{*A} z^B . \end{aligned} \tag{3.6}$$

This two-component Kinnersley's vector $\mathbf{z}$ will be used to generate new solutions.

In the remaining part of this section we investigate what solutions satisfy the condition (3.2).[9] Since the condition (3.2) is SU(2,1)-symmetric we may consider, without loss in generality, only three cases in which the vector $\mathbf{a}$ take the following forms

$$\begin{aligned} &\mathbf{a}_+ = (0,1,0), \quad \langle \mathbf{a}_+^*, \mathbf{a}_+ \rangle = 1, \\ &\mathbf{a}_0 = (0,1,1), \quad \langle \mathbf{a}_0^*, \mathbf{a}_0 \rangle = 0, \\ &\mathbf{a}_- = (0,0,1), \quad \langle \mathbf{a}_-^*, \mathbf{a}_- \rangle = -1. \end{aligned} \tag{3.7}$$

Then the vectors $\mathbf{u}$ which satisfy Eq. (3.2) become $(u^+, 0, w^+)$, (u^0, w^0, w^0) and $(u^-, v^-, 0)$.[10] If we choose the basis vectors $\mathbf{e}_A$ in the form

$$\begin{aligned} &\mathbf{e}_1^+ = (1,0,0), \quad \mathbf{e}_2^+ = (0,0,1), \\ &\mathbf{e}_1^0 = (1,0,0), \quad \mathbf{e}_2^0 = (0,1,1), \\ &\mathbf{e}_1^- = (1,0,0), \quad \mathbf{e}_2^- = (0,1,0), \end{aligned} \tag{3.8}$$

the metric $\tilde{\eta}_{AB}$ and the vector $\mathbf{z}$ are given by

$$\begin{aligned} &\| \tilde{\eta}_{AB}^+ \| = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \mathbf{z}^+ = (u^+, w^+), \\ &\| \tilde{\eta}_{AB}^0 \| = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{z}^0 = (u^0, w^0), \\ &\| \tilde{\eta}_{AB}^- \| = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{z}^- = (u^-, v^-). \end{aligned} \tag{3.9}$$

Since only the ratio of u, v, and w is significant, these solutions are uniquely described by the potentials

$$\xi^+ = \frac{u^+}{w^+}, \quad \xi^0 = \frac{u^0}{w^0}, \quad \xi^- = \frac{u^-}{v^-}. \tag{3.10}$$

By setting $w^+ = w^0 = v^- = 1$ in Eq. (3.5) we see that the potentials ξ^+, ξ^0, ξ^- must satisfy

$$(|\xi^+|^2 - 1)\nabla^2 \xi^+ = 2\xi^{+*} \nabla \xi^+ \cdot \nabla \xi^+, \tag{3.11}$$

$$\nabla^2 (1/\xi^0) = 0, \tag{3.12}$$

$$(|\xi^-|^2 + 1)\nabla^2 \xi^- = 2\xi^{-*} \nabla \xi^- \cdot \nabla \xi^- . \tag{3.13}$$

The equation (3.11) is the well-known Ernst equation[11] and its known solutions are those discovered by Weyl,[12] Kerr,[13] and Tomimatsu–Sato.[14] Since Eq. (3.12) is a linear equation, the potential ξ^0 determines a very wide class of solutions, which have been discussed by Perjés[15] and Israel–Wilson.[8] Some solutions of Eq. (3.13) have been obtained by the present author.[9] Hence we have many solutions which can be used as old solutions. Simple examples are the followings[9]:

$$\begin{aligned} &\xi^+ = px - iqy \quad (p^2 + q^2 = 1), \\ &\xi^0 = \hat{x} - iy, \\ &\xi^- = p\hat{x} - i\bar{q}y \quad (p^2 - \bar{q}^2 = 1), \end{aligned} \tag{3.14}$$

where (x, y) and $(\hat{x}, y)$ are the prolate and the oblate coordinates defined by[16]

$$\begin{aligned} &\rho = \kappa[(x^2 - 1)(1 - y^2)]^{1/2}, \quad z = \kappa x y, \\ &\rho = \hat{\kappa}[(\hat{x}^2 + 1)(1 - y^2)]^{1/2}, \quad z = \hat{\kappa} \hat{x} y. \end{aligned} \tag{3.15}$$

IV. NEW SOLUTIONS

We seek new solutions which fulfil the requirement (A) of the last section and whose Kinnersley's vector u^α is given by

$$u^\alpha = b^\alpha{}_{AB} z^A z^{*B} \tag{4.1}$$

where $b^\alpha{}_{AB}$ is constant. It is easy to see that the form (4.1) is a solution to Eq. (2.9) provided

$$\langle \mathbf{b}_{AB}^*, \mathbf{b}_{CD} \rangle = \sigma \tilde{\eta}_{AB} \tilde{\eta}_{DC} + \tau \tilde{\eta}_{AC} \tilde{\eta}_{DB}, \tag{4.2}$$

$$\tilde{\eta}_{AB} \mathbf{b}_{CD} + \tilde{\eta}_{DC} \mathbf{b}_{BA} - \tilde{\eta}_{AC} \mathbf{b}_{BD} - \tilde{\eta}_{DB} \mathbf{b}_{CA} = 0, \tag{4.3}$$

where σ and τ are real constant ($\sigma + \tau \neq 0$) and $\mathbf{b}_{AB}$ is a vector with components $(b^1{}_{AB}, b^2{}_{AB}, b^3{}_{AB})$. Using Eq. (4.2) we obtain

$$\langle \mathbf{u}^*, \mathbf{u} \rangle = (\sigma + \tau) \langle \mathbf{z}^*, \mathbf{z} \rangle^2 . \tag{4.4}$$

Because $\mathbf{u}$ can be replaced by $h^{-1}\mathbf{u}$ we may set

$$\sigma + \tau = \tfrac{1}{2}, \tag{4.5}$$

without loss in generality.[17] From Eqs. (4.1) and (2.10) we obtain

$$[R_{mn}(\gamma)]_{\text{new}} = 4\tau [R_{mn}(\gamma)]_{\text{old}} . \tag{4.6}$$

This gives, for the metric (3.1), the relation

$$\gamma_{\text{new}} = 4\tau \gamma_{\text{old}} . \tag{4.7}$$

Later we show that $\tau = 1$ [see Eqs. (4.10) and (4.11)].

Now we determine the vectors $\mathbf{b}_{AB}$ which correspond to the metrics $\tilde{\eta}^{\pm}_{AB}{}^{0}$ given in Eq. (3.9). The condition (4.3) implies the relations

$$\mathbf{b}^{+}_{11}=\mathbf{b}^{+}_{22},\quad \mathbf{b}^{0}_{22}=0,\quad \mathbf{b}^{-}_{11}=-\mathbf{b}^{-}_{22}, \tag{4.8}$$

for $(AB,CD)=(11,22)$, $(22,11)$, $(12,12)$, $(21,21)$ and it becomes the identity $0=0$ for other indices. Part of the condition for $\tilde{\eta}^{\pm}_{AB}$ implies

$$\sigma^{\pm}+\tau^{\pm}=\langle \mathbf{b}^{\pm *}_{11},\mathbf{b}^{\pm}_{11}\rangle=\pm\langle \mathbf{b}^{\pm *}_{11},\mathbf{b}^{\pm}_{22}\rangle=-\sigma^{\pm}$$

or

$$2\sigma^{\pm}+\tau^{\pm}=0. \tag{4.9}$$

Combining Eq. (4.9) with Eq. (4.5) we obtain

$$\sigma^{\pm}=-\tfrac{1}{2}\qquad \tau^{\pm}=1. \tag{4.10}$$

Since the condition (4.2) for $\tilde{\eta}^{0}_{AB}$ contains σ^{0} and τ^{0} only in the form $\sigma^{0}+\tau^{0}$, we may choose

$$\sigma^{0}=-\tfrac{1}{2},\quad \tau^{0}=1. \tag{4.11}$$

Using the values (4.10) and (4.11), we tabulate the condition (4.2) for each case:

$\langle \mathbf{b}^{+*}_{AB},\mathbf{b}^{+}_{CD}\rangle$	00	12	21
00	1	0	0
12	0	−1	0
21	0	0	−1

$\langle \mathbf{b}^{-*}_{AB},\mathbf{b}^{-}_{CD}\rangle$	00	12	21
00	1	0	0
12	0	1	0
21	0	0	1

$\langle \mathbf{b}^{0*}_{AB},\mathbf{b}^{0}_{CD}\rangle$	00	12	21
00	1	0	0
12	0	0	0
21	0	0	0

(4.12)

where

$$\mathbf{b}^{\pm}_{00}=\sqrt{2}\,\mathbf{b}^{\pm}_{11}=\pm\sqrt{2}\,\mathbf{b}^{\pm}_{22},\ \mathbf{b}^{0}_{00}=\sqrt{2}\,\mathbf{b}^{0}_{11}. \tag{4.13}$$

The conditions for $\mathbf{b}^{\pm}_{AB}$ cannot be satisfied by the usual complex vectors, because the inner product $\langle \mathbf{b}^{*},\mathbf{b}\rangle$ has the signature $(+,+,-)$. Fortunately, however, as is shown in the Appendix the vectors u^{α} can contain a symbol j with properties

$$j^{2}=-1,\quad j^{*}=j. \tag{4.14}$$

The symbol j in the vectors u^{α} may be interpreted either as an imaginary unit in the functions f, ω_{m}, γ_{mn}, $F_{\mu\nu}$ or as an auxiliary symbol which can be eliminated from the final expression. With the help of j we can obtain many sets of vectors which satisfy the condition (4.12). The followings are simple examples:

$$\mathbf{b}^{+}_{00}=\begin{bmatrix}1\\0\\0\end{bmatrix},\quad \mathbf{b}^{+}_{12}=\frac{1}{\sqrt{2}}\begin{bmatrix}0\\-ij\\1\end{bmatrix},\quad \mathbf{b}^{+}_{21}=\frac{1}{\sqrt{2}}\begin{bmatrix}0\\ij\\1\end{bmatrix},$$

$$\mathbf{b}^{0}_{00}=\begin{bmatrix}-1\\0\\0\end{bmatrix},\quad \mathbf{b}^{0}_{12}=\frac{1}{\sqrt{2}}\begin{bmatrix}0\\j+e^{i\varphi}\\j+e^{i\varphi}\end{bmatrix},\quad \mathbf{b}^{0}_{21}=\frac{1}{\sqrt{2}}\begin{bmatrix}0\\j-e^{i\varphi}\\j-e^{i\varphi}\end{bmatrix},$$

$$\mathbf{b}^{-}_{00}=\begin{bmatrix}1\\0\\0\end{bmatrix},\quad \mathbf{b}^{-}_{12}=\frac{1}{\sqrt{2}}\begin{bmatrix}0\\1\\ij\end{bmatrix},\quad \mathbf{b}^{-}_{21}=\frac{1}{\sqrt{2}}\begin{bmatrix}0\\1\\-ij\end{bmatrix}, \tag{4.15}$$

where φ is a real parameter. Then Eq. (4.1) becomes

Class I.

$$\begin{bmatrix}u\\v\\w\end{bmatrix}=\frac{1}{\sqrt{2}}\begin{bmatrix}|u^{+}|^{2}+|w^{+}|^{2}\\2j\,\mathrm{Im}(u^{+}w^{+*})\\2\,\mathrm{Re}(u^{+}w^{+*})\end{bmatrix},$$

Class II.

$$\begin{bmatrix}u\\v\\w\end{bmatrix}=\frac{1}{\sqrt{2}}\begin{bmatrix}-|u^{0}|^{2}\\2j\,\mathrm{Re}(u^{0}w^{0*})+2\exp(i\varphi)\,\mathrm{Im}(u^{0}w^{0*})\\2j\,\mathrm{Re}(u^{0}w^{0*})+2\exp(i\varphi)\,\mathrm{Im}(u^{0}w^{0*})\end{bmatrix},$$

Class III.

$$\begin{bmatrix}u\\v\\w\end{bmatrix}=\frac{1}{\sqrt{2}}\begin{bmatrix}|u^{-}|^{2}-|v^{-}|^{2}\\2\,\mathrm{Re}(u^{-}v^{-*})\\-2j\,\mathrm{Im}(u^{-}v^{-*})\end{bmatrix}. \tag{4.16}$$

Thus starting with the solutions described by the potentials ξ^{+}, ξ^{0}, ξ^{-} we can generate new solutions by the formulas (4.16). From the potentials given in Eq. (3.14) we obtain

I. $$\begin{bmatrix}u\\v\\w\end{bmatrix}=\frac{1}{\sqrt{2}}\begin{bmatrix}p^{2}x^{2}-\bar{q}^{2}y^{2}+1\\-2\bar{q}y\\2px\end{bmatrix},\quad (p^{2}-\bar{q}^{2}=1),$$

II. $$\begin{bmatrix}u\\v\\w\end{bmatrix}=\frac{1}{\sqrt{2}}\begin{bmatrix}x^{2}-y^{2}\\2x-2\exp(i\varphi)y\\2x-2\exp(i\varphi)y\end{bmatrix}, \tag{4.17}$$

III. $$\begin{bmatrix}u\\v\\w\end{bmatrix}=\frac{1}{\sqrt{2}}\begin{bmatrix}p^{2}\hat{x}^{2}-q^{2}y^{2}-1\\2p\hat{x}\\-2qy\end{bmatrix},\quad (p^{2}+q^{2}=1),$$

where we made the following replacements[18]:

I. $jq\to\bar{q}$, II. $j\hat{x}\to x$, III. $j\bar{q}\to -q$.

The expressions (4.17) do not contain the symbol j and, if desired, we can easily acertain, without recourse to j, that they are solutions of Eq. (2.9). The first vector in Eq. (4.17) leads to the Bonner—MPST solution and it turns out that the Class I transformation corresponds to the generating method discussed by Bonner[6] and Misra *et al.*[7] The Class II, III transformations are new. Some solutions obtained by our method are listed in a separate paper.[19]

APPENDIX

In this appendix we clarify the mathematical implication of the symbol j. At first sight the j seems to be a very curious symbol: its sign does not change by complex conjugation in spite of its imaginary nature. However this apparent difficulty disappears when we realize what is meant by complex conjugation *in a complex-potential formalism*. The point is that the imaginary unit whose sign must change by complex conjugation is only the one which is introduced by hand in the course of the reformulation of the Einstein—Maxwell equations. If the potentials contain an imaginary unit of a different origin, its sign had rather remain unchanged by complex conjugation. Such a case occurs when the coordinates x^{m} and the parameters

$c^N (N=1,2,\cdots)$ take complex values.[20] We set

$$x^m = x_1^m + jx_2^m, \quad c^N = c_1^N + jc_2^N, \tag{A1}$$

where a new symbol j is employed as an imaginary unit ($j^2=-1$) in order to indicate that it has a quite different origin from the imaginary unit i used for the complex-potential formalism. For the imaginary unit j we have

$$j^* = j. \tag{A2}$$

To gain an insight into the role of the property (A2) we analyse the complex-potential formalism of Ernst.[11] Consider the following system of differential equations:

$$f\nabla^2 f = \nabla f \circ \nabla f - \nabla\varphi \circ \nabla\varphi, \quad f\nabla^2\varphi = 2\nabla f \circ \nabla\varphi. \tag{A3}$$

Ernst introduced a complex potential

$$\mathcal{E} = f + i\varphi \tag{A4}$$

and obtained a single complex equation

$$\mathrm{Re}\,\mathcal{E}\,\nabla^2\mathcal{E} = \nabla\mathcal{E} \circ \nabla\mathcal{E}. \tag{A5}$$

Usually it is considered that Eq. (A5) is equivalent to the system (A3) only when f and φ are real functions. However it can be easily shown that the formal realities

$$f^* = f, \quad \varphi^* = \varphi \tag{A6}$$

are sufficient for Eq. (A5) to be equivalent to the system (A3). In fact, under the condition (A6) the equation (A5) and its conjugate equation become

$$\begin{aligned} f\nabla^2(f+i\varphi) &= \nabla(f+i\varphi)\circ\nabla(f+i\varphi), \\ f\nabla^2(f-i\varphi) &= \nabla(f-i\varphi)\circ\nabla(f-i\varphi). \end{aligned} \tag{A7}$$

The addition and the subtraction of these equations leads to the system (A3). The role of the property (A2) is to guarantee the formal realities (A6) even when the coordinates and the parameters are complex numbers. Accordingly, if the system (A3) is satisfied by complex functions f and φ, then Eq. (A5) is satisfied by a potentail $\mathcal{E}$ containing the symbol j, and vice versa.

Similar consideration applies to the complex-potential formalism of Kinnersley. The symbol j contained in u, v, and w is interpreted as the imaginary unit which makes the functions f, ω_m, γ_{mn}, and $F_{\mu\nu}$ be complex. If j is contained in the form (A1) we can eliminate j from u, v, and w by analytic continuation.

ACKNOWLEDGMENT

The author would like to express his thanks to Professor T. Takabayasi for his encouragement.

[1] J. Ehlers, in *Les theóries relativistes de la gravitation* (CNRS), Paris, 1959).
[2] B. K. Harrison, J. Math. Phys. **9**, 1744 (1968).
[3] F. J. Ernst, Phys. Rev. **168**, 1415 (1968).
[4] R. Geroch, J. Math. Phys. **12**, 918 (1968).
[5] W. Kinnersley, J. Math. Phys. **14**, 651 (1973).
[6] W. B. Bonner, Z. Physik **161**, 439 (1961); **190**, 444 (1966).
[7] R. M. Misra *et al.*, Phys. Rev. D **7**, 1587 (1973); **8**, 1942 (1973).
[8] W. Israel and G.A. Wilson, J. Math. Phys. **13**, 865 (1972).
[9] Y. Tanable, Prog. Theor. Phys. **57**, 840 (1977).
[10] In what follows we attach the superscripts +, 0, and − to various quantities associated with $\mathfrak{a}_+$, $\mathfrak{a}_0$, and $\mathfrak{a}_-$ respectively.
[11] F. J. Ernst, Phys. Rev. **167**, 1175 (1968).
[12] H. Weyl, Ann. Phys. (Leipzig) **54**, 117 (1917).
[13] R. P. Kerr, Phys. Rev. Lett. **11**, 237 (1963).
[14] A. Tomimatsu and H. Sato, Prog. Theor. Phys. **50**, 90 (1973).
[15] Z. Perjés, Phys. Rev. Lett. **27**, 1668 (1971).
[16] B. H. Voorhees, Phys. Rev. D **2**, 2119 (1970).
[17] If $\sigma+\tau<0$, choose $h = j|2(\sigma+\tau)|^{-1/2}$ using j defined in Eq. (4.14).
[18] Note that the replacements $\hat{x} \to -jx$, $\hat{\kappa} \to j\kappa$ bring the oblate coordinates to the prolate coordinates.
[19] Y. Tanabe, "Exact Solutions of Stationary Axisymmetric Einstein—Maxwell Equations," Prog. Theor. Phys. **60** (1978) (to be published).
[20] Note that a solution of the Einstein—Maxwell equations is still a solution even if the coordinates and the parameters take complex values.

Theorem on the representations of SO(n) groups

Kurt Fisher and Henry F. Ahner

Adelphi University, Garden City, New York 11530
(Received 28 February 1978)

It is shown that SO(n) groups ($n > 5$) that possess a c-number Lorentz Casimir operator F do not have infinite-dimensional representations. Accordingly these groups cannot support new positive-energy wave equations of Staunton type. In an appendix Staunton's spin-1/2 positive-energy wave equation is extended, in an analogous manner to that used by Rarita and Schwinger (RS) on Dirac's equation, to yield positive-energy wave equations for arbitrary spin. It is noted that analogous auxiliary conditions to those of RS do not hold. Stationary solutions for low spin are listed.

I. INTRODUCTION

Two spin-$\frac{1}{2}$ equations[1] exist that are representations of the SO(3,2) group, such that the Lorentz Casimir operator F is a c-number. One of these, Dirac's equation,[2] serves as a starting point for much of modern physics. It has been extended, by Rarita and Schwinger[3] among others,[4] to provide relativistic equations for other spins. The other equation, Staunton's positive-energy wave equation,[5] has only recently been discovered and while some of its properties are known,[6] its full range of possible uses in physics have not been explored.

Positive-energy equations for other spins are expected to exist.[7] It is the purpose of this paper to show that when the Lorentz Casimir operator F is a c-number there are only finite-dimensional representations of the SO(n, m) group ($n + m > 5$).[8] Accordingly one must seek higher spin positive-energy wave equations[9] by looking in other directions.[10] In an appendix we demonstrate the analog of the Rarita–Schwinger procedure[3] for obtaining a spin-$\frac{3}{2}$ positive-energy field from the known Staunton spin-$\frac{1}{2}$ equation. This is not the same thing as finding a positive-energy spin-$\frac{3}{2}$ equation of Staunton-type because auxiliary conditions must be imposed.

In Sec. II, Staunton's equation is briefly recalled. Section III contains our theorem on the SO(n, m) group. It shown that Staunton's spin-$\frac{1}{2}$ equation is the only one constructible as an infinite-dimensional representation of an SO(n, m) group ($n + m \geq 5$) that contains the Lorentz Casimir operator as a c-number. Appendix A, which contains the Rarita–Schwinger type extension, contains one new point: The subsidiary conditions are shown to be different from those in the Rarita–Schwinger case. Solutions for stationary cases are given in Appendix B.

II. STAUNTON'S SPIN-½ EQUATION

Staunton's spin-$\frac{1}{2}$ positive energy wave equation[5] is

$$T_\mu \Psi(x^\mu, q_1, q_2) = 0, \tag{1}$$

where Ψ is a space–time scalar function and a function of internal variables q_1 and q_2,

$$T_\mu = -P_\mu + iS_{\mu\nu}P^\nu + mV_\mu, \quad P_\mu = i\partial_\mu. \tag{2}$$

The operators $S_{\mu\nu}$ and V_μ satisfy the SO(3,2) Lie algebra commutator relations

$$\begin{aligned}
&[S_{\mu\nu}, S_{\rho\sigma}] = i(g_{\mu\rho}S_{\nu\sigma} - g_{\mu\sigma}S_{\nu\rho} + g_{\nu\sigma}S_{\mu\rho} - g_{\nu\rho}S_{\mu\sigma}),\\
&[V_\mu, S_{\rho\sigma}] = i(g_{\mu\sigma}V_\rho - g_{\mu\rho}V_\sigma),\\
&[V_\mu, V_\nu] = iS_{\mu\nu}.
\end{aligned} \tag{3}$$

The operators $S_{\mu\nu}$ and V_μ may be realized by the second-order operators:

$$\begin{aligned}
S_{23} &= -\tfrac{1}{2}(q_1q_2 + \eta_1\eta_2),\\
S_{31} &= -\tfrac{1}{4}(q_1^2 + \eta_1^2 - q_2^2 - \eta_2^2),\\
S_{12} &= -\tfrac{1}{2}(q_2\eta_1 - q_1\eta_2),\\
S_{10} &= \tfrac{1}{4}(q_1^2 - \eta_1^2 - q_2^2 + \eta_2^2),\\
S_{20} &= \tfrac{1}{2}(\eta_1\eta_2 - q_1q_2),\\
S_{30} &= \tfrac{1}{2}(q_1\eta_1 + \eta_2q_2),\\
V_0 &= \tfrac{1}{4}(q_1^2 + q_2^2 + \eta_1^2 + \eta_2^2),\\
V_1 &= \tfrac{1}{2}(-q_1\eta_1 + q_2\eta_2),\\
V_2 &= \tfrac{1}{2}(q_1\eta_2 + q_2\eta_1),\\
V_3 &= \tfrac{1}{4}(q_1^2 + q_2^2 - \eta_1^2 - \eta_2^2),
\end{aligned} \tag{4}$$

where $\eta_i = -i\partial/\partial q_i$.

The free-field momentum eigenvalue solution[5] is

$$\Psi = (Aq_1 + Bq_2)\Psi_0(q,p)\exp(-ip^\mu x_\mu), \tag{5}$$

where

$$\begin{aligned}
\Psi_0(q,p) = \exp\{&-\tfrac{1}{2}(p_0 + p_3)^{-1}[m(q_1^2 + q_2^2)\\
&+ ip_1(q_1^2 - q_2^2) - 2ip_2q_1q_2]\},
\end{aligned} \tag{6}$$

and A and B are arbitrary. A and B provide the two spin-$\frac{1}{2}$ degrees of freedom. Note also that m is the particles' rest mass. All these results follow logically from the representation condition that $F = \frac{1}{4}S_{\mu\nu}S^{\mu\nu}$ be a c-number.[1]

III. THEOREM ON SO(n)

Theorem: If the Lorentz Casimir operator, F, of the internal group SO(n), $n > 5$, is a c-number, then the group does not have any infinite-dimensional representations.

Proof: Consider an internal group SO(n)[11] whose generators satisfy the commutation relations

$$[S_{AB}, S_{CD}] = i(g_{AC}S_{BD} - g_{AD}S_{BC} + g_{BD}S_{AC} - g_{BC}S_{AD}), \tag{7}$$

where capital Latin indices range over the ($n' + 4$)-directions of the space ($n' > 1$) that are labeled by four-valued lower case Greek indices and the n'-valued

 0022-2488/78/1909-1812$1.00

lower case Latin indices. Then $S_{AB}=\{S_{\mu\nu},S_{\mu i},S_{ij}\}$. The metric g_{AB} is diagonal with components of magnitude one.

Taking the Lorentz Casimir operator $F\equiv\frac{1}{4}S_{\mu\nu}S^{\mu\nu}$ to be a c-number, evaluating the commutator of $S_{\mu i}$ with F, and also the commutator of $S_{i\sigma}$ and $S_{\mu}{}^{\sigma}$, one obtains

$$S_{i\sigma}S_{\mu}{}^{\sigma}=\tfrac{3}{2}iS_{i\mu}=-S_{\mu}{}^{\sigma}S_{i\sigma}. \tag{8}$$

Now define D_{ij} as

$$D_{ij}\equiv S_{i\mu}S_{j}{}^{\mu}. \tag{9}$$

The results of the evaluation of $[S_{i\mu},S_{j}{}^{\mu}]$ and of $[S_{i\mu},S_{j\nu}]S^{\mu\nu}$, using Eq. (8), may be compared to yield

$$D_{ij}=2iS_{ij}+\tfrac{4}{3}g_{ij}F. \tag{10}$$

Upon contraction of Eq. (10) with g^{ij} and with S^{ij}, we obtain

$$D\equiv D_{ij}g^{ij}=\tfrac{4}{3}n'F=S_{i\mu}S^{i\mu}, \tag{11}$$

$$S^{ij}D_{ij}=2iS^{ij}S_{ij}. \tag{12}$$

Since $[S_{AB}S^{AB},S_{\mu\nu}]=0$ (this follows from the Lie algebra), it follows from Eq. (11) that $[S_{\mu i},S_{jk}S^{jk}]=0$. Taking $[S_{\mu i},S_{jk}S^{jk}]=0$, and expanding the commutator $[S_{i\mu},S^{i}{}_{j}]$ we find

$$S_{j\mu}S^{j}{}_{i}=-\tfrac{1}{2}i(n'-1)S_{i\mu}=S_{i}{}^{j}S_{j\mu}. \tag{13}$$

Similarly to the above, define

$$D_{\mu\nu}\equiv S^{i}{}_{\nu}, \tag{14}$$

form $[S_{i\mu},S^{i}{}_{\nu}]$ and compare it to the evaluation of $[S_{i\mu},S_{j\nu}]S^{ij}$, using (13) to obtain

$$D_{\mu\nu}=\tfrac{1}{3}n'g_{\mu\nu}F+\tfrac{1}{2}in'S_{\mu\nu}, \tag{15}$$

$$S_{ij}S^{ij}=\tfrac{1}{3}n'(n'-1)F. \tag{16}$$

Since division by $n'-1$ is involved in obtaining (15), this is the place where a difference between the present work and the case considered by Staunton and Browne ($n'=1$) occurs.[1]

Compute $[S_{i\mu},S_{j\nu}]\,S^{\mu\alpha}$ and $[S_{k\mu},S_{i\alpha}]S^{kj}$ by two methods. Obtain

$$\begin{aligned}g_{ij}(S_{\mu\nu}S^{\mu\alpha}-\tfrac{3}{2}iS_{\nu}{}^{\alpha})=&-S_{ij}(S_{\nu}{}^{\alpha}+\tfrac{3}{2}i\delta_{\nu}^{\alpha})\\&+\delta_{\nu}^{\alpha}D_{ij}-S_{i\nu}S_{j}{}^{\alpha},\end{aligned} \tag{17}$$

and

$$\begin{aligned}g_{\nu\alpha}[S_{ki}S_{j}{}^{k}-\tfrac{1}{2}i(n'-1)S_{ij}]=&\,g_{ij}[D_{\nu\alpha}-\tfrac{1}{2}i(n'-1)S_{\nu\alpha}]\\&-S_{\nu\alpha}S_{ij}-S_{i\nu}S_{j\alpha}.\end{aligned} \tag{18}$$

Contract (17) with g^{ij} and S^{ij}. Find, using (16) that

$$S_{\mu\nu}S^{\mu\alpha}=\tfrac{4}{3}F\delta_{\nu}^{\alpha}-n'^{-1}D_{\nu}{}^{\alpha}+\tfrac{3}{2}iS_{\nu}{}^{\alpha}, \tag{19}$$

$$D_{\nu\alpha}=\frac{n'}{3}F(g_{\nu\alpha}+2i\,S_{\nu\alpha}). \tag{20}$$

Similarly contract (18) with $g^{\nu\alpha}$ and g^{ij}. These operations yield

$$S_{ki}S^{k}{}_{j}=\frac{n'}{3}Fg_{ij}-\tfrac{1}{4}D_{ij}+\tfrac{1}{2}i(n'-1)S_{ij}, \tag{21}$$

and

$$D_{\nu\alpha}=\tfrac{1}{3}n'Fg_{\nu\alpha}+\tfrac{1}{2}in'S_{\nu\alpha}. \tag{22}$$

Comparison of (20) and (22) indicates that

$$F=\tfrac{3}{4}, \tag{23}$$

and therefore $D=n'$, while $S^{ij}S_{ij}=\frac{1}{4}n'(n'-1)$. Accordingly, $S_{AB}S^{AB}=\frac{1}{4}(n'+3)(n'+4)$. Since SO(3,1) is a noncompact subgroup, by assumption, the representation considered is shown to be nonunitary and finite-dimensional. The theorem is proven.

In what follows we generate the connection between the representations considered and Dirac matrices.

The other Lorentz Casimir operator, G, is given by

$$G\equiv\tfrac{1}{8}\epsilon^{\mu\nu\alpha\beta}S_{\mu\nu}S_{\alpha\beta}. \tag{24}$$

The operator $\Omega_{i\mu}$ is defined by

$$\Omega_{i\mu}\equiv[S_{i\mu},G], \tag{25}$$

and is found to be

$$\Omega_{i\mu}=\tfrac{1}{2}\epsilon_{\mu\nu\rho\sigma}S^{\nu\rho}S_{i}{}^{\sigma}. \tag{26}$$

Squaring (26) and using (22) and (23) we obtain

$$\Omega_{i\mu}\Omega^{i\mu}=-9n'/4. \tag{27}$$

This ensures that G is not a c-number. If G were a c-number, then (25) would indicate that $\Omega_{i\mu}=0$, in contradiction with (27).

By utilizing (22) and (26), we establish that

$$S_{i\alpha}\Omega^{i}{}_{\beta}=\frac{in'}{2}g_{\alpha\beta}G-\frac{3n'}{8}\epsilon_{\alpha\beta\rho\sigma}S^{\rho\sigma}, \tag{28}$$

$$\Omega^{i}{}_{\beta}S_{i\alpha}=\frac{in'}{2}g_{\alpha\beta}G-\frac{3n'}{8}\epsilon_{\alpha\beta\rho\sigma}S^{\rho\sigma}. \tag{29}$$

Multiply (26) by $S_{k}{}^{i}$ and use (13) to obtain

$$S_{ki}\Omega^{i}{}_{\alpha}=-\tfrac{1}{2}i(n'-1)\Omega_{k\alpha}. \tag{30}$$

Now multiply (29) and (28) together, contract, and utilize (10), (18), and (30) to find

$$G^{2}=-\tfrac{9}{16}. \tag{31}$$

From (15) construct

$$D_{\mu\nu}+D_{\nu\mu}=S_{i\mu}S^{i}{}_{\nu}+S_{i\nu}S^{i}{}_{\mu}=\frac{n'}{2}g_{\mu\nu}. \tag{32}$$

The results embodied in (23), (31), and (32) indicate that a finite dimensional matrix representation exists for any SO group within the scope of this discussion. It is realizable in terms of finite matrices whose elements may be grouped as Dirac γ matrices. Note that the representations permitted are nonunitary. Also note the sign of G^2. Therefore, despite the large number of scalar operators constructible, no positive-energy infinite-dimensional wave functions exist for the groups being considered.

As an example of the matrix representations, consider the six-dimensional case. In six dimensions, with $i,\ j=5,6$ one may take

$$S_{5\mu}=\tfrac{1}{2}(g_{55})^{1/2}\begin{bmatrix}\gamma_{\mu} & 0\\ 0 & -\gamma_{\mu}\end{bmatrix},$$

$$S_{6\mu}=\tfrac{1}{2}(g_{66})^{1/2}\begin{bmatrix}0 & \gamma_{\mu}\\ \gamma_{\mu} & 0\end{bmatrix}.$$

Then

$$S_{56} = -\tfrac{1}{2} i (g_{55} g_{66})^{1/2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

$$G = -\tfrac{3}{4} \begin{bmatrix} \gamma_5 & 0 \\ 0 & \gamma_5 \end{bmatrix},$$

and

$$\Omega_{5\mu} = -\tfrac{3}{4} i (g_{55})^{1/2} \begin{bmatrix} \gamma_\mu \gamma_5 & 0 \\ 0 & -\gamma_\mu \gamma_5 \end{bmatrix},$$

$$\Omega_{6\mu} = -\tfrac{3}{4} i (g_{66})^{1/2} \begin{bmatrix} 0 & \gamma_\mu \gamma_5 \\ \gamma_\mu \gamma_5 & 0 \end{bmatrix}.$$

For the general case we include, for completeness, some other interesting relations that may be obtained:

$$\begin{aligned} &S_\nu{}^\mu \Omega_{i\mu} = \tfrac{1}{2} i \Omega_{i\nu} + 4 S_{i\nu} G, \\ &\Omega_{i\mu} S_\nu{}^\mu = -\frac{i}{2} \Omega_{i\nu} + 4 G S_{i\nu}, \\ &D_{ij} \Omega^{i\mu} = -(n'-2)\, \Omega_j{}^\mu = \Omega^{i\mu} D_{ji}, \\ &S^{ij}(S_{i\mu} S_{j\nu} + S_{j\nu} S_{i\mu}) = -\tfrac{1}{2} n'(n'-1) S_{\mu\nu}, \\ &\epsilon^{\mu\nu\alpha\beta} \Omega_{i\alpha} \Omega^i{}_\beta = n' S^{\mu\nu}. \end{aligned} \tag{33}$$

ACKNOWLEDGMENT

We are grateful for helpful comments from L.P. Staunton.

APPENDIX A

Consider the SL(2, C) generator $M_{\mu\nu}$ given by

$$M_{\mu\nu} = L_{\mu\nu} + \Sigma_{\mu\nu} + S_{\mu\nu}, \tag{A1}$$

where

$$L_{\mu\nu} \equiv i(x_\mu \partial_\nu - x_\nu \partial_\mu), \quad \Sigma_{\mu\nu} \equiv -\tfrac{1}{4} i (\gamma_\mu, \gamma_\nu),$$

and γ_μ are the Dirac matrices and $S_{\mu\nu}$ are given by Eq. (4). In some cases one or more of the terms in (A1) may not be used.[12]

Now let Ψ be a multicomponent wavefunction instead of a scalar. The multiple components may variously be labeled via tensor indices, Dirac spinor indices, or both. These indices will be suppressed for most of this appendix.

The wavefunction that describes the new field is required to separately satisfy Eq. (1) for each of its components. It also must satisfy

$$W_\mu W^\mu \Psi = -m^2 j(j+1) \Psi, \tag{A2}$$

where m is the mass and j the spin. Here W_μ is the Pauli–Lubanski operator for the particular $M_{\mu\nu}$ considered. By construction, Ψ transforms like a spin-j field under the action of the Lorentz group generated by $M_{\mu\nu}$.

Several points should be noted. First, the solutions may be obtained in all cases by use of Clebsch–Gordon coefficients for addition of angular momenta.[13] This means that one may immediately write solutions for all the interesting low spin cases. These are tabulated in Appendix B.

Secondly, only a few low spin possibilities occur. There are only a few types of field for each spin that may be generated by this method. (For spin 0 there are three; for $\frac{1}{2}$ integral spin there are four; for integral spin there are five.)

Third, auxiliary conditions exactly similar to those that hold in the Rarita–Schwinger spin-$\frac{3}{2}$ theory (i.e., $P^\mu \Psi_\mu = \gamma^\mu \Psi_\mu = 0$) can not be applied here. For example, in the case of a vector field, Ψ_μ, we have

$$T_\mu \Psi_\nu = 0 = (-P_\mu + i S_{\mu\rho} P^\rho + m V_\mu) \Psi_\nu, \tag{A3}$$

which implies on contraction with V^μ [5] the Majorana equation

$$(V^\mu P_\mu - m) \Psi_\nu = 0. \tag{A4}$$

Then if $P^\mu \Psi_\mu = 0$, upon contraction of (A3) using (A4) and (3) we obtain

$$(V^\mu P_\mu - 2m) V^\nu \Psi_\nu = 0. \tag{A5}$$

In the stationary case [adding a μ index to A and B in Eq. (5)] this equation is identically satisfied and of course $\Psi_0 = 0$ ($A_0 = B_0$).

However if $V^\mu \Psi_\mu = 0$, upon contraction of (A3) using (A4) and (3) we obtain

$$P^\mu \Psi_\mu = 0.$$

The condition $V^\mu \Psi_\mu = 0$, evaluated for the stationary case solution (equating coefficients of independent internal functions of q_i to zero) gives the following relations between A_i and B_i

$$\begin{aligned} &A_1 = -iA_3 = -B_2, \\ &B_1 = iB_3 = A_2. \end{aligned}$$

Thus there are only two independent components among the eight components of Ψ_μ in this case: not enough for a spin-$\frac{3}{2}$ particle. Accordingly the condition $V^\mu \Psi_\mu = 0$ may not be imposed.

Just as Staunton's procedure[5] offers an alternative method of taking the "square root" of the Klein–Gordon equation, it should be possible to use higher spin fields to find a "square root" of Einstein's gravitational equations via positive-energy wave equations. This would be analogous to the one found by Teitelboim using standard spin-$\frac{3}{2}$ fields.[14] All questions of supersymmetries within the framework of positive-energy wave equations, as well as those of interactions between fields each with positive-energy internal coordinates, to our knowledge, remain unexplored.

APPENDIX B

There are three methods by which spin-0 wave equations may be constructed. The first is the standard Klein–Gordon scalar field equation itself. The second method involves a field with internal coordinates q_i,

q_2 and a spinor index. The third involves a field with internal coordinates, a spinor and a vector index.

The spin-eigenvalue forms of Eq. (5) occur when $B=-iA\ (+\frac{1}{2})$ and when $B=iA\ (-\frac{1}{2})$. Denote these solutions of Staunton's equation by $|S+\rangle$ and $|S-\rangle$ respectively. Denote a Dirac column spinor, with eigenvalues of Σ_{12} that are $+\frac{1}{2}$, $-\frac{1}{2}$ by $|D+\rangle$ and $|D-\rangle$ respectively. Also denote a spacelike space—time vector field whose z components are $+1, 0$, and -1 by $|V+\rangle$, $|V0\rangle$, $|V-\rangle$, respectively.

Then using Clebsch—Gordon coefficients,[13] the two new spin-0 wavefunctions satisfying Eqs. (1) and (8) are $\Psi_{D,S,0}$ and $\Psi_{D,S,V,0}$ where

$$\Psi_{D,S,0}=\tfrac{1}{2}\sqrt{2}\,[\,|D+\rangle|S-\rangle-|D-\rangle|S+\rangle\,],$$

$$\Psi_{D,S,V,0}=\tfrac{1}{3}\sqrt{3}\,[\,|D+\rangle|S+\rangle|V-\rangle-\tfrac{1}{2}\sqrt{2}(|D+\rangle|S-\rangle + |D-\rangle|S+\rangle)|V0\rangle+|D-\rangle|S-\rangle|V+\rangle].$$

There are two ways to compose spin $\frac{1}{2}$ in addition to "pure" Dirac or Staunton fields. These are by adjoining to either field a space—time vector index. When the Dirac field is so acted upon, the Rarita—Schwinger case is obtained. It includes a spin-$\frac{1}{2}$ component. The only new case involves a field $\Psi_{S,V,1/2}$ given by

$$\Psi_{S,V,1/2}\ (\text{spin up})=+\tfrac{1}{3}\sqrt{3}\,[+|S+\rangle|V0\rangle-\sqrt{2}\,|S-\rangle|V+\rangle],$$

and

$$\Psi_{S,V,1/2}(\text{spin down})=\tfrac{1}{3}\sqrt{3}\,[+\sqrt{2}\,|S+\rangle|V-\rangle-|S-\rangle|V0\rangle].$$

The new spin-$\frac{3}{2}$ field obtained, $\Psi_{S,V,3/2}$, is given by

$$\Psi_{S,V,3/2}(+\tfrac{3}{2})=|S+\rangle|V+\rangle,$$

$$\Psi_{S,V,3/2}(+\tfrac{1}{2})=\tfrac{1}{3}\sqrt{3}[\sqrt{2}\,|S+\rangle|V0\rangle+|S-\rangle|V+\rangle],$$

$$\Psi_{S,V,3/2}(-\tfrac{1}{2})=\tfrac{1}{3}\sqrt{3}\,[\sqrt{2}\,|S-\rangle|V0\rangle+|S+\rangle|V-\rangle],$$

$$\Psi_{S,V,3/2}(-\tfrac{3}{2})=|S-\rangle|V-\rangle.$$

Some of these cases may allow a minimally coupled electromagnetic interaction. This is currently being investigated. The possibility of such an interaction exists because the fields are required to satisfy, for each component separately, Staunton's equation, which permits such a minimal coupling.

[1] L. P. Staunton and S. Browne, Phys. Rev. D **12**, 1026 (1975).
[2] P. A. M. Dirac, Proc. Roy. Soc. A **117**, 610 (1928); **118**, 351 (1928).
[3] W. Rarita and J. Schwinger, Phys. Rev. **60**, 61 (1941).
[4] See for example, H. Umezawa, *Quantum Field Theory* (Interscience, New York, 1956), Chaps. 4, 5, and 8.
[5] L. P. Staunton, Phys. Rev. D **10**, 1760 (1974). Note that the representation of the SO(3, 2) group employed in the construction of this positive-energy wave equation is unitary.
[6] For classical limit see Ref. 1; for interaction with gravitational field and Lagrangian see H. F. Ahner, Phys. Rev. D **11**, 3384 (1975); **13**, 250 (1976); for connection to strings see L. P. Staunton, Phys. Rev. D **13**, 3269 (1976).
[7] L. C. Biedenharn, M. Y. Han, and H. van Dam, Phys. Rev. D **8**, 1735 (1973); this has interpretation in terms of composite systems.
[8] Representations of the SO(3, 2) group are discussed in L. Jaffe, J. Math. Phys. **12**, 882 (1971).
[9] Wave equations on spaces other than the Minkowski space—time manifold have been treated by R. Y. Cusson and L. P. Staunton, Nuovo Cimento A **17**, 303 (1973) for the case of the Poincaré group. See also the many references listed there. For a detailed discussion of the properties of the matrices in SL(2, c) invariant wave equations, see M. Lorente, P. L. Huddleston, and P. Roman, J. Math. Phys. **14**, 1495 (1973). The cases considered here are not strongly interlocking, as theirs are.
[10] One possibility is symplectic groups. See L. P. Staunton (to appear). Another possibility is to consider SO(n, m) groups where the operator F is no longer a c-number. L. Gordon is investigating this case.
[11] The procedure followed in this paper is patterned after the SO(3, 2) derivation found in the Appendix of Ref. 1.
[12] Methods used in this Appendix are well known. They have been utilized [see S. Browne, Nucl. Phys. B **79**, 70 (1974)] for the case of Dirac's noninteracting spin-0 positive-energy wave equation. P. A. M. Dirac, Proc. Roy. Soc. A **322**, 435 (1971).
[13] All relations we need are tabulated simply and conveniently in J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics* (Wiley, New York, 1952), p. 792.
[14] C. Teitelboim, Phys. Rev. Lett. **38**, 1106 (1977).

Field equations and integrability conditions for special type *N* twisting gravitational fields

F. J. Ernst and I. Hauser

Department of Physics, Illinois Institute of Technology, Chicago, Illinois 60616
(Received 16 January 1978)

We let SNT designate a special kind of twisting type *N* gravitational field solution of the Einstein field equations, viz., one for which there exists a real scalar field χ which satisfies

$(\nabla_{[\alpha}k_{\beta}-k_{[\alpha}\nabla_{\beta}\chi)\,(\nabla_{\gamma]}k_{\delta})t^{\delta}=0,$

where k_α is a principal null vector, $t^{\alpha}t_{\alpha}=t^{\alpha}k_{\alpha}=0$, and $t^{\alpha}t_{\alpha}^{*}=1$. We obtain some theorems which provide necessary and sufficient criteria for an SNT to be the twisting type *N* metric discovered by I. Hauser, in Phys. Rev. Lett. **33**, 1112 (1974). Field equations, a hierarchy of integrability conditions, and analytic techniques which are applicable to the quest for a new SNT are given. Two tractable SNT subcases are considered in detail.

1. INTRODUCTION

We let NT denote any type *N* vacuum solution of the Einstein field equations such that the principal null rays are twisting.[1] For any given NT, there exists a null tetrad k, m, t, t* where k is a principal null vector, and the corresponding connection forms are given by[2]

$$\begin{aligned} &dx^{\alpha}t^{\beta}\nabla_{\alpha}k_{\beta}=z(A^{*}k+t)=d\zeta,\\ &dx^{\alpha}(m^{\beta}\nabla_{\alpha}k_{\beta}+t^{*\beta}\nabla_{\alpha}t_{\beta})=0,\\ &dx^{\alpha}t^{*\beta}\nabla_{\alpha}m_{\beta}=h\,d\zeta. \end{aligned} \tag{1}$$

z, A, ζ, and h are complex scalar fields, and $k:=dx^{\alpha}k_{\alpha}$ and $t:=dx^{\alpha}t_{\alpha}$ are 1-forms. The real part of z is the expansion and the imaginary part of z is the twist of the principal null rays.

By a *special* NT, to be abbreviated SNT, we shall mean any NT which satisfies the condition[3] that there exists a real scalar field χ such that

$$k\wedge(Ad\zeta+A^{*}d\zeta^{*})=k\wedge d\chi. \tag{2}$$

The only NT known to date[4] is an example of an SNT; this known solution will be designated as NT_1.

The *general* NT problem has been surprisingly resistant to the efforts of numerous investigators. The authors have, therefore, recently focused their attention on the problem of finding a new SNT. For reasons which will become clear in Sec. 2, the SNT field equations are reasonably amenable to analysis and seem to offer some chance of discovering a new NT, i.e., if there exists any SNT other than NT_1.

The main objective of this paper is to present the equations, techniques, and a few theorems which we have developed in our quest for a new SNT. As evidenced by the numerous requests for details which we have received, there is an expanding interest in the type *N* twisting problem. It is our hope that the results of this paper will be useful to the newcomers as well as to current researchers in the field. We would especially like to see some involvement by mathematicians who may not be versed in general relativity but who are attracted by a curious problem in differential-functional equations.

The NT field equations which we use have been derived in a preceding paper by one of the authors (I. H.).[2] In Sec. 2, these equations are specialized to SNT and are expressed in a modified form which is more suitable for deriving their integrability conditions. A hierarchy of four integrability conditions is derived in Sec. 3.

In Sec. 4, we consider the problem of finding those SNT which have a Killing vector. So far, our efforts have yielded only NT_1, and it is conceivable that NT_1 is the only SNT with a Killing vector. The work which remains to be done to settle this question is explained.

In Sec. 5, we pursue another likely subcase of the SNT problem, one which does not necessarily involve the existence of a Killing vector. For this subcase, two additional integrability conditions are derived.

In the summary of Sec. 6, we review current work in progress and suggest further lines of inquiry.

2. THE SNT EQUATIONS

As has already been discussed in a preceding paper by Hauser,[2,3] the assumption (2) implies the existence of a coordinate system ρ, σ, ζ, ζ^{*}[5] and of a complex scalar field $\Omega=\Omega(\zeta,\zeta^{*})$ *which does not depend on* ρ *and* σ such that, for any SNT,

$$\begin{aligned} &k=p(d\sigma+\Omega d\zeta+\Omega^{*}d\zeta^{*}),\\ &t=(\rho+i\tau)d\zeta-A^{*}k,\\ &m=d\rho-\tfrac{1}{2}[p^{-1}(DD^{*}+D^{*}D)p]k-i(D\tau-2A\tau)d\zeta+i(D\tau\\ &\qquad-2A\tau)^{*}d\zeta^{*}, \end{aligned} \tag{3}$$

where

$$\rho:=\mathrm{Re}(z^{-1}), \tag{4a}$$

$$\tau:=\mathrm{Im}(z^{-1}), \tag{4b}$$

$$D:=\frac{\partial}{\partial\zeta}-\Omega\,\frac{\partial}{\partial\sigma}\ , \tag{4c}$$

$$p:=\exp(-\chi), \tag{4d}$$

[a]Research supported in part by the National Science Foundation under Grant No. PHY75-08750.

 0022-2488/78/1909-1816$1.00

$$A := D\chi, \tag{4e}$$

$$\Delta := p^{-1}\tau = \tfrac{1}{2}i(D\Omega^* - D^*\Omega). \tag{4f}$$

The only nonzero null tetrad components of the Riemann tensor are

$$m^\alpha t^{*\beta} m^\gamma t^{*\delta} R_{\alpha\beta\gamma\delta} = [(\rho + i\tau)p]^{-1}\mathring{h} \tag{5}$$

and its complex conjugate, where we always use the overhead dot to designate differentiation with respect to σ, as in

$$\mathring{h} := \frac{\partial h}{\partial \sigma}.$$

Except for ρ itself, all of the scalar fields in the above expressions are ρ-independent and are to be computed from p, Ω, and h by using Eqs. (4d)—(5). The fields p, Ω, and h are to be found by solving *a triad of equations which constitute the crux of the problem* and which are given, for any SNT, by

$$\Delta D(\Delta^{-1}DP) = -(h+\beta)P, \tag{6}$$

$$DD^*P - (\gamma/4)P - i\Delta\mathring{P} = 0, \tag{7}$$

$$D^*h = 0, \tag{8}$$

where we have restricted our chart to a domain in which $\Delta > 0$ and where

$$P := \Delta^{1/2}p, \tag{9a}$$

$$\beta := \Delta^{1/2}D^2\Delta^{-1/2}, \quad \gamma := DD^*\ln\Delta. \tag{9b, c}$$

Equation (7) is not completely independent of Eqs. (6) and (8) in the sense that Eqs. (6) and (8) imply that the left side of Eq. (7) is independent of σ.[6] However, *this does not signify that we can ignore Eq. (7) in the analysis.* Equations (6)—(8) were obtained by specializing the general NT equations given by Hauser[2] to SNT, but we have modified his equations by using P instead of p as a dependent variable. The basic reason for this modification derives from a study of the invariance group of the preceding equations, and we now consider this group.

There is a residual arbitrariness in the null tetrad and the coordinate system. Specifically, all of the preceding equations remain invariant in form under the following group of transformations[2] in which $F = F(\zeta)$ is an arbitrary analytic function of ζ, $g(\zeta, \zeta^*)$ is an arbitrary real C^∞ function of ζ and ζ^*, and α is a positive real scaling parameter:

$$\begin{aligned}
&\mathrm{k} \to \mathrm{k}\exp(\tfrac{1}{2}F + \tfrac{1}{2}F^*),\\
&\mathrm{t} \to (\mathrm{t} + \tfrac{1}{2}F'^*\mathrm{k})\exp(\tfrac{1}{2}F - \tfrac{1}{2}F^*),\\
&\mathrm{m} \to (\mathrm{m} - \tfrac{1}{2}F'\mathrm{t} - \tfrac{1}{2}F'^*\mathrm{t}^* - \tfrac{1}{4}|F'|^2)\exp(-\tfrac{1}{2}F - \tfrac{1}{2}F^*);\\
&\rho \to \rho, \quad \sigma \to \alpha\sigma + g(\zeta, \zeta^*), \quad \zeta \to \int d\zeta \exp F.
\end{aligned} \tag{10}$$

Various scalar fields which we have defined undergo the following corresponding transformations[2]:

$$A \to (A - \tfrac{1}{2}F')\exp(-F), \tag{11a}$$

$$P \to \alpha^{-1/2}P, \tag{11b}$$

$$\Delta \to \alpha\Delta\exp(-F - F^*), \tag{11c}$$

$$\Omega \to \left[\alpha\Omega - \left(\frac{\partial g}{\partial \zeta}\right)\right]\exp(-F), \tag{11d}$$

$$h \to [h - \tfrac{1}{2}F'' + \tfrac{1}{4}(F')^2]\exp(-2F), \tag{11e}$$

$$\beta \to [\beta + \tfrac{1}{2}F'' - \tfrac{1}{4}(F')^2]\exp(-2F), \tag{11f}$$

$$\gamma \to \gamma\exp(-F - F^*). \tag{11g}$$

Observe that, except for a scaling factor, P is an invariant, and DP, DD^*P, and $\Delta D(\Delta^{-1}DP)$ are tensors under the group. This is what induced us to use P instead of p as a dependent variable, and this attention to the tensor character of the terms paid off in great simplifications of Eq. (7) and of the integrability conditions which will be derived in Sec. 3.

From Eq. (11d), it is clear that g can be chosen, with $\alpha = 1$ and $F = 0$, to transform to a new Ω which satisfies the condition

$$D\Omega^* + D^*\Omega = 0. \tag{12}$$

($D\Omega = \partial\Omega/\partial\zeta$, since $\partial\Omega/\partial\sigma = 0$.) Then, from Eqs. (4f), (11d), and (12), there exists a real scalar field $\kappa = \kappa(\zeta, \zeta^*)$ such that

$$\Omega = iD\kappa, \quad \Delta = DD^*\kappa; \tag{13}$$

whereupon the general solution of Eq. (8) is given by

$$h = h(\sigma - i\kappa, \zeta). \tag{14}$$

It is the neatness of this general solution of Eq. (8) which especially makes the SNT problem simple to analyze compared with the general NT problem. With the constraint of Eq. (12), the function $g(\zeta, \zeta^*)$ in Eqs. (10) and (11) is no longer arbitrary; instead, it is given by

$$g(\zeta, \zeta^*) = \frac{1}{2i}[f(\zeta) - f(\zeta)^*], \tag{15}$$

where $f(\zeta)$ is an arbitrary analytic function of ζ. The transformation law for κ corresponding to the group (10) is given by

$$\kappa \to \alpha\kappa + \tfrac{1}{2}[f(\zeta) + f(\zeta)^*]. \tag{16}$$

The arbitrary analytic function $F(\zeta)$ does not enter into the transformation of κ.

We next show exactly how NT_1 fits into the above scheme by proving the following theorem with the aid of Eqs. (9c), (11c), (11d), (11g), (13), and (15).

Theorem I: An SNT is an NT_1 if and only if $\gamma = 0$.

The proof proceeds by first noting that NT_1 as given by Hauser[2,4] is clearly an SNT such that $\Delta =$ constant, which implies $\gamma = 0$. As regards the converse, consider any SNT such that $\gamma = 0$. Equation (9c) implies the existence of a function $F_1(\zeta)$ such that $\Delta = \exp(F_1 + F_1^*)$. Apply the transformation (11c) with $F = F_1$ and $\alpha = 1$, whereupon $\Delta \to 1$, and Eqs. (13) imply $\Omega = i\zeta^* + F_2(\zeta)$. Apply the transformation (11d) with $f = \zeta^2 + 2i\int d\zeta F_2$, whereupon $\Omega \to i(\zeta + \zeta^*)$. This is precisely the choice of Ω which led to NT_1; i.e., any NT with this Ω is an NT_1. Q.E.D.

For any given choice of $\Omega(\zeta, \zeta^*)$, we regard Eqs. (6) and (7) as a pair of linear equation in P for which h plays a role analogous to an eigenvalue. The choices of Ω for which a solution exists and the corresponding h are to be determined by analyzing the successive integrability conditions[7] for Eqs. (6) and (7) as computed with the aid of the relation

$$DD^* - D^*D = 2i\Delta\left(\frac{\partial}{\partial\sigma}\right). \tag{17}$$

We next derive the first four integrability conditions.

3. FOUR INTEGRABILITY CONDITIONS

The manipulations of this section are guided by careful adherence to the tensor character of the various terms under the group (10), and reference to Eqs. (11) will help the reader to follow the calculations.

Operate on Eq. (7) with D, and apply Eqs. (17), (6), and the identity

$$\Delta D(\Delta^{-1}\gamma) = -2D^*\beta. \tag{18}$$

After some manipulation, we obtain the first integrability condition

$$3i\Delta\dot{DP} = \tfrac{1}{2}(D^*\beta)P - \tfrac{3}{4}\gamma DP + (h+\beta)D^*P. \tag{19}$$

Divide Eq. (19) through by Δ, and operate with D^*. Use Eqs. (6)—(8), (17), and (18). The imaginary part of the result is a mere identity, but the real part yields a second integrability condition

$$\begin{aligned}\Delta^2\ddot{P} = {}& \tfrac{1}{2}(D\beta^*)DP + \tfrac{1}{2}(D^*\beta)D^*P \\ & + [\tfrac{1}{12}\Delta D(\Delta^{-1}D\beta^*) + \tfrac{1}{12}\Delta D^*(\Delta^{-1}D^*\beta) \\ & - \tfrac{1}{3}|h+\beta|^2 - \tfrac{1}{16}\gamma^2]P.\end{aligned} \tag{20}$$

Divide Eq. (19) through by Δ^2, and operate with D. Use Eqs. (6), (7), and (18). We obtain a third condition

$$\begin{aligned}& 2(D^*\beta)DP + \{\Delta^2 D[\Delta^{-2}(h+\beta)]\}D^*P + 4i(h+\beta)\Delta\dot{P} \\ & + [\gamma(h+\beta) + \tfrac{1}{2}\Delta^2 D(\Delta^{-2}D^*\beta) + 3i\Delta\dot{h}]P = 0.\end{aligned} \tag{21}$$

Divide Eq. (21) through by Δ, and operate with D^*. Use Eqs. (6)—(8), (18), and (19). In the resulting DP term, use the identity,

$$D^*(\Delta^{-1}D^*\beta) = \Delta^{-1/2}(D^{*2}D^2\Delta^{-1/2}) - |D^2\Delta^{-1/2}|^2. \tag{22}$$

We then obtain the fourth condition

$$\begin{aligned}& [2\Delta D^*(\Delta^{-1}D^*\beta) - \tfrac{4}{3}|h+\beta|^2]DP + [\tfrac{3}{2}\Delta^2 D(\Delta^{-2}D^*\beta) + i\Delta\dot{h}]D^*P \\ & + 2i(D^*\beta)\Delta\dot{P} + \{-\Delta^2(h+\beta)^*D[\Delta^{-2}(h+\beta)] + \tfrac{3}{2}\gamma D^*\beta \\ & - \tfrac{8}{3}(h+\beta)D\beta^* + \tfrac{1}{2}\Delta D^*[\Delta D(\Delta^{-2}D^*\beta)]\}P = 0.\end{aligned} \tag{23}$$

The general SNT problem is still sufficiently complicated so that it is wise to start with the analysis of some subcases, and that is precisely what we did and are still doing. We recall that NT_1 has a Killing vector.[4] The first question which we studied in any detail was that of the existence of another SNT with a Killing vector. The techniques which we are using to answer this question and some partial results which we have obtained are covered in the next section.

4. PROBLEM OF AN SNT WITH A KILLING VECTOR

Throughout Sec. 4, we assume that a Killing vector $\mathbf{K}$ exists.[8] It is then convenient to replace the complex coordinates ζ, ζ^* by real coordinates[5]

$$\xi := \mathrm{Re}(\sqrt{2}\,\zeta), \quad \eta := \mathrm{Im}(\sqrt{2}\,\zeta). \tag{24}$$

In the Appendix, we show that the transformation (10) can be applied to select the null tetrad and coordinate system such that there exist a real constant b and real scalar fields $\lambda = \lambda(\xi)$, s, and $q = q(s, \xi)$ for which the following relations hold:

$$s := \sigma\exp(-b\eta), \tag{25a}$$

$$\kappa = \lambda(\xi)\exp(b\eta), \tag{25b}$$

$$\Delta = \delta(\xi)\exp(b\eta), \quad \delta := \tfrac{1}{2}(\lambda'' + b^2\lambda), \tag{25c}$$

$$h = [(s - i\lambda)\exp(ib\xi)], \tag{25d}$$

$$p = q(s,\ \xi)\exp(-b\eta), \tag{25e}$$

$$D = \frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial\xi} - i\frac{\partial}{\partial\eta} - i(\lambda' - ib\lambda)e^{b\eta}\frac{\partial}{\partial\sigma}\right), \tag{25f}$$

$$K^\alpha\frac{\partial}{\partial x^\alpha} = \frac{\partial}{\partial\eta} + b\sigma\frac{\partial}{\partial\sigma} \text{ (Ref. 9).} \tag{25g}$$

Observe that Eq. (25d) is consistent with Eqs. (14), (25a), and (25b). With the commitment to Eqs. (25), the only remaining arbitrariness in our null tetrad and coordinate system is given by Eqs. (10) and (15) with

$$F = \beta + i\epsilon\pi \quad f = (\sigma_0 + i\kappa_0)\exp(-i\sqrt{2}\,b\zeta), \tag{26}$$

where β, σ_0, κ_0 are real parameters, and $\epsilon = 0$ or 1. The Killing vector undergoes the corresponding transformation $\mathbf{K} \to \mathbf{K}\exp(-\beta - i\epsilon\pi)$.

The derivation of Eqs. (25) and (26) is relegated to an Appendix to avoid any distraction from our central task of describing how the integrability conditions are applied. The only integrability conditions which we need in this section are Eqs. (19) and (20). The first step is to use Eqs. (6), (7), (9a), (25c), and (25f) to compute expressions for $\Delta\dot{DP}$ and $\Delta^2\ddot{P}$ *as linear combinations of* DP, D^*P, and P, *with coefficients which depend at most on* $h+\beta$, $(h+\beta)^*$, δ, δ', γ, and the variable

$$x := \lambda' - bs. \tag{27}$$

These expressions are then substituted for $\Delta\dot{DP}$ and $\Delta^2\ddot{P}$ in Eqs. (19) and (20), and the results are expressed as follows:

$$\nu DP - (h+\Lambda)D^*P + (1/\sqrt{2}\,)[3\delta x^{-1}(h+\Lambda) + \nu_1]P = 0, \tag{28}$$

$$-(3/2\sqrt{2}\,)[\nu_2^*DP + \nu_2 D^*P] + (|h+\Lambda|^2 + \nu_3)P = 0, \tag{29}$$

where

$$\Lambda := \beta + \tfrac{3}{2}\delta^2 x^{-2}, \tag{30a}$$

$$\nu := \tfrac{3}{2}\delta^2 x^{-2} - \tfrac{3}{2}(\delta' - 2ib\delta)x^{-1} + \tfrac{3}{4}\gamma, \tag{30b}$$

$$\nu_1 := -\tfrac{9}{2}\delta^3 x^{-3} - \tfrac{3}{2}ib\delta^2 x^{-2} + \tfrac{3}{4}\gamma\delta x^{-1} - \tfrac{1}{2}\beta', \tag{30c}$$

$$\nu_2 := \delta(\delta' + 5ib\delta)x^{-2} + \beta', \tag{30d}$$

$$\begin{aligned}\nu_3 := {}& -(\tfrac{3}{2}\delta^2 x^{-2})^2 + \tfrac{1}{4}(3\gamma + 9b^2)\delta^2 x^{-2} \\ & - \tfrac{1}{4}\Delta D(\Delta^{-1}D\beta^*) - \tfrac{1}{4}\Delta D^*(\Delta^{-1}D^*\beta) + \tfrac{3}{16}\gamma^2.\end{aligned} \tag{30e}$$

Note that all of the above Λ, ν, ν_i are polynomials in x^{-1} *with coefficients which depend only on* ξ. In general, γ is real; β is real if $b = 0$ and complex if $b \neq 0$.

Equation (28), its complex conjugate, and Eq. (29) are linear and homogeneous in DP, D^*P, P. Therefore, for a solution to exist, it is necessary that the determinant formed from their coefficients vanish,

$$\begin{aligned}& \tfrac{3}{4}\{3\delta x^{-1}(\nu_2 + \nu_2^*)|h+\Lambda|^2 + (3\delta x^{-1}\nu + \nu_1)^*\nu_2^*(h+\Lambda) \\ & + (3\delta x^{-1}\nu + \nu_1)\nu_2(h+\Lambda)^* + \nu^*\nu_1\nu_2^* + \nu\nu_1^*\nu_2\} \\ & - (|h+\Lambda|^2 + \nu_3)(|h+\Lambda|^2 - |\nu|^2) = 0.\end{aligned} \tag{31}$$

In spite of the imposing appearance of the above equa-

tion, there is a systematic method of attacking it which offers a good chance of either giving us a new NT solution or enabling us to prove that the only SNT with a Killing vector is NT_1. The reader can gain a better understanding of our method by first studying its simpler applications to the proofs of two theorems. After we have given and discussed these theorems, we shall return to Eq. (31) and discuss our current efforts to analyze the equation.

The first theorem is motivated by a striking feature of NT_1,[4] viz., for the given h of NT_1, Eqs. (6) *and* (7) *have two real linearly independent solutions* P_1 *and* P_2. In fact, the only essential parameter of NT_1 is the mixing parameter μ which occurs in the linear combination $P=(\cos\mu)P_1+(\sin\mu)P_2$. The following theorem provides us with a converse statement pertaining to the set of all SNT with Killing vectors.

Theorem II: Suppose there exists a choice of h and a corresponding SNT with a Killing vector such that Eqs. (6) and (7) have two real linearly independent solutions for P (i.e., there is exactly one essential parameter in addition to those occurring in h). Then, $b=0$, and the SNT is NT_1.

The proof starts by observing that the premise of the theorem implies a rank less than 2 for Eq. (28) and its complex conjugate *regarded as two linear homogeneous equations in* DP, D^*P, P. Therefore,

$$|h+\Lambda|^2-|\nu|^2=0, \tag{32}$$

$$\nu(3\delta x^{-1}\nu+\nu_1)^*+(3\delta x^{-1}\nu+\nu_1)(h+\Lambda)^*=0. \tag{33}$$

In the rest of the proof, we distinguish between the cases $b=0$ and $b\neq 0$.

Suppose $b=0$. Then $\sigma=s$, $\kappa=\lambda$, $\Delta=\delta$, and $x=\kappa'$. Also, Λ, ν, and ν_i are real and depend at most on ξ. From Eq. (5), $\dot{h}$ cannot be zero. Therefore, Eq. (33) implies

$$3\delta x^{-1}\nu+\nu_1=0. \tag{34}$$

We substitute from Eqs. (30b) and (30c) into the above Eq. (34) and use Eqs. (9b), (9c), and (25f). Then, upon replacing ξ by x as our independent variable, Eq. (34) becomes a linear homogeneous differential equation in Δ with a solution of the form

$$\Delta=\Delta_0+c_1x^m+c_2x^n, \tag{35}$$

where Δ_0, c_1, c_2 are constants, and m and n are simple irrational numbers whose specific values are of no importance for the understanding of the general procedure. *The proof continues by applying a key technique to Eq.* (32). Solve Eq. (32) for h^*:

$$h^*=-\Lambda+\nu^2(h+\Lambda)^{-1}. \tag{36}$$

Then, operate on the above with D and use Eq. (8) and the relation

$$\sqrt{2}\,Dh=(-2i\kappa')\dot{h}, \quad \text{if} \quad b=0, \tag{37}$$

which follows from Eqs. (25d) and (25f). There results an equation which can be algebraically solved for $\dot{h}$ to give

$$2i\dot{h}=(\Lambda'/\nu^2\kappa')(h+\Lambda)^2-(2\nu'/\nu\kappa')(h+\Lambda)+(\Lambda'/\kappa'). \tag{38}$$

Operate on the above with D^*, and use Eq. (37) and the fact that $D^*\dot{h}=\partial(D^*h)/\partial\sigma=0$ for SNT. There results

$$0=(\Lambda'/\nu^2\kappa')'(h+\Lambda)^2+2[(\Lambda'/\nu^2\kappa')\Lambda'-(\nu'/\nu\kappa')'](h+\Lambda)-2(\nu'/\nu\kappa')\Lambda'+(\Lambda'/\kappa')'. \tag{39}$$

Since $\dot{h}\neq 0$, *the coefficients in the above polynomial in* $h+\Lambda$ *must vanish identically*. The result is a set of differential equations which can be solved for Λ and ν as functions of κ, whereupon Eq. (38) can be integrated to yield $h(\sigma-i\kappa)$. However, the full solution is not needed for our present purpose. All we need is the easily proven statement that constants b_1, b_2, b_3 exist such that

$$\Lambda=b_1 \quad \text{and} \quad \nu=b_2\exp(b_3\kappa), \tag{40a}$$

or

$$(\Lambda-b_1)^2-\nu^2=b_2. \tag{40b}$$

Next, substitute (32a) and (32b) into the above Eqs. (40), then substitute in the expression (35) which was previously obtained for Δ, and recall that $\Delta\neq 0$ since the twist is not zero. One finds that Eq. (40a) leads to a contradiction, and Eq. (40b) is satisfied by Eq. (35) if and only if $c_1=c_2=0$. In the latter case, $\gamma=0$, whereupon Theorem I tells us the SNT is NT_1.

We next sketch the proof when $b\neq 0$. Equations (27), (30b), and (30c) imply

$$3\delta x^{-1}\nu+\nu_1\neq 0$$

since $b\delta\neq 0$. Solve Eq. (33) for h^* and operate on the resulting expression with D. Then multiply through by $(3\delta x^{-1}\nu+\nu_1)^2$ and collect all terms on one side to obtain an equation of the form $X=0$. Take note of Eqs. (30) and the relation

$$\sqrt{2}\,D(x^{-1})=-2\delta x^{-2}+ibx^{-1},$$

which derives from Eqs. (25) and (27). It then becomes clear that X is a polynomial in x^{-1} of degree 7, and the coefficient of each power of x^{-1} must vanish since $\dot{x}\neq 0$. It is sufficient to inspect the x^{-7} term; its vanishing implies $b=0$, which contradicts our original hypothesis. Hence, *when* $b\neq 0$, *there is no* SNT *which satisfies the premise of the theorem*. Q. E. D.

Theorems I and II are based on two properties of NT_1. Another property of NT_1 is that its h satisfies a quadratic equation of the form $|h|^2+\Lambda_1(h+h^*)+\Lambda_2=0$ where Λ_1 and Λ_2 depend at most on ξ. The following theorem is the converse statement for the SNT with a $b=0$ Killing vector set.

Theorem III: Consider any given SNT with a Killing vector such that $b=0$. Suppose there exist scalar fields Λ_1 and Λ_2 which depend at most on ξ such that

$$|h|^2+\Lambda_1(h+h^*)+\Lambda_2=0. \tag{41}$$

Then the SNT is NT_1.

A detailed study of Eq. (31) when $b=0$ and when Eq. (41) holds reveals that there are only three distinct possibilities for which the premise of the above theorem is true:

(1) Equation (32) is true, whereupon we have the $b=0$ part of Theorem II.

(2) Equation (32) is not true, but $\nu_2=0$, whereupon Eq. (31) implies $|h+\Lambda|^2+\nu_3=0$.

(3) Equation (32) is not true, and $\nu_2\neq 0$, but $3\delta x^{-1}\nu + \nu_1=0$. Then Eq. (31) implies $|h+\Lambda|^2+\nu_3 - (9/2)\delta x^{-1}\nu_2=0$.

In any case, the proof of Theorem III proceeds along the same lines as the proof of Theorem II. The steps differ only in details from those of Eqs. (34)—(40). (For example, the solution of $\nu_2=0$ has the same form as the expression for Δ in Eq. (35). The irrational numbers m and n are different, but that does not alter the course of the proof.)

When the premise of Theorem III does not hold, we are faced with the analysis of Eq. (31). This is a much more difficult task than the analysis of Eq. (41), but we have made some progress by using similar techniques. For example, suppose $b=0$.[10] We first solve Eq. (31) for

$$h^*=G(h,\ \Lambda_1, \Lambda_2, \Lambda_3, \Lambda_4),$$

where Λ_1, Λ_2, Λ_3, Λ_4 are simple polynomial expressions in the functions Λ, ν, ν_i of ξ. Then we operate on the above with D and obtain [see Eq. (37)]

$$0=-2i\kappa' \dot{h}\frac{\partial G}{\partial h}+\Lambda_i'\frac{\partial G}{\partial \Lambda_i}.$$

We solve the above for $\dot{h}$ and apply the equality $D^*\dot{h}=0$. The result which we obtain is expressible in the form

$$P_9=P_7(P_3)^{1/2},$$

where P_3, P_7, P_9 are polynomials in $h+\Lambda$ of respective degrees $=3$, ≤ 7, ≤ 9, with coefficients which are polynomial expressions in the Λ_i and their first and second derivatives. Therefore,

$$P_7=P_9=0.$$

Upon equating the coefficients in the above polynomials to zero, we obtain a host of ordinary differential equations in the dependent variables Λ_i. Our immediate aim is to get relations between the functions Λ, ν, ν_i and to check these relations against Eqs (30). This is not a task which can be completed overnight since the differential equations are severely nonlinear. We have integrated two of them so far.

In the next section we do not assume that a Killing vector exists.

5. THE SUBCASE $D^*\beta = 0$ AND $\gamma \neq 0$

If a given STN is NT_1, then Theorem I and Eq. (18) implies $D^*\beta=0$. The converse is not necessarily true, and we are thus led to consider the possibility that there exists an SNT for which $D^*\beta=0$ and $\gamma\neq 0$. In the remainder of this section, we assume $D^*\beta=0$ and $\gamma\neq 0$ unless we explicitly state the contrary.

Since β is an analytic function of ζ when $D^*\beta=0$, the transformation (11f) can be used to make $\beta\to 0$ by an appropriate choice of $F(\zeta)$. However, we have found that it is wiser to hold this transformation in reserve, since we may want to choose $F(\zeta)$ to achieve simplification of another field such as $h+\beta$.

Equations (9c) and (18) imply the existence of a real constant ϵ such that

$$\gamma=-\epsilon\Delta \quad \text{if} \quad D^*\beta=0. \tag{42}$$

$\epsilon\neq 0$ if $\gamma\neq 0$. Therefore, we can always select the scaling parameter α in Eqs. (11) so that $\epsilon\to\pm 1$. However, we prefer not to do this now since another numerical value may suggest itself later.

Equations (9c), (13), and (42) imply the existence of $g(\zeta)$ such that

$$\ln\Delta=-\epsilon\kappa+g(\zeta)+g(\zeta)^*.$$

Therefore, for any given α and $F(\zeta)$, we can select $f(\zeta)$ in the transformations (11c) and (16) so that $g\to 0$. Then,

$$\Delta=\exp(-\epsilon\kappa). \tag{43}$$

As regards the dependence of Δ on ζ and ζ^*, let $u_1(\zeta)$ and $u_2(\zeta)$ be any given solutions of the equations

$$u_i''=\beta u_i, \quad u_1u_2'-u_2u_1'=1. \tag{44}$$

Then, the general solution of Eq. (9b) is

$$\Delta^{-1/2}=c_{ij}u_i^*u_j, \tag{45}$$

where $\{c_{ij}\}$ is constant Hermitian matrix such that

$$\det\{c_{ij}\}=\tfrac{1}{2}\epsilon.$$

It is apparent that the integrability conditions (21) and (23) greatly simplify when $D^*\beta=0$. That makes it easy to derive a fifth integrability condition from which P is completely eliminated. In the derivation, it is convenient to let

$$H:=\Delta^2(h+\beta)^{-1}. \tag{46}$$

Divide Eq. (23) through by Δ^2, operate with D^*, and use Eqs. (6) and (7). Then, with the help of Eq. (21), we can eliminate $\dot{P}$, DP, and D^*P and obtain the new integrability condition

$$|DH|^2-4\gamma|H|^2=4i\Delta(H\dot{H}^*-H^*\dot{H}). \tag{47}$$

From Eqs. (9c), (17), (46), (47), and the equality $D^*(h+\beta)=0$, one can prove that $DH\neq 0$, which is a fact of some relevance in the sequel.

Another useful relation is derived by noting that Eqs. (21), (23), and their complex conjugates are linear and homogeneous in DP, D^*P, $\Delta\dot{P}$, and P. The condition that the four by four determinant of their coefficients vanish yields after a suitable grouping of terms and factors:

$$|DH/H|^2|W-1|^2=-\tfrac{4}{3}\gamma(1-|W|^2), \tag{48}$$

where W is an invariant under the group (10) and is defined by

$$W:=(3i/4\Delta)(D^*H^*/DH)\dot{H}. \tag{49}$$

Equations (46) to (49) are also applicable when $\gamma=0$. From Eq. (48), $W=1$ if the SNT is NT_1, and it can be proven that (conversely) $W=1$ and $D^*\beta=0$ imply NT_1.

The conditions (47) and (48) still have to be fully exploited, and it remains an open question as to whether there exists any SNT such that $D^*\beta=0$ and $\gamma\neq 0$. One

significant point is made by the following theorem which is the $D^*\beta=0$ analog of Theorem II.

Theorem IV: For any given κ and h such that an SNT exists, $D^*\beta=0$, and $\gamma\neq 0$, it follows that Eqs. (6) and (7) have exactly one linearly independent real solution P.

We give only a rough sketch of the proof. Suppose Eqs. (6) and (7) have two linearly independent solutions. When $D^*\beta=0$, Eq. (23) and its complex conjugate are linear and homogeneous in DP, D^*P, P. The condition that the rank of these equations be less than 2 yields two relations involving H, DH, and $\dot{H}$. From these relations, the equation $D^*(h+\beta)=0$, and Eqs. (9c) and (17), we can establish a contradiction with Eq. (47) (by the method used to prove $DH\neq 0$). The details are left for the reader.

6. SUMMARY

The working equations to be used in the analysis of the general SNT problem are Eqs. (4c), (6)—(8), (13), (14), (17), and the equations in Sec. 3. Also, one must keep in mind the possiblility of using the transformations given by Eqs. (11) to obtain simplifications at appropriate points of the analysis.

There are some worthwhile subjects for investigation other than the obvious one of searching for a new NT. A key question concerns the degree of arbitrariness in h corresponding to a given Ω. For NT_1, h is uniquely determined[2] by the choice of Ω. How about other choices of Ω? This question is applicable to any NT, whether or not it is SNT.

Then there is the problem of generalizing Theorems II and IV. Can these theorems be extended to the set of all SNT (without the constraints that a Killing vector exist or that $D^*\beta=0$) or, even better, to the set of all NT?

As regards the set of SNT with Killing vectors, the key problem as we have formulated it is to crack Eq. (31) subject to Eqs. (25) and (30). Of course, there may not exist any *new* SNT with a Killing vector, but that remains to be proven and is a problem in itself.

Then there is the especially interesting subcase of the set of SNT such that $D^*\beta=0$, where the key problem as we have formulated it is to crack Eqs. (47) and (48). Our own analysis of these equations has already been initiated, and we expect that other investigators will want to try their hand. An open question is that of a suitable relaxation of the ansatz $D^*\beta=0$; Eqs. (18) and (21) can be helpful in that venture.

In closing, we stress that there are probably more elegant and more powerful techniques than any which we have devised, and it is our sincere hope that the problem will attract experts on such techniques, including some who are not currently working on the problem.

APPENDIX

For any SNT with a Killing vector **K**, we want to show that we can use the transformations given by Eqs. (10) and (11) to select our null tetrad and coordinates so that Eqs. (25) hold. The relevant null tetrad component forms of the Killing vector structural equations are given, e.g., by Collinson.[8] By using these equations, we can establish that the function $F(\zeta)$ in the transformation (10) can be chosen, with $\alpha=0$ and $g=0$, to yield a null tetrad such that

$$\pounds\zeta=i/\sqrt{2}\,. \tag{A1}$$

$$\pounds\mathbf{k}=\pounds\mathbf{m}=\pounds\mathbf{t}=0, \tag{A2}$$

where $\pounds$ *denotes the Lie derivation with respect to* **K**. We grant Eqs. (A1) and (A2), since the prooof does not differ significantly from similar proofs given by various people, e.g., Collinson. Then, from Eqs. (1), (4), and (A1),

$$\pounds\rho=\pounds\tau=\pounds A=\pounds h=\pounds d\zeta=0, \tag{A3}$$

From the relation,[2]

$$d\wedge k=k\wedge(Ad\zeta+A^*d\zeta^*)-2i\tau d\zeta\wedge d\zeta^*, \tag{A4}$$

and from Eqs. (2), (A2), (A3), and the vanishing of $\pounds d\wedge k$, we obtain

$$k\wedge d(\pounds\chi)=0$$

However, we are assuming nonzero twist. Therefore, the above equation implies $d(\pounds\chi)=0$. Therefore, there exists a real constant b such that

$$\pounds\chi=b. \tag{A5}$$

This is the constant b which occurs in Eqs. (25).

From Eqs. (4) and (A3),

$$\pounds\tau=(\pounds\Delta-b\Delta)\exp(-\chi)=0.$$

However, $\Delta=\Delta(\zeta,\zeta^*)$ for SNT. Therefore, from Eqs. (A1), (24), and the above equation, there exists $\delta=\delta(\xi)$ such that

$$\Delta=\delta(\xi)\exp(b\eta). \tag{A6}$$

From Eqs. (13), we then obtain the existence of $\mu=\mu(\zeta)$ and $\lambda=\lambda(\xi)$ such that

$$\kappa=\lambda(\xi)\exp(b\eta)+\mu(\zeta)+\mu(\zeta)^*,$$

$$\delta=\tfrac{1}{2}(\lambda''+b^2\lambda). \tag{A7}$$

We use the transformation (16) to make $\mu=0$, whereupon

$$\kappa=\lambda(\xi)\exp(b\eta). \tag{A8}$$

A careful analysis now reveals that the only residual freedom in our choice of a null tetrad and coordinates via Eqs. (10) and (11) is, in view of our choices (A1) and $\mu=0$, given by

$$f(\zeta)=\sqrt{2}c\zeta+\kappa_0+\sigma_0 \quad \text{if} \quad b=0,$$

$$f(\zeta)=(\kappa_0+ib^{-1}c)[\exp(-i\sqrt{2}b\zeta)-1]$$
$$+\kappa_0+\sigma_0 \quad \text{if } b\neq 0, \tag{A9}$$

$$F=\beta+i\epsilon\pi, \quad \epsilon:=0,1,$$

where c, κ_0, σ_0, β are real parameters, and $\mathbf{K}\to\mathbf{K}\exp(-\beta-i\epsilon\pi)$ under the residual group.

From the Eq. (3) for k and the relation $\pounds k=0$, we next prove that

$$d(\pounds\sigma-b\sigma)=0.$$

Therefore,

$$\mathcal{L}\sigma = b\sigma + a, \quad a = \text{const.}$$

We can show that, under the residual group (for $\alpha = \beta = \epsilon = 0$) $a \to a + c$. Therefore, by letting $c = -a$ we arrive at

$$\mathcal{L}\sigma = b\sigma. \tag{A10}$$

From Eqs. (A1), (A3), and (A10), we then obtain (25g). Also, we already have Eqs. (25b) and (25c) via (A6) to (A8).

The rest of the derivation of Eqs. (25) and (26) is fairly straightforward and is left for the reader.

[1]For a review and bibliography on algebraically special gravitational fields with twisting rays, in the context of a review of exact solutions, see W. Kinnersley, in *General Relativity and Gravitation*, edited by G. Shaviv and J. Rosen (Wiley, New York, 1975), pp. 109—35.

[2]I. Hauser, J. Math. Phys. **19**, 661 (1978). In this reference, TNT designated any type *N* twisting gravitational field. We are dropping the first *T*. Our signature is +2, and $k \cdot m = t \cdot t^* = 1$.

[3]This condition was discussed in Sec. 4C of Ref. 2. A special case of the condition was given in Ref. 4.

[4]I. Hauser, Phys. Rev. Lett. **33**, 1112 (1974). Equation (15) in this article has a typographical error: $1+y^2$ should be in the *denominator*. Also, omit the factor 4 on the right side of Eq. (16). The coordinate notations (u, ξ, ρ, σ) in this reference were replaced in Ref. 2 by $(\rho, \sigma, \xi, \eta)$, and A was interchanged with A^*. We follow Ref. 2.

[5]Our coordinate notations are consistent with those of I. Robinson and A. Trautman, Phys. Rev. Lett. **4**, 431 (1960); Proc. Roy. Soc. Ser. A **265**, 463 (1962).

[6]This is a special case of a theorem which holds for any NT and which is proven in Ref. 2.

[7]NT_1 as given in Ref. 4 was originally derived by computing the successive integrability conditions of the NT field equations for the special case $\Omega = i(\zeta + \zeta^*)$. The details of the original derivation are given in Ref. 2.

[8]No NT can have more than one Killing vector, as was shown by C.D. Collinson, J. Phys. A **2**, 621 (1969).

[9]If $b \neq 0$, we can (of course) select a null tetrad k', m', t', t'^* and a coordinate system ρ', σ', ζ', ζ'^* such that $K^\alpha(\partial/\partial x^\alpha) = \partial/\partial\eta'$ where $\eta' = \mathrm{Im}(\sqrt{2}\,\zeta')$. However, that would involve a transformation outside the group (10), and the resulting Ω' would be dependent on σ' (i.e., D' would not commute with $\partial/\partial\sigma'$). There appears to be no advantage in that choice here.

[10]For $b \neq 0$ we have only completed some of the details of the analysis of Eq. (31).

On the application of the generalized quantal Bohr–Sommerfeld quantization condition to single-well potentials with very steep walls

Nanny Fröman and Per Olof Fröman

Institute of Theoretical Physics, University of Uppsala, Uppsala, Sweden
(Received 7 December 1977)

This investigation concerns some model potentials with very steep walls, for which the quantal Bohr–Sommerfeld *half-integer* quantization condition, if necessary generalized to correspond to modified phase-integral approximations of arbitrary order, and used without or with higher-order corrections included, can be used for obtaining, very accurately, the energy eigenvalues of the bound states, apart possibly from the lowest ones. One of the cases treated is the potential proportional to $\cot^2 x$, for which a modified Bohr–Sommerfeld half-integer quantization condition yields the energy eigenvalues exactly. That the Bohr–Sommerfeld *half-integer* quantization condition is applicable to potentials with very steep walls may at first sight seem surprising in view of the well-known fact that the energy eigenvalues of a square-well potential with infinitely high walls are obtained exactly from the Bohr–Sommerfeld *integer* quantization condition, i.e., the quantization condition obtained by replacing $(s+1/2)\pi$ by $(s+1)\pi$ in the right-hand member of the Bohr–Sommerfeld half-integer quantization condition. From the study of a potential with horizontal bottom and linearly rising walls, which goes over into a square well when the inclination of the walls tends to zero, it can easily be understood, why the *half-integer* quantization condition is appropriate when the steep walls have a finite slope.

1. INTRODUCTION

The time-independent Schrödinger equation for the motion of a quantal particle with the energy E in a one-dimensional potential $V(z)$ is

$$\frac{d^2\psi}{dz^2} + Q^2(z)\,\psi = 0, \tag{1}$$

where, with obvious notations,

$$Q^2(z) = \frac{2m}{\hbar^2}[E - V(z)]. \tag{2}$$

The eigenvalue problem of finding the bound states of this Schrödinger equation can be treated by means of a method for solving connection problems which has been developed by the present authors[1] (cf. also a further extension by Fröman[2]). Using this method, we derived in Ref. 1 an exact quantization condition for the case of a single-well potential [cf. Eq. (10.20) in Ref. 1]. By deleting in this exact quantization condition a small term for which an upper bound was given, one obtains the Bohr—Sommerfeld (JWKB) *half-integer* quantization condition [cf. Eq. (10.22) in Ref. 1]. In Ref. 3 the extension of the treatment given in Ref. 1 to a new kind of higher-order phase-integral approximation (cf. also Ref. 3]. One arrives at a further generalization by tion was generalized correspondingly [cf. Eq. (13) in Ref. 3)]. One arrives at a further generalization by modifying the arbitrary-order phase-integral approximations as described in Ref. 5 and on pp. 126—31 in Ref. 6. In this way one obtains the exact quantization condition (7′) in Ref. 7, i.e.,

$$\frac{1}{2}\int_\Lambda q(z)\,dz = (s+\tfrac{1}{2})\,\pi + \arg\frac{F_{11}(-\infty,z)}{F_{11}(+\infty,z)}, \quad s=0,1,2,\cdots, \tag{3}$$

where $q(z)$ is given by Eqs. (12) and (13) in Ref. 5, and Λ is the contour shown in Fig. 1(c) of Ref. 7. When E is an eigenvalue, the last term in (3) is, according to Ref. 7, independent of z, which may thus be chosen as an arbitrary point in the complex z plane. By neglecting this term in (3) one obtains the Bohr—Sommerfeld *half-integer* quantization condition, generalized to modified phase-integral approximations of arbitrary order, [cf. Eq. (24) in Ref. 7],

$$\tfrac{1}{2}\int_\Lambda q(z)\,dz = (s+\tfrac{1}{2})\pi, \quad s=0,1,2,\cdots. \tag{4}$$

The question of the applicability of this approximate quantization condition can thus be settled by the evaluation of an upper bound or an approximate expression for the last term, i.e., the correction term $\arg[F_{11}(-\infty,z)/F_{11}(+\infty,z)]$, in the exact quantization condition (3). When the generalized classical turning points x' and x'', i.e., the two relevant real zeros of $Q^2_{\mathrm{mod}}(z)$, are well separated, one can obtain an upper bound which is given by Eq. (15) in Ref. 7. Another upper bound, which is useful not only when x' and x'' are well-separated but also when these points lie close together (whether the bottom of the potential well has approximately parabolic shape or not), is given by Eq. (18) in Ref. 7. In addition to those upper bounds for the correction term in the exact quantization condition (3) one can derive the approximate formula (23) in Ref. 7 for this term. In this way one obtains the approximate quantization condition (25) in Ref. 7, which may at first sight seem to be an essential improvement of the generalized quantal Bohr—Sommerfeld quantization condition. However, as remarked in Ref. 7, it seems to be preferable to use a higher order of the last mentioned quantization condition, i.e., (4) in the present paper, instead of using (25) in Ref. 7. In this connection we remark that, as already mentioned in Ref. 7, condition (25) in Ref. 7, used in the first-order approximation, is the same as (4) in the present paper used in the third-order approximation.

The above-mentioned upper bounds for the errors in

 0022-2488/78/1909-1823$1.00

the quantization conditions in Ref. 7, which are valid for unmodified as well as modified phase-integral approximations of arbitrary order, are expressed in terms of the so called μ integral, which is to be performed along paths with certain properties of monotonicity. When the μ integral pertaining to the first-order phase-integral approximation (which is the same as the first-order JWKB approximation) is small compared to unity, we can in general expect that the μ integral will decrease rapidly with increasing order of the phase-integral approximations until, at a certain optimal order, the μ integral starts increasing. Thus, if the first-order approximation works satisfactorily, the first few of the higher-order approximations should in general yield improved results. In many cases it is therefore sufficient to restrict the estimation of the appropriate μ integral to that pertaining to the first-order approximation.

For a sufficiently smooth single-well potential one has, as a particular case of (4), when the unmodified first-order JWKB approximation is used, the approximate quantal Bohr—Sommerfeld *half-integer* quantization condition

$$\int_{x'}^{x''} Q(z)\,dz = (s+\tfrac{1}{2})\pi, \quad s = 0, 1, 2, \cdots, \tag{5}$$

where x' and x'' are the classical turning points. For a square-well potential with infinitely high walls, one finds instead that the energy eigenvalues are obtained exactly from the Bohr—Sommerfeld *integer* quantization condition

$$\int_{x'}^{x''} Q(z)\,dz = (s+1)\pi, \quad s = 0, 1, 2, \cdots. \tag{6}$$

As was pointed out in a letter from Prof. G.H. Wannier to the present authors it is therefore interesting and important to investigate which quantization condition one shall use when a continuous potential with very (but not infinitely) steep walls looks similar to a square-well potential. For such a potential it seems plausible that the Bohr—Sommerfeld *integer* quantization condition (6) could be used for obtaining at least very rough approximations of the lowest energy eigenvalues, and from the results in Refs. 1 and 7 one may expect that the Bohr—Sommerfeld *half-integer* quantization condition (5), possibly modified, or its generalization (4), could be used for determining approximately the eigenvalues of the more highly excited bound states.

In the following sections we shall consider some single-well model potentials with very steep walls. We shall restrict ourselves to considering the eigenvalues, but it would also be of interest to calculate quantal expectation values and matrix elements without the use of wavefunctions (cf. Refs. 8 and 9) and to investigate the accuracy of the normalized (cf. Ref. 10) phase-integral eigenfunctions, especially inside the classically allowed region.

2. POTENTIAL PROPORTIONAL TO $\cot^2 z$

In this section we shall consider the eigenvalue problem for the one-dimensional Schrödinger equation when the potential is

$$V(z) = \frac{\hbar^2}{2m} B \cot^2 z, \tag{7}$$

where B is a positive constant, and the physical range of the variable z is $0 < z < \pi$. Introducing instead of the energy E the quantity A by putting

$$E = \frac{\hbar^2}{2m} A, \tag{8}$$

we can write (2) as follows,

$$Q^2(z) = A - B\cot^2 z. \tag{9}$$

To every given positive value of B there corresponds a discrete sequence of positive eigenvalues A for which the wavefunction tends to zero as z tends to $+0$ and to $\pi - 0$.

The exact solution of the Schrödinger equation in the case of the $\cot^2$ potential is considered in problem 12 on pages 2 and 72—4 in Ref. 11. From the treatment there one realizes that the condition that our differential equation (1), with $Q^2(z)$ given by (9), has a solution ψ which is equal to zero when $z = 0$ and $z = \pi$ is

$$\begin{aligned}
&\tfrac{1}{2}(A+B)^{1/2} - \tfrac{1}{2}[(B+\tfrac{1}{4})^{1/2} - \tfrac{1}{2}] = \tfrac{1}{2}(s+1), \quad s = 0, 1, 2, \cdots, \\
&\text{i.e.,} \\
&A = [s + \tfrac{1}{2} + (B+\tfrac{1}{4})^{1/2}]^2 - B \\
&\quad = \tfrac{1}{4} - B + [s + (B+\tfrac{1}{4})^{1/2}][s + (B+\tfrac{1}{4})^{1/2} + 1], \\
&s = 0, 1, 2, \cdots.
\end{aligned} \tag{10}$$

Evaluating the integral occurring in the left-hand member of (5) and (6), with $Q^2(z)$ given by (9), and using the exact expression (10) for the eigenvalues of A, we obtain

$$\begin{aligned}
&\frac{1}{\pi}\int_{x'}^{x''} Q(z)\,dz \\
&\quad = (A+B)^{1/2} - B^{1/2} = s + \tfrac{1}{2} + (B+\tfrac{1}{4})^{1/2} - B^{1/2} \\
&\quad = \begin{cases} s + \Delta, & \text{for all values of } B, \\ s + 1 + O(\sqrt{B}), & \text{when } B \ll \tfrac{1}{4}, \\ s + \tfrac{1}{2} + O(1/\sqrt{B}), & \text{when } B \gg \tfrac{1}{4}, \end{cases}
\end{aligned} \tag{11}$$

where $\Delta = \tfrac{1}{2} + (B+\tfrac{1}{4})^{1/2} - B^{1/2}$.

Thus, with $s + \Delta$, where $\Delta = \tfrac{1}{2} + (B+\tfrac{1}{4})^{1/2} - B^{1/2}$, in the right-hand member, the first-order unmodified JWKB-quantization condition would yield the eigenvalues exactly. We note, however, that the quantity Δ, here obtained from known exact results, is particular for the potential under consideration, and no definite conclusion can be drawn as to the existence of such a quantity (independent of s) and its value for a general steep-wall potential. See also the discussion in Ref. 7. We further realize from (11) that for the *first-order unmodified* JWKB approximation the *integer* Bohr—Sommerfeld quantization condition (6) can be used for all possible values of the quantum number s when $\sqrt{B} \ll 1$, whereas the *half-integer* Bohr—Sommerfeld quantization condition (5) (which is used in problem 18b on pages 4 and 78 in Ref. 11) can be used for all possible values of s when $\sqrt{B} \gg 1$. When a convenient modification is used, one can, however, as we shall presently show, obtain

the eigenvalues more accurately, and even exactly, by means of the *half-integer* Bohr-Sommerfeld quantization condition.

Let us now treat the eigenvalue problem for the potential well (7) by using the previously mentioned *modified* phase-integral approximations of arbitrary order. According to (9) the function $Q^2(z)$ is symmetric with respect to the point $z=\pi/2$, and for $Q^2_{\rm mod}(z)$ we shall choose a convenient function with the same symmetry, namely

$$Q^2_{\rm mod}(z)=a-b\cot^2 z, \tag{12}$$

where

$$b=B+\tfrac{1}{4}\quad(>\tfrac{1}{4}), \tag{13}$$

and a is a so far arbitrary constant, which is assumed to differ not too much from A. Because of (13) the phase-integral approximations are valid also at the points where $\cot^2 z=\infty$ and thus, in particular, at the end points $z=0$ and $z=\pi$ of the range of the physical variable z. The phase of $Q_{\rm mod}(z)$ is chosen to be positive on the upper lip of a cut along the real axis between the generalized classical turning points (cf. Fig. 1(b) in Ref. 7).

When phase-integral approximations of the order $2N+1$ are used, and when we integrate along the closed contour Λ encircling the generalized classical turning points x' and x'' (cf. Fig. 1(c) in Ref. 7), we get (cf. Eqs. (8) and (7) in Ref. 10)

$$\tfrac{1}{2}\int_\Lambda q(z)\,dz=\sum_{n=0}^{N} L^{(2n+1)} \tag{14}$$

with

$$L^{(2n+1)}=\tfrac{1}{2}\int_\Lambda Y_{2n}Q_{\rm mod}\,dz=\tfrac{1}{2}\int_\Lambda Z_{2n}Q_{\rm mod}\,dz, \tag{15}$$

where [see Eqs. (9a, b, c) and (6) in Ref. 10]

$$Z_0=1, \tag{16a}$$

$$Z_2=\tfrac{1}{2}\epsilon_0, \tag{16b}$$

$$Z_4=-\tfrac{1}{8}\epsilon_0^2, \tag{16c}$$

with

$$\epsilon_0=\frac{Q^2(z)-Q^2_{\rm mod}(z)}{Q^2_{\rm mod}(z)}+Q^{-3/2}_{\rm mod}(z)\,\frac{d^2}{dz^2}\,Q^{-1/2}_{\rm mod}(z). \tag{16d}$$

Putting $t=-\cot z$, and denoting by Λ_t the contour in the t plane which corresponds to the contour Λ in the z plane, we can write (15) as follows (note the direction of integration indicated in Fig. 1(c) of Ref. 7)

$$\begin{aligned}L^{(2n+1)}&=\frac{1}{2}\int_{\Lambda_t} Z_{2n}Q_{\rm mod}\frac{dt}{1+t^2}\\&=\pi i\left(\operatorname*{Res}_{t=+i}\frac{Z_{2n}Q_{\rm mod}}{1+t^2}+\operatorname*{Res}_{t=-i}\frac{Z_{2n}Q_{\rm mod}}{1+t^2}\right)\\&\quad+\tfrac{1}{2}\int_{\Lambda_t'} Z_{2n}Q_{\rm mod}\frac{dt}{1+t^2},\end{aligned} \tag{17}$$

where Λ_t' is a contour encircling in the negative sense the part of the real t axis corresponding to the generalized classically allowed region, as well as the two points $t=\pm i$, which lie outside of the contour Λ_t. Using (17), and noting that $Q_{\rm mod}$ is approximately equal to $-it\sqrt{b}$ for large values of $|t|$ (cf. Fig. 1(b) in Ref. 7) and that Z_{2n} tends to the Kronecker symbol $\delta_{n,0}$ as $|t|\to\infty$, we easily find that (17) can be written as follows [cf. (12)],

$$\begin{aligned}L^{(2n+1)}&=\tfrac{1}{2}\pi(a+b)^{1/2}(Z_{2n})_{t=+i}+\tfrac{1}{2}\pi(a+b)^{1/2}(Z_{2n})_{t=-i}\\&\quad-\pi\sqrt{b}\,\delta_{n,0},\end{aligned} \tag{18}$$

where $(a+b)^{1/2}$ denotes a positive quantity.

To calculate expressions for the quantities $(Z_{2n})_{t=\pm i}$ appearing in (18) we need formulas for $(\epsilon_0)_{t=\pm i}$ [cf. (16a, b, c)]. Since

$$d(\cot z)/dz=-(1+\cot^2 z)=-(1+t^2)=0\quad\text{when } t=\pm i,$$

it follows from (12) that the derivative $d^2Q^{-1/2}_{\rm mod}/dz^2$ is equal to zero when $t=\pm i$. From the definition (16d) of ϵ_0 and the formulas (9), (12), and (13) we therefore get

$$(\epsilon_0)_{t=\pm i}=\left(\frac{Q^2-Q^2_{\rm mod}}{Q^2_{\rm mod}}\right)_{t=\pm i}=\frac{c}{a+b}, \tag{19}$$

where

$$c=A-a-\tfrac{1}{4}. \tag{20}$$

With the aid of (16a, b, c) and (19) it follows from (18) that

$$L^{(1)}=\pi[(a+b)^{1/2}-\sqrt{b}\,], \tag{21a}$$

$$L^{(3)}=\tfrac{1}{2}\pi c(a+b)^{-1/2}, \tag{21b}$$

$$L^{(5)}=-\tfrac{1}{8}\pi c^2(a+b)^{-3/2}. \tag{21c}$$

For the *first-order* JWKB approximation, modified according to (12) with (13), we find from (14), (21a), (13), and (20), when use is also made of the exact expression (10) for the eigenvalues of A, that

$$\begin{aligned}\frac{1}{\pi}\int_{x'}^{x''} Q_{\rm mod}(z)\,dz&=(a+b)^{1/2}-\sqrt{b}\\&=(A+B-c)^{1/2}-(B+\tfrac{1}{4})^{1/2}\\&=\{[s+\tfrac{1}{2}+(B+\tfrac{1}{4})^{1/2}]^2-c\}^{1/2}-(B+\tfrac{1}{4})^{1/2}\\&=s+\tfrac{1}{2}+O\left(\frac{c}{s+\frac{1}{2}+(B+\frac{1}{4})^{1/2}}\right)\end{aligned} \tag{22}$$

if $|c|\ll s+\frac{1}{2}+(B+\frac{1}{4})^{1/2}$. When the first-order JWKB approximation is modified according to (12) with (13), the *half-integer* Bohr-Sommerfeld quantization condition is thus approximately valid for all possible values of s, when B is sufficiently large. Furthermore, by comparing (11) and (22), we realize that, for any (even very small) positive value of B, this quantization condition is more accurate than the unmodified (half-integer or integer) Bohr–Sommerfeld quantization conditions (5) or (6) for sufficiently highly excited states, i.e., for sufficiently large values of the quantum number s.

Using the modification (12) with (13), we shall now evaluate the quantization condition (4) in the first-, third-, and fifth-order approximations. In the first-order approximation the quantization condition (4) becomes [cf. (21a)]

$$\pi[(a+b)^{1/2}-\sqrt{b}\,]=(s+\tfrac{1}{2})\pi, \tag{23a}$$

which, with the use of (13) and (20), yields

$$A=[s+\tfrac{1}{2}+(B+\tfrac{1}{4})^{1/2}]^2\left(1+\frac{c}{[s+\frac{1}{2}+(B+\frac{1}{4})^{1/2}]^2}\right)-B. \tag{24a}$$

In the third-order approximation the quantization condition (4) becomes [cf. (21a, b)]

$$\pi[(a+b)^{1/2}-\sqrt{b}\,]+\tfrac{1}{2}\pi c(a+b)^{-1/2}=(s+\tfrac{1}{2})\pi, \tag{23b}$$

from which, using (13) and (20), we obtain

$$\begin{aligned}A&=\tfrac{1}{2}[s+\tfrac{1}{2}+(B+\tfrac{1}{4})^{1/2}]^2\\&\quad\underset{(-)}{+}\tfrac{1}{2}[s+\tfrac{1}{2}+(B+\tfrac{1}{4})^{1/2}]\{[s+\tfrac{1}{2}+(B+\tfrac{1}{4})^{1/2}]^2-2c\}^{1/2}\\&\quad-B+\tfrac{1}{2}c\\&=[s+\tfrac{1}{2}+(B+\tfrac{1}{4})^{1/2}]^2\left[1+O\left(\frac{c^2}{[s+\frac{1}{2}+(B+\frac{1}{4})^{1/2}]^4}\right)\right]-B.\end{aligned} \tag{24b}$$

In the fifth-order approximation the quantization condition (4) becomes [cf. (21a, b, c)]

$$\pi[(a+b)^{1/2}-\sqrt{b}\,]+\tfrac{1}{2}\pi c(a+b)^{-1/2}-\tfrac{1}{8}\pi c^2(a+b)^{-3/2}=(s+\tfrac{1}{2})\pi. \tag{23c}$$

When $|c|/[s+\frac{1}{2}+\sqrt{b}\,]^2\ll 1$ one can solve (23c) by expressing $(a+b)^{1/2}/[s+\frac{1}{2}+\sqrt{b}\,]$ as a power series in $c/[s+\frac{1}{2}+\sqrt{b}\,]^2$. In this way one obtains

$$(a+b)^{1/2}/[s+\tfrac{1}{2}+\sqrt{b}\,]=1-\frac{\frac{1}{2}c}{[s+\frac{1}{2}+\sqrt{b}\,]^2}-\frac{\frac{1}{8}c^2}{[s+\frac{1}{2}+\sqrt{b}\,]^4}+O\left(\frac{c^3}{[s+\frac{1}{2}+\sqrt{b}\,]^6}\right)$$

and hence, with the use of (13) and (20),

$$A=[s+\tfrac{1}{2}+(B+\tfrac{1}{4})^{1/2}]^2\left[1+O\left(\frac{c^3}{[s+\frac{1}{2}+(B+\frac{1}{4})^{1/2}]^6}\right)\right]-B. \tag{24c}$$

Comparing (24a, b, c) and (10), we see that with increasing order of the modified phase-integral approximations used [cf. (12) and (13)], the energy eigenvalues obtained from the quantization condition (4) rapidly approach the exact eigenvalues, provided that $|c|\ll[s+\frac{1}{2}+(B+\frac{1}{4})^{1/2}]^2$. If, in particular, we choose $c=0$, i.e., $a=A-\frac{1}{4}$ [cf. (20)], and hence

$$Q^2_{\rm mod}=A-\tfrac{1}{4}-(B+\tfrac{1}{4})\cot^2 z, \tag{25}$$

we realize that the quantization condition (4) yields the energy eigenvalues exactly in the first-, third-, and fifth-order approximations. This result is of particular interest, since the potential in question appears in connection with the free rotation of diatomic molecules and the theory of Legendre functions.

3. POTENTIAL PROPORTIONAL TO cosh (z/C)

In this section we shall consider the potential

$$V(z)=\frac{\hbar^2}{2m}B\cosh\frac{z}{C}, \tag{26}$$

where B and C are positive constants. If one chooses these constants conveniently, a graphical representation of the potential $V(z)$ for real values of z looks fairly like a square-well potential with infinitely high walls (cf. Fig. 1). Writing

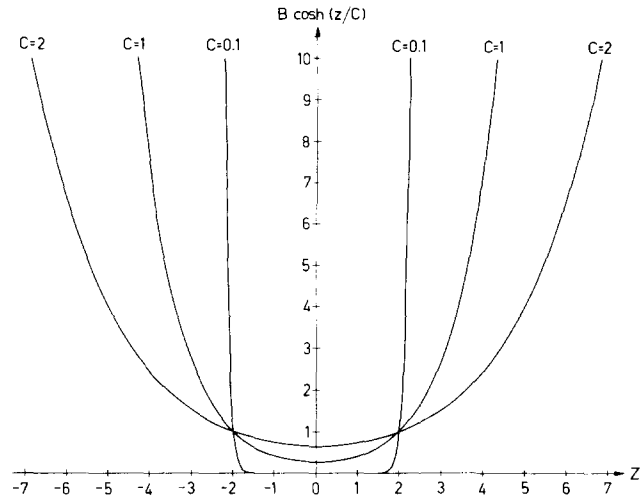


FIG. 1. Graphical representation of the function $B\cosh(z/C)$ for real values of z, when $C\,\mathrm{Arcosh}(1/B)$ is kept fixed and equal to 2, while C assumes various values. The draft of the figure was drawn automatically by means of a computer program written by T. Risch.

$$E=\frac{\hbar^2}{2m}A, \tag{27}$$

and inserting (26) and (27) into (2), we get

$$Q^2(z)=A-B\cosh(z/C). \tag{28}$$

The constant C is simply a scale factor, the constant B is a measure of the strength of the potential, and the constant A is to be determined by solving the eigenvalue problem. Since we are considering a bound state in a potential well, the function $Q^2(z)$ must be positive when z is real and sufficiently small in absolute value, which, according to (28), is possible only when $A>B$. When this condition is fulfilled, one finds from (28) that the zeros of the function $Q^2(z)$ are located at the points

$$z=\pm t+2\pi\nu Ci,\quad \nu=0,\ \pm1,\ \pm2,\cdots, \tag{29}$$

where

$$t=C\ln\left\{\frac{A}{B}+\left[\left(\frac{A}{B}\right)^2-1\right]^{1/2}\right\}. \tag{30}$$

The classical turning points, i.e., the zeros of $Q^2(z)$ which lie on the real z axis, are thus located at $z=\pm t$.

We shall use unmodified phase-integral approximations which means that we choose

$$Q^2_{\rm mod}(z)=Q^2(z). \tag{31}$$

By means of the definition (16d) of ϵ_0 and Eqs. (28) and (31) we obtain the formula

$$\epsilon_0(z)\,Q(z)=\frac{[(B/A)\cosh(z/C)]^2+4(B/A)\cosh(z/C)-5(B/A)^2}{16\sqrt{A}C^2[1-(B/A)\cosh(z/C)]^{5/2}}. \tag{32}$$

For the particular potential under consideration in

the present section it can be shown that, for the first-order approximation, the μ integral in the estimate (15) in Ref. 7 is considerably larger than the μ integral for a conveniently chosen path in the estimates (18) and (23) in Ref. 7. In the following we shall therefore use only the latter estimates. In these estimates the point z can be chosen to lie on the straight line $\mathrm{Im} z = C\pi$, which from (28) is seen to be an anti-Stokes line for the first-order phase-integral approximation, and the path for the μ integral, with the correct properties of monotonicity, can be chosen to be this straight line except for very large values of $|z|$ where the path shall approach the points $-\infty$ and $+\infty$ on the real axis. The value of the μ integral along this path is, however, the same as the value of the μ integral along the path $z = x + iC\pi$ from $x = -\infty$ to $x = +\infty$. Since $\cosh[(x+iC\pi)/C] = -\cosh[x/C]$ we therefore find, with the use of (32), that the μ integral corresponding to the first-order phase-integral approximation and the path described above is

$$\mu(-\infty,+\infty) = \int_{-\infty+iC\pi}^{+\infty+iC\pi} |\epsilon_0(z)\,Q(z)\,dz|$$

$$= \int_{-\infty}^{+\infty} \left| \frac{[(B/A)\cosh(x/C)]^2 - 4(B/A)\cosh(x/C) - 5(B/A)^2}{16C^2\sqrt{A}[1+(B/A)\cosh(x/C)]^{5/2}} \right| \times dx. \tag{33}$$

When $B/A \ll 1$, we can (for sufficiently small values of B/A) obtain a very accurate value of the integral (33) by replacing in the integrand $\cosh(x/C)$ by $\frac{1}{2}\exp\{x/C\}$. The result obtained after an elementary, straightforward calculation, in which further use of the assumption that $B/A \ll 1$ is also made, is

$$\mu(-\infty,+\infty) \approx \frac{5+2\sqrt{5}}{30C\sqrt{A}} \approx \frac{0.3}{C\sqrt{A}}, \quad A \gg B. \tag{34}$$

The first-order μ integral $\mu(-\infty,+\infty)$ appearing in (18) and (23) in Ref. 7, with a conveniently chosen path of integration, is thus small compared to unity when $A/B \gg 1$ and $C\sqrt{A} \gg 1$.

The smaller the constants B and C are, when $C\,\mathrm{Arcosh}(1/B)$ is kept fixed, the flatter will the bottom of the potential well become and the more square-well shaped the visual picture of the potential. In fact, when B and C tend to zero in such a way that $C\,\mathrm{Arcosh}(1/B)$ remains constant, the graphical representation of $V(z)$ for real values of z approaches a square-well potential with infinitely high walls located at $z = \pm C\,\mathrm{Arcosh}(1/B)$. For levels lying sufficiently low the energy eigenvalues can therefore be obtained from the Bohr—Sommerfeld *integer* quantization condition (6). (However, the generalization of the Bohr—Sommerfeld *integer* quantization condition to higher-order phase-integral approximations can be expected to give erroneous results for these low-lying levels due to the large values of the higher derivatives of the potential in the regions in which the slope of the potential changes rapidly from almost horizontal to almost vertical.) In spite of this fact it follows from (34) that for any positive values of B and C, however small, the Bohr—Sommerfeld *half-integer* quantization condition (5) becomes applicable when A is sufficiently large, i.e., when one considers sufficiently highly excited states. One can expect that, by using higher-order phase-integral approximations,[3,4] one should achieve a strong decrease in the value of the μ integral and hence a corresponding improvement of the eigenvalue in question obtained by means of the quantization condition (4).

In the left-hand member of the quantization condition (4) the contribution from the first-order approximation, with $Q^2(z)$ defined by (28) and the classical turning points $z = \pm t$ given by (30), is

$$L^{(1)} \equiv \tfrac{1}{2}\int_\Lambda Q(z)\,dz = \int_{-t}^{t} [A - B\cosh(z/C)]^{1/2}\,dz$$

$$= 2\int_0^t [A - B\cosh(x/C)]^{1/2}\,dx. \tag{35a}$$

To obtain the corresponding contribution from the third-order approximation: we use (32), note that $\frac{1}{2}\int_\Lambda$ can be replaced by $\int_{-\infty+iC\pi}^{+\infty+iC\pi}$, make the substitution $z = x + iC\pi$, note that $\cosh[(x+iC\pi)/C] = -\cosh[x/C]$ and that the integrand is an even function of x, getting

$$L^{(3)} \equiv \frac{1}{2}\int_\Lambda \tfrac{1}{2}\epsilon_0(z)\,Q(z)\,dz = \frac{1}{16C^2\sqrt{A}} \times \int_0^\infty \frac{[(B/A)\cosh(x/C)]^2 - 4(B/A)\cosh(x/C) - 5(B/A)^2}{[1+(B/A)\cosh(x/C)]^{5/2}}\,dx. \tag{35b}$$

When B/A is sufficiently small, one can easily calculate the integrals in (35a) and (35b) very accurately by replacing $\cosh(x/C)$ in the integrands by $\frac{1}{2}\exp\{x/C\}$. After elementary calculations, in which further use is also made of the assumption that $B/A \ll 1$, one obtains

$$L^{(1)} \approx 2C\sqrt{A}\left(\ln\frac{8A}{B} - 2\right), \quad A \gg B, \tag{36a}$$

$$L^{(3)} \approx -\frac{1}{12C\sqrt{A}}, \quad A \gg B. \tag{36b}$$

It is seen that the value obtained in (34) for the first-order μ integral is only somewhat larger than the absolute value of the third-order contribution $L^{(3)}$ obtained in (36b). Thus we realize that the μ integral in question yields a realistic upper bound for the error involved in the first-order Bohr—Sommerfeld half-integer quantization condition.

We shall conclude this section by reporting the results of some numerical calculations performed some years ago by Å. Nordlund (unpublished) which illustrate the accuracy obtainable by means of the quantization condition (4) when various orders of the phase-integral approximations are used. These calculations were performed for the parameter values $B = 1/\cosh 20 \approx 0.12\times10^{-9}$ and $C = \frac{1}{10}$ (cf. Fig. 1), and the results are shown in Table I. The exact eigenvalues, also given in Table I, were calculated by means of a modification of a program which had originally been written for quite a different purpose and could not be used for calculating the eigenvalues for even values of the quantum number s. This is the reason why the results in Table I have reference to only odd values of s. Table I shows clearly the great accuracy of the *half-integer* Bohr—Sommerfeld quantization condition [and in particular

TABLE I. Eigenvalues A for the Schrödinger equation corresponding to $Q^2(z)=A-B\cosh(z/C)$ with $B=1/\cosh 20\approx 0.12\times 10^{-9}$ and $C=0.1$ (see Fig. 1).

	Approximate eigenvalues A calculated by means of the quantization condition (4) for various orders of the phase-integral approximations					Exact eigenvalues A calculated by a numerical method
s	1st order	3rd order	5th order	7th order	9th order	
1	1.4	1.7856	1.93			1.7862
3	6.7	7.03	7.066	7.11	7.21	7.059
5	15.3	15.61	15.627	15.636	15.646	15.632
7	27.0	27.315	27.326	27.3283	27.3301	27.3290
9	41.7	42.004	42.0107	42.0119	42.01236	42.01242
11	59.2	59.576	59.5804	59.58091	59.58108	59.58118
13	79.6	79.954	79.9570	79.95726	79.95734	79.95738
15	102.8	103.076	103.0783	103.0785	103.0785	103.0785

its generalization (4) when higher-order phase-integral approximations are used] for the excited states. As has been shown in calculations performed by A. Hökback (unpublished) all figures in the columns corresponding to the first- and third-order approximations in Table I can also be reproduced by approximating the first- and third-order contributions in the left-hand member of the quantization condition (4) by the simple analytical expressions (36a) and (36b), respectively. The values of B and C chosen in the calculations now described are not small enough that the *integer* Bohr–Sommerfeld quantization condition (6) be valid for the first few lowest energy levels. One can, however, achieve this situation by choosing still smaller values of B and C connected by the relation $C\,\mathrm{Arcosh}(1/B)=2$, i.e., $B=1/\cosh(2/C)$; cf. Fig. 1.

4. POTENTIAL WITH HORIZONTAL BOTTOM AND LINEARLY RISING WALLS

It is illuminating to discuss also another example of a potential with steep walls, namely the potential

$$V(x)=\begin{cases}-\dfrac{\hbar^2}{2m}\beta(x+\alpha), & \text{when } x\leq-\alpha,\\ 0, & \text{when } -\alpha\leq x\leq\alpha,\\ \dfrac{\hbar^2}{2m}\beta(x-\alpha), & \text{when } \alpha\leq x,\end{cases}\tag{37}$$

where α and β are positive constants. Here we write x instead of z to emphasize that the whole discussion of this potential will be made on the real axis of the complex z plane. The potential (37), which is depicted in Fig. 2, approaches a square-well potential with infinitely high walls when $\beta\to+\infty$. For the Schrödinger equation corresponding to the potential (37) and the energy

$$E=\hbar^2k^2(2m)\tag{38}$$

the solution vanishing at $x=+\infty$ is, when $x\geq\alpha$, proportional to the Airy function $\mathrm{Ai}(\beta^{1/3}(x-\alpha-k^2/\beta))$. When $-\alpha\leq x\leq\alpha$ the solution is a linear combination of the functions $\sin(kx)$ and $\cos(kx)$. The ratio of the coefficients of these functions is obtained by matching at $x=\alpha$ the logarithmic derivative of the linear combination to the logarithmic derivative of the Airy function. By noting that wavefunctions of odd parity are equal to zero at $x=0$, while wavefunctions of even parity have their derivative equal to zero at $x=0$, one then obtains the exact quantization conditions

$$\tan(k\alpha)=\begin{cases}\gamma\dfrac{\mathrm{Ai}(-\gamma^2)}{\mathrm{Ai}'(-\gamma^2)} & \text{(odd parity)},\qquad (39a)\\ -\dfrac{1}{\gamma}\dfrac{\mathrm{Ai}'(-\gamma^2)}{\mathrm{Ai}(-\gamma^2)} & \text{(even parity)},\qquad (39b)\end{cases}$$

where

$$\gamma=k\beta^{-1/3}.\tag{40}$$

If, for fixed energy E, i.e., for fixed k, we let $\beta\to+\infty$, i.e., $\gamma\to+0$, the right-hand members of (39a, b) tend to $+0$ and $-\infty$, respectively, and the quantization conditions (39a, b) therefore give

$$2\alpha k=(s+1)\pi,\quad s=0,1,2,\cdots,$$

i.e., (39a, b) transform into the Bohr–Sommerfeld *integer* quantization condition (6) pertinent to the square-well potential approached by the potential (37) as $\beta\to+\infty$.

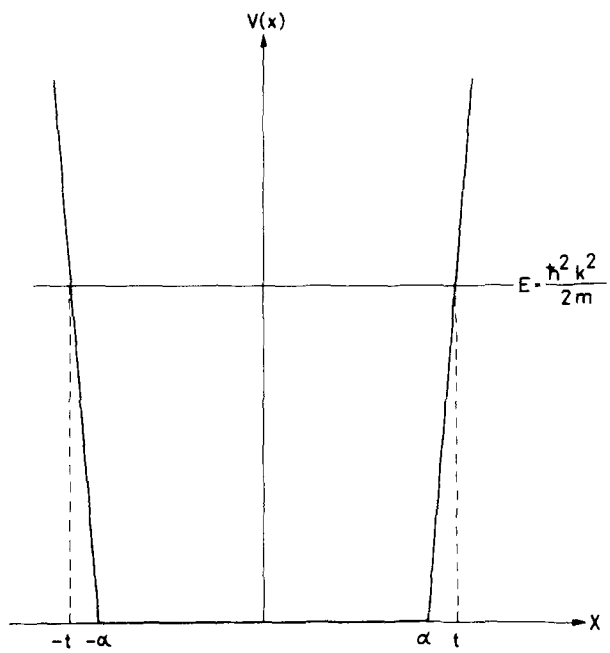


FIG. 2. Graphical representation of the potential defined by (37). The classical turning points corresponding to the energy $E=\hbar^2k^2/(2m)$ are denoted by $\pm t$.

If, however, the energy E is not kept fixed as $\beta \to +\infty$ but is made so large that $\gamma \gg 1$, the asymptotic expression (with only the dominating term retained) for the Airy function can be used in the quantization conditions (39a, b), and in this way the Bohr–Sommerfeld *half-integer* quantization condition (5) for the potential (37) is obtained.

Thus, if the walls of the potential (37), depicted in Fig. 2, are very steep, the eigenfunctions pertaining to the low-lying bound states are very close to zero at $x = \pm\alpha$, and the quantization condition is approximately the same as that for a square-well, and hence the Bohr–Sommerfeld *integer* quantization condition (6) is applicable. When the energy increases, the distance from the point $+\alpha$ or $-\alpha$, where the slope of the potential changes discontinuously, to the closest classical turning point ($+t$ or $-t$, respectively) increases and finally becomes so large that the appropriate connection formula for the first-order JWKB approximation of the wavefunction can be used at $x = \pm\alpha$. In this way one easily understands the reason why the *half-integer* Bohr–Sommerfeld quantization condition (5) becomes applicable for sufficiently excited bound states of the potential (37), however steep the walls may be (cf. Fig. 2), *if only the steepness is finite*. In a way one can therefore say that for the very highly excited states there is no continuous transition from a potential with β finite but very large to a potential with β infinite.

ACKNOWLEDGMENT

The question concerning the applicability of the Bohr–Sommerfeld quantization condition to the bound states of the potentials considered in Secs. 2 and 3 of the present paper was raised by Professor G. H. Wannier in a letter to us. We wish to thank Professor Wannier for an interesting and stimulating correspondence which initiated the investigations presented in the present paper.

[1] N. Fröman and P. O. Fröman, *JWKB-Approximation, Contributions to the Theory* (North-Holland, Amsterdam, 1965) (Russian translation: MIR, Moscow, 1967).
[2] N. Fröman, Ark. Fys. **31**, 445–51 (1966).
[3] N. Fröman, Ark. Fys. **32**, 541–48 (1966).
[4] N. Fröman, Ann. Phys. (N.Y.) **61**, 451–64 (1970).
[5] N. Fröman and P. O. Fröman, Ann. Phys. (N.Y.) **83**, 103–7 (1974).
[6] N. Fröman and P. O. Fröman, Nuovo Cimento B **20**, 121–32 (1974).
[7] N. Fröman and P. O. Fröman, J. Math. Phys. **19**, 1830 (1978).
[8] N. Fröman, Phys. Lett. A **48**, 137–9 (1974).
[9] N. Fröman and P. O. Fröman, J. Math. Phys. **18**, 903–6 (1977).
[10] P. O. Fröman, Ann. Phys. (N.Y.) **88**, 621–30 (1974).
[11] D. ter Haar (Ed.), *Selected Problems in Quantum Mechanics*, revised and augmented 2nd ed. of Gol'dman, Krivchenkov, Kogan, and Galitskii, *Problems in Quantum Mechanics* (Academic, New York, 1964).

On phase-integral quantization conditions for bound states in one-dimensional smooth single-well potentials

Nanny Fröman and Per Olof Fröman

Institute of Theoretical Physics, University of Uppsala, Uppsala, Sweden
(Received 7 December 1977)

Previous work by the present authors on phase-integral quantization conditions for single-well potentials is extended and generalized. Exact as well as approximate quantization conditions are considered and investigated more fully than previously. Not only their necessity but also their sufficiency is treated. An improvement of the estimate of the error of the generalized quantal Bohr–Sommerfeld quantization condition is also given.

1. INTRODUCTION

In the first few chapters of a previous publication[1] the present authors worked out a general method for handling connection problems for the first-order JWKB-approximation and modifications of it. In chapter 10 of Ref. 1 this method was used for proving that certain quantization conditions, obtained in exact form, are necessary for the existence of bound states.

In Ref. 2 N. Fröman showed that the sum of the odd-order terms in the usual JWKB-expansion can be expressed in terms of the sum of the even-order terms, and thus she obtained a formal solution of the time-independent Schrödinger equation containing only the even-order terms. By truncating the sum of the even-order terms occurring in the amplitude as well as in the phase of this solution, she then obtained higher-order phase-integral approximations that, in contrast to the higher-order JWKB-approximations, can be handled by the method developed in Ref. 1. The quantization conditions given in chapter 10 of Ref. 1 are therefore valid also for the new type of higher-order phase-integral approximations. The quantal Bohr—Sommerfeld quantization condition, with higher-order corrections included, was given explicitly in Ref. 2. Some general properties of the above-mentioned higher-order phase-integral approximations are discussed in another paper by N. Fröman.[3] The general problem concerning modifications of these higher-order phase-integral approximations, useful in situations in which the unmodified approximations would fail or be less convenient than modified ones, was treated in two papers by the present authors (see Ref. 4 and pp. 126—31 in Ref. 5).

In the present paper, which is based on the use of these phase-integral approximations of arbitrary order, we present an extension of the above-mentioned results on the bound-state problem. We shall thus assume that $q(z)$ is defined as in Ref. 4 [cf. Eqs. (12), (13), (7), (3), (4a) in Ref. 4]. This function $q(z)$ is approximately equal to $Q_{\mathrm{mod}}(z)$ except in the neighborhood of the zeros of $Q^2_{\mathrm{mod}}(z)$. At those points the function $q(z)$ has singularities, and in the neighborhood of each such singularity $q(z)$ has a certain number of zeros (cf. Ref. 3). Since we shall consider an ordinary quantal bound-state problem, the functions $Q^2(z)$ and $Q^2_{\mathrm{mod}}(z)$ are assumed to be real on the real axis. We shall frequently refer to equations in Ref. 1, assuming, however, that $q(z)$ is defined as stated above. To make the paper more easily readable we have therefore in an Appendix collected those formulas from Ref. 1 which we refer to. It is important to note that for the validity of the treatment given in chapter 10 of Ref. 1 it is sufficient to assume that the functions $q(z)$ and $Q(z)$ are similar in shape except in certain regions in the neighborhood of the generalized classical turning points (i.e., the zeros of Q^2_{mod} on the real axis) or possibly existing singularities of $Q^2(z)$, where the two functions $q(z)$ and $Q(z)$ differ considerably. The phase of the function $q^{1/2}(z)$ is chosen such that sufficiently far away from the generalized classical turning points the values of $q^{1/2}(z)$ on the different parts of the real axis are those indicated in Fig. 10.1(b) of Ref. 1 [cf. our Fig. 1(b)]. For the single-well potential under consideration and real values x of z, the function $w(x)$ is given by Eqs. (11a,b,c) in Ref. 2 with the contours of integration depicted in Fig. 1 of the present paper. As a consequence, the integral $\int_x^{x''}|q(x)|dx$, occurring in the formulas in chapter 10 of Ref. 1, shall be replaced by $\frac{1}{2}\int_\Lambda q(z)\,dz$, where Λ is the contour depicted in Fig. 1 of the present paper. When these changes are assumed to have been made, the equations in Ref. 1, to which we shall refer, become valid for the modified phase-integral approximations of arbitrary order used in the present paper. As in chapter 10 of Ref. 1 the eigenvalue problem is formulated for the interval $(-\infty, +\infty)$, but the treatment applies also to other ranges of the physical variable z. Thus, to treat a radial problem we simply replace $-\infty$ by $+0$ in our formulas.

In passing we remark that from (10.8a,b) in Ref. 1 it follows that, to a given value of E there corresponds at the most one eigenfunction (except for an arbitrary constant factor). This is in accordance with the well-known fact that one-dimensional potentials with certain properties of regularity have nondegenerate bound states.

The present paper provides a deeper treatment of what is discussed in Ref. 1 from the middle of p. 106 to the end of chapter 10.

2. THE EXACT QUANTIZATION CONDITION $F_{12}(-\infty, +\infty) = 0$

Up to the middle of p. 106 in Ref. 1 it has been shown that the condition (10.13) in Ref. 1, i.e.,

$$F_{12}(-\infty, +\infty) = 0, \tag{1}$$

is necessary for the existence of a bound state. (Of the

 0022-2488/78/1909-1830$1.00

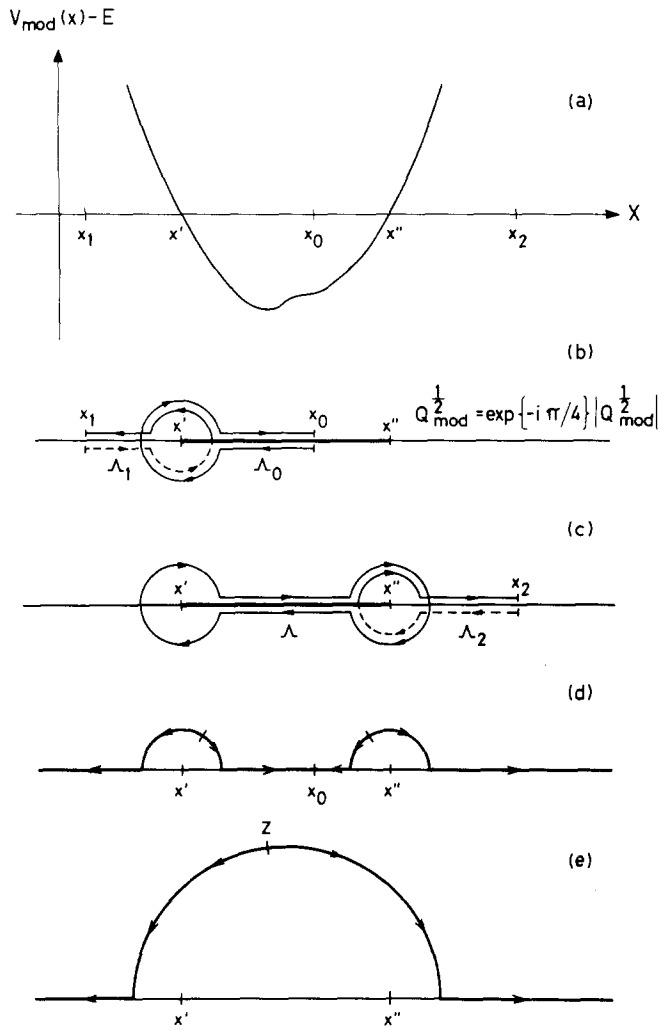


FIG. 1. (a) For the case of a bound state in a general single-well potential the figure shows the qualitative behavior of $V_{mod}(x)-E$, where $V_{mod}(x)$, called the modified potential, is related to $Q^2_{mod}(x)$ according to the formula $Q^2_{mod}(x)=(2m/\hbar^2)[E - V_{mod}(x)]$. (b), (c) Contours of integration on the Riemann surface to be used for obtaining $w(x)$. The parts of these contours which lie on the second Riemann sheet are indicated by dashed lines. The phase of $Q^{1/2}_{mod}(x)$ to the right of x'' on the real axis of the first Riemann sheet is indicated. The heavy line indicates a cut. (d), (e) The path of integration in the complex z plane for the μ integral in the estimates (15) and (18), respectively. The arrows indicate the directions in which $|\exp[i\int^z q(z)dz]|$ increases (in the nonstrict sense).

two proofs given in Ref. 1, the one given immediately below Eq. (10.13) in Ref. 1 is to be preferred.)

We shall now show that the above quantization condition (1) is also sufficient. Let us therefore suppose that this condition is fulfilled for a certain value E of the energy, and let us select a special solution $\psi(z)$ of the Schrödinger equation for which $a_1(+\infty)=0$ but $a_2(+\infty)\neq 0$. For this special solution the formulas (10.8a,b) in Ref. 1 are valid. Since the condition (1) above is assumed to be fulfilled, it follows from (10.8a) in Ref. 1 that $a_1(-\infty)=0$. As a consequence of this it follows, according to pp. 104—105 in Ref. 1, that the limit $a_2(-\infty)$ exists and is finite and that $a_1(z)$ and $a_2(z)$ are also given by formulas which one obtains by replacing $+\infty$ by $-\infty$ in (10.8a,b) in Ref. 1. Obviously $a_2(-\infty)\neq 0$, for otherwise the solution $\psi(z)$ would be identically equal to zero, in contradiction to the definition of this function. In the same way as on p. 105 in Ref. 1 we now obtain (10.12) in Ref. 1 and the corresponding formula with $a_2(+\infty)$ replaced by $a_2(-\infty)$. From these two formulas it follows that for the special solution $\psi(z)$ considered, the integral $\int_{-\infty}^{+\infty}|\psi(x)|^2dx$ is finite, provided that the integrals $\int_{-\infty}|f_2(x)|^2dx$ and $\int^{+\infty}|f_2(x)|^2dx$ are both convergent. Thus $\psi(z)$ is the wavefunction of a bound state, and we have proved that (10.13) in Ref. 1, i.e., (1) above, is also a sufficient condition for the existence of a bound state.

Let us note that if the modified phase-integral functions of arbitrary order defined by Eqs. (11a,b), (12), (13) in Ref. 4 were exact solutions of the Schrödinger equation (cf. pp. 11 and 110—112 in Ref. 1), the quantities ϵ and μ defined by Eqs. (18) and (19) in Ref. 5 would be identically zero, and the matrix $F(z,z_0)$ would be exactly a unit matrix for all values of z and z_0, and in particular we would have $F_{12}(-\infty,+\infty)=0$. Since this relation would be true for all values of E, the quantization condition (1) would be void of any information. This situation, however, cannot arise for the kind of problem which we are considering, for if $F_{12}(-\infty,+\infty)$ were equal to zero for every value of E, the boundary condition that $a_1(+\infty)=0$ and $a_2(+\infty)\neq 0$ would imply that also $a_1(-\infty)=0$ and $a_2(-\infty)\neq 0$, and hence that bound states would exist for all values of E, which is impossible for the eigenvalue problem under consideration. Thus the quantity ϵ, defined by Eq. (18) in Ref. 5, cannot be identically equal to zero in a bound-state problem, when $Q^2_{mod}(z)$ is chosen appropriately so that the corresponding modified potential has the properties shown in Fig. 1(a) of the present paper, which means that $q(z)$ has the properties shown in Fig. 10.1 of Ref. 1 with the alteration that for phase-integral approximations of higher order the function $q(z)$ has not simple zeros but higher-order poles at the generalized classical turning points x' and x'', which by definition are the real zeros of $Q^2_{mod}(z)$.

Letting z be an arbitrary point in the region of the complex z plane under consideration, we obtain, because of the multiplication rule (3.27) in Ref. 1, the identity

$$F_{12}(-\infty,+\infty) = F_{11}(-\infty,z)F_{12}(z,+\infty)+F_{12}(-\infty,z)F_{22}(z,+\infty), \quad (2)$$

which by means of the inversion formula (3.20) in Ref. 1 can be written

$$F_{12}(-\infty,+\infty) = -F_{11}(-\infty,z)F_{12}(+\infty,z)+F_{12}(-\infty,z)F_{11}(+\infty,z). \quad (3)$$

The quantities appearing in this identity exist and are finite according to (4.5a,b) in Ref. 1.

By means of the identity (3) the necessary and sufficient condition (1) for the existence of a bound state can be written [cf. p. 106, line 8 from above in Ref. 1 together with the inversion formula (3.20) in Ref. 1]

$$\frac{F_{11}(-\infty,z)}{F_{11}(+\infty,z)} = \frac{F_{12}(-\infty,z)}{F_{12}(+\infty,z)}. \quad (4)$$

This quantization condition can also be obtained from the formulas [cf. pp. 105–106 in Ref. 1 and the discussion below (1) in the present paper]

$$F_{12}(z,-\infty)a_2(-\infty)=a_1(z)$$
$$=F_{12}(z,+\infty)a_2(+\infty),$$
$$F_{22}(z,-\infty)a_2(-\infty)=a_2(z)$$
$$=F_{22}(z,+\infty)a_2(+\infty),$$

and the inversion formula (3.20) in Ref. 1.

3. THE GENERALIZED QUANTAL BOHR-SOMMERFELD QUANTIZATION CONDITION AND CLOSELY RELATED QUANTIZATION CONDITIONS

Let x_0 be a point on the real axis in the classically allowed region, lying far enough from the generalized classical turning points x' and x'' to be outside the regions around those points in which there are zeros of $q(z)$, when higher-order phase-integral approximations are used. With the phase of $Q_{\mathrm{mod}}^{1/2}(x)$ chosen as shown in Fig. 1(b) in the present paper (cf. Fig. 10.1(b) in Ref. 1, we then get from (5.9a) in Ref. 1 the symmetry relations

$$F_{12}(-\infty,x_0)$$
$$=i\exp[2\mathrm{Im}w(x_0)-2i\,\mathrm{Re}w(-\infty)]F_{11}^{*}(-\infty,x_0), \tag{5a}$$

$$F_{12}(+\infty,x_0)$$
$$=-i\exp[2\mathrm{Im}w(x_0)-2i\,\mathrm{Re}w(+\infty)]F_{11}^{*}(+\infty,x_0), \tag{5b}$$

which are valid irrespective of the choice of the lower limit in the integral (3.3) in Ref. 1 defining $w(z)$. If, for instance, this choice is made such as to correspond to x' being the lower limit in the integral when the first-order phase-integral approximation is used, the formulas (5a, b) become the same as (30a, b) in Ref. 6. Substituting (5a, b) into (3) with z chosen equal to x_0, we get the identity

$$F_{12}(-\infty,+\infty)=iF_{11}^{*}(-\infty,x_0)F_{11}(+\infty,x_0)$$
$$\times\exp[2\mathrm{Im}w(x_0)-2i\,\mathrm{Re}w(-\infty)]$$
$$\times\left\{1+\left[\frac{F_{11}(-\infty,x_0)}{F_{11}(+\infty,x_0)}\Big/\left(\frac{F_{11}(-\infty,x_0)}{F_{11}(+\infty,x_0)}\right)^{*}\right]\right.$$
$$\left.\times\exp\{-2i[\mathrm{Re}w(+\infty)-\mathrm{Re}w(-\infty)]\}\right\}$$

which, by the use of Eq. (3.3) in Ref. 1, defining $w(z)$, and the fact that the analytical function $q(z)$ is real in the interval (x',x'') of the real axis, can be written as follows [cf. Eq. (31) in Ref. 6],

$$F_{12}(-\infty,+\infty)$$
$$=iF_{11}^{*}(-\infty,x_0)F_{11}(+\infty,x_0)$$
$$\times\exp[2\mathrm{Im}w(x_0)-2i\mathrm{Re}w(-\infty)]$$
$$\times\left\{1-\exp\left[-2i\left(\tfrac{1}{2}\int_{\Lambda}q(z)\,dz-\tfrac{1}{2}\pi\right.\right.\right.$$
$$\left.\left.\left.-\arg\frac{F_{11}(-\infty,x_0)}{F_{11}(+\infty,x_0)}\right)\right]\right\}, \tag{6}$$

where Λ is the contour defined in Fig. 1(c) in the present paper. We emphasize that (6) is an identity, which is valid for every possible value of E provided that x_0 is located on the real axis between the points x' and x'', and that the phase of $Q_{\mathrm{mod}}^{1/2}(x)$ is chosen as described above. Inserting this identity into the quantization condition (1), and noting that $F_{11}(-\infty,x_0)$ and $F_{11}(+\infty,x_0)$ are different from zero, as is seen from the estimate (6.13a) in Ref. 1, valid when the appropriate μ integrals are small compared to unity, we get

$$\tfrac{1}{2}\int_{\Lambda}q(z)\,dz=(s+\tfrac{1}{2})\pi+\arg\frac{F_{11}(-\infty,x_0)}{F_{11}(+\infty,x_0)}, \tag{7}$$

where s is an integer which, because of the positive nature of $Q_{\mathrm{mod}}(x)$ in the classically allowed region of the real axis and the smallness of $\arg[F_{11}(-\infty,x_0)/F_{11}(+\infty,x_0)]$, cannot assume negative values. Since the quantization condition (1) is both necessary and sufficient, so also is the quantization condition (7), as is obvious from the derivation of it given above. Specializing (7) to the first-order phase-integral approximation, we obtain, because of the inversion formula (3.20) in Ref. 1, the quantization condition (10.20) in Ref. 1 with $z=x_0$.

From (3.18), (3.13), and (3.3) in Ref. 1 it follows that

$$\frac{\partial}{\partial z}F(z,z_0)=\tfrac{1}{2}i\epsilon(z)q(z)$$
$$\times\begin{pmatrix}1 & \exp[-2iw(z)]\\ -\exp[2iw(z)] & -1\end{pmatrix}F(z,z_0). \tag{8}$$

Using this formula and the formulas (3.20) and (3.27) in Ref. 1, and recalling also (4.5a, b) in Ref. 1, we obtain the identities

$$\frac{\partial}{\partial z}\frac{F_{11}(-\infty,z)}{F_{11}(+\infty,z)}=\frac{\tfrac{1}{2}i\epsilon(z)q(z)\exp[2iw(z)]}{[F_{11}(+\infty,z)]^2}F_{12}(-\infty,+\infty) \tag{9a}$$

and

$$\frac{\partial}{\partial z}\frac{F_{12}(-\infty,z)}{F_{12}(+\infty,z)}=\frac{\tfrac{1}{2}i\epsilon(z)q(z)\exp[-2iw(z)]}{[F_{12}(+\infty,z)]^2}F_{12}(-\infty,+\infty). \tag{9b}$$

Since it can be assumed that $\epsilon(z)$ is not identically equal to zero (cf. Sec. 2 of the present paper), it follows from (9a) that if, and only if, the quantization condition (1) is fulfilled, we have

$$\frac{\partial}{\partial z}\frac{F_{11}(-\infty,z)}{F_{11}(+\infty,z)}=0 \tag{10}$$

for every value of z. The validity of (10) is therefore, like condition (1), a necessary and sufficient quantization condition. Since this quantization condition means that the quotient $F_{11}(-\infty,z)/F_{11}(+\infty,z)$ is independent of z, we can, letting $z\to+\infty$ [cf. (4.5a) in Ref. 1], write it as

$$\frac{F_{11}(-\infty,z)}{F_{11}(+\infty,z)}=F_{11}(-\infty,+\infty),\quad z\text{ arbitrary}, \tag{11a}$$

or, inverting the quotient and letting $z \to -\infty$, as

$$\frac{F_{11}(+\infty,z)}{F_{11}(-\infty,z)} = F_{11}(+\infty,-\infty), \quad z \text{ arbitrary.} \tag{11b}$$

The quantization conditions (11a) and (11b), which have here been obtained as both necessary and sufficient quantization conditions, are seen to agree with the quantization conditions (10.14b) and (10.14a), respectively, in Ref. 1, when use is made of the inversion formula (3.20) in Ref. 1. In Ref. 1 we proved only the necessity of these quantization conditions.

According to (10.17b) in Ref. 1 we have a symmetry relation stating that

$$F_{22}(-\infty,+\infty)\exp\{i[\tfrac{1}{2}\textstyle\int_\Lambda q(z)\,dz - \tfrac{1}{2}\pi]\} \text{ is real.} \tag{12}$$

Using (3.20) in Ref. 1, we can write (12) as

$$\tfrac{1}{2}\textstyle\int_\Lambda q(z)\,dz = (s+\tfrac{1}{2})\pi - \arg F_{11}(+\infty,-\infty), \tag{12'}$$

where s is an unspecified integer. Combining this symmetry relation with the quantization condition (11b), we obtain the necessary and sufficient quantization condition [cf. (10.20) in Ref. 1 and (13) in Ref. 2]

$$\frac{1}{2}\int_\Lambda q(z)\,dz = (s+\tfrac{1}{2})\pi + \arg\frac{F_{11}(-\infty,z)}{F_{11}(+\infty,z)},$$
$$s = 0,1,2,\cdots, \tag{7'}$$

where z is any point in the complex plane. The reason why s in (7') is restricted to nonnegative integers is that $Q_{\text{mod}}(x)$ is positive in the classically allowed region and that $|\arg[F_{11}(-\infty,z)/F_{11}(+\infty,z)]| \ll 1$. We remark that one can also derive the quantization condition (7') by combining the quantization conditions (7) and (10). We emphasize that (12') is a symmetry relation, i.e., an identity valid for any value of E, while (7) and (7') are quantization conditions, by means of which the energy eigenvalues can be determined.

Excluding for a moment cases in which the generalized classical turning points x' and x'' lie close together, we shall estimate the last term in (7), i.e., the last term in (7') with the point z chosen as a point x_0 on the real axis between x' and x'' (cf. Fig. 1). Since the points x' and x'' are well separated, there exist paths from x_0 to $\pm\infty$, proceeding as shown in Figs. 6.1(b) and 6.2(b) in Ref. 1, for which $\mu(\pm\infty,x_0) \ll 1$, where in our present treatment μ is defined by Eqs. (19), (18), and (16) in Ref. 5. As has been mentioned in chapter 3 of Ref. 1, the diagonal elements of the F matrix do not depend on the choice of the fixed lower limit in the integral (3.3) in Ref. 1 defining $w(z)$. Independently of how this lower limit of integration has been chosen, we can therefore use the estimate (6.13a) in Ref. 1, generalized to the modified phase-integral approximations of arbitrary order (cf. Ref. 4 and pp. 126–31 in Ref. 5), i.e. (cf. p. 107 in Ref. 1),

$$|F_{11}(\pm\infty,x_0) - 1| \le \mu(\pm\infty,x_0)$$
$$+\text{higher powers of } \mu(\pm\infty,x_0), \tag{13}$$

where $x' < x_0 < x''$, and $\mu(\pm\infty,x_0)$ denotes the μ integral [cf. Eqs. (19), (18), and (16) in Ref. 5] from x_0 to $\pm\infty$ along such a path as is shown in our Fig. 1(d) [cf. also Figs. 6.1(b) and 6.2(b) in Ref. 1]. From (13) it follows that

$$\left|\frac{F_{11}(-\infty,x_0)}{F_{11}(+\infty,x_0)} - 1\right| \le \mu(-\infty,+\infty)$$
$$+\text{higher powers of } \mu(-\infty,+\infty) \tag{14}$$

and hence

$$\left|\arg\frac{F_{11}(-\infty,x_0)}{F_{11}(+\infty,x_0)}\right| \le \mu(-\infty,+\infty)$$
$$+\text{higher powers of } \mu(-\infty,+\infty), \tag{15}$$

where $\mu(-\infty,+\infty)$ denotes the μ integral from $-\infty$ to $+\infty$ along a path, the two parts of which, namely those from $-\infty$ to x_0 and from x_0 to $+\infty$ possess the properties of monotonicity described in Sec. 6.3 of Ref. 1. For this μ integral the integration is thus to be performed along a path in the complex z plane proceeding as shown in Fig. 1(d) of the present paper. The radii of the two semicircles in this figure shall be chosen such that $|\exp[i\int^z q(z)\,dz]|$ has precisely two minima on the path of integration from $-\infty$ to $+\infty$ [cf. Fig. 1(b) in the present paper].

If the generalized classical turning points x' and x'' are not well separated, which is the case for the lowest energy eigenstates, we cannot expect the quantity $\mu(-\infty,+\infty)$, occurring in the right-hand members of (14) and (15), to be very small. Therefore, we shall now also derive another estimate for the correction term in the exact quantization condition (7'), which is useful both when x' and x'' are well separated and when those points lie close together, whether the bottom of the potential well has an approximately parabolic shape or not. We shall thus estimate, and even approximately calculate, the last term in (7') by choosing the point z appearing there in a different way than in the previous estimate, i.e., (15). Considering, for the actual single-well potential, the picture of the Stokes' and anti-Stokes' lines corresponding to the first-order approximation (cf. Fig. 1 in Ref. 7), and recalling that far away from x' and x'' these Stokes' and anti-Stokes' lines are almost the same as those corresponding to higher-order phase-integral approximations, we realize that for a single-well potential we can always, for any conveniently chosen order of approximation, choose z such that there exists a path from z to $-\infty$ as well as a path from z to $+\infty$ along which $|\exp[i\int^z q(z)\,dz]|$ increases monotonically in the direction from z to $\pm\infty$ as shown by the arrows in Fig. 1(e). We can therefore use the basic estimate (4.3a) in Ref. 1 with $M=1$, generalized to modified phase-integral approximations of arbitrary order (cf. Ref. 4 and pp. 126–31 in Ref. 5), getting

$$|F_{11}(\pm\infty,z) - 1| \le \tfrac{1}{2}\{\exp[\mu(\pm\infty,z)] - 1\}, \tag{16}$$

where $\mu(\pm\infty,z)$ denotes the μ integral along the path of monotonicity from z to $\pm\infty$ [cf. Fig. 1(e)]. When $\mu(\pm\infty,z) \ll 1$ it easily follows from (16) that

$$\left|\frac{F_{11}(-\infty,z)}{F_{11}(+\infty,z)} - 1\right| \le \tfrac{1}{2}\mu(-\infty,+\infty)$$
$$+\text{higher powers of } \mu(-\infty,+\infty) \tag{17}$$

and hence

$$\left|\arg\frac{F_{11}(-\infty,z)}{F_{11}(+\infty,z)}\right| \le \tfrac{1}{2}\mu(-\infty,+\infty)$$

$$+\text{higher powers of } \mu(-\infty,+\infty), \tag{18}$$

where $\mu(-\infty,+\infty)$ now denotes the μ integral along a path [cf. Fig. 1(e)] from $-\infty$ via z to $+\infty$, on which $|\exp\{i\int^z q(z)\,dz\}|$ has a single minimum [at the point z appearing in the last term of (7′)].

In the same manner as (15), formula (18) gives only an upper bound for the correction term in the exact quantization condition [(7) or (7′)]. We can, however, make the derivation of (18) somewhat more sophisticated such that we obtain not only an upper bound but an approximate formula for the quantity in question. In fact, modifying the derivation of (16) slightly by also taking explicitly into account the second term in the series defining $F_{11}(\pm\infty,z)$ (cf. pp. 18, 26, and 27 in Ref. 1), we obtain

$$\begin{aligned}&\left|F_{11}(\pm\infty,z)-1-\tfrac{1}{2}i\int_z^{\pm\infty}\epsilon(z)\,q(z)\,dz\right|\\&\leq\tfrac{1}{2}\{\exp[\mu(\pm\infty,z)]-1-\mu(\pm\infty,z)\}\\&=\tfrac{1}{4}[\mu(\pm\infty,z)]^2\\&\quad+\text{higher powers of } \mu(\pm\infty,z).\end{aligned}\tag{19}$$

From this formula and the definition of the μ integral [cf. Eq. (19) in Ref. 5] it follows that

$$\begin{aligned}\frac{F_{11}(-\infty,z)}{F_{11}(+\infty,z)}&=1-\tfrac{1}{2}i\int_{-\infty}^{+\infty}\epsilon(z)q(z)\,dz\\&\quad+O[\mu^2(-\infty,+\infty)]\end{aligned}\tag{20}$$

and hence

$$\begin{aligned}\arg\frac{F_{11}(-\infty,z)}{F_{11}(+\infty,z)}&=-\operatorname{Re}\int_{-\infty}^{+\infty}\tfrac{1}{2}\epsilon(z)q(z)\,dz\\&\quad+O[\mu^2(-\infty,+\infty)],\end{aligned}\tag{21}$$

where $\mu(-\infty,+\infty)$ is the same μ integral as in the estimates (17) and (18) [cf. Fig. 1(e)], and the symbol $O(\mu^2)$ is used to denote a quantity which is at the most of the order of magnitude of μ^2. Since $Q^2(z)$ and $Q^2_{\text{mod}}(z)$ are assumed to be real on the real axis, it follows from Eq. (3) in Ref. 4 that $\epsilon_0(z)$ is real on the real axis, and from Eq. (7) in Ref. 4, with $Y_0=1$, and Eq. (4a) in Ref. 4 it then follows that the quantities Y_{2n} are real on the real axis. From Eqs. (16) and (18) in Ref. 5 one can then conclude that $q(z)/Q_{\text{mod}}(z)$ and $\epsilon(z)$ are real on the real axis. These functions are analytic functions which, in the region of the complex z plane under consideration, have singularities only at the zeros of $Q^2_{\text{mod}}(z)$, if a convenient modification is used. Since the complex z plane is assumed to be cut as in our Fig. 1, we now realize that

$$\operatorname{Re}\int_{-\infty}^{+\infty}\tfrac{1}{2}\epsilon q\,dz=\tfrac{1}{2}\int_\Lambda\tfrac{1}{2}\epsilon q\,dz,\tag{22}$$

where Λ is the contour shown in Fig. 1(c). Formula (21) can therefore be written as follows,

$$\begin{aligned}\arg\frac{F_{11}(-\infty,z)}{F_{11}(+\infty,z)}&=-\frac{1}{2}\int_\Lambda\tfrac{1}{2}\epsilon(z)\,q(z)\,dz\\&\quad+O[\mu^2(-\infty,+\infty)].\end{aligned}\tag{23}$$

On the right-hand side of (23) the last term may be expected to be of the order of magnitude of the square of the first term, unless large cancellations occur in the integral.

From the exact quantization condition (7) and the estimate (15), as well as from the exact quantization condition (7′) and the estimate (18), we obtain, if $\mu(-\infty,+\infty)\ll 1$, the approximate quantization condition

$$\tfrac{1}{2}\int_\Lambda q(z)\,dz=(s+\tfrac{1}{2})\pi,\tag{24}$$

where s is a nonnegative integer. The correction term, which has been deleted in obtaining the approximate quantization condition (24) from the exact quantization condition (7) or (7′), can be expected to be very small for high energy levels, i.e., when x' and x'' are well separated [since the μ integrals in the estimates (15) and (18) are then very small compared to unity]. It can also be expected to be small for low-lying energy levels, i.e., when x' and x'' are not well separated, especially if the bottom of the potential well has approximately parabolic shape [since the μ integral in the estimate (18) is then very small compared to untiy due to the absence of zeros or singularities of $Q^2_{\text{mod}}(z)$ in the relevant region of the complex z plane]. If the potential is not only approximately but exactly parabolic in shape (linear harmonic oscillator), with the minimum at $x=0$, we may choose the unmodified expression for the function $q(z)$ appearing in the phase-integral approximations of arbitrary order, and we may choose the particular point z in our Fig. 1(e) as $z=+iR$, where R is a positive number larger than x'' $(=-x')$. When $R\to\infty$, the μ integral corresponding to the first-order as well as to higher-order approximations for the path shown in our Fig. 1(e) then tends to zero. Hence the estimate (18) gives $\arg[F_{11}(-\infty,+i\infty)/F_{11}(+\infty,+i\infty)]=0$. Therefore, the last term on the right-hand side of (7′) is exactly equal to zero, and hence we have rigorously proved the well-known fact that one obtains the exact values of the energy levels of a harmonic oscillator from the quantization condition (24) already in the first-order approximation, the higher-order terms yielding no contributions, if $q(z)$ is chosen to correspond to unmodified phase-integral approximations. A corresponding proof restricted to the first-order approximation was given on p. 109 in Ref. 1.

When unmodified phase-integral approximations are used, i.e., when $Q_{\text{mod}}(z)=Q(z)$, the quantization condition (24) is the same as the generalized quantal Bohr–Sommerfeld quantization condition which was used by Dunham[8,9] in his treatment of the energy levels of diatomic molecules. The application of this unmodified quantization condition to such radial problems has been criticized and also remedied by Krieger and Rosenzweig[10] (cf. also a paper by Howard[11]). By using, in (24), modified phase-integral approximations obtained by choosing $Q^2_{\text{mod}}(z)=Q^2(z)-1/(4z^2)$, we get directly and in explicit form a quantization condition which is equivalent to the final result of the above-mentioned procedure devised by Krieger and Rosenzweig to correct Dunham's treatment of the energy levels of diatomic molecules. When the interatomic potential is more singular than the centrifugal barrier at small interatomic distances, the unmodified phase-integral approximations can also be used, and Dunham's approach[8,9] is then justified.

Using (23), we obtain from (7′) the approximate quantization condition

$$\tfrac{1}{2}\int_\Lambda q(z)\,dz = (s+\tfrac{1}{2})\,\pi - \tfrac{1}{2}\int_\Lambda \tfrac{1}{2}\epsilon(z)\,q(z)\,dz, \quad s=0,1,2,\cdots, \tag{25}$$

where the error can be expected to be of the order of magnitude of the square of the last integral in (25), if large cancellations do not occur in this integral. At first sight one might think that (25) represents an essential improvement of (24), but actually this does not seem to be the case, since, instead of applying (25) in a certain order of approximation, it should in general be preferable to apply (24) in a higher order. In this connection we remark that (25) in the first-order approximation is obtained from (24) in the third-order approximation simply by moving the third-order contribution from the left-hand to the right-hand member. For approximations of higher order than the third in (24), the connection between (24) and (25) is, however, more complicated.

The discussion below (25) indicates that the attempts, sometimes met with in the literature, to replace the first-order JWKB-quantization condition for a single-well potential by a condition of the type

$$\int_{x'}^{x''} Q(x)\,dx = (s+\Delta)\,\pi, \quad s=0,1,2,\cdots, \quad \text{with } \Delta \neq \tfrac{1}{2},$$

seem to be of limited value and rather unnatural within our present framework, in which the use of higher-order phase-integral approximations, consistently modified if necessary, is in general the most convenient way of increasing the accuracy. We shall now elucidate our point of view somewhat further. Consider for this sake a radial Schrödinger equation

$$\frac{d^2\psi}{dz^2} + Q^2(z)\,\psi = 0, \tag{26}$$

and let us assume that there exists a simple path Λ', as shown in Fig. 2(b), emerging from $+\infty$, encircling the classical turning points and the origin, and returning to $+\infty$, on which $Q^2(z)$ and $Q^2_{\text{mod}}(z)$ are approximately constant. (Note that in the complex z plane there is no need for any cut crossing Λ', but inside Λ' there is at least one cut, namely between the points x' and x''.) With the phase of $Q_{\text{mod}}(z)$ chosen as indicated in Fig. 2(b), the quantity $|\exp[-iw(z)]|$ is zero at $+\infty$ and increases monotonically when z moves on the path Λ', either in the lower or in the upper half of the complex z plane, from $+\infty$ to the point where Λ' crosses the negative real axis. Therefore, a solution of (26) which tends to zero as $z \to +\infty$ can, except for an arbitrary constant factor, be represented by the phase-integral function $q^{-1/2}(z)\exp[-iw(z)]$ on the part of the path Λ' lying in the lower half of the complex z plane as well as on the part of Λ' lying in the upper half of the complex z plane. In general the solution thus represented by the same phase-integral function on the whole path Λ' is not single-valued. If, however, the path Λ' encircles no singularities of $Q^2(z)$, except possibly for a pole at the origin of at the most the second order, and if we have the situation of a bound state, the corresponding solution of (26) is single-valued on the path Λ'. Since

$$w(z) = \int^z q(z)\,dz$$

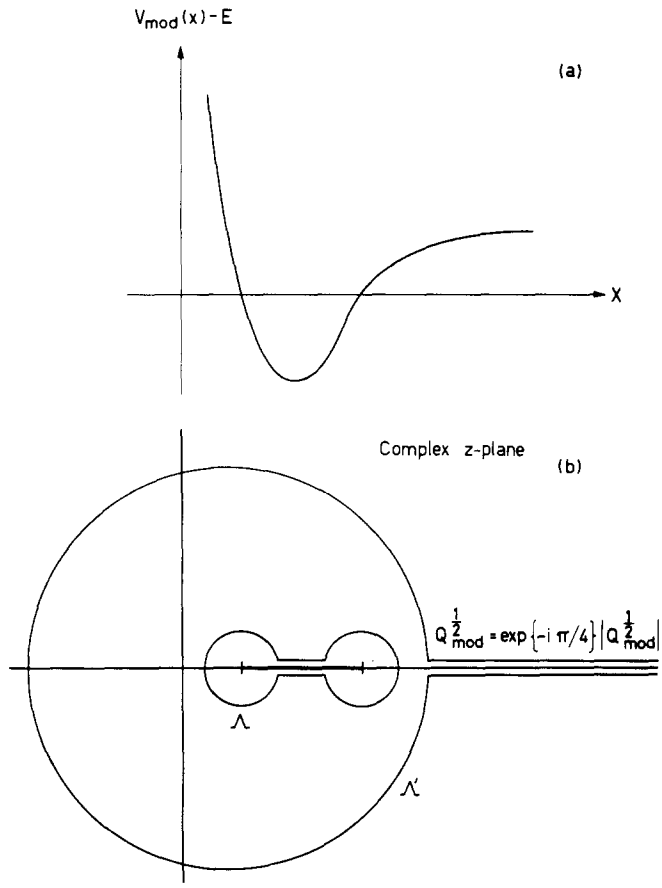


FIG. 2. (a) Qualitative behavior of $V_{\text{mod}}(x)-E$ for the case of a bound state in a hydrogenlike ion. (b) Contours of integration in the complex z plane used in the quantization conditions (27), (28), and (33). The heavy line indicates a cut. The phase of $Q^{1/2}_{\text{mod}}(x)$ is indicated for large positive values of x.

the single-valuedness of the bound-state solution represented by the previously mentioned phase-integral function on the whole path Λ' then implies that (cf. Ref. 12)

$$\tfrac{1}{2}\int_{\Lambda'} q(z)\,dz = \text{integer multiple of } \pi. \tag{27}$$

From the derivation of this quantization condition one immediately realizes that if one can choose the path Λ' such that the relative error of the phase-integral expression $q^{-1/2}(z)\exp[-iw(z)]$ for the bound-state wavefunction on Λ' can be made arbitrarily small, the quantization condition (27) is exact. Assuming now $q(z)$ to be single-valued at the origin [which excludes the case that $Q^2_{\text{mod}}(z)$ has a first-order pole at $z=0$], we can by means of residue calculus replace the original path of integration Λ' by another path Λ, as depicted in Fig. 2(b), which encircles the generalized classical turning points but not the origin. Thus one easily finds that the quantization condition (27) can be written

$$\frac{1}{2}\int_\Lambda q(z)\,dz = \left(\text{integer} - \operatorname*{Res}_{z=0}\frac{q(z)}{i}\right)\pi, \tag{28}$$

where the path of integration Λ is encircled in the negative sense. Let us now restrict the above treatment to a hydrogenlike ion for which

$$Q^2(z)=\frac{2m}{\hbar^2}\left(E+\frac{Ze^2}{z}\right)-\frac{l(l+1)}{z^2} \tag{29}$$

with well-known notations. We choose

$$Q^2_{\mathrm{mod}}(z)=\frac{2m}{\hbar^2}\left(E+\frac{Ze^2}{z}\right)-\frac{\alpha^2}{z^2}, \tag{30}$$

where α is a constant, which is assumed to be different from zero [in order that $q(z)$ be single-valued at the origin] but otherwise arbitrary. The conditions for the validity of (27) and (28) are then fulfilled. Since we can choose the contour Λ' arbitrarily large, so that the error of the phase-integral expression $q^{-1/2}(z)\times\exp[-iw(z)]$ for the bound-state wavefunction on Λ' can be made arbitrarily small, (27) and (28) are exact quantization conditions. The function $Q^2_{\mathrm{mod}}(z)$ has only two zeros (the generalized classical turning points) and one singularity (the second-order pole at the origin). Inserting (29) and (30) into the definition of ϵ_0 [cf. Eq. (11′) in Ref. 5], we get

$$(\epsilon_0)_{z=0}=((l+\tfrac{1}{2})^2-\alpha^2)/\alpha^2. \tag{31}$$

Using this formula and Eqs. (9a, b, c) in Ref. 2 and Eqs. (14′) and (16) in Ref. 5, and noting that close to the origin $Q_{\mathrm{mod}}(z)=i\alpha/z+\cdots$, if α is chosen to be positive, we get

$$\operatorname*{Res}_{z=0}\frac{q(z)}{i}=\begin{cases}\alpha\,, & \text{(first-order)} \quad (32a)\\ \alpha+\dfrac{1}{2\alpha}((l+\frac{1}{2})^2-\alpha^2)\,, & \text{(third-order)} \quad (32b)\\ \alpha+\dfrac{1}{2\alpha}((l+\frac{1}{2})^2-\alpha^2)-\dfrac{1}{8\alpha^3}((l+\frac{1}{2})^2-\alpha^2)^2. & \text{(fifth-order)} \quad (32c)\end{cases}$$

Inserting (32) into (28), and denoting the integer in (28) by $s+1+[\alpha]$, where s is a new integer, and $[\alpha]$ is the integer part of α, we get the *exact* quantization condition

$$\tfrac{1}{2}\textstyle\int_\Lambda q(z)\,dz=(s+\Delta)\pi, \tag{33}$$

where

$$\Delta=\begin{cases}[\alpha]+1-\alpha, & \text{(first-order)} \quad (34a)\\ [\alpha]+1-\alpha-\dfrac{1}{2\alpha}((l+\frac{1}{2})^2-\alpha^2), & \text{(third-order)} \quad (34b)\\ [\alpha]+1-\alpha-\dfrac{1}{2\alpha}((l+\frac{1}{2})^2-\alpha^2)+\dfrac{1}{8\alpha^3}((l+\frac{1}{2})^2-\alpha^2)^2. & \text{(fifth-order)} \quad (34c)\end{cases}$$

This result shows that the choice $\alpha=l+\frac{1}{2}$ gives the half-integer quantal Bohr–Sommerfeld quantization condition independently of the order of the phase-integral approximations used, while any other choice of α gives a more complicated quantization condition in which Δ depends on the order of approximation used.

We shall finally discuss the existence of energy values satisfying the exact quantization condition (7′) and their relation to the energy values determined from the corresponding approximate quantization condition (24). When the energy E increases continuously and monotonically from the value corresponding to the bottom of the potential well, the contour integral $\frac{1}{2}\int_\Lambda q(z)\,dz$ increases continuously and monotonically from approximately the value zero, as is easily realized from the fact that the behavior of this integral is essentially determined by its first and leading term, i.e., $\int_{x'}^{x''}|Q_{\mathrm{mod}}(x)|\,dx$. At the same time the expression $\arg[F_{11}(-\infty,z)/F_{11}(+\infty,z)]$ changes continuously with E, and when z is chosen conveniently it is seen that the absolute value of this expression does not exceed the smallest one of the two integrals $\mu(-\infty,+\infty)$ appearing in (15) and (18). If this smallest value of $\mu(-\infty,+\infty)$ is small compared to unity, it is therefore reasonable to assume that $\arg[F_{11}(-\infty,z)/F_{11}(+\infty,z)]$ depends only slightly on E. Since, furthermore, the quantity $\frac{1}{2}\int_\Lambda q(z)dz$ increases rather rapidly with E, we expect the quantity $\frac{1}{2}\int_\Lambda q(z)dz-\arg[F_{11}(-\infty,z)/F_{11}(+\infty,z)]$ to increase monotonically with E. Since this quantity is also a continuous function of E, we realize that for every nonnegative integer s (not exceeding a certain upper limit corresponding to the ionization limit) there is precisely one value of the energy E for which the exact quantization condition (7′) is fulfilled, and there is an adjacent value of the energy E for which the generalized quantal Bohr–Sommerfeld quantization condition, i.e., the approximate quantization condition (24) is fullfilled. (Close to the ionization limit there may occur exceptional cases.) Since (7′) is a necessary and sufficient quantization condition, we conclude that very close to each energy value which satisfies the generalized quantal Bohr–Sommerfeld quantization condition (24) there is actually a true energy eigenvalue. The energy eigenvalues can therefore be approximately determined by means of the approximate quantization condition (24). It should be emphasized, however, that this conclusion rests on the assumption that the appropriate μ integral is small compared to unity and only slightly energy-dependent. A similar discussion can obviously be carried through for the approximate quantization condition (25).

APPENDIX

In this appendix we collect those formulas from Ref. 1 which we refer to in the text. The numbering here of these formulas is the same as in Ref. 1. We are indebted to the referee for suggesting that we should add this appendix in order to make the paper more easily readable.

$$w(z)=\int^z q(\zeta)\,d\zeta, \tag{3.3}$$

$$M(w)=\tfrac{1}{2}i\epsilon\begin{pmatrix}1 & \exp\{-2iw\}\\ -\exp\{2iw\} & -1\end{pmatrix}, \tag{3.13}$$

$$\frac{\partial F(w,w_0)}{\partial w}=M(w)F(w,w_0), \tag{3.18}$$

$$\begin{pmatrix}F_{11}(w_0,w) & F_{12}(w_0,w)\\ F_{21}(w_0,w) & F_{22}(w_0,w)\end{pmatrix}=\begin{pmatrix}F_{22}(w,w_0) & -F_{12}(w,w_0)\\ -F_{21}(w,w_0) & F_{11}(w,w_0)\end{pmatrix}, \tag{3.20}$$

$$F(z_2,z_0)=F(z_2,z_1)F(z_1,z_0), \tag{3.27}$$

$$|F_{11}(w,w_0)-1|\le\frac{1}{2M}[\exp\{M\mu\}-1], \tag{4.3a}$$

$$\lim_{\substack{z\to Z\\ \text{along }\Lambda}}F_{11}(z,z_0)\ \text{exists and is finite}, \tag{4.5a}$$

$$\lim_{\substack{z\to Z \\ \text{along } \Lambda}} F_{12}(z,z_0) \text{ exists and is finite,} \tag{4.5b}$$

$$F_{12}(x_1,x_2)=\left(\frac{q(x_1)}{|q(x_1)|}\Big/\frac{q(x_2)}{|q(x_2)|}\right)\exp\{2\,\mathrm{Im}[w(x_2)] - 2i\,\mathrm{Re}[w(x_1)]\}F_{11}^{*}(x_1,x_2), \tag{5.9a}$$

$$|F_{11}(x_1,x_2)-1| \leq \frac{|F_{11}(x_1,z_0)-1|+|F_{11}(x_2,z_0)-1|+|F_{21}(x_2,z_0)|}{1-|F_{11}(x_2,z_0)-1|-|F_{21}(x_2,z_0)|} \leq \mu + \text{higher powers of } \mu, \text{ if } \mu \text{ is small compared to 1,} \tag{6.13a}$$

$$a_1(z)=F_{12}(z,+\infty)\,a_2(+\infty), \tag{10.8a}$$

$$a_2(z)=F_{22}(z,+\infty)\,a_2(+\infty), \tag{10.8b}$$

$$\psi(x)=a_2(+\infty)f_2(x)[1+0(\mu)] \text{ for } x''<x<+\infty, \text{ if } \mu=\mu(x,+\infty) \text{ is small compared to 1,} \tag{10.12}$$

$$F_{12}(-\infty,+\infty)=0, \tag{10.13}$$

$$F_{22}(-\infty,+\infty)=\frac{F_{22}(z,+\infty)}{F_{22}(z,-\infty)}, \tag{10.14a}$$

$$F_{22}(+\infty,-\infty)=\frac{F_{22}(z,-\infty)}{F_{22}(z,+\infty)}, \tag{10.14b}$$

$$F_{22}(x_1,x_2)\exp[i(\textstyle\int_{x'}^{x''}|q(x)|\,dx-\frac{1}{2}\pi)] \text{ is real,} \tag{10.17b}$$

$$\int_{x'}^{x''}|q(x)|\,dx=(s+\tfrac{1}{2})\,\pi-\arg\frac{F_{22}(z,+\infty)}{F_{22}(z,-\infty)}. \tag{10.20}$$

[1] N. Fröman and P.O. Fröman, *JWKR-Approximation, Contributions to the Theory* (North-Holland, Amsterdam, 1965), Russian translation: MIR, Moscow, 1967.
[2] N. Fröman, Ark. Fys. **32**, 541—8 (1966).
[3] N. Fröman, Ann. Phys. (N.Y.) **61**, 451—64 (1970).
[4] N. Fröman and P.O. Fröman, Ann. Phys. (N.Y.) **83**, 103—7 (1974).
[5] N. Fröman and P.O. Fröman, Nuovo Cimento B **20**, 121—32 (1974).
[6] S. Yngve, J. Math. Phys. **13**, 324—31 (1972).
[7] N. Fröman and P.O. Fröman, J. Math. Phys. **18**, 96—9 (1977).
[8] J.L. Dunham, Phys. Rev. **41**, 713—20 (1932).
[9] J.L. Dunham, Phys. Rev. **41**, 721—31 (1932).
[10] J.B. Krieger and C. Rosenzweig, Phys. Rev. **164**, 171—3 (1967).
[11] R.A. Howard, J. Chem. Phys. **54**, 4252—5 (1971).
[12] P.B. Bailey, J. Math. Phys. **5**, 1293—7 (1964).

The tensor virial theorem in quantum mechanics[a)]

Leon Cohen

Hunter College of the City University, New York, New York 10021
(Received 4 August 1977; revised manuscript received 31 January 1978)

A quantum mechanical generalization of the scalar virial theorem is derived and specialized to atoms and molecules in the Born–Oppenheimer approximation. The theorem is the quantum mechanical counterpart to Chandrasekhar's classical tensor virial theorem. The usual scalar virial equation follows by tensor contraction. One possible application is the introduction of more than one scale factor in a trial wavefunction. The scaling method proposed involves different stretchings for the different spatial coordinates. This is in contrast to the standard method of using the scalar virial theorem where the stretching is the same in all directions. An example is given where the introduction of multiple scale factors and the imposition of the tensor virial theorem yields a better result than the usual procedure of subjecting the wavefunction to a single scale transformation and imposing the scalar virial theorem.

I. INTRODUCTION

A classical mechanical generalization of the virial theorem has been given by Chandrasekhar and has been applied to the study of a number of problems.[1] Since the quantities that are related in Chandrasekhar's generalization are tensors the result is called the tensor virial theorem. Although the classical scalar virial theorem has a very well-known quantum mechanical form, it appears that no analogous quantum mechanical counterpart has been given for the classical tensor virial theorem. We shall do so below, and specialize the result to the case of atoms and molecules. We follow as closely as possible Chandrasekhar's notation.

Consider N mass points with Cartesian components $x_i^{(\alpha)}$. The Greek superscripts in parenthesis refer to the different particles and the Latin subscripts distinguish rectangular components, with analogous notation for the momenta $p_i^{(\alpha)}$.

The kinetic energy and potential energy tensors are defined, respectively, as

$$T_{ij}=\frac{1}{2}\sum_{\alpha}\left\langle\frac{P_i^{(\alpha)}P_j^{(\alpha)}}{m^{(\alpha)}}\right\rangle, \tag{1.1}$$

$$V_{ij}=\sum_{\alpha}\left\langle x_i^{(\alpha)}\frac{\partial V}{\partial x_j^{(\alpha)}}\right\rangle, \tag{1.2}$$

where V is the potential energy of the system.

The quantum mechanical tensor virial theorem for stationary states which we prove below is

$$2T_{ij}=V_{ij}. \tag{1.3}$$

II. DERIVATION

There are a number of ways of deriving Eq. (1.3), but perhaps the most straightforward is by using Heisenberg's equation of motion. The derivation that is often given for the scalar virial theorem[2] can be followed if one considers the operator $\sum_\alpha x_i^{(\alpha)}p_j^{(\alpha)}$ rather than the usual $\sum_{i\alpha}x_i^{(\alpha)}p_i^{(\alpha)}$ that is used in deriving the scalar virial theorem. For Hamiltonians of the form

$$H=\sum_{i\alpha}(p_i^{(\alpha)})^2+V,$$

we have for stationary states

$$\left\langle\left[H,\ \sum_\alpha x_i^{(\alpha)}p_j^{(\alpha)}\right]\right\rangle=0$$

$$=\left\langle\sum_\alpha\{[H,x_i^{(\alpha)}]p_j^{(\alpha)}+x_i^{(\alpha)}[H,p_j^{(\alpha)}]\}\right\rangle. \tag{2.1}$$

Since

$$[H,\ x_i^{(\alpha)}]=-\frac{i\hbar}{m^{(\alpha)}}p_i^{(\alpha)},$$

$$[H,\ p_j^{(\alpha)}]=i\hbar\frac{\partial V}{\partial x_j^{(\alpha)}}.$$

Equation (2.1) becomes

$$\sum_\alpha\left\langle-\frac{i\hbar}{m^{(\alpha)}}p_i^{(\alpha)}p_j^{(\alpha)}+i\hbar x_i^{(\alpha)}\frac{\partial V}{\partial x_j^{(\alpha)}}\right\rangle=0 \tag{2.2}$$

from which Eq. (1.3) follows directly. This is the most general form of the tensor virial theorem. The standard scalar virial theorem follows from (1.3) by taking the trace.

We now specialize the theorem to the case of atoms and molecules.

A. Atoms

For an N-electron atom the potential energy is

$$V=-Ze^2\sum_\gamma\frac{1}{|\mathbf{x}^{(\gamma)}|}+e^2\sum_{\gamma\beta}{}'\frac{1}{|\mathbf{x}^{(\gamma)}-\mathbf{x}^{(\beta)}|}. \tag{2.3}$$

Differentiating with respect to $x_j^{(\alpha)}$, we have

$$\frac{\partial V}{\partial x_j^{(\alpha)}}=Ze^2\frac{x_j^{(\alpha)}}{|\mathbf{x}^{(\alpha)}|^3}$$

$$-e^2\sum_\beta\frac{x_j^{(\alpha)}-x_j^{(\beta)}}{|\mathbf{x}^{(\alpha)}-\mathbf{x}^{(\beta)}|^3} \tag{2.4}$$

and that

$$V_{ij}=Ze^2\sum_\alpha\left\langle\frac{x_i^{(\alpha)}x_j^{(\alpha)}}{|\mathbf{x}^{(\alpha)}|^3}\right\rangle$$

[a)]Supported by a grant from the City University Faculty Research Award Program.

 0022-2488/78/1909-1838$1.00

$$-e^2\sum_{\alpha\beta}{}'\left\langle\frac{x_i^{(\alpha)}(x_j^{(\alpha)}-x_j^{(\beta)})}{|\mathbf{x}^{(\alpha)}-\mathbf{x}^{(\beta)}|}\right\rangle$$

$$=Ze^2\sum_{\alpha}\left\langle\frac{x_i^{(\alpha)}x_j^{(\alpha)}}{|\mathbf{x}^{(\alpha)}|^3}\right\rangle$$

$$-\tfrac{1}{2}e^2\sum_{\alpha\beta}{}'\left\langle\frac{(x_i^{(\alpha)}-x_i^{(\beta)})(x_j^{(\alpha)}-x_j^{(\beta)})}{|\mathbf{x}^{(\alpha)}-\mathbf{x}^{(\beta)}|^3}\right\rangle . \tag{2.5}$$

B. Molecules

The appropriate form of the scalar virial theorem for molecules in the Born-Oppenheimer approximation was first considered by Slater.[3] He showed that, due to the fact that nuclei are held fixed, an additional term must be added to the standard mode of presenting the theorem. Slater's result has been extensively used in the study of molecular structure and binding.

We now derive the analagous term for the tensor virial theorem. Taking the electronic potential energy to be

$$V_e=-\sum_{K\alpha}\frac{Z_Ke^2}{|\mathbf{x}^{(\alpha)}-\mathbf{x}^{(K)}|}$$

$$+\sum_{\alpha\beta}{}'\frac{e^2}{|\mathbf{x}^{(\alpha)}-\mathbf{x}^{(\beta)}|}, \tag{2.6}$$

where K denotes the nuclear coordinates, we have

$$V_{ij}=\sum_{K\alpha}\left\langle\frac{Z_Ke^2x_i^{(\alpha)}(x_j^{(\alpha)}-x_j^{(K)})}{|\mathbf{x}_j^{(\alpha)}-\mathbf{x}_j^{(K)}|^3}\right\rangle$$

$$-\tfrac{1}{2}e^2\sum_{\alpha\beta}{}'\left\langle\frac{(x_i^{(\alpha)}-x_i^{(\beta)})(x_j^{(\alpha)}-x_j^{(\beta)})}{|\mathbf{x}^{(\alpha)}-\mathbf{x}^{(\beta)}|^3}\right\rangle$$

$$=\sum_{K\alpha}\left\langle\frac{Z_Ke^2(x_i^{(\alpha)}-x_i^{(K)})(x_j^{(\alpha)}-x_j^{(K)})}{|\mathbf{x}^{(\alpha)}-\mathbf{x}^{(K)}|^3}\right\rangle$$

$$-\tfrac{1}{2}e^2\sum_{\alpha\beta}{}'\left\langle\frac{(x_i^{(\alpha)}-x_i^{(\beta)})(x_j^{(\alpha)}-x_j^{(\beta)})}{|\mathbf{x}^{(\alpha)}-\mathbf{x}^{(\beta)}|^3}\right\rangle$$

$$+\sum_{K\alpha}\left\langle\frac{Z_Ke^2x_i^{(K)}(x_j^{(\alpha)}-x_j^{(K)})}{|\mathbf{x}^{(\alpha)}-\mathbf{x}^{(K)}|^3}\right\rangle . \tag{2.7}$$

The last term can be expressed as

$$-\sum_K\left\langle x_i^{(K)}\frac{\partial V_e}{\partial x_j^{(K)}}\right\rangle=-\sum_K x_i^{(K)}\frac{\partial E}{\partial x_j^{(K)}}, \tag{2.8}$$

where E is the total electronic energy. This follows from the Hellman—Feynman theorem. At the equilibrium configuration this term is zero.

Hence, for molecules in the Born—Oppenheimer approximation

$$V_{ij}=\sum_{K\alpha}\left\langle Z_Ke^2\frac{(x_i^{(\alpha)}-x_i^{(K)})(x_j^{(\alpha)}-x_j^{(K)})}{|\mathbf{x}^{(\alpha)}-\mathbf{x}^{(K)}|^3}\right\rangle$$

$$-\tfrac{1}{2}e^2\sum_{\alpha\beta}\left\langle\frac{(x_i^{(\alpha)}-x_i^{(\beta)})(x_j^{(\alpha)}-x_j^{(\beta)})}{|\mathbf{x}^{(\alpha)}-\mathbf{x}^{(\beta)}|^3}\right\rangle$$

$$-\sum_K x_i^{(K)}\frac{\partial E}{\partial x_j^{(K)}}. \tag{2.9}$$

III. CONCLUSION

In the classical case the scalar and tensor virial theorems can be thought of as the first two moment equations. Chandrasekhar has also considered moment equations of higher order. In general they involve moments that mix position and velocity. The set of all moment equations is equivalent to the original equation from which moments are taken. Of course, the advantage of considering low order moment equations is that very often the solution of a moment equation is a good approximation to the exact equation being considered.

In the quantum mechanical case an interesting problem arises if one considers generalizing the classical higher order moment equations. As mentioned above higher order moments mix position and velocity, but there is no unique way of writing the quantum mechanical counterpart to operators which mix position and momentum. Indeed it has been shown that there are an infinite numer of so-called correspondence rules which allow the writing of quantum mechanical Hermitian operators from their classical counterparts.[4] It may be that higher order moment equations generated via any correspondence rule would be equivalent, but that is not clear.

One possible use of the tensor virial theorem derived above is to use it as a means of ascertaining the accuracy of approximate wavefunctions. The possibility of using the scalar virial theorem for that purpose was pointed out by Slater,[3] but Löwdin[5] has shown that any approximate molecular wavefunction can always be made by to satisfy the scalar virial theorem by simple scaling of the coordinates. This may also be true of the tensor virial theorem but the scaling in that case would clearly not be so simple. One possible measure of the accuracy of a wavefunction that could be used is

$$\mathrm{Tr}\{2T_{ij}-V_{ij}\}^2. \tag{3.1}$$

It would be of interest to take a particular molecule and calculate the quantity given by expression (3.1) for a variety of molecular wave functions to determine whether indeed there is a strong correlation with the accuracy of the wavefunction. H_2 would be a good candidate as there are a great number of approximate wave functions for it.

In certain cases it may be possible to use the tensor virial theorem to introduce a set of different scale factors to improve a given approximate wave function. The scalar virial theorem allows the introduction of only *one* scale factor. (Optimizing the energy with respect to the scale factor is equivalent to satisfying the scalar virial theorem and we suspect the same may be true for the tensor virial theorem.)

One may consider the standard scaling method[5] as transforming the coordinates via

$$\mathbf{r}\to\eta\mathbf{I}\mathbf{r}, \tag{3.2}$$

where η is the constant scale factor and I the unit matrix. This is equivalent to stretching each coordinate by the *same* relative amount. But the most general transformation compatable with the tensor virial theorem is

$$\mathbf{r} \to \mathbf{A}\mathbf{r}\,, \tag{3.3}$$

where **A** is a symmetric matrix. The diagonal terms of **A** are in general not equal and hence the stretching given by Eq. (3.3) can be different for each Cartesian coordinate. Furthermore, if the off diagonal terms are not zero the transformation will involve stretchings and rotations.

We give an example where the usual procedure of introducing a single scale factor and imposing the scalar virial theorem does not yield as good a result as the introduction of multiple scale factors and the imposition of the tensor virial theorem. Consider the two-dimensional harmonic oscillator with frequencies w_x and w_y. Suppose we take a wavefunction of the form

$$\psi(x,y) = \frac{1}{\pi^{1/2}} e^{-1/2(x^2+y^2)} \tag{3.4}$$

and introduce a single scale factor as in the standard procedure[5]

$$x \to \eta x \quad y \to \eta y. \tag{3.5}$$

Imposing the scalar virial theorem yields

$$\eta = \left[\frac{m^2}{\hbar^2}\,\frac{w_x^2 + w_y^2}{2}\right]^{1/4} \tag{3.6}$$

with a corresponding energy

$$E(\eta) = \hbar\left(\frac{w_x^2 + w_y^2}{2}\right)^{1/2}. \tag{3.7}$$

This is the best that can be had with one scale factor.

On the other hand if we introduce *two* scale factors

$$x \to \eta_x x \quad y \to \eta_y y. \tag{3.8}$$

A straightforward calculation shows that the imposition of the tensor virial theorem yields

$$\eta_x = \left(\frac{m w_x}{\hbar}\right)^{1/2}, \quad \eta_y = \left(\frac{m w_y}{\hbar}\right)^{1/2} \tag{3.9}$$

with a corresponding energy

$$E(\eta_x, \eta_y) = \tfrac{1}{2}\hbar w_x + \tfrac{1}{2}\hbar w_y, \tag{3.10}$$

which is the correct energy.

[1] S. Chandrasekhar, *Ellipsoidal Figures of Equilibrium* (Yale University Press, New Haven, 1969); and references therein.
[2] F. Merzbacher, *Quantum Mechanics* (Wiley, New York, 1961), pp. 162–3.
[3] J.C. Slater, J. Chem. Phys. **1**, 687 (1933).
[4] L. Cohen, J. Math. Phys. **7**, 781, (1966); **17**, 597 (1970).
[5] P.O. Löwdin, J. Molec. Spectr. **3**, 46 (1959).

Quantum theory and Hilbert space

A. R. Marlow

Loyola University, New Orleans, Louisiana 70118
(Received 11 October 1977)

Two theorems are proven which show that any orthomodular partially ordered set (hence, in particular, the orthologic of questions on a physical system) can be embedded in the lattice of closed subspaces of a Hilbert space in such a way that the standard trace formula of quantum theory can be used to calculate all probabilities. Possible conclusions from these results and from the existence of counterexamples to stronger conjectures are then discussed.

1. INTRODUCTION

Practically since the beginning of quantum theory, the question of the extent to which the Hilbert space language of the theory could be considered an adequate mathematical language for the study of arbitrary physical systems has been widely discussed. The search for counterexamples to any general claims has been conducted with as much diligence as, and somewhat more success than, the search for proofs of general validity of the language. The mathematical context of the problem can be described as the theory of Hilbert extensions of orthomodular partially ordered sets (orthoposets) and the probability measures (states) that can exist on such sets. General background information and any definitions not explicitly given in this paper can be found in our first few references.[1-3]

There seems to be general agreement that the language of orthoposets[1,5] and their states is capable of dealing with arbitrary physical situations, whether of classical, modern or some as yet undiscovered type. This agreement follows rather directly from the reasonable requirement of Einstein and Heisenberg that a theory, at least initially, stay as close as possible to the actual measurements made in laboratories. As currently applied, the requirement implies the structure of an orthoposet for the simple yes—no measurements in terms of which all observables can be expressed, with the ordering, orthocomplementation and orthogonality relations determined by physically measured probabilities. The area of disagreement involves the thorny problems of the possibility, necessity and/or usefulness of realizing the physical orthoposets and states in terms of orthoposets isomorphic to the lattices of closed subspaces (or, equivalently, projection operators) of Hilbert space. It is in this context that examples abound but counterexamples more abound.

Invariably in work on these problems, Boolean extensions of orthoposets—the main rivals of Hilbert extensions—enter in, if only as tools for constructing examples and counterexamples. These are realizations of orthoposets in terms of Boolean algebras (orthoposets isomorphic to lattices of subsets of a set). Zierler and Schlessinger[3] have studied what we call Boolean order preserving extensions (O-extensions), and by constructing for each orthoposet Q a unique minimal Boolean O-extension, have demonstrated the existence of such extensions for arbitrary orthoposets. We will refer to their construction as the ZS-extension of Q. Using this construction, they show that for a wide class of orthoposets (including those isomorphic to the set of projection operators on a Hilbert space) there cannot exist any Boolean O-extensions that also preserve all probability measures.

The situation for Hilbert order preserving extensions is no better. Greechie and Miller[4] have shown that the class of all orthoposets can be divided into three mutually exclusive subclasses: Those having no states, exactly one state, or an infinity of states, and they have constructed simple examples of each type. Then Wright[5] has pointed out that for a class of examples in the last category there exist states that cannot carry over to any Hilbert O-extension.

Since, for physical purposes, preservation of probability measures seems at least of equal importance with preservation of ordering, the present author was led to ask if there might exist measure preserving extensions (M-extensions) for arbitrary orthoposets, which do not completely preserve the ordering. One of the purposes of this paper is to give an affirmative answer to this question by constructing Boolean M-extensions for every orthoposet. Using these, it is then possible to construct Hilbert M-extensions, and finally, what we call QM-extensions, in which computation of probability can be carried out according to the standard quantum mechanical trace formula.

2. PRELIMINARY NOTIONS

After a brief review of the notion of an orthoposet to establish notation, we will give only those definitions not current in the literature.

D1. For Q a partially ordered set (poset) and $S \subset Q$, we write $\wedge S$ and $\vee S$ for the greatest lower and least upper bound elements for S when they exist (also referred to as the *meet* and *join* of S), with the standard alternative notations $p \wedge q$, $\vee_i q_i$, etc., for pairs and sequences. By an *orthocomplemented poset* we mean a poset Q with a greatest and least element, 1 and 0, and an operation $q \to q'$ (orthocomplementation) such that:

(1) $q'' = q$;

(2) $q \leq p \Rightarrow p' \leq q'$;

(3) $q \wedge q' = 0$; $q \vee q' = 1$.

Two elements p, q in an orthocomplemented poset will be said to be *orthogonal* ($p \perp q$) if $p \leq q'$, which, by (1) and (2), is also equivalent to $q \leq p'$. Then an *orthomodular poset* (or simply an *orthoposet*) can de defined

 0022-2488/78/1909-1841$1.00

as an orthocomplemented poset Q which satisfies:

(4) $\forall\ p,q\in Q,\ p\perp q\Rightarrow p\vee q\in Q$;

(5) $\forall\ p,q\in Q, p\leq q\Rightarrow q=p\vee(p\vee q')'$.

An orthoposet will be called an *ortholattice* if condition (4) is generalized to

(4′) $\forall\ p,q\in Q, p\vee q\in Q$. □

Since it is easily shown that if $p\vee q$ exists then $p'\wedge q'$ exists and is equal to $(p\vee q)'$, condition (5) is equivalent to

$$\forall\ p,q\in Q,\ p\leq q\Rightarrow q=p\vee(p'\wedge q),$$

and, similarly, (4) and (4′) can be replaced by equivalent conditions in terms of meets.

In the terminology of D1 we can specify the two main types of orthoposets we will need:

D2. A *Boolean algebra* is an ortholattice isomorphic to an ortholattice of subsets of some set, with the usual set theoretical inclusion, intersection, union and complementation providing the orthoposet ordering and operations. A *Hilbert ortholattice* is an ortholattice isomorphic to the ortholattice of all closed subspaces of a Hilbert space, with subspace inclusion, intersection, union and orthocomplementation providing the orthoposet ordering and operations. □

Of course it is well known that Boolean algebras can be given an intrinsic algebraic definition, but for purposes of physical insight D2 is useful: it allows the characterization of standard quantum mechanics as simply probability theory carried out with events as subsets of a set replaced by events as subspaces of a Hilbert space. Gleason's theorem[6] then specifies probabilities in terms of one-dimensional subspaces (hence, unit vectors) and their convex combinations.

D3. For Q an orthoposet, we write Φ_Q for the set of *probability measures* or *states* on Q, i.e., the set of functions $\varphi\colon Q\to[0,1]$ such that

(1) $\varphi(1)=1$;

(2) $\forall\ q,p\in Q, q\perp p\Rightarrow\varphi(q\vee p)=\varphi(q)+\varphi(p)$.

A state φ will be said to be *mixed* if it can be written as a finite or countably infinite convex combination of states $\varphi_i\neq\varphi$, i.e., if $\varphi=\sum_i\lambda_i\varphi_i$ where $\sum_i\lambda_i=1, 0\leq\lambda_i\leq 1$. Otherwise φ will be said to be *pure*. The set Φ_Q is said to be *full* if

$$\forall\ q,p\in Q,\ [\varphi(q)\leq\varphi(p),\ \forall\varphi\in\Phi_Q]\Rightarrow q\leq p.$$

We now define several terms not current in the literature.

D4. For S a subset of an orthoposet Q, we will write $\overline{S}$ for the set of all finite S-values sequences, and $q^n=(q_1^n,q_2^n,\ldots,q_n^n)$, $p^n=(p_1^n,p_2^n,\ldots,p_n^n),\cdots$, for particular sequences of length n. The symbol $\binom{n}{m}$ will mean the set of all selections of m distinct integers from the set $\{1,2,\ldots,n\}$. We will say that S is:

(1) *unitary* if $1\in S$;

(2) $\perp$-*exclusive* if, $\forall\ p\in S,\ q\in Q,\ q\perp p\Rightarrow q\notin S$;

(3) a *base* for Q if, $\forall\ p\in Q,\ \{p,p'\}\cap S\neq\emptyset$;

(4) *an M-base* (measure-base) if S is a unitary $\perp$-exclusive base for Q.

Moreover, if Q is also an ortholattice, we will say that S is

(5) $\wedge$-*perfect if*, $\forall\ q^m,p^n\in\overline{S},\ m\leq n$,

$$[\wedge_{i=1}^m q_i^m\leq\vee_{\alpha\in\binom{n}{m}}(\wedge_{j\in\alpha}p_j^n)]\Rightarrow$$
$$[\exists\ \alpha\in\tbinom{n}{m}\colon\forall\ j\in\alpha\,(\exists\, i\leq m\colon q_i^m\leq p_j^n)].\ \square$$

M-bases can be described perhaps more intuitively as subsets which exclude orthogonality but contain exactly one element of every pair of orthocomplements. We have now as our first lemma:

L1. There exist M-bases for every orthoposet. □

Proof: By Zorn's lemma there exist maximal $\perp$-exclusive subsets of any orthoposet Q, since it is easily seen that the union of any chain of $\perp$-exclusive subsets is itself $\perp$-exclusive. Assume S is such a maximal subset, and, for some $p\in Q$, assume $\{p,p'\}\cap S=\emptyset$. Then by maximality there exist $q,r\in S$ such that $q\perp p$ and $r\perp p'$ which implies $q\perp r$, contradicting $\perp$-exclusiveness for S. Hence, every maximal $\perp$-exclusive subset of Q is an M-base. Q.E.D.

3. EXTENSIONS OF ORTHOPOSETS

D5. Given an orthoposet Q, we say that a pair $\{L,\lambda\}$ is an *extension* of Q provided that L is an orthoposet and $\lambda\colon Q\to L$ is a mapping with the properties:

(1) $\lambda(1)=1\in L$;

(2) $\lambda(q')=\lambda(q)'\in L,\ \forall\ q\in Q$;

(3) for some base $S\subset Q$,

$$p\leq q\iff\lambda(p)\leq\lambda(q),\ \forall\ p,q\in S.$$

We call L the *range* of the extension, λ the *embedding* of the extension, and S a *base* for the extension. If Q itself is a base for an extension $\{L,\lambda\}$, we say that $\{L,\lambda\}$ is an O-*extension* (*order preserving extension*) of Q. An extension $\{L,\lambda\}$ will be called an M-*extension* (*measure preserving extension*) of Q if:

(4) $\forall\ \varphi\in\Phi_Q,\ \exists\ \hat{\varphi}\in\Phi_L\colon\hat{\varphi}(\lambda(q))=\varphi(q),\forall\ q\in Q$.

An O-extension satisfying (4) will be called an MO-*extension*. If the range of an extension is a Boolean algebra we say that the extension is a *Boolean extension*; if the range is a Hilbert ortholattice we call the extension a *Hilbert extension*, and we say that the Hilbert space underlying the ortholattice is the Hilbert space of the extension. We will generally use the notation $\{A,\alpha\}$ for a Boolean extension and $\{P,\pi\}$ for a Hilbert extension, where P is understood to be the ortholattice of projection operators on the underlying Hilbert space. A Hilbert extension $\{P,\pi\}$ of an orthoposet Q will be called a Boolean Hilbert extension if $\pi(Q)$ is a subset of a Boolean algebra of projection operators in P. □

The orthoposets Q of physical interest are, of course, those with nonempty (and usually even full) sets of states. It will be useful to have some terminology specifically constructed for the physically relevant situation—loosely, the QM or quantum mechanical situation, although all our definitions below will also

apply (sometimes trivially) to orthoposets with no states. Up to the present time physicists have generally required that the extensions used be O-extensions as well as M-extensions, and even, excluding superselection rules, that the embeddings be orthoposet isomorphisms. In light of the fact that some of the previously mentioned counterexamples of Greechie *et al.*[7] have full sets of states but cannot have any Hilbert MO-extensions, we believe these requirements are too strong, and that the following definitions spell out the minimum physically useful requirements that can be shown to hold in all cases.

D6. A pair $\{E,\eta\}$ will be said to be a QM-extension of an orthoposet Q if $E=P\cup T$, where P is the orthoposet of projection operators and T the set of nonnegative trace 1 self-adjoint operators on a Hilbert space, and $\eta: Q\cup\Phi_Q\to E$ is a mapping such that:

(1) $\{P,\pi\}$, $\pi=\eta\,|\,Q$, is a Hilbert M-extension of Q;

(2) $\forall\ \varphi\in\Phi_Q$, $q\in Q$, $\eta(\varphi)\in T$ and

$$\varphi(q)=\mathrm{trace}(\eta(q)\eta(\varphi)).$$

A QM-extension will be said to be:

Boolean if

(3) (P,π) is Boolean;

Pure if

(4) $\eta(\Phi_Q)\subset P\cap T$;

Convex if

(5) for all mixed states $\varphi=\sum_i\lambda_i\varphi_i\in\Phi_Q$, $\eta(\varphi)=\sum_i\lambda_i\eta(\varphi_i)$. □

Now Gleason's theorem[6] and its extension to arbitrary Hilbert spaces under the assumption of the continuum hypothesis,[8] guarantee that the set of states on the projection operators of any Hilbert space of dimension greater than 2 is isomorphic to the set T of nonnegative self-adjoint trace 1 operators in such a way that the computation of probability can proceed according to the standard quantum theoretical formula used in D6. This obviously implies the following result which we formalize as a lemma:

L2. For any Hilbert M-extension $\{P,\pi\}$ of an orthoposet Q such that the Hilbert space of the extension is of dimension $\geqslant 2$ (and assuming the continuum hypothesis if H is nonseparable), there exist QM-extensions $\{E,\eta\}$ of Q such that $E=P\cup T$ and $\pi=\eta\,|\,Q$. □

Our results will not depend on this lemma, however, since we will explicitly construct the operators representing states in our QM-extensions.

4. BOOLEAN *U*-EXTENSIONS

Starting with an arbitrary unitary base S of an orthoposet Q, we introduce a method (unrelated subsets) for constructing a Boolean extension $\{A(S),\alpha\}$ of Q such that S is a base for the extension. The method is very flexible, and by starting with different bases and dividing out suitably chosen ideals, a variety of extensions can be produced, including the ZS-extension if the initial base is chosen to be Q itself. In this latter case it is easy to check that the algebra $A(Q)$ is an O-extension of Q, and the results of Zierler and Schlessinger[3] show then that there is a homomorphism onto the ZS-extension. The precise advantage of the unrelated subset method, however, is that it also allows the construction and study of extensions which have useful properties not always shared by O-extensions. In particular, we will limit our constructions to those resulting from the initial choice of an M-base, since in this case we can show that an M-extension is produced.

D7. A subset S of an orthoposet Q will be said to be *unrelated* if, for every pair $\{p,q\}\subset S$, neither $p\leqslant q$ nor $q\leqslant p$ is true. (From this definition it is evident that the empty set $\emptyset$ and any singlet subset $\{q\}$ are examples of unrelated subsets). For $S\subset Q$, let $\mathcal{P}(S)$ be the power set of S and $\mathcal{P}_{\shortparallel}(S)$ the set of nonempty unrelated subsets of S, and for $N\subset S$ define:

$$N^{\shortparallel}=\{q\in N:\ \forall\ p\in N,\ p\leqslant q\Rightarrow p=q\}$$

For $M,N\in\mathcal{P}_{\shortparallel}(S)$, let $M*N=(M\cup N)^{\shortparallel}$. □

It is clear from this definition that $N^{\shortparallel}$ is always an unrelated set, and that $N^{\shortparallel}=\emptyset$ only if N contains an infinite chain with no lower bound in N. We have the lemma:

L3. For S a unitary subset of an orthoposet Q, the pair $\mathcal{P}^*_{\shortparallel}(S)=\{\mathcal{P}_{\shortparallel}(S),*\}$ is an Abelian, idempotent semigroup with unit $\{1\}$, where $*$ is the binary operation defined in D7. □

Proof: From D7 and the remarks immediately above, $\mathcal{P}_{\shortparallel}(S)$ is obviously closed under the operation $*$, since, for $M,N\in\mathcal{P}_{\shortparallel}(S)$, by unrelatedness the only possible chains in $M\cup N$ are of finite length 2. The checking of associativity—$[(M\cup N)^{\shortparallel}\cup K]^{\shortparallel}=[M\cup N\cup K]^{\shortparallel}$—is trivial, involving only the transitivity of the ordering $\leqslant$ on Q and the fact that the maximum length of chains in $M\cup N\cup K$ is 3. Hence, $\mathcal{P}^*_{\shortparallel}(S)$ is a semigroup. The abelian property for $*$ is obvious, and then the fact that $\{1\}$ is the unit element follows from the fact that 1 is an upper bound for every element of Q. Finally, idempotency follows from the easily checked fact that, $\forall\ N\in\mathcal{P}_{\shortparallel}(S)$, $N^{\shortparallel}=N$, and the lemma is established. Q.E.D.

D8. For S a unitary base of an orthoposet Q and $\mathcal{P}^*_{\shortparallel}(S)$ the semigroup given by L3, let $U^*(S)$ be the semigroup ring defined by $\mathcal{P}^*_{\shortparallel}(S)$ over the field $2=\{0,1\}$, i.e., $U^*(S)$ is the free vector space of all 2-valued functions on $\mathcal{P}^*_{\shortparallel}(S)$ having only a finite number of nonzero values, with the product $*$ extended by linearity.[9] □

L4. For S a unitary base of an orthoposet Q, the ring $U^*(S)$ defined in D8 is Boolean, and so uniquely defines a corresponding Boolean algebra $U(S)$. □

Proof: We need only check that every element of $U^*(S)$ is idempotent. But from D8 we have $P\in U^*(S)$ $\Rightarrow P=\sum_{i=1}^n N_i$, $N_i\in\mathcal{P}^*_{\shortparallel}(S)$. Hence $P*P=\sum_{i,j=1}^n N_i*N_j$ $=\sum_{i=1}^n N_i=P$, since $N_i*N_j+N_j*N_i=0$, where 0 of course, is the unique function in $U^*(S)$ having all values zero. Therefore, $U^*(S)$ is a Boolean ring, and we define the corresponding Boolean algebra $U(S)$ in the

standard way[10]: for $P, Q, \in U^*(S)$ we set

$$P \wedge Q = P * Q,$$

$$P \vee Q = P + Q + P * Q,$$

$$P' = \{1\} + P. \qquad \text{Q.E.D.}$$

It is at this point that we could, if we wished, produce a Boolean extension $\{A(S), \alpha\}$ of Q corresponding to an arbitrary unitary base $S \subset Q$ by letting $A(S) = U(S)/U_0(S)$, where $U_0(S)$ is the ideal generated in $U(S)$ by elements of the form $\{1\} + \{q\} + \{q'\} = (\{q\} \wedge \{q'\}') \vee (\{q\}' \wedge \{q'\}) \in U(S)$ for all $q \in Q$ such that $\{q, q'\} \subset S$, and taking for $\alpha(q)$ the equivalence class of $\{1\} + \{q'\}$ if $q \notin S$ (since then $q' \in S$, from the definition of a base). It is straightforward to show that $\{A(S), \alpha\}$ is an extension of Q with base S (the Boolean U-extension of Q with respect to S), and of course if $S = Q$, the extension will be an O-extension, but it is more direct for the purposes of this paper to restrict consideration to M-bases, in which case no quotient algebra is necessary at all. We have

L5. For S an M-base of an orthoposet Q, the pair $\{A(S), \alpha\}$ is a Boolean extension of Q if $A(S)$ is the Boolean algebra $U(S)$ generated by the nonempty unrelated subsets of S (as in L4) and α is the mapping defined by

$$\alpha(q) = \begin{cases} \{q\}, & q \in S \\ \{1\} + \{q'\}, & q \in Q - S. \end{cases} \quad \square$$

Proof: Conditions (1) and (2) of D5 are obviously satisfied from the definition of α. (Recall that an M-base is $\perp$-exclusive, so that $q \in S \Leftrightarrow q' \in Q - S$.) Then, from the definition of the product $*$ in D7, we have $q \leq p \Leftrightarrow \{q\} * \{p\} = \{q\}$, and so D5, (3) is also satisfied. Q.E.D.

D9. For S an M-base of an orthoposet Q, we will call the extension $\{A(S), \alpha\}$ given in L5 the Boolean U-extension of Q with respect to S. $\square$

L6. Under the conditions of L5, the image set $\alpha(S) \subset A(S)$ is $\wedge$-perfect. $\square$

Proof: We first note that a sequence $a^m \in \overline{\alpha(S)}$ (cf. D4) has singlet subsets $a_i^m = \{p_i\} \subset S$ as range elements, and then that $\alpha(S)$ is $\perp$-exclusive, since $\alpha(p) \perp \alpha(q)$, $p, q, \in S$, implies $\{p\} * \{q\} = 0 \in A(S)$, which is impossible. We have, then, $\wedge_{i=1}^m a_i^m = \wedge_{i=1}^m \{p_i\} = N \neq 0 \in A(S)$, where $N \in P_{\shortparallel}(S)$ is an unrelated subset consisting of some selection of the elements $p_i \in S$. Similarly, any join of the meets of sequences $b^n \in \overline{\alpha(S)}$ can be expressed as a sum $\sum_k N_k \in P_{\shortparallel}^*(S)$ of nonempty unrelated subsets N_k consisting of selections of elements $q_j \in S$ such that $\{q_j\} = b_j^n$. Then, from the unrelated subset construction, it follows easily that $N * (\sum_k N_k) = N$ only if $N * N_k = N$ for some k, and this last condition immediately implies D4,5), the definition of $\wedge$-perfect. Q.E.D.

5. EXISTENCE OF M-EXTENSIONS

Let A be a Boolean algebra and $S \subset A$ a unitary subset. Horn and Tarski[11] construct a relation $\dashv$ on the set $\overline{S}$ of finite S-valued sequences, and, in terms of this relation, define a *partial measure* on S as a nonnegative real-valued function φ such that:

(PM1) $\varphi(1) = 1$;

(PM2) $\forall\, a^m,\ b^n \in \overline{S},\ a^m \dashv b^n \Rightarrow \sum_{i=1}^m \varphi(a_i^m) \leq \sum_{j=1}^n \varphi(b_j^n)$.

They then prove their theorem[12]:

Theorem 1.22: If φ is a partial measure on a unitary subset S of a Boolean algebra A, there is a measure ψ on A which agrees with φ on S. If, moreover, φ assumes only the values 0 and 1, the same may be postulated for ψ.

We refer to the original paper for the details of the construction and proof; for our purposes we need only the theorem and the following specific properties of the relation $\dashv$:

$(\dashv 1)$ $a^m \dashv b^n \Rightarrow [a_i^m \leq \vee_{j=1}^n b_j^n],\ \forall\, i \leq m$;

$(\dashv 2)$ $a^m \dashv b^n \Rightarrow [\wedge_{i=1}^m a_i^n \leq \vee_{\alpha \in \binom{n}{m}} (\wedge_{j \in \alpha} b_j^n)]$.

We have now the lemma:

L7. Let $\varphi: S \to [0,1]$ be a function on a unitary $\wedge$-perfect subset S of a Boolean algebra A, such that

(1) $\varphi(1) = 1$;

(2) $\forall\, a, b \in S,\ a \leq b \Rightarrow \varphi(a) \leq \varphi(b)$.

Then there exists a measure ψ on A which agrees with φ on S. $\square$

Proof: By Theorem 1.22 above (Horn and Tarski), the lemma will be established if φ is a partial measure on S, and so we need to establish the (PM2) holds for φ. From $(\dashv 1)$ and the $\wedge$-perfect property applied to the elements $a_i^m \in S$, regarded as single element sequences, we have

(a) $a^m \dashv b^n \Rightarrow \forall\, a_i^m,\ \exists\, b_j^n \colon a_i^m \leq b_j^n$,

and from $(\dashv 2)$ and the $\wedge$-perfect property we get

(b) $a^m \dashv b^n \Rightarrow [\exists \alpha \in \binom{n}{m} \colon \forall\, j \in \alpha\, (\exists i \leq m : a_i^m \leq b_j^n)]$.

Then, by simple counting, (a) and (b) together with condition (2) of the lemma imply (PM2). Q.E.D.

Using L6 and L7 we can prove our basic theorem, establishing the existence of M-extensions for arbitrary orthoposets.

T1. The U-extension of an orthoposet Q with respect to an M-base $S \subset Q$ is a Boolean M-extension of Q. $\square$

Proof: The function $\bar{\varphi}: \alpha(S) \to [0,1]$ defined by $\bar{\varphi}(\alpha(q)) = \varphi(q)$, $\forall\, q \in S$, $\varphi \in \Phi_Q$, where α is the embedding of the U-extension of Q with respect to an M-base S, obviously satisfies the conditions of L7. The theorem then follows from L6. Q.E.D.

6. QUANTUM THEORY IN HILBERT SPACE FOR ARBITRARY ORTHOPOSETS

T2. There exist pure QM-extensions of every orthoposet. $\square$

Proof: By T1, there exist Boolean M-extensions $\{A, \alpha\}$ for every orthoposet Q, and, by the Stone representation theorem, we can always choose A to be the Boolean algebra of all clopen subsets of a totally dis-

connected compact Hausdorff space X.[13] For $q \in Q$, we define $\xi_q : X \to \{0,1\}$ to be the characteristic function of the subsets $\alpha(q) \subset X$. Since $\{A, \alpha\}$ is an M-extension of Q, every measure $\varphi \in \Phi_Q$ can be extended to a measure on A, and then further extended in a natural way to a σ-measure on the σ-ring generated by $A \subset \rho(X)$, i.e., to the σ-ring of Baire subsets of X.[14] For $\varphi \in \Phi_Q$, let H_φ be the Hilbert space of equivalence classes of real or complex functions on X, square-integrable with respect to such a σ-extension of φ.

Now we must take account of the possibility that Φ_Q may be empty or may contain an insufficient number of measures to guarantee the one-to-oneness we require of an extension. To handle these cases, let $\Psi_Q = \Phi_Q \cup \{\xi_1\}$, where ξ_1 is of course the characteristic function of X itself, and let H_{ξ_1} be the Hilbert space of square-integrable real or complex sequences on X. Let $H = \sum^{\oplus}_{\psi \in \Psi_Q} H_\psi$, and for $q \in Q$, let $\pi(q) = \sum^{\oplus}_{\psi \in \Psi_Q} \pi_\psi(q)$, where $\pi_\psi(q)$ is the projection operator defined on H in the standard way by the characteristic function ξ_q. It is obvious then from the construction that $\{P, \pi\}$ is a Boolean Hilbert extension of Q, where P is the set of projection operators on H. To complete the construction of a QM-extension, we simply assign to each $\varphi \in \Phi_Q$ the projection operator $\pi(\varphi) \in P$ corresponding to the one-dimensional subspace containing the vector $|\varphi\rangle = \sum^{\oplus}_{\psi \in \Psi_Q} |\varphi, \psi\rangle$, where

$$|\varphi, \psi\rangle = \begin{cases} \xi_1 \in H_\psi, & \psi = \varphi, \\ 0 \in H_\psi, & \psi \neq \varphi. \end{cases}$$

With this extension of the mapping π to include $Q \cup \Phi_Q$ in its domain, it is clear that $\{P, \pi\}$ is a pure QM-extension of Q. Q.E.D.

To obtain an easy corollary we define:

D10. The set of pure states on an orthoposet Q will be said to *span* Φ_Q if every state is a convex combination of the pure states. □

We have:

C1. If the pure states on an orthoposet Q span Φ_Q, then there exists a convex QM-extension of Q. □

Proof: Let Ψ be the set of pure states on Q, and let $\{P, \pi\}$ be, for, example, the pure QM-extension constructed in T2. Initially define the mapping $\eta: Q \cup \Psi \to P$ by $\eta = \pi | Q \cup \Psi$ and then extend η by convexity to include all of Φ_Q in its domain. If, as before, T is the set of nonnegative self-adjoint trace 1 operators on the Hilbert space of the extension $\{P, \pi\}$ and $E = P \cup T$, then $\{E, \eta\}$ is obviously a convex QM-extension of Q. Q.E.D.

7. CONCLUSIONS

Several implications of our main result (T2) seem to be worth some discussion. First, for any system of physical observables,[15] our theorem implies that we can always find Hilbert space models in terms of which all empirically measurable probabilities can be calculated in the standard quantum theoretical way. The price we have to pay to make this convenience available with complete generality is that macroscopically preparable observables and states may be insufficient to determine the ordering (and hence the orthogonality) relations of the Hilbert space mode. What this implies is that, if we insist on Hilbert space models, and if we can find no reasonable physical grounds for excluding the known examples of orthoposets which cannot have simultaneously order preserving and measurable preserving Hilbert extensions, we may be forced to accept models which introduce "hidden" i.e., not macroscopically preparable) observables and states which violate restrictions usually imposed on physical theories such as locality and causality. Of course, unobservable operators and states have been accepted in Hilbert space models since the discovery of superselection rules, and there are general logical and philosophical grounds for believing that any ultimate explanation of spacetime and its structure (i.e., gravitation) in terms of quantum theoretical entities, presumably more fundamental than spacetime itself, will be compelled to exempt such entities from the, by definition, inapplicable restrictions of locality and causality as we currently understand them. What we find interesting is that the pure mathematics of orthoposets and their probability measures seems to point in the same direction.

Then there are the implications of T2 for Boolean hidden variable theories in general. While the unrelated subset construction introduced in this paper demonstrates the existence of Boolean M-extensions for every orthoposet, and so should end any search for general proofs of the nonexistence of such models, the obviously contrived and vastly "overlarge" nature of the resulting extension in cases where the original orthoposet is not already Boolean should suggest that the *a priori* imposition of the Boolean requirement on physical theory in general would not be in accord with Occam's razor. On the contrary, it seems reasonable to expect that much more practical and economical non-Boolean models will be available for any real physical situation in which distributivity is not in evidence.

Finally, what about the requirement that physical theory be formulated in Hilbert space language? Basically, this is assuming a linear structure for physical theory, with a Euclidean inner product in terms of which probability can be computed. We have seen that this is always possible, but it might be in the same category as the Boolean requirement—an overcomplicating assumption that obscures rather than reveals more economical structures. We believe there are counterindications to this possibility. Piron's work[16] has shown that a wide class of orthoposets (atomic, of degree greater than or equal to four, satisfying the covering law and some additional assumption involving the field of scalars to be used) can be identified, up to isomorphisms, with Hilbert orthoposets. Because of difficulties in physically motivating all the assumptions used, we cannot justifiably call this class the set of physical orthoposets, but Piron's work does indicate the naturalness of the Hilbert space formulation for a broad range of physical situations. More recently, Dirac[17] has discussed the belief, based on his own experience, that the imposition of simple, clear and

rigorous mathematical requirements, such as linearity, on a theory may be more productive than the search for seemingly realistic physical conditions; it was precisely the requirement of linearity that led to the formulation of the Dirac equation.

In the end it may come down to this: whatever the merits of the Hilbert space language as a necessity of nature, it seems at present to be a necessity of thought, allowing us to approximate physical situations as closely as needed, and ultimately to arrive at simpler structures if they exist.

ACKNOWLEDGMENTS

The author gratefully acknowledges the support of a grant from the Research Corporation during the course of this work, and some very helpful discussions with R.J. Greechie, Ron Wright, and Carl Brans.

[1]George W. Mackey, *Mathematical Foundations of Quantum Mechanics* (Benjamin, New York, 1963).
[2]Max Jammer, *The Philosophy of Quantum Mechanics* (Wiley, New York, 1974).
[3]N. Zierler and M. Schlessinger, Duke Math. J. **32**, 251—62 (1965).
[4]R.J. Greechie and F.R. Miller, Technical Report No. 14, Dept. of Math., Kansas State U. (April, 1970).
[5]Private communication to author. (Ron Wright, Dept. of Math. and Stat., U. of Mass., Amherst, Mass. 01002.)
[6]A.M. Gleason, J. Math. Mech. **6**, 885—93 (1957).
[7]Ref. 4 and E.R. Gerelle, R.J. Greechie, F.R. Miller, "Weights on Spaces," edited by C.P. Enz and J. Mehra, in *Physical Reality and Mathematical Description* (Reidel, Dordrecht-Holland, 1974), particularly the example (Fig. V) on p. 186.
[8]M. Eilers and E. Horst, Int. J. Theor. P. **13**, 419—24 (1975).
[9]N. Jacobson, *Lectures in Abstract Algebra* (Van Nostrand, Princeton, New Jersey, 1951), pp. 1, 95.
[10]P.R. Halmos, *Lectures on Boolean Algebras* (Springer, New York, 1974).
[11]Alfred Horn and Alfred Tarski, Trans. Am. Math. Soc. **64**, 467—97 (1948). In the notation of Horn and Tarski (p. 470), $\dashv$ is symbolized by $\leq$, but we prefer the present notation both to avoid confusion with the ordering on the Boolean algebra and because the relation is not quite a partial ordering, since it lacks antisymmetry.
[12]Ref. 11, p. 477.
[13]Ref. 10, p. 77, ff.
[14]Roman Sikorski, *Boolean Algebras* (Springer, New York, 1969), 3rd ed., p. 203.
[15]R.J. Greechie has pointed out to the author that the construction and proofs used in this paper apply equally well when Q is simply an orthocomplemented partially ordered set and Φ_Q is a set of order preserving functions into $[0, 1]$. This opens the possibility that the existence of joins and the modularity condition may be more a requirement of the mathematical language of Boolean or Hilbert extensions rather than a requirement of physical situations. We hope to investigate this in a subsequent paper.
[16]C. Piron, Helv. Phys. Acta **37**, 439—68 (1964); also, *Foundations of Quantum Physics* (Benjamin, New York, 1976).
[17]P.A.M. Dirac, "The Mathematical Foundations of Quantum Theory," edited by A.R. Marlow, in *Mathematical Foundations of Quantum Theory* (Academic, New York, 1978).

On the Kerr–Tomimatsu–Sato family of solutions with nonintegral distortion parameter

Masatoshi Yamazaki

Department of Physics, Kanazawa University, Kanazawa 920, Japan
(Received 7 March 1978)

The generalization of the Kerr–Tomimatsu–Sato family of solutions for gravitational fields of spinning masses to the case of the arbitrary positive nonintegral distortion parameter δ is conjectured.

1. INTRODUCTION

It has been shown[1] that the Kerr[2] and the three Tomimatsu–Sato[3] spinning mass solutions, i.e., stationary axisymmetric, asymptotically flat exact solutions of Einstein's vacuum field equations, can be written in a concise closed form with arbitrary positive integer distortion parameter δ.

The purpose of the present paper is to generalize the Kerr–Tomimatsu–Sato family of spinning mass solutions with arbitrary positive integer distortion parameter δ to solutions[4] with arbitrary positive nonintegral distortion parameter δ.

In Sec. 2 the gravitational field equations for sources with axial symmetry and angular momentum will be expressed in the form of partial differential equations of only the first order.[5] In Sec. 3 the Kerr–Tomimatsu–Sato Family of spinning mass solutions with integral distortion parameter δ will be briefly outlined for the convenience of subsequent discussions. In Sec 4 the Kerr–Tomimatsu–Sato family of solutions with nonintegral distortion parameter δ will be conjectured. Finally, in Sec 5 we shall indicate the basis for our conjecture.

2. GRAVITATIONAL FIELD EQUATIONS

From Einstein's vacuum field equations $R_{\mu\nu}=0$ with the line element written down as

$$ds^2=f^{-1}[e^{2\gamma}(dz^2+d\rho^2)+\rho^2 d\phi^2]-f(dt-\omega d\phi)^2 \tag{1}$$

we obtain the following set of equations:

$$\operatorname{div}\left(\frac{1}{f}\operatorname{grad} f+\omega\frac{f^2}{\rho^2}\operatorname{grad}\omega\right)=0 \tag{2}$$

$$\operatorname{div}\left(\frac{f^2}{\rho^2}\operatorname{grad}\omega\right)=0 \tag{3}$$

$$\frac{\partial\gamma}{\partial z}-\frac{\rho}{2f^2}\left(\frac{\partial f}{\partial z}\frac{\partial f}{\partial\rho}+\frac{\partial\Omega}{\partial z}\frac{\partial\Omega}{\partial\rho}\right)=0 \tag{4}$$

$$\frac{\partial\gamma}{\partial\rho}+\frac{\rho}{4f^2}\left[\left(\frac{\partial f}{\partial z}\right)^2+\left(\frac{\partial\Omega}{\partial z}\right)^2-\left(\frac{\partial f}{\partial\rho}\right)^2-\left(\frac{\partial\Omega}{\partial\rho}\right)^2\right]=0 \tag{5}$$

where f,ω, amd γ are functions of z and ρ only, and Eqs. (6′) are used in Eqs. (4) and (5). We have

$$\rho=\kappa(x^2-1)^{1/2}(1-y^2)^{1/2},\quad z=\kappa xy,\quad \kappa=\frac{mp}{\delta},$$

$$a=x^2-1,\ b=y^2-1,\ \text{and}\ p^2+q^2=1.$$

The mass is m and the angular momentum is m^2q.

Because any source-free vector field may be expressed as a curl of another vector field, and now we treat two-dimensional vectors, it follows from Eq. (3) that there must exist a scalar Ω which satisfies

$$\frac{f^2}{\rho^2}\frac{\partial\omega}{\partial y}=\frac{\partial}{\partial x}\left(\frac{-\Omega}{\kappa b}\right),\quad \frac{f^2}{\rho^2}\frac{\partial\omega}{\partial x}=\frac{\partial}{\partial y}\left(\frac{-\Omega}{\kappa a}\right) \tag{6}$$

or[5]

$$\frac{\partial\omega}{\partial z}=\frac{\rho}{f^2}\frac{\partial\Omega}{\partial\rho},\ \text{and}\ \frac{\partial\omega}{\partial\rho}=-\frac{\rho}{f^2}\frac{\partial\Omega}{\partial z}. \tag{6′}$$

It follows also from Eq. (2) that there must exist another scalar P which satisfies

$$\begin{aligned}\frac{1}{f}\frac{\partial f}{\partial x}+\omega\frac{f^2}{\rho^2}\frac{\partial\omega}{\partial x}&=\frac{\partial}{\partial y}\left(\frac{-\omega\Omega}{\kappa a}+\frac{P}{a}\right),\\ \frac{1}{f}\frac{\partial f}{\partial y}+\omega\frac{f^2}{\rho^2}\frac{\partial\omega}{\partial y}&=\frac{\partial}{\partial x}\left(\frac{-\omega\Omega}{\kappa b}+\frac{P}{b}\right).\end{aligned} \tag{7}$$

From Eqs. (6) and (7) we obtain

$$\begin{aligned}\frac{1}{2f^2}\frac{\partial}{\partial x}\left(f^2+\Omega^2\right)&=\frac{\partial}{\partial y}\left(\frac{P}{a}\right),\\ \frac{1}{2f^2}\frac{\partial}{\partial y}\left(f^2+\Omega^2\right)&=\frac{\partial}{\partial x}\left(\frac{P}{b}\right).\end{aligned} \tag{8}$$

As there is a factor $f^2+\Omega^2$ in Eqs. (8), we choose a common denominator B and define A,B, and $2I$ by relations

$$f=\frac{A}{B}\ \text{and}\ \Omega=\frac{2I}{B}. \tag{9}$$

From three independent scalars A,B, and $2I$ [Ernst's[5] $\epsilon=(A+i2I)/B$] one may arrive at other three independent scalars G,H, and I[Ernst's[5] $\xi=(H+iI)/G$] through definitions

$$A=\frac{(H^2-G^2)+I^2}{G}\ \text{and}\ B=\frac{(H+G)^2+I^2}{G}, \tag{10}$$

or

$$B=A+2H+2G\ \text{and}\ H^2+I^2=AG+G^2. \tag{11}$$

Then Eqs. (8) become[1]

$$\begin{aligned}\frac{2a}{A^2}\left(H\frac{\partial}{\partial x}(A+2G)-(A+2G)\frac{\partial}{\partial x}H\right)&=\frac{\partial P}{\partial y},\\ \frac{2b}{A^2}\left(H\frac{\partial}{\partial y}(A+2G)-(A+2G)\frac{\partial}{\partial y}H\right)&=\frac{\partial P}{\partial x}.\end{aligned} \tag{12}$$

In a similar manner Eqs. (6) become[1]

 0022-2488/78/1909-1847$1.00

$$-\frac{1}{\kappa}\frac{\partial\omega}{\partial x}=\frac{2b}{A^2}\left(I\frac{\partial B}{\partial y}-B\frac{\partial I}{\partial y}\right)=\frac{\partial}{\partial x}\left(Q+R\right),$$

$$-\frac{1}{\kappa}\frac{\partial\omega}{\partial y}=\frac{2a}{A^2}\left(I\frac{\partial B}{\partial x}-B\frac{\partial I}{\partial x}\right)=\frac{\partial}{\partial y}\left(Q+R\right) \tag{13}$$

$$\frac{2b}{A^2}\left(I\frac{\partial}{\partial y}(A+2G)-(A+2G)\frac{\partial}{\partial y}I\right)=\frac{\partial Q}{\partial x},$$

$$\frac{2a}{A^2}\left(I\frac{\partial}{\partial x}(A+2G)-(A+2G)\frac{\partial}{\partial x}I\right)=\frac{\partial Q}{\partial y}, \tag{14}$$

and

$$\frac{4b}{A^2}\left(I\frac{\partial}{\partial y}H-H\frac{\partial}{\partial y}I\right)=\frac{\partial R}{\partial x},$$

$$\frac{4a}{A^2}\left(I\frac{\partial}{\partial x}H-H\frac{\partial}{\partial x}I\right)=\frac{\partial R}{\partial y}. \tag{15}$$

From Eqs. (13) we obtain

$$\omega=-\kappa(Q+R)+\text{constant}, \tag{16}$$

where the constant in Eq. (16) is determined so as to be $\omega=0$ when $q=0$ (zero angular momentum). Equations (12), (14), and (15) are the gravitational field equations for sources with axial symmetry and angular momentum in the form of partial differential equations of only first order.

Three Noether currents of the present SU(1, 1) symmetry[6] are

$$n\left(a^{-1/2}\frac{\partial P}{\partial y},\ -(-b)^{-1/2}\frac{\partial P}{\partial x}\right),$$

$$n\left(a^{-1/2}\frac{\partial Q}{\partial y},\ -(-b)^{-1/2}\frac{\partial Q}{\partial x}\right),$$

and

$$n\left(a^{-1/2}\frac{\partial R}{\partial y},\ -(-b)^{-1/2}\frac{\partial R}{\partial x}\right), \tag{17}$$

where the factor n is $(a-b)^{-1/2}$.

3. SOLUTIONS WITH INTEGER δ

In the present section the Kerr–Tomimatsu–Sato family of solutions with integral distortion parameter δ will be outlined for the convenience of subsequent discussions. The solutions are[1]

$$f=A/B, \tag{18}$$

$$\omega=-\kappa\left(Q+R-\frac{\delta}{pq}\right), \tag{19}$$

$$e^{2\gamma}=A/p^{2\delta}(a-b)^{\delta^2} \tag{20}$$

$$A=F(\delta^2)=\sum_{r=1}^{\delta}e(r)c(\delta,r)f(r)F(\delta^2-r), \tag{21}$$

$$B=A+2H+2G, \tag{22}$$

$$H=\sum_{r=1}^{\delta}d(r)pxa^{(r-1)}\sum_{r'=r}^{\delta}c(\delta,r')F(\delta^2-r') \tag{23}$$

$$I=\sum_{r=1}^{\delta}d(r)(-qy)b^{r-1}\sum_{r'=r}^{\delta}c(\delta,r')F(\delta^2-r'), \tag{24}$$

$$G=\sum_{r=1}^{\delta}c(\delta,r)F(\delta^2-r), \tag{25}$$

$$P=\frac{2qy}{pqA}\delta\sum_{r=1}^{\delta}\sum_{r'=1}^{\delta}p^2a^rb^{1-r'}g(\delta,r,r')F(\delta^2-r), \tag{26}$$

$$Q=-\frac{2px}{pqA}\delta\sum_{r=1}^{\delta}\sum_{r'=1}^{\delta}q^2b^ra^{1-r'}\,g(\delta,r,r')F(\delta^2-r), \tag{27}$$

and

$$R=\frac{\delta}{pqA}\sum_{r=1}^{\delta}\sum_{r'=1}^{\delta}\left(p^2a^rb^{1-r'}-q^2b^ra^{1-r'}\right)h(\delta,r,r')\times F(\delta^2-r). \tag{28}$$

Coefficients $c(\delta,r)$, $d(r)$, $e(r)$, $g(\delta,r,r')$, and $h(\delta,r,r')$, are[1]

$$c(\delta,r)=\frac{(\delta+r-1)!}{(\delta-r)!}\frac{2^{2r-1}\delta}{(2r)!},\quad r=1,2,\cdots,\delta, \tag{29}$$

$$d(r)=\frac{(-1)^{r-1}(2r-2)!}{\{2^{r-1}(r-1)!\}^2},\quad r=1,2,\cdots, \tag{30}$$

$$e(r)=-2d(r+1), \tag{31}$$

$$g(\delta,r,r')=\frac{re(r)c(\delta,r)}{\delta^2}\sum_{t=r'}^{\delta}\frac{td(t-r'+1)c(\delta,t)}{r+t-1}; \tag{32}$$

$$r,r'=1,2,\cdots,\delta,$$

and

$$h(\delta,r,r,')=\frac{rr'e(r)c(\delta,r)c(\delta,r')}{\delta^2(r+r'-1)},$$

$$r,r'=1,2,\cdots,\delta. \tag{33}$$

Functions $f(r)$ are[1]

$$f(r)=p^2a^r+q^2b^r,\quad r=1,2,3,\cdots. \tag{34}$$

Polynomials $F(\delta^2)$ and $F(\delta^2-r)$ are[1]

$$F(\delta^2)=\frac{\det(f(r+r'-1)/(r+r'-1))}{\det(1/(r+r'-1))},$$

$$r,r'=1,2,\cdots,\delta \tag{35}$$

and

$$F(\delta^2-r)=\frac{(-1)^{r-1}}{re(r)c(\delta,r)}\frac{\det(f(s+t-1)/(s+t-1))}{\det(1/(i+j-1))},$$

$$s=1,2,\cdots,r-1,\ r+1,\cdots,\delta,\ t=2,3,\cdots,\delta, \tag{36}$$

$$i,j=1,2,\cdots,\delta,\text{ and }r=1,2,\cdots,\delta.$$

The determinant form of $F(\delta^2)$ and $F(\delta^2-r)$, $r=1,2,\cdots,\delta$ comes from relations[1]

$$\sum_{r=1}^{\delta}h(\delta,r,r')f(r+r'-1)F(\delta^2-r)=\delta_{r'1}F(\delta^2),\qquad r'=1,2,\cdots,\delta,$$

which are obtained from $H^2+I^2=AG+G^2$. Polynomials $re(r)c(\delta,r)F(\delta^2-r)$, $r=1,2,\cdots,\delta$ are cofactors of $F(\delta^2)$ with respect to the row r, column $r'=1$ elements, respectively. The solutions given in Eqs. (18) ~ (36) satisfy[7] gravitational field Eqs. (12), (14), and (15).

4. SOLUTIONS WITH POSITIVE NONINTEGRAL δ

In the present section we shall give the conjectured Kerr–Tomimatsu–Sato family of solutions with nonintegral distortion parameter δ and we shall explain the

basis of our conjecture in the next section.

The solutions with positive nonintegral δ are obtained by means of the following four recipes from the solutions with integral δ given in Sec. 3:

I. Coefficient $c(\delta, r)$ is generalized into the form

$$c(\delta, r)=\frac{\Gamma(\delta+r)}{\Gamma(\delta-r+1)}\,\frac{2^{2r-1}\delta}{(2r)!},\quad r=1,\ 2,\ 3,\ \cdots, \tag{37}$$

where $\Gamma(\delta+r)$ are gamma functions.

II. Polynomials $F(\delta^2)$ of degree δ^2 in a and b and degree δ in p^2 and q^2 are replaced by

$$\begin{aligned}F(\delta^2)=&\sum_{s=0}^{\infty}\sum_{r_i=1}^{\infty}\cdots\sum_{r'_k=1}^{\infty}\cdots p^{2(\delta-s)}q^{2s}a^{\delta^2-n}b^n(-1)^{n+s}\frac{1}{s!}\\&\times\left(\frac{\Pi_{i>j=1}^{s}(r_i-r_j)\Pi_{k>l=1}^{s}(r'_k-r'_l)}{\Pi_{i,k=1}^{s}(r_i+r'_k-1)}\right)^2\\&\times\left(\prod_{i=1}^{s}\frac{\Gamma(\delta+r_i)}{\Gamma(\delta-r_i+1)\{(r_i-1)!\}^2}\right)\\&\times\left(\prod_{k=1}^{s}\frac{\Gamma(\delta+r'_k)}{\Gamma(\delta-r'_k+1)\{(r'_k-1)!\}^2}\right),\end{aligned} \tag{38}$$

where

$$n=\sum_{i=1}^{s}r_i+\sum_{k=1}^{\delta}r'_k-s.$$

III. Polynomials $F(\delta^2-r)$, $r=1,2,\cdots,\delta$, of degree δ^2-r in a and b and degree $\delta-1$ in p^2 and q^2 are replaced by

$$\begin{aligned}F(\delta^2-r)=&\sum_{s=0}^{\infty}\sum_{r_i=1}^{\infty}\cdots\sum_{r'_k=1}^{\infty}\cdots p^{2(\delta-1-s)}q^{2s}a^{\delta^2-r-n}b^n(-1)^{n+s}\frac{1}{s!}\\&\times\left(\frac{\Pi_{i=1}^{s}(r_i-r)\Pi_{k=1}^{s}(r'_k-1)}{\Pi_{i=1}^{s}r_i\Pi_{k=1}^{s}(r+r'_k-1)}\right)\\&\times\left(\frac{\Pi_{i>j=1}(r_i-r_j)\Pi_{k>l=1}^{s}(r'_k-r'_l)}{\Pi_{i,k=1}^{s}(r_i+r'_k-1)}\right)^2\\&\times\left(\prod_{i=1}^{s}\frac{\Gamma(\delta+r_i)}{\Gamma(\delta-r_i+1)[(r_i-1)!]^2}\right)\\&\times\left(\prod_{k=1}^{s}\frac{\Gamma(\delta+r'_k)}{\Gamma(\delta-r'_k+1)\{(r'_k-1)!\}^2}\right),\end{aligned} \tag{39}$$

where $r=1,\ 2,\ 3,\ldots$ and also

$$n=\sum_{i=1}^{s}r_i+\sum_{k=1}^{s}r'_k-s.$$

IV. All summations with respect to r, r', etc. in Eqs. (21)—(28), and (32) are replaced by summations to infinity instead of to δ.

It should be noted that the solutions which are obtained by means of these four recipes reduce to the solutions given in Eqs. (18)~(36) of Sec. 3 when the distortion parameter δ becomes a positive integer. Eqs. (38) and (39) for the functions $F(\delta^2)$ and $F(\delta^2-r)$ are valid when $|b/a|<1$. When $|a/b|<1$, p and q, and a and b must be simultanously interchanged in Eqs. (38) and (39).

5. DERIVATIONS OF NONINTEGRAL SOLUTIONS

In the present section we shall derive the nonintegral solutions given in the preceding Sec. 4 from the integral solutions given in Sec. 3.

Polynomials $F(\delta^2)$ and $F(\delta^2-r)$ with positive integral δ were written down in Eqs. (35) and (36) of Sec. 3 as determinants of degree δ and $\delta-1$, respectively. With the help of Laplace's expansion theorem Eq. (35) for $F(\delta^2)$ and Eq. (36) for $F(\delta^2-r)$ can be written down as

$$\begin{aligned}F(\delta^2)=&\sum_{s=0}^{\delta}\sum_{r_i=1}^{\delta}\cdots\sum_{r'_k=1}^{\delta}\cdots p^{2(\delta-s)}q^{2s}a^{\delta^2-n}b^n(-1)^{n+s}\frac{1}{s!}\\&\times\left(\frac{\Pi_{i>j=1}^{s}(r_i-r_j)\Pi_{k>l=1}^{s}(r'_k-r'_l)}{\Pi_{i,k=1}^{s}(r_i+r'_k-1)}\right)^2\\&\times\left(\prod_{i=1}^{s}\frac{(\delta+r_i-1)!}{(\delta-r_i)![(r_i-1)!]^2}\right)\\&\times\left(\prod_{k=1}^{s}\frac{(\delta+r'_k-1)!}{(\delta-r'_k)![(r'_k-1)!]^2}\right)\end{aligned} \tag{40}$$

and

$$\begin{aligned}F(\delta^2-r)=&\sum_{s=0}^{\delta-1}\sum_{r_i=1}^{\delta}\cdots\sum_{r'_k=1}^{\delta}\cdots p^{2(\delta-1-s)}a^{\delta^2-r-n}b^n(-1)^{n+s}\frac{1}{s!}\\&\times\left(\frac{\Pi_{i=1}^{s}(r_i-r)\Pi_{k=1}^{s}(r'_k-1)}{\Pi_{i=1}^{s}r_i\Pi_{k=1}^{s}(r+r'_k-1)}\right)\\&\times\left(\frac{\Pi_{i>j=1}^{s}(r_i-r_j)\Pi_{k>l=1}^{s}(r'_k-r'_l)}{\Pi_{i,k=1}^{s}(r_i+r'_k-1)}\right)^2\\&\times\left(\prod_{i=1}^{s}\frac{(\delta+r_i-1)!}{(\delta-r_i)![(r_i-1)!]^2}\right)\\&\times\left(\prod_{k=1}^{s}\frac{(\delta+r'_k-1)!}{(\delta-r'_k)![(r'_k-1)!]^2}\right),\end{aligned} \tag{41}$$

where

$$n=\sum_{i=1}^{s}r_i+\sum_{k=1}^{s}r'_k-s \text{ and } r=1,\ 2,\cdots,\delta.$$

It follows that Eq. (40) for $F(\delta^2)$, Eq. (41) for $F(\delta^2-r)$, and Eq. (29) for $c(\delta, r)$ contain the factors $(\delta+r-1)!/(\delta-r)!$. When integral δ is generalized into positive nonintegral δ, these factors $(\delta+r-1)!/(\delta-r)!$ may be generalized into gamma functions $\Gamma(\delta+r)/\Gamma(\delta-r+1)$. In this way, we arrive at the four recipes, given in Sec. 4 for obtaining nonintegral solutions. It will be left to a future study to show that these conjectured solutions with nonintegral δ actually satisfy the graviational field Eqs. (12), (14), and (15).

From the solutions given in Sec. 4 it is straightforward to obtain charged spinning mass solutions with nonintegral δ using the formulas which are given in the third paper cited in Ref. 1.

[1] M. Yamazaki, J. Math. Phys. **18**, 2502 (1977); Prog. Theor. Phys. **57**, 1951 (1977); J. Math. Phys. **19**, 1376 (1978).
[2] R. P. Kerr, Phys. Rev. Lett. **11**, 237 (1963).
[3] A. Tomimatsu and H. Sato, Phys. Rev. Lett. **29**, 1344 (1972); Prog. Theor. Phys. **50**, 95 (1973).
[4] C. M. Cosgrove, J. Phys. A **10**, 1481 (1977). Here an alternative approach to the construction of such solutions has been discussed.
[5] F. J. Ernst, Phys. Rev. **167**, 1175 (1968).
[6] R. Matzner and C. Misner, Phys. Rev. **154**, 1229 (1967).
[7] S. Hori, "On the exact solution of Tomimatsu-Sato family for an arbitrary integral value of the deformation parameter," Prog. Theor. Phys. (to be published).

Search for periodic Hamiltonian flows: A generalized Bertrand's theorem

E. Onofri and M. Pauri

Istituto di Fisica dell'Università di Parma, Sezione di Fisica Teorica, 43100 Parma, Italy
and Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Milano, Italy
(Received 8 August 1977)

A complete classification is given of the two-dimensional Hamiltonian systems (whose Hamilton–Jacobi equation separates in Cartesian or polar coordinates) which admit strictly periodic motions for open sets of initial conditions (completely degenerate systems). Any of the systems which are separable in Cartesian coordinates turn out to be canonically equivalent to some anisotropic harmonic oscillator. In the polar case our results provide a generalization of a celebrated theorem of Bertrand. It is proven that all the completely degenerate systems fall into two families. These families are characterized by the semiclassical inverse "spectral functions"

$n_1J_1+n_2J_2=\mathcal{J}(H)=\alpha(-H)^{-1/2}-\beta,$

$n_1J_1+n_2J_2=\mathcal{J}(H)=\alpha H-\beta$

(α, β real positive constants) and contain, as central symmetric cases, the Kepler system and the isotropic harmonic oscillator, respectively. Qualitative and higher symmetry properties of these systems are also discussed.

1. INTRODUCTION

In this paper we present a method to characterize strictly periodic Hamiltonian dynamical systems in classical mechanics. It is well known, since the last century, that the only strictly periodic systems with rotational symmetry are the isotropic harmonic oscillator and the Kepler problem (Bertrand's theorem).[1,2] The extension to noncentral systems has been considered in the recent literature,[3,4] the aim being to classify systems possessing a maximal dynamical symmetry. The technique employed to determine such systems was based on the alleged connection between complete degeneracy of the motion (i.e., strict periodicity) and the separability of the Hamilton–Jacobi equation in more than one system of coordinates. As we stressed in previous papers,[5,6] this connection is not entirely rigorous and the results which stem from this principle must be checked. This is due, essentially, to the local character of curvilinear coordinates: The separation constants in such coordinates[7] may fail to be isolating in phase space.[5] Only if the motion is confined to the regularity domain of the local chart can one be sure that such constants of the motion are isolating, but this fact can be checked only on the explicit solutions of the equations of motion. On the other hand, there exist completely degenerate systems which are separable in only one system of coordinates (anisotropic oscillators, for instance) and these are overlooked in this approach.

We shall adopt a more direct attack on the problem. Starting from the existence of action-angle variables in a suitable region in phase space, we shall impose the degeneracy condition directly in the form

$$\mathcal{J}(H)=n_1J_1+n_2J_2, \tag{1.1}$$

where H is the Hamiltonian, (J_1,J_2) are action variables, (n_1,n_2) are positive integers without common factors whose meaning is well known from analytical mechanics, and the function $\mathcal{J}(H)$ is *a priori* arbitrary—it is one of the unknown functions in our problem. This line has been followed by Enriotti and Faccini[4]; at a certain point in their analysis, however, they introduce an "ansatz" which severely limits the generality of their results. We shall be able, instead, to transform Eq. (1.1) into a system of integral equations where the potentials defining H enter as unknown functions. These equations can be reduced to linear integral equations of the simplest type (Abel's equation); as such they are easily solved. This appears to be a simple extension to the multidimensional case of the "inverse problem" studied in Ref. 2 (Sec. 12). It is well known that the classical inverse scattering problem can also be solved in this way.[8]

We shall limit ourselves to systems with two degrees of freedom which are separable in Cartesian coordinates or in polar coordinates. In these cases we obtain an exhaustive characterization of all the potentials which admit strictly periodic motions. Unlike the previous known examples which depend only on a finite number of parameters, our potentials are defined in terms of one (polar case) or two (Cartesian case) arbitrary functions. It is nevertheless a simple task to check whether or not a given system is completely degenerate using our classification. The Cartesian case is essentially trivial, being completely reducible to the calculation of Landau and Lifshitz. The result in this case is that H is canonically isomorphic to some harmonic oscillator. A more refined analysis is required for the polar case. We shall prove that the integral equations which follow from Eq. (1.1) admit solutions if and only if $\mathcal{J}$ is of the form

$$\mathcal{J}(H)=\begin{cases}\alpha(-H)^{-1/2}-\beta,\\ \alpha H-\beta\end{cases} \tag{1.2}$$

(α,β real positive constants). We find two families of strictly periodic Hamiltonian systems, then; the Kepler problem belongs to the first family, the harmonics oscillator to the second one. This is the *generalized Bertrand's theorem* to which the title of this paper refers. All these Hamiltonians (except for $\beta=0$ which correspond to the Kepler problem and to the harmonic oscillator) have roughly the same qualitative properties. The motion takes place around a minimum

 0022-2488/78/1909-1850$1.00

of the potential which is at a distance $r_0 \propto \beta^2$ from the origin with a depth $V_0 \propto \beta^{-2}$ for the first family and $r_0 \propto \sqrt{\beta}$, $V_0 \propto \beta$ for the second one.

Little is known in general about the global symmetry in phase space of these Hamiltonians. In particular instances we have proved the existence of a global SU(2) symmetry for the members of the first (Kepler) family with $n_1 = n_2 = 1$; the Schrödinger equation of the corresponding quantum systems has been solved and it shows the characteristic higher degeneracy of a SU(2)-symmetric Hamiltonian. The existence of a dynamical (noninvariance) group and the "geometric quantization" of these Hamiltonians are being investigated.[9]

The paper is organized as follows: In Sec. 2 we discuss the Cartesian case, essentially to show how Abel's equation works in the simplest case. In Sec. 3 the more complex problem of systems which are separable in polar coordinates is dealt with and we prove our generalized Bertrand's theorem. In Sec. 4 we summarize the results and discuss the qualitative features and the symmetry properties of these Hamiltonians. Finally, the elementary properties of Abel's equation (Euler transform) are quoted in the Appendix, for the reader's convenience. Three tables at the end of the paper summarize the results.

2. "RECTANGULAR" SYSTEMS

Let us consider a particle of mass m moving in a two-dimensional rectangular potential $U(x_1, x_2) = U_1(x_1) + U_2(x_2)$ so that the Hamiltonian is

$$H(x_1, x_2, p_1, p_2) = \frac{1}{2m}(p_1^2 + p_2^2) + U_1(x_1) + U_2(x_2). \tag{2.1}$$

The Hamilton–Jacobi equation is then separable in Cartesian coordinates and the constants of separation may be chosen to be the partial energies E_1, E_2 $(E_1 + E_2 = E)$. Assuming that the functions U_1, U_2 are sufficiently well behaved and possess a local minimum at x_{0_1}, x_{0_2} respectively, it follows that for certain intervals of values of E_1, E_2 the surface $\sum(E_1, E_2)$,

$$\sum(E_1, E_2) \equiv \left\{ \frac{1}{2m} p_i^2 + U_i(x_i) = E_i, \quad i = 1, 2 \right\}, \tag{2.2}$$

is compact and sufficiently regular. Actually in this case $\sum(E_1, E_2)$ is a two-dimensional torus[10] and it is possible to define the action-angle variables

$$J_i = \frac{1}{2\pi} \oint \{2m[E_i - U_i(x)]\}^{1/2} dx. \tag{2.3}$$

It is well known that the Hamiltonian is a function of these action variables only, and not of the conjugate angle variables w_1, w_2. It is also well known that the system admits strictly periodic motions if and only if the Hamiltonian depends on the action variables through the expression $n_1 J_1 + n_2 J_2$ with positive integers n_1, n_2,[11] i.e., if a continuous function $\mathcal{J}$ exists such that

$$\mathcal{J}(E) = n_1 J_1 + n_2 J_2 = \mathcal{J}(E_1 + E_2). \tag{2.4}$$

Let us choose the additive constant of the energy in such a way that $U_i(x_{0_i}) = 0$. Since J_i is a function of E_i alone and $J_i(0) = 0$, it follows that

$$\mathcal{J}(E_1 + E_2) = n_1 J_1(E_1) + n_2 J_2(E_2) = \mathcal{J}(E_1) + \mathcal{J}(E_2), \tag{2.5}$$

which, because of the continuity of $\mathcal{J}$, implies

$$\mathcal{J}(\lambda E) = \lambda \mathcal{J}(E), \tag{2.6}$$

i.e., $\mathcal{J}$ is linear,

$$\mathcal{J}(E) = E/\omega \tag{2.7}$$

(ω real positive constant).

Then we have

$$J_i(E_i) = \frac{1}{2\pi} \oint \sqrt{2m[E_i - U_i(x)]}\, dx = E_i/(n_i \omega). \tag{2.8}$$

We shall now regard Eq. (2.8) as a system of integral equations defining $U_i(x_i)$. The two cases $i = 1, 2$ can be solved separately; to simplify the notation, we drop the index i, for the moment, and rewrite the equation as

$$\oint \sqrt{2m[E - U(x)]}\, dx = \frac{2\pi}{n\omega} E. \tag{2.8'}$$

This equation can be linearized[2] by introducing the inverse function $x = x(U)$; since $U(x)$ has a minimum at $x = x_0$, $x(U)$ has two branches $x_+(U)$, $x_-(U)$ which merge at $U = 0$. Let $\delta(U) = x_+(U) - x_-(U)$ be the discontinuity along the cut. Equation (2.8′) is now transformed as follows,

$$\begin{aligned} \frac{2\pi}{n\omega} E &= 2\int_0^E \sqrt{2m(E-U)} \left(\frac{dx_+}{dU} - \frac{dx_-}{dU} \right) dU \\ &= 2\sqrt{2m} \int_0^E \sqrt{E-U}\, \frac{d\delta(U)}{dU}\, dU \\ &= \sqrt{2m} \int_0^E \frac{\delta(U)}{\sqrt{E-U}}\, dU. \end{aligned} \tag{2.9}$$

After an integration by parts, we have used the fact that $\delta(0) = 0$. The equation

$$\int_0^E \frac{\delta(U)}{\sqrt{E-U}}\, dU = \frac{2\pi}{\sqrt{2mn\omega}} E \tag{2.10}$$

is of Abel's type and the solution is found at once [see Eq. (A3)],

$$\delta(U) = \left(\frac{2}{m}\right)^{1/2} \frac{1}{n\omega} \int_0^U \frac{dE}{\sqrt{U-E}} = \frac{2}{n\omega} \left(\frac{2U}{m}\right)^{1/2}. \tag{2.11}$$

The most general solution $U(x)$ of Eq. (2.8′) is then obtained by writing

$$x(U) = \pm \tfrac{1}{2}\delta(U) + G(U), \tag{2.12}$$

where $G(U)$ is a single-valued function of U near $U = 0$. The simplest choice, $G = x_0 = \text{constant}$, corresponds to the harmonic oscillator potential

$$U(x) = \frac{m}{2}(n\omega)^2 (x - x_0)^2. \tag{2.13}$$

In order that $U(x)$ has the required properties (single valuedness, a local minimum at x_0), $G(U)$ must be restricted to verify

$$|G'(U)| < (n\omega \sqrt{2mU})^{-1}. \tag{2.14}$$

As an example, let us consider

$$G(U)=\alpha(U+\gamma)^{1/2} \tag{2.15}$$

with $\gamma>0$, $|\alpha|<\alpha_0\equiv(1/n\omega)(2/m)^{1/2}$. By inverting we find

$$U(x)=(\alpha_0^2-\alpha^2)^{-2}\{\alpha_0 x-\alpha[x^2+\gamma(\alpha_0^2-\alpha^2)]^{1/2}\}^2. \tag{2.16}$$

The choice $\alpha=\alpha_0$ gives rise to an infinite barrier at the origin $x=0$; the potential is

$$U(x)=(2\alpha_0)^{-2}\left(x-\frac{\alpha_0^2\gamma}{x}\right)^2, \tag{2.17}$$

which is well known to give oscillations with a period independent of the energy.

In conclusion, the most general Hamiltonian of the form (2.1) which satisfies Eq. (2.4) is constructed as follows: We choose any two functions $G_1(U_1)$, $G_2(U_2)$ satisfying Eq. (2.14); these are inserted into Eq. (2.12) which defines $U_1(x_1)$ and $U_2(x_2)$. Notice that all these Hamiltonians can be transformed into some anisotropic harmonic oscillator through a global canonical transformation.[10]

3. "POLAR" SYSTEMS

A. The classical inverse bound state problem

Let us now consider a particle of mass m moving in a potential $U(r,\varphi)$ so that the Hamiltonian is

$$H(r,\varphi,p_r,p_\varphi)=\frac{1}{2m}\left(p_r^2+\frac{p_\varphi^2}{r^2}\right)+V(r)+\frac{U(\varphi)}{r^2}. \tag{3.1}$$

The Hamilton–Jacobi equation is then separable in polar coordinates. We assume that for certain intervals of values of the separation constants E,λ the surface

$$\sum(E,\lambda)\equiv\begin{cases}H(r,p_r,\varphi,p_\varphi)=E,\\ p_\varphi^2+2mU(\varphi)=\lambda^2,\end{cases} \tag{3.2}$$

is compact and sufficiently regular. In this case, it is a two-dimensional torus and it is possible to define the action variables

$$J_r(E,\lambda)=\frac{1}{2\pi}\oint\{2m[E-V(r)]-\lambda^2/r^2\}^{1/2}dr,$$
$$J_\varphi(\lambda)=\frac{1}{2\pi}\oint[\lambda^2-2mU(\varphi)]^{1/2}d\varphi. \tag{3.3}$$

Let

$$J_r(E,\lambda)=\mathcal{J}(E,J_\varphi(\lambda)). \tag{3.4}$$

We shall solve the following inverse problem: Given the functions $\mathcal{J}(E,J_\varphi)$ and $J_\varphi(\lambda)$, determine the potentials $V(r)$ and $U(\varphi)$. Let us recall that $\mathcal{J}$ *and* $J_\varphi(\lambda)$ *determine the semiclassical spectrum of* H, so that the problem may have an independent interest in itself. The problem will be solved separately for $V(r)$ and $U(\varphi)$. The effective potential is defined as usual,

$$V(\lambda,r)=V(r)+\frac{\lambda^2}{2mr^2}. \tag{3.5}$$

In order to have bounded orbits on a compact $\sum(E,\lambda)$ the effective potential must have a minimum at some $r_0(\lambda)$; we define

$$V_0(\lambda)=V(\lambda,r_0(\lambda)). \tag{3.6}$$

In a neighborhood of $r_0(\lambda)$ we define the inverse function $r(\lambda,V)$ with two branches $r_1(\lambda,V)\geq r_2(\lambda,V)$. We have

$$J_r(E,\lambda)=\frac{1}{\pi}\int_{V_0}^{E}\sqrt{2m(E-V)}\,\frac{d}{dV}(r_1(\lambda,V)-r_2(\lambda,V))\,dV$$
$$=\frac{\sqrt{2m}}{\pi}\int_{V_0}^{E}\sqrt{E-V}\,\delta'(\lambda,V)\,dV$$
$$=\frac{\sqrt{2m}}{2\pi}\int_{V_0}^{E}\frac{\delta(\lambda,V)}{\sqrt{E-V}}\,dV, \tag{3.7}$$

where $\delta(\lambda,V)=r_1-r_2$. We are led to an Abel integral equation also in this case. Solving for δ [see Eq. (A3)] we obtain

$$\delta(\lambda,V)=\left(\frac{2}{m}\right)^{1/2}\int_{V_0(\lambda)}^{V}\frac{\mathcal{J}_E'(E,J_\varphi)}{\sqrt{V-E}}\,dE. \tag{3.8}$$

Now let us differentiate the first Eq. (3.3) with respect to λ; it follows that

$$\frac{\lambda}{2\pi}\oint\frac{d(1/r)}{\sqrt{2m(E-V)}}=\frac{\partial}{\partial\lambda}J_r(E,\lambda). \tag{3.9}$$

Then, by defining $\eta(\lambda,V)=1/r_2-1/r_1$ we have

$$\frac{\lambda}{\pi}\int_{V_0}^{E}\frac{\eta'(\lambda,V)}{\sqrt{2m(E-V)}}\,dV=\frac{\partial}{\partial\lambda}J_r(E,\lambda). \tag{3.10}$$

As usual, we obtain

$$\eta(\lambda,V)=\frac{\sqrt{2m}}{\lambda}\int_{V_0(\lambda)}^{V}\frac{\mathcal{J}_\lambda'(E,J_\varphi(\lambda))}{\sqrt{V-E}}\,dE. \tag{3.11}$$

Finally, from the knowledge of δ and η we can find $r(\lambda,V)$, namely

$$r=r(\lambda,V)=\pm\tfrac{1}{2}\delta(\lambda,V)+\left(\frac{1}{4}\delta^2(\lambda,V)+\frac{\delta(\lambda,V)}{\eta(\lambda,V)}\right)^{1/2}. \tag{3.12}$$

Equations (3.8), (3.11), and (3.12) completely solve the problem for the radial potential $V(r)$. Actually the function $V_0(\lambda)$ is readily obtained from the data $\mathcal{J}(E,J_\varphi)$ and $J_\varphi(\lambda)$, since it holds that

$$\mathcal{J}(V_0(\lambda),J_\varphi(\lambda))=0. \tag{3.13}$$

Notice that, unlike the Cartesian case, $V(r)$ *is uniquely determined.*

We can find $U(\varphi)$ by a similar procedure. Let U_0 be the minimum of $U(\varphi)$. In a neighborhood of U_0 we can define the inverse function $\varphi(U)$ with two branches $\varphi_1(U)\geq\varphi_2(U)$. For λ^2 sufficiently close to $2mU_0$ it must hold that

$$\sqrt{2m}\int_{U_0}^{\lambda^2/2m}\left(\frac{\lambda^2}{2m}-U\right)^{1/2}\frac{d}{dU}[\varphi_1(U)-\varphi_2(U)]\,dU=\pi J_\varphi(\lambda),$$
$$J_\varphi((2mU_0)^{1/2})=0. \tag{3.14}$$

This is again an Abel equation which can be solved to give

$$\delta\varphi=\varphi_1(U)-\varphi_2(U)=\left(\frac{2}{m}\right)^{1/2}\int_{U_0}^{U}\frac{(d/d\Lambda)J_\varphi(\sqrt{2m\Lambda})}{\sqrt{U-\Lambda}}\,d\Lambda. \tag{3.15}$$

As in Sec. 2, the most general $U(\varphi)$ is then obtained

by solving for U in the equation

$$\varphi(U) = \pm \tfrac{1}{2}\delta\varphi(U) + G(U), \tag{3.16}$$

where G is any regular function in a neighborhood of the real positive U axis ($U \geq U_0$).

This formulation of the problem is correct in the assumption that U_0 is an isolated minimum, but not every choice of $J_\varphi(\lambda)$ is compatible with this assumption. For example, $J_\varphi(\lambda) = \lambda$ does not make sense when inserted in Eq. (3.15); we would obtain $\delta\varphi(U) = \pi$ for every U which is incompatible with a local minimum of U. This fact raises a serious problem, which must also be solved for $J_r(E, \lambda)$, namely to determine the class of admissible initial data $J_r(E, \lambda)$, $J_\varphi(\lambda)$ for which a solution to the inverse problem exists. What may happen is that the potential $V(\lambda, r)$ which we obtain by solving for V in Eq. (3.12), fails to be of the form $V(r) + \lambda^2/(2mr^2)$. In this case all of the formalism breaks down, and the only conclusion would be that the data is incompatible with the particular model given by Eq. (3.1). This problem will be solved in Sec. 3C in the special case of completely degenerate systems.

B. Example: The "Kepler family"

An interesting example (which we shall prove to exhaust almost all the possibilities for the periodic case) is given by the following choice of initial data:

$$\begin{aligned} &J_r(E, \lambda) = \mathcal{J}(E) - qJ_\varphi(\lambda), \\ &\mathcal{J}(E) = mk(-2mE)^{-1/2} - \beta \quad (k, \beta \text{ real positive constants}), \\ &J_\varphi(\lambda) = J(\lambda), \quad \text{(to be determined)} \\ &q = \text{rational positive number.} \end{aligned} \tag{3.17}$$

By inserting this data into Eqs. (3.8), (3.11), and (3.13), we easily obtain:

$$\begin{aligned} &\delta(\lambda, V) = -\frac{k}{V}\left(1 - \frac{V}{V_0}\right)^{1/2}, \\ &\eta(\lambda, V) = 2\sqrt{2m}\, q[J'(\lambda)/\lambda](V - V_0(\lambda))^{1/2}, \\ &mk[-2mV_0(\lambda)]^{-1/2} = qJ(\lambda) + \beta. \end{aligned} \tag{3.18}$$

From Eq. (3.12) we then obtain

$$r(\lambda, V) = \pm \frac{k}{2V}\left(1 - \frac{V}{V_0}\right)^{1/2} + \left[\left(-\frac{k}{2V}\right)^2\left(1 - \frac{V}{V_0}\right) - \frac{k\lambda}{2\sqrt{-2mV_0}qJ'V}\right]^{1/2} \tag{3.19}$$

and, after some algebraic manipulations,

$$r^2V^2 + 2\left(\frac{\lambda(qJ+\beta)}{2mqJ'} - \frac{(qJ+\beta)^2}{m}\right) - k^2 + \left(\frac{\lambda}{2mr}\right)^2\left(\frac{qJ+\beta}{qJ'}\right)^2 = 0. \tag{3.20}$$

This shows that the solution $V = V(\lambda, r)$ is of the form $V(r) + \lambda^2/(2mr^2)$ *if and only if* $J(\lambda)$ *satisfies the differential equation*

$$-\frac{\lambda(qJ+\beta)}{qJ'} + 2(qJ+\beta)^2 = \lambda^2 + c_1, \tag{3.21}$$

whose general solution is

$$J(\lambda) = \frac{1}{\sqrt{2q}}\{\lambda^2 + c_1 \pm [(\lambda^2 + c_1)^2 + c_2]^{1/2}\}^{1/2} - \frac{\beta}{q}. \tag{3.22}$$

Correspondingly we obtain the following potential,

$$V(r) = c_1/2mr^2 - (k^2\gamma^2 - c_2(2m)^{-2})^{1/2}\, r^{-2}. \tag{3.23}$$

The constant c_1 is irrelevant since it can always be associated with $U(\varphi)$; it will be adjusted in order to have $U(\varphi) \geq 0$, as we have tacitly assumed in Eq. (3.2). Equation (3.22) gives the most general $J(\lambda)$ for which the radial inverse problem has solutions: only the plus sign survives when we solve for the angular potential. To simplify the calculation, let $c_1 = 0$, $c_2 = -\gamma^4$, then it holds that

$$J(\lambda) = \frac{1}{2q}[(\lambda^2 + \gamma^2)^{1/2} \pm (\lambda^2 - \gamma^2)^{1/2}] - \frac{\beta}{q}. \tag{3.24}$$

This is to be inserted into Eq. (3.15) to yield

$$\delta\varphi(U) = q^{-1}\tan^{-1}\left(\frac{U - U_0}{U_0 + \gamma^2/(2m)}\right)^{1/2} + q^{-1}\tan^{-1}\left(\frac{U - U_0}{U_0 - \gamma^2/(2m)}\right)^{1/2}, \tag{3.25}$$

where $U_0 = (1/2m)[\beta^2 + (\gamma^2/2\beta)^2]$. The minus sign in Eq. (3.22) would give a negative $\delta\varphi(U)$ and must be discarded. According to Eq. (3.16) we have now to choose a function $G(U)$; the simplest choice is $G = 0$ which gives

$$U(\varphi) = \frac{\beta^2}{m}\,\frac{1 + \alpha^2\cos^2(2q\varphi)}{1 + \cos(2q\varphi)[1 - \alpha^2\sin^2(2q\varphi)]^{1/2}} \tag{3.26}$$

which is defined for $|\varphi| < \pi/2q$, $\alpha = \frac{1}{2}(\gamma/\beta)^2 < 1$. The limiting case $\alpha \to 1$ gives a nonanalytic potential for which U_0 is not an isolated minimum; apart from that, the potential is continuous, with continuous gradient and it gives a strictly periodic motion,

$$U(\varphi)\big|_{\alpha=1} = \frac{\beta^2}{2m}\,\frac{1 + \cos^2(2q\varphi)}{1 + \cos(2q\varphi)\,|\cos(2q\varphi)|}. \tag{3.27}$$

C. The generalized Bertrand's theorem

We are now prepared to examine the basic problem of classifying all the Hamiltonians which admit separation in polar coordinates and whose time evolution is periodic for an open region in phase space. We shall make use of the formalism of Sec. 3 A and we shall choose

$$J_r(E, \lambda) \equiv \mathcal{J}(E) - qJ(\lambda), \quad J_\varphi(\lambda) \equiv J(\lambda), \tag{3.28}$$

where q is any positive rational number. The example of Sec. 3 B shows that *not every input* $\{\mathcal{J}(E), J(\lambda), q\}$ *is compatible with the structure of the Hamiltonian.* Therefore, we shall now determine the most general admissible input. Equations (3.8), (3.11), and (3.13) assume the form

$$\begin{aligned} &\text{(i)}\ \delta(\lambda, V) = \left(\frac{2}{m}\right)^{1/2}\int_{V_0(\lambda)}^{V}\frac{\mathcal{J}'(E)\,dE}{\sqrt{V - E}}, \\ &\text{(ii)}\ \eta(\lambda, V) = 2\sqrt{2mq}\,\frac{J'(\lambda)}{\lambda}\sqrt{V - V_0(\lambda)}, \\ &\text{(iii)}\ \mathcal{J}(V_0(\lambda)) = qJ(\lambda). \end{aligned} \tag{3.29}$$

By differentiation we find

$$\delta'_\lambda = -\left(\frac{2}{m}\right)^{1/2} \frac{qJ'(\lambda)}{\sqrt{V - V_0(\lambda)}} \tag{3.30}$$

and Eq. (3.12) is transformed into

$$r(\lambda, V) = \pm\tfrac{1}{2}\delta + \left(\frac{\delta^2}{4} - \frac{\lambda\delta\delta'_\lambda}{(2qJ')^2}\right)^{1/2}. \tag{3.31}$$

The condition that $V(\lambda, r)$, implicitly defined by Eq. (3.31), be of the form $V(r) + \lambda^2/(2mr^2)$ can be stated in differential terms as follows:

$$\left.\frac{\partial V}{\partial \lambda}\right|_{r=\text{const}} = -\frac{\partial r}{\partial \lambda} \Big/ \frac{\partial r}{\partial V} = \frac{\lambda}{mr^2}, \tag{3.32}$$

i.e.,

$$mr^2 \frac{\partial r}{\partial \lambda} + \frac{\partial r}{\partial V} = 0. \tag{3.33}$$

By inserting $r(\lambda, V)$ in terms of δ [Eq. (3.31)], we obtain *two* partial differential equations in the single unknown δ, namely

$$\delta''_{\lambda V} = -\Omega\delta'^3_\lambda, \tag{3.34}$$

$$\delta''_{\lambda\lambda} = \frac{\delta'_V}{\Omega\delta^2} - 2\frac{\delta'^2_\lambda}{\delta} - \frac{\Omega - 1 + \lambda\Omega'}{\lambda\Omega}\delta'_\lambda, \tag{3.35}$$

where $\Omega = [2qJ'(\lambda)]^{-2}$. [Notice that Eq. (3.34) is precisely the differential version of Eq. (3.30). In order to simplify the notation, we put $m = 1$ until Eq. (3.60)].

Since we have more equations than unknown functions, some compatibility conditions must be satisifed. We proceed as follows: By differentiating Eq. (3.34) with respect to λ, and Eq. (3.35) with respect to V we find a third equation, independent from the previous ones, namely

$$\delta''_{VV} = 2\frac{\delta'^2_V}{\delta} + 2\Omega^2\delta\delta'^4_\lambda - 5\Omega\delta'_V\delta'^2_\lambda + 2\Omega\left(\frac{\Omega - 1}{\lambda} + \frac{\Omega'}{2}\right)\delta^2\delta'^3_\lambda. \tag{3.36}$$

There is a second integrability condition between this latter and Eq. (3.34) which gives

$$12(\Omega\delta\delta'^2_\lambda - \delta'_V)^2 + 16\left(\frac{\Omega - 1}{\lambda} + \frac{\Omega'}{2}\right)\delta'_\lambda\delta^2(\Omega\delta\delta'^2_\lambda - \delta'_V) - \delta^4\delta'^2_\lambda\left(\Omega\Omega'' - 2\Omega'^2 + \lambda^{-1}(7\Omega' - 5\Omega\Omega') + \frac{(\Omega - 1)(6 - 8\Omega)}{\lambda^2}\right) = 0. \tag{3.37}$$

This implies that for some function $G(\lambda)$ it must hold that

$$\delta'_V = \Omega\delta\delta'^2_\lambda + G(\lambda)\,\delta'_\lambda\delta^2. \tag{3.38}$$

By comparing this latter equation with Eq. (3.34) we find

$$\delta\delta'^2_\lambda\left[3G - 2\left(\frac{\Omega - 1}{\lambda} + \frac{\Omega'}{2}\right)\right] + \delta'_\lambda\delta^2\left[G' + \frac{G^2}{\Omega} - \frac{G}{\Omega}\left(\frac{\Omega - 1}{\lambda} + \Omega'\right)\right] = 0. \tag{3.39}$$

We are forced now to conclude that both quantities in square brackets in Eq. (3.39) must vanish, since otherwise we would obtain that

$$\delta'_\lambda/\delta = \Phi(\lambda), \tag{3.40}$$

i.e., $\delta(\lambda, V)$ would be separated in a product of a function of λ times a function of V which in turn would imply [through Eq. (3.30)] that

$$V_0(\lambda) = \text{const} \Rightarrow J(\lambda) = \text{const} \tag{3.41}$$

which is nonsense. In conclusion we can go on with Eq. (3.38) where $G(\lambda)$ is a solution of the following equations:

$$G(\lambda) = \tfrac{1}{3}\left(\Omega' + 2\frac{\Omega - 1}{\lambda}\right), \qquad G'(\lambda) = \frac{G}{3\Omega}\left(2\Omega' + \frac{\Omega - 1}{\lambda}\right). \tag{3.42}$$

It follows that the function $\Omega(\lambda)$, which is related to our input function $J(\lambda)$, must satisfy the following second order equation

$$3\Omega\Omega'' - 2\Omega'^2 + \lambda^{-1}\Omega'(\Omega + 5) - 2\lambda^{-2}(\Omega - 1)(4\Omega - 1) = 0. \tag{3.43}$$

The general solution of this equation can be found to be

$$\Omega(\lambda) = \frac{(\lambda^2 + c_1)^2 + c_2}{2\lambda^2[\lambda^2 + c_1 \pm ((\lambda^2 + c_1)^2 + c_2)^{1/2}]}. \tag{3.44}$$

Actually it will be more convenient to use Eq. (3.42) as the formal definition of $G(\lambda)$. From Eq. (3.38) it follows that

$$\Delta'_V = \frac{1}{3}\left(5\Omega' + 4\frac{\Omega - 1}{\lambda}\right)\Delta^2 + 2\Omega\Delta\Delta'_\lambda, \tag{3.45}$$

where $\Delta = \delta\delta'_\lambda$; we can also derive an equation for Δ directly from Eq. (3.35); it holds that

$$\Delta'_V = \left(\Omega' + \frac{\Omega - 1}{\lambda}\right)\Delta^2 + \Omega\Delta\Delta'_\lambda. \tag{3.46}$$

Let us solve Eq. (3.45), (3.46) with respect to Δ'_λ and Δ'_V:

$$\Delta'_\lambda = -\frac{G'}{G}\Delta, \quad \Delta'_V = G\Delta^2. \tag{3.47}$$

The general solution is then

$$\Delta(\lambda, V) = -G(\lambda)^{-1}(V - a)^{-1} \tag{3.48}$$

The only integration constant a may be dropped, since it is an additive constant to the energy. The singular case $G(\lambda) \equiv 0$ must be treated separately (see below). From Eq. (3.48) we obtain

$$\delta = \frac{\Delta}{\delta'_\lambda} = \frac{\sqrt{2\Omega}\sqrt{V - V_0(\lambda)}}{VG(\lambda)}. \tag{3.49}$$

By differentiating with respect to λ and comparing to Eq. (3.30), it follows that

$$\frac{V'_0}{V_0} = \frac{G(\lambda)}{\Omega(\lambda)} = \frac{d}{d\lambda}\ln\frac{G(\lambda)^2}{\Omega}, \tag{3.50}$$

where the last step is a consequence of Eq. (3.42). In conclusion we have

$$V_0(\lambda) \propto G^2/\Omega. \tag{3.51}$$

Now by computing the derivative of $(-V_0)^{-1/2}$ we find

$$\frac{d}{d\lambda}[-V_0(\lambda)]^{-1/2} \propto \frac{d}{d\lambda}\left(\frac{\sqrt{\Omega}}{G}\right) = -\frac{1}{2\sqrt{\Omega}} \propto qJ', \tag{3.52}$$

hence

$$[-V_0(\lambda)]^{-1/2} \propto qJ(\lambda) + \text{const.} \tag{3.53}$$

By recalling Eq. (3.29iii) we conclude that the only admissible $\mathcal{J}(E)$ is given by

$$\mathcal{J}(E) = \alpha(-E)^{-1/2} - \beta. \tag{3.54}$$

As a consequence, the potentials found in Sec. 3B) exhaust all the possibilities, except for the singular case $G \equiv 0$, which we shall now examine.

Equation (3.47) with $G(\lambda) \equiv 0$ becomes

$$\Delta'_V = 0, \quad \Delta'_\lambda = -\frac{\Omega'}{2\Omega}\Delta \tag{3.55}$$

which imply $\Delta(\lambda, V) = \text{const} \times \Omega^{-1/2}$. It follows that

$$\delta(\lambda, V) = -\sqrt{2}C\sqrt{V - V_0} \quad (C \text{ a real constant}). \tag{3.56}$$

An argument similar to that used above shows that in this case

$$V_0'(\lambda) \propto \Omega^{-1/2} = 2qJ'(\lambda), \tag{3.57}$$

hence

$$\mathcal{J}(E) = \alpha E - \beta. \tag{3.58}$$

In conclusion the only admissible input functions are

$$\mathcal{J}(E) = \begin{cases} \alpha(-E)^{-1/2} - \beta, \\ \alpha E - \beta. \end{cases} \tag{3.59}$$

This is the result we refer to as a "generalized Bertrand's theorem."

We are left to consider in detail the case $\mathcal{J}(E) = \alpha E - \beta$. We have

$$\alpha V_0(\lambda) - \beta = qJ(\lambda),$$

$$\delta = \left(\frac{2}{m}\right)^{1/2} 2\alpha\sqrt{V - V_0},$$

$$r(\lambda, V) = \pm\alpha\left(\frac{2}{m}\right)^{1/2}\sqrt{V - V_0} + \left(\frac{2\alpha^2}{m}(V - V_0) + \frac{\alpha\lambda}{mqJ'}\right)^{1/2}. \tag{3.60}$$

By inverting we find

$$V(\lambda, r) = V_0 + \frac{m}{8\alpha^2}r^2 - \frac{\lambda}{4\alpha qJ'} + \frac{\lambda^2}{2m(2qJ')^2 r^2}. \tag{3.61}$$

The admissible $J(\lambda)$ are given by

$$(2qJ')^{-2} = 1 + c_1\lambda^{-2}. \tag{3.62}$$

Correspondingly the potential $V(r)$ is given by

$$V(r) = \frac{m}{8\alpha^2}r^2 + \frac{c_1}{2mr^2}. \tag{3.63}$$

The second term is trivial, since it can be reabsorbed into the angular part $U(\varphi)$; therefore, a linear $\mathcal{J}(E)$ corresponds to a harmonic oscillator radial potential. $U(\varphi)$ can be obtained as a special case of Eq. (3.25) with $\gamma = 0$; we do not find anything new with respect to the potentials already known, except for the arbitrariness in the choice of $G(U)$.

The discussion of the angular part $U(\varphi)$ is not complete until we have discussed the following situation. Suppose that $U(\varphi)$ is bounded and let

$$U_0 \leq U \leq U_1, \quad \lambda^2 > 2mU_1. \tag{3.64}$$

This means that the particle rotates around the origin. The equation defining $J_\varphi(\lambda)$ is the following,

$$J_\varphi(\lambda) = \frac{1}{2\pi}\int_0^{2\pi} d\varphi[\lambda^2 - 2mU(\varphi)]^{1/2}, \tag{3.65}$$

which *is not* equivalent to an integral equation of Abel's type. In general we do not have an explicit inversion formula for this case, however if we require that the flow is periodic for $\lambda^2 > 2mU_1$, we know that $J_\varphi(\lambda)$ is restricted to the form given in Eqs. (3.22) and (3.24). For this particular input we can solve Eq. (3.65). Let

$$\frac{1}{2\pi}\int_0^{2\pi} d\varphi[\lambda^2 - 2mU(\varphi)]^{1/2} = \frac{1}{2q}(\lambda^2 + \gamma^2)^{1/2} \pm \frac{1}{2q}(\lambda^2 - \gamma^2)^{1/2} - \frac{\beta}{q}. \tag{3.66}$$

An asymptotic estimate as $\lambda \to \infty$ shows that $\beta = 0$, $q = 1$, and only the plus sign is admissible. The equation

$$\frac{1}{2\pi}\int_0^{2\pi} d\varphi[\lambda^2 - 2mU(\varphi)]^{1/2} = \frac{1}{2q}(\lambda^2 + \gamma^2)^{1/2} + \frac{1}{2q}(\lambda^2 - \gamma^2)^{1/2} \tag{3.67}$$

admits an infinite number of solutions obtained as follows: Let A_1, A_2 be a measurable partition of the interval $(0, \ldots, 2\pi)$ such that $\mu(A_1) = \mu(A_2) = \pi$. The potential

$$U(\varphi) = \begin{cases} -\gamma^2/2m, & \varphi \in A_1, \\ +\gamma^2/2m, & \varphi \in A_2, \end{cases} \tag{3.68}$$

is a solution of Eq. (3.67). It is still to be shown that this is the general solution. To this aim, let us make an expansion in λ^{-1},

$$(2\pi)^{-1}\int_0^{2\pi} d\varphi \sum_n \binom{1/2}{n}(-2mU(\varphi))^n \lambda^{-2n} = \tfrac{1}{2}\sum_n \binom{1/2}{n}[\gamma^{2n} + (-\gamma)^{2n}]\lambda^{-2n}, \tag{3.69}$$

which implies ($a = \gamma^2/2m$)

$$(2\pi)^{-1}\int_0^{2\pi} U(\varphi)^n d\varphi = \tfrac{1}{2}[a^n + (-a)^n]. \tag{3.70}$$

It follows that

$$(2\pi)^{-1}\int_0^{2\pi}[U(\varphi)^2 - a^2]^2 d\varphi = 0. \tag{3.71}$$

$(U^2 - a^2)^2$ being nonnegative, it must vanish, which shows that Eq. (3.68) is the general solution. The other choice of c_2 [i.e., $J(\lambda)$ given by Eq. (3.22) with $c_2 > 0$] does not admit solutions of this type, since we would obtain an equation

$$(2\pi)^{-1}\int_0^{2\pi} U^{2n} d\varphi = (-1)^n a^{2n}, \tag{3.72}$$

which is impossible.

4. CONCLUDING REMARKS

Let us summarize the results obtained in previous sections and briefly comment on them. The most general Hamiltonian in two degrees of freedom with strictly periodic time evolution has been determined by

assuming separability in Cartesian coordinates *or* polar coordinates. The result is given in Table I (Cartesian case) and Table II′—II″ (polar case). It will be noted that previously known potentials are contained as special cases.[3,4] We have not listed discontinuous potentials which arise for some special choice of parameters (see Sec. 3B) nor the potentials given by Eq. (3.68) which also are discontinuous.

It would take too long to study the properties of these Hamiltonians in detail. We shall limit ourselves to a few remarks.

First of all, since all the Hamiltonians belonging to the same family have in common the same function $\mathcal{J}(H)$ apart from the additive constant β, one may ask whether it is possible to identify them by means of a canonical transformation. The simplest candidate is of course a translation on the action variables,

$$J \to \tilde{J} = J + \text{const}, \qquad w \to \tilde{w} = w. \tag{4.1}$$

However such a transformation cannot be everywhere defined since the action variables are positive definite. Even if we allow for more general transformations, it can be proven that Hamiltonians with different values of β are canonically inequivalent from a global point of view.[10] This fact is obvious in the case $\beta = 0$ for the "Kepler" family (H is unbounded from below) as compared to the case $\beta \neq 0$ (H is bounded from below). Still a transformation like Eq. (4.1) may be used *locally* to find the solution of Hamilton's equations for a Hamiltonian with $\beta > 0$ starting from the known solution of the Kepler problem [see for example the Hamiltonian defined by Eq. (4.3) below].

Another interesting point is the existence of a global symmetry [SO(3) or SU(2)] for the Hamiltonians we have classified. In the Cartesian case, with $n_1 = n_2$, it is fairly obvious that a suitable choice of $G_i(U_i)$ such that J_1 and J_2 are still defined in the whole range $(0, +\infty)$ gives a Hamiltonian essentially equivalent to an isotropic harmonic oscillator. We conclude, in this case, that the Hamiltonian is SU(2)-symmetric, the realization of the Lie algebra in terms of Poisson brackets being given by

$$\begin{aligned} M_1 &= (J_1 J_2)^{1/2} \cos(w_1 - w_2), \\ M_2 &= (J_1 J_2)^{1/2} \sin(w_1 - w_2), \\ M_3 &= \tfrac{1}{2}(J_1 - J_2), \\ I &= (M_1^2 + M_2^2 + M_3^2)^{1/2} = H/(2\omega). \end{aligned} \tag{4.2}$$

TABLE I. Hamiltonian in Cartesian coordinates (x_1, x_2, p_1, p_2).

$$H = \frac{1}{2m}(p_1^2 + p_2^2) + U_1(x_1) + U_2(x_2)$$

$$\mathcal{J}(H) = n_1 J_1 + n_2 J_2 = \frac{H}{\omega}$$

$$x_i(U_i) = \left(\frac{2}{m}\right)^{1/2} U_i^{1/2}(n_i\omega)^{-1} + G_i(U_i) \quad (i = 1, 2)$$

$G_i(U_i)$ single valued in a neighborhood of $U_i \geq 0$

$$|G_i'| < (2m n_i^2 \omega^2 U_i)^{-1/2}$$

Special cases:

$G_i = 0$ Harmonic oscillator

$G_i = \alpha(U_i + \gamma)^{1/2}$ Eq. (2.16)

$\alpha = (1/n_i\omega)\left(\frac{2}{m}\right)^{1/2}$ Eq. (2.17)

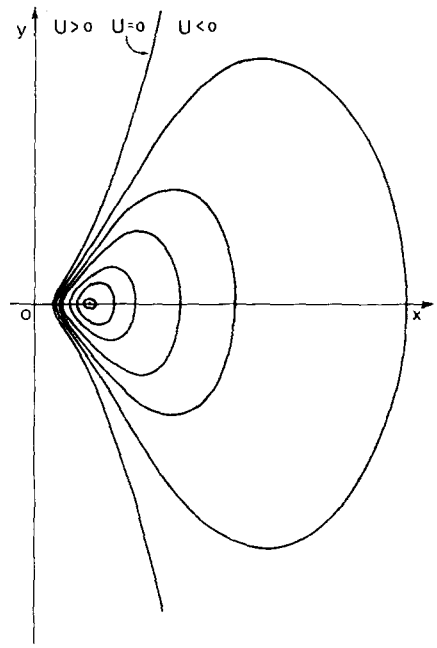


FIG. 1. $U(r, \varphi)$. $\gamma = \kappa = 1$, $m = 1/2$, $q = 1$, $\beta = (3)^{-1/4}$, $r_0 = (3)^{-1/2}$.

Consider now the polar case. Let H belong to the first family, $\beta > 0$, $G(U) = 0$, $q = 1$. The potential has a minimum at $\varphi = 0$, $r = r_0 \propto \beta^2$ with a depth $V_0 \propto \beta^{-2}$ (see Fig. 1). Unlike the Kepler problem, both J_r and J_φ are defined in the range $(0, +\infty)$. The motion takes place around the minimum and for a sufficiently low energy the system can be approximated by an isotropic harmonic oscillator. So we expect a SU(2) symmetry instead of the SO(3) symmetry of the regularized Kepler problem. If we defined M_i as in Eq. (4.2) with $J_1 = J_\varphi$ and $J_2 = J_r$ we have a realization of the Lie algebra of SU(2); in some cases it is easy to prove that this realization can be integrated to give a global realization of the group SU(2). This is the case of the Hamiltonian (first given in Ref. 3)[12]

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p^2}{r^2}\right) - \frac{k}{r} + \beta^2(2mr^2\cos^2\varphi)^{-1}. \tag{4.3}$$

A direct calculation of $M_i(J, w)$ shows that these functions are differentiable throughout the energy surface $H = E(-mk^2/(2\beta^2) < E < 0)$; this fact, together with the compactness of the energy surface, implies the integrability of the Lie algebra realization to a realization of SU(2) (according to a well-known theorem by Palais[13]). It is plausible that the other Hamiltonians in both families (at least for $q = 1$) also share the same SU(2) symmetry. The isotropic oscillator ($q = \frac{1}{2}$) and the Kepler problem emerge here as rather exceptional cases.

Finally, it may be interesting to investigate whether

TABLE II′. Hamiltonian in polar coordinates (r, φ, p_φ)—"Kepler" family.

$$H=\frac{1}{2m}(p_r^2+p_\varphi^2/r^2)+V(r)+U(\varphi)/r^2$$

$$\mathcal{J}(H)=J_r+qJ_\varphi=\frac{mk}{\sqrt{-2mH}}-\beta$$

$$J(\lambda)=\frac{1}{\sqrt{2}\,q}\{\lambda^2+c_1+[(\lambda^2+c_1)^2+c_2]^{1/2}\}^{1/2}-\frac{\beta}{q}$$

$$V(r)=\frac{c_1}{2mr^2}-\frac{1}{r^2}\left(k^2r^2-\frac{c_2}{(2m)^2}\right)^{1/2}$$

$$\varphi(U)=\pm\frac{1}{2q}\tan^{-1}\left(\frac{2\beta\sqrt{2m(U-U_0)}}{\Gamma-2m(U-U_0)}\right)+G(U)\quad(\beta>0),\quad U(\varphi)=0 \text{ if } \beta=0$$

$\beta\geq 0$, $q=n_1/n_2>0$, $k>0$ (c_1, c_2) real constants, $\beta=0\Longrightarrow c_1=c_2=0$, $q=1$ (Kepler)

$$U_0=(2m)^{-1}[\beta^2-c_2(2\beta)^{-2}-c_1]\geq 0,\quad \Gamma=\beta^2+c_2(2\beta)^{-2}$$

$$|G'(U)|<\beta\left(U+\frac{c_1}{2m}+\frac{c_2}{4m\beta^2}\right)[2m(U-U_0)]^{1/2}\left[\left(U+\frac{c_1}{2m}+\frac{c_2}{4m\beta^2}\right)^2+\frac{c_2}{4m\beta^2}\right]^{-1}$$

$c_1=0$, $c_2=-\gamma^4$, $G=0$: $V=-\frac{1}{r}\left[k^2+\left(\frac{\gamma^2}{2mr}\right)^2\right]^{1/2}$, $U(\varphi)=\frac{\beta^2}{m}\frac{1+\alpha^2\cos^2(2q\varphi)}{1+\cos(2q\varphi)[1-\alpha^2\sin^2(2q\varphi)]^{1/2}}$

$c_1=c_2=0$, $G=q^{-1}\tan^{-1}(U/U_1-1)^{1/2}$, $V=-\frac{k}{r}$, $U(\varphi)=\frac{\beta_1^2}{\cos^2(\frac{1}{2}q\varphi)}+\frac{\beta_2^2}{\sin^2(\frac{1}{2}q\varphi)}$

where

$$\alpha=\frac{\gamma^2}{2\beta^2}\quad \beta_1=\frac{\sqrt{U_0}-\sqrt{U_1}}{2},\quad \beta_2=\frac{\sqrt{U_0}+\sqrt{U_1}}{2}$$

a SU(2, 1) transitive realization can be defined in the negative energy portion of phase space; this would allow a geometric quantization[9] of these Hamiltonians.

For the time being, we checked that the Schrödinger quantization of the Hamiltonian of Eq. (4.3) shows the characteristic degeneracy of a SU(2)-symmetric Hamiltonian and the spectrum is the semiclassical one; both results should be confirmed in a geometric quantization.

APPENDIX: ABEL'S EQUATION

The equation

$$\int_a^x \frac{f(y)\,dy}{\sqrt{x-y}}=g(x) \tag{A1}$$

is a special case of Abel's equation (or Euler transform).[14] When $g(x)$ is absolutely continuous, the solution is given by

$$\int_a^x f(t)\,dt=\frac{1}{\pi}\int_a^y \frac{g(x)}{\sqrt{y-x}}\,dx \tag{A2}$$

or, equivalently, by

$$f(y)=\frac{1}{\pi}\frac{g(a)}{\sqrt{y-a}}+\frac{1}{\pi}\int_a^y \frac{g'(x)\,dx}{\sqrt{y-x}}\,. \tag{A3}$$

In order that $f(y)$ be bounded in a neighborhood of a it is necessary that

$$g(a)=0. \tag{A4}$$

In fact, from $|f(y)|<M$ $(a\leq y\leq x)$, it follows that

$$|g(x)|=\left|\int_a^x \frac{f(y)\,dy}{\sqrt{x-y}}\right|<2M\sqrt{x-a} \tag{A5}$$

and $g(a)=0$.

If moreover we require that $f(a)=0$, as in most of our applications, then $g(x)$ must behave like $(x-a)^\alpha$ with $\alpha>+\frac{1}{2}$ as $x\to a$. In fact

$$\begin{aligned} f(y)&\approx\frac{1}{\pi}\int_a^y\frac{\alpha(x-a)^{\alpha-1}}{\sqrt{y-x}}\\ &=\frac{\alpha}{\pi}\int_0^1 dt(y-a)^{\alpha-1/2}(1-t)^{-1/2}t^{\alpha-1}\\ &=\frac{\alpha}{\pi}(y-a)^{\alpha-1/2}B(\alpha,1/2). \end{aligned} \tag{A6}$$

TABLE II″. Hamiltonian in polar coordinates—"Oscillator" family.

$$\mathcal{J}(H)=J_r+qJ_\varphi=\frac{H}{2\omega}-\beta$$

$$J(\lambda)=\frac{1}{q}(\lambda-\beta)$$

$$V(r)=\tfrac{1}{2}m\omega^2r^2$$

$$\varphi(U)=\pm q^{-1}\tan^{-1}\left(\frac{U}{U_0}-1\right)^{1/2}+G(U)\quad(\beta>0),\quad U=0 \text{ if } \beta=0$$

See Table II′ with $c_1=c_2=0$; $\omega>0$

$G=0$: $U(\varphi)=U_0/\cos^2(q\varphi)$

$G=q^{-1}\tan^{-1}\left(\frac{U}{U_1}-1\right)^{1/2}$: $U(\varphi)$ as in Table II′

[1]J. Bertrand, Comptes Rendus **77**, 849—853 (1873); A. Wintner, *The Analytical Foundation of Celestial Mechanics* (Princeton U.P., Princeton, N.J., 1941), Chap. III, pp. 211—19; V. Arnold, *Les Méthodes Mathématiques de la mécanique classique* (Editions MIR, Moscow, 1976), Chap. 2, Sec. 8D.
[2]L. Landau and E. Lifshitz, *Mechanics* (Pergamon, New York, 1960), Chap. III.
[3]J. Fris, V. Mandrosov, Ya. A. Smorodinsky, M. Uhlir, and P. Winternitz, Phys. Lett. **16**, 354—6 (1965); A. A. Makarov, Ya. A. Smorodinsky, Kh. Valiev, and P. Winternitz, Nuovo Cimento **52A**, 1061—84 (1967).
[4]M. Enriotti and M. L. Faccini, Suppl. Nuovo Cimento, VI serie prima **4**, 1109—26 (1968); Nuovo Cimento **62A**, 561—80 (1969).
[5]E. Onofri and M. Pauri, J. Math. Phys. **14**, 1106—15 (1973).
[6]E. Onofri and M. Pauri, "On the Relations among Degeneration, Symmetry and Separability of the Dynamical Equations," Parma University, IFPR-T-039, Parma, 1974.
[7]E. Onofri and M. Pauri, Lett. Nuovo Cimento, Ser. I **2**, 607—13 (1969).
[8]R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966), Chap. V, Sec. 5.9.
[9]See for instance D. Simms, "Proceedings of the 4th International Colloquium on Group-Theoretical Methods in Physics, 1975," University of Nijmegen, The Netherlands (Lecture Notes in Physics, Vol. 50, Springer, 1976).
[10]V. I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics* (Benjamin, New York, 1968), Appendixes 26, 32, and 33.
[11]As is well known, n_i represents the number of times that the projection of an orbit of the system into the coordinate plane (q_i/p_i) is covered during one period of the motion.
[12]It is worthwhile to notice that the transformation

$$\tilde{r} = r,$$
$$\tilde{p}_r = p_r,$$
$$\tilde{\varphi} = \cos^{-1}\left[\left(\frac{p_\varphi^2 + \beta^2/\sin^2\varphi}{p_\varphi^2 + \beta^2\cot^2\varphi}\right)^{1/2}\cos\varphi\right],$$
$$\tilde{p}_\varphi = \left(p_\varphi^2 + \frac{\beta^2}{\sin^2\varphi}\right)^{1/2},$$

maps the system defined by Eq. (4.3) into a subsystem of the Kepler problem corresponding to $\tilde{p}_\varphi > \beta$. In terms of action-angle variables this transformation reads

$$J_\varphi \to J_\varphi = \tilde{J}_\varphi + \beta,$$
$$J_r \to \tilde{J}_r = J_r,$$
$$w_\varphi \to \tilde{w}_\varphi = w_\varphi,$$
$$w_r \to \tilde{w}_r = w_r,$$

and thus is of the form given by Eq. (4.1). Of course it does not represent a global canonical transformation and therefore it cannot relate the symmetry properties of the two systems. Actually the generators of the Kepler system[7] are transformed into

$$M_x = \frac{M_z}{(-2mH)^{1/2}(M_z^2 - \beta^2)^{1/2}} \times [-p_r p_\varphi \sin\varphi + (mk - M_z^2/r)\cos\varphi],$$
$$M_y = \frac{1}{(-2mH)^{1/2}(M_z^2 - \beta^2)^{1/2}} \times [M_z^2 p_r \cos\varphi + (mk - M_z^2/r)p_\varphi \sin\varphi],$$
$$M_z = (p_\varphi^2 + \beta^2/\sin^2\varphi)^{1/2},$$

which are clearly singular at $\tilde{p}_\varphi = M_z = \beta$.
[13]R. S. Palais, "A Global Formulation of the Lie Theory of Transformation Groups," Mem. Amer. Math. Soc. **22**, (1957).
[14]E. Goursat, *A Course in Mathematical Analysis*, Vol. III, Part 2: "Integral Equations, Calculus of Variations" (Dover, New York, 1964), Chap. VIII, I, *Tables of Integral Transforms*, edited by A. Erdélyi (McGraw-Hill, New York, 1954), Vol. II, Chap. XIII; L. Tonelli, Math. Ann. **99**, 183—199 (1928).

Light-cone finite normal products

John H. Lowenstein[a)]

Department of Physics, New York University, New York, New York 10003

Eugene R. Speer[b)]

Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08903
and Institute for Advanced Study, Princeton, New Jersey 08540
(Received 20 January 1978)

A graphical subtraction procedure for constructing the perturbative Green functions of light-cone finite, multiply localized products of fields is proposed. The existence of the Green functions as tempered distributions is proved, together with the properties of light-cone finiteness and localization on a line segment. The derivation of light cone expansions is sketched, but not treated in detail.

1. INTRODUCTION

A convenient operator expansion for displaying in concise form the light-cone singularities of products of fields would be one of the from[1]

$$A(x)B(y)=N^*[A(x)B(y)]+\sum_i F_i((x-y)^2)O_i(x,y), \qquad (1.1)$$

where $N^*[A(x)B(y)]$ is a light-cone finite bilocal field (normal product), the $O_i(x, y)$ are smeared out N^* normal products, and the $F_i(z^2)$ are complex valued functions which are singular for z^2 tending to zero. In Ref. 1, an expansion of the form (1.1) has been shown to exist in a relatively simple example, that of the product $A(x)A(y)$ in the perturbative A^4 model. In that case, the F_i are all powers of logarithms, and the O_i can all be expressed in terms of smeared-out light-cone finite normal products

$$N^*[\partial_{\mu_1}\cdots\partial_{\mu_a}A(x)\partial_{\nu_1}\cdots\partial_{\nu_b}A(y)].$$

In certain respects, the construction of Ref. 1 falls short of a completely satisfactory realization of the program sketched above. To motivate the present work, it is useful to review the major deficiencies of that construction.

(1) It depends on a rather unwieldy subtraction procedure, based on iterated application of Zimmermann's identities[2] relating short-distance finite normal products. Although the subtractions have a recursive structure reminiscent of renormalization, no prescription for removing light-cone singularities on a graph-by-graph basis is given.

(2) It does not establish uniform localization of the formally bilocal fields appearing in the light-cone expansion. A reasonable definition (presumably not the only one) of "bilocal" would require $O_i(x, y)$ to be localized on the line segment joining x to y, with $O_i(x, y)$ commuting with all z which are spacelike with respect to all points of that line segment. In the construction of Ref. 1, the formal bilocals are indeed localized on a segment of the line passing through x and y, but unbounded growth of the localization segment with increasing perturbative order is not excluded.

(3) The construction cannot be generalized to other products of fields, such as products of currents in a charged scalar theory, without encountering formidable complications. In Ref. 1, the analysis was considerably simplified by the limitations to logarithmic light-cone singularities and bilinear products of the basic fields and their derivatives. Such simplifications could not be expected to persist for general field products.

(4) There is no uniform (in all orders in the coupling constant) polynomial bound on the momentum-space growth at infinity of the vertex functions of the N^* products defined in Ref. 1. Hand in hand with this is the necessity of increasingly many subtractions as one proceeds to higher orders, just as one finds for ordinary vertex functions in a nonrenormalizable theory. This suggests that outside of perturbation theory, the N^* product may not be well-defined as a tempered distribution, requiring stronger large-momentum cut-offs than are provided by Schwartz class test functions. Moreover, there will be no renormalization group or Callan—Symanzik equations for the vertex functions of such normal products, making expansions such as (1.1) of only limited usefulness in the phenomenology of theories with asymptotic freedom.

In this paper we define light-cone finite normal products which avoid the first three of the enumerated difficulties. We develop a graph-by-graph subtraction scheme which allows one to define a quite general light-cone finite normal product $N^*[\prod_i \phi_i(x+\theta_i\xi)]$ localized on the minimal line segment containing all $x+\theta_i\xi$. Our method comes tantalizingly close to complete success. For low order diagrams, we are able to maintain the desired control over the numbers of subtractions. We are unable, however, to establish convergence of the Feynman-parameter integral for graphs of arbitrary complexity without making addition-

a)Research supported in part by National Science Foundation Grant No. PHY 74-21778 A02.
b)Research supported in part by National Science Foundation Grant No. MP 574-05783 A01.

 0022-2488/78/1909-1859$1.00

al subtractions. It is to be hoped that future investigations will yield the key to pushing the full program through to completion.

This article is organized as follows. In Sec. 2 the formal outlines of our subtraction scheme are developed in conjunction with a study of certain one-loop, two-loop and many-loop graphs. Later in Sec. 2, a precise formulation of the subtraction procedure is presented, and the corresponding convergence theorem is stated The proof of the theorem is given in Sec. 3. We conclude, in Sec. 4, with a brief discussion of light-cone expansions involving our normal products.

2. DEFINITION OF N^* PRODUCTS

We wish to define vacuum expectation values

$$\langle 0|T[N^*(\prod_{i=1}^{n}\phi_i(x+\theta_i\xi))\prod_{j=1}^{m}\tilde{\psi}_j(p_j)]|0\rangle, \tag{2.1}$$

where ϕ_i and ψ_j are interacting fields, and (2.1) is to be a tempered distribution in x, $p_1,\cdots,p_m\in\mathbb{R}^4$ which is a continuous function of $\xi\in\mathbb{R}$ and $\theta_i\in\mathbb{R}$. Note that, by taking various θ_i equal, we may define light-cone finite products of several currents.) Let $\hat{\Gamma}$ be a Feynman graph contributing to (2.1), with Γ the graph obtained by identifying all vertices $x+\theta_i\xi$, and let $\Delta_l(q)$ be the propagator for a line l of Γ, $\{k_i\}$ be loop momenta for Γ, and

$$q_l=\sum_i a_{li}k_i+\sum_j b_{lj}p_j$$

be line momenta for Γ. Then the unrenormalized amplitude for $\hat{\Gamma}$ is formally

$$\exp[ix\cdot(\sum_j p_j)]\int\prod_i dk_i\prod_l\Delta_l(q_l)\prod_{i=1}^{n}\exp(i\theta_i\xi\cdot q_{l_i}), \tag{2.2}$$

where l_i is the line of Γ which is incident on $x+\theta_i\xi$ in $\hat{\Gamma}$. We must introduce subtractions into (2.2).

A. A one loop example

Let A be a scalar field of mass m with A^4 interaction, and consider

$$\langle 0|T[A(x+\xi)A(x-\xi)\tilde{A}(p_1)\tilde{A}(p_2)]|0\rangle. \tag{2.3}$$

If Γ is the graph of Fig. 1 (the vertex V represents the coalescence of $x+\xi$ and $x-\xi$), the integral in (2.2) becomes formally

$$\int d^4k\exp(2ik\cdot\xi)\,[(k+\tfrac{1}{2}p)^2-m^2]^{-1}[(k-\tfrac{1}{2}p)^2-m^2]^{-1}$$

$$=\int_0^\infty\int_0^\infty d\alpha_1 d\alpha_2(\alpha_1+\alpha_2)^{-2}\exp i\left(\frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2}p^2\right.$$

$$\left.+\frac{\alpha_1-\alpha_2}{\alpha_1+\alpha_2}\,p\cdot\xi-\frac{\xi^2}{\alpha_1+\alpha_2}-(\alpha_1+\alpha_2)m^2\right). \tag{2.4}$$

For $\xi^2\neq 0$, (2.4) is regularized by the ξ^2 term in the exponential, but the integral diverges (at $\alpha_1=\alpha_2=0$) for $\xi^2=0$. Zimmermann's short distance finite normal product $N_2[A(x+\xi)A(x-\xi)]$ of Ref. 2 is defined for this diagram by subtracting from the integrand its value at $p=0$; using

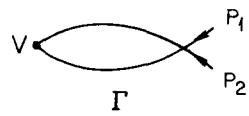


FIG. 1. One loop graph.

$$F(p)-F(0)=\int_0^1 d\tau\,\frac{d}{d\tau}F(\tau p)$$

we obtain

$$i\int d\alpha_1 d\alpha_2(\alpha_1+\alpha_2)^{-3}[2\tau\alpha_1\alpha_2 p^2+(\alpha_1-\alpha_2)p\cdot\xi]$$

$$\times\exp i\{\cdots\}. \tag{2.5}$$

The first term of (2.5) is convergent for all ξ, but the second still diverges for $\xi^2=0$ (although of course it is finite—in fact, zero,—when $\xi\to 0$ in nonlightlike directions). The example suggests a method for defining light-cone finite products: Working in terms of invariants p^2, $p\cdot\xi$, and ξ^2, one subtracts at $p^2=0$; here this gives only the first term of (2.5). Of course, in more complicated graphs, similar subtractions will be necessary for subgraphs as well.

B. Preliminary definition of N^* products

We now give a preliminary version of the renormalization operation needed to define N^* products; a final, precise version will be given below. Let Γ be a 1PI Feynman graph for which each line l has propagator

$$\Delta_l(q)=Z_l(q)\exp(i\xi_l\cdot q)(q^2-m_l^2+i0)^{-1},$$

with Z_l an invariant polynomial of degree ρ_l; we write $n(\Gamma)$, $N(\Gamma)$, and $m(\Gamma)$ for the number of loops, lines, and vertices, respectively. Suppose further that to each 1PI $\gamma\subset\Gamma$ we have assigned a subtraction index $\delta(\gamma)$. Let γ be a 1PI subgraph of Γ, let $\{k_i\}$ and $\{q_l\}$ be loop and line momenta for γ, and let $\mathcal{F}$ be forest of 1PI subgraphs of γ, with $\lambda_1,\cdots\lambda_R$ the maximal proper subgraphs of γ in $\mathcal{F}$ and $\bar{\gamma}=\gamma/\lambda_1\lambda_2\cdots\lambda_R$. Then define recursively

$$\bar{\mathcal{G}}_{\mathcal{F}}^{\gamma}=\int d^4k_1\cdots d^4k_{n(\bar{\gamma})}\prod_{i=1}^{R}\mathcal{G}_{\mathcal{F}}^{\lambda_i}(\tilde{\mathfrak{p}})\prod_{l\in\bar{\gamma}}[Z_l(q_l)$$

$$\times\exp i\,\xi_l\cdot q_l\exp i\,\alpha_l(q_{l0}^2-(1-i\epsilon)\vec{q}_l^{\,2}-m_l^2(1-i\epsilon))], \tag{2.6}$$

$$\mathcal{G}_{\mathcal{F}}^{\gamma}=\begin{cases}\bar{\mathcal{G}}_{\mathcal{F}}^{\gamma}, & \text{if } \gamma\notin\mathcal{F},\\ -t^{\delta(\gamma)}\bar{\mathcal{G}}_{\mathcal{F}}^{\gamma}, & \text{if } \gamma\in\mathcal{F}.\end{cases} \tag{2.7}$$

Here

$$t^{\delta}\prod_r p_{i_r}^{\mu_r}F(p_i\cdot p_j,\ p_i\cdot\xi_l,\ \xi_l\cdot\xi_m)$$

$$=\sum_{j=0}^{\delta}\left(\frac{d}{d\tau}\right)^j\prod_r(\tau p_{i_r}^{\mu_r})F(\tau^2 p_i\cdot p_j,\ p_i\cdot\xi_l,\ \xi_l\cdot\xi_m)\big|_{\tau=0} \tag{2.8}$$

for $\delta > 0$, $t^\delta = 0$ for $\delta < 0$. In (2.6) if j is a vertex of λ_i, $\tilde{p}_j$ denotes the total momentum entering this vertex from external and internal lines of γ. Finally

$$\mathcal{R}^*(p, \xi) = \lim_{\epsilon \to 0} \int_0^\infty \cdots \int_0^\infty \prod_{l \in \Gamma} d\alpha_l \sum_{\mathcal{J}} U_j^\Gamma .$$

Remark: (a) t^δ is in fact not well defined since its operand is not a covariant function for $\epsilon > 0$, and even if it were, it could not in general be written uniquely as a function of the invariants. Our refined definition below, however, gives an explicit formula for U_j^γ which avoids the difficulty.

(b) Ideally we would like to take $\delta(\gamma) = d(\gamma)$ where $d(\gamma) = 4n(\gamma) - 2N(\gamma) + \sum_{l \in \gamma} \rho_l$ is the superficial divergence. This choice, however, does not appear to give convergence; we try to illuminate the problems in the next section.

C. Two-loop and multi-loop examples

Consider first the contribution to (2.3) from the graph of Fig. 2. Here the forest formula reduces to iterated $(1-t)$ operations for γ and Γ; since γ does not involve V and $d(\gamma) = 0$ it is natural to choose $\delta(\gamma) = 0$. Then $\mathcal{R}^*$ has the form (ignoring ϵ dependence)

$$\int \prod d\alpha_l (1 - t^{\delta(\Gamma)}) [f_1(\alpha) + f_2(\alpha)\xi^2 + \sum_i f_{3i}(\alpha) p_i \cdot \xi + \sum_{i,j} f_{4ij}(\alpha) p_i \cdot p_j] \exp i \left[\sum_{i,j} V_{ij}(\alpha) p_i \cdot p_j + \sum_i Y_i(\alpha) \xi \cdot p_i + W(\alpha)\xi^2 - \sum_l \alpha_l m_l^2 \right]. \quad (2.9)$$

The α_1, α_2 subintegration is convergent because of the γ subtraction, but the overall integration of the $f_1, \cdots, f_4$ terms is respectively logarithmically, linearly, and logarithmically divergent, and convergent. If we take $\delta(\Gamma) = 0$ [note $d(\Gamma) = 0$], the first and third terms are rendered convergent (V is homogeneous of degree 1); the f_2 term appears to be logarithmically divergent but for $\xi^2 \neq 0$ it is regulated by the $W(\alpha)\xi^2$ term in the exponential, and for $\xi^2 \to 0$ it vanishes [compare $\int_0 \xi^2 t^{n-1} \exp(-\xi^2/t) dt$]. Thus subtractions of minimal degree suffice to give a light-cone finite normal product here.

Observe, however, that the subtraction of the subgraph γ has led to $p \cdot \xi$ factors in (2.9). Our subtraction procedure (2.8) ignores these, thus oversubtracting as far as momentum power counting is concerned. In general, oversubtraction at one level

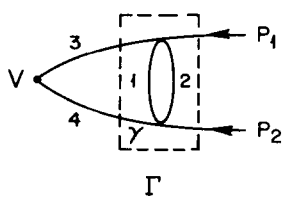


FIG. 2. Two loop graph.

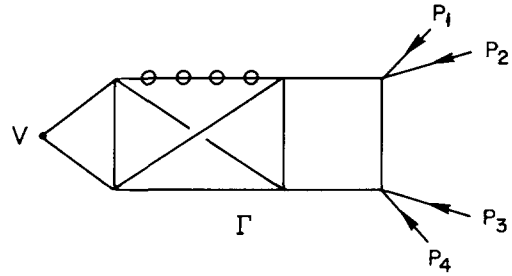


FIG. 3. Many loop graph.

necessitates higher subtraction degrees at higher levels.

This can happen to us. The contribution of Fig. 3 to

$$\langle 0 | T[A(x+\xi)A\,x-\xi)] \prod_{i=1}^4 \tilde{A}(p_i) | 0 \rangle$$

does not appear to yield a convergent amplitude if $\delta(\Gamma)$ is chosen to be minimal (i.e., negative). Another way to view this difficulty is as follows: the α-space integral for $\mathcal{R}^*$ in this case is similar to (2.9), but the function corresponding to $W(\alpha)$ vanishes in the interior of the integration region, and hence $\exp iW(\alpha)\xi^2$ no longer regulates the term corresponding to $f_2(\alpha)\xi^2$. We conclude that we cannot systematically take $\delta(\gamma) = d(\gamma)$, in fact, we will need subtraction degrees which increase without bound with the order in perturbation theory.

D. Final definition of N^* products

We will complete the definition of the $\mathcal{R}^*$ operation by giving an explicit formula similar to that of Appelquist[3] and Bergere–Zuber[4] for the BPH $\mathcal{R}$ operation. Take Γ as in Sec. 2B, with $\gamma \subset \Gamma$ a 1PI subgraph.

We introduce the standard combinatoric functions for γ, i.e., fixing a vertex k and letting i, j denote vertices and s, t lines of γ, we define

$$U^\gamma(\alpha) = \sum_T \prod_{l \notin T} \alpha_l ,$$

$$V_{ij}^\gamma(\alpha) = (U^\gamma)^{-1} \sum_{T_2} \prod_{l \notin T_2} \alpha_l ,$$

$$Y_{is}^\gamma(\alpha) = (U^\gamma)^{-1} \sum_T (\pm) \prod \alpha_l$$

$$X_{st}^\gamma(\alpha) = (U^\gamma)^{-1} \sum_{T^*} (\pm) \prod_{l \notin T^*} \alpha_l ,$$

the sums running respectively over all trees T of γ, all two trees T_2 of γ disconnecting i and j from k, all trees T of γ for which the path in T from k to i passes through s in the same (+) or opposite (−) direction as s, and all sets T^* formed by adding one line to a tree, such that the circuit in T^* contains both s and t, oriented coherently (+) or incoherently (−). If $\mathcal{J}$ is a forest for γ, we follow Appelquist by introducing variables τ_λ, $\lambda \in \mathcal{J}$, and writing $\bar{\alpha}_l = (\prod_{l \in \lambda \in \mathcal{J}} \tau_\lambda^2)\alpha_l$,

$\bar{U}(\alpha,\tau)=\prod_{\lambda\in\mathcal{J}}\tau_\lambda^{-2n(\lambda)}U(\bar{\alpha})$, and $\bar{T}(\alpha,\tau)=T(\bar{\alpha})$ for $T=V$, Y, or X. Finally,

$$\bar{W}^\lambda=\tau_\lambda^2\bar{X}^\lambda-\sum_\mu\bar{X}^\mu$$

for any $\lambda\in\mathcal{J}\cup\{\gamma\}$; the sum is over the maximal proper subgraphs μ of λ with $\mu\in\mathcal{J}$. Then we define

$$\begin{aligned}\mathcal{Y}_{\mathcal{J}}^{\gamma}=&\prod_{\lambda\in\mathcal{J}}(-t_{\tau_\lambda}^{\delta(\lambda)})\prod_{l\in\gamma}Z_l\left(\frac{1}{i}\frac{\partial}{\partial r_l}\right)\bar{U}(\alpha,\tau)^{-2}\\&\times\exp i\{p^T\bar{V}p+p^T\bar{Y}(\bar{r}+\xi)-\tfrac{1}{4}[\bar{r}\bar{X}^\gamma\bar{r}+\sum_{\lambda\in\{\gamma\}\cup\mathcal{J}}\\&\times(2\bar{r}\bar{W}^\lambda\xi\prod_{\mu\supset\lambda}\tau_\mu^{-1}+\xi\bar{W}^\lambda\xi)]\}\Big|_{\tau_\lambda=1}.\end{aligned}\tag{2.10}$$

Here t_τ^δ extracts the Taylor series in τ, centered at 0, to order δ (with $t_\tau^\delta=0$ if $\delta<0$), $\bar{r}_l=r_l\prod_{\lambda\ni l}\tau_\lambda$, $p^T\bar{V}p=\sum\bar{V}_{ij}p_i\cdot p_j$, etc., and $p_i\cdot p_j=p_{i0}p_{j0}-(1-i\epsilon)\mathbf{p}_i\cdot\mathbf{p}_j$, $p_i\cdot r_l=p_{i0}r_{l0}-\mathbf{p}_i\cdot\mathbf{r}_l$, $r_l\cdot r_m=r_{l0}r_{m0}-(1-i\epsilon)^{-1}\mathbf{r}_l\cdot\mathbf{r}_m$, and similarly with r_l, r_m replaced by ξ_l, ξ_m. (For us the condition $\mu\supset\lambda$ includes the possibility $\mu=\lambda$). We remark that (2.10), with $\xi_l=0$ for all l, is the standard formula[3,4] for the $\mathcal{R}$ operation restricted to a single forest.

We want to justify (2.10) by showing its relation to (2.6)–(2.8). Suppose then that either γ is minimal in $\mathcal{J}$ or that we have justified (2.10) for all subgraphs of γ. If we write

$$Z_l(q)=Z_l\left(\frac{1}{i}\frac{\partial}{\partial r_l}\right)\exp(iq\cdot r_l)\big|_{r_l=0}$$

the integral of (2.6) becomes a Gaussian; evaluation of this integral yields (2.10) with no t_{τ_γ} operator with $\tau_\gamma=1$. By (2.7) this completes the case $\gamma\in\mathcal{J}$. [We omit a detailed derivation of this result. The only difficulty is in evaluating certain combinations of matrix products which arise on completing the square in the exponential.

One observes, however, that these same expressions arise in a similar recursive evaluation of the $\mathcal{R}$ operation, and that in that case the amplitude may also be evaluated in one step by a rescaling of integration variables[4]; comparison of these formulas for the $\mathcal{R}$ operation then yields an evaluation of the needed expressions.]

Finally, if $\gamma\in\mathcal{J}$, we must apply the Taylor operator of (2.7), (2.8). It is quite complicated to apply (2.8) directly since factors $p_i\cdot p_j$ and $p_i\cdot\xi_l$ can both be generated by the spin terms; instead, to simplify our scheme, we scale with τ_γ the momentum variables in both the $p^T\bar{V}p$ and $p^T\bar{Y}\bar{r}$ (but not $p^T\bar{Y}\xi$) terms, leading immediately to (2.10). The effect of this choice is that for certain terms, involving overall $p\cdot\xi$ factors resulting from the r derivatives, we are undersubtracting in comparison with (2.8). However, the increased subtraction degrees we use suffice to give convergence despite this undersubtraction.

Definition 2.1: The *-renormalized amplitude for Γ is

$$\mathcal{R}_\Gamma^*(p;\xi)=\lim_{\epsilon\to0}\int_0^\infty\cdots\int_0^\infty\prod d\alpha_l\sum_{\mathcal{J}}\mathcal{Y}_{\mathcal{J}}^{\Gamma}(p_i;\xi;\alpha)\tag{2.11}$$

with $\mathcal{Y}_{\mathcal{J}}^{\Gamma}$ given by (2.10). Furthermore, (2.1) is defined by expanding in graphs and applying (2.11), with $\xi_l=\theta_i\xi$ if l is incident on $x+\theta_i\xi$ in $\hat{\Gamma}$; $\xi_l=0$ otherwise.

E. Statement of results

We want to investigate Definition 2.1 for the case (2.2) arising in the N^* product of several currents as well as for general ξ. Thus let V be a fixed vertex of Γ and $\{\mathcal{L}_1,\cdots\mathcal{L}_k\}$ a partition of the lines incident on V; let $\mathcal{H}$ denote the set of 1PI $\gamma\subset\Gamma$ such that γ intersects at most one $\mathcal{L}_j$, and suppose for convenience that all lines incident on V are oriented into V. Then we consider the hypothesis: (H) $\xi_l=\xi_j\in\mathbb{R}^4$ if $l\in\mathcal{L}_j(j=1,\cdots,k)$; $\xi_l=0$ otherwise.

The justification of Def. 2.1, proved in the next section, is:

Theorem 2.2: There exists a choice of subtraction degrees $\delta(\gamma)$ such that the integral in (2.11) is absolutely convergent for $\epsilon>0$, and defines a tempered distribution in $p_1,\cdots,p_n$ which is continuous in ϵ, ξ for $\epsilon\geq0$ and $\xi\in\mathbb{R}^{4N}$; if (H) holds we may take minimal subtraction degrees ($\delta(\gamma)=d(\gamma)$) for $\gamma\in\mathcal{H}$, and continuity then holds in $\epsilon>0$ and $\xi\in\mathbb{R}^{4k}$.

Remark 2.3: (a) In fact we give an explicit recursive formula for calculating $\delta(\gamma)$ but do not claim that the result is optimal.

(b) The subtractions for Γ itself are the terms of (2.11) with $\Gamma\in\mathcal{J}$. Now consider (2.2): from (2.10), a counterterm is a linear combination of integrals

$$\int d\alpha\,P(p_1,\cdots,p_m;\xi;\alpha)F(\xi^2)\exp i\sum_{j=1}^m p_j\cdot(x+c_j\xi),\tag{2.12}$$

where P is a polynomial in p. The Fourier transform in p of the integrand in (2.12) is supported at $x_j=x+c_j\xi$; it follows from the definition of Y_{is} that c_j is a convex combination of $\theta_1,\cdots,\theta_n$, and hence the counterterm is (formally) supported on the minimal line segment containing all points $x+\theta_i\xi$. Again, at least formally, this gives the same localization for $N^*(\prod\phi_i(x+\theta_i\xi))$. Some further discussion is given in Sec. 4.

3. CONVERGENCE OF N^* PRODUCTS

We will divide our proof of Theorem 2.2 into several sections for clarity.

A. Decomposition of integration region

We follow Breitenlohner and Maison.[5] Consider a triple $(\mathcal{C},\mathcal{B},\sigma)$, where $\mathcal{C}$ is a maximal forest of 1PI subgraphs of Γ, $\mathcal{B}\subset\mathcal{C}\setminus\{\Gamma\}$, and σ is a map assigning to each $\gamma\in\mathcal{C}$ a line $\sigma(\gamma)$ in $\bar{\gamma}\equiv\gamma/\lambda_1\cdots,\lambda_p$, with $\{\lambda_i\}$ the maximal subgraphs of γ in $\mathcal{C}$. The region $\mathcal{D}(\mathcal{C},\mathcal{B},\sigma)\subset\{\alpha\,|\,\alpha_l\geq0\}$ is defined by

$$\alpha_l=\begin{cases}\prod_{\gamma\subset\lambda\in J} t_\lambda, & \text{if } l=\sigma(\gamma),\\ \\ \beta_l\alpha_{\sigma(\gamma)}, & \text{if } l\in\bar\gamma,\ l\neq\sigma(\gamma),\end{cases} \tag{3.1}$$

and the restrictions

$$t_\Gamma\geqslant 0;\quad 1\geqslant t_\gamma\geqslant 0,\ \gamma\notin\beta,\ \gamma\neq\Gamma;\ t_\gamma\geqslant 1,\ \gamma\in\beta;1\geqslant\beta_l\geqslant 0. \tag{3.2}$$

We occasionally write $\beta_l\equiv 1$ if $l\in\sigma(C)$. Exactly as in Ref. 5, we show that (2.11) becomes

$$\lim_{\epsilon\to 0}\sum_{(C,\beta,\sigma)}\int_{D(C,\beta,\sigma)}\Pi d\alpha_l\, Y^\Gamma_{C,\beta}, \tag{3.3}$$

where Y^Γ_C is given by formula (2.10) for Y^Γ_C but with $-t^{\delta(\lambda)}_{\tau_\lambda}$ replaced by $(1-t^{\delta(\lambda)}_{\tau_\lambda})$ for $\lambda\notin\beta$.

B. Evaluation of γ and τ derivatives

We first observe that the combinatoric functions $\bar U(\equiv\bar U^\Gamma)$, $\bar V$, etc. satisfy

(*P*1) $\bar W^\mu$ is independent of τ_λ unless $\lambda\subset\mu$; $\bar W^\mu_{lm}=0$ unless lines l and m are in μ; if (*H*) holds, $\sum_m \bar W^\gamma_{lm}\xi_m=0$ for $\gamma\in H$.

(*P*2) for any $\lambda\in C$, we have homogeneity in the variables τ_λ^{-2} and $\{\alpha_l\,|\,l\in\lambda\}$, with deg $\bar U=n(\lambda)$, deg $\bar X=$deg $\bar Y=$deg $\bar V=$deg $\bar W^\mu=0$ unless $\mu\subset\lambda$; deg $\bar W^\mu=-1$ if $\mu\subset\lambda$.

(*P*3) $\bar Y$, $\bar V$ and $(\prod_{r,s\in\lambda}\tau_\lambda^2)\bar X_{rs}$ have the form $P(\alpha,\tau^2)\bar U^{-1}$, and $\bar W^\lambda$ the form $P(\alpha,\tau^2)\,(\bar U^\lambda\prod\bar U^{\lambda_i})^{-1}$, with P a polynomial and $\{\lambda_i\}$ the maximal subgraphs of λ in C. [On first inspection it appears that $\bar W^\lambda$ might contain factors τ_μ^{-1} for $\mu\subset\lambda$, but this possibility may be eliminated using $\bar U^\lambda|_{\tau_\mu=0}=\bar U^\mu\bar U^{\lambda/\mu}$ (Ref. 6), etc.]

(*P*4) $\bar W^\lambda_{st}$ is a positive semidefinite matrix.

To evaluate the spin terms it is convenient to use $\partial/\partial r_l=\prod_{\lambda\ni l}\tau_\lambda\partial/\partial\bar r_l$. Differentiating and setting $\bar r=0$, we see that ΠZ_l is replaced by a sum of terms

$$Q_0(p)\prod_{\gamma\in C}\tau_\gamma^{k(\gamma)}\mathcal{H}(\bar Y)\prod_{\gamma\in C\setminus H}\mathcal{H}_\gamma(\bar W^\gamma\xi)\prod_{l,m}\Big(\bar X_{lm}\prod_{\lambda\ni l,m}\tau_\lambda^2\Big)^{a_{lm}}, \tag{3.4}$$

where Q_0 is a polynomial and $\mathcal{H}(\bar Y)$, $\mathcal{H}_\gamma(\bar W^\gamma\xi)$ are monomials in $\{\bar Y_{is}\}$ and $\{(\bar W^\gamma\xi)_s=\sum_t\bar W^\gamma_{st}\xi_t\}$, respectively. If (3.4) arose from $\Pi(\partial/\partial\bar r_l)^{\rho'_l}$ (with $\rho'_l\leqslant\rho_l$) and if $c_\gamma=\deg\mathcal{H}_\gamma$, then

$$k(\gamma)\equiv\sum_{l\in\gamma}\rho'_l-2\sum_{l,m\in\gamma}a_{lm}-\sum_{\substack{\mu\subset\gamma\\ \mu\notin H}}c_\mu\geqslant 0; \tag{3.5}$$

note that we have used (P1) to restrict to $\gamma\notin H$ (taking $H=\emptyset$ by convention when (H) is not assumed), and have grouped the factors of τ as suggested by (P3). The only odd powers of τ now occur in $\tau_\gamma^{k(\gamma)}$, so we use

$$(-t^\delta_\tau)\tau^k F(\tau^2)\big|_{\tau=1}=\begin{cases}0, & \text{if } \delta<k,\\ \\ -t^{[(\delta-k)/2]}_\eta F(\eta)\big|_{\eta=1}, & \text{if } \delta\geqslant k;\end{cases}$$

here $[n]$ is $n/2$ [resp. $(n-1)/2$] if n is even (resp. odd). Thus, if we insert (3.4) into (3.3), the result vanishes unless $\delta(\lambda)\geqslant k(\lambda)$ for all $\lambda\in\beta$, and otherwise is a sum of terms

$$\int_{D(C,\beta,\sigma)}\Pi d\alpha\quad\int\prod_{\substack{\lambda\in C\setminus\beta\\ \delta(\lambda)\geqslant k(\lambda)}}d\mu(\eta_\lambda)\prod_{\lambda\in C}\Big(\frac{\partial}{\partial\eta_\lambda}\Big)^{e_\lambda}Q_0(p)\mathcal{H}(\hat Y)$$
$$\times\prod_{\gamma\notin H}\mathcal{H}_\gamma(\hat W^\gamma\xi)\prod_{l,m}(\hat X_{lm}\prod_{\lambda\ni l,m}\eta_\lambda)^{a_{lm}}\hat U^{-2} \tag{3.6}$$
$$\times\exp i\,[p^T\hat V p+p^T\hat Y\xi-\tfrac14\sum_\lambda\xi\hat W^\lambda\xi-\sum\alpha_l m_l^2(1-i\epsilon)]\big|_{\substack{\eta_\lambda=0,\lambda\in\beta\\ \eta_\lambda=1,\lambda\notin\beta\\ \delta(\lambda)<k(\lambda)}}$$

with $\hat U(\alpha,\eta)=\bar U(\alpha,\tau)\big|_{\tau^2_\lambda=\eta_\lambda}$ etc.;

$$\int d\mu(\eta_\lambda)=\{[\delta(\lambda)-k(\lambda)]!\}^{-1}\int_0^1[1-\eta(\lambda)]^{\delta(\lambda)-k(\lambda)}d\eta_\lambda,$$

and

$$e_\lambda\begin{cases}=[(\delta(\lambda)-k(\lambda))/2]+1, & \text{if } \delta(\lambda)\geqslant k(\lambda),\ \lambda\notin\beta,\\ =0,\ \delta(\lambda)<k(\lambda),\ \lambda\notin\beta,\\ \leqslant[(\delta(\lambda)-k(\lambda))/2],\ \lambda\in\beta.\end{cases} \tag{3.7}$$

We next carry out the η derivatives. Then, aside from the exponential, (3.6) becomes a sum of terms

$$Q'_0(p)\mathcal{D}(\hat U^{-2})\mathcal{H}_1(\mathcal{D}\hat Y)\mathcal{H}_2(\mathcal{D}\hat V)\prod_{\gamma\notin H}[\mathcal{H}''_\gamma(\mathcal{D}W^\gamma\xi)\mathcal{H}'_\gamma(\mathcal{D}\xi^T\hat W^\gamma\xi)]$$
$$\times\prod_{l,m}(\mathcal{D}\hat X_{lm}\Pi\eta)^{a_{lm}}. \tag{3.8}$$

Here (with some abuse of notation) $\mathcal{D}(\hat U)^{-2}$ denotes some product of $(\partial/\partial\eta)$ operators on $\hat U^{-2}$, $\mathcal{H}_1(\mathcal{D}\hat Y)$ a monomial in the η derivatives of $\{Y_{is}\}$, etc. We still have $c_\gamma=\deg\mathcal{H}''_\gamma$ and, if $\deg\mathcal{H}'_\gamma=c'_\gamma$, then from (P1)

$$\sum_{\gamma\subset\lambda}c'_\gamma\leqslant\sum_{\substack{\gamma\subset\lambda\\ \neq}}e_\gamma. \tag{3.9}$$

Further, (P3) implies that (3.8) has the form

$$\frac{Q(p,\xi)P(\alpha,\eta)}{\Pi(\hat U^\lambda)\hat p^{(\lambda)}} \tag{3.10}$$

with P a polynomial, and (P2) shows that (3.10) is homogeneous in η_λ^{-1}, $\{\alpha_l\,|\,l\in\lambda\}$ of degree

$$-2n(\lambda)-\sum_{\mu\subset\lambda}(c_\mu+c'_\mu)-\sum_{l,m\in\lambda}a_{lm}+e_\lambda. \tag{3.11}$$

[Note that each $\partial/\partial\eta_\lambda$ contributes $+1$ to (3.11).]

Now set $\eta_\gamma=0$, $\gamma\in\beta$. For $\mu\in C$ we let $\tilde\mu$ denote μ modulo its maximal proper subgraphs in in β; then

$$\hat U^\lambda\big|_{\eta_\gamma=0,\gamma\in\beta}=\hat U^{\tilde\lambda}\prod_{\substack{\gamma\in\beta\\ \gamma\subset\lambda\\ \neq}}\hat U^{\tilde\gamma},$$

so that (3.10) becomes

$$\frac{Q(p,\xi)R(\alpha,\eta)}{\Pi_\lambda\hat U^{\tilde\lambda})^{q(\lambda)}}, \tag{3.12}$$

where $R=p|_{\eta_\gamma=0,\,\gamma\in\beta}$. Moreover, (3.12) has degree (3.11) in

η_λ^{-1}, $\{\alpha_l \,|\, l\in\lambda\}$ for $\lambda\notin\beta$ and in $\{\alpha_l \,|\, l\in\lambda\}$ for $\lambda\in\beta$.

C. Change of integration variables

Introduce the variables $\{t_\lambda, \beta_l\}$ by (3.1). By standard arguments[7]

$$\hat{U}^{\tilde{\mu}} = \prod_{\nu\in C} t_\nu^{j(\nu)} E^{\tilde{\mu}}(\alpha, \tau),$$

where $E^{\tilde{\mu}}\geq 1$, and $j(\nu)=n(\nu\cap\tilde{\mu})$ is precisely the degree of homogeneity of $\hat{U}^{\tilde{\mu}}$ in $\{\alpha_l \,|\, l\in\nu\}$, if $\nu\in\beta$, or in η_ν^{-1}, $\{\alpha_l \,|\, l\in\nu\}$ if $\nu\notin\beta$. Thus the homogeneity of (3.12) implies that when multiplied by the Jacobian $\prod t_\lambda^{N(\lambda)-1}$ it becomes a sum of terms of the form

$$\prod_{\lambda\in C} t_\lambda^{b_\lambda-1} F(t, \beta, \eta) Q(p, \xi), \tag{3.13}$$

where $|F|\leq 1$ and

$$b_\lambda \geq N(\lambda) - 2n(\lambda) - \sum_{\mu\subset\lambda}(c_\mu + c'_\mu) - \sum_{l,m\in\lambda} a_{lm} + e_\lambda, \tag{3.14}$$

with equality in (3.14) if $\lambda\in\beta$.

In order to be able to do the t_Γ integration explicitly we write

$$\sum_{\gamma\in C} \xi\hat{W}^\gamma\xi\,\big|_{\alpha_l=\beta_l\Pi t_\lambda} \equiv -[t_\Gamma(1-i\epsilon)]^{-1}A(t, \beta, \eta, \xi)^2.$$

By (P2), A is independent of t_Γ; recall that $\xi_s\cdot\xi_t = (1-i\epsilon)^{-1}(\xi_s^0\xi_t^0(1-i\epsilon) - \vec{\xi}_s\cdot\vec{\xi}_t)$, so by (P4) $\operatorname{Im} A^2\geq 0$ and we take $\operatorname{Im} A>0$. Then

$$\exp\left(-\frac{i}{4}\sum_\gamma \xi\hat{W}^\gamma\xi\right) = \int_{-\infty}^{\infty} d\omega\,[-it_\Gamma(1-i\epsilon)/\pi]^{1/2} \times\exp\Big(-i\{t_\Gamma(1-i\epsilon)\omega^2 - \omega A\}\Big). \tag{3.15}$$

Since (3.15) is absolutely convergent, it suffices to prove absolute convergence and ϵ, ξ continuity of (3.6) with the substitution (3.15), i.e., using (3.13), of

$$\int d\omega\,\Pi d\mu(\eta_\lambda)\,\Pi d\beta_l \Pi t_\lambda^{b_\lambda-1}\,dt_\lambda\, t_\Gamma^{1/2} Q(p, \xi) F(t, \beta, \eta) \times\exp\{i(p^T\hat{Y}\xi + \omega A)\exp\{it_\Gamma[p^T\tilde{V}p - M(1-i\epsilon)]\}. \tag{3.16}$$

Here $\tilde{V}=t_\Gamma^{-1}\hat{V}$ is independent of t_Γ,

$$M=\omega^2+\sum_l m_l^2\beta_l \prod_{l\in\lambda\neq\Gamma} t_\lambda,$$

and the integration region for β, t is (3.2).

D. Estimates for convergence and continuity

To verify absolute convergence of (3.16) we note that F and the first exponential are bounded by 1, the second exponential by $\exp(-\epsilon t_\Gamma M)$. When these bounds are inserted in (3.16), the resulting t_Γ integral can be done explicitly if $b_\Gamma>0$ to give $\Gamma(b_\Gamma+\frac{1}{2})(\epsilon M)^{-(b_\Gamma+1/2)}$; thus we need only the estimate (verified below)

$$\Pi t_\lambda^{b_\lambda} M^{-(b_\Gamma+1/2)} \leq K(m_{\sigma(\Gamma)}^2+\omega^2)^{-a} \prod_{\lambda\neq\Gamma} t_\lambda^{a_\lambda}, \tag{3.17}$$

where $a>\frac{1}{2}$, $a_\lambda<0$ if $\lambda\in\beta$ and $a_\lambda>0$ if $\lambda\notin\beta$. On the other hand, we may evaluate the t_Γ integral in (3.16) directly to give

$$\Gamma(b_\Gamma+1/2)\,M^{-(b_\Gamma+1/2)}\,[p^T\tilde{V}p/M - (1-i\epsilon)]^{-(b_\Gamma+1/2)}.$$

It is easy to verify that $\tilde{V}/M$ is uniformly bounded; hence[7] for any Schwartz test function $\psi(p)$,

$$\int\psi(p)(p^T\tilde{V}p/M - (1-i\epsilon)]^{-(b_\Gamma+1/2)}dp$$

is a continuous bounded function of t, β, η and ϵ for $\epsilon\geq 0$. Using Lebesque dominated convergence and estimating as above we see that as a distribution (3.16) is continuous for $\xi\in\mathbb{R}^{4N}$ and $\epsilon\geq 0$ [or, under (H), for $\xi\in\mathbb{R}^{4k}$, $\epsilon\geq 0$].

It remains to verify that (3.17) and $b_\Gamma>0$ can be ensured by suitable choice of $\delta(\gamma)$. Write $\delta(\gamma)=d(\gamma)+D(\gamma)$, and take $D(\gamma)=0$ if (H) holds and $\gamma\in\mathcal{H}$; otherwise, choose $\delta(\gamma)$ recursively so that for any forest $\mathcal{F}$ of proper 1PI subgraphs of γ,

$$D(\gamma)\geq\sum_{l\in\gamma}\rho_l + \sum_{\lambda\in\mathcal{F}}[\delta(\lambda)+D(\lambda)+1]. \tag{3.18}$$

Then if $\gamma\in C$, $\gamma\notin\beta$ and $\mathcal{G}\subset\beta$ is a family of disjoint subgraphs of γ, (3.14), (3.7), (3.5), and (3.18) imply

$$2(b_\gamma - \sum_{\mathcal{G}} b_\lambda)\geq D(\gamma)+1 - \sum_{\gamma\in\mathcal{G}} D(\lambda) - \sum_{\mu\subset\gamma}(c_\mu+2c'_\mu)\geq 1 \tag{3.19}$$

(or ≥ 2 since the b's are the integers): note $\gamma\in\mathcal{H}$ implies $\mu\in\mathcal{H}$ for $\mu\subset\gamma$, hence $c_\mu=c'_\mu=0$.

In particular, with $\gamma=\Gamma$ and $\mathcal{G}=\emptyset$, (3.16) yields $b_\Gamma\geq 1$. For $\lambda\in\beta$, define f_λ inductively by

$$f_\lambda = \max(0,\, b_\lambda - \sum_{\mu\subset\lambda} f_\mu),$$

and note that for any $\gamma\in C$,

$$\sum_{\lambda\subset\lambda} f_\lambda \begin{cases} \geq b_\lambda, & \text{if } \gamma\in\beta, \\ = \sum_{\mu\in\mathcal{G}} b_\mu, & \text{if } \gamma\notin\beta, \end{cases} \tag{3.20}$$

where $\mathcal{G}\subset\beta$ is some family as in (3.19). Choose θ with $0<\theta<|\beta|^{-1}$, then

$$M^{(b_\Gamma+1/2)} \geq (m_{\sigma(\Gamma)}^2+\omega^2+\sum_{\lambda\in\beta} m_{\sigma(\lambda)}^2 \prod_{\gamma\supset\lambda} t_\gamma)^{(b_\Gamma+1/2)}$$

$$\geq (m_{\sigma(\Gamma)}^2+\omega^2)^{b_\Gamma+1/2-\sum_\beta(d_\lambda+\theta)} \prod_{\lambda\in\beta}(m_{\sigma(\lambda)}^2 \prod_{\gamma\supset\lambda} t_\gamma)^{d_\lambda+\theta}$$

$$= \tilde{K}(m_{\sigma(\Gamma)}^2+\omega^2)^{b_\Gamma+1/2-\sum_\beta(d_\lambda+\theta)} \prod_{\substack{\gamma\in\beta\\ \gamma\neq\Gamma}} t_\gamma^{\sum_{\lambda\subset\gamma}(d_\lambda+\theta)}.$$

Now (3.14) follows immediately, using (3.19) and (3.20). This completes the proof of Theorem 2.6.

Remark: The basic requirement on $\delta(\gamma)$ is that

(3.19) be satisfied; (3.18) can be modified to give somewhat smaller $\delta(\gamma)$ while maintaining (3.19).

4. DISCUSSION OF LIGHT-CONE EXPANSIONS

Having specified a graph-by-graph subtraction procedure for Green functions of light-cone finite, multilocal products of fields, we now consider the question of whether such field products can be used to construct a light-cone expansion of the form (1.1). That this can be done is guaranteed (at least formally) by the recursive nature of the subtractions, and the broad outlines of such a derivation will be presented below. A detailed, rigorous treatment would require considerable additional effort and is probably premature. At this point, higher priority should be given to the task of improving the subtraction scheme so as to control the large-momentum behavior of vertex functions in a "renormalizable" way.

As a caveat to future investigators in this field, it should be pointed out that in searching for a suitable definition of light-cone finite normal products one must always make provisiion (if only at the level of a plausibility argument) for an eventual light-cone expansion. This is because the latter, by relating normal products to ordinary products of fields, allows one to establish, almost immediately, the legitimacy of the normal products as localized, covariant operator fields.

As mentioned above, the crucial property of our subtraction scheme which leads to an expansion (1.1) is its recursive nature, expressed in Eq. (2.6). That formula must be understood in the following sense: $\overline{\mathcal{Y}}_{\mathcal{J}}^{\gamma}$ is to be written in the standard α-parametric form, namely (2.10) with γ omitted from the final product over $\lambda \in \mathcal{J}$, and $t^{\delta(\gamma)}$ is to be understood as a Taylor series in τ_γ. We see that $\mathcal{Y}_{\mathcal{J}}^{\gamma}$ will then have the general form

$$\mathcal{Y}_{\mathcal{J}}^{\gamma} = \exp[i\tilde{p}^T(\overline{Y}^{\gamma})_0 \xi][\sum_a M_a^{\gamma}(\tilde{p}, \xi) G_a^{\gamma}(\alpha, \xi^2)], \tag{4.1}$$

where $\{\tilde{p}_j\}$ is the set of external momenta of γ, $\xi_i = \theta_i \xi$ are assumed to be nonvanishing only for l incident on the normal-product vertex, M_a^{γ} is a monomial, and

$$(\overline{Y}^{\gamma})_0 = Y^{\gamma}\big|_{\tau_\lambda = 0,\, \lambda \in \mathcal{J}(\gamma)}.$$

Following Zimmermann[2] in his derivation of algebraic identities relating momentum-space integrands, we may iterate (2.6) to obtain

$$\mathcal{Y}_{\mathcal{J}}^{\Gamma} = \sum_a (\mathcal{Y}_{\mathcal{J}\setminus\mathcal{J}(\gamma)}^{\Gamma/\gamma})^a (\mathcal{Y}_{\mathcal{J}(\gamma)}^{\gamma})_a, \tag{4.2}$$

where γ is the smallest element of $\mathcal{J}$ containing the normal product vertex and $\mathcal{J}(\gamma)$ is $\mathcal{J}$ restricted to γ. The index a runs through the terms of (4.1), with the factor $M_a^{\gamma}(\tilde{p}, \xi)\exp[i\tilde{p}^T(\overline{Y}^{\gamma})_0\xi]$ incorporated in $(\mathcal{Y}_{\mathcal{J}\setminus\mathcal{J}(\gamma)}^{\Gamma/\gamma})_a$ and $G_a^{\gamma}(\alpha, \xi^2)$ included in $(\mathcal{Y}_{\mathcal{J}(\gamma)}^{\gamma})_a$. Integrating over α and summing over all Γ and $\mathcal{J}$, one obtains an identity for Green functions which one can write in the shorthand form

$$N_\delta^*(\prod_{i=1}^n \phi_i(x+\theta_i\xi) = \prod_{i=1}^n \phi_i(x+\theta_i\xi) + \sum_r \int d\eta\, F_r(\eta, \xi^2) \times N_{\delta r}^*(\prod_j \mathcal{D}_{r_j}\psi_{r_j}\ (x+\eta_{r_j}\theta_j\xi)) \tag{4.3}$$

where the derivative operator $\mathcal{D}_{r_j}$ may contain factors $\xi\cdot\partial/\partial x$ as well as $\partial/\partial x^\mu$, and the subtraction degrees, which may be greater than minimal.

Extracting powers of ξ^2 and $\ln\xi^2$ in $F_r(\eta, \xi^2)$, and applying the LSZ reduction formula to obtain an operator relation, one obtains a multilocal, light-cone expansion generalizing (1.1). Since, for any γ, $|\overline{Y}^{\gamma}| \leq 1$ [see Remark 2.3(b)], the support of $F_r(\eta, \xi^2)$ is contained in $|\eta| \leq 1$. The localization of $N^*(\prod_{i=1}^n \phi_i(x+\theta_i\xi))$ on the line segment connecting $x+\theta_{\min}\xi$ to $x+\theta_{\max}\xi$ then follows from the operator light-cone expansion by mathematical induction in the perturbative order.

ACKNOWLEDGMENT

One of us (E.R.S.) would like to thank Dr. Harry Woolf for hospitality at The Institute for Advanced Study, where part of this work was done.

[1] J. Lowenstein, Nucl. Phys. B **126**, 467—92 (1977).
[2] W. Zimmermann, in *Lectures on Elementary Particles and Quantum Field Theory*, edited by S. Deser, M. Grisaru, and H. Pendelton (MIT Press, Cambridge, 1971).
[3] T. Appelquist, Ann. Phys. **54**, 27—61 (1969).
[4] M. Bergere and J. Zuber, Commun. Math. Phys. **35**, 113—40 (1974).
[5] P. Breitenlohner and D. Maison, Commun. Math. Phys. **52**, 55—75 (1977).
[6] N. Nakanishi, *Graph Theory and Feynman Integrals* (Gordon and Breach, New York, 1971).
[7] E. Speer, *Generalized Feynman Amplitudes* (Princeton U. P., Princeton, N.J., 1969).

A time dependent generalization of the multiple scattering formalism[a)]

Joel I. Gersten

Department of Physics, City College of the City University of New York, New York, New York 10031
(Received 22 December 1977; revised manuscript received 10 March 1978)

The multiple scattering formalism is generalized to include the restricted class of time dependent potentials in which each scattering center is moving with an arbitrary, but constant, velocity.

The multiple scattering formalism[1,2] has proven to be a valuable tool for analyzing problems in atomic and nuclear collision theory and solid state physics. The theory, as generally applied, expresses the wavefunction for scattering from a collection of N targets in terms of the transition matrices of the individual targets. Previous formulations of the method have been restricted to the case of stationary scatterers. In this paper the multiple scattering formalism is generalized to a limited, but useful, class of time dependent problems in which each target may be moving with its own constant velocity.

Consider the interaction of a particle with N target centers each moving with constant velocity $\mathbf{v}_n$. The recoil of the targets will be neglected—an approximation that is warranted for heavy targets and low enough projectile energies. The nonrelativistic Schrödinger equation for the problem, in units where $\hbar = m = 1$, is

$$\left[i\frac{\partial}{\partial t} - \frac{p^2}{2} - \sum_{n=1}^{N} U_n(\mathbf{r} - \mathbf{v}_n t)\right]\Psi(\mathbf{r}, t) = 0, \tag{1}$$

where $[U_n(\mathbf{r})]$ are the targets potentials. Generally $U_n(\mathbf{r})$ will be centered around some arbitrary origin $\mathbf{r} = \mathbf{R}_n$.

Let us introduce the Galilean boost operators

$$B_n(t) = \exp(i\mathbf{v}_n \cdot \boldsymbol{\beta}), \tag{2}$$

where $\boldsymbol{\beta}$, the generator of the boost, is the Hermitian operator

$$\boldsymbol{\beta} = \frac{1}{\hbar}(\mathbf{p}t - m\mathbf{r}), \tag{3}$$

where we have temporarily restored atomic units. B_n has the properties

$$B_n \mathbf{r} B_n^{-1} = \mathbf{r} + \mathbf{v}_n t, \tag{4a}$$

$$B_n \mathbf{p} B_n^{-1} = \mathbf{p} + m\mathbf{v}_n, \tag{4b}$$

and

$$B_n i\hbar \frac{\partial}{\partial t} B_n^{-1} = i\hbar \frac{\partial}{\partial t} + \mathbf{v}_n \cdot \mathbf{p} + \frac{m}{2}\mathbf{v}_n^2. \tag{4c}$$

The function $B_n\Psi$ is simply the wavefunction of the system viewed from an inertial reference frame in which the nth scatterer is at rest. The potential of Eq. (1) may be written as

$$\sum_n U_n(\mathbf{r} - \mathbf{v}_n t) = \sum_n B_n^{-1}(t) U_n(\mathbf{r}) B_n(t). \tag{5}$$

Upon successively transforming one's frame of reference to different scatterers one may arrive at an intuitive set of multiple scattering equations. These transformations are illustrated symbolically in Fig. 1 where we show the scattering as viewed from the laboratory frame and from the frame in which target 1 is at rest. Similar diagrams may be drawn where other particles are at rest. Figure 1(a) depicts the incident and outgoing waves for each scatterer as viewed in the laboratory system. The incident wave for a given target is constructed by superimposing the outgoing waves from other targets and adding this to the free wave, Φ. In Fig. 1(b) we redraw this diagram as viewed from a frame in which 1 is at rest. It is in this frame that a simple relation exists connecting the scattered and incident waves. The hypothesized equations, which will be justified later, are

$$\Psi(t) = \Phi(t) + \sum_n B_n^{-1}(t) \int\int G_0^+(t, t') T_n(t', t'') \times \Psi_n(t'')\, dt'\, dt'', \tag{6}$$

which expresses the total wavefunction in terms of a free wave, Φ, and a sum of scattered waves from the various targets. Here $\Psi_n(t)$ is the scattered wave arriving at the nth target in a frame in which the target is at rest. Since we are using transition matrices in our formalism, it is clear that $\Psi_n(t)$ will propagate as a free wave. Physically $\Psi_n(t)$ is formed by superimposing all free waves in the problem. These include the original incident wave Φ plus those free waves emerging from the other scattering centers. It satisfies the

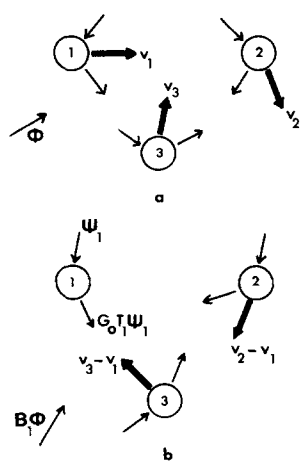


FIG. 1. The scattering process as viewed from the laboratory frame, (a), and from a frame in which target 1 is at rest, (b). In the latter frame the scattering off target 1 is simply described. The frames are related to each other through the Galilean boost operator, B_1. The small arrows associated with each scattering center depict the incident and scattered waves at that center.

[a)]Research supported by City University Faculty Research Award Program grant No. 11684.

 0022-2488/78/1909-1866$1.00

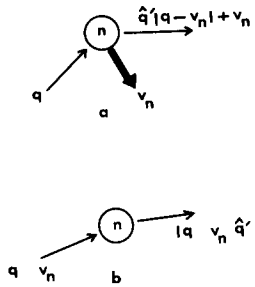


FIG. 2. A collision viewed from the laboratory frame, (a), and from a frame in which a target n is at rest, (b). The collision is elastic in the latter frame but not in the former.

equation

$$\Psi_n(t)=B_n(t)\Phi(t)+\sum_{\substack{m\\ m\neq n}} B_n(t)B_m^{-1}(t)\int\int G_0^+(t,t')T_m(t',t'')\times\Psi_m(t'')\,dt'\,dt''. \tag{7}$$

The free retarded propagator has been denoted by $G_0^+(t, t')$ and obeys the equation

$$\left(i\frac{\partial}{\partial t}-\frac{p^2}{2}\right)G_0^+(t,t')=\delta(t-t'). \tag{8}$$

The transition matrix, $T_m(t,t')$, appearing in Eqs. (6) and (7) is defined through the relation

$$T_n(t',t'')=U_n\delta(t'-t'')+U_n\int G_0^+(t',t)T_n(t,t'')\,dt. \tag{9}$$

It must be emphasized that in Eq. (9) U_n is not dependent on time. Thus the solution to the time dependent multiple scattering problem has been reduced to a knowledge of the ordinary transition matrices for a single target.

In the above equations we have elected to work in a time representation, as opposed to the more standard energy representation. Even if the incident wave, Φ, were taken to be monoenergetic, the various scattered waves would surely contain admixtures of states with many different energies. This is illustrated in Fig. 2 where a scattering event is viewed from the laboratory frame and from a frame in which the target is at rest. The target behaves as if it had infinite mass since its velocity is unchanged.

Equations (6) and (7) are the conjectured multiple scattering equations. Let us now prove that they satisfy Eq. (1). Inserting Eqs. (5) and (6) into the left-hand side of Eq. (1) and using the facts that $\Phi(t)$ obeys the free Schrödinger equation, and that $i(\partial/\partial t)-p^2/2$ commutes with $B_n^{-1}(t)$ leads to

$$\begin{aligned}Q\equiv&\left(i\frac{\partial}{\partial t}-\frac{p^2}{2}-\sum_n U_n(\mathbf{r}-\mathbf{v}_n t)\right)\Psi(t)\\ =&\sum_n B_n^{-1}(t)\Big[\int dt'\,T_n(t,t')\Psi_n(t')-U_nB_n(t)\Phi(t)\\ &-U_n\int\int G_0^+(t,t'')T_n(t'',t')\Psi_n(t')\,dt'\,dt''\Big]\\ &-\sum_{mn}{}' B_n^{-1}(t)U_nB_n(t)B_m^{-1}(t)\int\int G_0(t,t'')T_m(t'',t')\\ &\times\Psi_m(t')\,dt'\,dt'',\end{aligned}$$

where we have used Eq. (8). Using Eq. (7) and regrouping terms leads to

$$Q=\sum_n B_n^{-1}(t)\Big\{\int[T_n(t,t')-U_n\delta(t-t')]\Psi_n(t')\,dt'-U_n\int\int G_0^+(t,t')T_n(t',t'')\Psi_n(t'')\,dt'\,dt''\Big\}.$$

Finally from Eq. (9) one is able to show that Q vanishes identically, proving the validity of the equations.

As a simple check on the equations, in the limit where all velocities go to zero, the Galilean boost operators reduce to identity operators and Eqs. (7) and (8) collapse to the conventional time independent multiple scattering equations.[1,2]

In order to obtain an explicit solution to the above equations one must resort to a particular representation. Two representations are convenient. The first is the familiar plane wave representation. The matrix elements of the boost operator are simply

$$\langle\mathbf{q}\,|B_n(t)\,|\mathbf{q}'\rangle=(2\pi)^3\delta(\mathbf{v}_n-\mathbf{q}-\mathbf{q}'). \tag{10}$$

Alternatively one may employ the boost representation in which the boost operators are diagonal

$$\beta\phi_{\mathbf{b}}(\mathbf{r},t)=\mathbf{b}\phi_{\mathbf{b}}(\mathbf{r},t). \tag{11}$$

The eigenfunctions are obtained by integration of this differential equation and are given by

$$\phi_{\mathbf{b}}(\mathbf{r},t)=(2\pi t)^{-3/2}\exp\left[\frac{i}{t}\left(\frac{r^2}{2}+\mathbf{b}\cdot\mathbf{r}\right)\right]. \tag{12}$$

They form a complete orthonormal set, i.e.,

$$\int d\mathbf{b}\,\phi_{\mathbf{b}}^*(\mathbf{r},t)\phi_{\mathbf{b}}(\mathbf{r}',t)=\delta(\mathbf{r}-\mathbf{r}') \tag{13}$$

and

$$\int d\mathbf{r}\,\phi_{\mathbf{b}}^*(\mathbf{r},t)\phi_{\mathbf{b}'}(\mathbf{r},t)=\delta(\mathbf{b}-\mathbf{b}'). \tag{14}$$

The free propagator, in this representation is

$$\langle\mathbf{b},t\,|G_0^+|\,\mathbf{b}',t'\rangle=-i\Theta(t-t')\delta(\mathbf{b}-\mathbf{b}')\times\exp\left[\frac{ib^2}{2}\left(\frac{1}{t}-\frac{1}{t'}\right)\right]. \tag{15}$$

The matrix elements of the boost operator are simply

$$\langle\mathbf{b},t\,|B_n(t)\,|\mathbf{b}',t\rangle=\delta(\mathbf{b}-\mathbf{b}')\exp(i\mathbf{v}_n\cdot\mathbf{b}). \tag{16}$$

While it is not the purpose of this paper to present detailed applications of the time dependent multiple scattering formalism, let us outline some possible problems that are amenable to study using it. In atomic physics such problems as charge exchange between ions and/or atoms, and ionization and excitation by heavy projectiles can be approached by letting $N=2$. It is interesting to note, in this regard, that in the special (and nonphysical) case of one-dimensional delta functions some analytic solutions to these problems have recently been obtained.[3,4,5] In solid state physics such problems as beam foil excitation or stripping may be approached. Also processes in which there is an intense deposition of energy may be studied, as it may be assumed that the atoms of the solid will freely expand from some period of time.

The present exposition has been limited to the case of particle scattering, but it is clear that a parallel

development exists for electromagnetic scattering. Using techniques developed here it should be possible to discuss the scattering of electromagnetic waves from an exploding (or imploding) system of fragments—a problem of some current interest in the laser fusion problem.

[1]M. Lax, Rev. Mod. Phys. **23**, 287 (1951).
[2]K. Watson, Phys. Rev. **89**, 575 (1953).
[3]S. K. Zhdanov and A. S. Chikhachev, Sov. Phys.-Dokl. **19**, 696 (1975) [Dokl. Akad. Nauk. SSSR **218**, 1323 (1974)].
[4]W. Dappen, J. Phys. B. **10**, 2399 (1977).
[5]In dealing with bound state problems one simply omits the incident wave Φ from Eqs. (6) and (7).

Variational principles on rth order jets of fibre bundles in field theory [a]

Víctor Aldaya and José A. de Azcárraga

Física Matemática, Facultad de Ciencias, Salamanca, Spain
(Received 8 August 1977)

The Hamilton and the modified Hamilton variational principles in classical field theory are studied for physical systems described by Lagrangian and Hamiltonian densities depending on arbitrary order derivatives of the field. These principles are established on the fibre bundles $J^r(E)$, $J^1(J^{r-1}(E))$, $J^{1*}(J^{r-1}(E))$. This is accomplished by defining an appropriate Poincaré–Cartan form. This form is also required in the definition of the associated symmetry problem and in the explicit construction of the Noether currents.

1. INTRODUCTION AND MOTIVATION

The geometric formulation of the variational principles in field theory has aroused special interest on account of its applications to local field theories. Since the early work of Dedecker[1] on the variational calculus, the subject has attracted the attention of many authors.[2–16] Basically, the starting point of the geometric approach to the variational principles is the definition of the Lagrangian density as a function on a fibered manifold, the bundle $J^r(E)$ on the r-jets[17] of a (vector) bundle E. For the Lagrangian densities usually considered in field theory, it is sufficient to use the 1-jet formalism.[2–7, 9, 12, 13] The Lagrangian theory thus constructed leads to the customary equations of motion and is suitable for the study of its general invariance properties.[2–11, 13, 14, 16]

The usual ($r=1$) Lagrangian theory is, however, clearly inadequate for the study of more general situations where the Lagrangian density depends on higher order derivatives of the field and, even more so, for the study of theories which present a nonlocal character. These are not uncommon in physics; for instance, the recently introduced Melosh transformation,[18] which is required in hadronic matrix elements of otherwise local interactions (such as e.m. int.), has a nonlocal structure. This nonlocal character is a common feature of the so-called "canonical transformations" which may be used to introduce symmetries which are not apparent in the original (local) formulation.

With these questions in view, we shall perform in this paper a systematic study of the variational principles allowing the integrand in the action integral to depend on an *arbitrary number of derivatives* of the field. For a theory described by a Lagrangian density, the dependence on an infinite number of derivatives (i.e., $r \to \infty$) would give it a nonlocal character. In general, the infinite limit may not be reached, but we can expect that the necessity of introducing nonlocal operators will manifest itself in this scheme at lower orders. For instance, and in the case of a canonical transformation where the nonlocality is a consequence of taking it exactly, the symmetries introduced will retain an approximate validity at finite orders of its corresponding power expansion. This will be shown explicitly elsewhere,[19] where a specific example is considered.

Our analysis of the different variational principles starts with the *Hamilton principle* (Principle I) from which the Euler–Lagrange (EL) equations are obtained. The derivation of the simple EL equations for a Lagrangian density $\mathcal{L}^r$ depending on higher order derivatives of the field may be accomplished by using $J^r(E)$ as the definition space of $\mathcal{L}^r$; this formalism has already been considered (see, e.g., Refs. 11 and 12). The *modified Hamilton principle* (Principle II) makes use of the *Poincaré–Cartan form*. The use of the Poincaré–Cartan form is specially useful in the globalization of the EL equations and in symmetry considerations, as it was first emphasized by the work of García and Pérez-Rendón (see, e.g., Refs. 8 and 7). However, for general Lagrangians $\mathcal{L}^r$, Principle II requires a nontrivial *generalized Poincaré-Cartan form*. This is given in the present paper, where the regularity condition for the equivalence of both Hamiltonian principles is also established.

The Hamiltonian equations are also obtained in the general case from Principle II by writing the Poincaré–Cartan form in terms of a scalar Hamiltonian density. As a preliminary step, Principles I and II are applied to Lagrangian theories defined on $J^1(J^{r-1}(E))$, and the connection with the variational principles on $J^r(E)$ is exhibited by using the appropriate Lagrange multipliers. Then a Hamiltonian density is defined on the dual fiber bundle of $J^1(J^{r-1}(E)) \to J^{r-1}(E)$, $J^{1*}(J^{r-1}(E))$ which leads to the Hamilton equations. The Legendre transformation establishes the equivalence between the EL and Hamilton equations.

Finally, we devote particular attention to the Noether theorem, of fundamental importance in many physical applications. Its general formulation is now made possible by using the general Poincaré–Cartan form previously defined, and the explicit form of the conserved currents is given in the different formalisms.

This paper is organized as follows: In Sec. 2 we briefly review the Hamilton and the modified Hamilton principles for the usual (lowest) case $J^1(E)$ and set the notation. Section 3 is devoted to the study of the two variational principles on $J^r(E)$ and to the Noether theorem. Section 4 includes the Hamiltonian formalism

[a] Paper partially supported by the I.E.N., Madrid.

 0022-2488/78/1909-1869$1.00

and the Hamilton equations at lowest order and in Sec. 5 the previous variational principles, Poincaré–Cartan form, and Noether theorem are analyzed on $J^1(J^{r-1}(E))$ and $J^{1*}(J^{r-1}(E))$. The numbering of definitions and principles inside the different sections is arranged in a way which facilitates the comparison between the different formalisms introduced.

2. VARIATIONAL PRINCIPLES ON $J^1(E)$

Let (E, M, π) be the vector bundle whose module of cross sections $\Gamma(E)$ is composed of vector valued functions ψ with arguments on M, the Minkowski space. In particular, $\Gamma(E)$ includes the fields which describe the space–time evolution of the particles of the theory and provide the support of a unitary representation of the Poincaré group. Thus (our notation is similar to that of Refs. 4, 7, and 8) the points of the basis M are labeled by their coordinates x^μ, $\mu(\nu,\rho,\eta)=0,1,2,3$; the bundle E is the product $M\otimes R^\alpha\otimes R^i$, where α is the spinorial index of the field (for instance, $\alpha=1,2,3,4$ if ψ is the Dirac field). The index i will account for a possible additional inner symmetry [such as, for example, SU(3)]. Thus, the coordinates of ψ are $(x^\mu, y^{\alpha i})$; we shall omit the index i hereafter. Finally, π is the usual Cartesian projection of E onto M.

Definition 2.1: Let $(J^1(E), M, \pi^1)$ be the vector bundle of the first order jets associated to E $[J^0(E)\equiv E,\ \pi^0\equiv\pi]$. A *Lagrangian density density* $\mathcal{L}^1$ is a real valued function on $J^1(E)$ and accordingly of the arguments $(x^\mu, y^\alpha, y^\alpha_\mu)$ which constitute a system of canonical coordinates of $J^1(E)$. In general, and due to Poincaré invariance, $\mathcal{L}^1$ will not depend explicitly on x^μ.

Proposition 2.1: The 1-jet prolongation of ψ, $j^1(\psi)\equiv\bar\psi^1$ is the only section of $(J^1(E), M, \pi^1)$ such that j^1 is an injection of $\Gamma(E)$ into $\Gamma(J^1(E))$ and

$$\theta^\alpha|_{\bar\psi^1(M)}=0, \tag{2.1}$$

where θ^α is the set of 1-forms defined on $J^1(E)$ by

$$\theta^\alpha = dy^\alpha - y^\alpha_\mu dx^\mu. \tag{2.2}$$

In what follows we shall simply write $|_\psi$ to indicate the restriction to cross sections.

Proposition 2.2: Given a vector field X on E, $X\in\Gamma(T(E))$, its 1-jet prolongation through the injection j^1 is the only field $j^1(X)\equiv\bar X^1$ on $J^1(E)$ such that it is an infinitesimal contact transformation (ict), i.e.,

$$L_{\bar X^1}\theta^\alpha = A^\alpha_\beta\theta^\beta, \tag{2.3}$$

where $L_{\bar X^1}$ is the Lie derivative with respect to $\bar X^1$. Thus, the Pfaffian system (2.2) is stable under the one-parameter group generated by $\bar X^1$.

Thus, if X has the form

$$X = X^\mu\frac{\partial}{\partial x^\mu} + X^\alpha\frac{\partial}{\partial y^\alpha}, \tag{2.4}$$

$\bar X^1$ will be given by

$$\bar X^1 = X + \bar X^\alpha_\mu\frac{\partial}{\partial y^\alpha_\mu}. \tag{2.5}$$

Using (2.3), A^α_β and $\bar X^\alpha_\mu$ may be found to be

$$A^\alpha_\beta = \frac{\partial X^\alpha}{\partial y^\beta} - y^\alpha_\mu\frac{\partial X^\mu}{\partial y^\beta}, \tag{2.6}$$

$$\bar X^\alpha_\mu = \frac{\partial X^\alpha}{\partial x^\mu} - y^\alpha_\nu\frac{\partial X^\nu}{\partial x^\mu} + A^\alpha_\beta y^\beta_\mu = \frac{\partial X^\alpha}{\partial x^\mu} + \frac{\partial X^\alpha}{\partial y^\beta}y^\beta_\mu - y^\alpha_\nu\left(\frac{\partial X^\nu}{\partial x^\mu} + \frac{\partial X^\nu}{\partial y^\beta}y^\beta_\mu\right). \tag{2.7}$$

[Usually, $X^\nu\neq X^\nu(y^\alpha)$ and $\partial X^\nu/\partial y^\beta=0$.]

Definition 2.2: Given a Lagrangian density $\mathcal{L}^1$, the *Hamilton functional* I^1 is the application of $\Gamma(E)$ on R defined by

$$I^1(\psi)=\int_{j^1(\psi)(M)}\mathcal{L}^1(j^1(\psi))\pi^{1*}\omega\quad\forall\quad\psi\in\Gamma(E), \tag{2.8}$$

where $\pi^{1*}\omega$ is the prolongation of ω on $J^1(E)$ and $\omega = dx^0\wedge dx^1\wedge dx^2\wedge dx^3$ is the volume 4-form on M. Since $\pi^{1*}\omega$ only has components on M, we shall write just ω henceforth. $\mathcal{L}^1(j^1(\psi))$ is a function of y^α and its first order derivative, i.e., $y^\alpha_\mu=\partial_\mu y^\alpha$.

Principle 2.1 (*Hamilton*): The *critical sections* (*trajectories*) of the variational problem are the solutions of

$$(\delta I^1)_\psi(X)\equiv\frac{d}{d\lambda}I^1(\psi^\lambda)\Big|_{\lambda=0}=\int_{j^1(\psi)(M)}L_{\bar X^1}(\mathcal{L}^1(\bar\psi^1)\omega)=0$$

or

$$\int_M[j^1(\psi)]^*\{L_{\bar X^1}(\mathcal{L}^1\omega)\}=0\quad\forall\quad X\in\Gamma(T(E)), \tag{2.9}$$

where X is a vector field on E which generates a one-parameter transformation, λ its parameter, $\bar X^1$ the vector field on $(J^1(E))$ which is the 1-jet injection of X (in general, we denote jet prolongations with a bar), $L_{\bar X^1}$ the corresponding Lie derivative, and $(j^1(\psi))^*$ the dual of $j^1(\psi)$ acting on forms.

As is well known, (2.9) leads to

$$(EL)^1\psi=0, \tag{2.10}$$

where $(\mathrm{EL})^1$ is the usual Euler–Lagrange operator

$$\psi\overset{(\mathrm{EL})^1}{\to}[j^1(\psi)]^*\left(\frac{\partial\mathcal{L}^1}{\partial y^\alpha}-\frac{d}{dx^\mu}\left(\frac{\partial\mathcal{L}^1}{\partial y^\alpha_\mu}\right)\right). \tag{2.11}$$

The space of solutions of (2.10) will be called $\mathcal{U}^1_{\mathcal{L}}$.

Definition 2.3: Given a Lagrangian density, the Poincaré–Cartan (or Hilbert) form Θ^1 is defined by

$$\Theta^1=\Omega+\Omega', \tag{2.12}$$

where

$$\Omega=\mathcal{L}^1\omega,\quad \Omega'=\theta^\alpha\wedge\Omega_\alpha, \tag{2.13}$$

$$\Omega_\alpha\equiv\frac{\partial\mathcal{L}^1}{\partial y^\alpha_\mu}\theta_\mu,\quad \theta_\mu\equiv(-)^\mu dx^0\wedge\cdots\wedge d\hat x^\mu\wedge\cdots dx^3. \tag{2.14}$$

Because of (2.1), Θ^1 is equal to $\mathcal{L}^1\omega$ on 1-jet extensions $\bar\psi^1$, and thus (2.9) may be written using (2.3) as

$$(\delta I)_\psi(X)=\int_{\bar\psi^1(M)}L_{\bar X^1}\Theta^1=0. \tag{2.15}$$

Definition 2.4: Given the Poincaré–Cartan form Θ^1 on $J^1(E)$, the *modified Hamilton functional* I'^1 is the application of $\Gamma(J^1(E))$ on R defined by

$$I'^1(\psi^1)\equiv\int_{\psi^1(M)}\Theta^1\quad\forall\quad\psi^1\in\Gamma(J^1(E)). \tag{2.16}$$

Principle 2.2 (*Modified Hamilton principle*): A cross section is critical iff

$$(\delta I'^1)_{\psi^1}(X^1)=\int_{\psi^1(M)}L_{X^1}\Theta^1=0\quad\forall\quad X^1\in\Gamma(T(J^1(E))). \tag{2.17}$$

Since in the above expression [compare with (2.9) and (2.15)] X^1 is of the general form

$$X^1 = X^\mu \frac{\partial}{\partial x^\mu} + X^\alpha \frac{\partial}{\partial y^\alpha} + X^\alpha_\mu \frac{\partial}{\partial y^\alpha_\mu}. \tag{2.18}$$

(2.17) may be written in the form

$$i_{X^1} d\Theta^1 \big|_{\psi^1(M)} = 0, \tag{2.19}$$

where i_{X^1} denotes the inner product and the identity $L_X = i_X d + d i_X$ has been used.

A straightforward calculation from (2.19) leads to the known result that, for Lagrangians satisfying the *regularity condition*,

$$\det\left(\frac{\partial \mathcal{L}^1}{\partial y^\alpha_\mu \partial y^\beta_\nu}\right) \neq 0. \tag{2.20}$$

The space $\mathcal{U}'^1$ of the critical cross sections as defined by (2.17) is given by

$$\frac{d}{dx^\mu}\left(\frac{\partial \mathcal{L}^1}{\partial y^\alpha_\mu}\right) - \frac{\partial \mathcal{L}^1}{\partial y^\alpha} = 0 \tag{2.21}$$

with $y^\alpha_\mu = \partial_\mu y^\alpha$ (this condition comes from the coefficient of X^α_μ). Then $\mathcal{U}^1_L = \mathcal{U}'^1$ and the variational principles 2.1 and 2.2 are equivalent [for an intrinsic formulation of (2.20) see Ref. 7]. We shall always assume the regularity condition is fulfilled.

Definition 2.5: A transformation of the physical system described by $\mathcal{L}^1$ is a *symmetry* when it leaves the equations of the motion invariant. Thus, an ict $\bar{X}^1_s$ generates a symmetry transformation if

$$L_{\bar{X}^1}(\mathcal{L}\omega) - d\Delta = 0 \tag{2.22}$$

on any 1-jet prolongation $\bar{\psi}^1$, where $\Delta = \Delta^\mu \theta_\mu$ is a 3-form whose coefficients do not depend on y^α_μ (this guarantees that the original and the transformed Lagrangians differ in a four divergence which does not alter the EL equations). In particular, $d\Delta$ may be absent (this is the case for the Poincaré generators).

Theorem 2.1 (*Noether*): If $\bar{X}^1_s$ generates a symmetry transformation, then

$$d(i_{\bar{X}^1_s}\Theta^1 - \Delta)\big|_{\bar{\psi}^1} = 0 \tag{2.23}$$

for all sections $\psi \in \mathcal{U}^1$. The proof is immediate from (2.22) using (2.19).

Remark: When the two variational principles are not equivalent (the "regularity condition" is not satisfied) it is equally simple to define the symmetry transformation and the Noether current associated with it from Θ^1. To do this[7] it is only necessary to replace (2.22) by $(L_{X^1}\Theta^1 - d\omega_{X^1}) = 0$ and (2.23) by $d(i_{X^1}\Theta^1 - \omega_{X^1})\big|_{\psi^1} = 0$, where ω_{X^1} is a 3-form.

3. VARIATIONAL PRINCIPLES ON $J^r(E)$

We proceed now to extend the above considerations to generalized systems which depend on arbitrary order derivatives of the field.

Definition 3.1: Let $(J^r(E), M, \pi^r)$ be the vector bundle of the rth order jets associated with E. A generalized Lagrangian density $\mathcal{L}^r$ is a real valued function on $J^r(E)$, of arguments $(x^\mu, y^\alpha, y^\alpha_{\mu_1}, \ldots, y^\alpha_{\mu_1 \ldots \mu_r})$. Note that, as is customary, the usual covariance of the Lagrangian will require the use of nonindependent components on $J^r(E)$. We shall keep implicitly this trivial redundance in $\mathcal{L}^r$ and therefore in the expression of the tensor fields.

Proposition 3.1: The r-jet prolongation of ψ, $j^r(\psi) \equiv \bar{\psi}^r$, is the only section of $(J^r(E), M, \pi^r)$ such that j^r is an injection of $\Gamma(E)$ into $\Gamma(J^r(E))$ and

$$\theta^\alpha_{\mu_1 \ldots \mu_s}\big|_{\bar{\psi}^r} = 0, \quad s = 0, 1, \ldots, (r-1), \tag{3.1a}$$

where

$$\theta^\alpha_{\mu_1 \ldots \mu_s} \equiv dy^\alpha_{\mu_1 \ldots \mu_s} - y^\alpha_{\mu_1 \ldots \mu_s \mu}\, dx^\mu. \tag{3.1b}$$

Proposition 3.2: Given a vector field on E, its r-jet prolongation $\bar{X}^r$ through the injection j^r is the only vector field on $J^r(E)$ such that it is an ict, i.e.,

$$L_{\bar{X}^r}\theta^\alpha_{\mu_1 \ldots \mu_s} = \sum_{s' \leq s} A^{\alpha\, \nu_1 \ldots \nu_{s'}}_{\beta\, \mu_1 \ldots \mu_s}\, \theta^\beta_{\nu_1 \ldots \nu_{s'}}, \quad s = 0, 1, \ldots, (r-1). \tag{3.2}$$

For instance, if X is given by (2.4), its 2-jet prolongation is given by

$$\bar{X}^2 = X + \bar{X}^\alpha_\mu \frac{\partial}{\partial y^\alpha_\mu} + \bar{X}^\alpha_{\mu\nu} \frac{\partial}{\partial y^\alpha_{\mu\nu}} \tag{3.3}$$

and a calculation shows that $\bar{X}^\alpha_\mu$ is given by (2.7) and that

$$A^\alpha_{\beta\mu} = \frac{\partial \bar{X}^\alpha_\mu}{\partial y^\beta} - y^\alpha_{\mu\eta}\frac{\partial X^\eta}{\partial y^\beta}, \quad A^{\alpha\nu}_{\beta\mu} = \frac{\partial \bar{X}^\alpha_\mu}{\partial y^\beta_\nu} \tag{3.4}$$

from which $\bar{X}^\alpha_{\mu\nu}$ is given by

$$\bar{X}^\alpha_{\mu\nu} = \frac{\partial \bar{X}^\alpha_\mu}{\partial x^\nu} - y^\alpha_{\mu\eta}\frac{\partial X^\eta}{\partial x^\nu} + A^\alpha_{\beta\mu} y^\beta_\nu + A^{\alpha\eta}_{\beta\mu} y^\beta_{\eta\nu}. \tag{3.5}$$

Definition 3.2: Given a $\mathcal{L}^r$, the Hamilton functional I^r is the application of $\Gamma(E)$ on R defined by

$$I^r(\psi) \equiv \int_{j^r(\psi)(M)} \mathcal{L}^r(j^r(\psi))\omega. \tag{3.6}$$

Principle 3.1 (*Hamilton*): The critical cross sections are the solutions of

$$\int_{(M)} [j^r(\psi)]^* \{L_{\bar{X}^r}(\mathcal{L}^r\omega)\} = 0 \quad \forall\ X \in \Gamma(T(J^0(E))), \tag{3.7}$$

i.e., $\psi \in \mathcal{U}^r_L$ iff

$$[j^r(\psi)]^* \left\{ \sum_{s=0}^{r} (-)^s \frac{d^s}{dx^{\mu_1} \cdots dx^{\mu_s}} \frac{\partial \mathcal{L}^r}{\partial y^\alpha_{\mu_1 \cdots \mu_s}} \right\} = 0, \tag{3.8}$$

as may be derived from (3.7).

Definition 3.3: Given a Lagrangian density $\mathcal{L}^r$, the Poincaré–Cartan form (2.12) is now generalized to

$$\Theta^r = \mathcal{L}^r\omega + \Omega', \tag{3.9}$$

where

$$\Omega' = \sum_{s=0}^{r-1} \theta^\alpha_{\mu_1 \cdots \mu_s} \wedge \Omega^{\mu_1 \cdots \mu_s}_\alpha, \tag{3.10}$$

$$\Omega^{\mu_1 \cdots \mu_s}_\alpha = O^{\mu_1 \cdots \mu_s \mu}_\alpha \theta_\mu \equiv \left(\frac{\partial \mathcal{L}^r}{\partial y^\alpha_{\mu_1 \cdots \mu_s \mu}} + \lambda^{(r)\mu_1 \cdots \mu_s \mu}_\alpha\right)\theta_\mu, \tag{3.11}$$

$$s = 0, 1, \ldots, (r-1)$$

and the λ's satisfy the equations

$$\lambda_\alpha^{(r)\mu_1\cdots\mu_s}\big|_{\psi^r}\omega = d\Omega_\alpha^{\mu_1\cdots\mu_s}\big|_{\psi^r}, \quad s=1,\ldots,(r-1), \quad \forall\ \psi^r \tag{3.12}$$

$$\lambda_\alpha^{(r)\mu_1\cdots\mu_r}\big|_{\psi^r}=0 \quad (r \text{ indices}). \tag{3.13}$$

The above expressions define the $\Omega_\alpha^{\mu_1\cdots\mu_s}$ and Θ^r forms on cross sections of $J^r(E)$; this will allow us the definition of the modified Hamilton functional (3.17) which is defined on $\Gamma(J^r(E))$. However, one might ask whether it is possible to define Θ^r on each point of the bundle. This is easily done for the Hamilton principle, where it is sufficient to substitute y_μ^α, $y_{\mu\nu}^\alpha,\cdots$ for the terms $\partial y^\mu/\partial x^\mu$, $\partial y_\mu^\alpha/\partial x^\nu,\cdots$ which appear in (3.14) to obtain an expression for the λ's directly defined on the bundle.

On cross sections, and for $s=0,1,\ldots,(r-2)$,

$$\lambda_\alpha^{(r)\mu_1\cdots\mu_s\mu}\big|_{\psi^r} = -\frac{d}{dx^\nu}\left(\frac{\partial \mathcal{L}^r}{\partial y^\alpha_{\nu\mu_1\cdots\mu_s\mu}} + \lambda_\alpha^{\nu\mu_1\cdots\mu_s\mu}\right). \tag{3.14}$$

In particular, if $r=1$, no λ's appear in the definition of Θ^1[(2.13)]. For $r=3$, for instance,

$$\Theta^3 = \mathcal{L}^3\omega + \theta^\alpha\wedge\left(\frac{\partial \mathcal{L}^3}{\partial y_\mu^\alpha} + \lambda_\alpha^{(3)\mu}\right)\theta_\mu + \theta_\nu^\alpha\wedge\left(\frac{\partial \mathcal{L}^3}{\partial y_{\mu\nu}^\alpha} + \lambda_\alpha^{(3)\mu\nu}\right)\theta_\mu + \theta_{\nu\sigma}^\alpha\wedge\left(\frac{\partial \mathcal{L}^3}{\partial y_{\mu\nu\sigma}^\alpha}\right)\theta_\mu, \tag{3.15}$$

where

$$\lambda_\alpha^{(3)\mu\nu}\big|_{\psi^3} = -\frac{d}{dx^\sigma}\frac{\partial \mathcal{L}^3}{\partial y^\alpha_{\mu\nu\sigma}}; \qquad \lambda_\alpha^{(3)\mu}\big|_{\psi^3} = -\frac{d}{dx^\nu}\left(\frac{\partial \mathcal{L}^3}{\partial y^\alpha_{\mu\nu}} - \frac{d}{dx^\sigma}\frac{\partial \mathcal{L}}{\partial y^\alpha_{\mu\nu\sigma}}\right). \tag{3.16}$$

Definition 3.4a: Given the Poincaré–Cartan form Θ^r, the *modified Hamilton* functional I'^r is the application of $\Gamma(J^r(E))$ on R defined by

$$I'^r(\psi^r) \equiv \int_{\psi^r(M)}\Theta^r \quad \forall\ \psi^r\in\Gamma(J^r(E)). \tag{3.17}$$

Principle 3.2 (*Modified Hamilton principle*): A cross section is a trajectory iff

$$(\delta I'^r)_{\psi^r}(X^r) = \int_{\psi^r(M)} L_{X^r}\Theta^r = 0 \quad \forall\ X^r\in\Gamma(T(J^r(E))) \tag{3.18}$$

with

$$X^r = X^\mu\frac{\partial}{\partial x^\mu} + \sum_{s=0}^{r} X^\alpha_{\mu_1\cdots\mu_s}\frac{\partial}{\partial y^\alpha_{\mu_1\cdots\mu_s}}. \tag{3.19}$$

(3.18) may be written as

$$i_{X^r}d\Theta^r\big|_{\psi^r}=0. \tag{3.20}$$

A tedious but straightforward calculation shows that the first r equations (those corresponding to the coefficients of $X^\alpha_{\mu_1},\ldots,X^\alpha_{\mu_1\cdots\mu_r}$) may be written in the form

$$\begin{pmatrix} \frac{\partial}{\partial y^\beta_{\nu_1}}O^\mu_\alpha & \cdots & \frac{\partial}{\partial y^\beta_{\nu_1}}O^{\mu_1\cdots\mu_{r-1}\mu}_\alpha \\ \frac{\partial}{\partial y^\beta_{\nu_1\cdots\nu_r}}O^\mu_\alpha & \cdots & \frac{\partial}{\partial y^\beta_{\nu_1\cdots\nu_r}}O^{\mu_1\cdots\mu_{r-1}\mu}_\alpha \end{pmatrix} \times \begin{pmatrix} \frac{\partial y^\alpha}{\partial x^\mu} - y^\alpha_\mu \\ \frac{\partial y^\alpha_{\mu_1\cdots\mu_{r-1}}}{\partial x^\mu} - y^\alpha_{\mu_1\cdots\mu_{r-1}\mu} \end{pmatrix} = 0. \tag{3.21}$$

When this system admits only the trivial solution, i.e., the generalized "regularity condition" is satisfied [the above expression leads to (2.20) for $r=1$], the r-jet prolongation condition is implemented, and then the coefficient of X^α immediately reproduces the Euler–Lagrange equations (3.8). In this case $\mathcal{L}^r_l = \mathcal{L}'^r$ and again the variational Principles 3.1 and 3.2 are equivalent.

Definition 3.5: As in Definition 2.5, an ict $\bar{X}^r$ generates a symmetry transformation if, on any r-jet prolongation $\bar{\psi}^r$,

$$L_{\bar{X}^r}(\mathcal{L}^r\omega) - d\Delta = 0, \tag{3.22}$$

where $\Delta = \Delta^\mu\theta_\mu$ is a 3-form whose coefficients do not depend on $y^\alpha_{\mu_1\cdots\mu_r}$.

Theorem 3.1. (*Noether*): If $\bar{X}^r_s$ generates a symmetry transformation, then

$$d(i_{\bar{X}^r}\Theta^r - \Delta)\big|_{\bar{\psi}^r}=0 \tag{3.23}$$

for all sections $\psi\in\mathcal{L}^r_l$.

Proof:

$$L_{\bar{X}^r_s}(\mathcal{L}^r\omega) = L_{\bar{X}^r_s}(\Theta^r - \Omega') = L_{\bar{X}^r_s}\Theta^r - L_{\bar{X}^r_s}\Omega' = i_{\bar{X}^r_s}d\Theta^r + d i_{\bar{X}^r_s}\Theta^r - L_{\bar{X}^r_s}\Omega'. \tag{3.24}$$

Now, since the first term of (3.24) is zero for the r-jet prolongations of the elements of $\mathcal{L}^r_l$ [(3.20), (3.8)] and the third is also zero on r-jet extensions [(3.10), (3.2)], (3.23) follows from (3.22). ∎

Using

$$i_{\bar{X}^r}\theta^\alpha_{\mu_1\cdots\mu_s} = \bar{X}^\alpha_{\mu_1\cdots\mu_s} - y^\alpha_{\mu_1\cdots\mu_s\mu}X^\mu, \tag{3.25}$$

the conserved quantity $i_{\bar{X}^r}\Theta^r$ may be written in the form

$$i_{\bar{X}^r}\Theta^r = \sum_{s=0}^{r-1}(\bar{X}^\alpha_{\mu_1\cdots\mu_s} - y^\alpha_{\mu_1\cdots\mu_s\mu}X^\mu)\Omega_\alpha^{\mu_1\cdots\mu_s} + \mathcal{L}^r X^\mu\theta_\mu - \sum_{s=0}^{r-1}\theta^\alpha_{\mu_1\cdots\mu_s}\wedge i_{\bar{X}^r}\Omega^{\mu_1\cdots\mu_s}_\alpha; \tag{3.26}$$

the last term may be omitted since it disappears in $i_{\bar{X}^r}\Theta^r|_{\bar{\psi}^r}$.

(3.26) may be written in a perhaps more familiar form using the Hodge (dual) operator $*$. Defining

$$j = *(i_{\bar{X}^r}\Theta^r) \tag{3.27}$$

and taking its exterior codifferential $\delta\sim *d*$ we obtain

$$\delta j = *d(i_{\bar{X}^r}\Theta^r) = \mathrm{div} j = 0, \tag{3.28}$$

where the conserved current takes the form

$$j^\mu = \mathcal{L}^r X^\mu + \sum_{s=0}^{r-1}(\bar{X}^\alpha_{\mu_1\cdots\mu_s} - y^\alpha_{\mu_1\cdots\mu_s\nu}X^\nu)O^{\mu_1\cdots\mu_s\mu}_\alpha. \tag{3.29}$$

Of course, (3.27)–(3.29) are evaluated on the r-jet prolongations of the solutions of the Euler–Lagrange equations (space $\mathcal{L}^r_l$). An overall minus sign is sometimes conventionally introduced in (3.29).

4. FIRST ORDER HAMILTONIAN FORMALISM AND HAMILTON EQUATIONS

In this section we apply the vector bundle formalism to the usual Hamilton equations in classical field theory. To do this, we must first define the Hamiltonian den-

sity $\mathcal{H}$ in this formalism. The independent variables are now the fields and their conjugate momenta. Thus, the appropriate space for the definition of $\mathcal{H}$ will be $J^{1*}(E)$, dual of $J^1(E)\to E$, which for later convenience will be written as $J^{1*}(J^0(E))$. Given a coordinate system $(x^\mu, y^\alpha, y^\alpha_\mu)$ in $J^1(J^0(E))$ and the corresponding dual system $(x^\mu, y^\alpha, \pi^\mu_\alpha)$ for $J^{1*}(J^0(E))$, the *Legendre transformation* D_L

$$D_L : \psi^1 \to D_L(\psi^1), \quad \psi^1 \in J^1(J^0(E)) \tag{4.1a}$$

is defined by

$$\pi^\mu_\alpha(D_L(\psi^1)) = \frac{\partial \mathcal{L}}{\partial y^\alpha_\mu}\bigg|_{\psi^1}. \tag{4.1b}$$

As we have already done in previous expressions, we shall omit the argument when no confusion arises.

Given a Lagrangian $\mathcal{L}^1$ on $J^1(J^0(E))$, we may now proceed to define a Hamiltonian associated to it, which we shall do in a Lorentz invariant manner.

Definition 4.1: A "Hamiltonian" $\mathcal{H}^1$ is a scalar real function on $J^{1*}(J^0(E))$ defined by

$$\mathcal{H}^1 = y^\alpha_\mu \pi^\mu_\alpha - \mathcal{L}^1. \tag{4.2}$$

In terms of $\mathcal{H}^1$, the Poincaré–Cartan form (2.12) may be written as

$$\Theta^1 = \pi^\mu_\alpha dy^\alpha \wedge \theta_\mu - \mathcal{H}\omega. \tag{4.3}$$

[Of course, Θ^1 may be defined without explicit recourse to $\mathcal{L}$ from an $\mathcal{H}$ defined on $J^{1*}(J^0(E))$ and the Liouville form $\Lambda^\mu = \pi^\mu_\alpha dy^\alpha$ as $\Theta = \Lambda^\mu \wedge \theta_\mu - \mathcal{H}\omega$.]

With Θ^1 defined on $J^*(J^0(E))$ [(4.3)], (2.16) is replaced by

$$I'^1(\psi^{1*}) \equiv \int_{\psi^{1*}(M)} \Theta^1 \quad \forall\ \psi^{1*} \in \Gamma(J^{1*}(J^0(E))); \tag{4.4}$$

(2.17) by

$$(\delta I'^1)_{\psi^{1*}}(X^{1*}) = \int_{\psi^{1*}(M)} L_{X^{1*}}\Theta^1 \quad \forall\ X^{1*} \in \Gamma(T(J^{1*}(J^0(E)))); \tag{4.5}$$

and (2.19) by

$$i_{X^{1*}} d\Theta^1\big|_{\psi^{1*}} = 0. \tag{4.6}$$

Writing

$$X^{1*} = X^\mu \frac{\partial}{\partial x^\mu} + X^\alpha \frac{\partial}{\partial y^\alpha} + X^\mu_\alpha \frac{\partial}{\partial \pi^\mu_\alpha}, \tag{4.7}$$

a simple calculation shows that (4.6) is nothing but the Hamilton equations (as in the Lagrangian case, the coefficient of X^μ may be seen to lead to a trivial identity)

$$\frac{\partial \mathcal{H}}{\partial \pi^\mu_\alpha} = \frac{\partial y^\alpha}{\partial x^\mu}, \quad \frac{\partial \mathcal{H}}{\partial y^\alpha} = -\frac{\partial \pi^\mu_\alpha}{\partial x^\mu}. \tag{4.8}$$

When the Legendre transformation is a diffeomorphism, substitution of (4.2) into (4.8) leads again to the 1-jet prolongation $y^\alpha_\mu = \partial_\mu y^\alpha$ and to the Euler–Lagrange equations (2.10).

5. VARIATIONAL PRINCIPLES ON $J^1(J^{r-1}(E))$ AND $J^{1*}(J^{r-1}(E))$

We now proceed to extend the variational principles 2.1 and 2.2 from $J^1(J^0(E))$ to $J^1(J^{r-1}(E))$. Our aim here is to provide a new generalization which is suitable for the Hamilton equations since—as we have seen in the previous paragraph—this requires a formalism where coordinates and momenta can be adequately defined.

The space $J^1(J^{r-1}(E))$ is the vector bundle of the 1-jets associated to the fibre space $(J^{r-1}(E), M, \pi^{r-1})$. The coordinate system for $J^1(J^{r-1}(E))$ will be defined by

$$(x^\mu, y^\alpha, y^\alpha_{\mu_1}, \ldots, y^\alpha_{\mu_1\cdots\mu_{r-1}}; g^\alpha_{,\mu}, g^\alpha_{\mu_1,\mu}, \ldots, g^\alpha_{\mu_1\cdots\mu_{r-1},\mu}). \tag{5.1}$$

We now extend the results of Sec. 2 to this case.

Definition 5.1: A Lagrangian of type $\mathcal{L}^{1,r-1}$ is an application of $J^1(J^{r-1}(E))$ on R.

Proposition 5.1: The 1-jet prolongation of $\psi^{r-1}, j^1(\psi^{r-1}) \equiv \overline{\psi^{r-1}}^1$ is the only section of $(J^1(J^{r-1}(E)), M, \pi)$ such that j^1 is an injection of $\Gamma(J^{r-1}(E))$ into $\Gamma(J^1(J^{r-1}(E)))$ and

$$\theta^{1,r-1}{}^\alpha_{\mu_1\cdots\mu_s}\big|_{\overline{\psi^{r-1}}^1} = 0, \quad s = 0, 1, \ldots, (r-1), \tag{5.2}$$

where the structure forms $\theta^{1,r-1}$ (we shall simply write θ) are *now*

$$\theta^\alpha_{\mu_1\cdots\mu_s} = dy^\alpha_{\mu_1\cdots\mu_s} - g^\alpha_{\mu_1\cdots\mu_s,\mu} dx^\mu. \tag{5.3}$$

As is clear from (5.2), the jet prolongation condition requires

$$g^\alpha_{\mu_1\cdots\mu_s,\mu} = \partial_\mu y^\alpha_{\mu_1\cdots\mu_s}.$$

Proposition 5.2: Given a vector field on $J^{r-1}(E)$, $X^{r-1} \in \Gamma(T(J^{r-1}(E)))$, its 1-jet prolongation through the injection j^1 is the only field $j^1(X^{r-1}) \equiv \overline{X^{r-1}}^1$ such that it is an ict, i.e.,

$$L_{\overline{X^{r-1}}^1}\theta^\alpha_{\mu_1\cdots\mu_s} = \sum_{m=0}^{r-1} A^\alpha_{\mu_1\cdots\mu_s}{}^{;\nu_1\cdots\nu_m}_{;\beta}\theta^\beta_{\nu_1\cdots\nu_m}$$
$$s = 0, 1, \ldots, (r-1), \tag{5.4}$$

where the 1-jet extension of X^{r-1} is determined from the general form of $X^{1,r-1}$,

$$X^{1,r-1} = X^\mu \frac{\partial}{\partial x^\mu} + X^\alpha \frac{\partial}{\partial y^\alpha} + \cdots + X^\alpha_{\mu_1\cdots\mu_{r-1}} \frac{\partial}{\partial y^\alpha_{\mu_1\cdots\mu_{r-1}}}$$
$$+ X^\alpha_{,\mu} \frac{\partial}{\partial g^\alpha_{,\mu}} + \cdots + X^\alpha_{\mu_1\cdots\mu_{r-1},\mu} \frac{\partial}{\partial g^\alpha_{\mu_1\cdots\mu_{r-1},\mu}}. \tag{5.5}$$

(5.4) and (5.5) lead to [compare with (2.6) and (2.7)]

$$A^\alpha_{\mu_1\cdots\mu_s}{}^{;\nu_1\cdots\nu_{s'}}_{;\beta} = \frac{\partial X^\alpha_{\mu_1\cdots\mu_s}}{\partial y^\beta_{\nu_1\cdots\nu_{s'}}} - g^\alpha_{\mu_1\cdots\mu_s,\mu}\frac{\partial X^\mu}{\partial y^\beta_{\nu_1\cdots\nu_{s'}}} \tag{5.6}$$

so that, *for 1-jet prolongations*, the components $\bar{X}^\alpha_{\mu_1\cdots\mu_{r-1},\mu}$ of (5.5) are written

$$\bar{X}^\alpha_{\mu_1\cdots\mu_s,\mu} = \frac{\partial X^\alpha_{\mu_1\cdots\mu_s}}{\partial x^\mu} - g^\beta_{\mu_1\cdots\mu_s,\nu}\frac{\partial X^\nu}{\partial x^\mu}$$
$$+ \sum_{m=0}^{r-1} A^\alpha_{\mu_1\cdots\mu_s}{}^{;\nu_1\cdots\nu_m}_{;\beta} g^\beta_{\nu_1\cdots\nu_m,\mu}. \tag{5.7}$$

Definition 5.2: Given a Lagrangian density $\mathcal{L}^{1,r-1}$, the *Hamilton functional* $I^{1,r-1}$ is the application of $\Gamma(J^{r-1}(E))$ on R defined by

$$I^{1,r-1}(\psi^{r-1}) = \int_{j^1(\psi^{r-1})(M)} \mathcal{L}^{1,r-1}(j^1(\psi^{r-1}))\omega$$
$$\forall\ \psi^{r-1} \in \Gamma(J^{r-1}(E)). \tag{5.8}$$

Principle 5.1 (*Hamilton*): The trajectories of the variational problem are the solutions of

$$\int_M [j^1(\psi^{r-1})]^*\{L_{\overline{X^{r-1}}^1}(\mathcal{L}^{1,r-1}\omega)\}=0 \quad \forall\, X^{r-1}\in\Gamma(T(J^{r-1}(E)))\text{.} \tag{5.9}$$

(5.9) leads to a set of "disconnected" Euler–Lagrange equations of the form

$$\frac{\partial \mathcal{L}^{1,r-1}}{\partial y^\alpha_{\nu_1\cdots\nu_s}}-\frac{d}{dx^\mu}\left(\frac{\partial \mathcal{L}^{1,r-1}}{\partial g^\alpha_{\nu_1\cdots\nu_s,\,\mu}}\right)=0,\quad s=0,1,\ldots,(r-1), \tag{5.10}$$

with $g^\alpha_{\nu_1\cdots\nu_s,\,\mu}=\partial_\mu y^\alpha_{\nu_1\cdots\nu_s}$.

Definition 5.3: Given a $\mathcal{L}^{1,r-1}$, the Poincaré–Cartan form is given by

$$\Theta^{1,r-1}=\mathcal{L}^{1,r-1}\omega+\sum_{s=0}^{r-1}\theta^\alpha_{\mu_1\cdots\mu_s}\wedge\Omega^{\mu_1\cdots\mu_s}_\alpha, \tag{5.11}$$

where

$$\Omega^{\mu_1\cdots\mu_s}_\alpha\equiv\frac{\partial \mathcal{L}^{1,r-1}}{\partial g^\alpha_{\mu_1\cdots\mu_s,\,\mu}}\theta_\mu\equiv O^{\mu_1\cdots\mu_s,\,\mu}_\alpha\theta_\mu\text{.} \tag{5.12}$$

It should be noted that the definition given here for the Ω's is a straightforward generalization of (2.14) and no different from (3.11) since for 1-jet prolongations that formula would not contain λ's.

Definition 5.4: Given $\Theta^{1,r-1}$ on $J^1(J^{r-1}(E))$ the *modified Hamilton functional* $I^{1,r-1}$ is the application of $\Gamma(J^1(J^{r-1}(E)))$ on R defined by

$$I^{1,r-1}(\psi^{1,r-1})\equiv\int_{\psi^{1,r-1}}\Theta^{1,r-1}\quad\forall\,\psi^{1,r-1}\in\Gamma(J^1(J^{r-1}(E)))\text{.} \tag{5.13}$$

Principle 5.2 (*Modified Hamilton principle*): A cross section is critical iff

$$(\delta I^{1,r-1})_{\psi^{1,r-1}}(X^{1,r-1})=\int_{\psi^{1,r-1}(M)}L_{X^{1,r-1}}\Theta^{1,r-1}=0$$
$$\forall X^{1,r-1}\in\Gamma(T(J^1(J^{r-1}(E))))\text{.} \tag{5.14}$$

The critical sections are thus the solutions of

$$i_{X^{1,r-1}}d\Theta^{1,r-1}\big|_{\psi^{1,r-1}}=0\text{.} \tag{5.15}$$

A calculation from (5.15) leads to the following system of equations [which come from the coefficients of $X^\alpha_{\mu_1\cdots\mu_s,\,\mu}$, $s=0,1,\ldots,(r-1)$]:

$$\begin{pmatrix}\frac{\partial}{\partial g^\beta_{,\nu}}O^\mu_\alpha & \cdots & \frac{\partial}{\partial g^\beta_{,\nu}}O^{\mu_1\cdots\mu_{r-1},\,\mu}_\alpha\\ \frac{\partial}{\partial g^\beta_{\nu_1\cdots\nu_{r-1},\nu}}O^\mu_\alpha & \cdots & \frac{\partial}{\partial g^\beta_{\nu_1\cdots\nu_{r-1},\nu}}O^{\mu_1\cdots\mu_{r-1},\,\mu}_\alpha\end{pmatrix}$$
$$\times\begin{pmatrix}\frac{\partial y^\alpha}{\partial x^\mu}-g^\alpha_\mu\\ \frac{\partial y^\alpha_{\mu_1\cdots\mu_{r-1}}}{\partial x^\mu}-g^\alpha_{\mu_1\cdots\mu_{r-1},\,\mu}\end{pmatrix}=0. \tag{5.16}$$

When (5.16) admits only the trivial solution (regularity condition for $\mathcal{L}^{1,r-1}$), the 1-jet prolongation condition is satisfied, and then the remaining equations [from the coefficients of $X^\alpha_{\mu_1\cdots\mu_s}$, $s=0,1,\ldots,(r-1)$] reproduce (5.10). Thus, if this regularity condition is fulfilled, Principles 5.1 and 5.2 are equivalent.

Definition 2.5 may now be extended to $\mathcal{L}^{1,r-1}$, and the corresponding theorem is obtained:

Theorem 5.1 (*Noether*): If $\overline{X^r_s}^1$ generates a symmetry transformation $[L_{\overline{X^{r-1}}^1}(\mathcal{L}^{1,r-1}\omega)-d\Delta=0$, where $\Delta=\Delta^\mu\theta_\mu$ and Δ^μ does not depend on the g's], then

$$d(i_{\overline{X^{r-1}}^1}\Theta^{1,r-1}-\Delta)\big|_{\overline{\psi^{r-1}}^1}=0\quad\forall\,\psi^{r-1}\in\mathcal{U}^{1,r-1}_{\mathcal{L}}\text{.} \tag{5.17}$$

Following the already usual steps, the form of the conserved current turns out to be (on $\overline{\psi^{r-1}}^1$, $\psi^{r-1}\in\mathcal{U}^{1,r-1}_{\mathcal{L}}$)

$$j^{(1,r-1)\mu}=\mathcal{L}^{1,r-1}X^\mu+\sum_{s=0}^{r-1}(X^\alpha_{\mu_1\cdots\mu_s}-g^\alpha_{\mu_1\cdots\mu_s,\,\nu}X^\nu)O^{\mu_1\cdots\mu_s,\,\mu}_\alpha\text{.} \tag{5.18}$$

To derive the *Hamilton equations* which obey the variational principle 5.2 we introduce the definition space of the scalar Hamiltonian density, $J^{1*}(J^{r-1}(E))$. Its corresponding coordinate system is given by

$$(x^\mu,y^\alpha,y^\alpha_{\mu_1}\cdots y^\alpha_{\mu_1\cdots\mu_{r-1}};\pi^{\,\mu}_\alpha,\pi^{\mu_1,\,\mu}_\alpha\cdots\pi^{\mu_1\cdots\mu_{r-1},\,\mu}_\alpha)\text{.} \tag{5.19}$$

The Legendre transformation D_L: $J^1(J^{r-1}(E))\to J^{1*}(J^{r-1}(E))$ is now defined by

$$D_L:\psi^{1,r-1}\to D_L(\psi^{1,r-1}), \tag{5.20}$$

$$\pi^{\mu_1\cdots\mu_s,\,\mu}_\alpha(D_L(\psi^{1,r-1}))=\frac{\partial \mathcal{L}}{\partial g^\alpha_{\mu_1\cdots\mu_s,\,\mu}}\text{.} \tag{5.21}$$

Definition 5.5: A scalar Hamiltonian density is the real valued function on $J^{1*}(J^{r-1}(E))$ defined by

$$\mathcal{H}=\sum_{s=0}^{r-1}\pi^{\mu_1\cdots\mu_s,\,\mu}_\alpha g^\alpha_{\mu_1\cdots\mu_s,\,\mu}-\mathcal{L}^{1,r-1}\omega\text{.} \tag{5.22}$$

In terms of $\mathcal{H}$, the Poincaré–Cartan form (5.11) is given by

$$\Theta^{1,r-1}=\sum_{s=0}^{r-1}(\pi^{\mu_1\cdots\mu_s,\,\mu}_\alpha dy^\alpha_{\mu_1\cdots\mu_s})\wedge\theta_\mu-\mathcal{H}\omega\text{.} \tag{5.23}$$

To obtain the set of Hamilton equations it is now sufficient to adapt (5.13) to the new definition space. Substituting $\psi^{1*,r-1}$ for $\psi^{1,r-1}$ and

$$X^{1*,r-1}=X^\mu\frac{\partial}{\partial x^\mu}+\sum_{s=0}^{r-1}X^\alpha_{\mu_1\cdots\mu_s}\frac{\partial}{\partial y^\alpha_{\mu_1\cdots\mu_s}}$$
$$+\sum_{s=0}^{r-1}X^{\mu_1\cdots\mu_s,\,\mu}_\alpha\frac{\partial}{\partial\pi^{\mu_1\cdots\mu_s,\,\mu}_\alpha} \tag{5.24}$$

for $X^{1,r-1}$, direct calculation from (5.15) gives

$$\frac{\partial\mathcal{H}}{\partial\pi^{\mu_1\cdots\mu_s,\,\mu}_\alpha}=\frac{\partial y^\alpha_{\mu_1\cdots\mu_s}}{\partial x^\mu}, \tag{5.25a}$$

$$\frac{\partial\mathcal{H}}{\partial y^\alpha_{\mu_1\cdots\mu_s}}=-\frac{\partial\pi^{\mu_1\cdots\mu_s,\,\mu}_\alpha}{\partial x^\mu},\quad s=0,1,\ldots,(r-1)\text{.} \tag{5.25b}$$

When the Legendre transformation is a diffeomorphism, recombination of (5.25a) and (5.25b) again reproduces the set of equations (5.10).

Clearly, the set of Euler–Lagrange equations (5.10) is not equivalent to (3.8). The reason is obvious: Since $\mathcal{L}^{1,r-1}$ is defined on $J^1(J^{r-1}(E))$ it will not correspond to the usual generalized Lagrangian density depending on a field and its r first derivatives, but rather to a Lagrangian density depending on several tensor fields and their first derivatives. However it is not difficult

to relate $\mathcal{L}^r$ (Definition 3.1) to the Lagrangian density $\mathcal{L}^{1,r-1}$ (Definition 5.1) and to prove that both formalisms then lead to identical results. This can be performed by observing that the cross sections $\psi^{1,r-1} \in \Gamma(J^1(J^{r-1}(E)))$ can be projected onto those $\psi^r \in \Gamma(J^r(E))$ by means of the trivial projection

$$(x^\mu, y^\alpha, y^\alpha_{\mu_1}, \cdots, y^\alpha_{\mu_1\cdots\mu_{r-1}}; g^\alpha_{,\mu}, g^\alpha_{\mu_1,\mu} \cdots g^\alpha_{\mu_1\cdots\mu_{r-1},\mu})$$

$$\xrightarrow{P} (x^\mu, y^\alpha, g^\alpha_{,\mu} \cdots g_{\mu_1\cdots\mu_{r-1},\mu}) \tag{5.26}$$

so that a $\mathcal{L}^r$ may be redefined on $J^1(J^{r-1}(E))$ by identifying its components with the last part of (5.26) and injecting them into $J^1(J^{r-1}(E))$. We may now apply the variational principle 5.1 to $\mathcal{L}^r$ incorporating the constraint $y^\alpha_{\mu_1\cdots\mu_{r-1}} = g^\alpha_{\mu_1\cdots\mu_{r-2},\mu}$. The natural way to incorporate a constraint in a Lagrangian formalism is the use of the Lagrange multipliers.[20] Thus, we define on $J^1(J^{r-1}(E))$

$$\mathcal{L}' = \mathcal{L} + \sum_{s=0}^{r-1} \lambda^{\mu_1\cdots\mu_s}_\alpha (y^\alpha_{\mu_1\cdots\mu_s} - g^\alpha_{\mu_1\cdots\mu_{s-1},\mu_s}). \tag{5.27}$$

Substitution of $\mathcal{L}'$ into (5.10) gives a set of equations for $\mathcal{L}$ coupled by the Lagrange multipliers, and their elimination leads to (3.8), Q.E.D.

ACKNOWLEDGMENT

The authors wish to thank Professor P. L. García for helpful discussions on his work.[7]

[1] P. Dedecker, *Calcul des Variations, Formes Différentielles et Champs Géodésiques*, Coll. Int. de Géometrie Différentielle, Strasbourg 1953 (C.N.R.S., Paris, 1954); *Calcul des Variations et Topologie Algébrique*, Thèse d'Agrégation, Univ. de Liège, 1957. (Published in Mémoires de la Soc. Royale des Sciences de Liège, 4ème série, **XIX**, Fasc. 1.)
[2] A. Trautman, Commun. Math. Phys. **6**, 248 (1967). [For a review of fibre bundles associated with space–time see Rep. Math. Phys. **1**, 29 (1970).]
[3] P. L. García, "Geometría Simpléctica en la Teoría Clásica de Campos," Collect. Math. **19**, 73 (1968); "Estructura Compleja en la Teoría Clásica de Campos," **19**, 155 (1968).
[4] R. Hermann, *Vector Bundles in Mathematical Physics*, Vol. 1 (Benjamin, New York, 1970).
[5] J. Śniatycki, Proc. Cambridge Phil. Soc. **68**, 475 (1970).
[6] H. Goldschmidt and S. Sternberg, Ann. Inst. Fourier (Grenoble) **23**, 203 (1973).
[7] P. L. García, Symposia Math. **XIV**, 219 (1974).
[8] P. L. García and A. Pérez-Rendón, Archiv. Ration. Mech. Anal. **43**, 101 (1971).
[9] D. Krupka, Rep. Math. Phys. (Toruń) **5**, 355 (1974).
[10] D. Krupka and A. Trautman, Bull. Acad. Pol. Sci. **XXII**, 207 (1974).
[11] D. Krupka, Folia Fac. Sci. Nat. Univ. Purkinianae Brunensis (Physica) **XIV**, 1 (1975).
[12] D. Krupka, J. Math. Anal. Appl. **49**, 180 (1975).
[13] D. Krupka, J. Math. Anal. Appl. **49**, 469 (1975).
[14] W. Szczyrba, Ann. Pol. Math. **32**, 145 (1976).
[15] J. Kijowski and W. Szczyrba, Commun. Math. Phys. **46**, 183 (1976).
[16] D. Krupka, Arch. Math. 2, Scripta Fac. Sci. Nat. UJEP Brunensis **XII**, 99 (1976).
[17] M. C. Ehresmann, C. R. Acad. Sci. Paris **233**, 598, 777, 1081 (1951).
[18] H. J. Melosh, Ph. D. Thesis, California Institute of Technology, 1973; Phys. Rev. D **9**, 1095 (1974).
[19] V. Aldaya and J. A. de Azcárraga, "Vector bundles, rth order Noether invariants and canonical symmetries in Lagrangian field theory," J. Math. Phys. **19**, 1876 (1978).
[20] R. S. Palais, *Foundations of Global Nonlinear Analysis* (Benjamin, New York, 1968), p. 106.

Vector bundles, rth order Noether invariants and canonical symmetries in Lagrangian field theory

V. Aldaya and J. A. de Azcárraga

Física Matemática, Facultad de Ciencias, Universidad de Salamanca, Spain
(Received 24 August 1977)

We present a formulation of the "canonical" transformations of the Dirac theory by using the bundle of r-jets associated with the Dirac vector bundle. This allows us, by means of a variational principle previously introduced, to study in a natural way (through an appropriate Noether theorem) the role which these "canonical" theories play in the definition of "new" symmetries. The limit $r\to\infty$ corresponds to the equivalence between the canonical and the ordinary symmetries.

1. INTRODUCTION AND MOTIVATION

The analysis and motivation of the different transformations of the Dirac equation (and also of higher order spin equations) has been the subject of continuous study since Foldy and Wouthuysen[1] introduced their transformation allowing invariants to be associated with certain dynamical variables (spin, orbital angular momentum) which in their usual Dirac form are not constants of the motion. Several so-called "canonical" transformations have since been introduced, and their implications considered from a variety of viewpoints.[2] The last of these is the Melosh transformation,[3] which relates the "current" and "constituent" quarks[4] and which, in so doing, provides an expression for the generators of the $SU(6)_w$ symmetry of the strong interactions previously only known for those of its SU(3) subgroup through the conserved vector current hypotehesis.

Both the FW and the Melosh transformations present a feature in common: They lead to constants of the motion of the usual Dirac Lagrangian theory by starting with simple expressions which are constants of the motion in another theory, usually referred to in physical literature as the FW (or Melosh) "representation," this term indicating (Sec. 3), not only a mere change of basis but a new starting point used to define dynamical quantities. In this paper we assume the task of identifying what this new starting point is and, as a result, the meaning of the canonical transformations (as, for instance, the FW transformation) when taken up to rth order will become clear in the context of Lagrangian field theory.

This will be done by using the calculus of variations (Euler-Lagrange equations, Noether theorem) on $J^1(J^r(E))$,[5] where $J^r(E)$ is the bundle of the r-jets of the vector bundle $E\xrightarrow{\pi}M$, where in our applications E is the Dirac bundle $R^i\otimes R^\alpha\otimes M\xrightarrow{\pi}M$. We follow the same notation of Ref. 5; in particular, α is the spinorial index, i is a possible inner symmetry index [such as SU(3)] and M, the Minkowski space, is labeled by the coordinates x^μ. In general $\Gamma(E)$, the module of the cross sections of (E,M,π), will include the support space of a unitary representation of the Poincaré group.

This paper is organized as follows. Section 2 is devoted to the formulation of the variational principle and summarizes the results[5] which are used in the ensuing discussions. In Sec. 3 a Lagrangian density $\mathcal{L}^{1,r}$ on $J^1(J^r(E))$ is constructed whose space of solutions $\mathcal{U}_{\mathcal{L}}^{\infty}$ is unitarily related to $\mathcal{U}_{\mathcal{L}}$, the space of solutions of the ordinary Dirac equation. The additional symmetries will be introduced *through* the formulation of the *Noether theorem* for $\mathcal{L}^{1,\infty}$. This Lagrangian $\mathcal{L}^{1,\infty}$ will be precisely the transformed usual Dirac Lagrangian, and in this way our approach will exhibit the mechanism by which the "canonical" transformations may be used to define "new" symmetries. Finally Sec. 4 summarizes the results and includes some comments on a possible description of "extended" particles.[10]

2. HAMILTON VARIATIONAL PRINCIPLE ON $J^1(J^r(E))$

The reader is referred to Ref. 5 for all the results contained in this section.

The starting point is the vector bundle $E\xrightarrow{\pi}M$, with fibre of R^n type (here E will be a globally trivial bundle on M, π being the Cartesian projection). The coordinate system for E will be (x^μ,y^α), where x^μ is a coordinate system for M and y^α is the coordinate system of the fiber part of E.

In our variational principle the solutions will be given by cross sections of $J^r(E)$ (in the limiting case, r will be allowed to go to infinity; for a differentiable structure on $J^\infty(E)$ see Ref. 6) rather than by sections of E as is the case for the usual variational problem on $J^1(E)[\equiv J^1(J^0(E))]$. Accordingly, the Lagrangian density $\mathcal{L}^{1,r}$ defining the physical theory will be a real function on $J^1(J^r(E))$—the usual Lagrangian densities, involving only first order derivatives of the field, are defined on $J^1(E)$—and we shall perform a first order variational calculus on $J^r(E)$.

The coordinate system on $J^1(J^r(E))$ will be given by

$$(x^\mu,y^\alpha,y^\alpha_{\mu_1},\ldots,y^\alpha_{\mu_1\cdots\mu_r};\ g^\alpha_{,\mu},g^\alpha_{\mu_1,\mu},\ldots,g^\alpha_{\mu_1\cdots\mu_r,\mu}), \tag{2.1}$$

where ω is the volume form on M, $dx^0\wedge dx^1\wedge dx^2\wedge dx^3$. for $J^r(E)$ and $g^\alpha_{\mu_1\cdots\mu_s,\mu}$ are the "velocities" associated with $y^\alpha_{\mu_1\cdots\mu_s}$. Note that in this coordinate system $y^\alpha_{\mu_1\cdots\mu_s}$ is *not* necessarily $\partial_{\mu_s}y^\alpha_{\mu_1\cdots\mu_{s-1}}$ nor $g^\alpha_{\mu_1\cdots\mu_s,\mu}$ is $\partial_\mu y^\alpha_{\mu_1\cdots\mu_s}$. These relations are fulfilled only for jet prolongations: only for $\{j^r(\psi)\equiv\overline{\psi}^r\}\subset\{\psi^r\,|\,\psi^r\subset\Gamma(J^r(E))\}$ is the first condition satisfied and only for $\{j^1(\psi^r)\equiv\overline{\psi^r}^1\}\subset\{\psi^{r,1}\subset\Gamma(J^1(J^r(E)))\}$ is the second.

 0022-2488/78/1909-1876$1.00

Given a $\mathcal{L}^{1,r}$ on $J^1(J^r(E))$ the action functional $I^{1,r}$: $\Gamma(J^r(E)) \to R$ is defined by

$$I^{1,r}(\psi^r) = \int_{j^1(\psi^r)(M)} \mathcal{L}^{1,r}(j^1(\psi^r))\omega, \tag{2.2}$$

where ω is the volume form on M, $dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3$. The variation $\delta I^{1,r}$ of $I^{1,r}$ for a point of $\Gamma(J^r(E))$ is defined as the following linear application:

$$(\delta I^{1,r})_{\psi^r}: X^r \in \Gamma(T(J^r(E))) \to \int_M [j^1(\psi^r)]^* \{\mathcal{L}_{\overline{X}^{r1}}(\mathcal{L}^{1,r}\omega)\} \tag{2.3}$$

where $\overline{X}^{r1} \equiv j^1(X^r) \in \Gamma(T(J^1(E)))$ and $[j^1(\psi^r)]^*\mathcal{L} = \mathcal{L}(j^1(\psi^r))$. (2.2) and (2.3) allow us to define the following.

Hamilton principle: The trajectories of the variational problem constituting the space $\mathcal{U}^{1,r}$ are the solutions of

$$(\delta I^{1,r})_{\psi^r}(X^r) = 0 \quad \forall \quad X^r \in \Gamma(T(J^r(E))). \tag{2.4}$$

(2.4) leads to the set of Euler–Lagrange equations

$$\frac{\partial \mathcal{L}^{1,r}}{\partial y^\alpha_{\mu_1\cdots\mu_s}} - \frac{d}{dx^\mu}\left(\frac{\partial \mathcal{L}^{1,r}}{g^\alpha_{\mu_1\cdots\mu_s,\mu}}\right) = 0, \quad s = 1, \ldots, r, \tag{2.5}$$

where *in this expression* $g^\alpha_{\mu_1\cdots\mu_s,\mu} = \partial_\mu y^\alpha_{\mu_1\cdots\mu_s}$.

The Noether theorem may be formulated in a convenient form if we first introduce the Poincaré–Cartan form Θ to define the *modified Hamilton principle*, which states that the solutions of the variational principle are the solutions of the equation

$$(\delta I'^{1,r})_{\psi^{1,r}}(X^{1,r}) = 0 \quad \forall \quad X^{1,r} \in \Gamma(T(J^1(J^r(E)))), \tag{2.6}$$

where

$$I'^{1,r}(\psi^{1,r}) = \int_{\psi^{1,r}} \Theta^{1,r} \tag{2.7}$$

and

$$\Theta^{1,r} = \mathcal{L}^{1,r}\omega + \sum_{s=0}^{r} \theta^\alpha_{\mu_1\cdots\mu_s} \wedge \Omega^{\mu_1\cdots\mu_s}_\alpha. \tag{2.8}$$

In (2.8), $\theta^\alpha_{\mu_1\cdots\mu_s}$ is the structure form of the bundle $J^1(J^r(E))$ [Ref. 5, Eq. (5.3)] and $\Omega^{\mu_1\cdots\mu_s}_\alpha$ is the form used in defining the Legendre transformation [Ref. 5, Eq. (5.12)]. For "regular" Lagrangians, both principles lead to the same space of solutions and are thus equivalent.

The *Noether theorem* and the corresponding Noether current may now be established: If $\overline{X}^{r1}_s$ generates a symmetry transformation of the variational problem, i.e., $L_{\overline{X}^{r1}_s}(\mathcal{L}^{1,r}\omega) = d\Delta$ (Δ is a three form not depending on the g^αs), then

$$d(i_{\overline{X}^{r1}_s}\Theta^{1,r} - \Delta)\big|_{\overline{\psi}^{r1}} = 0 \quad \forall \quad \psi^r \in \mathcal{U}^{1,r}. \tag{2.9}$$

This leads, assuming $\Delta = 0$, to the following expression for the conserved $[\partial^\mu(j_\mu - \Delta_\mu) = 0]$ current (*on* $\overline{\psi}^{r1}$, $\psi^r \in \mathcal{U}^{1,r}$):

$$j^{(1,r)\mu} = \mathcal{L}^{1,r}X^\mu + \sum_{s=0}^{r}(X^\alpha_{\mu_1\cdots\mu_s} - g^\alpha_{\mu_1\cdots\mu_s,\nu}X^\nu)\frac{\partial \mathcal{L}^{1,r}}{\partial g^\alpha_{\mu_1\cdots\mu_s,\mu}}, \tag{2.10}$$

where X^μ, $X^\alpha_{\mu_1\cdots\mu_s}$ are the components of the vector field X_s on $J^r(E)$.

3. SYMMETRIES AND "CANONICAL" TRANSFORMATIONS

We now turn to the consideration of the symmetries of nonlocal type which are introduced by means of "canonical" transformations. For example, it is well known that, although

$$\mathbf{S} = \int d^3x \psi^\dagger_D \Sigma \psi_D \tag{3.1}$$

and

$$\{F\} = \int d^3x\, \psi^\dagger_D \{\mathbf{1}, \Sigma_x, \beta\Sigma_x, \beta\Sigma_y\} \otimes \tfrac{1}{2}\lambda^i \psi_D, \quad i = 0, 1, \ldots, 8, \tag{3.2}$$

which generate the SU(2) (spin) and the $U(6)_W$ (currents) algebras, are not symmetries of the Dirac theory, the use of the FW and Melosh transformations allows the definition of a conserved mean spin and $U(6)_W$ (strong) algebras. More precisely, we ask ourselves if it is possible to obtain—*via the Noether theorem*—nonlocal charges associated with simple (local) *transformations of the Dirac vector bundle* as, for instance, the vertical rotations on E. In this way, we shall show how the formalism of Sec. 2 may be used to describe a more general type of symmetries (i.e., those given by integrals of nonlocal operators) which appear to be relevant in the physical world and whose existence seems to be related to a nonlocal structure of the hadronic particles themselves. In fact, there is no compelling reason, for instance, which requires the symmetries of strong interaction physics to be generated by integrals of local current densities. (Of course, there is also no reason why a Lagrangian approach should necessarily be relevant in strong interaction physics.) On the other hand, the variational principle on $J^1(J^\infty(E))$ of Sec. 2 gives as solutions cross sections of $J^\infty(E)$—fields of the form $\{\psi^\alpha(x), \psi^\alpha_\mu(x), \ldots, \psi^\alpha_{\mu_1\cdots\mu_s}(x)\cdots\}$—which indicates that, to know the physical system in a point x_0 of M, we require, in principle, not only the knowledge of $\psi^\alpha(x_0)$ in that point, but the value of a certain field Φ^α of E in all the points of an open set containing x_0 [as is suggested by the interpretation of the cross-section of $J^\infty(E)$ as the coefficients of a Taylor expansion about x_0].

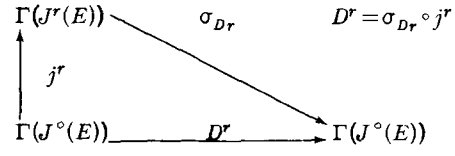
Let $\mathcal{L}^{1,0}(=\mathcal{L}^1)$ be a Lagrangian density whose dependence on the fields is of the usual form; it is then an application $\mathcal{L}^{1,0}: J^1(J^0(E)) \to R$. Consider now a differential operator D^r on $\Gamma(E)$ of order r, i.e., such that

$$D^r: \Gamma(E) \to \Gamma(E) \tag{3.3}$$

linearly and

$$D^r: \Gamma^x_{r+1} \to \Gamma^x_1(E), \tag{3.4}$$

where $\Gamma^x_s(E)$ is the submodule of $\Gamma(E)$ of the cross sections which are zero at order s at x (i.e., the sections and their t first derivatives, $t = 1, \ldots, s$). D^r admits the following factorization:



which allows us to define the Lagrangian $\mathcal{L}^{1,r}$ associated with D^r as

$$\mathcal{L}^{1,r}: J^1(J^r(E)) \to R \tag{3.5}$$

such that for the 1-jet prolongation of the cross sections [which is what is relevant in the Euler–Lagrange equations, see Eqs. (2.3) and (2.5)]

$$\mathcal{L}^{1,r}(j^1(\psi^r)) = \mathcal{L}^{1,0}(j^1(\sigma_{D^r}(\psi^r))). \tag{3.6}$$

(3.6) shows, in the limit $r \to \infty$, the relation existing between the ordinary Lagrangian and the canonically transformed one; in that limit, D^r $(r \to \infty)$ is the canonical transformation or, more exactly, its adjoint $T^\dagger$ [which is defined according to the scalar product $\Gamma(E)$] and both theories are unitarily related.

Let us apply the above scheme taking the Dirac Lagrangian and the Foldy–Wouthuysen transformation as an example. We shall perform the calculations in first order of the transformation, which will be sufficient to illustrate the theory. In the coordinate system for $J^1(J^0(E))$

$$(x^\mu,\ y^\alpha,\ y_{\alpha *};\ g^\alpha_{,\mu},\ g_{\alpha *,\mu}) \tag{3.7}$$

[an obvious modification of (2.1) to accomodate the complex conjugation] the Dirac Lagrangian is written

$$\mathcal{L}^{1,0} = i y_{\alpha *}(\gamma^0\gamma^\mu)^\alpha_\beta g^\beta_{,\mu} - m y_{\alpha *}(\gamma^0)^\alpha_\beta y^\beta. \tag{3.8}$$

The FW transformation[1] is defined at first order by the application

$$\sigma_{(\mathrm{FW}^1)^\dagger}\ (\equiv \sigma_1^\dagger): J^1(E) \to J^0(E) \tag{3.9}$$

such that

$$(x^\mu,\ y^\alpha,\ y^\alpha_\mu) \xrightarrow{\sigma_1^\dagger} (x^\mu,\ y^\alpha + (i/2m)(\gamma^k)^\alpha_\beta y^\beta_k), \tag{3.10}$$

$$(x^\mu,\ y_{\alpha *},\ y_{\alpha *\mu}) \xrightarrow{\sigma_1^\dagger} (x^\mu,\ y_{\alpha *} + (i/2m)\, y_{\beta *k}(\gamma^k)^\beta_\alpha) \tag{3.11}$$

$(\alpha, \beta, \gamma = 1,2,3,4;\ i,j,k,m = 1,2,3;\ \mu,\nu,\sigma = 0,1,2,3)$. Accordingly,

$$\begin{aligned}
&j^1(\psi) = (y^\alpha,\ \partial_\mu y^\alpha),\\
&j^1(\sigma_1^\dagger\psi) = (y^\alpha_\gamma + (i/2m)(\gamma^k)^\alpha_\beta y^\beta_k,\ \partial_\mu y^\alpha + (i/2m)(\gamma^k)^\alpha_\beta \partial_\mu y^\beta_k),\\
&g^\alpha_{,\mu}(j^1(\sigma_1^\dagger\psi)) = \partial_\mu y^\alpha + (i/2m)(\gamma^k)^\alpha_\gamma \partial_\mu y^\gamma_k\\
&\qquad = (g^\alpha_{,\mu} + (i/2m)(\gamma^k)^\alpha_\gamma g^\gamma_{k,\mu})\big|_{j^1(\psi)},
\end{aligned} \tag{3.12}$$

where $|_{j^1(\psi)}$ means "restriction to 1-jet prolongation," and $\mathcal{L}^{1,1}$ takes the form

$$\begin{aligned}
\mathcal{L}^{1,1} = {}& i y_{\alpha *}(\gamma^0\gamma^\mu)^\alpha_\beta \partial_\mu y^\beta - \frac{1}{2m} y_{\alpha *}(\gamma^0\gamma^\mu\gamma^m)^\alpha_\beta \partial_\mu y^\beta_m\\
&- \frac{1}{2m} y_{\beta *k}(\gamma^k\gamma^0\gamma^\mu)^\beta_\gamma \partial_\mu y^\gamma - (i/4m^2) y_{\beta *k}(\gamma^k\gamma^0\gamma^\mu\gamma^m)^\beta_\alpha \partial_\mu y^\alpha_m\\
&- m y_{\alpha *}(\gamma^0)^\alpha_\beta y^\beta - \tfrac{1}{2} i y_{\alpha *}(\gamma^0\gamma^m)^\alpha_\beta y^\beta_m\\
&- \tfrac{1}{2} i y_{\beta *k}(\gamma^k\gamma^0)^\beta_\alpha y^\alpha + (1/4m) y_{\beta *k}(\gamma^k\gamma^0\gamma^m)^\beta_\alpha g^\alpha_m.
\end{aligned} \tag{3.13}$$

The Euler–Lagrange equations may now be derived with the result

$$i\partial_\mu[\bar\psi - (i/2m)\bar\psi_k\gamma^k]\gamma^\mu + m[\bar\psi - (i/2m)\bar\psi_k\gamma^k] = 0, \tag{3.14a}$$

where $\bar\psi \equiv \psi^\dagger\gamma^0$ and $\psi^\dagger$ represents the generic cross section of components $y_{\alpha *}$. Proceeding at second order (we shall omit the explicit form of $\mathcal{L}^{1,2}$), we obtain

$$\begin{aligned}
&i\partial_\mu\left(\bar\psi - \frac{i}{2m}\bar\psi_k\gamma^k + \frac{1}{8m^2}\delta^{km}\bar\psi_{km}\right)\gamma^\mu\\
&\quad + m\left(\bar\psi - \frac{i}{2m}\bar\psi_k\gamma^k + \frac{1}{8m^2}\delta^{km}\psi_{km}\right) = 0.
\end{aligned} \tag{3.14b}$$

It is immediately observed that the solutions ψ of (3.14a), (3.14b) are the first and second order FW fields. In the limit $r \to \infty$ we would find the FW equation $(i\partial_0 - \omega\gamma^0)\psi = 0$.

To study the Poincaré invariance in $\mathcal{L}^{1,r}$ we require first that the action of the Poincaré generators on $J^r(E)$ be defined. This may be accomplished with the help of the operator D^r in the same way it was used for constructing $\mathcal{L}^{1,r}$. Nevertheless, the action of the rotation and the translation subgroups on $J^r(E)$ coincides with the jet prolongation of the usual action on $J^0(E)$. Thus, if $J_{(i)}$ and $P_{(\mu)}$ are the usual rotation and translation generators on $J^0(E)$, we define $\bar J^r{}_{(i)} = j^r(J_{(i)})$ and $\bar P^r{}_{(\mu)} = j^r(P_{(\mu)})$ on $J^r(E)$.

The Noether currents of the theory are obtained from the following general expression [compare (2.10)]:

$$\begin{aligned}
j^{(1,1)\mu}_{(a)} = {}& (X^\alpha_{(a)} - g^\alpha_{,\nu} X^\nu_{(a)})\frac{\partial \mathcal{L}^{1,1}}{\partial g^\alpha_{,\mu}} + (X_{(a)\alpha *} - g_{\alpha *,\nu} X^\nu_{(a)})\frac{\partial \mathcal{L}^{1,1}}{\partial g_{\alpha *,\mu}}\\
&+ (X^\alpha_{(a)\sigma} - g^\alpha_{\sigma,\nu} X^\nu_{(a)})\frac{\partial \mathcal{L}^{1,1}}{\partial g^\alpha_{\sigma,\mu}}\\
&+ (X_{(a)\alpha *\sigma} - g_{\alpha *\sigma,\nu} X^\nu_{(a)})\frac{\partial \mathcal{L}^{1,1}}{\partial g_{\alpha *\sigma,\mu}} + \mathcal{L}^{1,1} X^\mu_{(a)}.
\end{aligned} \tag{3.15}$$

In our case, with $\mathcal{L}^{1,1}$ given by (3.13), (3.15) reduces to the first and third terms since

$$\frac{\partial \mathcal{L}^{1,1}}{\partial g_{\alpha *,\mu}} = \frac{\partial \mathcal{L}^{1,1}}{\partial g_{\alpha *\sigma,\mu}} = 0, \tag{3.16}$$

Only the components on $J^1(E)$ of $\bar{\bar X}^1_{(a)}$ contribute to (3.15),

$$\begin{aligned}
\bar X^1_{(a)} = {}& X^\mu_{(a)}\frac{\partial}{\partial x^\mu} + X^\alpha_{(a)}\frac{\partial}{\partial y^\alpha} + X_{(a)\alpha *}\frac{\partial}{\partial y_{\alpha *}}\\
&+ \bar X^\alpha_{(a)\mu}\frac{\partial}{\partial y^\alpha_\mu} + \bar X_{(a)\alpha *\mu}\frac{\partial}{\partial y_{\alpha *\mu}},
\end{aligned} \tag{3.17}$$

where[5]

$$\begin{aligned}
&\bar X^\alpha_\nu = \frac{\partial X^\alpha}{\partial x^\nu} - y^\alpha_\mu\frac{\partial X^\mu}{\partial x^\nu} + A^\alpha_\beta y^\beta_\nu + A^{\alpha\beta *} y_{\beta *\nu},\\
&\bar X_{\alpha *\nu} = \frac{\partial X_{\alpha *}}{\partial x_\nu} - y_{\alpha *\mu}\frac{\partial X^\mu}{\partial x_\nu} + A_{\alpha *\beta} y^\beta_\nu + A^{\beta *}_{\alpha *} y_{\beta *\nu}
\end{aligned} \tag{3.18}$$

and

$$\begin{aligned}
&A^\alpha_\beta = \frac{\partial X^\alpha}{\partial y^\beta} - y^\alpha_\mu\frac{\partial X^\mu}{\partial y^\beta},\quad A_{\alpha *\beta} = \frac{\partial X_{\alpha *}}{\partial y_\beta} - y_{\alpha *\nu}\frac{\partial X^\nu}{\partial y^\beta}\\
&A^{\alpha\beta *} = \frac{\partial X^\alpha}{\partial y_{\beta *}} - y^\alpha_\mu\frac{\partial X^\mu}{\partial y_{\beta *}},\quad A^{\beta *}_{\alpha *} = \frac{\partial X_{\alpha *}}{\partial y_{\beta *}} - y_{\alpha *\nu}\frac{\partial X^\nu}{\partial y_{\beta *}},
\end{aligned} \tag{3.19}$$

since the part of X on $J^1(E)$, $\bar X^\alpha_\mu$, etc., corresponds to the jet lifting from the part acting on E.

As a first illustration we evaluate the Hamiltonian density. Since the 4-translations have only components

on the basis M, $X_{(\mu)}=\delta^\nu_\mu\,\partial/\partial X^\nu$ on E, on $J^1(E)$ and also on $J^1(J^1(E))$. Thus $X^\nu_{(\mu)}=\delta^\nu_\mu$ and the Hamiltonian density (whose charge generates the time translation) is given by (the minus sign is added for convenience)

$$-j^0_{(0)}=g^\alpha_{\cdot,0}\frac{\partial \mathcal{L}^{1,1}}{\partial g^\alpha_{\cdot,0}}+g^\alpha_{\sigma,0}\frac{\partial \mathcal{L}^{1,1}}{\partial g^\alpha_{\sigma,0}}-\mathcal{L}^{1,1}$$

$$=\{\psi^\dagger+\frac{i}{2m}\psi^\dagger_k\gamma^k\}\{\gamma^0\boldsymbol{\gamma}\mathbf{p}+m\gamma^0\}\{\psi+\frac{i}{2m}\gamma^m\psi_m\}. \tag{3.20}$$

Since the fields appearing in (3.20) are solutions of (3.14), (3.20) is the usual Hamiltonian in this approximation. In the limit $r\to\infty$ (3.20) could be written either in the form $\psi^\dagger_{\mathrm{FW}}\omega\gamma^0\psi_{\mathrm{FW}}$ or as $\psi^\dagger_D(\gamma^0\boldsymbol{\gamma}\mathbf{p}+m)\psi_D$ (both expressions are unitarily equivalent since $\psi_{\mathrm{FW}}=(\mathrm{FW})\psi_D$, the subindex FW defining the "FW representation").

For the rotations (corresponding to the total angular momentum $\mathbf{J}=\mathbf{L}+\mathbf{S}$) we have, on $J^0(E)$,

$$X^J_{(i)}=\epsilon_{ij}{}^\sigma_{\cdot}X^j\frac{\partial}{\partial x^\sigma}+\frac{i}{2}(\Sigma_i)^\alpha_\beta y^\beta\frac{\partial}{\partial y^\beta}-\frac{i}{2}y_{\beta*}(\Sigma_i)^\beta_\alpha\frac{\partial}{\partial y_{\alpha*}} \tag{3.21}$$

and its jet prolongation into $J^1(E)$ is given by

$$X^J_{(i)}+\overline{X}^\alpha_{(i)\nu}\frac{\partial}{\partial y^\alpha_\nu}+\overline{X}_{(i)\alpha*\nu}\frac{\partial}{\partial y_{\alpha*\nu}}, \tag{3.22}$$

where, for instance,

$$\overline{X}^\alpha_{(i)\nu}=-y^\alpha_\mu\epsilon_{ij}{}^\mu_{\cdot}\delta^j_\nu+y^\beta_\nu\tfrac{1}{2}i(\Sigma_i)^\alpha_\beta. \tag{3.23}$$

In the above expressions Σ is the spin matrix for the Dirac fields and ϵ the antisymmetric tensor ($\epsilon_{ij}{}^\mu_{\cdot}=0$ for $\mu=0$). Thus, the angular momentum charge densities are given by

$$-j^0_{(i)}=\{\psi^\dagger+\frac{i}{2m}\psi^\dagger_k\gamma^k\}\left(\frac{\Sigma_i}{2}+(\mathbf{x}\wedge\mathbf{p})i\right)\{\psi+\frac{i}{2m}\gamma^m\psi_m\}, \tag{3.24}$$

which for $r\to\infty$ reads $\psi^\dagger_{\mathrm{FW}}(\frac{1}{2}\Sigma_{\mathrm{FW}}+\mathbf{X}_{\mathrm{FW}}\wedge\mathbf{p})\psi_{\mathrm{FW}}$, FW representation of the familiar expression $\psi^\dagger_D(\frac{1}{2}\Sigma+\mathbf{x}\wedge\mathbf{p})\psi_D$ since $\Sigma_{\mathrm{FW}}=(\mathrm{FW})\Sigma(\mathrm{FW})^\dagger$ and $\mathbf{X}_{\mathrm{FW}}=(\mathrm{FW})\mathbf{x}(\mathrm{FW})^\dagger$. Note that the invariance of $\mathcal{L}^{1,1}$ under (3.21) requires $\psi_k=\partial_k\psi$, as is the case of the first order FW-transformed Lagrangian.

The conserved quantities, e.g., (3.24) and (3.20) are, of course, not new; they reflect the Poincaré invariance of the theory at the given order and, as has been shown, these expressions are nothing but the usual ones in the FW representation. However, as mentioned previously, $\mathcal{L}^{1,1}$ may be used to define conserved quantities at the given order of approximation. This is the case with the vertical rotations

$$(X^J)^V=\frac{i}{2}(\Sigma_i)^\alpha_\beta y^\beta\frac{\partial}{\partial y^\alpha}-\frac{i}{2}y_{\beta*}(\Sigma_i)^\beta_\alpha\frac{\partial}{\partial y_{\alpha*}}, \tag{3.25}$$

which act only in the fibers of the Dirac bundle E (hence their name). Although (3.25) do not lead to a constant of the motion of the $\mathcal{L}^{1,0}$ theory, it is simple to see that they are conserved in the $\mathcal{L}^{1,r}$ theory up to order $(1/m^2)^r$. To obtain the charge densities, we again apply the Noether theorem to $\mathcal{L}^{1,r}$ and obtain

$$-j^{V_0}_{(i)}=\tfrac{1}{2}\psi^\dagger\Sigma_i\psi+(\text{terms of order } 1/m^2\cdots) \tag{3.26}$$

and, when $r\to\infty$

$$-j^{V_0}_{(i)}=\tfrac{1}{2}\psi^\dagger\Sigma_i\psi. \tag{3.27}$$

since the fields of (3.27) are solutions of $(\gamma^0-\omega\gamma^0)\psi=0$, i.e., they are ψ_{FW} fields and we get, since $\psi_{\mathrm{FW}}=\mathrm{FW}\psi_D$,

$$-j^{V_0}_{(i)}=\psi^\dagger_D\Sigma^M_i\psi_D, \tag{3.28}$$

where Σ^M is the "mean spin" of Foldy and Wouthuysen, which is conserved (as it is $\mathbf{X}^M\wedge\mathbf{p}$).

Analogous considerations can be made to show how this scheme works for the Melosh transformation, taking

$$X_3=\tfrac{1}{2}i(\Sigma_3)^\alpha_\beta y^\beta\frac{\partial}{\partial y^\alpha} \tag{3.29}$$

$$X_\perp=\tfrac{1}{2}i(\Sigma_\perp\gamma^0)^\alpha_\beta y^\beta\frac{\partial}{\partial y^\alpha} \tag{3.30}$$

as the components of a "rotation" acting on E which, via the Noether theorem, will lead to the conserved "strong" charges. However, the motivation for introducing the new symmetry operators in this way is less clear from the physical point of view. In fact, the $U(6)_W$ symmetry has its full relevance in the infinite momentum limit and, in that limit, the corresponding charges are already a symmetry.[7] Thus, other motivations have been introduced (see the second paper quoted in Ref. 3 and, e.g., Ref. 8 and references therein) to justify the Melosh transformation. Some of these difficulties come from the fact that the infinite momentum procedure is a limit rather than a transformation, and accordingly it is not immediately clear how to account for it in the present framework. We hope to come back to this point in the future.

4. SUMMARY

The formulation of the variational calculus and the associated Noether problem on the vector bundle $J^1(J^r(E))$ and, in particular, on $J^1(J^\infty(E))$ has allowed us to give a precise definition of the "canonical transformations," often introduced in field theory, in a Lagrangian framework. This approach is specially relevant in showing the mechanism which allows the definition of new symmetries of the theory which were not apparent in its original formulation. As an illustration we have considered the case of the spin of a Dirac particle (3.1) which, by means of the nonlocal transformation $(\mathrm{FW})^*$ leads to the "mean spin"[1] $\mathbf{S}^M$

$$\mathbf{S}^M=\int d^3x\,\psi^\dagger_{\mathrm{FW}}\Sigma\psi_{\mathrm{FW}}=\int d^3x\,\psi^\dagger_D\Sigma^M\psi_D. \tag{4.1}$$

These new charges $\mathbf{S}^M$ cannot be identified with space–time generators acting on the Dirac bundle E on which the fields are defined. However, in our formulation they still admit the usual simple interpretation: they are the charges associated, through the Noether theorem for $\mathcal{L}^{1,r}$, with the generators $(X^J)^V$ corresponding to the vertical rotations which act on the fibre space E in the usual way.

In the case of Melosh a similar approach leads to the new conserved charges $SU(2)_W$ or $U(6)_W$ (bearing in mind the comments made at the end of the preceding Section).

Finally, we would like to comment that the association of a field Φ defined in an open set about x_0 (Sec. 3)

with the solutions (cross sections) Ψ^∞ of the variational problem might provide a description of an "extended" hadron particle. However, this association is not automatically guaranteed since the Whitney prolongation theorem[9] esatblishes that if a cross section Ψ^∞ of $J^\infty(E)$ is a Whitney function of class $C^\infty(K)$ on the compact K, then Φ is the same for all $x_0 \in K$, i.e., $\Psi^\infty = j^\infty(\Phi)$, in which case we would be led again to a local description of the physical system.

[1]L. L. Foldy and S. A. Wouthuysen, Phys. Rev. **78**, 29 (1950). See also J. D. Bjorken and S. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, New York, 1964), Chap. 4.

[2]An extensive list of references can by found in the paper by A. J. Kálnay, in *Problems in the Foundations of Physics*, edited by M. Bunge (Springer, 1971).

[3]H. J. Melosh, Ph. D. thesis (California Institute of Technology, 1973); Phys. Rev. D **9**, 1095 (1974).

[4]M. Gell-Mann, in *Elementary Particle Physics*, Acta Physica Austr. Suppl. **9**, 733 (1972).

[5]V. Aldaya and J. A. de Azcárraga, J. Math. Phys. **19**, xxx (1978).

[6]D. Krupka, Folia Fac. Scient. Nat. Univ. Purkinianae Brunensis (Physica) XIV, 1 (1975).

[7]S. P. de Alwis, Nucl. Phys. B **55**, 427 (1973).

[8]V. Aldaya and J. A. de Azcárraga, Phys. Rev. D **14**, 1049 (1976).

[9]See J. C. Tougeron, *Idéaux de Fonctions Différentiables*, Ergebnisse der Mathematik, Vol. 71 (Springer-Verlag, Berlin, 1972), p. 77.

[10]*Note added in proof*: For another approach to a theory of extended hadrons in terms of fibre bundles, see W. Drechsler, Fortschr. Phys. **23**, 607 (1975); *Group Theoretical Methods in Physics* (Nijmegen, 1975), (Lecture Notes in Physics **50**, Springer, Berlin, 1976); Phys. Lett. B **66**, 439 (1977); Nuovo Cimento A **41**, 597 (1977).

Nuclearity of some spaces of C^∞ vectors in induced representations

D. Arnal

Physique-Mathématique, Faculté des Sciences Mirande, 21000—Dijon, France
(Received 19 October 1977)

We give conditions on irreducible unitary induced representations of inhomogeneous Lie groups ensuring nuclearity of the space of differentiable vectors. We study the consistency of these conditions. Finally we show that our conditions hold almost everywhere with respect to the Plancherel measure for some generalizations of the Poincaré group.

INTRODUCTION

In this paper, we study the topology of some spaces of C^∞ vectors in unitary irreducible representations (UIR) U of inhomogeneous Lie groups $G \circ T$, G semisimple and T normal Abelian. In general $G \cdot T$ is a symmetry or a dynamical group of the system.

The space $\mathcal{D}$ of C^∞ vectors with its topology appears as a "good domain of states" for many important observables (essentially self-adjoint operators): for instance, those which are representatives of elements of the universal enveloping algebra $\mathcal{U}$ of $G \cdot T$, or those which commute with a symmetric elliptic element of $\mathcal{U}$. All these operators leave $\mathcal{D}$ invariant and the topology of $\mathcal{D}$ is the coarsest, making all maps $\mathcal{D} \to \mathcal{D}$ in dU ($\mathcal{U}$) continuous. It is a Fréchet topology ($\mathcal{D}$ is a complete countably Hilbert space). If $G \cdot T$ is the Poincaré group, it has been known for a long time[1,2] that the "non-zero-mass UIR" have nuclear spaces $\mathcal{D}$ while the spaces $\mathcal{D}$ of the "zero-mass UIR" (except the UIR induced from the trivial orbit) are nonnuclear.

Physically, it is important to know some criteria ensuring the nuclearity of $\mathcal{D}$. Indeed, when $\mathcal{D}$ is a nuclear space, we have much information on the spectrum of representative of elliptic elements of $\mathcal{U}$, but in principle we can use the nuclear spectral theorem and define a natural space containing all the generalized eigenvectors for each of the above observables, namely the topological dual $\mathcal{D}'$ of $\mathcal{D}$. In $\mathcal{D}'$, we can extend the representation dU of $\mathcal{U}$ and use "infinitesimal methods" rigorously; for instance, we can define the "continuous basis" of physical states which are not elements of the initial Hilbert space.

Another physical application of interest is to use the nonlinear representations of our symmetry group, these latter being connected with nonlinear differential equations. To build (and classify) these representations, we have to consider cohomology groups of $G \cdot T$ in some tensor products of linear representations of $G \circ T$; then, in the case of representations possessing nuclear $\mathcal{D}$ spaces, we can compute explicitly that cohomology.

Independently, it can be interesting to know if the nonzero mass condition for nuclearity in the case of the Poincaré group is a coincidence or on the contrary if this physical condition take a prominent part in the topological structure of $\mathcal{D}$. Here, we study some sequence of groups generalizing the case of the Poincaré group and establish that $\mathcal{D}$ is nonnuclear only in "zero-mass" representations.

From a mathematical point of view, a few conjectures have existed for a long time on this problem.[1,2] People thought to prove the nuclearity of $\mathcal{D}$ with hypothesis on the little group. We present here some examples which show that it is not the case. Therefore, we have searched for criteria on the inducing orbit.

If G is compact, $\mathcal{D}$ is always a nuclear space; thus one could think that this property remains almost everywhere true if G is only semisimple. We give here a counterexample, but we prove a similar result in many cases.

In the first section, we specify our notation and give many examples. We state our criteria in the second section. Then we study the Plancherel measure for $G \cdot T$, which shows that one of our assumptions on the representation holds almost everywhere; this forms the subject of the third section. Finally we apply in the last section our result in the case of the groups $SL(n, \mathbb{C}) \cdot \mathbb{R}^{n^2}$, $SO(p,q) \cdot \mathbb{R}^{p+q}$, generalizing the Poincaré group.

I. NOTATIONS AND EXAMPLES

Preliminaries

In this paper, we consider a Lie group $G \cdot T$ which is the semidirect product of a semisimple group G with a normal Abelian subgroup T. We assume that the action of G on T is regular, i.e., that it satisfies usual condition (see Ref. 3, p. 438, conditions O_1 and O_2, for instance). Then we obtain all the UIR of $G \cdot T$ by the induction theory of Mackey. We use the following notation:

We denote by $\mathfrak{g} + \mathfrak{t}$ the Lie algebra of $G \circ T$, $\mathcal{U}(\mathfrak{g}+\mathfrak{t})$ (or $\mathcal{U}$) its universal enveloping algebra. If u belongs to $\mathcal{U}$, we shall identify u with a right-invariant differential operator on $G \circ T$ and we shall denote by $u \circ f$ its action on a C^∞ function f on G.

If ξ_0 is an element of $\hat{T}$, the G orbit of ξ_0 will be Ω_{ξ_0} (or Ω). $\pi: G \to \Omega$ will be the canonical map and s a Borel section of Ω to G. The stability subgroup of ξ_0 will be $H_{\xi_0} \cdot T$ (or $H \cdot T$) and $\mathfrak{h} + \mathfrak{t}$ its Lie algebra. We choose a C^∞ function ρ_H on G such that

$$\rho_H(xh) = [\delta_H(h)/\delta_G(h)]\rho_H(x) \quad (x \in G,\ h \in H)$$

 0022-2488/78/1909-1881$1.00

(δ_H and δ_G are the modular functions of the groups G and H). Associated with this ρ_H is a quasi-invariant measure ν_H on G/H defined by

$$\int_G f(x)\ \rho_H(x)d_G(x)=\int_{G/H}d\nu_H(x)\int_H f(xh)d_H(h),$$

$f\in C_c(G)$ (Ref. 3, p. 475).

We can put ν_H on Ω, and then we shall denote ν_Ω this measure.

Let L be a UIR of H in a Hilbert space E; we denote by U^L the (unitary) representation induced by the representation $L\times\xi_0$ of $H\cdot T$ on E. The representation space of U^L is the space E^L of E-valued functions f on G such that

(i) for each $a\in E$, $x\to(f(x),a)$ is a Borel function on G,

(ii) for all $h\in H$, $x\in G$, $f(xh)=\rho_H(h)^{1/2}L(h^{-1})f(x)$,

(iii) the integral

$$\int_{G/H}\rho_H(x)^{-1}\,\|f(x)\|^2\,d_H(\dot{x})$$

is finite (see Ref. 3, p. 374). And we put

$$(U^L(y)f)(x)=f(y^{-1}x),\quad (U^L(t)f)(x)=(x\cdot\xi_0)(t)f(x),$$

$\forall y,x\in G$, $t\in T$, $f\in E^L$;

we shall denote by E_∞ the space of C^∞ vectors for L in E and by E^L_∞ the space of C^∞ vectors for U^L in E^L. We shall denote dU^L (dL) the representations of $\mathcal{U}(\mathfrak{g}+\mathfrak{t})$ ($\mathcal{U}(\mathfrak{h})$) on E^L_∞ (E_∞) obtained by differentiation of $U^L(L)$. Now, T being an Abelian Lie group, we can write

$$T=\mathbb{R}^N\times\mathbb{Z}^p\times\mathbf{T}^q\times F,$$

where $\mathbb{R}$ is the real line, $\mathbb{Z}$ the set of integers, $\mathbf{T}$ the one-dimensional torus, and F is a finite group. But E^L_∞ is exactly the space of C^∞ vectors for the restriction of U^L to the connected component $G_0\cdot T_0$ of the identity e in $G\cdot T$. We can assume that $T=\mathbb{R}^n\times\mathbf{T}^q$ or $\hat{T}=\mathbb{R}^n\times\mathbb{Z}^q$.

A G_0 orbit Ω in $\hat{T}$ being connected, Ω is a G_0 orbit in $\mathbb{R}^n\times\{\eta\}$, $\eta\in\mathbb{Z}^q$. Then if X belongs to the Lie algebra of $\mathbf{T}^q$,

$$dU^L(X)=i\eta(X)\cdot Id.$$

E^L_∞ is exactly the space of C^∞ vectors for the restriction of U^L to $G_0\cdot\mathbb{R}^n$.

In this paper, we assume that T is $\mathbb{R}^n$, we identify $\hat{T}$ with $\mathfrak{t}^*$ by

$$\xi(\exp X)=\exp[i\xi(X)]\quad\text{for all } X \text{ of } \mathfrak{t},\ \xi \text{ of } \hat{T}=\mathfrak{t}^*.$$

We shall also consider a maximal compact subgroup K of G and denote by $|\ |$ a K-invariant Euclidean norm on $\mathbb{R}^n$ (if G has an infinite center, take K in the quotient of G by the kernel of the representation of G in $\mathbb{R}^n$). Poulsen[4] characterized E^L_∞:

Proposition (*Poulsen*,[4] *Goodman*[5]): E^L_∞ is the space of C^∞ functions f of G in E such that $u\cdot f$ belongs to E^L for every u in $\mathcal{U}(\mathfrak{g})$ and $(1+|\pi(x)|^2)^m f(x)$ belongs to E^L for every positive integer m.

Proof: Following Poulsen, the first condition defines the set of C^∞ vectors for the restriction of U^L to G, the second one defines the space of C^∞ vectors for the restriction of U^L to T since, if (X_i) is an orthonormal basis of $\mathfrak{t}=\mathbb{R}^n$, $1+|\pi(x)|^2=dU^L(1+\sum x_i^2)$. We conclude with a result of Goodman.[5]

A few examples

(1) E^L_∞ is a nuclear space if G is a compact group.[2] If $G\cdot T$ is the Poincaré group, E^L_∞ is a nuclear space for representations corresponding to the trivial orbit $\Omega=\{0\}$, and to a nonzero mass orbit.[2,1]

We could generalize the positive-mass case in $SO(n,1)\cdot\mathbb{R}^{n+1}$; E^L_∞ is also nuclear.[6]

We shall reestablish all these results in the last section as immediate corollary of our criterion.

(2) If $G\circ T$ is the Poincaré group and Ω is a half light cone, E^L_∞ is nonnuclear.[2] The proof of this fact arises from the geometrical structure of Ω.

If $G\circ T$ is the group $SL(n,\mathbb{R})\cdot\mathbb{R}^n$, the action of $SL(n,\mathbb{R})$ on $\mathbb{R}^n$ being the canonical one, we see with the same proof that E^L_∞ is nuclear if and only if Ω is $\{0\}$. From the result of the third section, the set of UIR U^L such that E^L_∞ is nuclear is of Plancherel measure 0 for $SL(n,\mathbb{R})\cdot\mathbb{R}^n$.

(3) An example of nonnuclear space E^L_∞ with $H=\{e\}$ was given in Ref. 6. Here is another example of compact H with nonnuclear space E^L_∞: We consider the Lie group $SO(n,1)\cdot(\mathbb{R}^{n(n+1)/2}\times\mathbb{R}^{n+1})$, where $SO(n,1)$ acts on $\mathbb{R}^{n(n+1)/2}$ via the adjoint representation in $\mathfrak{so}(n,1)=\mathbb{R}^{n(n+1)/2}$ and on $\mathbb{R}^{n+1}$ canonically. We take a Cartan decomposition $\mathfrak{k}+\mathfrak{p}$ of $\mathfrak{so}(n,1)$, X a (nonnull) element of $\mathfrak{p}$ and ξ a vector in $\mathbb{R}^{n+1}$ such that

$$\exp tX\,\xi=\exp(t\lambda)\xi,\quad\lambda\neq0.$$

Then we consider the orbit Ω of (X,ξ) in $\mathbb{R}^{n(n+1)/2}\times\mathbb{R}^{n+1}$. The stability subgroup of (X,ξ) is evidently $SO(n-1)$. And we have, $\exp tX(X,\xi)=(X,\exp[t\lambda]\xi)$. So if (Y,η) belongs to Ω, $(Y,\mu\eta)$ $(\mu>0)$ belongs to Ω. Therefore, E^L_∞ is not a nuclear space in this case.[2]

Remark: In all these examples, the orbit Ω is closed when E^L_∞ is a nuclear space and conversely.

II. MAIN THEOREM

Here is our main criterion:

Theorem: Let us assume that

(1) the orbit Ω is closed,

(2) we can choose $\rho_H=1$ and $(1+|\pi(x)|^2)^{-l}\in L^1(G/H, d\nu_H)$ for a positive integer l,

(3) E_∞ is a nuclear space. Then E^L_∞ is a nuclear space.

Before stating the proof of the theorem, we establish two lemmas:

Lemma 1: Let α be the map defined by

$$\forall f\in E^L_\infty,\quad \alpha(f)=f(e)\quad (e \text{ is the unity of } G).$$

Then α is a E_∞ valued function and α is a continuous map from E^L_∞ to E_∞.

Proof: We have

$$f(h)=\rho_H(h)^{1/2}L(h^{-1})f(e),\quad\forall f\in E^L_\infty,\ \forall h\in H,$$

f and ρ_H being C^∞; $\alpha(f)$ belongs to E_∞. Moreover, if X is an element of $\mathfrak{h}$, we can write

$$dL(X)(f(e)) + X\cdot(\rho_H)^{1/2}\big|_e \cdot f(e)$$

$$= \lim_{t\to 0}\frac{1}{t}\{[\rho_H^{1/2}(\exp tX)L(\exp tX) - I]f(e)\}$$

$$= \lim_{t\to 0}\frac{1}{t}[f(\exp -tX) - f(e)]$$

$$= \alpha(dU^L(X)(f)).$$

This proves the continuity of α:

$$\forall\, u\in \mathcal{U}(\mathfrak{h}),\ \exists\, v,w\in\mathcal{U},\ M,M'>0\ \ni$$

$$\|dL(u)(\alpha(f))\|_E \le M\,\|\alpha(dU^L(v)f\|_E$$

$$\le M'\|dU^L(w)f\|_{E^L}\ \ \forall\ f\in E_\infty^L,$$

the last inequality being in Ref. 4.

Finally α is onto: Let v be a vector of E_∞, π the canonical map from G to G/H, s a C^∞ section from a compact neighborhood U of $\pi(e)$ in G/H to G such that $s(\pi(e)) = e$, V a compact neighborhood of e in G such that $\pi(V)\subset U$.

We define a C^∞ function p from VH to H by

$$x\in VH,\quad p(x) = s(\pi(x))^{-1}x.$$

Let φ be a complex valued C^∞ function with a support included in U and such that $\varphi(\pi(e)) = 1$. We put

$$\forall\, x\in VH,\quad \psi(x) = \rho_H^{1/2}(p(x))L(p(x)^{-1})\varphi(\pi(x))v.$$

ψ is a C^∞ function from G to E with compact support mod H, so ψ belongs to E_∞^L,[7] and we have

$$\alpha(\psi) = v.$$

Lemma 2: Let θ be a finite-dimensional representation of $\mathfrak{g}$ in a complex vector space V, $\|\ \|$ a norm in $\mathcal{L}(V)$. If Ω is closed, then there exists a positive C, an integer m and an element ξ_0 of Ω such that

$$\forall\ \xi\in\Omega\ \ \exists\ x\in G\qquad \xi = x\xi_0\ \text{ and }\ |\xi|^{2m}\ge C\,\|\theta x\|.$$

Proof: First, let us assume that Ω is noncompact. Let ξ_0 be a point of Ω, where $|\xi|$ is minimal. Let $\mathfrak{k}+\mathfrak{a}+\mathfrak{h}$ be a Iwasawa decomposition of $\mathfrak{g}$. Each element x of G can be written:

$$x = k'\exp Hk\ \text{ where } k,k'\in K,\ H\in\mathfrak{a}$$

(K is the analytic subgroup of G with Lie algebra $\mathfrak{k}$). Let (ξ_λ) be a basis of eigenvectors for all the H of $\mathfrak{a}$ in $\mathbb{R}^n = \mathfrak{t}^*$, let $\mathcal{X}$ be the compact set:

$$\mathcal{X} = \{\xi\in\Omega\ \ni\ (2n+1)\,|\xi_0|^2\le|\xi|^2\le(2n+2)\,|\xi_0|^2\}.$$

Let us consider the function ($|\ |$ being K invariant):

$$\varphi_H(t) = |k'\exp tHk\ \xi_0|^2 = \sum_\lambda |(k\xi_0,\xi_\lambda)|^2\exp[2t\lambda(H)].$$

If Ω is noncompact, there exists H such that $\varphi_H(t) > \varphi_H(0)$: then $\varphi_H(t)$ is a continuous unbounded real functions, and so $\mathcal{X}$ is not empty.

Now if $\xi = k'\exp Hk\,\xi_0$ belongs to $\mathcal{X}$, we have

$$|\xi|^2 = \sum_\lambda |(k\xi_0,\xi_\lambda)|^2\exp[2\lambda(H)\ge(2n+1)\,|\xi_0|^2$$

or

$$\sum_{\lambda/\lambda(H)\neq 0}|(k\xi_0,\xi_\lambda)|^2\exp[2\lambda(H)]$$

$$\ge(2n+1)\,|\xi_0|^2 - \sum_{\lambda/\lambda(H)=0}|(k\xi_0,\xi_\lambda)|^2\ge 2n\,|\xi_0|^2,$$

then, there exists λ_0 such that

$$|(k\xi_0,\xi_{\lambda_0})|^2\exp[2\lambda_0(H)]\ge 2\,|\xi_0|^2\ge 2\,|(k\xi_0,\xi_{\lambda_0})|^2.$$

Thus

$$\lambda_0(H)\ge\tfrac{1}{2}\log 2.$$

For each k of K, we put $L(k) = \{\lambda\ \ni\ (k\xi_0,\xi_\lambda)\neq 0\}$. For each L, let S_L be a subspace of $\mathfrak{a}$ such that

$$\mathfrak{a} = \bigcap_{\lambda\in L}\mathrm{Ker}\lambda\oplus S_L.$$

On S_L, we define a norm:

$$|H|_L = \sup_{\lambda\in L}|\lambda(H)|.$$

Let $|\ |$ a norm on $\mathfrak{a}$; thus, there exists a positive M such that

$$|H|\le M|H|_L\ \text{ for all } H \text{ in } S_L \text{ and for all } L.$$

If $\xi = k'\exp Hk\xi_0$, there exists a (unique) H_1 in S_L such that

$$\xi = k'\exp H_1 k\xi_0.$$

Then:

$$(2n+2)\,|\xi_0|^2\ge\sum_\lambda|(k\xi_0,\xi_\lambda)|^2\exp[2\lambda(H_1)]$$

Let μ the weight such that $|\mu(H_1)| = |H_1|_L$. If $\mu(H_1)$ is positive, we have

$$\exp(2|H_1|_L\le(2n+2)\frac{|\xi_0|^2}{\inf_{\lambda\in L}|(k\xi_0,\xi_\lambda)|^2} = (2n+2)M(k);$$

if $\mu(H_1)$ is negative, we write

$$0 = \varphi'_{H_1}(0) = \sum_{\lambda\in L}|(k\xi_0,\xi_\lambda)|^2\,\lambda(H_1).$$

or

$$|H_1|_L = -\mu(H_1)\le\sup_{\lambda(H_1)>0}\lambda(H_1)\ \frac{|\xi_0|^2}{\inf_{\lambda\in L}|(k\xi_0,\xi_\lambda)|^2}$$

$$\le\tfrac{1}{2}M(k)\log[(2n+2)M(k)].$$

Now, we can define a open covering of $K\cdot\xi_0$ by

$$K\cdot\xi_0\subset\bigcup_{k\in k}\{k_1\xi_0\ \ni\ L(k_1) = L(k)$$

$$\text{and }\inf_{\lambda\in L}|(k_1\xi_0,\xi_\lambda)|^2 > \tfrac{1}{2}\inf_{\lambda\in L}|(k\xi_0,\xi_\lambda)|^2\}.$$

$K\circ\xi_0$ being compact, we find in this covering a finite subcovering, and we find N such that

$$\forall\,\xi\in\mathcal{X},\quad \xi = k'\exp Hk\circ\xi_0\ \ (H\in S_{L(k)});$$

$$\text{then } |H|\le N.$$

Let m be a positive integer such that $m\log 2 > N$. Let ξ be in Ω such that

$$|\xi|^2\ge(2n+2)\,|\xi_0|^2,\quad \xi = k'\exp Hk\xi_0\ \ (H\in S_{L(k)});$$

then there exists $t\in\,]0,1]$ such that $k'\exp tHk\xi_0$ is in $\mathcal{X}$ so that $|H| < (1/t)N$ and

$$|\xi|^{2m}\ge(2\,|\xi_0|^2)^m\exp[2m\lambda_0(H)]$$

$$\ge(2\,|\xi_0|^2)^m\,[\exp(1/t)2m\lambda_0(tH)]$$

$$\ge(2\,|\xi_0|^2)^m\exp[(1/t)m\log 2]$$

or

$$|\xi|^{2m}\ge c\exp[|H|]\ \text{ and }\ \xi = k'\exp Hk\circ\xi_0.$$

Let $\mathcal{X}'$ be the set $\{\xi\in\Omega\ \ni\ |\xi|^2\le(2n+1)\,|\xi_0|^2\}$; $\mathcal{X}'$ is compact so that there exists a finite covering of $\mathcal{X}'$ by

compact balls B_i and, on each $B_i \cap \Omega$, a continuous section s_i from $B_i \cap \Omega$ to G. So there exists a compact set K in G such that

$$K \cdot \xi_0 = \chi'.$$

So we can extend the above relation to the entire orbit Ω. That is also true if Ω is compact. Now if θ is a finite-dimensional representation of G, there exists c', p such that

$$\|\theta(k' \exp Hk)\| \leq C' \quad \|\theta(\exp H)\| \leq c' \exp[p|H|],$$

and this proves the lemma.

Proof of the theorem

Following Poulsen,[4] there exists u in $\mathcal{U}(\mathfrak{g}+\mathfrak{t})$ and a positive c such that for all f in E_∞^L:

$$\|f(e)\|_E \leq c\,\|dU^L(u)f\|_{E^L}.$$

Then

$$\forall x \in G, \quad \|f(x)\|_E = \|U^L(x^{-1})f(e)\| \leq c\|dU^L(\mathrm{Ad}x(u))f\|_{E^L}.$$

Now, $\|f(x)\|_E$ is constant on left H cosets, so we can apply Lemma 2 and obtain

$$\|f(x)\|_E \leq c_2|\pi(x)|^{2m}\|dU^L(v)f\|_{E^L} \quad \forall x \in G, \forall f \in E_\infty^L.$$

Let $\Delta\mathfrak{t}$ be the Laplacian of $\mathfrak{t}$, we have

$$\|f(x)\|_E = |\pi(x)|^{-2m}\|(dU^L(\Delta_{\mathfrak{t}}^m)f)(x)\|_E \leq c_2\|dU^L(v\Delta_{\mathfrak{t}}^m)f\|_{E^L}.$$

Let us consider now the two functional spaces A, B:

A (resp. B) is the space of C^∞ E-valued functions on G such that

$$f(xh) = L(h^{-1})f(x) \quad \forall x \in G, \ h \in H,$$

and $(1+|\pi(x)|^2)^m\,\|(u \circ f)(x)\|$ belong to $L^1(G/H;\nu_H)$ (resp. $L^\infty(G/H,\nu_H)$) for every u in $\mathcal{U}$. We put on A (resp. B) a topology defined by the seminorms:

$$\|f\|_{A,p,u} = \int_{G/H}(1+|\pi(x)|^2)^p\|u \cdot f(x)\|\,d\nu_H(\dot{x}),$$

$$p \in \mathbb{N}, \ u \in \mathcal{U}$$

$$(\text{resp. } \|f\|_{B,p,u} = \sup_{\dot{x} \in G/H}(1+|\pi(x)|^2)^p\|u \cdot f(x)\|).$$

First we prove that the topological equalities $A = B = E_\infty^L$ hold. Indeed, if f belongs to E_∞^L, we have

$$\|f\|_{B,p,u} \leq C_2\|dU^L(v\Delta_{\mathfrak{t}}^m(1-\Delta_{\mathfrak{t}})^p)f\|_{E_L},$$

and if f belongs to B, we have

$$\forall u \in \mathcal{U}, \quad \|dU^L(U)f\|_{E_L} = [\int_{G/H}\|u \cdot f(x)\|^2\,d\nu_H(\dot{x})]^{1/2} \leq \|f\|_{B,l,u}[\int_{G/H}(1+|\pi(x)|^2)^{-2l}d\nu_H(\dot{x})]^{1/2},$$

which is finite with our hypothesis. The same proof holds for the equality $A = B$.

In order to prove the nuclearity of E_∞^L, we use the criterion of Ref. 8, Chap. 1, Sec. 3.3, Theorem 2.

If all simply absolutely convergent serie $\sum_{k=1}^\infty F_k$ in $(E_\infty^L)'$ is absolutely convergent, then E_∞^L is nuclear. But if $\sum F_k$ is simply absolutely convergent, there exists M, u, p such that

$$\sum_{k=1}^\infty |F_k(f)| \leq M\|f\|_{c,p,u} \ \forall \ f \in E_\infty^L$$

(Ref. 8, Chap. 1, Sec. 3.3, Lemma 1).

Choosing a Borel section s from G/H to G, and extending F_k to $L^1(G/H, E, (1+|\pi(x)|^2)^p d\nu_H(\dot{x}))$, we can use Ref. 9, Chap. 6, Sec. 2, no. 6, to define the functions G_k in $L^\infty(G/H, (1+|\pi(x)|^2)^p d\nu_H(\dot{x}))$ such that

$$F_k(\varphi) = \int_{G/H}(G_k(\dot{x}), \varphi(\dot{x}))(1+|\pi(x)|^2)^p d\nu_H(\dot{x}).$$

For all φ in $L^1(G/H, E, (1+|\pi(x)|^2)^p\,d\nu_H(\dot{x}))$. Now, let (e_i) be an orthonormal basis of E such that if Δh is the Laplacian of $\mathfrak{h}$ (the Lie algebra of H):

$$(1 - dL(\Delta h))^q(e_i) = \lambda_i e_i \quad \text{and} \quad \sum \lambda_i^{-1} < \infty.$$

Such a basis exists since E_∞ is nuclear. Let us put

$$G_{ki}(\dot{x}) = (G_k(\dot{x}), e_i).$$

Since for all φ in $L^1(G/H, E)$ $(e_i, \varphi)e_i$ is in $L^1(G/H, E)$, we have

$$\forall\, i, \ \forall\ \psi \in L^1(G/H, \mathbb{C}, (1+|\pi(x)|^2)^p d\nu_H(\dot{x}))$$

$$\sum_{k=1}^\infty \left| \int_{G/H} G_{ki}(\dot{x})\psi(\dot{x})(1+|\pi(x)|^2)^p d\nu_H(\dot{x})\right| \leq M\int_{G/H}|\psi(\dot{x})|(1+|\pi(x)|^2)^p d\nu_H(\dot{x}).$$

Now, we fix i, for all integer K, we consider the sets X in G/H where the functions $\mathrm{Re}G_{ki}(\dot{x})$, $\mathrm{Im}G_{ki}(\dot{x})$ have constant signs for all $k \leq K$; afterwards we take any positive function ψ in $L^1(G/H, \mathbb{C}, (1+|\pi(x)|^2)^p d\nu_H(\dot{x}))$, we write $\psi = \sum_X \psi_X$ with

$$\psi_X(\dot{x}) = \psi(\dot{x}) \ \text{ if } \dot{x} \in X, \ 0 \text{ elsewhere}$$

Then

$$\sum_{k=1}^K \int |G_{ki}(\dot{x})|\,\psi(\dot{x})(1+|\pi(x)|^2)^p\,d\nu_H(\dot{x}) = \sum_X \sum_{k=1}^K \left| \int_{G/H} G_{ki}\psi_X(1+|\pi|^2)^p d\nu_H\right| \leq M\int_{G/H}\psi(1+|\pi|^2)^p d\nu_H(\dot{x}).$$

So we extend this enaquility for all ψ in L^1, and we conclude that

$$\operatorname*{Ess\,sup}_{\dot{x} \in G/H} \sum_{k=1}^\infty |G_{ki}(\dot{x})| \leq 4M \ \forall\, i.$$

or

$$\operatorname*{Ess\,sup}_{\dot{x} \in G/H} \sum_{k=1}^\infty \|\sum_i \lambda_i^{-1} G_{ki}(\dot{x})e_k\| < \infty.$$

On the other hand, thanks to Lemma 1, we have

$$\|dL(1-\Delta_{\mathfrak{h}})^q(u \cdot f)(s(\dot{x}))\|_E \leq \|v \cdot U^L(x)u \cdot f\|_{E^L},$$

where v belongs to $\mathcal{U}$: so with the same argument as above, we can find w in $\mathcal{U}$ such that

$$\sup_{\dot{x}}(1+|\pi(x)|^2)^{p+l}\|dL(1-\Delta_{\mathfrak{h}})^q[(u \circ f)(s(\dot{x}))]\|_E \leq \|w \cdot f\|_{E^L} \quad \forall\, f \in E_\infty^L.$$

Now adapting the proof of Ref. 8, Chap. 1, Sec. 3, we conclude that

$$|F_k(f)| \leq \sup_{\dot{x}}(1+|\pi(x)|^2)^{p+l}\|dL(1-\Delta_{\mathfrak{h}})^q[(u \circ f)(x)]\|_E \times \int_{G/H}(1+|\pi(x)|^2)^{-l}\|\sum_i \lambda_i^{-1}G_{ki}(\dot{x})e_i\|d\nu_H \leq \|w \cdot f\|_{E^L}\int_{G/H}(1+|\pi(x)|^2)^{-l}\|\sum_i \lambda_i^{-1}G_{ki}(\dot{x})e_i\|d\nu_H.$$

So let us consider the completion of E_∞^L for the norm $\|w \cdot f\|_{E^L}$; in its dual we have

$$\sum_{k=1}^{\infty}\|F_k\| \leq \int_{G/H}(1+|\pi(x)|^2)^{-t}\sum_k\|\sum_i\lambda_i^{-1}G_{ki}(x)e_i\|d\nu_H<\infty.$$

Our theorem is now proved.

III. PLANCHEREL MEASURE AND HYPOTHESIS 2

Let us consider now the assumptions of the theorem. From Lemma 1, we see the following:

Proposition 1: The assumption (3) is a necessary condition.

Proof: Since E_∞ is a Frechet space, the application α of Lemma 1 is a strict morphism, so E_∞ is a nuclear space when E_∞^L is nuclear.

Now we study the validity of assumption (2), and we want to prove that (2) is almost everywhere true for the Plancherel measure on $(GT)\hat{}$, the dual of GT.

The class of the Plancherel measure μ [on $(GT)\hat{}$] is the class of measure associated with the left regular representation R of GT in $L^2(GT)$ (see Ref. 10, for instance). We suppose $\mu(GT)\hat{}$ finite. We denote by $\mathcal{Y}$ the set of orbits in $\mathbb{R}^n$ $(=t^*=\hat{T})$ with its Borel structure. With our hypothesis on GT, $\mathcal{Y}$ is isomorphic to a Borel subset of $\mathbb{R}^n$, so $\mathcal{Y}$ is standard.[11] Our notations are those of Ref. 10.

Lemma 3: There exists a Borel map from $(GT)\hat{}$ into $\mathrm{Rep}(T)$ such that

$$\Psi(U^L)\cong\aleph_0\int_\Omega\xi d\nu_\Omega(\xi)$$

if U^L is induced from the orbit Ω, ν_Ω being the above measure ν_H.

Proof: Let us consider the restriction map r from $\mathrm{Rep}(GT)$ into $\mathrm{Rep}(T)$. r is continuous for the topologies of these spaces.[10] Following Ref. 10, Prop. 4.6.2, there exists a Borel map s from $(GT)\hat{}$ into $\mathrm{Rep}(GT)$ such that $s(U^L)$ is in the class of U^L. Let us consider now the map φ from $\mathrm{Rep}(GT)$ into $\mathrm{Rep}(GT)$ defined by

$$\varphi(\pi)=\pi\otimes\mathrm{Id},$$

Id being the identity map of an infinite-dimensional Hilbert space. φ is a continuous map; then the lemma holds for the composition map:

$$\psi=s\circ\varphi\circ r$$

Lemma 4: We have a decomposition

$$\aleph_0\cdot R\big|_T\cong\int^{\oplus}_{\mathrm{Rep}(T)}\theta\,d\sigma(\theta),$$

where σ is the measure into which μ is mapped by Ψ.

Proof: Ψ being a Borel map, we can define σ and use Ref. 10, Chap. 6, Sec. 3, no. 1, Theorem 1. So we have a family of measure $\lambda_\theta(\theta\in\mathrm{Rep}T)$ such that

$$\mu=\int^{\oplus}\lambda_\theta d\sigma(\theta).$$

Now if $(\int^{\oplus}e_{i\theta}d\sigma(\theta))_i$ is a basis of the representation space of $\int^{\oplus}\theta d\sigma(\theta)$, for all t in T and i,j

$$\int_{\mathrm{Rep}'(t)}(\theta(t))e_{i\theta},e_{j\theta})d\sigma$$

$$=\int_{(GT)\hat{}}(\Psi(U^L)(t)e_{i\Psi(U^L)},e_{j\Psi(U^L)})d\mu(U^L)$$

holds.

But the second term defines a representation of T which is equivalent to

$$\int_{(GT)\hat{}}\aleph_0U^L\big|_T d\mu(U^L)\cong\aleph_0R\big|_T.$$

Proposition 2: Let $d\xi$ be the Lebesgue measure on $\mathbb{R}^n$, $\mathcal{X}$ be a subset of $\mathbb{R}^n$ which is a union of orbits and $\mathcal{X}'$ be the subset of $(GT)\hat{}$ defined by

$$\mathcal{X}'=\{U^L \text{ such that } U^L \text{ is induced from } \Omega \text{ and } \Omega\subset\mathcal{X}\}.$$

Then $\mathcal{X}'$ is a null set for μ if and only if $\mathcal{X}$ is a null set for $d\xi$.

Proof: We can decompose $\aleph_0R\big|_T$ by a Fourier transform on T (see Ref. 12) in

$$\aleph_0R\big|_T\cong\aleph_0\int_{\mathbb{R}^n}\xi d\xi.$$

But if

$$d\xi=\int_{\mathcal{Y}}\nu_\Omega d\lambda(\Omega)$$

is the decomposition of $d\xi$ with respect to the measurable equivalence relation "ξ_1 and ξ_2 belong to the same orbit,"[9] we have

$$\aleph_0R\big|_T=\aleph_0\int_{\mathcal{Y}}\int_\Omega\xi d\nu_\Omega(\xi)d\lambda(\Omega).$$

But the measure ν_Ω is λ-almost everywhere G invariant, so by unicity we can suppose that ν_Ω is the measure of Lemma 3. Then comparing this decomposition with the result of Lemma 4, we obtain two decompositions of the Lebesgue measure with respect to the above equivalence relation. $\mathrm{Rep}(T)$ and $\mathcal{Y}$ being standard Borel set, there exist two null sets N_1 and N_2 and a Borel isomorphism between $\mathcal{Y}-N_1$ and $\mathrm{Rep}(T)-N_2$ intertwining the classes of measure of λ and σ (Ref. 9, Chap. 6, Sec. 3, Theorem 4). The proposition follows from this isomorphism.

Corollary: Condition (2) of Theorem 1 is μ-almost everywhere true.

Of course, this assumption is almost everywhere true for the Lebesgue measure in $\mathbb{R}^n$.

Remark: Condition (1) is used in the proof of the theorem only via Lemma 2. Thus we can replace this condition by the little more general one:

(1′) There exists a nonnull invariant subspace V in $\mathbb{R}^n$ such that $P\Omega$ is closed, where P is the orthogonal projection on V.

IV. APPLICATIONS

Our theorem is useful if we characterize some classes of orbits by the values of continuous invariant functions on $\hat{T}$ since these orbits actually are closed. Here are two examples, generalizing the case of the Poincaré group where that happens:

(1) $G\cdot T=\mathrm{SO}_0(p,q)\cdot\mathbb{R}^{p+q}$: We suppose the action of $\mathrm{SO}_0(p,q)$ [the connected component of the identity in $\mathrm{SO}(p,q)$] in $\mathbb{R}^{p+q}$ is the canonical one. Then if $\xi=(\xi_1,\ldots,\xi_p,\xi_{p+1},\ldots\xi_{p+q})$ belongs to $\mathbb{R}^{p+q}$, the quantity $m^2=\sum_{i=1}^{p}\xi_i^2-\sum_{j=1}^{q}\xi_{p+j}^2$ is invariant. If m^2 is nonnull, the orbit of ξ is characterized by its value (and a sign if p or q is 1), the stabilizer of ξ is isomorphic to $\mathrm{SO}_0(p-1,q)$ or $\mathrm{SO}_0(p,q-1)$, following the sign of m^2; so we see immediately that the three conditions of our theorem hold and E_∞^L is a nuclear space in all this case. If m^2 is null, besides the trivial orbit $\{0\}$, for which E_∞^L of course, in nuclear, we have an orbit (or two orbits if

p or q is 1) which is a cone so that, thanks to the proof of Ref. 2, it is easy to see that E_∞^L is never nuclear.

(2) $G\Gamma = SL(n,C)\cdot \mathbb{R}^{n^2}$: This group was studied by Rideau and Angelopoulos.[13,1] We realize $\mathbb{R}^{n^2}$ as the group of $n\times n$ Hermitian matrices A, and we know that the orbits are characterized by the determinant of A and the numbers p,q,r of positive negative and null eigenvalues of A.[1]

Let us consider the set χ of the A with $\det A \neq 0$. χ is invariant dense and open in $\mathbb{R}^{n^2}$. The orbits in it are closed since eigenvalues are continuous functions of A. Condition (2) holds since χ is the finite union of the open subset $\chi_{p,q}$ of the A in χ with fixed number of positive eigenvalues.

Finally, the stabilizers are all semisimple: they are isomorphic to $SU(p,q)$. Our result can be applied, and all the spaces of C^∞ vectors are nuclear for this set of representations χ'. Moreover, thanks to Proposition 2, the complementary of χ' is a null set for the Plancherel measure.

ACKNOWLEDGMENTS

I should like to thank M. Flato for suggesting to me the subject as well as for many helpful discussions. I am also indebted to E. Angelopoulos and B. Nagel for remarks and exchange of views about this paper.

[1]E. Angelopoulos, "Reduction on the Lorentz subgroup of UIR's of the Poincaré group induced by a semisimple little group", J. Math. Phys. **15**, 155—65 (1974).
[2]B. Nagel, "Generalized eigenvectors in group representations," in *Studies in Mathematical Physics* (Reidel, Dordrecht, Boston, 1971).
[3]G. Warner, *Harmonic Analysis on Semi-Simple Lie Groups. I* (Springer-Verlag, Berlin, New York, 1972).
[4]N. Poulsen, "On C^∞-vectors and intertwining bilinear forms for representations of Lie groups," J. Funct. Anal. **9**, 87—120 (1972).
[5]R. Goodman: "Analytic and entire vectors for representations of Lie groups," Trans. Amer. Math. Soc. **143**, 55—76 (1969).
[6]D. Arnal, "Vecteurs différentiables dans les représentations unitaires des groupes de Lie inhomogènes," C.R. Acad. Sci. Paris Ser. A **285**, 203—06 (1977).
[7]F. Bruhat, "Sur les représentations induites des groupes de Lie," Bull. Soc. Math. France **84**, 97—205 (1956).
[8]I.M. Guelfand and N.Y. Vilenkin, *Les distributions, tome* 4 (Dunod, Paris, 1967).
[9]N. Bourbaki, *Livre VI, Intégration* (Hermann, Paris, 1959), Chap. 6.
[10]J. Dixmier, *Les C*-algèbres et leurs représentations* (Gauthier-Villars, Paris, 1964).
[11]G. Mackey, "Borel structure in groups and their duals," Trans. Amer. Math. Soc. **85**, 134—65 (1957).
[12]E. Angelopoulos, "Sur la réduction de la représentation régulière droite de $SL(nC)$ inhomogène" C.R. Acad. Sci. Paris Sér. A **263**, 400—02 (1966).
[13]G. Rideau, "On the reduction of the $r{:}r$: of Poincaré group," Preprint IHP (1966).

Slow motion approximation in predictive relativistic mechanics. I. Approximated dynamics up to order c^{-4} [a)]

J. Martin and J. L. Sanz

Departamento de Física Teórica, Universidad Autónoma de Madrid, Canto Blanco, Madrid-34 Spain
(Received 18 July 1977)

We obtain the most general two-body system relativistic invariant having a Newtonian limit up to c^{-4} order. In particular, it includes the scalar, gravitational, and other well-known dynamics up to c^{-2} order. It also includes to c^{-4} order the dynamics obtained recently by Lapiedra and Mas in predictive relativistic time symmetric electrodynamics.

1. INTRODUCTION

The predictive relativistic mechanics (PRM) of the isolated systems of N structureless point particles, in the manifestly predictive formalism,[1] is based on two fundamental principles:

(a) Principle of predictivity: the evolution of such a system is ruled as an ordinary second order system of differential equations over $\mathbb{R}^{3N}$.

(b) Principle of relativity: the set of trajectories of such a differential system is invariant as regards the Poincaré group.

These principles have been considered incompatible for some time. However, in 1961 Havas and Plebanski[2] demonstrated the compatibility of Principles (a) and (b). Later, Currie[3] and Hill[4] found the necessary conditions that the accelerations of the particles must satisfy in order for the system to be Poincaré invariant. Bel[5] proved that such conditions are also sufficient. These conditions that constitute a first order system of nonlinear partial differential equations which must be satisfied by the accelerations, are called in the literature "world line conditions." We shall call them Currie–Hill equations.

In this paper we study the isolated systems of two particles in an approximate framework (up to now no physically meaningful exact solution of the Currie–Hill equations has been obtained). We adopt the most commonly employed approximation in relativistic dynamics, i.e., the slow-motion approximation, which consists of an expansion of the accelerations in c^{-1}, where c is the velocity of light in vacuum. This perturbation technique permits the recurrent calculation of the accelerations by solving, at each order, a very simple first order system of linear partial differential equations.

We obtain the most general two-body system that satisfies the Currie–Hill equations, which is invariant by space inversion and possesses a Newtonian limit up to order c^{-4}. In particular, it includes to order c^{-2}, the following interactions deduced from the classical theory of fields: scalar (in its two versions), vectorial (in particular, electromagnetic), and gravitational interactions. It also includes to order c^{-4} the dynamics obtained by Lapiedra and Mas[6] for the two-body problem in predictive relativistic time symmetric electrodynamics.

2. MANIFESTLY PREDICTIVE FORMALISM

An isolated system of N structureless point particles is described by an ordinary second order system of differential equations[7] over $\mathbb{R}^{3N}$

$$\frac{dx_a^i}{dt}=v_a^i,\quad \frac{dv_a^i}{dt}=\mu_a^i(t,x_b^j,v_c^k), \tag{1}$$

where the functions μ_a^i must satisfy the following first-order system of nonlinear partial differential equations which state the invariance of the set of trajectories by the Poincaré group[3–5]:

$$\frac{\partial \mu_a^i}{\partial t}=0, \tag{2}$$

$$\epsilon_b \frac{\partial \mu_a^i}{\partial x_b^j}=0, \tag{3}$$

$$\eta_{jk}^l\left[x_b^k\frac{\partial \mu_a^i}{\partial x_b^l}+v_b^k\frac{\partial \mu_a^i}{\partial v_b^l}\right]=\eta_{jk}^i\mu_a^k, \tag{4}$$

$$c^{-2}v_b^k(x_b^j-x_a^j)\frac{\partial \mu_a^i}{\partial x_b^k}+[c^{-2}v_b^kv_b^j+c^{-2}(x_b^j-x_a^j)\mu_b^k-\epsilon_b\delta^{jk}]\frac{\partial \mu_a^i}{\partial v_b^k}$$
$$=c^{-2}(2v_a^j\mu_a^i+v_a^i\mu_a^j). \tag{5}$$

Equations (2) and (3) express that the functions μ_a^i, which we shall call accelerations, are invariant by the space–time translation group. Equation (4) states that, as regards the space rotation group, the functions μ_a^i are vectorial functions of vectorial variables, and Eq. (5) is associated with pure Lorentz transformations. We shall call them Currie–Hill equations.

Moreover, we assume that the dynamic system (1) is invariant under space inversion; this implies

$$\mu_a^i(t,-x_b^j,-v_c^k)=-\mu_a^i(t,x_b^j,v_c^k), \tag{6}$$

but we do not assume that the system is invariant under time inversion.

3. APPROXIMATED DYNAMICS

In order to solve Eqs. (2)–(6) in the case of two particles ($N=2$) we adopt the slow-motion approximation. We shall assume that the accelerations μ_a^i can be expanded into power series of c^{-1},

[a)]Research supported by the Instituto de Estudios Nucleares, Madrid, Spain.

 0022-2488/78/1909-1887$1.00

$$\mu_a^i = \sum_{n=0}^{\infty} c^{-n} \overset{n}{\mu}{}_a^i = \overset{0}{\mu}{}_a^i + c^{-1}\overset{1}{\mu}{}_a^i + c^{-2}\overset{2}{\mu}{}_a^i + c^{-3}\overset{3}{\mu}{}_a^i + c^{-4}\overset{4}{\mu}{}_a^i + O(c^{-5}), \tag{7}$$

where the functions $\overset{n}{\mu}{}_a^i$ are independent of c and verify the conditions

$$\overset{0}{\mu}{}_a^i = -m_a^{-1}\frac{\partial V(r)}{\partial x_i^a}, \tag{8a}$$

$$\overset{1}{\mu}{}_a^i = 0, \tag{8b}$$

where $m_a > 0$ is the mass of the particle a, $x_{aa'}^i \equiv x_a^i - x_{a'}^i$, $r \equiv +(x_{aa}^i x_{aa'i})^{1/2}$, and $V(r)$ is independent of c.

Assumption (8) is based on the following considerations: First, we want the dynamic system to constitute a Newtonian system of two particles in the limit $c \to \infty$; this justifies (8a). Second, it is well known that the dynamic systems corresponding to N interacting particles (scalar, electromagnetic, or gravitational) deduced from the classical theory of fields to order c^{-2} do not contain the c^{-1} term; this justifies (8b). We shall say that the dynamic systems that verify the property (8) have a Newtonian limit.

Trivially, (2) and (3) express that $\mu_a^i = \mu_a^i(x_{aa'}^j, v_b^k)$, that is to say, the functions μ_a^i are independent of t and dependent of the relative position. Equation (5) can be written, for $N=2$, as

$$-G_j\mu_a^i = c^{-2}\left[x_{aa'j}\left(v_{a'}^k\frac{\partial\mu_a^i}{\partial x^{a'k}} + \mu_{a'}^k\frac{\partial\mu_a^i}{\partial v^{a'k}}\right) - v_j^b R_b\mu_a^i + 2v_{aj}\mu_a^i + v_a^i\mu_{aj}\right], \tag{9}$$

where G_j and R_b are the linear operators

$$G_j \equiv \epsilon_b\frac{\partial}{\partial v_b^j}, \quad R_b \equiv v_b^k\frac{\partial}{\partial v^{kb}}. \tag{10}$$

By substituting the expansion (7) in (4), (6), and (9), and equating each term of the series, we obtain at each order n that the functions $\overset{n}{\mu}{}_a^i(x_{aa'}^j, v_b^k)$ must satisfy:

$$\eta_{jk}^l\left(x_b^k\frac{\partial\overset{n}{\mu}{}_a^i}{\partial x_b^l} + v_b^k\frac{\partial\overset{n}{\mu}{}_a^i}{\partial v_b^l}\right) = \eta_{jk}^i\overset{n}{\mu}{}_a^k, \tag{11a}$$

$$\overset{n}{\mu}{}_a^i(-x_{aa'}^j, -v_b^k) = -\overset{n}{\mu}{}_a^i(x_{aa'}^j, v_b^k), \tag{11b}$$

$$G_j\overset{0}{\mu}{}_a^i = G_j\overset{1}{\mu}{}_a^i = 0,$$

$$-G_j\overset{n}{\mu}{}_a^i = x_{aa'j}D_{a'}\overset{n-2}{\mu}{}_a^i - v_j^b R_b\overset{n-2}{\mu}{}_a^i + 2v_{aj}\overset{n-2}{\mu}{}_a^i + v_a^i\overset{n-2}{\mu}{}_{aj} + x_{aa'j}\sum_{\substack{r+s=n-2\\ r=1,\ldots,n-2\\ s=0,1,\ldots,n-3}}\overset{r}{\mu}{}_{a'}^k\frac{\partial\overset{s}{\mu}{}_a^i}{\partial v^{a'k}} \quad (n>1), \tag{11c}$$

where D_a is the linear operator

$$D_a \equiv v_a^k\frac{\partial}{\partial x^{ak}} + \overset{0}{\mu}{}_a^k\frac{\partial}{\partial v^{ak}}. \tag{12}$$

The compatibility of conditions (8) with Eqs. (11) at orders $n=0,1$ is easy to prove. It is also evident that Eqs. (11) constitute a recurrent method for the calculation of the different orders of the expansion (7), since the terms in the second members of (11c) are of lower order than those in the first member.

To each order, Eqs. (11a) and (11c) can be considered as a system of linear partial differential equations with the subsidiary condition (11b), whose general solution we shall write as

$$\overset{n}{\mu}{}_a^i = \overset{n}{\rho}{}_a^i + \overset{n}{\mu}{}_*^i{}_a, \tag{13}$$

where $\overset{n}{\mu}{}_*^i{}_a$ is the general solution of the homogenoues system $G_j\overset{n}{\mu}{}_*^i{}_a = 0$ satisfying (11a), (11b), and $\overset{n}{\rho}{}_a^i$ is a particular solution of the complete system (11).

It is easy to see that $\overset{n}{\mu}{}_*^i{}_a$ must be of the form

$$\overset{n}{\mu}{}_*^i{}_a = \overset{n}{f}{}_a^*x_{aa'}^i + \overset{n}{g}{}_a^*v_{aa'}^i, \tag{14}$$

$$(v_{aa'}^i \equiv v_a^i - v_{a'}^i, \quad s \equiv \tfrac{1}{2}x_{aa}^i v_{aa'i}, \quad v^2 \equiv v_{aa}^i v_{aa'i}),$$

where $\overset{n}{f}{}_a^*$, $\overset{n}{g}{}_a^*(r,s,v^2)$ are arbitrary functions of their arguments. In fact, $G_j\overset{n}{\mu}{}_*^i{}_a = 0$ expresses that the functions $\overset{n}{\mu}{}_*^i{}_a(x_{aa'}^j, v_{aa'}^k)$ are dependent of the relative velocity. Thus, in the generic case we can write

$$\overset{n}{\mu}{}_*^i{}_a = \overset{n}{f}{}_a^*(x_{aa'}^j, v_{aa'}^k)x_{aa'}^i + \overset{n}{g}{}_a^*(x_{aa'}^j, v_{aa'}^k)v_{aa'}^i + \overset{n}{h}{}_a^*(x_{aa'}^j, v_{aa'}^k)\eta_{lm}^i x_{aa'}^l v_{aa'}^m,$$

where $\overset{n}{f}{}_a^*$, $\overset{n}{g}{}_a^*$, and $\overset{n}{h}{}_a^*$ are arbitrary functions of the relative position and velocity.

On the other hand, (11a) trivially implies that the preceding functions are rotationally invariant, from which we deduce that they are arbitrary functions of three particular solutions, for example, r, s, and v^2. Finally, if we take into consideration (11b), it is obvious that $\overset{n}{h}{}_a^*$ must be identically null in order for the system to be invariant under space inversion. (14) is thus proven.

The problem is then reduced to finding a particular solutions of (11) at each order.

Defining

$$\overset{n}{F}{}_a^i \equiv m_a\overset{n}{\mu}{}_a^i \tag{15}$$

we shall have up to order c^{-1}

$$\overset{0}{F}{}_a^i \equiv \frac{W}{r^2}x_{aa'}^i, \quad W \equiv rV' \cdot \left(' \equiv \frac{d}{dr}\right), \quad \overset{1}{F}{}_a^i \equiv 0 \tag{16}$$

by hypothesis (but compatible with the Currie–Hill equations and invariant under space inversion).

Before continuing on to the next orders, we summarize for later usage the actions of some operators upon the variables (r, s, q, y_a, y):

$$q \equiv \tfrac{1}{2}\eta_a x_{aa'}^i(v_{ai} + v_{a'i}), \quad \eta_a\colon \eta_1 = +1, \ \eta_2 = -1, \qquad y_a \equiv v_a^i v_{ai} = v^2, \quad y \equiv v^2, \tag{17}$$

$$\mu \equiv \frac{m_1 m_2}{m_1 + m_2}, \quad m \equiv \frac{m_1 - m_2}{m_1 + m_2}, \quad D \equiv \epsilon^a D_a = v_b^k\frac{\partial}{\partial x_b^k} + \overset{0}{\mu}{}_b^k\frac{\partial}{\partial v_b^k} \tag{18}$$

(see Table I).

The proof of Table I is straightforward. If suffices to take into consideration the definitions of the operators and variables respectively. Hereinafter, we shall use it without explicit reference.

4. ORDER 2

Taking into account (15) and (16), Eq. (11c) can be written at this order

$$-G_j\overset{2}{F}{}_a^i = x_{aa'j}D_{a'}\overset{0}{F}{}_a^i + 2v_{aj}\overset{0}{F}{}_a^i + \overset{0}{F}{}_{aj}v_a^i. \tag{19}$$

It is easy to see that

TABLE I.

	r	s	q	y_a	$y_{a'}$	y
$\frac{\partial}{\partial x_i^a}$	$\frac{1}{r}x^i_{aa'}$	$\frac{1}{2}v^i_{aa'}$	$\frac{\eta_a}{2}(v^i_a+v^i_{a'})$	0	0	0
$\frac{\partial}{\partial v_i^a}$	0	$\frac{1}{2}x^i_{aa'}$	$\frac{\eta_a}{2}x^i_{aa'}$	$2v^i_a$	0	$2v^i_{aa'}$
D	$\frac{2}{r}s$	$\frac{1}{2}\left(y-\frac{W}{\mu}\right)$	$\frac{1}{2}\left(y_1-y_2+\frac{mW}{\mu}\right)$	$-\frac{2W}{m^a r^2}(s+\eta_a q)$	$-\frac{2W}{m^{a'} r^2}(s-\eta_a q)$	$-\frac{4W}{\mu r^2}s$
$D_{a'}$	$\frac{1}{r}(s-\eta_a q)$	$\frac{1}{4}\left(y-y_a+y_{a'}-\frac{2W}{m^{a'}}\right)$	$\frac{\eta_a}{2}\left(y-y_a-3y_{a'}+\frac{2W}{m^{a'}}\right)+\frac{2W}{m^{a'}}$	0	$-\frac{2W}{m^{a'} r^2}(s-\eta_a q)$	$-\frac{4W}{m^{a'} r^2}s$
R_a	0	$\frac{1}{2}(s+\eta_a q)$	$\frac{\eta_a}{2}(s+\eta_a q)$	$2y_a$	0	$y+y_a-y_{a'}$
G_j	0	0	$\eta_a x_{aa'j}$	$2v_{aj}$	$2v_{a'j}$	0

$$\overset{2}{F}{}^i_{\rho^a} \equiv \overset{2}{a}_{\rho^a} x^i_{aa'} + \overset{2}{b}_{\rho^a} v^i_{aa'}, \tag{20}$$

with

$$\overset{2}{a}_{\rho^a} \equiv \frac{W}{r^2}y_a + \frac{1}{2r}\left(\frac{W}{r^2}\right)^{\cdot}\eta_a q(2s-\eta_a q) + \frac{W}{4r^2}\left(\frac{2V-W}{m^a}+\frac{W}{m^{a'}}\right),$$

$$\overset{2}{b}_{\rho^a} \equiv \frac{W}{r^2}\eta_a q, \tag{21}$$

is a particular solution of (19). On the other hand, it is obvious that $\overset{2}{\mu}{}^i_{\rho^a} \equiv m_a^{-1}\overset{2}{F}{}^i_{\rho^a}$ satisfies (11a) since $\overset{2}{a}_{\rho^a}$ and $\overset{2}{b}_{\rho^a}$ are dependent on the variables (r,s,q,y_a) which are particular solutions of (11a). Analogously, it is obvious that $\overset{2}{\mu}{}^i_{\rho^a}$ satisfies (11b).

In short, taking into account (14), we conclude that the general solution of (11) at order two is

$$\overset{2}{F}{}^i_a \equiv m_a \overset{2}{\mu}{}^i_a = \left[\frac{W}{r^2}y_a + \frac{1}{2r}\left(\frac{W}{r^2}\right)^{\cdot}\eta_a q(2s-\eta_a q) + \frac{W}{4r^2}\left(\frac{2V-W}{m^a}+\frac{W}{m^{a'}}\right) + \overset{2}{a}{}^*_a(r,s,y)\right]x^i_{aa'} + \left[\frac{W}{r^2}\eta_a q + \overset{2}{b}{}^*_a(r,s,y)\right]v^i_{aa'}, \tag{22}$$

where $\overset{2}{a}{}^*_a$ and $\overset{2}{b}{}^*_a$ are arbitrary functions of their arguments.

5. ORDER 3

Taking into account (16), Eq. (11c) can be written at this order

$$G_j \overset{3}{\mu}{}^i_a = 0. \tag{23}$$

Therefore, the general solution of (11) is, according to (14),

$$\overset{3}{F}{}^i_a \equiv m_a \overset{3}{\mu}{}^i_a = \overset{3}{a}{}^*_a(r,s,y)x^i_{aa'} + \overset{3}{b}{}^*_a(r,s,y)v^i_{aa'}, \tag{24}$$

where $\overset{3}{a}{}^*_a$ and $\overset{3}{b}{}^*_a$ are arbitrary functions of their arguments.

6. ORDER 4

Taking into account (15) and (16), Eq. (11c) can be written at this order,

$$-G_j \overset{4}{F}{}^i_a = x_{aa'j}D_{a'}\overset{2}{F}{}^i_a - v^b_j R_b \overset{2}{F}{}^i_a + 2v_{aj}\overset{2}{F}{}^i_a + \overset{2}{F}_{aj}v^i_a. \tag{25}$$

First, we can prove that

$$\overset{4}{F}{}^i_{\rho^a} \equiv \overset{4}{a}_{\rho^a} x^i_{aa'} + \overset{4}{b}_{\rho^a} v^i_{aa'}, \tag{26a}$$

with

$$\overset{4}{a}_{\rho^a} \equiv \frac{1}{4r}\left(\frac{W}{r^2}\right)^{\cdot}(s^2+2q^2-3\eta_a sq)(y_a-y_{a'}) - y_a \times\left[\frac{W}{4r^2}\left(\frac{2V-W}{m^a}+\frac{W}{m^{a'}}\right)+\overset{2}{a}{}^*_a\right] + \tfrac{1}{4}\epsilon^b(s+\eta_b q)y_b\frac{\partial \overset{2}{a}{}^*_a}{\partial s} + \tfrac{1}{4}[(y_a-y_{a'})^2+2y(y_a+y_{a'})]\frac{\partial \overset{2}{a}{}^*_a}{\partial y} + \frac{W}{m^{a'}r^2}\eta_a q\left(\frac{W}{2r^2}\eta_a q + \overset{2}{b}{}^*_a\right) - \frac{\eta_a q}{8r}(2s-\eta_a q) \times\left[\frac{W}{r^2}\left(\frac{2V-W}{m^a}+\frac{W}{m^{a'}}\right)\right]^{\cdot} - \frac{1}{8r}\left[\frac{1}{r}\left(\frac{W}{r^2}\right)^{\cdot}\right]^{\cdot} q^2(4s^2+q^2-4\eta_a sq) - \frac{W}{2m^{a'}r}\left(\frac{W}{r^2}\right)^{\cdot} \times\eta_a q(s-\eta_a q) - \frac{1}{2r}\eta_a q(2s-\eta_a q)\frac{\partial \overset{2}{a}{}^*_a}{\partial r} + \frac{W}{2m^{a'}r^2}\eta_a q\left(r^2\frac{\partial \overset{2}{a}{}^*_a}{\partial s} + 8s\frac{\partial \overset{2}{a}{}^*_a}{\partial y}\right) - \frac{y}{4r}\eta_a q\left[s\left(\frac{W}{r^2}\right)^{\cdot} + r\frac{\partial \overset{2}{a}{}^*_a}{\partial s}\right], \tag{26b}$$

$$\overset{4}{b}_{\rho^a} \equiv \frac{W}{4r^2}(s-3\eta_a q)(y_a-y_{a'}) - \tfrac{1}{2}(2y_a-y_{a'})\overset{2}{b}{}^*_a + \tfrac{1}{4}\epsilon^b(s+\eta_b q)y_b\frac{\partial \overset{2}{b}{}^*_a}{\partial s} + \tfrac{1}{4}[(y_a-y_{a'})^2+2y(y_a+y_{a'})] \times\frac{\partial \overset{2}{b}{}^*_a}{\partial y} - \frac{1}{2r}\left(\frac{W}{r^2}\right)^{\cdot} q^2(2s-\eta_a q) - \frac{W}{4r^2}\eta_a q\left(y+\frac{2V-W}{m^a}+\frac{3W}{m^{a'}}\right) - \eta_a q\left[\overset{2}{a}{}^*_a + \frac{1}{4}\left(y-\frac{2W}{m^{a'}}\right)\frac{\partial \overset{2}{b}{}^*_a}{\partial s} - \frac{4W}{m^{a'}r^2}s\frac{\partial \overset{2}{b}{}^*_a}{\partial y}\right] - \frac{1}{2r}\eta_a q(2s-\eta_a q)\frac{\partial \overset{2}{b}{}^*_a}{\partial r}, \tag{26c}$$

is a particular solution of (25). In fact,

$$-G_j \overset{4}{F}{}^i_{\rho^a} = (-G_j \overset{4}{a}_{\rho^a})x^i_{aa'} + (-G_j \overset{4}{b}_{\rho^a})v^i_{aa'}. \tag{27}$$

On the other hand, we can write

$$\overset{2}{F}{}^i_a \equiv \overset{2}{a}_a x^i_{aa'} + \overset{2}{b}_a v^i_{aa'}, \quad \overset{2}{a}_a \equiv \overset{2}{a}_{\rho^a} + \overset{2}{a}{}^*_a, \quad \overset{2}{b}_a \equiv \overset{2}{b}_{\rho^a} + \overset{2}{b}{}^*_a,$$

$$\Rightarrow x_{aa'}D_{a'}\overset{2}{F}{}^i_a - v^b_j R_b \overset{2}{F}{}^i_a + 2v_{aj}\overset{2}{F}{}^i_a + \overset{2}{F}_{aj}v^i_a$$

$$= x^i_{aa'}\left[\left(D_{a'}\overset{2}{a}_a - \frac{W}{m^{a'}r^2}\overset{2}{b}_a\right)x_{aa'j} - (R_a\overset{2}{a}_a - 2\overset{2}{a}_a)v_{aj}\right.$$

$$\left. - (R_{a'}\overset{2}{a}_a)v_{a'j}\right] + v^i_{aa'}[(D_{a'}\overset{2}{b}_a + \overset{2}{a}_a)x_{aa'j}$$

$$- (R_a\overset{2}{b}_a - 2\overset{2}{b}_a)v_{aj} - (R_{a'}\overset{2}{b}_a + \overset{2}{b}_a)v_{a'j}]. \tag{28}$$

Therefore, we deduce from (27) and (28) that the following equations must be satisfied:

$$G_j\overset{4}{a}_{\rho a} = -\left(D_a\overset{2}{a}_a - \frac{W}{m^{a'}r^2}\overset{2}{b}_a\right)x_{aa'j} + (R_a\overset{2}{a}_a - 2\overset{2}{a}_a)v_{aj} + (R_{a'}\overset{3}{a}_a)v_{a'j},$$

$$G_j\overset{4}{b}_{\rho a} = -\ D_{a'}\overset{2}{b}_a + \overset{2}{a}_a)x_{aa'j} + (R_a\overset{2}{b}_a - 2\overset{2}{b}_a)v_{aj} + (R_{a'}\overset{2}{b}_a + \overset{2}{b}_a)v_{a'j}. \tag{29}$$

Moreover, we have

$$G_j\overset{4}{a}_{\rho a} = \eta_a\frac{\partial\overset{4}{a}_{\rho a}}{\partial q}x_{aa'j} + 2\frac{\partial\overset{4}{a}_{\rho a}}{\partial y^a}v_{aj} + 2\frac{\partial\overset{4}{a}_{\rho a}}{\partial y^a}v_{a'j}$$

and an analogous expression for $G_j\overset{4}{b}_{\rho a}$; therefore, according to (29),

$$-\eta_a\frac{\partial}{\partial q}\overset{4}{a}_{\rho a} = D_a\overset{2}{a}_a - \frac{W}{m^{a'}r^2}\overset{2}{b}_a,\quad 2\frac{\partial}{\partial y^a}\overset{4}{a}_{\rho a} = R_a\overset{2}{a}_a - 2\overset{2}{a}_a,$$

$$2\frac{\partial}{\partial y^{a'}}\overset{4}{a}_{\rho a} = R_{a'}\overset{2}{a}_a, \tag{30a}$$

$$-\eta_a\frac{\partial}{\partial q}\overset{4}{b}_{\rho a} = D_{a'}\overset{2}{b}_a + \overset{2}{a}_a,\quad 2\frac{\partial}{\partial y^a}\overset{4}{b}_{\rho a} = R_a\overset{2}{b}_a - 2\overset{2}{b}_a,$$

$$2\frac{\partial}{\partial y^{a'}}\overset{4}{b}_{\rho a} = R_{a'}\overset{2}{b}_a + \overset{2}{b}_a. \tag{30b}$$

After a tedious but straightforward calculation one easily verifies that Eqs. (30) are identically satisfied.

On the other hand, by a reasoning similar to that employed at order two, we conclude that $\overset{4}{\mu}{}^i_{\rho a} = m_a^{-1}\overset{4}{F}{}^i_{\rho a}$ satisfies (11a) and (11b).

Then, taking into account (14), the general solution of (11) at order four can be written

$$\overset{4}{F}{}^i_a \equiv m_a\overset{4}{\mu}{}^i_a = [\overset{4}{a}_{\rho a} + \overset{4}{a}{}^*_a(r,s,y)]x^i_{aa'} + [\overset{4}{b}_{\rho a} + \overset{4}{b}{}^*_a(r,s,y]v^i_{aa'}, \tag{31}$$

where $\overset{4}{a}_{\rho a}$ and $\overset{4}{b}_{\rho a}$ are the functions (26b) and (26c), respectively, and $\overset{4}{a}{}^*_a$ and $\overset{4}{b}{}^*_a$ are arbitrary functions of their arguments.

7. SUMMARY AND CONCLUSIONS

In brief: The most general dynamic system that can be expanded into power series of c^{-1}, satisfying the Currie–Hill equations, invariant under space inversion and possessing a Newtonian limit, is given up to order c^{-4} by

$$F^i_a \equiv m_a\mu^i_a = -\frac{W}{r^2}x^i_{aa'} + \frac{1}{c^2}\left\{\left[\frac{W}{r^2}v_a^2 + \frac{1}{2r}\left(\frac{W}{r^2}\right)^{\cdot}\eta_a q(2s - \eta_a q)\right.\right.$$

$$\left. + \frac{W}{4r^2}\left(\frac{2V-W}{m^a} + \frac{W}{m^{a'}}\right) + \overset{2}{a}{}^*_a\right]x^i_{aa'}$$

$$\left. + \left[\frac{W}{r^2}\eta_a q + \overset{2}{b}{}^*_a\right]v^i_{aa'}\right\} + \frac{1}{c^3}\left\{\overset{3}{a}{}^*_a x^i_{aa'} + \overset{3}{b}{}^*_a v^i_{aa'}\right\}$$

$$+ \frac{1}{c^4}\{(\overset{4}{a}_{\rho a} + \overset{4}{a}{}^*_a)x^i_{aa'} + (\overset{4}{b}_{\rho a} + \overset{4}{b}{}^*_a)v^i_{aa'}\} + O\left(\frac{1}{c^5}\right), \tag{32}$$

where $\overset{n}{a}{}^*_a$, $\overset{n}{b}{}^*_a$ (r,s,v^2) are arbitrary functions of their arguments and $\overset{4}{a}_{\rho a}$, $\overset{4}{b}_{\rho a}$ are given by (26b) and (26c).

In particular, included in (32) up to order c^{-2}, are the following interactions whose approximate dynamics are deduced from the classical theory of fields: scalar (in its two versions), vectorial (in particular, electromagnetic), and gravitational interactions.

The fact that a version of the scalar interaction (theory $\gamma = 1$),[8] in the same manner as the vectorial and gravitational interactions, satisfies the Curie–Hill equations to order c^{-2}, is well known in the literature.[9–13] We incorporate the new relevant fact that another scalar interaction version (theory $\gamma = 0$) also satisfies the Currie–Hill equations to order c^{-2}.

We remark that also included in our dynamics (32) are other approximated dynamics obtained from the following approximated Lagrangians to order c^{-2}: (i) the Bazañski Lagrangian[14] corresponding to the interaction of two masses with charge in general relativity, (ii) the Estabrook Lagrangian[15] corresponding to the interaction of two point masses in the Brans–Dicke theory, (iii) the Bopp[16] and Bagge[17] Lagrangians which are the generalizations of scalar and vectorial interactions, respectively, with $V(r)$ arbitrary, (iv) the Mas[9] and Kennedy[10] Lagrangians, (v) the Woodcock–Havas Lagrangian[11] obtained from Lorentz-invariant variational principles, and (vi) the Stachel–Havas Lagrangian[12] obtained through a canonical formalism.

Finally, Lapiedra and Mas have recently determined[6] the accelerations for the two-body problem in predictive relativistic electrodynamics up to order c^{-5}. They assume the Currie–Hill equations, certain symmetries (including invariance under time inversion), and the relativistic extension of Coulomb's law, so obtaining a unique dynamic that up to order c^{-4} has the form (32) with

$$V \equiv \frac{g}{r},\quad g \equiv e_1e_2,$$

$$\overset{2}{a}{}^*_a \equiv \frac{g^2}{4r^4}\left(\frac{3}{m^a} - \frac{5}{m^{a'}}\right) + \frac{g}{2r^3}\left(v^2 - \frac{3s^2}{r^2}\right),\quad \overset{2}{b}{}^*_a \equiv -\frac{g}{r^3}s,$$

$$\overset{4}{a}{}^*_a \equiv \frac{g}{r^5}s^2\left(\frac{15s^2}{8r^2} - \frac{3v^2}{4} - \frac{9g}{m^{a'}r}\right) + \frac{g^2}{4m^{a'}r^4}\left[v^2 + \frac{g}{r}\left(\frac{2}{m^a} + \frac{1}{m^{a'}}\right)\right]$$

$$\overset{4}{b}{}^*_a \equiv \frac{3g}{2r^4}s\left(\frac{s^2}{r} + \frac{3g}{m^{a'}}\right).$$

[1] Such a denomination of the formalism has been recently adopted by L. Bel.
[2] P. Havas and J.F. Plebanski, Bull. Am. Phys. Soc. 5, 433 (1961).
[3] D.G. Currie, Phys. Rev. 142, 817 (1966).
[4] R.N. Hill, J. Math. Phys. 8, 201 (1967).
[5] L. Bel, Ann. Inst. H. Poincaré 12, 307 (1970).
[6] R. Lapiedra and L. Mas, Phys. Rev. D 13, 2805 (1976).
[7] $i,j,k,\ldots = 1,2,3$; $a,b,c,a',\ldots = 1,\ldots,N$ where N is the number of particles, a' is always different from a; all indices follow the summation convention; η_{ijk} is the Levi–Civita pseudotensor normalized with $\eta_{123} = 1$; c is the velocity of light in vacuum; $\epsilon_a = \epsilon^a = +1$.

[8]The two scalar field theories mentioned have been distinguished in the literature by a parameter γ which can take the values 0 or 1 in the motion equations. See L. Bel and B. Lêaute, Ann. Inst. H. Poincarê **10**, 317 (1969); L. Bel. and J. Martín, Phys. Rev. D **9**, 2760 (1974).
[9]L. Mas, C.R. Acad. Sci. Paris A **271**, 206 (1970).
[10]F.J. Kennedy, Am. J. Phys. **40**, 63 (1972).
[11]H. W. Woodcock and P. Havas, Phys. Rev. D **6**, 3422 (1972).
[12]J. Stachel and P. Havas, Phys. Rev. D **13**, 1598 (1976).
[13]D. Hirondel, J. Math. Phys. **15**, 1471 (1974).
[14]S. Bazañski, Acta Phys. Polon. **16**, 423 (1957).
[15]F.B. Estabrook, Astrophys. J. **158**, 81 (1969).
[16]F. Bopp, Ann. Phys. **38**, 345 (1940); **42**, 573 (1943); Phys. Z. (1945).
[17]E. Bagge, Z. Naturforsch **1**, 361 (1946).

Stochastic quantization of wave fields and its application to dissipatively interacting fields

Kunio Yasue

Department of Physics, Nagoya University, Nagoya 464, Japan
(Received 12 September 1977)

The stochastic quantization procedure, proposed by Nelson, is further developed to include the problem of field quantization. To maintain the mathematical rigor in treating infinitely many degrees of freedom, basic notions of the nonstandard analysis are adopted. As an application of the present method, a quantum mechanical description of dissipatively interacting fields, such as the laser electric field in the lossy cavity, is investigated.

I. INTRODUCTION

Several years ago, Nelson[1] proposed a new framework of nonrelativistic quantum mechanics based on stochastic calculus which is now frequently called the stochastic quantization procedure. Although he succeeded in deriving the Schrödinger equation with mathematical rigor, many other parts of the subject matter of quantum theory such as spin, relativity, and field quantization were left uninvestigated.

Later on, those with spin and relativity were overcome satisfactorily: Dankel[2] has incorporated spin freedom by applying the stochastic quantization procedure to dynamical systems on finite dimensional C^∞ manifolds, and Caubet,[3] Pena-Auerbach,[4] and Lehr and Park[5] have generalized it to include relativistic quantum mechanics. There have been, on the other hand, no mathematically consistent generalizations of the stochastic quantization procedure to include the quantum field theory. This is because the probability theory based on the standard analysis can not provide profitable techniques in investigating the stochastic calculus on the infinite dimensional space of generalized functions.

However, in regarding the considerable result of the formal stochastic quantization of a real free field given by Guerra and Ruggiero,[6,7] the existence of a rigorous stochastic formulation of the quantum field theory has been expected.

In the present paper, we investigate such a stochastic formulation of the quantum field theory. We make use of basic notions of the nonstandard analysis for the purpose of treating infinitely many degrees of freedom consistently.

The most significant point of the present stochastic formulation is that it relies on neither Hamiltonian nor Lagrangian but on the field equation in the generalized sense. So it seems applicable to the wider class of wave fields, that is, not only to nondissipative fields but also to dissipative ones. To clarify the situation, we shall demonstrate the stochastic quantization of dissipatively interacting fields.

In Sec. II, we present the stochastic calculus on the infinite dimensional space of generalized functions. Section III is devoted to performing the stochastic quantization of nondissipative fields. Section IV treats the stochastic quantization of dissipative fields interacting with chaotic thermal environments. As an application of the present method, we investigate a quantum mechanical description of the laser electric field dissipatively interacting with the cavity in Sec. V.

II. STOCHASTIC CALCULUS ON KAWABATA SPACE $\mathcal{E}_{\{e_n\}}(R^d)$

This section is devoted to sketching the structure of Kawabata space and to develop stochastic calculus on it.

We fix a free ultrafilter F on N (N denotes the totality of natural numbers). Let ${}^*E = \Pi_{n\in N} R^n/F$ be the ultra-Euclidean space, that is, a quotient space of $\Pi_{n\in N}R^n$ with respect to the following equivalence relation

$$a \sim b \Leftrightarrow \{n\in N \mid a^{(n)} = b^{(n)}\} \in F, \tag{1}$$

where $a = \{a^{(n)}\}_{n\in N} = \{(a_1^{(n)}, \cdots, a_n^{(n)})\}_{n\in N}$ and $b = \{b^{(n)}\}_{n\in N} = \{(b_1^{(n)}, \cdots, b_n^{(n)})\}_{n\in N}$ belong to $\Pi_{n\in N}R^n$. The equivalence class which contains $a = \{a^{(n)}\}_{n\in N}$ should be denoted by $[a^{(n)}]_{n\in}{}^*E$.

It can be seen easily that the ultra-Euclidean space ${}^*E$ possesses the structure of linear space over the ultrareal field ${}^*R = R^N/F$, and the structure of Euclidean space with respect to the inner product

$$\langle [a^{(n)}]_n, [b^{(n)}]_n \rangle = [\langle a^{(n)}, b^{(n)} \rangle]_n = [\sum_{p\le n} a_p^{(n)} b_p^{(n)}]_n \in {}^*R. \tag{2}$$

According to Kawabata and Kurata[8] we can generalize the concept of real-valued functions on the d-dimensional Euclidean space R^d as follows.

Let $\mathcal{S}(R^d)$ be the Schwartz space over R^d and $\{e_n\}_{n\in N} \subset \mathcal{S}(R^d)$ be a complete normalized orthogonal system of $L_2(R^d)$ [$L_2(R^d)$ denotes the Hilbert space of real square integrable functions on R^d]. Then we can associate a unique ${}^*R$-valued function ϕ on R^d with each element $[a^{(n)}]_{n\in}{}^*E$;

$$\phi(\cdot) = [\phi_n(\cdot)]_n = [\sum_{p\le n} a_p^{(n)} e_p(\cdot)]_n. \tag{3}$$

Totality of such ${}^*R$-valued functions is denoted by ${}^*\mathcal{E}_{\{e_n\}}(R^d)$ and called the Kawabata space. It is homeomorphic to ${}^*E$ if we define the inner product

 0022-2488/78/1909-1892$1.00

$$\langle \phi, \psi \rangle = \langle [\sum_{p \le n} a_p^{(n)} e_p]_n, [\sum_{q \le n} b_q^{(n)} e_q]_n \rangle$$

$$= [\sum_{p,q \le n} a_p^{(n)} b_q^{(n)} \langle e_p, e_q \rangle]_n$$

$$= [\sum_{p \le n} a_p^{(n)} b_p^{(n)}]_n$$

$$= \langle [a^{(n)}]_n, [b^{(n)}]_n \rangle, \tag{4}$$

for any $\phi,\ \psi \in {}^*\mathcal{E}_{\{e_n\}}(R^d)$.

Note that any function

$$\phi = [\sum_{p \le n} a_p^{(n)} e_p]_n \in {}^*\mathcal{E}_{\{e_n\}}(R^d)$$

is locally differentiable and integrable in the sense that

$$\operatorname{grad} \phi(x) = [\sum_{p \le n} a_p^{(n)} \operatorname{grad} e_p(x)]_n \in {}^*R \tag{5}$$

and

$$\int_{R^d} \phi(x) d^d x = [\sum_{p \le n} a_p^{(n)} \int_{R^d} e_p(x) d^d x]_n \in {}^*R \tag{6}$$

hold. Local product of such ${}^*R$-valued functions ϕ and ψ can be defined to be

$$\phi(x)\psi(x) = [\sum_{q,p \le n} a_p^{(n)} b_q^{(n)} e_p(x) e_q(x)]_n \in {}^*R. \tag{7}$$

Now we shall proceed to develop the stochastic calculus on ${}^*\mathcal{E}_{\{e_n\}}(R^d)$.

Let $X^{(n)}(t) = (X_1^{(n)}(t), \cdots, X_n^{(n)}(t))$ be an n-dimensional diffusion process described by the stochastic differential equation

$$dX^{(n)}(t) = a^{(n)}(X^{(n)}(t), t)dt + B^{(n)}(dt), \tag{8}$$

where $B^{(n)}(t)$ denotes an n-dimensional Wiener process with diffusion constant β and $a^{(n)}(\cdot, t)$ a given vector field on R^n. Namely the probability distribution of $X^{(n)}(t)$

$$\operatorname{Prob}\{X^{(n)}(t) \in d^n x^{(n)}\} = \rho(x^{(n)}, t) d^n x^{(n)}, \tag{9}$$

satisfies the following Fokker—Planck equation.

$$\frac{\partial}{\partial t}\rho = -\sum_{p \le n} \frac{\partial}{\partial x_p^{(n)}}(a_p^{(n)}\rho) + \beta \sum_{p \le n} \frac{\partial^2}{\partial x_p^{(n)2}} \rho. \tag{10}$$

Starting with a family of such diffusion processes $\{X^{(n)}(t)\}_{n=1}^{\infty}$, we can construct a diffusion process $\Psi(t)$ on ${}^*\mathcal{E}_{\{e_n\}}(R^d)$

$$\Psi(t) = [\sum_{p \le n} X_p^{(n)}(t) e_p]_n \in {}^*\mathcal{E}_{\{e_n\}}(R^d). \tag{11}$$

The nonstandard probability distribution of $\Psi(t)$ is given by

$$^*\operatorname{Prob}\{\Psi(t) \in \delta\psi\} \equiv [\operatorname{Prob}\{X^{(n)}(t) \in d^n x^{(n)}\}]_n$$

$$= [\rho(x^{(n)}, t) d^n x^{(n)}]_n$$

$$= [\rho(x^{(n)}, t)]_n [d^n x^{(n)}]_n$$

$$= P(\psi, t)\delta\psi, \tag{12}$$

where $P(\psi, t) = [\rho(x^{(n)}, t)]_n$, $\psi = [\sum_{p \le n} x_p^{(n)} e_p]_n$, and $\delta\psi = [d^n x^{(n)}]_n$.

The nonstandard probability density $P(\psi, t)$ defined above satisfies the following second-order functional differential equation:

$$\frac{\partial}{\partial t}[\rho(x^{(n)}, t)]_n$$

$$= -[\sum_{p \le n} \frac{\partial}{\partial x_p^{(n)}}\{a_p^{(n)}(x^{(n)}, t)\rho(x^{(n)}, t)\}]_n$$

$$+ \beta[\sum_{p \le n} \frac{\partial^2}{\partial x_p^{(n)2}} \rho(x^{(n)}, t)]_n. \tag{13}$$

Equation (13) can be rewritten in a unified form

$$\frac{\partial}{\partial t}P(\psi, t) = -\int d^d x \frac{\delta}{\delta\psi(x)}\{(A_t\psi(x))P(\psi, t)\}$$

$$+ \beta \int d^d x \frac{\delta^2}{\delta\psi(x)^2} P(\psi, t), \tag{14}$$

where A_t denotes a transformation on ${}^*\mathcal{E}_{\{e_n\}}(R^d)$;

$$A_t;\psi = [\sum_{p \le n} x_p^{(n)} e_p]_n \mapsto A_t\psi$$

$$= [\sum_{p \le n} a_p^{(n)}(x^{(n)}, t) e_p]_n. \tag{15}$$

The functional derivative of a functional on ${}^*\mathcal{E}_{\{e_n\}}(R^d)$ of the type

$$F(\psi) = [F(x^{(n)})]_n \in {}^*R \tag{16}$$

is defined, following Kawabata and Kurata,[8] to be

$$\frac{\delta F}{\delta\psi(x)} = \left[\sum_{p \le n} \frac{\partial F(x^{(n)})}{\partial x_p^{(n)}} e_p(x)\right]_n \in {}^*\mathcal{E}_{\{e_n\}}(R^d). \tag{17}$$

Equivalence of Eq. (13) and Eq. (14) can be seen directly as follows:

$$\int d^d x \frac{\delta}{\delta\psi(x)}\{(A_t\psi(x))P(\psi, t)\} = \int d^d x \frac{\delta}{\delta\psi(x)}\left\{\left[\sum_{p \le n} a_p^{(n)}(x^{(n)}, t) e_p(x)\right]_n \left[\rho(x^{(n)}, t)\right]_n\right\}$$

$$= \int d^d x \frac{\delta}{\delta\psi(x)}[\sum_{p \le n} a_p^{(n)}(x^{(n)}, t)\rho(x^{(n)}, t) e_p(x)]_n$$

$$= \int d^d x \left[\sum_{p,q \le n} \frac{\partial}{\partial x_q^{(n)}}\{a_p^{(n)}(x^{(n)}, t)\rho(x^{(n)}, t)\} e_p(x) e_q(x)\right]_n$$

$$= \left[\sum_{p,q \le n} \frac{\partial}{\partial x_q^{(n)}}\{a_p^{(n)}(x^{(n)}, t)\rho(x^{(n)}, t)\}\langle e_p, e_q \rangle\right]_n = \left[\sum_{p \le n} \frac{\partial}{\partial x_p^{(n)}}\{a_p^{(n)}(x^{(n)}, t)\rho(x^{(n)}, t)\right]_n, \tag{18}$$

$$\int d^d x \frac{\delta^2}{\delta\psi(x)^2} P(\psi,t) = \int d^d x \left[\sum_{p,q\le n} \frac{\partial^2}{\partial x_p^{(n)} \partial x_q^{(n)}} \rho(x^{(n)},t) e_p(x) e_q(x) \right]_n$$

$$= \left[\sum_{p,q\le n} \frac{\partial^2}{\partial x_p^{(n)} \partial x_q^{(n)}} \rho(x^{(n)},t) \langle e_p, e_q \rangle \right]_n = \left[\sum_{p\le n} \frac{\partial^2}{\partial x_p^{(n)2}} \rho(x^{(n)},t) \right]_n. \tag{19}$$

The expectation value of a random variable of the type $F(\Psi(t),t) = [F(X^{(n)}(t),t)]_n$ is given by the following functional integral form:

$$E\{F(\Psi(t),t)\} = [E\{F(X^{(n)}(t),t)\}]_n = [\int F(x^{(n)},t)\rho(x^{(n)},t) d^n x^{(n)}]_n = \int_{*\mathcal{E}_{\{e_n\}}(R^d)} F(\psi,t) P(\psi,t) \delta\psi, \tag{20}$$

where the functional integral of a functional of the type (16) is defined, following Kawabata and Kurata,[8] to be

$$\int_{*\mathcal{E}_{\{e_n\}}(R^d)} F(\psi)\delta\psi = [\int F(x^{(n)}) d^n x^{(n)}]_n \in {}^*R. \tag{21}$$

For the later use, it is convenient to introduce the mean derivatives of a random variable of the above-mentioned type $F(\Psi(t),t)$;

$$DF(\Psi(t),t) = \lim_{h\downarrow 0} \frac{1}{h} E\{F(\Psi(t+h),t+h) - F(\Psi(t),t) \mid \Psi(t)\} = \lim_{h\downarrow 0} \frac{1}{h} [E\{F(X^{(n)}(t+h,t+h) - F(X^{(n)}(t),t) \mid X^{(n)}(t)\}]_n$$

$$= \left[\left(\frac{\partial}{\partial t} + \sum_{p\le n} a_p^{(n)} \frac{\partial}{\partial x_p^{(n)}} + \beta \sum_{p\le n} \frac{\partial^2}{\partial x_p^{(n)2}} \right) F(X^{(n)}(t),t) \right]_n = \left(\frac{\partial}{\partial t} + \int d^d x A_t \psi(x) \frac{\delta}{\delta\psi(x)} + \beta \int d^d x \frac{\delta^2}{\delta\psi(x)^2} \right) F(\Psi(t),t), \tag{22}$$

$$D_* F(\Psi(t),t) = \lim_{h\downarrow 0} \frac{1}{h} E\{F(\Psi(t),t) - F(\Psi(t-h),t-h) \mid \Psi(t)\} = \lim_{h\downarrow 0} \frac{1}{h} [E\{F(X^{(n)}(t),t) - F(X^{(n)}(t-h),t-h) \mid X^{(n)}(t)\}]_n$$

$$= \left[\left(\frac{\partial}{\partial t} + \sum_{p\le n} \hat{a}_p^{(n)} \frac{\partial}{\partial x_p^{(n)}} - \beta \sum_{p\le n} \frac{\partial^2}{\partial x_p^{(n)2}} \right) F(X^{(n)}(t),t) \right]_n = \left(\frac{\partial}{\partial t} + \int d^d x \hat{A}_t \psi(x) \frac{\delta}{\delta\psi(x)} - \beta \int d^d x \frac{\delta^2}{\delta\psi(x)^2} \right) F(\Psi(t),t), \tag{23}$$

where $E\{. \mid .\}$ denotes the conditional expectation and $\hat{A}_t$ a transformation on ${}^*\mathcal{E}_{\{e_n\}}(R^d)$ such that

$$\hat{A}_t;\psi = [\sum_{p\le n} x_p^{(n)} e_p]_n \mapsto \hat{A}_t\psi = [\sum_{p\le n} \hat{a}_p^{(n)}(x^{(n)},t) e_p]_n, \tag{24}$$

with

$$\hat{a}_p^{(n)} = a_p^{(n)} - 2\beta \frac{\partial}{\partial x_p^{(n)}} \log\rho. \tag{25}$$

Namely we have the following relation between A_t and $\hat{A}_t$:

$$\hat{A}_t\psi(x) = A_t\psi(x) - 2\beta \frac{\delta}{\delta\psi(x)} \log P. \tag{26}$$

These functional differential operators D and D_* are called as mean forward and mean backward derivatives generated by $\Psi(t)$, respectively.

III. STOCHASTIC QUANTIZATION OF NONDISSIPATIVE FIELDS

We present the basic ideas of the stochastic quantization of nondissipative fields.

Let us consider a real classical field $\psi(\mathbf{x},t)$ on R^3 with the following field equation. (The light velocity c is assumed to be unity in this section.)

$$\ddot{\psi}(\mathbf{x},t) = \nabla^2\psi(\mathbf{x},t) - \chi^2\psi(\mathbf{x},t), \tag{27}$$

where χ is a parameter with dimension L^{-1} and ∇ denotes the three-dimensional gradient. Equation (27) is equivalent to the following one for $\psi(\cdot,t) \in {}^*\mathcal{E}_{\{e_n\}}(R^3)$:

$$\ddot{\psi}(\mathbf{x},t) = -\frac{\delta}{\delta\psi(\mathbf{x},t)} \frac{1}{2} \int \{(\nabla\psi)^2 + \chi^2\psi^2\} d^3 y. \tag{28}$$

According to the first basic assumption of the stochastic quantization procedure, *the quantized field $\Psi(t)$ is assumed to be a diffusion process on ${}^*\mathcal{E}_{\{e_n\}}(R^3)$ with diffusion constant $\beta = \hbar/2$*, where $\hbar$ stands for Planck's constant divided by 2π. The probability density $P(\psi,t)$ of the quantized field can be determined through Eq. (14) with a given initial condition, say $P(\psi,0) = P_0(\psi)$.

To determine the unknown transformation A_t on ${}^*\mathcal{E}_{\{e_n\}}(R^3)$, we need the following second basic assumption; *the field equation* (28) *is valid with the substitution*

$$\begin{cases} \ddot{\psi}(\mathbf{x},t) \mapsto \frac{1}{2}(DD_* + D_*D)\Psi(t) \\ \psi(\mathbf{x},t) \mapsto \Psi(t). \end{cases} \tag{29}$$

Namely we have

$$\frac{1}{2}(DD_* + D_*D)\Psi(t) = -\frac{\delta}{\delta\Psi(t)} \frac{1}{2} \int \{(\nabla\Psi(t)^2 + \chi^2\Psi(t)^2\} d^3 x, \tag{30}$$

which can be written explicitly as

$$\frac{1}{2}\left\{\frac{\partial}{\partial t}\hat{A}_t\psi(\mathbf{x})+\int d^3y\,A_t\psi(\mathbf{y})\,\frac{\delta}{\delta\psi(\mathbf{y})}\hat{A}_t\psi(\mathbf{x})+\frac{\hbar}{2}\int d^3y\,\frac{\delta^2}{\delta\psi(\mathbf{y})^2}\hat{A}_t\psi(\mathbf{x})+\frac{\partial}{\partial t}A_t\psi(\mathbf{x})\right.$$
$$\left.+\int d^3y\hat{A}_t\psi(\mathbf{y})\,\frac{\delta}{\delta\psi(\mathbf{y})}A_t\psi(\mathbf{x})-\frac{\hbar}{2}\int d^3y\,\frac{\delta^2}{\delta\psi(\mathbf{y})^2}A_t\psi(\mathbf{x})\right\}\Big|_{\psi=\Psi(t)}=-\frac{\delta}{\delta\psi(\mathbf{x})}\frac{1}{2}\int\{(\nabla\psi(\mathbf{y})^2+\chi^2\psi(\mathbf{y})^2\}d^3y\Big|_{\psi=\Psi(t)}. \tag{31}$$

Finally, we make the following additional assumption on A_t and $\hat{A}_t$:

$$\tfrac{1}{2}(A_t+\hat{A}_t)\psi(\mathbf{x})=\hbar\,\frac{\delta}{\delta\psi(\mathbf{x})}S(\psi,t), \tag{32}$$

where $S(\cdot,t)$ is a functional on ${}^*\mathcal{E}_{\{e_n\}}(R^3)$ of the type $S(\psi,t)=[S(x^{(n)},t)]_n$.

It seems worthwhile to notice that Eqs. (14), (26), (31), and (32), with given initial conditions, are sufficient to characterize the quantized field $\Psi(t)$ completely.

Let us introduce the state functional

$$\Omega(\psi,t)=\sqrt{P(\psi,t)}\,\exp\{iS(\psi,t)\} \tag{33}$$

on ${}^*\mathcal{E}_{\{e_n\}}(R^3)$. Then Eqs. (14) and (31), with Eqs. (26) and (32), yield the following Schrödinger equation for the state functional $\Omega(\psi,t)$:

$$i\hbar\frac{\partial}{\partial t}\Omega(\psi,t)=\frac{1}{2}\int d^3x\left\{-\hbar^2\frac{\delta^2}{\delta\psi(\mathbf{x})^2}+[\nabla\psi(\mathbf{x})]^2+\chi^2\psi(\mathbf{x})^2\right\}\Omega(\psi,t). \tag{34}$$

Equation (34) with the given initial condition completely determines the behavior of the quantized field $\Psi(t)$.

Thus the stochastic quantization is shown to provide the same representation as the canonical one.

IV. QUANTIZATION OF DISSIPATIVE FIELDS

We study here a quantum mechanical description of dissipative fields interacting with chaotic thermal environments. (Quantum mechanical description of a particle dissipatively interacting with a thermal environment was investigated by the present author[9] and Skagerstam.[10])

Let us consider a classical acoustic wave field with field equation

$$\ddot{\psi}(\mathbf{x},t)=u^2\nabla^2\psi(\mathbf{x},t)-\lambda\dot{\psi}(\mathbf{x},t)+J(\mathbf{x},t), \tag{35}$$

where u denotes the sound velocity. Complicated interactions between the field and the thermal environment are renormalized phenomenologically into two terms; the friction coefficient $\lambda\geqq 0$ and the random source field $J(\mathbf{x},t)$. Note that the random source $J(\mathbf{x},t)$ is assumed to be a Gaussian white noise with mean 0 and covariance

$$E\{J(\mathbf{x},t)J(\mathbf{y},s)\}=J_0\delta^3(\mathbf{x}-\mathbf{y})\delta(t-s). \tag{36}$$

Because of the absense of a physically meaningful Lagrangian or Hamiltonian of such a dissipative field, neither canonical nor path-integral quantization can be performed. The only available quantization procedure seems to be the present stochastic one which demands no Lagrangian or Hamiltonian but the field equation in the generalized sense.

Quantization procedure is the same as the preceding nondissipative case:

(Q-1) *Quantized field $\Psi(t)$ is a diffusion process on* ${}^*\mathcal{E}_{\{e_n\}}(R^3)$ with $\beta=\hbar/2$.

(Q-2) *Field equation* (35) *is valid with the following substitution*:

$$\begin{cases}\ddot{\psi}(\mathbf{x},t)\mapsto\frac{1}{2}(DD_*+D_*D)\Psi(t),\\ \dot{\psi}(\mathbf{x},t)\mapsto\frac{1}{2}(D+D_*)\Psi(t),\\ \psi(\mathbf{x},t)\mapsto\Psi(t),\end{cases} \tag{37}$$

that is,

$$\frac{1}{2}(DD_*+D_*D)\Psi(t)=-\frac{\delta}{\delta\Psi(t)}\frac{u^2}{2}\int d^3x\{\nabla\psi(\mathbf{x})\}^2-\frac{\lambda}{2}(D+D_*)\Psi(t)+\frac{\delta}{\delta\Psi(t)}\int d^3x\,\psi(\mathbf{x})J(\mathbf{x},t). \tag{38}$$

(Q-3)

$$\frac{1}{2}(A_t+\hat{A}_t)\Psi(t)=\hbar\,\frac{\delta}{\delta\Psi(t)}S(\Psi(t),t) \tag{39}$$

(Q-4) *Probability amplitude $\Omega(\psi,t)$ is defined by*

$$\Omega(\psi,t)\equiv\sqrt{P(\psi,t)}\,\exp\{iS(\psi,t)\}. \tag{40}$$

Finally Eqs. (14) and (38), with Eqs. (26) and (39), give us the following Schrödinger type nonlinear equation.

$$i\hbar\frac{\partial}{\partial t}\Omega(\psi,t)=\frac{1}{2}\int d^3x\left\{-\hbar^2\frac{\delta^2}{\delta\psi(\mathbf{x})^2}+u^2[\nabla\psi(\mathbf{x})]^2\right\}\Omega(\psi,t)+\frac{i\lambda\hbar}{2}\log\frac{\bar{\Omega}(\psi,t)}{\Omega(\psi,t)}\cdot\Omega(\psi,t)-\int d^3x\,\psi(\mathbf{x})J(\mathbf{x},t)\Omega(\psi,t). \tag{41}$$

Equation (41) completely characterizes a quantum mechanical description of the dissipative acoustic wave field (35). Notice that the dissipative feature of the wave field appears in Eq. (41) through the nonlinear potential

$$\lambda S(\psi,t)=\frac{i\hbar}{2}\lambda\log\frac{\bar{\Omega}(\psi,t)}{\Omega(\psi,t)}. \tag{42}$$

That this is so can be seen as follows:

Neglecting the random potential term in Eq. (41), the expectation value of the total energy operator

$$H=\frac{1}{2}\int d^3x\left\{-\hbar^2\frac{\delta^2}{\delta\psi(\mathbf{x})^2}+u^2[\nabla\psi(\mathbf{x})]^2\right\} \tag{43}$$

with respect to the characteristic state $\Omega(\cdot,t)$ is given by

$$E(t)=\int\bar{\Omega}(\psi,t)\left[\frac{1}{2}\int d^3x\left\{-\hbar^2\frac{\delta^2}{\delta\psi(\mathbf{x})^2}+u^2[\nabla\psi(\mathbf{x})]^2\right\}\right]\Omega(\psi,t)\delta\psi=\int\bar{\Omega}(\psi,t)H\Omega(\psi,t)\delta\psi. \tag{44}$$

As the state functional $\Omega(\cdot,t)$ is a characteristic solution of Eq. (41) without the random potential, we can immediately calculate the rate of energy loss

$$\frac{d}{dt}E(t)=\int\left\{\frac{\partial\bar{\Omega}(\psi,t)}{\partial t}H\Omega(\psi,t)+\bar{\Omega}(\psi,t)H\frac{\partial\Omega(\psi,t)}{\partial t}\right\}\delta\psi$$

$$= -\frac{\lambda}{4}\hbar^2 \int \left\{ \int d^3x \left| \frac{\delta}{\delta\psi(\mathbf{x})} \log \frac{\bar{\Omega}(\psi, t)}{\Omega(\psi, t)} \right|^2 \left| \Omega(\psi, t) \right|^2 \right\} \delta\psi$$

$$< 0. \tag{45}$$

Thus we have shown that the Schrödinger type non-linear functional differential equation (41) provides a quantum mechanical description of the dissipative field.

V. APPLICATION TO THE LASER ELECTRIC FIELD IN THE LOSSY CAVITY

As a practical problem, there are two typical examples of dissipative fields; the viscous quantum fluid of the nucleonic matter and the laser electric field in the lossy cavity.

In the former case, the relative density disturbance of the nucleonic matter is considered as a viscous fluid. It is described by the Navier—Stokes equation

$$\rho\dot{\mathbf{v}} + \rho(\mathbf{v}\cdot\nabla)\mathbf{v} = -u^2\nabla\rho + \eta\nabla^2\mathbf{v}, \tag{46}$$

and the equation of continuity

$$\dot{\rho} + \nabla\cdot(\rho\mathbf{v}) = 0, \tag{47}$$

where u and η denote the sound velocity and the viscosity, respectively. As we are interested in the quantum fluctuation of the relative density disturbance, it is enough to consider the linearized equation

$$\ddot{\rho} - u^2\nabla^2\rho - \nu\nabla^2\dot{\rho} = 0, \tag{48}$$

where ν denotes the kinematic viscosity. Equation (48) may be utilized to analyze the nuclear dissipative phenomena such as the giant resonance and the heavy-ion collision. Quantum mechanical analysis of such phenomena needs the quantization of the dissipative field (48) by the present method.[11]

In the latter case, Maxwell's equations in mks units

$$\begin{aligned} &\nabla\cdot\mathbf{D} = 0, \qquad \mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P}, \\ &\nabla\cdot\mathbf{B} = 0, \qquad \mathbf{B} = \mu_0\mathbf{H}, \\ &\nabla\times\mathbf{E} = -\dot{\mathbf{B}}, \quad \nabla\times\mathbf{H} = \mathbf{J} + \dot{\mathbf{D}}, \end{aligned} \tag{49}$$

and the phenomenological Ohmic loss relation

$$\mathbf{J} = \sigma\mathbf{E} \tag{50}$$

give us the following dissipative field equation:

$$-\nabla^2\mathbf{E} + \mu_0\sigma\dot{\mathbf{E}} + \mu_0\epsilon_0\ddot{\mathbf{E}} = -\mu_0\ddot{\mathbf{P}}. \tag{51}$$

The laser electric field $\mathbf{E}$ is forced by the imposed polarization vector $\mathbf{P}$ and damped by the Ohmic energy loss in the cavity through Eq. (51). Equation (51) completely determines the laser electric field in the lossy cavity without detailed description of the mechanism of the cavity loss. The conventional treatment of such a lossy laser is to describe the atoms in a laser quantum mechanically and the electric field (51) classically.[12] Therefore, quantum mechanical description of the laser electric field in the lossy cavity seems to be needed. (A constructive approach to quantum mechanics of the lossy laser from a fundamental point of view was given by Hepp and Lieb.[13] They made use of the Heisenberg representation, whereas our standpoint, explained here, may be understood as the Schrödinger representation.)

In this section, we investigate quantum mechanics of the laser electric field described by the dissipative field equation (51) in much detail.

For simplicity, let us assume the laser electric field be linearly polarized, that is,

$$\mathbf{E}(\mathbf{x}, t) = \mathbf{e}E(x, t) \tag{52}$$

for some unit vector $\mathbf{e}$. Then Eq. (51) becomes

$$\ddot{E}(\mathbf{x}, t) = c^2\nabla^2 E(\mathbf{x}, t) - \frac{1}{\epsilon_0}\dot{E}(\mathbf{x}, t) - \frac{1}{\epsilon_0}\mathbf{e}\cdot\ddot{\mathbf{P}}(\mathbf{x}, t), \tag{53}$$

which is evidently of the same form as Eq. (35). Therefore, quantization of the laser electric field in the lossy cavity can be performed straightforwardly as is explained in the preceding section.

Quantum mechanical behavior of the laser electric field (53) is characterized by the following Schrödinger type non-linear functional differential equation.

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\Omega(E, t) &= \tfrac{1}{2}\int d^3x \left\{ -\hbar^2 \frac{\delta^2}{\delta E(\mathbf{x})^2} + c^2[\nabla E(\mathbf{x})]^2 \right\} \Omega(E, t) \\ &\quad + \frac{i\hbar\sigma}{2\epsilon_0}\log\frac{\bar{\Omega}(E, t)}{\Omega(E, t)}\cdot\Omega(E, t) \\ &\quad + \frac{1}{\epsilon_0}\int d^3x\, E(\mathbf{x})\mathbf{e}\cdot\mathbf{P}(\mathbf{x}, t)\Omega(E, t). \end{aligned} \tag{54}$$

To further simplify the analysis, we take the complete normalized orthogonal system $\{e_n\}_{n\in N} \subset \mathcal{S}(R^3)$ to be the eigenfunctions of the three-dimensional Laplacian ∇^2. Namely we have

$$\nabla^2 e_n(\mathbf{x}) = -k_n^2 e_n(\mathbf{x}), \quad n\in N, \tag{55}$$

where $-k_n^2$'s are the eigenvalues.

As the quantized electric field $E(\mathbf{x}) \in {}^*\mathcal{E}_{\{e_n\}}(R^3)$ and the state functional $\Omega(E, t)$ are decomposed as

$$E(\mathbf{x}) = \Big[\sum_{p\le n} a_p^{(n)} e_p(\mathbf{x})\Big]_n$$

and

$$\Omega(E, t) = [\Omega(a^{(n)}, t)]_n,$$

respectively, we can rewrite Eq. (54) as

$$\begin{aligned} &\left[i\hbar\frac{\partial}{\partial t}\Omega(a^{(n)}, t) \right]_n \\ &= \left[-\frac{\hbar^2}{2}\sum_{p\le n} \left\{ \frac{\partial^2}{\partial a_p^{(n)2}} + \frac{c^2k_p^2}{2} a_p^{(n)2} \right\} \Omega(a^{(n)}, t) \right]_n \\ &\quad + \left[\frac{i\hbar\sigma}{2\epsilon_0}\log\frac{\bar{\Omega}(a^{(n)}, t)}{\Omega(a^{(n)}, t)}\cdot\Omega(a^{(n)}, t) \right]_n \\ &\quad + \frac{1}{\epsilon_0}\left[\sum_{p\le n} a_p^{(n)}\ddot{P}_p(t)\Omega(a^{(n)}, t) \right]_n, \end{aligned} \tag{56}$$

where we have made the abbreviation $P_p(t) = \int e_p(\mathbf{x})\mathbf{e}\cdot\mathbf{P}(\mathbf{x}, t)d^3x$. To solve the functional differential equation (54), it is enough to consider its finite dimensional cross section ("cross section" denotes the finite dimensional element inside the bracket $[\ ]_n$)

$$i\hbar\frac{\partial}{\partial t}\Omega(a^{(n)},t)=-\frac{\hbar^2}{2}\sum_{p\le n}\left\{\frac{\partial^2}{\partial a_p^{(n)2}}+\frac{c^2k_p^2}{2}a_p^{(n)2}\right\}$$
$$\times\Omega(a^{(n)},t)+\frac{i\hbar\sigma}{2\epsilon_0}\log\frac{\bar{\Omega}(a^{(n)},t)}{\Omega(a^{(n)},t)}\cdot\Omega(a^{(n)},t)$$
$$+\frac{1}{\epsilon_0}\sum_{p\le n}a_p^{(n)}\ddot{P}_p(t)\Omega(a^{(n)},t). \tag{57}$$

If we put

$$\Omega(a^{(n)},t)=\prod_{p\le n}\Omega(a_p^{(n)},t), \tag{58}$$

Eq. (57) reduces to the following one-dimensional ones

$$i\hbar\frac{\partial}{\partial t}\Omega(a_p^{(n)},t)$$
$$=\left\{-\frac{\hbar^2}{2}\frac{\partial^2}{\partial a_p^{(n)2}}+\frac{c^2k_p^2}{2}a_p^{(n)2}\right\}\Omega(a_p^{(n)},t)$$
$$+\frac{i\hbar\sigma}{2\epsilon_0}\log\frac{\bar{\Omega}(a_p^{(n)},t)}{\Omega(a_p^{(n)},t)}\cdot\Omega(a_p^{(n)},t)$$
$$+\frac{1}{\epsilon_0}a_p^{(n)}\ddot{P}_p(t)\Omega(a_p^{(n)},t),\quad 1\le p\le n. \tag{59}$$

A characteristic solution of Eq. (59) is obtained by introducing the so-called photon coherent state $\Omega_c(a_p^{(n)};z)$, $z\in Z$ (Z denotes the complex plane). The coherent state is defined as an eigenstate of the annihilation operator[14];

$$\left(\frac{1}{2\hbar k_pc}\right)^{1/2}\left(\frac{\partial}{\partial a_p^{(n)}}+ck_pa_p^{(n)}\right)\Omega_c(a_p^{(n)};z)$$
$$=z\Omega_c(a_p^{(n)};z). \tag{60}$$

Namely, the coherent state

$$\Omega_c\left(a_p^{(n)};\alpha_p(t)-\frac{i}{ck_p}\dot{\alpha}_p(t)\right)$$
$$\times\exp\left[-\frac{1}{2\hbar ck_p}\dot{\alpha}_p^2(t)-\frac{i}{\hbar}\{\alpha_p(t)\dot{\alpha}_p(t)+g_p(t)\}\right] \tag{61}$$

solves Eq. (59), provided that $\alpha_p(t)$ satisfies the classical equation of motion

$$\ddot{\alpha}_p(t)=-\frac{\sigma}{\epsilon_0}\dot{\alpha}_p(t)-c^2k_p^2\alpha_p(t)-\frac{1}{\epsilon_0}\ddot{P}_p(t), \tag{62}$$

and $g_p(t)$ is related to $\alpha_p(t)$ as

$$\dot{g}_p(t)+g_p(t)=\frac{\hbar ck_p}{2}+\frac{\hbar c^2k_p^2}{2}\alpha_p(t)^2-\frac{1}{2}\dot{\alpha}_p(t)^2. \tag{63}$$

Correspondingly, quantum mechanical behavior of the laser electric field in the lossy cavity can be represented by the photon coherent state

$$\Omega_c(E,t)=[\Omega_c(a^{(n)},t)]_n$$
$$=\left[\prod_{p\le n}\Omega_c\left(a_p^{(n)};\alpha_p(t)-\frac{i}{ck_p}\dot{\alpha}_p(t)\right)\right.$$
$$\left.\times\exp\left[-\frac{1}{2\hbar ck_p}\dot{\alpha}_p(t)^2-\frac{i}{\hbar}\{\alpha_p(t)\dot{\alpha}_p(t)+g_p(t)\}\right]\right]_n. \tag{64}$$

Therefore, we may be allowed to mention that the quantized laser electric field in the lossy cavity fluctuates around its classical value $E(\mathbf{x},t)=[\sum_{p\le n}\alpha_p(t)e_p(\mathbf{x})]_n$ with minimum uncertainty. This provides a quantum theoretical background to the validity of the conventional semiclassical treatment of the lossy laser.

[1]E. Nelson, Phys. Rev. **150**, 1079 (1966); *Dynamical Theories of Brownian Motion* (Princeton U. P., Princeton, N.J., 1967).
[2]T. Dankel, Jr., Arch. Ration. Mech. Anal. **37**, 192 (1970); J. Math. Phys. **18**, 253 (1977).
[3]J-P. Caubet, *Le Mouvement Brownien Relativiste*, Lecture Notes in Mathematics No. 559 (Springer, Berlin, 1976).
[4]L. de la Pena-Auerbach, J. Math. Phys. **12**, 453 (1971).
[5]W.J. Lehr and J.L. Park, J. Math. Phys. **18**, 1235 (1977).
[6]F. Guerra and P. Ruggiero, Phys. Rev. Lett. **31**, 1022 (1973).
[7]F. Guerra, J. de Physique Suppl. Colloque **34**, C1—95 (1973); "On the Connection between Euclidean-Markov Field Theory and Stochastic Quantization," *C*-Algebras and Their Applications to Statistical Mechanics and Quantum Field Theory*, edited by D. Kastler (North-Holland, Amsterdam, 1976).
[8]R. Kurata, Su-ri Kagaku (Math. Sci.) **170**, 79 (1977) (in Japanese).
[9]K. Yasue, J. Stat. Phys. **16**, 113 (1977).
[10]B. Skagerstam, J. Math. Phys. **18**, 308 (1977).
[11]K. Yasue, Phys. Lett. B **64**, 239 (1976).
[12]D. Rogovin and M. Scully, Phys. Rep. C **25**, 175 (1976).
[13]K. Hepp and E. Lieb, "Constructive Macroscopic Quantum Electrodynamics," in *Constructive Quantum Field Theory*, edited by G. Velo and A. Wightman (Springer, Belin, 1973), Lecture Notes in Physics, Vol. 25.
[14]J.R. Klauder and E.C.G. Sudarshan, *Fundamentals of Quantum Optics* (Benjamin, New York, 1968).

Unified theory of direct interaction between particles, strings, and membranes

Patricio S. Letelier [a)]

Department of Physics, Boston University, Boston, Massachusetts 02215
(Received 5 August 1977)

A model of generalized relativistic membranes that contains as special cases, particles, geometric strings and geometric membranes and other new one- and two-dimensional objects is studied. The equations of motion for such objects in direct interaction are studied. The constraints on the interaction due to the freedom of gauge of the free model are solved. A scalar, a vectorial, and a tensorial type of interaction are discussed. Conservation theorems associated with Poincaré invariance of the action are studied, as well as the generalization for action-at-a-distance theories of the action and reaction law.

1. INTRODUCTION

Motivated by the recent applications of the quantized model of geometric strings to the study of elementary particles,[1] we studied a classical theory of direct interactions between geometric membranes.[2]

Analyzing the models of direct interaction between geometric strings[3] and geometric membranes,[2] we arrived at the conclusion that considerable improvements and new insights would be gained if these theories were cast in a manifestly gauge invariant form. When the previous program was accomplished,[4,5] we realized that the gauge invariant theories of direct interaction between particles,[6] between strings, and between membranes have several things in common, since the formalisms of these three theories are similar. In particular, the conservation theorems for these theories can be derived in a completely parallel way.[7] Thus, we concluded that a unification of these theories in a single formalism would provide us with a better understanding, as well as show us in a better way their differences and similarities. The realization of the above program is discussed in the present paper.

The generalization of the model of geometric string to include particles on it is as old a problem as the theory of geometric stings itself. The motivation to realize this program has been, mainly, to obtain a richer dual resonant model of particles[8] and to solve the problem of quark confinement.[9]

In this paper we study how particles and strings can be attached to a membrane, at the classical level, and in a general way. As a particular case we have the action for particles attached to one string. The general object of the theory is a membrane with a one-parameter family of strings and a two-parameter family of particles attached to it. By switching off the strings and the membrane, the particles and the membrane, the particles and the strings, we recover the action for particles, strings and a memberane, respectively.

In Sec. 2 we study the free membrane equation of motion; in particular we study the behavior of the edges of open objects.

In Sec. 3 we study the direct interaction between membranes; in particular we study and solve the constraints that are a consequence of the gauge freedom of the theory. As a particular case we reproduce the interaction between particles, between geometric stings, and between geometric membranes previously studied by us.[7]

In Sec. 4 we examine some particular cases of tensorial, vectorial, and scalar interactions.

In Sec. 5 the conservation theorems associated with Poincaré invariance of the action for the system of direct interacting membranes are studied. The conservation theorems generalize the conservation theorems previously studied by us.

In Sec. 6 we study the "action and reaction law" for different types of membranes in direct interaction.

In Sec. 7 we discuss some possible applications and generalizations of the previous results.

2. FREE MEMBRANES

In this section we discuss the action for a free relativistic membrane, with a family of strings and a "cloud" of particles moving on it. In particular we study the case where the particles and the strings are fixed on the membrane. The space–time where the different objects evolve in Minkowski space with metric $\eta_{\mu\nu}$ of signature -2.

First, we study the Lorentz scalar densities that can be formed at a point on the membrane with the 4-velocity of one particle $v^\mu(t)$; with the bivector that spans the string world sheet $\Sigma^{\mu\nu}(T^A)$; and with the trivector that spans the membrane world tube $\Sigma^{\mu\nu\lambda}(\xi^U)$. The Greek indices run from 0 through 3, A, B, C, $\cdots$, from 0 to 1 and U, V, W, $\cdots$, from 0 through 2. The parameters t, T^A, and ξ^U do not need to be Lorentz invariant. We will require that v^μ, $\Sigma^{\mu\nu}$, and $\Sigma^{\mu\nu\lambda}$ transforms as scalar densities under the reparametrization of t, T^A, and ξ^U respectively. This point will be studied in great detail later in this section.

A particle world line on a string world sheet can be described by a function α, such that $\alpha(T^A)=0$. Thus the conditions for a particle to be on a string can be implemented by

$$t=t(T^A), \quad \alpha(T^A)=0. \tag{2.1a}$$

A string world sheet inside a membrane world tube can

[a)] Present address: Departamento de Física, Universidade de Brasília 70.000 Brasília, Brazil

 0022-2488/78/1909-1898$1.00

be described by a function β such that $\beta(\zeta^U)=0$. Thus the condition for a string to be on a membrane can be implemented by

$$T^A = T^A(\zeta^U), \quad \beta(\zeta^U)=0. \tag{2.1b}$$

A particle world line inside a membrane world tube can be described by two functions r and s such that: $r(\zeta^U)=0$ and $s(\zeta^U)=0$. Thus the condition for a particle to be on the membrane can be implemented by

$$t = t(\zeta^U), \quad r(\zeta^U)=0, \quad s(\zeta^U)=0. \tag{2.1c}$$

If the particle is on the string and the string is on the membrane, we have

$$t = t[T^A(\zeta^U)], \quad \alpha[T^A(\zeta^U)]=0, \quad \beta(\zeta^U)=0. \tag{2.1d}$$

$\Sigma^{\mu\nu}$ and $\Sigma^{\mu\nu\lambda}$ are a simple bivector and a simple trivector,[10] respectively, so

$$\Sigma^{\mu[\nu}\Sigma^{\lambda\sigma]}=0, \tag{2.2a}$$

$$\Sigma^{[\mu\nu\lambda}\Sigma^{\rho]\sigma\pi}=0, \tag{2.2b}$$

$$\Sigma^{\mu[\nu\lambda}_{\ \ \nu}\Sigma^{\rho\sigma\pi]}=0, \tag{2.2c}$$

where $[\cdots]$ denotes antisymmetrization in the enclosed indices.

The particle and the string lie on the membrane, thus

$$v^{[\mu}\Sigma^{\rho\sigma\tau]}=0\,, \tag{2.3a}$$

$$\Sigma^{\mu[\nu}\Sigma^{\rho\sigma\tau]}=0. \tag{2.3b}$$

If the particle lies on the string, we have

$$v^{[\mu}\Sigma^{\rho\tau]}=0. \tag{2.4}$$

All the independent invariants that can be formed out of v^μ, $\Sigma^{\mu\nu}$, $\Sigma^{\mu\nu\lambda}$, $\eta_{\mu\nu}$, and $\epsilon_{\mu\nu\lambda\sigma}$ that satisfy (2.2) and (2.3) are

$$v^2 \equiv v^\mu v_\mu, \tag{2.5a}$$

$$\gamma \equiv \frac{1}{2!}\Sigma^{\mu\nu}\Sigma_{\mu\nu}, \tag{2.5b}$$

$$\sigma \equiv \frac{1}{3!}\Sigma^{\mu\nu\lambda}\Sigma_{\mu\nu\lambda}, \tag{2.5c}$$

$$\kappa \equiv (v^\mu v^\alpha \Sigma_{\mu\nu}\Sigma^\nu_\alpha)^{1/2}, \tag{2.5d}$$

$$\omega \equiv (v^\mu \Sigma_{\mu\alpha\beta}\Sigma^{\alpha\beta})^{2/3}. \tag{2.5e}$$

That (2.5) are the only independent invariants that satisfy (2.2) and (2.3) can be seen by forming all the possible scalar combinations of the elements previously described. Some useful identities are

$$\Sigma^{\mu\nu}\Sigma_{\nu\lambda}\Sigma^{\lambda\sigma} = 2\gamma\Sigma^{\sigma\mu}, \tag{2.6a}$$

$$\Sigma^{\mu\alpha\beta}\Sigma_{\alpha\beta\rho}\Sigma^{\rho\nu\lambda} = 3!\ \sigma\Sigma^{\mu\nu\lambda}, \tag{2.6b}$$

$$\Sigma^{\mu\nu\lambda}\Sigma_{\nu\lambda}\Sigma^{\alpha\beta}_{\mu}\Sigma_{\alpha\beta} = 4\sigma\gamma, \tag{2.6c}$$

$$^*\Sigma^{\mu\nu}{}^*\Sigma^\lambda_\nu = 2\gamma\eta^{\mu\nu} - \tfrac{1}{2}\Sigma^{\mu\nu}\Sigma^\lambda_\nu, \tag{2.6d}$$

$$(^*\Sigma^{\mu\nu} \equiv \epsilon^{\mu\nu\alpha\beta}\Sigma_{\alpha\beta}). \tag{2.6e}$$

Let us consider a 2-family of particles characterized by a vector $v^\mu = v^\mu[t(\zeta^U),\ \zeta^1,\ \zeta^2]$ and a density $\rho = \rho[t(\zeta^U),\ \zeta^U]$. Note that we have adapted the parmetrization of the family of particles to the parametrization of the membrane, so now the functions r and s that appear in (2.1c) can be taken as $\zeta^1 = \text{const}$ and $\zeta^2 = \text{const}$ and the different world lines of the particles will be described by different values of the constants.

Let us also consider a one parameter family of strings characterized by a bivector $\Sigma^{\mu\nu} = \Sigma^{\mu\nu}[T^A(\zeta^U), h]$, where h is either ζ^1 or ζ^2, and a density $\lambda = \lambda[T^A(\zeta^U), \zeta^U]$. Now the function β that appears in (2.1b) can be taken as $h = \text{const}$, described by different values of the constant.

We shall demand that v^μ, $\Sigma^{\mu\nu}$, and $\Sigma^{\mu\nu\lambda}$ transform like scalar densities under reparametrization as mentioned above. This means

$$v^\mu \xrightarrow{t} {}'v^\mu = \frac{\partial t}{\partial t'}v^\mu, \tag{2.7a}$$

$$\Sigma^{\mu\nu} \xrightarrow{T} {}'\Sigma^{\mu\nu} = J(T^A/T^{B'})\Sigma^{\mu\nu}, \tag{2.7b}$$

$$\Sigma^{\mu\nu\lambda} \xrightarrow{\zeta} {}'\Sigma^{\mu\nu\lambda} = J(\zeta^U/\zeta^{U'})\Sigma^{\mu\nu\lambda}, \tag{2.7c}$$

$$v^\mu \xrightarrow{\zeta} {}'v^\mu = v^\mu, \tag{2.8a}$$

$$\Sigma^{\mu\nu} \xrightarrow{\zeta} {}'\Sigma^{\mu\nu} = \Sigma^{\mu\nu} = \Sigma^{\mu\nu}, \tag{2.8b}$$

where J is the Jacobian of the indicated transformations. We shall always assume that all the transformations are nonsingular and that the transformation of the time coordinate t, T^0, and ζ^0 leave invariant the interval $(-\infty, +\infty)$.

A "representation" of v^μ, $\Sigma^{\mu\nu}$, and $\Sigma^{\mu\nu\lambda}$ is given by

$$v^\mu = \frac{dz^\mu}{dt}\,, \tag{2.9}$$

$$\Sigma^{\mu\nu} = \epsilon^{AB}y^\mu_A y^\nu_B, \tag{2.10}$$

$$\Sigma^{\mu\nu\lambda} = \epsilon^{UVW}x^\mu_U x^\nu_V x^\lambda_W, \tag{2.11}$$

$$y^\mu_A = \frac{\partial}{\partial T^A}y^\mu, \quad x^\mu_U = \frac{\partial}{\partial \zeta^U}x^\mu, \tag{2.12}$$

where ϵ^{AB} and ϵ^{UVW} are the two- and three-dimensional Levi-Civita symbols normalized as follows:

$$\epsilon^{01} = \epsilon^{012} = 1. \tag{2.13}$$

From (2.3) we obtain

$$v^\mu = b^U x^\mu_U, \tag{2.14}$$

$$y^\mu_A = c^U_A x^\mu_U. \tag{2.15}$$

If the particles are moving along the strings, we get

$$v^\mu = a^A y^\mu_A = a^A c^U_A x^\mu_U, \tag{2.16}$$

where a^A, b^U, and c^U_A are functions that transform as their indices indicate. Note that they are restricted by conditions like $v^\mu v_\mu = b^U b^U x^\mu_U x_{\mu U} > 0$, etc. The functions b^U and c^U_A may be used to describe the motion of the particles and strings relative to the membrane and the function a^A to describe the motion of particles on strings relative to the strings.

The action that describes the system of particles and strings moving on the membrane is taken as

$$\mathcal{A}_* = \int d^3\zeta\, \Omega_*[\rho^2 v^2,\ \lambda^2\gamma,\ \rho\lambda\kappa,\ (\rho\lambda)^{3/2}\omega,\ e^2\sigma], \tag{2.17}$$

where $d^3\zeta$ denotes $d\zeta^0 d\zeta^1 d\zeta^2$. The integration in ζ^0 runs from $-\infty$ to $+\infty$ and the range of ζ^1 and ζ^2 will depend on the type of object with which we are dealing; they will be specified later. e is a constant parameter.

The invariants (2.5) can be expressed in terms of x^μ_U, through (2.9)–(2.15). Thus, (2.17) will give an equation for x^μ_U alone. In order to have an action invariant under a reparametrization of all the param-

eters involved in the theory we shall require that Ω_* be homogeneous function of degree one-half in its variables and that the densities transform like

$$\rho \xrightarrow{\xi} {}'\rho = J\left(\frac{\xi^U}{\xi^{v'}}\right)\rho, \tag{2.18a}$$

$$\lambda \xrightarrow{\xi} {}'\lambda = J\left(\frac{\xi^U}{\xi^{v'}}\right)\lambda, \tag{2.18b}$$

$$\rho \xrightarrow{t} {}'\rho = \frac{\partial t'}{\partial t}\rho, \tag{2.19a}$$

$$\lambda \xrightarrow{T} {}'\lambda = J\left(\frac{T^{A'}}{T^B}\right)\lambda. \tag{2.19b}$$

Now we will specialize the Lagrangian density Ω_* in order to describe particles and strings attached to the membrane. In this case $\omega = 0$ and $\kappa = v(-2\gamma)^{1/2}$, so the dependence of the action on the invariants reduces to

$$\mathcal{A} = \int d^3\zeta\, \Omega(\rho^2 v^2, \lambda^2\gamma, e^2\sigma). \tag{2.20}$$

Furthermore, in this case the geometry of the problem gives us a natural parameterization. Let us choose the parameters ζ^0 and T^0 in such a way that x_0^μ and y_0^μ represent the velocity at each point of the membrane and at each point of the string respectively. Then, if the particles and the strings are fixed with respect to the membrane we will have

$$\frac{\partial y^\mu}{\partial T^0} \| \frac{\partial x^\mu}{\partial \zeta^0} \| \frac{\partial z^\mu}{\partial t}, \tag{2.21}$$

and without loss of generality we can demand

$$\frac{\partial y^\mu}{\partial T^0} = \frac{\partial x^\mu}{\partial \zeta^0} = \frac{\partial z^\mu}{\partial t}. \tag{2.22}$$

We can also choose the parameter ζ^1 and T^1 in such a way that one of the spacelike vectors of the membrane is parallel to the string spacelike vector. Moreover, without loss of generality we can impose

$$y_1^\mu = x_1^\mu. \tag{2.23}$$

When we choose this parametrization the family of strings will be described by the parameter ζ^2, i.e., $h = \zeta^2$.

From (2.22), (2.23), (2.5), and (2.9)–(2.15) we have that v^2 and γ are related to the metric tensor of the membrane world tube, σ_{UV}, as follows:

$$\sigma = \det\sigma_{UV}, \quad (\sigma_{UV} = x_U^\mu x_{\mu V}), \tag{2.24a}$$

$$v^2 = \sigma_{00}, \tag{2.24b}$$

$$\gamma = \det\sigma_{AB}. \tag{2.25c}$$

The identification (2.22) and (2.23) tells us that we need only one parameter $\tau \equiv t = T^0 = \zeta^0$ and one parameter $\xi = T^1 = \zeta^1$. We shall also introduce the notation $\eta \equiv \zeta^3$. In the general case $e \neq 0$, $\rho \neq 0$, and $\lambda \neq 0$ the freedom of reparametrization (gauge) left by these identifications

$$\tau \to \tau' = \tau'(\zeta^U), \tag{2.25a}$$

$$\xi \to \xi' = \xi'(\xi), \tag{2.25b}$$

$$\eta \to \eta' = \eta'(\eta), \tag{2.25c}$$

From (2.18) and (2.19) we find that ρ and λ transform under (2.25) as

$$\rho \to {}'\rho = J\left(\frac{\xi, \eta}{\xi', \eta'}\right)\rho, \tag{2.26}$$

$$\lambda \to {}'\lambda = \frac{\partial \eta}{\partial \eta'}\lambda. \tag{2.27}$$

Thus, in the adapted parametrization used, $\rho = \rho(\xi, \eta)$ and $\lambda = \lambda(\eta)$.

The action (2.20) describes free particles when $\lambda = e = 0$ and $\rho = m\delta(\xi)\delta(\eta)$, free geometric strings when $\rho = e = 0$ and $\lambda = M\delta(\eta)$, and free geometric membranes when $\rho = \lambda = 0$ and $e = \mathfrak{M}$, where m is the mass of the particle and M and $\mathfrak{M}$ are constants [see Refs. 4 and 5]. When $\rho = f(\xi)\delta(\eta)$, $\lambda = M\delta(\eta)$, and $e = 0$ we have that Ω describes a string with particles attached to it, i.e., a generalization of Takabayasi's "realistic" model of strings,[8] that we shall call a p-string. Takabayasi's model is the particular p-string with Lagrangian density strings,[8] that we shall call a p-string. Takabayasi's model is the particular p-string with Lagrangian density

$$\Omega = \delta(\eta)(m^2 v^2 - M^2\gamma)^{1/2}. \tag{2.28}$$

The gauge freedom of this model is $\tau \to \tau' = \tau(\xi, \tau)$ and $\xi \to \xi' = \xi + a$, where a is a constant.

When $\rho = 0$, $\lambda \neq 0$, and $e \neq 0$ we have the generalization of the p-string model to one dimension higher, i.e., an s-membrane. The general case will be named $p-s$-membrane or simply membrane.

For particles we have that the gauge freedom is

$$\tau \to \tau' = \tau'(\tau), \tag{2.29}$$

for geometric strings,

$$\zeta^A \to \zeta^{A'} = \zeta^{A'}(\zeta^B), \tag{2.30}$$

and for geometric membranes,

$$\zeta^U \to \zeta^{U'} = \zeta^{U'}(\zeta^V). \tag{2.31}$$

A discussion of this point can be found in Refs. 4 and 5.

The range of ξ is taken as $\bar{\lambda}^1 < \xi < \bar{\delta}^1$ for open strings and for open membranes in the direction $\partial/\partial\xi$; and as $\lambda^1 < \xi \leq \delta^1$ for closed strings and closed membranes in the same direction; and as $-\infty < \xi < +\infty$ for particles. The range of η is taken as $\bar{\lambda}^2 < \eta < \bar{\delta}^2$ for open membranes in the direction $\partial/\partial\eta$; as $\lambda^2 < \eta \leq \delta^2$ for closed membranes in the same direction; and as $-\infty < \eta < +\infty$ for strings and particles.

In the case of closed strings and closed membranes we require that the objects be closed in a geometrical sense, as well as being smooth, i.e., for strings we require

$$x^\mu(\tau, \lambda^1) = x^\mu(\tau, \delta^1), \tag{2.32a}$$

$$\left.\frac{\partial x^\mu}{\partial \xi}\right|_{\xi=\lambda^1} = \left.\frac{\partial x^\mu}{\partial \xi}\right|_{\xi=\delta^1}, \tag{2.32b}$$

and for membranes

$$x^\mu(\tau, \lambda^1, \eta) = x^\mu(\tau, \delta^1, \eta), \tag{2.33a}$$

$$x^\mu(\tau, \xi, \lambda^2) = x^\mu(\tau, \xi, \delta^2), \tag{2.33b}$$

$$\frac{\partial}{\partial \xi} x^\mu \Big|_{\xi=\lambda^1} = \frac{\partial x^\mu}{\partial \xi}\Big|_{\xi=\delta^1}, \tag{2.33c}$$

$$\frac{\partial x^\mu}{\partial \eta}\Big|_{\eta=\lambda^2} = \frac{\partial x^\mu}{\partial \eta}\Big|_{\eta=\delta^2}. \tag{2.33d}$$

In the case of open strings and open membranes the gauge freedom is further restricted by

$$\frac{\partial \xi^1}{\partial \tau}\Big|_{\xi=\bar{\lambda}^1,\bar{\delta}^1} = 0, \tag{2.34a}$$

and

$$\frac{\partial \xi^1}{\partial \tau}\Big|_{\xi=\bar{\lambda}^1,\bar{\delta}^1} = \frac{\partial \xi^1}{\partial \eta}\Big|_{\xi=\bar{\lambda}^1,\bar{\delta}^1} = 0, \tag{2.35a}$$

$$\frac{\partial \eta^1}{\partial \tau}\Big|_{\eta=\bar{\lambda}^2,\bar{\delta}^2} = \frac{\partial \eta^1}{\partial \xi}\Big|_{\eta=\bar{\lambda}^2,\bar{\delta}^2} = 0, \tag{2.35b}$$

respectively, in order that the boundaries remain invariant.

We also consider semi-open membranes. In this case the conditions are either (2.33a), (2.33c), and (3.35b) or (2.33b), (2.33d), and (3.35a) depending on which are the open edges. The conditions (2.32)—(2.35) were studied in Refs. 4 and 5. Note that for p-strings, p, s-membranes, and $p-s$-membranes, the conditions (2.34) and (2.35) are automatically satisfied.

Let us consider the variation of the membrane world tube

$$x^\mu \to x^{\mu\prime} = x^\mu + \delta x^\mu, \tag{2.36}$$

$$\delta x^\mu(\tau = \pm\infty, \xi, \eta) = 0, \tag{2.37a}$$

$$\delta x^\mu(\tau, \xi = \lambda^1, \bar{\lambda}^1, \delta^1, \bar{\delta}^1, \eta) = 0, \tag{2.37b}$$

$$\delta x^\mu(\tau, \xi, \eta = \lambda^2, \bar{\lambda}^2, \delta^2, \bar{\delta}^2) = 0. \tag{2.37c}$$

The equations of motion for the body of the free membrane are obtained by demanding that the action (2.20) be stationary under (2.36)—(2.37). They are

$$\frac{\partial}{\partial \xi^U}\frac{\partial}{\partial x^\mu_U}\Omega = 0. \tag{2.38}$$

Equation (3.38) can also be written as

$$\mathfrak{D} x_\mu = 0, \tag{2.39}$$

where

$$\tfrac{1}{2}\mathfrak{D} \equiv \frac{\partial}{\partial \tau}\left(\frac{\partial \Omega}{\partial(\rho^2 v^2)}\rho^2 \frac{\partial}{\partial \tau}\right) + \frac{\partial}{\partial \xi^A}\left(\frac{\partial \Omega}{\partial(\lambda^2\gamma)}\lambda^2\gamma\gamma^{AB}\frac{\partial}{\partial \xi^B}\right) + \frac{\partial}{\partial \xi^U}\left(\frac{\partial \Omega}{\partial \sigma}\sigma\sigma^{UV}\frac{\partial}{\partial \xi^V}\right). \tag{2.40}$$

The symbol γ^{AB} denotes the elements of the inverse matrix of $\|\sigma_{AB}\|$.

An open membrane will be closed in a physical sense if no momentum crosses the boundaries of the world tube. This condition is achieved by requiring

$$p^1_\mu\,|_{\xi\to\bar{\lambda}^1,\bar{\delta}^1} = 0, \tag{2.41a}$$

$$p^2_\mu\,|_{\eta\to\bar{\lambda}^2,\bar{\delta}^2} = 0, \tag{2.41b}$$

where

$$p^U_\mu \equiv \frac{\partial \Omega}{\partial x^\mu_U}$$

$$= 2\left(\frac{\partial \Omega}{\partial(\rho^2 v^2)}\rho^2 v_\mu \delta^U_0 + \frac{\partial \Omega}{\partial(\lambda^2\gamma)}\lambda^2\gamma\gamma^{AB}x_{B\mu}\delta^U_A + \frac{\partial \Omega}{\partial \sigma}\sigma\sigma^{UV}x_{V\mu}\right). \tag{2.42}$$

For semi-open membranes only one of the conditions (2.41) applies depending on which are the open edges. Conditions (2.41) in the case of geometric objects imply that the open strings' end points as well as the open and semi-open membranes' edges travel with the speed of light,[4,5] for p-strings, p-membranes, s-membranes, and $p-s$-membrances (2.41) do not necessarily imply the same properties for the strings' end points and membranes' edges, e.g., for the Takabayasi model of p-strings one has solutions where the strings' end points travel with a velocity[8] less than c.

The multiplication of Eq. (2.39) by $\dot{x}^\mu$ yields

$$\dot{x}^\mu \mathfrak{D} x_\mu \equiv 0. \tag{2.43}$$

Furthermore, in the case of *geometric objects* we have

$$x'^\mu \mathfrak{D} x_\mu \equiv 0, \tag{2.44a}$$

$$\hat{x}^\mu \mathfrak{D} x_\mu \equiv 0, \tag{2.44b}$$

where we have introduced the notation $(\dot{x}^\mu, x'^\mu, \hat{x}^\mu) \equiv (x^\mu_0, x^\mu_1, x^\mu_2)$. The previous identities are consequences of the invariance of the respective actions under the reparametrizations $\tau \to \tau' = \tau'(\tau, \xi, \eta)$, $\xi \to \xi' = \xi'(\tau, \xi, \eta)$, and $\eta \to \eta' = \eta'(\tau, \xi, \eta)$, respectively. Also, we find in the general case

$$\dot{x}^\mu p^1_\mu \equiv 0, \tag{2.45a}$$

$$\dot{x}^\mu p^2_\mu \equiv 0, \tag{2.45b}$$

and for *geometric objects*

$$\hat{x}^\mu p^1_\mu \equiv 0, \tag{2.46a}$$

$$x'^\mu p^2_\mu \equiv 0. \tag{2.46b}$$

Another important concept associated with the free membrane is the energy—momentum energy tensor denisty defined on the membrane world tube. It can be easily obtained using the expression

$$T^{\mu\nu} = 2\frac{\partial \Omega}{\partial \eta_{\mu\nu}}. \tag{2.47}$$

Thus,

$$T^{\mu\nu} = \frac{\partial \Omega}{\partial(\rho^2 v^2)}\rho^2 \dot{x}^\mu \dot{x}^\nu + \frac{\partial \Omega}{\partial(\lambda^2\gamma)}\lambda^2\gamma\gamma^{AB}x^\mu_A x^\nu_B + \frac{\partial \Omega}{\partial \sigma}\sigma\sigma^{\mu\nu}x^\mu_U x^\nu_V. \tag{2.48}$$

We also have that

$$T^{\mu\nu}\dot{x}_\mu = \Omega \dot{x}_\nu, \tag{2.49a}$$

$$T^{\mu\nu}\dot{x}_\mu \dot{x}_\nu = \Omega v^2 \geq 0, \tag{2.49b}$$

$$p^U_\mu \dot{x}^\mu = \Omega \delta^U_0. \tag{2.50}$$

Thus, from (2.49b) we find that Ω must be nonnegative function of its arguments.

The energy—momentum tensor defined in all the space is

$$T^{\mu\nu}(y^\lambda) = \int d^3\xi\, T^{\mu\nu}(\xi^U)\delta[y^\lambda - x^\lambda(\xi^U)]. \tag{2.51}$$

From the equation of motion we find that

$$\partial_\mu T^{\mu\nu} = 0, \tag{2.52}$$

i.e., the energy and the momentum are conserved. Note that if one postualtes (2.48), Eqs. (2.52) yield the equation of motion (2.39). A discussion of this point for geometric objects can be found in Ref. 11.

We shall close this section by giving two particular cases of free Lagrangian densities

$$\Omega = m_1 \delta(\xi - \bar{\lambda}^1) v + m_2 \delta(\xi - \bar{\delta}^1) v + m\sqrt{-\gamma}, \tag{2.53}$$

$$\Omega = m\delta(\eta - \bar{\lambda}^2)\sqrt{-\gamma} + \mathfrak{M}\sqrt{\sigma}. \tag{2.54}$$

(2.53) describes an open string with two particles of masses m_1 and m_2 attached to its end points, and (2.54) describes an open membrane with a string located on its boundary. In the latter case $\eta = \bar{\lambda}^2$ is taken as the equation of the boundary. Thus, (2.53) is a classical analog to the problem of attaching quarks[9] to the strings' end points, and (2.54) represents the same geometrical problem in one dimension higher.

3. INTERACTING MEMBRANES

In this section we consider a system of N direct interacting membranes, interacting via a two-body type of force. The action for this system is taken as

$$\mathcal{A}_T = \sum_p \int d^3\zeta_p \, \Omega_p + \sum_{p<q} \sum \int\int d^3\zeta_p \, d^3\zeta_q \, \mathfrak{R}_{pq}, \tag{3.1a}$$

where

$$\mathfrak{R}_{pq} = \mathfrak{R}_{pq}(\dot{x}^\mu_p, \Sigma^{\mu\nu}_p, \Sigma^{\mu\nu\lambda}_p, x^\mu_p; \dot{x}^\mu_q, \Sigma^{\mu\nu}_q, \Sigma^{\mu\nu\lambda}_q, x^\mu_q; \bar{\rho}_p, \bar{\rho}_q, \bar{\lambda}_p, \bar{\lambda}_q), \tag{3.1b}$$

$p, q = 1, 2, \ldots, N,$

and $\bar{\rho}$ and $\bar{\lambda}$ are densities that transform like ρ and λ, respectively.

In general we do not require that $\mathfrak{R}_{pq}$ be symmetric under the interchange of membranes. Note that we have restricted the dependence of $\mathfrak{R}_{pq}$ to x^μ and its first derivatives. This has been done in order to end up with second order integro-differential equations. This restriction leaves us with enough generality in the interaction term to describe a large class of interesting interactions. Another restriction on $\mathfrak{R}_{pq}$ is that its dependence on x^μ_p and x^μ_q be such that when the membranes p and q are far apart the interaction be zero This condition can be implemented by requiring that the dependence of $\mathfrak{R}_{pq}$ on x^μ_p and x^μ_q be such that it produces the desired property, e.g., $\mathfrak{R}_{pq} \propto \delta[(x_p - x_q)^2]$ (See Ref. 5).

Now let us perform the variation of each membrane world tube in the action (3.1); we get

$$\begin{aligned}\delta\mathcal{A} = \sum_p \int d^3\zeta_p \frac{\partial \Omega_p}{\partial x^\mu_{pU}} \delta x^\mu_{pU} + \sum_p \int d^3\zeta_p \Bigg[\sum_{p>q} \int d^3\zeta_q \\ \times \left(\frac{\partial \mathfrak{R}_{pq}}{\partial x^\mu_p} \delta x^\mu_p + \frac{\partial \mathfrak{R}_{pq}}{\partial x^\mu_{pU}} \delta x^\mu_{pU} \right) + \sum_{q<q} \int d^3\zeta_q \\ \times \left(\frac{\partial \mathfrak{R}_{pq}}{\partial x^\mu_p} \delta x^\mu_p + \frac{\partial \mathfrak{R}_{qp}}{\partial x^\mu_{pU}} \delta x^\mu_{pU} \right) \Bigg].\end{aligned} \tag{3.2}$$

From (2.37), (3.2), and the variational principle we obtain

$$\frac{\partial}{\partial \xi^U_p} \frac{\partial \Omega_p}{\partial x^\mu_{pU}} = \left(\frac{\partial}{\partial x^\mu_p} - \frac{\partial}{\partial \xi^U_p} \frac{\partial}{\partial x^\mu_{pU}} \right) \Upsilon_p \tag{3.3a}$$

$$\Upsilon_p \equiv \sum_{q>p} \int d^3\zeta_q \mathfrak{R}_{pq} + \sum_{q<p} \int d^3\zeta_q \mathfrak{R}_{qp}. \tag{3.3b}$$

Equation (3.3a) can be written in an equivalent form as follows,

$$\mathfrak{D}_p x_{p\mu} = \left(\frac{\partial}{\partial x^\mu_p} - \frac{\partial}{\partial \xi^U_p} \frac{\partial}{\partial x^\mu_{pU}} \right) \Upsilon_p. \tag{3.4}$$

For the subset of closed membranes the conditions (2.33) also apply in the present case, but for the subset of open or semi-open membranes we must replace (2.41) with

$$\mathfrak{P}^1_{p\mu}\big|_{\zeta_p = \bar{\lambda}^1_p, \bar{\delta}^1_p} = 0, \tag{3.5a}$$

$$\mathfrak{P}^2_{p\mu}\big|_{\eta_p = \bar{\lambda}^2_p, \bar{\delta}^2_p} = 0, \tag{3.5b}$$

where

$$\mathfrak{P}^U_{p\mu} = P^U_{p\mu} + \frac{\partial}{\partial x^\mu_{pU}} \Upsilon_p \tag{3.6}$$

From the identities (2.43) and (2.45), and the equations of motion we obtain the following constraints on the interaction:

$$\frac{\partial}{\partial \tau_p}\left(1 - \dot{x}^\mu_p \frac{\partial}{\partial \dot{x}^\mu_p}\right) - \frac{\partial}{\partial \xi_p}\left(\dot{x}^\mu_p \frac{\partial}{\partial x'^\mu_p}\right) - \frac{\partial}{\partial \eta_p}\left(\dot{x}^\mu_p \frac{\partial}{\partial \hat{x}^\mu_p}\right) \Upsilon_p = 0, \tag{3.7}$$

$$\dot{x}^\mu_p \frac{\partial}{\partial \dot{x}^\mu_p} \Upsilon_p \Big|_{\zeta_p = \bar{\lambda}^1_p, \bar{\delta}^1_p} = 0, \tag{3.8a}$$

$$\dot{x}^\mu \frac{\partial}{\partial \hat{x}^\mu_p} \Upsilon_p \Big|_{\eta_p = \bar{\lambda}^2_p, \bar{\delta}^2_p} = 0. \tag{3.8b}$$

The condition (3.7) can be obtained also by demanding that the action (3.1) be stationary under the change of parametrization $\tau_p \to \tau'_p = \tau_p + \delta\tau_p(\tau_p, \xi_p, \eta_p)$.

For geometric objects we have that (2.44a), (2.44b), (2.45a), and (2.45b) give us the following constraints on the interaction

$$\left[\frac{\partial}{\partial \xi_p}\left(1 - x'^\mu_p \frac{\partial}{\partial x'^\mu_p}\right) - \frac{\partial}{\partial \eta_p}\left(x'^\mu_p \frac{\partial}{\partial \hat{x}^\mu_p}\right) - \frac{\partial}{\partial \tau_p}\left(x'^\mu_p \frac{\partial}{\partial \dot{x}^\mu_p}\right)\right] \Upsilon_p = 0, \tag{3.9a}$$

$$\left[\frac{\partial}{\partial \eta_p}\left(1 - \hat{x}^\mu_p \frac{\partial}{\partial \hat{x}^\mu_p}\right) - \frac{\partial}{\partial \tau_p}\left(\hat{x}^\mu \frac{\partial}{\partial \dot{x}^\mu_p}\right) - \frac{\partial}{\partial \xi_p}\left(\hat{x}^\mu_p \frac{\partial}{\partial x'^\mu_p}\right)\right] \Upsilon_p = 0, \tag{3.9b}$$

$$\hat{x}^\mu_p \frac{\partial}{\partial \hat{x}^\mu_p} \Upsilon_p \Big|_{\zeta_p = \bar{\lambda}^1_p, \bar{\delta}^1_p} = 0, \tag{3.10a}$$

$$\dot{x}^\mu_p \frac{\partial}{\partial \hat{x}^\mu_p} \Upsilon_p \Big|_{\eta_p = \bar{\lambda}^2_p, \bar{\delta}^2_p} = 0. \tag{3.10b}$$

The constraints (3.9a) and (3.9b) can also be found by demanding that the action be stationary under $\xi \to \xi' = \xi + \delta\xi(\tau, \xi, \eta)$ and $\eta \to \eta' = \eta + \delta\eta(\tau, \xi, \eta)$, respectively.

The constraints (3.8) and (3.10) are automatically satisfied as a consequence of the identities:

$$\dot{x}^\alpha \frac{\partial}{\partial \hat{x}^\alpha} \sum^{\mu\nu} = 0, \tag{3.11a}$$

$$\dot{x}^\alpha \frac{\partial}{\partial \hat{x}^\alpha}\sum^{\mu\nu\lambda}=0, \tag{3.11b}$$

$$\hat{x}^\alpha \frac{\partial}{\partial \hat{x}^\alpha}\sum^{\mu\nu\lambda}=0, \tag{3.11c}$$

$$\hat{x}^\alpha \frac{\partial}{\partial \hat{x}^\alpha}\sum^{\mu\nu\lambda}=0. \tag{3.11d}$$

To solve the constraint (3.7) we realize that this constraint comes from the fact that the free action is invariant under (2.25), i.e., that Ω transforms like a scalar density of weight 1 under this transformation. So the constraints will be solved if $\mathfrak{R}_{pq}$ transforms like a product of scalar densities, i.e., like $\Omega_p\Omega_q$ under (2.25). First we construct out of the argument of $\mathfrak{R}_{pq}$ all the independent quantities that transform as a scalar density under (2.25); they are $\bar{\rho}\dot{x}^\mu$, $\bar{\lambda}\sum^{\mu\nu}$, and $\bar{e}\sum^{\mu\nu\lambda}$, where $\bar{e}$ is a constant parameter. Now the general solution to the problem is a function

$$\mathfrak{R}_{pq}=\mathfrak{R}_{pq}(\bar{\rho}_p\dot{x}^\mu_p,\bar{\lambda}_p\textstyle\sum^{\mu\nu}_p,\bar{e}_p\sum^{\mu\nu\lambda}_p;\bar{\rho}_q\dot{x}^\mu_q,\bar{\lambda}_q\sum^{\mu\nu}_q,\bar{e}_q\sum^{\mu\nu\lambda}_q;$$
$$x^\mu_p,x^\mu_q), \tag{3.12}$$

homogeneous of the first degree in the variables $\bar{\rho}_p\dot{x}^\mu_p$, $\bar{\lambda}_p\sum^{\mu\nu}_p$, $\bar{e}_p\sum^{\mu\nu\lambda}_p$; homogeneous of the first degree in the variables $\bar{\rho}_q\dot{x}^\mu_q$, $\bar{\lambda}_q\sum^{\mu\nu}_q$, $\bar{e}_q\sum^{\mu\nu\lambda}_q$; and arbitrary in the variables x^μ_p, x^μ_q. Now it is easy to check that the constraint (3.7) is solved.

When $\bar{\rho}=\delta(\xi)\delta(\eta)$ and $\bar{\lambda}=\bar{e}=0$, Eq. (3.12) describes the interaction between particles, then

$$\mathfrak{R}^{(p)}_{pq}=\delta(\xi_p)\delta(\eta_p)\delta(\xi_q)\delta(\eta_q)\mathfrak{R}_{pq}(\dot{x}^\mu_p,\dot{x}^\mu_q,x^\mu_p,x^\mu_q)$$
$$=\delta(\xi_p)\delta(\eta_p)\delta(\xi_p)\delta(\eta_q)v_pv_q$$
$$\times R(\dot{x}_p/v_p,\dot{x}_q/v_q,x^\mu_p,x^\mu_q). \tag{3.13}$$

Note that the arguments $\dot{x}^\mu_p/v_p$ and $\dot{x}^\mu_q/v_q$ are homogeneous functions of zeroth degree in $\dot{x}^\mu_p$ and $\dot{x}^\mu_q$, respectively, thus $R_{pq}\equiv{}^{(p)}_{pq}(\dot{x}^\mu_p/v_p,\dot{x}^\mu_q/v_q,x^\mu_p,x^\mu_q)$ can be taken as an arbitrary function of all its variables.

When $\bar{\rho}=\bar{e}=0$ and $\bar{\lambda}=\delta(\eta)$, and $\bar{\rho}=\bar{\lambda}=0$ and $\bar{e}=1$, (3.12) reproduces the interaction between geometric strings and between geometric membranes, respectively. In these cases the constraints (3.9) are automatically satisfied.[4,5]

An important particular case of interaction is obtained when the N objects are described as follows:

$$\rho_i=m_i\delta(\xi_i)\delta(\eta_i)=m_i\bar{\rho}_i, \tag{3.14a}$$

$$\lambda_i=\bar{\lambda}_i=e_i=\bar{e}_i=0, \tag{3.14b}$$

$$x^\mu_i=z^\mu_i,\quad i,j=1,\ldots,P; \tag{3.14c}$$

$$\lambda_a=M_a\delta(\eta_a)=M_a\bar{\lambda}_a, \tag{3.15a}$$

$$\rho_a=\bar{\rho}_a=e_a=\bar{e}_a=0, \tag{3.15b}$$

$$x^\mu_a=x^\mu_a,\quad a,b=P+1,\ldots,P+S; \tag{3.15c}$$

$$e_r=\mathfrak{M}_r=\mathfrak{M}_r\bar{e}_r, \tag{3.16a}$$

$$\bar{\lambda}_r=\lambda_r=\bar{\rho}_r=\rho_r=0, \tag{3.16b}$$

$$x^\mu_r=y^\mu_r,\quad r,s=P+S+1,\ldots,P+S+M=N. \tag{3.16c}$$

In this case the action (3.1) reduces to

$$\mathcal{A}_T=\sum_i\int m_ir_id\tau_i+\sum_a\int d^2\tau_aM_a\,(-\gamma_a)^{1/2}$$
$$+\sum_r\int d^3\zeta_r\,\mathfrak{M}_r(\sigma_r)^{1/2}$$
$$+\sum_i\sum_{<J}\int d\tau_id\tau_Jv_iv_JR_{iJ}$$
$$+\sum_a\sum_{<b}\int d^2\tau_ad^2\tau_b(-\gamma_a)^{1/2}(-\gamma_b)^{1/2}R_{ab}$$
$$+\sum_r\sum_{<s}\int d^3\zeta_rd^3\zeta_s(\sigma_r)^{1/2}(\sigma_s)^{1/2}R_{rs}$$
$$+\sum_i\sum_a\int d\tau_id^2\tau_av_i(-\gamma_a)^{1/2}R_{ia}$$
$$+\sum_i\sum_r\int d\tau_id^3\zeta_rv_i(\sigma_i)^{1/2}R_{ir}$$
$$+\sum_a\sum_r\int d^2\tau_ad^3\zeta_r\,(-\gamma_a)^{1/2}(\sigma_i)^{1/2}R_{ar}, \tag{3.17}$$

where we have introduced the notation $d^2\tau=d\tau d\xi$. Equation (3.17) describes the action for a system of P particles, S geometric strings, and M geometric membranes interacting via a two-body type of direct interaction between objects of the same class as well as objects belonging to different classes.

For the system under consideration, the equations of motion (3.4) and (3.5) give us

$$m_i\frac{d}{d\tau_i}\frac{v^\mu_i}{v_i}=\mathcal{L}^\mu_i(V_i+U_i+W_i), \tag{3.18}$$

$$M_a\Box_ax^\mu_a=\mathcal{L}^\mu_a(\phi_a+\chi_a+H_a), \tag{3.19a}$$

$$\mathfrak{M}_r\Box_ry^\mu_r=\mathcal{L}^\mu_r(\phi_r+\Theta_r+\Xi_r), \tag{3.20a}$$

$$[M_a(-\gamma_a)^{1/2}x^1_{a\mu}+\beta^1_{a\mu}(\phi_a+\chi_a+H_a)]_{\xi_a\to\bar{\lambda}^1_a,\bar{\delta}^1_a}=0, \tag{3.19b}$$

$$[\mathfrak{M}_r(\sigma_r)^{1/2}y^1_{r\mu}+\beta^1_{r\mu}(\phi_r+\Theta_r+\Xi_r)]_{\xi_r\to\bar{\lambda}^1_r,\bar{\delta}^1_r}=0, \tag{3.20b}$$

$$[\mathfrak{M}_r(\sigma_r)^{1/2}y^2_{r\mu}+\beta^2_{r\mu}(\phi_r+\Theta_r+\Xi_r)]_{\eta_r\to\bar{\lambda}^2_r,\bar{\delta}^2_r}=0, \tag{3.20c}$$

where the definition of the symbols used in (3.17)–(3.20) can be found in Ref. 7.

4. PARTICULAR CASES OF INTERACTIONS

The gravitational interaction between p-s-membranes in the weak field approximation is given by

$$\mathfrak{R}_{pq}=\mu_p\mu_qT^{\mu\nu}_pT_{q\mu\nu}\delta[(x_p-x_q)^2], \tag{4.1}$$

where μ_p and μ_q are constant and $T^{\mu\nu}$ is given by (2.48). $T^{\mu\nu}$ is a homogeneous function of degree 1 in the variables $\rho^{\nu\mu}$, $\lambda\sum^{\mu\nu}$, and $e\sum^{\mu\nu\lambda}$. This can be seen from (2.48), (2.5), and the identities

$$\textstyle\sum^{\mu\nu}\sum^{\lambda}_{\nu}=-\gamma\gamma^{AB}x^\mu_Ax^\lambda_B, \tag{4.2}$$

$$\sum^{\mu\alpha\beta}\sum^{\nu}_{\alpha\beta} = 2\sigma\sigma^{UV}x^{\mu}_{U}x^{\nu}_{V}. \tag{4.3}$$

The study of this interaction for the case of geometric membranes can be found in Ref. 5. The gravitational field associated with (4.1), in the case of geometric strings and Takabayashi's realistic strings, has been studied by Stachel and the author.[12]

The "electromagnetic" interaction between p-s-membranes is given by

$$\mathfrak{R}_{pq} = \bar{\rho}_p\bar{\rho}_q\dot{x}^{\mu}_{p}\dot{x}_{q\mu}\delta[(x_p - x_q)^2]. \tag{4.4}$$

In the case of particles $[\bar{\rho} = Q\delta(\xi)\delta(\eta)]$, this interaction reduces to the Fokker principle of electrodynamics.[13] For geometric strings and geometric membranes (4.4) is not allowed ($\bar{\rho} = 0$).

From (3.4), (3.6), (3.3), and (4.4) we find

$$\mathfrak{P}^1_{\mu} = P^1_{\mu}, \tag{4.5a}$$

$$\mathfrak{P}^2_{\mu} = P^2_{\mu}, \tag{4.5b}$$

$$\mathfrak{D}x_{p\mu} = \bar{\rho}_p\dot{x}^{\nu}_{p}\sum_{q\neq p}F_{q\mu\nu}, \tag{4.5c}$$

where

$$F_{\mu\nu} \equiv \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}, \tag{4.6}$$

$$A_{\mu}(y) \equiv \int d^3\zeta\,\bar{\rho}\dot{x}_{\mu}\delta\{[y - x(\xi^{\nu})]^2\}. \tag{4.7}$$

Equations (4.5a) and (4.5b) tell us that the open edges of open or semi-open membranes as well as the end points of open strings do not interact with the "field" $F^{\mu\nu}$.

Another interesting interaction between p-s-membranes is the one given by

$$\mathfrak{R}_{pq} = \bar{\lambda}_p\bar{\lambda}_q\sum^{\mu\nu}_{p}\sum_{q\mu\nu}\delta[(x_p - x_q)^2]. \tag{4.8}$$

This interaction is not allowed for particles and geometric membranes ($\bar{\lambda} = 0$). Let us consider only two closed membranes interacting via (4.8), then the equation of motion (3.4) reduces to

$$\mathfrak{D}_p x_{p\lambda} = \bar{\lambda}_p\sum^{\mu\nu}_{p}F_{q\lambda\mu\nu}, \quad p,q = 1,2, \tag{4.9}$$

where

$$F_{\lambda\mu\nu} \equiv \partial_{\lambda}\psi_{\mu\nu} + \partial_{\mu}\psi_{\nu\lambda} + \partial_{\nu}\psi_{\lambda\mu}, \tag{4.10}$$

$$\psi_{\mu\nu}(y) \equiv \int d^3\zeta\,\lambda\sum_{\mu\nu}(\zeta^U)\delta\{[y - x(\zeta^U)]^2\}. \tag{4.11}$$

Note that $F_{\lambda\mu\nu}$ is a totally antisymmetric tensor. From (2.33), (4.10), and (4.11) we find

$$\partial_{\lambda}F^{\lambda\mu\nu} = -4\pi J^{\mu\nu}, \tag{4.12}$$

where

$$J^{\mu\nu}(y^{\rho}) \equiv \int d^3\zeta\,\lambda\sum^{\mu\nu}\delta[y^{\rho} - x^{\rho}(\zeta^{\nu})]. \tag{4.13}$$

Equation (4.10) can also be written in terms of differential forms as[14]

$$\mathbf{F} = d\psi, \quad \psi \equiv \frac{1}{3!}\psi_{\mu\nu}dx^{\mu}\wedge dx^{\nu}. \tag{4.14}$$

The Poincaré lemma tells us that

$$d\mathbf{F} = 0. \tag{4.15}$$

Equation (4.15) written in components is

$$\partial_{\sigma}F_{\lambda\mu\nu} + \partial_{\lambda}F_{\mu\nu\sigma} + \partial_{\mu}F_{\nu\sigma\lambda} + \partial_{\nu}F_{\sigma\lambda\mu} = 0. \tag{4.16}$$

At this point we can recast the action-at-a-distance formalism in terms of the field $F_{\mu\nu\lambda}$ that obeys the field equation (4.12) and (4.16). In particular we can study the free field ($J^{\mu\nu} = 0$). For a discussion of this point see Refs. 4 and 5.

Let us define the dual of $F_{\mu\nu\lambda}$, i.e., F_{ρ} given by

$$\epsilon^{\mu\nu\lambda\rho}F_{\rho} = F^{\mu\nu\lambda}. \tag{4.17}$$

From (4.12), (4.17), and the condition $J^{\mu\nu} = 0$ we get

$$\partial_{\mu}F_{\nu} - \partial_{\nu}F_{\mu} = 0, \tag{4.18}$$

thus

$$F_{\nu} = \partial_{\nu}\phi. \tag{4.19}$$

(4.17), (4.19), and (4.16) gives us

$$\Box\phi = 0.$$

So the free field theory associated with the interaction (4.8) is a massless scalar field. This interaction was first studied by Kalb and Ramond[3] in the context of the theory of direct interactions between geometric strings. Recently Lund and Regge[15] found that (4.16) and (4.12) can be used to study vortices in a superfluid. The "generalization" of (4.8) for the case of geometric membranes can be found in Ref. 2.

Many other interactions that can be "factorized" like (4.1), (4.4), and (4.8) could be written. The study of field theories associated with a given "factorizable" interaction between p-s-membranes is a trivial extension of the one that we carried out for geometric strings and geometric membranes.[4,5]

5. CONSERVATION THEOREMS

In this section the conservation theorems—energy momentum conservation and center-of-mass theorem and angular momentum conservation—are derived as a consequence of the Poincaré invariance of the action $\mathcal{A}_T$.

Let us define the reduced action related to (3.1) by "reducing" the range of the time evolution parameter τ and keeping the same ranges of the spatial parameters as before, i.e.,

$$\mathcal{A}^{**} \equiv \mathcal{A}_T[\tau^*_p \leq \tau_p \leq \tau^{**}_p; \tau^*_q \leq \tau_q \leq \tau^{**}_q], \tag{5.1}$$

where τ^*_p, τ^{**}_p, τ^*_q, and τ^{**}_q are arbitrary constants restricted only by the requirement that τ^{**}_p and τ^{**}_q refer to a later time than τ^*_p and τ^*_q, respectively.

The variation of $\mathcal{A}^{**}$ gives us

$$\delta\mathcal{A}^{**} = \sum_{p}\int^{**}_{*} d^3\zeta_p \times\left[\frac{\partial\Omega_p}{\partial x^{\mu}_{pU}}\delta x^{\mu}_{pU} + \frac{\partial}{\partial x^{\mu}_{p}}\Upsilon^*_p\delta x^{\mu}_{p} + \frac{\partial}{\partial x^{\mu}_{pU}}\Upsilon^*_p\delta x^{\mu}_{pU}\right], \tag{5.2}$$

where Υ^x_p is defined by changing the range of τ_q in (3.3b) as follows,

$$\Upsilon^*_p \equiv \Upsilon_p[\tau^*_q \leq \tau_q \leq \tau^{**}_q]. \tag{5.3}$$

We also define the variation $\delta\mathcal{A}^*$ by letting in (5.1) $\tau_q^* \to -\infty$, $\tau_q^{**} \to +\infty$, i.e.,

$$\delta\mathcal{A}^* \equiv \delta\mathcal{A}^{**}[\tau_q^* \to -\infty, \tau_q^{**} \to +\infty]. \tag{5.4}$$

From (5.4), (5.3), and (5.2) we get

$$\delta\mathcal{A}^* = \sum_p \int_*^{**} d^3\zeta_p \left[\frac{\partial}{\partial x^\mu_{pU}} \delta x^\mu_{pU} + \frac{\partial}{\partial x^\mu_p} \Upsilon_p \delta x^\mu_p + \frac{\partial}{\partial x^\mu_{pU}} \Upsilon_p \delta x^\mu_{pU}\right]. \tag{5.5}$$

Equations (3.3a) and (5.5) yield

$$\delta\mathcal{A}^* = \sum_p \int_*^{**} d^3\zeta_p \frac{\partial}{\partial \zeta_p^U} \left\{\left[\frac{\partial}{\partial x^\mu_{pU}} (\Omega_p + \Upsilon_p)\right] \delta x^\mu_p\right\}. \tag{5.6a}$$

Stoke's theorem and (5.6a) give us

$$\delta\mathcal{A}^* = \tfrac{1}{2}\sum_p \int_{\partial D_p^*} d\xi_p^U \wedge d\xi_p^V \epsilon_{UVW} \left[\frac{\partial}{\partial x^\mu_{pw}} (\Omega_p + \Upsilon_p)\right] \delta x^\mu_p, \tag{5.6b}$$

where ∂D^* denotes the boundary of the membrane world tube limited by the surfaces $\tau = \tau^*$ and $\tau = \tau^{**}$.

From (2.33), (3.5), and (5.6b) we have that

$$\delta\mathcal{A}^* = \sum_p \int_{C_p} d^2\xi_p \delta x^\mu_p \frac{\partial}{\partial \dot{x}^\mu_p} (\Omega_p + \Upsilon_p)\Big|_{\tau_p = \tau_p^{**}} - \sum_p \int_{C_p} d^2\xi_p \delta x^\mu_p \frac{\partial}{\partial \dot{x}^\mu_p} (\Omega_p + \Upsilon_p)\Big|_{\tau_p = \tau_p^{*}}, \tag{5.7}$$

where we have introduced the notation $d^2\xi_p \equiv d\xi_p d\eta_p$. For membranes, C_p denotes the portion of the surface $\tau_p = \text{constant}$, limited by the corresponding membrane world-tube. For strings, C_p denotes the Cartesian product of the line defined by the intersection of $\tau_p = \text{const}$ with the corresponding string world sheet by $(-\infty < \eta_p < +\infty)$. For particles, C_p is $(-\infty < \xi_p < +\infty) \times (-\infty < \eta_p < +\infty)$.

Equation (5.7) is gauge invariant because (5.5), (2.33) (2.33), and (3.5) are gauge invariant expressions, even for the most general gauge, i.e., the gauge associated with geometric membranes. For a discussion of this point see Ref. 7.

A. Energy-momentum conservation

Let us impose on $\mathcal{A}^{**}$ the condition of being invariant under arbitrary infinitesimal translations in time and space, i.e., under the transformation (2.36) with

$$\delta x^\mu_p \equiv \epsilon^\mu. \tag{5.8}$$

Note that invariance of $\mathcal{A}^{**}$ under any transformation implies the invariance of $\mathcal{A}^*$, $\mathcal{A}$ and $\mathfrak{R}_{pq}$ under the same transformation, due to the arbitrariness of τ_p^*, τ_q^*, etc.

Equations (5.7) and (5.8) suggest that we define the total four-momentum for the system of interacting membranes as the limit

$$P_\mu \equiv \sum_p \int_{C_p} d^2\xi_p \frac{\partial}{\partial \dot{x}^\mu_p} (\Omega_p + \Upsilon_p)\Big|_{\tau_p \to \infty}. \tag{5.9}$$

Note that the definition (5.9) is gauge invariant, because we can always cast (5.9) in a similar form to (5.6) adding to the RHS of (5.9) integrals that have the value zero. For a discussion of this point see Ref. 7.

From (5.9), the equations of motion (3.3a), Stoke's theorem, (3.5), and (2.33), we get

$$P_\mu = \sum_p \int_{C_p} d^2\xi_p \frac{\partial}{\partial \dot{x}^\mu} (\Omega_p + \Upsilon_p) + \sum_p \int_{\tau_p}^{\infty} d\tau_p \int_{C_p} d^2\xi_p \frac{\partial}{\partial x^\mu_p} \Upsilon_p. \tag{5.10}$$

From (5.10), (3.4), and the invariance of $\mathfrak{R}_{pq}$ under (2.36) and (5.8) we obtain

$$P_\mu = \sum_p \int_{C_p} d^2\xi_p \frac{\partial}{\partial \dot{x}^\mu} (\Omega_p + \Upsilon_p) + \sum_{p<q}\sum \left[\int_{\tau_p}^{\infty} \int_{-\infty}^{\tau_q} - \int_{-\infty}^{\tau_p} \int_{\tau_q}^{\infty}\right] d\tau_p d\tau_q \times \int_{C_p} \int_{C_q} d^2\xi_p d^2\xi_q \frac{\partial \mathfrak{R}_{pq}}{\partial x^\mu_{pq}}, \tag{5.11}$$

where

$$x^\mu_{pq} = x^\mu_p - x^\mu_q. \tag{5.12}$$

Note that

$$P_\mu = P_\mu(\tau_1, \tau_2, \ldots, \tau_N), \tag{5.13}$$

$$\frac{\partial P_\mu}{\partial \tau_p} = 0, \quad 1 \le p \le N. \tag{5.14}$$

The conservation laws (5.14) are a trivial consequence of definition (5.9). Also, (5.14) can be proved directly by making use of the equations of motion (3.3a) and (3.5) and Eqs. (2.33).

When we specialize Ω_p and $\mathfrak{R}_{pq}$ by imposing the conditions (3.14)–(3.16), Eq. (5.11) yields

$$P^\mu_{(psm)} = P^\mu_{(p)} + P^\mu_{(s)} + P^\mu_{(m)} + P^\mu_{(ps)} + P^\mu_{(pm)} + P^\mu_{(sm)}. \tag{5.15}$$

Expression (5.15) represents the total 4-momentum for a system of P particles, S geometric strings, and M geometric membranes in direct interaction. The explicit form of the terms in the RHS of (5.15) can be found in Ref. 7.

B. Center-of-mass-theorem and angular momentum conservation

Let us impose invariance of $\mathcal{A}^{**}$ under arbitrary infinitesimal rotations in Minkowski space–time, i.e., invariance under (2.36) with

$$\delta x_{p\mu} = \epsilon_{\mu\nu} x^\nu_p; \quad \epsilon_{\mu\nu} = -\epsilon_{\nu\mu}. \tag{5.16}$$

The invariance of $\mathcal{A}^{**}$ under infinitesimal rotation and translations in Minkowski space implies

$$x^{[\mu}_{pq} \eta^{\nu]\alpha} \frac{\partial}{\partial x^\alpha_{pq}} \mathfrak{R}_{pq} + x^{[\mu}_{pU} \eta^{\nu]\alpha} \frac{\partial}{\partial x^\alpha_{pU}} \mathfrak{R}_{pq} + x^{[\mu}_{qU} \eta^{\nu]\alpha} \frac{\partial}{\partial x^\alpha_{qU}} \mathfrak{R}_{pq} = 0, \tag{5.17}$$

where

$$a^{[\mu} b^{\nu]} \equiv a^\mu b^\nu - a^\nu b^\mu. \tag{5.18}$$

Equation (5.17) will be satisfied when the dependence of $\mathfrak{R}_{pq}$ is through Poincare invariant combinations of its

arguments. These Poincare invariant arguments are known as the fundamental invariants of a given interaction.[16] The invariants that can be formed with the arguments of $\mathfrak{R}_{pq}$ are essentially the same as the fundamental invariants for two-body types of interactions between particles, geometric strings, and geometric membranes that we discuss in Refs. 4 and 5.

Equations (5.7) and (5.16) suggest that we define the total angular momentum bivector for the system of interacting membranes as the limit

$$L^{\mu\nu} \equiv \sum_p \int_{C_p} d^2\xi_p x_p^{[\mu}\eta^{\nu]\alpha} \frac{\partial}{\partial \dot{x}_p^\alpha}(\Omega_p + \Upsilon_p)\Big|_{\tau_p \to -\infty}. \tag{5.19}$$

The equations of motion (3.3a) give us the identity

$$\frac{\partial}{\partial \xi_p^U}\left[x_p^{[\mu}\eta^{\nu]\alpha}\frac{\partial}{\partial x_{pU}^\alpha}(\Omega_p + \Upsilon_p)\right.$$
$$\left. = x_p^{[\mu}\eta^{\nu]\alpha}\frac{\partial}{\partial x_p^\alpha}\Upsilon_p + x_{pU}^{[\mu}\eta^{\nu]\alpha}\frac{\partial}{\partial x_{pU}^\alpha}\Upsilon_p. \right. \tag{5.20}$$

To prove (5.20) we have made use of the identity

$$x_{pU}^{[\mu}\eta^{\nu]\alpha}\frac{\partial}{\partial x_{pU}^\alpha}\Omega_p \equiv 0; \tag{5.21}$$

That is a consequence of the manifest Poincaré invariance of the free action.

From (5.17)—(5.20), (3.5), (2.33), and Stoke's theorem we find

$$L^{\mu\nu} = \sum_p \int d^2\xi_p x_p^{[\mu}\eta^{\nu]\alpha}\frac{\partial}{\partial \dot{x}_p^\alpha}(\Omega_p + \Upsilon_p)$$
$$+ \sum_{p<q}\sum\left[\int_{\tau_p}^{\infty}\int_{-\infty}^{\tau_q} - \int_{-\infty}^{\tau_p}\int_{\tau_q}^{\infty}\right]d\tau_p d\tau_q \int_{C_p}\int_{C_q} d^2\xi_p d^2\xi_q$$
$$\times\left[x_p^{[\mu}\eta^{\nu]\alpha}\frac{\partial}{\partial x_{pq}^\alpha}\mathfrak{R}_{pq} + x_{pU}^{[\mu}\eta^{\nu]\alpha}\frac{\partial}{\partial x_{pU}^\alpha}\mathfrak{R}_{pq}\right. . \tag{5.22}$$

Note that

$$L^{\mu\nu} = L^{\mu\nu}(\tau_1, \tau_2, \ldots, \tau_N), \tag{5.23}$$

$$\frac{\partial L^{\mu\nu}}{\partial \tau_p} = 0, \quad 1 \le p \le N. \tag{5.24}$$

Equation (5.24) can be proved in a similar way to (5.14). The conservation of L^{0i} is known in the center-of-mass-theorem and the conservation of L^{ij} as angular momen momentum conservation.

When we specialize Ω_p and $\mathfrak{R}_{pq}$ by imposing the restrictions (3.14)—(3.16), Eqs. (5.22) can be written as

$$L^{\mu\nu}_{(psm)} = L^{\mu\nu}_{(p)} + L^{\mu\nu}_{(s)} + L^{\mu\nu}_{(m)} + L^{\mu\nu}_{(ps)} + L^{\mu\nu}_{(pm)} + L^{\mu\nu}_{(sm)}. \tag{5.25}$$

Expression (5.25) represents the total conserved angular momentum bivector for the system under consideration. The explicit form of the terms that appear in the RHS of (5.25) can be found in Ref. 7.

6. ACTION AND REACTION

Let us consider two objects, p and q, of the same type and topology, e.g., two closed s-membrane, two semi-open "realistic" p-membranes, interacting via a symmetric interaction, i.e.,

$$\mathfrak{R}_{pq} = \mathfrak{R}_{qp}, \quad p, q = 1, 2. \tag{6.1}$$

In this case the equation of motion (3.3) can be written as

$$\mathfrak{D}_p x_p^\mu = \int d^3\zeta_q \left[\frac{\partial}{\partial x_{pq}^\mu} - \frac{\partial}{\partial \zeta_p^U}\frac{\partial}{\partial x_{pU}^\mu}\right]\mathfrak{R}_{pq}. \tag{6.2}$$

In writing (6.2) we have used the fact that Poincaré invariance of the action implies translation invariance of $\mathfrak{R}_{pq}$.

Stoke's theorem gives us

$$\int d^3\zeta_q \frac{\partial}{\partial \zeta_q^U}\frac{\partial}{\partial x_q^\mu}\mathfrak{R}_{pq} = \tfrac{1}{2}\int d\zeta_q^U \wedge d\zeta_q^V \epsilon_{UVW}\frac{\partial}{\partial x_{qW}^\mu}\mathfrak{R}_{pq}. \tag{6.3}$$

From (6.2) and (6.3) we find

$$\mathfrak{D}_p x_{p\mu} = \int d^3\zeta_q \mathcal{J}_{\mu pq}$$
$$- \tfrac{1}{2}\int d\zeta_q^U \wedge d\zeta_q^V \epsilon_{UVW}\frac{\partial \mathfrak{D}_{pq}}{\partial x_{qw}^\mu}, \tag{6.4}$$

where

$$\mathcal{J}_{\mu pq} \equiv \frac{\partial \mathfrak{R}_{pq}}{\partial x_{pq}^\mu} - \frac{\partial}{\partial \zeta_p^U}\frac{\partial}{\partial x_{pU}^\mu}\mathfrak{R}_{pq} + \frac{\partial}{\partial \zeta_q^U}\frac{\partial \mathfrak{R}_{pq}}{\partial x_{qU}^\mu}. \tag{6.5}$$

Note that under a gauge transformation $\mathcal{J}_{pq}^\mu$ transforms like $\Omega_p\Omega_q$, and

$$\mathcal{J}_{pq}^\mu = -\mathcal{J}_{qp}^\mu. \tag{6.6}$$

Now we shall further restrict the interation by imposing the condition

$$\frac{\partial \mathfrak{R}_{pq}}{\partial x_p^\mu}\Big|_{\tau_p \to \pm\infty} = 0, \quad p = 1, 2. \tag{6.7}$$

For particles, this condition is enough to assure that the second term in the RHS of Eq. (6.4) vanishes. For closed objects also, (6.7) makes zero the surface integral in (6.4). But for open or semi-open objects we must add extra conditions to obtain the same result, e.g.,

$$\frac{\partial}{\partial \dot{x}_p^\mu}\mathfrak{R}_{pq}\Big|_{\xi_p - \bar{x}_p^1, \bar{\delta}_p^1} = 0, \quad p = 1, 2, \tag{6.8a}$$

$$\frac{\partial}{\partial \dot{x}_p^\mu}\mathfrak{R}_{pq}\Big|_{\eta_p - \bar{x}_p^2, \bar{\delta}_p^2} = 0, \quad p = 1, 2. \tag{6.8b}$$

These conditions apply only to the open boundaries of open or semi-open objects. Equations (3.5) tell us that the conditions (6.8) kill the interaction at the boundaries, i.e., the boundaries move as in the free case.

The conditions (6.8) are gauge covariant, even under the most general transformation given by (2.31). Equations (6.7) are covariant only under the gauge transformation (2.25). If the two objects in consideration have more gauge freedom than (2.25) we replace (6.7) by

$$\frac{\partial \mathfrak{R}_{pq}}{\partial x_{pU}^\mu}\Big|_{\tau_p \to \pm\infty} = 0, \quad p = 1, 2. \tag{6.9}$$

These conditions kill the surface integral in (6.4) as well as being covariant even under (2.31).

If the interaction is such that it satisfies (6.1), and (6.7)—(6.9), depending on the type and topology of the

object as stated before, we have that (6.4) reduces to

$$\mathfrak{D}_p x^\mu_p = \int d^3\tau_q R^\mu_{pq}. \tag{6.10}$$

The fact that under the previous assumption (6.2) can be cast as (6.10) is known as the "action-and-reaction law." See Ref. 7. Now we can interpret the function R^μ_{pq} as a function that describes a stream of impulses coming from the antisymmetry of R^μ_{pq} in p and q tells us that the impulses emanating from p and q respectively are equal in magnitude with opposite signs.

There is an evident asymmetry between closed and open or semi-open objects, because of (6.8). In general, one of the most important properties of open objects[1] is that they can interact through their open edges or end points; Eq. (6.8) does not allow such types of interactions. In the case of geometric objects, they are too strong due to the fact that the gauge freedom of these objects imposes strong conditions on the dependence of $\mathfrak{R}_{pq}$ on its arguments, as discussed in Sec. 4. But for other types of objects it is possible to satisfy all the above mentioned conditions, e.g., the interaction given by (4.4) satisfies all of them, even for open objects. Other examples of particular interactions that can be written like (6.10) can be found in Ref. 7.

From (6.10), (6.6), (6.8), (3.5), and (2.33) we get

$$\sum_{p=1}^{2}\int_{C_p} d^2\xi_p \frac{\partial\Omega_p}{\partial\dot{x}^\mu_p}\bigg|_{\tau_p=-\infty} = \sum_{p=1}^{2}\int d^2\xi_p \frac{\partial\Omega_p}{\partial\dot{x}^\mu_p}\bigg|_{\tau_p=+\infty}, \tag{6.11}$$

Equation (6.11) can also be derived from (5.11) under the same assumptions. Equation (6.11) tells us that (6.7)–(6.9) can also be regarded as the conditions for two objects being asymptotically free.

The evolution of the angular-momentum bivector can be cast in a similar way to (6.10), as follows: From (5.20) and (6.1) we obtain

$$\begin{aligned}\frac{\partial}{\partial\xi^U_p}\left[x^{[\mu}_p\eta^{\nu]\alpha}\frac{\partial}{\partial x^\alpha_{pU}}\Omega_p\right] &= \int d^3\xi_q\left[x^{[\mu}_p\eta^{\nu]\alpha}\frac{\partial}{\partial x^\alpha_{pq}}\mathfrak{R}_{pq}\right.\\ &\left. + x^{[\mu}_{pU}\eta^{\nu]\alpha}\frac{\partial}{\partial x^\alpha_{pU}}\mathfrak{R}_{pq} - \frac{\partial}{\partial\xi^U_p}\,x^{[\mu}_p\eta^{\nu]\alpha}\frac{\partial}{\partial x^\alpha_{pU}}\mathfrak{R}_{pq}\right].\end{aligned} \tag{6.12}$$

Equations (6.11) and (5.17) give us

$$\begin{aligned}\frac{\partial}{\partial\xi^U_p}\left[x^{[\mu}_p\eta^{\nu]\alpha}\frac{\partial}{\partial x^\alpha_{pU}}\Omega_p\right] &= \int d^3\tau_q \mathcal{J}^{\mu\nu}_{pq} - \tfrac{1}{2}\int d\xi^U_q\wedge d\xi^V_q\\ &\times\epsilon_{UVW}x^{[\mu}_q\eta^{\nu]\alpha}\frac{\partial}{\partial x^\alpha_q}\mathfrak{R}_{pq},\end{aligned} \tag{6.13}$$

where

$$\begin{aligned}\mathcal{J}^{\mu\nu}_{pq} \equiv{}& \tfrac{1}{2}(x^{[\mu}_p\eta^{\nu]\alpha} + x^{[\mu}_q\eta^{\nu]\alpha})\frac{\partial\mathfrak{R}_{pq}}{\partial x^\alpha_{pq}} + \tfrac{1}{2}x^{[\mu}_{pU}\eta^{\nu]\alpha}\frac{\partial}{\partial x^\alpha_{pU}}\mathfrak{R}_{pq}\\ &- \tfrac{1}{2}x^{[\mu}_q\eta^{\nu]\alpha}\frac{\partial}{\partial x^\alpha_{qU}}\mathfrak{R}_{pq} - \frac{\partial}{\partial\xi^U_p}\left(x^{[\mu}_p\eta^{\nu]\alpha}\frac{\partial}{\partial x^\alpha_{pU}}\mathfrak{R}_{pq}\right.\\ &\left. + \frac{\partial}{\partial\xi^U_q}\,x^{[\mu}_q\eta^{\nu]\alpha}\frac{\partial}{\partial x^\alpha_{rU}}\mathfrak{R}_{pq}\right).\end{aligned} \tag{6.14}$$

The function $R^{\mu\nu}_{pq}$ has the following properties,

$$R^{\mu\nu}_{pq} = -R^{\mu\nu}_{qp} = -R^{\nu\mu}_{pq}. \tag{6.15}$$

The surface integral in (6.13) is zero for particles and closed strings if

$$x^{[\mu}_p\eta^{\nu]\alpha}\frac{\partial\mathfrak{R}_{pq}}{\partial\dot{x}^\alpha_p}\bigg|_{\tau_p=\pm\infty} = 0, \quad p=1,2. \tag{6.16}$$

Equation (6.16) is only invariant under (2.25). In order to make it invariant under (2.31), it can be modified as follows,

$$x^{[\mu}_p\eta^{\nu]\alpha}\frac{\partial\mathfrak{R}_{pq}}{\partial x^\alpha_{pU}}\bigg|_{\tau_p=\pm\infty} = 0, \quad p=1,2. \tag{6.17}$$

For open and semi-open objects neither the condition (6.16) nor (6.17) kills the surface integral in (6.13). In this case we add the two extra gauge covariant conditions

$$x^{[\mu}_p\eta^{\nu]\alpha}\frac{\partial\mathfrak{R}}{\partial\dot{x}^\alpha_p}_{pq}\bigg|_{\xi_p=\bar{x}^1_p,\delta^1_p} = 0, \quad p=1,2, \tag{6.18a}$$

$$x^{[\mu}_p\eta^{\nu]\alpha}\frac{\partial\mathfrak{R}}{\partial\hat{x}^\alpha_p}_{pq}\bigg|_{\eta_p=\bar{x}^2_p,\delta^2_p} = 0, \quad p=1,2. \tag{6.18b}$$

Conditions (6.18) are too strong for geometric objects, but for other types of objects it can be satisfied easily, e.g., the interaction given by (4.4) satisfies all the requirements to kill the surface term in (6.13).

From (6.13) and (6.16)–(6.18) it follows that

$$\frac{\partial}{\partial\xi^U_p}\left[x^{[\mu}_p\eta^{\nu]\alpha}\frac{\partial}{\partial x^\alpha_{pU}}\Omega_p\right] = \int d^3\xi_q\,\mathcal{J}^{\mu\nu}_{pq}. \tag{6.19}$$

Equation (6.19) is a generalization of the classical mechanics equations that relates the change of angular momentum to the torque. The function $\mathcal{J}^{\mu\nu}_{pq}$ describes a series of "angular impulses" coming from the object q that modifies the angular momentum of the object p. The antisymmetry of $\mathcal{J}^{\mu\nu}_{pq}$ with respect to p and q tells us that the impulses coming from p are equal in magnitude but opposite in sense to the impulses coming from q. Also, (6.19) tells us that the objects are asymptotically free, i.e.,

$$\begin{aligned}&\sum_{p=1}^{2} x^{[\mu}_p\eta^{\nu]\alpha}\frac{\partial}{\partial x_{pU}}\Omega_p\bigg|_{\tau_p=+\infty}\\ &= \sum_{p=1}^{2} x^{[\mu}_p\eta^{\nu]\alpha}\frac{\partial}{\partial x^\alpha_{pU}}\Omega_p\bigg|_{\tau_p=-\infty}.\end{aligned} \tag{6.20}$$

Equation (6.20) can be proved in a similar way to (6.11).

7. DISCUSSION

In this paper we have unified in a single formalism all the main features of the theory of direct interaction between particles, strings, and membranes previously developed by us.[2,4,5,7] Also, we have generalized some of our previous results, e.g., the study of the action and reaction law studied in Sec. 6 generalizes the work done in Ref. 7 in several respects: The objects in interaction are more complex and the type of inter-

action analyzed is more general. The methodology used to achieve these results does not depend on the particular type of action; it rather depends on its symmetries.

It is interesting to point out that the relativistic geometric membranes has not yet been quantized, one of the difficulties being that the general solution for the classical free membrane is not known. (For a discussion of this point see Refs. 2 and 5.) This problem of quantization presents some similarities with the quantization of the gravitational field, of course, being in principle much simpler.

The spinning string model of Neveu and Schwarz[17] has recently been derived[18] by a procedure analogous to the one used by Dirac to obtain from the Klein–Gordon equations the equations for electrons.[19] It seems possible to use the same procedure to generalize the Neveu–Schwarz model in two different directions, first by using a more general model of string, e.g., p-string, and second by using a higher dimensional model of string, e.g., geometric membranes or other types of membranes. In the latter case one might obtain a model bearing some similarities with the SLAC model[20] of hadrons. If this program can be realized, the study of the different objects discussed in this thesis might have some applications to describe properties of hadrons.

Our model of interacting objects does not include properties like classical spin or isospin. Preliminary work shows that it is possible to include these properties in an analogous form to what has been done[21] for the theory of direct interactions between particles.

The models of "bags" (closed membranes) studied in this thesis are models of empty bags, i.e., they do not have fields inside. By filling the bags it seems possible to construct new models of hadrons. This program presents some technical difficulties, e.g., the boundary value problem associated with the field inside the bags seems rather formidable.

ACKNOWLEDGMENT

I wish to thank Professor J. Stachel for the many enlightening discussions that we have had about parametrization of surfaces and the use of densities, and for a careful reading of the manuscript.

[1] A good review of the string model can be found in J. Scherk, Rev. Mod. Phys. **47**, 123 (1975).
[2] P. S. Letelier, J. Math. Phys. **17**, 923 (1976).
[3] M. Kalb and P. Ramond, Phys. Rev. D **9**, 2273 (1974).
[4] P. S. Letelier, Phys. Rev. D **15**, 1055 (1977).
[5] P. S. Letelier, Phys. Rev. D **16**, 322 (1977).
[6] H. W. Woodcock and P. Havas, Phys. Rev. D **6**, 3422 (1972).
[7] P. S. Letelier, "Conservation laws for a system of particles, strings, and membranes," Phys. Rev. D, to appear.
[8] T. Akabayashi, in *Quantum Mechanics, Determinism, Causality, and Particles*, edited by M. Flato *et al.* (Reidel, Dordrecht-Holland, 1976).
[9] Y. Nambu, in *Symmetries and Quark Models*, edited by R. Chand. (Gordon and Breach, New York, 1970); L. Susskind, Nuovo Cimento A **69**, 457 (1970); S.-H. H. Tye, Phys. Rev. D **13**, 3416 (1976).
[10] J. A. Schouten, *Tensor Analysis for Physicists* (Oxford U. P., London, 1951), p. 24.
[11] M. Gürses and F. Gürsey, Phys. Rev. D **11**, 967 (1975); C. Aragone and S. Deser, Nucl. Phys. B **92**, 327 (1975).
[12] P. S. Letelier and J. Stachel, paper in preparation.
[13] A. D. Fokker, Z. Phys. **58**, 386 (1929), see also R. P. Feynman and J. A. Wheeler, Rev. Mod. Phys. **21**, 425 (1949).
[14] D. Lovelock and H. Rund, in *Tensors, Differential Forms, and Variational Principles* (Wiley, New York, 1975), Chap. 5.
[15] F. Lund and T. Regge, Phys. Rev. D **14**, 1729 (1976).
[16] P. Havas, in *Problems in the Foundation of Physics*, edited by M. Bunge (Springer, New York, 1971).
[17] A. Neveu and J. H. Schwarz, Nucl. Phys. B **31**, 86 (1971).
[18] C. Teitelboim, Phys. Rev. Lett. **38**, 1106 (1977).
[19] P. A. M. Dirac, Proc. R. Soc. London A**117**, 610, A**118**, 351 (1928).
[20] W. A. Bardeen *et al.*, Phys. Rev. D **11**, 1094 (1975).
[21] W. Herman, Temple University thesis, 1976 (unpublished).

General algebraic theory of identical particle scattering[a)]

Gy. Bencze

Central Research Institute for Physics, [b)] H-1525 Budapest 114, P.O.B. 49, Hungary
and Department of Physics and Astronomy, University of Maryland, College Park, Maryland 20742

Edward F. Redish[c)]

Laboratory for Astronomy and Solar Physics, NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771
and Department of Physics and Astronomy, [b)] University of Maryland, College Park, Maryland 20742
(Received 30 January 1978)

We consider the nonrelativistic N-body scattering problem for a system of particles in which some subsets of the particles are identical. We demonstrate how the particle identity can be included in a general class of linear integral equations for scattering operators or components of scattering operators. The Yakubovskii, Yakubovskii–Narodestkii, Rosenberg, and Bencze–Redish–Sloan equations are included in this class. Algebraic methods are used which rely on the properties of the symmetry group of the system. Operators depending only on physically distinguishable labels are introduced and linear integral equations for them are derived. This procedure maximally reduces the number of coupled equations while retaining the connectivity properties of the original equations.

I. INTRODUCTION

Numerous methods for writing down mathematically well-behaved equations for the scattering operators in the nonrelativistic N-body problem have been developed in the past few years. [1–9] In most of these methods the particles are initially considered to be distinguishable, and any effects of particle identity are put in only when specific problems are treated. Since the number of physical situations and equation types is large, we here present a general method for constructing reduced equations which include the effect of particle identity.

In the bound state case, the treatment of particle identity is generally considered a technical problem. One must solve the Schrödinger eigenvalue problem on the subspace of properly symmetrized[10] vectors of the Hilbert space. Such problems can be typically handled either by the elaborate mathematical machinery of the shell model[11] or with hyperspherical harmonics. [12] Both methods rely heavily on the theory of the symmetric group.

On the other hand the inclusion of exchange symmetry in a scattering problem is far from trivial. The reason is that when exchange scattering is present, the asymptotic form of the scattering wavefunction includes both incoming and scattered waves in different regions of the many-body configuration space. There is no longer a single channel Hamiltonian which describes the asymptotic waves so all the well-known problems of describing rearrangement channels arise. One conventional procedure is simply to solve the scattering problem for distinguishable particles and afterwards to sum the contributions of physically indistinguishable channels to the cross section.

In some of the exact formulations of N-particle scattering the exchange symmetry can be incorporated directly into the scattering equations with a resulting decrease in the number of coupled equations. Such a procedure has been carried out by Lovelace in the three-body problem[13] and by Kharchenko and Kuzmichev[14] for the four-body Faddeev–Yakubovskii equations. The first explicit treatment of the scattering of an arbitrary number of identical particles was carried out by the authors in Ref. 15 (henceforth referred to as I). In this paper abstract group theoretic methods were used to incorporate exchange symmetry into the Bencze–Redish–Sloan (BRS) N-particle scattering equations. [5,6,16,17]

The treatment presented in I made use of the specific properties of the BRS equations. In this paper we develop a general algebraic method of including exchange symmetry which can be applied to a large variety of N-particle scattering equations and allows the treatment of an arbitrary number of different kinds of identical particles, which may be bosons or fermions. Specifically we consider two important classes of N-particle equations, the channel coupling class of equations, which are written in terms of transition operators, and the chain coupling class which are written for components of the N to N transition operator. The former class includes the BRS equations, the set of equations described by Bencze and Tandy, [19] and the equations of Chandler and Gibson. [9] The latter include the Rosenberg, [1] Yakubovskii, [3] and Yakubovskii–Narodestkii[4] equations. For the sake of completeness, it should be mentioned that there also exist N-particle formalisms intermediate between the chain and channel coupling classes. In these formalisms[20–22] integral equations are written for operators labelled by chains, but these operators are components of the physical transition operators which are labelled by two partitions. While it is straightforward to apply our algebraic method to these equations, for simplicity of discussion we restrict our considerations to equations of the chain and channel coupling classes.

The paper is organized as follows. Section II contains a discussion of the general form of the N-particle

a)Work supported in part by U. S. Department of Energy.
b)Permanent address.
c)N. R. C.—N. A. S. Sr. Res. Res. Assoc.

 0022-2488/78/1909-1909$1.00

equations and the basic group theoretic results associated with the treatment of identical particles. The symmetrization of the equations for the case of arbitrary numbers of identical particles is carried out in Sec. III and the mathematical properties of the symmetrized kernel are discussed. Section IV contains some applications and a study of the associated combinatorial problems. The results are summarized and conclusions presented in Sec. V.

II. PERMUTATION SYMMETRY IN N-PARTICLE SYSTEMS

There exist numerous formulations of N-particle scattering theory. Although they are all exact in principle, the N-particle dynamics is handled in different ways in the different theories. A common property of these exact theories is that they obtain coupled equations for operators or wave functions which are labelled by partitions or by chains of partitions[3] of the N-particle system.

Equations of the channel coupling class are written for the transition operators

$$T^{ab} = V^a + V^a G V^b, \tag{1}$$

where a and b label partitions of the system. Equations of the chain coupling class are written for the quantities T^A, where T^A is a component of the N to N transition operator

$$T^{00} \equiv T = \sum_A T^A. \tag{2}$$

We have used the notation 0 to indicate the N-cluster partition. The index A may be either a single partition or a set of partitions satisfying certain internal conditions (a chain). We will use Greek letters α, β, $\gamma, \cdots$ to indicate either a partition or a chain of partitions, and we write the set of possible labels as $\mathcal{L}$. The general equation then takes the form

$$T^\alpha = I^\alpha + \sum_{\beta \in \mathcal{L}} K^{\alpha\beta} T^\beta. \tag{3}$$

If Eqs. (3) refer to transition operators, the quantities T^α and I^α also carry a second label, γ, which indicates the initial state of the scattering process, i.e., $T^\alpha \to T^{\alpha\gamma}$, and similarly for I^α. For the sake of simplifying the notation we suppress this index except where it is relevant.

Let us now assume that the N-particle system contains some particles which are identical. In this case, the permutations of the identical particles form a finite symmetry group whose elements commute with the exact N-particle Hamiltonian. If all the N particles are identical, the symmetry group would be the full symmetric group on N objects, S_N. In the case that there are k different kinds of identical particles with n_i particles of the ith type ($N = n_1 + n_2 + \cdots + n_k$), then the symmetry group of the system, S, will be isomorphic to the direct product of symmetric groups, S_{n_i}, viz.,

$$S \cong S_{n_1} \times S_{n_2} \times \cdots \times S_{n_k}. \tag{4}$$

If the particles of the system are permuted by some element of S, then the system labels will in general also be affected. The group S induces a transformation group of the label set, $\mathcal{L}$, into itself. Symbolically we write

$$P\alpha = \alpha', \quad P \in S. \tag{5}$$

Depending on the system of labelling, certain elements of the symmetry group will leave a given label invariant. For example, if the label is a partition, the interchange of a pair of identical particles within a single cluster of the partition or the exchange of two identical clusters will produce a new partition which is identified as being identical to the original one. The set of elements $P \in S$ for which

$$P\alpha = \alpha \tag{6}$$

forms a subgroup of S which we denote as S_α.

The reduction of the integral equations (3) by the transformation group S is made possible by the transformation properties of the inhomogeneous term and the kernel under S. Specifically we assume that

$$PI^\alpha P^{-1} = I^{P\alpha} \tag{7}$$

and

$$PK^{\alpha\beta}P^{-1} = K^{P\alpha P\beta}. \tag{8}$$

In the case of the channel coupling equations, the suppressed index of the inhomogeneous term also is assumed to transform, viz.,

$$PI^{\alpha\beta}P^{-1} = I^{P\alpha P\beta}. \tag{9}$$

We describe any labelled quantity which transforms via a relation analogous to (7) or (8) as *label transforming*. Specifically, this means that the transform of the labelled quantity is that same quantity with only its labels transformed.

We now demonstrate that the quantities T^α are label transforming. In general, the kernel $K^{\alpha\beta}$ is a function of a complex parameter z which is taken to be equal to the scattering energy approached from above in the complex plane. We further assume that the kernel is compact or precompact in the complex z plane cut along the positive real axis. The analytic Fredholm theorem[23] then implies that Eq. (3) has a unique solution everywhere in the complex cut plane except possibly on a discrete set D. We then have the following:

Theorem I: If I^α and $K^{\alpha\beta}$ transform by Eq. (7) and (8) and $E \notin D$, then

$$PT^\alpha P^{-1} = T^{P\alpha}. \tag{10}$$

In the channel coupling case the second index of T transforms like (9).

Proof: Apply P to the left of (3) and P^{-1} to the right. Introduce $P^{-1}P$ between the K and the T. Using (7) and (8) gives

$$(PT^\alpha P^{-1}) = I^{P\alpha} + \sum_{\beta \in \mathcal{L}} K^{P\alpha P\beta}(PT^\beta P^{-1}).$$

For any $P \in S$ as β runs over $\mathcal{L}$, so does $P\beta$. Replacing the unknown vector of operators $PT^\alpha P^{-1}$ by the operators $T^{P\alpha}$ yields a solution by (3). Since $E \notin D$ the solution is unique. Q.E.D.

The situations in which these transformation pro-

perties are relevant can easily be seen. In the formalisms discussed above the dependence on a particular particle is entirely through a potential involving that particle. In that case it is easy to see that a renumbering of the particles produced by any permutation simply changes the resulting labels as required. For example, in Ref. 19 classes of equations are considered in which $I^{\alpha\gamma}=C_\gamma V^\alpha_\gamma$ and $I^{\alpha\gamma}=V^\alpha W_{\alpha\gamma}$. In the first case, the coefficients C_γ depend only on the number of clusters in each partition and are therefore invariant under a transformation ($C_{P_a}=C_a\ \forall P\in S$). The inhomogeniety is therefore label transforming. In the second case (Kouri—Levin[7] couplings) $W_{\alpha\gamma}$ is a numerical matrix whose elements sum to unity along each row and column. The value of a matrix element in this case depends on specific partitions, so in general $W_{P_aP_b}\neq W_{ab}$. This coupling scheme therefore does not satisfy transformation property (9).

Consider any label α. Since the permutation of identical particles does not change any physical property of the system, labels related by $\alpha'=P\alpha$ are physically equivalent. We therefore introduce a binary relation

$$\alpha' \mathcal{R} \alpha \quad \text{iff} \quad \exists\, P\in S\in P\alpha=\alpha'.$$

Lemma: $\mathcal{R}$ is an equivalence relation on $\mathcal{L}$.

The proof is straightforward.

The relation $\mathcal{R}$ therefore splits the set $\mathcal{L}$ into disjoint equivalence classes. Since the quantities which depend on the label set are in some sense physically equivalent for all labels in a given equivalence class, we expect that class functions can be constructed carrying all the physically relevant information. These class functions will be constructed in Sec. III.

We now construct the projection operators on states of proper symmetry. In the physical description of many-particle systems containing identical particles, the wavefunction of the system must transform by a one-dimensional irreducible representation of the symmetry group, namely

$$P\psi=\pm\psi,$$

the plus or minus being chosen according to whether an even or odd number of fermions is exchanged by the operator P. Since S is a product of permutation groups S_i, and since every element of S_i can be written as a product of transpositions, it follows that every element of S can be written as the product of transpositions of identical particles. We define f_P to be $+1$ if, when written as the product of transpositions of identical particles, P contains an even number of fermion transpositions, and -1 if it contains an odd number of fermion transpositions. For convenience we write $\hat{P}=f_pP$ and $|S|$ for the order of the group S. The symmetrizer on states of appropriate symmetry is[24]

$$R=\frac{1}{|S|}\sum_{P\in S}\hat{P}. \tag{11}$$

It satisfies the following properties:

$$R^2=R \tag{12}$$

and

$$\hat{Q}R=R\ \forall\ Q\in S. \tag{13}$$

One may obtain a useful factorization of the operator R, which generalizes the one given in I. Consider a particular element of the label set, $\alpha\in\mathcal{L}$. Those elements of S which leave α invariant form a subgroup of S which we label S_α. We then have

$$P\alpha=\alpha\ \forall\ P\in S_\alpha\subset S. \tag{14}$$

If we label the equivalence class by $a=[\alpha]$ and if N_a denotes the number of elements of $\mathcal{L}$ in a, then by Lagrange's theorem[24] we have

$$N_a=|S|/|S_\alpha|. \tag{15}$$

We then have:

Theorem: For any element α of the equivalence class a, if

$$R_\alpha=\frac{1}{|S_\alpha|}\sum_{P\in S_\alpha}\hat{P} \tag{16}$$

then

$$R=\frac{1}{N_\alpha}\sum_{\alpha'\in a}\hat{P}_{\alpha'\alpha}R_\alpha, \tag{17}$$

$$R=\frac{1}{N_a}\sum_{\alpha'\in a}R_\alpha\hat{P}_{\alpha\alpha'}, \tag{18}$$

where $P_{\alpha'\alpha}$ is any permutation in S which maps α into α'.

Proof: For each $\alpha'\in a$, $\hat{P}_{\alpha'\alpha}R_\alpha$ is a sum of elements of S (together with their phases) which exhausts a single coset of S_α. The summation over α' produces a sum over cosets. Since the cosets are disjoint, when the sum in (17) is expanded using (16), the resulting sum exhausts the entire group. The correctness of the normalization follows from (15). Similar arguments yield Eq. (18). Q. E. D.

Finally, we observe the result.

Lemma: For any label transforming quantity, A^α

$$R_\alpha A^\alpha=A^\alpha R_\alpha. \tag{19}$$

Proof: By the label transforming property, we have

$$PA^\alpha=A^\alpha P\ \ \forall\ P\in S_\alpha.$$

The result follows upon inserting the explicit expression for R_α, Eq. (16).

III. INTEGRAL EQUATIONS FOR IDENTICAL PARTICLE SCATTERING

In this section we construct operators labelled by the equivalence classes of the label set. These operators carry the complete content of the permutation symmetry of the N-particle system. Integral equations for these quantities are derived, reducing the number of coupled equations and the number of quantities required for the description of real processes.

Our first task is to define appropriate operators labelled by equivalence classes of labels. We review briefly the results of I to clarify the procedure. There, the quantities considered were transition operators for a system of N identical particles. Their matrix elements between fully symmetrized channel states gave the transition probabilities. Class operators are constructed by defining

$$T^{ab} = \sqrt{N_b/N_a} \sum_{\alpha \in a} R_{\alpha_0} \hat{P}_{\alpha_0} T^{\alpha \beta_0}, \tag{20}$$

where N_b and N_a are constants associated with the normalization of the channel wavefunctions. Here α_0 and β_0 denote fixed, but arbitrary, representatives of the equivalence classes a and b, respectively. These are referred to as *canonical labels*. The transition probability for physical processes is given by the on-shell matrix element

$$\mathcal{T}^{ab} = \langle \Phi_{\alpha_0} | T^{ab} | \Phi_{\beta_0} \rangle, \tag{21}$$

where the states Φ_{α_0} and Φ_{β_0} are channel wavefunctions satisfying the symmetry internal to the bound clusters of the channel, i.e.,

$$R_{\alpha_0} \Phi_{\alpha_0} = \Phi_{\alpha_0}, \quad R_{\beta_0} \Phi_{\beta_0} = \Phi_{\beta_0}. \tag{22}$$

The class function in this case was constructed by defining a canonical initial label and symmetrizing on the left. This is sufficient due to the well-known property that one may symmetrize either the initial state or the final state in a many-body scattering matrix element.[25] It is not necessary to symmetrize on both sides. The internal symmetry of the wavefunctions used to calculate matrix elements was retained because of the fact that bound states of the proper symmetry are then required.

We will construct class operators for our two general classes of scattering equations in a similar way. We first demonstrate that Eq. (3) can be shown to give an equation for the class operator constructed for the case of a general symmetry group. Then we show how the physical matrix elements are related to those of the class operators in the two cases. Finally we demonstrate that the connectivity structure of the equation is not destroyed in the transition to an equation for the class operators.

The class function we construct is

$$T^a = \sum_{\alpha \in a} R_{\alpha_0} \hat{P}_{\alpha_0 \alpha} T^{\alpha}, \tag{23}$$

for the chain coupling class and

$$T^{ab} = \sum_{\alpha \in a} R_{\alpha_0} \hat{P}_{\alpha_0 \alpha} T^{\alpha \beta_0} \sqrt{N_b/N_a} \tag{24}$$

for the channel coupling class. In the channel coupling case we take the initial index to be β_0 and suppress it and the numerical factor $(N_b/N_a)^{1/2}$ until the end of the discussion. The numerical factors will be chosen to yield the proper normalization of the physical matrix element.

Multiplying Eq. (3) by $R_{\alpha_0} \hat{P}_{\alpha_0 \alpha}$ and summing on the index α gives the equation

$$T^a = I^a + \sum_{\beta} X^{a\beta} T^{\beta}, \tag{25}$$

where we have defined

$$I^a = \sum_{\alpha \in a} R_{\alpha_0} \hat{P}_{\alpha_0 \alpha} I^{\alpha} \tag{26}$$

and

$$X^{a\beta} = \sum_{\alpha \in a} R_{\alpha_0} \hat{P}_{\alpha_0 \alpha} K^{\alpha \beta}. \tag{27}$$

The appropriate permutation operators must be extracted from $X^{a\beta}$ on the right. This is made possible by the following.

Proposition:

$$X^{a\beta} P_{\beta\beta'} = X^{a\beta'}. \tag{28}$$

Proof: This property follows directly from the transformation property of K. From the definition of X and the transformation property of K, we have

$$X^{a\beta} \hat{P}_{\beta\beta'} = \sum_{\alpha \in a} R_{\alpha_0} \hat{P}_{\alpha_0 \alpha} K^{\alpha\beta} P_{\beta\beta'}$$

$$= \sum_{\alpha \in a} R_{\alpha_0} \hat{P}_{\alpha_0 \alpha} \hat{P}_{\beta\beta'} K^{Q\alpha Q\beta},$$

where for simplicity of notation we have written

$$Q = P^{-1}_{\beta\beta'}.$$

Now $P^{-1}_{\beta\beta'} = P_{\beta'\beta}$ so

$$Q\beta = P_{\beta'\beta}\beta = \beta'$$

and

$$\hat{P}_{\alpha_0 \alpha} \hat{P}_{\beta\beta'} = \hat{P}_{\alpha}{}^{-1}_{\alpha_0} \hat{Q}^{-1} = (\hat{Q} \hat{P}_{\alpha\alpha_0})^{-1}$$

$$= (\hat{P}_{Q\alpha\alpha_0})^{-1} = \hat{P}_{\alpha_0 Q\alpha}.$$

Therefore, we have

$$X^{a\beta} \hat{P}_{\beta\beta'} = \sum_{\alpha \in a} R_{\alpha_0} \hat{P}_{\alpha_0 Q\alpha} K^{Q\alpha\beta'}.$$

As α runs over all the elements of a, so does $Q\alpha$ for any Q. Therefore, the sum may be taken over $Q\alpha$ instead of over α, and relabelling the dummy index $Q\alpha$ as α gives the result. Q.E.D.

Corollary:

$$X^{a\beta} R_{\beta} = X^{a\beta}. \tag{29}$$

This result follows immediately from the proposition as all the terms $\hat{P}$ contained in the R_β sum leave β invariant.

These two propositions allow us to make the right index on the $X^{a\beta}$ in Eq. (27) canonical by pulling out the appropriate permutation operator. The resulting operator is a class operator and the sum over all β may be broken up into a sum over b and a sum over $\beta \in b$. Using the proposition, Eq. (25) becomes

$$T^a = I^a + \sum_{\beta} X^{a\beta_0} \hat{P}_{\beta_0 \beta} T^{\beta}. \tag{30}$$

By the corollary the internal symmetrizer may be extracted to give

$$T^a = I^a + \sum_{\beta} X^{a\beta_0} R_{\beta_0} \hat{P}_{\beta_0 \beta} T^{\beta}. \tag{31}$$

Finally, breaking up the sum gives the result

$$T^a = I^a + \sum_{b} \mathcal{K}^{ab} T^b \tag{32}$$

where we have written

$$\mathcal{K}^{ab} = X^{a\beta_0} = \sum_{\alpha \in a} R_{\alpha_0} \hat{P}_{\alpha_0 \alpha} K^{\alpha \beta_0} \tag{33}$$

for the symmetrized kernel. Equation (32) is an integral equation for the class operators as desired. In the channel coupling case the right-hand sides of Eqs. (26) and (33) require the additional factor $(N_b/N_a)^{1/2}$.

We now consider the relation of the physical matrix elements to the matrix elements of the class operators. In the chain coupling case the physical matrix elements desired could be the matrix elements of the full T operator between some arbitrary initial state, ϕ; and a set of final noninteracting states of all the particles, ϕ_0. At least one of these states must have the correct symmetry. Because our equations (3) were chosen to have the T on the right, we have symmetrized from the left. We therefore require that the left wavefunction have the proper symmetry. The physical matrix element is therefore

$$T=\langle\tilde{\phi}_0|T|\phi'\rangle, \tag{34}$$

where[25]

$$\langle\tilde{\phi}_0|=\langle\phi_0|R[n_1!n_2!\cdots n_k!]^{1/2}. \tag{35}$$

The state ϕ_0 may be a set of plane waves, for example, or a coordinate space state in which each of the particles is at a particular point. Using (2), the physical matrix elements may be written

$$T=[n_1!n_2!\cdots n_k!]^{1/2}\langle\phi_0|R\sum_\alpha T^\alpha|\phi'\rangle. \tag{36}$$

By properties (13) and (16), the required permutation operators may be introduced before the T^α. We may therefore write

$$T=[n_1!n_2!\cdots n_k!]^{1/2}\langle\phi_0|R\sum_\alpha R_{\alpha_0}\hat{P}_{\alpha_0\alpha}T^\alpha|\phi'\rangle. \tag{37}$$

Decomposing the sum into a sum over classes, a, and a sum over the elements in those classes, $\alpha\in a$ gives

$$T=[n_1!\cdots n_k]^{1/2}\langle\phi_0|R\sum_a T^a|\phi'\rangle \tag{38}$$

so

$$T=\langle\tilde{\phi}_0|\sum_a T^a|\phi'\rangle. \tag{39}$$

For the channel coupling case the physical matrix elements are transition matrix elements. The indices of the T operators are partitions, and one takes matrix elements between states of the appropriate partition Hamiltonians. Physical channels are labelled by class indices instead of partitions. The transition matrix element between a state of class b and one of class a will be

$$T'_{ab}=\langle\tilde{\phi}_a|U|\tilde{\phi}_b\rangle, \tag{40}$$

where U is the Ekstein transition operator [see Eq. (2.5) of I] and $|\tilde{\phi}_a\rangle$ and $|\tilde{\phi}_b\rangle$ are final and initial states of the proper symmetry. Following I, we express both final and initial states in terms of symmetrizations of canonical states. We therefore set

$$|\tilde{\phi}_b\rangle=\sqrt{N_b}\,R|\phi_{\beta_0}\rangle \tag{41}$$

and similarly for $|\tilde{\phi}_a\rangle$. The initial state consists of the bound clusters of the representative partition β_0 and plane waves for their relative motion. As in I we assume that the wavefunctions of the bound clusters in both the initial and final states have the proper symmetry for each cluster, but have not been symmetrized between members of different clusters. This allows the straightforward insertion of standard bound state theories (which have the correct symmetry) into the scattering equations.

The operator T'_{ab} may now be expressed in terms of matrix elements of the T operator (1) as follows. Using the fact that R commutes with U and $R^2=R$, and inserting (41) into (40) gives

$$T'_{ab}=\sqrt{N_aN_b}\langle\phi_{\alpha_0}|RU|\phi_{\beta_0}\rangle. \tag{42}$$

Expanding R by Eq. (11) yields

$$T'_{ab}=\sqrt{N_aN_b}\,\frac{1}{|S|}\sum_{P\in S}\langle\phi_{\alpha_0}|P^{-1}U|\phi_{\beta_0}\rangle$$

$$=\frac{\sqrt{N_aN_b}}{|S|}\sum_{P\in S}\langle\phi_{P\alpha_0}|U|\phi_{\beta_0}\rangle. \tag{43}$$

Since the matrix element of U agrees with that of T^{ab} on-shell, we may replace T'_{ab} by

$$T_{ab}=(\sqrt{N_aN_b}/|S|)\sum_{P\in S}\langle\phi_{P\alpha_0}|T^{P\alpha_0\beta_0}|\phi_{\beta_0}\rangle. \tag{44}$$

Using Eq. (22), the definition of S_α, and the fact that

$$\langle\phi_\alpha|=\langle\hat{P}_{\alpha\alpha_0}\phi_{\alpha_0}|=\langle\phi_{\alpha_0}|\hat{P}^\dagger_{\alpha\alpha_0}=\langle\phi_{\alpha_0}|\hat{P}_{\alpha_0\alpha}, \tag{45}$$

we obtain

$$T_{ab}=(\sqrt{N_aN_b}/|S|)|S_a|\sum_{\alpha\in a}\langle\phi_{\alpha_0}|R_{\alpha_0}\hat{P}_{\alpha_0\alpha}T^{P\alpha_0\beta_0}|\phi_{\beta_0}\rangle. \tag{46}$$

It is now only necessary to observe the following:

Lemma:

$$N_a=|S|/|S_a|. \tag{47}$$

This follows very directly from the arguments of Ref. 25. Assuming the asymptotic states $|\tilde{\phi}_a\rangle$ and $|\tilde{\phi}_b\rangle$ correspond to normalizable wavepackets, we must have

$$\langle\tilde{\phi}_a|\tilde{\phi}_a\rangle=1$$

so

$$N_a=(\langle\phi_{\alpha_0}|R|\phi_{\alpha_0}\rangle)^{-1}.$$

Expanding out R by (11) gives

$$\langle\phi_{\alpha_0}|R|\phi_{\alpha_0}\rangle=(1/|S|)\sum_{P\in S}\langle\phi_{\alpha_0}|\hat{P}|\phi_{\alpha_0}\rangle.$$

If we choose wavepackets such that all the clusters are well separated, interchanging between different clusters gives 0 and within the same cluster gives 1 (assuming $\langle\phi_{\alpha_0}|\phi_{\alpha_0}\rangle=1$). Therefore, we have

$$\langle\phi_{\alpha_0}|R|\phi_{\alpha_0}\rangle=|S_a|/|S|=1/N_a.$$

We therefore have

$$T_{ab}=\sqrt{N_b/N_a}\sum_{\alpha\in a}\langle\phi_\alpha|T^{\alpha\beta_0}|\phi_{\beta_0}\rangle. \tag{48}$$

Introducing the class operator given by (24) the physical matrix element is given by

$$T_{ab}=\langle\phi_{\alpha_0}|T^{ab}|\phi_{\beta_0}\rangle. \tag{49}$$

We observe that all the complex counting considered in Paper I [including the δ term of Eq. (2.35)] is obtained from the straightforward group theoretic considerations of Eq. (47).

We now consider the connectivity structure of the resulting equations for the class operators, Eq. (32). We have the following:

Theorem: If the nth power of the operator $K^{\alpha\beta}$ is completely connected, then so is the nth power of the operator K^{ab}.

Proof: Let us consider the mth power of the operator K. The theorem follows immediately from the following.

Lemma:

$$(K^m)^{ab} = \sum_{\alpha \in a} R_{\alpha_0} \hat{P}_{\alpha_0 \alpha} (K^m)^{\alpha \beta_0}. \tag{50}$$

To prove the lemma, we observe that

$$(K^m)^{ab} = \sum_c (K^{m-1})^{ac} K^{cb}$$

$$= \sum_c (K^{m-1})^{ac} \sum_{\gamma \in c} R_{\gamma_0} \hat{P}_{\gamma_0 \gamma} K^{\gamma \beta_0} \tag{51}$$

by the definition of matrix multiplication and of K [Eq. (33)]. We now note that if we have (where $j = m - k$)

$$(K^m)^{ab} = \sum_c (K^j)^{ac} \sum_{\gamma \in c} R_{\gamma_0} \hat{P}_{\gamma_0 \gamma} (K^k)^{\gamma \beta_0} \tag{52}$$

with $j > 1$, we can write

$$(K^m)^{ab} = \sum_{d, c, \gamma \in c} (K^{j-1})^{ad} K^{dc} R_{\gamma_0} \hat{P}_{\gamma_0 \gamma} (K^k)^{\gamma \beta_0}$$

$$= \sum_{d, \gamma} (K^{j-1})^{ad} X^{d \gamma_0} R_{\gamma_0} \hat{P}_{\gamma_0 \gamma} (K^k)^{\gamma \beta_0}.$$

Using the proposition and corollary [Eqs. (28) and (29)] gives

$$= \sum_{d, \gamma} (K^{j-1})^{ad} X^{d \gamma} (K^k)^{\gamma \beta_0},$$

which by the definition of the operator X [Eq. (27)] becomes

$$= \sum_{d, \gamma} (K^{j-1})^{ad} \sum_{\delta \in d} R_{\delta_0} \hat{P}_{\delta_0 \delta} K^{\delta \gamma} (K^k)^{\gamma \beta_0}$$

$$= \sum_d (K^{j-1})^{ad} \sum_{\delta \in d} R_{\delta_0} \hat{P}_{\delta_0 \delta} (K^{k+1})^{\delta \beta_0}. \tag{53}$$

This equation has the same structure as Eq. (52) with k incremented by one. We may therefore induce on k beginning with $k = 1$ [Eq. (51)] and bring the operator over to the left until only a single power of K remains. The final step follows in a similar manner. Q.E.D.

We have therefore demonstrated that the symmetrization (23)—(24) defines class functions for both the chain and channel coupling cases, that equations for these class operators can be defined [Eq. (32)] with all the symmetry carried in the structure of the inhomogeneity (26) and the kernel (33), that the physical matrix elements are given simply in terms of matrix elements of the class operators [Eqs. (39) and (49)], and finally that the connectivity structure of the unsymmetrized equation is maintained by the symmetrization.

We conclude this section with a few comments about the KLT equations.[7,8] As we remarked above, the KLT method uses a numerical matrix in the kernel of their equations so that, although the KLT equations have the structure of (3), neither the kernel nor the resulting T operator are label transforming. There are many T operators corresponding to different choices of the W matrix. These T operators have different off-shell continuations, although the exact operators agree on-shell. The permutation transformation $T^{\alpha\beta} \to PT^{\alpha\beta}P^{-1}$ maps one T operator into a T operator corresponding to a different coupling scheme. This prevents the application of our procedure to this case. An attempt to symmetrize these equations was made by Tobocman.[18] In this work the equations are symmetrized before coupling in contrast to our method which symmetrizes after the coupled equations have been constructed. Tobocman's method does not produce an equation with connected kernel.

IV. APPLICATIONS AND ASSOCIATED COMBINATORIAL PROBLEMS

The results presented in the previous section make it straightforward to construct properly symmetrized N-particle integral equations once the symmetry group of the system is given. Only the basic combinatorial problems[26] of ennumerating equivalence classes needs to be solved to make the equations ready for practical applications.

In the following we present two specific applications in order to show precisely what problems arise and how they are dealt with. We consider first the nuclear physics example of Z identical protons and N identical neutrons described by channel coupling class equations employing minimal (two-cluster) coupling. As a second example we consider the problem of a homonuclear diatomic molecule with $2Z$ electrons and two identical spin zero (boson) nuclei, treated by the Yakubovskii equations,[3] a chain coupling case.

For the first example, N identical neutrons and Z identical protons, the permutation group of the system is isomorphic to

$$S \cong S_N \times S_Z$$

so the order of the group is $|S| = N!Z!$. We consider a formulation in which the indices are two-cluster partitions. An example of a formulation of this type is that of BRS which has a kernel whose first iterate is completely connected.

To construct the physical quanitites needed, one must determine the equivalence classes of the labels and the number of elements in each equivalence class. The number of distinct labels for $N+Z$ distinguishable particles is known[6] to be $2^{N+Z-1} - 1$. For the system with identical particles, all the distinct physical two-cluster partitions may be characterized by a pair of positive integers, (n, z), which give the number of neutrons and protons in the smaller fragment. Since the number of neutrons and protons are both fixed, the second fragment must contain $(N-n, Z-z)$ neutrons and protons, respectively.

A pair of integers, (N, Z), where N and Z cannot both vanish simultaneously, is known in the theory of combinatorics as a bipartite number.[27] The equivalence classes of the two-cluster partitions correspond precisely to all the partitions of a bipartite number into two bipartite numbers. The number of such partitions was given by Macmahon[27] in closed form as

$$R_{NZ} = \left\lfloor \frac{(N+1)(Z+1) - 1}{2} \right\rfloor. \tag{54}$$

This is therefore the number of equivalence classes

and therefore the number of coupled equations in Eq. (32) for this example.

We now must determine the number of elements in each equivalence class in order to construct the operators R_{α_0}. Given a particular two-cluster partition of the form, $(n,z)-(N-n,Z-z)$, how many distinct partitions are there in the same equivalence class? If one applies all the permutations in S, one would obtain $N!Z!$ partitions. Not all of them are distinct, however, as exchanging protons and/or neutrons in a single cluster does not lead to a new partition. We have to divide by the number of ways of permuting the protons and neutrons in each of the two clusters. This reduces the number of elements to $N!Z!/(n!(N-n)!z!(Z-z)!)$ or $\binom{N}{n}\binom{Z}{z}$. If the two clusters of the partition are identical, then the complete interchange of the two clusters does not produce a new partition. This requires reducing the number of distinct partitions produced by an addition factor of two. We therefore obtain

$$N_a = N_{(n,z)} = \binom{N}{n}\binom{Z}{z}\frac{1}{1+\delta_{N/2\,n}\delta_{Z/2\,z}} . \tag{55}$$

These numbers fix the normalization of the operators R_{α_0} which will be used in constructing the inhomogeneous term by Eq. (26) and the kernel by Eq. (33). This is also the number of terms in the $\alpha \in a$ summations [Eqs. (23) and (24)].

Let us now investigate how the Yakubovskii equations have to be symmetrized for the case of a diatomic homonuclear molecule. The labels for these equations are chains of partitions, $a_1 \supset a_2 \supset a_3 \supset \cdots \supset a_{N-1} \supset a_N$, where a_j represents a partition of the N-particles into j-clusters and the relation $a_j \supset a_{j+1}$ means the partition a_{j+1} can be obtained by breaking a single one of the clusters of a_j. The initial and final partitions, a_1 and a_N, being unique are usually omitted.

For distinguishable particles, the number of complete (maximal) chains can be ennumerated by elementary methods. This can be done because the set of partitions forms a semimodular, relatively complemented lattice.[28] When the symmetry group of the system is applied, the set partitions and therefore the set of labels is split up into disjoint equivalence classes. The equivalence classes of the partitions can still be regarded as a partially ordered set, since the ordering of the partitions is preserved by the mapping of each partition into its equivalence class. However, the lattice property of the set of partitions is in general not preserved. This

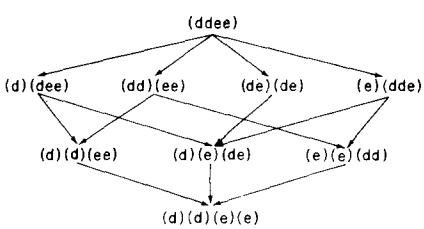


FIG. 1. The partition set for a four-body system consisting of two pairs of identical particles. Arrows indicate inclusion. A complete maximal chain is a path leading from the top element *(ddee)* to the bottom *(d)(d)(e)(e)*. There are seven such distinct chains.

TABLE I. Number of coupled integral equations for $N=3$, 4, and 5 particles in the Yakubovskii (Y) and Bencze–Redish–Sloan (BRS) methods. The labels a, b, c, $\cdots$ indicate distinct species of particles.

	$N=3$		
	aaa	*aab*	*abc*
Y	1	2	3
BRS	1	2	3

	$N=4$				
	aaaa	*aaab*	*aabb*	*aabc*	*abcd*
Y	2	5	7	11	18
BRS	2	3	4	5	7

	$N=5$						
	aaaaa	*aaaab*	*aaabb*	*aaabc*	*aabbc*	*aabcd*	*abcde*
Y	4	15	26	45	61	105	180
BRS	2	4	5	7	8	11	15

hinders one considerably in trying to find a closed form expression for the number of equivalence classes of maximal chains. The problem is equivalent to a currently unsolved problem in graph theory; however, an algorithm can be given for ennumerating the number of equivalence classes of chains without explicitly symmetrizing the original equations. The explicit symmetrization procedure has been used in all previously considered examples, viz., the works of Kharchenko and Kuzmichev[14] and that of Roy-Choudhury, *et al.*[29]

Under the symmetrization group, each partition will map into a multipartite number. Therefore, each chain will map into a chain of multipartite numbers. As a specific example we consider the system D_2 consisting of two deuterons and two electrons, labelled $1=d$, $2=d$, $3=e$, and $4=e$. All the possible distinct partitions are indicated in Fig. 1 with the possible inclusions shown by arrows. One can easily read off from this diagram that the number of maximal chains is seven. The number of coupled Yakubovskii equations for this system will therefore be seven. The kernels and inhomogeneous terms for the equations coupling the operators labelled by these chains can easily be constructed once the subgroups for each chain are determined. The number of Yakubovskii and BRS equations for the various cases of different numbers of identical particles in the three-, four- and five-body problem are given in Table I.

To be specific, the full symmetry group is generated by the operators P_{12} (interchange of the deuterons) and P_{34} (interchange of the electrons). Taking into account the fact that the deuterons are bosons and the electrons are fermions yields the full symmetrizer

$$R = \tfrac{1}{4}(1 + P_{12} - P_{34} - P_{12}P_{34}).$$

Consider the chain $a = (dd)(ee) \supset (d)(d)(ee)$. A canonical chain for this equivalence class is the chain $\alpha_0 = (12)(34) \supset (1)(2)(34)$. The full symmetry group leaves this chain invariant. Therefore, we construct the operator

$$R_{\alpha_0} = R.$$

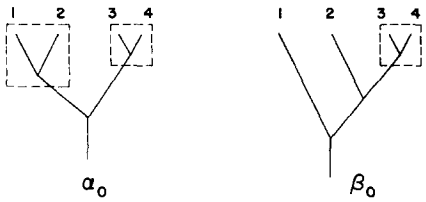


FIG. 2. Tree representations of the canonical chains in the four-body problem *ddee*. The chains are read off from the tree by following the successive connections down from the top of the tree. Particles 1 and 2 are type *d* and 3 and 4 are type *e*. The dotted boxes indicate parts of the tree which are invariant under an element of the symmetry group.

There is only a single term in the α sum in Eqs. (23), (26), and (33). For the chain $b = (d)(dee) \supset (d)(d)(ee)$ a canonical chain is $\beta_0 = (1)(234) \supset (1)(2)(34)$. Only the group elements 1 and P_{34} leave this chain invariant. Therefore, we have

$$R_{\beta_0} = \tfrac{1}{2}(1 - P_{34}).$$

There are two chains in this equivalence class, β_0, and the chain $(2)(134) \supset (1)(2)(34)$. There will therefore be two elements in the sums $\beta \in b$.

In general, each chain can be associated with a tree. The trees for the chains α_0 and β_0 in the above example are shown in Fig. 2. The invariance subgroup of a particular chain is determined by the number of indistinguishable particles which are connected directly together rather than being joined to a cluster. The internal symmetry operator for each chain must be determined by considering the structure of the specific chain.

Upon the construction of these operators, the equations coupling the symmetrized operators, (32), may be written down directly for any case without the need of beginning with the original equations (3).

V. CONCLUSIONS

The treatment of identical particles in N-particle scattering gives rise to nontrivial mathematical problems. In this paper we demostrate that a general algebraic treatment can be developed for a large class of N-particle scattering theories. This treatment relies on the properties of the permutation group of the system. These theories include channel coupling equations such as those of BRS or Chandler and Gibson (but explicitly excluding theories of KLT type) and chain coupling equations such as those of Rosenberg, Yakubovskii, and Narodestkii and Yakubovskii.

There are two features of basic importance which make the general algebraic treatment possible. First, in all the N-particle theories the quantities to be determined are labelled by partitions or (incomplete or complete) chains of partitions of the N-particle system. Consequently, the number of coupled equations is determined by the way of labelling rather than the treatment of N-body dynamics. Second, the permutation group of the system generates an equivalence relation on the set of labels and accordingly splits it into disjoint equivalence classes. Only the equivalence classes and operator valued functions of these classes have physical meaning. This statement is just an abstract formulation of the property that due to the indistinguishability of some of the particles the amplitudes of the physical processes can be expressed as coherent sums of "direct" and "exchange" processes. In our general algebraic treatment we construct operators which are sums of permutation operators acting on the scattering operators of the relevant theory. These class operators given by Eqs. (23) and (24) are labelled by the equivalence classes of the set of original labels. If the inhomogeneous term and the kernel of the relevant set of N-particle equations satisfy certain general conditions, i.e., if they are label transforming, it follows that the N-particle equations can be reformulated in terms of the class operator, with a simultaneous reduction in the number of coupled equations. Furthermore, the physical matrix elements are expressed in terms of matrix elements of the class operators. The entire burden of the symmetry is carried by the symmetrization of the inhomogeneous term and the kernel. We also demonstrate that the connectivity properties of the kernel are passed on to the symmetrized equation.

The most remarkable fact exhibited by the considerations of this paper is that the treatment of identical particles, i.e., exchange effects, is actually independent of the N-particle dynamics for a large class of N-particle scattering theories. This is the consequence of the nondynamical nature of the permutation symmetry of the system.

ACKNOWLEDGMENTS

The authors are grateful to the National Science Foundation (U.S.A.) and the Institute of Cultural Relations (Hungary) for providing the exchange grant which made this collaboration possible. The authors would like to thank the Central Research Institute for Physics (Budapest) and the University of Maryland (College Park) for their hospitality. We would also like to thank Professor F.S. Levin for pointing out the difficulties with the symmetrization of the KLT equations.

[1] L. Rosenberg, Phys. Rev. **140**, B217 (1965).
[2] O.A. Yakubovskii, Yad. Fiz. **5**, 1312 (1967) [Sov. J. Nucl. Phys. **5**, 937 (1967)].
[3] P. Grassberger and W. Sandhas, Nucl. Phys. B **2**, 181 (1967); E.O. Alt, P. Grassberger, and W. Sandhas, Dubna preprint.
[4] I.M. Narodetskii and O.A. Yakubovskii, Yad. Fiz. **14**, 315 (1971) [Sov. J. Nucl. Phys. **14**, 178 (1972)].
[5] Gy. Bencze, Nucl. Phys. A **210**, 568 (1973).
[6] E.F. Redish, Nucl. Phys. A **225**, 16 (1974).
[7] D.F. Kouri and F.S. Levin, Nucl. Phys. A **253**, 395 (1975).
[8] W. Tobocman, Phys. Rev. C **9**, 2466 (1974).
[9] C. Chandler and A. Gibson, J. Math. Phys. **14**, 1328 (1973); **18**, 2336 (1977).
[10] We here and henceforth use the word "symmetrized" to mean symmetrized when the particles are bosons and antisymmetrized when the particles are fermions.
[11] A. deShalit and I. Talmi, *Nuclear Shell Theory* (Academic, New York, 1963).
[12] Yu.A. Simonov, Yad. Fiz. **7**, 1210 (1968) [Sov. J. Nucl. Phys. **7**, 722 (1968)].

[13]C. Lovelace, Phys. Rev. **135**, B1225 (1964).
[14]V. F. Kharchenko and V. E. Kuzmichev, Nucl. Phys. A **183**, 606 (1972).
[15]Gy. Bencze and E. F. Redish, Nucl. Phys. A **238**, 240 (1975).
[16]I. H. Sloan, Phys. Rev. C **6**, 1945 (1972).
[17]A similar treatment of the Kouri—Levin—Tobocman (KLT) equations[7,8] was attempted by Tobocman.[18] Difficulties with this approach will be discussed below.
[18]W. Tobocman, Phys. Rev. C **12**, 1146 (1975).
[19]Gy. Bencze and P. Tandy, Phys. Rev. C **16**, 564 (1977).
[20]W. Sandhas, in *Few Body Dynamics*, edited by A. Mitra, I. Slaus, V. Bhasin, and V. Gupta (North-Holland, Amsterdam, 1976), p. 540.
[21]B. R. Karlsson and E. M. Zeiger, Phys. Rev. D **9**, 1761 (1974); D **10**, 1291 (1974).
[22]V. Vanzani, Lett. Nuovo Cimento **16**, 1 (1976).
[23]M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. I: Functional Analysis* (Academic, New York, 1972).
[24]H. Boerner, *Representations of Groups* (North-Holland, Amsterdam, 1963); C. W. Curtis and I. Reiner, *Representation Theory of Finite Groups and Associative Algebras* (Interscience, New York, 1962).
[25]M. Goldberger and K. M. Watson, *Collision Theory* (Wiley, New York, 1964).
[26]Gy. Bencze, Phys. Lett. B **72**, 155 (1977).
[27]P. A. Macmahon, *Combinatory Analysis* (Chelsea, New York, 1960).
[28]G. Szâsz, *Introduction to Lattice Theory* (Academic, New York, 1963).
[29]H. Roy-Choudhury, V. Gautam, and D. Sural, Phys. Rev. C. **16**, 465 (1977).

Energy–momentum tensor symmetries and concomitant conservation laws. I. Einstein-massless-scalar (meson) field

Gerald H. Katzin[a] and Jack Levine

Department of Physics [a] *and Department of Mathematics, North Carolina State University, Raleigh, North Carolina 27650*
(Received 10 April 1978)

Symmetries of energy–momentum tensors $\mathcal{T}$ in a Riemannian space–time are defined by infinitesimal mappings $\bar{x}^i = x^i + \xi^i(x)\delta a$ where the mapping vector ξ^i is determined by the symmetry condition $\mathcal{L}_\xi(g^{w/2}\mathcal{T}) = 0$, ($g^{w/2}\mathcal{T}$ is a relative tensor of weight w, $g \equiv$ absolute value of the metrical determinant, and $\mathcal{L}_\xi$ is the Lie derivative with respect to the vector ξ^i). The existence of such symmetry vectors ξ^i leads to concomitant conservation laws in the form of conserved vector currents J^i for both special and general relativity. The currents J^i will be explicit functions of the energy momentum tensor $\mathcal{T}$ and the symmetry vector ξ^i. The symmetries and conservation laws so obtained will in general differ from the familiar Trautman formulation. The theory is applied to obtain symmetries and conserved currents for a class of conformally flat solutions of the Einstein-massless-scalar (meson) field equations.

I. INTRODUCTION

Let $g^{w/2}\mathcal{T}$ be a relative tensor of weight w where $\mathcal{T}$ is an absolute tensor of valence (p,q) $(p,q \geqslant 0)$, in a Riemannian space–time V_4 with metric g_{ij} (signature -2). An infinitesimal mapping

$$\bar{x}^i = x^i + \xi^i(x)\delta a \tag{1.1}$$

is said to define a symmetry of the relative tensor if

$$\mathcal{L}_\xi(g^{w/2}\mathcal{T}) = 0,$$

$$g \equiv \text{absolute value det} g_{ij}, \tag{1.2}$$

where $\mathcal{L}_\xi$ (also written as $\mathcal{L}$ for brevity) is the Lie derivative[1] with respect to the vector ξ^i of (1.1). Such vectors ξ^i which satisfy Eq. (1.2) are called symmetry vectors, and the corresponding mappings Eq. (1.1) are called symmetry mappings.

In this paper we shall consider symmetries of the type (1.2) where the absolute tensor $\mathcal{T}$ is taken to be an energy–momentum tensor. We shall show that such symmetries lead to conservation laws in the form of conserved vector currents for both special and general relativity. The functional structure of the conserved currents will be explicit functions of the energy–momentum tensor and the symmetry vector. [See for example Eq. (2.5D).] These symmetries are of additional interest since they are interrelated with the kinematical properties of the matter and/or fields associated with the energy–momentum tensor.

For the case of general relativity these symmetries may be given an alternative formulation and/or interpretation in terms of the metrical quantities of the left side of the Einstein field equations. An analysis of these symmetries based on such an alternative formulation will be included in a later paper which will also include a study of symmetries defined by Eq. (1.2) where $\mathcal{T}$ need not be an energy–momentum tensor.[2] For the case of special relativity the abovementioned alternative formulation will of course not be present.

The forms of the conservation laws associated with the symmetries Eq. (1.2) will be similar to those generalizations of the Trautman[3] conservation law developed by Katzin–Levine–Davis [Ref. 2(a)] and Collinson.[4] Due to the differences in the symmetries defined by Eq. (1.2) and those used by the abovementioned writers the actual conserved vector currents associated with Eq. (1.2) will in general differ from those previously obtained by these writers. We give below a brief discussion of the abovementioned earlier work insofar as it relates to our present work.

From Trautman's work[3] it is well known that if a Riemannian space–time V_4 admits a Killing vector ξ^i (motion) defined by[5]

$$\mathcal{L}g_{ij} \equiv \xi_{i;j} + \xi_{j;i} = 0, \tag{1.3}$$

then there exists a conservation law of the form

$$[T^i_j\xi^j]_{;i} \equiv g^{-1/2}[g^{1/2}T^i_j\xi^j]_{,i} = 0, \tag{1.4}$$

where $\xi^j = g^{ji}\xi_i$, and where $T^i_j (\equiv g_{jk}T^{ik})$ is an energy–momentum tensor for which

$$T^{ij} = T^{ji}, \tag{1.5}$$

$$T^{ij}_{\ \ ij} = 0. \tag{1.6}$$

It also follows from Trautman's work that if a space–time V_4 admits a conformal motion defined by

$$\mathcal{L}g_{ij} = 2\sigma(x)g_{ij}, \tag{1.7}$$

then a conservation law of the form (1.4) will exist based upon the symmetry vector ξ^i which satisfies Eq. (1.7) provided $T = 0$, where

$$T \equiv g_{ij}T^{ij}. \tag{1.8}$$

The generalization of Trautman's work by Katzin, Levine, and Davis [Ref. 2(a)] showed if a Riemannian space–time of general relativity with scalar curvature $R = 0$ admitted a symmetry (which they called a Ricci collineation)[6] defined by a vector ξ^j which satisfied[7]

$$\mathcal{L}R_{ij} = 0, \tag{1.9}$$

then it followed that the space–time admitted the vanishing vector divergence

$$(R^i_j\xi^j)_{;i} = 0. \tag{1.10}$$

 0022-2488/78/1909-1918$1.00

By use of the Einstein field equations

$$R^i_j - \tfrac{1}{2}R\delta^i_j = -T^i_j \tag{1.11}$$

(with $R=0$ and hence $T=0$) it then followed that a conservation law of the Trautman form Eq. (1.4) again resulted but based upon the Ricci collineation symmetry vector ξ^i defined by Eq. (1.9).

Collinson[4] pointed out that the Katzin–Levine–Davis equation (1.10) based upon Eq. (1.9) was also valid for $R\neq 0$. From Eq. (1.11) expressed in the form

$$R^i_j = (T^i_j - \tfrac{1}{2}T\delta^i_j) \tag{1.12}$$

it then followed that (1.10) could be expressed in the form[4]

$$[g^{1/2}(T^i_j - \tfrac{1}{2}T\delta^i_j)\xi^j]_{;i} = 0. \tag{1.13}$$

It also follows by means of the field equations (1.11) expressed in the form (1.12) that the symmetry condition Eq. (1.9) implies that the energy–momentum tensor satisfies the symmetry condition

$$\mathcal{L}(T_{ij} - \tfrac{1}{2}Tg_{ij}) = 0. \tag{1.14}$$

It is easily shown by contraction of Eq. (1.14) with g^{ij} and use of Eq. (1.5) and (1.6) that the associated conservation law Eq. (1.13) is again obtained (as would be expected).

By means of the Bianchi identity

$$R_{\ j} \equiv 2R^i_{j;i} \tag{1.15}$$

and the expansion of $\mathcal{L}R_{ij}$ it follows that

$$2(R^i_j\xi^j)_{;i} \equiv g^{ij}\mathcal{L}R_{ij}. \tag{1.16}$$

From Eq. (1.16) it is immediate that the existence of a Ricci collineation Eq. (1.9) is a sufficient but not a necessary condition for the existence of a conservation law of the form (1.13).[8]

By means of the identity Eq. (A10) of Appendix A with $S_{ij}=R_{ij}$ and (1.16) it follows that

$$2(R^i_j\xi^j)_{;i} \equiv g^{ij}\mathcal{L}R_{ij} \equiv -g^{-1/2}g_{ij}\mathcal{L}[g^{1/2}(R^{ij} - \tfrac{1}{2}Rg^{ij})]. \tag{1.17}$$

From Eq. (1.17) we observe that an alternative to the Ricci collineation Eq. (1.9) as a sufficient condition to obtain the vanishing divergence Eq. (1.10) [and hence to obtain a conservation law of the form (1.13)] is the symmetry condition

$$\mathcal{L}[g^{1/2}(R^{ij} - \tfrac{1}{2}Rg^{ij})] = 0. \tag{1.18}$$

By means of the Einstein field equations (1.11) the symmetry defined by Eq. (1.18) implies

$$\mathcal{L}(g^{1/2}T^{ij}) = 0, \tag{1.19}$$

which is of the form (1.2). We are thus led (as mentioned in the beginning of this section) to the study of symmetries of this form.

In Sec. II we formulate three types of symmetry conditions of the form (1.2) based upon covariant (T_{ij}), contravariant (T^{ij}), and mixed (T^i_j) forms, respectively, of the energy–momentum tensor. Although T_{ij}, T^{ij}, and T^i_j may each be considered as representations of the same physical energy–momentum tensor, the respective symmetry conditions based on these three forms of the tensor do not in general define equivalent symmetry mappings. For each type of symmetry condition of the form (1.2) we obtain for the two cases weight $w=1$ or $\mathcal{L}T=0$ (w arbitrary) a conservation law (in the form of a conserved current) concomitant with the existence of a symmetry mapping. For general weight w we also show that each type of symmetry mapping leads to a conservation law involving an additional vector A^i defined by $\mathcal{L}_\xi T = A^i_{;i}$. Such vectors A^i will always exist (locally).

For conformally related space–times a necessary and sufficient condition is obtained in order that conserved currents in the respective spaces be conformally related.

In Sec. III we apply the theory of Sec. II to a class of conformally flat solutions of the Einstein massless-scalar (meson) field equations for which the energy–momentum tensor has the form $T_{ij} = \psi_{,i}\psi_{,j} - \tfrac{1}{2}\Lambda g_{ij}$, $\Lambda \equiv g^{ij}\psi_{,i}\psi_{,j}$ (where ψ is a scalar field satisfying $g^{ij}\psi_{;ij}=0$. We determine the symmetries which satisfy $\mathcal{L}_\xi(g^{w/2}T_{ij})=0$ and obtain the concomitant conserved vector currents which in general are not of the Trautman form. For the case $\Lambda=0$ (which implies $\psi^i \equiv g^{ij}\psi_{,j}$ is a null vector), the general solution for the symmetry vector ξ^i is found along with the concomitant conserved current vectors. It is shown that the conserved current vectors J^i are null vectors which are parallel displaced with respect to the null geodesic congruences defined by ψ^i, and that the J^i themselves define null geodesic congruences. For the case $\Lambda\neq 0$ several types of symmetry solutions are obtained depending on the choice of parameters which occur in the metric g_{ij} and the scalar field ψ and on the assumed value of the weight w in the symmetry condition. Of particular interest when $w=1$ is a ten-parameter group of symmetries G_{10} which contains a seven-parameter subgroup of conformal motions which in turn contains a six-parameter subgroup of motions. The conserved current vectors concomitant with this G_{10} group of symmetries are shown to be spacelike and also have the interesting property of being conserved currents in the Minkowski space–time which is conformally related to the underlying conformally flat space–time. In a suitably chosen basis of this G_{10} it is shown that the conserved currents concomitant with the three nonconformal symmetries have a conformal scale relationship to the conserved currents associated with three of the motion symmetries of the G_{10}.

II. CONSERVATION LAWS BASED ON SYMMETRIES OF THE ENERGY–MOMENTUM TENSOR

Each of the three associated forms of the energy–momentum tensor T_{ij}, T^{ij}, T^i_j when used in (1.2) for $\mathcal{T}$ will determine a symmetry equation. We write these in the respective forms (for generality we assume a Riemannian space–time)

$$\mathcal{L}_D(g^{w/2}T_{ij}) = 0, \tag{2.1D}$$

$$\mathcal{L}_U(g^{w/2}T^{ij}) = 0, \tag{2.1U}$$

$$\mathcal{L}_M(g^{w/2}T^i_j) = 0. \tag{2.1M}$$

We denote by ξ^i_D, ξ^i_U, and ξ^i_M the three symmetry vectors

associated with the three symmetry equations (2.1D), (2.1U), (2.1M), respectively. As mentioned in the previous section these vectors will in general be distinct. The notation $\mathcal{L}_D$, for example, indicates Lie differentiation with respect to the vector $\xi^i{}_D$.

We now show how to obtain conservation laws, in the form of conserved currents, which are associated with the existence of each of the abovementioned symmetry vectors. With this in mind, we consider first (2.1D) which expands to[1]

$$\mathcal{L}_D T_{ij} + w\xi^k{}_{D;k} T_{ij} = 0. \tag{2.2D}$$

Further expansion of (2.2D) gives

$$T_{ij;k}\xi^k{}_D + T_{kj}\xi^k{}_{D;i} + T_{ik}\xi^k{}_{D;j} + w\xi^k{}_{D;k} T_{ij} = 0. \tag{2.3D}$$

Contraction of (2.3D) with g^{ij} and use of Eqs (1.5), (1.6), and (1.8) leads to

$$2\left[\left(T^i_j + \frac{w}{2} T\delta^i_j\right)\xi^j{}_D\right]_{;i} - (w-1)T_{,i}\xi^i{}_D = 0. \tag{2.4D}$$

In a similar manner we obtain from (2.1D) the condition

$$2\left[\left(T^i_j - \frac{w}{2} T\delta^i_j\right)\xi^j{}_U\right]_{;i} + (w-1)T_{,i}\xi^i{}_U = 0. \tag{2.4U}$$

By inspection of Eqs. (2.4D) and (2.4U) we may state the following theorem

Theorem 2.1 D,U: With respect to a symmetric energy–momentum tensor T^{ij} ($T^{ij}{}_{;j} = 0$, $T \equiv g^{ij}T_{ij}$) of a Riemannian space–time:

(D) If there exists an infinitesimal symmetry mapping $\bar{x}^i = x^i + \xi^i{}_D(x)\delta a$ such that $\mathcal{L}_D(g^{w/2}T_{ij}) = 0$, then there exists a concomitant conservation law

$$[(T^i_j + \tfrac{1}{2}wT\delta^i_j)\xi^j{}_D]_{;i} = 0, \tag{2.5D}$$

provided $w = 1$ or $\mathcal{L}_D T = 0$ (w arbitrary).

(U) If there exists an infinitesimal symmetry mapping $\bar{x}^i = x^i + \xi^i{}_U(x)\delta a$ such that $\mathcal{L}_U(g^{w/2}T^{ij}) = 0$, then there exists a concomitant conservation law

$$[(T^i_j - \tfrac{1}{2}wT\delta^i_j)\xi^j{}_U]_{;i} = 0, \tag{2.5U}$$

provided $w = 1$ or $\mathcal{L}_U T = 0$ (w arbitrary).

Remark 1: If in (2.4D) the term $(w-1)T_{,i}\xi^i{}_D \neq 0$, then since (locally) there always exist vectors $A^i{}_D(x)$ such that

$$T_{,i}\xi^i{}_D = \mathcal{L}_D T = A^i{}_{D;i}, \tag{2.6D}$$

we have concomitant with the symmetry (2.1D) the conservation law

$$\left[\left(T^i_j + \frac{w}{2} T\delta^i_j\right)\xi^j{}_D - \frac{(w-1)}{2} A^i{}_D\right]_{;i} = 0. \tag{2.7D}$$

In a like manner if in Eq. (2.4U) the term $(w-1)T_{,i}\xi^i{}_U \neq 0$ we have concomitant with the symmetry (2.1U) the conservation law

$$\left[\left(T^i_j - \frac{w}{2} T\delta^i_j\right)\xi^j{}_U + \frac{(w-1)}{2} A^i{}_U\right]_{;i} = 0. \tag{2.7U}$$

Note that in Eqs. (2.7D) and (2.7U) the vectors $A^i{}_D$ and $A^i{}_U$ are not unique.

Remark 2: Note that if $T = 0$ the conservation laws (2.5D) and (2.5U) reduce to the Trautman form (1.4).

Remark 3: Note that the conservation law (2.5U) based upon the symmetry condition Eq. (2.1U) with $w = 1$ is of the same form as the conservation law (1.13) based upon the existence of the Ricci collineation (1.9) [if use be made of the Einstein field equations in the form (1.12)].

Remark 4: Note the conservation law (2.5D) concomitant with the symmetry (2.1D) with $w = 1$ differs in form from (1.13) [based upon (1.9) or equivalently[9] (1.14)] and also from (2.5U) [based upon (2.1U) with $w = 1$ or equivalently[9] (1.18)].

For the case of general relativity it is immediate that the symmetry (2.1D) $w = 1$ can be given the equivalent form

$$\mathcal{L}_D[g^{1/2}(R_{ij} - \tfrac{1}{2}Rg_{ij})] = 0, \tag{2.8D}$$

and the associated conservation law (2.5D) may expressed in the form

$$[(R^i_j - R\delta^i_j)\xi^j{}_D]_{;i} = 0. \tag{2.9D}$$

This latter equation should be compared with (1.10) [which holds for both the symmetry vectors defined by (1.9) and (2.1U) with $w = 1$].

The following identity is derived in Appendix B,

$$\mathcal{L}(g^{1/2}R) \equiv (g^{1/2}R^j_i\xi^i)_{,j} - \tfrac{1}{2}g^{ij}\mathcal{L}[g^{1/2}(R_{ij} - \tfrac{1}{2}Rg_{ij})]. \tag{2.10}$$

It follows by use of (2.10) that with respect to a symmetry vector $\xi^i{}_D$ [satisfying (2.8D)] the Lie derivative of the Lagrange density of the gravitational field is represented as a divergence in the form

$$\mathcal{L}_D(g^{1/2}R) = (g^{1/2}R^j_i\xi^i{}_D)_{;j} = (g^{1/2}R^j_i\xi^i{}_D)_{,j}. \tag{2.11D}$$

We finally consider the case in which the symmetry mapping (1.1) is defined by (2.1M). Expansion of (2.1M) followed by contraction on i and j leads to

$$(wT\xi^i{}_M)_{;i} - (w-1)T_{,i}\xi^i{}_M = 0. \tag{2.12M}$$

By inspection of (2.12M) we may state the following theorem.

Theorem 2.2M: If a Riemannian space–time admits an infinitesimal symmetry mapping $\bar{x}^i = x^i + \xi^i{}_M(x)\delta a$ such that $\mathcal{L}_M(g^{w/2}T^i_j) = 0$, where T^i_j is the mixed form of the energy–momentum tensor, then there exists a concomitant conservation law

$$(T\xi^i{}_M)_{;i} = 0 \tag{2.13M}$$

provided $w = 1$ or $\mathcal{L}_M T = 0$ ($w \neq 0$ but otherwise arbitrary).

Remark 5: If in (2.12M) the term $(w-1)T_{,i}\xi^i{}_M \neq 0$, then in a manner similar to that discussed in Remark 1 following Theorem 2.1D, U we may obtain the conservation law

$$(wT\xi^i{}_M - (w-1)A^i{}_M)_{;i} = 0, \tag{2.14M}$$

concomitant with the symmetry (2.1M).[10]

For the case of general relativity it is easily shown by means of the Einstein field equations that the symmetry conditions (Eqs. 2.1D,U,M) and the equation $\mathcal{L}T=0$ are always satisfied by any motion vector admitted by the space–time. It follows from Theorems 2.1D,U and 2.2M that the conservation laws (2.5D,U) and (2.13M) hold for the motion symmetry vectors. Hence (2.13M) with either (2.5D) or (2.5U) give Trautman's conservation law (1.4) for motion symmetries. Note that these remarks imply that a motion symmetry vector leads to the existence of two (in general) conservation laws $(T^i_j\xi^j)_{;i}=0$, $(T\xi^i)_{;i}=0$.

We now prove a theorem concerning conserved current vectors in conformally related space–times. While this theorem is of interest in itself it will have application in the example considered in the next section. Consider then a space–time V_4 with metric $g_{ij}(x)$ which is conformally related to a space–time $\tilde{V}_4$ with metric $\tilde{g}_{ij}(x)$ in that

$$\tilde{g}_{ij}(x)=e^{2\sigma}g_{ij}(x),$$
$$\tilde{g}=e^{8\sigma}g,\ \sigma=\sigma(x). \tag{2.15}$$

Let $J^i(x)$ be a conserved current in the V_4 in that

$$J^i_{;i}=0. \tag{2.16}$$

In $\tilde{V}_4$ define a vector $\tilde{J}^i(x)$ conformally related to $J^i(x)$ by

$$\tilde{J}^i(x)=e^{2\sigma}J^i(x). \tag{2.17}$$

We have by use of (2.15) and (2.17) that[11]

$$\tilde{J}^i_{;i}\equiv\tilde{g}^{-1/2}(\tilde{g}^{1/2}\tilde{J}^i)_{,i}=e^{2\sigma}[J^i_{;i}+6\sigma_{,i}J^i]. \tag{2.18}$$

By inspection of (2.18) we may state the following theorem.

Theorem 2.3: Given two conformally related Riemannian space–times V_4 and $\tilde{V}_4$ with respective metrics $g_{ij}(x)$ and $\tilde{g}_{ij}(x)=e^{2\sigma(x)}g_{ij}(x)$ such that V_4 admits the conserved current $J^i(x)$ in that $J^i_{;i}=0$, then a necessary and sufficient condition that $\tilde{V}_4$ admit the conformally related conserved current $\tilde{J}^i(x)\equiv e^{2\sigma}J^i(x)$ in that $\tilde{J}^i_{;i}=0$ is $\sigma_{,i}J^i=0$.

Remark: The condition $\sigma_{,i}J^i=0$ implies the current vector J^i is perpendicular to the normal to the hypersurface $\sigma=$constant.

III. APPLICATION TO CONFORMALLY FLAT SOLUTIONS OF EINSTEIN MASSLESS-SCALAR (MESON) FIELD EQUATIONS

As an application of the theory of Sec. II we shall determine symmetries and concomitant conserved current vectors for a class of conformally flat solutions of the Einstein massless-scalar (meson) field equations. Based on the energy–momentum tensor

$$T_{ij}=\psi_{,i}\psi_{,j}-\tfrac{1}{2}\Lambda g_{ij},\quad \Lambda\equiv g^{ij}\psi_{,i}\psi_{,j} \tag{3.1}$$

for a massless scalar field ψ Gürses[12] obtained solutions of the Einstein field equations (1.11) assuming the space–time to be conformally flat. For our application we consider one class of Gürses' solutions where the conformally flat space–time is given by

$$\Phi=\frac{1}{U^2}\sum e_i(dx^i)^2,$$

$$g_{ij}=\frac{e_i\delta_{ij}}{U^2},$$
$$g^{ij}=e_i\delta_{ij}U^2, \tag{3.2}$$
$$e_1=e_2=e_3=-1,\ e_4=1,$$

where,

$$\frac{1}{U^2}\equiv N+\alpha_0,\ N\equiv\alpha_i x^i\neq 0,\ \alpha_i,\alpha_0=\text{constants}. \tag{3.3}$$

For the metric (3.2) the scalar field ψ is given by[13]

$$\psi=\beta_0\ln(\alpha_i x^i+\alpha_0),\quad \beta_0=\text{constant}\neq 0 \tag{3.4}$$

and satisfies the scalar wave equation (massless Klein–Gordon equation)

$$g^{ij}\psi_{;ij}=\psi^i_{;i}=0,$$
$$\psi^i\equiv g^{ij}\psi_j. \tag{3.5}$$

From Eqs. (3.1)–(3.4) we may write

$$\Lambda=\beta_0^2B_0U^6,\ B_0\equiv\sum e_i\alpha_i^2, \tag{3.6}$$

and

$$T_{ij}=A_{ij}U^4,\ A_{ij}\equiv\beta_0^2(\alpha_i\alpha_j-\tfrac{1}{2}B_0e_i\delta_{ij}). \tag{3.7}$$

We shall now determine symmetry vectors ξ^i which satisfy (2.1D) $[\mathcal{L}_\xi(g^{w/2}T_{ij})=0]$ where T_{ij} is given by (3.7). From Eqs. (2.3D) and (3.7) the necessary and sufficient condition for the symmetry mapping may be written in the form

$$A_{im}\xi^m_j+A_{mj}\xi^m_{,i}+A_{ij}Y=0, \tag{3.8}$$

where

$$Y\equiv w\xi^m_{;m}-2U^2Z, \tag{3.9a}$$

$$Z\equiv\alpha_m\xi^m. \tag{3.9b}$$

With reference to (3.1) and (3.6) we consider the two cases (A) $\Lambda=0$ (ψ^i is a null vector), (B) $\Lambda\neq 0$ (ψ^i is not a null vector).

A. The case $\Lambda=0$

From (3.6) and (3.7) we find

$$B_0=\sum e_i\alpha_i^2=0, \tag{3.10}$$

and

$$A_{ij}=\beta_0^2\alpha_i\alpha_j,\ T_{ij}=\beta_0^2\alpha_i\alpha_jU^4. \tag{3.11}$$

For this case $\Lambda=0$, it follows from (3.3) and (3.10) that at least two of the α_i are not zero. By use of (3.9) and (3.11) in (3.8) we obtain

$$\alpha_iZ_{,i}+\alpha_jZ_{,i}+\alpha_i\alpha_jY=0. \tag{3.12}$$

If we put $i=j$ in (3.12) the result may be expressed as

$$2Z_{,i}+\alpha_iY=0. \tag{3.13}$$

If $Z_{,i}$ from (3.13) is used in (3.12) the result is identically zero. This implies that (3.13) is necessary and sufficient for the symmetry mapping.

From (3.13) it follows that

$$\alpha_j Z_{,i} - \alpha_i Z_{,j} = 0. \tag{3.14}$$

The solution to (3.14) is $Z = Z(N)$, where $Z(N)$ is arbitrary and N is given by (3.3). Hence from (3.9) we have

$$Z = \alpha_m \xi^m = Z(N). \tag{3.15}$$

By use of (3.15) we find (3.13) reduces to[14]

$$Y + 2Z' = 0. \tag{3.16}$$

By means of (3.2), (3.3), and (3.9) in (3.16) we obtain

$$w\xi^m_{,m} + 2(w-1)U^2 Z + 2Z' = 0. \tag{3.17}$$

Equations (3.15) and (3.17) are necessary and sufficient conditions which the symmetry vector ξ^i must satisfy. We proceed with their solution.

There is no loss of generality in assuming

$$\alpha_1 = -1, \alpha_2 = \alpha_3 = 0, \alpha_4 = 1, N = x^4 - x^1,$$

$$\frac{1}{U^2} \equiv N + \alpha_0. \tag{3.18}$$

Hence from (3.15) we have

$$\xi^4 - \xi^1 = Z(N), \tag{3.19}$$

and (3.17) takes the form

$$w\xi^2_{,2} = F,$$

$$F(x^1, x^2, x^3, x^4) \equiv -[w(\xi^1_{,1} + \xi^1_{,4} + \xi^3_{,3} + (w+2)Z' + 2(w-1)U^2 Z]. \tag{3.20}$$

Subcase $w \neq 0$: For the subcase $w \neq 0$, the general solution of (3.15) and (3.17) for the symmetry vector ξ^i may be written in the form

$$\xi^1 = \xi^1(x^1, x^2, x^3, x^4),$$

$$\xi^2 = \frac{1}{w}\int F dx^2 + G(x^1, x^3 x^4),$$

$$\xi^3 = \xi^3(x^1, x^2, x^3, x^4), \xi^4 = \xi^1 + Z(N) \ (w \neq 0), \tag{3.21}$$

where ξ^1, ξ^3, G, and Z are arbitrary functions of their indicated arguments.

Subcase $w = 0$: For the subcase[15] $w = 0$, (3.17) reduces to

$$Z' - U^2 Z = 0. \tag{3.22}$$

From the form of U given in (3.18), the solution to (3.22) is

$$Z(N) = c_0(N + \alpha_0), c_0 = \text{arbitrary constant } (w = 0). \tag{3.23}$$

For this case the general solution for ξ^i is expressible as follows

ξ^1, ξ^2, ξ^3 arbitrary functions of x^1, x^2, x^3, x^4,

$$\xi^4 = \xi^1 + c_0(N + \alpha_0) \ (w = 0). \tag{3.24}$$

For the case $\Lambda = 0$ we have from (3.1) that ψ^i defines a null vector from which it follows that $\psi^i_{;k}\psi^k = 0$ and hence ψ^i defines a null geodesic congruence. It also follows from (3.1) that $T = 0$, and hence the conservation law (2.5D) takes the Trautman form[16] $(T^i_j \xi^j)_{;i} = 0$. Thus when $\Lambda = 0$ the conserved current vectors for both $w = 0$ and $w \neq 0$ subcases has the form

$$J^i = T^i_j \xi^j = (\psi_{,j}\xi^j)\psi^i \equiv \rho\psi^i,$$

$$\psi^i \equiv g^{ij}\psi_{,j}, \rho \equiv \psi_{,j}\xi^j = \beta_0 U^2 Z. \tag{3.25}$$

When $w = 0$ (in which case the symmetry vector ξ^j defines a Ricci collineation) we obtain from Eqs. (3.18), (3.23), and (3.25) that $\rho = \beta_0 c_0$ and the conservation law $J^i_{;i} = 0$ reduces to the Klein–Gordon equation (3.5).

When $w \neq 0$, it follows from Eqs. (3.2)–(3.4), (3.18), and (3.25) that the scalar factor ρ and g^{ij} appearing in the conserved current J^i are functions of N which is essentially the phase of the scalar wave ψ and hence the current J^i is a function of the phase of the wave.

Since ψ^i is a null vector it follows from (3.25) that the conserved current J^i is also a null vector. From $\psi^i_{;k}\psi^k = 0$ and $J^i_{;i} = (\psi_{,j}\xi^j\psi^i)_{;i} = \psi_{,ji}\xi^j\psi^i + \xi_{j;i}\psi^j\psi^i + \psi_{,j}\xi^j\psi^i_{;i} = 0$ it follows that

$$J^i_{;k}\psi^k = -J^i g^{jk}\psi_{,jk} = -J^i\psi^k_{;k}. \tag{3.26}$$

By use of the Klein–Gordon equation in the form $g^{jk}\psi_{,jk} = 0$ we have from (3.26) that the conserved null vector is parallel displaced with respect to the null geodesic congruence defined by ψ^i. Alternatively with the Klein–Gordon equation expressed in the form $\psi^k_{;k} = 0$ we may interpret this parallel displacement of the conserved current J^i as a consequence of the kinematical property that the congruence defined by ψ^i has vanishing expansion $\theta \equiv \frac{1}{2}\psi^k_{;k}$.[17] It also follows from Eqs. (3.25), (3.26) that

$$J^i_{;k}J^k = 0. \tag{3.27}$$

Hence J^i also defines a null geodesic congruence.

B. The case $\Lambda \neq 0$ and at least two $\alpha_i \neq 0$

We consider first the case when in the expression for N of (3.3) at least two of the $\alpha_i \neq 0$. In this case there is no loss of generality in assuming $\alpha_2 = \alpha_3 = 0$. From (3.7) we obtain

$$A_{11} = A_{44} = \tfrac{1}{2}\beta_0^2(\alpha_1^2 + \alpha_4^2),$$
$$A_{22} = A_{33} = \tfrac{1}{2}\beta_0^2(\alpha_4^2 - \alpha_1^2)$$
$$A_{14} = \tfrac{1}{2}\beta_0^2\alpha_1\alpha_4, \text{ other } A_{ij} = 0. \tag{3.28}$$

We thus take

$$N = \alpha_1 x^1 + \alpha_4 x^4, \tag{3.29}$$

where from Eq. (3.6) we must have $\alpha^2_4 - \alpha^2_1 \neq 0$.

Subscase $w = 1$: For this choice of N it can be shown that if $w = 1$ a particular solution of the symmetry equation (2.1D) in which the symmetry vector is assumed to be of the form $\xi^i = \xi^i(x^i, x^4)$ is given by

$$\xi^i = b^i - ae_i(A_{i1}x^4 + A_{i4}x^4),$$
$$b^i, a = \text{arbitrary constants.} \tag{3.30}$$

This leads to a five-parameter group of symmetries

$G_5=[P_1, P_2, P_3, P_4, Q]$, where

$$P_C=\partial_C,\ \xi^i_{(C)}=\delta^i_C,$$
$$c=1,\ldots,4,$$
$$Q=-\sum e_i(A_{i1}x^4+A_{i4}x^1)\partial_i, \tag{3.31}$$
$$\xi^i_{(Q)}=-e_i(A_{i1}x^4+A_{i4}x^1).$$

This G_5 contains a $G_4=[P_1, P_2, P_3, P_4]$ subgroup of conformal motions which contains a $G_2[P_2, P_3]$ subgroup of motions.

For the case $w=1$, the conservation law (2.5D) leads to conserved currents J^i of the form

$$J^i=(T^i_j+\tfrac{1}{2}T\delta^i_j)\xi^j. \tag{3.32}$$

The five conserved current vectors concomitant with the five symmetry vectors ξ^i of (3.31) are obtained from (3.32). The currents are proportional to the phase of the ψ waves and take the form

$$J^i_{(C)}=U^6e_iA_{iC},\ C=1,\ldots,4, \tag{3.33}$$
$$J^i_{(Q)}=\tfrac{1}{4}\beta_0^2U^6[x^1\alpha_1\alpha_4(\alpha_1^2-\alpha_4^2)-x^4\alpha_4^2(\alpha_1^2+2\alpha_4^2),0,0,$$
$$x^1\alpha_1^2(\alpha_4^2-2\alpha_1^2)-x^4\alpha_1\alpha_4(\alpha_1^2-\alpha_4^2)], \tag{3.34}$$

where $U^6=(\alpha_1x^1+\alpha_4x^4+\alpha_0)^{-3}$. [See (3.3) and (3.29).]

Remark: It should be noted that the space–time defined by the metric (3.2), (3.3), and (3.29) also admits the motion symmetry defined by $\xi^i=[0,x^3,-x^2,0]$. This motion vector is a solution of the symmetry equation $\mathcal{L}_\xi(g^{w/2}T_{ij})=0$, w arbitrary. (Refer to the last paragraph of Sec. 2.)

C. The case $\Lambda\neq0$ and three $\alpha_i=0$

We continue with the case $\Lambda\neq0$. We now take $\alpha_1=\alpha_2=\alpha_3=0$, $\alpha_4=1$ and assume w is arbitrary. For this choice we have from Eq. (3.2) and (3.3)

$$\frac{1}{U^2}=x^4+\alpha_0,\ g_{ij}=e_i\delta_{ij}(x^4+\alpha_0),$$
$$e_1=e_2=e_3,\ e_4=1. \tag{3.35}$$

From (3.6) we note $B_0=1$ and $\Lambda=\beta_0^2U^6$. For the above choice of α's we also find from (3.7) that

$$A_{ij}=\tfrac{1}{2}\beta_0^2\delta_{ij},$$
$$T_{ij}=\tfrac{1}{2}\beta_0^2\delta_{ij}U^4. \tag{3.36}$$

The symmetry condition (3.8) now takes the form

$$\xi^i_{,j}+\xi^j_{,i}+\delta_{ij}Y=0, \tag{3.37}$$

where [refer to (3.9)]

$$Y=w\xi^m_{,m}-2U^2\xi^4, \tag{3.38a}$$
$$Z=\xi^4. \tag{3.38b}$$

If $i\neq j$ (3.37) reduces to

$$\xi^i_{,j}+\xi^j_{,i}=0,\ i\neq j. \tag{3.39}$$

This equation has the known solution[18]

$$\xi^i=b^i+\omega^i_jx^j+\tau x^i-\tfrac{1}{2}a_iS,\ \omega^i_j=-\omega^j_i,\quad a_0,a_i,b_i,\omega^i_j=\text{constants}, \tag{3.40}$$

where

$$\tau\equiv a_0+a_mx^m,\ S\equiv\sum(x^j)^2. \tag{3.41}$$

If $i=j$ in (3.37) and use be made of (3.40) and (3.41) we obtain

$$Y+2\tau=0. \tag{3.42}$$

From (3.35), (3.38), (3.40), and (3.41) used in (3.42) we obtain

$$3wa_mx^mx^4-\tfrac{1}{2}(w-1)a_4\sum(x^m)^2$$
$$+3wa_0x^4+\alpha_0(2w+1)a_mx^m$$
$$\times(w-1)\omega^4{}_mx^m+\alpha_0(2w+1)a_0$$
$$+b^4(w-1)=0. \tag{3.43}$$

Equation (3.43) must hold identically in the x^i. This implies

$$wa_1=0,\ wa_2=0,\ wa_3=0,\ (w-1)a_4=0,\ (5w+1)a_4=0, \tag{3.44}$$
$$(2w+1)\alpha_0a_4+3wa_0=0,\ (2w+1)\alpha_0a_\beta+(w-1)\omega^4_\beta=0,\ \beta=1,2,3, \tag{3.45}$$
$$(2w+1)\alpha_0a_0+(w-1)b^4=0. \tag{3.46}$$

We consider the three cases $w=1$, $w=0$, $w\neq0,1$.

Subcase $w=1$: In this case (3.44)–(3.46) imply $a_0=a_1=a_2=a_3=a_4=0$. Hence the symmetry vector ξ^i of (3.40) reduces to

$$\xi^i=b^i+\omega^i_jx^j,$$
$$\omega^i_j=-\omega^j_i,\ (w=1), \tag{3.47}$$

where b^i and ω^i_j are then arbitrary constants. It can be shown that the symmetry vector (3.47) generates a ten-parameter group whose generating vectors are

$$\xi^i_{(A)}=\delta^i_A,$$
$$\xi^i_{(AB)}=\delta^i_Bx^A-\delta^i_Ax^B,$$
$$A,B=1,\ldots,4. \tag{3.48}$$

The group symbols are then given by[17] $P_A\equiv\xi^i_{(A)}\partial_i=\partial_A$, $S_{AB}\equiv\xi^i_{(AB)}\partial_i\equiv x^A\partial_B-x^B\partial_A$. The group may be represented in the form $G_{10}=[P_A, S_{AB}]$. It is recognized from (3.47) that this G_{10} is a group of motions in a four-dimensional Euclidean space. This G_{10} is also a group of conformal motions in the conformally flat space defined by (3.35) but with all $e_i=+1$. In the actual space–time (3.35) under consideration this G_{10} is not a group of conformal motions since the three symbols $S_{\alpha4}$, $\alpha=1,2,3$ do not define conformal motions for the metric (3.35). However in the actual space–time (3.35) there exists a seven-parameter conformal motion subgroup of this G_{10}. This G_7 is represented by $[P_A, S_{\alpha\beta}]$, $A=1,\ldots,4$, $\alpha,\beta=1,2,3$. This G_7 in turn contains a subgroup of motions G_6 represented by $[P_\alpha, S_{\alpha\beta}]$.

Corresponding to each of the vectors ξ^i of the 10-parameter group of symmetries defined by (3.48) we obtain [from (2.5D) with $w=1$] ten conserved currents J^i of the form (3.32). These ten current vectors can be arranged into

two sets. The current vectors corresponding to $\xi^i_A=\delta^i_A$ of (3.48) are obtained from (3.32), (3.35), and (3.36) in the form

$$J^i_{(A)}=\tfrac{1}{2}\beta^2_0U^6(e_i-1)\delta^i_A. \tag{3.49}$$

Associated with the six symmetry vectors $\xi^i_{(AB)}$ of (3.48) we have the corresponding current vectors

$$J^i_{(AB)}=\tfrac{1}{2}\beta^2_0U^6(e_i-1)(\delta^i_Bx^A-\delta^i_Ax^B). \tag{3.50}$$

Note from (3.49) that the conservation law corresponding to the conformal motion symmetry $\xi^i_{(4)}=\delta^i_4$ is trivial in that $J^i_{(4)}\equiv 0$. Since $e_4=1$, the fourth component $J^4=0$ for each of the nine remaining nontrivial current vectors $J^i_{(\alpha)}$, $J^i_{(AB)}$, $\alpha=1,2,3$, $A,B=1,...,4$, given by (3.49), (3.50). It thus follows that these nine conserved currents are each spacelike. For convenience we give the expanded form of these nine conserved currents associated with the symmetry $\mathscr{L}_\xi(g^{1/2}T_{ij})=0$

$$J^i_{(1)}=W[1,0,0,0],\ J^i_{(2)}=W[0,1,0,0],$$
$$J^i_{(3)}=W[0,0,1,0], \tag{3.51}$$
$$J^i_{(12)}=W[-x^2,x^1,0,0],\ J^i_{(13)}=W[-x^3,0,x^1,0],$$
$$J^i_{(23)}=W[0,-x^3,x^2,0], \tag{3.52}$$
$$J^i_{(\alpha 4)}=-x^4J^i_{(\alpha)},\ \alpha=1,2,3, \tag{3.53}$$

where $W\equiv-\beta^2{}_0U^6$.

From the discussion following Eq. (3.48) it is seen that the six conserved currents $J^i_{(\alpha)}$, $J^i_{(\alpha\beta)}$, $\alpha=1,2,3$ given by Eqs. (3.51), (3.52) are concomitant with the six motion symmetries $\xi^i_{(\alpha)}$, $\xi^i_{(\alpha\beta)}$ (which comprise the $G_6=[P_\alpha, S_{\alpha\beta}]$). However the remaining three conserved vectors $J^i_{(\alpha 4)}$ of (3.53) are concomitant with the three nonconformal motion symmetries $\xi^i_{(\alpha 4)}$. It is of particular interest to determine further information about the currents which arise from the noncomformal motion symmetries

With this objective in mind we consider the following change in basis of the group G_{10}: $\xi^{*i}_{(A)}=\xi^i_{(A)}$, $\xi^{*i}_{(\alpha\beta)}=\xi^i_{(\alpha\beta)}$, $\xi^{*i}_{(\alpha 4)}=\alpha_0\xi^i_{(\alpha)}-\xi^i_{(\alpha 4)}$. It is evident by inspection that in the new basis the $\xi^{*i}_{(\alpha)}$, $\xi^{*i}_{(\alpha\beta)}$ define the same G_6 of motions, and the $\xi^{*i}_{(\alpha 4)}$ is a noncomformal motion vector. In the new basis the associated currents J^{*i} are given by

$$J^{*i}_{(A)}=J^i_{(A)},$$
$$J^{*i}_{(\alpha\beta)}=J^i_{(\alpha\beta)},$$
$$J^{*i}_{(\alpha 4)}=\frac{1}{U^2}J^i_{(\alpha)}, \tag{3.54}$$

where the $J^{*\,i}_{(\alpha 4)}$ are now the conserved currents concomitant with the nonconformal symmetries. The last equation of (3.54) may be interpreted as a conformal relation between the conserved currents $J^{*\,i}_{(\alpha 4)}$ and $J^i_{(\alpha)}$ in the two conformally related spaced C_4 [defined by (3.35)] and the associated Minkowski space M_4 (with metric $e_i\delta_{ij}$). To see this we note that may consider $J^i_{(\alpha)}$ to be conserved currents in M_4 in that $J^i_{(\alpha),i}=0$,[19] and the conformal factor relating the C_4 Eq. (3.35) and M_4 is $1/U^2$. This conformally related interpretation of $J^{*\,i}_{(\alpha 4)}$ is an illustration of Theorem 2.3 in which we choose $V_4=M_4$, $\tilde{V}_4=C_4$, and $e^{2\sigma}=1/U^2$ [refer to Eq. (3.35)]. Since $\sigma_{,i}J^i_{(\alpha)}=0$, Theorem 2.3 applies and hence $\tilde{J}^i_{(\alpha)}=e^{2\sigma}J^i_{(\alpha)}=(1/U^2)J^i_{(\alpha)}=J^{*\,i}_{(\alpha 4)}$ [see (3.54)] is a conserved current in C_4.

Subcase $w=0$: In this case (3.44)–(3.46) imply $a_4=0$, $b^4=a_0$, $\omega^4_\alpha=a_\alpha$. Hence of the 15 constants appearing in the symmetry vector ξ^i (3.40) there are ten essential constants. These may be chosen to be b^i and ω^i_j so that (3.40) takes the form

$$\xi_i=b^i+\omega^i_jx^j+(b^4+\omega^4_mx^m)x^i-\tfrac{1}{2}\omega^4_iS,$$
$$S\equiv\sum(x^j)^2. \tag{3.55}$$

The vector ξ^i of (3.55) can be shown to generate a ten-parameter group G^+_{10} which can be represented by

$$G^+_{10}=[P_\alpha, W+P_4, S_{\alpha\beta}, V_\alpha], \tag{3.56}$$

where

$$W\equiv x^i\partial_i,\ V_\alpha\equiv x^\alpha W+S_{\alpha 4}-\tfrac{1}{2}SP_\alpha. \tag{3.57}$$

This G^+_{10} contains a seven-parameter subgroup of conformal motions $G^+_7=[P_\alpha, W+P_4, S_{\alpha\beta}]$ which contains a $G^+_6=[P_\alpha, S_{\alpha\beta}]$ of motions.

Since $w\neq 1$, it follows from Theorem 2.1D that in order for the symmetry vector ξ^i (3.55) to induce a conserved current vector of the form defined by (2.5D) the symmetry vector must satisfy $\mathscr{L}_\xi T=0$. A simple calculation shows that of the ten symmetry vectors of G^+_{10} the only ones satisfying this condition are given by G^+_6. Since these define motions (and $w=0$) the associated conserved currents could have been obtained by the Trautman formula (1.4). For the remaining four nonmotion symmetry vectors of this G^+_{10} concomitant conserved currents of the type defined by (2.7D) may be constructed by solving (2.6D) for vectors A^i_D (which always exist locally).

Subcase $w\neq 0,\neq 1$: In this case (3.44)–(3.46) imply that $a_i=0, a_0=0, b^4=0, \omega^4{}_\alpha=0$. The remaining six constants b^α, ω^α_β lead to a six-parameter group of motion G^+_6 mentioned in the $w=1$ case above.

For this case $\mathscr{L}T=0$ and by Theorem 2.1D there will exist six conserved current vectors obtained from (2.5D). In this connection see the last paragraph in Sec. II.

IV. DISCUSSION

We have shown that symmetries of an energy–momentum tensor lead to conserved currents which are expressible in terms of the energy–momentum tensor and the symmetry vector. In general these conserved currents do not follow from the well-known Trautman formulation. Since energy–momentum tensors are of considerable interest in relativistic theories we believe a study of their symmetries and concomitant conserved currents would be of importance. To better understand what new physical insight this new class of symmetries and their concomitant conservation laws may reveal will require additional experience with several well-understood physical systems. Our further work along these lines has indicated that there is a close connection between the existence of the abovementioned symmetries, their concomitant conserved currents, and the kinema-

tical properties of the matter and/or fields associated with the energy-momentum tensor.[17]

APPENDIX A: PROOF OF IDENTITY EQ. (1.17)

Consider a second order tensor $S^{ij}=S^{ji}$ in a Riemannian space V_4. Define

$$S=g_{ij}S^{ij}=g^{ij}S_{ij}. \tag{A1}$$

From (A1) we obtain

$$2\mathcal{L}S=(\mathcal{L}g^{ij})S_{ij}+(\mathcal{L}g_{ij})S^{ij} +g^{ij}\mathcal{L}S_{ij}+g_{ij}\mathcal{L}S^{ij}. \tag{A2}$$

By use of (A2) and

$$\mathcal{L}g^{ij}=-g^{ia}g^{jb}\mathcal{L}g_{ab}, \tag{A3}$$

we obtain

$$2\mathcal{L}S-g_{ij}\mathcal{L}S^{ij}=g^{ij}\mathcal{L}S_{ij}. \tag{A4}$$

From the formula for the Lie derivative of a relative tensor[1] we obtain

$$g_{ij}\mathcal{L}[g^{1/2}(S^{ij}-\tfrac{1}{2}g^{ij}S)] =g^{1/2}g_{ij}[\mathcal{L}(S^{ij}-\tfrac{1}{2}g^{ij}S)+\xi^{k}_{;k}(S^{ij}-\tfrac{1}{2}g^{ij}S]. \tag{A5}$$

Use of (A3) and (A4) in (A5) results in the identity

$$g_{ij}\mathcal{L}[g^{1/2}(S^{ij}-\tfrac{1}{2}g^{ij}S)]\equiv -g^{1/2}g^{ij}\mathcal{L}S_{ij}. \tag{A6}$$

APPENDIX B: PROOF OF IDENTITY EQ. (2.10)

From the definition of the Lie derivative we have

$$\mathcal{L}(g^{1/2}G_{ij})\equiv g^{1/2}[(G_{ij}\xi^{k})_{;k}+G_{kj}\xi^{k}_{;i}+G_{ik}\xi^{k}_{;j}],$$

$$G_{ij}\equiv R_{ij}-\tfrac{1}{2}Rg_{ij}. \tag{B1}$$

By contraction of (B1) with g^{ij} and use of the doubly-contracted Bianchi identity and $\mathcal{L}(g^{1/2}R)\equiv g^{1/2}(R\xi^{j})_{;j}$ we obtain

$$\mathcal{L}(g^{1/2}R)\equiv g^{1/2}(R^{j}_{k}\xi^{k})_{;j}-\tfrac{1}{2}g^{ij}\mathcal{L}[g^{1/2}(R_{ij}-\tfrac{1}{2}Rg_{ij})] \tag{B2}$$

ACKNOWLEDGEMENT

The authors wish to thank Mr. L.K. Norris for his helpful comments.

[1]For general discussion of Lie derivatives see K. Yano, *The Theory of Lie Derivatives and Its Applications* (North-Holland, Amsterdam, 1957).

[2]For a general discussion of symmetries defined by Eq. (1.2) with weight $w=0$ refer to (a) G.H. Katzin, J. Levine, and W.R. Davis, J. Math. Phys. **10**, 617 (1969); (b) G.H. Katzin and J. Levine, Colloquim Mathematicum (Wroclaw, Poland) **26**, 21 (1972).

[3]A. Trautman, "Conservation laws in general relativity," in *Gravitation: An Introduction to Current Research*, edited by L. Witten (Wiley, New York, 1962).

[4]C.D. Collinson, Gen. Rel. Grav. **1**, 137 (1970).

[5]Unless otherwise noted lower-case Latin indices run from 1 to 4, lower-case Greek indices have the range 1,2,3, and the Einstein summation convention will be used. Partial differentiation will be indicated by a comma (,) and covariant differentiation by a semicolon(;).

[6]Ricci collineations are quite general in that they include curvature collineations, affine collineations, homothetic motions, etc. For details see Ref. 2.

[7]We take for the definitions of the Riemann curvature and Ricci tensor those given by L.P. Eisenhart, *Riemannian Geometry* (Princeton U.P., Princeton, N.J.,1960).

[8]In a recent series of papers [(a) D.R. Oliver, Jr. and W.R. Davis, J. Math. Phys. **17**, 1790 (1976); (b) L.H. Green, L.K. Norris, D.R. Oliver, Jr., and W.R. Davis, Gen. Rel. Grav. (in press 1976); (c) L.K. Norris, L.H. Green, and W.R. Davis, J. Math. Phys. **18**, 1305 (1977); (d) W.R. Davis, Nuovo Cimento (in press 1977)] equations of the form $\mathcal{L}R_{ij}=H_{ij}=0$, were studied. These equations were said to define a family of contracted Ricci collineations. The infinitesimal mapping vectors ξ^{i} defined by these equations are also sufficient to formulate conservation laws of the form (1.13).

[9]This equivalence depends upon use of the Einstein field equations (1.11).

[10]A remark similar to that following (2.7U) applies.

[11]The notation tilde, semicolon ($\tilde{;}$) indicates covariant differentiation with respect to the $\tilde{g}_{ij}$ metric.

[12]M. Gürses, Phys. Rev. D **15**, 2731 (1977). See also R.V. Penny, Phys. Rev. D **14**, 910 (1976) and D. Ray, J. Math. Phys. **18**, 1899 (1977).

[13]The constant $\beta_0^2=3/2$. This is based on the choice of $K=1$ in the Einstein field equations usd by Gürses. See Ref. 12.

[14]Primes denote differentiation with respect to the indicated argument.

[15]Note that since $A=0$ implies $T=0$, which in turn implies $R=0$, the case $w=0$ corresponds to a Ricci collineation symmetry. Refer to Eq. (1.9).

[16]Although the conservation law has reduced to the Trautman form (1.4) it should be stressed that the symmetry vector ξ^{i} in general is not a conformal motion as is required for Trautman's formulation of the conservation law.

[17]As a further illustration of the interrelation of the kinematical properties of matter and/or fields and its energy–momentum tensor symmetries and concomitant conserved vector currents consider the energy–momentum tensor $T^{ij}\equiv\rho u^{i}u^{j}$ of noninteracting inchoherent matter [$\rho(x)$ is the matter density and u^{i} is a timelike vector which defines a geodesic congruence]. For the 4-momentum density ρu^{i} to be a symmetry vector in that $\mathcal{L}_{U}(g^{1/2}T^{ij})=0$, $\xi^{i}_{U}\equiv\rho u^{i}$, it can be shown that it is necessary and sufficient that the geodesic congruence defined by the matter 4-velocity u^{i} have vanishing expansion, i.e., $u^{i}_{;i}=0$.

[18]J. Levine, Bull. Amer. Math. Soc. **45**, 766 (1939).

[19]We note that all of the J^{i} [(3.51)–(3.53)] in the original basis ξ^{i} may be considered as conserved currents in M_4 in that they each satisfy $J^{i}_{;i}=0$. A similar interpretation holds for each of the J^{*i} (3.54) expressed in terms of the new basis

Symmetries of the stationary Einstein–Maxwell field equations. III[a)]

William Kinnersley and D. M. Chitre

Department of Physics, Montana State University, Bozeman, Montana 59717
(Received 15 December 1977)

We further investigate the infinite hierarchy of potentials and nonlinear symmetry transformations given in a previous paper. We outline a general method of explicitly calculating the potentials for a given spacetime. We show that some of the transformations can be exponentiated to finite values of their parameters, making them available as a means of generating new solutions. In particular, we show how our transformations may be used to generate all static solutions (including Schwarzschild), starting from nothing but flat space.

1. INTRODUCTION

This paper is the third in a series[1] on the stationary axially symmetric Einstein—Maxwell field. We are considering coupled electromagnetic and gravitational fields in the absence of matter, under the assumption that the fields are independent of two of the coordinates, t and φ. For this situation, we have shown that the reduced Einstein—Maxwell equations possess a remarkably large invariance group **K**, containing an infinite number of parameters.

In paper I, we focused primary attention on two subgroups $\mathbf{G}\subset\mathbf{K}$ and $\mathbf{H}\subset\mathbf{K}$. **G** is the group of linear coordinate transformations in t and φ. **H** is an eight-parameter group isomorphic to SU(2, 1), and consists of certain nonlinear transformations of the field variables. The transformations of **H** are all related to the Killing vector in the t direction (they would continue to exist even if that were the *only* Killing vector[2]). Amalgamation of **G** and **H** yields the entire group **K**.

In paper I, we also began an analysis of **K** using an SU(2, 1) formulation. We found a set of new field variables upon which the action of **H** was linear. The new quantities form SU(2, 1) tensors, and thus the entire problem can be written in a form which is manifestly **H**-covariant. However the resulting equations are somewhat awkward to handle. The SU(2, 1) coupling coefficients f_{ijk} and d_{ijk} appear in the equations, and we constantly had to invoke identities relating products of f's and d's to each other. We now believe that this approach should be saved for the case in which only one Killing vector is present, since then the rest of **K** would not be available. At any rate, linearizing **H** does not lead to an immediate solution of the problem, because now it is the elements of **G** whose transformations become nonlinear.

In paper II, we turned to a **G**-covariant formalism, in which the basic objects of discussion are SL(2, R) tensors, possessing nonlinear transformations under H. We introduced a hierarchy of fields $\overset{n}{\phi}_A$, $\overset{n}{H}_{AB}$, A, $B=1$, 2, $n=1$, 2, $\cdots$, all obeying the simple field equations

$$\nabla\overset{n}{\phi}_A=-i\rho^{-1}f_A{}^X\tilde{\nabla}\overset{n}{\phi}_X, \tag{1.1}$$

$$\nabla\overset{n}{H}_{AB}=-i\rho^{-1}f_A{}^X\tilde{\nabla}\overset{n}{H}_{XB}, \tag{1.2}$$

and a hierarchy of potentials defined by

$$\nabla\overset{mn}{K}=\overset{m}{\varphi}{}^*_X\nabla\overset{n}{\varphi}{}^X, \tag{1.3}$$

$$\nabla\overset{mn}{L}_B=\overset{m}{\varphi}{}^*_X\nabla\overset{n}{H}{}^X{}_B, \tag{1.4}$$

$$\nabla\overset{mn}{M}_A=\overset{m}{H}{}^*_{XA}\nabla\overset{n}{\varphi}{}^X, \tag{1.5}$$

$$\nabla\overset{mn}{N}_{AB}=\overset{m}{H}{}^*_{XA}\nabla\overset{n}{H}{}^X{}_B. \tag{1.6}$$

(For further details, please see II.) We then showed that under the action of all infinitesimal elements of **K**, these objects go into themselves, again via nonlinear transformations.

In the present paper we have several objectives. First, in order to help clarify the meaning behind our approach, we will show how the quantities we use can be related to certain components of familiar four-dimensional objects. Also we will examine what gauge freedom is present in the definition of the potentials. We will show that for many of the generators of **K**, it is possible to perform explicit "exponentiation." Thus, some of the transformations can be written down in closed form for finite values of their parameters. This makes them immediately available for practical purposes in generating new Einstein—Maxwell solutions.

We next discuss a new method of calculating the entire set of potentials for several simple spacetimes, starting with Minkowski space. We then see what the finite transformations of **K** do to those spacetimes. One motivation for carrying out such an exercise is to try to understand how **K** affects asymptotic flatness. So far we cannot claim to have found a nontrivial element of **K** which always preserves asymptotic flatness. In fact, on the face of it, **K** would appear to be *the group which ruins asymptotic flatness in all possible ways*! However there is a bright side to this, and **K** has already proved useful: one can "cure" spacetimes which were not asymptotically flat to begin with.[3] (Actually, we have every reason to believe that "good" transformations *can* be constructed from elements of **K** which are infinite linear combinations of our simple ones.)

Finally, we will turn to static metrics, a case which of course is already well understood. But we ask to what extent **K** relates the various static solutions to each other. Using identities which hold only in the static case, we can write more transformations of **K** in closed form. We find that *the* $\overset{k}{\gamma}_{12}$*'s generate only other static solutions, and together they suffice to generate all static solutions from any one.*

a)Supported by National Science Foundation Grant PHY 76-12246.

 0022-2488/78/1909-1926$1.00

Geroch once conjectured[4] that the complete group K would suffice to generate all stationary solutions. We have taken the first steps toward deciding whether that conjecture is actually true.

2. FOUR-DIMENSIONAL FORMULATION

We would like to understand why the Einstein–Maxwell equations, which are relatively complicated, can be reduced to such a simple form as Eqs. (1.1), (1.2). Consider first the Maxwell equations. In terms of the usual 4-potential $F_{\mu\nu}=A_{\nu,\mu}-A_{\mu,\nu}$, half of Maxwell's equations are satisfied identically, and the remaining ones are

$$\Box A_\mu = 0\,. \tag{2.1}$$

Conversely the second half may be satisfied by introducing another 4-potential B_μ related to $F^*_{\mu\nu}$. Writing the identities

$$F_{\mu\nu} = -\tfrac{1}{2}\epsilon_{\mu\nu}{}^{\sigma\tau}F^*_{\sigma\tau} \tag{2.2}$$

$$F^*_{\mu\nu} = \tfrac{1}{2}\epsilon_{\mu\nu}{}^{\sigma\tau}F_{\sigma\tau} \tag{2.3}$$

in terms of the potentials

$$(A_{\nu,\mu}-A_{\mu,\nu}) = -\tfrac{1}{2}\epsilon_{\mu\nu}{}^{\sigma\tau}(B_{\tau,\sigma}-B_{\sigma,\tau}), \tag{2.4}$$

$$(B_{\nu,\mu}-B_{\mu,\nu}) = \tfrac{1}{2}\epsilon_{\mu\nu}{}^{\sigma\tau}(A_{\tau,\sigma}-A_{\sigma,\tau})\,, \tag{2.5}$$

and adding Eq. (2.4) to i times Eq. (2.5), gives exactly our formulation, Eq. (1.1). By introducing the redundant potential B_μ we have doubled the number of variables, and replaced Eq. (2.1) by a set of first-order equations.

Now consider the Einstein equations, and let $\overset{a}{\xi}_\mu$, $a=1$, 2 be the two Killing vectors, From each Killing vector we can construct the Killing bivector

$$\overset{a}{\xi}_{\mu\nu} = \overset{a}{\xi}_{\nu;\mu} - \overset{a}{\xi}_{\mu;\nu}\,. \tag{2.6}$$

It is well known[5] that in vacuum, $\overset{a}{\xi}_{\mu\nu}$ satisfies Maxwell's equations,

$$\overset{a}{\xi}{}^{\mu\nu}{}_{;\nu} = R^{\mu\nu}\overset{a}{\xi}_\nu = 0 \tag{2.7}$$

and that this is equivalent to a portion of Einstein's equations. As a consequence, the dual bivector, $\overset{a}{\xi}{}^*_{\mu\nu}$ is also derivable from a vector potential, say

$$\overset{a}{\xi}{}^*_{\mu\nu} = \overset{a}{\eta}_{\nu;\mu} - \overset{a}{\eta}_{\mu;\nu}\,. \tag{2.8}$$

By repeating the same reasoning employed in Eqs. (2.1)–(2.4), and identifying the components of $\overset{a}{\xi}_\mu + i\overset{a}{\eta}_\mu$ with H_{AB}, we are led to Eq. (1.2). Thus each of our fields $\overset{n}{\phi}_A$, $\overset{n}{H}_{AB}$ corresponds, in four-dimensional terms, to a *bivector solution of Maxwell's equations*.

Next, consider the first one of our sequence of potentials,

$$\nabla K = \varphi^*_X \nabla \varphi^X\,. \tag{2.9}$$

As shown in II, the imaginary part of K is algebraically expressible in terms of φ_A, so consider just the real part. The electromagnetic invariant $F^*_{\mu\nu}F^{\mu\nu}$ can be written in two ways in terms of the potentials

$$\begin{aligned} F^*_{\mu\nu}F^{\mu\nu} &= \tfrac{1}{2}\epsilon^{\mu\nu\sigma\tau}A_{\nu;\mu}A_{\tau;\sigma} \\ &= -\tfrac{1}{2}\epsilon^{\mu\nu\sigma\tau}B_{\nu;\mu}B_{\tau;\sigma}\,. \end{aligned}$$

Hence

$$\epsilon^{\mu\nu\sigma\tau}(A_{\nu;\mu}A_{\tau;\sigma} + B_{\nu;\mu}B_{\tau;\sigma}) = 0$$

which is expressible as a total divergence

$$J^\sigma{}_{;\sigma} = 0\,, \tag{2.10}$$

where

$$J^\sigma = \epsilon^{\mu\nu\sigma\tau}(A_{\nu;\mu}A_\tau + B_{\nu;\mu}B_\tau)\,. \tag{2.11}$$

Equation (2.10) implies the existence of a bivector $K^{\sigma\tau}$ such that

$$J^\sigma = K^{\sigma\tau}{}_{;\tau}\,. \tag{2.12}$$

Comparing Eqs. (2.9), (2.11), and (2.12), we see that our potential K corresponds to the particular component K^{34}.

Likewise, the other potentials we use may be derived by replacing $F_{\mu\nu}$ with $\overset{a}{\xi}_{\mu\nu}$ in the above discussion.

3. GAUGE TRANSFORMATIONS

Each of the potentials was defined by Eqs. (II.2.10)–(II.2.13) only up to an arbitrary integration constant. One might therefore want to consider the effect of gauge transformations which change these constants. However not all of the potentials are algebraically independent. We have chosen to impose certain relations between them, Eqs. (II.2.16)–(II.2.22), and these amount to constraints among some of the constants. To see how many potentials remain independent, arrange them in towers, e.g.,

$$\begin{array}{ccc} \overset{01}{N}_{AB} & & \\ \overset{02}{N}_{AB}, & \overset{11}{N}_{AB} & \\ \overset{03}{N}_{AB}, & \overset{12}{N}_{AB}, & \overset{21}{N}_{AB} \\ \cdots & \cdots & \cdots \end{array}$$

with $m+n \equiv q = \text{const}$ on the qth level. Given all the potentials in the first q levels, *one* new potential of each type (say $\overset{1q}{K}$, $\overset{1q}{L}_B$, $\overset{1q}{M}_A$, $\overset{1q}{N}_{AB}$) is sufficient to determine the others on the $(q+1)$st level, via Eqs. (II.2.19)–(II.2.22). The few remaining constants are further restricted by Eqs. (II.2.16)–(II.2.18). All we can do at last is to add a real constant σ to K, a complex constant c_A to L_A and M_B, and a Hermitian constant γ_{AB} to N_{AB}. But these gauge transformations are nothing new. They are precisely the ones which already occur in K, with one exception. In K, γ_{AB} was real and symmetric. The new part now permitted will be imaginary and antisymmetric,

$$\gamma_{AB} = i\tau\epsilon_{AB}\,. \tag{3.1}$$

The new transformation is

$$\begin{aligned} \overset{mn}{K} &\to \overset{mn}{K} + 2i\overset{k}{\tau}\overset{m+k,\,n}{K} - 2i\overset{m,\,n+k}{\tau K} - 4\overset{k}{\tau}\Big(\sum_s \overset{ms\;k-s+1,\,n}{K\;K}\Big) \\ &\quad + i\overset{k}{\tau}\Big(\sum_s \overset{ms\;k-s,\,n}{L_X M^X}\Big), \\ \overset{mn}{L_B} &\to \overset{mn}{L_B} + 2i\overset{k}{\tau}\overset{m+k,\,n}{L_B} + i\overset{m,\,n+k}{\tau L_B} - 4\overset{k}{\tau}\Big(\sum_s \overset{ms\;k-s+1,\,n}{K\;L_B}\Big) \\ &\quad + i\overset{k}{\tau}\Big(\sum_s \overset{ms\;k-s,\,n}{L_X N^X{}_B}\Big), \\ \overset{mn}{M_A} &\to \overset{mn}{M_A} - i\overset{k}{\tau}\overset{m+k,\,n}{M_A} - 2i\overset{m,\,n+k}{\tau M_A} - 4\overset{k}{\tau}\Big(\sum_s \overset{ms\;k-s+1,\,n}{M_A K}\Big) \\ &\quad + i\overset{k}{\tau}\Big(\sum_s \overset{ms\;k-s,\,n}{N_{AX} M^X}\Big). \end{aligned} \tag{3.2}$$

$$N_{AB}^{mn} \to N_{AB}^{mn} - i\,{}^{k}_{\tau} N_{AB}^{m+k,\,n} + i\,{}^{k}_{\tau} N_{AB}^{m,\,n+k}$$
$$- 4\,{}^{k}_{\tau}\left(\sum_s M_A^{ms}\, L_B^{k-s+1,\,n}\right) + i\,{}^{k}_{\tau}\left(\sum_s N_{AX}^{ms}\, N^{X}{}_{B}^{k-s,\,n}\right).$$

We should add a word of explanation why τ was originally overlooked in II. By repeated application of the recursion relations, Eqs. (II.2.19)—(II.2.22), one can prove that for $k \geq 1$,

$$\begin{aligned}
K^{m,\,n+k} - K^{m+k,\,n} &= \sum_s \left(2iK^{ms}\, K^{k-s+1,\,n} - L^{X\,ms}\, M^{X\,k-s,\,n}\right),\\
L_B^{m,\,n+k} - L_B^{m+k,\,n} &= \sum_s \left(2iK^{ms}\, L_B^{k-s+1,\,n} - L_X^{ms} N^{X\,k-s,\,n}{}_{B}\right),\\
M_A^{m,\,n+k} - M_A^{m+k,\,n} &= \sum_s \left(2iM_A^{ms}\, K^{k-s+1,\,n} - N_{AX}^{ms} M^{X\,k-s,\,n}\right),\\
N_{AB}^{m,\,n+k} - N_{AB}^{m+k,\,n} &= \sum_s \left(2iM_A^{ms}\, L_B^{k-s+1,\,n} - N_{AX}^{ms} N^{X\,k-s,\,n}{}_{B}\right).
\end{aligned} \tag{3.3}$$

Inserting this in Eq. (II.3.3) one finds that ${}^{k}_{\sigma}$ and ${}^{k}_{\tau}$ are equivalent, *but only for* $k \geq 1$. For $k \leq 0$ the quadratic terms in ${}^{k}_{\sigma}$, ${}^{k}_{\tau}$ vanish, leaving linear terms. For $k = 0$ even the linear terms coincide, but for $k \leq -1$, ${}^{k}_{\sigma}$ and ${}^{k}_{\tau}$ are different.

4. FINITE TRANSFORMATIONS

The finite transformations of **K** can be obtained from the infinitesimal ones given in paper II by "exponentiation." However this process is not completely straightforward for the group we are dealing with, since the infinitesimal transformations are nonlinear in the potentials. The exponentiation will be an infinite power series in the group parameter, with coefficients of ever increasing complexity. The transformation may eventually involve the *entire* infinite set of potentials (or possibly an infinite subset of them.) In that case, there would seem to be little hope of summing the series to get a useful closed form.[6] That is exactly what does happen for the transformations ${}^{k}_{\gamma 12}$ and ${}^{k}_{\sigma}$. Fortunately, the other transformations ${}^{k}_{\gamma 11}$, ${}^{k}_{\gamma 22}$, ${}^{k}_{c 1}$, ${}^{k}_{c 2}$ each involve only a finite number of potentials at once, and we have succeeded in writing some of them down in closed form.

From Eq. (II.3.1), the infinitesimal transformation for ${}^{k}_{\gamma 22}$ is

$$N_{11}^{mn} \to N_{11}^{mn} + \gamma \sum_s N_{11}^{ms}\, N_{11}^{k-s,\,n}, \tag{4.1}$$

$$N_{12}^{mn} \to N_{12}^{mn} - \gamma N_{11}^{m,\,n+k} + \gamma \sum_s N_{11}^{ms} N_{12}^{k-s,\,n}, \tag{4.2}$$

$$N_{21}^{mn} \to N_{21}^{mn} - \gamma N_{11}^{m+k,n} + \gamma \sum_s N_{21}^{ms}\, N_{11}^{k-s,\,n}, \tag{4.3}$$

$$N_{22}^{mn} \to N_{22}^{mn} - \gamma N_{12}^{m+k,\,n} - \gamma N_{21}^{m,\,n+k} + \gamma \sum_s N_{21}^{ms} N_{12}^{k-s,\,n}, \tag{4.4}$$

$$M_1^{mn} \to M_1^{mn} + \gamma \sum_s N_{11}^{ms} M_1^{k-s,\,n}, \tag{4.5}$$

$$M_2^{mn} \to M_2^{mn} - \gamma M_1^{m+k,\,n} + \gamma \sum_s N_{21}^{ms} M_1^{k-s,\,n}, \tag{4.6}$$

$$L_1^{mn} \to L_1^{mn} + \gamma \sum_s L_1^{ms} N_{11}^{k-s,\,n}, \tag{4.7}$$

$$L_2^{mn} \to L_2^{mn} - \gamma L_1^{m,\,n+k} + \gamma\left(\sum_s L_1^{ms} N_{12}^{k-s,\,n}\right), \tag{4.8}$$

$$K^{mn} \to K^{mn} + \gamma\left(\sum_s L_1^{ms} M_1^{k-s,\,n}\right). \tag{4.9}$$

We see from the summations that it is convenient to treat m, n as matrix indices. For given k, we define matrices N_{11}, N_{12}, L_1, etc. as

$$(N_{11})_{mn} \equiv N_{11}^{k-m,\,n}$$

and so on. Note that $n = 1, 2, \ldots, \infty$, $m = k, k-1, \ldots, 0, -1, \ldots, -\infty$ so that these matrices are infinite. However the sums run only over $s = 1, \ldots, k$. It turns out to be possible to write the following results in a form in which *only the square submatrices* $m, n = 1, \ldots, k$ *actually participate in the matrix operations.* Indices outside this range can occur only as unsummed indices, at the beginning and end of each term.

To illustrate this, consider Eq. (4.1), which is

$$N_{11} \to N_{11} + \gamma N_{11} N_{11}.$$

The exponentiation is

$$\begin{aligned}
N_{11} &\to N_{11} + \gamma N_{11} N_{11} + \gamma^2 N_{11} N_{11} N_{11} + \cdots\\
&= \gamma^{-1}[(I - \gamma N_{11})^{-1} - I].
\end{aligned}$$

As written, this expression is meaningful and valid only for $m, n = 1, \ldots, k$. However, written in the equivalent form

$$N_{11} \to N_{11} + \gamma N_{11}(I - \gamma N_{11})^{-1} N_{11} \tag{4.10}$$

it is valid for all m, n. The inverse still involves only a $k \times k$ matrix. (To check such a result, one need only show that it forms a one-parameter group, and has the right infinitesimal limit.) The finite transformation for $k = 2$ written out in complete detail is

$$N_{11}^{mn} \to N_{11}^{mn} + \gamma \frac{N_{11}^{m1}(1 - \gamma N_{11}^{02}) N_{11}^{1n} + N_{11}^{m2}(\gamma N_{11}^{01}) N_{11}^{1n} + N_{11}^{m1}(\gamma N_{11}^{12}) N_{11}^{0n} + N_{11}^{m2}(1 - \gamma N_{11}^{11}) N_{11}^{0n}}{1 - \gamma(N_{11}^{11} + N_{11}^{02}) + \gamma^2(N_{11}^{02} N_{11}^{11} - N_{11}^{12} N_{11}^{01})}. \tag{4.11}$$

For general k, the transformed quantity of particular interest is N_{11}^{01}:

$$N_{11}^{01} \to P_{k-2}/Q_k, \tag{4.12}$$

where P_k, Q_k are polynomials in γ of degree k:

$$\begin{aligned}
P_k &= N_{11}^{01} + \gamma(\quad) + \cdots + \gamma^k \Delta_{k+1},\\
Q_k &= 1 + \gamma(\quad) + \cdots + \gamma^k \Delta_k,\\
\Delta_k &= \det(N_{11}).
\end{aligned} \tag{4.13}$$

The transformations for all the other potentials can be written in a similar notation. Let

$$\begin{aligned}
('N_{11})_{mn} &= N_{11}^{2k-m,\,n},\\
(N'_{11})_{mn} &= N_{11}^{k-m,\,n+k},\\
('N'_{11})_{mn} &= N_{11}^{2k-m,\,n+k}.
\end{aligned} \tag{4.14}$$

We have obtained the following results:

$$
\begin{aligned}
&N_{12} \to (N_{12} - \gamma N'_{11}) + \gamma(N_{11})(I - \gamma N_{11})^{-1}(N_{12} - \gamma N'_{11}),\\
&N_{21} \to (N_{21} - \gamma' N_{11}) + \gamma(N_{21} - \gamma' N_{11})(I - \gamma N_{11})^{-1}(N_{11}),\\
&N_{22} \to (N_{22} - \gamma' N_{12} - \gamma N'_{21} + \gamma^2 N'_{11})\\
&\qquad + \gamma(N_{21} - \gamma' N_{11})(I - \gamma N_{11})^{-1}(N_{12} - \gamma N'_{11}),\\
&M_1 \to M_1 + \gamma(N_{11})(I - \gamma N_{11})^{-1}(M_1),\\
&M_2 \to (M_2 - \gamma' M_1) + \gamma(N_{21} - \gamma' N_{11})(I - \gamma N_{11})^{-1}(M_1),\\
&L_1 \to L_1 + \gamma(L_1)(I - \gamma N_{11})^{-1}(N_{11}),\\
&L_2 \to (L_2 - \gamma L'_1) + \gamma(L_1)(I - \gamma N_{11})^{-1}(N_{12} - \gamma N'_{11}),\\
&K \to K + \gamma(L_1)(I - \gamma N_{11})^{-1}(M_1).
\end{aligned}
\tag{4.15}
$$

The transformations $\overset{k}{\gamma}_{11}$ may be written down also, simply by interchanging indices, $1 \leftrightarrow 2$. Likewise, any transformation for which $\overset{k}{\gamma}_{AB}$ is a *null* parameter ($\overset{kXY}{\gamma}\overset{k}{\gamma}_{XY} = 0$) may be just as easily exponentiated. However the $\overset{k}{\gamma}_{12}$ transformation is not null in this sense, and cannot be included in the results. We postpone its discussion until Sec. 7.

5. CALCULATION OF POTENTIALS

We will now discuss a method for directly calculating the hierarchy of potentials for some particular spacetime. As indicated in Sec. 2, it is sufficient to calculate just $\overset{n}{H}_{AB}$. The two relevant equations are

$$\nabla \overset{1n}{N}_{AB} = H^*_{XA} \nabla \overset{n}{H}{}^X{}_B, \tag{5.1}$$

$$\overset{n+1}{H}_{AB} = i(\overset{1n}{N}_{AB} + H_{AX}\overset{n}{H}{}^X{}_B). \tag{5.2}$$

Eliminating $\overset{1n}{N}_{AB}$, we get

$$\nabla \overset{n+1}{H}_{AB} = i(H_{AX} + H^*_{XA})\nabla \overset{n}{H}{}^X{}_B + i\overset{n}{H}{}^X{}_B \nabla H_{AX} \tag{5.3}$$

which is the same as Eq. (II.4.3). To find the solution we introduce a generating function

$$F_{AB}(t) = \sum_{n=0}^{\infty} t^n \overset{n}{H}_{AB}. \tag{5.4}$$

Multiplying Eq. (5.3) by t^{n+1} and summing, we get

$$\nabla F_{AB} = it[(H_{AX} + H^*_{XA})\nabla F^X{}_B + F^X{}_B \nabla H_{AX}]$$

or

$$[\delta^A{}_X - it(H^A{}_X + H^{*A}_X)]\nabla F^X{}_B = it(\nabla H^A{}_X)F^X{}_B. \tag{5.5}$$

Denote the bracket on the lhs by $G^A{}_B(t)$. Since $iF^A{}_B(0) = G^A{}_B(0) = \delta^A{}_B$, the matrix inverses of F and G exist (at least as formal power series in the neighborhood of $t = 0$). Using matrix notation, Eq. (5.5) may therefore be rewritten as

$$(\nabla F)F^{-1} = itG^{-1}(\nabla H). \tag{5.6}$$

Using the explicit form of G, and the field equation, Eq. (1.2), this may be further simplified to

$$(\nabla F)F^{-1} = itS^{-2}[\nabla H - 2t(z\nabla H + \rho\tilde{\nabla}H)], \tag{5.7}$$

where

$$
\begin{aligned}
S^2 &= \det G\\
&= 1 - 2it(H^X{}_X + H^*{}_X{}^X) + t^2(H^{XY} + H^{*YX})(H_{XY} + H^*{}_{YX})\\
&= 1 - 4tz + 4t^2(\rho^2 + z^2).
\end{aligned}
\tag{5.8}
$$

The procedure for a particular spacetime will be to first calculate H_{AB} the integration of Eq. (II.3.14). From H_{AB}, the rhs of Eq. (5.7) may be determined. Further integration then yields $F(t)$, which may be expanded to produce the hierarchy.

In fact, a general "first integral" of Eq. (5.7) exists, and leads us to an important identity. Let $\beta = \det F$. Taking the trace of Eq. (5.7), we obtain

$$
\begin{aligned}
&\beta^{-1}\nabla\beta = -S^{-1}\nabla S\\
&\quad \Rightarrow \beta = c(t)\,S^{-1},
\end{aligned}
\tag{5.9}
$$

where $c(t)$ is an integration constant. Now arbitrary integration constants do appear in each H_{AB}, as discussed in Sec. 2. We are therefore free to choose $c(t)$ to be as simple as possible, consistent with the first two terms, $\overset{0}{H}_{AB}$ and $\overset{1}{H}_{AB}$. Using the definition of F, Eq. (5.4),

$$\beta = -1 - 2tz + O(t^2).$$

Hence we can choose $c(t) = -1$.

Expanding in powers of t,

$$
\begin{aligned}
S^{-1} &= \sum (2rt)^n P_n(\cos\theta),\\
\beta &= \sum t^{m+n}(\overset{m}{H}{}^1{}_1\overset{n}{H}{}^2{}_2 - \overset{m}{H}{}^1{}_2\overset{n}{H}{}^2{}_1)\\
&= \tfrac{1}{2}\sum t^n\left(\sum \overset{n-k}{H}{}^{XY}\overset{k}{H}_{XY}\right),
\end{aligned}
$$

where

$$r = (\rho^2 + z^2)^{1/2} \quad \cos\theta = z/r.$$

We therefore find the identity

$$\sum_k \overset{n-k}{H}{}^{XY}\overset{k}{H}_{XY} = -2(2r)^n P_n(\cos\theta) \tag{5.10}$$

which will find application in Sec. 7.

As an example, we now complete the calculation for flat space. We have

$$
H = \begin{bmatrix} 0 & -\rho^2 \\ -1 & -2iz \end{bmatrix},
$$

$$
(\nabla F)F^{-1} = itS^{-2}\begin{pmatrix} 0 & -2\rho\nabla\rho + 4t\rho(z\nabla\rho - \rho\nabla z) \\ 0 & -2i\nabla z + 4it(\rho\nabla\rho + z\nabla z) \end{pmatrix}.
$$

To solve, let

$$
F = \begin{bmatrix} A & B \\ C & D \end{bmatrix}, \quad AD - BC = \beta,
$$

$$
(\nabla F)F^{-1} = \beta^{-1}\begin{pmatrix} D\nabla A - C\nabla B & A\nabla B - B\nabla A \\ D\nabla C - C\nabla D & A\nabla D - B\nabla C \end{pmatrix}.
$$

Equating elements in the first column, we find necessary conditions:

$$C = aD, \quad A = aB + b, \quad bD = \beta. \tag{5.11}$$

To first order,

$$
F = \begin{pmatrix} -i & -t\rho^2 \\ -t & -i - 2itz \end{pmatrix} + O(t^2) \tag{5.12}
$$

so the simplest choice is

$$a = -it; \quad b = -i. \tag{5.13}$$

The elements in the second column now yield the solution

$$B = -(z - \tfrac{1}{2}t^{-1})S^{-1} + \tfrac{1}{2}it^{-1}. \tag{5.14}$$

Expansion of A, B, C, D in powers of t leads to the flat-space potentials

$$
\begin{aligned}
&\overset{n}{H}_{11} = (2r)^{n-1} P_{n-1}(\cos\theta),\\
&\overset{n}{H}_{12} = i(2r)^{n} P_{n}(\cos\theta)\\
&\overset{n}{H}_{21} = \tfrac{1}{2} i n^{-1} (2r)^{n} \sin\theta P^{1}_{n-1}(\cos\theta),\\
&\overset{n}{H}_{22} = -\tfrac{1}{2}(n+1)^{-1}(2r)^{n+1}\sin\theta P^{1}_{n}(\cos\theta).
\end{aligned}
\tag{5.15}
$$

These are equivalent to the forms given in Eqs. (II.4.7), (II.4.16), (II.4.17), but the functions involved are much more familiar.

In fact, the $\overset{n}{H}_{AB}$ have turned out to be solutions of Laplace's equation, for which there is a ready explanation. As a general consequence of Eq. (1.2),

$$
\nabla\cdot(\rho^{-2} f_A{}^X \nabla \overset{n}{H}_{XB}) = 0. \tag{5.16}
$$

This reduces to Laplace's equation for flat space. The same equations are also obeyed by $F_{AB}(l)$ itself:

$$
\begin{aligned}
&\nabla F_{AB} = -i\rho^{-1} f_A{}^X \tilde{\nabla} F_{XB}\\
&\nabla\cdot(\rho^{-2} f_A{}^X \nabla F_{XB}) = 0.
\end{aligned}
\tag{5.17}
$$

One more curious result involving flat-space potentials might be mentioned. When the $\overset{mn}{N}_{11}$ are calculated, one finds that $\overset{m1}{N}_{11}$ with $m \geq 1$ all vanish.

As an easy generalization, one can also derive the potentials for static cylindrical metrics,

$$
f_{11} = \rho^{2\alpha};\quad f_{12} = 0;\quad f_{22} = -\rho^{2-2\alpha}, \tag{5.18}
$$

where α is an arbitrary constant. We find

$$
\begin{aligned}
&\overset{n}{H}_{11} = (2\alpha)!_{n+\alpha-1} C_{n-1} (2r)^{n-1} (\sin\theta)^{\alpha} P^{-\alpha}_{n-1+\alpha}(\cos\theta),\\
&\overset{n}{H}_{12} = i_{n-\alpha} C_{n} (2r)^{n} (\sin\theta)^{\alpha} P^{\alpha}_{n-\alpha}(\cos\theta),\\
&\overset{n}{H}_{21} = -\tfrac{1}{2} i (n+\alpha)^{-1} (2\alpha)!_{n+\alpha} C_{n} (2r)^{n} (\sin\theta)^{1-\alpha}\ P^{1-\alpha}_{n-1+\alpha}(\cos\theta),\\
&\overset{n}{H}_{22} = -\tfrac{1}{2}{}_{n-\alpha} C_{n-1} (2r)^{n+1} (\sin\theta)^{1-\alpha} P^{\alpha-1}_{n-\alpha}(\cos\theta),
\end{aligned}
\tag{5.19}
$$

where ${}_nC_m$ is the binomial coefficient. Potentials for the Schwarzschild, Kerr, and Tomimatsu–Sato solutions will be derived in paper IV of this series.[7]

6. ASYMPTOTIC BEHAVIOR

We now ask what happens when the transformations of **K** are applied to flat space. For the infinitesimal transformations we find

$$
\begin{aligned}
&\overset{k}{\gamma}_{11}: H_{11} \to 1 + i\gamma (2r)^{k+1} P_{k+1}(\cos\theta),\\
&\overset{k}{\gamma}_{12}: H_{11} \to 1 - 2\gamma (2r)^{k} P_{k}(\cos\theta),\\
&\overset{k}{\gamma}_{22}: H_{11} \to 1 - i\gamma (2r)^{k-1} P_{k-1}(\cos\theta).
\end{aligned}
\tag{6.1}
$$

The weak gravitational fields thus generated are static and stationary multipoles of all orders, but of the "wrong" variety. They are all inner solutions, in the sense that the sources which produce them lie at infinity. None of them are asymptotically flat. However, as was pointed out in II, the asymptotic behavior resulting from a *finite* transformation may be entirely different.

Using the results of Sec. 4, we find

$$
\overset{1}{\gamma}_{22}: H_{11} \to (1 + i\gamma)^{-1}, \tag{6.2}
$$

$$
\overset{2}{\gamma}_{22}: H_{11} \to (1 + 2i\gamma z - \gamma^2\rho^2)^{-1}, \tag{6.3}
$$

$$
\overset{3}{\gamma}_{22}: H_{11} \to \frac{1 - i\gamma\rho^2}{1 + i\gamma(4z^2 - 3\rho^2) - \gamma^2(\rho^4 + 12\rho^2 z^2) - i\gamma^3\rho^6}. \tag{6.4}
$$

For general k, the structure is

$$
\overset{k}{\gamma}_{22}: H_{11} \to i P_{k-2}/Q_k, \tag{6.5}
$$

where P_k, Q_k are polynomials of degree k in the argument $i\gamma\rho^{k+1}$, with coefficients that are functions of z/ρ. Note in the examples above that the leading coefficients are particularly simple. It would appear that for flat space the determinant Δ_k defined in Eq. (4.13) has the value

$$
\Delta_k = (-i\rho^{k-1})^k \tag{6.6}
$$

although we have not been able to prove this in general.

In the asymptotic limit $\rho \to \infty$, the leading terms dominate, and

$$
H_{11} \approx (-)^{k+1} \gamma^{-2} \rho^{-2(k-1)}. \tag{6.7}
$$

Thus the new solution is asymptotically a static cylinder metric, and not asymptotically flat.

What if we started with a cylinder metric? The generalization is straightforward and we find

$$
H_{11} \to i\rho^{\alpha} P_{k-2}/Q_k,
$$

where P_k, Q_k are now polynomials in $i\gamma\rho^{k-\alpha+1}$, and

$$
\Delta_k = (-i\rho^{k+\alpha-1})^k
$$

In the limit $\rho \to \infty$, two cases must be distinguished. For $k + \alpha - 1 > 0$,

$$
H_{11} \approx (-)^{k+1} \gamma^{-2} \rho^{-2(k+\alpha-1)}
$$

but for $k + \alpha - 1 \leq 0$, the leading terms no longer dominate, and

$$
H_{11} \approx \rho^{\alpha}.
$$

In either case, the metric generated is asymptotically a static cylinder, not asymptotically flat.

7. STATIC METRICS

We now ask what general conclusions can be drawn if the metric under consideration is static. We would have

$$
f_{11} = f,\quad f_{12} = 0,\quad f_{22} = -\rho^2 f^{-1} \tag{7.1}
$$

and from Eq. (I.3.14) it would follow that

$$
\psi_{11} = \psi_{22} = 0. \tag{7.2}
$$

Thus H_{11}, H_{22} become purely real, and H_{12}, H_{21} purely imaginary. Examination shows this to hold for all of the potentials. In matrix form,

$$
\begin{aligned}
&\overset{n}{H} \sim \begin{pmatrix} \text{Real} & \text{Imag} \\ \text{Imag} & \text{Real} \end{pmatrix},\\
&\overset{mn}{N} \sim \begin{pmatrix} \text{Imag} & \text{Real} \\ \text{Real} & \text{Imag} \end{pmatrix}.
\end{aligned}
\tag{7.3}
$$

The number of independent functions is thus effectively reduced by half. A few interesting identities follow at once from this. For example, using Eq. (II.2.18) we can show that some potentials factor:

$$
\begin{aligned}
&\overset{nn}{N}_{11} = -\overset{0n}{N}_{11}\overset{0n}{N}_{21},\\
&\overset{nn}{N}_{22} = \overset{0n}{N}_{22}\overset{0n}{N}_{12}.
\end{aligned}
$$

The main point of interest, however, will be to obtain in closed form the action of $\overset{k}{\gamma}_{12}$. From Eq. (II.3.1), the infinitesimal transformation on $\overset{01}{N}_{11}$ is

$$\overset{01}{N}_{11} \to \overset{01}{N}_{11} - \gamma[\overset{k1}{N}_{11} + \overset{0,k:1}{N}_{11} + \sum_s (\overset{0s}{N}_{11}\overset{k-s,1}{N}_{21} + \overset{0s}{N}_{12}\overset{k-s,1}{N}_{11})]. \quad (7.4)$$

As written, the summation runs over $s = 1, \ldots, k$. The simplest approach is to use the option mentioned in II, of extending the sum over *all* values of s, both positive and negative. The linear terms then need no longer be written separately, since they result from $s = 0$ and $s = k+1$.

Using Eq. (II.2.15) and the reality conditions for static potentials, we can show that

$$\overset{m1}{N}_{11} = -\overset{0,m+1}{N}_{11} + \overset{0m}{N}_{11}(\overset{01}{N}_{12} - \overset{01}{N}_{21}) - 2\overset{01}{N}_{11}\overset{0m}{N}_{21},$$

$$\overset{m1}{N}_{21} = \overset{0,m+1}{N}_{12} - \overset{0m}{N}_{12}(\overset{01}{N}_{12} - \overset{01}{N}_{21}) + 2\overset{01}{N}_{11}\overset{0m}{N}_{22}.$$

The bracket in Eq. (7.4) simplifies to

$$-2\overset{01}{N}_{11}\sum_s (\overset{0s}{N}_{12}\overset{0,k-s}{N}_{21} - \overset{0s}{N}_{11}\overset{0,k-s}{N}_{22})$$

since the remaining terms cancel. This sum is equivalent to the expression we evaluated in Eq. (5.10). Our infinitesimal transformation thus becomes

$$\overset{01}{N}_{11} \to \overset{01}{N}_{11}(1 - 4\gamma(2r)^k P_k(\cos\theta)). \quad (7.5)$$

Since it is no longer nonlinear, the exponentiation is easy to do:

$$\overset{01}{N}_{11} \to \overset{01}{N}_{11} \exp(-4\gamma(2r)^k P_k(\cos\theta)). \quad (7.6)$$

It is well known that the static metrics can be solved completely by letting $f = \exp(2\chi)$, where χ obeys Laplace's equation. What we have now shown is that under $\overset{k}{\gamma}_{12}$

$$\chi \to \chi - 2\overset{k}{\gamma}(2r)^k P_k(\cos\theta). \quad (7.7)$$

By using each $\overset{k}{\gamma}_{12}$ in succession, $k = 0, 1, \cdots$, we may add to χ any harmonic function which is regular at the origin. Thus if we start with flat space, $\chi = 0$, *any* Weyl solution may be generated, provided only that we choose the origin to be a point where the desired χ is nonsingular. To generate Schwarzschild, for example, we want for χ the potential of a finite rod. We take an origin displaced away from the rod, say some distance out along the z axis. The expansion coefficients of χ about this origin provide the necessary $\overset{k}{\gamma}_{12}$'s. The metric will be given as a sum over k which converges only out to the nearest singularity (the end of the rod) and not to spatial infinity. Its asymptotic flatness will not be apparent until we do the sum and extend the result analytically.

[1] W. Kinnersley, J. Math. Phys. 18, 1529 (1977) (referred to as I); W. Kinnersley and D.M. Chitre, J. Math. Phys. 18, 1538 (1977) (referred to as II).
[2] W. Kinnersley, J. Math. Phys. 14, 651 (1973).
[3] E. Herlt, preprint (1977).
[4] R. Geroch, J. Math. Phys. 13, 394 (1972).
[5] J. Synge, *Relativity: The General Theory* (North-Holland, Amsterdam, 1960), p. 236.
[6] R. Geroch (private communication) has suggested it might be possible to exponentiate such transformations by using the generating function (cf. Sec. 5).
[7] W. Kinnersley and D.M. Chitre, J. Math. Phys. (to be published).

Quantization on a manifold with connection

J. Underhill

Department of Applied Mathematics and Theoretical Physics, The University of Liverpool, Liverpool, England
(Received 26 July 1977)

We give a rule Q which generalizes the Weyl correspondence to systems whose configuration space is equipped with an affine connection. We show that $[Q(f), Q(g)] = iQ(\{f, g\})$ for arbitrary f provided g has the form $X^i p_i$, where X^i is an affine vector field.

Dirac's idea[1] that quantization of a classical system means finding a representation of the Lie algebra of classical observables by operators in Hilbert space has been the starting point for most work in this subject. Since the work of van Hove[2] it has been clear that the operators of such a representation have no direct interpretation as quantum mechanical observables. Extracting such observables by means of a polarization of the classical phase space (i.e., passing from prequantization to quantization in the terminology of Kostant[3]) seems to lose one the ability to quantize most classical observables, as well as requiring, usually, a great deal of symmetry in the system.

One might try to weaken Dirac's requirement for a quantization by imposing the condition

$$[Q(f),\ Q(g)] = iQ(\{f,g\}) \tag{1}$$

($\{\ \}$ is the Poisson bracket, Q denotes quantization, $\hbar = 1$) where the classical observable f is arbitrary while g lies in some restricted class of observables. Weyl's correspondence[4] provides a quantization in this sense; in this case phase space is R^{2n} and (1) is satisfied for all f provided g is a quadratic function of the coordinates $x^1, \ldots, x^n, p_1, \ldots, p_n$. In particular g may have the form $\langle p, X\rangle = X^i(x)p_i$, where $p = p_i dx^i$ and the components of the vector field $X = X^i(\partial/\partial x^i)$ on configuration space R^n are first-degree polynomials in $x^1, \ldots, x^n$. Vector fields of this type are precisely the affine (or affine Killing) vector fields on R^n, i.e., the fields satisfying $\mathcal{L}_x \nabla_0 = 0$, where ∇_0 is the canonical flat affine connection on R^n.

In this paper we consider a system whose phase space is the cotangent bundle T^*M of an n-dimensional C^∞ manifold M equipped with a symmetric affine connection ∇. A quantization Q which reduces to Weyl's in the case $M = R^n$ is defined and it is shown that Q also singles out the affine vector fields on M in the way described above. Q assigns differential operators to functions which are sums of terms of the form f_X, where X is a symmetric contravariant tensor field on M of order m and

$$f_X(x,p) = \langle \overbrace{p \otimes \cdots \otimes p}^{m \text{ factors}},\ X(x)\rangle. \tag{2}$$

The highest m occurring is then the degree of the polynomial. It is convenient to restrict Q to such polynomial functions, since this allows us to avoid assumptions about the completeness of ∇.

DEFINITION OF Q

The Weyl correspondence assigns to the function f on R^{2n} the linear operator in $L^2(R^n)$ written formally

$$W(f) = \int \hat{f}(\alpha_1, \ldots, \alpha_n, \beta^1, \ldots, \beta^n) \times \exp\left(i\alpha_i x^i + \beta^j \frac{\partial}{\partial x^j}\right) d\alpha_1 \cdots d\beta^n,$$

where $\hat{f}$ is the Fourier transform of f:

$$f(x,p) = \int \hat{f}(\alpha,\beta) \exp(i\alpha_j x^j + i\beta^j p_j)\, d\alpha_1 \cdots d\beta^n.$$

A formal calculation shows that the matrix element

$$\begin{aligned}(\varphi, W(f)\psi) &= \int \overline{\varphi(x)}(W(f)\psi)(x)\, dx^1 \cdots dx^n \\ &= \pi^{-n} \int \overline{\varphi(x-v)}\psi(x+v)\, e^{-2i\langle p, v\rangle} f(x,p) \\ &\quad \times dv^1 \cdots dx^1 \cdots dp_1 \cdots dp_n,\end{aligned} \tag{3}$$

for $\varphi, \psi \in L^2(R^n)$.

The quantization Q is defined by trying to make sense of (3) when R^n is replaced by M. M is not provided with a natural measure, so we let the Hilbert space of quantum states be the completion of $\mathcal{D}_0(M)$, the space of C^∞ complex scalar $\frac{1}{2}$-densities on M of compact support. Then, given $\psi \in \mathcal{D}_0$, we wish to use the exponential map E_x (corresponding to the connection ∇) to pull ψ back to a $\frac{1}{2}$-density on T_xM. We avoid difficulties which arise because E_x is defined only near O in T_xM by introducing a cutoff function χ on TM which, for each $x \in M$, vanishes outside a normal neighborhood of $O \in T_xM$ symmetric with respect to reflection in O, and is unity in some smaller neighborhood of O. Then the product $\psi_x^* = \chi E_x^* \psi$ is a $\frac{1}{2}$-density $\in \mathcal{D}_0(T_xM)$, nonzero only in supp $\chi(x, \cdot)$ and defined there as $\chi(x,v)$ times the pull back of ψ at $E_x v$. Similarly we define ψ_x^-, simply replacing E_x by the map $v \to E_x(-v)$.

Since the product of two $\frac{1}{2}$-densities is a density, given $\varphi, \psi \in \mathcal{D}_0(M)$, the expression

$$\int e^{-2i\langle p, v\rangle} \psi_x^* \overline{\varphi_x^-} = \Phi(x,p) \tag{4}$$

defines a function Φ on T^*M, of rapid decrease in $\|p\|$ with respect to any Riemannian structure on M and, provided χ is chosen suitably, compactly supported in x. (For example, let S be any compact set containing supp ψ in its interior, then given a Riemannian structure on M, $\exists\ t > 0$ such that $E_x v \in$ supp ψ, $\|v\| < t \Rightarrow x \in S$. Choose χ so that $\chi(x,v) = 0$ whenever $\|v\| \geq t$, then $\Phi(x,p) = 0$ if $x \notin S$). Hence given $f \in C^\infty(T^*M)$, polynomial in the momenta, we may define

 0022-2488/78/1909-1932$1.00

$$(\varphi, Q(f)\psi) = \pi^{-n}\int \Phi(x,p) f(x,p)\, d\mu(x,p), \tag{5}$$

where $d\mu(x,p)$ is the canonical volume element on T^*M.

Property 1. $Q(f)$ is independent of the choice of χ and is a differential operator of order equal to the degree of f.

Proof. It is sufficient to consider the case $f = f_X$ [see (2)], where X vanishes outside some coordinate neighborhood $U \subset M$, (the general case follows by a partition of unity argument). If $x^1, \ldots, x^n$ are coordinates in U and $v^1, \ldots, v^n$, $p, \ldots, p_n$, the induced coordinates in T_xU, T_x^*U,

$$(\varphi, Q(f_x)\psi) = \pi^{-n}\int dx^1 \cdots dx^n \int dp_1 \cdots dp_n \times \Phi(x,p) X^{i_1 \cdots i_m}(x) p_{i_1} \cdots p_{i_m}.$$

Now

$$\int dp_1 \cdots dp_n \Phi(x,p) p_{i_1} \cdots p_{i_m} = (-2i)^{-m} \iint \frac{\partial^m}{\partial v^{i_1} \cdots \partial v^{i_m}} e^{-2i\langle p, v\rangle} D(x,v)\, dv^1 \cdots dv^n dp_1 \cdots dp_n$$

[where $D(x,v)\,dv^1 \cdots dv^n$ is the expression of $\psi_x^* \bar{\varphi}_x^-$ in local coordinates],

$$= (2i)^{-m} \iint e^{-2i\langle p, v\rangle} \frac{\partial^m}{\partial v^{i_1} \cdots \partial v^{i_m}} D(x,v)\, dv^1 \cdots dv^n dp_1 \cdots dp_n$$

$$= (2i)^{-m}\pi^n \frac{\partial^m}{\partial v^{i_1} \cdots \partial v^{i_m}} D(x,v)\Big|_{v=0}.$$

So

$$(\varphi, Q(f_x)\psi) = (2i)^{-m}\int dx^1 \cdots dx^n X^{i_1 \cdots i_m}(x) \frac{\partial^m}{\partial v^{i_1} \cdots \partial v^{i_m}} D(x,v)\Big|_{v=0}. \tag{6}$$

For each x, $\chi(x,v) = 1$ near $v = 0$ so that (6) is independent of the choice of χ.

Now suppose $\varphi = \tilde{\varphi}(x)(dx^1 \cdots dx^n)^{1/2}$, $\psi = \tilde{\psi}(x) \times (dx^1 \cdots dx^n)^{1/2}$ in local coordinates; let $\xi_\pm = E_x(\pm v)$ and let $J_\pm$ be the Jacobians $\det(\partial \xi_\pm^i / \partial v^j)$, then near $v = 0$

$$D(x,v) = \tilde{\psi}(\xi_+)\overline{\tilde{\varphi}(\xi_-)}(|J_+ J_-|)^{1/2}.$$

Hence $\partial^m D(x,v)/\partial v^{i_1} \cdots \partial v^{i_m}|_{v=0}$ is a sum of products of derivatives of $\bar{\tilde{\varphi}}$ and $\tilde{\psi}$ of total order $\leq m$. Integration by parts in (6) to remove derivatives of $\bar{\tilde{\varphi}}$ shows that $Q(f_x)$ is a differential operator of order $\leq m$. The highest derivative arises from the term

$$\sum_{k=0}^{m} \binom{m}{k} \frac{\partial^k \psi(x)}{\partial x^{i_1} \cdots \partial x^{i_k}} \frac{\partial^{m-k} \overline{\tilde{\varphi}(x)}}{\partial x^{i_{k+1}} \cdots \partial x^{i_m}} (-1)^{m-k}$$

in $\partial^m D(x,v)/\partial v^{i_1} \cdots dv^{i_m}|_{v=0}$, which contributes

$$(2i)^{-m} X^{i_1 \cdots i_m}(x) \sum_{k=0}^{m} \binom{m}{k} \frac{\partial^m}{\partial x^{i_1} \cdots \partial x^{i_m}} = i^{-m} X^{i_1 \cdots i_m}(x) \frac{\partial^m}{\partial x^{i_1} \cdots \partial x^{i_m}}$$

to $Q(f_x)$.

Examples

$m = 0$:

$$f_x(x,p) = X(x).$$

$$(\varphi, Q(f_x)\psi) = \int \overline{\tilde{\varphi}(x)} X(x) \tilde{\psi}(x)\, dx^1 \cdots dx^n = \int \overline{\varphi(x)} X(x)\, \psi(x),$$

$$Q(f_x) = X. \tag{7}$$

$m = 1$:

$$f_x(x,p) = X^i(x) p_i.$$

$$\frac{\partial}{\partial v^i} D(x,v)\Big|_{v=0} = \frac{\partial \tilde{\psi}}{\partial x^i}(x)\overline{\tilde{\varphi}(x)} - \tilde{\psi}(x)\frac{\partial}{\partial x^i}\overline{\tilde{\varphi}(x)},$$

so

$$(\varphi, Q(f_x)\psi) = (2i)^{-1}\int X^i(x)\left(\frac{\partial \tilde{\psi}}{\partial x^i}(x)\overline{\tilde{\varphi}(x)} - \tilde{\psi}(x)\frac{\partial}{\partial x^i}\overline{\tilde{\varphi}(x)}\right) dx^1 \cdots dx^n$$

$$= (2i)^{-1}\int \overline{\tilde{\varphi}(x)}\left[X^i(x)\frac{\partial \tilde{\psi}}{\partial x^i}(x) + \frac{\partial}{\partial x^i}(X^i(x)\tilde{\psi}(x))\right] dx^1 \cdots dx^n$$

$$= -i\int \overline{\varphi(x)}\,[X^i(x)\nabla_i \psi(x) + \tfrac{1}{2}(\nabla_i X^i(x))\,\psi(x)].$$

Thus

$$Q(f_x) = -i[X^i \nabla_i + \tfrac{1}{2}(\nabla_i X^i)], \tag{8}$$

where ∇_i denotes covariant derivative.

If X is a vector field and Y a function on M, then (7) and (8) show that

$$[Q(Y),\ Q(X)] = iX(Y) = iQ(X(Y)),$$

so that the canonical commutation relations are satisfied.

$m = 2$:

$$f_x(x,p) = X^{ij}(x) p_i p_j.$$

The calculation here is similar but more tedious, since $J_\pm$ contribute to the result, which is

$$Q(f_x) = (-i)^2[X^{ij}\nabla_i \nabla_j + (\nabla_i X^{ij})\nabla_j + \tfrac{1}{4}(\nabla_i \nabla_j X^{ij}) + \tfrac{1}{12} R_{ij} X^{ij}],$$

where R_{ij} is the Ricci tensor. In particular if ∇ is the Riemannian connection on a Riemannian space M, then $Q(\|p\|^2) = (-i)^2(\Delta + \frac{1}{12}R)$, where Δ is the Laplacian and R the scalar curvature.

Finally, it is easily seen that

$$(\varphi, Q(f)\psi) = (Q(\bar{f})\varphi, \psi),$$

so that Q assigns to real functions on T^*M densely defined symmetric differential operators in the Hilbert space of quantum mechanical states.

AFFINE VECTOR FIELDS

A vector field X on M is affine if it satisfies $\mathcal{L}_X \nabla = 0$.[5] The result proved in this section, namely that if X is affine then

$$[Q(f),\ Q(f_x)] = iQ(\{f, f_x\})$$

for all observables f, depends on the simple way in which affine fields vary along geodesics in M. Thus, let $\xi(t,x,v)$: $\exp_x(vt)$ and suppose v close enough to 0 so

that for $t\in[-1,1]$, ξ lies in a coordinate neighborhood $U\subset M$.

Lemma: Let X be affine, then

$$X^i(\xi)=X^j(x)\frac{\partial\xi^i}{\partial x^j}+\frac{\partial X^j}{\partial x^k}(x)\,v^k\frac{\partial\xi^i}{\partial v^j}. \tag{9}$$

Proof: (9) clearly holds when $t=0$ and furthermore

$$\frac{\partial}{\partial t}\left(X^j(x)\frac{\partial\xi^i}{\partial x^j}+\frac{\partial X^j}{\partial x^k}(x)\,v^k\frac{\partial\xi^i}{\partial v^j}\right)\Big|_{t=0}$$
$$=\frac{\partial X^j}{\partial x^k}(x)\,v^k\delta^i_j=\frac{\partial X^i}{\partial x^k}(x)\,v^k=\frac{\partial X^i}{\partial t}(\xi)\Big|_{t=0}.$$

The result follows when we have shown that both sides of (9) satisfy the same second order linear differential equation in t. Differentiating

$$\frac{\partial^2\xi^i}{\partial t^2}+\Gamma^i_{jk}(\xi)\frac{\partial\xi^j}{\partial t}\frac{\partial\xi^k}{\partial t}=0 \tag{10}$$

with respect to x^j and v^j successively, we see that the right hand side Z^i of (9) satisfies

$$\frac{\partial^2 Z^i}{\partial t^2}+2\Gamma^i_{jk}(\xi)\frac{\partial\xi^j}{\partial t}\frac{\partial Z^k}{\partial t}+\frac{\partial\Gamma^i_{jk}}{\partial\xi^l}(\xi)\frac{\partial\xi^j}{\partial t}\frac{\partial\xi^k}{\partial t}Z^l=0.$$

Now, in terms of coordinates $\angle_X\nabla=0$ becomes $\nabla_j\nabla_k X^i+R^i_{ljk}X^l=0$,[5] or when written out in full

$$\frac{\partial^2 X^i}{\partial\xi^j\partial\xi^k}(\xi)=\Gamma^i_{jk}(\xi)\frac{\partial X^i}{\partial\xi^l}(\xi)-\Gamma^i_{kl}(\xi)\frac{\partial X^l}{\partial\xi^j}(\xi)$$
$$-\Gamma^i_{jl}(\xi)\frac{\partial X^l}{\partial\xi^k}(\xi)-\frac{\partial\Gamma^i_{jk}}{\partial\xi^l}(\xi)X^l(\xi),$$

so that

$$\frac{\partial^2}{\partial t^2}X^i(\xi)=\frac{\partial X^i}{\partial\xi^j}(\xi)\frac{\partial^2\xi^j}{\partial t^2}+\frac{\partial^2 X^i}{\partial\xi^j\partial\xi^k}(\xi)\frac{\partial\xi^j}{\partial t}\frac{\partial\xi^k}{\partial t}$$
$$=-\frac{\partial X^i}{\partial\xi^j}(\xi)\Gamma^j_{kl}(\xi)\frac{\partial\xi^k}{\partial t}\frac{\partial\xi^l}{\partial t}+\Gamma^i_{jk}(\xi)\frac{\partial X^i}{\partial\xi^l}(\xi)\frac{\partial\xi^j}{\partial t}\frac{\partial\xi^k}{\partial t}$$
$$-2\Gamma^i_{kl}(\xi)\frac{\partial X^l}{\partial\xi^j}(\xi)\frac{\partial\xi^j}{\partial t}\frac{\partial\xi^k}{\partial t}-\frac{\partial\Gamma^i_{jk}}{\partial\xi^l}(\xi)X^l(\xi)\frac{\partial\xi^j}{\partial t}\frac{\partial\xi^k}{\partial t}$$
$$=-2\Gamma^i_{kl}(\xi)\frac{\partial\xi^k}{\partial t}\frac{\partial X^l}{\partial t}(\xi)-\frac{\partial\Gamma^i_{jk}}{\partial\xi^l}\frac{\partial\xi^j}{\partial t}\frac{\partial\xi^k}{\partial t}X^l(\xi).$$

Q. E. D.

Corollary: Let $J(t,x,v)=\det(\partial\xi^i/\partial v^j)$, then

$$X^j(x)\frac{\partial J}{\partial x^j}+\frac{\partial X^j}{\partial x^k}(x)\,v^k\frac{\partial J}{\partial v^j}=\left(\frac{\partial X^i}{\partial\xi^i}(\xi)-\frac{\partial X^i}{\partial x^i}(x)\right)J.$$

Property 2: If X is an affine vector field on M, then

$$[Q(f),Q(f_X)]=iQ(\{f,f_X\})$$

for arbitrary observables f.

Proof: Given $\varphi,\psi\in\mathcal{D}_0(M)$,

$$(\varphi,[Q(f),Q(f_X)]\,\psi)=(\varphi,Q(f)\,Q(f_X)\,\psi)$$
$$-(Q(f_{\bar X})\,\varphi,Q(f)\,\psi),$$

(where $\bar X$ is the affine vector field complex conjugate to X)

$$=\pi^{-n}\int\Psi(x,p)f(x,p)\,d\mu(x,p), \tag{11}$$

where

$$\Psi(x,p)=\int\exp(-2i\langle p,v\rangle)[(\xi)(f_X)\psi)^+_x\overline{\varphi^-_x}-\psi^+_x\overline{(Q(f_{\bar X})\,\varphi)^-_x}]. \tag{12}$$

If we assume that for v such that $\chi(x,v)\neq0$, $\xi(t,x,v)$ lies in a coordinate neighborhood U of x for $t\in[-1,1]$, then

$$(Q(f_X)\,\psi)^+_x=-i\left[\left(X^i(\xi)\frac{\partial\tilde\psi}{\partial\xi^i}(\xi)\right.\right.$$
$$\left.\left.+\frac{1}{2}\frac{\partial X^i}{\partial\xi^i}(\xi)\tilde\psi(\xi)\right)\sqrt{|J|}\right]\Big|_{t=1}\chi(x,v)(dv^1\cdots dv^n)^{1/2}$$
$$=-i\left[\left(X^j(x)\frac{\partial}{\partial x^j}\tilde\psi(\xi)+\frac{\partial X^j}{\partial x^k}(x)\,v^k\frac{\partial}{\partial v^j}\tilde\psi(\xi)\right.\right.$$
$$\left.\left.+\frac{1}{2}\frac{\partial X^i}{\partial\xi^i}(\xi)\tilde\psi(\xi)\right)\sqrt{|J|}\right]\Big|_{t=1}\chi(dv^1\cdots dv^n)^{1/2},$$

by the lemma, and from the corollary we get

$$=-i\left[X^j(x)\frac{\partial}{\partial x^j}(\tilde\psi(\xi_+)\sqrt{|J_+|})+\frac{\partial X^j}{\partial x^k}(x)\,v^k\frac{\partial}{\partial v^j}(\tilde\psi(\xi_+)\sqrt{|J_+|})\right.$$
$$\left.+\frac{1}{2}\frac{\partial X^i}{\partial x^i}(x)\tilde\psi(\xi_+)\sqrt{|J_+|}\right]\chi(dv^1\cdots dv^n)^{1/2}.$$

Similarly

$$\overline{(Q(f_{\bar X})\,\varphi)^-_x}$$
$$=i\left[X^j(x)\frac{\partial}{\partial x^j}(\overline{\tilde\varphi(\xi_-)}\sqrt{|J_-|})+\frac{\partial X^j}{\partial x^k}(x)\,v^k\frac{\partial}{\partial v^j}(\overline{\tilde\varphi(\xi_-)}\sqrt{|J_-|})\right.$$
$$\left.+\frac{1}{2}\frac{\partial X^i}{\partial x^i}(x)\,\overline{\tilde\varphi(\xi_-)}\sqrt{|J_-|}\right]\chi(dv^1\cdots dv^n)^{1/2}.$$

Substituting into (12) and integrating by parts, we get

$$\Psi(x,p)=-i\left(X^j(x)\frac{\partial}{\partial x^j}-\frac{\partial X^j}{\partial x^k}(x)\,p_j\frac{\partial}{\partial p_k}\right)\Phi(x,p)$$
$$+i\int\exp(-2i\langle p,v\rangle)\left((X^j(x)\frac{\partial}{\partial x^j}(\chi^2)\right.$$
$$\left.+\frac{\partial X^j}{\partial x^k}(x)\,v^k\frac{\partial}{\partial v^j}(\chi^2)\right)\tilde\psi\,\bar{\tilde\varphi}(\sqrt{|J_+J_-|})\,dv^1\cdots dv^n. \tag{13}$$

The integral in (13) contributes nothing when we integrate over T^*M, since its integrand vanishes in a neighborhood of $v=0$ and, as in Prop. 1, its behavior elsewhere is irrelevant. Finally, assume that f has x support in U (general case by partition of unity), then

$$(\varphi,[Q(f),Q(f_X)]\,\psi)$$
$$=i\int\left(X^j\frac{\partial f}{\partial x^j}-\frac{\partial X^j}{\partial x^k}p_j\frac{\partial f}{\partial p_k}\right)\Phi(x,p)\,d\mu(x,p)$$
$$=i(\varphi,Q(\{f,f_X\})\,\psi).$$

Q. E. D.

REMARKS

The exponential map on a manifold M with affine connection provides an unlimited number of ways of assigning linear operators to functions on T^*M. None of these ways can be expected to solve the "Dirac problem," so it seems sensible to look for weaker conditions which might serve to single out certain quantizations as having some physical interest, such as the canonical commutation relations:

$$[Q(Y),Q(X)]=i\hbar Q(X(Y)),$$

(X a vector field, Y a function on M), or a "correspondence principle" version of Dirac's condition, say

$$[Q(f), Q(g)] = i\hbar Q(\{f,g\}'), \tag{14}$$

where

$$\{f,g\}' = \{f,g\} + O(\hbar). \tag{15}$$

It is easily checked that the quantization defined in this paper satisfies these rather weak requirements [with $O(\hbar^2)$ replacing $O(\hbar)$ in (15)].

Property 2, imposed as a condition on Q, appears to be more restrictive. In fact, if we insist also that Q generalizes Weyl quantization in a manner uniformly applicable to arbitrary (M, ∇), then it seems likely that Q must be as defined here. Of course, whether or not Prop. 2 provides a serious restriction on quantizations of a particular T^*M depends on how many independent affine vector fields M possesses.

The physical significance of Prop. 2 is not very clear. If ∇ is the Riemannian connection on a Riemannian space M then affine vector fields have a simple relation to the free particle motion on M, corresponding to the Hamiltonian $H(x,p) = \frac{1}{2}\|p\|^2$. Namely, if X is affine then $f_X(x(t), p(t))$ varies linearly with the time. [In fact

$$\frac{d^2}{dt^2} f_X(x,p) = (\nabla_i \nabla_j X^k) g^{il} g^{jm} p_k p_l p_m = 0$$

since $g^{jm} g^{il} \nabla_i \nabla_j X^k + (\text{cyclic permutations of } k,l,m) = 0$ as a consequence of $\nabla_i \nabla_j X^k + R^k_{lij} X^l = 0$ and properties of the Riemann tensor.] Killing fields, defined by $\mathcal{L}_X g = 0$, have the property that f_x is a constant of the classical free-particle motion. Killing fields are affine fields so Prop. 2 implies that $Q(f_x)$ is a constant of the motion with Hamiltonian $Q(H)$.

We observed that the Weyl correspondence satisfies $[W(f), W(g)] = iW(\{f,g\})$ for all f also when $g = \frac{1}{2}\|p\|^2 = p_1^2 + \cdots + p_n^2$, a condition with clearer physical meaning, since it implies that the classical and quantum dynamics of a free particle in R^n are identical. It is easily checked that Q does not have this property even in the simplest (nonflat) cases, e.g., spaces of constant curvature. For such spaces the definition of Q can be modified slightly so as to guarantee a weaker property, that constants of the classical and quantum free particle motions correspond.[6] Bloore *et al.*[7] have shown that no quantization of a certain class (including our Q) can have the stronger property unless M is of constant curvature or Ricci flat.

[1] P. A. M. Dirac, *Lectures on Quantum Mechanics* (Publisher, New York, 1964).
[2] L. van Hove, Acad. Roy. Belg. Cl. Sci. Mem. Coll. in 8° **26**, No. 6 (1951).
[3] B. Kostant, *Quantization and unitary representations, Lecture Notes in Mathematics* (Springer, Berlin, 1970), Vol. 170.
[4] H. Weyl, *Group Theory and Quantum Mechanics* (Dover, Princeton, N.J., 1950).
[5] K. Yano, *Integral Formulas in Riemannian Geometry* (Dekker, New York, 1970).
[6] S. Taraviras, thesis 1975, University of Liverpool.
[7] F. J. Bloore, M. Assimakopoulos, and I. Ghobrial, "Quantization as deformation theory," J. Math. Phys. **17**, 1034 (1976).

Generalization of the inversionlike integral equations and application to nonlinear partial differential equations. II

Henri Cornille

Rockefeller University, New York, New York 10021
and D.P.H.T. Saclay, BP n° 2, 91190 Gif-sur-Yvette, France
(Received 6 May 1977)

Using pedestrian algebraic method, in the case of a coupled system of n linear first order differential equations $(\Delta_0 - i\phi\Lambda - Q)\psi = 0$, $\Lambda = (\delta_{ij}\lambda_i)$, $\Delta_0 = (\delta_{ij}\mu_i(x)(\partial/\partial x))$, $Q = (q_j^i)$, we deduce inversionlike integral equations (from which we can construct a class of potentials without introducing the data). Assuming, like in the Zakharov–Shabat theory, that the kernels of the linear integral inversion equation satisfy linear partial differential equations, we deduce, for the solution of the integral equation, the corresponding nonlinear partial differential equation. We show that the validity of the whole formalism depends mainly on the set of $\{\lambda_i\}$ of the associated differential linear system. We discuss fully the different possibilities which occur depending on different sets of $\{\lambda_i\}$.

I. INTRODUCTION

Since the pioneering works of Gardner, Greene, Kruskal and Miura,[1a] Lax,[1b] Zakharov and Shabat,[1c] Ablowitz, Kaup, Newell, and Segur[1d] the so-called "inverse scattering method" is considered as a useful tool in order to explicitly construct the solutions of some nonlinear partial differential equations[1e] (nlpde).

In a previous paper,[2] we have realized that in order to get integral equations (IE) such that we can generate both a class of solutions and of potentials of a differential system (and the related nlpde), it was not necessary to study the properties of the Jost solutions in the complex eigenvalue plane, and that these investigations could be replaced by pedestrian algebraic methods. In this way we deduced the IE (and some associated nlpde) corresponding to a linear first order differential system $I\mu(x)(\partial/\partial x)$ of n coupled equations with a second-order eigenvalue problem. However, other physically interesting nlpde cannot be deduced by this scheme. Manakov and Zakharov,[3a] and Kaup,[3b] have considered a third-order problem and a generalization has been done by Zakharov and Shabat.[3c] These authors[3c] postulate an inversionlike integral equation without establishing explicitly the link with the associated differential system. Moreover in their formal work they do not discuss the boundary conditions to be satisfied by the kernels of the integral equations in order that the formalism be correct. As we shall see here the discussion turns around the values of the constant term of the associated system and the sufficient conditions to be considered are far from being obvious. Ablowitz and Haberman[3d] have written down the nlpde which could be associated with a general linear system where the nonlinear parts of the nlpde can include triads, multitriads, quartets, multiquartets, cubic self-modal, self–self interactions, ° ° °. However the derivation of the general IE as well as the explicit construction from it of these nlpde was missing. It is the aim of this paper to first derive the IE associated to a linear system and secondly, introducing, like in the Zakharov–Shabat theory, lpde for the kernels of the IE, to explicitly construct Ablowitz and Haberman's nlpde.[3d]

Let us consider the $n \times n$ linear system

$$(\Delta_0(x) - i\phi\Lambda - Q(x))\,\psi(x) = 0, \tag{1}$$

with

$$\Delta_0 = \left(\delta_{ij}\,\mu_i(x)\,\frac{\partial}{\partial x}\right),\quad \Lambda = (\delta_{ij}\lambda_i),\quad Q = \begin{pmatrix} q_1^1 \circ \circ \circ q_1^n \\ \vdots \\ q_n^1 \cdots q_n^n \end{pmatrix},$$

ψ a column vector, $\mu_i(x) > 0$, $\lim_{x\to\infty}(\int^x [\mu_i(u)]^{-1}\,du) = +\infty$, the λ_i being fixed numbers where p of them $p \leq n$ have different values.

For $\mu_i \equiv 1$, $n = 3$ and three different λ_i, Kaup has[3] studied the analytical properties of the direct and inverse scattering problem. However[3] it does not seem easy to extend his method for $n > 3$. Our aim in Sec. II is to obtain from straightforward algebraic methods an integral equation (IE) such that we can both generate a class of solutions and of potentials Q of (1). We shall not try to define (as it must be for a true inverse formalism) the data on the basis of which Q could be reconstructed, nor to get from these data the kernel of the IE. Thus we want to provide a potential framework for an inversionlike procedure. Studying in Sec. II the validity of the procedure it appears convenient to associate to each kernel F_j^i of this IE the sign of the product $\lambda_i\lambda_j$. If this sign is negative then F_j^i can be chosen square integrable whereas this is not possible if this sign is positive. Then either we put to zero these badly behaving kernels (reducing the case of reconstructed potentials) or if we introduce such kernels we have to introduce the rules in order to avoid a breakdown of the formalism.

The Zakharov–Shabat general formalism[3c] ignores these conditions on the kernels of the IE generating the potentials so that it is necessary to study how the associated nlpde can still be deduced when the correct boundary conditions on the F_j^i are introduced.

In Sec. III we consider a simplified case where the $(p-1)$ first λ_i have a sign different from the remaining ones which are equal and we put $\mu_i \equiv 1$. In Sec. III A we construct explicitly the nlpde whose linear parts correspond to the operator constant $\partial/\partial t - \partial/\partial x$ and in

 0022-2488/78/1909-1936$1.00

Sec. III B to those corresponding to the constant $\partial/\partial t - \partial^2/\partial x^2$. In all these subsections we first consider the case where only the $\lambda_i\lambda_j < 0$ kernels F_j^i are present and we later study the modification of the formalism when the $\lambda_i\lambda_j > 0$ kernels F_j^i are introduced.

Before we generalize the IE in the matrix case, using algebraic methods, we recall how it works in the classical scalar case where it was previously done.[4] Let us consider

$$\left[-\gamma + \mu^2(x)\left(\frac{\partial^2}{\partial x^2} - Q(x)\right)\right]\psi(x) = 0, \qquad [a]$$

where μ is arbitrary. We associate (Q_0 and ψ^0 are known)

$$\left[-\gamma + \mu^2(x)\left(\frac{\partial^2}{\partial x^2} - Q_0(x)\right)\right]\psi^0(x) = 0, \quad D_0 = \mu^2\frac{\partial^2}{\partial x^2} - Q_0 \ .$$

We postulate a representation of ψ (c is a constant),

$$\psi(x) = \psi^0(x) + \int_x^a \mu^{-2}\psi^0(t)\,\overline{K(x,t)}\,dr\,. \qquad [b]$$

We put [b] into [a] and get $\{D = D_0 + 2\mu(d/dx) \times [K(x,x)\,\mu^{-1}(x)]\}$

$$0 = \mu^2(x)\,\psi(x)\left[Q_0 - Q - \left(\frac{2}{\mu}\frac{d}{dx}\frac{\overline{K}(x,x)}{\mu(x)}\right)\right] + \int_x^a \frac{\psi^0}{\mu^2}(t)[D(x) - D_0(t)]\,\overline{K}(x,t)\,dt\,.$$

If we find a kernel $\overline{K}$ such that

$$(D(x) - D_0(t))\,\overline{K}(x,t) = 0, \qquad [c]$$

then

$$Q = Q_0 - 2\mu^{-1}\frac{d}{dx}\left(\overline{K}(x,x)\,\mu^{-1}(x)\right). \qquad [d]$$

Now if we define

$$K(x,y) = F(x,y) + \int_x^a F(s,y)\,K(x,s)\,\mu^{-2}(s)\,ds \qquad [e]$$

and assume $(D_0(x) - D_0(y))\,F = 0$ [f], then K, the solution of [e], satisfies [c]. It follows that if in [b] we choose $\overline{K} = K$ then the potential Q is given by [d]. We can consider $a = 0, \infty, \cdots$ and we have to specify the boundary conditions on F and μ in order that all the derivations be correct.

II. INVERSIONLIKE EQUATIONS

A. Statement of the problem

First we write down a representation of a set of solutions of the whole system as transforms of the reduced set of solutions when the "potential" is switched off. Secondly we put this representation into the differential system and we get (boundary conditions appear at this stage) that these transforms must satisfy well-defined nlpde. Thirdly, we guess an integral equation (IE) such that the solutions satisfy the above nlpde. In this way we get that the kernels of the IE must satisfy well-defined partial differential equations and also boundary conditions in order that all the derivations be correct. A more conventional mathematical presentation would be to reverse this order and start from well specified assumptions on the kernels of the IE, but the insight of the method could be lost.

Let us formally write a set of n functions [which we would like to be solutions of Eq. (1)] with the following representation,

$$\psi_j = (v_i^j) = \left(\delta_{ij}u_j(x) + \int_x^\infty \mu_j^{-1}(y)\,u_j(y)\,\overline{K}_i^j(x,y)\,dy\right), \int \left(\mu_j(x)\frac{\partial}{\partial x} - i\lambda_j\phi\right)u_j(x) = 0, \quad u_j(x) = \exp(i\lambda_j\phi)\int^x \mu_j^{-1}(u)\,du. \qquad (2)$$

B. nlpde satisfied by $(K_i^j(x,y))$

We assume the following boundary conditions,

$$\lim_{y\to\infty} u_j(y)\,\overline{K}_i^j(x,y) = 0. \qquad (3)$$

Let us introduce the following lemma.

Lemma: If the boundary conditions (3) are satisfied, then we get:

$$\left(\mu_i(x)\frac{\partial}{\partial x} - i\phi\lambda_i\right)\int_x^\infty \mu_j^{-1}(y)\,u_j(y)\,\overline{K}_i^j(x,y)\,dy = \left(\frac{\lambda_i}{\lambda_j} - \frac{\mu_i(x)}{\mu_j(x)}\right)u_j\hat{\overline{K}}_i^j + \int_x \mu_j^{-1}u_j O_i^j\overline{K}_i^j\,dy\,,$$

$$\hat{\overline{K}}_i^j = \overline{K}_i^j(x,x), \quad O_i^j = \mu_j(x)\frac{\partial}{\partial x} + \frac{\lambda_j}{\lambda_i}\mu_i(y)\frac{\partial}{\partial y}\,. \qquad (4)$$

We assume that the ψ_j given by (2) are solutions of (1),

$$(\Delta_0 - i\Lambda\phi - Q)(\psi_1, \psi_2, \ldots, \psi_n) = (0),$$
$$\left(\mu_i(x)\frac{\partial}{\partial x} - i\lambda_i\phi\right)v_i^j - \sum_{m=1}^n q_i^m v_m^j = 0, \qquad (5)$$

we assume that the conditions of the lemma are satisfied, and taking into account Eq. (4), the result Eq. (5) is written down with scalar quantities

$$u_j(x)\left[-q_i^j + \hat{\overline{K}}_i^j\left(\frac{\lambda_i}{\lambda_j} - \frac{\mu_i(x)}{\mu_j(x)}\right)\right] + \int_x^\infty \mu_j^{-1}(y)\,u_j(y)[O_i^j\overline{K}_i^j - \sum_m q_i^m\overline{K}_m^j]\,dy = 0. \qquad (5')$$

It follows that if the transforms $\{\overline{K}_j^i\}$ of Eq. (2) satisfy the nlpde

$$O_j^i\overline{K}_j^i = \sum_{m\neq j}\overline{K}_m^i\hat{\overline{K}}_j^m\left(\frac{\lambda_j}{\lambda_m} - \frac{\mu_j(x)}{\mu_m(x)}\right), \qquad (6)$$

then the set of potentials $\{q_j^i\}$ of Eq. (1)

$$q_j^i = \left(\frac{-\mu_j(x)}{\mu_i(x)} + \frac{\lambda_j}{\lambda_i}\right)\hat{\overline{K}}_j^i, \quad q_i^i = 0, \qquad (7)$$

are such that the lhs of Eq. (5′) is zero.

C. A conjecture of an inversion equation of Eq. (1) [1]

Let us consider the following integral equation, $F_i^i = 0$ and

$$K_j^i(x,y) = F_j^i(x,y) + \sum_{m\neq i}^{m=n}\int_x^\infty F_m^i(s,y)\,K_j^m(x,s)\,\mu_m^{-1}(s)\,ds\,. \qquad (8)$$

We could also introduce the $F_i^i \neq 0$ kernels, however it will turn out that when we introduce the boundary conditions, they must be zero. In fact as we shall see later, other kernels must also be zero.

For each scalar kernel F_j^i we assume the following boundary condition,

$$\lim_{y\to\infty} F_j^i(x,y)=0, \quad \lim_{s\to\infty} F_j^i(s,y)K_l^j(x,s)=0 \tag{9}$$

and we require of course that the solution of Eq. (8) exists.

Property: If we assume that each kernel F_j^i satisfies both the boundary condition Eq. (9) and the partial differential equation

$$O_j^i F_j^i = \left(\mu_j(x)\frac{\partial}{\partial x}+\frac{\lambda_j}{\lambda_i}\mu_i(y)\frac{\partial}{\partial y}\right)F_j^i(x,y)=0, \tag{10}$$

then the solution of Eq. (8) satisfies the nlpde [Eq. (6)]

$$O_j^i K_j^i = \sum_{m\neq j}\left(\frac{\lambda_j}{\lambda_m}\frac{-\mu_j(x)}{\mu_m(x)}\right)K_m^i\hat{K}_j^m, \quad \hat{K}_j^m = K_j^m(x,x). \tag{6a}$$

For the proof we apply O_j^i to both sides of Eq. (8),

$$O_j^i K_j^i = O_j^i F_j^i - \sum_m \frac{\mu_j}{\mu_m}F_m^i\hat{K}_j^m + \sum\int_x \mu_m^{-1}F_m^i\mu_j(x)\frac{\partial}{\partial x}K_j^m\,ds$$

$$+\sum_m\int_x \mu_m^{-1}K_j^m\frac{\lambda_j}{\lambda_i}\mu_i(y)\frac{\partial}{\partial y}F_m^i\,ds. \tag{11}$$

Using relation (10), the rhs of Eq. (11) can be written

$$\sum_m\left(\frac{-\mu_j(x)}{\mu_m(x)}+\frac{\lambda_j}{\lambda_m}\right)F_m^i\hat{K}_j^m+\sum_m\int\mu_m^{-1}F_m^i O_j^m K_j^m\,ds,$$

and comparing it with the solution of Eq. (8), result (6) follows. Let us define $K(x,y)=(K_j^i(x,y))$ and

$$\mathcal{J}(x,y)=(F_j^i(x,y)), \quad \tilde{\mathcal{J}}(s,y)=(F_j^i(s,y)\mu_j^{-1}(s)\theta(s-x)).$$

Then Eq. (8) can be written in matrix form,

$$K(x,y)=\mathcal{J}(x,y)+\int_{-\infty}^{+\infty}\tilde{\mathcal{J}}(s,y)K(x,s)\,ds. \tag{8a}$$

In conclusion, if the kernel $\mathcal{J}$ satisfies both Eqs. (9) and (10), if we substitute the solution K of Eq. (8) into the representation (2) $\bar{K}_j^i=K_j^i$, and further if condition (3) is satisfied, then Eq. (2) are solutions of our starting partial differential system (1) and consequently Eq. (8) will be an associated IE.

D. Conditions on the kernel

We must provide the sufficient conditions on the kernel $\mathcal{J}$ in order that all the above derivations in subsections IIB and IIC be correct. We must verify (3), (9),

$$F_j^i(x,y)\underset{y\to\infty}{\to}0, \quad K_j^i(x,y)\underset{y\to\infty}{\to}0, \quad F_j^i(s,y)K_1^j(x,s)\underset{s\to\infty}{\to}0,$$

and that the solution of the IE [Eq. (8)] exists.

Each kernel F_j^i must satisfy the linear partial differential equation (10). However, if for each F_j^i we associate the product $\lambda_i\lambda_j$, we find two classes of different behaviors for these scalar kernels.

$\lambda_i\lambda_j<0$: We can have square integrable scalar kernels with a continuum as well as a discrete part,

$$F_j^i=\int d\nu_j^i\alpha_j^i(\nu_j^i)\exp i\nu_j^i\left(\int^x\mu_j^{-1}(u)\,du-\frac{\lambda_i}{\lambda_j}\int^y\mu_i^{-1}(u)\,du\right) \tag{12a}$$

or

$$F_j^i=\sum_m\alpha_{j,m}^i\exp(-\nu_{j,m}^i)\left(\int^x\mu_j^{-1}\,du-\frac{\lambda_i}{\lambda_j}\int^y\mu_i^{-1}\,du\right), \quad \nu_{j,m}^i>0. \tag{12b}$$

$\lambda_i\lambda_j>0$: For simplicity we consider discrete kernels: They are not square integrable,

$$F_j^i=\sum_m\alpha_{j,m}^i\exp\nu_{j,m}^i\left(\int^x\mu_j^{-1}\,du-\frac{\lambda_i}{\lambda_j}\int^\mu\mu_i^{-1}\,du\right), \quad \nu_{j,m}^i>0. \tag{12c}$$

Whereas in (12b), $F_j^i(x,y)\to 0$ either when $x\to\infty$ or $y\to\infty$, on the contrary in (12c), $F_j^i(x,y)_{y\to\infty}\not\to 0$ and they behaved badly when $x\to\infty$: $F_j^i(x,y)_{x\to\infty}\to\infty$. It follows that these kernels given by Eq. (12c) must be compensated by other kernels in order that both the solutions of the IE exist and the conditions of Eq. (9) hold.

Let us call $\mathcal{J}^+(\mathcal{J}^-)$ the set of (F_j^i) kernels with $\lambda_i\lambda_j>0$ $(\lambda_i\lambda_j<0)$. With $\mathcal{J}^-$ are associated square integrable scalar kernels $F_j^i(x,y)$ whereas with $\mathcal{J}^+$ are associated badly behaving kernels as in Eq. (12c). Although our aim is not to perform a complete mathematical study of integral equations when square integrable kernels $(\mathcal{J}^-)$ are mixed with badly behaved kernels $(\mathcal{J}^+)$ we would like at least to know if we can introduce some of these $F_j^i\in\mathcal{J}^+$ without violating the conditions (3), (9) or the existence of the solution of Eq. (8). Our study will be at two levels. First we consider very simple examples where we can write the solutions of Eq. (8) in closed form. If we find a counterexample then we exclude that case, if another example satisfies all the required properties, then we keep the possibility to introduce such corresponding kernels. Secondly, from the iteration of the solution of Eq. (8) we try to find some more general criteria.

1. Some very simple examples

We assume that the kernels are discrete and degenerate of the most simple form $F_j^i(x,y)=g_j^i(x)h_j^i(y)$ given by (12b) and (12c).

(i) We consider $n=p=3$, $\lambda_1<0$, $\lambda_2<0$, $\lambda_3>0$, and $\mathcal{J}$ of the type

$$\mathcal{J}=\begin{bmatrix}0 & F_2^1 & 0\\ F_1^2 & 0 & F_3^2\\ F_1^3 & 0 & 0\end{bmatrix},$$

$$DK_2^1=g_2^1(x)h_2^1(y), \quad A_j^i=\int_x^\infty\mu_k^{-1}g_k^ih_j^k\,ds,$$

$$D=1-\int_x^\infty\mu_1^{-1}g_1^2h_2^1\,ds\int_x^\infty\mu_2^{-1}g_2^1h_1^2\,ds-A_3^1A_2^3A_1^2.$$

All $h_j^i(s)$, $g_j^i(s)$ go to zero when $s\to\infty$, except $g_2^1(s)\to\infty$, $g_1^2(s)\to\infty$. The solution K_2^1 of Eq. (8) is written down. We must have both $g_1^2(s)h_2^1(s)\mu_1^{-1}(s)\to 0$, $g_2^1(s)h_1^2(s)\mu_2^{-1}(s)\to 0$, which are contradictory conditions. Thus if $F_j^i\in\mathcal{J}^+$, then both F_j^i, F_i^j kernels cannot be present.

(ii) We put $F_1^2 \equiv 0$, $F_2^1 \neq 0$ in the example of (i) and the solutions are

$$\mathcal{J} = \begin{bmatrix} 0 & F_2^1 & 0 \\ 0 & 0 & F_3^2 \\ F_1^3 & 0 & 0 \end{bmatrix},$$

$$DK = \begin{bmatrix} h_2^1(y) & & 0 \\ & h_3^2(y) & \\ 0 & & h_1^3(y) \end{bmatrix} \begin{bmatrix} A_1^2 A_3^1, & 1, & A_3^1 \\ A_1^2, & A_2^3 A_1^2, & 1 \\ 1, & A_2^3, & A_3^1 A_2^3 \end{bmatrix} \times \begin{bmatrix} g_1^3(x_1) & & 0 \\ & g_2^1(x_2) & \\ 0 & & g_3^2(x_3) \end{bmatrix}.$$

$D = 1 - A_3^1 A_2^3 A_1^2$. All $g_j^i(s)$, $h_j^i(s)$ go to zero when $s \to \infty$, except $g_2^1(s) \to \infty$. If $g_2^1(s) h_3^2(s) \mu_2^{-1} \to 0$ when $s \to \infty$, then the determinant D as well as all the solutions K_j^i exist, all $K_j^i(x, y) \underset{y \to \infty}{\to} 0$ all $F_j^i(s, y) K_l^j(x, s) \underset{s \to \infty}{\to} 0$. Thus we can introduce some $F_j^i \in \mathcal{J}^+$ but the expression (12c) alone is not sufficient to guarantee the validity of the formalism, such F_j^i kernels have to satisfy extra conditions linking for instance the parameters of (12b) and (12c).

(iii) $n = p = 4$, $\lambda_1 < 0$, $\lambda_2 < 0$, $\lambda_3 < 0$, $\lambda_4 > 0$. Let us still consider $F_j^i = g_j^i(x) h_j^i(y)$ and

$$\mathcal{J} = \begin{bmatrix} 0 & F_2^1(x, y) & 0 & F_4^1(x, y) \\ 0 & 0 & F_3^2(x, y) & 0 \\ F_1^3(x, y) & 0 & 0 & 0 \\ F_1^4(x, y) & 0 & 0 & 0 \end{bmatrix}.$$

All $g_j^i(s)$, $h_j^i(s) \to 0$ when $s \to \infty$ except $g_2^1(s) \to \infty$, $g_3^2 \to \infty$, $g_1^3 \to \infty$. The Fredholm determinant of the solution of Eq. (8) contains the terms $\int_x^\infty \mu_1^{-1} g_1^3 h_2^1 ds$, $\int_x^\infty \mu_2^{-1} g_2^1 h_3^2 ds$, $\int_x^\infty \mu_3^{-1} g_3^2 h_1^3 ds$ which must exist and this leads to contradictory conditions. Thus if F_j^i, F_k^j, F_i^k kernels belong to $\mathcal{J}^+$ they cannot all be present, one of them must be put equal to zero.

(iv) We can go on with $n = p = 5$, $\lambda_1 < 0$, $\lambda_2 < 0$, $\lambda_3 < 0$, $\lambda_4 < 0$, $\lambda_5 > 0$ and assume F_2^1, F_3^2, F_4^3, F_1^4 are present. We find that they lead to inconsistent conditions unless we put to zero one of them.

Finally from these examples we have learned for $\mathcal{J}^+$ that we cannot have both F_j^i and F_i^j, $F_j^i F_k^j F_i^k$ all present, and more generally $F_{i_2}^{i_1} F_{i_3}^{i_2} \cdots F_{i_n}^{i_{n-1}} F_{i_1}^{i_n}$ all present.

2. Iterations of the solutions of the IE

Another way to find the constraints which must be satisfied by the kernels F_j^i with $\lambda_i \lambda_j > 0$, is to consider the iterations of the IE Eq. (8). We limit our discussion to discrete kernels of the type (12b), (12c), each degenerate scalar kernel having only a finite number of terms: $F_j^i(x, y) = \sum_{m=1}^{m_0} g_{j,m}^i(x) h_{j,m}^i(y)$, m_0 finite. In order to determine the constraints, as we shall see, the first iteration plays a central role. On the other hand, can we hope to remove some of the above restrictions on $F_j^i \in \mathcal{J}^+$ by the introduction of more than one term in the degenerate kernels? We rewrite the IE, Eq. (8) in the form

$$K(x, y) = \sum_0^\infty K_n(x, y), \quad K_0 = \mathcal{J},$$
$$K_1 = \int_{-\infty}^{+\infty} \tilde{\mathcal{J}}(s, y) \mathcal{J}(x, s)\, ds, \tag{13}$$
$$(K_1)_j^i = \int_{-\infty}^{+\infty} \sum_l ds\, F_l^i(s, y) F_j^l(x, s)\, \theta(s - x_l)\, \mu_l^{-1}(s),$$

where, for the first iteration K_1, we have still written down the ith row, jth column term. We must require that all these terms be integrable when $s \to \infty$ or

$$\forall\ (i, j), \quad \sum_l F_l^i(s, y) F_j^l(x, s)\, \mu_l^{-1}(s) \underset{s \to \infty}{\to} 0 \tag{13'}$$

sufficiently quickly in order that the integral exists and we must check the compatibility of these conditions.

In this way, we again find the above rules for $F_j^i \in \mathcal{J}^+$. From the elements (i, i), (j, j) we get $\int_x^\infty \mu_j^{-1} F_j^i(s, y) F_i^j(x, s)\, ds < \infty$, $\int_x^\infty \mu_i^{-1} F_i^j(s, y) F_j^i(x, s)\, ds \leq \infty$ which leads to contradictory conditions for the considered kernels (12c).

On the other hand, even if we have more than one term for these kernels, they carry different x and y dependence and so do not appear to be cancellable non-integrable terms. From the elements (i, k), (k, j), (j, i) we get that $\int_x^\infty \mu_j^{-1} F_j^i(s, y) F_k^j(x, s)\, ds < \infty$, $\int_x^\infty \mu_i^{-1} F_i^k(s, y) F_i^j(x, s)\, ds < \infty$, $\int_x^\infty \mu^{-1} F_k^j(s, y) F_k^k(x, s)\, ds < \infty$, also lead to contradictory conditions if F_j^i, F_k^j, F_i^k belong to $\mathcal{J}^+$ and are all present; and so on.

Once those $F_j^i \in \mathcal{J}^+$ which cannot be present have been eliminated, we have still to find the conditions for the remaining ones $F_j^i \in \mathcal{J}^+$. From Eq. (13') we also get the constraints on the coefficients $\nu_{j,m}^i$ of these nonzero $F_j^i \in \mathcal{J}^+$ by requiring that all (i, j) elements of Eq. (13) be integrable. In this way, we find that these $(\nu_{j,m}^i)$ corresponding to $F_j^i \in \mathcal{J}^+$ [Eq. (12c)] are restricted by the values of those $(\nu_{j,m}^i)$ corresponding to $F_j^i \in \mathcal{J}^-$ [Eq. (12b)].

There is an interesting property that we get from the iteration of the IE. Let us first consider

$$\int_{-\infty}^{+\infty} \tilde{\mathcal{J}}(s, v) \tilde{\mathcal{J}}(w, s)\, ds$$
$$= \int_{-\infty}^{+\infty} ds \sum_l F_l^i(s, y) F_j^l(w, s)\, \theta(s - x_l)\, \theta(w - x_j)\, \mu_l^{-1}(s)\, \mu_j^{-1}(w) \tag{14}$$

for v and w fixed. We remark that for degenerate kernels (with m_0 finite) as those considered in this section, the integrability conditions when $s \to \infty$ are both the same in Eq. (13') and Eq. (14). K_n can be written

$$\int \tilde{\mathcal{J}}(u_1, y) \tilde{\mathcal{J}}(u_2, u_1)\, du_1 \cdots \int \tilde{\mathcal{J}}(u_{i+1}, u_i) \times \tilde{\mathcal{J}}(u_{i+2}, u_{i+1})\, du_{i+1} \cdots \int \tilde{\mathcal{J}}(u_n, u_{n-1})\, \mathcal{J}(x, u_n)\, du_n.$$

We see that the same integrability condition Eq. (13') or Eq. (14) is repeated in the general case and so we have no new condition in order that K_n exists. After the conditions which guarantee the existence of the iterations of the IE are obtained we still have to look after the other conditions Eq. (3), (9). We remark that our

choice of the y dependence of the kernels F_j^i in Eq. (12b) and Eq. (12c) is such that $\lim_{y\to\infty}\tilde{\mathcal{J}}(s,y)=0$, further y appears in $K_n(x,y)$ only from the first matrix $\tilde{\mathcal{J}}(u_1,y)$. It follows that $K_n(x,s)\to 0$ when $y\to 0$. Similarly the behavior of $\tilde{\mathcal{J}}(s,y)K_n(x,s)$ when s is large is given by $\tilde{\mathcal{J}}(s,y)\,\tilde{\mathcal{J}}(u_1,s)$ and so goes to zero when $s\to\infty$.

At the end we add some remarks concerning a more simple case $\mu_i(x)\equiv 1$ in Eq. (1), which, for simplicity, will be considered in the next section.

(i) This case is equivalent to $\mu_i(x)\equiv\mu(x)$ because this last case by a change of variable $\bar{x}=\int^x\mu^{-1}(u)\,du$ is reduced to the system (1) with $\Delta_0=(\delta_{ij}\partial/\partial\bar{x})$.

(ii) If in Eq. (1), ψ becomes $\psi\exp(i\phi\text{const.}\ x)$ then $\Lambda\to\Lambda+\text{const.}\ I$.

(iii) If further, $\lambda_i=\lambda_j$ for some (i,j) values, then from Eq. (7) we see that $q_j^i=0$. We will put $F_j^i=0$ for the scalar kernels in Eq. (8) corresponding to the (i,j) values such that $\lambda_i=\lambda_j$.

III. NLPDE FOR THE SOLUTIONS OF THE IE WHEN THE KERNELS SATISFY LPDE AND $\mu_i\equiv 1$ IN Eq. (1)

We consider $\lambda_i<0$, $i=1,\ldots,p-1$ and $\lambda_i=\lambda_p>0$, $i=p+1,\ldots,0$ where all the λ_i, except the last one λ_p, have the same sign. We rewrite $\mathcal{J}$ and K.

(i) We can write K as a sum of two terms: $K_{\rm I}$ where all the elements of the rows $p,p+1,\ldots,n$ are zero and $K_{\rm II}$ where all the elements of the first row, the second row,..., the $(p-1)$th row are zero,

$$K_{\rm I}=\begin{pmatrix} K_1^1\cdots K_n^1\\ \cdots\cdots\cdots\\ K_1^{p-1}\circ\circ\circ K_n^{p-1}\\ \bigcirc\end{pmatrix},\quad K_{\rm II}=\begin{pmatrix}\bigcirc\\ K_1^p\circ\circ\circ K_n^p\\ \cdots\cdots\cdots\\ K_1^n\cdots K_n^n\end{pmatrix}.$$

(ii) We can also write $\mathcal{J}$ as a sum of two terms: $\mathcal{J}_{\rm I}$ where all the elements of the rows $p,p+1,\ldots,n$ are zero and $\mathcal{J}_{\rm II}$ where all the elements of the first row, the second row,..., the $(p-1)$th row are zero.

Let us go on with our two component formulation $K_{\rm I}$ and $K_{\rm II}$, $\mathcal{J}_{\rm I}$, $\mathcal{J}_{\rm II}$ and define $\mathcal{J}_{\rm I}^{\tilde{}}$, $\mathcal{J}_{\rm II}^{\tilde{}}$ when $\lambda_i\lambda_j<0$, $\mathcal{J}_{\rm I}^{+}$, $\mathcal{J}_{\rm II}^{+}$ when $\lambda_i\lambda_j>0$ ($F_j^i,F_k^i,\ldots,F_l^i$ are not all present). We get:

$\mathcal{J}_{\rm II}^{+}=0$ because the corresponding F_j^i have $i\geq p$, $j\geq p$ and so $\lambda_i/\lambda_j=1$,

$\mathcal{J}_{\rm I}=\mathcal{J}_{\rm I}^{+}+\mathcal{J}_{\rm I}^{\tilde{}}$, $\mathcal{J}_{\rm II}=\mathcal{J}_{\rm II}^{\tilde{}}$,

$\mathcal{J}_{\rm I}^{\tilde{}}K_{\rm I}=\mathcal{J}_{\rm I}^{+}K_{\rm II}=\mathcal{J}_{\rm II}^{\tilde{}}K_{\rm II}=(0)$,

$$\mathcal{J}_{\rm II}=\begin{bmatrix}\bigcirc\\ F_1^p\circ\circ\circ F_{p-1}^p\\ \vdots\quad\vdots\\ F_1^n\circ\circ\circ F_{p-1}^n\end{bmatrix},\quad \mathcal{J}_{\rm I}^{+}=\begin{bmatrix}0 & F_2^1 & \cdots & F_{p-1}^1\\ \vdots & & & \vdots\\ & & & F_{p-1}^{p-2}\\ F_1^{p-1} & \cdots & F_{p-2}^{p-1} & 0\\ & \bigcirc & & \end{bmatrix}$$

$$\mathcal{J}_{\rm I}^{\tilde{}}=\begin{bmatrix}\bigcirc & F_p^1 & \cdots & F_n^1\\ & \vdots & & \vdots\\ & F_p^{p-1} & \cdots & F_n^{p-1}\\ & \bigcirc & & \end{bmatrix}.$$

The IE is written

$$\begin{bmatrix}K_{\rm I}\\ K_{\rm II}\end{bmatrix}=\begin{bmatrix}\mathcal{J}_{\rm I}\\ \mathcal{J}_{\rm II}\end{bmatrix}+\int\begin{bmatrix}\mathcal{J}_{\rm I}^{+} & \mathcal{J}_{\rm I}^{\tilde{}}\\ \mathcal{J}_{\rm II} & 0\end{bmatrix}\begin{bmatrix}K_{\rm I}\\ K_{\rm II}\end{bmatrix}.\tag{8b}$$

The method closely follows the one used for a second-order eigenvalue problem.[2] We introduce a parameter t: $\mathcal{J}=\mathcal{J}(x,y;t)$ and take the derivative with respect to t of the IE (8b),

$$\begin{bmatrix}K_{{\rm I},t}\\ K_{{\rm II},t}\end{bmatrix}-\int\begin{bmatrix}\mathcal{J}_{\rm I}^{+} & \mathcal{J}_{\rm I}^{\tilde{}}\\ \mathcal{J}_{\rm II} & 0\end{bmatrix}\begin{bmatrix}K_{\rm I}\\ K_{\rm II}\end{bmatrix}=\begin{bmatrix}\mathcal{J}_{{\rm I},t}\\ \mathcal{J}_{{\rm II},t}\end{bmatrix}+\int\begin{bmatrix}\mathcal{J}_{{\rm I},t}^{+} & \mathcal{J}_{{\rm I},t}^{\tilde{}}\\ \mathcal{J}_{{\rm II},t} & 0\end{bmatrix}\begin{bmatrix}K_{\rm I}\\ K_{\rm II}\end{bmatrix}.\tag{15}$$

We will assume a LPDE with respect to the three variables $(x,y;t)$ for $\mathcal{J}$. Firstly we will get an integral equation for $(K_{{\rm I},t},K_{{\rm II},t})$ and secondly add the integral equation where the derivatives with respect to x and y are present. In this way we will get nlpde with respect to the three variables $(x,y;t)$ and at the end restrict to the case $x=y$. We put $\mathcal{J}_{\rm I}^{+}\equiv 0$ in Sec. III A (1) and reintroduce it in III A (2).

A. LPDE for $\mathcal{J}$ of the first order in $\partial/\partial x+\partial/\partial y$

(1) $\mathcal{J}_{\rm I}^{+}=0$: We define Γ and assume for the kernel,

$$\Gamma=\begin{bmatrix}\gamma_1 & & & & \\ & \gamma_2 & & & \\ & & \ddots & & \\ & & & \gamma_{p-1} & \\ & & & & \gamma\\ & & & & & \ddots\\ & & & & & & \gamma\end{bmatrix},$$

$$\Gamma^{-1}\gamma_p^{-1}\mathcal{J}_{{\rm I},t}^{\tilde{}}=O_s\mathcal{J}_{\rm I}^{\tilde{}},\quad O_s=\frac{\partial}{\partial x}+\frac{\partial}{\partial y},$$
$$\mathcal{J}_{{\rm II},t}\gamma_p^{-1}\Gamma^{-1}=O_s\mathcal{J}_{\rm II}\tag{16}$$

where $\gamma_1,\ldots,\gamma_p$ are fixed numbers whereas γ is an arbitrary value.

In Appendix A, Secs. A_0–A_1, we get an integral equation proportional to the IE, and comparing with Eq. (8b) we finally get

$$\begin{bmatrix}K_{{\rm I},t}\Gamma^{-1}\gamma_p^{-1}-\left(\frac{\partial}{\partial x}K_{\rm I}+\Gamma(I-\Lambda^{-1}\lambda_p)\frac{\partial}{\partial y}K_{\rm I}\Gamma^{-1}+\Lambda^{-1}\lambda_p\frac{\partial}{\partial y}K_{\rm I}\right)\gamma\\ K_{{\rm II},t}\Gamma^{-1}\gamma_p^{-1}-O_sK_{\rm II}\end{bmatrix}$$
$$+\begin{bmatrix}K_{\rm I}\\ K_{\rm II}\end{bmatrix}(I-\Lambda^{-1}\lambda_p)(\Gamma\hat{K}_{\rm I}\Gamma^{-1}-\hat{K}_{\rm I})=(0)\tag{17}$$

where $\hat{K}_{\rm I}=K_{\rm I}(x,x)$.

TABLE I. Any p value, $n \geq p$, $\hat{K}^i_j = K^i_j(x, x; t)$.

$i \leq p-1, \quad p \leq j \leq n$

$$\left((\gamma_i\gamma_p)^{-1}\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right)\begin{pmatrix}\hat{K}^i_j \\ \hat{K}^i_i\end{pmatrix} = \sum_{l\neq i}^{l=p-1}\left(\frac{\lambda_p}{\lambda_l}-1\right)\left(\frac{\gamma_l}{\gamma_i}-1\right)\begin{pmatrix}\hat{K}^l_j & \hat{K}^l_i \\ \hat{K}^l_l & \hat{K}^l_j\end{pmatrix}$$

$i \leq p-1, \quad j \leq p-1, \quad i \neq j$

$$A^i_j = \gamma_i\gamma_j(\lambda_i - \lambda_j) + \gamma_j\gamma_p(\lambda_j - \lambda_p) + (\lambda_p - \lambda_i)\gamma_p\gamma_i$$

$$0 = \left(\frac{1}{\gamma_i\gamma_j}\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right)\hat{K}^i_j + \frac{\gamma_p}{\gamma_i}\left(\frac{\lambda_p}{\lambda_i}-1\right)^{-1}\left[-\frac{A^i_j\hat{K}^i_{j,x}}{\lambda_i\gamma_j\gamma_p} + \sum_{\substack{l=n \\ l\neq i \\ \neq j}}\left(\frac{\lambda_p}{\lambda_l}-1\right)\left(\frac{\gamma_i}{\gamma_j}-1\right)\left(\frac{\lambda_j}{\lambda_l}-1\right)\hat{K}^i_l\hat{K}^l_j - \sum_{\substack{l=p-1 \\ l\neq j}}\left(\frac{\lambda_p}{\lambda_l}-1\right)\left(\frac{\lambda_j}{\lambda_i}-1\right)\left(\frac{\gamma_i}{\gamma_j}-1\right)\hat{K}^l_i\hat{K}^j_l\right]$$

In Table I for arbitrary p and $n \geq p$ we write the nlpde corresponding to the "potentials" $\hat{K}^i_j$ ($i \neq j$ or not both i and j higher or equal to p). The evolution equations are written down for the two variables $(x, y = x; t)$.

In both Secs. III A and III B the method is similar. We put the K^i_j in three groups. In the first one, $K^i_j \in K_{II}$ and we get directly the nlpde. In the second group, $K^i_j \in K_I$ but $j \geq p$ and we fix our free parameter γ to be $\pm\gamma_i$. In both cases, the corresponding $F^i_j \neq 0$ and the linear part of the evolution equation is the same as that of F^i_j. In the last group, with $j \leq p-1$, we must consider a linear combination with the solution of the IE: If we restrict ourselves to $x = y$ (Table I), $O_s K^i_j$ is replaced by $\partial/\partial x(\hat{K}^i_j)$, K^i_j by $\hat{K}^i_j$ and we see that the nonlinear part of the equations represent triad and multitriad interactions.

(2) In Appendix A_2 we introduce $\mathcal{J}^+_I \neq 0$, with $\Gamma^{-1}\mathcal{J}^+_{I,t}\Gamma^{-1} - O_s\mathcal{J}^+_I = 0$, into the above formalism. We must have a relation $A^i_j = 0$, A^i_j being defined in Table I, which transforms the linear part of the $\hat{K}^i_j$ equations for $i \leq p-1$, $j \leq p-1$ into just

$$\left(\frac{1}{\gamma_i\gamma_j}\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right)\hat{K}^i_j.$$

B. LPDE for $\mathcal{J}$ of the second order in $\partial/\partial x + \partial/\partial y$

(1) $\mathcal{J}^+_I \equiv 0$: We consider the same Γ matrix as above in Eq. (16) but now the kernels $\mathcal{J}$ will satisfy lpde of the second order in (x, y),

$$\Gamma^{-1}\gamma_p^{-1}\mathcal{J}^-_{I,t} = O_s^2\mathcal{J}^-_I, \quad O_s^2 = \left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right)^2, \tag{18}$$

$$\mathcal{J}_{II,t}\gamma_p^{-1}\Gamma^{-1} = -O_s^2\mathcal{J}_{II}.$$

In Appendix B_1 finally we get an integral equation which represents the IE multiplied by some constant and we deduce easily the solution: $X = Y + Z$,

$$X = \begin{pmatrix}K_{I,t}\Gamma^{-1}\gamma_p^{-1} - (I - \lambda_p\Lambda^{-1})^2\,\Gamma\frac{\partial^2 K_I}{\partial y^2}\,\Gamma^{-1} + (O_{\Lambda_p})^2 K_I\end{pmatrix}, \qquad K_{II,t}\Gamma^{-1}\gamma_p^{-1} + O_s^2 K_{II}$$

$$Y = \begin{pmatrix}O_{\Lambda_p}K_I \\ O_s K_{II}\end{pmatrix}(I - \Lambda^{-1}\lambda_p)(-\hat{K}_I + \Gamma\hat{K}_I\Gamma^{-1}),$$

$$Z = \begin{pmatrix}K_I \\ K_{II}\end{pmatrix}(-I + \lambda_p\Lambda^{-1})\left[\frac{d}{dx}\hat{K}_I + (O_{\Lambda_p}K_I)_{y=x}\right.$$

$$+ (I - \Lambda^{-1}\lambda_p)\Gamma\left(\frac{\partial}{\partial y}K_I\right)_{y=x}\Gamma^{-1}$$

$$\left. + \hat{K}_I(I - \Lambda^{-1}\lambda_p)(\hat{K}_I - \Gamma\hat{K}_I\Gamma^{-1})\right] \tag{19}$$

where

$$O_{\Lambda_p} = \frac{\partial}{\partial x} + \lambda_p\Lambda^{-1}\frac{\partial}{\partial y}.$$

X represents the linear part of the equations while Y and Z are the nonlinear parts. This is our fundamental result from which we can get nlpde whose linear part contains second order derivatives $\partial^2/\partial x^2$. We could, as in the previous section, consider any p and n values, however, for simplicity we make explicit in Appendix B_1 only the cases $p = 3$ and 4 for any $n \geq p$.

If we restrict ourselves to $y = x$ (Table II), O_s^2 must be replaced by $\partial^2/\partial x^2$, O_s by $\partial/\partial x$, and K^i_j by $\hat{K}^i_j$. We see that the nonlinear terms contain cubic and triad interactions (of a type different than that of subsection A because derivatives of $\hat{K}^i_j$ are present).

(2) In Appendix B_2 we introduce $\mathcal{J}^{(+)}_I$, with $\Gamma^{-1}\mathcal{J}^+_{I,t}\Gamma^{-1} - O_s\mathcal{J}^+_I = 0$. We must have a relation $\gamma_i\gamma_j(\lambda_i - \lambda_j)^2 - \gamma_i\gamma_p(\lambda_i - \lambda_p)^2 + \gamma_j\gamma_p(\lambda_j - \lambda_p)^2 = 0$ which transforms the linear part of the $\hat{K}^i_j$ equations for $i \leq p-1$, $j \leq p-1$ into

$$\left((\gamma_i\gamma_j)^{-1}\frac{\partial}{\partial t} - \frac{\partial^2}{\partial x^2}\right)\hat{K}^i_j.$$

C. Representations of the scalar kernels F^i_j

At the end of this section we want to make explicit the different representations of the scalar kernels F^i_j in subsections A and B, where they have to satisfy either Eq. (16) or Eq. (18). As usual F^i_j can include a continuum part as well as a discrete part,

$$F^i_j = \int d\nu^i_j\alpha^i_j(\nu^i_j)\exp i\nu^i_j[x - \lambda_i\lambda_j^{-1}y + \epsilon t(1 - \lambda_i\lambda_j^{-1})^n\gamma^i_j] \tag{20a}$$

or

$$F^i_j = \sum_m \alpha^i_{j,m}\exp\eta\nu^i_{j,m}[x - \lambda_i\lambda_j^{-1}y + \epsilon t(1 - \lambda_i\lambda_j^{-1})^n\gamma^i_j], \quad \nu^i_{j,m} > 0. \tag{20b}$$

Also here we always have λ_i (or λ_j) $= \lambda_p$ if i (or j) $\geq p$.

Section III A (1): $n = 1$, $\epsilon = +1$, $\gamma^i_j = \gamma_i\gamma_p$ if $F^i_j \in \mathcal{J}^-_I$. $\gamma^i_j = \gamma_j\gamma_p$ if $F^i_j \in \mathcal{J}_{II}$, $\eta = -1$.

TABLE II. $n \geq p$, $p=3$ and 4, $\hat{K}_j^i = K_j^i(x,x;t)$.

$l=1,2,\ldots,p-1, \quad j=p,p+1,\ldots,n$

$$(\gamma_p\gamma_l)^{-1}\begin{pmatrix}\hat{K}_{l,t}^j\\ -\hat{K}_{j,t}^l\end{pmatrix}+\frac{\partial^2}{\partial x^2}\begin{pmatrix}\hat{K}_l^j\\ \hat{K}_j^l\end{pmatrix}=2\left(\frac{\lambda_p}{\lambda_l}-1\right)\left(\frac{\lambda_l}{\lambda_p}-1\right)\sum_{q=p}^{n}\begin{pmatrix}\hat{K}_l^j & \hat{K}_q^l & \hat{K}_l^q\\ \hat{K}_j^l & \hat{K}_q^q & \hat{K}_l^q\end{pmatrix}$$

$$+\sum_{\substack{m=1\\ m\neq l}}^{p-1}\left[\left(\frac{\lambda_l}{\lambda_m}-1\right)\begin{pmatrix}\hat{K}_m^m & \hat{K}_{\lambda m,x}^l\\ \hat{K}_m^j & K_{j,x}^m\end{pmatrix}+\left(\frac{\lambda_p}{\lambda_l}-1\right)\left(\frac{\lambda_l}{\lambda_m}-1\right)\begin{pmatrix}\hat{K}_l^j & \hat{K}_m^l & \hat{K}_l^m\\ \hat{K}_j^l & \hat{K}_l^m & \hat{K}_m^l\end{pmatrix}+\left(\frac{\lambda_l}{\lambda_p}-1\right)\left(\frac{\lambda_p}{\lambda_m}-1\right)\sum_{q=p}\begin{pmatrix}\hat{K}_m^j & \hat{K}_q^m & \hat{K}_l^q\\ \hat{K}_m^m & \hat{K}_m^q & \hat{K}_q^l\end{pmatrix}\right]$$

$p=3, \quad l=1,2, \quad l'=1,2, \quad l'\neq l, \quad j=3,4,\ldots,n,$

$p=4, \quad l=1,2,3, \quad l'=1,2,3, \quad l''=1,2,3, \quad l\neq l'\neq l''\neq l, \quad j=4,5,\ldots,n$

$$\delta_{l'}^{l}=1+\frac{\lambda_p}{\lambda_l}-\frac{\lambda_{l'}}{\lambda_l}-\frac{\lambda_p}{\lambda_{l'}}, \quad \nu_{l'}^{l}=2-\frac{\lambda_l}{\lambda_{l'}}-\frac{\lambda_{l'}}{\lambda_l},$$

$$\left(\delta_l^{l''}\frac{\partial^2}{\partial x^2}+\nu_l^{l''}(\gamma_l\gamma_p)^{-1}\frac{\partial}{\partial t}\right)\hat{K}_l^{l''}=\delta_l^{l''}\sum_{q=p}^{n}\hat{K}_q^l\hat{K}_{q,x}^{l''}\left(\frac{\lambda_l}{\lambda_p}-1\right)-\delta_{l'}^{l}\sum_{q=p}^{n}\hat{K}_l^{l''}\hat{K}_{l,x}^{q}\left(\frac{\lambda_l}{\lambda_p}-1\right)+\hat{K}_l^{l''}\hat{K}_{l''\,,x}^{l''}\left(\frac{\lambda_l}{\lambda_{l''}}-1\right)\delta_l^{l''}$$

$$+\hat{K}_{l,x}^{l''}\hat{K}_{l''}^{l'}\left[-\delta_l^{l'}\left(\frac{\lambda_l}{\lambda_{l''}}-1\right)+\nu_{l'}^{l}\left(\frac{\lambda_p}{\lambda_{l''}}-1\right)+\delta_l^{l'}\left(\frac{\lambda_l}{\lambda_{l'}}-1\right)\hat{K}_l^{l'}\left[\sum_{\substack{m=1\\ m\neq l'}}^{p-1}\hat{K}_m^{l'}\hat{K}_{l'}^{m}\left(\frac{\lambda_{l'}}{\lambda_m}-1\right)+\sum_{q=p}^{n}\hat{K}_q^{l'}\hat{K}_{l'}^{q}\left(\frac{\lambda_{l'}}{\lambda_l}-1\right)\right.\right.$$

$$\left.\left.+\nu_{l'}^{l}\hat{K}_l^{l'}\left(\frac{\lambda_p}{\lambda_l}-1\right)\left[\sum_{m\neq l}^{p-1}\hat{K}_m^{l}\hat{K}_l^{m}\left(\frac{\lambda_l}{\lambda_m}-1\right)+\sum_{q=p}^{n}\hat{K}_q^{l}\hat{K}_l^{q}\left(\frac{\lambda_l}{\lambda_q}-1\right)\right]\right]\right]$$

Section III B (1): $n=2$, if $F_j^i \in \mathcal{J}_I$, then $\epsilon=+1$, $\gamma_j^i=\gamma_i\gamma_p$. If $F_j^i \in \mathcal{J}_{II}$, then $\epsilon=-1$, $\gamma_j^i=\gamma_j\gamma_p$, $\eta=-1$.

In Sec. III A (2) [or III B (2)] we only consider (20b) with $\eta=+1$ and the other values are those in Sec. III A (1) [or III B (1)]. Further, the $\nu_{j,m}^i$ must be constrained (see Sec. II). For instance, $\int_x^\infty \mathcal{J}(s,y)\mathcal{J}(x,s)\,ds$ must exist. Further, if $F_j^i \in \mathcal{J}_I^*$, $\gamma_j^i=\gamma_i\gamma_j$, in Sec. III A (2), $\epsilon=+1$, in Sec. III B (2), $\epsilon=+1$ if $i<j$, $\epsilon=-1$ if $i>j$.

IV. CONCLUSION

In this paper, as well as in the preceding one, our motivation was to show that for a linear differential matrix case a purely algebraic method settled previously in the scalar case (see Ref. 4) could provide a derivation of the IE without any study of the complex eigenvalue plane. Due to the complexity of the general case we think that the interest of the method is more convincing here. In our approach, going from a three to a four order eigenvalue case is not an important complication whereas this is not so clear from Kaup's[4] results, working in the complex eigenvalue plane. The jump is really from two to three, not in the algebraic aspect of the problem, but due to the boundary conditions to be satisfied by the scalar kernels of the IE.

We have seen that there are two classes of such kernels $F_j^i(x,y)$ defined by the sign of their associated $\lambda_i\lambda_j$ product. In the first class $\lambda_i\lambda_j<0$, F_j^i goes to zero either when x or y goes to infinity and we recover all the nice properties of the previous case.[2] For the second class ($\lambda_i\lambda_j>0$), for simplicity and in order to see more clearly the drawbacks of the theory, we restrict our study to discrete kernels where they are going to zero when $y\to\infty$ but to infinity when $x\to-\infty$. In the iteration of the IE these diverging terms must be compensated in order that the solution exist. We find that this is not possible if in this class both F_j^i, F_l^j are present or F_j^i, F_k^j, F_l^k are present. Even if we satisfy these requirements, the parameters of these kernels have to be constrained by parameters of other kernels belonging to the same class or to the other class ($\lambda_i\lambda_j<0$).

For the derivation of the evolution equation we first consider only the kernels F_j^i of the first class ($\lambda_i\lambda_j<0$) and we get the expected nlpde. Introducing the other kernels does not change the equations, but only the velocities of the corresponding (same i,j as F_j^i) potentials K_j^i.

This work complements the Ablowitz–Haberman[3d] paper because it provides both the IE and it constructs explicitly the nlpde's that these authors have written down [in particular for the nlpde which are studied in Sec. II B and which were not considered in Ref. 3(c)].

Zakharov and Shabat[3c] have previously shown that purely algebraic method were sufficient to derive the nlpde and they got that the kernels of their integral equations have to satisfy lpde. However they postulate an inversionlike equation and the connection with the associated linear system is not expressed explicitly. In their formal work, no conditions about the $\{\lambda_i\}$ of the associated linear system appear whereas in the present work it is the main ingredient (and the main difficulty) of the formalism, concerning its validity or its breakdown.

In a recent work Newell and Redekopp[5] said that the breakdown of the Zakharov–Shabat theory is linked to the problem with more than one spatial dimension.

Recently[6a] we have obtained a generalization of the IE, with more than one coordinate, in the case where in Eq. (1) the differential operator $\delta_{ij}\mu_j(x)(\partial/\partial x)$ is replaced by the partial differential one $\delta_{ij}(\partial/\partial x_i)$. However concerning the validity of the formalism we get the

same discussion as we get here, depending upon whether $\lambda_i\lambda_j < 0$ or $\lambda_i\lambda_j > 0$. Moreover[6b] we have now shown in this multidimensional case that there exists another IE which avoids these limitations due to $\lambda_i\lambda_j > 0$ although another type of distinction occurs in both cases where the number of different coordinates is two or more than two. Unfortunately this other IE coincides with the previous one in the one coordinate case.

APPENDIX A

A_0: Properties of the solution of the IE when $\mu_j \equiv 1$, $\lambda_i < 0$ for $i < p$, $\lambda_i = \lambda_p > 0$ for $i \geq p$.

We establish a property, for the solution of Eq. (8), which will be useful later in the derivation of the nlpde. Let us define $\Lambda_t = \lambda_t I$ and

$$
\begin{aligned}
&O_{\Lambda_t} = \frac{\partial}{\partial x} + \Lambda^{-1}\Lambda_t \frac{\partial}{\partial y}\,, \quad (O_{\Lambda_t}K)\hat{} = (O_{\Lambda_t}K(x,y))_{y=x}\,,\\
&O_{\Lambda_t}K = O_{\Lambda_t}\mathcal{J} + \mathcal{J}(-I+\Lambda^{-1}\Lambda_t)\hat{K} + \int \mathcal{J} O_{\Lambda_t}K,\\
&O^2_{\Lambda_t}K = O^2_{\Lambda_t}\mathcal{J} + O_{\Lambda_t}(\mathcal{J}(-I+\Lambda^{-1}\Lambda_t)\hat{K})\\
&\qquad + \mathcal{J}(-I+\Lambda^{-1}\Lambda_t)(O_{\Lambda_t}K)\hat{} + \int \mathcal{J} O^2_{\Lambda_t}K.
\end{aligned}
\tag{$A_0$1}
$$

In particular, if λ_p is the only positive eigenvalue, Eq. ($A_0$1) can be written in a two-component formulation, noting that $O_{\Lambda_p}\mathcal{J}_I = 0$, $O_{\Lambda_p}K_{II} = O_s K_{II}$, $(I-\Lambda^{-1}\Lambda_p)K_{II} = 0$, we get

$$
\begin{aligned}
\begin{pmatrix} O_{\Lambda_p} & K_I \\ O_s & K_{II}\end{pmatrix} &= \begin{pmatrix} O_{\Lambda_p} & \mathcal{J}_I^+ \\ O_s & \mathcal{J}_{II}\end{pmatrix} + \begin{pmatrix} \mathcal{J}_I^+ \\ \mathcal{J}_{II}\end{pmatrix}(-I+\Lambda^{-1}\Lambda_p)\hat{K}_I\\
&+ \int \begin{pmatrix} \mathcal{J}_I^+, & \mathcal{J}_I^- \\ \mathcal{J}_{II}, & 0\end{pmatrix}\begin{pmatrix} O_{\Lambda_p} & K_I \\ O_s & K_{II}\end{pmatrix},
\end{aligned}
\tag{$A_0$2}
$$

where

$$
\begin{aligned}
&K = \left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right)K,\\
&\begin{pmatrix} O^2_{\Lambda_p} & K_I \\ O^2_s & K_{II}\end{pmatrix} = \begin{pmatrix} O^2_{\Lambda_p} & \mathcal{J}_I^+ \\ O^2_s & \mathcal{J}_{II}\end{pmatrix} + \begin{pmatrix} \mathcal{J}_I \\ \mathcal{J}_{II}\end{pmatrix}(-I+\Lambda^{-1}\lambda_p)\\
&\quad \times \left((O_{\Lambda_p}K_I)\hat{} + \frac{d}{dx}\hat{K}_I\right) + \begin{bmatrix} O_{\Lambda_p} & \mathcal{J}_I^+ \\ O_s & \mathcal{J}_{II}\end{bmatrix}\\
&\quad \times (-I+\Lambda^{-1}\lambda_p)\hat{K}_I + \int \begin{bmatrix} \mathcal{J}_I, & \mathcal{J}_I^- \\ \mathcal{J}_{II}, & 0\end{bmatrix}\begin{bmatrix} O^2_{\Lambda_p} & K_I \\ O^2_s & K_{II}\end{bmatrix}
\end{aligned}
\tag{$A_0$3}
$$

A_1: Let us define $O_\Lambda \mathcal{J} = (\partial/\partial x)\mathcal{J} + \Lambda^{-1}(\partial/\partial y)\mathcal{J}\Lambda$. From Eq. (10), $O_\Lambda \mathcal{J} = 0$ and taking into account the structure of $\mathcal{J}$, Eq. (16) can be rewritten:

$$
\begin{aligned}
&\mathcal{J}^-_{I,x} + \Lambda^{-1}\lambda_p \mathcal{J}^-_{I,y} = 0, \quad \Gamma^{-1}\gamma_p^{-1}\mathcal{J}^-_{I,t} = (I-\Lambda^{-1}\lambda_p)\mathcal{J}^-_{I,\gamma},\\
&\mathcal{J}_{II,x} + \lambda_p^{-1}F_{II,y}\Lambda = 0.
\end{aligned}
\tag{$A_1$1}
$$

At the rhs of Eq. (15) we substitute for $\mathcal{J}^-_{I,t}$, $\mathcal{J}_{II,t}$ the expressions written down in Eq. ($A_1$1). In this way we get free terms plus an integral type term where $\mathcal{J}$ is the kernel. We get the following integral equation (here $\mathcal{J}_I^+ \equiv 0$),

$$
\begin{aligned}
&\begin{bmatrix} K_{I,t} \\ K_{II,t}\end{bmatrix} + \Gamma\gamma_p(-I+\Lambda^{-1}\lambda_p)K_{I,y}\Big]\\
&= \begin{pmatrix} 0 \\ \mathcal{J}_{II,t}\end{pmatrix} - \begin{pmatrix} 0 \\ \mathcal{J}_{II}\end{pmatrix}(-I+\Lambda^{-1}\lambda_p)\Gamma\gamma_p\hat{K}_I\\
&\quad + \int \begin{pmatrix} 0 & \mathcal{J}_I^- \\ \mathcal{J}_{II} & 0\end{pmatrix}\begin{pmatrix} K_{I,t} + \Gamma\gamma_p(-I+\Lambda^{-1}\lambda_p)K_{I,y} \\ K_{II,t}\end{pmatrix},
\end{aligned}
\tag{$A_1$2}
$$

which has a structure similar to the IE Eq. (8b) except the presence of $\mathcal{J}_{II,t}$ in the free terms. Multiplying Eq. ($A_1$2) at the right by $\Gamma^{-1}\gamma_p^{-1}$ and subtracting Eq. ($A_0$2), due to our assumption Eq. (16), $\mathcal{J}_{II,t}\gamma_p^{-1}\Gamma^{-1} - O_s\mathcal{J}_{II}$ disappears. In this way we get an integral equation proportional to the IE and comparing with Eq. (8b) we finally get Eq. (17).

A_2: We introduce $\mathcal{J}_I^{(*)} \neq 0$ into the formalism. In supplement of Eq. (16) we assume

$$
\Gamma^{-1}\mathcal{J}^+_{I,t}\Gamma^{-1} = O_s\mathcal{J}_I^+. \tag{$A_2$1}
$$

Trying to get integral equations for quantities linked to $(K_{I,t}, K_{II,t})$ with $\mathcal{J}$ as kernel, it turns out that the following identity must be satisfied,

$$
\begin{aligned}
&\Gamma(\mathcal{J}_I^+ - \Lambda\,\mathcal{J}_I^+\Lambda^{-1})\,\Gamma + \Gamma\gamma_p(\Lambda\,\mathcal{J}_I^+\Lambda^{-1} - \lambda_p\,\mathcal{J}_I^+\Lambda^{-1})\\
&\quad - \mathcal{J}_I^+\Gamma\gamma_p(I-\Lambda^{-1}\lambda_p) = (0),
\end{aligned}
$$

or

$$
\gamma_i\gamma_j(\lambda_j - \lambda_i) + \gamma_i\gamma_p(\lambda_i - \lambda_p) - \gamma_j\gamma_p(\lambda_j - \lambda_p) = 0 \tag{$A_2$2}
$$

for (i row, j column) $\in \mathcal{J}_I^+$.

We get

$$
\begin{aligned}
&\begin{pmatrix} K_{I,t} - \gamma_p\Gamma(-I+\Lambda^{-1}\lambda_p)K_{I,y} \\ K_{II,t}\end{pmatrix}\\
&\quad - \int \begin{bmatrix} \mathcal{J}_I^+ & \mathcal{J}_I^- \\ \mathcal{J}_{II} & 0\end{bmatrix}\begin{bmatrix} K_{I,t} - \gamma_p\Gamma(-I+\Lambda^{-1}\lambda_p)K_{I,y} \\ K_{II,t}\end{bmatrix}\\
&\quad - \begin{pmatrix} \mathcal{J}_I^+ \\ \mathcal{J}_{II}\end{pmatrix}\Gamma\gamma_p(-I+\Lambda^{-1}\lambda_p)\hat{K}_I\\
&\quad - \begin{pmatrix} \mathcal{J}^+_{I,t} - \Gamma\gamma_p(I-\Lambda^{-1}\lambda_p)\mathcal{J}^+_{I,y} \\ \mathcal{J}_{II,t}\end{pmatrix} = (0).
\end{aligned}
\tag{$A_2$3}
$$

We multiply ($A_2$3) at the left by $\gamma_p^{-1}\Gamma^{-1}$, subtract Eq. ($A_0$2), and get

$$
\begin{aligned}
&\begin{bmatrix} K_{I,t}\gamma_p^{-1}\Gamma^{-1} + \Gamma(-I+\Lambda^{-1}\lambda_p)K_{I,y} - O_{\Lambda_p}K_I \\ K_{II,t}\gamma_p^{-1}\Gamma^{-1} - O_s K_{II}\end{bmatrix}\\
&\quad - \int \begin{pmatrix} \mathcal{J}_I^+ & \mathcal{J}_I^- \\ \mathcal{J}_{II} & 0\end{pmatrix}\begin{pmatrix} K_{I,t}\gamma_p^{-1}\Gamma^{-1} + \Gamma(-I+\Lambda^{-1}\lambda_p)K_{I,y} - O_{\Lambda_p}K_I \\ K_{II,t}\gamma_p^{-1}\Gamma^{-1} - O_s K_{II}\end{pmatrix}\\
&\quad - \begin{pmatrix} \mathcal{J}_I^+ \\ \mathcal{J}_{II}\end{pmatrix}(I-\Lambda^{-1}\lambda_p)(\hat{K}_I - \Gamma\hat{K}_I\Lambda^{-1})\\
&= \begin{bmatrix} \mathcal{J}^+_{I,t}\gamma_p^{-1}\Gamma^{-1} - \Gamma(I-\Lambda^{-1}\lambda_p)\mathcal{J}^+_{I,y}\Gamma^{-1} - O_{\Lambda_p}\mathcal{J}_I^+ \\ \mathcal{J}_{II,t}\gamma_p^{-1}\Gamma^{-1} - O_s\mathcal{J}_{II}\end{bmatrix}.
\end{aligned}
\tag{$A_2$4}
$$

Taking into account ($A_2$1) and ($A_2$2) for the first row at the rhs of ($A_2$4) and Eq. (16) for the second row, we get that this rhs is zero. It follows that the solution of ($A_2$4) is the same as the corresponding one in Sec. III A (1), i.e., Eq. (17). Finally, taking into account ($A_2$2) in

Table I for $\hat{K}_j^i$ with $i \leq p-1$, $j \leq p-1$ we see that the linear part of the nlpde has the wanted form $[(\gamma_i\gamma_j)^{-1}\hat{K}^i_{j,t} - \hat{K}^i_{j,x}]$.

APPENDIX B

B_1: $\mathcal{J}_I^+ \equiv 0$, Eq. (18) can be rewritten

$$\Gamma^{-1}\gamma_p^{-1}\mathcal{J}_{I,t}^- = (I-\Lambda^{-1}\lambda_p)^2\frac{\partial^2}{\partial y^2}\mathcal{J}_I^-,$$
$$\mathcal{J}_{II,t} = -\frac{\partial^2}{\partial y^2}\mathcal{J}_{II}(I-\Lambda^{-1}\lambda_p)^2\Gamma\gamma_p. \qquad (B_1 1)$$

Substituting at the rhs of Eq. (15) these expressions for $\mathcal{J}_t$, we get

$$H = \begin{pmatrix} K_{I,t} - \Gamma\gamma_p(-I+\Lambda^{-1}\lambda_p)^2\frac{\partial^2}{\partial y^2}K_I \\ K_{II,t} \end{pmatrix},$$

$$H - \int\begin{pmatrix} 0 & \mathcal{J}_I^- \\ \mathcal{J}_{II} & 0 \end{pmatrix}H = \begin{pmatrix} 0 \\ \mathcal{J}_{II,t} \end{pmatrix} - \begin{pmatrix} 0 \\ \mathcal{J}_{II} \end{pmatrix}(-I+\lambda_p\Lambda^{-1})^2\gamma_p\Gamma \times (K_{I,y})_{x=y} + \begin{pmatrix} 0 \\ \mathcal{J}_{II,x} \end{pmatrix} (-I+\Lambda^{-1}\lambda_p)^2\gamma_p\Gamma\hat{K}_I. \qquad (B_1 2)$$

Equation ($B_1$2) has the same kernel $\mathcal{J}$ as the IE, however the free terms contain $\mathcal{J}_{II,t}$ and $\mathcal{J}_{II,x}$, not present in the IE. We multiply Eq. ($B_1$2) at the right by $\Gamma^{-1}\gamma_p^{-1}$ and add Eq. ($A_0$3). It follows that $\mathcal{J}_{II,t}\Gamma^{-1}\gamma_p + O_s^2\mathcal{J}_{II}$ disappears in the free terms. In this way we have still in the free terms, one proportional to $O_s\mathcal{J}_{II}$ which is eliminated by comparing with Eq. ($A_0$2). Finally we get an integral equation which represents the IE multiplied by some constant and we deduce easily the solution Eq. (19).

$p=3$: In that case the kernels $\mathcal{J}_I^-$ and $\mathcal{J}_{II}$ are:

$$\mathcal{J}_I^- = \begin{bmatrix} 0 & 0 & F_3^1 & \cdots & F_n^1 \\ 0 & 0 & F_3^2 & \cdots & F_n^2 \\ & & \text{(0)} & & \end{bmatrix}, \quad \mathcal{J}_{II} = \begin{bmatrix} \text{(0)} & \\ F_1^3, & F_2^3 \\ \vdots & \vdots \\ F_1^n, & F_2^n \end{bmatrix},$$

$$\mathcal{J}_I^+ = \begin{bmatrix} 0 & F_2^1 & \\ F_1^2 & 0 & \\ & \text{(0)} & \end{bmatrix} = (0)$$

and satisfy for $p=3$:

$$\left[(\gamma_I\gamma_p)^{-1}\frac{\partial}{\partial t} + O_s^2\right]F_I^j = [(\gamma_I\gamma_p)^{-1} - O_s^2]F_j^I = 0, \quad I = 1,2,\ldots,p-1,\ j = p,p+1,n.$$

It turns out that the eigenvalues λ_i and the velocities γ_i must be linked in such a way that

$$\left(\frac{\lambda_p}{\lambda_I}-1\right)\left(1-\frac{\gamma_I}{\gamma_{I'}}\right) = -1+\frac{\lambda_{I'}}{\lambda_I} \qquad (B_1 3)$$

for $p=3$, $l=1$ ($l'=2$) or $2(l'=1)$ which gives in fact only one independent relation. $p=4$: $\mathcal{J}_I^-$ and $\mathcal{J}_{II}$ satisfy Eq. ($B_1$1) for $p=4$ and are

$$\mathcal{J}_I^- = \begin{bmatrix} 0 & 0 & 0 & F_4^1 & \cdots & F_n^1 \\ 0 & 0 & 0 & F_4^2 & \cdots & F_n^2 \\ 0 & 0 & 0 & F_4^3 & \cdots & F_n^3 \\ & & & \text{(0)} & & \end{bmatrix}, \quad \mathcal{J}_I^- = \begin{bmatrix} \text{(0)} & \\ F_1^4 & F_2^4 & F_3^4 \\ \cdots & \cdots & \text{(0)} \\ F_1^n & F_2^n & F_3^n \end{bmatrix},$$

$$\mathcal{J}_I^- = \begin{bmatrix} 0 & F_2^1 & F_3^1 & \\ F_1^2 & 0 & F_3^2 & \text{(0)} \\ F_1^3 & F_2^3 & 0 & \\ & \text{(0)} & & \end{bmatrix} = (0) \text{ in Appendix } B_1.$$

Equation ($B_1$3) must be satisfied for $l=1$ ($l'=2$ or 3), $l=2$ ($l'=1$ or 3), $l=3$ ($l'=2$ or 1), which gives three relations among the (λ_i, γ_i).

B_2: $\mathcal{J}_I^+ \neq 0$. We assume Eq. (18) and

$$\Gamma^{-1}\mathcal{J}_{I,t}^+\Gamma^{-1} = \epsilon O_s^2\mathcal{J}_I^+, \qquad (B_2 1)$$

where for $F_j^i \in \mathcal{J}_I^+$, $\epsilon = +1$ if $i<j$ and $\epsilon = -1$ if $i>j$. In the following, for simplicity we restrict ourselves to $\epsilon = +1$, $i<j$ but the extension for the other case, $i>j$, is trivial. We get an integral equation for $(K_{I,t}, K_{II,t})$ with $\mathcal{J}$ as kernel if

$$\Gamma(\Lambda^2\mathcal{J}_I^+\Lambda^{-1}\Lambda^{-1})\Gamma - 2\Gamma\Lambda\mathcal{J}_I^+\Lambda^{-1}\Gamma - \Gamma\gamma_p(I-\Lambda^{-1}\lambda_p)^2\Lambda^2\mathcal{J}_I^+\Lambda^{-1}\Lambda^{-1} + \mathcal{J}_I^+(I-\Lambda^{-1}\lambda_p)^2\Gamma\gamma_p = (0)$$

or

$$\gamma_i\gamma_j(\lambda_i-\lambda_j)^2 - \gamma_i\gamma_p(\lambda_i-\lambda_p)^2 + \gamma_j\gamma_p(\lambda_j-\lambda_p)^2 = 0. \qquad (B_2 2)$$

We get

$$W = \begin{pmatrix} K_{I,t} - X\frac{\partial^2}{\partial y^2}K_I \\ K_{II,t} \end{pmatrix}, \quad X = (I-\lambda_p\Lambda^{-1})^2\gamma_p\Gamma$$

$$W - \int\mathcal{J}W = \begin{pmatrix} \mathcal{J}_{I,t}^+ - X\frac{\partial^2}{\partial y^2}\mathcal{J}_I^+ \\ \mathcal{J}_{II,t} \end{pmatrix} + \begin{bmatrix} \mathcal{J}_{I,x}^+ \\ \mathcal{J}_{II,x} \end{bmatrix}X\hat{K}_I - \begin{bmatrix} \mathcal{J}_I^+ \\ \mathcal{J}_{II} \end{bmatrix}X\left(\frac{\partial}{\partial y}K_I(x,y)\right)_{y=x}. \qquad (B_2 3)$$

We multiply ($B_2$3) at the left by $(\gamma_p\Gamma)^{-1}$ and add Eq. ($A_0$3),

$$X = \begin{pmatrix} K_{I,t}(\gamma_p\Gamma)^{-1} - (I-\lambda_p\Lambda^{-1})^2\Gamma\frac{\partial^2}{\partial y^2}K_I\Gamma^{-1} + O_{\Lambda_p}^2K_I \\ K_{II,t}(\gamma_p\Gamma)^{-1} + O_s^2K_{II} \end{pmatrix},$$

$$X - \int\mathcal{J}X - \begin{pmatrix} O_{\Lambda_p}\mathcal{J}_I^+ \\ O_s\mathcal{J}_{II} \end{pmatrix}(I-\Lambda^{-1}\lambda_p)(-\hat{K}_I + \Gamma\hat{K}_I\Gamma^{-1}) + \begin{pmatrix} \mathcal{J}_I \\ \mathcal{J}_{II} \end{pmatrix}X\left(\frac{\partial}{\partial y}K_I(x,y)\right)_{y=x}(\Gamma\gamma_p)^{-1} \qquad (B_2 4)$$

$$+(-I+\Lambda^{-1}\lambda_p)[\hat{K}_{\mathrm{I},x}-(O_{\Lambda_p}K(x,y))_{y=x}]$$

$$=\begin{pmatrix}\mathcal{J}^{+}_{\mathrm{I},t}(\gamma_p\Gamma)^{-1}-(I-\Lambda^{-1}\lambda_p)^2\Gamma\ \mathcal{J}^{+}_{\mathrm{I}}\Gamma^{-1}+O^2_{\Lambda_p}\mathcal{J}^{+}_{\mathrm{I}}\\ \mathcal{J}_{\mathrm{II},t}(\gamma_p\Gamma)^{-1}+O^2_{s}\mathcal{J}_{\mathrm{II}}\end{pmatrix}$$

Taking into account ($B_2$2), and Eq. (18), the rhs of ($B_2$4) is zero. It follows that we have the same solution of ($B_1$4) as that written down in Eq. (19). Finally, taking into account the relation ($B_2$2) we get for the expressions written in Table II,

$$\left(\delta^{l'}_{l}\frac{\partial^2}{\partial x^2}+\gamma^{l'}_{l}(\gamma_l\gamma_p)^{-1}\frac{\partial}{\partial t}\right)\hat{K}^{l'}_{l}=\delta^{l'}_{l}\left(\frac{\partial^2}{\partial x^2}-\epsilon(\gamma_l\gamma_{l'})^{-1}\frac{\partial}{\partial t}\right)\hat{K}^{l'}_{l}.$$

[1](a) C.S. Gardner, J.M. Greene, M.D. Kruskal, R.M. Miura, Phys. Rev. Lett. **19**, 1095 (1967); (b) P.D. Lax, Commun. Pure Appl. Math. **21**, 467 (1968); (c) V.E. Zakharov and A.B. Shabat, JETP **34**, 62 (1972); **37**, 823 (1973); (d) M.J. Ablowitz, D.J. Kaup, A.C. Newell, M. Segur, Phys. Rev. Lett. **30**, 1962 (1973); **31**, 125 (1973), Stud. Appl. Math. **53**, 249 (1974); G.L. Lamb, Phys. Rev. Lett. **31**, 197 (1973); (e) A.C. Scott *et al.*, IEEE **61**, 1443 (1973); *Dynamical Systems Theory and Applications*, edited by J. Moser (Lecture Notes in Physics, Berlin, Heidelberg, New York).

[2]H. Cornille, J. Math. Phys. **18**, 1855 (1977).

[3](a) S.V. Manakov and V.E. Zakharov, JETP Lett. **18**, 243 (1973); (b) D.J. Kaup, Stud. Appl. Math. **55**, 9 (1976); (c) V.E. Zakharov and A.B. Shabat, Funct. Anal. Appl. **8**, 226 (1974); (d) M.J. Ablowitz and R. Haberman, J. Math. Phys. **16**, 2301 (1975).

[4]H. Cornille, J. Math. Phys. **17**, 2143 (1976).

[5]A.C. Newel and L.G. Redekoff, PRL **38**, 377 (1977).

[6](a) H. Cornille, J. Math. Phys. **19**, 1463 (1978); (b) H. Cornille, to appear in J. Phys. A.

Random phase wave: A soluble non-Markovian system[a)]

R. L. Dewar

Plasma Physics Laboratory, Princeton University, Princeton, New Jersey 08540
(Received 15 March 1978)

The averaged propagator and the corresponding mass operator (non-Markovian particle-wave collision operator) of a particle being accelerated by a random potential are constructed explicitly in a model system. The model consists of an ensemble of monochromatic waves of random phase, such as arises in narrow-bandwidth plasma turbulence, and is particularly interesting as a system exhibiting strong trapping. An essential simplifying feature is that the propagator is evaluated in oscillation-center picture, which greatly simplifies the momentum-space operators occurring in the problem, and leads to a remarkable factorization of the mass operator. General analyticity and symmetry properties are derived using a projection-operator method, and verified for the solution of the model system. The nature of the memory exhibited by the mass operator is briefly examined.

1. INTRODUCTION

In a turbulent collisionless plasma a reasonable way to understand the particle dynamics is to model the plasma waves by an *external* random potential. Poisson's equation is applied only after all statistical averages have been taken. Although this is not fully self-consistent, it greatly simplifies the problem since only the single-particle Hamiltonian need be used.

The problem attacked in this paper is to find such a model system in which the Dyson equation for the averaged single-particle propagator can be constructed, and which exhibits the effects of strong particle trapping. This phenomenon occurs in narrow bandwidth plasma turbulence. Such a system is particularly interesting because it is the case least amenable to perturbation theory. In the opposite limit of broad bandwidth, Dupree[1] showed some time ago that the Dyson equation can be approximated by a Fokker–Planck equation. This can be improved upon somewhat by use of renormalized perturbation theory, the end result typically being the direct-interaction approximation,[2] which puts in some non-Markovian behavior but does not treat particle trapping. Also there is strong evidence that even the renormalized perturbation series is divergent.[3,4] There is thus a real need for an analytically soluble model of the type we propose to treat.

There is also a need for general symmetry and analyticity results, such as those obtained by Forster[5] in the theory of liquids. These properties provide a test for the reasonableness of approximate forms of the mass operator, or a check on the correctness of exact solutions. In this paper (Appendix) we derive analogous results for the random acceleration problem using the projection operator method of Weinstock,[6] and verify that our explicit solution satisfies these requirements (Sec. 7).

The simple statistical ensemble which comprises our model is presented in Sec. 2, and in Sec. 3 an additional simplifying element is introduced: We make a canonical transformation to oscillation-center variables.[7,8] This means that the propagator which we average is not precisely the propagator for the distribution function in the usual variables. However, because the oscillation-center transformation preserves the qualitative nature of particle orbits, we believe that the oscillation-center propagator gives an equally valid "picture" of the statistical behavior of the system. The averaged oscillation-center propagator is calculated in Secs. 4 and 5.

The great advantage of working in oscillation-center variables is that the momentum-space operators occurring in the problem are simplified from general Hilbert-space operators to 2×2 matrices. This enables us in Sec. 6 to work backwards from the known average propagator to the "mass operator" in the Dyson equation. An unexpected factorization of the mass operator into a simple momentum-space operator and a scalar memory function is found to occur. The general properties derived in the Appendix are verified in Sec. 7, where we also show the behavior of the memory function with respect to time delay.

2. STATISTICAL MODEL

The model consists of an ensemble of one-dimensional systems in each realization of which there is a single stationary electrostatic wave present whose phase α is random over the ensemble. This models narrow bandwidth plasma turbulence as seen from a frame moving with the average phase velocity. Thus, the single-particle Hamiltonians are of the form

$$H_\alpha = \frac{p^2}{2m} - e\phi_0 \cos(k_0 x - \alpha), \tag{1}$$

where e and m are the particle charge and mass, respectively, and k_0 and ϕ_0 are the wavenumber and amplitude (assumed to be the same in each realization). The average of any quantity ϕ_α depending on α is

$$\overline{\phi} \equiv \langle\phi\rangle \equiv \int_0^{2\pi} \frac{d\alpha}{2\pi}\phi_\alpha. \tag{2}$$

We can, following O'Neil,[9] find the orbits governed by Hamiltonian (1) in terms of elliptic functions. Given the orbits the exact propagator $U_\alpha(\emptyset|0)$ can be written down. [Notationally, we closely follow an earlier paper,[7] with $\emptyset$ denoting the unsubscripted argument list

[a)]Work commenced in Department of Theoretical Physics, Research School of Physical Sciences, The Australian National University, and completed at Princeton University under U.S. ERDA Contract EY-76-C-02-3073.

 0022-2488/78/1909-1946$1.00

(x,p,t) and 0 denoting the list (x_0,p_0,t_0).] The average propagator $\bar{U}(\emptyset|0)$ can then be evaluated by applying Eq. (2), and is found to depend only on the differences $x-x_0$ and $t-t_0$, so the spatial and temporal dependence can be removed by Fourier transformation. Unfortunately, $\bar{U}_{k,\omega}(p,p_0)$ remains a complicated function of the initial and final momenta, thus making inversion impractical and precluding the possibility of explicitly constructing a Dyson equation for the average propagator. In order to obtain a more tractable momentum-space operator, we transform to an "oscillation-center" picture.

3. OSCILLATION-CENTER PICTURE

In Ref. 7 it was pointed out that the propagator

$$G_\alpha(\emptyset|0)\equiv C_\alpha(\emptyset)\,U_\alpha(\emptyset|0)\,C_\alpha^{-1}(0), \tag{3}$$

with C_α a unitary linear operator acting on phase space functions, is dynamically equivalent to U_α, being the propagator in a new set of canonical variables: the oscillation-center coordinates. Specifically, we define the oscillation-center transformation to be such that the oscillation-center momentum of an untrapped particle is a constant of the motion, while for trapped particles we require the oscillation-center momentum to be a constant of the motion *almost everywhere*. Because this transformation is constructed as the limiting case of a family of diffeomorphisms from phase space onto itself,[8] the topology of the orbits is unchanged. Thus, the sign of the oscillation-center momentum of a trapped particle must reverse periodically in order to keep the orbit topologically circular, the reversals being caused by collisions with thin potential barriers located at $k_0x=(2n+1)\pi+\alpha$, with n an integer. In Ref. 8 it was shown that such a canonical transformation can indeed be constructed. The oscillation-center orbit of a trapped particle in a particular realization of the ensemble is depicted in Fig. 1, while the graphs of position and momentum *versus* the dimensionless time parameter $\tau\equiv(1/2)k_0v_0(t-t_0)$ are shown in Fig. 2. Here $v_0\equiv\partial\bar{K}/\partial p_0$ is the oscillation-center velocity, K being the new Hamiltonian. In terms of the energy-dependent bounce frequency $\omega_b(|p_0|;e\phi_0/m,k_0)$ we have

$$\tau=\omega_b(p_0)\,\mathrm{sgn}(p_0)(t-t_0). \tag{4}$$

If we denote by $\hat{x}_\alpha(t|0)$ and $\hat{p}_\alpha(t|0)$ the oscillation-center position and momentum of a particle as a function of its initial position x_0 and momentum p_0 at initial time t_0, the exact retarded propagator in the oscillation-center picture is

$$G_\alpha(\emptyset|0)=\delta[x-\hat{x}_\alpha(t|0)]\delta[p-\hat{p}_\alpha(t|0)]\theta(t-t_0), \tag{5}$$

where $\theta(t)$ is the Heaviside step function.

For untrapped particles we have $\hat{p}_\alpha=p_0$ for all times. Furthermore, $\hat{x}_\alpha=x_0+v_0(t-t_0)$ is also independent of α, so that G_α is just the unperturbed propagator for a particle with velocity $v_0\equiv\partial\bar{K}/\partial p_0$. As a corollary, the averaged propagator is also the unperturbed propagator. For trapped particles the propagator is not so trivial, and requires a detailed discussion of the orbit functions $\hat{x}_\alpha$ and $\hat{p}_\alpha$.

4. ORBIT FUNCTIONS (TRAPPED PARTICLES)

In order to simplify the analytic representation of $\hat{p}(t|0)$ and $\hat{x}(t|0)$ in the trapping region, we define a phase β relative to the initial position by

$$\beta=\pi+\alpha-k_0x_0,\quad \text{modulo } 2\pi. \tag{6}$$

It is clear that the integral over α in Eq. (2) can be replaced by the integral from 0 to 2π over β.

We also define a triangular waveform function, $T(\tau)$, as depicted in Fig. 3(a). Both $\hat{p}$ and $d\hat{x}/dt$ can now be written in terms of the derivative T', which is a square waveform. We have

$$\hat{p}=p_0T'(\tau-\tfrac{1}{2}\beta) \tag{7}$$

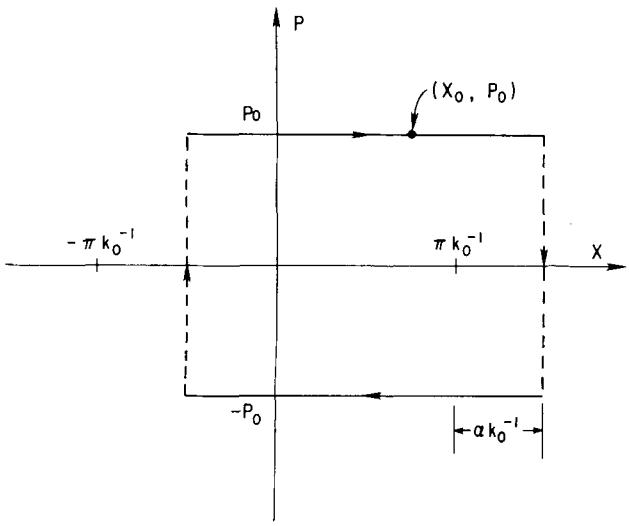


FIG. 1.

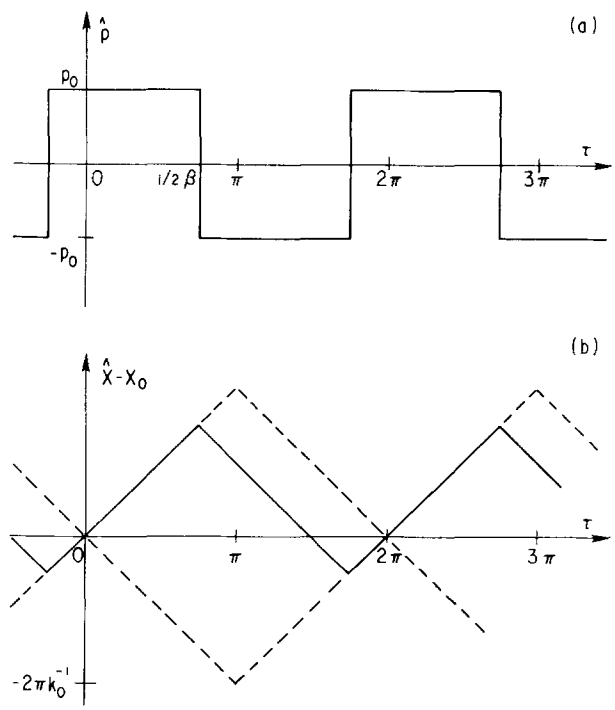


FIG. 2.

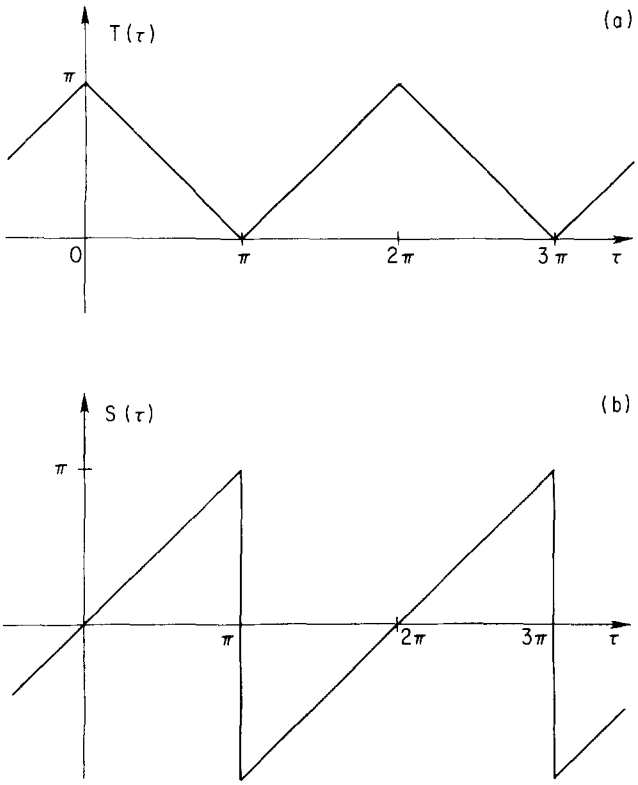


FIG. 3.

and

$$\frac{d\hat{x}}{dt}=v_0T'(\tau-\tfrac{1}{2}\beta), \tag{8}$$

where, as before, v_0 is $2\omega_b(p_0)\,\mathrm{sgn}(p_0)/k_0$. Integrating Eq. (8) we get

$$\hat{x}=x_0+2k_0^{-1}[T(\tau-\tfrac{1}{2}\beta)-T(\tfrac{1}{2}\beta)]. \tag{9}$$

If we denote τ(modulo 2π) by $\tau'\equiv\tau-2n\pi$, with n an integer chosen so that τ' is in the interval $[0,2\pi)$, then four cases may be distinguished:

Case 1: $0\le\tau'<\pi$, $0\le\frac{1}{2}\beta<\tau'$,

$$\hat{p}=-p_0,\quad \hat{x}=x_0+2k_0^{-1}(\beta-\tau').$$

Case 2: $0\le\tau'<\pi$, $\tau'\le\frac{1}{2}\beta<\pi$,

$$\hat{p}=p_0,\quad \hat{x}=x_0+2k_0^{-1}\tau'.$$

Case 3: $\pi\le\tau'<2\pi$, $0\le\frac{1}{2}\beta<\tau'-\pi$,

$$\hat{p}=p_0,\quad \hat{x}=x_0+2k_0^{-1}(\tau'-2\pi)$$

Case 4: $\pi\le\tau'<2\pi$, $\tau'-\pi\le\frac{1}{2}\beta<\pi$,

$$\hat{p}=-p_0,\quad \hat{x}=x_0+2k_0^{-1}(\beta-\tau').$$

It is convenient to define a sawtooth waveform function $S(\tau)$ as depicted in Fig. 3(b). In Cases 2 and 3, $\hat{x}$ is then simply $x_0+2k_0^{-1}S(\tau)$.

5. AVERAGE PROPAGATOR

It is now a simple matter to perform the integrations over β to calculate $G(\emptyset\,|\,0)$. First observe that in Cases 2 and 3 the integration is trivial because the integrand is independent of β over the ranges indicated. In Cases 1 and 4 the delta function containing $\hat{x}$ is removed by the integration, bearing in mind that there is one and only one solution of the equation $\hat{x}=x$ when x is in the range such that

$$|x-x_0|<2k_0^{-1}|S(\tau)|.$$

Outside this range there is no solution. Also observe that $|\partial\hat{x}/\partial\beta|=2k_0^{-1}$ in Cases 1 and 4. The result of the averaging is thus

$$\bar{G}(\emptyset\,|\,0)=\pi^{-1}T(\tau)\theta(t-t_0)\delta[x-x_0-2k_0^{-1}S(\tau)]\delta(p-p_0) \tag{10}$$
$$+(4\pi)^{-1}k_0\theta(t-t_0)\theta(2k_0^{-1}|S(\tau)|-|x-x_0|)\delta(p+p_0).$$

Note that, for $t>t_0$, $\bar{G}$ is a periodic function of t, the initial state recurring every bounce period. There is no irreversible phase mixing because p_0 defines a unique bounce frequency in the oscillation-center picture (provided k_0 and ϕ_0 are the same in all realizations of the ensemble).

Note that Eq. (10) is also valid for $p_0<0$, with τ defined as in Eq. (4). Thus, Eq. (10) expresses the fact that there is a nonvanishing probability of transitions only between the states with $p=\pm|p_0|$. To make this "coupled level" picture more explicit, we introduce a matrix notation for $\bar{G}$, and also choose to Fourier transform $\bar{G}$ at this time:

Case A: $p_0=+|p_0|$

$$\bar{G}(\emptyset\,|\,0)=\iint\frac{dk\,d\omega}{(2\pi)^2}\exp[ik(x-x_0)-i\omega(t-t_0)] \tag{11a}$$
$$\times i[\delta(p-|p_0|)\bar{G}^{++}_{k,\omega}(|p_0|)+\delta(p+|p_0|)\bar{G}^{-+}_{k,\omega}(|p_0|)],$$

Case B: $p_0=-|p_0|$

$$G(\emptyset\,|\,0)=\iint\frac{dk\,d\omega}{(2\pi)^2}\exp[ik(x-x_0)-i\omega(t-t_0)]$$

$$\times i[\delta(p-|p_0|)\,\bar{G}^{+-}_{k,\omega}(|p_0|)+\delta(p+|p_0|)\bar{G}^{--}_{k,\omega}(|p_0|)], \tag{11b}$$

where a factor i has been taken out of $\bar{G}^{\rho\sigma}_{k,\omega}$ (Ref. 7) to make it asymptotically equal to $(\omega-kpv_0)^{-1}\delta_{\rho,\sigma}$ as $\mathrm{Im}\omega\to+i\infty$.

If σ denotes + or −, the diagonal matrix elements are given by

$$\bar{G}^{\sigma\sigma}_{k,\omega}=\frac{1}{\omega-\sigma kv_0}$$
$$-2\omega_b\frac{\sin\{[\pi(\omega+kv_0)]/2\omega_b\}\sin\{[\pi(\omega-kv_0)]/2\omega_b\}}{\pi\sin(\pi\omega/\omega_b)(\omega-\sigma kv_0)^2}, \tag{12}$$

while the off-diagonal matrix elements are given by

$$\bar{G}^{+-}_{k,\omega}=\bar{G}^{-+}_{k,\omega}$$
$$=2\omega_b\frac{\sin\{[\pi(\omega+kv_0)]/2\omega_b\}\sin\{[\pi(\omega-kv_0)]/2\omega_b\}}{\pi\sin(\pi\omega/\omega_b)(\omega^2-k^2v_0^2)}. \tag{13}$$

Here v_0 denotes $2\omega_b/k_0$, and ω is to be interpreted as $\omega+i0$ near zeros of the denominators.

6. DYSON EQUATION

It is shown in the Appendix that a mass operator Σ should exist such that $\bar{G}$ obeys the Dyson equation

$$(\partial_t+L_{\bar{K}})\bar{G}(\emptyset\,|\,0)+\int d1\,\Sigma(\emptyset\,|\,1)\bar{G}(1\,|\,0)=\delta(\emptyset\,|\,0). \tag{14}$$

For untrapped particles, Σ vanishes identically, while from Eq. (11) it is clear that Σ can couple only the momenta $\pm p_0$, so we can use a matrix representation for Σ similar to that for $\bar{G}$. Using a representation for Σ precisely similar to Eq. (11) for $\bar{G}$ (including the factor of i), Eq. (14) becomes

$$(\mathbf{G}^{(0)-1}_{k,\omega} - \boldsymbol{\Sigma}_{k,\omega})\,\bar{\mathbf{G}}_{k,\omega} = \mathbf{I}\ , \tag{15}$$

with $\mathbf{I}$ the unit 2×2 matrix. Here $\bar{\mathbf{G}}_{k,\omega}$ is the matrix

$$\bar{\mathbf{G}}_{k,\omega} \equiv \begin{bmatrix} \bar{G}^{++}_{k,\omega} & \bar{G}^{+-}_{k,\omega} \\ \bar{G}^{-+}_{k,\omega} & \bar{G}^{--}_{k,\omega} \end{bmatrix}. \tag{16}$$

The inverse unperturbed propagator is given by

$$\mathbf{G}^{(0)-1}_{k,\omega} = \begin{bmatrix} \omega - kv_0 & 0 \\ 0 & \omega + kv_0 \end{bmatrix}. \tag{17}$$

Given $\bar{\mathbf{G}}_{k,\omega}$ from Eqs. (12) and (13) it is an easy matter to solve Eq. (15) for $\boldsymbol{\Sigma}_{k,\omega}$. We find that the momentum-space operator part of $\boldsymbol{\Sigma}_{k,\omega}$ factorizes out in a very simple manner

$$\boldsymbol{\Sigma}_{k,\omega} = \Sigma_{k,\omega} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \tag{18}$$

with the scalar memory function $\Sigma_{k,\omega}$ being given by

$$\Sigma_{k,\omega} = -2\omega_b^2 \frac{\sin\{[\pi(\omega + kv_0)]/2\omega_b\}\sin\{[\pi(\omega - kv_0)]/2\omega_b\}}{\pi^2\omega F[(\pi\omega/2\omega_b),(k/k_0)]}. \tag{19}$$

The function F in the denominator is defined by

$$F(\zeta,\kappa) \equiv \frac{\sin 2\zeta}{2\zeta} - \frac{\sin(\zeta+\pi\kappa)\sin(\zeta-\pi\kappa)}{\zeta^2 - \pi^2\kappa^2}. \tag{20}$$

In the denominator of Eq. (19), ω is to be read as $\omega + i0$.

7. PROPERTIES OF THE MASS OPERATOR

A. Symmetries

The following general symmetries are readily verified for both $\bar{G}_{k,\omega}$ and $\Sigma_{k,\omega}$,

$$\Sigma_{k,\omega}(p_1,p_2) = -\Sigma^*_{-k,-\omega}(p_1,p_2). \tag{21}$$

This is just the reality condition, the change of sign being a consequence of the factor of i taken out in the definitions of $\bar{G}_{k,\omega}$ and $\Sigma_{k,\omega}$ [Eq. (11)];

$$\Sigma_{k,\omega}(p_1,p_2) = \Sigma_{-k,\omega}(-p_1,-p_2). \tag{22}$$

This is a consequence of symmetry (on average) under spatial inversion. Finally, symmetry under interchange of momenta

$$\Sigma_{k,\omega}(p_1,p_2) = \Sigma_{k,\omega}(p_2,p_1) \tag{23}$$

can be derived from time reversal and spatial inversion invariance.

There is also an antisymmetry of the matrix in Eq. (18) about a horizontal axis. This symmetry is a kind of detailed balance and ensures conservation of probability at $k=0$.

B. Causality

We must show that $\Sigma_{k,\omega}$ obeys the causality condition of being an analytic function of ω in the upper half plane. It suffices to show that $F(\zeta,\kappa)$ has no zeros if ζ lies in the upper half plane.

To prove this by the Nyquist technique, it is sufficient to show that the hodograph of $F(\zeta,\kappa)$ does not encircle the origin as ζ traces out the contour shown in Fig. 4. The contour runs along the real axis from $-R$ to $+R$, except for small semicircular sections passing above the real zeros of $F(\zeta,\kappa)$, and then the contour returns to $(-R,0)$ along a large semicircle in the upper half plane.

For large $|\zeta|$, $F(\zeta,\kappa)$ can be approximated by $\sin(2\zeta)/2\zeta$, so that the hodographs of $F(\zeta,\kappa)$ and $\sin(2\zeta)/2\zeta$ must be essentially identical as ζ traverses the large semicircle, looping the origin $+n$ times (say). Each time the contour crosses two consecutive zeros on the real axis, the hodographs loop the origin -1 times in the counter-clockwise sense. Since $\sin(2\zeta)/2\zeta$ has no zeros in the upper (or lower) half plane, there must be $2n$ zeros in the interval of the real axis $(-R, R)$, so that the net number of times the hodograph encircles the origin is $-n+n=0$. To show that $F(\zeta,\kappa)$ also has no zeros in the upper half ζ plane then, all we need do is to show that it has the same number $(2n)$ of *real* zeros as $\sin(2\zeta)/2\zeta$ in any large interval $(-R,R)$ of the real axis, so that its hodograph also loops the origin $-n+n=0$ times.

By graphing the two terms which subtract to form $F(\zeta,\kappa)$, it is found that the number of crossings of the two graphs (zeros of F) is the same as the number of zeros of $\sin(2\zeta)/2\zeta$ in a long interval on the real line. This is found to hold for all κ, thus proving that the causality condition is satisfied by $\Sigma_{k,\omega}$. In the Appendix it is shown that this is a general property, following from the anti-self-adjointness of the Liouville operator.

C. Spectral representation

As shown in the Appendix, anti-self-adjointness of the Liouville operator implies the existence of the spectral representation[5]

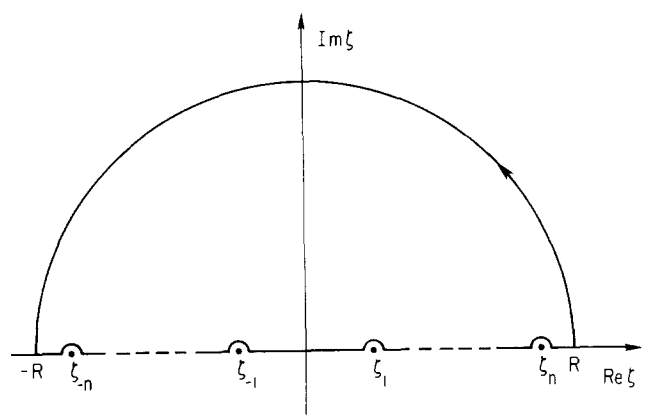


FIG. 4.

$$\Sigma_{k,\omega}(p_1,p_2)=\int_{-\infty}^{\infty}\frac{\Gamma_{k,\omega'}(p_1,p_2)}{\omega-\omega'+i0}, \tag{24}$$

where the damping matrix $\Gamma_{k,\omega}$ is a positive semidefinite operator. That is,

$$\int\int dp_1 dp_2 \Psi^*(p_1)\Gamma_{k,\omega}(p_1,p_2)\Psi(p_2)\geq 0 \tag{25}$$

for arbitrary functions Ψ.

In the 2×2 matrix representation of Eq. (11), we have

$$\Gamma_{k,\omega}=\Gamma_{k,\omega}\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}, \tag{26}$$

where the scalar damping function $\Gamma_{k,\omega}$ is given by

$$\Gamma_{k,\omega}=\frac{4\omega_b^2}{\pi}\left[\frac{\sin^2(\pi\kappa)}{F(0,\kappa)}\,\delta(\omega)-\sum_{i=-\infty}^{\infty}\frac{\sin(\zeta_i+\pi\kappa)\sin(\zeta_i-\pi\kappa)}{\zeta_i\partial F(\zeta_i,\kappa)/\partial\zeta_i}\,\delta\left(\omega-\frac{2\zeta_i\omega_b}{\pi}\right)\right] \tag{27}$$

where $\kappa\equiv k/k_0$ and the ζ_i are the zeros of $F(\zeta,\kappa)$. The matrix part of Eq. (26) is positive-semidefinite [with a null vector (1,1)], so one simply has to show that the scalar $\Gamma_{k,\omega}$ is greater than or equal to 0. One finds that the coefficients of the δ functions are always positive, so the requirement is satisfied.

Note that $\Gamma_{k,\omega}$ does not go to zero as $|\omega|\to\infty$, the asymptotic behavior being a "row of δ functions" with weights alternating between two different positive constants.

D. Temporal behavior

Inverting the Fourier transform of the memory function, we find its temporal behavior

$$\Sigma_k(t)=\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\,i\Sigma_{k,\omega}\exp(-i\omega t)=\theta(t)\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\,\Gamma_{k,\omega}\exp(-i\omega t). \tag{28}$$

The function $\Sigma_k(t)$ is graphed in Fig. 5 for the case $k=0.2\,k_0$. It is seen that the asymptotic behavior of $\Gamma_{k,\omega}$ at large ω leads to the occurrence of a δ function spike at time $t=0$, as if there were a component of the spectrum with zero autocorrelation time. However, these spikes recur periodically at later times, and

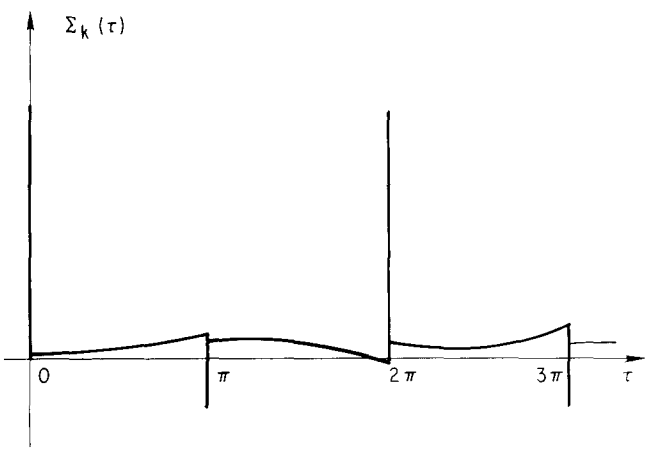


FIG. 5.

there is also a piecewise-continuous aperiodic component.

8. DISCUSSION

The two most striking features of the oscillation-center mass operator are the separation of the mass operator into a product of a scalar memory function and a momentum-space transition probability matrix, and the occurrence of δ functions in the temporal behavior of $\Sigma_k(t)$. Because the Hamiltonian system is integrable in each realization of the ensemble, the only randomness in the problem is that introduced by the ensemble. In more typical (and less tractable) cases the particle motion itself is highly unpredictable in most regions of phase space, thus damping the quasi-periodic tail of $\Sigma_k(t)$.

It is hoped that similar simplifications will be achieved when the oscillation-center transformation is applied to such nonintegrable systems. To treat such problems, it will be necessary to develop a convergent perturbation theory for the Hamilton–Jacobi equation for the Lie generating function.[8]

It might be taken as an objection to the use of the oscillation-center transformation to simplify the propagator that we end up calculating the propagator for the average oscillation-center distribution function rather than that for the average of the true distribution function. However, because the nature of the transformation is such that the discrepancy between the oscillation-center orbit and the corresponding exact orbit has no secular component, the longtime behavior of the two propagators must be similar. This would be especially true in systems where there are no trapped particles since the ratio of the distance between the exact and oscillation-center positions to the displacement from the initial point tends to zero at large times in such systems. The average distribution functions do differ by a nonsecular component related to the "fake diffusion" occurring in the conventional approach,[10] but the information contained in this component can be obtained from the known transformation operator C_α, which gives the coherent response of the plasma to the potential. It is in fact an advantage of the oscillation-center approach that this coherent response is separated from the secular behavior.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge useful conversations with Dr. J.A. Krommes, Dr. I. Lerche, and Dr. J.M. Greene.

APPENDIX

In order to apply the projection operator[6] technique, we define operators on the space of functions of x, p, and α. The inner product in this space is defined by

$$\langle f,\ g\rangle\equiv\int\int dx\,dp\int_0^{2\pi}d\alpha\, f_\alpha^*(x,p)g_\alpha(x,p). \tag{A1}$$

We now define the averaging operator A such that

$$\langle f,\ Ag\rangle\equiv\int\int dx dp\int_0^{2\pi}d\alpha\int_0^{2\pi}\frac{d\alpha'}{2\pi}\,f_\alpha^*(x,p)g_{\alpha'}(x,p), \tag{A2}$$

for all f and g. Clearly, A is a self-adjoint projection operator. We also define the complementary projector B by

$$B \equiv 1 - A. \tag{A3}$$

For time-independent systems, such as the one studied in this paper, the equation for the Fourier-transformed exact propagator $\hat{G}_\omega$ (regarded as an operator in the space defined in this Appendix) is

$$(\omega + i\bar{L})\hat{G}_\omega + i\tilde{L}\hat{G}_\omega = 1, \tag{A4}$$

where $\bar{L}$ is the average part $L_{\bar{K}}$ of the Liouville operator and $\tilde{L}$ is the residual part $L_{\tilde{K}}$. Provided the space of functions of x and p is suitably defined, $i\bar{L}$ and $i\tilde{L}$ are self-adjoint operators.

By multiplying Eq. (A4) on the left by A and on the right by A and B we obtain two coupled equations for $A\hat{G}A$ and $B\hat{G}A$. Eliminating the latter we have

$$(\omega + i\bar{L})A\hat{G}_\omega A + A\tilde{L}(\omega + i\bar{L} + iB\tilde{L}B)^{-1}\tilde{L}A\hat{G}_\omega A = 1, \tag{A5}$$

where we have used the fact that $A\tilde{L}A = 0$.

It is clearly consistent with Eq. (2) to define $\bar{G}_{k,\omega}$ by

$$2\pi\delta(k - k')\bar{G}_{k,\omega}(p,p') = \langle k,p|\hat{G}_\omega|k',p'\rangle, \tag{A6}$$

where the kets $|k',p'\rangle$ are defined componentwise by

$$\langle x,p,\alpha|k',p'\rangle \equiv (2\pi)^{-1/2}\exp(ik'x)\delta(p - p'). \tag{A7}$$

Note that $A|k,p\rangle = |k,p\rangle$. Thus, Eq. (A5) becomes

$$(\omega - kv)\bar{G}_{k,\omega}(p,p') - \int dp''\,\Sigma_{k,\omega}(p,p'')\bar{G}_{k,\omega}(p'',p') = \delta(p - p'), \tag{A8}$$

where $v \equiv \partial\bar{K}/\partial p$ and

$$2\pi\delta(k - k')\Sigma_{k,\omega}(p,p') = \langle k,p|\tilde{L}^\dagger(\omega + i\bar{L} + iB\tilde{L}B)^{-1}\tilde{L}|k',p'\rangle, \tag{A9}$$

$\tilde{L}^\dagger(= -\tilde{L})$ being the adjoint of $\tilde{L}$.

Since $\bar{L} + B\tilde{L}B$ is anti-self-adjoint, it can be written in the spectral representation[11]

$$\bar{L} + B\tilde{L}B = i\int_{-\infty}^{\infty}\lambda dE(\lambda),$$

where the projection operators $E(\lambda)$ are the spectral family of $\bar{L} + B\tilde{L}B$, and the integral runs over *real* λ. In this representation Eq. (A9) becomes

$$2\pi\delta(k - k')\Sigma_{k,\omega}(p,p') = \int_{-\infty}^{\infty} d\lambda\,\frac{d\langle k,p|\tilde{L}^\dagger E(\lambda)\tilde{L}|k',p'\rangle/d\lambda}{\omega - \lambda + i0} \tag{A10}$$

Thus the existence of Eq. (24) is guaranteed, and hence so is the causality property. Also, from the definition of $E(\lambda)$[11] it is clear that $\Sigma_{k,\omega}$ is a positive semidefinite operator.

[1] T. H. Dupree, Phys. Fluids **9**, 1773 (1966).
[2] S. A. Orszag and R. H. Kraichnan, Phys. Fluids **10**, 1720 (1967).
[3] S. A. Orszag, Phys. Fluids **18**, 271 (1975).
[4] I. G. Enting, Phys. Lett. A **59**, 417 (1977).
[5] D. Forster, Phys. Rev. A **9**, 943 (1974).
[6] J. Weinstock, Phys. Fluids **12**, 1045 (1969).
[7] R. L. Dewar, J. Phys. A **9**, 2043 (1976).
[8] R. L. Dewar, J. Phys. A **11**, 9 (1978).
[9] T. M. O'Neil, Phys. Fluids **8**, 2255 (1965).
[10] R. L. Dewar, Phys. Fluids **16**, 1102 (1973).
[11] P. Roman, *Some Modern Mathematics for Physicists and Other Outsiders* (Pergamon, New York, 1975), Vol. 2, p. 633.

Quantum generalization of Kolmogorov entropy

M. D. Srinivas

Department of Theoretical Physics, University of Madras, Guindy Campus, Madras-600025, India
(Received 19 September 1977)

The framework of quantum probability theory is employed in the investigation of ergodic properties of quantum dynamical systems. In the process certain basic notions of quantum information theory are developed. These are used to demonstrate the existence of a quantum Kolmogorov entropy which has the same operational significance as in classical theory. The notion of a quantum K-system is also introduced.

INTRODUCTION

The ergodic theory of classical dynamical systems has been successfully developed in the last 45 years[1-6] with a view to provide a rigorous mathematical foundation for classical statistical mechanics. In particular the entropy theory of dynamical systems initiated by Kolmogorov[7] has shown how classical dynamical systems can exhibit some sort of a stochastic behavior. In recent years, there have been several investigations (Refs. 8--11, and references cited therein) on a noncommutative generalization of classical ergodic theory to cover situations that are encountered in quantum statistical mechanics. In all these investigations the Abelian algebra of the random variables of classical theory is replaced by a non-Abelian algebra—usually assumed to be either a C^*-algebra or a von Neumann algebra. The problem of generalizing the notion of Kolmogorov (or dynamical) entropy to a noncommutative von Neumann algebra framework has been investigated recently by Emch (Ref. 12 and references cited therein) and Connes and Størmer.[13] Emch has also introduced the notion of a generalized (or non-Abelian) K-flow and has shown that every nonsingular generalized K-flow[12] has a strictly positive dynamical entropy.

The notion of Kolmogorov entropy in classical theory is intimately connected with the statistical uncertainty in the outcome of an experiment (with a finite number of outcomes) at time t, when the results of the same experiment performed at times $\{t-\tau,\ t-2\tau,\ t-3\tau, \ldots\}$ for some $\tau > 0$ are all assumed to be known. In extending such a notion meaningfully to quantum theory, it is therefore necessary to employ an operational framework in which statistical relations between successive observations can be discussed. However, in both the C^*-algebra and von Neumann algebra approaches (in the form they are usually employed in quantum statistical mechanics[14]) the observables of the theory are associated with the self-adjoint elements of the algebra, and no specification is made concerning the measurement transformations associated with their observation. In such a framework it is not possible to discuss the probability connections between the elements of a sequence of experiments performed on a system. Hence the von Neumann algebra model is not suitable for a generalization of the notion of Kolmogorov entropy into quantum theory in a physically meaningful way. Moreover, the concept of Kolmogorov entropy (in classical theory) arose out of the application of information theory to the study of dynamical systems; the need for adopting an operational approach in constructing a quantum information theory has been clearly demonstrated recently by Ingarden.[16,17]

Recent investigations[18-22] have shown that the appropriate framework for analyzing the statistical relations among successive observations in quantum theory is that of a quantum probability theory in which the space of events is identified with the set of all *operations* (or measurement transformations of the conventional theory), and the observables or random variables are defined as *operation valued measures* on their value spaces (See Sec. 1). *In the present investigation we shall employ the framework of quantum probability theory to obtain a quantum generalization of the notion of Kolmogorov entropy in such a way that it also retains its operational significance.* For this, we first introduce the notion of a quantum dynamical system in Sec. 1, and also consider quantum analogs of concepts like ergodicity and mixing for such a system. In Sec. 2, we shall develop certain basic notions of quantum information theory. These are used in Sec. 3 for demonstrating the existence of a quantum Kolmogorov entropy and also for the introduction of the concept of a quantum K-system.

1. QUANTUM DYNAMICAL SYSTEMS

In Ref. 20 we arrived at the following framework of quantum probability theory which is essentially a generalization of the approach developed by Davies and Lewis[18] and Edwards[19] into a probabilistic scheme:

(a) The set of all *operations* (or *events* of the theory) is the set $\mathcal{O} = \mathcal{L}_1^+(V)$ of all positive elements of the unit ball of $\mathcal{L}(V)$, the set of all bounded linear mappings of a complete base normed space V (with a base K for the closed generating cone V^+) into itself. Then the set Σ of all the maximal elements in $\mathcal{O}$ [under the partial order induced in $\mathcal{O}$ by the cone $\mathcal{L}^+(V)$ closed under the strong topology] is a uniformly closed convex semigroup.

(b) The set of all *states* (or *measures*) is the set of all linear functionals

$$\mu : \mathcal{O} \to [0,1],$$

which are continuous in the strong topology on $\mathcal{O}$ and satisfy

$$\text{(i)}\quad \mu(\xi) = 1, \quad \text{for all } \xi \in \Sigma, \tag{1.1}$$

$$\text{(ii)}\quad \mu(\xi A) = \mu(A), \quad \text{for all } \xi \in \Sigma \text{ and } A \in \mathcal{O}.$$

(c) An *observable* (or a *random variable*) with value space $(S, \mathcal{B}(S))$ (usually a standard Borel space) is a

 0022-2488/78/1909-1952$1.00

map $X: \mathcal{P}(S) \to \mathcal{O}$, which satisfies the following requirements:

(i) $X(S) \in \Sigma$; (1.3)

(ii) If $\{E_i\}$ is any disjoint sequence of elements of $\mathcal{P}(S)$, then $\sum_i X(E_i)$ converges to an element of $\mathcal{O}$ in the strong operator topology and

$$X\left(\bigcup_i E_i\right) = \sum_i X(E_i). \tag{1.4}$$

The following remarks may be made on the interpretation of the above formalism: Statistical physical theories are usually formulated as probabilitistic structures founded on a "logic of experimentally verifiable propositions." Each operation typically represents a "measurement transformation" which completely characterizes the experimental procedure associated with some experimentally verifiable proposition. In fact, with each $A \in \mathcal{O}$, we can associate a linear positive mapping of the set of all unnormalized states [i.e., strongly continuous positive linear functionals on $\mathcal{O}$ which satisfy (1.2)] into itself, by the rule $A: \mu \to \mu'$, where

$$\mu'(B) = \mu(BA) \tag{1.5}$$

for all $B \in \mathcal{O}$.

The *conjunction* $A \wedge B$ of two operations, $A, B, \in \mathcal{O}$ is defined by the equation

$$A \wedge B = BA, \tag{1.6}$$

and corresponds to the experimental procedure in which the system is subjected to the sequence $\{A, B\}$ of experimental procedures corresponding to A and B, in that order. If $A, B \in \mathcal{O}$ are such that $A + B \in \mathcal{O}$ also, then A and B are said to be *mutually disjoint* and their *disjunction* $A \vee B$ is defined by

$$A \vee B = A + B. \tag{1.7}$$

$A \vee B$ corresponds in some sense to a "fusing" or physical adjoining of the experimental procedures corresponding to A, B. $\mathcal{O}$ contains a *unique null element* which may be denoted as θ; however, the set Σ of *maximal elements* in $\mathcal{O}$ contains (in general) several elements apart from the identity operation I.

From (1.3) it is clear that each maximal operation corresponds to an experimentally verifiable proposition of the form: "In an experiment to measure the observable X with value space $(S, \beta(S))$ the outcome is found to lie in S." *Apart from the noncommutativity of the conjunction* (1.6), *it is the fact that the set* Σ *of maximal operations is nontrivial* (*i.e.*, $\Sigma \neq \{I\}$), *which constitutes the most important nonclassical feature of quantum probability theory—and gives rise to the so-called "quantum interference of probabilities."*[20,22]

Given an observable X with value space $(S, \beta(S))$ and $E \in \beta(S)$ and a state μ, $\mu(X(E))$ is the probability that when a measurement of X is made on (an ensemble of identical copies of) a system in state μ, the outcome will be found to lie in E. It can easily be shown that $E \to \mu(X(E))$ defines a probability measure on $(S, \beta(S))$ for each state μ. A more significant result proven by Davies and Lewis[18] is that if the value spaces $(S_1, \beta(S_1))$, $(S_2, \beta(S_2))$ of two observables X_1, X_2 are standard Borel spaces, then there exists a unique observable X on the product Borel space $(S_1 \times S_2, \beta(S_1 \times S_2))$ such that for any $E_1 \in \beta(S_1)$ and $E_2 \in \beta(S_2)$

$$X(E_1 \times E_2) = X_1(E_1) \wedge X_2(E_2), \tag{1.8}$$

i.e., the observable X is such that its observation corresponds to the composite observation: X_1 followed by X_2. With each element ρ of the base K of the complete base normed space V, we can uniquely associate a state μ_ρ on $\mathcal{O}$, via the identification

$$\mu_\rho(A) = ||A\rho||, \tag{1.9}$$

for all $A \in \mathcal{O}$. However, there does not seem to exist any characterization of the set of all states on $\mathcal{O}$, even in the *canonical model associated with conventional quantum mechanics where* V *is chosen to be the Banach space* $\mathcal{T}_S(\mathcal{H})$ *of all self-adjoint trace class operators on a separable Hilbert space* $\mathcal{H}$. The base K, in this case, would be the set of all density operators on $\mathcal{H}$.

The assumption (a) that the set of events $\mathcal{O}$ is to be identified with the whole of $\mathcal{L}_1^+(V)$ may not in general be quite justified. It would definitely be more realistic to stipulate that the set of events $\mathcal{O}$ has the structure

$$\mathcal{O} = \mathcal{L}_1^+(V) \cap \mathcal{C} \tag{1.10}$$

where $\mathcal{C}$ is a Banach subalgebra of $\mathcal{L}(V)$. We will now introduce an abstract mathematical framework which seems to be the appropriate generalization of the theories of the type (1.10). The generalization involved is quite similar to the one performed when one passes from an algebra of operators on a Hilbert space into a C^*-algebra.

(A) The set $\mathcal{O}$ of all operations (i.e., the quantum event space) is the set of all positive elements $\mathcal{U}_1^+$ in the unit ball of an ordered Banach algebra $\mathcal{U}$ with identity I and a closed cone $\mathcal{U}^+$. Σ is the uniformly closed convex set of all the maximal elements of $\mathcal{O}$ in the partial order on $\mathcal{O}$ induced by $\mathcal{U}^+$.

(B), (C) Same as (b) and (c) above, with the only exception being that the requirements of continuity (and convergence) under the strong operator topology are replaced, in each case, by the requirements of continuity (and convergence) under the norm topology.

When $\mathcal{O}$ is taken to be $\mathcal{L}_1^+(V)$ and we adopt the framework (A)—(C) instead of (a)—(c), the only difference will be that the class of allowed states will now be enhanced and the class of observables will be correspondingly reduced in such a way that each state continues to induce a probability measure on the value space of every observable. Before one seriously considers the framework (A)—(C) as a prototype for all operational approaches to quantum theory, it is necessary to strengthen the requirements (A)—(C) in such a way that every representation of the theory will have properties akin to (a)—(c), and possibly with a further restriction that V is of type $\mathcal{T}_S(\mathcal{H})$. In other words it is necessary to supplement the requirements (A)—(C) in such a way that every representation of $\mathcal{O}$ resembles the space of operations associated with the canonical model de-

scribed above. We shall not discuss this question here as *all the results which are obtained in this paper on the basis of the framework (a)–(c), will continue to hold in the abstract framework (A)–(C) also, provided in each definition and theorem we substitute "uniform topology" for the "strong operator topology."*

After this discussion of the basic framework we shall employ, we proceed to the definition of a quantum dynamical system. We denote by Aut$\mathcal{O}$ the group of automorphisms of $\mathcal{O}$ which are also continuous in the strong operator topology: i.e., $\mathcal{O} \in \text{Aut}\mathcal{O}$ is a one-to-one isometry of $\mathcal{O}$ onto itself which is continuous in the strong operator topology and satisfies

$$\varphi(A \wedge B) = \varphi(A) \wedge \varphi(B), \tag{1.11}$$

$$\varphi(A + B) = \varphi(A) + \varphi(B). \tag{1.12}$$

A *quantum dynamical system* shall be an aggregate $(\mathcal{O}, \mu, \varphi(R))$, where $\mathcal{O}$ is a quantum event space, μ is a state on $\mathcal{O}$, and $t \to \varphi_t$ is a homomorphism of the real line R into Aut$\mathcal{O}$ such that the conditions (QD1) and (QD2) are satisfied:

(QD1)

$$\mu \circ \varphi_t = \mu, \tag{1.13}$$

for all $t \in R$, where

$$(\mu \circ \varphi_t)(A) = \mu(\varphi_t A), \tag{1.14}$$

for all $A \in \mathcal{O}$;

(QD2) The function $t \to \varphi_t(A)$ is continuous in the strong operator topology on $\mathcal{O}$.

It is probably necessary to supplement the requirements (QD1) and (QD2) by a condition of the following form.

(QD3) For all $A, B \in \mathcal{O}$, the function $t \to \mu(A \wedge \varphi_t B)$ is continuous.

However, (QD3) does not play any role in our investigations.

We now introduce a class of states $\omega(\mu)$ associated with a given state μ, and which corresponds in some sense to the class of measures absolutely continuous with respect to a given measure in classical probability theory. The set $\omega(\mu)$ is also analogous to the family of normal states in the GNS representation associated with a given state in the C^*-algebra framework.[8-11,23] $\omega(\mu)$ is composed of all those states μ' which are given by

$$\mu'(B) = \mu(A \wedge B)/\mu(A), \tag{1.15}$$

for all $B \in \mathcal{O}$, where $A \in \mathcal{O}$ is such that $\mu(A) \neq 0$. We should note that unlike in classical theory, $\mu(B) = 0$ need not imply $\mu'(B) = 0$ for a $\mu' \in \omega(\mu)$.

In the rest of this section we shall briefly indicate how some of the basic notions of classical ergodic theory can be extended so as to be applicable for quantum dynamical systems.

1. *Ergodicity*: $(\mathcal{O}, \mu, \varphi(R))$ shall be said to be ergodic iff

$$\mu' \in \omega(\mu) \text{ and } \mu' \circ \varphi_t = \mu \text{ for all } t \in R \Rightarrow \mu' = \mu. \tag{1.16}$$

We have the following immediate consequence:

Lemma 1.1: If $(\mathcal{O}, \mu, \varphi(R))$ is ergodic and $A \in \mathcal{O}$ is an invariant operation (i.e., $\varphi_t A = A$ for all $t \in R$), then

$$\mu(A \wedge B) = \mu(A)\mu(B), \tag{1.17}$$

for all $B \in \mathcal{O}$.

2. *Asymptotic Independence*: A dynamical system $(\mathcal{O}, \mu, \varphi(R))$ is said to be asymptotically independent, if for any two random variables X_1, X_2, the sequence of random variables $\{\varphi_{-t} X_1, X_2\}$ becomes statistically independent in the limit $t \to \infty$; i.e., if $(S_1, \beta(S_1))$ and $(S_2, \beta(S_2))$ are the corresponding value spaces, we have the following relation which follows from the definition of statistical independence in quantum theory[22]:

$$\lim_{t \to \infty} \left| \mu(\varphi_{-t} X_1(E_1) \wedge X_2(E_2)) - \mu(X_1(E_1))\mu(\varphi_{-t} X_1(S_1) \wedge X_2(E_2)) \right| = 0, \tag{1.18}$$

for all $E_1 \in \beta(S_1)$ and $E_2 \in \beta(S_2)$. This is also equivalent to the requirement

$$\lim_{t \to \infty} \left| \mu(\varphi_{-t} A_1 \wedge A_2) - \mu(A_1)\,\mu(\varphi_{-t} \xi_1 \wedge A_2) \right| = 0, \tag{1.19}$$

for all A_1, $A_2 \in \mathcal{O}$ and for every $\xi_1 \in \Sigma$ such that $A_1 \leq \xi_1$.

3. *Asymptotic noninterference*: A dynamical system $(\mathcal{O}, \mu, \varphi(R))$ is said to be asymptotically noninterfering iff

$$\lim_{t \to \infty} \mu(\varphi_{-t} \xi \wedge A) = \mu(A), \tag{1.20}$$

for all $A \in \mathcal{O}$ and $\xi \in \Sigma$. We have referred to (1.20) as a noninterference property because it implies that an act of measurement performed on a system in state μ does not show up features of the "interference of probabilities" provided the measurement was made in the infinite past—i.e., a measurement on a system in state μ in the infinite past becomes equivalent to a "classical" or noninterfering measurement. Thus the notion of asymptotic noninterference is quite similar to the asymptotic commutation properties ("weak asymptotic Abelian-ness," etc.) which have been widely employed[11,23] in theories of C^*-dynamical systems in showing approach to equilibrium.

Both the notions of asymptotic independence and asymptotic noninterference are important for understanding the physical significance of the notion of mixing in quantum theory.

4. *Mixing*: A dynamical system $(\mathcal{O}, \mu, \varphi(R))$ is said to be mixing iff

$$\lim_{t \to \infty} \mu(\varphi_{-t} A \wedge B) = \mu(A)\mu(B), \tag{1.21}$$

for all $A, B \in \mathcal{O}$.

The importance of the notion of mixing stems from the following approach (or "return") to equilibrium property exhibited by a mixing system:

Lemma 1.2: Let $(\mathcal{O}, \mu, \varphi(R))$ be a dynamical system. Then the following are equivalent:

(i) $(\mathcal{O}, \mu, \varphi(R))$ is mixing;

(ii) $\lim_{t \to \infty} \mu' \circ \varphi_t = \mu$, for all $\mu' \in \omega(\mu)$; (1.22)

(iii) $(\mathcal{O}, \mu, \varphi(R))$ is asymptotically independent and asymptotically noninterfering.

It can be easily seen that mixing implies ergodicity. It may also be noted that the definitions we have provided above are more or less direct generalizations of the corresponding classical definitions in the following form[24]:

(a) Ergodicity $\Leftrightarrow$ the only invariant measure absolutely continuous with respect to μ is μ itself.

(b) Mixing $\Leftrightarrow$ any measure absolutely continuous with respect to μ converges to μ for infinite times.

In classical theory the notions of mixing and asymptotic independence coincide as every classical measurement is noninterfering.

2. ELEMENTS OF QUANTUM INFORMATION THEORY

The central concept in this section is that of a partition which formalizes the notion of an experiment with a finite number of outcomes. After all, most of the realistic experiments are of this type and, in addition, the way the various outcomes are labelled is not particularly relevent for studying the statistical relations between such experiments. Hence a partition may be defined as the set of events associated with an experiment with a finite number of outcomes in the following way: A *partition* α in a quantum event space $\mathcal{O}$ is a *finite collection* $\{A_i \mid 1 \leq i \leq m\}$ $(m \in Z^+)$ of *nonnull* elements $A_i \in \mathcal{O}$, such that

$$\xi^\alpha \equiv \sum_{i=1}^{m} A_i \in \Sigma. \tag{2.1}$$

Thus, given a partition $\alpha = \{A_i \mid 1 \leq i \leq m\}$, we can associate with it a random variable X_α, with value space $\{1, 2, \ldots, m\}$ such that

$$X_\alpha(\{i\}) = A_i. \tag{2.2}$$

All other random variables associated with α can be reduced to X_α by relabelling their value spaces as $\{1, 2, \ldots, m\}$.

If α is a partition, we shall denote by ξ_α the partition $\{\xi^\alpha\}$ with just the single element $\xi^\alpha \in \Sigma$. We refer to such single element partitions as *trivial partitions*. If $\alpha = \{A_i \mid 1 \leq i \leq m\}$ and $\beta = \{B_j \mid 1 \leq j \leq n\}$ are partitions, we shall write $\alpha \leq \beta$, whenever β can be subdivided into mutually disjoint subsets $\beta_k = \{B_{kr} \mid 1 \leq r \leq m_k\}$ $(1 \leq k \leq m)$, such that

$$A_k = \sum_{r=1}^{m_k} B_{kr},$$

for each $1 \leq k \leq m$. We also introduce the weaker relation $\alpha \leq_\mu \beta$, for each state μ on $\mathcal{O}$, if instead of (2.3) we have

$$\mu(A_k) = \sum_{r=1}^{m_k} \mu(B_{kr}), \tag{2.4}$$

for each $1 \leq k \leq m$. If $\alpha \leq \beta$, then $\alpha \leq_\mu \beta$ for each state μ on $\mathcal{O}$.

If $\alpha = \{A_i \mid 1 \leq i \leq m\}$ and $\beta = \{B_j \mid 1 \leq j \leq n\}$ are partitions, then $\alpha \vee \beta$ denotes the partition

$$\alpha \vee \beta = \{A_i \wedge B_j \mid 1 \leq i \leq m,\ 1 \leq j \leq n\}; \tag{2.5}$$

i.e., $\alpha \vee \beta$ corresponds to the composite experiment – α followed by β. *In quantum theory partitions* $\alpha \vee \beta$ *and* $\beta \vee \alpha$ *are in general different.* Also we *do not* have $\alpha \leq \alpha \vee \beta$ or $\beta \leq \alpha \vee \beta$ in general. However, for any partitions α, β, γ, δ, we have the following relations:

$$\alpha \vee \xi_\beta \leq \alpha \vee \beta; \tag{2.6a}$$

$$\xi_\alpha \vee \beta \leq \alpha \vee \beta; \tag{2.6b}$$

$$\alpha \leq \gamma \text{ and } \beta \leq \delta \Rightarrow \alpha \vee \beta \leq \gamma \vee \delta; \tag{2.6c}$$

$$\alpha \leq_\mu \alpha \vee \beta, \tag{2.6d}$$

for all states μ on $\mathcal{O}$.

Hereafter we shall work in a fixed *quantum probability space* $(\mathcal{O}, \mu)$, i.e., a quantum event space $\mathcal{O}$ together with a state μ on it. We would like to introduce the notion of the entropy $H(\alpha)$ of a partition $\alpha = \{A_i \mid 1 \leq i \leq m\}$ as a function of the probabilities $\{\mu(A_i)\}$. $H(\alpha)$ should be a measure of the average amount of information gained by finding the outcome of (an experiment X_α corresponding to) α—or, equivalently, a measure of the average uncertainty in the outcome of α. For this, we will have to determine the corresponding numerical function $H(p_1, p_2, \ldots, p_m)$, (for each $m \geq 1$) of a set of numbers $p_i = \mu(A_i) \geq 0$, which satisfy $\sum_{i=1}^{m} p_i = 1$. This function should clearly have the following properties (QI1) and (QI2):

(QI1) $H(p_1, p_2, \ldots, p_m)$ takes its largest value for $p_k = 1/m$ $(1 \leq k \leq m)$, i.e., the entropy $H(\alpha)$ of an experiment α is maximum when each of its outcomes are equally likely to occur.

(QI2) $H(p_1, p_2, \ldots, p_m, 0) = H(p_1, p_2, \ldots, p_m)$.

Another reasonable requirement on the entropy is the following: If the experiment $\beta = \{B_j \mid 1 \leq j \leq n\}$ is performed after the experiment $\alpha = \{A_i \mid 1 \leq i \leq m\}$ and $\{\alpha, \beta\}$ are statistically independent, then the uncertainty $H(\alpha \vee \beta)$ in the compound experiment "β followed by α" should be the sum of the uncertainty $H(\alpha)$ in the outcome of α and the *uncertainty* $H(\xi_\alpha \vee \beta)$ *in the outcome of experiment* β *when the system is subjected to the sequence of experiments* $\{\alpha, \beta\}$. It may be recalled[22] that when a system in state μ is subjected to the sequence of experiments $\{\alpha, \beta\}$, the physically meaningful probability $\Pr\{B_j\}$ for the outcome $1 \leq j \leq n$ in the experiment β is given by

$$\Pr\{B_j\} = \mu(\xi^\alpha \wedge B_j). \tag{2.7a}$$

Also, the experiments α, β may be said to be statistically independent when the system in state μ is subjected to the sequence of experiments $\{\alpha, \beta\}$ iff[22]

$$\mu(A_i \wedge B_j) = \mu(A_i)\mu(\xi^\alpha \wedge B_j), \tag{2.7b}$$

for all $1 \leq i \leq m$, $1 \leq j \leq n$. If we now write $\mu(A_i) = p_i$ and $\mu(\xi^\alpha \wedge B_j) = q_j$ and if (2.7b) is satisfied, then we have

$$H(\alpha \vee \beta) = H(p_1 q_1, p_1 q_2, \ldots, p_m q_n),$$

$$H(\alpha) = H(p_1, p_2, \ldots, p_m),$$

$$H(\xi_\alpha \vee \beta) = H(q_1, q_2, \ldots, q_n).$$

Thus the above requirement of the *additivity of uncertainty* (or *information*) *for statistically independent ex-*

periments can be expressed in the following form:

(QI3) $H(p_1q_1, p_1q_2, \ldots, p_mq_m)$

$$= H(p_1, p_2, \ldots, p_m) + H(q_1, q_2, \ldots, q_n),$$

for $m, n \geq 1$.

If we now add to (QI1), (QI2), (QI3) another requirement that, for each $m \geq 1$, $H(p_1, p_2, \ldots, p_m)$ be continuous in all its arguments, then the theorem of Khunchin[25] shows that, for each m, $H(p_1, p_2, \ldots, p_m)$ should be of the form

$$H(p_1, p_2, \ldots, p_m) = -\lambda \sum_{i=1}^{m} p_i \log p_i,$$

where λ is a positive constant.

The preceding discussion shows that a reasonable measure of the average uncertainty of (the outcome of an experiment corresponding to) a partition $\alpha = \{A_i \mid 1 \leq i \leq m\}$ is given by the *entropy* $H(\alpha)$ if we define $H(\alpha)$ by the relation

$$H(\alpha) = -\sum_{i=1}^{m} \mu(A_i) \log \mu(A_i), \tag{2.8}$$

where all logarithms are taken to base 2 and $0 \cdot \log 0 = 0$. A similar argument shows that if $\alpha = \{A_i \mid 1 \leq i \leq m\}$ and $\beta = \{B_j \mid 1 \leq j \leq m\}$ are any two partitions, then a reasonable measure of the average conditional uncertainty in the outcome of the experiment β given the outcome of experiment α, *when a system in state* μ *is, subjected to the sequence of experiments* $\{\alpha, \beta\}$, is given by the *conditional entropy* $H(\beta \mid \alpha)$ defined as follows:

$$H(\beta/\alpha) = \sum_{i=1}^{m} \mu(A_i) \left[-\sum_{j=1}^{m} \frac{\mu(A_i \wedge B_j)}{\mu(A_i)} \times \log \frac{\mu(A_i \wedge B_j)}{\mu(A_i)} \right]. \tag{2.9}$$

Since the conditional probabilities satisfy the relation[22]

$$0 \leq \mu(A_i \wedge B_j)/\mu(A_i) \leq 1, \tag{2.10}$$

$H(\beta/\alpha)$ given by Eq. (2.9) is well defined.

We now collect all the elementary properties of the entropy and conditional entropy in the following theorem:

Theorem 2.1: Let $(\mathcal{O}, \mu)$ be a quantum probability space and $\alpha = \{A_i \mid 1 \leq i \leq m\}$, $\beta = \{B_j \mid 1 \leq j \leq n\}$, and $\gamma = \{C_k \mid 1 \leq k \leq r\}$ be partitions. Then we have the following properties:

(i) $0 = H(\xi_\alpha) \leq H(\alpha) \leq \log m$, (2.11a)

$$H(\alpha) = \log m \iff \mu(A_i) = 1/m, \tag{2.11b}$$

for all $1 \leq i \leq m$;

(ii) $\alpha \leq \beta \Rightarrow \alpha \leq_\mu \beta \Rightarrow H(\alpha) \leq H(\beta)$; (2.12)

(iii) $H(\alpha \vee \xi_\beta) = H(\alpha)$, (2.13a)

$$H(\xi_\alpha \vee \beta) = H(\beta \mid \xi_\alpha); \tag{2.13b}$$

(iv) $H(\alpha \vee \beta) = H(\alpha) + H(\beta/\alpha)$, (2.14a)

$$H(\beta \vee \gamma/\alpha) = H(\beta/\alpha) + H(\gamma/\alpha \vee \beta); \tag{2.14b}$$

(v) $H(\alpha) \leq H(\alpha \vee \beta)$, (2.15a)

$$H(\beta/\alpha) \leq H(\beta/\xi_\alpha) = H(\xi_\alpha \vee \beta) \leq H(\alpha \vee \beta); \tag{2.15b}$$

(vi) $H(\beta/\alpha) \leq H(\beta \vee \gamma/\alpha)$, (2.16a)

$$H(\gamma/\alpha \vee \beta) \leq H(\beta \vee \gamma/\alpha); \tag{2.16b}$$

(vii) $H(\alpha \vee \beta) \leq H(\alpha) + H(\xi_\alpha \vee \beta)$, (2.17)

where the equality holds iff the experiments (corresponding to) α and β are statistically independent when a system in state μ is subjected to the sequence of experiments $\{\alpha, \beta\}$;

(viii) $H(\gamma/\alpha \vee \beta) \leq H(\gamma/\alpha \vee \xi_\beta)$; (2.18)

(ix) $H(\gamma/\alpha \vee \beta) \leq H(\gamma/\xi_\alpha \vee \beta)$; (2.19)

(x) $H(\beta \vee \gamma/\alpha) \leq H(\beta/\alpha) + H(\xi_\beta \vee \gamma/\alpha)$. (2.20)

Proof: (*i*)–(*iv*): These follow immediately from the definitions (2.8), (2.9).

(*v*)–(*vi*): These follow directly from (iv), except for the first inequality $[H(\beta/\alpha) \leq H(\beta/\xi_\alpha)]$ in (2.15b), which follows from (vii) which will be proved below.

(*vii*): To prove this, we make use of the well-known inequality

$$\sum_{i=1}^{m} \sum_{j=1}^{n} r_{ij} \log \frac{r_{ij}}{p_i q_j} \geq 0, \tag{2.21}$$

where $\{r_{ij}\}$ are such that, $r_{ij} \geq 0$,

$$\sum_{i=1}^{m} \sum_{j=1}^{n} r_{ij} = 1, \quad \text{and} \quad p_i = \sum_{j=1}^{n} r_{ij}, \quad q_j = \sum_{i=1}^{m} r_{ij}.$$

Also, in (2.21), the equality holds iff

$$r_{ij} = p_i q_j, \tag{2.22}$$

for all $1 \leq i \leq m$, $1 \leq j \leq n$. If we now set $r_{ij} = \mu(A_i \wedge B_j)$, then we obtain

$$p_i = \mu(A_i) \quad \text{and} \quad q_j = \mu(\xi^\alpha \wedge B_j).$$

If we substitute these in (2.21), (2.22), then (vii) follows immediately once the definition of statistical independence as given by (2.7b) is employed.

(*viii*): We shall first derive an inequality that is essential for the proof of (viii) and (ix).

Let $\{q_{jk} \mid 1 \leq j \leq n, \ 1 \leq k \leq r\}$ be a set of numbers which satisfy $q_{jk} \geq 0$, $\sum_{j=1}^{n} \sum_{k=1}^{r} q_{jk} = 1$ and let $p_j = \sum_{k=1}^{r} q_{jk}$. Then as a consequence of the Jensen's inequality,[1] we have

$$-\sum_{j=1}^{n} \sum_{k=1}^{r} q_{jk} \log \frac{q_{jk}}{p_j} = \sum_{j=1}^{n} p_j \sum_{k=1}^{r} \left(\frac{-q_{jk}}{p_j} \right) \log \frac{q_{jk}}{p_j} \leq -\sum_{k=1}^{r} \left(\sum_{j=1}^{n} q_{jk} \right) \log \left(\sum_{j=1}^{n} q_{jk} \right). \tag{2.23}$$

If we now substitute

$$q_{jk} = \mu(A_i \wedge B_j \wedge C_k)/\mu(A_i)$$

in (2.23), we obtain

$$-\sum_{j=1}^{n}\sum_{k=1}^{r}\frac{\mu(A_i\wedge B_j\wedge C_k)}{\mu(A_i)}\log\frac{\mu(A_i\wedge B_j\wedge C_k)}{\mu(A_i)}$$

$$\leq -\sum_{k=1}^{r}\frac{\mu(A_i\wedge \xi^{\beta}\wedge C_k)}{\mu(A_i)}$$

$$\times\log\frac{\mu(A_i\wedge \xi^{\beta}\wedge C_k)}{\mu(A_i)}. \tag{2.24}$$

If we multiply both sides of (2.24) by $\mu(A_i)$ and sum over i from 1 to m, then we obtain (viii) immediately.

(*ix*): To prove this, we shall again make use of the inequality (2.23) with $n=m$. If we substitute

$$q_{jk}=\mu(A_j\wedge B_i\wedge C_k)/\mu(\xi^{\alpha}\wedge B_i)$$

in (2.23), then we obtain

$$-\sum_{j=1}^{n}\sum_{k=1}^{r}\frac{\mu(A_j\wedge B_i\wedge C_k)}{\mu(\xi^{\alpha}\wedge B_i)}\log\frac{\mu(A_j\wedge B_i\wedge C_k)}{\mu(A_j\wedge B_i)}$$

$$\leq -\sum_{k=1}^{r}\frac{\mu(\xi^{\alpha}\wedge B_i\wedge C_k)}{\mu(\xi^{\alpha}\wedge B_i)}$$

$$\times\log\frac{\mu(\xi^{\alpha}\wedge B_i\wedge C_k)}{\mu(\xi^{\alpha}\wedge B_i)}. \tag{2.25}$$

By multiplying both sides of (2.25) by $\mu(\xi^{\alpha}\wedge B_i)$ and summing over i from 1 to n, we are led to (ix) directly.

(*x*): This follows easily from (iv) and (viii) once we notice that we always have

$$H(\gamma/\alpha\vee\xi_{\beta})=H(\xi_{\beta}\vee\gamma/\alpha). \tag{2.26}$$

The results contained in the above theorem can be easily generalized to cases where more number of partitions are involved. For the sake of completeness we just note the following result:

Corollary 2.2: Let $(\mathcal{O},\mu)$ be a quantum probability space and $\alpha_1,\alpha_2,\ldots,\alpha_n$, $\beta_1,\beta_2,\ldots,\beta_m$ be partitions. Then we have the following properties (xi)–(xv):

(xi) $H(\alpha_1\vee\alpha_2\vee\ldots\vee\alpha_n)$

$$\geq H(\alpha_1\vee\ldots\vee\alpha_{i-1}\vee\xi_{\alpha_i}\vee\alpha_{i+1}\vee\ldots\vee\alpha_n), \tag{2.27}$$

for all $1\leq i\leq n$;

(xii) $H(\alpha_1\vee\alpha_2\vee\ldots\vee\alpha_n)$

$$\leq H(\alpha_1)+H(\xi_{\alpha_1}\vee\alpha_2)+\ldots+H(\xi_{\alpha_1}\vee\ldots\vee\xi_{\alpha_{n-1}}\vee\alpha_n), \tag{2.28}$$

$$H(\alpha_1\vee\beta_1\vee\alpha_2\vee\beta_2\vee\ldots\vee\alpha_r\vee\beta_r)$$

$$\leq H(\xi_{\alpha_1}\vee\beta_1\vee\xi_{\alpha_2}\vee\beta_2\vee\ldots\vee\xi_{\alpha_r}\vee\beta_r)$$

$$+H(\alpha_1\vee\xi_{\beta_1}\vee\alpha_2\vee\xi_{\beta_2}\vee\alpha_r\vee\xi_{\beta_r}); \tag{2.29}$$

(xiii) $H(\beta_1\vee\beta_2\vee\ldots\vee\beta_m\,|\,\alpha_1\vee\alpha_2\vee\ldots\vee\alpha_n)$

$$\geq H(\beta_1\vee\ldots\vee\beta_{j-1}\vee\xi_{\beta_j}\vee\beta_{j+1}\vee\ldots\beta_m/\alpha_1\vee\ldots\vee\alpha_n), \tag{2.30}$$

for all $1\leq j\leq n$;

(xiv) $H(\beta_1\vee\beta_2\vee\ldots\vee\beta_m/\alpha_1\vee\alpha_2\vee\ldots\vee\alpha_n)$

$$\leq H(\beta_1\vee\ldots\vee\beta_m/\alpha_1\vee\ldots\vee\alpha_{j-1}\vee\xi_{\alpha_{j+1}}\vee\ldots\vee\alpha_n), \tag{2.31}$$

for all $1\leq j\leq n$;

(xv) $H(\beta_1\vee\ldots\vee\beta_m/\alpha_1\vee\ldots\vee\alpha_n)$

$$\leq H(\beta_1/\alpha_1\vee\ldots\vee\alpha_n)+H(\xi_{\beta_1}\vee\beta_2/\alpha_1\vee\ldots\vee\alpha_n)+\ldots$$

$$+H(\xi_{\beta_1}\vee\ldots\vee\xi_{\beta_{m-1}}\vee\beta_m/\alpha_1\vee\ldots\vee\alpha_n). \tag{2.32}$$

We shall now make a few remarks on the interpretation of some of the above properties. The properties (i)–(xv) differ from their classical counterparts mainly in:

(a) the appearance of the partitions ξ_{α}, ξ_{β}, etc. in most of these relations and:

(b) the fact that we have to pay attention to the order in which partitions are combined under the operation $\vee$.

In fact, if we replace all the partitions ξ_{α}, ξ_{β}, etc. by the partition $\{I\}$ with just the identity operation, then the relations (i)–(xv) will be reduced to the corresponding classical relations. This, of course, is equivalent to stipulating that the corresponding experiments $\alpha=\{A_i\}$, $\beta=\{\beta_i\}$, etc., are "classical" (or "noninterfering") in the sense that $A_i\leq I$, $B_j\leq I$, etc.

The appearance of the partitions ξ_{α}, ξ_{β}, etc., in most of (i)–(xv) is essentially due to the "quantum interference of probabilities." To see this more clearly, let us consider, for example, the relation $H(\beta/\alpha)\leq H(\beta/\xi_{\alpha})$. $H(\beta/\alpha)$ is the uncertainty in the outcome of β, given the outcome of experiment α, when the system in state μ is subjected to the sequence of experiments $\{\alpha,\beta\}$. It is therefore very reasonable that $H(\beta/\alpha)$ is less than or equal to $H(\beta/\xi_{\alpha})$, the uncertainty in the outcome of β when the system in state μ is subjected to the same sequence of experiments $\{\alpha,\beta\}$. However, we do not have the classical relation $H(\beta/\alpha)\leq H(\beta)$, because $H(\beta)$ is the uncertainty in the outcome of β when the system in state μ is subjected to the experiment β; i.e., $H(\beta)$ *refers to a totally different experimental situation* as compared to $H(\beta/\alpha)$ or $H(\beta/\xi_{\alpha})$ and hence $H(\beta)$ is entirely different from (and in general unrelated to) $H(\beta/\xi_{\alpha})$—unless α turns out to be "classical" i.e., $\xi^{\alpha}=I$), as noted above. Therefore, *the notion of the uncertainty (or the information contained) in the outcome of an experiment (like the notion of statistical independence*[22]) becomes meaningful *iff the sequence of experiments performed on the system is also specified.*

The preceding discussion has been essentially concerned with the notion of entropy (or uncertainty in the outcome) of a random variable—that too with a finite value space. From the point of view of information theory, a more interesting quantity would be

$$I(\alpha,\beta)=H(\alpha)+H(\xi_{\alpha}\vee\beta)-H(\alpha\vee\beta), \tag{2.33}$$

which may be called as the *mutual (or correlation) information* between α and β. From (2.17) we have the relation

$$I(\alpha,\beta)\geq 0. \tag{2.34}$$

From (2.33) and (2.34) it is also clear that $I(\alpha,\beta)$ *can be interpreted as the information conveyed about the outcome of* α *by the outcome of* β *(which is also the information conveyed about the outcome of* β *by the outcome of* α*), when a system in state* μ *is subjected*

to the sequence of experiments $\{\alpha, \beta\}$. This is quite different from $I(\beta, \alpha)$ given by

$$I(\beta, \alpha) = H(\beta) + H(\xi_\beta \vee \alpha) - H(\alpha \vee \beta), \tag{2.35}$$

which is also nonnegative and refers to the mutual information between α and β when a system in state μ is subjected to the sequence of experiments $\{\beta, \alpha\}$. Another nonclassical property of the quantum mutual information is that $I(\alpha, \alpha)$ and $H(\alpha)$ are, in general, quite different because *the measurements in quantum theory are not constrained to be repeatable.*[18,20]

Finally we would like to note that the notion of mutual information $I(\alpha, \beta)$ can be directly generalized to the case of random variables whose value spaces are not finite. In fact, we can define the mutual information $I(X, Y)$ of any two random variables X and Y with standard Borel value spaces $(S_1, \mathcal{B}(S_1))$, $(S_2, \mathcal{B}(S_2))$ in the same way as in classical theory,[26] by making use of the joint probability measure given by (1.8) as follows:

$$I(X, Y) = \operatorname{Sup} \sum_{i,j} \mu(X(E_i) \wedge Y(F_j)) \times \log\left[\frac{\mu(X(E_i) \wedge Y(F_j))}{\mu(X(E_i))\mu(X(S_1) \wedge Y(F_j))}\right], \tag{2.36}$$

where the supremum is taken over all paritions $\{E_i\}$ of S_1 and $\{F_j\}$ of S_2. Again, it is clear that $I(X, Y)$ and $I(Y, X)$ are not equal and actually refer to different experimental situations.

3. QUANTUM KOLMOGOROV ENTROPY

We now make use of the information theoretic concepts outlined in Sec. 2, for the study of quantum dynamical systems. If $(\mathcal{O}, \mu, \varphi(R))$ is a quantum dynamical system and $\alpha = \{A_i \mid 1 \leq i \leq n\}$ is a partition, then, for each t, the partition $\varphi_t\alpha$ is given by

$$\varphi_t\alpha = \{\varphi_t A_i \mid 1 \leq i \leq m\}, \tag{3.1}$$

and corresponds to the same experiment as α, but conducted at a "time" t later. It is easy to see that we have the following relations for all partitions α, and any time "t."

$$\varphi_t \xi_\alpha = \xi_{\varphi_t \alpha}; \tag{3.2a}$$

$$\alpha \leq \beta \Rightarrow \varphi_t\alpha \leq \varphi_t\beta; \tag{3.2b}$$

$$\varphi_t(\alpha \vee \beta) = \varphi_t\alpha \vee \varphi_t\beta; \tag{3.3}$$

$$H(\varphi_t\alpha) = H(\alpha); \tag{3.4}$$

$$H(\varphi_t\beta / \varphi_t\alpha) = H(\beta/\alpha). \tag{3.5}$$

Our first objective will be to arrive at a reasonable measure for the "*entropy of* φ_t $(t > 0)$ *with respect to the partition* α"—a quantity which shall be denoted as $h(\alpha, \varphi_t)$. Motivated by the classical theory, we would like to define $h(\alpha, \varphi_t)$ as the limit as $n \to \infty$ (if it exists), of the average uncertainty in the outcome of α at $t = 0$, given the outcomes of the experiments $\{\varphi_{-nt}\alpha, \varphi_{-(n-1)t}\alpha, \ldots, \varphi_{-t}\alpha\}$ when the system in state μ has been subjected to the infinite sequence of experiments $\{\ldots, \varphi_{-rt}\alpha, \varphi_{-(r-1)t}\alpha, \ldots, \varphi_{-t}\alpha, \alpha\}$, i.e., when the experiment α has been repeated at interval t, from the infinite past. Before discussing the question as to whether the above limit exists, it is necessary to first restrict ourselves to the class of all those experiments (with a finite number of outcomes) for which such an infinite repetetion does make sense in our theory. We are thus led to the following notion of an admissible partition: α is said to be an *admissible partition* of a dynanical system $(\mathcal{O}, \mu, \varphi(R))$ iff

$$\xi_\alpha^{r,t} = \lim_{n \to \infty}[\varphi_{-(r+n)t}\xi_\alpha \vee \varphi_{-(r+n-1)t}\xi_\alpha \vee \ldots \vee \varphi_{-rt}\xi_\alpha] \tag{3.6}$$

exists for all $t > 0$, where the right-hand side is required to converge in the strong operator topology (when we replace ξ_α by ξ^α). Since we have that φ_τ (for each $\tau \in R$) is continuous in the strong operator topology on $\mathcal{O}$, we have the relations

$$\varphi_{mt}\xi_\alpha^{r,t} = \xi_\alpha^{r-m,t} \tag{3.7a}$$

$$\varphi_\tau \xi_\alpha^{r,t} = \xi_{\varphi_\tau \alpha}^{r,t}, \tag{3.7b}$$

$$\xi_\alpha^{r,t} = \xi_\alpha^{r+1,t} \vee \varphi_{-rt}\xi_\alpha. \tag{3.7c}$$

We are now in a position to show that $h(\alpha, \varphi_t)$ exists for every admissible partition α.

Theorem 3.1: Let α be an admissible partition of a dynamical system $(\mathcal{O}, \mu, \varphi(R))$. Then the entropy $h(\alpha, \varphi_t)$ of φ_t with respect to partition α, given by

$$h(\alpha, \varphi_t) = \lim_{n \to \infty} H(\alpha / \xi_\alpha^{n,t} \vee \varphi_{-(n-1)t}\alpha \vee \ldots \vee \varphi_{-t}\alpha), \tag{3.8}$$

exists for all $t > 0$. Also

$$h(\alpha, \varphi_t) = \lim_{n \to \infty} \frac{1}{n} H(\xi_\alpha^{n,t} \vee \varphi_{-(n-1)t}\alpha \vee \ldots \vee \alpha) \tag{3.9}$$

for all $t > 0$.

Proof: Let us write

$$s_n = H(\alpha / \xi_\alpha^{n,t} \vee \varphi_{-(n-1)t}\alpha \vee \ldots \vee \varphi_{-t}\alpha). \tag{3.10}$$

Then $s_n \geq 0$ and from (2.31), (3.2) and (3.7), we obtain

$$s_n = H(\alpha / \xi_\alpha^{n+1,t} \vee \xi_{\varphi_{-nt}\alpha} \vee \varphi_{-(n-1)t}\alpha \vee \ldots \vee \varphi_{-t}\alpha) \geq H(\alpha / \xi_\alpha^{n+1,t} \vee \varphi_{-nt}\alpha \vee \varphi_{-(n-1)t}\alpha \vee \ldots \vee \varphi\alpha_t\alpha) = s_{n+1}. \tag{3.11}$$

Thus $\{s_n\}$ is a monotonic decreasing sequence of nonnegative numbers and hence

$$h(\alpha, \varphi_t) = \lim_{n \to \infty} s_n$$

exists. Now if we set

$$h_n = H(\xi_\alpha^{n,t} \vee \varphi_{-(n-1)t}\alpha \vee \ldots \vee \alpha),$$

then, because of (3.3), (3.4), and (3.7), we have the relation

$$h_n - h_{n-1} = s_n. \tag{3.12a}$$

In addition, since

$$h_1 = H(\xi_\alpha^{1,t} \vee \alpha) = H(\alpha / \xi_\alpha^{1,t}) = s_1,$$

we obtain

$$h_n = \sum_{k=1}^{n} s_k. \tag{3.12b}$$

Since we have already shown that $\{s_n\}$ tends to

$h(\alpha, \varphi_t)$ as $n \to \infty$, we just have to apply Cesaro's mean convergence theorem,[1] to obtain from (3.12b) the required result

$$h(\alpha, \varphi_t) = \lim_{n\to\infty} \frac{1}{n} H(\xi_\alpha^{n,t} \vee \varphi_{-(n-1)t}\alpha \vee \cdots \vee \alpha).$$

From the definition (3.8), it is clear that for any admissible partition α, $h(\alpha, \varphi_t)$ has the interpretation (which we sought for in the first place) of being the mean uncertainty in the outcome of α given the outcomes of the infinite sequence of experiments $\{\cdots, \varphi_{-rt}\alpha, \varphi_{-(r-1)t}\alpha, \ldots, \varphi_{-t}\alpha\}$. From (3.9), we see that $h(\alpha, \varphi_t)$ is also equal to

$$\lim_{n\to\infty} \frac{1}{n}\left[\begin{array}{l}\text{Information gained when the outcomes of}\\ \text{experiments } \{\varphi_{-(n-1)t}\alpha, \varphi_{-(n-2)t}\alpha, \ldots, \alpha\}\\ \text{are specified}\end{array}\right]$$

when the system in state μ is subjected to the infinite sequence of experiments $\{\cdots, \varphi_{-rt}\alpha, \varphi_{-(r-1)t}\alpha, \ldots, \alpha\}$. In other words, $h(\alpha, \varphi_t)$ is the rate of average gain of information per single observation, when the experiment is repeated at interval "t" from the infinite past.

As a matter of fact, our definition of the entropy $h(\alpha, \varphi_t)$ of φ_t with respect to partition α coincides with the notion of the *entropy of a quantum information source* which in this case is composed of the "stationary sequence" $\{\varphi_{mt}\alpha \mid m = 0, \pm 1, \pm 2, \cdots\}$ of random variables $\varphi_{mt}\alpha$ with a finite value space (i.e., a finite set of alphabets)—strictly speaking, we should replace the partition α by a corresponding random variable X_α [cf. (2.2)]. The successive "outputs" from our information source are nothing but the outcomes of the experiment α performed at intervals of time "t". The *source entropy* $h(\alpha, \varphi_t)$ can be easily calculated (to obtain results analogous to the classical theory[16,25]) in the following cases:

1. *Bernoulli source*: For each $r \geq 1$ and any integer m, the sequence of random variables $\{\varphi_{mt}\alpha, \varphi_{(m+1)t}\alpha, \ldots, \varphi_{(m+r)t}\alpha\}$ is statistically independent[22] when the system in state μ is subjected to the infinite sequence of experiments $\{\cdots, \varphi_{-nt}\alpha, \varphi_{-(n-1)t}\alpha, \ldots\}$. We therefore have, for each $r \geq 1$,

$$\begin{aligned} &H(\xi_\alpha^{r,t} \vee \varphi_{-(r-1)t}\alpha \vee \cdots \vee \alpha)\\ &= \sum_{m=1}^{r} H(\xi_\alpha^{m,t} \vee \varphi_{-(m-1)t}\alpha)\\ &= rH(\xi_\alpha^{1,t} \vee \alpha), \end{aligned} \tag{3.13}$$

where the last step follows from (3.3), (3.4), and (3.7). Hence the entropy of Bernoulli source is given by the relation

$$h(\alpha, \varphi_t) = H(\xi_\alpha^{1,t} \vee \alpha). \tag{3.14}$$

2. *Markov source*: Here the infinite sequence of random variables $\{\ldots, \varphi_{-rt}\alpha, \varphi_{-(r-1)t}\alpha, \ldots\}$ forms a quantum Markov chain.[22] It can then be shown that for all $r \geq 2$

$$\begin{aligned} &H(\alpha/\xi_\alpha^{r,t} \vee \varphi_{-(r-1)}\alpha \vee \cdots \vee \varphi_{-t}\alpha)\\ &= H(\alpha/\xi_\alpha^{2,t} \vee \varphi_{-t}\alpha). \end{aligned} \tag{3.15}$$

Hence the entropy of a Markov source is given by the relation

$$h(\alpha, \varphi_t) = H(\alpha/\xi_\alpha^{2,t} \vee \varphi_{-t}\alpha). \tag{3.16}$$

In the following theorem, we shall collect some of the properties of the entropy $h(\alpha, \varphi_t)$ of φ_t with respect to an admissible partition α of a general dynamical system.

Theorem 3.2: Let α, β be admissible partitions of a dynamical system $(\mathcal{O}, \mu, \varphi(R))$. Then the following properties (i)—(v) are valid for each $t > 0$:

(i) $$h(\alpha, \varphi_t) = \lim_{n\to\infty} \frac{1}{n} H(\xi_\alpha^{1,t} \vee \alpha \vee \varphi_t\alpha \vee \cdots \vee \varphi_{(n-1)t}\alpha); \tag{3.17}$$

(ii) $(1/n)H(\xi_\alpha^{n,t} \vee \varphi_{-(n-1)t}\alpha \vee \cdots \vee \alpha)$ decreases monotonically to the limit $h(\alpha, \varphi_t)$. As a consequence

$$h(\alpha, \varphi_t) \leq H(\xi_\alpha^{1,t} \vee \alpha); \tag{3.18}$$

(iii) If $\alpha \leq \beta$, then

$$h(\alpha, \varphi_t) \leq h(\beta, \varphi_t); \tag{3.19}$$

(iv) If $\alpha \vee \beta$ is admissible, then $\alpha \vee \xi_\beta$ and $\xi_\alpha \vee \beta$ are also admissible and

$$h(\alpha \vee \beta, \varphi_t) \leq h(\alpha \vee \xi_\beta, \varphi_t) + h(\xi_\alpha \vee \beta, \varphi_t); \tag{3.20}$$

(v) If m is a positive integer, then

$$\begin{aligned} &h(\varphi_{-(m-1)t}\alpha \vee \varphi_{-(m-2)t}\alpha \vee \cdots \vee \alpha, \varphi_{mt})\\ &= mh(\alpha, \varphi_t) \end{aligned} \tag{3.21}$$

Proof: (i) This follows directly from (3.3), (3.4), and (3.7).

(ii) From (3.10)—(3.12) we have

$$\begin{aligned} h_n &= s_1 + s_2 + \cdots + s_n\\ &\geq ns_{n+1}\\ &= n(h_{n+1} - h_n). \end{aligned}$$

Therefore

$$h_n/n \geq h_{n+1}/(n+1), \tag{3.22}$$

for all $n \geq 1$, which is the desired result.

(iii) If $\alpha \leq \beta$, then it is easy to see that $\xi_\alpha = \xi_\beta$. Hence we have from (2.6c) and (3.2b) that

$$\begin{aligned} &\xi_\alpha^{n,t} \vee \varphi_{-(n-1)t}\alpha \vee \cdots \vee \alpha\\ &\leq \xi_\beta^{n,t} \vee \varphi_{-(n-1)t}\beta \vee \cdots \vee \beta, \end{aligned}$$

from which (3.19) can be deduced easily by making use of (2.12) and (3.9).

(iv) Since we have

$$\xi_{\alpha\vee\beta} = \xi_{\alpha\vee\xi_\beta} = \xi_{\xi_\alpha\vee\beta} = \xi_\alpha \vee \xi_\beta, \tag{3.23}$$

and $\alpha \vee \beta$ is given to be admissible, we can easily show that

$$\xi_{\alpha\vee\xi_\beta}^{r,t} = \xi_{\xi_\alpha\vee\beta}^{r,t} = \xi_{\alpha\vee\beta}^{r,t}. \tag{3.24}$$

If in the inequality (2.29), we take

$$\alpha_1 = \xi_{\alpha\vee\beta}^{r,t},$$

$$\alpha_i = \varphi_{-(r-i+1)t}\alpha,$$

for $i > 1$ and $\beta_j = \varphi_{-(r-j)t}\beta$ for all j, then we get

$$H(\xi^{r,t}_{\alpha\vee\beta}\vee\varphi_{-(r-1)t}(\alpha\vee\beta)\vee\cdots\vee(\alpha\vee\beta))$$

$$\leq H(\xi^{r,t}_{\xi_\alpha\vee\beta}\vee\varphi_{-(r-1)t}(\xi_\alpha\vee\beta)\vee\cdots\vee(\xi_\alpha\vee\beta))$$

$$+H(\xi^{r,t}_{\alpha\vee\xi_\beta}\vee\varphi_{-(r-1)t}(\alpha\vee\xi_\beta)\vee\cdots\vee(\alpha\vee\xi_\beta)), \tag{3.26}$$

from which (3.20) follows immediately.

(v) If we write

$$\gamma=\varphi_{-(m-1)t}\alpha\vee\varphi_{-(m-2)t}\alpha\vee\cdots\vee\alpha, \tag{3.27}$$

then

$$\xi_\gamma=\varphi_{-(m-1)t}\xi_\alpha\vee\varphi_{-(m-2)t}\xi_\alpha\vee\cdots\vee\xi_\alpha, \tag{3.28}$$

and a simple calculation shows that γ is also admissible and that

$$\xi^{n,mt}_\gamma=\xi^{nm,t}_\alpha. \tag{3.29}$$

We therefore have

$$h(\gamma,\varphi_{mt})$$

$$=\lim_{n\to\infty}\frac{1}{n}H(\xi^{n,mt}_\gamma\vee\varphi_{-(n-1)mt}\gamma\vee\cdots\vee\varphi_{-mt}\gamma\vee\gamma)$$

$$=m\lim_{n\to\infty}\frac{1}{nm}H(\xi^{nm,t}_\alpha\vee\varphi_{-(nm-1)t}\alpha\vee\varphi_{-(nm-2)t}\alpha\vee\cdots\vee\varphi_{-t}\alpha\vee\alpha)$$

$$=mh(\alpha,\varphi_t), \tag{3.30}$$

which is the required result.

The properties (i), (ii), (iii), and (v) are also valid in classical theory. The property (iv) is different from the corresponding classical relation $h(\alpha\vee\beta,\varphi_t)\leq h(\alpha,\varphi_t)+h(\beta,\varphi_t)$. Another interesting nonclassical feature is that $h(\alpha,\varphi_t)$ need not be zero even if φ_t reduces to the identity transformation.

From our interpretation of $h(\alpha,\varphi_t)$ outlined earlier, it is clear that if for some admissible partition α we have $h(\alpha,\varphi_t)>0$, then our dynamical system exhibits some form of "stochastic behavior" as even after infinite repetitions of the experiment α (at fixed interval "t"), the outcome of the next experiment is still uncertain. In order to obtain a criterion as to when such a stochastic behavior may be expected, we shall now introduce the notion of the quantum Kolmogorov entropy.

Let $(\mathcal{O},\mu,\varphi(R))$ be a quantum dynamical system. For each $t>0$, the *entropy* $h(\varphi_t)$ *of the automorphism* φ_t is defined as

$$h(\varphi_t)=\sup h(\alpha,\varphi_t), \tag{3.31}$$

where the supremum is taken over all the admissible partitions α of the dynamical system $(\mathcal{O},\mu,\varphi(R))$. The *(quantum) Kolmogorov entropy* $h(\varphi)$ of the dynamical system $(\mathcal{O},\mu,\varphi(R))$ can now be defined as

$$h(\varphi)=\sup_{t>0}(1/t)h(\varphi_t). \tag{3.32}$$

For classical dynamical systems, it is a well-known result due to Abramow[2] that $(1/t)h(\varphi_t)$ is a constant so that $h(\varphi)=h(\varphi_1)$. At present we do not know whether a similar result can be obtained for quantum dynamical systems under the conditions (QD1)–(QD3) of Sec. 1.

Finally, we may note that the notion of a K-system can also be introduced in a manner analogous to the one employed in classical theory. As we have already remarked, a positive value for the Kolmogorov entropy $h(\varphi)$ indicates that the dynamical system exhibits stochastic behavior. A far more stronger statistical property is the following K-property: A dynamical system $(\mathcal{O},\mu,\varphi(R))$ is said to be a *K-system* iff, for each $t>0$ and every nontrivial admissible partition α,

$$h(\alpha,\varphi_t)>0. \tag{3.33}$$

Thus a K-system, has the property of "essential randomness" such that *no experiment (with a finite number of outcomes) on the system is deterministic*[27] in the sense that no matter how many times any given experiment is performed on the system (at any fixed interval of time t) the outcome of the next experiment is still uncertain.

[1] V.I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics* (Benjamin, New York, 1968).
[2] J.G. Sinai, Theory of Dynamical Systems, Lecture notes, Matematisk Institut, Aarhus Universitet (1970).
[3] A.S. Wightman, "Statistical mechanics and ergodic theory," in *Statistical Mechanis at the Turn of the Decade*, edited by E.G.D. Cohen (Dekker, New York, 1971).
[4] Lectures of Y.G. Sinai, D. Ruelle, and O.E. Lanford, in *Boltzman Equation*, edited by E.G.D. Cohen (Springer-Verlag, Berlin, 1973).
[5] J.L. Lebowitz, "Ergodic theory and Statistical mechanics," in *Transport Phenomena*, edited by G. Kriczenow and J. Marro (Springer-Verlag, Berlin, 1974).
[6] D.S. Ornstein, *Ergodic Theory, Randomness and Dynamical Systems* (Yale U.P., New Haven, 1973).
[7] A.N. Kolmogorov, "A new metric invariant of transient dynamical systems and automorphisms of Lebesgue spaces," Dokl. Akad. Nauk SSSR **119**, 861 (1958).
[8] D. Ruelle, *Statistical Mechanics* (Benjamin, New York, 1969), Chaps. 6, 7.
[9] G.L. Sewell, "Ergodic theory in algebraic statistical mechanics," in *Lectures in Theoretical Physics, Vol. XIV B'*, edited by W. Brittin (Colorado Ass. U.P., Boulder, 1973).
[10] G.G. Emch, *Algebraic Methods in Statistical Mechanics and Quantum Field Theory* (Wiley, New York, 1972).
[11] D.W. Robinson, *C**-algebras and quantum statistical mechanics, Varenna Lecture notes (1973).
[12] G.G. Emch, "Positivity of the K-entropy on non-Abelian K-flows," Z. Wahrscheinlichkeits theorie **29**, 241 (1974); "Generalised K flows," Commun. Math. Phys. **49**, 191 (1976).
[13] A Connes and E. Størmer, "Entropy for automorphisms of II_1 von Neumann algebras," Acta. Math. **134**, 289 (1975).
[14] These remarks certainly do not apply to the framework introduced by Haag and Kastler,[15] where the pure operations of a physical theory were associated with the set of elements in the unit ball of a *C**-algebra. Then, of course, the observables of the theory which will have to be defined as operation valued measures cannot be identified with the self-adjoint elements of the algebra.
[15] R. Haag and D. Kastler, "An algebraic approach to quantum field theory," J. Math. Phys. **5**, 848 (1964).
[16] R.S. Ingarden, "Information Theory and Thermodynamics, Part I—Information Theory," Chap. I—"Classical Theory," Torun, Chap. II—"Quantum Theory," Torun preprint (1974), pp. 270, 275 resp.
[17] R.S. Ingarden, "Quantum information theory," Rep. Math. Phys. **10**, 43 (1976).

[18]E. B. Davies and J. T. Lewis, "An operational approach to quantum probability theory," Commun. Math. Phys. **17**, 239 (1970).
[19]C. M. Edwards, "The operational approach to algebraic quantum theory I," Commun. Math. Phys. **16**, 207 (1970).
[20]M. D. Srinivas, "Foundations of a quantum probability theory," J. Math. Phys. **16**, 1672 (1975) [to be reprinted in *The Logico Algebraic Approach to Quantum Theory, Vol. II*, edited by C. A. Hooker (Reidel, Dordrecht, 1978)].
[21]M. D. Srinivas, "Quantum Counting processes" J. Math. Phys. **18**, 2138 (1977).
[22]M. D. Srinivas, "Conditional probabilities and statistical independence in quantum theory," J. Math. Phys. **19**, 1705 (1978).
[23]D. Kastler, "Equilibrium states of matter and operator algebras," in *Symposia Mathematica* (Academic, New York) (to be published).
[24]O. E. Lanford, "Time evolution of large classical systems," in *Dynamical Systems, Theory and Applications*, edited by J. Moser (Springer-Verlag, Berlin, 1975).
[25]A. I. Khinchin, *Mathematical Foundations of Information Theory* (Dover, New York, 1957).
[26]M. S. Pinsker, *Information and Information Stability of Random Variables and Processes* (Holden-Day, San Francisco, 1964).
[27]D. S. Ornstein, "What does it mean for a mechanical system to be isomorphic to the Bernoulli flow?", in *Dynamical Systems, Theory and Applications*, edited by J. Moser (Springer-Verlag, Berlin, 1975).

Spherical delta functions and multipole expansions

E. G. Peter Rowe[a)]

Department of Mathematics, University of Jos, Jos, Nigeria
(Received 5 May 1977)

The Cartesian Taylor series for an analytic function in three dimensions is rewritten as a series of solid spherical harmonics. A discussion of the distribution theory definition of singular spherical harmonics is given, which leads to a definition of spherical delta functions. An expansion of source functions in spherical delta functions and their derivatives leads to multipole expansions for the fields which, in a distribution theory sense, are valid everywhere.

1. SPHERICAL HARMONICS

The solid spherical harmonics in three dimensions are a complete set of linearly independent polynomial solutions of Laplace's equation. Those of degree l may be written

$$Y_{lm}(\mathbf{x}) = 2^{-l}(l!)^{-1}[(2l+1)/4\pi]^{1/2}[(l-m)!\,(l+m)!\,]^{1/2}\, W_{lm}(\mathbf{x}), \tag{1}$$

where the W_{lm} are coefficients in the expansion

$$[l^{-1}(x_1 - ix_2) + 2x_3 - l(x_1 + ix_2)]^l = \sum_{m=-l}^{+l} l^m W_{lm}(\mathbf{x}). \tag{2}$$

An equation of this form is given in Ref. 1, p. 248. By applying the Laplacian operator to the left-hand side of (2) one sees that the W_{lm} satisfy Laplace's equation. The numerical coefficient in (1) is chosen so that the $Y_{lm}(\mathbf{x})$ will have convenient properties. These properties are summarized in the Condon and Shortley phase convention.[2] A sketch of the connection between this definition of Y_{lm} and the usual one is given in Appendix A.

The surface spherical harmonics $Y_{lm}(\theta, \phi) \equiv Y_{lm}(\hat{\mathbf{x}})$ differ from $Y_{lm}(\mathbf{x})$ by the factor r^l, where $r \equiv |\mathbf{x}|$:

$$Y_{lm}(\mathbf{x}) = r^l Y_{lm}(\hat{\mathbf{x}}). \tag{3}$$

We have

$$Y_{lm}(-\mathbf{x}) = (-1)^l Y_{lm}(\mathbf{x}), \tag{4}$$

$$Y^*_{lm}(\mathbf{x}) = (-1)^m Y_{l-m}(\mathbf{x}), \tag{5}$$

and

$$\int d\Omega\, Y_{lm}(\mathbf{x})\, Y^*_{l'm'}(\mathbf{x}) = r^{2l}\delta_{ll'}\delta_{mm'}, \tag{6}$$

where $d\Omega = \sin\theta\, d\theta d\phi$ is the differential surface element on the unit sphere.

2. SPHERICAL TAYLOR SERIES

If $\psi(\mathbf{x})$ is a real analytic function, that is, if it can be expanded in a Taylor series in a ball around the origin, then we may write the uniformly convergent Cartesian Taylor expansion in the symbolic form

$$\psi(\mathbf{x}) = e^{\mathbf{x}\cdot\nabla}\psi(\mathbf{0}) = \sum_{n=0}^{\infty} \frac{1}{n!}(\mathbf{x}\cdot\nabla)^n \psi(\mathbf{0}). \tag{7}$$

The meaning of this symbolism is clear. The nth term in the series is

$$\frac{1}{n!}(\mathbf{x}\cdot\nabla)^n\psi(\mathbf{0}) \equiv \frac{1}{n!} x_{i_1}\cdots x_{i_n} \left.\frac{\partial^n \psi}{\partial x_{i_1}\cdots\partial x_{i_n}}\right|_{\mathbf{x}=\mathbf{0}}, \tag{8}$$

in which a summation convention is used.

The terms in the uniformly convergent "exponential" series (7) may be regrouped. Provided it is remembered that ∇ and ∇^2 finally "act" on a function ψ, as in (8), we can treat ∇ as a commuting algebraic quantity. Understanding ∇ in this sense, we may rearrange the terms of the exponential series to form a series of solid spherical harmonics whose coefficients are power series in $r^2\nabla^2$, in which $(r^2\nabla^2)^K$ is interpreted as $r^{2K}(\nabla^2)^K$. The expression we get is

$$e^{\mathbf{x}\cdot\nabla} = \sum_{l=0}^{\infty}\sum_{m=-l}^{+l} 4\pi\, Y_{lm}(\mathbf{x}) S_l(r^2\nabla^2)\, Y^*_{lm}(\nabla), \tag{9}$$

where the function S_l is given by

$$S_l(x) = \sum_{K=0}^{\infty} \frac{x^K}{2^K K!\,(2l+2K+1)!!} \tag{10}$$

with the double factorial notation

$$(2n+1)!! = (2n+1)(2n-1)(2n-3)\cdots 5\cdot 3\cdot 1.$$

The formal algebraic expression (9) is developed in Appendix B, where some useful properties of the functions S_l will also be found.

When the Cartesian Taylor series is regrouped according to the symbolic expansion (9), we may call it a spherical Taylor series:

$$\psi(\mathbf{x}) = \exp(\mathbf{x}\cdot\nabla)\psi(\mathbf{0}) = \sum_{lm} 4\pi\, Y_{lm}(\mathbf{x}) S_l(r^2\nabla^2)\, Y^*_{lm}(\nabla)\psi(\mathbf{0}). \tag{11}$$

I have not seen this formula in the literature, but it is the sort of relation that would appeal to late nineteenth century taste, in particular to Hobson's. It is equivalent to a formula of Hobson [below, Eqs. (12) and (14)].

The usual expansion of a continuous function $\psi(\mathbf{x})$ in spherical harmonics is

$$\psi(\mathbf{x}) = \sum_{lm} \psi_{lm}(r)\, Y_{lm}(\mathbf{x}), \tag{12}$$

where

$$\psi_{lm}(r) = r^{-2l}\int d\Omega \psi(\mathbf{x})\, Y^*_{lm}(\mathbf{x}). \tag{13}$$

If $\psi(\mathbf{x})$ is analytic, we may compare (11) and (12). In this case $\psi_{lm}(r)$ is the even power series in r given by

a) Permanent address: Department of Mathematics, University of Durham, Durham, England.

 0022-2488/78/1909-1962$1.00

$$\psi_{lm}(r) = 4\pi S_l(r^2\nabla^2) Y^*_{lm}(\nabla)\psi(0). \tag{14}$$

This formula is given by Hobson (Ref. 3, p. 161).

The special case of (13) and (14) for $l=0$ is known as the mean value formula of Pizetti[4]:

$$\int d\Omega\psi(\mathbf{x}) = 4\pi S_0(r^2\nabla^2)\psi(0)$$

$$= 4\pi \sum_{K=0}^{\infty} \frac{r^{2K}\nabla^{2K}}{(2K+1)!}\psi(0). \tag{15}$$

If $\nabla^2\psi = 0$ in a region around the origin, the right-hand side reduces to $4\pi\psi(0)$.

3. SINGULAR SPHERICAL HARMONICS

If the point $\mathbf{x}$ is closer to the origin than a point $\mathbf{y}$, the generating function for the Legendre polynomials gives

$$\frac{1}{|\mathbf{x}-\mathbf{y}|} = \sum_{l=0}^{\infty} \frac{r^l}{y^{l+1}} P_l(\hat{\mathbf{x}}\cdot\hat{\mathbf{y}}), \quad r < |\mathbf{y}| \equiv y. \tag{16}$$

Using the addition theorem for spherical harmonics, we get

$$\frac{1}{|\mathbf{x}-\mathbf{y}|} = \sum_{lm} \frac{4\pi}{2l+1} Y_{lm}(\mathbf{x}) \frac{Y^*_{lm}(\mathbf{y})}{y^{2l+1}}. \tag{17}$$

Equation (17) is the spherical Taylor series for $|\mathbf{x}-\mathbf{y}|^{-1}$ in the ball $r < y$. Identifying terms in (11) and (17) gives

$$\frac{4\pi}{2l+1} \frac{Y^*_{lm}(\mathbf{y})}{y^{2l+1}} = 4\pi S_l(r^2\nabla^2) Y^*_{lm}(\nabla) \frac{1}{|\mathbf{x}-\mathbf{y}|}\bigg|_{\mathbf{x}=0}. \tag{18}$$

Since $\nabla^2 |\mathbf{x}-\mathbf{y}|^{-1} = 0$ for $r < y$, only the first term, $S_l(0) = [(2l+1)!!]^{-1}$, in the series for $S_l(r^2\nabla^2)$ gives a non-zero contribution. Taking the complex conjugate and replacing $\mathbf{y}$ with $\mathbf{x}$, we get

$$\frac{Y_{lm}(\mathbf{x})}{r^{2l+1}} = \frac{1}{(2l-1)!!} Y_{lm}(-\nabla)\frac{1}{r} \quad (r \neq 0). \tag{19}$$

In this equation $(2l-1)!! = (2l+1)!!/(2l+1)$, so, for $l=0$, $(-1)!! = 1$.

We call $Y_{lm}(\mathbf{x})/r^{2l+1}$ a singular spherical harmonic; it is defined by (19) for $r \neq 0$, and in this region it satisfies Laplace's equation. The formula (19) is given by Hobson (Ref. 3, p. 127); the method used here to get it may be generalized.[5] Because $\nabla^2(1/r) = 0$ for $r \neq 0$, there are many alternative ways of writing the singular spherical harmonics as derivatives of $1/r$ for $r \neq 0$. A well-known expression is

$$\frac{Y_{lm}(\mathbf{x})}{r^{2l+1}} = C_{lm} \left(\frac{\partial}{\partial x_3}\right)^{l-m} \left(\frac{\partial}{\partial x_1} + i\frac{\partial}{\partial x_2}\right)^m \frac{1}{r} \quad (m \geq 0), \tag{19'}$$

in which the C_{lm} are constants. Formula (19') is given in Hobson (Ref. 3, p. 134) and quoted by Erdélyi (Ref. 1, p. 251).

If we are interested in the extension of (19) and (19') to distribution theory formulas, in which the derivatives are interpreted in the sense of distribution theory, we must note that they differ, in general, by derivatives of the delta function. The simplest example is for $l=2$, $m=0$. Using

$$Y_{20}(\hat{\mathbf{x}}) = \left(\frac{5}{4\pi}\right)^{1/2} \left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right)$$

we have

$$Y_{20}(-\nabla) = \left(\frac{5}{4\pi}\right)^{1/2} \frac{3}{2}\left(\frac{\partial^2}{\partial x_3^2} - \frac{1}{3}\nabla^2\right).$$

The generalized form of (19) gives

$$\frac{Y_{20}(\mathbf{x})}{r^5} = \frac{1}{3}\left(\frac{5}{4\pi}\right)^{1/2} \frac{3}{2}\left(\frac{\partial^2}{\partial x_3^2}\frac{1}{r} + \frac{4\pi}{3}\delta(\mathbf{x})\right),$$

whereas (19') gives only the first term. We shall find that it is the distribution theory form of (19) which agrees with an attractive alternative distribution theory definition of the singular spherical harmonics.

We define the singular spherical harmonics as generalized functions or distributions, by a relation of the form (19) but in which the derivatives are understood in the sense of distribution theory:

$$\frac{Y_{lm}(\mathbf{x})}{r^{2l+1}} \equiv \frac{1}{(2l-1)!!} Y_{lm}(-\nabla)\frac{1}{r}. \tag{20}$$

In (20), $1/r$ is the regular functional corresponding to the function $1/r$.

Using the notation of Gel'fand and Shilov,[6] the value of the functional (20) on a test function $\psi(\mathbf{x})$ is

$$\left(\frac{Y_{lm}}{r^{2l+1}}, \psi\right) = \frac{1}{(2l-1)!!}\left(Y_{lm}(-\nabla)\frac{1}{r}, \psi\right)$$

$$= \frac{1}{(2l-1)!!}\left(\frac{1}{r}, Y_{lm}(\nabla)\psi\right)$$

$$= \int d\mathbf{x} \frac{1}{(2l-1)!!} \frac{Y_{lm}(\nabla)\psi(\mathbf{x})}{r}. \tag{21}$$

In order to make possible some manipulations below, we must suppose that the test functions ψ are real and analytic everywhere, that is, that (7) and (11) are valid for all $\mathbf{x}$. We also presume that ψ, and its derivatives of all orders, vanish faster than r^{-N} for any integer N as $r \to \infty$. The simplest example of such a function is $\exp(-r^2)$. Our test functions ψ belong to the space Z (restricted to real argument) used by Gel'fand and Shilov in their discussion of Fourier transforms.

Under the assumptions of the previous paragraph we can write a spherical Taylor series (11) for the function $Y_{lm}(\nabla)\psi(\mathbf{x})$:

$$Y_{lm}(\nabla)\psi(\mathbf{x})$$

$$= 4\pi \sum_{l'm'} Y_{l'm'}(\mathbf{x}) S_{l'}(r^2\nabla^2) Y^*_{l'm'}(\nabla) Y_{lm}(\nabla)\psi(0). \tag{22}$$

Inserting (22) in (21) and integrating over angles, we find

$$\left(\frac{Y_{lm}(\mathbf{x})}{r^{2l+1}}, \psi\right) = \int r^2 dr \frac{4\pi}{(2l-1)!!}\frac{1}{r} S_0(r^2\nabla^2) Y_{lm}(\nabla)\psi(0). \tag{23}$$

The integrand is a convergent power series for all r, but, of course, we cannot integrate term by term over the infinite range. However, using (B11) and then (14), we can re-express the right-hand side of (23) as

$$\int dr \frac{4\pi}{(2l-1)!!}\left[\frac{1}{r}\frac{d}{dr}\right]^l [r^{2l+1}S_l(r^2\nabla^2)]Y_{lm}(\nabla)\psi(0)$$

$$=\int_0^\infty dr \frac{1}{(2l-1)!!}\left[\frac{1}{r}\frac{d}{dr}\right]^l [r^{2l+1}\psi^*_{lm}(r)]. \tag{24}$$

We may integrate by parts l times. The boundary terms vanish at $r=0$ because ψ_{lm} and its derivatives are finite there, and at $r=\infty$ because ψ belongs to the space Z. We get finally

$$\left(\frac{Y_{lm}}{r^{2l+1}},\psi\right) = \int_0^\infty dr\, r\psi^*_{lm}(r). \tag{25}$$

Using (25), we may show that the distribution theory definition (20) for the singular spherical harmonics coincides with the alternative integral definition

$$\left(\frac{Y_{lm}}{r^{2l+1}},\psi(\mathbf{x})\right) = \int d\mathbf{x}\theta(r-\epsilon)\frac{Y_{lm}}{r^{2l+1}}\psi(\mathbf{x}). \tag{26}$$

In (26), θ is the unit step function

$$\theta(\xi)=1, \quad \xi > 0,$$

$$\theta(\xi)=0, \quad \xi < 0,$$

and it serves to exclude the region $0 \le r < \epsilon$ from the range of integration. The limit $\epsilon \to 0$ is understood. In view of (13), the right-hand side of (26) may be rewritten

$$\int_\epsilon^\infty r^2 dr \frac{1}{r}\psi^*_{lm}(r) = \int_0^\infty r dr\, \psi^*_{lm}.$$

Therefore, the derivative definition (20) is the same as the integral definition (26). We may express this equality in the form

$$\frac{Y_{lm}(\mathbf{x})}{r^{2l+1}} \equiv \frac{1}{(2l-1)!!}Y_{lm}(-\nabla)\frac{1}{r} = \theta(r-\epsilon)\frac{Y_{lm}(\mathbf{x})}{r^{2l+1}}. \tag{27}$$

The notation in the right-hand term is meant to recall the integral distribution definition (26); it is not a product.

4. SPHERICAL DELTA FUNCTIONS

Using the important distribution theory formula

$$\nabla^2\frac{1}{r} = -4\pi\delta(\mathbf{x}), \tag{28}$$

we find that the singular spherical harmonics (20) satisfy the Poisson equations

$$\nabla^2\frac{Y_{lm}(\mathbf{x})}{r^{2l+1}} = -\frac{4\pi}{(2l-1)!!}Y_{lm}(-\nabla)\delta(\mathbf{x}) = -4\pi\delta_{lm}(\mathbf{x}), \tag{29}$$

where the spherical delta functions are defined by

$$\delta_{lm}(\mathbf{x}) \equiv \frac{1}{(2l-1)!!}Y_{lm}(-\nabla)\delta(\mathbf{x}). \tag{30}$$

The definition (30) is not, of course, restricted to the space of analytic test functions, but if ψ does belong to Z then the value of the distribution $\delta_{lm}(\mathbf{x})$ on ψ is, using (14),

$$(\delta_{lm},\psi) = \frac{1}{(2l-1)!!}(\delta, Y_{lm}(\nabla)\psi)$$

$$= \frac{2l+1}{4\pi}\psi^*_{lm}(0). \tag{31}$$

Similarly, the value of $\nabla^{2K}\delta_{lm}(\mathbf{x})$ on $\psi(\mathbf{x})$ is

$$(\nabla^{2K}\delta_{lm},\psi) = \frac{1}{(2l-1)!!}\nabla^{2K}Y_{lm}(\nabla)\psi(0)$$

$$= \frac{2l+1}{4\pi}\frac{d^K}{d(r^2)^K}\psi^*_{lm}(0)\frac{2^K(2l+2K+1)!!}{(2l+1)!!}. \tag{32}$$

If the Fourier transform of a function $F(\mathbf{x})$ is defined by

$$\tilde{F}(\mathbf{k}) = \int d\mathbf{x}F(\mathbf{x})\exp(-i\mathbf{k}\cdot\mathbf{x}), \tag{33}$$

then using the elementary transforms $\tilde{\delta}=1$, $(1/r)\tilde{} = 4\pi/\mathbf{k}^2$, we find

$$(\delta_{lm})\tilde{} = \frac{(-i)^l}{(2l-1)!!}Y_{lm}(\mathbf{k}) \tag{34}$$

and

$$\left(\frac{Y_{lm}}{r^{2l+1}}\right)\tilde{} = \frac{4\pi(-i)^l}{(2l-1)!!}\frac{Y_{lm}(\mathbf{k})}{\mathbf{k}^2}. \tag{35}$$

5. MULTIPOLE FIELDS AND SOURCES

In this section the language and symbols of electrostatics (steady magnetic fields) are used to express the relations between curl-free fields (divergence-free fields) and their point sources. The notation and units are those of Jackson.[7]

A static electric field satisfies

$$\nabla\cdot\mathbf{E} = 4\pi\rho, \quad \nabla\times\mathbf{E} = 0 \tag{36}$$

so that

$$\mathbf{E} = -\nabla\phi, \quad \nabla^2\phi = -4\pi\rho. \tag{37}$$

If

$$\phi(\mathbf{x}) = Y_{lm}(\mathbf{x})/r^{2l+1}, \tag{38}$$

the corresponding source (charge density) is

$$\rho(\mathbf{x}) = \delta_{lm}(\mathbf{x}), \tag{39}$$

and the corresponding electric field is

$$\mathbf{E}(\mathbf{x}) = -\nabla Y_{lm}(\mathbf{x})/r^{2l+1}. \tag{40}$$

A steady magnetic field satisfies

$$\nabla\times\mathbf{B} = (4\pi/c)\mathbf{j}(\mathbf{x}), \quad \nabla\cdot\mathbf{B} = \nabla\cdot\mathbf{j} = 0, \tag{41}$$

so that

$$\mathbf{B} = \nabla\times\mathbf{A}, \quad \nabla\cdot\mathbf{A} = 0, \quad \nabla^2\mathbf{A} = -(4\pi/c)\mathbf{j}. \tag{42}$$

In view of the identity

$$\nabla\times(\mathbf{x}\times\nabla) = \mathbf{x}\nabla^2 - \nabla\left(1 + r\frac{\partial}{\partial r}\right), \tag{43}$$

we have, for $r > 0$, $l > 0$,

$$-\frac{1}{l}\nabla\times(\mathbf{x}\times\nabla)\frac{Y_{lm}(\mathbf{x})}{r^{2l+1}} = -\nabla\frac{Y_{lm}(\mathbf{x})}{r^{2l+1}}. \tag{44}$$

Consequently, the magnetic field defined by the distribution theory formula

$$\mathbf{B}(\mathbf{x}) = -\frac{1}{l}\nabla\times(\mathbf{x}\times\nabla)\frac{Y_{lm}(\mathbf{x})}{r^{2l+1}}, \tag{45}$$

is the same as the electrostatic multipole field (40) for $r > 0$, $l > 0$. The source (current density) of this mag-

netic field is

$$\mathbf{j}(\mathbf{x}) = \frac{c}{4\pi}\nabla\times\mathbf{B}$$

$$= -\frac{c}{4\pi l}\nabla\times[\nabla\times(\mathbf{x}\times\nabla)]\frac{Y_{lm}(\mathbf{x})}{r^{2l+1}}$$

$$= \frac{c}{4\pi l}\nabla^2(\mathbf{x}\times\nabla)\frac{Y_{lm}}{r^{2l+1}} = -\frac{c}{l}(\mathbf{x}\times\nabla)\delta_{lm}(\mathbf{x}). \tag{46}$$

Because $(1+\mathbf{x}\cdot\nabla)Y_{lm}/r^{2l+1} = -lY_{lm}/r^{2l+1}$ even in distribution theory, the difference between (45) and (40) is a "point" distribution

$$\mathbf{B}(\mathbf{x}) - \mathbf{E}(\mathbf{x}) = -\frac{1}{l}\mathbf{x}\nabla^2\frac{Y_{lm}(\mathbf{x})}{r^{2l+1}} = \frac{4\pi}{l}\mathbf{x}\delta_{lm}(\mathbf{x}),$$

which satisfies

$$\nabla\times\left(\frac{4\pi}{l}\mathbf{x}\delta_{lm}\right) = -\frac{4\pi}{l}\mathbf{x}\times\nabla\delta_{lm}$$

$$\nabla\cdot\left(\frac{4\pi}{l}\mathbf{x}\delta_{lm}\right) = -4\pi\delta_{lm}.$$

To prove the second relation, we use the definition (30) of the spherical delta functions:

$$\left(\nabla\cdot\left(\frac{4\pi}{l}\mathbf{x}\delta_{lm}\right), \psi\right)$$

$$= -\frac{4\pi}{l}(\delta_{lm}, \mathbf{x}\cdot\nabla\psi)$$

$$= -\frac{4\pi}{l(2l-1)!!}(\delta(\mathbf{x}), Y_{lm}(\nabla)\mathbf{x}\cdot\nabla\psi)$$

$$= -\frac{4\pi}{l(2l-1)!!}(\delta, \mathbf{x}\cdot\nabla Y_{lm}(\nabla)\psi + lY_{lm}(\nabla)\psi)$$

$$= -4\pi(\delta_{lm}, \psi).$$

The formulas (39), (40) and (45), (46) for general l do not reveal the simplicity of the dipole case. The magnetic equations in particular are unnecessarily complicated. The dipole equations can be expressed more simply as follows.

An electric dipole $\mathbf{p}$ has a charge density $\rho = -\mathbf{p}\cdot\nabla\delta(\mathbf{x})$ and a corresponding electric field

$$\mathbf{E} = \nabla(\mathbf{p}\cdot\nabla)\frac{1}{r}. \tag{47}$$

One can show, as in Ref. 8, that

$$\left(\nabla(\mathbf{p}\cdot\nabla)\frac{1}{r}, \psi\right) = \int d\mathbf{x}\frac{1}{r}\nabla(\mathbf{p}\cdot\nabla)\psi$$

$$= \int d\mathbf{x}\theta(r-\epsilon)\psi\nabla(\mathbf{p}\cdot\nabla)\frac{1}{r} - \frac{4\pi}{3}\mathbf{p}\psi(0).$$

Therefore,

$$\mathbf{E} = \theta(r-\epsilon)\nabla(\mathbf{p}\cdot\nabla)\frac{1}{r} - \frac{4\pi}{3}\mathbf{p}\delta(\mathbf{x}). \tag{48}$$

This is the form in which the dipole field is written by Jackson.[7]

A magnetic dipole $\mathbf{m}$ produces a field $\mathbf{B}$ with the same form as (47) outside the source

$$\mathbf{B} = \nabla\times(\nabla\times\mathbf{m})\frac{1}{r} = \nabla(\mathbf{m}\cdot\nabla)\frac{1}{r} + 4\pi\mathbf{m}\delta(\mathbf{x}), \tag{49}$$

$$\mathbf{B} = \theta(r-\epsilon)\nabla(\mathbf{m}\cdot\nabla)\frac{1}{r} + \frac{8\pi}{3}\mathbf{m}\delta(\mathbf{x}). \tag{50}$$

The corresponding current is

$$\mathbf{j} = \frac{c}{4\pi}\nabla\times\mathbf{B} = c\nabla\times\mathbf{m}\delta(\mathbf{x}), \tag{51}$$

which is the standard form (Refs. 8, 9) for an infinitesimal current loop with magnetic moment $\mathbf{m}$.

To apply equations (45) and (46) to the magnetic dipole case, we must take $l=1$ and replace $-Y_{lm}/r^{2l+1} \to \mathbf{m}\cdot\nabla(1/r)$. Since

$$(\mathbf{m}\cdot\nabla)[\nabla\times(\mathbf{x}\times\nabla)]\frac{1}{r} = 0$$

$$= \nabla\times(\mathbf{m}\times\nabla)\frac{1}{r} + \nabla\times(\mathbf{x}\times\nabla)\mathbf{m}\cdot\nabla\frac{1}{r},$$

we regain (49) for $\mathbf{B}$; and since

$$(\mathbf{m}\cdot\nabla)(\mathbf{x}\times\nabla)\delta(\mathbf{x}) = 0 = \mathbf{m}\times\nabla\delta(\mathbf{x}) + \mathbf{x}\times\nabla(\mathbf{m}\cdot\nabla)\delta,$$

we regain (51) for $\mathbf{j}$.

The point sources $\rho = \delta_{lm}$ and $\mathbf{j} = -(c/l)(\mathbf{x}\times\nabla)\delta_{lm}$ generate fields (40) and (45) that extend beyond their sources. The fields are not point fields; they are nonzero however large r is. But the electrostatic field corresponding to the point source

$$\rho = \nabla^{2K}\delta_{lm}(\mathbf{x}) \quad (K \geq 1) \tag{52}$$

is zero outside an infinitesimal neighborhood of the point $\mathbf{x}=\mathbf{0}$:

$$\phi = -4\pi\nabla^{2K-2}\delta_{lm}(\mathbf{x}) \tag{53}$$

$$\mathbf{E} = 4\pi\nabla\nabla^{2K-2}\delta_{lm}(\mathbf{x}).$$

Similarly, the point current

$$\mathbf{j} = +(\mathbf{x}\times\nabla)\nabla^{2K}\delta_{lm}(\mathbf{x}) \quad (K \geq 1) \tag{54}$$

produces a point magnetic field

$$\mathbf{B} = -\frac{4\pi}{c}\nabla\times(\mathbf{x}\times\nabla)\nabla^{2K-2}\delta_{lm}(\mathbf{x}). \tag{55}$$

For $l=0$, both (54) and (55) vanish because $\mathbf{x}\times\nabla\delta = 0$.

In addition, the independent point currents

$$\mathbf{j}(\mathbf{x}) = \nabla\times(\mathbf{x}\times\nabla)\nabla^{2K}\delta_{lm}(\mathbf{x}) \quad (K \geq 0) \tag{56}$$

generate the point fields

$$\mathbf{B}(\mathbf{x}) = +\frac{4\pi}{c}(\mathbf{x}\times\nabla)\nabla^{2K}\delta_{lm}(\mathbf{x}). \tag{57}$$

The point fields will appear in Secs. 6 and 7 after we have introduced a formal expansion for finite extended sources in a series of spherical delta functions and their derivatives. The cumulative effect of an infinite number of point fields will be a field which is nonzero in an extended region when we interpret the series as a distribution on the space of analytic test functions.

6. ELECTROSTATIC MULTIPOLE EXPANSION

Suppose the electrostatic source $\rho(\mathbf{x})$ is a piecewise differentiable density function that vanishes outside some bounded region. We write (see Ref. 6, p. 160)

$$\rho(\mathbf{x}) = \int d\mathbf{x}'\rho(\mathbf{x}')\delta(\mathbf{x}-\mathbf{x}')$$

$$= \int d\mathbf{x}'\rho(\mathbf{x}')\exp(-\mathbf{x}'\cdot\nabla)\delta(\mathbf{x})$$

$$= \sum_{n=0}^{\infty}(-1)^n\frac{\rho_{i_1\cdots i_n}}{n!}\nabla_{i_1}\cdots\nabla_{i_n}\delta(\mathbf{x}), \tag{58}$$

where

$$\rho_{i_1\cdots i_n} \equiv \int d\mathbf{x}'\rho(\mathbf{x}')x'_{i_1}\cdots x'_{i_n}.$$

Equation (58) appears to be only a formal expansion. However, regarding $\rho(\mathbf{x})$ as a distribution on the space Z of analytic test functions, we have

$$(\rho,\psi) = \int d\mathbf{x}'\rho(\mathbf{x}')\psi(\mathbf{x}').$$

The integral is over the bounded region in which ρ is nonzero. We may expand ψ in its uniformly convergent Taylor series and integrate term by term, getting

$$(\rho,\psi) = \int d\mathbf{x}'\rho(\mathbf{x}')\exp(\mathbf{x}'\cdot\nabla)\psi(\mathbf{0})$$

$$= \sum_{n=0}^{\infty} \frac{\rho_{i_1\cdots i_n}}{n!}\nabla_{i_1}\cdots\nabla_{i_n}\psi(\mathbf{0})$$

$$= \sum_{n=0}^{\infty} \frac{\rho_{i_1\cdots i_n}}{n!}(-1)^n(\nabla_{i_1}\cdots\nabla_{i_n}\delta,\psi).$$

This is the implication of Eq. (58).

Using the expansion (9) to regroup the terms in (58), we get the corresponding decomposition of ρ into spherical delta functions and their derivatives:

$$\rho(\mathbf{x}) = 4\pi\sum_{lm}\int d\mathbf{x}'\rho(\mathbf{x}')Y^*_{lm}(\mathbf{x}')S_l(r'^2\nabla^2)Y_{lm}(-\nabla)\delta(\mathbf{x}). \quad (59)$$

Defining ρ_{lm} by (13) and using the definition (30), we have

$$\rho(\mathbf{x}) = 4\pi\sum_{lm}(2l-1)!!\int r'^{2l+2}dr'\rho_{lm}(r')S_l(r'^2\nabla^2)\delta_{lm}(\mathbf{x}). \quad (60)$$

This expansion of ρ is to be interpreted in the same way that the Cartesian expansion (58) was interpreted. We can make explicit the decomposition of ρ into terms (39) leading to external fields, and terms (52) leading to "point" fields:

$$\rho(\mathbf{x}) = \sum_{lm}\frac{4\pi}{2l+1}\int r'^{2l+2}\rho_{lm}(r')dr'\,\delta_{lm}(\mathbf{x})$$

$$+ 4\pi\sum_{lm}(2l-1)!!\sum_{K=1}^{\infty}\int\frac{r'^{2l+2K+2}}{2^K K!(2l+2K+1)!!}$$

$$\times\rho_{lm}(r')dr'\nabla^{2K}\delta_{lm}(\mathbf{x}).$$

Although highly formal, the decomposition of ρ is not entirely without physical application. Consider a simple case, a uniformly charged shell of radius a and total charge Q. Its density is

$$\rho(\mathbf{x}) = \frac{Q\delta(r-a)}{4\pi a^2} = Q\left[\delta(\mathbf{x}) + \sum_{K=1}^{\infty}\frac{a^{2K}\nabla^{2K}\delta(\mathbf{x})}{2^K K!(2K+1)!!}\right].$$

The electrostatic energy of such a shell in an external field $\phi_{\mathrm{EXT}}(\mathbf{x})$ is therefore

$$\int d\mathbf{x}\,\rho(\mathbf{x})\phi_{\mathrm{EXT}}(\mathbf{x}) = Q\phi_{\mathrm{EXT}}(\mathbf{0}) + \frac{Qa^2}{6}\nabla^2\phi_{\mathrm{EXT}}(\mathbf{0}) + \cdots.$$

The complete expansion is an example of the Pizetti formula (15). The "correction" to the point contribution has the form of the Darwin term in quantum electrodynamics (see, for example, Ref. 10).

Returning to the general case, we find, by comparing (29) and (37), the electrostatic potential ϕ to which ρ gives rise:

$$\phi(\mathbf{x}) = 4\pi\sum_{lm}(2l-1)!!\int[r'^{2l+2}dr'$$

$$\times\rho_{lm}(r')S_l(r'^2\nabla^2)Y_{lm}(\mathbf{x})/r^{2l+1}]. \quad (61)$$

The corresponding electric field is

$$\mathbf{E}(\mathbf{x}) = -4\pi\sum_{lm}(2l-1)!!\int[r'^{2l+2}dr'$$

$$\times\rho_{lm}(r')S_l(r'^2\nabla^2)\nabla Y_{lm}(\mathbf{x})/r^{2l+1}]. \quad (62)$$

Outside the source we may treat ϕ as an ordinary function rather than a distribution. In this region $\nabla^2[Y_{lm}/r^{2l+1}] = 0$, and Eq. (61) reduces to the usual multipole expansion

$$\phi(\mathbf{x}) = \sum_{lm}\frac{4\pi}{2l+1}\left[\int dr'\,\rho_{lm}(r')r'^{2l+2}\right]\frac{Y_{lm}(\mathbf{x})}{r^{2l+1}}. \quad (63)$$

In addition to the "external" terms (63), Eq. (61) contains, for each l, an infinite series of point fields (53) whose cumulative effect is an extended field. We make this explicit by using the series (10) for S_l together with the Poisson equations (29):

$$\phi(\mathbf{x}) = \sum_{lm}\frac{4\pi}{2l+1}\int r'^{2l+2}dr'\rho_{lm}(r')\frac{Y_{lm}(\mathbf{x})}{r^{2l+1}}$$

$$- (4\pi)^2\sum_{lm}\int r'^{2l+2}dr'\rho_{lm}(r')$$

$$\times\sum_{K=1}^{\infty}\frac{r'^{2K}\nabla^{2K-2}\delta_{lm}(\mathbf{x})}{K!\,2^K(2l+2K+1)!!/(2l-1)!!}. \quad (64)$$

This equation is to be understood as the formal expression of a distribution whose value on a decreasing analytic test function ψ is

$$(\phi,\psi) = \sum_{lm}\frac{4\pi}{2l+1}\int r'^{2l+2}\rho_{lm}(r')dr'\int r\psi^*_{lm}(r)dr$$

$$- (4\pi)^2\sum_{lm}\int r'^{2l+2}\rho_{lm}(r')dr'$$

$$\times\sum_{K=1}^{\infty}\frac{r'^{2K}\nabla^{2K-2}Y_{lm}(\nabla)\psi(\mathbf{0})}{K!\,2^K(2l+2K+1)!!}. \quad (65)$$

We have used (25) and (32) to obtain (65).

To confirm (65), we may compare it with the value of (ϕ,ψ), where ϕ is the integral solution of Poisson's equation:

$$\phi(\mathbf{x}) = \int d\mathbf{x}'\,\rho(\mathbf{x}')/|\mathbf{x}-\mathbf{x}'|. \quad (66)$$

Inserting $1 = \theta(r-r') + \theta(r'-r)$ to split up the range of integration and using (17) in the two appropriate forms, we have

$$\phi(\mathbf{x}) = \sum_{lm}\frac{4\pi}{2l+1}Y_{lm}(\mathbf{x})\int r'^2dr'$$

$$\times\left[\frac{\theta(r-r')r'^{2l}}{r^{2l+1}} + \frac{\theta(r'-r)}{r'}\right]\rho_{lm}(r').$$

Applying ϕ to a decreasing analytic test function ψ, we get

$$(\phi,\psi)=\sum_{lm}\frac{4\pi}{2l+1}\int r^2dr\,r'^2dr'\psi^*_{lm}(r)\rho_{lm}(r')$$
$$\times\left[\frac{r'^{2l}}{r}\theta(r-r')+\frac{r^{2l}}{r'}\theta(r'-r)\right].$$

The function $\theta(r-r')$ may now be replaced by $1-\theta(r'-r)$ to obtain

$$(\phi,\psi)=\sum_{lm}\frac{4\pi}{2l+1}\int rdr\psi^*_{lm}(r)\int r'^{2l+2}dr'\rho_{lm}(r')$$
$$+4\pi\sum_{lm}\int r\,dr\,\psi^*_{lm}(r)\frac{\theta(r'-r)}{2l+1}r'dr'$$
$$\times\rho_{lm}(r')[r^{2l+1}-r'^{2l+1}]$$

By introducing the explicit form (14) for ψ^*_{lm} in the second term and doing the r integral over its finite range, we get agreement with (65).

7. MAGNETIC MULTIPOLE EXPANSION

We suppose that a current $\mathbf{j}(\mathbf{x})$ is differentiable and of limited spatial extent. Because it is conserved it may be decomposed into a toroidal part and a poloidal part, determined by scalar fields $C(\mathbf{x})$ and $D(\mathbf{x})$ respectively:

$$\mathbf{j}(\mathbf{x})=\mathbf{x}\times\nabla C(\mathbf{x})+\nabla\times(\mathbf{x}\times\nabla)D(\mathbf{x}). \tag{67}$$

An illuminating proof of this formula is given by Moses in Ref. 11, where further references may be found.

The scalar functions C and D may be expanded in spherical delta functions in exactly the same way that the electrostatic source ρ was expanded in (60):

$$C(\mathbf{x})=4\pi\sum_{lm}(2l-1)!!\int r'^{2l+2}dr'\,C_{lm}(r')S_l(r'^2\nabla^2)\delta_{lm}(\mathbf{x}) \tag{68}$$

$$D(\mathbf{x})=4\pi\sum_{lm}(2l-1)!!\int r'^{2l+2}dr'\,D_{lm}(r')S_l(r'^2\nabla^2)\delta_{lm}(\mathbf{x}) \tag{69}$$

Because ∇^2 commutes with $\mathbf{x}\times\nabla$, we may now write the current

$$\mathbf{j}(\mathbf{x})=4\pi\sum_{lm}(2l-1)!!\int r'^{2l+2}dr'$$
$$\times C_{lm}(r')S_l(r'^2\nabla^2)[\mathbf{x}\times\nabla\delta_{lm}(\mathbf{x})]$$
$$+4\pi\sum_{lm}(2l-1)!!\int r'^{2l+2}dr'$$
$$\times D_{lm}(r')S_l(r'^2\nabla^2)[\nabla\times(\mathbf{x}\times\nabla)\delta_{lm}(\mathbf{x})]. \tag{70}$$

The $l=0$ terms in the two series are zero ($\mathbf{x}\times\nabla\delta=0$).

The magnetic field to which $\mathbf{j}$ gives rise is identified term by term according to the relation in Sec. 5. Outside the source only the $K=0$ terms in relations the toroidal part of the current contribute:

$$\mathbf{B}=\frac{4\pi}{c}\sum_{lm}\frac{l}{2l+1}\int dr'\,r'^{2l+2}C_{lm}(r')\nabla\frac{Y_{lm}(\mathbf{x})}{r^{2l+1}}. \tag{71}$$

Moses[11] shows that the toroidal part of $\mathbf{j}$ may be written

$$(\mathbf{x}\times\nabla)C(\mathbf{x})=\sum_{lm}f_{llm}(r)\frac{(-i)\mathbf{x}\times\nabla}{[l(l+1)]^{1/2}}Y_{lm}(\hat{\mathbf{x}}),$$

where

$$f_{llm}(r)=\int d\Omega\frac{[i(\mathbf{x}\times\nabla)Y^*_{lm}(\hat{\mathbf{x}})]\cdot\mathbf{j}(\mathbf{x})}{[l(l+1)]^{1/2}}$$

and the $l=0$ term is absent. In our notation, this gives

$$C_{lm}(r)=r^{-2l}\int d\Omega\frac{[(\mathbf{x}\times\nabla)Y^*_{lm}(\mathbf{x})]\cdot\mathbf{j}(\mathbf{x})}{l(l+1)}$$
$$=\frac{r^{-2l}}{l(l+1)}\int d\Omega\,Y^*_{lm}(\mathbf{x})\nabla\cdot(\mathbf{x}\times\mathbf{j}(\mathbf{x})). \tag{72}$$

Substituting (72) into (71), we get a representation of the external magnetic field in terms of a scalar potential.[12]

APPENDIX A

Expanding the trinomial in Eq. (2), we get an explicit form for $W_{lm}(\mathbf{x})$:

$$W_{lm}(\mathbf{x})=\sum_{c-a=m}\frac{l!}{a!b!c!}(x_1-ix_2)^a(2x_3)^b(-1)^c(x_1+ix_2)^c, \tag{A1}$$

in which the sum is over the nonnegative integers a, b, c such that $a+b+c=l$ and $c-a=m$.

If the operator $\mathbf{L}=-i\mathbf{x}\times\nabla$ is expressed in terms of the variables $\xi=x_1+ix_2$, $\eta=x_1-ix_2$, $z=x_3$, then its components are

$$L_z=L_3=\xi\frac{\partial}{\partial\xi}-\eta\frac{\partial}{\partial\eta},$$
$$L_+=L_1+iL_2=-\xi\frac{\partial}{\partial z}+2z\frac{\partial}{\partial\eta}, \tag{A2}$$
$$L_-=L_1-iL_2=+\eta\frac{\partial}{\partial z}-2z\frac{\partial}{\partial\xi}.$$

One can then show that the W_{lm} of (A1) satisfy

$$L_zW_{lm}=mW_{lm},$$
$$L_+W_{lm}=(l+m+1)W_{lm+1}\quad(m<l), \tag{A3}$$
$$L_-W_{lm}=(l-m+1)W_{lm-1}\quad(m>-l),$$
$$L_+W_{ll}=0=L_-W_{l-l}.$$

Using (A3) and the definition (1) for $Y_{lm}(\mathbf{x})$ in terms of $W_{lm}(\mathbf{x})$, we get

$$L_zY_{lm}=mY_{lm},$$
$$L_\pm Y_{lm}=[(l\mp m)(l\pm m+1)]^{1/2}Y_{lm\pm1}, \tag{A4}$$

which show that the harmonics with fixed l are related in phase by the Condon and Shortley convention.[2,7] For the specific case $m=-l$, (1) and (A1) give

$$Y_{l-l}(\mathbf{x})=\frac{1}{2^ll!}\left(\frac{2l+1}{4\pi}\right)^{1/2}[(2l)!]^{1/2}(\sin\theta)^lr^l\exp(-il\phi),$$

which is in agreement with equation (2.5.5) of Ref. 2.

APPENDIX B

From the calculus of special functions we get the physically important Rayleigh formula (Ref. 2, p. 81)

$$\exp(i\mathbf{x}\cdot\mathbf{y})=\sum_{l=0}^{\infty}i^l(2l+1)j_l(xy)P_l(\hat{\mathbf{x}}\cdot\hat{\mathbf{y}})$$

$$= \sum_{lm} 4\pi i^l \frac{j_l(xy)}{(xy)^l} Y_{lm}(\mathbf{x}) Y^*_{lm}(\mathbf{y}), \tag{B1}$$

where $x = |\mathbf{x}|$, $y = |\mathbf{y}|$, and j_l is the spherical Bessel function

$$j_l(x) = \left(\frac{\pi}{2x}\right)^{1/2} J_{l+1/2}(x) = x^l \sum_{K=0}^{\infty} \frac{(-1)^K x^{2K}}{K!\, 2^K (2l+2K+1)!!}. \tag{B2}$$

We are interested in finding a formal series expansion for $\exp(\mathbf{x}\cdot\mathbf{y})$. Replacing $\mathbf{x}$ by $-i\mathbf{x}$ in (B1) (which entails $x \to \pm ix$) and using the fact that $j_l(xy)/(xy)^l$ is an even power series in xy, we get

$$\exp(\mathbf{x}\cdot\mathbf{y}) = \sum_{lm} 4\pi [j_l(ixy)/(ixy)^l] Y_{lm}(\mathbf{x}) Y^*_{lm}(\mathbf{y}). \tag{B3}$$

The function in square brackets is related to the modified Bessel function $I_\nu(x) = (-i)^\nu J_\nu(ix)$. In fact,

$$\frac{[(\pi/2x)]^{1/2} I_{l+1/2}(x)}{x^l} = \sum_{K=0}^{\infty} \frac{x^{2K}}{K!\, 2^K (2l+2K+1)!!}$$

$$= \frac{j_l(ix)}{(ix)^l} = S_l(x^2), \tag{B4}$$

in which the last equality follows from definition (10) of the text. We now write (B3)

$$\exp(\mathbf{x}\cdot\mathbf{y}) = \sum_{lm} 4\pi S_l(x^2 y^2) Y_{lm}(\mathbf{x}) Y^*_{lm}(\mathbf{y}), \tag{B5}$$

which is the form quoted in Eq. (9) of Sec. 2.

From its definition,

$$S_l(x) = \sum_{K=0}^{\infty} \frac{x^K}{2^K K!\,(2l+2K+1)!!}, \tag{10}$$

the following properties of S_l can be deduced directly:

$$S_0(x) = \sum_{K=0}^{\infty} \frac{x^K}{(2K+1)!}, \tag{B6}$$

$$S_l(x) = 2^l \frac{d^l}{dx^l} S_0(x) = 2^l S_0^{(l)}(x), \tag{B7}$$

$$S_{l-1}(x) - (2l+1) S_l(x) = 2x S_l'(x) = x S_{l+1}(x), \tag{B8}$$

$$x \frac{d}{dx} S_l(x^2 a^2) = S_{l-1}(x^2 a^2) - (2l+1) S_l(x^2 a^2), \tag{B9}$$

$$\left(\frac{1}{x}\frac{d}{dx}\right) [x^{2l+1} S_l(x^2 a^2)] = x^{2l-1} S_{l-1}(x^2 a^2), \tag{B10}$$

$$\left(\frac{1}{x}\frac{d}{dx}\right)^l [x^{2l+1} S_l(x^2 a^2)] = x S_0(x^2 a^2). \tag{B11}$$

In (B9), (B10), and (B11), a^2 is a constant.

[1] *Higher Transcendental Functions, Vol. II*, edited by A. Erdélyi (McGraw-Hill, New York, 1953).
[2] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton U. P., Princeton, N. J., 1957).
[3] E. W. Hobson, *The Theory of Spherical and Ellipsoidal Harmonics* (Cambridge U. P., Cambridge, 1931).
[4] R. Courant and D. Hilbert, *Methods of Mathematical Physics, Vol. II* (Interscience, New York, 1962), p. 287.
[5] C. Müller, *Spherical Harmonics*, Lecture Notes in Mathematics **17**, (Springer-Verlag, Berlin, 1966), pp. 34—36.
[6] I. M. Gel'fand and G. E. Shilov, *Generalized Functions, Vol. I* (Academic, New York, 1964).
[7] J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975), 2nd ed.
[8] R. Gagnon, Am. J. Physics **38**, 879 (1970).
[9] R. H. Good and T. J. Nelson, *Classical Theory of Electric and Magnetic Fields* (Academic, New York, 1971), p. 226.
[10] L. L. Foldy and S. A. Wouthuysen, Phys. Rev. **78**, 29 (1950).
[11] H. E. Moses, J. Math. Phys. **17**, 1821 (1976).
[12] J. B. Bronzan, Am. J. Physics **39**, 1357 (1971).

Lower energy bounds for translation-invariant atomlike systems

Richard L. Hall

Department of Mathematics, Concordia University, Montreal, Quebec, Canada
(Received 16 March 1978)

A system of N identical particles with masses m_1 interact with each other via a pair potential V and with an additional particle of mass m_0 by the pair potential U, and the system obeys nonrelativistic quantum mechanics. A general *lower* bound formula is derived for the ground-state energy, which reduces to Coleman's result in the special case $m_0 = \infty$. The lower bound is compared with the exact energy for the harmonic oscillator problem $V(x) = k_1^2 x^2$ and $U(x) = k_0^2 x^2$, which has recently been solved exactly (Hall).

1. INTRODUCTION

We consider a system of $(N+1)$ particles labeled $(0, 1, 2, \ldots, N)$ consisting of a particle with mass m_0 and N identical particles with masses m_1. The identical particles interact with each other via a pair potential V and they interact with the remaining particle (labeled 0) by the pair potential U. Assuming that the system obeys nonrelativistic quantum mechanics we may write the translation-invariant Hamiltonian H as follows,

$$H = \left\{ \frac{\mathbf{p}_0^2}{2m_0} + \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_1} - \frac{(\mathbf{p}_0 + \sum_{i=1}^{n} \mathbf{p}_i)^2}{2m} + \sum_{i=1}^{N} U_{0i} + \sum_{1 \le i < j}^{N} V_{ij} \right\}, \tag{1.1}$$

where the total mass $m = m_0 + Nm_1$.

If U and V are both Hooke's law pair potentials, we have shown (Hall[1,2]) that the Hamiltonian H separates in suitable relative coordinates and consequently the "harmonic atom" is an exactly soluble problem. In the present paper we determine a *lower bound* E_1^L to the ground-state energy E_1 of H for general U and V. The bound $E_1^L \le E_1$ is expressed in terms of the exact eigenvalues of a suitable "reduced" three-particle Hamiltonian $\mathcal{H}$ which is related to H by the equation

$$\langle H \rangle = \langle \mathcal{H} \rangle. \tag{1.2}$$

The Hamiltonian $\mathcal{H}$ is chosen so that Eq. (1.2) is valid for any normalized translation-invariant $(N+1)$-particle wavefunction that is either symmetric or antisymmetric in the N identical-particle indices $(1, 2, \ldots, N)$.

There have been two independent lines of investigation leading to energy *lower* bounds via equations like (1.2). The first was the "equivalent two-body method" initiated by Wigner[3] in 1933 for the triton problem ($m_0 = \infty$, $U = 0$). A rigorous argument of Post[4] can be used to show that in many instances Wigner's method leads to lower energy bounds (for a very brief history see Hall *et al.*[5]). The second approach was introduced by Bopp,[6] using density matrices, for atomic systems (with $m_0 = \infty$, $U \neq 0$). Rigorous results were obtained in this case by Coleman *et al.*[7] Whereas in nuclear physics translation invariance is essential, in atomic physics it is an acceptable first approximation to assume an infinitely massive nucleus ($m_0 = \infty$). This difference has made the first line of investigation, the translation-invariant case, more complicated because the symmetry requirements of the Pauli principle must be expressed in terms of functions of the *relative* coordinates. It turns out that the lower bound in this case is sensitive to the choice of relative coordinates and we have had to optimize with respect to the class of coordinates which allow its derivation (Hall[8]).

In this article we treat the type of system considered by Bopp and Coleman but we assume m_0 is *finite* and we keep translation invariance throughout. Our main result (the theorem in Sec. 2) is a generalization of Coleman's results (Ref. 7) and reduces to the latter when $m_0 \to \infty$. We could also treat, by the same arguments, systems consisting of, say, S *groups* of identical particles interacting by pair potentials, as we have done (Hall[9]) *exactly* for the harmonic oscillator. However, we present here the simplest case in which two types of particle are present. Also, although we consider only *scalar* particles, the introduction of spin and isotopic spin is straightforward (cf. Hall *et al.*[5,8,10]).

Calogero *et al.*[11] have considered the case $U \neq 0$, $m_0 = \infty$, and they have constructed a large variety of reduced Hamiltonians $\mathcal{H}$ for this problem. For example, their Theorem 1 Corollary 1.1 yields a reduced Hamiltonian which is equivalent to Coleman's (Ref. 7), and, of course, to ours for $m_0 = \infty$. However, the problem of incorporating the maximum amount of permutation-symmetry restriction into the lower bound needs to be solved for each choice of $\mathcal{H}$. Furthermore, in our problem in which, m_0 is finite, the reduced Hamiltonian $\mathcal{H}$ only has the *form* of a three-body Hamiltonian in relative coordinates: In individual-particle coordinates $\mathcal{H}$ is an $(N+1)$-particle operator; this will be the case whenever the appropriate relative coordinates are not orthogonal (cf. Hall[8]). Other contributions and developments in this field have been made by Carr *et al.*,[12] Savchenko,[13] Balbutsev *et al.*,[14] and Weidemann[15] who applied density matrix methods to the few-nucleon problem.

2. THE MAIN RESULT

We define new coordinates $\rho = B\mathbf{r}$ and momenta

 0022-2488/78/1909-1969$1.00

$\pi = (B^T)^{-1}\mathbf{p}$, $\pi_j = -i\hbar\nabla_{\rho j}$, by

$$B = \begin{bmatrix} \frac{m_0}{m} & \frac{m_1}{m} & \frac{m_1}{m} & \cdots & \frac{m_1}{m} \\ -1 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & 0 \cdots & 0 \\ -1 & 0 & 0 & 1\ 0 \cdots & 0 \end{bmatrix} \quad \text{etc.} \tag{2.1}$$

Thus ρ_0 is the center-of-mass coordinate and $\rho_i = (\mathbf{r}_i - \mathbf{r}_0)$, $i = 1, 2, \ldots, N$. It is necessary for our argument (cf. Hall[8]) that (23) $\rho_2 = \rho_3$, where (23) permutes the *individual-particle* indices 2 and 3; Jacobi orthogonal relative coordinates are therefore *not* admissible. Apart from this constraint the choice has been made so that our theorem will reduce to Coleman's result[7] for $m_0 \to \infty$, $\mathbf{r}_0 \to 0$.

If we denote by K the *total* kinetic energy operator and by M the diagonal matrix of the masses, then we have $2K = \mathbf{p}^T M^{-1}\mathbf{p} = \pi^T B M^{-1} B^T \pi$. It follows from (2.1) that

$$\langle K\rangle = \tfrac{1}{2}\left\langle \left[\frac{\pi_0^2}{m} + N\left(\frac{1}{m_0} + \frac{1}{m_1}\right)\pi_1^2 + N(N-1)\frac{\pi_1\cdot\pi_2}{m_0}\right]\right\rangle , \tag{2.2}$$

where expectations are taken with respect to normalized translation-invariant $(N+1)$-particle functions which are symmetric or antisymmetric in the individual-particle indices $(1, 2, \ldots, N)$. Hence

$$\langle H\rangle = \langle \mathcal{H}\rangle, \tag{2.3}$$

where

$$\mathcal{H} = \frac{N}{2}\left[\frac{1}{2}\left(\frac{1}{m_1} + \frac{1}{m_0}\right)\left(\pi_1^2 + \pi_2^2\right) + \frac{N-1}{m_0}\pi_1\cdot\pi_2 + U_{01} + U_{02} + (N-1)V_{12}\right]. \tag{2.4}$$

The Hamiltonian $\mathcal{H}$ is $N/2$ times the translation-invariant Hamiltonian for a three-body problem [i.e., (2.4) with $N = 2$] consisting of two identical particles with masses m_1' and a third particle with mass m_0', where

$$m_1' = \left(\frac{1}{m_1} - \frac{N-2}{m_0}\right)^{-1}$$

and (2.5)

$$m_0' = \frac{m_0}{(N-1)}.$$

This interpretation fails if m_0 is too small, for it is necessary to require

$$m_0 > (N-2)m_1. \tag{2.6}$$

We suppose that $\mathcal{H}$ has n bound states $[\phi_i(\rho_1, \rho_2)]$ with a corresponding nondecreasing sequence of eigenvalues $\mathcal{E}_1 \leq \mathcal{E}_2 \leq \cdots \leq \mathcal{E}_n$. If the N identical particles are fermions, then we *only* consider the eigenstates of $\mathcal{H}$ which are antisymmetric in (ρ_1, ρ_2); similarly for bosons we restrict to symmetric states. With this notation and assumptions we prove the following

Theorem:

Bosons or fermions: $E_1 \geq \mathcal{E}_1 = E_1^L$. (2.7)

Fermions:

all N: $E_1 \geq \frac{1}{N-1}\sum_{i=1}^{N-1}\mathcal{E}_i = E_1^L$, (2.8a)

N odd: $E_1 \geq \frac{1}{N}\sum_{i=1}^{N}\mathcal{E}_i = E_1^L$, (2.8b)

where, if the number of bound (fermion) states of $\mathcal{H}$ is $n-$, the sums in (2.8) run up to $i = \min(n-, N-1)$ or $i = \min(n-, N)$, respectively.

Proof of the Theorem: Suppose the normalized ground state of H is the function $\psi(\rho_1, \rho_2, \ldots \rho_N)$ which is symmetric or antisymmetric in the variables $(\rho_1, \rho_2, \ldots, \rho_N)$ because it has this symmetry in $(\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_N)$. Then we can analyze ψ in terms of the normalized eigenfunctions $[\phi_i(\rho_1, \rho_2)]$ of $\mathcal{H}$ and we have

$$\psi(\rho_1, \rho_2, \ldots, \rho_N) = \sum_i C_i\phi_i(\rho_1, \rho_2)\psi_i(\rho_3, \rho_4, \ldots, \rho_N), \tag{2.9}$$

where $\|\phi_i\| = \|\psi_i\| = 1$, and $\|\psi\|^2 = \sum_i |C_i|^2 = 1$; and the summation is understood to include (if necessary) integration over the unbound states of $\mathcal{H}$. We now apply (2.9) to Eq. (2.3) to give

$$E_1 = \sum_i |C_i|^2\mathcal{E}_i. \tag{2.10}$$

If we put the entire "weight" in the first coefficient (corresponding to the lowest eigenstate of $\mathcal{H}$ with the appropriate symmetry), then (2.10) immediately yields (2.7).

In the case of fermions specifically, we apply an argument exactly analogous to that of Hall[8] and, for the bound-state coefficients, we obtain

$$|C_i|^2 \leq (N-1)^{-1}. \tag{2.11}$$

The inequality (2.11) yields the first bound (a) of (2.8). The more stringent bound (b) for N odd is obtained by Coleman's argument[7] which is applicable since the many-body state ψ is normalized. This establishes the theorem.

For $N = 2$ (the three-body problem) of course $E_1 = \mathcal{E}_1$. For $m_0 \to \infty$ and $\mathbf{r}_0 \to 0$, we have $\rho_i = \mathbf{r}_i$ and the reduced Hamiltonian $\mathcal{H} = NK/2$, where K is the two-body operator in Coleman's theory.[7] Thus (2.8) is exactly Coleman's result if $m_0 = \infty$.

The bound (2.8a) appears formally the same as the bound we obtained[8] for a system of N identical fermions ($U = 0$, $m_0 = \infty$). However there is an important difference which makes it difficult to derive our earlier result from (2.8a): In the earlier bound *all* the eigenstates of the reduced Hamiltonian must be counted, *not* just the fermion states. Suppose for example that in the earlier bound (the reader will need to consult Hall[8] in order to follow this argument) we use a suboptimal set of relative coordinates $\boldsymbol{\rho}_i = (\mathbf{r}_1 - \mathbf{r}_i)/\sqrt{2}$, $i \geq 2$, for which $\lambda = 2(N-1)/N$, then we get for the harmonic oscillator in one dimension $E_1^L/E_1 = N^{1/2}(N-1)^{1/2}/(N+1)\sqrt{2}$; however, the conjecture that only the fermion two-body states need to be included yields $E_1^L/E_1 = \sqrt{2}(N-\frac{1}{2})N^{1/2}/(N+1)(N-1)^{1/2}$ which is clearly false. The reason for this lies in the expansion (2.9) above: The factors ϕ_i and ψ_i are *necessarily* antisymmetric in $(1, 2)$ and $(3, 4, \ldots, N)$ respectively; in the correspond-

ing expansion of Hall[8] the functions ϕ_i [Eq. (3) in that reference] need not be antisymmetric in (1, 2) because this permutation alters the other relative coordinates. By choosing new coordinates in (2.9) and taking the special case $m_0 \to \infty$, $U \to 0$, it is possible to derive (2.11) and therefore (2.8a) appropriately, but it is much more convenient to treat the cases $U = 0$ and $U \neq 0$ separately.

3. THE HARMONIC OSCILLATOR

Since we have solved (Hall[1,2,9]) the translation-invariant harmonic atom (and harmonic matter) problem exactly for all N, we can use this system to compare E_1^L with E_1. It has become a tradition in the translation-invariant work to test results on exactly soluble oscillator problems. In fact we have proved (Hall[16,17]) that for N bosons with $U = 0$, the bound[5] corresponding to (2.7) yields the *exact* energy E_1 if and only if the pair-potential V has the form $V(\mathbf{r}) = k\mathbf{r}^2$ (orthogonal relative coordinates, $\lambda = 1$, are required for this result; see also Hall *et al.*[5,8]). For $U \neq 0$ we note that whenever we choose $m_0 = \infty$, the harmonic-oscillator example provides a test for Coleman's bound,[7] a test which has not been discussed before.

We consider scalar particles in one dimension and we suppose that U and V are given by

$$U(x) = k_0 x^2, \quad V(x) = k_1 x^2. \tag{3.1}$$

The ground-state energy of H is given[1,2,9] in this case by

$$E_1 = \frac{\hbar k_0}{\sqrt{2}} \left(\frac{1}{m_1} + \frac{N}{m_0} \right)^{1/2} + (N^q - 1)\hbar \left(\frac{N k_1^2 + k_0^2}{2m_1} \right)^{1/2}, \tag{3.2}$$

where $q = 1$ for bosons and $q = 2$ for fermions. Likewise the exact eigenvalues of H (a translation-invariant three-body Hamiltonian) are given by

$$\mathcal{E}(n, m) = \frac{N\hbar}{2\sqrt{2}} \left[k_0 \left(\frac{1}{m_1} + \frac{N}{m_0} \right)^{1/2} (1 + 2n) + \left[2(N-1)k_1^2 + k_0^2 \right]^{1/2} \left(\frac{1}{m_1} - \frac{N-2}{m_0} \right)^{1/2} \times (1 + 2m) \right], \tag{3.3}$$

where n and m are any positive or zero integers; for fermion states, m must be odd. With the exact results (3.2) and (3.3) the theorem proved in Sec. 2 can be tested in detail. We shall limit the discussion to a few illustrations.

A. The bound $E_1 \geq \mathcal{E}_1 = E_1^L$ applied to bosons

For the bound (2.7) we have $\mathcal{E}_1 = \mathcal{E}(0, 0) = E_1^L$. Suppose $k_1 = 0$ and $m_0 = (N-2)m_1$, the infimum of allowed values, then E_1^L/E_1 decreases from $\sqrt{2}/(1+\sqrt{2}) \approx 0.59$, at $N = 3$, to $\frac{1}{2}$ at $N = \infty$; for $m_0 = \infty$, $E_1^L/E_1 = 1$ for all N. The poorer results for m_0 finite may be related to our use of suboptimal[8] relative coordinates for the lower bound method; this difficulty, which persists in the fermion case, has not been overcome as yet. For $m_0 = \infty$ and $k_0 = 0$, we get $E_1^L/E_1 = [N/2(N-1)]^{1/2}$ in contrast to our earlier Method I[5] (for N identical particles) which yields $E_1^L/E_1 = 1$.

B. The bound $E_1 \geq (N-1)^{-1} \sum_{i=1}^{N-1} \mathcal{E}_i = E_1^L$ for fermions

Since the ordering of the eigenvalues $\mathcal{E}(n, m)$ of H depends on the masses, the coupling constants, and on N, we cannot give an explicit general recipe for E_1^L; we do have to remember that m is odd. If m_0 is sufficiently close to the extreme value $(N-2)m_1$, then the first $(N-1)$ energies are given by $\mathcal{E}_i = \mathcal{E}(0, i)$ for $i = 1, 3, 5, \ldots, (2N-3)$; in this extreme case the ratio E_1^L/E_1 is given (for $N > 2$) by

$$\frac{E_1^L}{E_1} = \tfrac{1}{2} N k_0 \{ k_0 + (N+1)[(N-2)(N-1)/2]^{1/2} \times (N k_1^2 + k_0^2)^{1/2} \}^{-1}. \tag{3.4}$$

Suppose, continuing this illustration, we set $k_1 = 0$ and $N = 3$, then from (3.4) we have $E_1^L/E_1 = \frac{3}{10}$; other examples of this extreme case will give worse results unless k_1^2 is negative[2,9] (repulsion between like particles). However, as m_0 increases the quality of the bound improves.

We now suppose $m_0 = \infty$ and also $k_1 = 0$ (no interaction between the "harmonic electrons") and we find

$$\frac{E_1^L}{E_1} = [N(N-1)]^{-1} \sum_{i=1}^{N-1} [1 + (n+m)_i], \tag{3.5}$$

where the sum is over the first $(N-1)$ terms, with m odd. By choosing special values of N so that the sum in (3.5) can be calculated exactly we find

$$N = (\nu+1)^2 + 1, \quad \nu = 0, 1, 2, \ldots, \qquad \frac{E_1^L}{E_1} = \frac{8\nu^2 + 19\nu + 12}{6(\nu+1)[(\nu+1)^2+1]}. \tag{3.6}$$

A few numerical values are shown in Table I.

By contrast, our earlier bound[8] for the case $U = 0$ $(k_0 = 0)$ yields $E_1^L/E_1 = N^{1/2}(N-1)^{1/2}/(N+1)\sqrt{2}$ with pair-distance coordinates $[\lambda = 2(N-1)/N]$, and $E_1^L/E_1 = \sqrt{3}(N-1)/2(N+1)$ with optimized coordinates $(\lambda = \frac{4}{3})$. This comparison suggests that it should be possible to improve the more general bound (2.8).

4. CONCLUSION

We have established a general lower bound on the ground-state energy of translation-invariant atomlike many-particle systems. In the special case $m_0 = \infty$ the bound yields Coleman's earlier result.[7] Various generalizations of the theorem in Sec. 2 are possible. For example, we can prove a similar result for a system composed by S *groups* of identical particles; we already have a test case for this type of theorem because we have solved[9] such problems *exactly* for Hooke's law potentials. Furthermore, by constructing a model for the lower bound, we can use the argument

TABLE I. The harmonic atom with noninteracting "electrons" $(k_1 = 0)$: The ratio E_1^L/E_1 of the lower bound to the exact energy is asymptotically like $N^{-1/2}$.

N	2	5	10	17	26	37	∞
$\frac{E_1^L}{E_1}$	1	0.65	0.46	0.35	0.28	0.23	0

of Hall[18] to invert the Rayleigh—Ritz principle and yield lower bounds on the entire bound-state spectrum.

We feel that such generalizations of our theorem would be premature at the present time. For bosons, the bound (2.7) does sometimes yield exact solutions, as we found in Sec. 3; but the potentially more interesting fermion bound (2.8) may be intrinsically poor for large N. In the case of small atoms the bound (2.8) appears to be quite good[7] (if we can accept arguments involving experimental data). However, the applications to the harmonic oscillator which we discussed in Sec. 3 demonstrate very clearly that not enough of the constraint of the Pauli principle has been kept in this lower bound.

The interesting question which arises here is whether it is possible to construct a new reduced Hamiltonian $\mathcal{H}$ which would allow small values of m_0. We could then allow $m_0 = m_1$ and $U = V$ and if the new bound would reduce *mutatis mutandis* to our earlier result[8] for $(N+1)$ identical fermions, some of the improvement might be retained for $U \neq V$ and m_0 large. It would also of course be theoretically much more satisfactory if both types of system could be treated comfortably within the same general formulation. The present paper which discusses atomlike systems in the center-of-mass frame is a step in this direction.

[1]R. L. Hall, Phys. Lett. A **65**, 85-6 (1978).
[2]R. L. Hall, "Some exact solutions to the translation-invariant N-body problem," J. Phys. A Math. Gen. (to be published).
[3]E. Wigner, Phys. Rev. **43**, 252—7 (1933).
[4]H. R. Post, Proc. Phys. Soc. A **66**, 649—52 (1953).
[5]R. L. Hall and H. R. Post, Proc. Phys. Soc. A **90**, 381—96 (1967).
[6]F. Bopp, Z. Phys. **156**, 348—59 (1959).
[7]A. J. Coleman, Rev. Mod. Phys. **35**, 668—89 (1963).
[8]R. L. Hall, Proc. Phys. Soc. A **91**, 16—22 (1967).
[9]R. L. Hall, "Exact solutions of Schrödinger's equation for translation-invariant harmonic matter," J. Phys. A Math. Gen. (to be published).
[10]R. L. Hall, J. Phys. A Math. Nucl. Gen. **1**, 468—74 (1968).
[11]F. Calogero and C. Marchioro, J. Math. Phys. **10**, 562—9 (1969).
[12]R. J. M. Carr and H. R. Post, J. Phys. A Math. Gen. **10**, L59—61 (1977).
[13]O. Ya. Savchenko, Sov. J. Nucl. Phys. **21**, 381—3 (1975).
[14]E. B. Balbutsev, B. F. Irgasiev, L. Majling, I. N. Mikhailov, and J. Řizek, J. Phys. G. Nucl. Phys. **2**, 681—8 (1976).
[15]E. G. Weidemann, Nucl. Phys. **65**, 559—76 (1965).
[16]R. L. Hall, Aequ. Math. **8**, 281—6 (1972).
[17]R. L. Hall, Can. J. Phys. **50**, 305—11 (1972).
[18]R. L. Hall, Phys. Lett. B **30**, 320—1 (1969).

Symmetries of the 6*j* coefficient

K. Venkatesh

Department of Physics, University of Mysore, Manasagangotri, Mysore 570 006, India
(Received 31 August 1977)

The equivalence between the two sets of ${}_4F_3(1)S$ for the 6*j* coefficient is explained through two sets of equalities. It is shown that the 144 symmetries can be partitioned into 12 sets of 12 each, depending on the number of terms in the series representation for the 6*j* coefficient. The 12-element group, recently discussed by Lockwood, is shown to account for only 12 symmetries of the 6*j* coefficient.

The 6*j* coefficient can be written[1] in the following symmetrical form:

$$(6J)=\left(\prod_{i=1}^{3}\prod_{k=1}^{4}\frac{(\beta_i-\alpha_k)!}{(\alpha_k+1)!}\right)^{\frac{1}{2}} \times\sum_t(-1)^t(t+1)!\left(\prod_{i=1}^{3}\prod_{k=1}^{4}(\beta_i-t)!(t-\alpha_k)!\right)^{-1}, \tag{1}$$

where $\beta_1=a+b+c+d$, $\beta_2=a+d+e+f$, $\beta_3=b+c+e+f$, and $\alpha_1=a+b+e$, $\alpha_2=c+d+e$, $\alpha_3=a+c+f$, $\alpha_4=b+d+f$. Let β_0 be the minimum of the β's and α_0 be the maximum of the α's. The number of terms in the above series representation, Eq. (1), is determined from: $n=\beta_0-\alpha_0$.

In our earlier studies, we have shown that the 144 symmetries of the 6*j* coefficient can be accounted for, either by a set of three ${}_4F_3(1)S$ (Set I),[2] each accounting for 48 symmetries, or by a set of four ${}_4F_3(1)S$ (Set II),[3] each accounting for 36 symmetries. From the study of the parameters and the convergence condition[4] of the ${}_4F_3(1)S$, the upper limit of the summation index in the ${}_4F_3(1)S$, n was determined as the minimum of the negative value of the negative parameters in the numerator. Since the number of terms in the Eq. (1) and the ${}_4F_3(1)S$ for the 6*j* coefficient are the same, we have the following sets of equalities:

Set I:

$$\text{MIN}(c+d-e,\ a+b-e,\ b+d-f,\ a+c-f)=\beta_1-\alpha_0, \quad \text{(i)}$$
$$\text{MIN}(d+f-b,\ a+f-c,\ d+e-c,\ a+e-b)=\beta_2-\alpha_0, \quad \text{(ii)}$$
$$\text{MIN}(c+f-a,\ b+f-d,\ b+e-a,\ c+e-d)=\beta_3-\alpha_0, \quad \text{(iii)}$$

Set II:

$$\text{MIN}(c+d-e,\ d+f-b,\ c+f-a)=\beta_0-\alpha_1, \quad \text{(i)}$$
$$\text{MIN}(a+b-e,\ a+f-c,\ b+f-d)=\beta_0-\alpha_2, \quad \text{(ii)}$$
$$\text{MIN}(b+d-f,\ d+e-c,\ b+e-a)=\beta_0-\alpha_3, \quad \text{(iii)}$$
$$\text{MIN}(a+c-f,\ a+e-b,\ c+e-d)=\beta_0-\alpha_4. \quad \text{(iv)}$$

Considering the equality (i) of Set I, if the minimum on the left-hand side is fixed, say $c+d-e$, then, obviously, $\alpha_0=\alpha_1$. From the way in which the two sets of ${}_4F_3(1)S$ are derived,[2,3] it is evident that all those symmetries of the 6*j* coefficient which are accounted for by the permutations of the parameters of the ${}_4F_3(1)$ series (i) of Set I are simultaneously accounted for by the ${}_4F_3(1)$ series (i) of Set II and both the ${}_4F_3(1)S$ will have $c+d-e+1$ number of terms. In fact, only 12 symmetries are accounted to by the ${}_4F_3(1)$ series (i) of Set I when it has $c+d-e+1$ number of terms. As an illustration of the proof, the list of these 12 symmetries is given in Table I. Hence the 48 symmetries which are accounted for by a single ${}_4F_3(1)$ series of Set I can be partitioned into four sets of 12 each, depending on the number of terms in the ${}_4F_3(1)$ series. Therefore, from the above arguments, it is straightforward that each one of the ${}_4F_3(1)$ series in Set II will simultaneously account for the set of 12 symmetries which are accounted for by a single ${}_4F_3(1)$ series of Set I and both the ${}_4F_3(1)S$ will have the same number of terms.

Starting from the Set II of equalities, it is easy to see that the converse is also true. The 36 symmetries which are accounted for by a single ${}_4F_3(1)$ series of Set II can be partitioned into three sets of 12 each. Each one of the three ${}_4F_3(1)S$ in Set I will simultaneously account for the set of 12 symmetries which are accounted for by a single ${}_4F_3(1)$ series of Set II and both the ${}_4F_3(1)S$ will have the same number of terms.

From the above argument, it is clear that the 144 symmetries can be partitioned into 12 sets of 12 each, depending on the number of terms in the series representation for the 6*j* coefficient. Since the number of terms in the series representation, Eq. (1), takes 12 different values[3] as the 6*j* coefficient goes through it's 144 symmetries, there should be 12 sets of canonical parameters[1] for the 6*j* coefficient: $(n_i; a_L^i, b_L^i, c_L^i; d_L^i, e_L^i; i=1,12)$.[6] It is elementary to calculate the 12 sets of canonical parameters. While the series representation, Eq. (1), is invariant for all values of n, the expression for the 6*j* coefficient in terms of the canonical parameters, Eq. (5) of Ref. 1, is invariant for the three values of n that are possible for a $\alpha_0(q_z$ of Ref. 1). Since the parameters d_L and e_L change their values as n takes different values, only 12 symmetries are accounted for by the 12-element symmetry group of permutations of $6J(n; a_L, b_L, c_L; d_L, e_L)$. Thus the set of 12 symmetries of the 6*j* coefficient corresponding to an n_i will become a group by itself in the canonical parametrization.

By making the substitution $t=\beta_0-s$ in Eq. (1), the 6*j* coefficient may be expressed as follows:

$$(6J)=PRT, \tag{2}$$

TABLE I: The list of 12 symmetries (Ref. 5) accounted to by the ${}_4F_3(1)$ series when it has $c+d-e+1$ number of terms.

$\begin{Bmatrix} a & b & e \\ d & c & f \end{Bmatrix}$	$\begin{Bmatrix} a_1 & b_1 & e_1 \\ d_1 & c_1 & f_1 \end{Bmatrix}$	$\begin{Bmatrix} e_2 & b_2 & a_2 \\ f_2 & c_2 & d_2 \end{Bmatrix}$
$\begin{Bmatrix} b & a & e \\ c & d & f \end{Bmatrix}$	$\begin{Bmatrix} b_1 & a_1 & e_1 \\ c_1 & d_1 & f_1 \end{Bmatrix}$	$\begin{Bmatrix} b_2 & e_2 & a_2 \\ c_2 & f_2 & d_2 \end{Bmatrix}$
$\begin{Bmatrix} a_3 & b_3 & e_3 \\ d_3 & c_3 & f_3 \end{Bmatrix}$	$\begin{Bmatrix} a_4 & e_4 & b_4 \\ d_4 & f_4 & c_4 \end{Bmatrix}$	$\begin{Bmatrix} e_5 & b_5 & a_5 \\ f_5 & c_5 & d_5 \end{Bmatrix}$
$\begin{Bmatrix} b_3 & a_3 & e_3 \\ c_3 & d_3 & f_3 \end{Bmatrix}$	$\begin{Bmatrix} e_4 & a_4 & b_4 \\ f_4 & d_4 & c_4 \end{Bmatrix}$	$\begin{Bmatrix} b_5 & e_5 & a_5 \\ c_5 & f_5 & d_5 \end{Bmatrix}$

 0022-2488/78/1909-1973$1.00

where

$$P=(-1)^{a_L+b_L+c_L+d_L+e_L}, \tag{3}$$

$$T=\sum_{s=0}^{n}(-1)^{-s}(4n+a_L+b_L+c_L+d_L+e_L-s+1)! \times[s!(s+d_L)!(s+e_L)!(n-s)!(n+a_L-s)! \times(n+b_L-s)!(n+c_L-s)!]^{-1}, \tag{4}$$

and R is given by Eq. (7) of Ref. 1.

Equation (2) is invariant under any permutation of a_L, b_L, and c_L or of d_L and e_L. Thus a 12-element symmetry group is evident, which is equivalent to that defined by Lockwood. Equation (2) is invariant for the four values of n that are possible for a β_0.

The conclusion is, that the two sets of equalities involving the parameters of the $6j$ coefficient explain, as to how the two sets of ${}_4F_3(1)S$ are equivalent in describing the 144 symmetries. The advantage of canonical parameters of the $6j$ coefficient is that the set of 12 symmetries corresponding to an n_i $(i=1,12)$ will form a group.

Note added in proof: Recently, the author has come to know of the article by Dr. K. Srinivasa Rao in the Proceedings of the VI Int. Colloquium on Group Theoretical Methods in Physics, Tubingen (1977). In this article Dr. Rao has also discussed the canonical parametrization of the $6j$ coefficient introduced by Lockwood.

ACKNOWLEDGMENTS

The author is grateful to Dr. G. Ramachandran for an interesting discussion and to Professor B. Sanjeevaiah for providing facilities for research. The author acknowledges with thanks the CSIR (India) for financial assistance.

[1]Loren A. Lockwood, J. Math. Phys. **18**, 45 (1977).
[2]K. Srinivasa Rao, T. S. Santhanam, and K. Venkatesh, J. Math. Phys. **16**, 1528 (1975).
[3]K. Srinivasa Rao and K. Venkatesh, Proc. Fifth Int. Colloquium on Group Theoretical Methods in Physics, Montreal (1976).
[4]M. E. Rose, *Multipole Fields* (Wiley, New York, 1955), Appendix B, p. 92.
[5]Subscripts in the parameters indicate that they are one of the five Regge symmetries. See T. Regge, Nuovo Cimento **11**, 116 (1959).
[6]The canonical parameters of Lockwood are written as a_L,b_L,c_L,d_L, and e_L.

On the equivalence of quantum mechanics and a certain class of Markov processes

Mark Davidson

Lawrence Berkeley Laboratory, Berkeley, California 94720
(Received 29 August 1977)

Some new results in the stochastic model of quantum mechanics are derived. The origin of the noncommutative algebra of quantum mechanics is given a simple interpretation within the stochastic model. Quantum expectations of mixed functions of coordinates and momenta are related to stochastic averages.

I. INTRODUCTION

It is well known that under certain restrictions, quantum mechanics can be interpreted as a Markov process for classical coordinates. Various authors[1-10] have derived Schrödinger's equation within the context of classical diffusion theory. These works interpret quantum statistics or quantum indeterminacy as originating from the interaction of nonrelativistic point particles with random forces. These forces must be present even in the vacuum. This view does not conflict with present understanding of nature, since vacuum fluctuations in quantum field theory are known to exist and are well verified experimentally. The random forces experienced by particles in the stochastic interpretation may be viewed as coming about due to random fluctuations of the various fields in the vacuum.

The techniques used to analyze stochastic models of quantum mechanics have their origins in the early works of Einstein and Somoluchowski who pioneered the modeling of Brownian motion. A more recent review has been given by Chandrasekhar.[11] For the more mathematically inclined, a number of texts deal with the general problem of stochastic processes.[12,13] A review of the stochastic interpretation has been given by Jammer.[14]

The most rigorous and precise derivation of Schrödinger's equation has been given by Nelson.[2,3] He considered a certain class of Markov process typical of a diffusion theory. He found that if the diffusion constant was proportional to $\hbar/m$, then the nonlinear diffusion equations could be transformed into a linear Schrödinger's equation. For the class of models Nelson studied, the sample trajectories are continuous but not differentiable. It is not possible to attribute an instantaneous velocity to a particle in his theory. In this paper we shall show that these properties allow a certain temporal ordering to be defined which allows one to relate arbitrary quantum operator expectations to stochastic averages. Noncommuting operators arise in a natural and physically appealing way. Finally, a way of extending Nelson's work to slightly complex times will be presented which makes the quantum theory and stochastic theory related in a very simple fashion. These results add credence to the stochastic interpretation of quantum mechanics, and may also prove to be of some use in diffusion theory.

In the next section some properties of the Wiener process, the archetypical diffusion process, are presented. These results are generalized to Nelson's model and their implications discussed in Sec. III. Conclusions are presented in Sec. IV.

II. A PROPERTY OF THE WIENER PROCESS

Wiener processes are discussed in numerous texts.[15-17] Let $W(t)$ denote such a process with zero drift. That is $W(t)-W(s)$ independent of $W(r)$ for $r<s, t$, and $E(W(t))=0$. Let us suppose that $W(0)=0$ identically. The following results have been shown:

(a) $W(t)-W(s)$ has a Gaussian distribution.

(b) Some expectations are:

$$E(W(t_1)W(t_2))=\nu\min(t_1,t_2)=\overline{W(t_1)W(t)}, \tag{1}$$

where ν shall be referred to here as the diffusion constant,

$$I_n=E\exp[J_1W(t_1)]\times\exp[J_nW(t_n])$$
$$=\exp\left(\tfrac{1}{2}\sum_{i,j=1}^{n}J_iJ_j\,\overline{W(t_i)W(t_j)}\right). \tag{2}$$

For n even we find

$$E(W(t_1)\times\cdots\times W(t_n))$$
$$=\frac{1}{2^{n/2}(n/2!)}\sum_p\overline{W(t_1)W(t_2)}\;\overline{W(t_3)W(t_4)}$$
$$\times\cdots\times\overline{W(t_{n-1})W(t_n)}, \tag{3}$$

where Σ_p denotes a sum over all permutations. For n odd, 3 vanishes.

The following result is derived from 2 and 1 if $t_k\neq t_l$, $l\neq k$,

$$\lim_{\substack{t_n-t_{n-1}\to 0,\\ J_n,J_{n-1}\to 0}}\left[\frac{1}{J_{n-1}}\frac{\partial}{\partial J_n}\frac{\partial}{\partial t_{n-1}}-\frac{1}{J_n}\frac{\partial}{\partial J_{n-1}}\frac{\partial}{\partial t_n}\right]I_n=\nu I_{n-2}. \tag{4}$$

Let us use the notation

$$x_1(t)=W(t),\quad x_2(t)=\dot{W}(t),\quad O_1(t)=1,\quad O_2(t)=\frac{\partial}{\partial t}. \tag{5}$$

This notation is not meant to suggest that $\dot{W}$ exists as a random variable. It does not. We define for $j_i=1$ or 2

$$\overline{x_{j_1}(t)\times\cdots\times x_{j_n}(t)}$$
$$=\lim_{\substack{t_{i+1}-t_i\to 0,\\ t_1=t}}\prod_{i=1}^{n}O_{j_i}(t_i)E(W(t_1)\times\cdots\times W(t_n)), \tag{6}$$

 0022-2488/78/1909-1975$1.00

It turns out that order matters for these functions. For example,

$$\overline{W(t)\dot{W}(t)} = \lim_{t_2-t_1\to 0.} \frac{\partial}{\partial t_2} E(W(t_1)W(t_2)) = 0,$$

$$\overline{\dot{W}(t)W(t)} = \lim_{t_2-t_1\to 0.} \frac{\partial}{\partial t_1} E(W(t_1)W(t_2)) = \nu.$$

If we extend the definition to sums by

$$\overline{A+B} = \overline{A}+\overline{B}, \quad \overline{cA} = c\overline{A}, \tag{7}$$

then we have

$$\overline{\dot{W}(t)W(t) - W(t)\dot{W}(t)} = \nu. \tag{8}$$

This is quite suggestive of quantum mechanics. The following much more general results follow from Eq. (4) and (6),

$$\overline{f(W(t),\dot{W}(t))(\dot{W}(t)W(t) - W(t)\ \dot{W}(t))g(W(t),\dot{W}(t))}$$
$$= \nu\overline{f(W(t),\dot{W}(t))g(W(t),\dot{W}(t))}, \tag{9}$$

for arbitrary ordered polynominal functions f and g. In other words, whenever $[\dot{W},W]$ appears in expressions defined by (4), (5), and (6) it may be replaced by ν. Loosely speaking we have

$$[\dot{W}(t),W(t)] = \nu. \tag{10}$$

This should be compared with the quantum mechanical result

$$[m\dot{x},x] = -i\hbar. \tag{11}$$

The diffusion constant ν must be real, but still these results are extremely suggestive.

The Wiener process also has a completeness relation. With the density ρ given by

$$\rho(W,t) = E(\delta(W - W(t))), \tag{12}$$

with δ the Dirac delta function, we have for $t_i < t$, $s_i > t$:

$$\int dW \frac{1}{\rho} E(W(t_1)\times\cdots\times W(t_n)\delta(W-W(t)))$$
$$\times E(W(s_1)\times\cdots\times W(s_m)\delta(W-W(t))) \tag{13}$$
$$= E(W(t_1)\times\cdots\times W(t_n)W(s_1)\times\cdots\times W(s_n)).$$

For the functions defined in (4), (5), and (6) this becomes

$$\int dw \frac{1}{\rho(W,t)} \overline{f(W(t),\dot{W}(t))\delta(W-W(t))}$$
$$\times \overline{\delta(W-W(t))g(W(t),\dot{W}(t))}$$
$$= \overline{f(W(t),(\dot{t}))g(W(t),(\dot{t}))}. \tag{14}$$

In the next section we consider a more general process for which these relations remain true. The restriction $W(0)=0$ can be relaxed without changing any of these results.

III. A STOCHASTIC MODEL OF QUANTUM MECHANICS

Nelson[2,3] has shown that a certain class of Markov processes yields a Schrödinger type equation. Only processes with a single degree of freedom will be considered here. The processes he looked at were defined by the difference equation

$$dx(t) = b(x(t),t)dt + dW(t), \tag{15}$$

which in integral form is

$$x(t) = x(0) + \int_0^t dt' b(x(t'),t') + W(t) - W(0), \tag{16}$$

where b is a smooth function of its arguments and $W(t)$ is a Wiener process. This class of processes is also discussed by Doob[12] and Breiman.[15]

Analogous to Eq. (6), time ordered functions can be defined $[x_1(t)=x(t),\ x_2(t)=\dot{x}(t)]$

$$\overline{x_{j_1}(t)\times\cdots\times x_{j_n}(t)}$$
$$= \lim_{\substack{t_{i+1}-t_i\to 0. \\ t_1=t}} \prod_{i=1}^{n} O_{j_i}(t_i)E(x(t_1)\times\cdots\times x(t_n)) \tag{17}$$

$$\overline{A+B} = \overline{A}+\overline{B}, \quad \overline{cA} = c\overline{A} \tag{18}$$

The following completeness results would be satisfied in a reasonable theory (They will be assumed true here.):

$$\int dx \frac{1}{\rho(x,t)} \overline{f(x(t),\dot{x}(t))\delta(x-x(t))}\ \overline{\delta(x-x(t))g(x(t),\dot{x}(t))}$$
$$= \overline{f(x(t),\dot{x}(t))g(x(t),\dot{x}(t))}, \tag{19}$$

$$p(x,t) = \overline{\delta(x-x(t))}, \tag{20}$$

The following result follows from Eq. (15),

$$\lim_{\substack{\Delta\to 0. \\ \epsilon\to 0. \\ \Delta/\epsilon\to 0}} \frac{1}{\Delta} E((x(t+\Delta)-x(t))x(t+\epsilon)$$
$$-(x(t+\epsilon+\Delta)-x(t+\epsilon))x(t)|x(t)=x)$$

$$\lim_{\Delta\to 0.} \frac{1}{\Delta}((W(t+\Delta)-W(t))^2)$$
$$= \nu. \tag{21}$$

It follows in particular from this that

$$\overline{\dot{x}(t)x(t) - x(t)\dot{x}(t)} = \nu, \tag{22}$$

and in general that

$$\overline{f(x(t),\dot{x}(t))[\dot{x}(t),x(t)]g(x(t),\dot{x}(t))}$$
$$= \nu\overline{f(x(t),\dot{x}(t))g(x(t),\dot{x}(t))} \tag{23}$$

so that once again we have the relation

$$[\dot{x}(t),x(t)] = \nu. \tag{24}$$

Analogous to (15) is a backwards equation

$$dx = b_*(x,t)\,dt + dW_*, \tag{25}$$

where

$$b_*(x,t)=\lim_{\Delta\to 0}\frac{1}{\Delta}E(x(t)-x(t-\Delta)\,|\,x(t)=x), \tag{26}$$

$$b(x,t)=\lim_{\Delta\to 0}\frac{1}{\Delta}E(x(t+\Delta)-x(t)\,|\,x(t)=x). \tag{27}$$

Two Fokker–Planck equations may be derived. They are

$$\frac{\partial\rho}{\partial t}=-\frac{\partial}{\partial x}(\rho b)+\frac{\nu}{2}\rho'', \tag{28}$$

$$\frac{\partial\rho}{\partial t}=-\frac{\partial}{\partial x}(\rho b_*)-\frac{\nu}{2}\rho'', \tag{29}$$

or

$$\frac{\partial\rho}{\partial t}=-\frac{\partial}{\partial x}(\rho V),\quad V=\frac{b+b_*}{2}, \tag{30}$$

$$U=\frac{\nu}{2}\frac{\partial}{\partial x}\ln(\rho),\quad U=\frac{b-b_*}{2}. \tag{31}$$

Equations (30) and (31) constitute two equations for three unknowns: b, b_*, and ρ. An additional equation is necessary to define a solution. Nelson supplemented these with an equation similar to Newton's $F=ma$. First he defined forward and backward derivatives

$$\begin{aligned}Df(x,t)&=\lim_{\Delta\to 0}\frac{1}{\Delta}E(f(x(t+\Delta),t+0)-f\ (x(t),t)\,|\,x(t)=x)\\&=\left(\frac{\partial}{\partial t}+b\frac{\partial}{\partial x}+\frac{\nu}{2}\frac{\partial^2}{\partial x^2}\right)f(x,t),\end{aligned} \tag{32}$$

$$\begin{aligned}D_*f(x,t)&=\lim_{\Delta\to 0}\frac{1}{\Delta}E(f(x(t),t)-f(x(t-\Delta),\ t-\Delta\ )|\,x(t)=x)\\&=\left(\frac{\partial}{\partial t}+b\frac{\partial}{\partial x}-\frac{\nu}{2}\frac{\partial^2}{\partial x^2}\right)f(x,t).\end{aligned} \tag{33}$$

He defines mean acceleration by

$$a=\tfrac{1}{2}(DD_*+D_*D)x, \tag{34}$$

and he compeletes the system of equations by

$$ma=F=-\frac{\partial}{\partial x}\phi(x), \tag{35}$$

where $\phi(x)$ is supposed to be the potential function for external forces. This assumption has been criticized,[18] but a rebuttal has also been given.[19] We shall accept Nelson's model for the remainder of this paper.

For this set of equations [(30),(31),(35)], Nelson found the following solution. Writing

$$\rho=\psi^*\psi,\quad \psi=e^{R+iS}, \tag{36}$$

he showed

$$\frac{\partial R}{\partial x}=\frac{1}{\nu}U,\quad \frac{\partial S}{\partial x}=\frac{1}{\nu}V \tag{37}$$

and

$$\frac{\partial\psi}{\partial t}=i\frac{\nu}{2}\frac{\partial^2\psi}{\partial x^2}-\frac{i}{m\nu}\phi\psi. \tag{38}$$

If $\nu=\hbar/m$, with $\hbar$ Planck's constant, then (38) takes the form

$$i\hbar\frac{\partial\psi}{\partial t}=\left(i\hbar\frac{\partial}{\partial x}\right)^2\psi+\phi\psi, \tag{39}$$

which is Schrödinger's equation for a one-dimensional particle in a potential ϕ. It follows that

$$\overline{f(x(t))}=\int dx\ \psi^*(x,t)f(x)\psi(x,t). \tag{40}$$

Nelson argued that (39) contains all the measurable information of quantum theory (when generalized to higher number of degrees of freedom) because only approximate positions of particles are ultimately measured in an experiment. It is still desirable to have a physical interpretation of expectations involving momenta.

The quantum expectation of a general operator f is

$$\langle\psi\,|f|\,\psi\rangle=\int dx\ e^{R-iS}f\left(x,-i\hbar\frac{\partial}{\partial x}\right)e^{R+iS}. \tag{41}$$

We will now show the following result for $\nu=\hbar/m$,

$$\overline{f(x,m\dot{x})}=\int dx\ e^{R-S}f\left(x,\hbar\frac{\partial}{\partial x}\right)e^{R+S}. \tag{42}$$

Consider the commutation rule under the bar sign: $[\dot{x},x]=\nu$. $\nu(\partial/\partial x)$ satisfies the same relation, $[\nu(\partial/\partial x),x]=\nu$. This suggests that we seek an operator for $m\dot{x}$ of the form $m\nu(\partial/\partial x)+F(x)$, where $F(x)$ is an as yet undetermined function of x. Let us try the choice

$$\overline{f(x,m\dot{x})}=\int dx\ \psi^*f\left(x,\nu m\frac{\partial}{\partial x}+F(x)\right)\psi. \tag{43}$$

In order that (43) be real for f real, it is necessary that

$$F(x)=-iS'+F'_R, \tag{44}$$

where F_R is real. We then have

$$\overline{f(x,m\dot{x})}=\int dx\ e^{R-F_R}f\left(x,\nu m\frac{\partial}{\partial x}\right)e^{R+F_R}. \tag{45}$$

We now show that $F_R=S$. From (40) we have the result

$$\frac{d}{dt}\overline{x^n}=\frac{d}{dt}\langle\psi|\,x^n\,|\psi\rangle \tag{46}$$

or

$$\overline{\dot{x}x^{n-1}+\cdots+x^{n-1}\dot{x}}=\left\langle\psi\left|\frac{p}{m}x^{n-1}+\cdots+x^{n-1}\frac{p}{m}\right|\psi\right\rangle. \tag{47}$$

Substituting (45) into (47) yields

$$F_R=S, \tag{48}$$

from which (42) follows. This derivation is by no means rigorous. What we can say is that if (42) is satisfied, then the correct commutation rules emerge, and the relations between time derivatives are correct.

Using (42), we derive the completeness relation [Eq. (19)]. This follows by substituting (42) into (19) and integrating. Thus, Eq. (45) is consistent with the completeness relation for Markov processes. Although our derivation is

not rigorous, it is the author's opinion that (42) must be correct. It can be rewritten

$$\overline{f(x,m\dot{x})} = \int dx\, \psi^* e^{iS-S} f(x,ip) e^{S-iS}\psi. \tag{49}$$

These relations relate arbitrary quantum expectations to stochastic expectations. Some simple examples are

$$\overline{m^2\dot{x}^2} = m^2 \int dx\, b(x) b_*(x) \rho(x), \tag{50}$$

$$\langle\psi|p^2|\psi\rangle = \overline{m^2\dot{x}^2} + 2\int dx\, \rho U^2. \tag{51}$$

In general, if

$$\langle| f(p,x) |\psi\rangle = H(S,R), \tag{52}$$

where H is a functional, then

$$\overline{f(-im\,\dot{x},x)} = H(iS,R)\,. \tag{53}$$

These results extend the connection between stochastic processes and quantum mechanics. They may be of practical use in two ways: First they suggest a new way of calculating quantum expectations in terms of Markov processes, second they suggest new ways to study certain classical processes in terms of quantum expectations. There should be no problem to generalizing these results to higher numbers of degree of freedom.

We now consider a modification of Nelson's model which makes the connection between quantum mechanics and Markov processes more transparent. We shall consider coordinates and momenta at complex times $t+is$ in the limit $s\to 0$. Suppose that for t fixed, $x(t+is)$ satisfies

$$dx = -ids\, b(x,t+is) - i\,[W_t(s+ds) - W_t(s)], \tag{54}$$

where $W_t(s)$ is a Wiener process in s for t fixed, and $x(t)$ is real for t real.

Analogous to Eqs. (6) and (17) we define an ordered expectation

$$\overline{x_{j_1}\times\cdots\times x_{j_n}} = \lim_{\epsilon\to 0} \overline{x_{j_1}(t+i\epsilon)\times\cdots\times x_{j_n}(t+in\epsilon)}\,. \tag{55}$$

We find in this case the commutation rules

$$[\dot{x},x] = -i\nu, \quad [x,x] = [\dot{x},\dot{x}] = 0. \tag{56}$$

We also find that

$$[x(t_1),x(t_2)] \neq 0. \tag{57}$$

As an example consider

$$[x(t+\Delta),x(t)] = -i\nu\Delta. \tag{58}$$

Additional postulates are needed to achieve Schrödinger's equation for this case. For example, a consistent theory might be possible with the additional assumption for t real,

$$dx = b(x,t)\,dt + dW(t). \tag{59}$$

Equation (59) leads to Schrödinger's equation via Nelson's argument. When coupled with (54) and (55) we find

$$\overline{f(x,m\dot{x})} = \int dx\psi^*\left(x, -im\nu\frac{d}{dx}\right)\psi. \tag{60}$$

With $\nu = \hbar/m$ the similarity to quantum mechanics is striking.

Equations (54) and (59) may not be mutually compatible. It might be possible to derive Schrödinger's equation from Eq. (54) alone along with a suitable assumption relating mean acceleration to a potential. We shall defer a more detailed study of these complex processes to a future work.

IV. CONCLUSION

It has been shown that a noncommutative algebra can be defined for Markov processes of the type which lead to Schrödinger's equation. Quantum expectations involving both coordinates and momenta are thereby related to stochastic expectations. These results add credence to the stochastic interpretation of quantum mechanics. It is difficult to believe that the similarity between these two theories is coincidental.

A method of extending the stochastic model of Nelson has also been given which continues the model to slightly complex times. The correspondence between quantum theory and the stochastic model are extremely close in this case.

Given the incomplete nature of nonrelativistic quantum mechanics, the stochastic models must be extended to relativistic fields. Only then are new results likely to follow.

[1] L.de la Peña-Auerbach, Rev. Mex. Fis. **19**, 133-45 (1970); Phys. Lett. A **24**, 603-4 (1967); **27**, 594-5 (1968).
[2] E. Nelson, *Dynamical Theories of Brownian Motion* (Princeton U.P., Princeton, New Jersey, 1967).
[3] E. Nelson, Phys. Rev. **150**, 1079-85 (1966).
[4] Imre Fényes, Zeits. Phys. **132**, 81-106 (1952).
[5] D. Bohm and J.P. Vigier, Phys. Rev. **96**, 208-16 (1954).
[6] R.E. Collins, Found. Phys. **7**, 475-95 (1976).
[7] J.E. Moyal, Proc. Cambridge Philos. Soc. **45**, 99-124 (1949).
[8] L. DeBroglie, Ann. Inst. Henri Poincaré **1**, 1 (1964).
[9] W. Garczyński, Acta Phys. Austriaca Suppl. **6**, 501-17 (1969).
[10] G.G. Comisar, Phys. Rev. **138**, B1332-7 (1965).
[11] S. Chandrasekhar, Rev. Mod. Phys. **15**, 1-89 (1943).
[12] J.L. Doob, *Stochastic Processes* (Wiley, New York, 1953).
[13] K. Ito, *Lectures on Stochastic Processes* (Tata, Bombay, 1961).
[14] Max Jammer, *The Philosophy of Quantum Mechanics* (Wiley, New York, 1974).
[15] L. Breiman, *Probability* (Addison-Wesley, Reading, Mass., 1968).
[16] W. Feller, *An Introduction to Probability Theory and Its Applications* (Wiley, New York, 1950).
[17] B.V. Gnedenko, *The Theory of Probability* (Chelsea, New York, 1962).
[18] A. Kracklauer, Phys. Rev. D **10**, 1358 (1974).
[19] W. Lehr, thesis (Washington State University, 1976).

On critical probabilities in percolation theory[a)]

John C. Wierman

School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455
(Received 8 August 1977)

Several versions of the concept of critical percolation probability are discussed in the bond percolation problem on the square lattice. Critical probabilities are also employed as technical devices in the proofs of two new results. First, there is a critical probability p_T below which all moments of the cluster size are finite. Secondly, an infinite connected cluster of open bonds exists with positive probability if and only if any angular sector contains an infinite connected cluster of open bonds with positive probability. An expression is derived for the expected number of open clusters per bond in the percolation model, relating to the problem of rigorously justifying a critical probability result of Sykes and Essam.

1. INTRODUCTION

This paper discusses the concept of a critical percolation probability in the bond percolation problem on the simple quadratic lattice. A bond percolation process is a mathematical model of the random spread of a fluid through a medium, where the random mechanism is that each bond in the medium is open with probability p, independently of all other bonds. Numerous examples of percolation processes are supplied by Frisch and Hammersley[1] and Shante and Kirkpatrick.[2]

Various definitions of critical probabilities, arising from different interpretations of penetration of the medium by the fluid, are discussed in Sec. 2. Theoretical relationships and bounds for the values of these critical probabilities are discussed.

Three new results concerning critical probabilities are presented, following definitions and preliminaries in Sec. 3. The first result, in Sec. 4, is that if the expected open cluster size is finite, then all moments of the cluster size are finite. The proof is obtained by defining a family of critical probabilities, then showing that they are actually identical. The critical percolation probabilities of the entire lattice and of any positive angular sector are shown to be identical in Sec. 5. In Sec. 6, an expression is derived for the expected number of open clusters per bond in the percolation model. An unverified assumption in the Sykes and Essam critical probability identification for the bond problem on the square lattice relates to this clusters-per-bond function.

2. CRITICAL PROBABILITIES

Imagine the fluid being introduced at a single source site at the origin. If p is small, few bonds will be open, so the fluid will travel only a short distance before becoming completely blocked. If p is nearly one, the fluid may flow indefinitely. The intuitive concept of the critical probability is the threshold value of p above which the fluid penetrates the lattice, and below which the fluid spreads only locally.

There are, however, several versions of the critical probability appearing in the literature, which results in some confusion. Three definitions will be introduced in this section, and others later, in proofs, for technical purposes.

Let $\mathscr{P}_n(p)$ denote the probability that at least n bonds are wetted by fluid from the origin. Since clearly $\mathscr{P}_n(p)$ is monotonically decreasing in n, the limit

$$\mathscr{P}(p)=\lim_{n\to\infty}\mathscr{P}_n(p) \tag{2.1}$$

exists for all p. Evidently $\mathscr{P}(p)$, the *percolation probability*, represents the probability that fluid spreads from a single source site to wet an infinite set of open bonds. The most commonly accepted definition of critical probability is

$$p_H=\inf\{p:\mathscr{P}(p)>0\}. \tag{2.2}$$

which distinguishes between the existence and nonexistence of infinite connected clusters of open bonds.

Let C denote the size of the cluster of open bounds containing the origin, that is, the number of bonds that are wetted by fluid introduced at the origin. Then the expected cluster size, with probability p that each bond is open, is given by

$$E_p(C)=\sum_{n=1}^{\infty}\mathscr{P}_n(p). \tag{2.3}$$

A critical probability p_T is defined by

$$p_T=\inf\{p:E_p(C)=\infty\}, \tag{2.4}$$

and thus represents the threshold value of p above which the expected number of wet bonds becomes infinite.

Since $\mathscr{P}(p)>0$ implies that the expected cluster size $E_p(C)$ is infinite, clearly $p_T\leqslant p_H$. In fact, it is known that for the planar square lattice

$$p_T\leqslant\tfrac{1}{2}\leqslant p_H, \tag{2.5}$$

the lower bound for p_H due to Harris,[3] and the upper bound for p_T a consequence of Lemma 2.2 of Wierman and Reh.[4] (Throughout the remainder of this paper, all results should be interpreted to apply to the planar square lattice only.) Theoretical bounds due to Hammersley,[5,6] show that

$$p_H\leqslant 0.646790 \quad \text{and} \quad p_T\geqslant 0.353210 \tag{2.6}$$

by using arguments involving the numbers of self-avoiding paths on the lattice. An interesting recent development is the discovery of the exact relationship

$$p_T+p_H=1 \tag{2.7}$$

by Seymour and Welsh.[7] It is widely conjectured that in fact

$$p_T=p_H=\tfrac{1}{2}. \tag{2.8}$$

[a)]Research supported in part by University of Minnesota Graduate School Research Grant 494-0350-4909-02 and by National Science Foundation Grant MCS74-05786 A02.

 0022-2488/78/1909-1979$1.00

Sykes and Essam[8] consider a function describing the average number of open clusters per bond on the lattice, and define a critical probability p_E as the location of a singularity of this function. Using several unverified assumptions, it is concluded that $p_E=\frac{1}{2}$ for bond percolation on the square lattice. An attempt to substantiate one of these assumptions is discussed in Sec. 6.

The Sykes and Essam result has led to a folklore belief that the critical probability problem is completely solved for the bond problem on the square lattice. However, by the nature of the definition of p_E, there is no known theoretical relationship between p_E and either p_H or p_T. Thus, even if the Sykes and Essam assumptions are correct and $p_E=\frac{1}{2}$, this provides no information concerning p_H or p_T.

3. SPONGES AND ANNULI

The $m\times n$ *sponge* is the subgraph $T(m,n)$ of the square lattice contained in the rectangle $\{(x,y)\colon 1\leqslant x\leqslant n,\ 1\leqslant y\leqslant m\}$. Each of the points $(1,y)$, $1\leqslant y\leqslant m$, on the left side is regarded as a fluid source, from which fluid percolates through open bonds in the sponge. The sponge crossing probability $S_p(m,n)$ is the probability that at least one of the points (n,k), $1\leqslant k\leqslant m$, on the right side becomes wet by fluid.

Define the *dual sponge* $T^*(m,n)$ by placing a site at each of the points $(i+\frac{1}{2},j+\frac{1}{2})$, $i=1,\ldots,n-1$, $j=0,1,\ldots,m+1$, and connecting these sites with bonds parallel to the coordinate axes. A bond in $T^*(m,n)$ crosses exactly one bond of the square lattice, and is open if and only if the bond crossed is open. A fundamental property of this dual relationship is that either a connected path of open bonds crosses $T(m,n)$ from left to right, or a connected path of closed bonds crosses $T^*(m,n)$ from top to bottom.

In the percolation model, since the events of existence of these crossing paths are mutually exclusive, the probabilities of the events add to unity. Each bond in the dual $T^*(m,n)$ is closed with probability $1-p$, so the probability that a closed path crosses the $(n-1)\times(m+1)$ sponge $T^*(m,n)$ is $S_{1-p}(n-1,m+1)$. Thus, the duality leads to the conclusion

$$S_p(m,n)+S_{1-p}(n-1,m+1)=1. \tag{3.1}$$

Seymour and Welsh[7] introduced the *critical sponge probability* p_s, defined by

$$p_s=\inf\Big\{p:\limsup_{n\to\infty} S_p(n,n+1)>0\Big\}. \tag{3.2}$$

As a tool in proving their main result, it was shown that $p_s=p_T\leqslant\frac{1}{2}$.

Let $R(n)$ denote the annulus of the square lattice bounded by the square C_n on the outside and D_n on the inside, where C_n consists of portions of the lines

$$y=-3n+1,\quad x=3n,\quad y=3n,\quad \text{and}\quad x=-3n+1, \tag{3.3}$$

while D_n consists of portions of the lines

$$y=-n+1,\quad x=n+1,\quad y=n+1,\quad \text{and}\quad x=-n. \tag{3.4}$$

As with the sponge, a dual annulus $R^*(n)$ consists of the sites $(i+\frac{1}{2},j+\frac{1}{2})$ which are connected by bonds crossing the bonds of $R(n)$, with each bond open or closed depending on the bond crossed in $R(n)$.

A circuit is a path of bonds which originates and terminates at the same point. Duality properties imply that either there is a circuit of closed bonds in the dual annulus $R^*(n)$ surrounding D_n or there is a path of open bonds from C_n to D_n in $R(n)$.

The key to the main theorem of Seymour and Welsh is the following result.

Lemma 3.1: The probability that there exists an open circuit in the annuls $R(n)$ which encloses D_n and is enclosed by C_n is at least

$$S_p(2n,2n+1)^{12}\big[1-\sqrt{1-S_p(2n,2n+1)}\,\big]^{64}. \tag{3.5}$$

Harris used a similar result concerning the existence of open circuits in annuli; however, he did not obtain a relationship between the annulus size and the probability of the existence of a circuit.

4. MOMENTS OF CLUSTER SIZE

In this section, the concept of critical probabilities leads to an interesting result about the higher moments of the cluster size. We show that p_T distinguishes between more than just the existence or nonexistence of the expected cluster size, by showing that all moments greater than the first are finite if $p<p_T$. A precise statement follows.

Theorem 4.1: Let C denote the size of the cluster of open bonds containing the origin. Then either $E(C^\alpha)<\infty$ for all $\alpha\geqslant 1$, or $E(C^\alpha)=\infty$ for all $\alpha\geqslant 1$.

In the proof of this result, a critical probability is defined for each moment, representing the value of p where the moment becomes infinite. The technique of Seymour and Welsh can then be used to show that they are actually all equal.

Begin by defining, for each $\alpha\geqslant 1$, the critical probability

$$p_{T^\alpha}=\inf\{p:E_p(C^\alpha)=\infty\}. \tag{4.1}$$

The key to the proof is the following Lemma, a modified version of Lemma 6 of Seymour and Welsh.

Lemma 4.2: If $\epsilon>0$ and $p>p_{T^\alpha}$, $\alpha\geqslant 1$, then for infinitely many values of n,

$$\begin{aligned}&[1-S_p(2n,2n+1)]^{12}\big[1-\sqrt{S_p(2n,2n+1)}\,\big]^{64}\\ &\quad\leqslant 1-1/3^{2\alpha}+\epsilon.\end{aligned} \tag{4.2}$$

The lemma is proved by contradiction. Supposing the conclusion to be false, choose an integer N so large that for all $l\geqslant 3^N$, the probability that there is a closed circuit in the dual annulus $R^*(l)$ is at least $1-1/3^{2\alpha}+\epsilon$. However, a closed circuit in the dual means that there is no open path across $R(l)$ from interior to exterior. Letting D_k denote the event that there is an open path crossing $R(3^k)$ from the inside to the outside, we have

$$P(D_k)\leqslant 1/3^{2\alpha}-\epsilon,\quad \forall k\geqslant N. \tag{4.3}$$

By considering the sequence of disjoint annuli $R(3^k)$,

the events D_k are independent. For an open path to reach from the origin to the outside of $R^*(3^n)$, all the events D_k, $k=1,2,\ldots,n$, must occur. Thus by counting the bonds inside $R(3^k)$'s outer boundary,

$$E(C^\alpha) \leqslant (4\times 3^{2N})^\alpha + \sum_{n\geqslant N} (4\times 3^{2n})^\alpha \prod_{k=N}^{n-1} P(D_k)$$
$$\leqslant 4^\alpha \times 3^{2N\alpha} + \sum_{n\geqslant N} 4^\alpha \times 3^{2n\alpha}(1/3^{2\alpha}-\epsilon)^{n-N} \quad (4.4)$$
$$< \infty,$$

which contradicts $p > p_{T^\alpha}$.

An equivalent statement of Lemma 4.2 is that $p > p_{T^\alpha}$ implies

$$\limsup_{n\to\infty} S_p(2n,2n+1) \geqslant \delta_\alpha > 0. \quad (4.5)$$

By the definition of p_s, it is clear that $p \geqslant p_s$, if $p > p_{T^\alpha}$. Hence we may conclude that $p_{T^\alpha} \geqslant p_s$ for all $\alpha \geqslant 1$.

Since $E_p(C) = \infty$ implies $E_p(C^\alpha) = \infty$ for all $\alpha \geqslant 1$, clearly $p_{T^\alpha} \leqslant p_T = p_s$ for all $\alpha \geqslant 1$.

Together, these conclusions show that $p_{T^\alpha} = p_T = p_s$ for all $\alpha \geqslant 1$, from which Theorem 4.1 follows.

5. INFINITE CLUSTERS IN QUADRANTS AND SECTORS

The critical probability p_H also represents more than just the threshold between existence and nonexistence of an infinite cluster of open bonds. Harris considered the question of whether an infinite cluster exists in a quadrant of the square lattice. Actually, when $p > p_H$, with probability one there exists a connected infinite cluster of open bonds in any angular sector, i.e., $\{(r,\theta): r \geqslant 0,\ 0 \leqslant \theta \leqslant a\}$ in polar coordinates for any $a > 0$. The demonstration given here is for a quadrant, for convenience in notation.

Consider the following sequence of sponges in the first quadrant. If k is an even integer, let B_k denote the $2^{k+1}\times 2^k$ sponge $T(2^{k+1},2^k)$. If k is odd, B_k will denote $T(2^k,2^{k+1})$. Each sponge is twice as long as it is wide, but the orientation with respect to the coordinate axes alternates with k.

The method of proof is to show that for all sufficiently large k, the sponge B_k is crossed lengthwise by a path of open bonds. Since such open paths in successive sponges B_k and B_{k+1} must cross, an infinite collection of these open paths will then link together to form an infinite cluster of open bonds in the first quadrant.

Using the duality properties, it is sufficient to show that for only finitely many k does a path of closed bonds cross the dual sponge B_k^* across its width, when $p > p_H$. Notice that, by Theorem 4.1, $p > p_H$ implies $1-p < p_T$, so the cluster size for closed bonds has all moments finite.

The dual sponge B_k^* is a $(2^k+1)\times(2^{k+1}-1)$ sponge. From each of the $2^{k+1}-1$ sites on one side, the probability of crossing the sponge B_k^* with closed bonds is no larger than the probability that the given site is in a cluster of at least 2^k+1 closed bonds. Hence the probability that the dual sponge is crossed by closed bonds is at most

$$(2^{k+1}-1)P(C \geqslant 2^k+1) \leqslant 2^{k+1}E(C^2)/(2^k+1)^2 \leqslant \frac{E(C^2)}{2^{k-1}}. \quad (5.1)$$

using Markov's inequality.

Therefore, the series of probabilities of these events converges,

$$E(C^2)\sum_{k=1}^{\infty} \frac{1}{2^{k-1}} < \infty, \quad (5.2)$$

so by the Borel–Cantelli Lemma, with probability one only finitely many of the events occur.

6. CLUSTERS PER BOND

Given a finite region of the square lattice, the number of open clusters per bond is simply the number of distinct clusters of open bonds in the region divided by the number of bonds in the region. One of the assumptions of Sykes and Essam was that the expected number of open clusters per bond converged to a limiting value (dependent on p) as the region expands. Grimmett[9] has provided a rigorous justification for this assumption. This section treats Grimmett's result and derives an expression for the limit function.

Let R_{mn} be the rectangular region in the first quadrant bounded by the lines $x=0$, $x=m$, $y=0$, and $y=n$. The bonds on the lines $x=m$ and $y=n$ are not included in R_{mn}, but those on $x=0$ and $y=0$ are, so there are $2mn$ bonds in R_{mn}.

$C_{mn}(k)$ is the number of open clusters of size k contained in R_{mn}.

$C(k)$ is the number of open clusters of size k in the entire lattice which have at least one bond in R_{mn}.

If B is a particular bond, $N_{mn}(B)$ is the number of bonds in R_{mn} which are in the open cluster containing B. [$N_{mn}(B)=0$ if B is closed.] $N(B)$ is the number of bonds in the open cluster containing B.

Finally,

$$C_{mn} = \sum_{k=1}^{2mn} C_{mn}(k)$$

and

$$C = \sum_{k=1}^{\infty} C(k).$$

Notice that an open cluster in the lattice may give rise to several distinct open clusters in R_{mn}, but each open cluster in R_{mn} corresponds to just one open cluster in the lattice. Therefore, for all configurations ω of open and closed bonds, one has

$$C(\omega) \leqslant C_{mn}(\omega). \quad (6.1)$$

A weakened form of Grimmett's result, in the bond formulation, may be stated as follows:

Theorem 6.1: $\lim_{m,n\to\infty} [C_{mn}(\omega)/2mn] = \theta(p)$ almost surely and in L^r for all $r > 0$, where

$$\theta(p) = \inf_{m,n} \frac{E_p(C_{mn})}{2mn} = \lim_{m,n\to\infty} \frac{E_p(C_{mn})}{2mn}. \quad (6.2)$$

Grimmett proved that $\theta(p)$ is continuous in p, and that θ is differentiable at $p=0$ with $\theta'(0)=1$. The following identification of $\theta(p)$ provides additional information.

Theorem 6.2: $\theta(p)=\sum_{i=1}^{\infty}(1/i)P_p(N(B)=i)$.

To see this, first notice that C_{mn} cannot exceed C by more than the number of open bonds in the boundary of R_{mn}, which is defined to consist of all bonds lying outside of the rectangle defined by the lines $x=0$, $x=m$, $y=0$, and $y=n$ which have exactly one site on the boundary of this rectangle, plus the bonds along the boundaries $x=m$ and $y=n$ of this rectangle. Thus

$$C_{mn}\leqslant C+(3m+3n+4). \tag{6.3}$$

By inequality (6.1),

$$C_{mn}\geqslant C=\sum_{i=1}^{\infty}C(i)\geqslant\sum_{i=1}^{\infty}\frac{1}{i}\sum_{B\in R_{mn}}l_{[N(B)=i]}. \tag{6.4}$$

where l_A is the indicator random variable of the event A, which is 1 if A occurs and zero otherwise. Then

$$\frac{E_p(C_{mn})}{2mn}\geqslant\sum_{i=1}^{\infty}\frac{1}{i}P_p[N(B)=i]. \tag{6.5}$$

For an inequality in the opposite direction, note that

$$\begin{aligned}C_{mn}&=\sum_{i=1}^{2mn}C_{mn}(i)\\&=\sum_{i=1}^{\infty}\frac{1}{i}\sum_{B\in R_{mn}}l_{[N_{mn}(B)=i]}\\&=\sum_{B\in R_{mn}}\sum_{i=1}^{\infty}\frac{1}{i}l_{[N_{mn}(B)=i]}.\end{aligned} \tag{6.6}$$

Fix a configuration ω of open and closed bonds. For any bond B in a cluster which is totally contained in R_{mn}, we have

$$\sum_{i=1}^{\infty}\frac{1}{i}l_{[N_{mn}(B)=i]}(\omega)=\sum_{i=1}^{\infty}\frac{1}{i}l_{[N(B)=i]}(\omega). \tag{6.7}$$

The bonds contained in clusters which are not entirely in R_{mn} can form no more clusters than the number of open bonds in the boundary of R_{mn}, since each cluster must leave R_{mn} through some open bond in the boundary. Letting ∂R denote the boundary of R_{mn}, for bonds of this type, we obtain

$$\sum_{B}\sum_{i=1}^{\infty}\frac{1}{i}l_{[N_{mn}(B)=i]}(\omega)\leqslant\sum_{B\in\partial R}2\,l_{[B\text{ is open}]}(\omega). \tag{6.8}$$

This is true because the sum on the left-hand side over bonds in any particular R_{mn} cluster is unity, but one or more of the open bonds in ∂R serves as an exit from R_{mn}. The factor of two enters because a single open boundary bond may be an exit for at most two R_{mn} clusters. By (6.6) and (6.8),

$$C_{mn}(\omega)\leqslant\sum_{B\in R_{mn}}\sum_{i=1}^{\infty}\frac{1}{i}l_{[N(B)=1]}(\omega)+2\sum_{B\in\partial R}l_{[B\text{ is open}]}(\omega), \tag{6.9}$$

so by (6.3),

$$\frac{E_p(C_{mn})}{2mn}\leqslant\sum_{i=1}^{\infty}\frac{1}{i}P_p(N(B)=i)+2p\left(\frac{3m+3n+4}{2mn}\right).$$

Letting both m and n tend to $+\infty$ yields

$$\sum_{i=1}^{\infty}\frac{1}{i}P_p(N(B)=i)=\theta(p), \tag{6.10}$$

proving the representation is valid.

Applying Abel summation to this expression yields

$$\theta(p)=p-\sum_{i=2}^{\infty}\frac{1}{i(i-1)}P_p(N(B)\geqslant i) \tag{6.11}$$

Since $P_p(N(B)\geqslant i)$ is monotone increasing in p, this shows that $\theta(p)$ is differentiable almost everywhere with $\theta'(p)\leqslant 1$.

For values of p near zero, the series in (6.11) is clearly $o(p)$, so $\theta'(0)=1$.

Noticing that $P_p(N(B)\geqslant i)\geqslant p^i$ for any p and i, one obtains

$$\theta(p)\leqslant p-\sum_{i=2}^{\infty}\frac{1}{i(i-1)}p^i. \tag{6.12}$$

A calculation gives the bound

$$\theta(p)\leqslant-(1-p)\ln(1-p), \tag{6.13}$$

so that in particular, for all $p\in[0,1]$,

$$\theta(p)\leqslant 1/e. \tag{6.14}$$

[1]H. L. Frisch and J. M. Hammersley, J. SIAM **11**, 894 (1963).
[2]V. K. S. Shante and Scott Kirkpatrick, Adv. Phys. **20**, 325 (1971).
[3]T. E. Harris, Proc. Camb. Phil. Soc. **56**, 13 (1960).
[4]John C. Wierman and Wolfgang Reh, Ann. Prob. (to be published).
[5]J. M. Hammersley, Proc. 87th Intern. Colloq., CNRS, Paris, 17 (1957).
[6]J. M. Hammersley, J. Roy. Stat. Soc. Ser. B **28**, 491 (1966).
[7]P. D. Seymour and D. J. A. Welsh, "Percolation probabilities on the square lattice," *Proceedings, Cambridge Conference on Combinational Theory* (1977).
[8]M. F. Sykes and J. W. Essam, J. Math. Phys. **5**, 1117 (1964).
[9]G. R. Grimmett, J. London Math. Soc. **2**, 346 (1976).

Pointwise ergodic theory on operator algebras

Charles Radin

Mathematics Department, University of Texas, Austin, Texas 78712
(Received 12 August 1977)

We extend Birkhoff's pointwise ergodic theorem from classical mechanics to the overlap with quantum mechanics.

1. INTRODUCTION

Ergodic theory was invented to elucidate the dynamical behavior of nonrelativistic many-body systems as described in a classical mechanical formalism. Soon thereafter it also was called upon to perform similar service for the same systems treated with quantum mechanics.

We will demonstrate below what we believe is the first nontrivial analog for quantum systems of Birkhoff's pointwise ergodic theorem[1]—nontrivial in that it does not require any special properties of the dynamics, in particular it does not require discrete spectrum of the Hamiltonian.[2] As there have been many misdirected versions of quantum ergodic theory, we feel it appropriate to state our objectives and results carefully.

In order to describe problems in classical and quantum mechanics in a parallel fashion it is convenient to use the algebraic formalism[3,4] in which the observables of the physical system are represented by the self-adjoint elements of a norm-separable C^*-algebra A with unit I, and the physical states by a subset of the set S of (mathematical) states of A; $f(a)$, for f in S and a in A, then represents the expected value of a when the system is in the state f. The classical mechanics of a system is described with an Abelian A, and the quantum mechanics with a non-Abelian A. The detailed structure of the dynamics is only imperfectly understood at present for interesting physical systems, but we know from simple examples[4-6] that it cannot be grossly misleading to assume that time evolution is represented by a one-parameter group $\{\alpha^t \mid t\in\mathbf{R}\}$ of *-automorphisms α^t of A. Thus the expectation value at time zero, $f(a)$, would evolve in time t in the Heisenberg picure to $f(\alpha^t a)$ or, equivalently, in the Schrödinger picture to $\tilde{\alpha}^t f(a)$, where $\tilde{\alpha}^t$ is the dual of α^t.

Unfortunately, this structure for the dynamics, which we will call "C^*-dynamics," has been shown[7,8] to be physically untenable in many imporant cases, while there is evidence to support hope that a certain modification might be generally acceptable.[9,10] This more general form, which we will term W^*-dynamics, requires certain (physically) distinguished states f_β, which are time invariant and such that in each of their GNS representations, π_β, of A the dynamics is represented by one-parameter groups $\{\alpha_\beta^t \mid t\in\mathbb{R}\}$ of *-automorphisms of the W^*-algebras $\pi_\beta(A)''$. Therefore, for a state f which is the restriction to A of a state (also denoted f) in the predual of some $\pi_\beta(A)''$, the dynamics is again of the form $f(a)\to f(\alpha^t a)$ or $f(a)\to\tilde{\alpha}^t f(a)$. It is, however, now an important problem to make sense of the evolution "$\tilde{\alpha}^t f$" if f is not of the above type. This problem was considered in Refs. 7 and 11 but only partially solved, in a sense described in later sections.

For simplicity, in the remainder of this section we will ignore the evidence of the last paragraph and assume not only a C^*-dynamics on A but also the appropriateness of a discrete time variable, n, so that the "orbit" of an initial state f would be $\{\tilde{\alpha}^n f \mid n\in\mathbb{Z}\}$. (There are certainly some physically interesting models where this is justified.[4,6])

A common approach to understanding the gross dynamical behavior of physical systems (particularly in statistical mechanics) is through "time average" quantities of the form

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\tilde{\alpha}^n f(a).$$

Of course, before such quantities can be convincingly manipulated, it is necessary to prove their existence. Historically, ergodic theory was invented to solve what we call the

Primary ergodic problem: Prove, for as many states f in S and observables a in A as possible, the existence of

$$\lim_{N\to\infty}\frac{1}{N}\sum_{N=0}^{N-1}\tilde{\alpha}^n f(a).$$

The most significant results in this direction actually correspond to the special form of the

Secondary ergodic problem: Prove, for as many states f in S as possible, the existence of

$$w^*\text{-}\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\tilde{\alpha}^n f. \tag{1}$$

In particular, the results referred to fall in two classes, "mean" and "pointwise" theories. For both one assumes given a state $\bar{f}$ in S which is a fixed point of $\tilde{\alpha}$. Mean ergodic theorems then prove (1) for a relatively small class of states f, mathematically and physically similar to $\bar{f}$, namely for f in L_1, the norm closure of

$$\bigcup_{n>1}\{g\in S \mid g\leqslant n\bar{f}\};$$

Pointwise (or individual) ergodic theorems prove (1) for "$\bar{f}$-almost every" state f in S.

Mean ergodic theory in the above sense was developed for Abelian A only (i.e., for classical mechanical systems) in Refs. 12 and 13 and in the general non-Abelian (i.e., quantum mechanical) setting in Refs. 14 and 15. The pointwise theory was developed for Abelian A in Ref. 1, and it is the effort to obtain a satisfactory noncommutative generalization of this result (i.e., satisfactory quantum version) which is the subject of this paper. For concreteness it is perhaps convenient to keep in mind the above problems for the three-

 0022-2488/78/1909-1983$1.00

dimensional Heisenberg model on an infinite lattice,[9] a highly nontrivial model where the above considerations are easily formulated and of definite interest.

2. NOTATION AND AN EXTENSION OF LANCE'S THEOREM

Throughout this section, M will denote a W^*-algebra, $\bar{f}$ a faithful state in the predual M_* of M, and A a σ-weakly dense, norm-separable sub-C^*-algebra of M containing the unit I. Further, α will denote a *-automorphism of M such that its dual $\tilde{\alpha}$ has $\bar{f}$ for a fixed point.

In this notation, Lance has proven[16] (see also Ref. 17):

Theorem I: There exists a norm continuous linear projection $T: M \to M$ and, for each finite subset A' of A, *a sequence* $\{P_j | j \in \mathbb{N}\}$ of projections in M (the sequence dependent on A') such that $\lim_{n\to\infty} \bar{f}(P_n) = 1$ and

$$\lim_{N\to\infty} \left\| \left[\frac{1}{N} \sum_{n=0}^{N-1} \alpha^n(a) - Ta \right] P_j \right\| = 0$$

for each a in A' and j in $\mathbb{N}$.

As our first step we prove the following related result.

Theorem II: If we further assume that $\bar{f}$ is tracial, then the above sequence $\{P_j | j \in \mathbb{N}\}$ can be chosen independent of $A' \subseteq A$.

Proof: Let $\{a_k | k \in \mathbb{N}\}$ be a norm dense subset of A. For A' being the singleton $\{a_k\}$, let $\{P_j(k) | j \in \mathbb{N}\}$ be the sequence guaranteed by Theorem I. By choosing a subsequence if necessary for each k, we can assume that

$$\bar{f}[P_j(k)] > 1 - 1/2^{k+j}.$$

Define

$$Q_j = \bigwedge_{k\geq 1} \bigwedge_{n\geq j} P_n(k).$$

Therefore

$$I - Q_j = \bigvee_{k\geq 1} \bigvee_{n\geq j} (I - P_n(k)),$$

so using the normality of $\bar{f}$ and (Ref. 18;2.1.5) (which is perhaps *the* crucial step for which we seem to need $\bar{f}$ to be tracial),

$$\bar{f}(I - Q_j) \leq \sum_{k\geq 1} \sum_{n\geq j} \bar{f}(I - P_n(k)) \leq \sum_{k\geq 1} \sum_{n\geq j} 1/2^{k+n} = 1/2^j \to 0 \quad \text{as } j \to \infty$$

i.e., $\lim_{j\to\infty} \bar{f}(Q_j) = 1$. Since $Q_j \leq P_j(k)$ for all j and k, it is clear that

$$\lim_{N\to\infty} \left\| \left[\frac{1}{N} \sum_{n=0}^{N-1} (\alpha^n - T) a_k \right] Q_j \right\| = 0$$

for every k.

Now given any fixed a in A, j in $\mathbb{N}$ and $\epsilon > 0$, one can choose k such that $\|a - a_k\| < \epsilon/4$, and N_1 such that

$$\left\| \left[\frac{1}{N} \sum_{n=0}^{N-1} (\alpha^n - T) a_k \right] Q_j \right\| < \epsilon/2$$

for all $N > N_1$.

Therefore

$$\left\| \left[\frac{1}{N} \sum_{n=0}^{N-1} (\alpha^n - T) a \right] Q_j \right\|$$

$$\leq \left\| \frac{1}{N} \sum_{n=0}^{N-1} (\alpha^n - T)(a - a_k) \right\| + \left\| \left[\frac{1}{N} \sum_{n=0}^{N-1} (\alpha^n - T) a_k Q_j \right] \right\|$$

$$< \epsilon/2 + \epsilon/2 = \epsilon$$

The advantage of having $\{P_j | j \in \mathbb{N}\}$ independent of A' is to avail ourselves of the following ideas based on Segal's noncommutative integration theory.[19]

3. APPLICATION TO THE ERGODIC PROBLEM

Throughout this section we use the notation of Sec. 2 with the further assumption that $\bar{f}$ is tracial. We will also need the following notation from Refs. 7 and 11.

Definition. A sequence $\{P_j | j \in \mathbb{N}\}$ of projections P_j in M will be called an *exhaustion* if: (a) $P_{j+1} \supseteq P_j$ for all j, and (b) $\lim_{j\to\infty} \bar{f}(P_j) = 1$.

Definition: A subset S' of the state space S of A is said to contain *$\bar{f}$-almost every* state, or to be *of full $\bar{f}$-measure*, if there exists an exhaustion $\{Q_j | \in \mathbb{N}\}$ such that

$$S' \supseteq S(\{Q_j | j \in \mathbb{N}\}) \equiv \bigcup_{j\geq 1} \overline{\{f^r | f \in M_*, \operatorname{supp} f \subseteq Q_j\}},$$

where f^r is the restriction to A of f, and the closure is with respect to the w^*-topology of S. The complement of a set of full $\bar{f}$-measure is *of $\bar{f}$-measure zero*.

As demonstrated in Ref. 11, the collection $\hat{S}$ of all sets of full $\bar{f}$-measure is closed under countable intersection. Also, it is proven in Ref. 7 that if A (and therefore M) is Abelian and X is the set of pure states on A in the w^*-topology, so that $A \simeq C(X)$ and $\bar{f}$ is (integration with respect to) a regular Borel probability measure on X, then every set in $\hat{S}$ contains "$\bar{f}$-almost every" point of X in the usual measure theoretic sense. The main result of Ref. 11, which we need at this point, is the noncommutative generalization of the von Neumann–Maharam theorem, namely that any *-automorphism of M, for example α, is implemented by or induces a canonical point transformation on S defined $\bar{f}$-almost everywhere. (The point transformation is "essentially unique" in that any two such transformations would have to agree $\bar{f}$-almost everywhere.) Since the transformation is canonical we will use for it the intuitive notation $f \to \tilde{\alpha} f$.

The generalization of the above from a single *-automorphism to a group of *-automorphisms (in particular to

the cyclic group $\{\alpha^n | n\in\mathbb{Z}\}$) is contained in Ref. 7. Thus given α on M, we have a canonical "orbit" $\{\tilde{\alpha}^n f | n\in\mathbb{Z}\}$ defined for $\bar{f}$-almost every f in S, defining a set $\tilde{S}\subseteq S$. Let $\tilde{\tilde{S}}\subseteq S$ be the set of full $\bar{f}$-measure defined by the exhaustion of Theorem II.

With the above notation we now prove

Theorem III: For $\bar{f}$-almost every f in S, the following limit exists:

$$w^*-\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\tilde{\alpha}^n f. \tag{2}$$

Proof: Let $\{P_j | j\in\mathbb{N}\}$ be an exhaustion corresponding to $\tilde{S}\cap\tilde{\tilde{S}}$ (defined above) and let $f=w^*-\lim_\gamma f_\gamma$, where $f_\gamma\in M_*$ and $\mathrm{supp} f_\gamma\subseteq P_j$, j fixed. For each a in A and integers N_1, N_2,

$$\begin{aligned}
&\left|\frac{1}{N_1}\sum_{n=0}^{N_1-1}\tilde{\alpha}^n f(a)-\frac{1}{N_2}\sum_{n=0}^{N_2-1}\tilde{\alpha}^n f(a)\right| \\
&\quad=\lim_\gamma\left|\frac{1}{N_1}\sum_{n=0}^{N_1-1}\tilde{\alpha}^n f_\gamma(a)-\frac{1}{N_2}\sum_{n=0}^{N_2-1}\tilde{\alpha}^n f_\gamma(a)\right| \\
&\quad=\lim_\gamma\left|\frac{1}{N_1}\sum_{n=0}^{N_1-1}[\tilde{\alpha}^n f_\gamma(a)]-f_\gamma(Ta)\right| \\
&\qquad-\frac{1}{N_2}\sum_{n=0}^{N_2-1}[\tilde{\alpha}^n f_\gamma(a)-f_\gamma(Ta)]\Bigg| \\
&\quad=\lim_\gamma\left|f_\gamma\left[\frac{1}{N_1}\sum_{n+0}^{N_1-1}(\alpha^n a-Ta)\right.\right. \\
&\qquad\left.\left.-\frac{1}{N_2}\sum_{n=0}^{N_2-1}(\alpha^n a-Ta)\right]\right| \\
&\quad\to 0 \quad \text{as } N_1,N_2\to\infty.
\end{aligned}$$

The existence of the limit (2) is then evident from the completeness of S in the w^*-topology.

4. PHYSICAL ASPECTS

Lance's Theorem I is an ergodic theorem concerned with "time averages" of operators; Theorem III is an analogous result (but under the added assumption that $\bar{f}$ be tracial) for time averages of states. As was emphasized in the introduction, aside from their inherent interest there is an added significance for results of the latter form, determined by their widespread utility in physics. We need to comment further on this point.

Assume a C^*-dynamics on A, with a continuous or discrete time variable t. For simplicity further assume $\alpha^t=\gamma^t$ for all approriate t, where $\{\gamma^s | s\in\mathbb{R}\}$ is a strongly continuous one-parameter group of *-automorphisms of A, let $\bar{f}$ be a faithful state on A fixed by all α^t, and let π be the GNS representation of A associated with $\bar{f}$. If $\bar{f}$ is tracial (or KMS as defined below) then the α^t extend by continuity to *-automorphisms of the W^*-algebra $M=\pi(A)''$ and $\bar{f}$ extends by continuity to a faithful, normal, tracial (or KMS) state on M.

The "usefulness" of Theorem III then rests solely on the appropriateness of the assumption that $\bar{f}$ be tracial. Here however there is some difficulty. In any form of measure theory the concept of a set being "of $\bar{f}$-measure zero" is only useful to the extent that such sets are in practice negligable, which depends essentially on the particular $\bar{f}$. For quantum mechanical applications then, before Theorem III can be used effectively one must determine a physically relevant *tracial* $\bar{f}$, one for which sets "of $\bar{f}$-measure zero" would be convicingly small or negligable in relevant calculations. Such states must (and do) occur corresponding to the infinite temperature state, as can be seen by taking the limit $\beta\to 0$ in the canonical ensemble. Physically such a state can clearly be used to study qualitative features at high temperature. As this is the regime where classical and quantum mechanics coincide, our results describe, physically as well as mathematically, the overlap between classical and quantum behavior.

Finite temperature states, which are not tracial, therefore represent the next frontier, and it would be of great value if the condition that $\bar{f}$ be tracial could be dropped from Theorem III; but this also does not seem highly probable at present.

One intermediate problem however, which does not seem entirely hopeless and the solution of which would be of definite physical interest, would be to prove the results of this paper for $\bar{f}$ being a KMS state, where we define this latter notation as follows. For each a, b in A and $d>0$, we define the functions

$$F_{ab}: s\in\mathbb{R}\to\bar{f}(b\gamma^s a)$$
$$G_{ab}: s\in\mathbb{R}\to\bar{f}([\gamma^s a]b)$$
$$e_d: s\in\mathbb{R}\to\exp(ds/2).$$

Then we say $\bar{f}$ is KMS if there exists $\beta>0$ such that $\hat{F}_{ab}=e_{2\beta}\hat{G}_{ab}$, where $\hat{F}_{ab}$ and $\hat{F}_{ab}$ are the Fourier transforms of the respective functions considered as tempered distributions.[20-22] Noting that a trace state could be considered a KMS state corresponding to $\beta=0$, we conclude with the question: Can Theorem II be extended by replacing the assumption that $\bar{f}$ be tracial with the assumption that $\bar{f}$ be KMS?

[1]G.D. Birkhoff, Proc. Natl. Acad. Sci. **17**, 650 (1931).
[2]S. Golden and H.C. Longuet-Higgens, J. Chem. Phys. **33**, 1479 (1960).
[3]I.E. Segal, Ann. Math. **48**, 930 (1947).
[4]G.G. Emch, *Algebraic Methods in Statistical Mechanics and Quantum Field Theory* (Wiley-Interscience, New York, 1972).
[5]R.V. Kadison, Topology **3** Suppl. 2, 177 (1965).
[6]D.A. Dubin, *Solvable Models in Algebraic Statistical Mechanics* (University Press, Oxford, 1974).
[7]C. Radin, Commun. Math. Phys. **33**, 283 (1973).
[8]C. Radin, Commun. Math. Phys. **50**, 69 (1977).
[9]D.A. Dubin and G.L. Sewell, J. Math. Phys. **11**, 2990 (1970).
[10]D. Ruelle, Helv. Phys. Acta **45**, 215 (1972).
[11]C. Radin, Proc. Amer. Math. Soc. **39**, 343 (1973).
[12]G. Birkoff, Proc. Natl. Acad. Sci. **24**, 154 (1938).
[13]S. Kakutani, Ann. Math. **42**, 523 (1941).
[14]C. Radin, Commun. Math. Phys. **21**, 291 (1971).
[15]C. Radin, Advan. Math. **21**, 110 (1976).
[16]E.C. Lance, Invent. Math. **37**, 201 (1976).
[17]Ya.G. Sinai and V.V. Anshelevich, Russian Math. Surveys **31**, 157 (1976).
[18]S. Sakai, *C*-Algebras and W*-Algebras* (Springer-Verlag, New York, 1970).
[19]I.E. Segal, Ann. Math. **57**, 401 (1953) and **58**, 595 (1953).
[20]D. Kastler, J.C.T. Pool, and E.T. Poulsen, Commun. Math. Phys. **12**, 175 (1969).
[21]R. Haag, N.M. Hugenholtz, and M. Winnink, Commun. Math. Phys. **5**, 215 (1976).
[22]M.A. Rieffel and A. Van Daele, Pacific J. Math. **69**, 187 (1977).

Generalized *C*-metric

Frederick J. Ernst

Department of Physics, Illinois Institute of Technology, Chicago, Illinois 60616
(Received 8 June 1977)

We describe a procedure for generalizing static axially symmetric solutions of the vacuum Einstein field equations, and employ this procedure to append an external gravitational field to the *C*-metric black hole solution. We conjecture that if the strength of the appended gravitational field is chosen appropriately, our generalized *C*-metric will serve as a better model for an accelerating black hole than the original *C*-metric.

In spite of the fact that sixty years have elapsed since Weyl[1] showed how *all* static axially symmetric solutions of Einstein's vacuum field equations might be constructed proceeding from solutions of Laplace's equation, surprisingly little effort has been devoted to the identification of those solutions which might possess some physical relevance. As recently as 1970 Kinnersley and Walker[2] succeeded in showing that the *C*-metric, discovered in 1918 by Levi-Civita,[3] is very close to being the field of an accelerating black hole. The present author[4] showed that the undesirable nodal singularity of the charged *C*-metric can be eliminated by employing a Harrison-type transformation in order to append a suitably chosen electric field, suggesting that Kinnersley and Walker were correct in presuming that the nodal singularity was due to the neglect of a physical origin for the acceleration.

Recently we have turned our attention to the analogous problem for the uncharged *C*-metric, speculating that the nodal singularity might be eliminated if a suitably chosen *gravitational* field were appended. We discovered that the procedure of Weyl, when adapted to this type of problem, takes on an unexpectedly elegant form, to which we should like to draw the reader's attention.

After developing the general procedure we shall illustrate it using the *C*-metric as our example. Naturally, the procedure can be applied to *any* static axially symmetric vacuum solution, but it would be very difficult to justify publication of most solutions so generated. However, for one particular choice of the appended gravitational field strength we conjecture that our generalized *C*-metric will serve as a better model for an accelerating black hole that the *C*-metric itself, since a physical origin of the acceleration will have been provided.

I. THE GENERAL PROCEDURE

We begin by considering any static axially symmetric vacuum field, choosing the coordinate system so that

$$ds^2=h[(dx^1)^2+(dx^2)^2]+l(dx^3)^2-f(dx^4)^2, \tag{1}$$

where h, l, and f are functions of x^1 and x^2 alone. It is convenient to introduce a complex differential operator

$$\nabla=\frac{\partial}{\partial x^1}+i\frac{\partial}{\partial x^2} \tag{2}$$

and two auxilliary real fields L and F defined (up to additive constants) by the relations

$$\nabla L=i\nabla \ln l, \tag{3}$$

$$\nabla F=i\nabla \ln f, \tag{4}$$

where

$$\rho=(lf)^{1/2} \tag{5}$$

has the geometrical significance of being the norm of the Killing bivector. That such fields L and F exist is guaranteed by the vacuum field equations!

Weyl himself showed[1] that one can always choose $x^1=\rho$ and $x^2=z$, where z is given (up to an additive constant) by the Cauchy–Riemann equations,

$$\nabla z=i\nabla\rho. \tag{6}$$

However, it is generally rather inconvenient to choose x^1 and x^2 in this fashion. Nevertheless, comparing Eqs. (3) and (4) with Eq. (6), we see that one may always write

$$z=\tfrac{1}{2}(L+F). \tag{7}$$

Naturally this in itself would scarcely justify the introduction of the two real fields L and F. However, we shall see that those fields are more generally useful.

Following Weyl's procedure we introduce auxilliary fields ψ and γ such that

$$h=e^{2(\gamma-\psi)},\quad l=\rho^2e^{-2\psi},\text{ and } f=e^{2\psi}. \tag{8}$$

The field ψ satisfies Laplace's equation, while the field γ is determined by the relation

$$(\nabla\gamma)(\nabla\rho)=\rho(\nabla\psi)^2. \tag{9}$$

[Weyl actually employed the canonical $\rho-z$ coordinates when he wrote out relations equivalent to our Eq. (9).]

It is easy to see that the assignments

$$\psi=kz,\quad \gamma=-\tfrac{1}{2}k^2\rho^2, \tag{10}$$

provide a solution of the vacuum field equations. Therefore, taking advantage of the linearity of Laplace's equation, we may modify our original metric by replacing the field ψ by $\psi+kz$. Since, however, Eq. (9) is not linear in the field γ, the modified field γ will not be simply the original field γ minus $\frac{1}{2}k^2\rho^2$. There will be in addition a contribution $\delta\gamma$, where

$$(\nabla\delta\gamma)(\nabla\rho)=2k\rho(\nabla\psi_{\text{original}})(\nabla z). \tag{11}$$

However, Eq. (6) may be employed in order to obtain the simpler equation

 0022-2488/78/1909-1986$1.00

$$\nabla\delta\gamma = 2ik\rho(\nabla\psi_{\text{original}}). \tag{12}$$

Finally, comparison of Eq. (12) and Eq. (4) yields the remarkable result that

$$\delta\gamma = kF + \text{const.} \tag{13}$$

The requirement of "elementary flatness" in the neighborhood of the symmetry axis fixes the additive constant; if the original metric (1) satisfied this criterion, then the constant in Eq. (13) should be set equal to zero.

Summing up our results, we see that given any static axially symmetric vacuum field (1) we can generate another such solution with the line element given by

$$ds^2 = he^{k(F-L)}e^{-k^2 lf}[(dx^1)^2 + (dx^2)^2] + le^{-k(L+F)}(dx^3)^2 - fe^{k(L+F)}(dx^4)^2. \tag{14}$$

The constant k is a measure of the strength of the appended gravitational field, which is roughly speaking in the negative z direction.

II. THE GENERALIZED C-METRIC

In order to apply the procedure described in the preceding section to the uncharged C-metric,

$$ds^2 = (x+y)^{-2}\left[\frac{(dx)^2}{A(x)} + \frac{(dy)^2}{B(y)} + A(x)(dz)^2 - B(y)(dt)^2\right], \tag{15}$$

it is merely necessary to replace Eq. (15) by

$$ds^2 = (x+y)^{-2}\left[e^{k(F-L)}e^{-k^2(x+y)^{-4}AB}\left(\frac{dx^2}{A} + \frac{dy^2}{B}\right) + Ae^{-k(L+F)}dz^2 - Be^{k(L+F)}dt^2\right]. \tag{16}$$

(The coordinate z employed in the C-metric has nothing to do with Weyl's canonical coordinate. Rather it is an angular coordinate.)

The fields L and F are readily evaluated using Eqs. (3) and (4). Observing that

$$A(x) = 1 - x^2 - 2ma\,x^3, \qquad B(y) = -1 + y^2 - 2ma\,y^3, \tag{17}$$

one obtains the remarkably simple results

$$L = -[(x+y)^{-2}B(y) + 2ma\,y], \qquad F = (x+y)^{-2}A(x) + 2ma\,x. \tag{18}$$

In Ref. 2 the C-metric was interpreted as the field of an accelerating "point particle," with a singularity which "in every way resemble(s) the Schwarzschild singularity." Actually, in Sec. IV of Ref. 2 it was found that the singularity structure of the C-metric is not quite what one would expect for an accelerating black hole; there exists a nodal singularity except when the acceleration parameter vanishes. In Ref. 4 we established that this nodal singularity could be eliminated from the charged C-metric if a suitable electric field is introduced in order to give a physical basis for the acceleration.

The method of analysis employed in Sec. IV of Ref. 2 can also be applied to our generalized C-metric (16). We conjecture that the nodal singularity will disappear when $k = 1$. This conjecture is based upon the expectation that for $k = 1$ the appended gravitational field should provide a physical basis for the acceleration of the "point particle" source of the C-metric.

Regardless of whether our conjecture is proved or disproved, we hope that our method of construction of the generalized C-metric will be of some interest.

ACKNOWLEDGMENT

While the author conceived of this project back in 1975, the actual work was only initiated while he was visiting the University of Waterloo, Waterloo, Ontario, during March and April of 1977. He is grateful to that institution and to J. Wainwright, in particular, for the hospitality which was accorded to him there.

[1] H. Weyl, Ann. Phys. (Leipz.) **54**, 117 (1917).
[2] W. Kinnersley and M. Walker, Phys. Rev. D **2**, 1359 (1970).
[3] T. Levi-Civita, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Nat. Rend. **27**, 343 (1918).
[4] F. J. Ernst, J. Math. Phys. **17**, 515 (1976).

Ideal vortex motion in two dimensions: Symmetries and conservation laws

David M. F. Chapman[a)]

Department of Physics, University of British Columbia, Vancouver, British Columbia, Canada, V6T 1W5
(Received 25 August 1977)

Each of the conservation laws of the motion of a system of vortices in a two-dimensional ideal fluid are shown to be uniquely associated with one of the symmetry transformations of the system, one of which can be a scale transformation. In the case of an infinite fluid with no boundaries the five symmetry transformations are: (a) two independent translations in space, (b) spatial rotations, (c) translation in time, and (d) the scale transformation $t' = e^{2\eta}t$, $z'_k = e^{\eta}z_k$, where η is real and z_k are the positions of the vortices in the z plane. Two cases of reduced symmetry are examined: the upper half-plane and the interior of the unit circle.

I. INTRODUCTION

Interest in the motion of vortices in a two-dimensional, ideal fluid was initiated in the late 19th century by mathematicians and physicists such as Helmholtz, Kirchhoff, Stokes, Routh, Kelvin, and Thomson. Much of this early work has become classic and is covered in the standard textbooks on hydrodynamics by Lamb[1] and Milne-Thomson,[2] as well as in other less well known works.[3] A treatise on the motion of vortices in two dimensions in an ideal fluid with arbitrary boundaries has been produced by Lin,[4] who also reviews some of the early literature.

It was recognized by the early workers that the equations of motion of a system of vortices could be put into canonical form by introducing a function which served as the Hamiltonian of the system and defining appropriate canonical coordinates and momenta. They also found five constants of the motion of vortex systems in an infinite fluid. Most of the literature deals with finding solutions to the equations of motion with various boundaries and initial conditions.

This paper introduces a Lagrangian formalism to discuss vortex motion, whereby the equations of motion of a system of vortices may be viewed as arising from the principle of least action involving a suitable Lagrangian function. The use of this formalism allows the association of the constants of the motion with the symmetry transformations of the system in the usual way, and also reveals a nonstandard result: There is a conservation law of vortex motion associated with a symmetry transformation which is a transformation of scale. Many physical systems have symmetry transformations which are scale transformations, but it is not always possible to associate a conservation law with this symmetry. A consequence of this is that, in most of the literature on Noether's Theorem,[5] scale transformations are dismissed as having no importance vis-á-vis conservation laws, if they are discussed at all.

The existence of this non-standard result makes it necessary to briefly review the least action principle of dynamics in Sec. II, including the definition of symmetry transformations. The usual test for a symmetry transformation is quoted, along with the statement of the associated conservation law. Symmetry transformations involving a scaling of the Lagrangian must be treated separately, and the appendix demonstrates that there exists a class of Lagrangians for which a conservation law may be associated with this type of transformation. The equations of motion, conservation laws, and symmetries of a system of vortices in an infinite fluid are discussed in Sec. III, using the results of Sec. II. Two systems of lower symmetry are examined in Sec. IV.

II. THE PRINCIPLE OF LEAST ACTION, SYMMETRY TRANSFORMATIONS, AND CONSERVATION LAWS

The principle of least action states that, for a dynamical system whose Lagrangian is the function $L(q_k,\dot{q}_k,t)$ $(k=1,2,...,N)$, the variation of the action

$$J=\int_{t_1}^{t_2} L(q_k,\dot{q}_k,t)dt \tag{1}$$

is zero with respect to infinitesimal variations δq_k in that path of integration $q_k(t)$ which is the actual trajectory of the system between $q_k(t_1)$ and $q_k(t_2)$, with $\delta q_k(t_1)=\delta q_k(t_2)=0$. The variation of J depends upon the functional form of L, and the principle of least action ultimately requires that the trajectory of the system be a solution of the differential equations given by the Euler–Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right)-\frac{\partial L}{\partial q_k}=0 \quad (k=1,2,...N) \tag{2}$$

regarding q_k and $\dot{q}_k$ as independent variables for the process of differentiation.

If the variables are subjected to the general transformation

$$t'=t'(t)$$

$$q'_k=q'_k(q_m,t) \quad (k,m=1,2,...N) \tag{3}$$

$$\dot{q}'_k=dq'_k/dt',$$

[a)]Present address: Defense Research Establishment Atlantic, P.O. Box 1012, Dartmouth, Nova Scotia, Canada, B2Y 3Z7.

 0022-2488/78/1909-1988$1.00

the functional form of the Lagrangian (expressed in terms of the new variables) must change in order that the numerical value of the action be invariant so the new Lagrangian L' is defined by

$$L'(q'_k,\dot{q}'_k,t')dt'=L(q_k,\dot{q}_k,t)dt. \tag{4}$$

L' is obtained by substituting the inverse of the transformation (3) into the RHS of (4). Since, in general, L' is of a functional form different from L, the differential equations of motion in the new variables $q'_k,\dot{q}'_k$, and t' will also have a form different from the differential equations in the variables $q_k,\dot{q}_k$, and t. However, if L' is of the form

$$L'(q'_k,\dot{q}'_k,t')=\alpha L(q'_k,\dot{q}'_k,t')+\frac{d}{dt'}\Lambda(q'_k,t') \tag{5}$$

in which α is a real constant and Λ is a scalar function, then the equations of motion are covariant and the transformation (3) is said to be a symmetry transformation of the system. A symmetry transformation maps a solution of the equations of motion into another possible solution. The test for an infinitesimal symmetry transformation for which $\alpha=1$ has been shown by Hill[6] to be

$$\left[L\frac{d}{dt}(\delta t)+\delta t\frac{\partial L}{\partial t}+\sum_k\left(\delta q_k\frac{\partial L}{\partial q_k}+\delta\dot{q}_k\frac{\partial L}{\partial\dot{q}_k}\right)\right]$$

$$=-\frac{d}{dt}\delta\Lambda \tag{6}$$

in which $\delta t=t'-t$, $\delta q_k=q'_k-q_k$, and $\delta\dot{q}_k=\dot{q}'_k-\dot{q}_k$ are infinitesimal quantities and $\delta\Lambda$ is an infinitesimal function of the q_k and t. The LHS must be expressible as the total time derivative of some infinitesimal function of the coordinates and time for the transformation to be a symmetry transformation. If $\delta\Lambda=0$, the Lagrangian is said to be invariant under the transformation. The conservation law associated with such a symmetry transformation reads, from Hill,[6]

$$\frac{d}{dt}\left[\left(L-\sum_k\dot{q}_k\frac{\partial L}{\partial\dot{q}_k}\right)\delta t+\sum_k\delta q_k\frac{\partial L}{\partial\dot{q}_k}+\delta\Lambda\right]=0. \tag{7}$$

A similar test exists for an infinitesimal symmetry transformation which scales L, but the associated conservation law, if one exists, does not read as simply as (7). The test equation results from adding the term $-\beta L$ to the LHS of (6), in which $\beta=\alpha-1$ is an infinitesimal real constant. It is not clear to the author what the general criteria are for such a symmetry transformation to be uniquely associated with a conservation law; however, there does exist a class of Lagrangians, as shown in the Appendix, for which such an association is possible. These Lagrangians are of the form

$$L(q_k,\dot{q}_k)=I(q_k,\dot{q}_k)+F(q_k) \tag{8}$$

in which I is an antisymmetric, bilinear function of the q_k and $\dot{q}_k$ of the form

$$I(q_k,\dot{q}_k)=(1/2)\sum_k\sum_n A_{kn}\dot{q}_k q_n \quad (A_{nk}=-A_{kn}) \tag{9}$$

and F is a function of the q_k alone. The scale transformation

$$t'=e^{\epsilon}t, \quad q'_k=e^{\eta}q_k, \quad \dot{q}'_k=e^{\eta-\epsilon}\dot{q}_k \tag{10}$$

is a symmetry transformation with $\alpha=e^{-2\eta}$ provided that F is the solution of the differential equation

$$\sum_k q_k\frac{\partial F}{\partial q_k}-mF=\lambda \tag{11}$$

for some values of λ and $m=2-\epsilon/\eta$. If such is the case, then the associated conservation law is the conservation of I since, along the trajectory,

$$I=-\tfrac{1}{2}\sum_k q_k\frac{\partial F}{\partial q_k}=-\tfrac{1}{2}(\lambda+mF). \tag{12}$$

F itself is a conserved quantity since L is not an explicit function of t. The conservation of F and I are independent laws, since, although F is always conserved, I is conserved only for those functions F which are solutions of (11).

This brief discourse has relevance to vortex motion since the Lagrangian for a system of vortices in an infinite fluid, to be introduced in Sec. III, is precisely of the form (8).

III. THE MOTION OF VORTICES IN AN INFINITE FLUID

The equations of N vortices situated at the position $z_k(t)$ $(k=1,2,...N)$ in the complex z plane with strengths γ_k $(k=1,2,...N)$ are the N first order, ordinary, nonlinear differential equations

$$i\dot{\bar{z}}_n=\sum_k{}'\frac{\gamma k}{z_n-z_k} \quad (n=1,2,\cdots N) \tag{13}$$

in which the dot represents differentiation with respect to time and the primed summation symbol means the sum over all k excluding $k=n$. The derivation of these equations of motion can be found in any of the classic works[1-3] along with solutions with various initial conditions. Note that the equations, being of first order in time, require only the initial positions as integration constants, the initial velocities being determined by the equations of motion and the initial positions. The presence of boundaries in the fluid would require the addition of extra terms to the RHS of (13).

It has been recognized for some time that these equations of motion have the following first integrals (constants) of the motion[7]:

(a) Center of circulation (two independent constants)

$$Z_0=\gamma_0^{-1}\sum_k\gamma_k z_k \quad (\gamma_0=\sum_k\gamma_k) \tag{14a}$$

(b) Moment of circulation

$$\Theta_0=\sum_k\gamma_k z_k\bar{z}_k \tag{14b}$$

(c) Vortex stream function

$$\Psi_0=-\sum_n\sum_k{}'\gamma_n\gamma_k\ln|z_n-z_k| \tag{14c}$$

(d) Angular moment of circulation

$$I_0=\frac{1}{2i}\sum_k \gamma_k(\bar{z}_k\dot{z}_k-z_k\dot{\bar{z}}_k)$$

$$=\tfrac{1}{2}\sum_n{\sum_k}' \gamma_n\gamma_k \tag{14d}$$

That these are indeed constants of the motion can be verified easily by direct substitution of the equations of motion.

The equations of motion can be put into canonical form by employing the vortex stream function Ψ_0 and by defining the canonical coordinates and momenta

$$q_k=z_k, \quad p_k=\gamma_k\bar{z}_k \tag{15}$$

and writing

$$i\dot{q}_k=\frac{\partial\Psi_0}{\partial p_k}, \quad i\dot{p}_k=-\frac{\partial\Psi_0}{\partial q_k}. \tag{16}$$

The time derivative of any dynamical variable $A(q_k,p_k,t)$ which is not a function of the $\dot{q}_k$ or $\dot{p}_k$ may be written, using (16) and the chain rule of differentiation,

$$i\frac{dA}{dt}=\{A,\psi_0\}+i\frac{\partial A}{\partial t}, \tag{17}$$

where

$$\{A,\Psi_0\}=\sum_k\left\{\frac{\partial A}{\partial q_k}\frac{\partial\Psi_0}{\partial p_k}-\frac{\partial A}{\partial p_k}\frac{\partial\Psi_0}{\partial q_k}\right\}. \tag{18}$$

This provides an alternate method of proving the invariance of Z_0, Θ_0, and Ψ_0, but not of I_0.

It is possible to view the equations of motion as arising from the principle of least action involving the Lagrangian

$$L(z_k,\bar{z}_k,\dot{z}_k,\dot{\bar{z}}_k)=\frac{1}{2i}\sum_k\gamma_k(\bar{z}_k\dot{z}_k-z_k\dot{\bar{z}}_k)$$

$$-\tfrac{1}{2}\sum_n{\sum_k}'\gamma_n\gamma_k\ln(z_n-z_k)(\bar{z}_n-\bar{z}_k)$$

$$=I_0+\Psi_0. \tag{19}$$

The equations of motion (13) and their complex conjugates are given by the Euler–Lagrange equations involving this Lagrangian, regarding z_k and $\bar{z}_k$ as independent coordinates.

This Lagrangian is of the form discussed in Sec. II, with a slight change in notation, and displays some unusual features: the velocities do not appear in quadrature, but in bilinear combination with the coordinates, ensuring that the equations of motion are of first order in time; the Lagrangian is the sum of two constants of the motion, so its value is constant along the trajectory; the Lagrangian cannot be split up into a "kinetic energy" part and a "potential energy" part, as is the case in the mechanics of point particles.

Now that the Lagrangian of the vortex system has been identified, the machinery of Sec. II may be employed to investigate the system for its symmetry transformations and conservation laws. Since the latter are already know, what is of interest is: With which symmetry transformation is each conservation law associated? The test for a symmetry transformation which does not scale L reads

$$L\frac{d(\delta t)}{dt}+\frac{1}{2i}\sum_k\gamma_k(\dot{z}_k\delta\bar{z}_k-\dot{\bar{z}}_k\delta z_k+\bar{z}_k\delta\dot{z}_k-z_k\delta\dot{\bar{z}}_k)$$

$$-\tfrac{1}{2}\sum_n{\sum_k}'\gamma_n\gamma_k\left(\frac{\delta z_n-\delta z_k}{z_n-z_k}+\frac{\delta\bar{z}_n-\delta\bar{z}_k}{\bar{z}_n-\bar{z}_k}\right)=-\frac{d}{dt}\delta\Lambda \tag{20}$$

and the associated conservation law is

$$\frac{d}{dt}\left[\Psi_0\delta t+\frac{1}{2i}\sum_k(\bar{z}_k\delta z_k-z_k\delta\bar{z}_k)+\delta\Lambda\right]=0. \tag{21}$$

Straightforward calculation provides the following symmetry transformations and conservation laws of vortex motion:

(a) Space translation symmetry and conservation of Z_0. An infinitesimal translation in space of the system, represented by

$$\delta t=0, \quad \delta z_k=a, \quad \delta\dot{z}_k=0 \ (a \text{ complex and infinitesimal}) \tag{22}$$

is a symmetry transformation which does not leave L invariant, but which requires the introduction of the function

$$\delta\Lambda=\frac{1}{2i}\left(a\sum_k\gamma_k\bar{z}_k-\bar{a}\sum_k\gamma_k z_k\right). \tag{23}$$

The associated conservation law follows from allowing a to be any complex number. Specifying a to be either purely real or purely imaginary results in the conservation of the imaginary part or real part, respectively, of the center of circulation Z_0, from (21).

(b) Rotational symmetry and conservation of Θ_0. An infinitesimal rotation of the system, represented by

$$\delta t=0, \quad \delta z_k=i\phi z_k, \quad \delta\dot{z}_k=i\phi\dot{z}_k \ (\phi \text{ real and infinitesimal}) \tag{24}$$

is a symmetry transformation which leaves L invariant. The associated conservation law in this case, from (20), is the conservation of the moment of circulation Θ_0.

(c) Time translation symmetry and conservation of Ψ_0. An infinitesimal translation of the system in time, represented by

$$\delta t=\tau, \ \delta z_k=\delta\dot{z}_k=0 \ (\tau \text{ real and infinitesimal}) \tag{25}$$

is a symmetry transformation which leaves L invariant, associated with the conservation of vortex stream function Ψ_0.

This exhausts the standard symmetry transformations of the system. There remains one conservation law, the conservation of I_0, and the symmetry transformation $t'=e^{2\eta}t$, $z'_k=e^{\eta}z_k$, $\dot{z}'_k=e^{-\eta}\dot{z}_k$, with η real.

(d) Scale transformation symmetry and conservation of I_0. Substitution of the canonical equations of motion (16) into I_0 gives the value of I_0 along the trajectory to be

$$I_0=-\tfrac{1}{2}\sum_k\left(z_k\frac{\partial\Psi_0}{\partial z_k}+\bar{z}_k\frac{\partial\Psi_0}{\partial\bar{z}_k}\right). \tag{26}$$

Following (11), it is observed that Ψ_0 is the solution of

$$\sum_k \left(z_k \frac{\partial \Psi_0}{\partial z_k} + \bar{z}_k \frac{\partial \Psi_0}{\partial \bar{z}_k} \right) - m\Psi_0 = \lambda \tag{27}$$

for $m=0$ and $\lambda = -\Sigma_n \Sigma'_k \gamma_n \gamma_k$, so the scale transformation

$$t' = e^{2\eta} t$$

$$z'_k = e^{\eta} z_k \quad (\eta \text{ real}) \tag{28}$$

$$\dot{z}'_k = e^{-\eta} \dot{z}_k,$$

is a symmetry transformation and is the one associated with the conservation of I_0, whose value is, from (26) and (27),

$$I_0 = \tfrac{1}{2} \sum_n \sum_k{}' \gamma_n \gamma_k. \tag{29}$$

In summary, each of the five conservation laws of vortex motion in an infinite, two-dimensional, ideal fluid have been shown to be uniquely associated with one of the five symmetry transformations of the system. One of the symmetry transformations is a scale transformation.

IV. VORTEX MOTION OF REDUCED SYMMETRY

The equations of motion, Lagrangian, symmetry transformations, and conservation laws of a system of vortices discussed so far have been of a system of vortices in an infinite fluid. It is illuminating to examine a couple of systems of which the symmetry is reduced due to the presence of boundaries. In each case the loss of a symmetry transformation due to the addition of symmetry-breaking terms in the Lagrangian is accompanied by the loss of the associated conservation law.

(1) Vortex motion in the upper half plane. The addition of the term

$$L' = \sum_n \sum_k \gamma_n \gamma_k \ln |z_n - \bar{z}_k| \tag{30}$$

to L results in the altered equations of motion

$$i\dot{\bar{z}}_n = \sum_k{}' \frac{\gamma_k}{z_n - z_k} - \sum_k \frac{\gamma_k}{z_n - \bar{z}_k} \quad (n=1,2,...,N). \tag{31}$$

The presence of the extra terms may be interpreted as each vortex at z_k of strength γ_k having induced an "image" at $\bar{z}_k$ of strength $-\gamma_k$. Such a configuration ensures that the real axis is a streamline of the fluid flow at all times, so (32) are the equations of motion of a system of N vortices situated in the upper half plane.[8]

Examining the Lagrangian $L+L'$ for its symmetry transformations reveals:

(a) An infinitesimal translation in space of the form (22) is a symmetry transformation of $L+L'$ only if $a=\bar{a}$, that is, translation along the real axis is a symmetry transformation, translation along the imaginary axis is not. Consequently, only the imaginary part of Z_0 is conserved.

(b) An infinitesimal rotation of the form (24) is not a symmetry transformation of $L+L'$, and Θ_0 is not conserved.

(c) An infinitesimal translation in time of the form (25) is a symmetry transformation of $L+L'$, associated with the conservation of $\Psi'_0 = \Psi_0 + L'$, the new vortex stream function.

(d) The function Ψ'_0 is a solution to (27) with $m=0$ and $\lambda = -\Sigma_k \gamma_k^2$, so the scale transformation (28) is a symmetry transformation of $L+L'$ associated with the conservation of I_0, whose value along the trajectory is

$$I_0 = -\tfrac{1}{2} \sum_k \gamma_k^2. \tag{32}$$

The constants of the motion in the upper half-plane are $(Z_0 - \bar{Z}_0)$, Ψ'_0 and I_0.

(2) Vortex motion inside the unit circle. The addition of the term

$$\dot{L}'' = \sum_n \sum_k \gamma_n \gamma_k \ln |1 - z_n \bar{z}_k| \tag{33}$$

to L gives the altered equations of motion

$$i\dot{\bar{z}}_n = \sum_k{}' \frac{\gamma_k}{z_n - z_k} - \sum_k \frac{\gamma_k}{z_n - 1/\bar{z}_k} \quad (n=1,2,...,N) \tag{34}$$

in analogy with the previous case, the images are induced at the reciprocal points $1/\bar{z}_k$ with opposite strengths $-\gamma_k$, and the unit circle $|z|=1$ is a streamline of the fluid flow, so these are the equations of motion of a system of vortices inside the unit circle.[9]

Investigation of the symmetry transformations of the Lagrangian $L+L''$ reveals:

(a) An infinitesimal translation in space of the form (22) is not a symmetry transformation, and Z_0 is not conserved.

(b) An infinitesimal rotation in space of the form (24) is a symmetry transformation, associated with the conservation of Θ_0.

(c) An infinitesimal translation in time of the form (25) is a symmetry transformation, associated with the conservation of $\Psi''_0 = \Psi_0 + L''$.

(d) The new vortex stream function Ψ''_0 is not a solution of (27) for any values of m or λ, so no scale transformation [including (28)] is a symmetry transformation, and I_0 is not conserved.

The constants of the motion inside the unit circle are Θ_0 and Ψ''_0.

ACKNOWLEDGMENTS

I thank F. A. Kaempffer for introducing to me the sub-

ject of vortex motion, and my co-workers G. Mertz and D. Hally for the suggestions they forwarded during our many discussions. I thank the National Research Council of Canada for their financial support in the form of an NRC Postgraduate Scholarship.

APPENDIX

Consider a dynamical system whose Lagrangian is

$$L(q_k,\dot{q}_k,t)=I(q_k,\dot{q}_k)+F(q_k), \tag{A1}$$

where

$$I=1/2\sum_k\sum_n A_{kn}\dot{q}_k q_n \tag{A2}$$

with $A_{nk}=-A_{nk}$, and $F(q_k)$ some scalar functions of the q_k alone.[10] The equations of motion of the system are

$$\sum_n A_{kn}\dot{q}_n=\frac{\partial F}{\partial q_k} \tag{A3}$$

or, assuming that A has an inverse A^{-1} such that $\Sigma_l A_{kl}A^{-1}_{ln}=\Sigma_l A^{-1}_{kl}A_{ln}=\delta_{kn}$,

$$\dot{q}_n=\sum_k A^{-1}_{nk}\frac{\partial F}{\partial q_k}. \tag{A4}$$

The value of I evaluated along the trajectory is then

$$I=-1/2\sum_k q_k\frac{\partial F}{\partial q_k}. \tag{A5}$$

Since L does not involve t explicitly, an infinitesimal time translation, represented by $\delta t=\tau$, $\delta q_k=\delta\dot{q}_k=0$, is a symmetry transformation, according to (6), and the associated conservation law is the conservation of F, from (8).

Under the scale transformation

$$t'=e^{\epsilon}t,\quad q'_k=e^{\eta}q_k,\quad \dot{q}'_k=e^{(\eta-\epsilon)}\dot{q}_k \tag{A6}$$

the new Langrangian is, from (4)

$$\begin{aligned}L'(q'_k,\dot{q}'_k)&=e^{-\epsilon}L(e^{-\eta}q'_k,e^{(\epsilon-\eta)}\dot{q}'_k)\\&=e^{-2\eta}L(q'_k,\dot{q}'_k)-e^{-2\eta}F(q'_k)\\&\quad+e^{-\epsilon}F(e^{-\eta}q'_k)\end{aligned} \tag{A7}$$

so (A6) is a symmetry transformation with $\alpha=e^{-2\eta}$ provided that

$$e^{-2\eta}F(q'_k)-e^{-\epsilon}F(e^{-\eta}q'_k)=\text{constant}. \tag{A8}$$

[Since F has been postulated to be a function of the coordinates alone, the RHS of (A8) can only be the total derivative of a function Λ which cannot be a function of the coordinates and, at most, is a linear function of time.]

Condition (A8) may be expressed in differential form by letting ϵ and η be infinitesimal and expanding (A8) up to first order terms in small quantities, resulting in the condition that, for (A7) to be a symmetry transformation, F must be a solution to the differential equation

$$\sum_k q_k\frac{\partial F}{\partial q_k}-mF=\lambda \tag{A9}$$

for some values of the real constants λ and $m=2-(\epsilon/\eta)$. If such is the case, it is clear from (A5) that the value of I along the trajectory is

$$I=-\tfrac{1}{2}(\lambda+mF) \tag{A10}$$

and is a conserved quantity of the motion associated with the symmetry transformation.

[1]H. Lamb, *Hydrodynamics* (Cambridge U.P., Cambridge, 1932), Chap. VII.
[2]L.M. Milne-Thomson, *Theoretical Hydrodynamics* (MacMillan, New York, 1968), Chap. XIX.
[3]See, for example, H. Villat, *Lecons sur la theorie des tourbillons* (Gauthier-Villars, Paris, 1930), Chap. III; N.E. Kochin, I.A. Kibel', and N.V. Roze, *Theoretical Hydrodynamics* (Interscience, New York, 1964), translated by B. Boyanovitch, Chap. V; M. Lagally, *Handbuch der Physik* (Springer Verlag, Berlin, 1927), Band III; J.J. Thomson, "Motion of Vortex Rings" (Adams Prize Essay, 1883).
[4]C.C. Lin, *On the Motion of Vortices in Two Dimensions* (University of Toronto Press, Toronto, 1943).
[5]E. Noether, Nachr. Kgl. Ges. Wiss. Gottingen, 171 (1918): there is considerable literature on this subject. A concise treatment of Noether's theorem as applied to classical mechanics has been written by E.A. Desloge and R.I. Karch, Am. J. Phys. **45**, 336 (1977). Also see C. Palmieri and B. Vitale, Nuovo Cimento A **66**, 299 (1970), E.L. Hill (Ref. 6) and literature cited therein.
[6]E.I. Hill, Rev. Mod. Phys. **23**, 253 (1951).
[7]The terminology and notation may not be conventional.
[8]See Ref. 2, p. 360.
[9]See Ref. 2, p. 364.
[10]It is not necessary that A_{kn} be antisymmetric, although only the antisymmetric part $A_{[kn]}$ would appear in the equations of motion and then the antisymmetric part of I would be conserved. The symmetric part of I could be written as $(d/dt)(\frac{1}{4}\Sigma_k\Sigma_n A_{(kn)}q_k q_n)$, where $A_{(kn)}$ is the symmetric part of A_{kn}, so the two Lagrangians would be equivalent, in the sense that they both supply the same differential equations of motion.

Hydrodynamic fluctuation theories

Ronald Forrest Fox

School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332
(Received 3 March 1978, revised manuscript received 17 April 1978)

Incompatible theories for hydrodynamic fluctuations far from full equilibrium are described and compared. The two types of approach are compared with the well established near equilibrium theory of hydrodynamic fluctuations. They are also checked for internal consistency and it is shown that one of the approaches checks out as inconsistent. The other approach is recommended as a valid approaach.

I. INTRODUCTION

The theory of fluctuations for hydrodynamic systems which are close to full equilibrium is firmly established.[1-3] Near full equilibrium, the equations may be linearized, and the Onsager–Machlup theory of thermodynamic fluctuations in linear systems may be extended to include the linearized hydrodynamic problem as a special case.[2] Away from full equilibrium, the hydrodynamic equations are nonlinear, and it is still not completely settled how the fluctuations should be included. Incompatible theories for the nonlinear regime have been presented.[4-6] In this paper it will be shown that one of these approaches[4,6] is also inconsistent internally and must be eliminated while the other approach[5] is almost certainly of the correct structure even if a rigorous mathematical derivation is still wanting.

The study of fluctuations for nonlinear systems has been developed in several different contexts, one of which is hydrodynamics. Early work by van Kampen[7] deals with the transition from microscopic master equations into macroscopic equations in the "macroscopic limit." The resulting macroscopic description is in general nonlinear. Kubo *et al.*[8] have reviewed this method and have carefully analyzed the structure of the resulting equations. A similar result has also been achieved in a separate but closely related study of fluctuations in homogeneous phase chemical reactions. McQuarrie's[9] master equation for such reactions has been studied in the macroscopic limit where equations for the average values[10] and for the fluctuations about the averages[11] have been obtained. Both of these investigations of master equation limits lead to equations, generally nonlinear, for the average values, *and* equations for the fluctuations around the average values. Therefore, the result is that one gets two sets of equations. One set is for the time evolution of the average values and is often called the deterministic set of equations. The other set is for the fluctuations around the deterministic values and can always be rendered as a nonstationary, Gaussian, Markov process, which is in fact a linear system of equation. In chemical reactions, the deterministic equations are the familiar mass action equations. The Gaussian nature of the fluctuation equations reflects the similarity of these results with the central limit theorem of probability theory.[12] Another context for the theory of fluctuations for nonlinear systems is the Ito–Stratonovich calculus[13] for stochastic differential equations. In this approach, fluctuations are simply inserted into the nonlinear macroscopic equations and the theorems of Ito and Stratonovich are then applied. For example, this approach has been used to describe fluctuations for the Bénard instability.[14] Other investigators[4,6] have used similar approaches without any parallel to the Ito–Stratonovich calculus being drawn. Mori[6] has in fact presented a derivation of his equations in this form. Keizer[5] has built a theory of hydrodynamic fluctuations, structured in the same way as the equations which arise in the master equation approach. Consequently, he gets nonlinear hydrodynamic equations for the average values, and a nonstationary, Gaussian, Markov process for the fluctuations around the averages. The fluctuations end up obeying *linear* equations.

Comparison with equilibrium and near equilibrium fluctuations[1-3] will be made here. It will be shown that only Keizer's approach to the hydrodynamics problem is consistent, internally, as well as with near equilibrium fluctuations. The presentation will proceed by first presenting the hydrodynamic equations in their nonlinear form. Following this, both types of fluctuation theory will be exhibited, and a comparison with the near equilibrium theory will be shown. Finally, the inconsistency of inserting the fluctuations directly into the nonlinear equations will be proved.

II. NONLINEAR HYDRODYNAMICS

The Navier–Stokes hydrodynamic equations[2] for the space–time variation of the mass density, $\rho(\vec{r},t)$, the velocity field $\vec{u}(\vec{r},t)$, and energy per gram, $\epsilon(\vec{r},t)$, are

$$\frac{\partial}{\partial t}\rho+\vec{\nabla}\cdot(\rho\vec{u})=0, \tag{1}$$

$$\rho\left(\frac{\partial}{\partial t}u_\alpha+(\vec{u}\cdot\vec{\nabla})u_\alpha\right)=-\frac{\partial}{\partial x_\beta}P_{\alpha\beta}, \tag{2}$$

$$\rho\left(\frac{\partial}{\partial t}\epsilon+(\vec{u}\cdot\vec{\nabla})\epsilon\right)=-\frac{\partial}{\partial x_\alpha}q_\alpha-P_{\alpha\beta}D_{\alpha\beta}, \tag{3}$$

in which repeated indices imply summations, $x_1=x$, $x_2=y$, and $x_3=z$, and $D_{\alpha\beta}$ is defined by

$$D_{\alpha\beta}\equiv\frac{1}{2}\left(\frac{\partial u_\alpha}{\partial x_\beta}+\frac{\partial u_\beta}{\partial x_\alpha}\right). \tag{4}$$

$P_{\alpha\beta}(\vec{r},t)$ is the pressure tensor defined by

$$P_{\alpha\beta}(\vec{r},t)\equiv p(\vec{r},t)\delta_{\alpha\beta}-2\eta\left(D_{\alpha\beta}-\frac{1}{3}D_{\gamma\gamma}\delta_{\alpha\beta}\right)-\xi D_{\gamma\gamma}\delta_{\alpha\beta}, \tag{5}$$

 0022-2488/78/1909-1993$1.00

in which $p(\vec{r},t)$ is the pressure field, η is the shear viscosity, and ξ is the bulk viscosity. $q_\alpha(\vec{r},t)$ is the heat flux vector field and is defined by

$$q_\alpha(\vec{r},t)=-K\frac{\partial}{\partial x_\alpha}T(\vec{r},t), \tag{6}$$

in which $T(\vec{r},t)$ is the temperature field and K is the heat conductivity.

Counting the vector components of $\vec{u}$ as three separate quantities, Eqs. (1), (2), and (3) characterize the space–time behavior of five quantities in terms of seven quantities which include the two additional quantities, p and T. Two of these seven quantities may be eliminated through application of two, local equilibrium, thermodynamic equations of state:

$$p(\vec{r},t)=p(\rho(\vec{r},t),T(\vec{r},t))$$

and (7)

$$\epsilon(\vec{r},t)=\epsilon(\rho(\vec{r},t),T(\vec{r},t)).$$

The first of these equations of state may be used to eliminate p from (2) and (3) in favor of ρ and T. The second equation of state may be used in conjunction with the first law of thermodynamics,

$$d\epsilon=Tds+\frac{p}{\rho^2}d\rho, \tag{8}$$

where s is the entropy per gram, *and* a Maxwell relation

$$\left(\frac{\partial s}{\partial\rho}\right)_T=-\frac{1}{\rho^2}\left(\frac{\partial p}{\partial T}\right)_\rho. \tag{9}$$

To obtain

$$\rho^2\left(\frac{\partial\epsilon}{\partial\rho}\right)_T=p-T\left(\frac{\partial p}{\partial T}\right)_\rho, \tag{10}$$

which permits elimination of ϵ in favor of ρ and T, Eq. (3) can be written

$$\rho\left(\frac{\partial}{\partial t}+\vec{u}\cdot\vec{\nabla}\right)\epsilon$$
$$=\rho\left(\frac{\partial\epsilon}{\partial\rho}\right)_T\left(\frac{\partial}{\partial t}+\vec{u}\cdot\vec{\nabla}\right)\rho+\rho\left(\frac{\partial\epsilon}{\partial T}\right)_\rho\left(\frac{\partial}{\partial t}+\vec{u}\cdot\vec{\nabla}\right)T$$
$$=-\frac{\partial}{\partial x_\alpha}q_\alpha-P_{\alpha\beta}D_{\alpha\beta}. \tag{11}$$

With (1) and the definition of the heat capacity per gram a constant density,

$$C_\rho\equiv\left(\frac{\partial\epsilon}{\partial\rho}\right)_\rho, \tag{12}$$

Eq. (11) becomes

$$\rho C_\rho\left(\frac{\partial}{\partial t}T+(\vec{u}\cdot\vec{\nabla})T\right)$$
$$=\rho\left(\frac{\partial\epsilon}{\partial\rho}\right)_T\rho\vec{\nabla}\cdot\vec{u}-\frac{\partial}{\partial x_\alpha}q_\alpha-P_{\alpha\beta}D_{\alpha\beta}. \tag{13}$$

From (10) it follows that

$$\rho\left(\frac{\partial\epsilon}{\partial\rho}\right)_T\rho\vec{\nabla}\cdot\vec{u}=p\vec{\nabla}\cdot\vec{u}-T\left(\frac{\partial p}{\partial T}\right)_\rho\vec{\nabla}\cdot\vec{u}. \tag{14}$$

From the equation of state (7), $(\partial p/\partial T)_\rho$ is determinable and will be replaced by the symbol B. From (4) and (5) we also get

$$P_{\alpha\beta}D_{\alpha\beta}=p\vec{\nabla}\cdot\vec{u}-2\eta D_{\alpha\beta}D_{\alpha\beta}-(\xi-\frac{2}{3}\eta)(D_{\gamma\gamma})^2. \tag{15}$$

Combining (14) and (15) with (13) yields

$$\rho C_\rho\left(\frac{\partial}{\partial t}T+(\vec{u}\cdot\vec{\nabla})T\right)$$
$$=-\frac{\partial}{\partial x_\alpha}q_\alpha-BT\vec{\nabla}\cdot\vec{u}+2\eta D_{\alpha\beta}D_{\alpha\beta}+(\xi-\tfrac{2}{3}\eta)(D_{\gamma\gamma})^2. \tag{16}$$

Therefore, we will take (1), (2), and (16) as our system of equations for five quantities, ρ, **u**, and T, and understand that the p in $P_{\alpha\beta}$ in Eq. (2) has been eliminated in terms of ρ and T by the equation of state (7). This implies the replacement

$$-\frac{\partial}{\partial x_\alpha}p=-A\frac{\partial}{\partial x_\alpha}\rho-B\frac{\partial}{\partial x_\alpha}T, \tag{17}$$

where A and B are defined by

$$A\equiv\left(\frac{\partial p}{\partial\rho}\right)_T\quad\text{and}\quad B=\left(\frac{\partial p}{\partial T}\right)_\rho. \tag{18}$$

III. HYDRODYNAMIC FLUCTUATIONS

In their book, *Fluid Mechanics*, Landau and Lifshitz[1,15] have argued that in addition to the deterministic space–time evolution of the hydrodynamic quantities, there are also fluctuations in these same quantities. They conclude that the pressure tensor should be augmented by a fluctuating contribution $\tilde{S}_{\alpha\beta}(\vec{r},t)$ which is Gaussian and has first and second moments

$$\langle\tilde{S}_{\alpha\beta}(\vec{r},t)\rangle=0$$

and (19)

$$\langle(\tilde{S}_{\alpha\beta}(\vec{r},t)\tilde{S}_{\mu\nu}(\vec{r}',t')\rangle=2k_BT\Delta_{\alpha\beta\mu\nu}\delta(\vec{r}-\vec{r}')\delta(t-t')$$

in which k_B is Boltzmann's constant, T is the temperature and $\Delta_{\alpha\beta\mu\nu}$ is the correlation tetratic

$$\Delta_{\alpha\beta\mu\nu}=\eta(\delta_{\alpha\mu}\delta_{\beta\nu}+\delta_{\alpha\nu}\delta_{\beta\mu})+(\xi-\tfrac{2}{3}\eta)\delta_{\alpha\beta}\delta_{\mu\nu}, \tag{20}$$

in which η and ξ are the viscosity coefficients in (5). They also find that the heat flux vector should also be augmented by a fluctuating contribution $\tilde{g}_\alpha(\mathbf{r},t)$ which is Gaussian and has first and second moments

$$\langle\tilde{g}_\alpha(\vec{r},t)\rangle=0$$

and (21)

$$\langle\tilde{g}_\alpha(\vec{r},t)\tilde{g}_\beta(\vec{r}',t')\rangle=2K\,k_BT^2\delta(\vec{r}-\vec{r}')\delta(t-t')\,\delta_{\alpha\beta},$$

in which K is the heat conductivity in (6). These expressions were subsequently rigorously established on the basis of an extension[2] of the Onsager–Machlup theory for near equilibrium fluctuations. However, this connection is only valid for the linearized hydrodynamic equations. Moreover, Landau and Lifshitz confine all of their example to the linearized cases.

For later comparison with the nonlinear expressions, the linearized hydrodynamic fluctuation equations are given below. The linearization is with respect to the equilibrium values of $\rho, \vec{u}$, and T which are respectively $\rho_{eq}, \vec{0}$, and T_{eq} and the derivations around these equilibrioum values are denoted by $\Delta\rho$, $\Delta\vec{u}$, and ΔT:

$$\frac{\partial}{\partial t}\Delta\rho+\rho_{eq}\vec{\nabla}\cdot\Delta\vec{u}=0, \tag{22}$$

$$\rho_{eq}\frac{\partial}{\partial t}\Delta u_\alpha=-A_{eq}\frac{\partial}{\partial x_\alpha}\Delta\rho-B_{eq}\frac{\partial}{\partial x_\alpha}\Delta T + \frac{\partial}{\partial x_\rho}[2\eta\Delta D_{\alpha\beta}(\xi-\tfrac{2}{3}\eta)\Delta D_{\gamma\gamma}\delta_{\alpha\beta}] + \frac{\partial}{\partial x_\beta}\tilde{S}_{\alpha\beta}, \tag{23}$$

$$\rho_{eq}C_{\rho\,eq}\frac{\partial}{\partial t}\Delta T=-B_{eq}T_{eq}\vec{\nabla}\cdot\Delta\vec{u} + K\vec{\nabla}^2\Delta T+\frac{\partial}{\partial x_\alpha}\tilde{g}_\alpha, \tag{24}$$

where $\Delta D_{\alpha\beta}=[(\partial\Delta u_\alpha/\partial x_\beta)+(\partial\Delta u_\beta/\partial x_\alpha)]$ and $A_{eq}\equiv(\partial p/\partial\rho)_{eq}$, $B_{eq}\equiv(\partial p/\partial T)_{eq}$, and $C_{\rho\,eq}\equiv(\partial\epsilon/\partial T)_{\rho\,eq}$. These equations, or their equivalent, have been derived by various methods by several authors.[2,3,6]

Several authors[4,6,14] have suggested that $\tilde{S}_{\alpha\beta}$ and $\tilde{g}_\alpha$ should accompany $P_{\alpha\beta}$ and $q_{\alpha\beta}$ everywhere $P_{\alpha\beta}$ and q_α occur in the nonlinear hydrodynamic equations (1), (2), and (16). The effect with regard to the q_α and $\tilde{g}_\alpha$ terms is of no special consequence because these terms appear only linearly even in the nonlinear hydrodynamic equation (16). However, the effect with regard to $\tilde{S}_{\alpha\beta}$ is crucial as will emerge in the sequel. The resulting equations are

$$\frac{\partial}{\partial t}\rho+\vec{\nabla}\cdot(\rho\vec{u})=0, \tag{25}$$

$$\rho\left(\frac{\partial}{\partial t}u_\alpha+(\vec{u}\cdot\vec{\nabla})u_\alpha\right)=-\frac{\partial}{\partial x_\beta}P_{\alpha\beta}+\frac{\partial}{\partial x_\beta}\tilde{S}_{\alpha\beta}, \tag{26}$$

$$\rho C_\rho\left(\frac{\partial}{\partial t}T+(\vec{u}\cdot\vec{\nabla})T\right)=-\frac{\partial}{\partial x_\alpha}q_\alpha+\frac{\partial}{\partial x_\alpha}\tilde{g}_\alpha - BT\vec{\nabla}\cdot\vec{u}+2\eta D_{\alpha\beta}D_{\alpha\beta} +(\xi-\tfrac{2}{3}\eta)(D_{\gamma\gamma})^2 +\tilde{S}_{\alpha\beta}D_{\alpha\beta}. \tag{27}$$

The unusual new term is the last term in (25), $\tilde{S}_{\alpha\beta}D_{\alpha\beta}$.

Using a terminology introduced earlier by the author[16] in another context, the fluctuations in Eqs. (22)–(24) are "additive" while in (25)–(27) they are both "additive" and, in the instance of $\tilde{S}_{\alpha\beta}D_{\alpha\beta}$, also "multiplicative." This "multiplicative" fluctuation will be shown to lead to serious difficulties, in the next section of this paper. It will also be seen that such a term is absent from the alternative formulation of fluctuations for nonlinear hydrodynamics which is presented in the following paragraphs.

The structure of Keizer's theory parallels the structure of results obtained in the master equation treatments of nonlinear fluctuations discussed in the introduction. To be concrete, suppose that the time evolution of a single quantity, n, is governed by a master equation of the form

$$\frac{\partial}{\partial t}P(n,t)=\int W(n,n')P(n',t)dn' - \int W(n',n)P(n,t)\,dn', \tag{28}$$

where $P(n,t)$ is the probability distribution for the values of n at time t. Denote the average value of n by $\bar{n}(t)$ and the fluctuation around $\bar{n}$ by Δn. If $W(n,n')$ satisfies certain conditions which will not be entered into here,[7,8,17] then the result is that

$$\frac{\partial}{\partial t}\bar{n}(t)=K_1(\bar{n}) \tag{29}$$

and

$$\frac{\partial}{\partial t}P(\Delta n,t)=-\frac{\partial}{\partial\Delta n}(K'_1(\bar{n})\Delta nP(\Delta n,t)) + \frac{1}{2}\frac{\partial^2}{\partial\Delta n^2}(K_2(\bar{n})P(\Delta n,t)), \tag{30}$$

where $K_1(\bar{n})\equiv\int dn'(n'-\bar{n})W(n',\bar{n})$, $K_2(\bar{n})\equiv\int dn'(n'-\bar{n})^2\times W(n',\bar{n})$, $K'_1(\bar{n})=(d/d\bar{n})K_1(\bar{n})$, and $P(\Delta n,t)$, as *distinct* from $P(n,t)$, is the probability distribution for the fluctuations, Δn, around the average $\bar{n}$. Equation (30) describes a nonstationary, Gaussian, Markov process which acquires its nonstationarity from the implicitly time dependent coefficients: $K'_1(\bar{n})$ and $K_2(\bar{n})$ which acquire theirs from $\bar{n}(t)$, the solution to (29). Generally, (29) is highly nonlinear because $K_1(\bar{n})$ is nonlinear, *but* it is a functional of $\bar{n}(t)$ alone. It is always possible[18] to render (30) in the equivalent Lagevin-like form

$$\frac{\partial}{\partial t}\Delta n=K'_1(\bar{n})\Delta n+\tilde{f}(t), \tag{31}$$

where $\langle\tilde{f}(t)\rangle=0$, $\langle\tilde{f}(t),\tilde{f}(s)\rangle=K_2(\bar{n}(t))\delta(t-s)$, and $\tilde{f}(t)$ is Gaussian. Equation (31) is clearly linear in Δn, although nonstationary and even though $\bar{n}$ satisfies a nonlinear equation.

In order to apply this theory to hydrodynamics, Keizer[5] has had to introduce a master equation basis from which to extract the results, especially the fluctuating force correlation $K^2(\bar{n})$ above. This he achieved by considering the hydrodynamic fluid to be comprised of myriads of tiny cells, and by analyzing a variety of "elementary processes." I doing so, he obtained a unique set of results as are exhibited below. Clearly, two sets of equations will be required. One set is for the average values of the hydrodynamic quantities, ρ, **u**, and T, which will be denoted by $\bar{\rho}$, $\bar{\vec{u}}$, and $\bar{T}$. The other

set will be for the fluctuations of these quantities *around* the averages and will be given in terms $\Delta\rho$, $\Delta\vec{u}$, and ΔT.

In precise analogy with the results obtained for the mass action laws[10] based upon the McQuarrie master equation, the equations for the average values are exactly the usual *nonlinear* hydrodynamic equations:

$$\frac{\partial}{\partial t}\bar{\rho}+\vec{\nabla}\cdot(\bar{\rho}\bar{\vec{u}})=0, \tag{32}$$

$$\bar{\rho}\left(\frac{\partial}{\partial t}\bar{u}_\alpha+(\bar{\vec{u}}\cdot\vec{\nabla})\bar{u}_\alpha\right)=-A\frac{\partial}{\partial x_\alpha}\bar{\rho}-B\frac{\partial}{\partial x_\alpha}\bar{T}+2\eta\frac{\partial}{\partial x_\beta}\bar{D}_{\alpha\beta}+(\xi-\tfrac{2}{3}\eta)\frac{\partial}{\partial x_\alpha}\bar{D}_{\gamma\gamma}, \tag{33}$$

$$\bar{\rho}C_\rho\left(\frac{\partial}{\partial t}\bar{T}+(\bar{\vec{u}}\cdot\vec{\nabla})\bar{T}\right)=K\nabla^2\bar{T}-B\bar{T}\vec{\nabla}\cdot\bar{\vec{u}}+2\eta\bar{D}_{\alpha\beta}\bar{D}_{\alpha\beta}+(\xi-\tfrac{2}{3}\eta)(\bar{D}_{\gamma\gamma})^2. \tag{34}$$

In (31) it is seen that the fluctuations satisfy a rate equation in which the nonstationary coefficient, $K'_1(\bar{n},(t))$, is derived from the nonlinear rate equation (29) by one differentiation. This amounts to a *linearization around* the average behavior! Because linearization and differentiation with respect to space–time are commutable, the equations are

$$\frac{\partial}{\partial t}\Delta\rho+(\bar{\vec{u}}\cdot\vec{\nabla})\Delta\rho+\Delta\rho\vec{\nabla}\cdot\bar{\vec{u}}+\bar{\rho}\vec{\nabla}\cdot\Delta\vec{u}+(\Delta\vec{u}\cdot\vec{\nabla})\bar{\rho}=0, \tag{35}$$

$$\bar{\rho}\frac{\partial}{\partial t}\Delta u_\alpha+\bar{\rho}(\Delta\vec{u}\cdot\vec{\nabla})\bar{u}_\alpha+\bar{\rho}(\bar{\vec{u}}\cdot\vec{\nabla})\Delta u_\alpha+\Delta\rho\frac{\partial}{\partial t}\bar{u}_\alpha+\Delta\rho(\bar{\vec{u}}\cdot\vec{\nabla})\bar{u}_\alpha=-A\frac{\partial}{\partial x_\alpha}\Delta\rho-B\frac{\partial}{\partial x_\alpha}\Delta T+2\eta\frac{\partial}{\partial x_\alpha}\Delta D_{\alpha\beta}+(\xi-\tfrac{2}{3}\eta)\Delta D_{\gamma\gamma}+\frac{\partial}{\partial x_\beta}\tilde{S}_{\alpha\beta}, \tag{36}$$

$$\bar{\rho}C_\rho\frac{\partial}{\partial t}\Delta T+\bar{\rho}C_\rho(\bar{\vec{u}}\cdot\vec{\nabla})\bar{T}+\bar{\rho}C_\rho(\Delta\vec{u}\cdot\vec{\nabla})\bar{T}+\Delta\rho C_\rho\frac{\partial}{\partial t}\bar{T}+\Delta\rho C_\rho(\bar{\vec{u}}\cdot\vec{\nabla})\bar{T}=K\nabla^2\Delta T-B\Delta T\vec{\nabla}\cdot\vec{u}-B\bar{T}\vec{\nabla}\cdot\Delta\vec{u}+4\eta\bar{D}_{\alpha\beta}\Delta D_{\alpha\beta}+(\xi-\tfrac{2}{3}\eta)2(D_{\gamma\gamma})(\Delta D_{\gamma\gamma})+\frac{\partial}{\partial x_\alpha}\tilde{g}_\alpha+\tilde{S}_{\alpha\beta}\bar{D}_{\alpha\beta}. \tag{37}$$

In (33) and (34), and in (36) and (37), as in (23) and (24) $\bar{D}_{\alpha\beta}$ and $\Delta D_{\alpha\beta}$ are defined in the natural manner: $\bar{D}_{\alpha\beta}\equiv\frac{1}{2}(\partial\bar{u}_\alpha/\partial x_\beta+\partial\bar{u}_\beta/\partial x_\alpha)$ and $\Delta D_{\alpha\beta}\equiv\frac{1}{2}(\partial\Delta u_\alpha/\partial x_\beta+\partial\Delta u_\beta/\partial x_\alpha)$. The important observations regarding (35)–(37) are that they are linear in $\Delta\rho$, $\Delta\vec{u}$ and ΔT, with nonstationary coefficients depending upon $\bar{\rho}$, $\vec{u}$, and $\bar{T}$ which are determined by (32)–(34). Especially notice that the $\tilde{S}_{\alpha\beta}D_{\alpha\beta}$ term in (27) is $\tilde{S}_{\alpha\beta}\bar{D}_{\alpha\beta}$ in (37) which means it is an "additive" fluctuation in the $\Delta\rho$, $\Delta\vec{u}$, and ΔT equations, which has a nonstationary modulator, $\bar{D}_{\alpha\beta}$, determined by (32)–(34). The fluctuations $\tilde{S}_{\alpha\beta}$ and $\tilde{g}_\alpha$ are identical with those given by (19)–(21). However in the near equilibrium, linearized theory in which (19)–(21) were originally derived, the temperature T in the correlations in (19)–(21) is the equilibrium temperature $T_{\rm eq}$. In the far rom equilibrium, nonlinear regime, the temperature T in (19)–(21) is instead $T(\mathbf{r},t)$. The possibility that K, η, and ξ acquire $\mathbf{r}$ and t dependence far from equilibrium may also be considered.[5]

IV. COMPARISON WITH EQUILIBRIUM FLUCTUATIONS

The comparisons between (25)–(27) and (22)–(24) and between (32)–(37) and (22)–(24) will proceed by making the latter comparison first since it is more easily achieved. Near equilibrium, the average values, $\bar{\rho}$, $\bar{\vec{u}}$ and $\bar{T}$ become the equilibrium values, $\rho_{\rm eq}$, $\vec{0}$, $T_{\rm eq}$, which are uniform in space–time so that all space–time derivatives of them vanish. If these values and their derivatives are put into (32)–(34), it is found that each equation reduces to simply $0=0$. When this substitution is made in (35)–(37), all that remains is precisely (22)–(24). In particular, $\tilde{S}_{\alpha\beta}\bar{D}_{\alpha\beta}=0$ near equilibrium because $\bar{D}_{\alpha\beta}=0$ there. Consequently, Keizer's theory goes over into the well established near equilibrium theory very nicely.

The situation is much more problematical with regard to comparing (25)–(27) with (22)–(24). Near equilibrium, nonlinear terms should be linearized around the equilibrium values. The result is

$$\frac{\partial}{\partial t}\Delta\rho+\rho_{\rm eq}\vec{\nabla}\cdot\Delta\vec{u}=0, \tag{38}$$

$$\rho_{\rm eq}\frac{\partial}{\partial t}\Delta u_\alpha=-A_{\rm eq}\frac{\partial}{\partial x_\alpha}\Delta\rho-B_{\rm eq}\frac{\partial}{\partial x_\alpha}\Delta T+\frac{\partial}{\partial x_\beta}\left[2\eta\Delta D_{\alpha\beta}+(\xi-\frac{2}{3}\eta)\Delta D_{\gamma\gamma}\delta_{\alpha\beta}\right]+\frac{\partial}{\partial x_\beta}\tilde{S}_{\alpha\beta}, \tag{39}$$

$$\rho_{\rm eq}C_\rho\frac{\partial}{\partial t}\Delta T=-B_{\rm eq}T_{\rm eq}\vec{\nabla}\cdot\Delta\vec{u}+K\nabla^2\Delta T+\frac{\partial}{\partial x_\alpha}\tilde{g}_\alpha+\tilde{S}_{\alpha\beta}\Delta D_{\alpha\beta}. \tag{40}$$

This is almost total agreement with (22)–(24) except for the $\tilde{S}_{\alpha\beta}\Delta D_{\alpha\beta}$ term which is kept because it is *linear* in the variable $\Delta\vec{u}$. It may be suggested that it must be small since it is bilinear in $\tilde{S}_{\alpha\beta}$ and $\Delta D_{\alpha\beta}$, both of which are supposedly small, so that it is negligible compared with other terms in the equation. However, as will be shown subsequently, its "multiplicative" structure leads to a "large" effect. In fact, generally, if a term is really negligibly small,

then if it is kept along in the analysis it should make a negligibly small effect, so no harm accrues by keeping it.

In Mori's analysis[6] this term is not seen in his near equilibrium results which essentially agree with (22)–(24). However, he carries out two separate derivations: one for near equilibrium and one for far from equilibrium. He does not show that his *results* for far from equilibrium go over into his results for near equilibrium in a natural manner such as exhibited above. Putterman[4,19] on the other hand definitely keeps the $\tilde{S}_{\alpha\beta}\Delta D_{\alpha\beta}$ term in (40). In his recent provacative analysis[19] of $1/f$ noise in voltage fluctuations, it is precisely this term which provides the $1/f$ spectrum.

In order to demonstrate the difficulty associated with the $\tilde{S}_{\alpha\beta}\Delta D_{\alpha\beta}$ term, (38)–(40) are rewritten in a form which normalizes the dimensionality of all the variables.

Define

$$a(\vec{r},t)\equiv\rho_{\rm eq}^{-1/2}\Delta\rho(\vec{r},t), \tag{41}$$

$$a_\alpha(\vec{r},t)\equiv(\rho_{\rm eq}/A)^{1/2}u_\alpha(\vec{r},t), \qquad \alpha=2,3,4, \tag{42}$$

$$a_5(\vec{r},t)\equiv(\rho_{\rm eq}C/T_{\rm eq}A)^{1/2}\Delta T(\vec{r},t), \tag{43}$$

and define the three "matrices," $A_{ij}(\vec{r},\vec{r}')$, $S_{ij}(\vec{r},\vec{r}')$, and $\widetilde{M}_{ij}(\vec{r},\vec{r}')$ by

$$A_{ij}(\vec{r},\vec{r}')\equiv\begin{pmatrix}0 & A_{1\alpha}(\vec{r},\vec{r}') & 0\\ A_{\alpha 1}(\vec{r},\vec{r}') & 0 & A_{\alpha 5}(\vec{r},\vec{r}')\\ 0 & A_{5\alpha}(\vec{r}-\vec{r}') & 0\end{pmatrix},$$

$$\begin{cases}A_{1\alpha}(\vec{r},\vec{r}')=A_{\alpha 1}(\vec{r},\vec{r}')\equiv A^{1/2}\dfrac{\partial}{\partial x_\alpha}\delta(\vec{r},\vec{r}'),\\[2ex] A_{5\alpha}(\vec{r},\vec{r}')=A_{\alpha 5}(\vec{r},\vec{r}')\equiv\dfrac{B}{\rho_{\rm eq}}\left(\dfrac{T_{\rm eq}}{C}\right)^{1/2}\dfrac{\partial}{\partial x_\alpha}\delta(\vec{r},\vec{r}').\end{cases} \tag{44}$$

Because $\alpha=2,3,4$, this is really a 5×5 matrix

$$S_{ij}(\vec{r},\vec{r}')\equiv\begin{pmatrix}0 & 0 & 0\\ 0 & S_{\alpha\beta}(\vec{r},\vec{r}') & 0\\ 0 & 0 & S_{55}(\vec{r},\vec{r}')\end{pmatrix},$$

$$\begin{cases}S_{\alpha\beta}(\vec{r},\vec{r}')=S_{\beta\alpha}(\vec{r},\vec{r}')\equiv\dfrac{1}{\rho_{\rm eq}}\dfrac{\partial}{\partial x_\mu\,\partial x'_\nu}\delta(\vec{r}-\vec{r}')\Delta_{\alpha\beta\mu\nu},\\[2ex] S_{55}(\vec{r},\vec{r}')\equiv\dfrac{K}{\rho_{\rm eq}C}\dfrac{\partial^2}{\partial x_\mu\,\partial x'_\nu}\delta(\vec{r}-\vec{r}')\delta_{\mu\nu},\end{cases} \tag{45}$$

where $\Delta_{\alpha\beta\mu\nu}\equiv(\delta_{\alpha\beta}\delta_{\mu\nu}+\delta_{\alpha\nu}\delta_{\beta\mu})+(\xi-\frac{2}{3}\eta)\delta_{\alpha\mu}\delta_{\beta\nu}$

$$\widetilde{M}_{ij}(\vec{r},\vec{r}')\equiv\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & \widetilde{M}_{5\alpha}(\vec{r},\vec{r}') & 0\end{pmatrix},$$

$$\widetilde{M}_{5\alpha}(\vec{r},\vec{r}')\equiv\tilde{S}_{\alpha\mu}\frac{\partial}{\partial x_\mu}\delta(\vec{r}-\vec{r}'), \qquad \alpha=2,3,4. \tag{46}$$

In this notation, Eqs. (22)–(24) become

$$\frac{\partial}{\partial t}a_i(\vec{r},t)+\int A_{ij}(\vec{r},\vec{r}')a_j(\vec{r}',t)\,d^3\vec{r}' +\int S_{ij}(\vec{r},\vec{r}')a_j(\vec{r}',t)\,d^3\vec{r}'$$

$$=\int\widetilde{M}_{ij}(\vec{r},\vec{r}')a_j(\vec{r}',t)\,d^3\vec{r}'+\widetilde{F}_i, \tag{47}$$

where

$$\widetilde{F}_\alpha(\vec{r},t)\equiv\left(\frac{1}{\rho_{\rm eq}A}\right)^{1/2}\frac{\partial}{\partial x_\beta}\tilde{S}_{\alpha\beta}(\vec{r},t), \qquad \widetilde{F}_1(\vec{r},t)\equiv0,$$

and

$$F_5(\vec{r},t)\equiv\left(\frac{1}{\rho_{\rm eq}T_{\rm eq}AC}\right)^{1/2}\frac{\partial}{\partial x_\alpha}\tilde{g}_\alpha(\vec{r},t).$$

Equation (47) is simultaneously an "additive" stochastic process and a "multiplicative" stochastic process.[16]

Let $G_{ij}(\vec{r},\vec{r}')\equiv A_{ij}(\vec{r},\vec{r}')+S_{ij}(\vec{r},\vec{r}')$ so that (47) may be rendered in the "short-hand" form

$$\frac{d}{dt}\vec{a}(t)+\mathbf{G}\vec{a}(t)=\widetilde{\mathbf{M}}(t)\vec{a}(t)+\vec{F}(t) \tag{48}$$

wherein $\vec{a}(t)$ is considered a vector labelled by *i and* and "continuous index" $\vec{r}$, and $\mathbf{G}$ denotes a matrix labelled by i and j *as well as* by $\vec{r}$ and $\vec{r}'$. Similarily, we may interpret $\widetilde{\mathbf{M}}(t)$ and $\widetilde{F}(t)$ as stochastic objects with *both* discrete indices and *continuous* indices. "Summation" over the $\vec{r}$ variables is integration.

The solution to (48) can be found exactly. Define $\vec{b}(t)$ by

$$\vec{a}(t)\equiv\exp(-\mathbf{G}t)\vec{b}(t). \tag{49}$$

Therefore,

$$\frac{d}{dt}\vec{b}(t) =\exp(\mathbf{G}t)\widetilde{\mathbf{M}}(t)\exp(-\mathbf{G}t)\vec{b}(t)+\exp(\mathbf{G}t)\widetilde{\vec{F}}(t). \tag{50}$$

The solution to this is

$$\vec{b}(t)=\underset{\leftarrow}{T}\exp\left[\int_0^t\exp(\mathbf{G}s)\widetilde{\mathbf{M}}(s)\exp(-\mathbf{G}s)\,ds\right]\vec{b}(0) +\int_0^t\underset{\leftarrow}{T}\exp\left[\int_{t'}^t\exp(\mathbf{G}s)\widetilde{\mathbf{M}}(s)\exp(-\mathbf{G}s)\,ds\right] \times\exp(\mathbf{G}t')\widetilde{\vec{F}}(t')\,dt', \tag{51}$$

in which $\underset{\leftarrow}{T}\exp(\cdots)$ denotes the "time ordered" exponential.

Suppose that we now look at $\langle\vec{b}(t)\rangle$. From (51), this involves two expressions, one of which is

$$\left\langle\underset{\leftarrow}{T}\exp\left[\int_0^t\exp(\mathbf{G}s)\widetilde{\mathbf{M}}(s)\exp(-\mathbf{G}s)\,ds\right]\right\rangle\vec{b}(0) =\exp\left\{\tfrac{1}{2}\int_0^t ds\int_0^t ds'\exp(\mathbf{G}s) \langle\widetilde{\mathbf{M}}(s)\exp[\mathbf{G}(s'-s)]\widetilde{\mathbf{M}}(s')\rangle \times\exp(-\mathbf{G}s')\right\}\vec{b}(0). \tag{52}$$

The equality follows because the $\widetilde{\mathbf{M}}$ has a delta function correlation which, for a Gaussian $\widetilde{\mathbf{M}}$, guarantees that the second cumulant is exact.[20] Now notice that

$$\langle\widetilde{\mathbf{M}}(s)\exp[\mathbf{G}(s'-s)]\widetilde{\mathbf{M}}(s')\rangle=\langle\widetilde{\mathbf{M}}(s)\widetilde{\mathbf{M}}(s')\rangle \tag{53}$$

because the delta function $\delta(s-s')$, implicit in this average, reduces $\exp[\mathbf{G}(s'-s)]$ to the identity matrix. Finally, from the definition of $\tilde{\mathbf{M}}$ it is clearly so that the product $\tilde{\mathbf{M}}(s)\tilde{\mathbf{M}}(s')$ is identically zero; i.e., $\tilde{\mathbf{M}}$ is nilpotent! Therefore, this part of (51) is just $\vec{b}(0)$:

$$\left\langle \underset{\leftarrow}{T} \exp\left[\int_0^t \exp(\mathbf{G}s)\tilde{\mathbf{M}}(s)\exp(-\mathbf{G}s)\,ds\right]\right\rangle \equiv 1,$$

The other expression in $\langle\vec{b}(t)\rangle$ is (54)

$$\left\langle \int_0^t dt' \underset{\leftarrow}{T} \exp\left\{\int_{t'}^t ds\, \exp(\mathbf{G}s)\tilde{\mathbf{M}}(s)\exp(-\mathbf{G}s)\right\} \times \exp(\mathbf{G}t')\vec{\tilde{F}}(t')\right\rangle. \tag{55}$$

From their definitions, it is clear that $\tilde{\mathbf{M}}$ and $\vec{\tilde{F}}$ *are* correlated since $\tilde{F}_\alpha$ contains $(\partial/\partial x_\beta)\tilde{S}_{\alpha\beta}$. All odd order products in the expansion of the exponential will *possibly* give nonzero results for the average when they are multiplied by $\vec{\tilde{F}}$, but even order terms in the exponential expansion will give zero when multiplied by $\vec{\tilde{F}}$ and averaged as a result of Gaussianness and zero-valued averages as in (19)–(21):

$$\begin{aligned}
&\left\langle \int_0^t dt' \underset{\leftarrow}{T} \exp\left[\int_{t'}^t ds\, e^{\mathbf{G}s}\tilde{\mathbf{M}}(s)e^{-\mathbf{G}s}\right] e^{\mathbf{G}t'}\vec{\tilde{F}}(t')\right\rangle \\
&= \left\langle \int_0^t dt'\, e^{\mathbf{G}t'}\vec{\tilde{F}}(t')\right\rangle \\
&\quad + \left\langle \int_0^t dt' \int_{t'}^t ds\, e^{\mathbf{G}s}\tilde{\mathbf{M}}(s)e^{-\mathbf{G}s}e^{\mathbf{G}t'}\vec{\tilde{F}}(t)\right\rangle \\
&\quad + \left\langle \int_0^t dt' \int_{t'}^t ds \int_{t'}^s ds' \right. \\
&\quad \left. e^{\mathbf{G}s}\tilde{\mathbf{M}}(s)e^{\mathbf{G}(s'-s)}\tilde{\mathbf{M}}(s')e^{-\mathbf{G}s'}e^{\mathbf{G}t'}\vec{\tilde{F}}(t')\right\rangle \\
&\quad + \left\langle \int_0^t dt' \int_{t'}^t ds \int_{t'}^s ds' \int_{t'}^{s'} ds'' \right. \\
&\quad \left. e^{\mathbf{G}s}\tilde{\mathbf{M}}(s)e^{\mathbf{G}(s'-s)}\tilde{\mathbf{M}}(s'')e^{-\mathbf{G}s''}e^{\mathbf{G}t'}\vec{\tilde{F}}(t')\right\rangle \\
&\quad + \text{higher order terms.}
\end{aligned} \tag{56}$$

Clearly, the first and third terms of the right-hand side vanish because they are odd order in TS. Any odd order, higher order terms will also vanish. Even the fourth term vanishes because the time delta correlation for TS and the time ordered integrals lead to the identities

$$\begin{aligned}
&\left\langle \int_0^t dt' \int_{t'}^t ds \int_{t'}^s ds' \int_{t'}^{s'} ds''\, e^{\mathbf{G}s}\tilde{\mathbf{M}}(s) \right. \\
&\quad \left. \times e^{\mathbf{G}(s'-s)}\tilde{\mathbf{M}}(s')e^{\mathbf{G}(s''-s')}\tilde{\mathbf{M}}(s'')e^{-\mathbf{G}s''}e^{\mathbf{G}t'}\vec{\tilde{F}}(t')\right\rangle \\
&= \int_0^t dt' \int_{t'}^t ds \int_{t'}^s ds' \int_{t'}^{s'} ds''\, e^{\mathbf{G}s}\langle\tilde{\mathbf{M}}(s) \\
&\quad \times e^{\mathbf{G}(s'-s)}\tilde{\mathbf{M}}(s'))e^{\mathbf{G}(s''-s')}\langle\tilde{\mathbf{M}}(s'') \\
&\quad \times e^{\mathbf{G}(t'-s'')}\vec{\tilde{F}}(t')\rangle = 0
\end{aligned} \tag{57}$$

without any "overlap" correlations. We get zero as with the first part of (51), from the nilpotence of $\tilde{\mathbf{M}}$. All higher *even* order forms will also vanish. Only the second term on the right-hand side of (56) is nonvanishing. It is equivalent to

$$\int_0^t dt' \int_{t'}^t ds\, e^{\mathbf{G}s}\langle\tilde{\mathbf{M}}(s)\vec{\tilde{F}}(t')\rangle. \tag{58}$$

We evaluate this as follows:

$$\begin{aligned}
&\langle\tilde{\mathbf{M}}(s)\vec{\tilde{F}}(t')\rangle \\
&= \int d^3r' \left\langle \tilde{S}_{\alpha\mu}(\vec{r},s)\frac{\partial}{\partial x_\mu}\delta(\vec{r}-\vec{r}') \right. \\
&\quad \left. \times \frac{1}{(\rho_{eq}A)^{1/2}}\frac{\partial}{\partial x'_\beta}\tilde{S}_{\alpha\beta}(\vec{r}',t')\right\rangle \\
&= \frac{2k_BT}{(\rho_{eq}A)^{1/2}} \\
&\quad \times \int d^3r' \left(\frac{\partial}{\partial x_\mu}\delta(\vec{r}-\vec{r}')\right)\left(\frac{\partial}{\partial x'_\beta}\delta(\vec{r}-\vec{r}')\right) \\
&\quad \times [\eta(\delta_{\alpha\alpha}\delta_{\mu\beta}+\delta_{\alpha\beta}\delta_{\mu\alpha})+(\xi-\tfrac{2}{3}\eta)\delta_{\alpha\mu} \\
&\quad \times\delta_{\alpha\beta}]\delta(s-t') \\
&= \frac{2k_BT}{(\rho_{eq}A)^{1/2}}\delta(t'-s)(\tfrac{10}{3}\eta+\xi) \\
&\quad \times \int d^3r' \left(\frac{\partial}{\partial x_\mu}\delta(\vec{r}-\vec{r}')\right)\left(\frac{\partial}{\partial x_\mu}\delta(\vec{r}-\vec{r}')\right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\int_0^t dt' \int_{t'}^t ds\, e^{\mathbf{G}s}\langle\tilde{\mathbf{M}}(s)\vec{F}(t')\rangle \\
&= \left(\int_0^t dt' \tfrac{1}{2}e^{\mathbf{G}t'}\right)\frac{2k_BT}{(\rho_{eq}A)^{1/2}}(\tfrac{10}{3}\eta+\xi) \\
&\quad \times \int d^3r' \left(\frac{\partial}{\partial x_\mu}\delta(\vec{r}-\vec{r}')\right) \\
&\quad \times\left(\frac{\partial}{\partial x'_\mu}\delta(\vec{r}-\vec{r}')\right) = \infty.
\end{aligned} \tag{59}$$

This singularity results from the double Dirac delta functions in the correlation function.

At this stage, the view that a theory with a structure of the type which arises from master equation treatments is both natural *and* necessary for a proper formulation of hydrodynamic fluctuations for the nonlinear regime seems justified. Keizer's[5] theory is such a theory. It remains to be seen if predictions for scattering measurements lead to quantitative agreement with actual measurements in the nonlinear regime.

ACKNOWLEDGMENT

This work has been supported by a grant from the National Science Foundation.

[1]L.D. Landau and E.M. Lifshitz, *Fluid Mechanics* (Pergamon, London, 1959), Chap. 17.
[2]R.F. Fox and G.E. Uhlenbeck, Phys. Fluids **13**, 1893, 2881 (1970).
[3]M. Bixon and R. Zwanzig, Phys. Rev. **187**, 267 (1969).
[4]S.J. Putterman, *Superfluid Hydrodynamics* (North-Holland, Amsterdam, 1974), Chap. 4.
[5]J.E. Keizer, Phys. Fluids **21**, 198 (1978); J. Chem. Phys. **64**, 1679 (1976); **63**, 5037 (1975).

[6]K.T. Mashiyama and H. Mori (to appear in J. Stat. Phys. 1978).
[7]N.G. van Kampen, Can. J. Phys. **39**, 551 (1961).
[8]R. Kubo, K. Matzuo, and K. Kitahara, J. Stat. Phys. **9**, 51 (1973).
[9]D.A. McQuarrie, J. Appl. Prob. **4**, 413 (1967).
[10]I. Oppenheim, K.E. Shuler, and G.H. Weiss, J. Chem. Phys. **50**, 460 (1969).
[11]T. Kurtz, J. Chem. Phys. **57**, 2976 (1972).
[12]T. Kurtz, J. Appl. Prob. **7**, 49 (1970); **8**, 344 (1971).
[13]L. Arnold, *Stochastic Differential Equations* (Wiley-Interscience, New York, 1974).
[14]R. Graham, Phys. Rev. A **10**, 1762 (1974).
[15]L.D. Landau and E.M. Lifshitz, Zh. Eksp. Teor. Fiz. **32**, 618 (1957) [Sov. Phys. JETP **5**, 512 (1957)].
[16]R.F. Fox, J. Math. Phys. **13**, 1196 (1972).
[17]R.F. Fox and M. Kac, BioSystems **8**, 187 (1977).
[18]R.F. Fox, J. Math. Phys. **19**, 127 (1978).
[19]S.J. Putterman, Phys. Rev. Lett. **39**, 585 (1977).
[20]R.F. Fox, J. Math. Phys. **17**, 1148 (1976).

A time-symmetric tachyon universe

Sushil K. Srivastava and Mahendra P. Pathak

Department of Mathematics, University of Gorakhpur, Gorakhpur 273001, India
(Received 28 March 1977)

The momentum flux (p) and contrast momentum flux ($K=\delta\rho/\rho$) are investigated in an open time-symmetric model of the tachyon universe containing different types of tachyon fluid, when $\eta\ll\eta_0$. It has also been shown that pressure turns negative when $\eta\ll\eta_0$. In the light of this approximation, how the physical situation changes in the model has been discussed. The tetrad technique has been used as a mathematical tool.

1. INTRODUCTION

A. Background

Velocity consideration has played a very important role in relativistic physics, and it is this consideration which has given birth to a new class of particles, called tachyons, moving faster than light. Sometime ago, on the basis of the theory of relativity proposed by Einstein in 1905,[1] there was the general belief that this class of particles cannot exist. But nowadays against this old view the existence of such particles is widely believed on theoretical grounds due to concerted efforts of some scientists as Bilaniuk, Deshpande, and Sudarshan,[2] Feinberg[3], Schmidt,[4] Terletskii,[5] Tanaka,[6] Recami and Mignani,[7] Antippa and Everett,[8] and Antippa.[9] Hence, in the present paper we consider tachyons as real particles despite of the fact that they have not been experimentally detected up till now. Though experimental detection of the tachyons has not come to a reality, still, while believing in their existence, it is very natural to think about their production. Therefore, here we will assume that at the epoch of big bang, along with the elementary particles of the other matter, tachyons were produced, and they, being superluminal and gravitationally repelled by the bradyons[7-10] (particles moving slower than light), soon rushed out of the region of bradyons, forming another block of the universe called tachyon universe or Meta-Universe. Many scientists, such as Foster and Ray,[11] Gott,[12] and the authors,[13,14] have proposed models for an open model consistent with the quantum mechanical effects related to the prevention of singularities proposed by Gott. Here, to avoid formation of casual loops, we consider the motion of tachyons unidirectional in space as the motion of bradyons is taken unidirectional in time. For our investigations we have considered our model containing tachyon fluid because a realistic model should include pressure also.

In subsection 1B we have given the tetrad technique to solve sophisticated problems of gravitational field equations in this paper. In Sec. 2 we have got the expressions for variation of momentum flux in the tachyon universe containing tachyon fluid. In the next section, contrast momentum flux ($\delta\rho/\rho$) of the model is investigated. The last section includes discussion of the results obtained in Secs. 2 and 3.

B. Notation

In this paper the space–time is represented as a four-dimensional Riemannian space with metric tensor g_{ij} of signature $(+,+,+,-)$. Covariant differentiation is indicated by a semicolon (;) and covariant differentiation along $(\theta, \phi, \alpha)=$const by a prime over the variable. Round brackets around the indices indicate symmetrization and square brackets antisymmetrization. Throughout the paper the velocity of light and the quantity $8\pi G/C^4$ have been taken as units.

$$R_{ab}-\tfrac{1}{2}g_{ab}R+\Lambda g_{ab}=(\rho+p)u_a u_b-pg_{ab}, \qquad (1.1)$$

where $p=(\gamma-1)\rho$. Here u^a are spacelike 4-velocities for the tachyon fluid so that $u^a u_a=+1$. The acceleration of the fluid is

$$\dot{u}_i u_{i;j}u^j, \text{ where the dot denotes } \partial/\partial t. \qquad (1.2)$$

The velocity gradient may be further split up as

$$u_{i;j}=W_{ij}+\sigma_{ij}+\tfrac{1}{2}\Theta H_{ij}-\dot{u}_i u_j,$$

where $\Theta=u^i{}_{;i}$ is the expansion scalar.

$$\sigma_{ij}=u_{(i;j)}+\dot{u}_{(i}u_{j)}-1/3\Theta H_{ij}$$

is the trace-free shear tensor.

$W_{ij}=u_{[i;j]}+\dot{u}_{[i}u_{j]}$ is the vorticity tensor.

Here H_{ij} is a tensor which projects a quantity from $\eta=$const to $(\theta, \phi, \alpha)=$const defined by

$$H_{ij}=g_{ij}-u_i u_j \qquad (1.3)$$

i.e.,

$$H_{ij}u^j=0, \quad H^i_i=3.$$

Here g_{ij} and u_i have got their usual meaning. The Ricci rotation coefficients are defined by

$$\Gamma_{abc}=e_a\cdot\nabla_b e_c=e^i_a e_{ci;j}e^j_b$$

so that

$$\Gamma_{abc}+\Gamma_{cba}=0.$$

 0022-2488/78/1909-2000$1.00

$\{e_a\}$ are four orthonormal vectors hereafter called tetrads of vectors, which are in general not always remaining the same. The Lie derivative of e_b with respect e_a is

$$\langle e_a, e_b\rangle = \gamma^c_{\cdot ab}\, e_c, \quad \gamma^c_{\cdot ab} = \gamma^c_{\cdot [ab]}.$$

It follows that $\gamma^c_{\cdot ab}$ and $\Gamma^c_{\cdot ab}$ are linearly dependent:

$$\gamma^c_{\cdot ab} = \Gamma^c_{\cdot ab} - \Gamma^c_{\cdot ba}$$

$$\Gamma_{abc} = \tfrac{1}{2}(\gamma_{abc} + \gamma_{cab} - \gamma_{bca}).$$

Now the Einstein field equations (1.1) can be written down in the tetrad form as

$$R_{db} = \partial_d \Gamma^c_{\cdot cb} - \partial_c \Gamma^c_{\cdot db} - \Gamma^c_{\cdot cg}\,\Gamma^g_{\cdot db}\Gamma^g_{\cdot cb}\Gamma^c_{\cdot gd}$$

$$= -\left(\Lambda - \frac{\rho}{2} - \frac{3p}{2}\right) H_{bd} - \left(\Lambda + \frac{\rho}{2} - \frac{p}{2}\right) u_b\, u_d. \tag{1.4}$$

$\partial_{[d}\gamma^f_{\cdot cb]} + \gamma^g_{[dc}\,\gamma^f_{b]g} = 0$ is the Jacobi identity. (1.5)

The tetrads are so chosen that the spacelike vector e_1 is the tachyon fluid vector u_1 so that

$$u^a = \delta^a_1, \quad u_a = \delta^1_a.$$

In a cosmological model filled with tachyonic perfect fluid, the lines of flow are spacelike geodesics and the contracted Bianchi identities are

$$\rho' + (\rho + p)\Theta = 0, \text{ where } \Theta = \theta_2 + \theta_3 + \theta_4. \tag{1.6}$$

Suppose the perturbation of the model results in the formation of momentum flux $\rho + \delta\rho$, so that the ratio of increase in momentum flux to the model is $K = \delta\rho{:}\rho$ and the relative expansion in this region is $-\delta\Theta$. Perturbation of (1.6) gives

$$\partial_1(\delta\rho) + \Theta\delta(\rho + p) + (\rho + p)\delta\Theta = 0. \tag{1.7}$$

Therefore, substituting $p = (\gamma - 1)\rho$, we have

$$\partial_1\left(\frac{\delta\rho}{\rho}\right) = \frac{\partial_1(\delta\rho)}{\rho} - \frac{\delta\rho}{\rho^2}\rho' = -\gamma\cdot\delta\Theta. \tag{1.8}$$

This gives the ratio of growth of K with respect to η in the condensation. Here we have considered the motion along the line $(\theta, \phi, \alpha) = \text{const}$ and have characterized x^i $(i = 1,2,3,4)$ by ψ, θ, ϕ, and α respectively.

2. MOMENTUM FLUX OF THE TACHYON UNIVERSE

Let us consider the model proposed by Gott,

$$ds^2 = \frac{d\eta^2}{1 + 8\pi\rho_0\eta_0^3/3\eta} + \eta^2[\cosh^2\alpha(d\theta^2 + \sin^2\theta d\phi^2 - d\alpha^2], \tag{2.1}$$

where

$$x = \eta\sin\theta\cos\phi\cosh\alpha,$$

$$y = \eta\sin\theta\sin\phi\cosh\alpha,$$

$$z = \eta\cos\theta\cosh\alpha,$$

$$t = \eta\sinh\alpha,$$

as the background model.

Now the metric (2.1) can be re-written as

$$ds_2 = d\psi^2 + \eta^2[\cosh^2\alpha(d\theta^2 + \sin^2\theta d\phi^2) - d\alpha^2], \tag{2.2}$$

where

$$\psi = \int \frac{d\eta}{(1 + 8\pi\rho_0\eta_0^3/3\eta)^{1/2}} = \begin{cases} \eta & \text{when } n \gg \eta_0, \\ \eta^{3/2}/(6\pi\rho_0\eta_0^3)^{1/2} & \text{when } n \ll \eta_0. \end{cases} \tag{2.3}$$

Also in the case of tachyons we can take $t = 0$ which implies $\alpha = 0$ showing that α is the measure of time t in this model.

Further, the nonvanishing tetrad components of the fundamental tensor g_{ij} in the line element (2.2) are given by

$$(e_1{}^1)_{\alpha=0} = 1, \quad (e_2{}^2)_{\alpha=0} = (1/\eta), \quad (e_3{}^3)_{\alpha=0} = (1/\eta)\sin\theta$$

and

$$(e_4{}^4)_{\alpha=0} = (1/\eta).$$

The components of $\gamma^a_{\cdot bc}$ are given as

$$[\theta_1]_{\alpha=0} = -\gamma^1_{\cdot 11} = -\frac{\partial}{\partial\psi}(\log e_1{}^1)$$

$$[\theta_2]_{\alpha=0} = -\gamma^2_{\cdot 12} = -\frac{\partial}{\partial\psi}(\log e_2{}^2)$$

$$[\theta_3]_{\alpha=0} = -\gamma^3_{\cdot 13} = -\frac{\partial}{\partial\psi}(\log e_3{}^3)$$

$$[\theta_4]_{\alpha=0} = -\gamma^4_{\cdot 14} = -\frac{\partial}{\partial\psi}(\log e_4{}^4)$$

and other components of $\gamma^a_{\cdot bc}$ vanish.

The tetrad field equations (1.4) are given as

$$\begin{aligned} D\theta_2 + \theta_2(\theta_2 + \theta_3 + \theta_4) &= -\Lambda + \frac{\rho}{2} - \frac{p}{2}, \\ D\theta_3 + \theta_3(\theta_2 + \theta_3 + \theta_4) &= -\Lambda + \frac{\rho}{2} - \frac{p}{2}, \\ D(\theta_2 + \theta_3 + \theta_4) + (\theta_2^2 + \theta_3^2 + \theta_4^2) &= -\Lambda - \frac{\rho}{2} - \frac{3p}{2}, \\ D\theta_4 + \theta_4(\theta_2 + \theta_3 + \theta_4) &= -\Lambda + \frac{\rho}{2} - \frac{p}{2}, \end{aligned} \tag{2.5}$$

where $D \equiv \partial/\partial\psi$. Further substituting from (2.4), $\theta_2 = \theta_3 = \theta_4 = \theta_0$ (say) and $\Lambda = 0$ in (2.5), we have

$$\begin{aligned} D\theta_0 + 3\theta_0^2 &= \frac{\rho}{2} - \frac{p}{2}, \\ 3D\theta_0 + 3\theta_0^2 &= \frac{\rho}{2} - \frac{p}{2}, \end{aligned} \tag{2.6}$$

Case I: $\eta \gg \eta_0$

The tetrad field equations (2.6) yield

$$\rho = 3/\psi^2 = 3/\eta^2 \tag{2.7}$$

and

$$p = -1/\psi^2 = -1/\eta^2. \tag{2.8}$$

The results (2.7) and (2.8) together imply that inward pressure increases with the expansion of the open universe containing perfect fluid. Also the result (2.8), showing the negative pressure in the model considered, is consistent with its openness.

Case II: $\eta \ll \eta_0$

In this case the model consists of different types of perfect fluids for which pressure is not negative.

1. Model containing tachyon dust or tachyonic neutrinos

Dust is characterized as a pressureless fluid and also neutrinos are collisionless particles. According to kinetic theory of gases, pressure is created due to collision, hence the neutrinos also characterized as pressureless perfect fluid. Hence $p=0$ for tachyon dust or tachyonic neutrinos. Thus in this case field equations (2.6) are

$$\begin{aligned} D\theta_0 + 3\theta_0^2 &= \rho/2, \\ 3D\theta_0 + 3\theta_0^2 &= -\rho/2, \end{aligned} \tag{2.9}$$

which are easily integrable yielding

$$\theta_0 = 2/3\psi \tag{2.10}$$

and

$$\rho = 4/3\psi^2 = 8\pi\rho_0(\eta_0/\eta)^3. \tag{2.11}$$

2. Model containing tachyonic radiation (collision-dominated)

In this case of collision-dominated radiation, $p=\rho/3$. Hence the field equations (2.6) are

$$\begin{aligned} D\theta_0 + 3\theta_0^2 &= \rho/3, \\ 3D\theta_0 + 3\theta_0^2 &= -\rho, \end{aligned} \tag{2.12}$$

which gives on integration

$$\theta_0 = 1/2\psi \tag{2.13}$$

and

$$\rho = 3/4\psi^2 = 9/2\pi\rho_0(\eta_0/\eta)^3. \tag{2.14}$$

3. Mode containing tachyonic fluid in superdense state

The superdense state of fluid is characterized by the equation of state $p=\rho$. Hence in this case field equations (2.6) are rewritten as

$$D\theta_0 + 3\theta_0^2 = 0, \quad 3D\theta_0 + 3\theta_0^2 = -2\rho; \tag{2.15}$$

on solving these equations we have

$$\theta_0 = 1/3\psi \tag{2.16}$$

and

$$\rho = 1/3\psi^2 = 2\pi\rho_0(\eta_0/\eta)^3. \tag{2.17}$$

Thus from the above results we find that momentum flux of the model decreases as η^{-3} with increasing η which measures the proper length in the case of tachyons and also the space–time is singular at $\eta=0$. But at a particular value of η, except $\eta=0$, the momentum flux in the case of pressureless fluid is largest, and, in the case of superdense state, it is the smallest, showing that the tachyonic fluid also does not have a tendency to remain in the superdense state like bradyons in bradyon cosmology.

3. PERTURBATION OF THE MOMENTUM FLUX IN THE MODEL

The spacelike counterpart of Raychaudhuri's equation which gives the field equations in the simplest form is written as

$$\Theta' + 1/3\Theta^2 + 2(\sigma^2 - W^2) + \tfrac{1}{2}(\rho - 3p) + \Lambda = 0. \tag{3.1}$$

For the sake of simplicity we take $W=0$, σ, W and vanishing. Now the field equation (3.1) is reduced to

$$\Theta' + 1/3\Theta^2 + \tfrac{1}{2}(\rho - 3p) = 0. \tag{3.2}$$

After a slight perturbation, Eq. (3.2) becomes

$$D^2\Theta + 2\theta_0 D\Theta + \tfrac{1}{2}(D\rho - 3Dp) = 0. \tag{3.3}$$

Now with the help of (1.8) and $p=(\gamma-1)\rho$, (3.3) is written as

$$D^2K + 2\theta_0 DK - \tfrac{1}{2}\gamma\rho(4-3\gamma)K = 0 \tag{3.4}$$

where $K=\delta\rho/\rho$.

Now we will discuss different cases.

Case I: $\eta \gg \eta_0$

With the help of (2.4), (2.7), (2.8), and $p=(\gamma-1)\rho$, we have the perturbed field equation (3.4) as

$$D^2K + (2/\psi)DK - 2K/\psi^2 = 0. \tag{3.5}$$

Solving this equation we have

$$\begin{aligned} K &= C_1\psi + C_2\psi^{-2} \\ &= C_1\eta + C_2\eta^{-2} \end{aligned} \tag{3.6}$$

The two parts of the solution (3.6) are

$$K_1 = C_1\eta \tag{3.7}$$

and

$$K_2 = C_2\eta^{-2}, \tag{3.8}$$

where C_1 and C_2 are integration constants.

Case II: $\eta \gg \eta_0$

1. Model containing tachyon dust or tachyonic neutrinos

As discussed above in Sec. 2 we have $p=0$ in this case. Hence the perturbed field equation is written as

$$D^2K + (4/3\psi)DK - (2/3\psi^2)\cdot K = 0, \tag{3.9}$$

which on integration gives

$$K=A_d\psi^{2/3}+B_d\psi^{-1}$$
$$=A'_D\eta+B'_d\eta^{-3/2}, \quad (3.10)$$

where $A'_d=A_d/(6\pi\rho_0\eta_0^3)^{1/3}$ and $B'_d=B_d(6\pi\rho_0\eta_0^3)^{1/2}$ are constants. A_d and B_d are integration constants. Thus the two parts of this solution are

$$K_1=A'_d\eta \quad (3.11)$$

and

$$K_2=B'_d\eta^{-3/2}. \quad (3.12)$$

2. Model containing tachyonic radiation (collision-dominated)

Substituting $p=\rho/3$ in Eq. (3.4), we have

$$D^2K+(1/\psi)DK=0, \quad (3.13)$$

which yields

$$K=A_r\log\psi+B_r$$
$$=A'_r\log\eta+B'_r, \quad (3.14)$$

where $A'_r=3/2\,A_r$ and $B'_r=B_r-\frac{1}{2}\log(6\pi\rho_0\eta_0^3)$ are constants. Here A_r and B_r are integration constant.

3. Model containing tachyonic fluid in superdense state

Substituting $p=\rho$ in Eq. (3.4), we have

$$D^2K+(2/3\psi)DK+(2/3\psi^2)K=0, \quad (3.15)$$

which gives on integration

$$K=\psi^{1/6}A_s\cos[(\sqrt{23}/6)\log\psi+B_s]$$
$$=\eta^{1/4}A'_s\cos[(\sqrt{23}/6)\log\eta+B'_s], \quad (3.16)$$

where $A'_s=A_s/(6\pi\rho_0\eta_0^3)^{1/12}$ and $B'_s=B_s(23/8)^{1/2}\times\log(6\rho\rho_0\eta_0^3)$ are constants and A_s and B_s are integration constants.

4. DISCUSSION

From the results obtained in Secs. 2 and 3, we find that the approximation in two cases when $\eta\gg\eta_0$ and when $\eta\ll\eta_0$ leads to a very interesting situation. In the former case results (2.7) and (2.8) imply that the model contains a fluid for which pressure is negative, meaning thereby that pressure increases with increasing spacelike quantity η, i.e., pressure increases with expansion of the model. This shows that pressure is checking the very cause of its increase. But, in this case, the model becomes asymptotically flat; hence this case does not arouse our interest. In the latter case, the situation of pressure of being negative does not arise and some important results are obtained. Results (2.11), (2.14), and (2.17) show that momentum flux decreases as η^{-3} with increasing η. Also we find that ρ is constant at $\eta=\eta_0$ and space–time is singular at $\eta=0$ in both cases. Thus from above it is implied that the entire physical situation of the model changes on crossing the hypersurface $\eta=\eta_0$.

In Sec. 3 on making perturbation of the model also the above approximation yields very important results which show that condensation of tachyons grows as η increases. Equations (3.8), (3.12) are trivial because they do not support condensation of tachyons in the model, which goes against the fact that a tachyon attracts tacyhons.[14] From Eq. (3.7) and (3.11), it is learned that the condensation in the model when $\eta\gg\eta_0$ grows almost in the same way as model contains tachyon dusts or tachyonic neutrinos (when $\eta\ll\eta_0$). From Eq. (3.14), it is evident that when model contains tachyonic radiation (collison-dominated) K grows with η logarithmatically and at $\eta=0$ it is indeterminate. Results (3.16) also shows that growth of condensation with η increasing when the model is filled with tachyons in superdense. The condensation is maximum when

$$\log\eta=(4/\sqrt{23})[\tan^{-1}(4/23)-B'_s].$$

ACKNOWLEDGMENTS

The authors are thankful to Dr. V.B. Johri, Department of Mathematics, University of Gorakhpur for encouragement and important suggestions. They are also thankful to the referee for his valuable comments. Financial supports from U.G.C. and C.S.I.R., New Delhi (India) are also acknowledged.

[1]A. Einstein, Ann. Physik **17**, 891 (1905). An English translation is to be found in Einstein *et al.*, *The Principle of Relativity* (Dover, New York, 1923).
[2]O.M.P. Bilaniuk, V.K. Deshpande, and E.C.G. Sudarshan, Am. J. Phys. **30**, 718 (1962).
[3]G. Feinberg, Phys. Rev. **159**, 1089 (1967).
[4]H. Schmidt, Z. Phys. **151**, 365 (1958).
[5]Ya.P. Terletskii, Sov. Phys. Dokl. **5**, 782 (1960).
[6]S. Tanaka, Prog. Theor. Phys. **24**, 171 (1960).
[7]E. Recami and R. Mignami, Nuovo Cimento **4**, 204 (1974).
[8]A.F. Antippa and A.E. Everett, Phys. Rev. D **4**, 2198–2203 (1971).
[9]A.F. Antippa, Phys. Rev. D **2**, 724–39 (1972).
[10]A.K. Raychaudhuri, J. Math. Phys. **15**, 856 (1974).
[11]J.C. Foster, Jr. and J.R. Roy, J. Math. Phys. **13**, 979 (1972).
[12]J.R. Cott II, Astrophys. J. **187**, 1–3 (1974).
[13]Sushil K. Srivastava and Mahendra P. Pathak, J. Math. Phys. **18**, 483 (1977).
[14]Sushil K. Srivastava, J. Math. Phys. (in press).
[15]K.H. Mariwalla, Am. J. Phys. **37**, 1231 (1969).

Comments on quantum systems subject to random pulses[a)]

M. Blume

Department of Physics, Brookhaven National Laboratory, Upton, New York 11973
and State University of New York, Stony Brook, New York 11794
(Received 28 February 1978)

An error in a recent paper on this subject by Gzyl is corrected. His results are then identical to those obtained earlier by Clauser and Blume.

In a paper in this Journal, Gzyl[1] discussed the Schrödinger equation for a system subject to random pulses. This problem had previously been treated by Clauser and Blume,[2] with the equation

$$i\frac{\partial\psi}{\partial t}=[H_0+\sum_i V_i\delta(t-t_i)]\psi \tag{1}$$

for the wavefunction $\psi(t)$. Here the V_i are random operators, and the times t_i are random instants, assumed Poisson distributed.

Gzyl states that Clauser and Blume integrated this equation incorrectly and that his solution "differs considerably" from theirs. It is the purpose of this note to show that Gzyl uses an unreasonable boundary condition for the discontinuity in the wavefunction at the time of the impulse and that, consequently, his expression is simply the Clauser—Blume result expanded to first order in the operators V_i. He also draws the incorrect conclusion that probability is not conserved when δ-function pulses are applied. Reasonable boundary conditions in fact lead to such conservation as long as the V_i are Hermitian.

Gzyl's error can be seen by considering the equation

$$\frac{d\psi}{dt}=-iV\delta(t-t_0)\psi. \tag{2}$$

Gzyl integrates this from t_0-h to t_0+h to obtain

$$\psi(t_0+)-\psi(t_0-)=-iV\psi(t_0-), \tag{3}$$

which is an improper treatment of $\psi(t_0)$ on the right side of the equation. The correct handling of the singularity can be seen by considering

$$\frac{d\psi}{dt}=-iVf(t)\psi(t), \tag{4}$$

with

$$f(t)=\begin{cases}\epsilon^{-1}, & t_0-\epsilon/2<t<t_0+\epsilon/2,\\ 0, & \text{otherwise},\end{cases}$$

and integrating the equation before letting $\epsilon\to 0$. We find

$$\psi(t)=\exp[-iV\textstyle\int_0^t f(t')dt']\psi(0),$$

and

$$\psi(t_0+h)=\exp[-iV]\psi(t_0-h),$$

with $h>\epsilon/2$, so that, as $\epsilon\to 0$,

$$\psi(t_0+)=\exp[-iV]\psi(t_0-), \tag{5}$$

the result used by Clauser and Blume. Gzyl's boundary condition (3) is the limit of $d\psi/dt=-iVf(t)\,\psi(t-h)$ as $\epsilon\to 0$. This equation is non-Hermitian, even if V is real, and leads to the conclusion that probability is not conserved. Since δ-function pulses are introduced in a model as a limit of pulses with finite width, Eq. (5) yields a physically correct choice of boundary conditions. Gzyl's result (3) follows by expansion of (5) to first order in V. Clauser and Blume correctly treat this singularity and their result [their Eqs. (3)—(6)] shows that with V Hermitian) the *phase* of the wavefunction is discontinuous, while the amplitude is continuous. Thus probability is conserved, i.e., $(d/dt)(|\psi(t)|^2)=0$. If V is non-Hermitian, this is no longer the case. It is the non-Hermitian character of V which simulates the effect of a heat bath. The nonconservation of probability in Gzyl's paper is seen to occur in second order in V and is again a consequence of his first-order expansion in V of the discontinuity in the wavefunctions.

[a)]Work at Brookhaven supported by the U.S. Department of Energy under contract EY-76-C-02-0016; work at State University of New York, Stony Brook supported by the National Science Foundation.

[1]H. Gzyl, J. Math. Phys. **18**, 1327 (1977).
[2]M.J. Clauser and M. Blume, Phys. Rev. B **3**, 583 (1971).

 0022-2488/78/1909-2004$1.00

ERRATA

Erratum: Inverse Gaussian transforms: General properties and application to Slater-type orbitals with noninteger and integer *n* in the coordinate and momentum representations [J. Math. Phys. 19, 52 (1978)]

William J. Taylor

Department of Chemistry, The Ohio State University, Columbus, Ohio 43210
(Received 5 June 1978)

In Eqs. (36b) and (38b) the allowed values of n should read: $n=l,\ l+1, \cdots$.

Erratum: Large time behavior of the superfluorescent decay [J. Math. Phys. 19, 619 (1978)]

V. Benza and E. Montaldi

Universita Degli Studi Di Milano, Instituto di Scienze Fisiche "Aldo Pontremoli," 20133 Milano, Italy
(Received 5 June 1978)

Reference 6 should read: R. Bonifacio, M. Gronchi, L.A. Lugiato, and A.M. Ricca, "Maxwell–Bloch Equations and Mean-Field Theory for Superfluorescence—R. Sanders and R.K. Bullough, Theory of FIR Superfluorescence," in *Cooperative Effects in Matter and Radiation*, edited by C.M. Bowden, D.W. Howgate, and H.R. Robl (Plenum, New York, 1977).

 0022-2488/78/1909-2005$0.00