Representations of a local current algebra: Their dynamical determination*

Ralph Menikoff and David H. Sharp

Theoretical Division, Los Alamos Scientific Laboratory, University of California, Los Alamos, New Mexico 87544
(Received 14 May 1975)

Local currents are used to describe nonrelativistic many-body quantum mechanics in the thermodynamic limit. The problem of determining a representation of the local currents corresponding to a given Hamiltonian is studied. We formulate the dynamics in such a way that one solves simultaneously for the ground state and the representation of the local currents. This leads to two coupled functional equations relating the generating functional to a functional which describes the ground state. Together these functionals determine a representation of the local currents in which the Hamiltonian is a well-defined operator. The functional equations are equivalent to a set of integro–differential equations for expansion coefficients of the two functionals.

1. INTRODUCTION

In this paper we employ the number density of particles, $\rho(\mathbf{x})$, and the particle flux density $\mathbf{J}(\mathbf{x})$ as variables to describe an interacting system of identical, spinless, nonrelativistic particles. We concentrate on systems in the thermodynamic limit, in which the number of particles N and the volume V approach infinity while the average density $\bar{\rho}=N/V$ approaches a finite constant. Our aim is to develop techniques for handling quantum mechanical systems having an infinite number of degrees of freedom. The techniques are developed here using the local currents $\rho(\mathbf{x})$ and $\mathbf{J}(\mathbf{x})$; however, they may find application in other contexts as well.

A quantum mechanical system is associated with a representation of the local currents in a Hilbert space $\mathcal{H}$. When the form of the two-body potential[1] $U(|\mathbf{x}|)$ is specified, a formal expression can be written for the Hamiltonian H governing the dynamics of the system. For systems having a finite number of particles, the currents can be represented on the fixed Hilbert space $\mathcal{H}=L^2(\mathbb{R}^{3N})$, regardless of the choice of $U(|\mathbf{x}|)$. This is a result[2–4] of a theorem due to von Neumann[5]:

Theorem: Let $\{x_j,p_j;j=1,2,\ldots,N\}$ be a set of self-adjoint operators with a common dense domain satisfying the canonical commutation relations

$$[x_j,x_k]=0=[p_j,p_k] \quad \text{and} \quad [x_j,p_k]=i\delta_{jk}.$$

Suppose the exponentiated operators

$$U(\boldsymbol{\alpha})=\exp(i\boldsymbol{\alpha}\cdot\mathbf{x}) \quad \text{and} \quad V(\boldsymbol{\beta})=\exp(i\boldsymbol{\beta}\cdot\mathbf{p})$$

satisfy the multiplication law of the Weyl group:

$$U(\boldsymbol{\alpha})V(\boldsymbol{\beta})=\exp(-i\boldsymbol{\alpha}\cdot\boldsymbol{\beta})V(\boldsymbol{\beta})U(\boldsymbol{\alpha}).$$

Then every representation of the Weyl group is unitarily equivalent to a direct sum of Schrödinger representations. Thus, for systems having a finite number of degrees of freedom, there is a unique representation of the local currents which is determined without solving the dynamics.

This simplicity does not carry over to systems having an infinite number of degrees of freedom. First, the representation of the local currents, like that of the more familiar canonical fields, is not unique.[6] Second, it follows from a nonrelativistic version of Haag's theorem[7] that the Hamiltonian of an interacting system will not make sense as an operator in Fock space, which is used for representing the fields or currents in a free theory. Consequently, we are faced with the problem of finding that representation of the local currents, from among the many unitarily inequivalent non-Fock representations, in which the formal expression for the Hamiltonian associated with a particular interacting system can be interpreted as a self-adjoint operator bounded from below. For this to be possible, the currents must be represented on a Hilbert space which contains the ground state vector of the interacting system. In this way the problem of representing the local currents and the dynamical problem of finding the ground state become interlocked: To find the ground state, one must have a representation of the currents in which the Hamiltonian is a well-defined operator, but to find this representation, one must know the ground state.[8]

In the following paragraphs we will briefly outline our approach to this problem, which is to formulate the dynamics in such a way that one solves simultaneously for the ground state and the representation of the local currents.

Our starting point is the algebra formed by the commutation relations of the local currents $\rho(\mathbf{x})$ and $\mathbf{J}(\mathbf{x})$. Attention is restricted to representations of the current algebra in which the ground state is cyclic for the smeared field $\rho(f)=\int\rho(\mathbf{x})f(\mathbf{x})\,d^3x$, since these correspond to systems of physical interest.[9] Then a representation of the $\rho(f)$'s is determined by the generating functional $L(f)$, the ground state expectation value of $\exp[i\rho(f)]$, while the action of $J(\mathbf{g})=\int\mathbf{J}(\mathbf{x})\cdot\mathbf{g}(\mathbf{x})\,d^3x$ can be specified by a multiplicative functional of $\rho(\mathbf{x})$, $\mathbf{A}(\mathbf{x},\rho)$, defined by[9]

$$\mathbf{K}(\mathbf{x})\Omega=\mathbf{A}(\mathbf{x},\rho)\Omega, \tag{1.1}$$

where

$$\mathbf{K}(\mathbf{x})=\nabla\rho(\mathbf{x})+2i\mathbf{J}(\mathbf{x}) \tag{1.2}$$

and Ω is the ground state.

The functionals $L(f)$ and $\mathbf{A}(\mathbf{x},\rho)$ are determined by two functional differential equations. One equation follows from Eq. (1.1) combined with time reversal invariance and reads[9]

$$[\nabla_{\mathbf{x}}-i\nabla f(\mathbf{x})]\frac{1}{i}\frac{\delta L(f)}{\delta f(\mathbf{x})}=\mathbf{A}\left(\mathbf{x},\frac{1}{i}\frac{\delta}{\delta f}\right)L(f). \tag{1.3}$$

The form of this equation depends only on algebraic properties of the theory. It can be thought of as determining $L(f)$ if $\mathbf{A}(\mathbf{x},\rho)$ is known. For example, it turns out that $\mathbf{A}(\mathbf{x},\rho)\equiv 0$ for a free Bose gas in the thermodynamic limit. In this case it can be shown[10] that $L(f)$ is determined uniquely by the functional equation

$$[\nabla_{\mathbf{x}} - i\nabla f(\mathbf{x})]\frac{1}{i}\frac{\delta L(f)}{\delta f(\mathbf{x})} = 0, \tag{1.4}$$

together with suitable boundary conditions.

In general, however, a second equation is needed which determines $\mathbf{A}(\mathbf{x},\rho)$ when an interaction potential $U(|\mathbf{x}|)$ is specified, and thereby relates the representation to the dynamics. This equation is arrived at as follows. The Hamiltonian of the system can be written formally as[11,12]

$$H = \tfrac{1}{8}\int d^3x\, K_i^{\dagger}(\mathbf{x})\frac{1}{\rho(\mathbf{x})}K_i(\mathbf{x}) + \tfrac{1}{2}\iint d^3x\, d^3y : \rho(\mathbf{x})\rho(\mathbf{y}) : U(|\mathbf{x}-\mathbf{y}|) - E_0\int \rho(\mathbf{x})\, d^3x, \tag{1.5}$$

where E_0 is the ground state energy per particle. Thus we have chosen the zero of energy so that $H\Omega = 0$. In nonrelativistic quantum mechanics we expect both the (kinetic energy/particle) and the (potential energy/particle) to be finite. We believe that subtracting out the ground state energy is the only modification necessary in order for the formal Hamiltonian to be well defined in some representation. A functional equation for $\mathbf{A}(\mathbf{x},\rho)$ is obtained by using Eq. (1.1) to eliminate $\mathbf{K}(\mathbf{x})$ from the equation

$$(\Omega, \exp[i\rho(f)]H\Omega) = 0. \tag{1.6}$$

The resulting equation [Eq. (3.15) in the text] can be thought of as determining $\mathbf{A}(\mathbf{x},\rho)$ if the potential $U(|\mathbf{x}|)$ and $L(f)$ are known. Since this equation for $\mathbf{A}(\mathbf{x},\rho)$ involves $L(f)$, it is a representation-dependent formulation of the operator equation $H\Omega = 0$. Thus it must be solved together with Eq. (1.3) for $L(f)$.

Assuming current conservation and time reversal invariance, one can show that the Hamiltonian for a nonrelativistic system of spinless interacting particles can be written in the general form[9]

$$H = \tfrac{1}{8}\int d^3x\, \tilde{K}_i(\mathbf{x})\frac{1}{\rho(\mathbf{x})}\tilde{K}_i(\mathbf{x}), \tag{1.7}$$

where $\tilde{\mathbf{K}}(\mathbf{x}) = \mathbf{K}(\mathbf{x}) - \mathbf{A}(\mathbf{x},\rho)$ and $\mathbf{A}(\mathbf{x},\rho)$ is the same quantity introduced in Eq. (1.1). Once a functional $\mathbf{A}(\mathbf{x},\rho)$ satisfying (1.3) and (1.6) has been determined, Eq. (1.7) provides a well-defined expression for the Hamiltonian as a Hermitian form in the representation defined by $L(f)$. It is an important characteristic of this approach that the coupled functional equations for $L(f)$ and $\mathbf{A}(\mathbf{x},\rho)$ must be solved simultaneously. We believe (but have not proved) that these equations, when supplemented by appropriate boundary conditions, determine completely a representation of the local currents in which there exists a cyclic vector Ω satisfying the condition $H\Omega = 0$.

Techniques for studying functional differential equations directly are not available except in very special cases. However, by expanding $L(f)$ in a functional power series, whose coefficients are correlation functions, and by expanding $\mathbf{A}(\mathbf{x},\rho)$ in powers of $\rho(\mathbf{x})$, the functional equations can be replaced by two sets of coupled integro—differential equations for the correlation functions and the coefficients in the expansion of $\mathbf{A}(\mathbf{x},\rho)$. The set of equations obtained from the first functional equation, Eq. (1.3), can be looked at as a generalization of the BBGKY hierarchy of equations, familiar from classical statistical mechanics, to the case of many-body potentials. The second set of equations is similar to one derived recently by starting with an extended Jastrow wavefunction for N particles and minimizing the ground state energy.[13] Equations of this kind have found application, for example, in the study of liquid helium.[14] Whether the functional equations are studied directly, or by using correlation function expansions, they will no doubt have to be applied to physical problems in conjunction with a suitable approximation scheme. One such approximation method is discussed in the following paper.[15]

The local ρ, $\mathbf{J}$ current algebra is the same for bosons and fermions. Bosons can nevertheless be distinguished from fermions in that they belong to unitarily inequivalent representations of the current algebra. We believe the equations we have derived for $L(f)$ and $\mathbf{A}(\mathbf{x},\rho)$ to be valid for both Bose and Fermi systems. For example, in one space dimension one can calculate $L(f)$ and $\mathbf{A}(\mathbf{x},\rho)$ explicitly for free bosons and fermions. We have checked in both cases that these quantities satisfy the coupled functional differential equations with $U(|\mathbf{x}|) = 0$. Thus, for a given potential, the functional equations may have more than one solution. This implies that in using the equations, one will have to add constraints that distinguish a Bose solution from a Fermi solution. We believe it is sufficient to impose a constraint on $\mathbf{A}(\mathbf{x},\rho)$ alone. We shall discuss this point briefly in a later section.

The paper is organized as follows. In Secs. 2A and 2B we review the N-particle representations of the current algebra and some previously established properties of the generating functional $L(f)$. Normal ordering for products of $\rho(f)$'s is introduced in Sec. 2C. The properties of the functional $\mathbf{A}(\mathbf{x},\rho)$ are discussed in Sec. 2D and in Appendix A. Section 3 contains the derivation of a functional differential equation for $\mathbf{A}(\mathbf{x},\rho)$. The integro—differential equations relating the correlation functions and the expansion coefficients of $\mathbf{A}(\mathbf{x},\rho)$ are derived in Secs. 4A and 4B. In Sec. 4C we verify that these equations are satisfied for a simple case: that of the N/V limit of a free Fermi gas in one space dimension. In Sec. 4D we discuss the problem of imposing constraints which select Bose from Fermi solutions of the functional equations. Appendix B contains a proof that the N-particle Fermi ground state is cyclic for the $\rho(f)$'s, a result that is used in the discussion. Finally, Sec. 5 discusses an analogy between our formulation of the dynamics of a quantum mechanical many-body system and classical statistical mechanics.

2. PRELIMINARIES

Our starting point is the local current algebra formed

by the commutation relations between $\rho(\mathbf{x})$ and $\mathbf{J}(\mathbf{y})$. In terms of the smeared currents

$$\rho(f)=\int \rho(\mathbf{x})f(\mathbf{x})\,d^3x \tag{2.1}$$

and

$$J(\mathbf{g})=\int \mathbf{J}(\mathbf{x})\cdot\mathbf{g}(\mathbf{x})\,d^3x, \tag{2.2}$$

the commutation relations are given by[11]

$$[\rho(f_1),\rho(f_2)]=0, \tag{2.3}$$

$$[\rho(f),J(\mathbf{g})]=i\rho(\mathbf{g}\cdot\nabla f), \tag{2.4}$$

and

$$[J(\mathbf{g}_1),J(\mathbf{g}_2)]=iJ(\mathbf{g}_2\cdot\nabla\mathbf{g}_1-\mathbf{g}_1\cdot\nabla\mathbf{g}_2). \tag{2.5}$$

We are mainly interested in representations of the currents corresponding to the thermodynamic limit of infinite Bose and Fermi systems. As an aid to understanding these, we first review N-particle representations of the current algebra.

A. N-Particle representations of a local current algebra

Quantum mechanical systems of N identical spinless particles (either bosons or fermions) are described by "N-particle representations" of the current algebra.[3,4] In these representations, the current operators are represented on the fixed Hilbert space:

$$\mathcal{H}=\begin{cases} L^2_+(\mathbb{R}^{3N})=\text{symmetric } L^2 \text{ functions of } N \text{ vector variables for bosons} \\ L^2_-(\mathbb{R}^{3N})=\text{antisymmetric } L^2 \text{ functions of } N \text{ vector variables for fermions.} \end{cases} \tag{2.6}$$

For both bosons and fermions, the action of $\rho(\mathbf{x})$ and $\mathbf{J}(\mathbf{x})$ on $\Psi(\mathbf{x}_1,\ldots,\mathbf{x}_N)\in\mathcal{H}$ is given by

$$[\rho(\mathbf{x})\Psi](\mathbf{x}_1,\ldots,\mathbf{x}_N)=\sum_{m=1}^{N}\delta(\mathbf{x}-\mathbf{x}_m)\Psi(\mathbf{x}_1,\ldots,\mathbf{x}_N) \tag{2.7}$$

and

$$[\mathbf{J}(\mathbf{x})\Psi](\mathbf{x}_1,\ldots,\mathbf{x}_N)=(-i/2)\sum_{m=1}^{N}[2\delta(\mathbf{x}-\mathbf{x}_m)\nabla_{\mathbf{x}_m}-(\nabla\delta)(\mathbf{x}-\mathbf{x}_m)]\Psi(\mathbf{x}_1,\ldots,\mathbf{x}_N). \tag{2.8}$$

Let Ω denote the ground state of the system. If Ω is a well-behaved N-particle wavefunction (i.e., if Ω is not singular and if the set of points on which Ω vanishes is of measure zero), then the ground state is cyclic for the smeared field $\rho(f)$. We assume that the ground state is also cyclic for $\rho(f)$ in representations corresponding to the thermodynamic limit of physical systems.[16] As a result the Hilbert space used for representing the currents is determined by the ground state expectation value of $\exp[i\rho(f)]$.[3] This quantity is called the generating functional for the representation. We denote it by $L(f)$ and write

$$L(f)=(\Omega,\exp[i\rho(f)]\Omega). \tag{2.9}$$

In the next subsection we mention some previously established properties of $L(f)$.

B. The generating functional $L(f)$[9]

The generating functional for an N-particle representation, which we denote by $L_N(f)$, is given by

$$L_N(f)=\int d^3x_1\cdots\int d^3x_N\Omega^*(\mathbf{x}_1,\ldots,\mathbf{x}_N)\times\exp\left(\sum_{j=1}^{N}if(\mathbf{x}_j)\right)\Omega(\mathbf{x}_1,\ldots,\mathbf{x}_N). \tag{2.10}$$

For purposes of considering the thermodynamic limit it is useful to write $L_N(f)$ in the form

$$L_N(f)=\sum_{n=0}^{N}\frac{1}{n!}\int d^3x_1\cdots\int d^3x_N\prod_{j=1}^{n}F(\mathbf{x}_j)R_n^{(N)}(\mathbf{x}_1,\ldots,\mathbf{x}_n), \tag{2.11}$$

where

$$F(\mathbf{x})=\exp[if(\mathbf{x})]-1$$

and the correlation functions $R_n^{(N)}$ are defined by

$$R_n^{(N)}(\mathbf{x}_1,\ldots,\mathbf{x}_n)=\frac{N!}{(N-n)!}\times\int d^3x_{n+1}\cdots\ d^3x_N\,|\Omega(\mathbf{x}_1,\ldots,\mathbf{x}_N)|^2. \tag{2.12}$$

The physical interpretation of the correlation functions is that

$$\frac{1}{n!}R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)=\begin{pmatrix}\text{the probability in the ground state of finding}\\ n \text{ particles at the positions } \mathbf{x}_1,\ldots,\mathbf{x}_n\\ \text{regardless of the positions of the remaining}\\ \text{particles}\end{pmatrix}. \tag{2.13}$$

In the thermodynamic limit, we expect $L_N(f)$ to converge to $L(f)$ and $R_n^{(N)}$ to converge to R_n in such a way that

$$L(f)=\sum_{n=0}^{\infty}\frac{1}{n!}\int d^3x_1\cdots\int d^3x_n\prod_{j_i=1}^{n}F(\mathbf{x}_j)R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n). \tag{2.14}$$

Such convergence has been demonstrated in a few simple cases.[9]

In order to completely determine a representation of the currents, one needs to know how the $J(\mathbf{g})$'s act. To discuss this, it is useful first to define normal ordering for the $\rho(f)$'s.

C. Normal ordering products of the field $\rho(f)$

In an N-particle representation, an expression for the product of n fields $\rho(f)$ can be obtained from Eq. (2.7). One finds

$$\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_n)=\sum_{j_1=1}^{N}\cdots\sum_{j_n=1}^{N}\delta(\mathbf{x}_1-\mathbf{x}_{j_1})\cdots\delta(\mathbf{x}_n-\mathbf{x}_{j_n}). \tag{2.15}$$

We define the normal ordered product of n fields $\rho(f)$ in the same manner except that we require the points $\mathbf{x}_1,\ldots,\mathbf{x}_n$ to be distinct. Thus we write

$$:\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_n):=\sum_{\{j_1\cdots j_n\}}\delta(\mathbf{x}_1-\mathbf{x}_{j_1})\cdots\delta(\mathbf{x}_n-\mathbf{x}_{j_n}), \tag{2.16}$$

where $\sum$ is the sum over the set $\{j_1,\ldots,j_n,\ 1\le j_p\le N,\ j_p\ne j_q \text{ if } p\ne q\}$.

The normal ordered product (2.16) can be expressed as the sum of products of $\rho(f)$'s and δ functions as follows:

$$:\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_n):=\prod_{j=1}^{n}\left(\rho(\mathbf{x}_j)-\sum_{k=1}^{j-1}\delta(\mathbf{x}_j-\mathbf{x}_k)\right). \tag{2.17}$$

By expanding the product, (2.17) can be written as

$$:\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_n):=(-1)^n\sum_G(-1)^m\prod_{j=1}^m(\#j-1)!\,G_j(\rho), \tag{2.18}$$

where G is a partition of $(1,2,\ldots,n)$ into distinct subsets $(G_1,G_2,\ldots,G_m)$. Also, if $G_j=\{1,2,\ldots,k\}$, then $\#j=k$, and $G_j(\rho)=\rho(\mathbf{x}_1)\prod_{p=2}^k\delta(\mathbf{x}_1-\mathbf{x}_p)$. For example,

$$:\rho(\mathbf{x})\rho(\mathbf{y}):=\rho(\mathbf{x})[\rho(\mathbf{y})-\delta(\mathbf{x}-\mathbf{y})]$$

and

$$\begin{aligned}:\rho(\mathbf{x})\rho(\mathbf{y})\rho(\mathbf{z}):=&\rho(\mathbf{x})\rho(\mathbf{y})\rho(\mathbf{z})-\rho(\mathbf{x})\rho(\mathbf{y})\delta(\mathbf{x}-\mathbf{z})\\&-\rho(\mathbf{x})\rho(\mathbf{z})\delta(\mathbf{y}-\mathbf{z})-\rho(\mathbf{y})\rho(\mathbf{z})\delta(\mathbf{x}-\mathbf{y})\\&+2\rho(\mathbf{x})\delta(\mathbf{x}-\mathbf{y})\delta(\mathbf{y}-\mathbf{z}).\end{aligned} \tag{2.19}$$

Equation (2.17) can be inverted. One finds

$$\rho(\mathbf{x}_1)\circ\cdots\circ\rho(\mathbf{x}_n)=\sum_G\prod_{j=1}^m:G_j(\rho):. \tag{2.20}$$

Remarks: (1) When $\rho(\mathbf{x})$ can be written in terms of canonical field operators, obeying either canonical commutation relations or canonical anticommutation relations, as

$$\rho(\mathbf{x})=\psi^\dagger(\mathbf{x})\psi(\mathbf{x}),$$

then

$$:\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_n):=\psi^\dagger(\mathbf{x}_1)\cdots\psi^\dagger(\mathbf{x}_n)\,\psi(\mathbf{x}_n)\cdots\psi(\mathbf{x}_1). \tag{2.21}$$

(2) Since the arguments $\mathbf{x}_1,\ldots,\mathbf{x}_n$ of the currents in a normal ordered product must be distinct, the product $:\rho(\mathbf{x}_1)\circ\cdots\circ\rho(\mathbf{x}_n):$ vanishes when applied to an N-particle state, if $n>N$. That is,

$$:\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_n):\Omega_N=0\quad\text{if } n>N. \tag{2.22}$$

This equation was named an "n-particle identity" in Ref. 4, where it was used to select irreducible representations of the current algebra (2.3)—(2.5) corresponding to systems of n-particles. Normal ordering can be defined for products of the fields $\rho(f)$ and $J(\mathbf{g})$ in a manner similar to that used above. We will not need the resulting formulas, which have been derived by Grodnik,[17] in this paper.

(3) Equations (2.18) and (2.20) resemble cluster expansions, but they are in fact different and should not be confused with them. Cluster functions are defined so that when a pair of points get far apart they vanish. The normal ordered product is defined so that it vanishes when two points coincide.

(4) Equation (2.17) provides a definition of normal ordering which has all the desired properties in each N-particle representation. We take (2.17) as the definition of normal ordering in the thermodynamic limit as well.

The matrix element of a normal ordered product of densities $\rho(f)$ can easily be expressed in terms of correlation functions if one remembers that the points are distinct. For example, using Eq. (2.15), one obtains

$$\begin{aligned}&(\Omega,\exp[i\rho(f)]:\rho(\mathbf{x}_1)\circ\cdots\circ\rho(\mathbf{x}_n):\Omega)\\&=:\frac{1}{i}\frac{\delta}{\delta f(\mathbf{x}_1)}\cdots\frac{1}{i}\frac{\delta}{\delta f(\mathbf{x}_n)}:L(f)\\&=\sum_{m=n}^\infty\frac{1}{(m-n)!}\prod_{k=1}^n\exp[if(\mathbf{x}_k)]\int d^3x_{n+1}\cdots\int d^3x_m\\&\quad\times\prod_{j=n+1}^m F(\mathbf{x}_j)R_m(\mathbf{x}_1,\ldots,\mathbf{x}_m).\end{aligned} \tag{2.23}$$

In particular,

$$(\Omega,:\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_n):\Omega)=R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n). \tag{2.24}$$

Finally, an important property of the normal ordered product which follows directly from Eq. (2.17) is that $:\rho(\mathbf{x})\cdots\rho(\mathbf{z}):$ is proportional to $\rho(\mathbf{x})$.

We now return to explaining how the current $J(\mathbf{g})$ acts in a representation.

D. $J(\mathbf{g})$ and the functional $\mathbf{A}(\mathbf{x},\rho)$

The action of $J(\mathbf{g})$ can be determined by a multiplicative functional $\mathbf{A}(\mathbf{x},\rho)$ defined by the equation

$$\mathbf{K}(\mathbf{x})\Omega=\mathbf{A}(\mathbf{x},\rho)\Omega, \tag{2.25}$$

where $\mathbf{K}(\mathbf{x})=\nabla\rho(\mathbf{x})+2i\mathbf{J}(\mathbf{x})$ and Ω is the ground state. For example, to compute the ground state expectation value of a product of currents which includes a factor $J(\mathbf{g})$, one uses the commutation relations to commute $J(\mathbf{g})$ through the factors of $\rho(f)$ until $J(\mathbf{g})$ acts on the ground state. One then uses Eq. (2.25) to eliminate $J(\mathbf{g})$ from the expectation value. The process is repeated until all factors $J(\mathbf{g})$ have been eliminated. One is left with a matrix element containing only the fields $\rho(f)$. A matrix element of this kind can be calculated from $L(f)$ by taking functional derivatives.

From this point on, we explicitly assume the system is time reversal invariant so that $\Omega=\Omega^*$. Using Eqs. (2.7) and (2.8) to compute the action of $\rho(\mathbf{x})$ and $\mathbf{J}(\mathbf{x})$ on the ground state in an N-particle representation and using the fact that $\mathbf{A}(\mathbf{x},\rho)$ is a multiplicative operator, one finds

$$\mathbf{A}(\mathbf{x},\rho)=\sum_{k=1}^N\delta(\mathbf{x}-\mathbf{x}_k)(\nabla_{\mathbf{x}_k}\ln\Omega^2)(\mathbf{x}_1,\ldots,\mathbf{x}_N). \tag{2.26}$$

Introducing the normal ordered products defined in Eq. (2.16), one can write Eq. (2.26) as

$$\mathbf{A}(\mathbf{x},\rho)=\int d^3x_2\cdots\int d^3x_N:\rho(\mathbf{x})\rho(\mathbf{x}_2)\circ\cdots\circ\rho(\mathbf{x}_N):\nabla_{\mathbf{x}}\ln\Omega^2. \tag{2.27}$$

This expression for $\mathbf{A}(\mathbf{x},\rho)$ is inconvenient for discussing the thermodynamic limit since the number of variables increases with N. We would like to decompose (2.27) into a sum of terms such that the number of variables in each term remains fixed as N increases. That is, we would like to express $\mathbf{A}(\mathbf{x},\rho)$ as a polynomial in $\rho(f)$:

$$\begin{aligned}\mathbf{A}(\mathbf{x},\rho)=\sum_{k=0}^{N-1}\frac{1}{k!}\int d^3y_1\cdots\int d^3y_k:\rho(\mathbf{x})\rho(\mathbf{y}_1)\cdots\rho(\mathbf{y}_k):\\\times\mathbf{A}_{k+1}(\mathbf{x},\mathbf{y}_1,\ldots,\mathbf{y}_k).\end{aligned} \tag{2.28}$$

To be useful in considering the thermodynamic limit, one would like the series (2.28) to converge term by term. For this to happen, the expansion coefficients $\mathbf{A}_{k+1}$ must fall off sufficiently fast for large argument. We believe, but have not proved, that this behavior is a consequence of the cluster decomposition property. In Appendix A we derive Eq. (2.28) and show how a suitable set of expansion coefficients $\mathbf{A}_{k+1}$ can be uniquely defined.

Finally, we note that the expansion (2.28) for $\mathbf{A}(\mathbf{x},\rho)$ is equivalent to writing the N particle ground state as an extended Jastrow wavefunction[13]

$$\Omega=\sum_{n=1}^{N}\left\{\prod_{j_1>j_2>\cdots>j_n}\phi_n(\mathbf{x}_{j_1},\ldots,\mathbf{x}_{j_n})\right\}, \tag{2.29}$$

in which case $\mathbf{A}_{k+1}$ is given by

$$\mathbf{A}_{k+1}(\mathbf{x},\mathbf{y}_1,\ldots,\mathbf{y}_k)=\nabla_{\mathbf{x}}\ln\phi^2_{k+1}(\mathbf{x},\mathbf{y}_1,\ldots,\mathbf{y}_k). \tag{2.30}$$

Thus, for N particles, Eqs. (2.28)–(2.30) provide a correspondence between Ω and $\mathbf{A}(\mathbf{x},\rho)$. In the thermodynamic limit, the ground state may not be expressible as a wavefunction. Nevertheless, $\mathbf{A}(\mathbf{x},\rho)$ exists and has the general properties suggested by Eqs. (2.28)–(2.30). Namely, $\mathbf{A}_k$ can be written as the gradient of a scalar function (thus $\nabla\times\mathbf{A}_k=0$) and each coefficient $\mathbf{A}_{k+1}(\mathbf{x},\mathbf{y}_1,\ldots,\mathbf{y}_k)$ is symmetric in all variables but $\mathbf{x}$.

Remarks: (1) Since the normal ordered product $:\rho(\mathbf{x})\cdots\rho(\mathbf{z}):$ is proportional to $\rho(\mathbf{x})$, it follows from (2.28) that $\mathbf{A}(\mathbf{x},\rho)$ is proportional to $\rho(\mathbf{x})$. Hence we can introduce a functional $\mathcal{A}(\mathbf{x},\rho)$ defined by

$$\mathbf{A}(\mathbf{x},\rho)=\rho(\mathbf{x})\mathcal{A}(\mathbf{x},\rho). \tag{2.31}$$

This fact will be of importance in Sec. 3, where we compute matrix elements of the Hamiltonian.

(2) We wish to emphasize that the existence of the multiplicative operator $\mathbf{A}(\mathbf{x},\rho)$ depends mainly on the property that the ground state is cyclic for $\rho(f)$. This property has been established for both bosons and fermions, and so we expect a functional $\mathbf{A}(\mathbf{x},\rho)$ can be introduced for systems obeying either form of statistics.

(3) In Appendix A, the coefficients $\mathbf{A}_k$ are expressed as gradients of integrals over $\ln\Omega^2$. As a result, the $\mathbf{A}_k$'s are independent of the normalization of Ω. Furthermore, Ω may vanish at some points (as it indeed does for fermions) since logarithmic singularities are integrable.

In the next section we derive an equation which determines the functional $\mathbf{A}(\mathbf{x},\rho)$ associated with a given physical system.

3. AN EQUATION FOR THE FUNCTIONAL $\mathbf{A}(\mathbf{x},\rho)$

To obtain a functional equation for $\mathbf{A}(\mathbf{x},\rho)$, we start with an expression for the Hamiltonian describing a nonrelativistic system of spinless particles. If the particles interact through a central two-body potential $U(|\mathbf{x}|)$, the Hamiltonian can be written formally as[11,12]

$$H=\tfrac{1}{8}\int d^3x K_i^\dagger(\mathbf{x})\frac{1}{\rho(\mathbf{x})}K_i(\mathbf{x}) + \tfrac{1}{2}\int\int d^3x\,d^3y:\rho(\mathbf{x})\rho(\mathbf{y}):U(|\mathbf{x}-\mathbf{y}|)-\left\langle\frac{E}{N}\right\rangle\int d^3x\rho(\mathbf{x}). \tag{3.1}$$

The first term in the above expression has been shown to be the Hamiltonian for a free Bose gas, and thus we interpret it as the kinetic energy. The second term is the potential energy. Normal ordering has been introduced to account only for the potential between distinct pairs of particles. The quantity $\langle E/N\rangle$ is the average ground state energy per particle. The last term in (3.1) subtracts out the ground state energy, thus defining the origin of energy so that the Hamiltonian vanishes on the ground state Ω:

$$H\Omega=0. \tag{3.2}$$

For systems in the thermodynamic limit, the ground state expectation value of each term in the Hamiltonian (3.1) is typically infinite (it may be zero in special cases). However, in an appropriate representation, the three terms add together to form a well-defined operator.

The representation dependence of the Hamiltonian can be illustrated by the following important example. It has been shown that the Hamiltonian for a free Bose gas is[12]

$$H_B=\tfrac{1}{8}\int d^3x K_i^\dagger(\mathbf{x})\frac{1}{\rho(\mathbf{x})}K_i(\mathbf{x}), \tag{3.3}$$

while that for a free Fermi gas is

$$H_F=\tfrac{1}{8}\int d^3x\,K_i^\dagger(\mathbf{x})\frac{1}{\rho(\mathbf{x})}K_i(\mathbf{x})-\left\langle\frac{E}{N}\right\rangle\int d^3x\,\rho(\mathbf{x}). \tag{3.4}$$

Since the ground state energy is infinite, the free Fermi Hamiltonian is not well defined in the representation which makes the free Bose Hamiltonian well defined, and vice versa. As a result, the Bose and Fermi Hamiltonians are associated with unitarity inequivalent representations of the local current algebra.[18]

The desired equation for $\mathbf{A}(\mathbf{x},\rho)$ is obtained by using the formal Hamiltonian (3.1) to write the equation

$$(\Omega,\exp[i\rho(f)]H\Omega)=0 \tag{3.5}$$

in terms of $\mathbf{A}(\mathbf{x},\rho)$, $L(f)$, and $U(|\mathbf{x}|)$.

First we evaluate the kinetic energy term. Using Eq. (2.25), we find

$$\begin{aligned}&(\Omega,\exp[i\rho(f)](\text{kinetic energy})\Omega)\\&=\tfrac{1}{8}\int d^3x\left(\Omega,\exp[i\rho(f)]K_i^\dagger(\mathbf{x})\frac{1}{\rho(\mathbf{x})}K_i(\mathbf{x})\Omega\right)\\&=\tfrac{1}{8}\int d^3x\left(\Omega,\exp[i\rho(f)]K_i^\dagger(\mathbf{x})\frac{1}{\rho(\mathbf{x})}A_i(\mathbf{x},\rho)\Omega\right)\\&=\tfrac{1}{8}\int d^3x\left(\Omega,\exp[i\rho(f)]\left\{2\nabla_i\rho(\mathbf{x})\frac{1}{\rho(\mathbf{x})}A_i(\mathbf{x},\rho)\right.\right.\\&\quad\left.\left.-\left[K_i(\mathbf{x}),\frac{1}{\rho(\mathbf{x})}A_i(\mathbf{x},\rho)\right]-A_i(\mathbf{x},\rho)\frac{1}{\rho(\mathbf{x})}A_i(\mathbf{x},\rho)\right\}\Omega\right).\end{aligned} \tag{3.6}$$

Next, using the commutation relations (2.3) and (2.4), one can show that

$$[\exp[i\rho(f)],K_i^\dagger(\mathbf{x})]=2i\rho(\mathbf{x})\nabla_i f(\mathbf{x})\exp[i\rho(f)]. \tag{3.7}$$

With the help of Eq. (3.7) we can write the kinetic energy term in the alternative form:

$$\begin{aligned}&(\Omega,\exp[i\rho(f)](\text{kinetic energy})\Omega)\\&=\tfrac{1}{8}\int d^3x\left(\Omega,\exp[i\rho(f)]A_i(\mathbf{x},\rho)\right.\\&\quad\left.\times\left\{\frac{1}{\rho(\mathbf{x})}A_i(\mathbf{x},\rho)+2i\nabla_i f(\mathbf{x})\right\}\Omega\right).\end{aligned} \tag{3.8}$$

Combining Eqs. (3.6) and (3.8), we obtain

$(\Omega,\exp[i\rho(f)](\text{kinetic energy})\Omega)$

$$= -\int \tfrac{1}{8} d^3x(\Omega, \exp[i\rho(f)]\{-i\nabla f(\mathbf{x}) \circ \mathbf{A}(\mathbf{x},\rho) + \rho(\mathbf{x})\nabla \cdot \mathcal{A}(\mathbf{x},\rho) + \tfrac{1}{2}[K_i(\mathbf{x}), \mathcal{A}_i(\mathbf{x},\rho)]\}\Omega), \quad (3.9)$$

where

$$\mathbf{A}(\mathbf{x},\rho) = \rho(\mathbf{x})\mathcal{A}(\mathbf{x},\rho). \quad (3.10)$$

The fact that $\mathbf{A}(\mathbf{x},\rho)$ is proportional to $\rho(\mathbf{x})$ was discussed in Sec. 2D. [see Eq. (2.31)].

By using Eq. (1.3), the term in Eq. (3.9) containing $\nabla f(\mathbf{x})$ can be written

$$\int d^3x\, \nabla_i f(\mathbf{x})(\Omega, \exp[i\rho(f)]A_i(\mathbf{x},\rho)\Omega) = \int d^3x \nabla f(\mathbf{x}) \circ (\nabla - i\nabla f)(\mathbf{x})(\Omega, \exp[i\rho(f)]\rho(\mathbf{x})\Omega). \quad (3.11)$$

Again using the commutation relations (2.3) and (2.4), one can show that

$$[K_i(\mathbf{x}), \mathcal{A}_i(\mathbf{x},\rho)] = 2\rho(\mathbf{x})[\nabla\delta/\delta\rho(\mathbf{x})] \cdot \mathcal{A}(\mathbf{x},\rho). \quad (3.12)$$

Remark: Care must be taken in evaluating the expression $\rho(\mathbf{x})[\nabla\delta/\delta\rho(\mathbf{x})] \cdot \mathcal{A}(\mathbf{x},\rho)$, since $\mathcal{A}(\mathbf{x},\rho)$ depends on the parameter $\mathbf{x}$. In particular, $[\nabla\delta/\delta\rho(\mathbf{x})]$ should be treated as one operator, and the expression $[\nabla\delta/\delta\rho(\mathbf{x})] \cdot \mathcal{A}(\mathbf{x},\rho)$ does not have the same meaning as $\nabla \cdot \{[\delta/\delta\rho(\mathbf{x})] \times \mathcal{A}(\mathbf{x},\rho)\}$. For example, let $\mathcal{A}(\mathbf{x},\rho) = \int d^3y\, \rho(\mathbf{y}) f(\mathbf{x},\mathbf{y})$. Then

$$[\nabla\delta/\delta\rho(\mathbf{x})]\, \mathcal{A}(\mathbf{x},\rho) = \int d^3y [\nabla_{\mathbf{x}}\delta(\mathbf{x}-\mathbf{y})] f(\mathbf{x},\mathbf{y}) = \nabla_{\mathbf{y}} f(\mathbf{x},\mathbf{y})\big|_{\mathbf{y}=\mathbf{x}},$$

while

$$\nabla\{[\delta/\delta\rho(\mathbf{x})]\mathcal{A}(\mathbf{x},\rho)\} = \nabla_{\mathbf{x}}[\int d^3y\, \delta(\mathbf{x}-\mathbf{y}) f(\mathbf{x},\mathbf{y})] = \nabla f(\mathbf{x},\mathbf{x}). \quad (3.13)$$

Now, letting

$$B(\mathbf{x},\rho) = \rho(\mathbf{x})\nabla \cdot \mathcal{A}(\mathbf{x},\rho) + \rho(\mathbf{x})[\nabla\delta/\delta\rho(\mathbf{x})] \cdot \mathcal{A}(\mathbf{x},\rho), \quad (3.14)$$

and using Eqs. (3.9)—(3.12), we can write Eq. (3.5) as a functional differential equation:

$$\left\{-\tfrac{1}{8}\int d^3x \left[B\left(\mathbf{x}, \frac{1}{i}\frac{\delta}{\delta f}\right) - i\nabla f \circ (\nabla - i\nabla f)(\mathbf{x}) \frac{1}{i}\frac{\delta}{\delta f(\mathbf{x})}\right] + \tfrac{1}{2}\int d^3x \int d^3y U(|\mathbf{x}-\mathbf{y}|) : \frac{1}{i}\frac{\delta}{\delta f(\mathbf{x})} \frac{1}{i}\frac{\delta}{\delta f(\mathbf{y})} : \right\} L(f)\Big|_0^f = 0, \quad (3.15)$$

where by $L(f)|_0^f$ we mean that, after the functional derivatives have been taken, one takes the difference of the resulting expression evaluated at f and at $f=0$. This has the effect of subtracting out the ground state energy.

We have previously shown[9] that $L(f)$ and $\mathbf{A}(\mathbf{x},\rho)$ also satisfy the equation

$$[\nabla - i\nabla f(\mathbf{x})]\frac{1}{i}\frac{\delta L(f)}{\delta f(\mathbf{x})} = \mathbf{A}\left(\mathbf{x}, \frac{1}{i}\frac{\delta}{\delta f}\right) L(f). \quad (3.16)$$

Equations (3.15) and (3.16) form a system of coupled functional differential equations which we believe can be solved simultaneously for $L(f)$ and $\mathbf{A}(\mathbf{x},\rho)$. However, to determine these quantities uniquely, it is to be expected that these equations will have to be supplemented by appropriate boundary conditions. For a free Bose gas $\mathbf{A}(\mathbf{x},\rho)=0$, Eq. (3.15) is trivially satisfied and Eq. (3.16) reduces to the simple form

$$[\nabla - i\nabla f(\mathbf{x})]\frac{1}{i}\frac{\delta L(f)}{\delta f(\mathbf{x})} = 0. \quad (3.17)$$

In Ref. 10 boundary conditions were identified which determined a unique solution to Eq. (3.17). Corresponding results for the more complicated system (3.15) and (3.16) have not been established.

Since we have defined the Hamiltonian so that $H\Omega = 0$, the question arises as to how the ground state, as opposed to an excited energy eigenstate, is selected. This can be done in the following way. In the thermodynamic limit, it is physically reasonable to require our system to display translation invariance and to satisfy the cluster decomposition property (particles become uncorrelated as their separation becomes large). These requirements can be imposed directly on $L(f)$. Thus we require:

(i) Translation invariance: $L(f_{\mathbf{a}}) = L(f)$, (3.18)

where $f_{\mathbf{a}}(\mathbf{x}) = f(\mathbf{x}-\mathbf{a})$;

(ii) Cluster decomposition: $\lim_{\lambda\to\infty}[L(f+h_{\lambda\mathbf{a}}) - L(f)L(h_{\lambda\mathbf{a}})] = 0$, (3.19)

where $h_{\lambda\mathbf{a}}(\mathbf{x}) = h(\mathbf{x}-\lambda\mathbf{a})$.

It can be shown[9,19] that a system with these properties has a unique translational invariant state, which is the ground state. Thus by imposing conditions (3.18) and (3.19) on $L(f)$, one selects out the solutions of Eqs. (3.15) and (3.16) for which Ω may be identified as the ground state.

In the next section we will use functional power series expansions of $L(f)$ and $\mathbf{A}(\mathbf{x},\rho)$ to study Eqs. (3.15) and (3.16) further.

4. EQUATIONS RESULTING FROM FUNCTIONAL POWER SERIES EXPANSIONS OF $L(f)$ AND $\mathbf{A}(\mathbf{x},\rho)$

So far we have found two coupled functional equations which, when supplemented by appropriate boundary conditions, we believe will determine a representation of the local currents associated with a given interacting system of particles. General techniques for solving such functional equations are not available. However, a number of important properties of these equations can be established by studying solutions which can be expanded in functional power series. Thus in this section we expand the generating functional $L(f)$ in a series whose coefficients are the correlation functions R_n, and we expand $\mathbf{A}(\mathbf{x},\rho)$ in powers of $\rho(\mathbf{x})$ with coefficients $\mathbf{A}_n$. Each functional equation can then be replaced by a set of coupled integro—differential equations relating R_n and $\mathbf{A}_n$. This procedure is analogous to solving a differential equation by means of a power series. Substituting the power series into the differential equation results in a recursion relation for the coefficients. The coupled equations relating R_n and $\mathbf{A}_n$ can be thought of as playing the role of such a recursion relation.

A. Equations for the expansion coefficients R_n and $\mathbf{A}_n$: Expansion of Eq. (3.16)

We first consider Eq. (3.16). Substituting into this equation the functional power series expansion for $L(f)$, Eq. (2.14), and that for $\mathbf{A}(\mathbf{x},\rho)$, Eq. (2.28), one obtains

$$\sum_{n=1}^{\infty} \frac{\exp[if(\mathbf{x}_1)]}{(n-1)!} \int d^3x_2 \cdots \int d^3x_n F(\mathbf{x}_2)$$

$$\times\cdots F(\mathbf{x}_n)\nabla_{\mathbf{x}_1}R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)$$

$$=\sum_{m=1}^{\infty}\frac{1}{(m-1)!}\sum_{n=m}^{\infty}\frac{1}{(n-m)!}\int d^3x_2\cdots\int d^3x_n$$

$$\times\prod_{j=1}^{m}\exp[if(\mathbf{x}_j)]\mathbf{A}_m(\mathbf{x}_1,\ldots,\mathbf{x}_m)\prod_{q=m+1}^{n}F(\mathbf{x}_q)$$

$$\times R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n), \tag{4.1}$$

where

$$F(\mathbf{x})=\exp[if(\mathbf{x})]-1.$$

In order to compare coefficients, the right-hand side must be expressed as a power series in $F(\mathbf{x})$. One can show that

$$\exp[if(\mathbf{x}_2)]\cdots\exp[if(\mathbf{x}_m)]=\sum_{r=0}^{m-1}\sum_{G_r}\left\{\prod_{j\in G_r}F(\mathbf{x}_j)\right\}, \tag{4.2}$$

where G_r is a subset of $\{2,3,\ldots,m\}$ containing r elements. Substituting Eq. (4.2) into Eq. (4.1) and rearranging terms, one finds that

[right hand side of Eq. (4.1)]

$$=\sum_{m=1}^{\infty}\frac{\exp[if(\mathbf{x}_1)]}{(m-1)!}\int d^3x_2\cdots\int d^3x_n\prod_{q=2}^{n}F(\mathbf{x}_q)$$

$$\times\sum_{m=0}^{\infty}\sum_{r=\max(0,m+1-n)}^{m}\frac{1}{r!\,(m-r)!}$$

$$\times\sum_{j_1\neq j_2\cdots\neq j_{m-r}=2}^{n}\int d^3x_{n+1}\cdots\int d^3x_{n+r}$$

$$\times\mathbf{A}_{m+1}(\mathbf{x}_1,\mathbf{x}_{j_1},\ldots,\mathbf{x}_{j_{m-r}},\mathbf{x}_{n+1}\cdots\mathbf{x}_{n+r})R_{n+r}. \tag{4.3}$$

Equation (4.1) is valid for all $\mathbf{x}_1$ and all test functions $f(\mathbf{x})$. Since $F(\mathbf{x})$ can be made arbitrarily small by choosing $f(\mathbf{x})$ sufficiently small, and the integrands are symmetric functions of the variables $\mathbf{x}_2,\ldots,\mathbf{x}_n$, the coefficients of powers of $F(\mathbf{x})$ in Eq. (4.1) can be equated with those of like powers of $F(\mathbf{x})$ in Eq. (4.3). This procedure yields a coupled set of equations relating R_n and $\mathbf{A}_n$:

$$\nabla_{\mathbf{x}_1}R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)=\sum_{m=0}^{\infty}\sum_{r=\max(0,m+1-n)}^{m}\frac{1}{r!\,(m-r)!}$$

$$\times\sum_{j_1\neq j_2\cdots\neq j_{m-r}=2}^{n}\int d^3x_{n+1}\cdots\int d^3x_{n+r}$$

$$\times\mathbf{A}_{m+1}(\mathbf{x}_1,\mathbf{x}_{j_1}\cdots\mathbf{x}_{j_{m-r}},\mathbf{x}_{n+1}\cdots\mathbf{x}_{n+r})R_{n+r},$$

$$n=1,2,\cdots. \tag{4.4}$$

The structure of this system of equations is similar to that of the BBGKY hierarchy familiar from classical statistical mechanics. This analogy is discussed further in Sec. 5.

B. Equations for the expansion coefficients R_n and $\mathbf{A}_n$: Expansion of Eq. (3.15)

Next we consider the functional equation derived from the Hamiltonian, Eq. (3.15). We begin by substituting the expansion for $\mathbf{A}(\mathbf{x},\rho)$, Eq. (2.28), into the expression for $B(\mathbf{x},\rho)$, defined by Eq. (3.14). One finds

$$B(\mathbf{x},\rho)=\sum_{m=1}^{\infty}\frac{1}{(m-1)!}\int d^3x_2\cdots\int d^3x_m:\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_m):$$

$$\times:\nabla_{\mathbf{x}_1}\cdot\mathbf{A}_m(\mathbf{x}_1,\ldots,\mathbf{x}_m)$$

$$+\sum_{m=1}^{\infty}\frac{1}{(m-1)!}\int d^3x_2\cdots\int d^3x_m\rho(\mathbf{x}_1)$$

$$\times\left\{\left[\nabla_{\mathbf{x}_1}+\left(\nabla_{\mathbf{x}_1}\frac{\delta}{\delta\rho(\mathbf{x}_1)}\right)\right]\frac{1}{\rho(\mathbf{x}_1)}:\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_m):\right\}$$

$$\times\mathbf{A}_m(\mathbf{x}_1,\ldots,\mathbf{x}_m). \tag{4.5}$$

The second term in the above equation vanishes. To see this, we use the expression for a normal ordered product of $\rho(\mathbf{x})$'s, Eq. (2.17), to write

$$\left[\nabla_{\mathbf{x}_1}+\left(\nabla_{\mathbf{x}_1}\frac{\delta}{\delta\rho(\mathbf{x}_1)}\right)\right]\frac{1}{\rho(\mathbf{x}_1)}:\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_n):$$

$$=\sum_j[\rho(\mathbf{x}_2)-\delta(\mathbf{x}_2-\mathbf{x}_1)]\cdots\left\{\left[\nabla_{\mathbf{x}_1}+\left(\nabla_{\mathbf{x}_1}\frac{\delta}{\delta\rho(\mathbf{x}_1)}\right)\right]\right.$$

$$\left.\times\left(\rho(\mathbf{x}_j)-\sum_{k=1}^{j-1}\delta(\mathbf{x}_j-\mathbf{x}_k)\right)\right\}\cdots\left(\rho(\mathbf{x}_m)-\sum_{k=1}^{m-1}\delta(\mathbf{x}_m-\mathbf{x}_k)\right). \tag{4.6}$$

This vanishes since

$$\left[\nabla_{\mathbf{x}_1}+\left(\nabla_{\mathbf{x}_1}\frac{\delta}{\delta\rho(\mathbf{x}_1)}\right)\right]\left(\rho(\mathbf{x}_j)-\sum_{k=1}^{j-1}\delta(\mathbf{x}_j-\mathbf{x}_k)\right)=0. \tag{4.7}$$

We can now write Eq. (3.15) in the form

$$\left(\Omega,\exp[i\rho(f)]\left(\int d^3x\tfrac{1}{8}i\nabla f\cdot(\nabla-i\nabla f)(\mathbf{x})\rho(\mathbf{x})\right.\right.$$

$$-\tfrac{1}{8}\sum_{m=1}^{\infty}\frac{1}{(m-1)!}\int d^3x_1\cdots\int d^3x_m:\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_m):$$

$$\times\nabla_{\mathbf{x}_1}\cdot\mathbf{A}_m(\mathbf{x}_1,\ldots,\mathbf{x}_m)+\tfrac{1}{2}\int\hat{d}^3x$$

$$\times\int d^3yU(|\mathbf{x}-\mathbf{y}|):\rho(\mathbf{x})\rho(\mathbf{y}):$$

$$\left.\left.-\int d^3x\langle E/N\rangle\rho(\mathbf{x})\right)\Omega\right)=0. \tag{4.8}$$

It proves convenient to use the relation

$$(\Omega,\exp[i\rho(f)]\int d^3x(-i\nabla f)\cdot(\nabla-i\nabla f)(\mathbf{x})\rho(\mathbf{x})\Omega)$$

$$=(\Omega,\exp[i\rho(f)]\int d^3x(\nabla-i\nabla f)^2(\mathbf{x})\rho(\mathbf{x})\Omega). \tag{4.9}$$

This equation is valid if

$$(\nabla-i\nabla f)(\mathbf{x})(\Omega,\exp[i\rho(f)]\rho(\mathbf{x})\Omega)$$

goes to zero faster than $1/|\mathbf{x}|^2$ as $|\mathbf{x}|\to\infty$. This condition can be related to one on the correlation functions; ∇R_n must go to zero faster than $1/|\mathbf{x}|^2$ as $|\mathbf{x}|\to\infty$. Combined with the cluster decomposition property, this implies that

$$|R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)-\bar{\rho}R_{n-1}(\mathbf{x}_2,\ldots,\mathbf{x}_n)|\to 0 \tag{4.10}$$

faster than $1/|\mathbf{x}_1|$ as $|\mathbf{x}_1|\to\infty$. Physically, the rate at which the cluster limit is approached should be related to the range of the potential. For quantum systems with short range forces, we expect Eq. (4.9) to be valid.

The procedure of the preceding subsection, applied to Eq. (4.8), now leads to the following set of equations relating R_n and $\mathbf{A}_n$:

$$\tfrac{1}{8}\sum_{j=1}^{n}\nabla_{\mathbf{x}_j}^2R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)=-\tfrac{1}{8}\int d^3x_{n+1}\nabla_{\mathbf{x}_{n+1}}^2$$

$$\times[R_{n+1}(\mathbf{x}_1,\ldots,\mathbf{x}_{n+1})-R_1(\mathbf{x}_{n+1})R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)]$$

$$+\sum_{m=1}^{\infty}\Bigg[\sum_{r=\max(m-n,0)}^{m}\Bigg(\frac{1}{r!\,(m-r)!}\int d^3x_{n+1}\cdots\int d^3x_{n+r}$$

$$\times\sum_{j_1\neq j_2\cdots\neq j_{m-r}=1}^{n} U_m(\mathbf{x}_{j_1}\cdots\mathbf{x}_{j_{m-r}},\mathbf{x}_{n+1}\cdots\mathbf{x}_{n+r})$$

$$\times R_{n+r}(\mathbf{x}_1,\ldots,\mathbf{x}_{n+r})\Bigg)-\frac{1}{m!}\int d^3x_{n+1}\cdots\int d^3x_{n+m}$$

$$\times U_m(\mathbf{x}_{n+1},\ldots,\mathbf{x}_{n+m})\,R_m(\mathbf{x}_{n+1},\ldots,\mathbf{x}_{n+m})\,R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)\Bigg],\tag{4.11}$$

where

$$U_2(\mathbf{x},\mathbf{y})=-\tfrac{1}{8}[(\nabla\circ\mathbf{A}_2)(\mathbf{x},\mathbf{y})+(\nabla\circ\mathbf{A}_2)(\mathbf{y},\mathbf{x})]+U(|\mathbf{x}-\mathbf{y}|),\tag{4.12}$$

and, for $n\neq 2$,

$$U_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)=-\tfrac{1}{8}\frac{1}{(n-1)!}\sum_{\text{perm}}(\nabla_{\mathbf{x}_1}\circ\mathbf{A}_n)(\mathbf{x}_{\mathbf{x}_1},\ldots,\mathbf{x}_{\mathbf{x}_n}).\tag{4.13}$$

Equation (4.11) is an identity for $n=0$, since in defining the Hamiltonian we subtracted out the ground state energy. This means that the ground state energy per unit volume, $\langle(E/V)(\mathbf{x})\rangle$, is given by the expression:

$$\langle(E/V)(\mathbf{x})\rangle=-\tfrac{1}{8}\nabla^2R_1(\mathbf{x})+\sum_{n=1}^{\infty}(1/n!)\times\int d^3x_2\cdots\int d^3x_n U_nR_n(\mathbf{x},\mathbf{x}_2,\ldots,\mathbf{x}_n).\tag{4.14}$$

Remarks: (1) To ensure that these equations describe the ground state, we impose translation invariance. This condition implies that $R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)$ and $\mathbf{A}_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)$ are translationally invariant. In particular, $R_1(\mathbf{x})=\bar{\rho}$, the average density, and $\mathbf{A}_1(\mathbf{x})=0$.

(2) The first term on the right-hand side of Eq. (4.11) vanishes upon integrating by parts. In addition, if the system is translationally invariant, $\nabla^2R_1(\mathbf{x})=0$. We have included this term because sometimes one wants to consider the possibility that translational invariance is broken, for example, by an external potential. In such cases the $\nabla^2R_1(\mathbf{x})$ term in Eq. (4.14) affects the energy density, but not the total energy, since $\int d^3x\nabla^2R_1(\mathbf{x})=0$. When translation invariance is broken the ground state must be selected by minimizing the energy/particle, which is

$$\lim_{V\to\infty}(1/V)\int_V d^3x\langle(E/V)(\mathbf{x})\rangle.$$

(3) The coefficients $\mathbf{A}_n$ appear in Eq. (4.11) only in the the form $\nabla\cdot\mathbf{A}_n$. Therefore, even if the R_n were known, one could not solve (4.11) for $\mathbf{A}_n$. However, for an N-particle representation, Eq. (2.30) implies that $\nabla\times\mathbf{A}_n=0$. Thus it is reasonable to impose the condition $\nabla\times\mathbf{A}_n=0$ when considering representations in the thermodynamic limit.

(4) One can generalize the foregoing results so as to include many-body potentials V_n. To do this one adds the term

$$\frac{1}{n!}\int d^3x_1\cdots\int d^3x_n{:}\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_n){:}V_n(\mathbf{x}_1\cdots\mathbf{x}_n)\tag{4.15}$$

to the Hamiltonian. Equation (4.11) remains the same except that U_n is replaced by U_n+V_n.

(5) For an N-particle representation, Eq. (4.11) can be interpreted in terms of wavefunctions. It is equivalent to the equation

$$E_0R_n=(N!/(N-n)!)\int d^3x_{n+1}\cdots\int d^3x_N\Omega H\Omega,\tag{4.16}$$

where

$$H=\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2+\frac{1}{2}\sum_{j\neq k}U(|\mathbf{x}_j-\mathbf{x}_k|)$$

and E_0 is the ground state energy. Here the ground state Ω has been written as an extended Jastrow wavefunction, (2.29) and the coefficients $\mathbf{A}_n$ are defined by Eq. (2.30).

(6) Campbell[13] has studied the ground state of an N-particle system using an extended Jastrow wavefunction. By varying the function ϕ_n [introduced in Eq. (2.29)] to minimize the energy, he derived the equation

$$(\Omega,(H-E_0)[{:}\tilde{\rho}(\mathbf{k}_1)\cdots\tilde{\rho}(\mathbf{k}_n){:}+{:}\tilde{\rho}(-\mathbf{k}_1)\cdots\tilde{\rho}(-\mathbf{k}_n){:}]\Omega)=0,\tag{4.17}$$

where $\tilde{\rho}(\mathbf{k})$ is the Fourier transform of $\rho(\mathbf{x})$. This is similar to Eq. (4.11), which can be written as

$$(\Omega,(H-E_0){:}\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_n){:}\Omega)=0.\tag{4.18}$$

The Hamiltonian is parity ($\mathbf{x}\to-\mathbf{x}$) invariant. We expect the ground state to be even under parity. Under this assumption

$$(\Omega,(H-E_0)[{:}\tilde{\rho}(\mathbf{k}_1)\cdots\tilde{\rho}(\mathbf{k}_n){:}-{:}\tilde{\rho}(-\mathbf{k}_1)\cdots\tilde{\rho}(-\mathbf{k}_n){:}]\Omega)=0$$

by symmetry. Equation (4.17) then implies

$$(\Omega,(H-E_0){:}\tilde{\rho}(\mathbf{k}_1)\cdots\tilde{\rho}(\mathbf{k}_n){:}\Omega)=0.$$

This is just the Fourier transform of Eq. (4.18) and hence is equivalent to it.

(7) If we had not subtracted the ground state energy from the Hamiltonian, then by translation invariance Eq. (4.11) would contain infinite terms of the form

$$\int d^3x_{n+1}\int d^3x_{n+2}U_2(\mathbf{x}_{n+1}-\mathbf{x}_{n+2})R_{n+2}(\mathbf{x}_1,\ldots,\mathbf{x}_{n+2}).\tag{4.19}$$

When the ground state energy is subtracted out, (4.19) becomes

$$\int d^3x_{n+1}\int d^3x_{n+2}U_2(\mathbf{x}_{n+1}-\mathbf{x}_{n+2})\times[R_{n+2}(\mathbf{x}_1,\ldots,\mathbf{x}_{n+2})-R_2(\mathbf{x}_{n+1}-\mathbf{x}_{n+2})R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)].\tag{4.20}$$

By the cluster decomposition property,

$$[R_{n+2}(\mathbf{x}_1,\ldots,\mathbf{x}_{n+2})-R_2(\mathbf{x}_{n+1}-\mathbf{x}_{n+2})R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)]\to 0$$

as $\mathbf{x}_{n+1},\mathbf{x}_{n+2}\to\infty$. This allows the integral to converge.

C. Example: The free Fermi gas in one dimension

As an aid to understanding Eqs. (4.4) and (4.11), we discuss the example of a free Fermi gas in one dimension in this section. For this example it has been shown[18] that

$$A(x,\rho)=2\int dy{:}\rho(x)\rho(y){:}(x-y)^{-1},\tag{4.21}$$

and

$$R_n(x_1,\ldots,x_n)=\det_{n\times n}G(x_j-x_k),\tag{4.22}$$

where

$$G(x)=\sin(\pi\bar{\rho}x)/\pi x\tag{4.23}$$

and $\bar{\rho}$ is the average density. Thus, in this case

$$U(|x|)=0, \quad A_2(x,y)=2(x-y)^{-1} \tag{4.24}$$

and

$$A_n=0, \quad n>2. \tag{4.25}$$

With these simplifications, Eq. (4.4) takes the form

$$(\partial/\partial x_1)R_n(x_1,\ldots,x_n)=\sum_{j=2}^{n}2(x_1-x_j)^{-1}R_n(x_1,\ldots,x_n) + \int dx_{n+1}\,2(x_1-x_{n+1})^{-1} \times R_{n+1}(x_1,\ldots,x_{n+1}). \tag{4.26}$$

This equation has been explicitly verified in Ref. 18.

Introducing (4.24) and (4.25) into Eq. (4.11), one obtains

$$\begin{aligned}&\tfrac{1}{8}\sum_{j=1}^{n}\partial_{x_j}^2 R_n(x_1,\ldots,x_n)\\ &=\tfrac{1}{2}\sum_{j\neq k=1}^{n}U_2(x_j,x_k)R_n(x_1,\ldots,x_n)+\int dx_{n+1}\\ &\times\sum_{j=1}^{n}U_2(x_j,x_{n+1})R_{n+1}(x_1,\ldots,x_n)+\tfrac{1}{2}\int dx_{n+1}\int dx_{n+2}\\ &\times U_2(x_{n+1},x_{n+2})[R_{n+2}(x_1,\ldots,x_{n+2})\\ &-R_2(x_{n+1},x_{n+2})R_n(x_1,\ldots,x_n)],\end{aligned} \tag{4.27}$$

where

$$U_2(x,y)=\tfrac{1}{2}(x-y)^{-2}. \tag{4.28}$$

In particular, for $n=1$, Eq. (4.27) becomes

$$0=\tfrac{1}{2}\int dx_2\,R_2(x_1-x_2)/(x_1-x_2)^2 + \tfrac{1}{4}\int dx_2\int dx_3[R_3(x_1,x_2,x_3)-\bar{\rho}R_2(x_2,x_3)]/(x_2-x_3)^2, \tag{4.29}$$

while for $n=2$ we obtain

$$\begin{aligned}&\tfrac{1}{8}(\partial_{x_1}^2+\partial_{x_2}^2)R_2(x_1,x_2)\\ &=\tfrac{1}{2}R_2(x_1,x_2)/(x_1-x_2)^2+\tfrac{1}{2}\int dx_3R_3(x_1,x_2,x_3)\\ &\times[(x_1-x_3)^{-2}+(x_2-x_3)^{-2}]+\tfrac{1}{4}\int dx_3\int dx_4\\ &\times[R_4(x_1,\ldots,x_4)-R_2(x_3,x_4)R_2(x_1,x_2)]/(x_3-x_4)^2.\end{aligned} \tag{4.30}$$

A laborious calculation shows explicitly that Eqs. (4.29) and (4.30) are satisfied when $R_n(x_1,\ldots,x_n)$ is given by Eq. (4.22), and we expect that (4.27) is valid for all n.

Equations (4.4) and (4.11) have another solution for the potential $U(|x|)=0$, namely

$$A(x,\rho)=0, \quad R_n=\bar{\rho}^n. \tag{4.31}$$

This solution describes a free Bose gas. Thus, for a given potential, we expect Eqs. (4.4) and (4.11) to have two distinct solutions, describing Bose and Fermi systems. This point is discussed further in Sec. 4D.

The example we have considered in this section can be generalized. In Ref. 18, it was indicated that

$$A(x,\rho)=2\lambda\int dy{:}\rho(x)\rho(y){:}(x-y)^{-1} \tag{4.32}$$

for a one-dimensional system interacting via the potential $U(x)=\lambda(\lambda-1)/x^2$, with $\lambda\geq 1$. We expect this system to be described by Eqs. (4.26)–(4.27), if the right-hand side of each equation is multiplied by a factor of λ.

D. Selecting solutions having definite statistics

The example in the previous subsection suggests that Eqs. (4.4) and (4.11) are valid for spinless fermions as well as bosons. In this section we discuss the Fermi case in more detail. First we consider a representation for N particles in a volume V and use the correspondence with wave mechanics. This representation can be defined on the Hilbert space $\mathcal{H}=L_A^2(V^N)$, the square integrable antisymmetric functions of N vector variables. Let $\Omega_F=$ the free Fermi ground state wavefunction. We stress that Ω_F is cyclic for $\rho(f)$. (This is proved in Appendix B). As a result, $\mathbf{A}(\mathbf{x},\rho)$ can be defined for free fermions. Furthermore, the expansion coefficients $\mathbf{A}_n$ defined in Appendix A involve only integrals over $\ln|\Omega_F|^2$. Thus, even though Ω_F vanishes at some points, the functions $\mathbf{A}_n$ exist since logarithmic singularities are integrable. We expect the $\mathbf{A}_n$ computed from Ω_F and the free Fermi correlation functions to satisfy Eqs. (4.4) and (4.11) with $U(|\mathbf{x}|)=0$. (This was established in one space dimension in the preceding subsection. In three space dimensions, the R_n are known,[18] but the coefficients $\mathbf{A}_n$ have not been evaluated explicitly.)

An interacting system of N fermions in a volume V can be represented on the same Hilbert space. Therefore, the ground state for a system interacting through the potential $U(|\mathbf{x}|)$ can be written as

$$\Omega=\Omega_U\Omega_F, \tag{4.33}$$

where Ω_U is a symmetric function and the free Fermi ground state carries the antisymmetry. It is easy to show that

$$\mathbf{K}(\mathbf{x})\Omega=[\mathbf{K}(\mathbf{x})\Omega_U]\Omega_F+\Omega_U[\mathbf{K}(\mathbf{x})\Omega_F]. \tag{4.34}$$

Therefore, one can write

$$\mathbf{A}(\mathbf{x},\rho)=\mathbf{A}_U(\mathbf{x},\rho)+\mathbf{A}_F(\mathbf{x},\rho), \tag{4.35}$$

where $\mathbf{A}_U(\mathbf{x},\rho)$ and $\mathbf{A}_F(\mathbf{x},\rho)$ are both multiplicative operators. One can then expand this $\mathbf{A}(\mathbf{x},\rho)$ in powers of $\rho(f)$ with coefficients

$$\mathbf{A}_n=(\mathbf{A}_U)_n+(\mathbf{A}_F)_n \tag{4.36}$$

These $\mathbf{A}_n$ and the corresponding correlation functions R_n satisfy (4.4) and (4.11). Thus these equations can describe an interacting system of N bosons or fermions.

We expect Eqs. (4.4) and (4.11) to remain true in the thermodynamic limit. Furthermore, we conjecture that they are sufficient to determine a representation of the currents. If they are, then statistics can be selected in the following way. We expect both $(\mathbf{A}_F)_n$ and $(\mathbf{A}_U)_n$ to converge in the thermodynamic limit. The coefficients $(\mathbf{A}_F)_n$ describe the Fermi statistics and the $(\mathbf{A}_U)_n$ arise from the interaction. We define $\mathcal{A}_B$ as the class of $\mathbf{A}(\mathbf{x},\rho)$ corresponding to Bose systems (or to the thermodynamic limit of symmetric wavefunction). At present we do not know how to characterize this class mathematically. The class of functionals $\mathbf{A}(\mathbf{x},\rho)$ corresponding to Fermi systems can then be written in the form

$$\mathcal{A}_F=\mathbf{A}_F(\mathbf{x},\rho)+\mathcal{A}_B, \tag{4.37}$$

where $\mathbf{A}_F(\mathbf{x},\rho)$ is the functional computed for free fermions. By seeking solutions to Eqs. (4.4) and (4.11) for which $\mathbf{A}(\mathbf{x},\rho)$ is restricted to either the class $\mathcal{A}_B$ or

the class $\mathcal{A}_F$, a representation of the currents having either Bose or Fermi statistics can be determined.

5. THE ANALOGY WITH CLASSICAL STATISTICAL MECHANICS

There is an interesting analogy between our presentation of a quantum many-body system in terms of currents and classical statistical mechanics. This analogy is briefly sketched below.

Using the canonical ensemble for a classical statistical system of N particles, correlation functions are defined by

$$R_n^{(N)}(\mathbf{x}_1, \ldots, \mathbf{x}_N) = Z_N^{-1} \frac{1}{(N-n)!} \int d^3x_{n+1} \cdots \times \int d^3x_N \exp(-\beta V_N), \tag{5.1}$$

where Z_N is the partition function,

$$Z_N = (1/N!) \int d^3x_1 \cdots \int d^3x_N \exp(-\beta V_N), \quad \beta = (1/kT), \tag{5.2}$$

and $V_N(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ is the potential for the interaction. A generating functional can be defined by substituting these correlation functions into Eq. (2.11). For a two-body potential,

$$V_N(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \tfrac{1}{2} \sum_{j \neq k}^{N} V(\mathbf{x}_j - \mathbf{x}_k), \tag{5.3}$$

and the correlation functions satisfy the equations

$$-kT\nabla_{\mathbf{x}_1} R_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \sum_{j=2}^{n} \nabla V(\mathbf{x}_1 - \mathbf{x}_j) R_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) + \int d^3x_{n+1} \nabla V(\mathbf{x}_1 - \mathbf{x}_{n+1}) R_{n+1}(\mathbf{x}_1, \ldots, \mathbf{x}_{n+1}). \tag{5.4}$$

These are known as the BBGKY (Bogoliubov–Born–Green–Kirkwood–Yvon) equations. Observe that the ground state correlation functions for a quantum system, Eq. (2.12), and the correlation functions for a classical statistical mechanics system, Eq. (5.1), would be the same if the ground state wavefunction and the classical potential were related by

$$|\Omega|^2 = \exp(-\beta V_N). \tag{5.5}$$

Thus, using an extended Jastrow wavefunction for the ground state, Eq. (2.29), and assuming a classical multibody potential of the form

$$V_N(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \sum_m \frac{1}{m!} \sum_{j_1 \neq j_2 \cdots \neq j_m} V_m(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_m}), \tag{5.6}$$

we find that the BBGKY equations generalize to Eq. (4.4) if we make the identification

$$\mathbf{A}_m = -\beta \nabla V_m. \tag{5.7}$$

We also remark that the BBGKY hierarchy can be summarized in a functional equation for the generating function,[20,21] as in the quantum mechanical case.

To summarize, in order to determine a representation of the current algebra associated with a quantum system, two sets of equations must be solved simultaneously: Eq. (4.11), which determines the coefficients $\mathbf{A}_n$, and Eq. (4.4), which determines the correlation functions R_n. For a classical statistical mechanical system, the coefficients $\mathbf{A}_n$ can be related directly to the potential. Therefore, only Eq. (4.4) need be solved.

The example of Sec. 4C can be used to illustrate this analogy. For a one-dimensional quantum system interacting via the two-body potential

$$U(x) = \lambda(\lambda - 1)/x^2, \quad \lambda \geq 1, \tag{5.8}$$

it has been shown that[18]

$$A(x, \rho) = 2\lambda \int dy : \rho(x)\rho(y) : (x - y)^{-1}. \tag{5.9}$$

By using Eq. (5.7), a classical system with the two-body potential

$$V(x) = -\ln|x|, \tag{5.10}$$

at a temperature

$$kT = (2\lambda)^{-1}, \tag{5.11}$$

would have the same correlation functions.[22]

Such a classical system can be given a simple physical interpretation. It corresponds to a system of identical parallel line charges constrained to lie in a plane, with the total charge balanced to zero by a uniform charge background of opposite polarity. This system is the thermodynamic limit of the Coulomb gas studied by Dyson.[23]

APPENDIX A

For an N-particle representation, $\mathbf{A}(\mathbf{x}, \rho)$ is given by [Eq. (2.27)]

$$\mathbf{A}(\mathbf{x}, \rho) = \int d^3x_2 \cdots \int d^3x_N : \rho(\mathbf{x}_1) \cdots \rho(\mathbf{x}_N) : \nabla_{\mathbf{x}_1} \ln|\Omega|^2, \tag{A1}$$

where Ω is the ground state wavefunction. In considering the thermodynamic limit, it is convenient to express $\mathbf{A}(\mathbf{x}, \rho)$ in the form

$$\mathbf{A}(\mathbf{x}, \rho) = \sum_{n=1}^{N} \frac{1}{(n-1)!} \int d^3x_2 \cdots \times \int d^3x_n : \rho(\mathbf{x}_1) \cdots \rho(\mathbf{x}_n) : \mathbf{A}_n(\mathbf{x}_1, \ldots, \mathbf{x}_n). \tag{A2}$$

In this section we show how the coefficients $\mathbf{A}_n$ can be defined uniquely in terms of $\ln|\Omega|^2$.

From the definition of normal ordering, Eqs. (A1) and and (A2) amount to expanding $\ln|\Omega|^2$ as follows:

$$\nabla_{\mathbf{x}_1} \ln|\Omega|^2(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \sum_j \mathbf{A}_1(\mathbf{x}_j) + \sum_{i<j} \mathbf{A}_2(\mathbf{x}_i, \mathbf{x}_j) + \sum_{i<j<k} \mathbf{A}_3(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) + \cdots. \tag{A3}$$

To evaluate the functions $\mathbf{A}_n(\mathbf{x}_1, \ldots, \mathbf{x}_n)$, we will need to use the following lemma (stated by Campbell[13] and discussed in more detail by Feenberg[14]):

Lemma: Let $F(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ be a symmetric function of N vector variables. Then

$$F(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \sum_{n=0}^{N} \sum_{1 \leq j_1 < j_2 < \cdots < j_n \leq N} F_n(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_n}), \tag{A4}$$

where

$$F_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) = (-1)^n \sum_{m=0}^{n} (-1)^m \sum_{j_1 < j_2 < \cdots < j_m} c_m(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_m}) \tag{A5}$$

and

$$c_m(\mathbf{x}_1, \ldots, \mathbf{x}_m) = \int \frac{d^3x_{m+1}}{V} \cdots \int \frac{d^3x_N}{V} F(\mathbf{x}_1, \ldots, \mathbf{x}_N). \tag{A6}$$

Proof: We introduce Fourier transforms as follows:

$$F(\mathbf{x}_1, \ldots, \mathbf{x}_N) = (1/V)^N \sum_{\mathbf{k}_1} \cdots \sum_{\mathbf{k}_N} \tilde{F}(\mathbf{k}_1, \ldots, \mathbf{k}_N) \times \prod_{j=1}^{N} \exp(i\mathbf{k}_j \cdot \mathbf{x}_j) \tag{A7}$$

and

$$\tilde{F}(\mathbf{k}_1, \ldots, \mathbf{k}_N) = \int_V d^3y_1 \cdots \int_V d^3y_N F(\mathbf{y}_1, \ldots, \mathbf{y}_N) \times \prod_{j=1}^{N} \exp(-i\mathbf{k}_j \cdot \mathbf{y}_j). \tag{A8}$$

We can rearrange the summations in Eq. (A7) in such a way that

$$\sum_{\mathbf{k}_1} \cdots \sum_{\mathbf{k}_N} = \sum_{n=0}^{N} \sum_{1 \le j_1 < j_2 < \cdots < j_n \le N} \left(\sum_{\mathbf{k}_{j_1} \neq 0} \cdots \sum_{\mathbf{k}_{j_n} \neq 0} \right) \Bigg|_{\mathbf{k}_j = 0 \text{ if } j \neq j_1, \ldots, j_n} \tag{A9}$$

Thus,

$$F(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \sum_{n=0}^{N} \sum_{1 \le j_1 < j_2 < \cdots < j_n \le N} F_n(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_n}), \tag{A10}$$

where

$$\begin{aligned} F_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) &= (1/V)^N \sum_{\mathbf{k}_1 \neq 0} \cdots \sum_{\mathbf{k}_n \neq 0} \tilde{F}(\mathbf{k}_1, \ldots, \mathbf{k}_n, 0, \ldots, 0) \\ &\times \prod_{j=1}^{n} \exp(i\mathbf{k}_j \cdot \mathbf{x}_j) \\ &= \int d^3y_1 \cdots \int d^3y_n \prod_{j=1}^{n} \left(\frac{1}{V} \sum_{\mathbf{k}_j \neq 0} \exp[i\mathbf{k}_j \cdot (\mathbf{x}_j - \mathbf{y}_j)] \right) \\ &\times \int \frac{d^3y_{n+1}}{V} \cdots \int \frac{d^3y_N}{V} F(\mathbf{y}_1, \ldots, \mathbf{y}_N). \end{aligned} \tag{A11}$$

Since $(1/V)\sum_{\mathbf{k}\neq 0} \exp[i\mathbf{k} \cdot (\mathbf{x} - \mathbf{y})] = \delta(\mathbf{x} - \mathbf{y}) - (1/V)$, $F_n(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ can be written in the form shown in Eqs. (A5) and (A6).

Finally, applying this lemma to the function $\ln|\Omega|^2(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ and using Eq. (A3), we obtain

$$\mathbf{A}_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \nabla_{\mathbf{x}_1} \left((-1)^n \sum_{m=0}^{n} (-1)^m \sum_{j_1 < j_2 < \cdots < j_m} c_m(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_m}) \right), \tag{A12}$$

with

$$c_m(\mathbf{x}_1, \ldots, \mathbf{x}_m) = \int \frac{d^3x_{m+1}}{V} \cdots \int \frac{d^3x_N}{V} \ln|\Omega|^2(\mathbf{x}_1, \ldots, \mathbf{x}_N). \tag{A13}$$

As an aid to understanding this procedure we consider the example of a system of N particles on a ring of length L interacting via the two body potential

$$U(x) = \lambda(\lambda - 1)[(L/\pi)\sin(\pi x/L)]^{-2}. \tag{A14}$$

Sutherland has shown that the exact ground state for this system is given by[22]

$$\Omega^{(N)}(x_1, \ldots, x_N) = (\text{const}) \prod_{N \ge j > k \ge 1} |\sin[\pi(x_j - x_k)/L]|^{\lambda}. \tag{A15}$$

In accord with Eq. (A13), let

$$c_n^{(N)} = \int_{-L/2}^{L/2} \frac{dx_{n+1}}{L} \cdots \int_{-L/2}^{L/2} \frac{dx_N}{L} \ln|\Omega|^2. \tag{A16}$$

It is not difficult to show that

$$\begin{aligned} c_0^{(N)} &= \ln(\text{const}) + \tfrac{1}{2}N(N-1)c, \\ c_1^{(N)} &= \tfrac{1}{2}N(N-1)c, \end{aligned} \tag{A17}$$

and, for $n \ge 2$,

$$c_n^{(N)} = 2\lambda \sum_{j<k} \ln|\sin[\pi(x_j - x_k)/L]| + \tfrac{1}{2}(N-n)(N+n-1)c,$$

where

$$c = 2\lambda \int_{-L/2}^{L/2} \frac{dx}{L} \ln|\sin(\pi x/L)|. \tag{A18}$$

Furthermore, using Eq. (A12), one finds that

$$\begin{aligned} A_0 &= \partial_{x_1}[\ln(\text{const}) + \tfrac{1}{2}N(N-1)c] = 0, \\ A_2 &= \partial_{x_1}[2\lambda \ln|\sin[\pi(x_1 - x_2)/L]| - c] \\ &= 2\lambda(\pi/L)\cot[\pi(x_1 - x_2)/L], \end{aligned} \tag{A19}$$

and, for $n \neq 0$ or 2,

$$A_n = \partial_{x_1}\{0\} = 0.$$

In the thermodynamic limit, the potential (A14) approaches the limit

$$U(x) = \lambda(\lambda - 1)/x^2, \tag{A20}$$

while

$$A_2(x_1, x_2) \to 2\lambda/(x_1 - x_2) \quad \text{as } N, V \to \infty \text{ and } N/V \to \bar{\rho}, \tag{A21}$$

and, for $n \neq 2$, $A_n = 0$. Thus, $A(x, \rho)$ exists in the thermodynamic limit, even though we do not have a wavefunction in this limit.

APPENDIX B

A system of N fermions in a box of volume V is described in the Hilbert space of square integrable antisymmetric functions of N vector variables, $\mathcal{H} = L_A^2(V^N)$. The free Fermi groundstate is an element of $\mathcal{H}$, and it can be verified directly that it vanishes only on a set of measure zero in this space. In this appendix we show th that any wavefunction $\Omega \in \mathcal{H}$ which vanishes only on a set of measure zero is cyclic for the fields $\rho(f)$.

Let $\psi \in \mathcal{H}$. Formally,

$$\int d^3y_1 \cdots \int d^3y_N (\psi/\Omega)(\mathbf{y}_1, \ldots, \mathbf{y}_N) : \rho(\mathbf{y}_1) \cdots \rho(\mathbf{y}_N) \tag{B1}$$

is the operator of multiplication by $(\psi/\Omega)(\mathbf{x}_1, \ldots, \mathbf{x}_N)$, and

$$\psi = \int d^3y_1 \cdots \int d^3y_N (\psi/\Omega)(\mathbf{y}_1, \ldots, \mathbf{y}_N) : \rho(\mathbf{y}_1) \cdots \rho(\mathbf{y}_N) : \Omega. \tag{B2}$$

Equation (B2) holds if (ψ/Ω) is a valid test function. However, this may not be the case. As a result, we need to show that (ψ/Ω) can be expressed as a limit of test functions and that Eq. (B.2) holds in the limit that the test functions converge to (ψ/Ω).

Let

$$\psi_N(\mathbf{x}) = \begin{cases} \psi & \text{if } |\psi(\mathbf{x})| \leq N \\ 0 & \text{if } |\psi(\mathbf{x})| > N. \end{cases} \tag{B3}$$

By Lebesque's dominated convergence theorem, $\psi_N \to \psi$. Let $V_\epsilon = \{\mathbf{x}; |\Omega(\mathbf{x})| < \epsilon\}$ and let $\mu(V)$ be the Lebesque measure of the set V. Then $\mu(V_\epsilon) \to 0$ as $\epsilon \to 0$ since we have assumed that Ω vanishes only on a set of measure zero. Let

$$\chi_{\epsilon,N}(\mathbf{x}) = \begin{cases} (\psi_N/\Omega)(\mathbf{x}) & \text{for } \mathbf{x} \notin V_\epsilon \\ 0 & \text{for } \mathbf{x} \in V_\epsilon. \end{cases} \tag{B4}$$

The function $\chi_{\epsilon,N}$ is bounded and measurable. As a result it is a valid test function. Choose $\epsilon(N)$ such that

$$N^2 \mu(V_{\epsilon(N)}) \to 0 \quad \text{as } N \to \infty. \tag{B5}$$

We can now show that

$$\chi_N = \int d^3y_1 \cdots \int d^3y_N \chi_{\epsilon(N),N}(\mathbf{y}_1, \ldots, \mathbf{y}_N): \times :\rho(\mathbf{y}_1) \cdots \rho(\mathbf{y}_N):\Omega \to \psi \quad \text{as } N \to \infty$$

as follows. Note that

$$\|\chi_N - \psi\| \leq \|\chi_N - \psi_N\| + \|\psi_N - \psi\|. \tag{B6}$$

Since

$$\|\chi_N - \psi_N\|^2 = \int_{V_\epsilon} d^3x \, |\psi_N(\mathbf{x})|^2 \leq N^2 \mu(V_\epsilon) \to 0, \quad \text{as } N \to \infty \tag{B7}$$

and

$$\|\psi_N - \psi\| \to 0, \quad \text{as } N \to \infty, \tag{B8}$$

it follows that

$$\|\chi_N - \psi\| \to 0, \quad \text{as } N \to \infty. \tag{B9}$$

Hence, any $\psi \in \mathcal{H}$ can be expressed as the limit of a polynomial in the $\rho(f)$'s acting on Ω, and thus Ω is cyclic for the field $\rho(f)$.

*Work supported by the Energy Research and Development Administration.

[1]Throughout this paper we assume that the particles interact through a central two-body potential.

[2]N-particle representations of the currents are discussed in Sec. 2A of this paper and in Ref. 3 and 4.

[3]G. A. Goldin, J. Math. Phys. 12, 462 (1971).

[4]J. Grodnik and D. H. Sharp, Phys. Rev. D 1, 1531 (1970).

[5]J. von Neumann, Math. Ann. 104, 570 (1931). For a textbook discussion of this theorem see M. Reed and B. Simon, *Methods of Modern Mathematical Physics. I. Functional Analysis* (Academic, New York, 1972), p. 275.

[6]R. Haag, "Canonical Commutation Relations in Field Theory and Functional Integration," in *Lectures in Theoretical Physics, Vol. III*, edited by W. E. Brittin, B. W. Downs and J. Downs (Interscience, New York, 1961), pp. 353–381.

[7]R. Haag, Dan. Mat. Fys. Medd. 29, No. 12 (1955). An excellent discussion of this theorem and its implications can be found in: A. S. Wightman, "Relativistic Dynamics of Quantized Fields," in *High Energy Electromagnetic Interactions and Field Theory*, edited by M. Levy (Gordon and Breach, New York, 1967), Sec. VI.

[8]This point is familar to workers in quantum field theory. It is made clearly, for example, in A. S. Wightman, "Progress in the Foundations of Quantum Field Theory," in *Proceedings of the 1967 International Conference on Particles and Fields*, edited by C. R. Hagen, G. Guralnik and V. S. Mathur (Interscience, New York, 1967), pp. 188–218.

[9]R. Menikoff, J. Math. Phys. 15, 1138 (1974).

[10]G. A. Goldin, J. Grodnik, R. T. Powers, and D. H. Sharp, J. Math. Phys. 15, 88 (1974).

[11]R. F. Dashen and D. H. Sharp, Phys. Rev. 165, 1857 (1968). We work in units in which $m = 1$ and $\hbar = 1$. For a definition of $1/\rho(x)$, see Ref. 12.

[12]G. A. Goldin and D. H. Sharp, "Lie Algebras of Local Currents and their Representations," in *Group Representations in Mathematics and Physics, Battelle Seattle 1969 Rencontres*, edited by V. Bargmann (Springer, New York, 1970), pp. 300–11.

[13]C. E. Campbell, Phys. Lett. 44A, 471 (1973).

[14]For example, see: E. Feenberg, Ann. Phys. (N. Y.) 48, 128 (1974); C. Woo, Phys. Rev. A 6, 2312 (1972).

[15]R. Menikoff and D. H. Sharp, J. Math. Phys. 16, xxxx (1975).

[16]Cyclicity for the ρ's has been proven for free bosons in the thermodynamic limit in Ref. 10.

[17]J. Grodnik, Ph. D. thesis, submitted to the University of Pennsylvania (1969) (unpublished).

[18]R. Menikoff, J. Math. Phys. 15, 1394 (1974).

[19]H. Araki, J. Math. Phys. 1, 492 (1960).

[20]A generating functional was introduced into statistical mechanics by N. N. Bogluibov, J. Phys. (U. S. S. R.) 10, 256 (1946). A translation of this paper appears as "Problems of a Dynamical Theory in Statistical Physics," in *Studies in Statistical Mechanics*, edited by V. I. J. de Boer and G. Uhlenbeck (North-Holland, Amsterdam, 1962), pp. 11–50. In the notation of the present paper, Bogoliubov defined the generating functional as

$$L_B(f) = \sum_{n=0} \frac{1}{n!} \int d^3x_1 \cdots \int d^3x_n f(\mathbf{x}_1) \cdots f(\mathbf{x}_n) \bar{\rho}^{-n} R_n(\mathbf{x}_1, \ldots, \mathbf{x}_n).$$

This corresponds to writing

$$L_B(f) = (\Omega, :\exp[\rho(f/\bar{\rho})]:\Omega)$$

in the current algebra case. Note that the normal ordered exponential is no longer a unitary operator.

[21]The similarity between the correlation functions of statistical mechanics and the ground state expectation values of products of field operators has been noted by many authors; for example, M. Green, "Some Applications of the Generating Functional of the Molecular Distribution Functions," in *Lectures in Theoretical Physics, Vol. III*, edited by W. E. Brittin, B. W. Downs and J. Downs (Interscience, New York, 1961), pp. 195–220.

[22]A similar correspondence has been pointed out by B. Sutherland, Phys. Rev. A 4, 2019 (1971) in the case of N particles in a volume V.

[23]F. J. Dyson, J. Math. Phys. 3, 157 (1962).

Approximate representations of a local current algebra*

Ralph Menikoff and David H. Sharp

Theoretical Division, Los Alamos Scientific Laboratory, University of California, Los Alamos, New Mexico 87544
(Received 14 May 1975)

An approximate method for dealing with nonrelativistic many-body quantum systems having short range interactions is developed using local currents. The scheme is based on determining approximate representations of subalgebras of the local currents. This mathematical framework is used to discuss several approximation schemes.

1. INTRODUCTION

In a previous paper[1] we used local currents as variables to describe nonrelativistic quantum systems in the thermodynamic limit. We derived two coupled functional equations which determine simultaneously the representation of the local currents and the ground state associated with a given Hamiltonian. Each functional equation is equivalent to an infinite set of coupled integro–differential equations which form a hierarchy of a type frequently occurring in the study of systems having an infinite number of degrees of freedom. An approximation method is needed which somehow replaces this infinite set of equations by a finite set, which can then be solved. In this paper we outline one method for doing this, which can be interpreted as a scheme for constructing approximate representations of a local current algebra, in a sense to be defined later.

The approximation method to be discussed here is aimed at describing systems with a large number of particles interacting through a potential having a short range and a repulsive core. The first step in the approximation is to restrict the representation of the current algebra to a finite volume V. This results in a Fock representation for the local currents. The number of particles in the volume is not fixed but the average number is determined by the average density. We may then restrict the number of degrees of freedom since physically we expect the probable number of particles in V to be peaked about the average. The second step is to cut off the interaction, which is done by including only the interactions of particles within the volume V. This is motivated by the short range of the potential which, physically, we expect to result in short range correlations and limited clustering of particles. By restricting our attention to a finite volume we are in effect considering a subalgebra of the local currents. Our scheme determines an approximate representation of the subalgebra. By considering larger and larger volumes we can obtain a representation of the current algebra as an "inductive limit" of the approximate representations of the subalgebras.

The approximations we develop are similar to those used by other physicists[2] in studying quantum liquids and gases in which the ground state is expressed as an extended Jastrow wavefunction. The local currents and their representations form a mathematical framework for discussing such approximation schemes, and we also believe that they aid in justifying and understanding the physical nature of these approximations. The setup is also suggestive of other more general approximations. In addition, it may be possible to carry this approach over to relativistic field theory in which similar problems arise as a result of the infinite number of degrees of freedom. It may then be possible to adapt the approximations developed here for many-body systems to particle physics.

2. CURRENT ALGEBRA FRAMEWORK

The appropriate currents for describing nonrelativistic quantum systems are the number density of particles $\rho(\mathbf{x})$ and the particle flux density $\mathbf{J}(\mathbf{x})$. The smeared currents $\rho(f)=\int \rho(\mathbf{x})f(\mathbf{x})\,d^3x$ and $J(\mathbf{g})=\int \mathbf{J}(\mathbf{x})\cdot\mathbf{g}(\mathbf{x})\,d^3x$ form a local current algebra. Representations of this algebra are discussed in detail in Refs. 3–7. The results of our previous paper[1] concerning the Hamiltonian and the dynamical determination of current algebra representations are summarized below.

A. The Hamiltonian and generating functional

The Hamiltonian for a system of identical spinless particles interacting via a two-body potential $U(|\mathbf{x}|)$ is formally given by[8]

$$H=\tfrac{1}{8}\int d^3x\,K_i^\dagger(\mathbf{x})\frac{1}{\rho(\mathbf{x})}K_i(\mathbf{x}) + \tfrac{1}{2}\int d^3x\int d^3y:\rho(\mathbf{x})\rho(\mathbf{y}):U(|\mathbf{x}-\mathbf{y}|), \tag{2.1}$$

where

$$\mathbf{K}(\mathbf{x})=\nabla\rho(\mathbf{x})+2i\mathbf{J}(\mathbf{x}). \tag{2.2}$$

In order for the formal expression for H to exist, an appropriate representation for the currents must be found. In such a representation, the Hamiltonian can be written in a well-defined manner as[5]

$$H=\tfrac{1}{8}\int d^3x\,\tilde{K}_i^\dagger(\mathbf{x})\frac{1}{\rho(\mathbf{x})}\tilde{K}_i(\mathbf{x}), \tag{2.3}$$

where

$$\tilde{\mathbf{K}}(\mathbf{x})=\mathbf{K}(\mathbf{x})-\mathbf{A}(\mathbf{x},\rho)$$

and $\mathbf{A}(\mathbf{x},\rho)$ is a multiplicative functional defined by its action on the ground state Ω, given by

$$\mathbf{K}(\mathbf{x})\Omega=\mathbf{A}(\mathbf{x},\rho)\Omega. \tag{2.4}$$

For later reference, we note that $\mathbf{A}(\mathbf{x},\rho)$ can be expanded as follows:

$$\mathbf{A}(\mathbf{x},\rho)=\sum_{m=0}^{\infty}\frac{1}{m!}\int d^3y_1\cdots\int d^3y_m:\rho(\mathbf{x})\rho(\mathbf{y}_1)\cdots\rho(\mathbf{y}_m):\times\mathbf{A}_{m+1}(\mathbf{x},\mathbf{y}_1,\ldots,\mathbf{y}_m). \tag{2.5}$$

We also remark that the coefficients $\mathbf{A}_n$ can be written in the form

$$\mathbf{A}_{n+1}(\mathbf{x},\mathbf{y}_1,\ldots,\mathbf{y}_n)=\nabla_{\mathbf{x}}\ln\phi_{n+1}^2(\mathbf{x},\mathbf{y}_1,\ldots,\mathbf{y}_n), \tag{2.6}$$

where each ϕ_n is symmetric in its arguments, so that $\nabla\times\mathbf{A}_n=0$.

A representation of the current algebra can be defined by the functional $\mathbf{A}(\mathbf{x},\rho)$ and the generating functional

$$L(f)=(\Omega,\exp[i\rho(f)]\Omega). \tag{2.7}$$

It will also prove useful to have a series expansion for $L(f)$. In terms of ground state correlation functions $R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)$, such an expansion is given by

$$L(f)=\sum_{n=0}^{\infty}\frac{1}{n!}\int d^3x_1\cdots\int d^3x_n\,F(\mathbf{x}_1)\cdots F(\mathbf{x}_n)R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n) \tag{2.8}$$

where

$$F(\mathbf{x})=\exp[if(\mathbf{x})]-1$$

and $R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)$ is the nth ground state correlation function.

The representation of the local currents associated with a given Hamiltonian is determined by two coupled functional equations relating $\mathbf{A}(\mathbf{x},\rho)$ and $L(f)$. When the expansions (2.5) and (2.8) are substituted into these equations, the following set of equations relating the expansion coefficients $\mathbf{A}_n$ and R_n result[1]:

$$\nabla_{\mathbf{x}_1}R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)=\sum_{m=0}^{\infty}\sum_{r=\max(0,m+1-n)}^{m}\frac{1}{r!\,(m-r)!}\times\sum_{j_1\neq j_2\neq\cdots\neq j_{m-r}=2}^{n}\int d^3x_{n+1}\cdots\int d^3x_{n+r}\times\mathbf{A}_{m+1}(\mathbf{x}_1,\mathbf{x}_{j_1}\cdots\mathbf{x}_{j_{m-r}},\mathbf{x}_{n+1}\cdots\mathbf{x}_{n+r})R_{n+r},\quad n=1,2,\cdots, \tag{2.9}$$

and

$$\tfrac{1}{8}\sum_{j=1}^{n}\nabla_{\mathbf{x}_j}^2R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)=-\tfrac{1}{8}\int d^3x_{n+1}\nabla_{\mathbf{x}_{n+1}}^2\{R_{n+1}(\mathbf{x}_1,\ldots,\mathbf{x}_{n+1})-R_1(\mathbf{x}_{n+1})R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)\}+\sum_{m=1}^{\infty}\Bigg[\sum_{r=\max(m-n,0)}^{m}\Bigg(\frac{1}{r!\,(m-r)!}\int d^3x_{n+1}\cdots\int d^3x_{n+r}\times\sum_{j_1\neq j_2\neq\cdots\neq j_{m-r}=1}^{n}U_m(\mathbf{x}_{j_1}\cdots\mathbf{x}_{j_{m-r}},\mathbf{x}_{n+1}\cdots\mathbf{x}_{n+r})\times R_{n+r}(\mathbf{x}_1,\ldots,\mathbf{x}_{n+r})\Bigg)-\frac{1}{m!}\int d^3x_{n+1}\cdots\int d^3n_{n+m}\times U_m(\mathbf{x}_{n+1},\ldots,\mathbf{x}_{n+m})R_m(\mathbf{x}_{n+1},\ldots,\mathbf{x}_{n+m})\times R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)\Bigg] \tag{2.10}$$

where

$$U_2(\mathbf{x},\mathbf{y})=-\tfrac{1}{8}[(\nabla\cdot\mathbf{A}_2)(\mathbf{x},\mathbf{y})+(\nabla\cdot\mathbf{A}_2)(\mathbf{y},\mathbf{x})]+U(|\mathbf{x}-\mathbf{y}|), \tag{2.11}$$

and, for $n\neq 2$,

$$U_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)=-\frac{1}{8}\frac{1}{(n-1)!}\sum_{\text{perm}}(\nabla_1\circ\mathbf{A}_n)(\mathbf{x}_{\pi_1},\ldots,\mathbf{x}_{\pi_n}). \tag{2.12}$$

A quantity of particular physical interest is the ground state energy per unit volume, $\langle(E/V)(\mathbf{x})\rangle$. This is given by

$$\langle(E/V)(\mathbf{x})\rangle=-\tfrac{1}{8}\nabla^2R_1(\mathbf{x})+\sum_{n=1}^{\infty}(1/n!)\times\int d^3x_2\cdots\int d^3x_n U_nR_n(\mathbf{x},\mathbf{x}_2,\ldots,\mathbf{x}_n). \tag{2.13}$$

We remark that in the thermodynamic limit one can ensure that Eqs. (2.9) and (2.10) describe the ground state by imposing translation invariance and the cluster decomposition property.

The main difficulty in solving these equations is that they form an infinite hierarchy in which an equation for one correlation function involves higher correlation functions. In the following, we shall consider how this hierarchy might be truncated in a physically reasonable manner.

B. Restricting the representation to a finite volume

We restrict the representation of the local currents to a volume V by considering the subalgebra of smeared currents with test functions having support in V. If supp $f\subset V$, the expansion for the generating functional, Eq. (2.8), can be written as[5]

$$L(f)=\sum_{n=0}^{\infty}\frac{1}{n!}\int_V d^3x_1\cdots\int_V d^3x_n\times\prod_{j=1}^{n}\exp[if(\mathbf{x}_j)]P_n(V;\mathbf{x}_1,\ldots,\mathbf{x}_n), \tag{2.14}$$

where

$$P_n=\sum_{j=0}^{\infty}\frac{(-1)^j}{j!}\int_V d^3x_{n+1}\cdots\int_V d^3x_{n+j}R_{n+j}(\mathbf{x}_1,\ldots,\mathbf{x}_{n+j}). \tag{2.15}$$

The coefficients P_n have the interpretation that

$$(1/n!)P_n(V;\mathbf{x}_1,\ldots,\mathbf{x}_n)=\begin{pmatrix}\text{the probability in the ground state for finding}\\ n\text{ particles at the points }\mathbf{x}_1,\ldots,\mathbf{x}_n\text{ in the}\\ \text{volume }V\text{ with the remaining particles outside }V\end{pmatrix}.$$

By comparison, the interpretation of the coefficients R_n is that:

$$(1/n!)R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)=\begin{pmatrix}\text{the probability in the ground state for finding }n\\ \text{particles at the points }\mathbf{x}_1,\ldots,\mathbf{x}_n\text{ regardless of}\\ \text{the positions of the remaining particles}\end{pmatrix}.$$

For $\mathbf{x}_1, \ldots, \mathbf{x}_n \in V$, the R_n's can be expressed in terms of the P_n's as follows:

$$R_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \sum_{j=0}^{\infty} \frac{1}{j!} \int_V d^3x_{n+1} \cdots \int_V d^3x_{n+j} P_{n+j}(\mathbf{x}_1, \ldots, \mathbf{x}_{n+j}). \quad (2.16)$$

The restricted representation is the direct sum of N particle representations (Fock space). The number of particles in the volume V is not fixed. The average number of particles is

$$N_{av} = \bar{\rho} V, \quad (2.17)$$

where $\bar{\rho}$ is the average density. For systems having short range forces with a repulsive core we expect short correlation lengths and little clustering together of particles. Consequently, only the coefficients P_n with n near N_{av} should be important. This suggests that we can approximate Eqs. (2.9) and (2.10) taking into account only these P_n's.

C. Truncating the interaction

The R_n's can be written in terms of P_n only if the arguments $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are inside the volume V, Eq. (2.16). As a result, it is necessary to truncate the interaction if one is going to approximate Eqs. (2.9)–(2.10) using the P_n's.

Inside V, the particles interact through the two-body potential $U(|\mathbf{x}|)$. In addition, the potential results in interactions between particles inside V and those outside. Since we are assuming the potential has a short range, these are limited to particles near the boundary. Let us estimate the number of interactions within V, N_I, compared to the number between particles inside V with those outside, N_I'. Let r = the range of the interaction and S = the surface area of V. The number of particles within the interaction range of a given particle is $n \sim \bar{\rho} r^3$. Thus the number of interactions within V is $N_I \sim n\bar{\rho} V$, while the number of interactions between particles inside V and those outside V is $N_I' \sim n\bar{\rho} r S$. Hence the ratio $(N_I'/N_I) \sim (rS/V)$ goes to zero as the volume gets large, provided the volume is not too pathological in shape. As a result, it is a reasonable approximation to ignore the interaction between particles inside and outside of V, provided V is sufficiently large. We can incorporate this approximation into our system of equations by limiting all integrals to the volume V in Eqs. (2.9)–(2.13). Substituting Eq. (2.16) into Eq. (2.9), we then obtain

$$\mathcal{O}_n P_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) = -\sum_{r=1}^{\infty} \frac{1}{r!} \int_V d^3x_{n+1} \cdots \int_V d^3x_{n+r} \mathcal{O}_{n+r} P_{n+r}(\mathbf{x}_1, \ldots, \mathbf{x}_{n+r}), \quad (2.18)$$

where

$$\mathcal{O}_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \left(\nabla_{\mathbf{x}_1} - \sum_{m=1}^{n} \frac{1}{(m-1)!} \sum_{(j; j_1=1)}^{n} \mathbf{A}_m(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_m}) \right) \quad (2.19)$$

and $\sum_{(j; j_1=1)}^{n}$ is the sum over the set of indices

$$\{1 \leq j_1, j_2, \ldots, j_m \leq n; j_p \neq j_q \text{ if } p \neq q \text{ and } j_1 = 1\}.$$

Using Eq. (2.16) together with Eq. (2.10), we also obtain

$$\{Q_n - E_V\} P_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) = -\sum_{r=1}^{\infty} \frac{1}{r!} \int_V d^3x_{n+1} \cdots \int_V d^3x_{n+r} \{Q_{n+r} - E_V\} \times P_{n+r}(\mathbf{x}_1, \ldots, \mathbf{x}_{n+r}), \quad (2.20)$$

where

$$Q_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \left(\tfrac{1}{8} \sum_{j=1}^{n} \nabla_{\mathbf{x}_j}^2 - \sum_{m=1}^{n} \frac{1}{m!} \sum_{(j)}^{n} U_m(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_m}) \right), \quad (2.21)$$

and E_V is the total energy in the volume V, given by

$$E_V = \sum_{n=1}^{\infty} \frac{1}{n!} \int_V d^3x_1 \cdots \int_V d^3x_n Q_n P_n(\mathbf{x}_1, \ldots, \mathbf{x}_n). \quad (2.22)$$

These equations have a simple structure. Each P_m is coupled to the higher P_n's by multiplying by an operator $\mathcal{O}_n$ or Q_n and then integrating out (or averaging over) the extra variables. This suggests that we can interpret the surroundings outside the volume V as a resevoir of particles which provides the source of the coupling among the P_n's. Thus, even though we have truncated the interaction, there remains an influence of the outside surroundings through the particles supplied to the volume V by the reservoir.

D. Analysis of Eq. (2.18)

We can find a simple solution to Eq. (2.18) in the following way. Suppose

$$\left(\nabla_{\mathbf{x}_1} - \sum_{m=1}^{n} \frac{1}{(m-1)!} \sum_{(j; j_1=1)}^{n} \mathbf{A}_m(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_m}) \right) \times P_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) = 0. \quad (2.23)$$

A set of P_n's satisfying Eq. (2.23) will also satisfy Eq. (2.18). Furthermore, the different P_n's in such a set are not independent because they are related to each other through their dependence on the $\mathbf{A}_n$'s. Now, to solve Eq. (2.23), we let[1]

$$\mathbf{A}_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) = 2\nabla_{\mathbf{x}_1} \ln \phi_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) \quad (2.24)$$

and

$$\psi_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) = c_n \prod_{m=1}^{n} \prod_{(j)}^{n} \phi_m(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_m}), \quad (2.25)$$

where c_n is a constant. Then Eq. (2.23) and hence Eq. (2.18) is satisfied if

$$P_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) = |\psi_n(\mathbf{x}_1, \ldots, \mathbf{x}_n)|^2. \quad (2.26)$$

We can justify this form for the P_n's and determine the constants c_n by reference to the thermodynamic limit. Consider a large volume ω containing N particles. We express the ground state of the system as an extended Jastrow wavefunction,

$$\Omega_\omega = (\eta_\omega / \omega^N)^{1/2} \prod_{m=1}^{N} \prod_{(j)}^{N} \phi_m^{(\omega)}(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_m}), \quad (2.27)$$

where η_ω is a normalization constant. The functions $\phi_n^{(\omega)}$ are symmetric and can be uniquely defined by imposing the conditions[1,9]

$$\int_\omega d^3x_n \ln|\phi_n^{(\omega)}(\mathbf{x}_1,\ldots,\mathbf{x}_n)|^2 = 0. \tag{2.28}$$

Here we only consider systems having interactions of short range, r. The condition (2.28) then implies that $\phi_n^{(\omega)}(\mathbf{x}_1,\ldots,\mathbf{x}_n)\approx 1$ when $|\mathbf{x}_j-\mathbf{x}_k|>r$.

We next determine the form of the P_n's for a subvolume V small compared to the volume ω but large compared to the interaction volume r^3. The P_n's are defined by

$$P_n^{(\omega)}(V;\mathbf{x}_1,\ldots,\mathbf{x}_n) = [N!/(N-n)!](\eta_\omega/\omega^N) \times \int_{\omega-V} d^3x_{n+1}\cdots\int_{\omega-V} d^3x_N |\Omega_\omega(\mathbf{x}_1,\ldots,\mathbf{x}_N)|^2. \tag{2.29}$$

We can split Ω_ω into three factors

$$\psi_n(\mathbf{x}_1,\ldots,\mathbf{x}_n) = \prod_{m=1}^{n}\prod_{\{j\}}\phi_m^{(\omega)}(\mathbf{x}_{j_1},\ldots,\mathbf{x}_{j_m}),$$

$$\psi'_{N-n}(\mathbf{x}_{n+1},\ldots,\mathbf{x}_N) = \prod_{m=1}^{N-n}\prod_{\{j;j_k>n\}}^{N}\phi_m^{(\omega)}(\mathbf{x}_{j_1},\ldots,\mathbf{x}_{j_m}), \tag{2.30}$$

and

$$\chi(\mathbf{x}_1,\ldots\mathbf{x}_N) = \prod_{m=1}^{N}\prod_{\{j;j_k\le n, j_{k'}>n\}}^{N}\phi_m^{(\omega)}(\mathbf{x}_{j_1},\ldots,\mathbf{x}_{j_m}).$$

If the variables $\{\mathbf{x}_j; j\le n\}$ lie within the subvolume V and are a distance r or greater away from its boundary, then $\chi\approx 1$ when the variables $\{\mathbf{x}_j; j>n\}$ lie outside of V. The function ψ_n can be considered an "approximate" wavefunction for n particles in the subvolume V. This is a result of the cluster decomposition property; as particles get far apart, their interaction diminishes and they become independent. Consequently, when the n particles in the subvolume V are far from the boundary, the effect of the particles outside can be neglected, and the wavefunction should factorize. Therefore,

$$P_n^{(\omega)}(V;\mathbf{x}_1,\ldots,\mathbf{x}_n) \approx (N!/(N-n)!\,(1/\omega^n)(1-V/\omega)^{N-n}|\psi_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)|^2 \times \int_{\omega-V} d^3x_{n+1}\cdots\int_{\omega-V} d^3x_N[\eta_\omega/(\omega-V)^{N-n}] \times |\psi'_{N-n}(\mathbf{x}_{n+1},\ldots,\mathbf{x}_N)|^2. \tag{2.31}$$

In the thermodynamic limit ($N\to\infty$, $\omega\to\infty$ such that $N/\omega\to\bar\rho$, the average density) we expect

$$\phi_n^{(\omega)}(\mathbf{x}_1,\ldots,\mathbf{x}_n)\to\phi_n(\mathbf{x}_1,\ldots,\mathbf{x}_n) \tag{2.32}$$

and

$$\eta_\omega\to\eta(\bar\rho). \tag{2.33}$$

As a result,

$$P_n^{(\omega)}(\mathbf{x}_1,\ldots,\mathbf{x}_n)\to P_n(\mathbf{x}_1,\ldots,\mathbf{x}_n) \approx \bar\rho^n\exp(-\bar\rho V)\prod_{m=1}^{n}\prod_{\{j\}}\phi_m^2(\mathbf{x}_{j_1},\ldots,\mathbf{x}_{j_m}), \tag{2.34}$$

for $\mathbf{x}_1,\ldots,\mathbf{x}_n$ contained in V and farther from the boundary than a distance r. As V gets larger, we expect the boundary effects to be less important and Eq. (2.34) to become a better approximation for P_n. This should result in the proper normalization for the P_n's, namely,

$$\sum_{n=0}^{\infty}\frac{1}{n!}\int_V d^3x_1\cdots\int_V d^3x_n P_n(V;\mathbf{x}_1,\ldots,\mathbf{x}_n)\to 1,$$

as $V\to\infty$. However, for finite V we must normalize the P_n's. The normalized P_n's are given by

$$P_n(\mathbf{x}_1,\ldots,\mathbf{x}_n) = |c_n|^2\prod_{m=1}^{n}\prod_{\{j\}}\phi_m^2(\mathbf{x}_{j_1},\ldots,\mathbf{x}_{j_m}) \tag{2.35}$$

where

$$|c_n|^2 = \bar\rho^n\exp(-\bar\rho V)\Big/\Big(\sum_{n=0}^{\infty}\frac{1}{n!}\bar\rho^n\exp(-\bar\rho V) \times\int_V d^3x_1\cdots\int_V d^3x_n\prod_{m=1}^{n}\prod_{\{j\}}\phi_m^2(\mathbf{x}_{j_1},\ldots,\mathbf{x}_{j_m})\Big). \tag{2.36}$$

This is the form for the P_n's [Eqs. (2.25) and (2.26)] that we obtained from Eq. (2.23).

E. Analysis of Eq. (2.20)

We have related the functions P_n and $\mathbf{A}_n$ to a set of functions ϕ_n. These are determined by substituting Eqs. (2.24)–(2.26) into Eq. (2.20). This results in the set of equations

$$\psi_n(H_n-E_V)\psi_n = -\sum_{r=1}^{\infty}\frac{1}{r!}\int d^3x_{n+1}\cdots\int d^3x_{n+r}\,\psi_{n+r}(H_{n+r}-E_V)\psi_{n+r}, \tag{2.37}$$

where H_n is the Hamiltonian for n particles interacting via the potential $U(|\mathbf{x}|)$,

$$H_n = -\tfrac{1}{2}\sum_{j=1}^{n}\nabla_{\mathbf{x}_j}^2 + \tfrac{1}{2}\sum_{j\ne k}U(|\mathbf{x}_j-\mathbf{x}_k|), \tag{2.38}$$

and E_V is the ground state energy associated with the volume V,

$$E_V = \sum_{n=1}^{\infty}\frac{1}{n!}\int_V d^3x_1\cdots\int_V d^3x_n\,\psi_n H_n\psi_n. \tag{2.39}$$

The set of equations (2.37) is equivalent to the variational equations

$$\delta E_V/\delta\phi_n = 0. \tag{2.40}$$

Thus Eqs. (2.37) and (2.39) determine a local extremum of the energy and play the same role as Schrödinger's equation. In the thermodynamic limit the ground state can be selected by imposing translation invariance and the cluster decomposition property on the functions R_n. For finite volumes, these conditions are replaced by selecting the P_n's which minimize the energy E_V.

In order to obtain Eq. (2.37) from the variational equations (2.40), we need to impose a boundary condition on the functions ϕ_m. The most reasonable condition to impose is that the normal derivative of ϕ_m vanishes on the boundary of V:

$$\hat n\cdot\nabla_{\mathbf{x}_j}\phi_m(\mathbf{x}_1,\ldots,\mathbf{x}_m)\big|_{\mathbf{x}_j\in\partial V} = 0. \tag{2.41}$$

This is consistent with the condition

$$\phi_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)\to 1 \text{ as } |\mathbf{x}_j-\mathbf{x}_k|\to\infty, \tag{2.42}$$

which we expect to hold for systems with a short range interaction.

For the thermodynamic limit to exist we need to impose the constraints given by Eq. (2.28). We can in-

corporate these constraints into the variational equations using Lagrange multipliers. Thus, instead of E_V, we minimize

$$\tilde{E}_V = E_V - \sum_{n=1}^{\infty} \int_V d^3x_1 \cdots \int_V d^3x_n \times \sum_{\text{perm}} \lambda_n(\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}) \ln|\phi_n(\mathbf{x}_1, \ldots, \mathbf{x}_n)|^2. \quad (2.43)$$

Then, from

$$\delta\tilde{E}_V/\delta\phi_n = 0, \quad (2.44)$$

we obtain

$$\sum_{\text{perm}} \lambda_n(\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}) = \sum_{r=0}^{\infty} \frac{1}{r!} \int_V d^3x_{n+1} \cdots \int_V d^3x_{n+r} \psi_{n+r}(H_{n+r} - E_V)\psi_{n+r}, \quad (2.45)$$

so that Eqs. (2.45) and (2.28) replace Eq. (2.37).

F. Discussion

Since we have cut off the interaction in order to obtain equations to determine the P_n's, we can expect reasonable results only away from the boundary of the volume V. There are several quantities that can be examined to estimate the accuracy of the approximation. The P_n's can be used to calculate $R_1(\mathbf{x})$, for $\mathbf{x} \in V$. By translation invariance, the exact value of $R_1(\mathbf{x})$ is $\bar{\rho}$. Thus we can use the quantity $|R_1(\mathbf{x}) - \bar{\rho}|/\bar{\rho}$ as a measure of the degree of accuracy of the approximation. Similarly translation invariance requires $\mathbf{A}_1(\mathbf{x}) = 0$ and the energy density (2.13) to be constant. These quantities can also be used to test the accuracy of the approximation.

If a knowledge of the representation of the currents over a larger region of space is desired, or if one needs greater accuracy over a given region of space, then a larger volume V can be considered. We expect that, as $V \to \infty$, the equations discussed in this section determine the representation of the currents exactly, at least for systems having short correlation lengths.

3. TRUNCATING THE NUMBER OF DEGREES OF FREEDOM

So far, by truncating the interaction to the volume V we have obtained a set of equations for the ground state probability distributions P_n and the functions $\mathbf{A}_n$ characterizing the ground state. Both sets of functions $\{P_n\}$ and $\{\mathbf{A}_n\}$ can be expressed in terms of a set of functions $\{\phi_n\}$ by means of Eqs. (2.24) and (2.35). These functions are to be determined by minimizing the energy in the volume V, Eq. (2.39), subject to the conditions (2.28) and (2.41). We still have an infinite set of functions to solve for, since there may be an arbitrarily large number of particles in the volume V. However, we expect the probability for finding a large number of particles to be small, and the degrees of freedom they represent to be insignificant. Therefore, as a further approximation we will truncate the P_n's by setting $P_n = 0$ for $|n - N|$ large, where N is the average number of particles, $\bar{\rho}V$. Since, according to Eqs. (2.25)—(2.26), the P_n's can be expressed in terms of ϕ_m's having $m \leq n$, when we truncate the P_n's we are in effect also truncating the ϕ_m's and the $\mathbf{A}_n$'s. This is physically reasonable since the ϕ_n's (or $\mathbf{A}_n$'s) describe correlations in the ground state. If we limit the number of particles in V to N, then there can be at most N-body correlations.

As a result of these approximations the system, restricted to a volume V, is described by a finite set of functions ϕ_n. Variations of this general scheme are considered in the next subsections.

A. Recovering the *N/V* limit

The probability of finding n particles in the volume V is given by

$$\rho_n = \frac{1}{n!} \int_V d^3x_1 \cdots \int_V d^3x_n P_n(\mathbf{x}_1, \ldots, \mathbf{x}_n). \quad (3.1)$$

Using Eq. (2.35), we find that

$$\rho_n \propto \frac{1}{n!} \bar{\rho}^n \exp(-\bar{\rho}V) \int_V d^3x_1 \cdots \int_V d^3x_n \prod_{m=1}^{n} \prod_{\{j\}}^{n} \phi_m^2(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_m}). \quad (3.2)$$

Since $\phi_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) \approx 1$ when $|\mathbf{x}_j - \mathbf{x}_k| \gg r$, we expect the most important ρ_n's to be those for which n is peaked about the average number of particles $N = \bar{\rho}V$, owing to the multiplicative factor $(1/n!)\bar{\rho}^n \exp(-\bar{\rho}V)$.

We now consider the approximation which results from keeping only ρ_N. The energy of the volume V, Eq. (2.39), then reduces to

$$E_V = \frac{\int_V d^3x_1 \cdots \int_V d^3x_N \psi_N H_N \psi_N}{\int_V d^3x_1 \cdots \int_V d^3x_N |\psi_N|^2}, \quad (3.3)$$

where

$$\psi_N(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \prod_{m=1}^{N} \prod_{\{j\}}^{N} \phi_m(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_m}). \quad (3.4)$$

Varying the functions ϕ_m so as to minimize E_V subject to the constraint (2.28) leads to the equations

$$\int_V d^3x_{n+1} \cdots \int_V d^3x_N \psi_N(H_N - E_V)\psi_N = 0, \quad (3.5)$$

for $n = 0, 1, \ldots, N$. This can be written in terms of correlation functions and is equivalent to Eqs. (2.10)—(2.12) when the integrals extend over the volume V, with

$$R_n(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \begin{cases} \dfrac{(N!/(N-n)!)\int_V d^3x_{n+1} \cdots \int_V d^3x_N |\psi_N(\mathbf{x}_1, \ldots, \mathbf{x}_N)|^2}{\int_V d^3x_1 \cdots \int_V d^3x_N |\psi_N(\mathbf{x}_1, \ldots, \mathbf{x}_N)|^2}, \\ \qquad \text{for } 0 \leq n \leq N, \\ 0, \quad \text{for } n > N. \end{cases} \quad (3.6)$$

and

$$\mathbf{A}_n = 2\nabla\phi_n/\phi_n. \quad (3.7)$$

These R_n also satisfy Eq. (2.9) under the same restrictions. By the Rayleigh—Ritz principle, these equations are equivalent to Schrödinger's equation for N particles in a volume V,

$$H_N\psi_N = E_V\psi_N. \quad (3.8)$$

Similar equations have been derived previously by Campbell.[9]

To determine a representation of the currents in all space, we would have to consider larger and larger

volumes V. Thus this approximation just recovers the thermodynamic limit, in which $N \to \infty$ and $V \to \infty$ in such a way that $N/V \to \bar{\rho}$. In the next section we consider additional approximations which simplify this scheme.

B. Approximations fo the Jastrow wavefunction type

Since we are considering potentials with a short range r, for a given average density $\bar{\rho}$ every particle interacts on the average with $N_I \approx \bar{\rho} r^3$ other particles. Thus we expect the correlations to be described fairly well by the ϕ_n's with $n \leq N_I$. This amounts to approximating ψ_N, Eq. (3.4) in the previous section, by

$$\psi_N(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \prod_{m=1}^{N_I} \prod_{\{j\}}^{N} \phi_m(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_m}). \tag{3.9}$$

We again determine the ϕ_m by varying them so as to minimize the energy (3.3). This leads to the set of equations (3.5) with $N \leq N_I$.

In the case when $N_I = 2$, ψ_N is known as a Jastrow wavefunction. These have been used by many physicists in the study of the ground state of liquid helium and other problems.[2] They start with Eq. (2.9) (in all space) which can be written

$$\nabla_{\mathbf{x}_1} R_2(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{A}_2(\mathbf{x}_1, \mathbf{x}_2) R_2(\mathbf{x}_1, \mathbf{x}_2) + \int_{-\infty}^{\infty} \mathbf{A}_2(\mathbf{x}_1, \mathbf{x}_3) R_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)\, d^3x_3 \tag{3.10}$$

and

$$E = \tfrac{1}{2} \int_{-\infty}^{\infty} d^3x_2 [-\tfrac{1}{4} \nabla \cdot \mathbf{A}_2(\mathbf{x}_1, \mathbf{x}_2) + U(|\mathbf{x}_1 - \mathbf{x}_2|)] R_2(\mathbf{x}_1, \mathbf{x}_2), \tag{3.11}$$

where

$$\mathbf{A}_2(\mathbf{x}_1, \mathbf{x}_2) = 2 \nabla_{\mathbf{x}_1} \ln \phi_2(\mathbf{x}_1, \mathbf{x}_2). \tag{3.12}$$

The system of equations is closed by relating R_3 to R_2, for example, by means of the Kirkwood superposition approximation:

$$R_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \bar{\rho}^{-3} R_2(\mathbf{x}_1, \mathbf{x}_2) R_2(\mathbf{x}_2, \mathbf{x}_3) R_2(\mathbf{x}_1, \mathbf{x}_3). \tag{3.13}$$

By substituting Eq. (3.12) into Eq. (3.10) one can solve numerically for R_2 given $\mathbf{A}_2$, and then compute E as a functional of $\mathbf{A}_2$. Then $\mathbf{A}_2$ is varied to minimize E. Recently several physicists[9-11] have considered going beyond this approximation. They use a wavefunction of the form (3.9) with $N_I > 2$. The resulting scheme is known as an extended Jastrow approximation.

The point of view adopted in this paper is quite different. We limit ourselves to a finite volume and truncate the equations by using the P_n's instead of the R_n's, as opposed to working in an infinite volume, keeping only R_2 and relating R_3 to R_2 by the Kirkwood superposition approximation.

C. An approximation scheme allowing for particle fluctuations

For a given volume V we expect to obtain a better approximation by allowing the number of particles to fluctuate about the average value. Since the probability of finding n particles in a volume V is peaked about $N = \bar{\rho} V$, we consider those P_n with n centered about N, i.e., with $N - \Delta \leq n \leq N + \Delta$ for some integer Δ. All the other P_n's are approximated by zero. We also take into account only those ϕ_n with $n \leq N_I = \bar{\rho} r^3$.

Let

$$\psi_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) = [\bar{\rho}^n \exp(-\bar{\rho} V)]^{1/2} \prod_{m=1}^{N_I} \prod_{\{j\}}^{n} \phi_m(\mathbf{x}_1, \ldots, \mathbf{x}_{j_n}). \tag{3.14}$$

Then P_n is given by

$$P_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) = |\psi_n(\mathbf{x}_1, \ldots, \mathbf{x}_n)|^2 \Big/ \sum_{n=N-\Delta}^{n=N+\Delta} (1/n!) \times \int_V d^3x_1 \cdots \int_V d^3x_n |\psi_n(\mathbf{x}_1, \ldots, \mathbf{x}_n)|^2. \tag{3.15}$$

In this approximation, the energy in the volume V is

$$E_V = \frac{\sum_{n=N-\Delta}^{N+\Delta} (1/n!) \int_V d^3x_1 \cdots \int_V d^3x_n\, \psi_n H_n \psi_n}{\sum_{n=N-\Delta}^{N+\Delta} (1/n!) \int_V d^3x_1 \cdots \int_V d^3x_n |\psi_n|^2}. \tag{3.16}$$

Varying ϕ_n, constrained by Eq. (2.28), to minimize E_V now leads to the equations

$$\sum_{\text{perm}} \lambda_n(\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}) = \sum_{r=N-\Delta-n}^{N+\Delta-n} \frac{1}{r!} \int_V d^3x_{n+1} \cdots \int_V d^3x_{n+r} \psi_{n+r} (H_{n+r} - E_V) \psi_{n+r} \tag{3.17}$$

for $n \leq N_I$ and where λ_n is a Lagrange multiplier. This equation, together with Eq. (2.28), can then be used to solve for λ_n and ϕ_n.

D. Discussion

The main point is that approximate information about a small part of an infinite system can be obtained by considering a relatively small number of degrees of freedom. The degree of translational invariance can be used as an indication of the accuracy of the approximation. If more accurate results are required, then either more P_n's and ϕ_n's can be taken into account or a larger volume V may be considered. By allowing the number of particles to fluctuate we expect to get reasonable results with a smaller volume (and hence a fewer number of variables) than would be needed for the Jastrow wave function type approximation discussed in Sec. 3B.

4. APPROXIMATING REPRESENTATIONS OF A LOCAL CURRENT ALGEBRA

The representations of the local ρ, J current algebra provide a mathematical framework for describing nonrelativistic many body systems. For systems having an infinite number of degrees of freedom, the problem of finding a representation of the local currents appropriate to a given interaction is interlocked with that of solving the dynamics. In this section we discuss how the approximation schemes presented in Sec. 3 are related to approximating representations of the local currents. The definition of an approximating representation combines two mathematical concepts: "inductive limit" and "approximating Hilbert spaces."

A. The inductive limit

For nonrelativistic systems of spinless particles, the

representations of the local currents of physical interest are cyclic for the ρ's. Consequently, the Hilbert space for these representations is determined by a representation of $\rho(\mathbf{x})$.

Let $\mathfrak{A}$ be the algebra formed by the smeared currents $\rho(f)$. Let $\{V_n\}$ be a sequence of larger and larger volu volumes, $V_n \supset V_m$ if $n > m$, tending to infinity, $\cup_n V_n = R^3$. Let $\mathfrak{A}_n$ be the subalgebra formed by the smeared currents $\rho(f)$ with the test functions restricted to have support in V_n. Clearly, $\mathfrak{A}_m \subset \mathfrak{A}_n$ if $m < n$. The sequence of subalgebras $\{\mathfrak{A}_n\}$ converges to an algebra, called the inductive limit of the sequence. We now describe the sense in which the inductive limit is just the algebra $\mathfrak{A}$.

Suppose Π is a representation of the algebra $\mathfrak{A}$. It is defined by the generating functional $L(f)$, Eqs. (2.7) and (2.8), which in turn is determined by a set of correlation functions $\{R_n\}$. Subrepresentations Π_n can be constructed by restricting Π to the subalgebra $\mathfrak{A}_n$. The generating functionals for these $L_n(f)$, Eq. (2.14), are determined by a set of probability distribution functions $\{P_m(V_n)\}$.

Conversely, suppose we are given a consistent set of representations, one Π_n for each $\mathfrak{A}_n$. By consistent we mean that $L_n(f) = L_m(f)$ if $n > m$ and $\text{supp} f \subset V_m$. From the probability distributions for the volume V_n we can compute the correlation functions $R_m(\mathbf{x}_1, \ldots, \mathbf{x}_m)$ for $\mathbf{x}_j \in V_n$, Eq. (2.16). The consistency condition insures that the same R_m is obtained from the probability distributions for overlapping volumes. Since the volumes V_n form an increasing sequence tending towards infinity, we can compute the correlation functions in all space. These determine a generating functional $L(f)$ and hence a representation Π of the algebra $\mathfrak{A}$.

Remark: The inductive limit we have defined is similar to one used in classical statistical mechanics.[12] There, a state on a C^*-algebra is considered, rather than a representation of a local current algebra. The inductive limit is then expressed as a weak-$*$ convergent sequence of states on increasing subalgebras. A representation of the local currents can be used to define a C^*-algebra formed by the exponentiated currents[3]; norm closure of $\mathfrak{A} = \text{span}\{\exp[i\rho(f)], \exp[iJ(g)]$; for all test functions f and $g\}$, where the norm on $\mathfrak{A}$ is defined by the representation. The C^*-algebra defined from different representations is not necessarily the same since the norm, and hence the closure may be different. Since systems with different interactions are described by different representations, the C^*-algebra for a given system is not known *a priori*.[13]

B. Generalized inductive limit

In order to obtain a set of equations for the distribution functions $P_n(V_m)$, we had to truncate the interaction to the volume V_m. Therefore, the equations only determine the function P_n approximately. Consequently, in order to determine a representation of the local currents we need to generalize the concept of an inductive limit. Let Π_n be the approximate subrepresentation of the currents for the volume V_n determined by the approximation scheme of Sec. 3. Let $\Pi_{m,n}$ be the subrepresentation formed by restricting Π_n to the subalgebra $\mathfrak{A}_m$ (where $n > m$). We expect the approximate subrepresentations to converge in some sense as $n \to \infty$,

$$\Pi_{m,n} \to \tilde{\Pi}_m, \tag{4.1}$$

since physically the surface effects we neglect become negligible as the volume increases if the correlations are short range. Furthermore, we expect $\tilde{\Pi}_m$ to form a consistent set of subrepresentations and their inductive limit to determine a representation of the current algebra.

The convergence of the subrepresentations $\Pi_{m,n}$ can be defined using their generating functionals $L_{m,n}(f)$. Each generating functional defines a set of probability distribution functions $\{P_{k,n}(V_m),\ k = 1, 2, \cdots\}$. We expect these functions to converge

$$P_{k,n}(V_m) \to \tilde{P}_k(V_m) \quad \text{as } n \to \infty. \tag{4.2}$$

The set of functions $\{\tilde{P}_k(V_m)\}$ then defines a generating functional and hence a representation $\tilde{\Pi}_m$.

Combining these ideas with the inductive limit, we see that the representation of the current algebra can be determined from the approximate subrepresentations Π_n as follows. The probability distributions associated with the subrepresentation Π_n determine an approximate set of correlation functions $R_k^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_k)$ for $\mathbf{x}_j \in V_n$, Eq. (2.16). The same correlation functions are not obtained for overlapping volumes as would be the case for a consistent set of subrepresentations. However, we can form a generalized inductive limit since for a given volume the correlation functions converge; $R_k^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_k) \to R_k(\mathbf{x}_1, \ldots, \mathbf{x}_k)$ for $\mathbf{x}_j \in V_m$ as $n \to \infty$. By considering larger and larger volumes V_m we can obtain the correlation functions in all space. These can be used to define a generating functional and hence a representation of the currents.

C. Approximating Hilbert spaces

It is instructive to consider the convergence of the subrepresentations $\Pi_{m,n} \to \tilde{\Pi}_m$ in terms of approximating Hilbert spaces.[14] These are defined as follows. Let H be a Hilbert space with inner product $(\cdot\,,\cdot)$ and let H_n be a sequence of Hilbert spaces with inner product $(\cdot\,,\cdot)_n$ together with the linear maps T_n from H to H_n. The spaces H_n approximate H if

(i) $\|T_n\| \leq \text{const}$

and

(ii) $\lim_{n\to\infty} (T_n\psi, T_n\psi)_n = (\psi, \psi)$, for all $\psi \in H$.

We say a sequence of vectors $\{\psi_n \in H_n\}$ converges to $\psi \in H$ if

$$\lim_{n\to\infty} \|\psi_n - T_n\psi\|_n = 0, \tag{4.3}$$

and we say a sequence of self-adjoint operators ρ_n on H_n converges to a self-adjoint operator ρ on H, if for all $s \in R$ and $\psi \in H$

$$\lim_{n\to\infty} \|\exp(is\rho_n)T_n\psi - T_n\exp(is\rho)\psi\|_n = 0. \tag{4.4}$$

The cyclic representations $\Pi_{m,n}$ of the subalgebra $\mathfrak{A}_m$ are defined on the vector spaces $H_{m,n} = \{\text{span } \rho(f_1) \cdots \rho(f_k)\Omega_{m,n}$; supp $f_j \in V_m$ with $\Omega_{m,n}$ a cyclic vector (ground state)$\}$ together with an inner product which is defined by the generating functional $L_{m,n}(f)$.[15]

Let us define the linear mapping $T_n: H_{m,n} \to H_m$ by

$$T_n\rho(f_1)\cdots\rho(f_k)\Omega_{m,n}=\rho(f_1)\cdots\rho(f_k)\Omega_m. \tag{4.5}$$

Under these mappings the spaces $\mathcal{H}_{m,n}$ approximate $\mathcal{H}_m$. This concept can be used as the definition of convergence of the subrepresentations $\Pi_{m,n}\to\tilde{\Pi}_m$.

D. Summary

We believe the concept of approximating representations may be useful in other contexts. We conclude by stating an abstract definition of this concept. By a sequence of approximating representations of the algebra $\mathfrak{A}$ we mean the sequence consisting of a representation Π_n for each subalgebra $\mathfrak{A}_n$ such that

(i) $\mathfrak{A} = \cup_n \mathfrak{A}_n$,

(ii) $\mathfrak{A}_{n+1} \supset \mathfrak{A}_n$,

(iii) $\mathfrak{A}\Pi_{m,n}\to\tilde{\Pi}_m$ as $n\to\infty$ in the sense of approximating Hilbert spaces,

(iv) the $\tilde{\Pi}_m$ are a consistent set of subrepresentations,

(v) the inductive limit of $\tilde{\Pi}_m$ is a representation Π of $\mathfrak{A}$.

*Work supported by the Energy Research and Development Administration.

[1]R. Menikoff and D. H. Sharp, J. Math. Phys. 16, xxxx (1975).
[2]For a review see E. Feenberg, *Theory of Quantum Fluids* (Academic, New York, 1969).
[3]G. A. Goldin, J. Math. Phys. 12, 462 (1971).
[4]G. A. Goldin, J. Grodnik, R. T. Powers, and D. H. Sharp, J. Math. Phys. 15, 88 (1974).
[5]R. Menikoff, J. Math. Phys. 15, 1138 (1974).
[6]R. Menikoff, J. Math. Phys. 15, 1394 (1974).
[7]Brief reviews of the representation theory of local current algebras can be found in G. A. Goldin and D. H. Sharp, "Lie Algebras of Local Currents and their Representations," in *Group Representations in Mathematics and Physics, Battelle Seattle* 1969 *Rencontres*, edited by V. Bargmann (Springer, New York, 1970), pp. 300—11, and D. H. Sharp, "What we have learned about representing local nonrelativistic current algebras," in *Local Currents and Their Applications*, edited by D. H. Sharp and A. S. Wightman (North-Holland, 1974), pp. 85—98.
[8]R. F. Dashen and D. H. Sharp, Phys. Rev. 165, 1857 (1968). We work in units in which $m=1$ and $\hbar=1$, and $:\rho\rho:$ means normal ordering. See Ref. 1.
[9]C. E. Campbell, Phys. Lett. **44A**, 471 (1973).
[10]E. Feenberg, Ann. Phys. (N.Y.) 48, 128 (1974).
[11]C. Woo, Phys. Rev. A 6, 2312 (1972).
[12]For example, see D. Ruelle, *Statistical Mechanics* (Benjamin, New York, 1969).
[13]*A* C^*-algebra can be formed by exponentiating the canonical fields ϕ and π. In this case it has been shown by Segal [I. E. Segal, Dan. Mat. Fys. Medd. **31**, No. 12 (1959)] that the C^*-algebra is independent of the representation of ϕ and π. An equivalent theorem for the local currents ρ and J is not known. For further discussion, see R. Haag, "Canonical Commutation Relations in Field Theory and Functional Integration," in *Lectures in Theoretical Physics, VIII*, edited by W. E. Brittin, B. W. Downs, and J. Downs (Interscience, New York, 1961), pp. 353—81.
[14]H. F. Trotter, Pacific J. Math. 8, 887 (1958).
[15]This is just the Gel'fand—Naimark—Segal construction. A clear exposition of the GNS construction can be found in M. C. Reed, "The GNS Construction—A Pedagogical Example," in *Lectures on Elementary Particles and Quantum Field Theory*, edited by S. Deser, M. Grisaru, H. Pendleton (M. I. T. Press, Cambridge, Mass., 1970).

The Weyl tensor and shear-free perfect fluids

E. N. Glass*

Department of Physics, University of Windsor, Windsor, Ontario N9B 3P4, Canada
(Received 3 June 1975)

It is proved that a necessary and sufficient condition for a shear-free perfect fluid to be irrotational is that the Weyl tensor be pure electric type. For shear-free isentropic flow with unit tangent u^α, we find the conservation law $\nabla_\alpha(n^{1/3}i\omega\, u^\alpha)=0$, where i is the relativistic specific enthalpy, n is the conserved particle number density, and ω is the vorticity scalar.

1. INTRODUCTION

Shear-free perfect fluids are studied as models for astrophysical systems when differential rotation can be neglected. They appear in equilibrium models for stellar interiors when the flow vector is chosen as a constant sum of Killing vectors (stationary and axially symmetric for instance), in spherically symmetric collapse models, and in many cosmological models.

In this paper we prove the theorem that a shear-free perfect fluid is irrotational if and only if the Weyl tensor is pure electric type. The idea for the theorem proved here came from earlier work[1] in which it was shown that a necessary and sufficient condition for a stationary vacuum space—time to be static is that the Weyl tensor be pure electric type. The timelike Killing congruence is a shear-free, expansion-free vector field, and the theorem was essentially due to the shear-free property (a similar theorem holds for conformal timelike Killing vectors).

In addition, for shear-free isentropic flow, we show that the scalar $n^{-2/3}i\omega$ is constant along the streamlines, and derive the conservation law $\nabla_\alpha(n^{1/3}i\omega\, u^\alpha)=0$ for angular momentum density.

All necessary kinematic relations are collected in appendix.

2. PERFECT FLUID

A perfect fluid is described by the energy—momentum tensor

$$T_{\mu\nu}=wu_\mu u_\nu-p\gamma_{\mu\nu}, \tag{1}$$

where $u^\mu u_\mu=1$, and $\gamma^\mu{}_\nu$ is the projection operator onto the 3-space quotient to the streamlines:

$$\gamma_{\mu\nu}:=g_{\mu\nu}-u_\mu u_\nu. \tag{2}$$

Projection will be abreviated by $\perp$ which projects all free indices, i.e.,[2]

$$\perp A_{\alpha\beta;}{}^{\beta\nu}:=\gamma_\alpha{}^\mu\gamma_\lambda{}^\nu A_{\mu\beta;}{}^{\beta\lambda}.$$

The kinematic quantities which characterize the streamlines are the acceleration

$$a^\mu:=u^\alpha\nabla_\alpha u^\mu=:\dot{u}^\mu,$$

the expansion[3]

$$\theta:=\nabla_\mu u^\mu,$$

the shear

$$\sigma_{\mu\nu}:=\perp[u_{(\mu;\nu)}-\tfrac{1}{3}\theta\gamma_{\mu\nu}],$$

and the vorticity (or rotation) bivector

$$\omega_{\mu\nu}:=\perp u_{[\mu;\nu]}.$$

It follows from these definitions that

$$u_{\mu;\nu}=a_\mu u_\nu+\omega_{\mu\nu}+\sigma_{\mu\nu}+\tfrac{1}{3}\theta\gamma_{\mu\nu}. \tag{3}$$

The vorticity (twist, rotation) vector

$$\omega^\mu:=\tfrac{1}{2}\eta^{\mu\nu\alpha\beta}u_\nu u_{\alpha;\beta},$$

where $\eta^{0123}:=-(-g)^{-1/2}$, has a dual relation with the vorticity bivector

$$\overset{*}{\omega}_{\mu\nu}=2\omega_{[\mu}u_{\nu]}.$$

The shear and vorticity scalars are

$$2\sigma^2:=\sigma_{\alpha\beta}\sigma^{\alpha\beta}\geq 0,$$

$$2\omega^2:=\omega_{\alpha\beta}\omega^{\alpha\beta}=-2\omega_\alpha\omega^\alpha\geq 0.$$

It will also be useful to define the alternating tensor in the quotient space

$$\eta^{\mu\nu\alpha}:=\eta^{\mu\nu\alpha\beta}u_\beta.$$

The equations of motion $T^{\mu\nu}{}_{;\nu}=0$ have components

$$\dot{w}+(w+p)\theta=0, \tag{4}$$

$$(w+p)a_\mu=p_{,\alpha}\gamma^\alpha{}_\mu. \tag{5}$$

The Einstein field equations $G_{\mu\nu}=-\kappa T_{\mu\nu}$ determine the Ricci tensor:

$$R_{\mu\nu}=-(\kappa/2)[(w+3p)u_\mu u_\nu+(p-w)\gamma_{\mu\nu}]. \tag{6}$$

To conclude this section, we note that the condition for a fluid to be shear-free is $\sigma_{\mu\nu}=0$, which is equivalent to

$$\underset{u}{\mathcal{L}}\gamma_{\mu\nu}=\tfrac{2}{3}\theta\gamma_{\mu\nu}. \tag{7}$$

3. WEYL TENSOR

The electric and magnetic parts of the Weyl tensor are defined by

$$E_{\mu\alpha}:=C_{\mu\nu\alpha\beta}u^\nu u^\beta, \tag{8a}$$

$$B_{\mu\alpha}:=\overset{*}{C}_{\mu\nu\alpha\beta}u^\nu u^\beta. \tag{8b}$$

Both tensors are symmetric, trace-free, and orthogonal to u^α.

Theorem: A necessary and sufficient condition for a shear-free perfect fluid to be irrotational is that the Weyl tensor be pure electric type.

Proof: The "necessary" part of the theorem follows

when the vorticity and shear are set equal to zero in Eq. (A4). To prove sufficiency, consider Eq. (A8):

$$\nabla^\mu B_{\mu\beta} = -\kappa(w+p)\omega_\beta - a^\alpha B_{\alpha\beta} - 3\omega^\alpha E_{\alpha\beta} \tag{9}$$

where the shear-free condition has been imposed and the field equations (6) have been used to evaluate the Ricci tensor term. Setting $B_{\mu\beta}$ equal to zero yields

$$\omega^\alpha[E_{\alpha\beta} + (\kappa/3)(w+p)\gamma_{\alpha\beta}] = 0. \tag{10}$$

When the Weyl tensor is either pure electric or pure magnetic, it must be[4] Petrov type I, D, or 0. $\omega^\alpha = 0$ follows immediately for type 0. For types I and D

$$\det[E_{\alpha\beta} + (\kappa/3)(w+p)\gamma_{\alpha\beta}] \neq 0, \tag{11}$$

since if the determinant were zero, then $E_{\alpha\beta}$ would have three equal eigenvalues, which is not possible. Thus $\omega^\alpha = 0$. ■

One might also expect that a shear-free perfect fluid with vanishing electric type Weyl tensor must be conformally flat (particularly since such a theorem holds for a shear-free timelike congruence in vacuum), but it is simply not true. Equation (A7) yields

$$3\omega^\alpha B_{\alpha\beta} + (\kappa/3)w_{,\alpha}\gamma_{\alpha\beta} = 0 \tag{12}$$

for a perfect fluid with $E_{\alpha\beta} = 0$. The proof of flatness for the vacuum case follows directly, but neither Eq. (12) nor any of the other Ricci or Bianchi identity equations require that ω^α and $B_{\alpha\beta}$ vanish when a shear-free perfect fluid is present.

Equation (A3) does show that *geodesic* flow with $E_{\mu\nu} = \sigma_{\mu\nu} = 0$ must be irrotational and hence conformally flat. This agrees with a result for which Ellis[5] credits the "if" part to Trümper: A perfect fluid space–time with an equation of state has a flow with $a^\mu = \omega^\mu = \sigma^{\mu\nu} = 0$, and a Robertson–Walker metric, iff it is conformally flat.

4. CONSERVED QUANTITY

The relativistic specific enthalpy is

$$i := 1 + \epsilon + p/\rho,$$

where ϵ is the specific internal energy of the fluid and ρ the proper density with

$$w = \rho(1+\epsilon).$$

The tensor $i\omega_{\mu\nu}$ is constant along the streamlines of an isentropic perfect fluid, and it is known[6] that this constancy represents the relativistic form of the Newtonian vorticity conservation law. We will derive this conservation law here and then use it in the shear-free case to construct another conserved quantity.

Isentropic flow implies

$$di = dp/\rho, \tag{13}$$

from which it follows that

$$\frac{dp}{(w+p)} = \frac{di}{i} =: df. \tag{14}$$

The equation of motion (5) can now be written

$$a_\mu = f_{,\alpha}\gamma^\alpha{}_\mu. \tag{15}$$

Direct computation yields

$$2a_{[\mu;\nu]} = -f_{,\alpha}u^\alpha\omega_{\mu\nu}. \tag{16}$$

The kinematic equation (A6) combined with Eq. (16) results in

$$\underset{u}{\mathcal{L}}(e^f\omega_{\mu\nu}) = 0, \tag{17}$$

and from the definition of f in Eq. (14), we have the result

$$\underset{u}{\mathcal{L}}(i\omega_{\mu\nu}) = 0, \tag{18}$$

which holds for all isentropic flow.

It is possible to find a scalar which is constant along shear-free streamlines and which is a generalization of Rayner's result[7,8] $\underset{u}{\mathcal{L}}\,\omega = 0$ for rigid motions.

First note that shear-free motion is conformally rigid, and with the definition

$$\underset{u}{\mathcal{L}}\log n := -\theta, \tag{19}$$

the rigid spatial metric is $n^{2/3}\gamma_{\mu\nu}$, since Eq. (7) becomes

$$\underset{u}{\mathcal{L}}(n^{2/3}\gamma_{\mu\nu}) = 0. \tag{20}$$

The equation of motion (4) determines $\log n$ when an equation of state $p = p(w)$ exists,

$$\log n = \int \frac{dw}{w+p}, \tag{21}$$

and Eq. (4) can then be replaced by Eq. (19) in its standard form:

$$(nu^\alpha)_{;\alpha} = 0.$$

n can be interpreted as the conserved particle number density.

The scalar which is constant along the streamlines is constructed by writing

$$(n^{-2/3}\gamma^{\alpha\mu})(n^{-2/3}\gamma^{\beta\nu})\,\underset{u}{\mathcal{L}}(i^2\omega_{\alpha\beta}\omega_{\mu\nu}) = 0, \tag{22}$$

which follows from Eq. (18). Note that

$$\underset{u}{\mathcal{L}}(n^{-2/3}\gamma^{\alpha\mu}) = -2n^{-2/3}a^{(\alpha}u^{\mu)},$$

so that Eq. (22) reduces to

$$\underset{u}{\mathcal{L}}(n^{-4/3}i^2\omega^2) = 0,$$

or finally

$$\underset{u}{\mathcal{L}}(n^{-2/3}i\omega) = 0. \tag{23}$$

To interpret the conserved scalar, we will proceed to reformulate Eq. (23) in the standard form of a conservation law.

The rule for the Lie derivative of a scalar density ψ of weight W is

$$\underset{u}{\mathcal{L}}\,\psi = \psi_{;\alpha}u^\alpha + W\psi u^\alpha{}_{;\alpha}.$$

Thus Eq. (19) has the equivalent form

$$\underset{u}{\mathcal{L}}[(-g)^{1/2}n] = 0. \tag{24}$$

Equation (23) can now be written as

$$\underset{u}{\mathcal{L}}[(-g)^{1/2}n^{1/3}i\omega] = 0$$

or equivalently

$$(n^{1/3}i\omega u^\alpha)_{;\alpha} = 0. \tag{25}$$

Equation (25) can be interpreted as conservation of angular momentum density for shear-free isentropic flow.

5. SUMMARY AND COMMENTS

A theorem has been proved which states that a necessary and sufficient condition for a shear-free perfect fluid to be irrotational is that the Weyl tensor be pure electric type. The magnetic part of the Weyl tensor has no Newtonian analog, and there are shear-free, rigidly rotating, fluid solutions (such as Maclaurin spheroids) in Newtonian hydrodynamics. It is now clear that any relativistic counterpart of such Newtonian configurations must have a nonzero magnetic Weyl tensor.

It is interesting to note that when a perfect fluid *is* shear-free and irrotational, there are only three possibilities: the fluid is either static, type D, or conformally flat.[9]

The possibility was left open that perfect fluid metrics exist with $\sigma_{\mu\nu}=E_{\mu\nu}=0$, where $B_{\mu\nu}\neq 0$ and $\omega^{\mu}\neq 0$ whenever a^{μ} is nonzero. Such solutions would be rather unphysical stellar models since they would not possess a Newtonian limit.

For shear-free isentropic flow, it was shown that the scalar $n^{-2/3}i\omega$ is constant along the streamlines. The conservation law $\nabla_{\alpha}(n^{1/3}i\omega u^{\alpha})=0$ was constructed from the scalar and interpreted as conservation of angular momentum density.

ACKNOWLEDGMENTS

It is a pleasure to thank R.O. Hansen, B. Mashhoon, and J. Winicour for useful remarks.

APPENDIX: KINEMATIC RELATIONS

With field equations unspecified, the Ricci identities provide kinematic propagation equations for the vorticity, shear, and expansion of a fluid congruence. The Bianchi identities provide the kinematic equations for the electric and magnetic parts of the Weyl tensor. Sign conventions are fixed by Ricci's identity

$$2u_{\nu;[\alpha\beta]}\equiv u^{\mu}R_{\mu\nu\alpha\beta},$$

and $R_{\nu\alpha}:=R^{\beta}{}_{\nu\alpha\beta}$. The Riemann and Weyl tensors are related by

$$R^{\mu\nu}{}_{\alpha\beta}=C^{\mu\nu}{}_{\alpha\beta}-2\delta^{[\mu}{}_{[\alpha}R^{\nu]}{}_{\beta]}+\tfrac{1}{3}\delta^{\mu}{}_{[\alpha}\delta^{\nu}{}_{\beta]}R.$$

Ricci's identity yields the following kinematic relations:

$$R_{\alpha\beta}u^{\alpha}u^{\beta}=\dot{\theta}+\tfrac{1}{3}\theta^2+2\sigma^2-2\omega^2-a^{\alpha}{}_{;\alpha}, \tag{A1}$$

$$\perp u^{\beta}R_{\alpha\beta}=\perp(\omega_{\alpha\beta;}{}^{\beta}-\sigma_{\alpha\beta;}{}^{\beta}+\tfrac{2}{3}\theta_{,\alpha})-a^{\beta}(\sigma_{\alpha\beta}+\omega_{\alpha\beta}), \tag{A2}$$

$$E_{\mu\nu}=\tfrac{1}{2}(\perp R_{\mu\nu}-\tfrac{1}{3}\gamma_{\mu\nu}R_{\alpha\beta}\gamma^{\alpha\beta})+\perp a_{(\mu;\nu)}-a_{\mu}a_{\nu}$$
$$-\perp\dot{\sigma}_{\mu\nu}-\tfrac{2}{3}\theta\sigma_{\mu\nu}-\sigma_{\mu\alpha}\sigma_{\nu}{}^{\alpha}+\omega_{\mu\alpha}\omega_{\nu}{}^{\alpha}$$
$$+\tfrac{1}{3}\gamma_{\mu\nu}(2\sigma^2-2\omega^2-a^{\alpha}{}_{;\alpha}), \tag{A3}$$

$$B_{\mu\nu}=-\eta^{\alpha\beta}{}_{(\mu}\gamma_{\nu)}{}^{\lambda}[\omega_{\lambda\alpha;\beta}+\sigma_{\lambda\alpha;\beta}]-2a_{(\mu}\omega_{\nu)}, \tag{A4}$$

$$\omega^{\alpha}{}_{;\alpha}+2a_{\alpha}\omega^{\alpha}=0, \tag{A5}$$

$$\perp a_{[\mu;\nu]}=\perp\dot{\omega}_{\mu\nu}+\tfrac{2}{3}\theta\omega_{\mu\nu}-2\omega_{\alpha[\mu}\sigma_{\nu]}{}^{\alpha}$$
$$=\pounds_{u}\omega_{\mu\nu}. \tag{A6}$$

The Bianchi identities

$$\nabla^{\mu}C_{\mu\nu\alpha\beta}=R_{\nu[\alpha;\beta]}-\tfrac{1}{6}g_{\nu[\alpha}R_{,\beta]},$$

and the Weyl tensor expressed in terms of its electric and magnetic parts ($g_{\mu\nu\alpha\beta}:=g_{\mu\alpha}g_{\nu\beta}-g_{\mu\beta}g_{\nu\alpha}$),

$$C_{\mu\nu\rho\sigma}=(g_{\mu\nu\alpha\beta}g_{\rho\sigma\lambda\gamma}-\eta_{\mu\nu\alpha\beta}\eta_{\rho\sigma\lambda\gamma})u^{\alpha}u^{\lambda}E^{\beta\gamma}$$
$$-(g_{\mu\nu\alpha\beta}\eta_{\rho\sigma\lambda\gamma}+\eta_{\mu\nu\alpha\beta}g_{\rho\sigma\lambda\gamma})u^{\alpha}u^{\lambda}B^{\beta\gamma},$$

yield the kinematic relations for $E^{\beta\gamma}$ and $B^{\beta\gamma}$ by straightforward calculation. In this paper we will only need the following divergence relations:

$$\nabla^{\mu}E_{\mu\alpha}=R_{\nu[\alpha;\beta]}u^{\nu}u^{\beta}+\tfrac{1}{12}\gamma_{\alpha}{}^{\beta}R_{,\beta}-a^{\beta}E_{\alpha\beta}$$
$$+3\omega^{\beta}B_{\alpha\beta}-u_{\alpha}\sigma^{\mu\nu}E_{\mu\nu}, \tag{A7}$$

$$\nabla^{\mu}B_{\mu\alpha}=\tfrac{1}{2}\eta_{\alpha}{}^{\mu\nu}u^{\beta}R_{\beta\mu;\nu}-a^{\beta}B_{\alpha\beta}-3\omega^{\beta}E_{\alpha\beta}-u_{\alpha}\sigma^{\mu\nu}B_{\mu\nu}. \tag{A8}$$

*Supported in part by NRC of Canada Grant A9343.

[1]E.N. Glass, J. Math. Phys. 15, 1930 (1974).

[2]Covariant derivatives are denoted by semicolons and ∇_{μ}; partial derivatives by commas. Parentheses around indices denote symmetrization and brackets around indices denote antisymmetrization. The dual operator is $\frac{1}{2}\eta^{\mu\nu\alpha\beta}$ and the dual is denoted by an asterisk. $\pounds_{u}$ symbolizes the Lie derivative with respect to u^{α}.

[3]We adhere to conventional usage and drop the prefix "rate-of" from expansion, shear, and rotation.

[4]P. Jordan, J. Ehlers, and W. Kundt, Akad. Wiss. Mainz, Abh. Math.-Nat. Kl., No. 2 (1960).

[5]G.F.R. Ellis, "Relativistic Cosmology," in *General Relativity and Cosmology*, Course 47, edited by R.K. Sachs (Academic, New York, 1971), p. 136.

[6]S.W. Hawking and G.F.R. Ellis, *The Large Scale Structure of Space–Time* (Cambridge U.P., Cambridge, 1973), p. 84.

[7]C.B. Rayner, C.R. Acad. Sci. Paris 248, 929 (1959).

[8]R.H. Boyer, Proc. Roy. Soc. Lond. A **283**, 343 (1965).

[9]A. Barnes, Gen. Relativity Gravitation 4, 105 (1973).

Singularities in nonsimply connected space–times

Dennis Gannon

Department of Mathematics, University of California, Davis, Davis, California 95616
(Received 14 July 1975)

Space–times with asymptotically flat nonsimply connected spacelike slices are shown to possess enough intrinsic geometric structure to guarantee the existence of singularities under conditions usually considered insufficient. In particular, it is shown that if the normal geodesics to the spacelike slice are converging on a suitable compact set, and the space–time satisfies a standard energy condition, then it is timelike geodesically incomplete. A similar result holds if the space–time satisfies the chronology and generic conditions.

INTRODUCTION

Scattered throughout the literature on general relativity one finds frequent reference to space—times that admit a spacelike hypersurface that is asymptotically flat but that might also have a tubular bridge or "wormhole" connecting separate regions of that hypersurface. Wheeler[1] frequently discusses such a possibility in connection with his imaginative approach to the electric dipole as line of force trapped in a multiply-connected topology. Still other workers encounter similar possibilities when considering the spacelike slices through models of black holes.[2,3] (In each of the Reissner—Nordström, Schwarzschild, and Kerr families of solutions there are hypersurfaces that possess nontrivial topological structures.)

Brill and Deser[4] also consider spaces of non-Euclidean topology when speaking of the extremal gravitation energy problem. As much of their very important analysis depends on the existence of nonsingular solutions, the question of incompleteness in such cases is certainly significant.

In an earlier work, this author[5] considers the case of a bridge connecting two asymptotically flat spaces and shows that when the hypersurface is Cauchy, the space—time is singular if the energy condition holds. This condition, known more commonly as the weak-energy condition, is the statement that the energy density is nonnegative. More formally, $T_{ab}W^aW^b \geq 0$ for any timelike vector W.

In the case where the bridge is connected to one asymptotically flat space, the surface is no longer simply connected. Because of its rich geometric structure, the nonsimply connected, asymptotically-flat 3-geometry is an ideal situation in which to look for singularities in the developing space—time, with a minimum amount of sturcture on the initial Cauchy data.

The goal of this paper will be to show that for most of the standard singularity theorems that require a compact partial Cauchy surface or a closed trapped surface, one may arrive at the same conclusions by assuming the existence of an asymptotically flat, nonsimply connected hypersurface.

To be more specific, we shall say that a spacelike hypersurface S is regular near infinity if it satisfies the following three conditions:

(1) $S=\cup_{i=1}^{\infty}W_i$, $W_i\subset W_{i+1}$, where W_i is a compact 3-manifold whose boundary ∂W_i is homeomorphic to a sphere.

(2) S-interior W_i is homeomorphic to $\partial W_i\times R^+$, where R^+ denotes the nonnegative real line.

(3) The inward directed null geodesics orthogonal to ∂W_i are converging everywhere on ∂W_i. Here inward means those geodesics which locally project orthogonally into W_i.

The last condition is meant to reflect the fact that in an asymptotically flat hypersurface in an asymptotically flat space—time, one should be able to find sphere S of sufficiently large radius to put them in the nearly flat region so that the ingoing null geodesics are converging. To say that S is nonsimply connected shall mean the fundamental group $\pi_1(S)$ is nontrivial. (This is loosely interpreted as saying that there is a loop in S that cannot be shrunk to a point in S.) Figure 1 illustrates a two-dimensional cross section of such a surface.

S spacelike hypersurface without edge is said to be a partial Cauchy surface. If, in addition, S has the property that every nonspacelike curve in the space—time (M, g) meets S, then S is said to be a Cauchy surface. The first result of this article will be to show that if (M, g) satisfies the energy condition and M admits a Cauchy surface that is nonsimply connected and regular near infinity, then (M, g) is not complete.

This result is analogous to the theorem of Penrose[6] that states any space—time satisfying the energy condition with a Cauchy surface and a closed trapped surface is singular.

Without the global assumption that S is a Cauchy surface, one must specify more of the initial data on S in order to arrive at any definitive result. Hawking[7,8] accomplishes this by showing that if S is a compact partial Cauchy surface in a space—time satisfying the energy

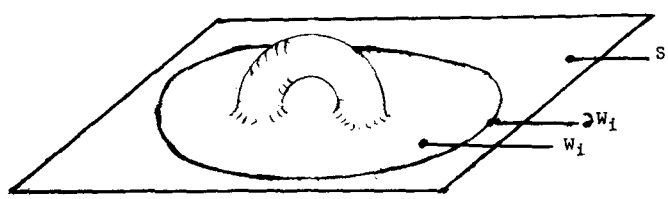


FIG. 1.

condition, and if the timelike normals are converging everywhere on S, then (M, g) is incomplete. In the case where S is not compact, an important result is given by Hawking and Penrose.[9,10] They show that if (M, g) satisfies (a) the energy condition (b) the chronology condition (no timelike loops), (c) a technical but reasonable generic condition guaranteeing the existence of conjugate points, and (d) some form of future or past trapped set, then (M, g) is incomplete.

Briefly then, the results herein will be to show that if (M, g) admits a partial Cauchy surface nonsimply connected and regular near infinity, then in the above results one need not assume anything about compactness or the existence of trapped sets.

1. PRELIMINARIES

Much of the notation here that is more or less standard will be according to the definitions given in the book by Hawking and Ellis[11] (see also the monograph by Penrose[12]). Important among these will be: $J^+(X)$, the causal future of X; $I^+(X)$, the chronological future of X; $E^+(X)$, which is $J^+(X) - I^+(X)$. The future domain of dependence of a set X, $D^+(X)$ is defined to be the set of points p in M such that every past directed nonspacelike through p meets X.[13]

From geometric topology we shall need the concept of the universal covering space $\bar{X}$ of a space X. First, a covering is a space Y together with a surjective map π: $Y \to X$ that is locally a homeomorphism. Then Y is said to be a covering space for X. A covering space $\bar{X}$ is said to be universal if $p : \bar{X} \to X$ is a covering with the property that if $\pi : Y \to X$ is any other covering then there is a map $f : \bar{X} \to Y$ such that $\pi \circ f = p$. A fundamental fact about such objects is that a universal covering space exists for every space X, and it is unique and simply connected.[14]

In proving singularity theorems, a standard useful tool is to consider congruences of null geodesics. Let ξ^i be the tangent vector field to such a congruence, and e_a, $a = 1, 2$, be a basis for a spacelike cross section V of this congruence. The convergence c of the congruence is defined to be $-\nabla_a \xi^a$. The importance of positive convergence lies in the following fact.

Focusing Lemma: Let (M, g) satisfy the energy condition and let $c = c_0 > 0$ at x in V. Then, within an affine distance of $2/c_0$ there is a point conjugate to V along the geodesic through x.

Proof: See the excellent survey by Geroch.[15]

By variational methods it is not hard to show that a point beyond such a conjugate lies in $I^+(V)$.

Before proceeding to the proof of the first main result, one more technical fact must be considered. Let S be a simply connected partial Cauchy surface for a space–time (M, g). Let V be a 2-surface in S. Then the boundary of the causal future of V, $\partial J^+(V)$, is an acronal 3-manifold in M. Furthermore, $\mathrm{ad}D^+(S)$ is causally simple, $\partial J^+(V) \cap D^+(S)$ is generated by null geodesics with past endpoints on V. There are two such families orthogonal to V.

Lemma 1.1: With the notation as above, a null generator of one family cannot meet a null generator of the other family in $\partial J^+(V) \cap D^+(S)$.

Sketch of Proof[16]: By assuming the existence of such an occurrence, one can construct a loop in $\partial J^+(V) \cap D^+(S)$ that meets V only once. $D^+(S)$ admits a timelike vector field which can be used to project this loop into S. But as S is assumed to be simply connected, this projected loop is contractable in S. From geometric intersection theory, it follows that this projected loop must meet V in at least one other point. Following the integral curves of the vector field from this point back to the original loop yields a timelike curve from V to $\partial J^+(V) \cap D^+(S) \subset E^+(V)$. But this is impossible because $E^+(V) = \partial J^+(V) - I^+(V)$.

The following proposition will serve as the primary tool in proving the singularity results that follow.

Proposition 1.2: Let (M, g) be a space–time satisfying the energy condition and have a partial Cauchy surface nonsimply connected and regular near infinity. Then $D(S)$ is incomplete.

Proof: By a theorem of Geroch,[17] $D(S)$ is homeomorphic to $S \times R$. This means in particular $D(S)$ may be retracted onto S, and hence

$$\pi_1(S) \approx \pi_1(D(S)).$$

Let $\overline{D(S)}$, $\bar{S}$ be the universal covering spaces for $D(S)$ and S respectively. The metric g on $D(S)$ lifts to a metric $\bar{g}$ for $\overline{D(S)}$. It is easily seen that $\bar{S}$ is a Cauchy surface for $\overline{D(S)}$. Furthermore, any local notations satisfied by g at a point in $D(S)$ will be satisfied by $\bar{g}$ at a point in $\pi^{-1}(p)$ in $\overline{D(S)}$, where $\pi : \overline{D(S)} \to D(S)$ is the covering map. This follows because π is a local diffeomorphism and $\bar{g}$ is chosen to be locally isometric to g via π. It will be shown that $\overline{D(S)}$ contains a null geodesic which cannot be extended beyond a certain value of its affine parameter. Then projecting this curve into $D(S)$ results in the same situation, and the proof is complete.

Let $B = \partial W$, W being a compact nonsimply connected submanifold of S whose existence is guaranteed by the assumption that S is regular and nonsimply connected. Furthermore, we have that B is a sphere whose ingoing null geodesics are everywhere converging. The projection map π, restricted to $\pi^{-1}(B)$, yields a covering fibration for B. We must observe that $\pi^{-1}(B)$ is a disjoint union of connected components each homeomorphic to B. To see this, let A be one component of $\pi^{-1}(B)$. Considered by itself, A is a covering space for B. But, as B is a sphere, it is simply connected so serves as its own universal covering space. Therefore, by the universality property of covering spaces there is a map $f : B \to A$ such that $\pi \circ f = \text{identity}$. This implies π is a one-to-one homeomorphism when restricted to A. For any x in S, $\pi^{-1}(x)$ contains as many elements as there are elements in the group $\pi_1(S)$. Hence, if $\pi_1(S) \neq 0$, there is more than one component to $\pi^{-1}(B)$.

As $B = \partial W$, $\pi^{-1}(B) = \partial \pi^{-1}(W)$. But, as W is a retract of S, $\pi^{-1}(W)$ is a retract of $\bar{S}$ and $\pi^{-1}(W)$ is the universal covering space $\bar{W}$ for W. It follows that if A is a component of $\pi^{-1}(B)$, then A is not the boundary of a compact 3-manifold in $\bar{S}$. But as $\bar{S}$ is a retract of $D(\bar{S})$, A does not bound a 3-manifold in $D(\bar{S}) = \overline{D(S)}$ either.

To complete the proof, focus attention on $\partial J^+(A)$ in $D(\bar{S})$. This is an acronal 3-manifold in $D(\bar{S})$ generated by null geodesics with past end points on A. By Lemma 1.1, the ingoing null geodesics do not meet the outgoing null geodesics off A, and hence both families form proper submanifolds of $D(\bar{S})$. Furthermore, the ingoing null curves for A must be converging because they must project to the converging ingoing null curves for B. By the Focusing Lemma, these geodesics must leave $\partial J^+(A)$ within a finite affine distance. This implies that if these geodesics leave $\partial J^+(A)$, the submanifold determined by them would be compact with boundary A. But by the above observations, this is impossible. Hence, there exists some geodesic from A that cannot be extended to every value of its affine parameter. QED

In what follows it shall be important to observe the geodesic that is inextendable in $D(\bar{S})$ projects to an inextendable geodesic in $\partial J^+(B) \cap D(S)$. Or, in particular, when $D(S) \not\subset M$, then there is a null geodesic in $\partial J^+(B)$ that either meets $H^+(S)$, the future Cauchy horizon of S or is inextendable and M is incomplete.

This last aspect can be summarized in the following.

Corollary 1.2. Let (M, g) satisfy the energy condition and have a Cauchy surface S that is nonsimply connected and regular near infinity. Then (M, g) is null geodesically incomplete.

Proof. If S is a Cauchy surface for M, then $D(S) = M$. QED

2. THE SINGULARITY THEOREMS

We now consider the case when S is not a Cauchy surface for M. The first theorem will consider the case when the timelike normal vector to S are converging everywhere on S. Convergence c is defined in much the same way here as was the case for null geodesics. Let ξ^a be the unit normal field to S; then $c = -\nabla_a \xi^a$, where $a = 1, 2, 3$. The version of the Focusing Theorem that applies here is much the same as the previous case. If the energy condition holds, there is a conjugate point to S along any geodesic with positive convergence c_0 at S. And this conjugate point occurs within an affine distance of $c/3$. Furthermore, for any point x in $D^+(S)$ there is a geodesic orthogonal to S of maximal length from S to x with no conjugate in A between x and S.[18]

Theorem 2.1: Let (M, g) satisfy the energy condition and have a partial Cauchy surface S nonsimply connected and regular near infinity. Then if $c > 0$ everywhere on S, (M, g) is not timelike geodesically complete.

Proof: Let W be a compact manifold such that ∂W is a sphere whose inward null generators are converging. By Proposition 1.2, there is a null generator γ for $\partial J^+(\partial W)$ that must intersect $H^+(S)$ or else it cannot be extended beyond some value of its affine parameter in M. Furthermore, this geodesic is inward directed and hence forms a generator of the boundary of $D^+(W)$. As W is compact, the convergence c takes some minimum value $c_0 > 0$ on W. As stated previously, every timelike geodesic orthogonal to S must encounter a focal point within an affine distance $3/c_0$ from S, and every point p in $D^+(W)$ must lie on a geodesic orthogonal to S with no focal point in the interior.

Construct V to be the set of points on geodesics orthogonal to W and within a distance $3/c_0$ from W. Then (M, g) is timelike complete, V is compact, and $D^+(W) \subset V$. Therefore, $D^+(W)$ is compact and is contained in V.

Now consider the curve γ. Assume γ never meets $H^+(S)$. Then $\gamma \subset$ interior of $D^+(S)$ which is globuly hyperbolic and satisfies the strong causality condition. This means every nonspacelike curve in $D^+(S)$ must leave every compact set it meets in $D^+(S)$.[19] In particular, γ must leave $D^+(W)$. But this is impossible because every point on γ lies in $\partial J^+(\partial W)$ along an interior directed null generator and is therefore not accessible from $S-W$ from a nonspacelike curve. Hence γ meets $H^+(S)$.

Let σ be the past directed null generator of $H^+(S)$ that meets γ. Let q be a point on σ and τ be a nonspacelike path from S to q. The combined path $\tau\sigma$ is then a nonspacelike path from S to γ, which by the above observations can only occur if q is in W. It follows then that σ is a null generator for $H^+(W)$. The proof now proceeds along the lines of Hawking's theorem quoted in the Introduction. Let $d(S, p)$ be the least upper bound of the lengths of nonspacelike curves from p to S, if it exists. Let $d(S, p)$ be infinity if there is no least upper bound and zero if there is no timelike separation. It is not hard to show $d(S, p)$ is lower semicontinuous on M. Let p be a point in σ, p_i in $J^-(p)$, with p_i converging to p. Then each $p_i \in V$ so by compactness $p \in V$, and furthermore there is a curve τ from W to p of maximal length. Let q be any point to the past of p along σ. $d(S, q)$ must be less than $d(S, p)$, for otherwise the maximal path from S to q together with the part of σ between q and p could be deformed to a longer path from S to p than τ. But this implies that $d(S, q)$ must be decreasing on σ, and the compactness of W and the lower semicontinuity of $d(S, p)$ implies there is a minimum somewhere on σ. This could only occur at a past endpoint for σ which does not exist because σ is a generator for $H^+(S)$. This contradiction implies our assumption that (M, g) was timelike complete was incorrect. QED

In Theorem 2.1, it should be observed that $c > 0$ could be replaced by the condition $c < 0$ as the proof could have been carried out in $D^-(S)$ just as easily. Furthermore, it is only necessary to have $c \neq 0$ on the compact set W.

In the following we replace the convergence condition by a condition known as the generic condition for nonspacelike geodesics. This condition, described extensively by Hawking and Ellis,[20] can be seen to imply that every nonspacelike geodesic contains a pair of conjugate points. Analytically, if K is the direction of the geodesic, the condition reads

$$K^c K^d K_{[a} R_{b]cd[e} K_{f]} \neq 0$$

at some point along the geodesic. Furthermore, we shall require that the chronology condition shall hold. That is, there are no closed timelike loops in M. Then the following result can be seen as an extension of the theorem of Hawking and Penrose.[21]

Theorem 2.2. Let (M, g) satisfy the energy, generic, and chronology conditions. Then, if S is a partial Cauchy surface nonsimply connected and regular near infinity, M cannot be geodesically complete.

Proof. Consider $\partial J^+(\partial W) \cap H^+(S)$. If this is empty, we are done, by the corollary to Proposition 1. If it is nonempty, then, as in Theorem 2.1, $H^+(S) \cap D^+(W) \subset H^+(W)$. Therefore, $H^+(W)$ is not compact, because if it were a generator for $H^+(S)$ would have an accumulation point. Using the generic condition, one could then find a timelike variation of this generator which would be closed.

Put a timelike vector field on the closure of $D^+(W)$. The integral curves of this field must meet W. If every integral curve also met $H^+(W)$, then this would define a map from W onto $H^+(W)$. This is impossible as W is compact. Therefore, it follows that there is some future inextendable timelike curve γ in $D^+(W)$ that never meets $H^+(W)$. Furthermore, $J^-(\gamma) \cap S \subset W$; hence $J^-(\gamma) \cap S$ is compact.

But, as γ is future inextendable, the null generators for $\partial J^-(\gamma)$ have no future end points. By employing the generic condition, they must then have past end points. Therefore, each generator for $E^-[J^-(\gamma) \cap S]$ must terminate within a finite affine distance. Hence, $E^-(B)$ is a past trapped set. But this is exactly the condition needed to satisfy the theorem of Hawking and Penrose mentioned previously. QED

3. REMARKS AND COMMENTS

1. Another approach to the question of singularities in nonsimply connected asymptotically flat space—times would be to look for a minimal 2-surface in the spacelike slice. This surface could be homologically nontrivial in M, and under the right extrinsic conditions could be shown to be a trapped surface. The now classical techniques of Penrose[22] might be employed to prove the existence of a singularity. Furthermore, this might lead to a more geometric picture of the singularity as the "point" where the "wormhole" collapses as the surface evolves in the space—time M. The main problem with this method lies in the existence problem for the minimal surface. To the best of this author's knowledge this is unsolved.

2. Concerning Theorem 2.1, one should consider the possibility that the convergence condition cannot be attained on a nonsimply connected regular surface. To answer this, one must know a good deal more about the solution space of the initial value equations governing the relation of the curvature of S to the extrinsic curvature of S in M.

In fact, by considering $S \times R$ with a trivial product metric, one can conclude from Proposition 1.1 that S cannot be Ricci flat. This is, of course, a well-known consequence of the relation of the curvature forms on S to the nontrivial homology of S, but it is pleasing to see this come out of our more geometric considerations.

3. It should also be noted that the condition of regularity cannot be weakened very much without Proposition 1.1 failing to be valid. In particular, Minkowski space admits orbit spaces with spacelike slices that can be compact and nonsimply connected, $S^1 \times S^1 \times S^1$, or nonsimply connected and infinite, $S^1 \times S^1 \times R$. (In the latter case $S^1 \times S^1 \times R$ does not have imbedded spheres near infinity.)

[1] J. A. Wheeler, *Gemetrodynamics* (Academic, New York, 1962).
[2] G. W. Gibbons, Commun. Math. Phys. **27**, 87 (1972).
[3] C. W. Misner and J. A. Wheeler, Ann. Phys. (N.Y.) **2**, 525 (1957).
[4] D. R. Brill and S. Deser, Ann. Phys. (N.Y.) **50**, 548 (1968).
[5] D. Gannon, "On the topology of spacelike hypersurfaces, singularities, and black holes," to appear in General Relativity Gravitation.
[6] R. Penrose, Phys. Rev. Lett. **14**, 57 (1965).
[7] S. W. Hawking, and G. F. R. Ellis, *The Large Scale Structure of Space-Time* (Cambridge U. P., Cambridge, 1973).
[8] S. W. Hawking, Proc. Roy. Soc. Lond. A **300**, 187 (1967).
[9] S. W. Hawking and R. Penrose, Proc. Roy. Soc. Lond. A **315**, 529 (1970).
[10] See Ref. 7.
[11] See Ref. 7.
[12] R. Penrose, *Techniques of Differential Topology in Relativity* (Soc. Ind. Appl. Math., Philadelphia, 1972).
[13] R. P. Geroch, J. Math. Phys. **11**, 437 (1970).
[14] E. Spanier, *Algebraic Topology* (McGraw-Hill, New York, 1966).
[15] R. P. Geroch, "Singularities," in *Relativity*, edited by M. Carmeli, S. Fickler, and L. Witten (Plenum, New York, 1970).
[16] See Ref. 5.
[17] See Ref. 13.
[18] See Ref. 7.
[19] See Ref. 7.
[20] See Ref. 7.
[21] See Ref. 9.
[22] See Ref. 6.

Ergodicity of quantum mechanical systems

Koichiro Matsuno

Central Research Laboratories, Nippon Electric Company, Ltd., Kawasaki 213, Japan
(Received 24 February 1975)

The ergodicity of pure quantum states is maintained in the space of orthonormal quantum states which diagonalizes an observable. The ergodicity of mixed quantum states, which is met in quantum statistical mechanics admitting an ensemble of many similar systems, is identical to the principle of equal *a priori* probabilities.

1. INTRODUCTION

The quasi-ergodic theorem of von Neumann is applied to an arbitrary Hamiltonian system in quantum mechanics under two constraints.[1,2] One is that the Hamiltonian system may admit an introduction of a set of coarse-grained quantum states in terms of which both the Hamiltonian and a macro-observable are diagonalized, in which resonances are prohibited. The other constraint is that *a priori* probability of each microscopic state belonging to a coarse-grained state may be equal. Although the first constraint might seem acceptable from a physical point of view, the second one needs verification, since the logical foundation of equal *a priori* probabilities is not found in the proposition itself. Henceforth, such an attempt seems to be desirable that the ergodicity of pure quantum states in quantum mechanics may be proved on a more sound basis with less arbitrariness, if possible.

In the present paper, we prove the ergodicity of pure quantum states such that in any Hamiltonian system in quantum mechanics the time average of an arbitrary observable, which can be measured and identified at an arbitrary time, equals its average over all the available quantum states which diagonalize the observable. The Hamiltonian is not diagonalized in this representation. The ergodicity is shown to be a natural consequence of the proved equal *a priori* probability of expectation, which says that an observer, who is not informed of any past history of events preceding the expectation, expects equally any quantum state, which is an eigenstate of the observable, at an arbitrary time.

In order to complete the proof of the ergodicity of pure quantum states, we first discuss the role of observation in quantum mechanics in Sec. 2. The proof is found in Sec. 3. In addition to the ergodicity of pure quantum states, one can sometimes think of the ergodicity of mixed quantum states if quantum statistical mechanics is concerned. A comparison of the two ergodicities which are independent of each other is given in Sec. 4.

2. OBSERVATION IN QUANTUM MECHANICS

The Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\Psi = H\Psi \tag{2.1}$$

determines a probabilistic evolution of pure quantum states in terms of the wavefunction Ψ and the Hamiltonian H of a given quantum mechanical system, where t is time and $\hbar$ is Planck's constant divided by 2π. If one concentrates only on identifying an arbitrary observable

$$A, \tag{2.2}$$

which is, of course, an Hermitian operator, in the course of quantum mechanical evolution subjected to the equation of motion

$$i\hbar\frac{\partial}{\partial t}A = [A, H], \tag{2.3}$$

the associated observation will take place in the space of orthonormal quantum states $\{\varphi_n\}$ $(n = 1, 2, \ldots, N)$, which diagonalizes the observable A as follows:

$$A\varphi_n = a_n\varphi_n, \quad n = 1, 2, \ldots, N, \tag{2.4}$$

where a_n is the eigenvalue of A corresponding to the eigenfunction φ_n and the total number of the states N does not necessarily remain finite. The observer who has made a measurement of A in the space of orthonormal set $\{\varphi_n\}$ at an arbitrary time identifies only one particular state of A and its associated eigenvalue. It never occurs that the observer would identify more than two states belonging to $\{\varphi_n\}$ at each measurement, since the observable A is diagonalized. Otherwise, the observable would no longer remain to be observable. This is because any one of the observables is so defined that each quantitative measurement of them does not admit any of indefiniteness. One readily notes that such an observable, which could be identified at each instantaneous time, and the Hamiltonian H in (2.1) do not commute with each other.

The equation of motion (2.3) predicts the probability of finding the observable A in an arbitrary state among $\{\varphi_n\}$ of (2.4) at an arbitrary time, with the use of the knowledge of the state identified at a previous time. The expected value of the observable A at time t can be expressed as

$$\langle A(t)\rangle = \langle\Psi(t)\,|\,A\,|\,\Psi(t)\rangle, \tag{2.5}$$

where $\Psi(t)$ is the solution of the Schrödinger equation (2.1). This reduces to

$$\langle A(t)\rangle = \sum_{n=1}^{N} c_n^* c_n a_n \tag{2.6}$$

with

$$\Psi(t) = \sum_{n=1}^{N} c_n\varphi_n, \tag{2.7}$$

$$\sum_{n=1}^{N} |c_n|^2 = 1, \tag{2.8}$$

where the wavefunction $\Psi(t)$ is decomposed in terms of

the orthonormal set $\{\varphi_n\}$. The quantity $\langle A(t)\rangle$, however, is an expected value of A at time t and is by no means the value of A which is really identified at time t. The value of the observable A identified at time t is always one out of $\{a_n\}$ of (2.4).

The Hamiltonian H in (2.1) is also diagonalized as

$$H\phi_n = E_n\phi_n, \quad n=1, 2, \ldots, \tag{2.9}$$

in terms of an orthonormal set $\{\phi_n\}$. However, the observer concerns himself only with identifying the observable A at an arbitrary time and not with the Hamiltonian. The identification of the Hamiltonian is impossible to such an observer since it would take an infinite time interval. This yields the statement that the Hamiltonian is not diagonalized in the space of the orthonormal set $\{\varphi_n\}$.

Suppose that an observer measures and identifies a state out of N states $\{n\}$ at each interval of τ starting from $t=t_0$. If the state identified at $t=t_0+r\tau$ ($r=0, 1, 2, \cdots$) is n, the contracted wavefunction at $t=t_0+r\tau$ will turn out to be

$$\Psi(t_0+r\tau)=\sum_{l=1}^{N} c_l(t_0+r\tau)\varphi_l \tag{2.10}$$

with

$$|c_l(t_0+r\tau)|^2=\begin{cases}1, & l=n,\\ 0, & l\neq n.\end{cases} \tag{2.11}$$

The quantum mechanical evolution

$$\frac{\partial}{\partial t}c_m(t)=\sum_{l=1}^{N}\frac{1}{i\hbar}H_{ml}^{(\varphi)}c_l(t) \tag{2.12}$$

with

$$H_{ml}^{(\varphi)}\equiv\langle\varphi_m|H|\varphi_l\rangle \tag{2.13}$$

subjected to the initial condition (2.11) determines the probability of expectation $P(m,\ t_0+(r+1)\tau|n,\ t_0+r\tau)$ that the state identified with n at $t=t_0+r\tau$ is followed by the state m at $t=t_0+(r+1)\tau$. The result is

$$P(m,\ t_0+(r+1)\tau|n,\ t_0+r\tau)=|c_m(t_0+(r+1)\tau)|^2, \tag{2.14}$$

which follows from (2.12) under the constraint (2.11).

Since the state is identified at each interval of τ and the Hamiltonian is a constant of motion, the translational invariance

$$P(m,\ t_0+(r+1)\tau|n,\ t_0+r\tau)=P(m,\ t_0+(r'+1)\tau|n,\ t_0+r'\tau) \tag{2.15}$$

with $r, r'=0, 1, 2, \cdots$

is always kept, where m and n represent arbitrary states in $\{n\}$ of (2.4). The transition probability defined in (2.14) also gives the transitivity

$$\begin{aligned}P(m,\ t_0+(q+r+s)\tau|n,\ t_0+s\tau)&=\sum_{l=1}^{N}P(m,\ t_0+(q+r+s)\tau|l,\ t_0+(r+s)\tau)\\&\quad\times P(l,\ t_0+(r+s)\tau|n,\ t_0+s\tau)\end{aligned} \tag{2.16}$$

with $q, r, s=0, 1, 2, \ldots,$

since the present stochastic process with discrete time is Markovian as shown in (2.15). Here, it should be stressed that even if the transition probability

$$P(m,\ t_0+(q+r+s)\tau|n,\ t_0+s\tau)$$

with $(q+r)\geq 2$ is referred to, the observer is supposed to measure and identify a state out of $\{n\}$ of (2.4) at each interval of τ.

The stochasticity presented in the Markov process (2.16) is of purely quantum-mechanical nature. Each state appearing in the expressions (2.14) ~ (2.16) is a pure quantum state.

3. ERGODICITY OF PURE QUANTUM STATES

The Hamiltonian H in (2.1) and the observable A in (2.2) must not commute,

$$[A, H]\neq 0, \tag{3.1}$$

in order that the observable A may be measured and identified at each instantaneous time. In addition to it, we shall suppose that the matrix

$$\|H_{mn}^{(\varphi)}\| \tag{3.2}$$

is irreducible, since if the Hamiltonian is further reducible, the kinetics would simply be a superposition of that of each sub-Hamiltonian which is irreducible. As a result, the Markov process (2.16) turns out to be irreducible.

To begin with, it is noted that the Markov process (2.16) can fix the *a priori* probability of expectation that an observer, who is not informed of any past history of events preceding the measurement, identifies an arbitrary quantum state out of N states given in (2.4) at an arbitrary time. One can thus prove the following Theorem[3]:

Theorem: Suppose an irreducible Markov process with N states $\{n\}$ $(n=1, 2, \ldots, N)$. Then, an observer who is not informed of any past history of events expects an arbitrary Markovian state n at an arbitrary time with the *a priori* probability[4]

$$\alpha_n=1/N, \quad n=1, 2, \ldots, N. \tag{3.3}$$

The present theorem determines the time average of the observable A of the form

$$\bar{A}=\lim_{T\to\infty}\frac{1}{T}\int_{t_0}^{t_0+T}A_{ob}(t)\,dt \tag{3.4}$$

with

$A_{ob}(t)=a_n$ if the quantum state identified by the observer at time t is n.

Here, one should distinguish the expression in (3.4) from another time average

$$\langle A\rangle=\lim_{T\to\infty}\frac{1}{T}\int_{t_0}^{t_0+T}\langle\Psi(t)|A|\Psi(t)\rangle\,dt \tag{3.5}$$

in which the wavefunction $\Psi(t)$ evolves through (2.1) without undergoing any sort of contraction such as caused by real measurements. The distinction comes from the observation that $\langle A\rangle$ in (3.5) is the expected time average of A which is not accompanied with identi-

fying quantum states at each instantaneous time in the interval T and that real measurements never enter in (3.5). On the other hand, the time average $\bar{A}$ in (3.4) is the average of a series of collected data in which the observer really identifies quantum states at each instantaneous time in the course of quantum-mechanical evolution.

Suppose that an observer makes an identification of quantum states belonging to $\{n\}$ of (2.4) at each interval of small τ starting from $t=t_0$. Thus, a set of identified quantum states is prepared. The probability of expectation that an arbitrary state belonging to the set is n turns out to be α_n $(=1/N)$ because of the definition of the *a priori* probability of expectation. An examination of the case with a strong inequality

$$T/\tau \gg N \quad (T \to \infty \text{ and } \tau \to +0) \tag{3.6}$$

gives the fact that the integral of the measured quantities in (3.4) reduces to

$$\bar{A} = \sum_{n=1}^{N} \alpha_n a_n \tag{3.7}$$

$$= \frac{1}{N} \sum_{n=1}^{N} a_n, \tag{3.8}$$

since if the integral in the sense of Lebesgue converges,[3] its asymptotic value cannot differ from the expected value (3.7).

The expression in (3.8) shows that the time average $\bar{A}$ in (3.4) is identical to the observable A averaged over all the available quantum states $\{n\}$ of (2.4). This gives the ergodicity of pure quantum states, in the sense that when the time average is taken the observable is really measured and identified at each instantaneous time. The wavefunction in (2.1) contracts at each identification of quantum state on the part of the observer. However, the expression given in the time average (3.5) does not exhibit such a contraction since real measurements are never employed. If one thinks of a time average of a real observable, it is understood that the procedure of measurement also goes along with it. Consequently, the time average of measured observables should be emphasized over the expected average, which does not collaborate with real measurements.

4. ERGODICITY OF MIXED QUANTUM STATES

Although the stochastic property presented in (2.16) is of truly quantum-mechanical nature, the equation of transitivity must be generalized into the equation of motion of the density matrix and interpreted in the context of quantum statistical mechanics instead of pure quantum mechanics if only limited measurements providing insufficient information are available in comparison with the former quantum mechanical case. The statistical mechanical description necessarily introduces the concept of ensemble in terms of which the degree of insufficient information is expressed. The probability of expectation as a diagonal element of the density matrix

$$\rho_n (= P(n, t \mid l, t_0)) \tag{4.1}$$

is now understood as the probability of event that a representative system out of many similar systems, each of which follows the stochastic evolution presented in (2.16), finds itself in the state n of (2.4). Therefore, the quantum state n presented in the probability of expectation (4.1) is not purely quantum-mechanical. The state should be a mixed quantum state, characteristic of a representative system belonging to an ensemble in statistical mechanics.[5,6]

The ensemble average, which has nothing common with the average of an observable over all the available pure quantum states, will be expressed in terms of the asymptotic probability of ensemble

$$\rho_n^{(\text{asy})} = \lim_{(t-t_0)\to\infty} P(n, t \mid l, t_0) \tag{4.2}$$

if the right-hand side of (4.2) certainly converges for an arbitrary state n and if random *a priori* phases are employed for the ensemble. The average of the observable A over the aged ensemble satisfying (4.2) leads to

$$\langle A \rangle_{\text{ens}} = \sum_{n=1}^{N} \rho_n^{(\text{asy})} a_n. \tag{4.3}$$

On the other hand, the time average of the observable A with respect to a single quantum-mechanical system has already been given in (3.8). The ergodicity of mixed quantum states, which says that the time average $\bar{A}$ in (3.8) should be equal to the ensemble average $\langle A \rangle_{\text{ens}}$ of a quantum statistical ensemble, gives

$$\rho_n^{(\text{asy})} = 1/N. \tag{4.4}$$

The condition that the ergodicity of mixed quantum states, by means of which quantum statistical mechanics could be justified, may hold is that each quantum state in a given aged ensemble has equal *a priori* probability. One must, however, be careful to the point that the principle of equal *a priori* probabilities presented in (4.4) is a consequence of the ergodicity of mixed quantum states. Furthermore, the ergodicity of mixed quantum states would be satisfied if and only if all of the asymptotic values of the probabilities of ensemble in (4.2) exist and are equal to each other.

5. CONCLUDING REMARKS

Given an arbitrary Hamiltonian system in quantum mechanics, the ergodicity of pure quantum states is maintained in the space of orthonormal quantum states which diagonalizes an observable. The corresponding Hamiltonian and the observable must not commute, since the observable is supposed to be measured and identified at each instantaneous time in the interval of taking its time average.

The present ergodicity in quantum mechanics should be compared with the similar one proved in metric-transitive dynamical systems.[2] However, the ergodicity of pure quantum states does not give a direct justification of quantum statistical mechanics, since only mixed quantum states are concerned in the latter. The ergodicity of mixed quantum states, which necessarily introduces an ensemble of many similar systems, will be justified only if the principle of equal *a priori* probabilities is maintained with respect to the ensemble. The principle is just an identical expression to the ergodicity of mixed quantum states. If this ergodicity does not hold, one must leave quantum statistical mechanics admitting both mixed quantum states and the associated ensemble for pure quantum mechanics in which, however, the ergodicity of pure quantum states persists.

[1]J. von Neumann, Z. Phys. 57, 30 (1929).
[2]D. ter Haar, *Elements of Statistical Mechanics* (Holt, Rinehart and Winston, New York, 1961), Appendix I and references cited therein.
[3]K. Matsuno, J. Math. Phys. 16, 604 (1975).
[4]Here, we understand that there exist at least two kinds of probability, each of which may follow the axioms of probability. One is the logical probability of expectation which is specific to a subjective observer. The other is the statistically inductive probability as relative frequency which is frequently used in ensemble theory of stochastic process, although it would not be certain whether the probability certainly converges as absolute frequency increases. The equal *a priori* probability of expectation should be distinguished from the principle of equal *a priori* probabilities (cf. Ref. 2 and Sec. 4 of this paper). The *a priori* probability of expectation is specific to an observer. However, *a priori* probabilities used in the context of the principle of equal *a priori* probabilities are not quite clear with respect to whether they refer to the logical probabilities of expectation or to the statistical probabilities as relative frequencies. In order to avoid unnecessary complexities, we use the term the principle of (*a priori*) equal weight in Ref. 3 in place of the principle of equal *a priori* probabilities.
[5]L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon, London, 1969).
[6]The time evolution of the density matrix is free from the contraction of wavefunctions. Hence, one is not sure whether any observable averaged over a statistical ensemble could certainly be physically measurable, since the contraction of wavefunctions, which is inevitable for real measurements, is absent in this scheme of quantum statistics.

Note on the computation formula of the boost matrices of SO(n-1,1) and continuation to the d matrices of SO(n)

Takayoshi Maekawa

Department of Physics, Kumamoto University, Kumamoto, Japan
(Received 5 May 1975)

The general formula of computing the boost matrices of SO($n-1,1$), which is valid for the single- and double-valued representations and is similar to that of Vilenkin and Wolf, is given. It is noted that a phase factor of a unit magnitude in the boost matrices must be taken into account in analytic continuation to the d matrices of SO(n), and then the formula of computing the d matrices of SO(n) is given. It is remarked that the d matrices of SO(n) are expressed in terms of those of SO($n-1$).

In previous papers,[1,2] the d matrices of SO(n) and the boost matrices of SO($n-1,1$) are studied and recursion relations among the matrix elements are obtained. From the relations for SO(n), it is shown that the D matrices can be determined completely and be expressed in terms of the lowering operators which consist of the differential operators with respect to the Euler angles. For SO($n-1,1$), however, the boost matrices cannot be computed in a general form.

The formula for computing these matrices is given by Vilenkin[3] and Wolf[4] as an integral of the d matrices of SO($n-1$). Vilenkin's formula is given independent of the single- and double-valued representations but does not contain a summation over some numbers, contrary to Wolf's formula, which is derived for a single-valued representation. Both of them do not take into account continuation of a factor which may be equated to 1 in the boost matrices but plays an important role in the d matrices of SO(n), and the d matrices without the factor do not satisfy the unitarity condition as easily seen in special cases. The purpose of this note is to give the general formulas for computing the unitary irreducible representation matrices of SO($n-1,1$) (principal series) and SO(n).

Let us introduce the Gel'fand and Tsetlin bases[5] of the unitary irreducible representation of SO(n) which are characterized by the canonical chain of subgroups $SO(n)\supset SO(n-1)\supset\cdots\supset SO(2)$,

$$|m_{jk}\rangle=|\lambda_n,\lambda_{n-1},\ldots,\lambda_2\rangle, \tag{1}$$

where λ_j stands for the row $(m_{j1},m_{j2},\ldots,m_{j[j/2]})$, all λ_j are written in a row, and $[j/2]$ is the largest integer smaller than or equal to $j/2$. The numbers m_{jk} take integers or half-integers simultaneously with restrictions

$$\begin{aligned}&m_{2j+1,i}\geq m_{2j,i}\geq m_{2j+1,i+1}\quad(i=1,2,\ldots,j),\\&m_{2j,i}\geq m_{2j-1,i}\geq m_{2j,i+1}\quad(i=1,2,\ldots,j-1),\\&m_{2j+1,j}\geq|m_{2j,j}|.\end{aligned} \tag{2}$$

The dimension of the unitary irreducible representation of SO(n) is determined by the numbers m_{nj} ($j=1,2,\ldots,[n/2]$) subjected to the restrictions (2), and given as follows[2] for n odd,

$$N(\lambda_n)=\left((n-2)!\,(n-4)!!\prod_{j=1}^{(n-5)/2}(n-3-2j)!\right)^{-1}\prod_{i=1}^{(n-1)/2}(2l_{ni}+1)\times\prod_{j=1}^{(n-3)/2}\prod_{k=j+1}^{(n-1)/2}\{l_{nj}(l_{nj}+1)-l_{nk}(l_{nk}+1)\}, \tag{3}$$

and for n even,

$$N(\lambda_n)=2^{(n/2)-1}\left((n-2)!\,(n-4)!!\prod_{j=1}^{(n-4)/2}(n-3-2j)!\right)^{-1}\times\prod_{j=1}^{(n-2)/2}\prod_{k=j+1}^{n/2}(l_{nj}^2-l_{nk}^2), \tag{4}$$

where

$$l_{nj}=m_{nj}+[n/2]-j\quad(j=1,2,\ldots,[n/2]).$$

The representation matrix of SO(n) may be parametrized in terms of the Euler angles as follows[1,3,4]:

$$D^{(n)}(\{\theta_n\})=D^{(n-1)}(\{\theta_{n-1}\})\left\{\prod_{j=n}^{4}R_{jj-1}(\theta_{nn-j+1})\right\}\times R_{31}(\theta_{nn-2})R_{12}(\theta_{nn-1}), \tag{5}$$

where

$$\{\theta_n\}\equiv(\theta_{21},\ldots,\theta_{nn-1}),\quad 0\leq\theta_{jk}\leq\pi,\quad k=1,2,\ldots,j-2,\quad 0\leq\theta_{jj-1}<2\pi,\quad j=1,2,\ldots,n.$$

$D^{(n-1,1)}$ of SO($n-1,1$) is obtained from (5) by substituting $R_{nn-1}(\theta_{n1})$ into $R^b_{nn-1}(\zeta)$ ($0\leq\zeta<\infty$), where $R^b_{nn-1}(\zeta)$ is the boost in the $(n-1)$th direction through ζ.

The invariant measure of SO(n) is given as follows:

$$\begin{aligned}&dV_n=dS_n\,dV_{n-1},\quad dV_2=d\theta_{21},\\&dS_n=\prod_{j=1}^{n-1}\{(\sin\theta_{nj})^{n-j-1}\,d\theta_{nn-j}\},\end{aligned} \tag{6}$$

where dS_n is the surface element of a sphere in a n-dimensional space. The volume of SO(n), V_n, is given by $V_n=V_{n-1}\,2\pi^{n/2}/\Gamma(n/2)$ and $V_2=2\pi$.

The representation D matrices of SO(n) are calculated through

$$D^{(\lambda_n)}_{\bar\lambda_{n-1}\bar\lambda'_{n-1}}(\{\theta_n\})=\langle\lambda_n\bar\lambda_{n-1}|\,D^{(n)}(\{\theta_n\})\,|\lambda_n\bar\lambda'_{n-1}\rangle, \tag{7}$$

where the notations such as $\bar\lambda_{n-1}=(\lambda_{n-1},\ldots,\lambda_2)$ are introduced. The d matrices of SO(n) and the boost matrices of SO($n-1,1$) are defined as follows:

$$d^{(\lambda_n)}_{\lambda_{n-1}(\lambda_{n-2})\lambda'_{n-1}}(\theta)=\langle\lambda_n\lambda_{n-1}\bar\lambda_{n-2}|\,R_{nn-1}(\theta)\,|\lambda_n\lambda'_{n-1}\bar\lambda_{n-2}\rangle, \tag{8}$$

$${}^bd^{(\lambda_n)}_{\lambda_{n-1}(\lambda_{n-2})\lambda'_{n-1}}(\zeta)=\langle\lambda_n\lambda_{n-1}\bar\lambda_{n-2}|\,R^b_{nn-1}(\zeta)\,|\lambda_n\lambda'_{n-1}\bar\lambda_{n-2}\rangle. \tag{9}$$

Here, the same notations for the bases are used in both

the d and boost matrices, since the Gel'fand and Tsetlin bases for $SO(n-1,1)$ are classified by the canonical chain $SO(n-1,1)\supset SO(n-1)\supset\cdots\supset SO(2)$ and can be written in the form (1) in which all numbers except m_{n1} ($\equiv\rho_n$) are subject to the restrictions (2).[6] The ρ_n takes the value $(2-n)/2+i\nu$, ν real, for the principal series of the unitary irreducible representation of $SO(n-1,1)$.[7]

The orthogonality and completeness relations for the D matrices of $SO(n)$ are given as follows[4,8]:

$$\int_{SO(n)} dV_n \overline{D^{(\lambda_n)}_{\bar\lambda_{n-1}\bar\lambda'_{n-1}}(\{\theta_n\})}\, D^{(\lambda'_n)}_{\bar\lambda_{n-1}\bar\lambda'_{n-1}}(\{\theta_n\}) = \delta_{\bar\lambda_{n-1}\bar\lambda'_{n-1}}\delta_{\bar\lambda'_{n-1}\bar\lambda''_{n-1}}\delta_{\lambda_n\lambda'_n}V_n/N(\lambda_n), \tag{10}$$

$$\sum_{\lambda_n}\frac{N(\lambda_n)}{V_n}\sum_{\bar\lambda_{n-1}\bar\lambda'_{n-1}}\overline{D^{(\lambda_n)}_{\bar\lambda_{n-1}\bar\lambda'_{n-1}}(\{\theta_n\})}D^{(\lambda_n)}_{\bar\lambda_{n-1}\bar\lambda'_{n-1}}(\{\theta'_n\}) = \delta_{(n)}(\{\theta_n\},\{\theta'_n\}), \tag{11}$$

where $\delta_{\lambda\lambda'}\cdots$ stand for a product of Kronecker δ's in the individual indices. The expression for $\delta_{(n)}(\cdot,\cdot)$ is given as follows:

$$\delta_{(n)}(\{\theta_n\},\{\theta'_n\}) = \left\{\prod_{j=1}^{n-1}(\sin\theta_{nj})^{j-n+1}\delta(\theta_{nj}-\theta'_{nj})\right\}\times\delta_{(n-1)}(\{\theta_{n-1}\},\{\theta'_{n-1}\}), \tag{12}$$

$$\delta_{(2)}(\{\theta_2\},\{\theta'_2\}) = \delta(\theta_{21}-\theta'_{21}).$$

From (10) and (11), we obtain the following relations for the d matrices[4]:

$$\sum_{\lambda_{n-2}}N(\lambda_{n-2})\int_0^\pi d\theta\sin^{n-2}\theta\,\overline{d^{(\lambda_n)}_{\lambda_{n-1}\lambda(n-2)\lambda'_{n-1}}(\theta)}\,d^{(\lambda'_n)}_{\lambda_{n-1}(\lambda_{n-2})\lambda'_{n-1}}(\theta) = \delta_{\lambda'_n\lambda_n}\frac{\sqrt{\pi}\Gamma((n-1)/2)N(\lambda_{n-1})N(\lambda'_{n-1})}{\Gamma(n/2)N(\lambda_n)}, \tag{13}$$

$$\sum_{\lambda_n}N(\lambda_n)\overline{d^{(\lambda_n)}_{\lambda_{n-1}(\lambda_{n-2})\lambda'_{n-1}}(\theta)}\,d^{(\lambda_n)}_{\lambda_{n-1}(\lambda''_{n-2})\lambda''_{n-1}}(\theta') = \delta_{\lambda_{n-2}\lambda''_{n-2}}\frac{\delta(\theta-\theta')\sqrt{\pi}\Gamma((n-1)/2)N(\lambda_{n-1})N(\lambda'_{n-1})}{(\sin\theta)^{n-2}\Gamma(n/2)N(\lambda_{n-2})}. \tag{14}$$

The relations (13) and (14) hold for the d matrices of the single- and double-valued representations.

The d matrices of $SO(n)$ and the boost matrices of $SO(n-1,1)$ can be determined by the d matrices of $SO(n-1)$.[3,4] We can give the expression for the boost matrices of the principal series in the form[3,4]

$${}^bd^{(\lambda_n)}_{\lambda_{n-1}(\lambda''_{n-2})\lambda_{n-1}}(\xi) = \frac{P([n-1]/2)}{\sqrt{\pi}\Gamma([n-2]/2)}\frac{\sqrt{N(\lambda_{n-1})N(\lambda'_{n-1})}}{N(\lambda_{n-2})N(\lambda'_{n-2})}\omega_1(\lambda_{n-1},\lambda'_{n-1})$$
$$\times\sum_{\lambda_{n-3}}N(\lambda_{n-3})\int_0^\pi d\theta\sin^{n-3}\theta\,\overline{d^{(\lambda_{n-1})}_{\lambda_{n-2}(\lambda_{n-3})\lambda'_{n-2}}(\theta)}(\cosh\xi-\cos\theta\sinh\xi)^{\rho_n}$$
$$\times d^{(\lambda'_{n-1})}_{\lambda_{n-2}(\lambda_{n-3})\lambda'_{n-2}}(\theta'), \tag{15}$$

where $\lambda_n\equiv(\rho_n,\lambda_{n-2})$ and

$$\cos\theta' = \frac{\cos\theta\cosh\xi-\sinh\xi}{\cosh\xi-\cos\theta\sinh\xi},$$

$$\omega_1(\lambda_{n-1},\lambda'_{n-1}) = \left(\prod_{j=1}^{[(n-1/2)]}\frac{\Gamma(\rho_n+m_{n-1j}+n-j-1)\Gamma(m'_{n-1j}-\rho_n-j+1)}{\Gamma(m_{n-1j}-\rho_n-j+1)\Gamma(\rho_n+m'_{n-1j}+n-j-1)}\right)^{1/2}.$$

$\omega_1(\lambda_{n-1},\lambda'_{n-1})$ is a phase factor which may be equated to 1 for $\rho_n=(2-n)/2+i\nu$, but becomes an important factor in the d matrices of $SO(n)$.

In fact, it is easy to show that (15) satisfies the group properties and the unitarity condition, i.e.,

$$\sum_{\lambda''_{n-1}}{}^bd^{(\lambda_n)}_{\lambda_{n-1}(\lambda_{n-2})\lambda''_{n-1}}(\xi_1)\,{}^bd^{(\lambda_n)}_{\lambda''_{n-1}(\lambda_{n-2})\lambda'_{n-1}}(\xi_2) = {}^bd^{(\lambda_n)}_{\lambda_{n-1}(\lambda_{n-2})\lambda'_{n-1}}(\xi_1+\xi_2),$$
$${}^bd^{(\lambda_n)}_{\lambda_{n-1}(\lambda_{n-2})\lambda'_{n-1}}(0) = \delta_{\lambda_{n-1}\lambda'_{n-1}}, \tag{16}$$
$${}^bd^{(\lambda_n)}_{\lambda_{n-1}(\lambda'_{n-2})\lambda'_{n-1}}(\xi) = {}^bd^{(\lambda_n)}_{\lambda'_{n-1}(\lambda'_{n-2})\lambda_{n-1}}(-\xi).$$

(13) and (14) are used to derive the first and second relations, and the third is derived under the condition $\rho_n+\rho_n^*+n-2=0$, which is satisfied by $\rho_n=(2-n)/2+i\nu$, ν real. Therefore, it follows that the expression (15) provides the boost matrices for the principal series of the unitary irreducible representation of $SO(n-1,1)$, and we may conclude that formula (15) holds for both the single- and the double-valued representations, for the relations (13) and (14) are used in the proof of (16).

The d matrices of $SO(n)$ will be obtained from (15) by analytic continuation, i.e., $\xi\to -i\theta_{n1}$ and $\rho_n\to m_{n1}$, integer or half-integer. We must, however, take into account the continuation of the phase factor $\omega_1(\lambda_{n-1},\lambda'_{n-1})$ which does not change (16), but may give rise to the result different from 1 for the d matrices. Actually, we obtain the following expression for the d matrices:

$$d^{(\lambda_n)}_{\lambda_{n-1}(\lambda'_{n-2})\lambda'_{n-1}}(\theta_1) = (\omega_1\to\omega_2,\ \xi\to -i\theta_1,\ \rho_n\to m_{n1}\text{ in (15)}), \tag{17}$$

where

$$\lambda_n=(m_{n1},\ \lambda_{n-2}),$$
$$\cos\theta' = \frac{\cos\theta\cos\theta_1+i\sin\theta_1}{\cos\theta_1+i\cos\theta\sin\theta_1},$$

$$\omega_2(\lambda_{n-1},\lambda'_{n-1}) = i^{\sum_{j=1}^{[(n-1)/2]}(m_{n-1j}-m'_{n-1j})}$$
$$\times\left(\prod_{j=1}^{[(n-1)/2]}\frac{\Gamma(m_{n1}+m_{n-1j}+n-j-1)\Gamma(m_{n1}-m_{n-1j}+j)}{\Gamma(m_{n1}+m'_{n-1j}+n-j-1)\Gamma(m_{n1}-m'_{n-1j}+j)}\right)^{1/2}.$$

It is easy to see that ω_1 and ω_2 are connected with each other by the continuation. The expressions for ω_1 and ω_2 are obtained by considering the infinitesimal transformation and the explicit forms of D_{nn-1}.[5,6,1]

Formulas (15) and (17) will be useful for the computation of the boost and the d matrices of $SO(n-1,1)$ and $SO(n)$. For $SO(3)$, it is necessary, in order to make the result coincide with the usual one, that we divide the right-hand side of (17) by 2, take the integral from 0 to 2π, and change a factor i in $\omega_2(\lambda_2,\lambda'_2)$ into $-i$, because of the usual definition of the d matrices with $R_{31}(\theta)$ instead of $R_{32}(\theta)$. For $SO(n)$, it is sufficient for us to compute the d matrices at $\theta_1=\pi/2$, for we have the relation[9]

$$R_{nn-1}(\theta) = UR_{n-1\,n-2}(\theta)U^{-1},$$
$$U\equiv R_{21}(\tfrac{\pi}{2})R_{31}(\tfrac{\pi}{2})\prod_{j=4}^{n}R_{j\,j-1}(\tfrac{\pi}{2}). \tag{18}$$

The d matrices of SO(n) can, therefore, be expressed in terms of the finite sums of those of SO($n-1$), i.e.,

$$d^{(\lambda_n)}_{\lambda_{n-1}(\lambda_{n-2})\lambda'_{n-1}}(\theta) = \sum_{\lambda''_{n-1}\lambda''_{n-2}\lambda'_{n-2}\lambda'_{n-3}} a(\lambda''_{n-1}\lambda''_{n-2}\lambda'_{n-2}\lambda'_{n-3})\, d^{(\lambda''_{n-1})}_{\lambda_{n-2}(\lambda'_{n-3})\lambda''_{n-2}}(\theta), \quad (19)$$

where

$$a(\lambda''_{n-1}\lambda''_{n-2}\lambda'_{n-2}\lambda'_{n-3}) = \sum_{\lambda^1_2 \cdots \lambda'_{n-4}} \langle \lambda_n \lambda_{n-1} \overline{\lambda}_{n-2} | \cup | \lambda_n \lambda''_{n-1} \overline{\lambda}'_{n-2} \rangle \times \langle \lambda_n \lambda''_{n-1} \lambda''_{n-2} \overline{\lambda'}_{n-3} | \cup^{-1} | \lambda_n \lambda'_{n-1} \overline{\lambda}_{n-2} \rangle.$$

Note added in proof: Finally, it is noted that the ${}^I d$ matrices for the ISO($n-1$) group (inhomogenous orthogonal group), which denote the matrices corresponding to a translation along the $(n-2)$th direction, can be obtained from (15) by contraction, i.e., $\rho_n \to i\infty$ with $i\rho_n\xi = \gamma\xi$, γ real and $0 \leq \xi < \infty$. It is seen that the expression for the ${}^I d$ matrices is given by replacing $\omega_1(\lambda_{n-1}, \lambda'_{n-1}) \to \exp(\frac{1}{2}\pi i \sum(m_{n-1\,j} - m'_{n-1\,j}))$, $\theta' \to \theta$ and $(\cosh\zeta - \cos\theta \times \sinh\zeta)^{\rho_n} \to \exp(i\gamma\xi\cos\theta)$ in (15), and the result obtained by this procedure agrees with that in Ref. 4 except for a phase factor.

[1] T. Maekawa, J. Math. Phys. 16, 334 (1975).
[2] T. Maekawa (unpublished).
[3] N. Ja. Vilenkin, Dokl. Akad. Nauk SSSR 113, 16 (1957) (Russian).
[4] K. B. Wolf, J. Math. Phys. 12, 197 (1971).
[5] I. M. Gel'fand and M. L. Tsetlin, Dokl. Akad. Nauk SSSR 71, 1017 (1950) (Russian).
[6] A. Chakrabarti, J. Math. Phys. 9, 2087 (1968).
[7] F. Schwartz, J. Math. Phys. 12, 131 (1971); M. Ikeda and T. Maekawa, Math. Jpn. 16, 173 (1971).
[8] N. Ja. Vilenkin, *Special Functions and the Theory of Group Representations*, Math. Monographs Vol. 22, translated by V. N. Singh (American Mathematical Society, Providence, R.I., 1968).
[9] T. Maekawa, Math. Jpn. 19, 133 (1974).

Retarded multipole fields and the inhomogeneous wave equation*

W. E. Couch and R. J. Torrence

Department of Mathematics, University of Calgary, Calgary, Alberta, Canada T2N 1N4
(Received 20 June 1975)

The inhomogeneous wave equation for a class of driving terms that arise in certain physical problems is analyzed by introducing a kind of "inner product" g with respect to which the 2^l-pole solutions, ψ_l, of the homogeneous wave equation are an orthogonal basis. This allows the condition that δ, the Lth multipole part of the driving term, will give rise to a nonspreading solution to be expressed as $g(\psi_L, r^2\delta) = 0$. The complete solution is found in terms of its spreading and nonspreading parts, and the backscattered radiation is calculated from the spreading part.

1. INTRODUCTION

In a recent paper[1] the authors considered the axially symmetric inhomogeneous scalar wave equation

$$\Box\Phi = \Delta \tag{1.1}$$

for a particular class of physically interesting driving terms. Spherical coordinates (r, θ, φ) were used along with the expansions

$$\Phi = \sum_{L=0}^{\infty} \phi_L Y_{L0},$$

$$\Delta = \sum_{L=0}^{\infty} \delta_L Y_{L0},$$

in spherical harmonics Y_{L0}. In what follows we will fix L and omit the subscript on δ_L and ϕ_L. We restricted ourselves to driving terms which are of the form

$$\delta = \sum_{\alpha=3}^{T+1} \frac{d_\alpha(u)}{r^\alpha}, \tag{1.2}$$

where T is a positive integer and $u = t - r$ is retarded time, and which satisfy certain other requirements,[2] which will be discussed later. For this class of driving terms, a condition on δ necessary and sufficient for the corresponding ϕ to be nonspreading[3] was derived. A solution was found to be nonspreading if and only if it is a finite series of the form

$$\phi = \sum_{\alpha=0}^{T-2} \frac{f_\alpha(u)}{r^{\alpha+1}}, \tag{1.3}$$

and the condition on δ equivalent to this was shown to be

$$\sum_{\alpha=0}^{T-L-2} \frac{(-2)^\alpha(\alpha+L)!}{\alpha!(\alpha+2L+1)!} \frac{d^\alpha}{du^\alpha} d_{\alpha+L+3}(u) = 0. \tag{1.4}$$

In this paper we introduce a method of treating Eq. (1.1) which uses the retarded 2^l-pole solutions, ψ_l, of the homogeneous wave equation as a basis in which functions of u and r that are finite series in $1/r$ may be expanded. In particular, the quantities $r^2\delta$ and ϕ of the type given by Eqs. (1.2) and (1.3), respectively, can always be written as a unique finite linear combination of multipole fields, ψ_l, with appropriately chosen moment functions. An "inner product", g, is introduced on the relevant space of functions of two variables in such a way that, for example, the coefficients of expansion for ϕ are obtained in the usual way, that is, from $g(\psi_l, \phi)$. This inner product is given by a sum, rather than being given by an integral over the functions involved.

We use the inner product structure to accomplish several things. We give a suggestive derivation of the condition for nonspreading; in fact, it will be easy to see that in terms of g the condition for nonspreading is that $g(\psi_L, r^2\delta) = 0$, i.e., that the $l = L$ term in the expansion of $r^2\delta$ in terms of the ψ_l's be absent. Thus our method and the resulting condition are analogous to well-known theorems concerning the existence of solutions to an inhomogeneous Sturm—Liouville problem in the case of ordinary differential equations.[4] In addition we give the solution, ϕ, in the form of a nonspreading part plus a spreading part and use g to obtain a simple form for the former.

The spreading part of ϕ also has a simple form, and we use it to calculate the backscattered radiation. The form of the spreading part of ϕ also enables us to give a simple statement of the requirements which along with Eq. (1.2) define the class of driving terms that we consider.

2. THE BASIS OF MULTIPOLE FIELDS

Let ψ be the scalar field in Minkowski space, and expand ψ in spherical harmonics. The coefficients, $\psi_l(u, r)$, will satisfy

$$2\frac{\partial^2\psi_l}{\partial u\partial r} - \frac{\partial^2\psi_l}{\partial r^2} + \frac{2}{r}\left(\frac{\partial\psi_l}{\partial u} - \frac{\partial\psi_l}{\partial r}\right) + \frac{l(l+1)\psi_l}{r^2}$$
$$\equiv \Box_l\psi_l = 0, \quad l \geq 0, \tag{2.1}$$

where $u = t - r$. If we introduce the infinite lower triangular matrix

$$C_l{}^\alpha = \begin{cases} (l+\alpha)!/2^\alpha\alpha!(l-\alpha)!, & \alpha \leq l, \\ 0, & \alpha > l, \end{cases} \tag{2.2}$$

and the related matrix of differential operators

$$\mathcal{C}_l{}^\alpha = C_l{}^\alpha \frac{d^{l-\alpha}}{du^{l-\alpha}}, \tag{2.3}$$

then Eq. (2.1) is satisfied by the finite series

$$\psi_l = \sum_{\alpha=0}^{l} \frac{\mathcal{C}_l{}^\alpha a_l(u)}{r^{\alpha+1}}, \tag{2.4}$$

where $a_l(u)$, the lth multipole moment of the retarded field, is any sufficiently differentiable function of u.

Let Φ be the space of functions of two variables of the form

$$\phi=\sum_{\alpha=0}^{T-2}\frac{f_\alpha(u)}{r^{\alpha+1}},$$

where $f_\alpha(u)$ are bounded and sufficiently smooth functions. We introduce a kind of "basis" in Φ with elements defined by

$$R_\alpha\equiv 1/r^{\alpha+1},\quad \alpha\geq 0. \tag{2.5}$$

Obviously any member of Φ can be expanded in this basis; however, the expansion coefficients will be function of u. We now introduce a second basis in Φ. If we define the operators

$$\Psi_l\equiv\sum_{\alpha=0}^{l}C_l{}^\alpha R_\alpha, \tag{2.6}$$

then the set of Ψ_l can also serve as a basis in Φ. In particular $\psi_l=\Psi_l a_l(u)$, where, since the Ψ_l are operators on their coefficients, we shall always write their coefficients on their right. Then Eq. (2.6) is the transformation between bases.

We next introduce an "inner product" g on Φ by defining a bilinear map from $\Phi\times\Phi$ to the set of functions of u:

$$g(\Psi_l a_l(u),\Psi_{l'}a_{l'}(u))\equiv g_{ll'}\equiv\delta_{ll'}\bar{a}_l(u)a_{l'}(u). \tag{2.7}$$

If we have ϕ, $\chi\in\Phi$ with $g(\phi,\chi)=0$, we shall say that ϕ and χ are orthogonal, so we have made the Ψ_l's into an orthogonal basis. If we transform back to the R_α basis, we find that

$$g(R_\alpha f_\alpha(u),R_\beta f_\beta(u))\equiv g_{\alpha\beta}=\sum_{l=0}^{\infty}(C^{-1}{}_\alpha{}^l\bar{f}_\alpha)(C^{-1}{}_\beta{}^l f_\beta), \tag{2.8}$$

where

$$C^{-1}{}_\alpha{}^l=C^{-1}{}_\alpha{}^l\frac{d^{\alpha-l}}{du^{\alpha-l}}, \tag{2.9}$$

with

$$C^{-1}{}_\alpha{}^l=\begin{cases}\dfrac{(-1)^{\alpha+l}2^\alpha\alpha!(2l+1)}{(\alpha-l)!(\alpha+l+1)!}, & l\leq\alpha,\\ 0, & l>\alpha.\end{cases} \tag{2.10}$$

The matrix $C^{-1}{}_\alpha{}^l$ is lower triangular and is the inverse of $C_l{}^\alpha$. This follows indirectly from Ref. 1. Likewise the matrix of differential operators, $C^{-1}{}_\alpha{}^l$, is inverse to $C_l{}^\alpha$.

3. SOLUTION OF THE INHOMOGENEOUS EQUATION

Consider the equation

$$\Box_L\phi=\delta \tag{3.1}$$

with $\delta\in\Phi$ and given by Eq. (1.2). We now derive an expression for ϕ which (1) is a simple form for the solution, (2) clearly reveals requirements on δ that are sufficient for the solution to exist, (3) provides an expression for the backscattered radiation when the field is spreading, and (4) shows that when δ meets the requirements of (2) the condition given by Eq. (1.4) is equivalent to ϕ being nonspreading.

The driving term naturally occurs in the R_α basis as in Eq. (1.2), and we will ultimately express ϕ and the nonspreading condition in terms of the $d_\alpha(u)$. The necessary calculations are most efficiently done in the Ψ_l basis using the inner product g.

The expansion for $r^2\delta$ in the Ψ_l basis is

$$r^2\delta=\sum_{l=0}^{T-2}\Psi_l A_l(u), \tag{3.2}$$

where the expansion coefficients $A_l(u)$ are found from $g(\psi_l,r^2\delta)$. We have

$$g(\psi_l,r^2\delta)=\sum_{\alpha=0}^{\infty}\sum_{\beta=0}^{\infty}\sum_{n=0}^{\infty}C^{-1}{}_\alpha{}^n(\bar{\psi}_l)^\alpha C^{-1}{}_\beta{}^n(r^2\delta)^\beta, \tag{3.3}$$

where $(\psi_l)^\alpha$ denotes the α-component of ψ_l in the R_α basis, and similarly for $(r^2\delta)^\beta$. These components are given by $(\psi_l)^\alpha=C_l{}^\alpha a_l(u)$ and $(r^2\delta)^\beta=d_{\beta+3}(u)$, and when they are used in Eq. (3.3) we obtain

$$g(\psi_l,r^2\delta)=\bar{a}_l(u)\sum_{\beta=0}^{T-2}C^{-1}{}_\beta{}^l d_{\beta+3}(u). \tag{3.4}$$

Hence it is seen from Eq. (2.7) that Eq. (3.2) is the expansion of $r^2\delta$ if the $A_l(u)$ are given by

$$A_l(u)=\sum_{\beta=0}^{T-2}C^{-1}{}_\beta{}^l d_{\beta+3}(u). \tag{3.5}$$

Since

$$\Box_L\psi_l=[L(L+1)-l(l+1)]r^{-2}\psi_l, \tag{3.6}$$

we see that every term in $r^2\delta$ except the one for which $l=L$ can be generated by $\Box_L[L(L+1)-l(l+1)]^{-1}\Psi_l A_l$; hence ϕ may be taken to be a series in $\Psi_l A_l(u)$, $l\neq L$, which solves Eq. (3.1) for all except the $l=L$ part of $r^2\delta$, plus a term Λ, which solves the $l=L$ part. That is, we have

$$\phi=\sum_{\substack{l=0\\ l\neq L}}^{T-2}[L(L+1)-l(l+1)]^{-1}\Psi_l A_l(u)+\Lambda, \tag{3.7}$$

where Λ must satisfy

$$\Box_L\Lambda=\frac{\Psi_L A_L}{r^2}=\sum_{\alpha=0}^{L}\frac{C_L{}^\alpha A_L}{r^{\alpha+3}}, \tag{3.8}$$

or, for $L\geq 1$,

$$\Box_L\Lambda=\sum_{\alpha=0}^{L-1}\frac{C_L{}^\alpha A_L}{r^{\alpha+3}}+\frac{C_L{}^L A_L}{r^{L+3}}\equiv\delta_\Lambda, \tag{3.9}$$

with A_L given by Eq. (3.5) with $l=L$. The case $T-2<L$ is trivially nonspreading with ϕ given by Eq. (3.7) and $A_L=\Lambda=0$, and in what follows we assume $T-2\geq L$.

The driving term δ_Λ is a member of Φ, and only the last term in $r^2\delta_\Lambda$ has an $l=L$ part in its expansion in the Ψ_l basis. We find Λ by applying to Eq. (3.9) the process which led to Eq. (3.7). The expansion coefficients of

$$r^2\sum_{\alpha=0}^{L-1}\frac{C_L{}^\alpha A_L}{r^{\alpha+3}}$$

are, by Eq. (3.5), seen to be

$$\sum_{\beta=0}^{L-1}C^{-1}{}_\beta{}^l C_L{}^\beta A_L=\delta_L{}^l A_L-C^{-1}{}_L{}^l C_L{}^L\frac{d^{L-l}}{du^{L-l}}A_L(u),$$
$$l=0,1,\ldots,(L-1). \tag{3.10}$$

Hence we have

$$\phi=\sum_{\substack{l=0\\ l\neq L}}^{T-2}[L(L+1)-l(l+1)]^{-1}\Psi_l A_l(u)$$
$$-\sum_{l=0}^{L-1}[L(L+1)-l(l+1)]^{-1}$$
$$\times C^{-1}{}_L{}^l C_L{}^L\Psi_l\frac{d^{L-l}}{du^{L-l}}A_L(u)+S, \tag{3.11}$$

where S must satisfy

$$\Box_L S = C_L{}^L A_L(u)/r^{L+3}, \quad L \geq 0. \tag{3.12}$$

For $L=0$, ϕ is given by Eq. (3.11) with the second series omitted.

By direct substitution into Eq. (3.12) it can be seen that a solution for S is

$$S = \frac{(-1)^{L+1}}{2^L(2L+1)} r^L \frac{\partial^L}{\partial r^L}\left(\frac{1}{r^{L+1}} \int_{-\infty}^{u} \frac{A_L(x)\,dx}{u+2r-x}\right). \tag{3.13}$$

We eliminate the freedom of adding to S a term

$$r^L \frac{\partial^L}{\partial r^L}\left(\frac{b(v)}{r^{L+1}}\right),$$

where $b(v)$ is an arbitrary function of $v \equiv u+2r$ by the physical requirement that radiation coming in from infinity is not present.

Equation (3.11) is our desired expression for ϕ with S given by Eq. (3.13). The moment functions $A_l(u)$ are given by Eq. (3.5) for all l; in particular, for $A_L(u)$ we have, after shifting the sum by $\beta \equiv \alpha + L$,

$$A_L(u) = 2^L(2L+1) \sum_{\alpha=0}^{T-L-2} \frac{(-2)^\alpha (\alpha+L)!}{\alpha!(\alpha+2L+1)!} \frac{d^\alpha}{du^\alpha} d_{\alpha+L+3}(u). \tag{3.14}$$

It is now seen from Eq. (3.13) that a requirement on δ sufficient to insure that the solution ϕ be given by Eq. (3.11) is that the integral

$$\int_{-\infty}^{u} \frac{A_L(x)\,dx}{(u+2r-x)}$$

and a sufficient number of its derivatives exist. We assume that the original driving term δ given by Eq. (1.2) satisfies this requirement.

In order to obtain a simple form for the backscattered radiation, consider the special case that all the $d_\alpha(u)$ have compact support in some interval $[u_1, u_2]$. Then all the $A_l(u)$ have this same support and ϕ vanishes for $u < u_1$ and in the region $u > u_2$ ϕ is given by

$$\phi = r^L \frac{\partial^L}{\partial r^L}\left(\frac{B(v)}{r^{L+1}}\right), \tag{3.15a}$$

$$B(v) \equiv \frac{(-1)^{L+1}}{2^L(2L+1)} \int_{u_1}^{u_2} \frac{A_L(x)\,dx}{(u+2r-x)}. \tag{3.15b}$$

This is, in the region $u > u_2$, a purely advanced radiation field with advanced moment function $B(v)$; this gives the backscattered radiation.

It is seen from Eqs. (3.11) and (3.13) that S is the only part of ϕ that has a nonlocal dependence on δ; hence ϕ is nonspreading if and only if $S=0$. But $S=0$ if and only if $A_L=0$. Hence the condition on δ such that ϕ will be a nonspreading field is $g(\psi_L, r^2\delta)=0$, and we see from Eq. (3.14) that this is the same as Eq. (1.4), the original condition derived in Ref. 1.

4. CONCLUSION

A proper vector space structure could have been introduced on Φ in Sec. 2 by introducing such a structure onto the set of coefficients $f_\alpha(u)$; then a proper inner product could have been defined by

$$g(\Psi_l a_l(u), \Psi_{l'} a_{l'}(u)) = \delta_{ll'} \int_{-\infty}^{\infty} \bar{a}_l(u)\, a_{l'}(u)\,du. \tag{4.1}$$

This more conventional and complicated approach was unnecessary in the context of this paper, but a more elaborate approach may be justified by other applications which we are considering, for example, a derivation of the formula for the Clebsch–Gordan coefficients of the Lorentz group.

The presence of functional coefficients in Φ is a direct result of our refusal to do a harmonic analysis with respect to u. Since the problem can be done without it, such an analysis seems to obscure the physical problem without a compensating benefit. It is of interest that a harmonic analysis with respect to u would leave us with finite (Bessel) series in $1/r$, while one with respect to t would yield infinite series (Bessel functions). This simple distinction has been utilized by other authors,[5,6] but without the explicit comment it deserves.

A related fact is the nature of the inner product defined here. Orthogonality of Bessel functions of different order can be defined with respect to an inner product defined by an integral.[7] By working with u instead of t, the orthogonality defined in this paper is between Bessel *series* of different order, and as demonstrated here, is with respect to an inner product defined by a sum.

ACKNOWLEDGMENTS

The authors wish to thank D.R. Westbrook for comments which prompted the approach to the condition for nonspreading which is developed here and A.I. Janis for comments concerning the class of driving terms to which the results apply.

*This work has been supported by the National Research Council of Canada.

[1] W.E. Couch and R.J. Torrence, J. Math. Phys. 16, 857 (1975); 16, 2180 (1975), Erratum.

[2] These requirements were considered in the Erratum of Ref. 1. The results of this paper enable us to give a more precise statement of these requirements.

[3] By the term nonspreading we mean that the driving term $\delta(u,r)$, for u in any small interval Δu, does not require the presence of any advanced radiation in the corresponding solution ϕ at retarded times outside Δu. In Ref. 1 the definition is motivated by considering pulses of radiation; this is a heuristic convenience, but unnecessary, as the locality of the resulting condition shows. For further discussion of nonspreading see Ref. 1 and also W. Kundt and E.T. Newman, J. Math. Phys. 12, 2193 (1968). These latter authors give a detailed analysis of nonspreading for homogeneous hyperbolic differential equations in two dimensions.

[4] See, for example, B. Friedman, *Principles and Techniques of Applied Mathematics* (Wiley, New York, 1956), p. 170, Theorem 3.6.

[5] S. Persides, J. Math. Phys. 14, 1017 (1973).

[6] A.I. Janis and E.T. Newman, J. Math. Phys. 6, 902 (1965).

[7] E.T. Copson, *Theory of Functions of a Complex Variable* (Oxford U.P., Oxford, 1935), p. 341.

On the computation of the prolate spheroidal radial functions of the second kind

B. P. Sinha and R. H. MacPhie

Department of Electrical Engineering, University of Waterloo, Waterloo, Ontario, Canada
(Received 14 July 1975)

The series expansion for the prolate spheroidal radial function of the second kind (or its derivative) is found to be slowly convergent when the eccentricity of the spheroid is large (a thin spheroid). To overcome this difficulty, a method is presented in which a small finite number of terms are summed in the conventional manner, and then the infinite remainder series is approximated by an integral of a continuous function. The validity of the method is confirmed by comparing the computed Wronskian with the theoretical. Satisfactory agreement (three to five significant figures) and a very substantial reduction in computation time are achieved.

I. INTRODUCTION

The differential equation[1-3] for the prolate spheroidal radial function is given by

$$\frac{d}{d\xi}\left((\xi^2-1)\frac{d}{d\xi}R_{mn}(h,\xi)\right)-\left(E_{mn}-h^2\xi^2+\frac{m^2}{\xi^2-1}\right)R_{mn}(h,\xi)=0, \tag{1}$$

where ξ is the radial coordinate, $1\leq\xi<\infty$, $m=0,1,2,\cdots$, n is an integer $\geq m$, $h=(2\pi/\lambda)F$, with F being the semi-interfocal distance of the spheroid and λ being the wavelength of the interacting wave, and E_{mn} is the eigenvalue (or separation constant) for which a fast accurate method of computation is described in Refs. 4 and 5.

Two independent solutions of Eq. (1) are[1]

$$R_{mn}^{(1)}(h,\xi)=\left(\frac{\xi^2-1}{\xi^2}\right)^{m/2}\sum_{r=0,1}^{\infty}{}' a_r^{mn}(h)\,j_{m+r}(h\xi) \tag{2}$$

and

$$R_{mn}^{(2)}(h,\xi)=\left(\frac{\xi^2-1}{\xi^2}\right)^{m/2}\sum_{r=0,1}^{\infty}{}' a_r^{mn}(h)\,n_{m+r}(h\xi) \tag{3}$$

where $a_r^{mn}(h)$ are the normalized expansion coefficients, $j_{m+r}(h\xi)$ and $n_{m+r}(h\xi)$ are respectively the spherical Bessel and spherical Neumann functions, and $\sum'$ denotes the summation over even or odd values of r according as $(n-m)$ is even or odd. $R_{mn}^{(1)}(h,\xi)$ and $R_{mn}^{(2)}(h,\xi)$ are called prolate radial functions of the first and second kind, respectively.

The derivatives of $R_{mn}^{(1)}(h,\xi)$ and $R_{mn}^{(2)}(h,\xi)$ are given respectively by

$$\begin{aligned}R_{mn}^{(1)\prime}(h,\xi)&=\frac{d}{d\xi}R_{mn}^{(1)}(h,\xi)\\&=\left(\frac{\xi^2-1}{\xi^2}\right)^{m/2}\left(h\sum_{r=0,1}^{\infty}{}' a_r^{mn}(h)\frac{d}{d(h\xi)}j_{m+r}(h\xi)\right)\\&\quad+\frac{m}{\xi(\xi^2-1)}R_{mn}^{(1)}(h,\xi)\end{aligned} \tag{4}$$

and

$$\begin{aligned}R_{mn}^{(2)\prime}(h,\xi)&=\frac{d}{d\xi}R_{mn}^{(2)}(h,\xi)\\&=\left(\frac{\xi^2-1}{\xi^2}\right)^{m/2}\left(h\sum_{r=0,1}^{\infty}{}' a_r^{mn}(h)\frac{d}{d(h\xi)}n_{m+r}(h\xi)\right)\\&\quad+\frac{m}{\xi(\xi^2-1)}R_{mn}^{(2)}(h,\xi).\end{aligned} \tag{5}$$

The theoretical[1-3] value of Wronskian of $R_{mn}^{(1)}(h,\xi)$ and $R_{mn}^{(2)}(h,\xi)$ is obtained from the asymptotic forms of these functions and is given by

$$W_t=1/h(\xi^2-1). \tag{6}$$

The computed Wronskian W_c is obtained from the computed values of the radial functions and their derivatives by the relation

$$W_c=[R_{mn}^{(1)}(h,\xi)R_{mn}^{(2)\prime}(h,\xi)-R_{mn}^{(1)\prime}(h,\xi)R_{mn}^{(2)}(h,\xi)]_{\text{computed}}. \tag{7}$$

Since $R_{mn}^{(1)}(h,\xi)$ and $R_{mn}^{(1)\prime}(h,\xi)$ can be computed very accurately due to rapid convergence of the series (2) and (4) without any difficulty, the accuracy of computation of $R_{mn}^{(2)}(h,\xi)$ and $R_{mn}^{(2)\prime}(h,\xi)$ (also the degree of convergence of the series representing these functions) is determined by matching the computed and theoretical Wronskians W_t and W_c.

Flammer[2] (p. 32) asserts that "the series of Eq. (3) for $R_{mn}^{(2)}(h,\xi)$ converges very slowly, if at all, when $h\xi$ is small." He also cites Morse and Feshbach[3] (p. 1506), according to whom "the series does not converge well for $h\xi$ small, in fact it is an asymptotic series not being absolutely convergent for any finite value of $h\xi$." Weeks[6] (p. 46), who used a series similar to that in Eq. (3), asserts that it converges very slowly in the neighborhood of $\xi=1$. In fact he used the series for computation of $R_{mn}^{(2)}(h,\xi)$ by directly adding a sufficient number of terms and checking the accuracy of the result by comparing computed and theoretical Wronskians. He also discussed the suitability of the series representation and its use for computation even though the convergence is very slow. Flammer[2] (p. 33–34) gives an alternative series for $R_{mn}^{(2)}(h,\xi)$ which he claims to converge well, so long as h is not too large, for small values of ξ, that is, for values of ξ near, but not including, unity. But the computation of $R_{mn}^{(2)}$ by using this series is much more complicated and time consuming due to the inclusion of two sets of expansion coefficients, a joining factor between radial and angle functions and two kinds of associated Legendre functions as compared to a single set of expansion coefficients and spherical Neumann functions used in series (3).

In actual computation by using the series (3) and (5) for the cases of small $h\xi$ or $\xi=1$, it is found that the series does converge very slowly as verified by comparing W_c and W_t. For large values of $h\xi$, and ξ of

course, the series are found to converge much more rapidly.

When the series (3) and (5) converge slowly, the computation of $R_{mn}^{(2)}(h,\xi)$ for a reasonable accuracy (determined by the matching of the Wronskians) consumes large amounts of computer storage and time, and further we encounter serious scaling problems[6] when the number of terms of the series becomes large. In this paper the difficulties involving slow convergence and scaling are overcome by summing a small finite number of terms in the usual manner and then by approximating the infinite remainder series by an integral of a continuous function. This method yields excellent matching (three to five significant figures) of the Wronskians for various values of h and ξ used in Ref. 7, indicating thereby the correctness of the results.

II. INTEGRATION METHOD

In this method, the first few terms of the series (3) are summed up to the order $r=N$ say, where N is a moderately large number such that the quantities $a_N^{mn}(h)$ and $n_{m+N}(h\xi)$ remain within the machine range. The remainder series is then replaced by an integral of a continuous function of r. $R_{mn}^{(2)'}(h,\xi)$ is computed in a similar way.

To this end we write Eq. (3) as

$$R_{mn}^{(2)}(h,\xi)=\left(\frac{\xi^2-1}{\xi^2}\right)^{m/2}S, \tag{8}$$

where

$$S=\sum_{r=0,1}^{\infty}{}' t_r \tag{9}$$

and

$$t_r=a_r^{mn}(h)n_{m+r}(h\xi). \tag{10}$$

The ratio of successive terms can be expanded (see the Appendix) in inverse powers of $\nu=m+r$ for large r. That is,

$$\frac{t_{r+2}}{t_r}=\left(1+\sum_{k=1}^{\infty}\frac{H_k}{\nu^k}\right)\frac{1}{\xi^2}. \tag{11}$$

The quantities H_k are derived in the Appendix for $k=1,\ldots,6$.

When truncated to six terms, (11) is a very good approximation to the actual value of the ratio t_{r+2}/t_r even if r is as low as 20, in which case the agreement is good to five or six significant figures.

Now let us consider a function

$$\phi(r)=c\exp(-\alpha\nu)\nu^{\beta}\exp\left(\sum_{k=1}^{\infty}\frac{\delta_k}{\nu^k}\right), \tag{12}$$

where c, α, β, and the δ_k's are constants. Then evidently

$$\frac{\phi(r+2)}{\phi(r)}=\exp(-2\alpha)\left(1+\sum_{k=1}^{6}\frac{I_k}{\nu^k}+O(\nu^{-7})\right), \tag{13}$$

where the I_k's are functions of β and $\delta_1,\delta_2,\ldots,\delta_{k-1}$ and the terms of order $O(\nu^{-7})$ are considered negligible. Now by setting $I_k=H_k$, $k=1,2,\ldots,6$, and $\exp(-2\alpha)=1/\xi^2$, the series (13) can be completely matched with the series (11) up to the sixth inverse power of ν. Thus the series (13) is made to give the same degree of approximation to t_{r+2}/t_r as the series (11). We also get all the unknowns α,β, and δ_k's, up to $k=5$ (see the Appendix) except c. Now we can write an approximate explicit expression for t_r for large r as a continuous function $t(r)$ as

$$t(r)=c\exp[-\alpha(m+r)](m+r)^{\beta}\exp\left(\sum_{k=1}^{5}\frac{\delta_k}{(m+r)^k}\right), \tag{14}$$

where c is deduced from (14) by knowing the exact value of t_r at $r=N$.

The sum S of Eq. (9) can be broken into two parts,

$$S=S_1+S_2, \tag{15}$$

where $S_1=\sum_{r=0,1}^{N'}t_r$ and is obtained by direct addition and $S_2=\sum_{r=N+2}^{\infty'}t_r$, which is replaced by integration by using the functional form of $t(r)$ and the trapezoidal rule. Now from Fig. 1

$$\int_{N+2}^{N+4}t(r)\,dr=2t_{N+2}-\epsilon_2=2t_{N+4}+\epsilon_1. \tag{16}$$

Therefore,

$$t_{N+2}+t_{N+4}=\int_{N+2}^{N+4}t(r)\,dr+\tfrac{1}{2}(\epsilon_2-\epsilon_1). \tag{17}$$

If $(\epsilon_2-\epsilon_1)\to 0$, which strongly holds in the case of a slowly varying $t(r)$, we get

$$t_{N+2}+t_{N+4}\approx\int_{N+2}^{N+4}t(r)\,dr,$$

$$t_{N+4}+t_{N+6}\approx\int_{N+4}^{N+6}t(r)\,dr,$$

. .

. .

. .

and hence

$$S_2=\sum_{r=N+2}^{\infty'}t_r\approx\tfrac{1}{2}\int_{N+2}^{\infty}t(r)\,dr+\tfrac{1}{2}t_{N+2}. \tag{18}$$

By suitable transformations we can write

$$\int_{N+2}^{\infty}t(r)\,dr=\frac{c\exp[-\alpha(m+N+2)]}{\alpha}\times\int_0^{\infty}\exp(-y)\left(\frac{y}{\alpha}+N+m+2\right)^{\beta}\times\exp\left(\sum_{k=1}^{5}\frac{\delta_k}{(y/\alpha+N+m+2)^k}\right)dy. \tag{19}$$

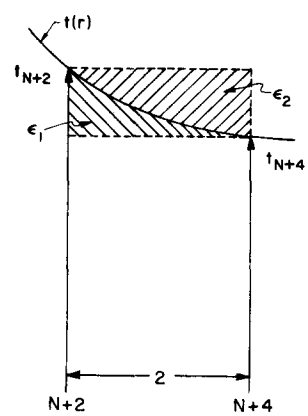


FIG. 1. Integration by trapezoidal rule.

TABLE I. Comparison of direct sum method with integration method: $h=2$, $m=2$, $n=2$, $\xi=1.005$; theoretical Wronskian $=0.4987531\times10^2$.

Direct sum method					Integration method				
No. of Terms	Computer time	$R_{mn}^{(2)}(h,\xi)$	$R_{mn}^{(2)\prime}(h,\xi)$	Computed Wronskian	No. of Terms	Computer time	$R_{mn}^{(2)}(h,\xi)$	$R_{mn}^{(2)\prime}(h,\xi)$	Computed Wronskian
2500	4.94 sec	-0.485191×10^2	0.972934×10^4	0.498541×10^2	10	0.85 sec	-0.485198×10^2	0.973691×10^4	0.498725×10^2
					40	0.89 sec	-0.485224×10^2	0.973694×10^4	0.4987529×10^2

Some radial functions of the second kind and their derivatives: $h=2$, $m=2$, $\xi=1.005$, no. of terms employed $=40$; theoretical Wronskian $=0.4987531\times10^2$

n	$R_{mn}^{(2)}(h,\xi)$	$R_{mn}^{(2)\prime}(h,\xi)$	Computed Wronskian
2	-0.485224×10^2	0.973694×10^4	0.4987529×10^2
3	-0.172812×10^3	0.352247×10^5	0.4987527×10^2
4	-0.879076×10^3	0.182367×10^6	0.4987515×10^2
5	-0.599301×10^4	0.126816×10^7	0.4987495×10^2

This integral can be evaluated numerically by a convenient and accurate scheme such as the Gauss—Laguerre quadrature formula.[8]

Similarly the derivative $R_{mn}^{(2)\prime}(h,\xi)$ is determined from the relation obtained from Eq. (3) as

$$R_{mn}^{(2)\prime}(h,\xi)=\left(\frac{\xi^2-1}{\xi^2}\right)^{m/2} h\sum_{r=0,1}^{\infty\prime} t'_r+\frac{m}{\xi(\xi^2-1)}R_{mn}^{(2)}(h,\xi), \tag{20}$$

where

$$t'_r=a_r^{mn}(h)\frac{d}{d(h\xi)}n_{m+r}(h\xi). \tag{21}$$

The sum

$$S'=\sum_{r=0,1}^{\infty\prime} t'_r\approx\sum_{r=0,1}^{N\prime} t'_r+\tfrac{1}{2}\int_{N+2}^{\infty} t'(r)\,dr+\tfrac{1}{2}t'_{N+2} \tag{22}$$

is obtained in a similar fashion as the evaluation of S.

After computing $R_{mn}^{(2)}(h,\xi)$ and $R_{mn}^{(2)\prime}(h,\xi)$, the Wronskian W_c is computed from the knowledge of $R_{mn}^{(1)}(h,\xi)$ and $R_{mn}^{(1)\prime}(h,\xi)$ and compared with W_t in order to estimate accuracy.

III. RESULTS

The direct summation and integration methods are compared in Table I. It is seen that about 2500 terms are needed for a three significant figures matching of the computed Wronskian with the theoretical Wronskian; a better matching is obtained even if the number of terms for the direct sum in the integration method is as low as 10. It is also noted that computer time required by the integration method is very small as compared with that required by the direct summation.

Table I also contains the values of some radial functions and their derivatives computed by the integration method.

IV. CONCLUSIONS

The simplified method presented in this paper for the determination of the prolate spheroidal radial functions and their derivatives eliminates one of the major stumbling blocks in the computation of the prolate spheroidal wave functions. In fact, the present paper and the methods for computation of eigenvalues discussed in earlier articles[4,5] drastically simplify the determination of the entire set of prolate spheroidal wavefunctions.

APPENDIX

The ratio of the successive terms t_r of the series (9) is given by

$$t_{r+2}/t_r=A_{r+2}B_{r+2}, \tag{A1}$$

where

$$A_{r+2}=\frac{a_{r+2}^{mn}(h)}{a_r^{mn}(h)} \quad\text{and}\quad B_{r+2}=\frac{n_{m+r+2}(h\xi)}{n_{m+r}(h\xi)}.$$

Substituting (2) in Eq. (1), using the recurrence relations of Neumann functions, collecting the coefficients of n_{m+r}, and equating to zero, we get

$$A_{r+2}=\gamma_{r+2}/[\alpha_{r+2}(E_{mn}-\beta_{r+2})-A_{r+4}],$$

where

$$\alpha_{r+2}=-\frac{(2\nu+7)(2\nu+9)}{(\nu-m+4)(\nu-m+3)h^2},$$

$$\beta_{r+2}=(\nu+2)(\nu+3)+\frac{2\nu^2+10\nu+(11-2m)^2}{(2\nu+3)(2\nu+7)}h^2,$$

$$\gamma_{r+2}=\frac{(\nu+m+1)(\nu+m+2)(2\nu+7)(2\nu+9)}{(2\nu+1)(2\nu+3)(\nu-m+4)(\nu-m+3)},$$

and $\nu=m+r$.

Since $A_r\to0$ when $r\to\infty$, for large r we have

$$A_{r+2}\approx\frac{\gamma_{r+2}}{\alpha_{r+2}}\frac{1}{E_{mn}-\beta_{r+2}}, \quad\text{neglecting } A_{r+4}. \tag{A2}$$

With the help of the binomial theorem and the use of approximate relation (A2), A_{r+2} is expanded in inverse powers of ν as

$$A_{r+2}=\left(1+\sum_{k=1}^{6}\frac{F_k}{\nu^k}+O(\nu^{-7})\right)\left[\left(1+\frac{1}{2\nu}\right)\left(1+\frac{3}{2\nu}\right)\right]^{-1}\left(\frac{h^2}{4\nu^2}\right), \tag{A3}$$

where

$$F_1 = 2(m-1),$$

$$F_2 = [u + (4m^2 - 28m + 24)]/2^2,$$

$$F_3 = [(4m-14)u - (40m^2 - 184m + 144)]/2^3,$$

$$F_4 = [u^2 + (4m^2 - 68m + 140)u + 2H + (304m^2 - 1168m + 864)]/2^4,$$

$$F_5 = [(4m-24)u^2 - (80m^2 - 768m + 1208)u + (8m-48)H - (2080m^2 - 7264m + 5184)]/2^5, \quad \text{(A4)}$$

$$F_6 = [u^3 + (4m^2 - 108m + 356)u^2 + (1008m^2 - 7216m + 9584)u + 4uH + (8m^2 - 216m + 718)H + (13504m^2 - 44608m + 31104)]/2^6$$

with $u = 4E_{mn} - 2h^2$ and $H = (4m^2 - 1)h^2$.

Using the expansion[8]

$$n_r(h\xi) = -\frac{1\cdot 3\cdot 5\cdots(2r-1)}{(h\xi)^{r+1}} \times \left(1 + \frac{(\frac{1}{2}h^2\xi^2)}{1!\,(2r-1)} + \frac{(\frac{1}{2}h^2\xi^2)^2}{2!\,(2r-1)(2r-3)} + \cdots\right),$$

and the binomial theorem, we obtain a series for B_{r+2} in inverse powers of ν as

$$B_{r+2} = \left(1 + \sum_{k=1}^{6} \frac{G_k}{\nu^k} + O(\nu^{-7})\right).\left[\left(1+\frac{1}{2\nu}\right)\left(1+\frac{3}{2\nu}\right)\right]\left(\frac{4\nu^2}{h^2\xi^2}\right), \quad \text{(A5)}$$

where

$$G_1 = 0, \quad G_2 = -w, \quad G_3 = w,$$

$$G_4 = -(w^2 + 7w)/4, \quad G_5 = -(4w^2 - 20w)/8, \quad \text{(A6)}$$

$$G_6 = -(4w^3 + 14w^2 + 61w)/16$$

with $w = \frac{1}{2}h^2\xi^2$.

Neglecting powers higher than $(1/\nu)^6$ in the product of (A3) and (A5), we get approximately

$$\frac{t_{r+2}}{t_r} = A_{r+2}B_{r+2} \approx \left(1 + \sum_{k=1}^{6} \frac{H_k}{\nu^k}\right)\frac{1}{\xi^2}, \quad \text{(A7)}$$

where

$$H_i = F_i + F_{i-1}G_1 + \cdots + F_1G_{i-1} + G_i, \quad i = 1, 2, \ldots, 6.$$

From Eq. (12)

$$\frac{\phi(r+2)}{\phi(r)} = \exp(-2\alpha)\left(1+\frac{2}{\nu}\right)^{\beta}\exp\left\{\sum_{k=1}^{5}\frac{\delta_k}{\nu^k}\left[\left(1+\frac{2}{\nu}\right)^{-k} - 1\right]\right\}. \quad \text{(A8)}$$

Expanding (A8) by using the binomial theorem and exponential expansion in power series of $(1/\nu)$ and equating the coefficients of the resulting series with those of the series in (A7) for equal powers of $(1/\nu)$ from the first to the sixth, we get finally

$$\alpha = \ln(\xi), \quad \beta = H_1/2,$$

$$\delta_1 = [2\beta(\beta-1) - H_2]/2,$$

$$\delta_2 = [\tfrac{4}{3}\beta(\beta-1)(\beta-2) - 4\beta\delta_1 + 4\delta_1 - H_3]/4,$$

$$\delta_3 = [\tfrac{2}{3}\beta(\beta-1)(\beta-2)(\beta-3) - 4\beta(\beta-1)\delta_1 + 8\beta(\delta_1 - \delta_2) - (8\delta_1 - 2\delta_1^2 - 12\delta_2) - H_4]/6,$$

$$\delta_4 = [\tfrac{4}{15}\beta(\beta-1)\cdots(\beta-4) - \tfrac{8}{3}\beta(\beta-1)(\beta-2)\delta_1 + 8\beta(\beta-1) \times (\delta_1 - \delta_2) - 2\beta(8\delta_1 - 2\delta_1^2 - 12\delta_2 + 6\delta_3) + (16\delta_1 - 8\delta_1^2 + 8\delta_1\delta_2 - 32\delta_2 + 24\delta_3) - H_5]/8,$$

$$\delta_5 = [\tfrac{4}{45}\beta(\beta-1)\cdots(\beta-5) - \tfrac{4}{3}\beta(\beta-1)\cdots(\beta-3)\delta_1 + \tfrac{16}{3}\beta(\beta-1)(\beta-2)(\delta_1 - \delta_2) - 2\beta(\beta-1) \times (8\delta_1 - 2\delta_1^2 - 12\delta_2 + 6\delta_3) + 2\beta(16\delta_1 - 8\delta_1^2 + 8\delta_1\delta_2 - 32\delta_2 + 24\delta_3 - 8\delta_4) - (32\delta_1 - 24\delta_1^2 + 40\delta_1\delta_2 - 80\delta_2 - 8\delta_2^2 - 12\delta_1\delta_3 + 80\delta_3 - 40\delta_4 + \tfrac{4}{3}\delta_1^2) - H_6]/10.$$

In the case of the derivative

$$t'_{r+2}/t'_r = A_{r+2}B'_{r+2},$$

where

$$B'_{r+2} = \frac{d}{d\xi}n_{r+2}(h\xi)\Big/\frac{d}{d\xi}n_r(h\xi),$$

and this can be obtained by application of the binomial expansion as a power series

$$B'_{r+2} = \left(1 + \sum_{k=1}^{6}\frac{G_k}{\nu^k} + O(\nu^{-7})\right)\left[\left(1+\frac{1}{2\nu}\right)\left(1+\frac{3}{2\nu}\right)\right]\left(\frac{4\nu^2}{h^2\xi^2}\right), \quad \text{(A9)}$$

where

$$G_1 = 2, \quad G_2 = -(w+2),$$

$$G_3 = (3w+2), \quad G_4 = -(w^2 + 19w + 8)/4,$$

$$G_5 = (24w^2 + 56w + 16)/8,$$

$$G_6 = -(4w^3 + 62w^2 + 153w + 32)/16.$$

The values of H_i's and δ_i's are then obtained in similar fashion as discussed before.

*Work supported by Grant No. A-2176 from the National Research Council, Ottawa, Canada.

[1] J. A. Stratton, P. M. Morse, L. J. Chu, D. C. Little, and F. J. Carbato, *Spheroidal Wave Functions* (Wiley, New York, 1956).
[2] C. Flammer, *Spheroidal Wave Functions* (Stanford U. P., Stanford, Calif., 1957).
[3] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953).
[4] B. P. Sinha, R. H. MacPhie, and T. Prasad, IEEE Trans. Antennas Propagation, AP-21, 406 (1973).
[5] D. B. Hodge, J. Math. Phys. 11, 2308 (1970).
[6] W. L. Weeks, "Prolate Spheroidal Wave Functions for Electromagnetic Theory," University of Illinois Antenna Laboratory, Tech. Report No. 38, 1959.
[7] B. P. Sinha and R. H. MacPhie, IEEE Trans. Antennas Propagation, **AP-23**, 676 (1975).
[8] M. Abramowitz and A. I. Stegum, *Handbook of Mathematical Functions* (Dover, New York, 1968).

A Bäcklund transformation in two dimensions*

Hsing-Hen Chen†

Institute for Advanced Study, Princeton, New Jersey 08540
(Received 30 May 1975)

Bäcklund transformation method is applied to find solutions of a nonlinear evolution equation. This equation describes weakly dispersive nonlinear shallow water wave in two space dimensions.

1. INTRODUCTION

Recently, inverse scattering method[1–5] and Bäcklund transformation method[6–9] have been widely studied to obtain solutions of many nonlinear evolution equations with physical applications. Typical examples are Kortweg—de Vries equations[10] describing shallow water wave and ion acoustic wave, nonlinear Schrödinger equation describing Laser self-focusing,[2] sine—Gordon equation[3] describing a nonlinear model field theory, etc. A common feature of these equations is the existence of solitary wave solutions.[11] These are stable pulses balanced between nonlinear and dispersive effects. They maintain their identities even after colliding with each other and, therefore, represent new degrees of freedom associated with the nonlinear media.

However, successful these methods are. They were restricted, up to now, to one-space dimensional problems only.[11] It is certainly very desirable to extend these methods to higher dimensions in accord to the physical reality.

Dryuma[12] recently presented a Lax equation for a two-dimensional Kortweg—de Vries equation. It provides hopes that one can solve a two-dimensional problem by using either inverse scattering or Bäcklund transformation method. We will show in this paper that Bäcklund transformation can indeed be applied to this particular problem. Soliton solutions can be constructed similar to the one-dimensional case. However, solitons in two dimensions are not localized. They are straight lines of infinite extent. Multisoliton solutions can also be constructed through a superposition formula. They are also of infinite extent, constituting parts of wedges and polygons.

In this paper, we study the two-dimensional Kortweg—de Vries equation, first obtained by Kadomtsev and Petviashvili[13] in describing disturbances in a weakly dispersive, weakly nonlinear medium:

$$q_{xt} \pm q_{yy} + q_{xx} + (3q^2)_{xx} + q_{xxxx} = 0. \tag{1.1}$$

The positive sign refers to negative dispersion while the negative sign refers to the positive dispersion. Note that Eq. (1.1) reduces to a KdV equation in the xt plane and to a Boussinesq equation in the xy plane.

2. GENERALIZED LAX CRITERION

If we are given two linear Hermitian operators $\hat{L}[\partial_{\mathbf{x}}, \mathbf{q}(\mathbf{x})]$ and $\hat{A}[\partial_{\mathbf{x}}, \mathbf{q}(\mathbf{x})]$ ($\mathbf{x}$ and $\mathbf{q}$ denote the totality of $x_1, x_2, \ldots, x_n$ and $q_1, q_2, \ldots, q_m$ respectively) that commute with each other, i.e.,

$$[\hat{L}, \hat{A}] = 0, \tag{2.1}$$

then there exists a simultaneous set of eigenfunctions $\{\psi(\mathbf{x})\}$ such that

$$\hat{L}\psi = \alpha\psi \quad \text{and} \quad \hat{A}\psi = \beta\psi, \tag{2.2}$$

where α and β are eigenvalues independent of $\mathbf{x}$. Equation (2.1) will give us the evolution equations of interest and Eqs. (2.2) are the basic equations from which we derive our Bäcklund transformations for the evolution equations obtained from Eq. (2.1).

We will clarify these points by showing an example adapted from Dryuma[12] in the following:

$$\hat{L} = \partial_x^2 + q(x, y, t) - b\partial_y, \tag{2.3}$$

$$\hat{A} = 4i\partial_x^3 + 3i(q\partial_x + \partial_x q) + 3ib\int^x q_y dx + i\partial_x + i\partial_t \tag{2.4}$$

By choosing $b = \pm i/\sqrt{3}$, we get, from Eqs. (2.3), (2.4), the equation

$$q_{xt} - q_{yy} + q_{xx} + (3q^2)_{xx} + q_{xxxx} = 0. \tag{2.5}$$

On the other hand, choosing $b = \pm 1/\sqrt{3}$, we get

$$q_{xt} + q_{yy} + q_{xx} + (3q^2)_{xx} + q_{xxxx} = 0. \tag{2.6}$$

In both cases, we have two choices of b in getting the same equation. We will concentrate first on Eq. (2.6). The transformation $y \to iy$ automatically carries it to Eq. (2.5).

3. BÄCKLUND TRANSFORMATION

From Eqs. (2.2), (2.3) and (2.4), we have

$$\begin{aligned} &\partial_x^2\psi + q\psi \pm (1/\sqrt{3})\partial_y\psi = \alpha\psi, \\ &4i\partial^3\psi + 3i(q\partial_x + \partial_x q)\psi \pm 3i\int^x q_y dx\,\psi + i\partial_x\psi + i\partial_t\psi = \beta\psi. \end{aligned} \tag{3.1}$$

This set of equations can be understood as a transformation equation between q and ψ, where q is a solution to Eq. (2.6). We will show in the following that ψ is related to another solution q' of Eq. (2.6). Therefore, Eq. (3.1) provides a relation between two solutions q and q' of the same Eq. (2.6), that is, a Bäcklund transformation.

Let $\phi \equiv \ln\psi$ and $w_x \equiv q$; we can rewrite Eq. (3.1) as

$$\begin{aligned} &\phi_{xx} + \phi_x^2 + w_x \mp (1/\sqrt{3})\phi_y = \alpha, \\ &4i(\phi_{xxx} + 3\phi_x\phi_{xx} + \phi_x^3) + 3i(2w_x\phi_x + w_{xx}) \\ &\pm \sqrt{3}iw_y + i\phi_x + i\phi_t = \beta. \end{aligned} \tag{3.2}$$

Eliminating w in Eq. (3.2), we get a nonlinear evolution equation for ϕ:

$$\begin{aligned} &\phi_{xxx} - 2\phi_x^3 + 6\alpha\phi_x \pm 2\sqrt{3}\int^x \phi_y\phi_{xx}dx \\ &+ \int^x \phi_{yy}dx + \phi_x + \phi_t = -i\beta. \end{aligned} \tag{3.3}$$

It means that, for every solution ϕ of Eq. (3.3), we can construct a solution $w_x \equiv q$ of Eq. (2.6). Now, it is easy to observe that for every (ϕ, β) pair satisfying Eq. (3.3) we have $(-\phi, -\beta)$ satisfying the same equation with a sign change in the fourth term. For this new pair $(-\phi, -\beta)$, there is a corresponding solution $q' \equiv w'_x$ of Eq. (2.6) such that

$$\begin{aligned} &-\phi_{xx} + \phi_x^2 + w'_x \mp (1/\sqrt{3})\phi_y = \alpha, \\ &4i(-\phi_{xxx} + 3\phi_x\phi_{xx} - \phi_x^3) + 3i(-2w'_x\phi_x + w'_{xx}) \\ &\mp \sqrt{3} i w'_y - i\phi_x - i\phi_t = -\beta. \end{aligned} \tag{3.4}$$

Taking the difference and sum of Eq. (3.2) and Eq. (3.4), we get

$$\begin{aligned} &2\phi_{xx} = w'_x - w_x, \\ &8i(\phi_{xx} + \phi_x^3) + 6i\phi_x(w + w')_x + 3i(w - w')_{xx} \\ &\pm \sqrt{3}\, i(w + w')_y + 2i\phi_x + 2i\phi_t = 2\beta. \end{aligned} \tag{3.5}$$

and

$$\begin{aligned} &2\phi_x^2 + (w + w')_x \mp (2/\sqrt{3})\phi_y = 2\alpha, \\ &24\phi_x\phi_{xx} + 6i\phi_x(w - w')_x + 3i(w + w')_{xx} \pm \sqrt{3}\, i(w - w')_y = 0. \end{aligned} \tag{3.6}$$

From (3.5) we have $\phi = \frac{1}{2}\int^x (w' - w)\,dx + \phi_0(y, t)$, where ϕ_0 is an arbitrary function of y and t. For convenience, we can choose $\phi_0 = \mp 3\alpha y - i\beta t$. This will cancel out the α and β in the Bäcklund transformation and implies that these two constants are not essential in constructing solutions. Now, by substituting the relation between ϕ and w' into Eqs. (3.5) and (3.6), we get the Bäcklund transformation

$$\begin{aligned} &(w' - w)^2 + 2(w' + w)_x - \epsilon(2/\sqrt{3}) \int^x (w' - w)_y\,dx = 0, \\ &4(w' - w)_{xx} + (w' - w)^3 + 3(w' + w)_x(w' - w) + 3(w - w')_{xx} \\ &+ \sqrt{3}\epsilon(w + w')_y + (w' - w)_x + \int^x (w' - w)_t\,dx = 0. \end{aligned} \tag{3.7}$$

The double signs appearing in the transformation equations reflect the fact that we can have wave propagating in both positive and negative direction in y. It is therefore convenient to replace them by a single parameter $\epsilon = \pm 1$.

From Eq. (3.7) we can proceed on to construct specific solutions of Eq. (2.6). Starting from a known solution (say $w = 0$), we can get many solutions w' from Eq. (3.7) by integrating it directly, or more conveniently by going back to Eq. (3.1),

$$\begin{aligned} &\partial_x^2\psi - (\epsilon/\sqrt{3})\partial_y\psi = 0, \\ &4i\partial_x^3\psi + i\partial_x\psi + i\partial_t\psi = 0; \end{aligned} \tag{3.8}$$

the most general solution of ψ is

$$\psi = \sum_k a_k \exp[ikx - \sqrt{3}\epsilon k_y^2 + i(4k^3 - k)t] \equiv \sum_k a_k \exp(\xi_k). \tag{3.9}$$

The summation runs over all complex values of k and a_k is a spectral function. The relation, $q' = w'_x = 2\phi_{xx} = (2\psi_x/\psi)_x$, then implies that

$$q' = 2\,\frac{\left(-\sum_k a_k k^2 \exp(\xi_k)\right)\left(\sum_k a_k \exp(\xi_k)\right) + \left(\sum_k k a_k \exp(\xi_k)\right)^2}{\left(\sum_k a_k \exp(\xi_k)\right)^2} \tag{3.10}$$

is a solution of Eq. (2.6). Special choice of a_k will result in special solutions. For example, let $a_k = \delta_{k,-ik_1} + \delta_{k,-ik_2}$; then

$$\begin{aligned} q' = {} & \tfrac{1}{2}(k_1 - k_2)^2 \operatorname{sech}^2[\tfrac{1}{2}(k_1 - k_2)\,x + (\sqrt{3}\epsilon/2)(k_1^2 - k_2^2)y \\ & - \tfrac{1}{2}(4k_1^3 - 4k_2^3 + k_1 - k_2)t]. \end{aligned} \tag{3.11}$$

It is a two-dimensional solition with amplitude $\frac{1}{2}(k_1 - k_2)^2$ and velocity

$$v_x = -4(k_1^2 + k_2^2 + k_1k_2) - 1, \quad v_y = \frac{\epsilon}{\sqrt{3}}\,\frac{v_x}{k_1 + k_2}.$$

Changing the sign of k_1 and k_2 will change the sign between v_y and v_x. There is then no need to keep ϵ in the Bäcklund transformation. We set, therefore, $\epsilon = 1$ in Eq. (3.7).

Singular solution can also be obtained from Eq. (3.10). The choice $a_k = \delta_{k,-ik_1} - \delta_{k,-ik_2}$ results in

$$\begin{aligned} q'_s = {} & -\tfrac{1}{2}(k_1 - k_2)^2 \operatorname{csch}^2[\tfrac{1}{2}(k_1 - k_2)x + \tfrac{1}{2}\sqrt{3}\epsilon\,(k_1^2 - k_2^2)y \\ & - \tfrac{1}{2}(4k_1^3 - 4k_2^3 + k_1 - k_2)t]. \end{aligned} \tag{3.12}$$

Care must be taken in choosing a_k to generate regular solution from Bäcklund transformation.

4. SUPERPOSITION FORMULA

The advantage of Bäcklund transformation is the possibility of deriving a superposition formula for solutions. This superposition enables us to construct more complicated solutions by algebraic means only. No more integration quadrature is needed then. It also implies that solutions obtained in the last section are stable. They do not lose their identities after colliding with each other.

To derive the superposition formula, we let w_1 be a solution generated by Bäcklund transformation from a known solution w_0 with a certain spectral function $a_{1,k}$, w_2 be a second solution generated from w_0 with spectral function $a_{2,k}$, and w_3 a third solution generated from w_1 with spectral function $a_{2,k}$. From the definition and Bäcklund transformation Eq. (3.7), we have

$$(w_1 - w_0)^2 + 2(w_1 + w_0)_x - (2/\sqrt{3}) \int^x (w_1 - w_0)_y\,dx = 0, \tag{4.1}$$

$$(w_2 - w_0)^2 + 2(w_2 + w_0) - (2/\sqrt{3}) \int^x (w_2 - w_0)_y\,dx = 0, \tag{4.2}$$

$$(w_3 - w_1)^2 + 2(w_3 + w_1)_x - (2/\sqrt{3}) \int^x (w_3 - w_1)_y\,dx = 0. \tag{4.3}$$

It is easily shown that

$$w_3 + w_0 = w_1 + w_2 + 2(w_1 - w_2)_x/(w_1 - w_2) \tag{4.4}$$

satisfy the above three equations and is therefore the superposition formula we are searching for.

Starting from $w_0 = 0$, we have w_1 and w_2 as given in Eq. (3.10) with spectral functions $a_{1,k}$ and $a_{2,k}$ respectively. They can be called single-spectrum solutions. Insert them into formula (4.4); we get a solution w_3 contains two spectral functions (both $a_{1,k}$ and $a_{2,k}$). It can be called a two-spectra solution. In particular, when we choose $a_{1,k}$, $a_{2,k}$, the special form that generates 2D solitary waves in Eqs. (3.11) and (3.12), we get w_3 a two-soliton solution in two dimensions. This solution constitutes a two-straight-line wavefront. The two straight lines intersect at a certain point. Each one

is broken at the intersecting point, reconnecting with the other one, forming an angular wedge. The two wedges, opposite to each other, travel together with a group velocity

$$\vec{v} = \vec{v}_1 + \hat{n} \times \hat{e}_1 \, (\cot\theta \, v_1 - \csc\theta \, v_2),$$

where $\hat{e}_1 = \vec{v}_1 / v_1$, $\hat{n} = \vec{v}_2 \times \vec{v}_1 / |\vec{v}_1 \times \vec{v}_2|$,

$$\cos\theta = \vec{v}_1 \cdot \vec{v}_2 / v_1 v_2, \quad \text{and } v_1 > v_2.$$

It behaves therefore like a wedged soliton. When three-line soliton solutions are superposed, we will observe relative motions between three wedges and an inner triangle. Higher rank solutions behave similarly. As remarked in the Introduction, these solutions are not localized. If we do have an initial localized wavepacket, we conjecture that this wavepacket will spread and develop into extended solutions of the kind described above.

ACKNOWLEDGMENT

The author thanks Professor K.M. Case for many helpful discussions.

*Work supported in part by AEC Grant AT(11-1) 3237.

†New address: Department of Physics and Astronomy, University of Maryland, College Park, Maryland 20742.

[1]C. Gardner, J. Greene, M. Kruskal, and R. Miura, Phys. Rev. Lett. 19, 1095 (1967).
[2]V. Zakharov and A. Shabat, Zh. Eksp. Teor. Phys. 61, 118 (1971) [Sov. Phys. JETP 34, 62 (1972)].
[3]M. Ablowitz, P. Kaup, A. Newell, and H. Segur, Phys. Rev. Lett. 30, 1262 (1973).
[4]M. Wadati, J. Phys. Soc. Jpn. 34, 1289 (1973).
[5]H. Flaschka, Prog. Theor. Phys. 51, 703 (1974).
[6]G. Lamb, Rev. Mod. Phys. 43, 99 (1971).
[7]H. Wahlquist and F. Estabrook, Phys. Rev. Lett. 31, 1386 (1973).
[8]H. Chen, Phys. Rev. Lett. 33, 925 (1974).
[9]H. Chen and C. Liu, J. Math. Phys. 16, 1428 (1975).
[10]M. Kruskal and N. Zabusky, Phys. Rev. Lett. 15, 240 (1965).
[11]A. Scott, F. Chu, and D. McLaughlin, Proc. IEEE 61, 1443 (1973).
[12]V. Dryuma, Zh. Eksp. Teor. Phys. Pis. Red. 19, 753 (1974) [Sov. Phys. JETP Lett. 19, 387 (1974)].
[13]B. Kadomtsev and V. Petviashvili, Sov. Phys. Dokl. 15, 539 (1970).

Lorentz covariant treatment of the Kerr–Schild geometry*

Metin Gürses† and Feza Gürsey

Physics Department, Yale University, New Haven, Connecticut 06520
(Received 26 December 1974)

It is shown that a Lorentz covariant coordinate system can be chosen in the case of the Kerr–Schild geometry which leads to the vanishing of the pseudo energy–momentum tensor and hence to the linearity of the Einstein equations. The retarded time and the retarded distance are introduced and the Liénard–Wiechert potentials are generalized to gravitation in the case of world-line singularities to derive solutions of the type of Bonnor and Vaidya. An accelerated version of the de Sitter metric is also obtained. Because of the linearity, complex translations can be performed on these solutions, resulting in a special relativistic version of the Trautman–Newman technique and Lorentz covariant solutions for spinning systems can be derived, including a new anisotropic interior metric that matches to the Kerr metric on an oblate spheroid.

1. INTRODUCTION

In general relativity, the field equations are often simplified when we deal with algebraically special or degenerate metrics. The degeneracy of the metrics is linked with the multiplicity of the Debever–Penrose directions.[1,2] One of the important examples for the algebraically special metrics is the Kerr–Schild[2] metric which is given as

$$g_{\mu\nu}=\eta_{\mu\nu}-2V\lambda_\mu\lambda_\nu \tag{1.1}$$

where $\eta_{\mu\nu}=(1,-1,-1,-1)$ is the Minkowski metric, V is a scalar function, and λ_μ is a light like vector both with respect to $g_{\mu\nu}$ and $\eta_{\mu\nu}$:

$$g_{\mu\nu}\lambda^\mu\lambda^\nu=\eta_{\mu\nu}\lambda^\mu\lambda^\nu=0. \tag{1.2}$$

This null vector is also geodesic both with respect to $g_{\mu\nu}$ and $\eta_{\mu\nu}$, that is,

$$\lambda^\mu\partial_\mu\lambda_\nu=\lambda^\mu\nabla_\mu\lambda_\nu=0, \tag{1.3}$$

where ∂_μ and ∇_μ are covariant derivatives with respect to $\eta_{\mu\nu}$ and $g_{\mu\nu}$, respectively. These two properties of λ_μ, Eqs. (1.2), (1.3), make it a shear free double Debever–Penrose vector. If the scalar function V is a constant, λ_μ becomes a Killing vector.

The Kerr–Schild metric has been studied by several authors by using either the tetrad formalism[2] or the direct procedure[3] in solving the field equations. We use the second method in a special relativistic covariant way and find the energy–momentum tensor (e.m.t.) of the matter and the field other than the gravitational field. The mixed form of the e.m.t. is linear in the function V and also it is divergence free in the ordinary sense, that is,

$$\partial_\mu T^\mu{}_\nu=0. \tag{1.4}$$

Therefore, the pseudo-energy–momentum tensor (p.e.m.t.) of the gravitational field must also be conserved. We find that it vanishes in this coordinate system. Vanishing of the p.e.m.t. makes the field equations linear. Because of this fact, the gravitational field is not its own source in this coordinate system. If a metric can be thrown into the Kerr–Schild form by a coordinate transformation, the gravitational energy and momentum are cancelled by this coordinate transformation which represents some kind of acceleration according to the equivalence principle.

A method, which leads to a new metric from an old one, is based on making a complex translation along one of the coordinates without changing the physical character of the source. Such a complex translation is allowed in classical electrodynamics and in linearized general relativity because of their linearity[4] of the equations. In exact general relativity complex translation was used several years ago by Newman and Janis[5] to obtain the Kerr metric from the Schwarzschild metric and by Newman *et al.*[6] to obtain the charged Kerr metric from the Reissner–Nordstrom metric. Recently, Adler *et al.*[3] used complex translation and reobtained the Kerr metric in the Kerr–Schild coordinate system without drawing attention to the linearity of the field equations. Now it becomes clear that complex translation is allowed in general relativity whenever we can find a coordinate system in which the p.e.m.t. vanishes or the Einstein equations are linear in this coordinate system. This is of course not true for an arbitrary metric. It happens to be true in the algebraically special Kerr–Schild geometry.

Another advantage of the Kerr–Schild metric is the following. When we take λ_μ as the gradient of the retarded time and V as a function of the retarded distance for an accelerated system (particles, charges, etc.) we get simply the result of Bonnor and Vaidya,[7] generalizing the Liénard–Wiechert potential in electromagnetism to the retarded gravitational potential. In addition to their result we also find the accelerated version of the de Sitter metric.

In Sec. 2, we find the Einstein tensor of the Kerr–Schild metric and show the linearity of the field equations. We also prove that λ_μ is a double Debever–Penrose vector for any e.m.t. In Sec. 3 we find the field of accelerated systems, especially of the charged particle in a de Sitter universe. In Sec. 4 we complexify the solutions discussed in the previous section for non-accelerated systems. We find the e.m.t. for this case. The Kerr[8] and charged Kerr metrics[6] are special cases of this e.m.t. For the interior metric, this tensor is shown to correspond to the e.m.t. of an anisotropic perfect fluid and to match the Kerr metric on an oblate spheroid. In Sec. 5, we show the resemblence between the linearized field equations obtained from an approximation procedure and the field equations obtained for the Kerr–Schild metric.

2. THE KERR-SCHILD GEOMETRY

The light like character of the four vector λ_μ greatly simplifies the calculations. Because of this property, it can be raised and lowered with both $\eta_{\mu\nu}$ and $g_{\mu\nu}$, and we also have

$$\sqrt{-g}=1, \tag{2.1}$$

so that we have

$$g^{\mu\nu}=\eta^{\mu\nu}+2V\lambda^\mu\lambda^\nu. \tag{2.2}$$

The Riemann–Christoffel symbols and the curvature tensor, for this metric, are

$$\Gamma^\alpha{}_{\mu\nu}=-[(l^\alpha l_\nu)_{,\mu}+(l^\alpha l_\mu)_{,\nu}-\eta^{\gamma\alpha}(l_\mu l_\nu)_{,\gamma}+4Al^\alpha l_\mu l_\nu], \tag{2.3}$$

$$\begin{aligned}R^\gamma{}_{\mu\nu\alpha}&=\partial_\nu\Gamma^\gamma{}_{\mu\alpha}-\partial_\alpha\Gamma^\gamma{}_{\mu\nu}+\Gamma^\gamma{}_{\beta\nu}\Gamma^\beta{}_{\mu\alpha}-\Gamma^\gamma{}_{\beta\alpha}\Gamma^\beta{}_{\mu\nu}\\&=l^\gamma l_{[\alpha}\psi_{\nu]\mu}-\eta^{\gamma\tau}l_\mu l_{[\nu}\psi_{\alpha]\tau}+2Al_\mu l_{[\alpha}\Theta_{\nu]}{}^\gamma\\&\quad-2Al^\gamma l_{[\nu}\Theta'_{\alpha]\mu}-l^\gamma\Phi_{\nu\mu\alpha}+l^\gamma\Phi_{\alpha\mu\nu}\\&\quad-l_\nu\Phi_\mu{}^\gamma{}_\alpha+l_\mu\Phi_\alpha{}^\gamma{}_\nu-l_\mu\Phi_\nu{}^\gamma{}_\alpha+\eta^{\gamma\tau}l_\nu\Phi_{\tau\mu\alpha}\\&\quad-\eta^{\gamma\tau}l_\alpha\Phi_{\tau\mu\nu}-\eta^{\beta\sigma}(l^\gamma l_\nu)_{,\beta}(l_\mu l_\alpha)_{,\sigma}+\eta^{\beta\sigma}(l^\gamma l_\alpha)_{,\beta}(l_\mu l_\nu)_{,\sigma}\\&\quad-(l_\alpha l^\gamma)_{,\mu\nu}+(l_\nu l^\gamma)_{,\mu\alpha}+\eta^{\gamma\beta}(l_\mu l_\alpha)_{,\beta\nu}-\eta^{\gamma\beta}(l_\mu l_\nu)_{,\beta\alpha}\\&\quad-4(Al^\gamma l_\mu l_\alpha)_{,\nu}+4(Al^\gamma l_\mu l_\nu)_{,\alpha},\end{aligned} \tag{2.4}$$

where

$$\begin{aligned}&l_\mu=\sqrt{V}\lambda_\mu,\\&\psi_{\mu\nu}=l_{\beta,\nu}l^\beta{}_{,\mu},\\&\Theta^\gamma{}_\mu=l^\gamma{}_{,\mu}-l_\mu{}^\gamma,\\&\Theta'_{\mu\nu}=l_{\mu,\nu}+l_{\nu,\mu},\\&\Phi_{\mu\nu\alpha}=l^\beta{}_{,\mu}(l_\nu l_\alpha)_{,\beta},\end{aligned}$$

and

$$\begin{aligned}&l_{[\alpha}\psi_{\beta]\mu}=l_\alpha\psi_{\beta\mu}-l_\beta\psi_{\alpha\mu},\\&l_{[\alpha}\Theta^\gamma{}_{\nu]}=l_\alpha\Theta^\gamma{}_\nu-l_\nu\Theta^\gamma{}_\alpha.\end{aligned}$$

We note that

$$l^\gamma\psi_{\gamma\beta}=l^\gamma\Phi_{\alpha\gamma\beta}=l^\gamma\Phi_{\alpha\beta\gamma}=0, \tag{2.5}$$

$$l_\gamma\Theta^\gamma{}_\mu=Al_\mu,\quad A=-\lambda^\gamma(V^{1/2})_{,\gamma} \tag{2.6}$$

$$l^\gamma\Phi_{\alpha\gamma\beta}=2Al_\gamma l_\beta, \tag{2.7}$$

$$\begin{aligned}l_\gamma R^\gamma{}_{\mu\nu\alpha}&=-l_\gamma(l_\alpha l^\gamma)_{,\mu\nu}+l_\gamma(l_\nu l^\gamma)_{,\mu\alpha}\\&\quad+l^\beta(l_\mu l_\alpha)_{,\beta\nu}-l^\beta(l_\mu l_\nu)_{,\beta\alpha},\end{aligned} \tag{2.8}$$

$$\begin{aligned}l_\gamma l^\nu R^\gamma{}_{\mu\nu\alpha}&=l^\beta l^\gamma(l_\mu l_\alpha)_{,\beta\gamma}\\&=[(V_{,\gamma}\lambda^\gamma)_{,\beta}\lambda^\beta V]\lambda_\mu\lambda_\alpha.\end{aligned} \tag{2.9}$$

We can find the Ricci tensor by using the identities (2.5)–(2.7), and letting $\gamma=\nu$ in Eq. (2.4). It reads

$$\begin{aligned}R_{\mu\alpha}&=-(l_\alpha l^\gamma)_{,\mu\gamma}+\eta^{\gamma\beta}(l_\mu l_\alpha)_{,\beta\gamma}-(l_\mu l^\gamma)_{,\gamma\alpha}\\&\quad+2[\eta^{\gamma\tau}l^\beta{}_{,\gamma}l_{\beta,\tau}-(K\lambda^\gamma)_{,\gamma}+(A^2-L^2)]l_\mu l_\alpha,\end{aligned} \tag{2.10}$$

where

$$L=-l^\alpha{}_{,\alpha}=-(V^{1/2}\lambda^\alpha)_{,\alpha} \tag{2.11a}$$

$$K=(A+L)V^{1/2}=-(V\lambda^\mu)_{,\mu}. \tag{2.11b}$$

The Ricci tensor with mixed components $R^\mu{}_\alpha$ and curvature scalar are,

$$R^\mu{}_\alpha=(K\lambda^\mu)_{,\alpha}+\eta^{\mu\gamma}(K\lambda_\alpha)_{,\gamma}+\Box(V\lambda^\mu\lambda_\alpha), \tag{2.12}$$

$$R=2(K\lambda^\mu)_{,\mu},\quad \Box=\eta^{\mu\nu}\partial_\mu\partial_\nu.$$

Hence the Einstein tensor reads

$$G^\mu{}_\alpha=(K\lambda^\mu)_{,\alpha}+\eta^{\mu\gamma}(K\lambda_\alpha)_{,\gamma}+\Box(V\lambda^\mu\lambda_\alpha)-\delta^\mu{}_\alpha(K\lambda^\gamma)_{,\gamma}, \tag{2.13}$$

with K given by (2.11b).

The algebraic classification of space–times is done by means of Weyl's conformal tensor which is defined by

$$\begin{aligned}C^\gamma{}_{\mu\nu\alpha}&=R^\gamma{}_{\mu\nu\alpha}+\tfrac{1}{2}g_{\mu\alpha}R^\gamma{}_\alpha-\tfrac{1}{2}g_{\mu\nu}R^\gamma_\alpha+\tfrac{1}{2}g_{\mu\alpha}\delta^\gamma{}_\nu\\&\quad-\tfrac{1}{2}R_{\mu\nu}\delta^\gamma{}_\alpha-\tfrac{1}{6}(g_{\mu\alpha}\delta^\gamma{}_\nu-g_{\mu\nu}\delta^\gamma{}_\alpha)R.\end{aligned} \tag{2.14}$$

We can easily find that

$$\lambda_\gamma\lambda^\nu C^\gamma{}_{\mu\nu\alpha}=H\lambda_\mu\lambda_\alpha, \tag{2.15}$$

where

$$H=V(V_{,\mu}\lambda^\mu)_{,\nu}\lambda^\nu+A^2-L^2+\eta^{\gamma\beta}l^\tau{}_{,\beta}l_{\tau,\gamma}-\tfrac{2}{3}(K\lambda^\mu)_{,\mu},$$

with A and L given by (2.6) and (2.11a), respectively. Equation (2.15) tells us that λ_μ is a double Debever–Penrose vector, thus space–time is algebraically degenerate.

Now let us show that Einstein's tensor is divergence free in the ordinary sense, that is,

$$\partial_\mu G^\mu{}_\nu=0. \tag{2.16}$$

From the Bianchi identities we know that $G^\mu{}_\nu$ is conserved covariantly,

$$\nabla_\mu G^\mu{}_\alpha=\partial_\mu G^\mu{}_\alpha+\Gamma^\mu{}_{\mu\beta}G^\beta{}_\alpha-\Gamma^\beta{}_{\mu\alpha}G^\mu{}_\beta=0; \tag{2.17}$$

from Eq. (2.1), we have

$$\Gamma^\mu{}_{\mu\beta}=\partial_\beta\sqrt{-g}=0;$$

and it is also straightforward to show that

$$\Gamma^\beta{}_{\mu\alpha}G^\mu{}_\beta=0.$$

Thus, we obtain Eq. (2.16). In general we know that the conservation law for the total energy–momentum tensor is given as

$$\partial_\mu(T^\mu{}_\nu+t^\mu{}_\nu)=0$$

where $t^\mu{}_\nu$ is the p.e.m.t. of the gravitational field. This p.e.m.t. is given in different forms, i.e., the Einstein[9] and the Landau[10] forms. In our case these two forms are the same because of Eq. (2.1) and they both vanish. The total e.m.t. $T_{\mu\nu}+t_{\mu\nu}$ is given by

$$T^\mu{}_\nu+t^\mu{}_\nu=\frac{-1}{2G}\partial_\rho f^{\rho\mu}{}_\nu, \tag{2.18}$$

where G is the gravitational constant and $f^{\rho\mu}{}_\nu$ is defined as

$$f^{\mu\rho}{}_\nu=-f^{\rho\mu}{}_\nu=\mathcal{G}_{\nu\sigma}\partial_\lambda(\mathcal{G}^{\sigma\mu}\mathcal{G}^{\rho\lambda}-\mathcal{G}^{\sigma\rho}\mathcal{G}^{\mu\lambda}),$$

where

$$\mathcal{G}^{\sigma\mu}=\sqrt{-g}g^{\sigma\mu},$$

$$\mathcal{G}_{\sigma\mu}=\frac{1}{\sqrt{-g}}g_{\sigma\mu}.$$

For the Kerr–Schild metric $f^{\mu\rho}{}_\nu$ becomes

$$f^{\mu\rho}{}_{\nu}=2\partial_{\beta}[V\delta^{\mu}{}_{\nu}\lambda^{\rho}\lambda^{\beta}+V\eta^{\rho\beta}\lambda_{\nu}\lambda^{\mu}-V\eta^{\mu\beta}\lambda^{\rho}\lambda_{\nu}-V\delta^{\rho}{}_{\nu}\lambda^{\mu}\lambda^{\beta}]. \tag{2.19}$$

We find that

$$G(T^{\mu}{}_{\nu}+t^{\mu}{}_{\nu})=G^{\mu}{}_{\nu}. \tag{2.20}$$

Hence, from the Einstein equations

$$t^{\mu}{}_{\nu}=0. \tag{2.21}$$

Using this fact and Gupta's equation[11] which reads

$$\partial_{\alpha}\partial_{\beta}(\eta^{\alpha\beta}\mathcal{G}^{\mu\nu}-\eta^{\mu\alpha}\mathcal{G}^{\nu\beta}-\eta^{\nu\alpha}\mathcal{G}^{\mu\beta}+\eta^{\mu\nu}\mathcal{G}^{\alpha\beta})$$
$$=2G\eta^{\mu\nu}(T^{\nu}{}_{\lambda}+t^{\nu}{}_{\lambda}) \tag{2.22}$$

and using the metric in Eq. (2.2) we recover Eq. (2.13). The absence of $t^{\nu}{}_{\lambda}$ in Eq. (2.22) makes the field equations linear, because the total energy–momentum tensor $(T^{\nu}{}_{\lambda}+t^{\nu}{}_{\lambda})$ can only depend on the metric itself not on its derivatives. Thus in the Kerr–Schild geometry Einstein's equations take the linear form

$$\partial_{\alpha}\partial_{\beta}(\eta^{\alpha\beta}g^{\mu\nu}-\eta^{\mu\alpha}g^{\nu\beta}-\eta^{\nu\alpha}g^{\mu\beta}+\eta^{\mu\nu}g^{\alpha\beta})=2G\eta^{\mu\lambda}T^{\nu}{}_{\lambda} \tag{2.23}$$

where $T^{\nu}{}_{\lambda}$ is the energy–momentum tensor of matter and radiation excluding the gravitational field.

3. GRAVITATIONAL FIELD OF ACCELERATED SYSTEMS (NONSPINNING CASE)

Assume that any element of the system under consideration is on a geodesic Γ which is described by an affine parameter τ. Construct a light cone from the observation point x^{μ}, which intersects the geodesic Γ at any point $Z^{\mu}(\tau)$. The velocity of the element of the system is

$$\dot{Z}_{\mu}=\frac{dZ_{\mu}}{d\tau},$$

with

$$\eta^{\mu\nu}\dot{Z}_{\mu}\dot{Z}_{\nu}=\epsilon$$

where $\epsilon=1$ and $\epsilon=0$ correspond to the timelike and lightlike cases, respectively. We define a retarded distance R by

$$R=\dot{Z}^{\mu}(x_{\mu}-Z_{\mu}(\tau_0)),$$

for the value τ_0 of τ for which the distance between the point $Z^{\mu}(\tau_0)$ and the point x^{μ} is lightlike, that is,

$$\eta_{\mu\nu}(x^{\mu}-Z^{\mu}(\tau_0))(x^{\mu}-Z^{\mu}(\tau_0))=0. \tag{3.1}$$

From now on we shall use $\dot{Z}^{\mu}$ to denote $\dot{Z}^{\mu}(\tau_0)$. Differentiation of Eq. (3.1) with respect to x^{μ} gives us

$$\partial_{\mu}\tau_0=[x_{\mu}-Z_{\mu}(\tau_0)]/R\ .$$

Now, we can define the lightlike 4-vector λ_{μ} as

$$\lambda_{\mu}=\partial_{\mu}\tau_0. \tag{3.2}$$

It is straightforward to show that λ_{μ} satisfies Eqs. (1.2), (1.3). In order to find the e.m.t. we need the following identities:

$$\partial_{\alpha}R=\dot{Z}_{\alpha}-\lambda_{\alpha}(\epsilon-R\ddot{Z}_{\beta}\lambda^{\beta}), \tag{3.3a}$$

$$\lambda^{\alpha}\partial_{\alpha}R=\lambda^{\alpha}\dot{Z}_{\alpha}=1, \tag{3.3b}$$

$$\lambda_{\mu,\nu}=\frac{1}{R}[\eta_{\mu\nu}-\lambda_{\mu}\dot{Z}_{\nu}-\lambda_{\nu}\dot{Z}_{\mu}+\lambda_{\mu}\lambda_{\nu}(\epsilon-R\ddot{Z}^{\alpha}\lambda_{\alpha})], \tag{3.3c}$$

$$K=-\frac{1}{R}(RV'-2V), \tag{3.3d}$$

$$V'=\frac{dV}{dR}.$$

Here, we assume that the scalar function V is only a function of the retarded distance R. Using Eq. (3.3a)–(3.3d) and Eq. (2.13), we get the e.m.t. as[12]

$$T^{\mu}{}_{\nu}=\left(V''+\frac{2V'}{R}\right)\delta^{\mu}{}_{\nu}+\left(-V''+\frac{2V}{R^2}\right)(\dot{Z}^{\mu}\lambda_{\nu}+\dot{Z}_{\nu}\lambda^{\mu})$$
$$+[\epsilon V''-2zV'+2(-\epsilon+2zR)(V/R^2)]\lambda^{\mu}\lambda_{\nu} \tag{3.4}$$

where

$$z=\ddot{Z}^{\alpha}\lambda_{\alpha}. \tag{3.5}$$

This e.m.t. in Eq. (3.4) has some simple forms for some special V's. When

$$V=m/R-e^2/2R^2\ \ (e \text{ and } m \text{ are constants}),$$

we get the Bonner and Vaidya[7] solution. When

$$V=(\rho_0/6)R^2,\ \ \rho_0 \text{ is const},$$

we get a new solution corresponding to the gravitational field generated by a de Sitter space in accelerated motion, i.e., the interior solution corresponds to a finite matter free space–time region with a cosmological constant, so that

$$G^{\mu}{}_{\nu}-\rho_0\delta^{\mu}{}_{\nu}=0.$$

We verify that the only vacuum solution with zero cosmological constant is the Schwarzschild metric with $z=0$ for a nonspinning system.

4. COMPLEXIFICATION: FIELD OF UNACCELERATED SPINNING SYSTEMS

The gravitational field of an accelerated spinning system can be found either by solving the Einstein equations given in Eq. (2.13) or by complexifying the solutions found in the previous section. We choose the second method because of its simplicity and use special relativistic spinor representations of the four vectors.

In our method we simply make a complex translation along x^{μ} and find the real and imaginary parts of every four vectors and scalar functions. The new complex quantities are

$$x'^{\mu}=x^{\mu}+ia^{\mu},$$
$$\tau'=\tau_1+i\tau_2,$$
$$Z'^{\mu}=Z_{1\mu}+iZ_{2\mu},$$
$$\dot{Z}'^{\mu}=v_{1\mu}+iv_{2\mu},$$

where a_{μ} is a constant spacelike 4-vector and[13]

$$v_{1\mu}=\frac{\partial Z_{1\mu}}{\partial\tau_1}=\frac{\partial Z_{2\mu}}{\partial\tau_2},$$

$$v_{2\mu}=-\frac{\partial Z_{1\mu}}{\partial\tau_2}=\frac{\partial Z_{2\mu}}{\partial\tau_1}.$$

Instead of Eq. (3.1) we have the following two equations:

$$(x_{\mu}-Z_{1\mu})(x^{\mu}-Z_1{}^{\mu})-(a_{\mu}-Z_{2\mu})(a^{\mu}-Z_2{}^{\mu})=0, \tag{4.1}$$

$$(x_\mu - Z_{1\mu})(a^\mu - Z_2{}^\mu) = 0. \tag{4.2}$$

We shall look for solutions that can be expressed in terms of the complex null vector λ'_μ

$$\lambda'_\mu = \partial_\mu \tau' = \mu_\mu + i\nu_\mu, \tag{4.3}$$

where

$$\mu_\mu = \partial_\mu \tau_1 = \frac{1}{r_1{}^2 + r_2{}^2}[r_1(x_\mu - Z_{1\mu}) + r_2(a_\mu - Z_{2\mu})], \tag{4.4}$$

$$\nu_\mu = \partial_\mu \tau_2 = \frac{1}{r_1{}^2 + r_2{}^2}[-r_2(x_\mu - Z_{1\mu}) + r_1(a_\mu - Z_{2\mu})], \tag{4.5}$$

and the complex retarded distance is

$$R' = r_1 + ir_2, \tag{4.6}$$

where

$$r_1 = v_1{}^\mu(x_\mu - Z_{1\mu}) - v_2{}^\mu(a_\mu - Z_{2\mu}),$$
$$r_2 = v_2{}^\mu(x_\mu - Z_{1\mu}) + v_1{}^\mu(a_\mu - Z_{2\mu}).$$

Instead of the identities in Eqs. (3.3a)–(3.3d), we have the following:

$$r_{1,\alpha} = v_{1\alpha} + k\mu_\alpha + l\nu_\alpha, \tag{4.7a}$$

$$r_{2,\alpha} = v_{2\alpha} - l\mu_\alpha + k\nu_\alpha, \tag{4.7b}$$

$$\mu_\alpha v_1{}^\alpha - v_2{}^\alpha \nu_\alpha = 1, \tag{4.7c}$$

$$v_{1\alpha}\nu^\alpha + v_{2\alpha}\mu^\alpha = 0, \tag{4.7d}$$

$$\mu_\alpha \mu^\alpha = \nu_\alpha \nu^\alpha, \tag{4.7e}$$

$$\mu_\alpha \nu^\alpha = 0, \tag{4.7f}$$

$$\mu_{\alpha,\beta} = \frac{1}{r_1{}^2 + r_2{}^2}[r_1\eta_{\alpha\beta} + (r_2 l - r_1 k)\mu_\alpha\mu_\beta - (r_2 l - r_1 k)\nu_\alpha\nu_\beta - (r_1 l + r_2 k)(\mu_\alpha\nu_\beta + \mu_\beta\nu_\alpha) - r_2(\mu_\alpha v_{2\beta} + \mu_\beta v_{2\alpha} + \nu_\alpha v_{1\beta} + \nu_\beta v_{1\alpha}) + r_1(-\mu_\alpha v_{1\beta} - v_{1\alpha}\mu_\beta + \nu_\alpha v_{2\beta} + \nu_\beta v_{2\alpha})], \tag{4.7g}$$

$$\nu_{\alpha,\beta} = \frac{1}{r_1{}^2 + r_2{}^2}[-r_2\eta_{\alpha\beta} + (r_1 l + r_2 k)\mu_\alpha\mu_\beta - (r_1 l + r_2 k)\nu_\alpha\nu_\beta + (-r_1 k + r_2 l)(\mu_\alpha\nu_\beta + \mu_\beta\nu_\alpha) - r_1(\nu_\alpha v_{1\beta} + \nu_\beta v_{1\alpha} + \mu_\alpha v_{2\beta} + \mu_\beta v_{2\alpha}) + r_2(\mu_\alpha v_{1\beta} + \mu_\beta v_{1\alpha} - v_{2\alpha}\nu_\beta - v_{2\beta}\nu_\alpha)], \tag{4.7h}$$

where

$$k = -1 + r_1 a_1{}^\alpha \mu_\alpha - r_2 \nu_\alpha a_1{}^\alpha - r_2 a_2{}^\alpha \mu_\alpha - r_1 a_2{}^\alpha \nu_\alpha, \tag{4.8}$$

$$l = -r_1 a_2{}^\alpha \mu_\alpha + r_2 a_2{}^\alpha \nu_\alpha - r_2 a_1{}^\alpha \mu_\alpha - r_1 a_1{}^\alpha \nu_\alpha, \tag{4.9}$$

$$a_{1\alpha} = \frac{\partial v_{1\alpha}}{\partial \tau_1} = \frac{\partial v_{2\alpha}}{\partial \tau_2}, \tag{4.10}$$

$$a_{2\alpha} = -\frac{\partial v_{1\alpha}}{\partial \tau_2} = \frac{\partial v_{2\alpha}}{\partial \tau_2}. \tag{4.11}$$

Now in order to find the real null 4-vector λ_μ we use the spinor representation of the 4-vectors. If A_μ is a 4-vector its spinor equivalent is given as

$$A = \sigma_\mu A^\mu \quad (A^\mu = \sigma^\mu{}_{\alpha\dot{\beta}} A^{\alpha\dot{\beta}};\ \alpha, \dot{\beta} = 1, 2),$$

where

$$\sigma_\mu = (\sigma_0, \boldsymbol{\sigma}).$$

σ_0 is two-dimensional identity matrix and $\boldsymbol{\sigma}$'s are the Pauli spin matrices. They satisfy the following anti-commutation relation,

$$\sigma_\mu \bar{\sigma}_\nu + \sigma_\nu \bar{\sigma}_\mu = 2\eta_{\mu\nu}. \tag{4.12}$$

where

$$\bar{\sigma}_\mu = \sigma_2(\sigma_\mu)^T\sigma_2 = (\sigma_0, -\boldsymbol{\sigma}),$$

and $\sigma_\mu{}^T$ denotes the transpose of σ_μ. Using the spinor representations of the complex vectors λ'_μ and $\dot{Z}'_\mu$ we get the following identities:

$$\lambda'\bar{\dot{Z}}' + \dot{Z}'\bar{\lambda}' = 2, \tag{4.13a}$$

$$\lambda'\bar{\lambda}' = 0, \tag{4.13b}$$

$$\dot{Z}'\bar{\dot{Z}}' = 1. \tag{4.13c}$$

Then we define the spinor representation of the real null vector λ_μ as

$$\lambda = \lambda'\bar{v}_1\lambda'^\dagger[\mathrm{Tr}(\lambda')/\mathrm{Tr}(\lambda'\bar{v}_1\lambda'^\dagger)], \tag{4.14}$$

where $v_{1\mu}$ is the real part of Z'_μ. Note that when λ'_μ is real ($a_\mu = 0$), Eq. (4.14) becomes an identity. Since, in this section, we are only interested in the fields of the systems with uniform velocity, the procedure outlined above becomes simpler. The scalar functions k and l in Eqs. (4.8) and (4.9) become -1 and 0, respectively, and

$$r_{1,\alpha} = n_\alpha - \mu_\alpha,$$
$$r_{2,\alpha} = -\nu_\alpha,$$

where

$$\dot{Z}'_\mu = v_{1\mu} = n_\mu, \quad v_{2\mu} = 0.$$

Then, the null vector λ_μ can be found from Eq. (4.14) as

$$\lambda_\alpha = \frac{\mu^\beta\mu_\beta n_\alpha - \mu_\alpha + \epsilon_{\alpha\gamma\pi\beta}n^\gamma\mu^\pi\nu^\beta}{\mu^\beta\mu_\beta - 1}. \tag{4.15}$$

The derivative of λ_μ with respect to the coordinates x^μ is found as[14]

$$\lambda_{\mu,\nu} = \frac{1}{r_1{}^2 + r_2{}^2}[r_1(\eta_{\mu\nu} + \lambda_{\mu\nu} - n_\mu\lambda_\nu - n_\nu\lambda_\mu) + r_2\epsilon_{\mu\nu\alpha\beta}n^\alpha\lambda^\beta] \tag{4.16}$$

Here, we notice that the velocity vector n_μ is a Killing vector, because it satisfies the equations

$$n^\alpha\lambda_\alpha = 1,$$
$$n^\alpha\lambda_{\beta,\alpha} = 0,$$

and since it is a timelike vector, it is always possible to bring it to its rest frame $\delta_\mu{}^0$ by a Lorentz transformation. Thus

$$n^\mu\partial_\mu g_{\alpha\beta} = \partial_0 g_{\alpha\beta} = 0.$$

To find the e.m.t. we make a further choice for the form of V by taking

$$V = f(r_1)/(r_1{}^2 + r_2{}^2). \tag{4.17}$$

With this assumption, the Kerr–Schild metric can be transformed into the Boyer–Lindquist[15] coordinate system, which reads

$$ds^2 = \left(1 - \frac{2f}{\Sigma}\right)d\bar{t}^2 - \frac{\Sigma}{\Delta}dr_1^2 - \Sigma d\Theta^2 - \frac{B}{\Sigma}\sin^2\Theta\, d\bar{\phi}^2 + \frac{4af}{\Sigma}\sin^2\Theta\, d\bar{\phi}\, dt,$$

where

$$\Sigma = r_1^2 + a^2\cos^2\Theta,$$

$$\Delta = r_1^2 + a^2 - 2f,$$

$$B = (r_1^2 + a^2)^2 - a^2\Delta\sin^2\Theta.$$

Transformations from Kerr–Schild coordinates into the Boyer–Lindquist coordinates are

$$(r_1 + ia)e^{i\phi}\sin\Theta = x + iy,$$

$$r_1\cos\Theta = z,$$

$$dt = d\bar{t} + \frac{2f}{\Delta}dr_1,$$

$$d\phi = d\bar{\phi} + \frac{a}{\Delta}dr_1.$$

In the new coordinate system λ_μ and the e.m.t. $T^\mu{}_\nu$ take the forms

$$\lambda_\mu = \left(1, \frac{\Sigma}{\Delta}, 0, -a\sin^2\Theta\right),$$

$$T_{\mu\nu} = (D + 4h)u_\mu u_\nu - (D + 4h)\frac{\Sigma}{\Delta}m_\mu m_\nu - (D + 2h)g_{\mu\nu},$$

where

$$u_\mu = \frac{\Delta}{\Sigma}(1, 0, 0, -a\sin^2\Theta),$$

$$m_\mu = (0, -1, 0, 0),$$

$$D = -f_{r_1r_1}/\Sigma \quad \left(f_{r_1} = \frac{df}{dr_1}\right)$$

$$h = \frac{r_1 f_{r_1} - f}{\Sigma^2}.$$

The Kerr and the charged Kerr metrics correspond to the vanishing of $f_{r_1r_1}$. For an interior metric, the e.m.t. in Eq. (4.18) corresponds to the e.m.t. of an anisotropic perfect fluid distribution. Isotropy is destroyed in the radial direction. We note that the deviation from a perfect fluid distribution can also be regarded as arising from the contribution of a moving Nambu string.[16] Such anisotropic energy–momentum tensors have also been discussed recently by Bowers and Liang.[17] This interior metric matches to the Kerr metric on an oblate spheroid, $r_1 = r_0$ the equation of this surface being

$$r_0^4 - r_0^2(r^2 - a^2) - a^2z^2 = 0$$

where

$$r^2 = x^2 + y^2 + z^2$$

and the function $f(r_1)$ satisfies the following boundary conditions

$$f(r_0) = mr_0,$$

$$\left.\frac{d}{dr_1}f(r_1)\right|_{r_1=r_0} = m,$$

m being the total mass.

5. LINEARIZED GENERAL RELATIVITY AND TRAUTMAN'S COMPLEX TRANSLATION

The linearized theory of the gravitational field can be developed by regarding the actual Riemannian space—time as a first order perturbation of flat space—time. Here, in contrast to many authors we take $\mathcal{G}^{\mu\nu} = \sqrt{-g}\,g^{\mu\nu}$ as the gravitational field and assume that[11,18]

$$\sqrt{-g}\,g^{\mu\nu} = \eta^{\mu\nu} + 2\epsilon\phi^{\mu\nu}, \tag{5.1}$$

where ϵ is a constant and $\phi^{\mu\nu}$ is a symmetric tensor. In the linearized theory, we neglect terms of all but the first order in ϵ. Hence

$$(-\det\mathcal{G}^{\mu\nu})^{1/2} = (-\det g_{\mu\nu})^{1/2} = (-g)^{1/2} = 1 + \epsilon\phi, \tag{5.2}$$

where

$$\phi = \eta^{\mu\nu}\phi_{\mu\nu}. \tag{5.3}$$

Field equations, in terms of $\phi^{\mu\nu}$, follow as

$$GT^\mu{}_\nu = +\epsilon(-\eta^{\mu\beta}\phi^\gamma{}_{\nu,\beta\gamma} - \eta^{\alpha\beta}\phi^\mu{}_{\alpha,\beta\nu} + \Box\phi^\mu{}_\nu - \delta^\mu{}_\nu\phi^{\alpha\beta}{}_{,\alpha\beta}). \tag{5.4}$$

Without any choice of gauge, it is easy to show that

$$\partial_\mu T^\mu{}_\nu = 0,$$

and, of course, the pseudo-energy–momentum tensor vanishes in this approximation.[11]

It is remarkable that the field equation (5.4) is exactly the same as the one (2.33) which was obtained for the Kerr–Schild metric. All the metrics which are in the Kerr—Schild class are also the solution of linearized field equations (5.4), but the reverse is not true in general.

Trautman[4] has developed a method of constructing classes of new solutions to linear special relativistic partial differential equations. In particular, he used the method to produce null curling solutions of Maxwell's equations and he stated that the same method can also be used in linearized Einstein's equations. Now it becomes very clear that Newman's complex translation is nothing but Trautman's complex translation.

6. CONCLUSION

To obtain linear gravitational field equations there are two possible methods. In the first one we use an approximation procedure which leads to linearized general relativity. In the second one we put some constraints on the symmetric tensor $\phi^{\mu\nu}$ in Eq. (5.1), in such a way that the pseudo-energy—momentum tensor vanishes. In this work we showed the existence of the second possibility. It is an open question whether the Kerr—Schild coordinate system is the only coordinate system in which Einstein's equations become linear for a special geometry.

We have further obtained the gravitational field of accelerated nonspinning particles and unaccelerated

spinning particles. It is also possible to obtain the gravitational field of accelerated spinning particles. Work on the latter type solution is in progress.

If we take the null vector λ_μ as a constant null vector, the Kerr–Schild metric describes gravitational waves such that plane fronted waves are in this class of metrics with nonvanishing Weyl tensor. The corresponding space–time is of Petrov-type N.

As another possible application of our method the following remark is in order. Quantization of general relativity becomes simple for the linearized approximate theory. Since in the case of special geometries Einstein's theory becomes exactly linear in the Kerr–Schild coordinates, the same quantization procedure could also be applied in these special cases.

ACKNOWLEDGMENTS

We thank Dr. J. Bardeen, Dr. J. Friedman, and Dr. D. Eardley for helpful discussions.

*Research (Yale Report C00-3075-86) supported by the US Atomic Energy Commission under contract A1(11-1)-3075.

†On leave from the Middle East Technical University, Ankara, Turkey.

[1]R. Penrose, Ann. Phys. (N.Y.) **10**, 171 (1960); F.A. Pirani, "Introduction to Gravitational Radiation Theory" in *Lectures on General Relativity* (Prentice-Hall, Englewood Cliffs, N.J., 1965).

[2]G.C. Debney, R.P. Kerr, and A. Schild, J.Math. Phys. **10**, 1842 (1969); R. Kerr and A. Schild, "Applications of nonlinear partial differential equations in mathematical physics," *Proceedings of Symposia in Applied Mathematical* (Amer. Math. Soc., Providence, R.I., 1965), Vol. XVII, p. 199.

[3]M.M. Schiffer, R.J. Adler, J. Mark, and C. Sheffield, J. Math. Phys. **14**, 52 (1973); R.J. Adler, M. Field, and C. Sheffield, preprint (May 1974).

[4]A. Trautman, Proc. Roy. Soc. A**270**, 326 (1962); E. Newman, J. Math. Phys. **14**, 102 (1973).

[5]E.T. Newman and A.I. Janis, J. Math. Phys. **6**, 915 (1965).

[6]E.T. Newman, E. Couch, K. Chinnapared, A. Exton, A. Prakash, and K. Torrence, J. Math. Phys. **6**, 918 (1965).

[7]W.B. Bonnor and P.C. Vaidya, "General Relativity," papers in honor of J.L. Synge, edited by L. O'Raifeartaigh (Dublin Institute for Advanced Studies) p. 119 (1972).

[8]R.P. Kerr, Phys. Rev. Lett. **11**, 237 (1963).

[9]A. Einstein, Ann. Phys. **49**, 769 (1916).

[10]L.D. Landau and E.M. Lifschitz, *The Classical Theory of Fields* (Addison-Wesley, Reading, Mass., 1971), p. 306.

[11]S.N. Gupta, Proc. Phys. Soc. (Lond.) A 65, 608 (1952); Phys. Rev. 96, 1683 (1954).

[12]In this and following spections we take the gravitational constant G as unity.

[13]In this section we use only time like velocity vector, that is we take $\epsilon = 1$. Later in this section we take the imaginary part of x_μ in the third direction.

[14]Although this equation has been found by Adler *et al.* (Ref. 3) in a noncovariant way they did not draw attention to the fact that n^μ is the velocity vector of the system which generates the gravitational field.

[15]R.H. Boyer and R.W. Linquist, J. Math. Phys. **8**, 265 (1967).

[16]M. Gürses and F. Gürsey, Phys. Rev. D **11**, 967 (1975).

[17]R.L. Bowers and E.P.T. Liang, Ap. J. **188**, 657 (1974).

[18]M. Gürses, Nuovo Cimento 20A, 363 (1974).

Direct use of Young tableau algebra to generate the Clebsch-Gordan coefficients of $SU(2)$*

Jack Nachamkin

Los Alamos Scientific Laboratory, University of California, Los Alamos, New Mexico 87544
(Received 3 February 1969; final revision received 7 April 1975)

Although it is well known that the irreducible representations of the $SU(N)$ groups may be generated by using Young tableau algebra, this technique seems to have found little use for the derivation of closed algebraic expressions for the Clebsch–Gordan coefficients of these groups. A frontal attack on the derivation of these coefficients using tableau symmetrizers is described. As an example, the $SU(2)$ group illustrates the fundamental ideas behind the process.

I. INTRODUCTION

This paper describes how the Young tableau algebra technique is applied to find explicit algebraic formulas for the Clebsch—Gordan coefficients (CGC), with attention focused on the $SU(2)$ group. Physicists are familiar with this group because it is locally isomorphic to the three-dimensional rotation group R_3.

In the $SU(2)$, or R_3, case, spin operators are used to define an "internal" symmetry: Young tableau symmetrizers are applied directly to generate operators R_{JJ} which transform the state

$$|j_1, j_2; J' = j_1 + j_2,\ M' = J'\rangle$$

into the state $|j_1, j_2;\ J, M = J\rangle$. If the angular momentum vectors $\mathbf{J}_1$ and $\mathbf{J}_2$ operate on states with angular momentum j_1 and j_2, respectively, the R_{JJ} appear as

$$R_{JJ} = \sum_{\gamma=0}^{n} g_{J\gamma} J_{1-}^{\gamma} J_{2-}^{n-\gamma} \tag{1}$$

where the $g_{J\gamma}$ depend only on j_1, j_2, J, and γ. The $J_{i\xi}$'s are components of J_i, with

$$n = j_1 + j_2 - J. \tag{2}$$

In Sec. II we review some of the properties of states made up of N spin-$\frac{1}{2}$ particles coupled to a total spin S. The key property is that all states with a given S transform among themselves under the permutations on particle numbers, thus forming a basis for the $(\frac{1}{2}N+S,\ \frac{1}{2}N-S)$ irreducible representation of the symmetric group of N indices.

In Sec. III the Young tableau symmetrizers are defined and used to generate R_{JJ} from an unsymmetrized product of spin operators, identifying sums of spin operators as J_{1-} and J_{2-}. After obtaining R_{JJ}, the operator

$$R_{JM} \propto (J_{1-} + J_{2-})^{J-M} R_{JJ} \tag{3}$$

is used to find an expression for CGC:

$$C^{j_1 j_2 J}_{m_1 m_2 M} = \langle j_1 m_1, j_2 m_2 | j_1 j_2 J M\rangle.$$

II. SOME PROPERTIES OF SPIN STATES AND OPERATORS

When using a set of N coupled spin-$\frac{1}{2}$ particles, it is well known that the states with total spin S have simple properties under the group of particle number permutations. This may be seen immediately by considering the operator

$$Q = \sum_{i,j=1}^{n} [(s_{i+}s_{j-} + s_{i-}s_{j+})/2 + s_{i0}s_{j0}], \tag{4}$$

where the spin operators s_{i+}, s_{i0}, and s_{i-} operate on the spin states ϕ_{i+} "spin-up" and ϕ_{i-} "spin-down":

$$s_{i\pm}\phi_{i\pm} = 0, \tag{5a}$$

and

$$s_{i\pm}\phi_{i\mp} = \phi_{i\pm}, \tag{5b}$$

$$s_{i0}\phi_{i\pm} = \pm\tfrac{1}{2}\phi_{i\pm}. \tag{5c}$$

These operators not only obey the usual angular momentum commutation rules

$$[s_{i+}, s_{i-}] = 2s_{i0} \tag{6a}$$

and

$$[s_{i0}, s_{i\pm}] = s_{i\pm}, \tag{6b}$$

but from Eqs. (5a) and (5b)

$$(s_{i\pm})^2 = 0. \tag{7}$$

Returning to Eq. (4), note that Q is invariant under any permutation of particle numbers. Defining the components of the total spin vector S

$$S_\xi = \sum_{i=1}^{N} s_{i\xi},\quad \xi = +,\ 0,\ -. \tag{8}$$

It is also evident that Q is the dot product

$$Q = \mathbf{S}\cdot\mathbf{S}. \tag{9}$$

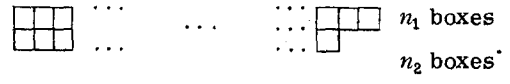
From the above, we may infer that eigenstates of Q are those for which the total spin S is a good quantum number, and thus form the bases for the irreducible representations of the symmetric group on N particles (Refs. 1, 2). The irreducible representation of the symmetric group corresponding to spin S is denoted by the two-rowed Young talbeau with n_1 boxes in the first row and n_2 boxes in the second row (Ref. 1),

$$n_1 = \tfrac{1}{2}N + S \tag{10a}$$

and

$$n_2 = \tfrac{1}{2}N - S. \tag{10b}$$

Graphically this looks like



(11)

Such diagrams are referred to by the partition of N into

n_1 into n_2, $[n_1, n_2]$. The state with the highest spin $S = \frac{1}{2}N$ corresponds to the partition $[N, 0]$, usually abbreviated to $[N]$.

In matrix notation the operators and states may be replaced by

$$s_{i+} \to \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}_i, \quad s_{i-} \to \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}_i, \quad s_{i0} \to \tfrac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}_i, \tag{12}$$

$$\phi_{i+} \to \begin{pmatrix} 1 \\ 0 \end{pmatrix}_i \text{ and } \phi_{i-} \to \begin{pmatrix} 0 \\ 1 \end{pmatrix}_i. \tag{13}$$

From Eqs. (7) and (8), it can be shown by induction that

$$(S_\pm)^r = r! \sum_{i_1 < i_2 \cdots i_r N} s_{i_1 \pm} \cdots s_{i_r \pm}. \tag{14}$$

Because the sum in Eq. (9) contains $\binom{N}{r}$ terms, where

$$\binom{N}{r} = \binom{N}{N-r} = \frac{N!}{r!\,(N-r)!}, \tag{15}$$

and because $(S_\pm)^r$ applied to a state with definite S changes only the magnetic quantum number, applying $(S_-)^r$ to the state $|S = \frac{1}{2}N, M_s = \frac{1}{2}N\rangle$, which is made up of only one term, yields

$$(S_-)^r |S = \tfrac{1}{2}N, M = \tfrac{1}{2}N\rangle = r!\binom{N}{r}^{1/2} \Big| S = \tfrac{1}{2}N, M = \tfrac{1}{2}N - r\Big\rangle. \tag{16a}$$

This may be generalized immediately to the usual angular momentum stepping operator J_-, i.e.,

$$(J_-)^r |J, M = J\rangle = r! \binom{2J}{r}^{1/2} \Big| J, M = J - r\Big\rangle, \tag{16b}$$

where

$$J_- |J, M\rangle = \sqrt{(J+M)(J-M+1)}\,|J, M-1\rangle. \tag{17}$$

In effect, the operators S_ξ may be regarded in the same way as the angular momentum operators J_ξ, because they have the same commutation properties (Ref. 3). The only difference is that the S_ξ's have an internal structure that is not quite evident in the J_ξ's but is used in Sec. III.

Before proceeding, we break the S_ξ's into sums, each containing $2j_i$ terms

$$S_\xi = \sum_i s_{i\xi}, \tag{18a}$$

where

$$S_{1\xi} = \sum_{k=1}^{2j_1} s_{k\xi}, \tag{18b}$$

$$S_{2\xi} = \sum_{k=2j_1+1}^{2j_1+2j_2} s_{i\xi}, \text{ etc.} \tag{18c}$$

Note that now

$$(S_{i\pm})^{2j_i+1} = 0. \tag{18d}$$

These sums will now represent angular momentum operators.

III. TABLEAU AND ANGULAR MOMENTUM SHIFTING OPERATORS

Given the sums

$$S = S_{1\xi}\left(\equiv \sum_{k=1}^{2j_1} s_{k\xi}\right) + S_{2\xi}\left(\equiv \sum_{k=2j_1+1}^{2j_1+2j_2=N} s_{k\xi}\right), \tag{19a}$$

and drawing the analogy to

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2, \tag{19b}$$

we wish to generate a state with a total spin

$$J - \tfrac{1}{2}N - n = j_1 + j_2 - n, \tag{20}$$

while preserving the property that the first $2j_1$ spin states are coupled to spin j_1 while the last $2j_2$ spin states are coupled to j_2. This state would then correspond to one of two coupled angular momenta,

$$|j_1, j_2; J, M\rangle. \tag{21}$$

Once again, expressed in terms of spins, there are "internal" symmetries that are not immediately evident in Eq. (21).

The desired state corresponds to the Young tableau $[N-n, n]$. There is a very simple way for generating such states (Refs. 2, 4): Fill in the boxes of the Young tableau with the numbers $1, 2, \cdots, N = 2j_1 + 2j_2$ in any order. Next form the operators that symmetrize the terms belonging to the particle numbers found in separate rows, and then for the operators that antisymmetrize the particle numbers in each column.

For example, consider the partition $[F] = [3, 2]$ with the numbers $1, \cdots, 5$ written in order:

1	2	3
4	5	

(22)

The operator that symmetrizes the subscripts 1, 2, and 3 is denoted by $[[1, 2, 3]]$, and the operator symmetrizing 4 and 5 is denoted by $[[4, 5]]$. Antisymmetrizers are denoted by two curly brackets, such as $\{1, 4\}$ and $\{2, 5\}$. Symmetrizers $[[a_1, a_2, \cdots, a_r]]$ in terms of permutations are

$$[[a_1, a_2, \cdots, a_r]] = \sum_{\gamma=1}^{r!} P_\gamma. \tag{22a}$$

Similarly,

$$\{i, j\} = 1 - P_{ij}, \tag{22b}$$

where P_{ij} permutes the subscripts i and j.

A tableau symmetrizer for $SU(2)$ may be defined as a product of the two $[[\ ]]$ operators and the n $\{\}$ operators, taken in any order. The most convenient symmetrizer is usually determined by the specific problem, but we will adhere to the convention that all the $\{\}$'s will be to the left of the $[[\ ]]$'s. Thus, Eq. (22) corresponds to

$$Y_{[F]} = \{1, 4\}\{2, 5\}[[1, 2, 3]][[4, 5]]. \tag{22c}$$

If, in addition, we want to symmetrize the suscripts 1, 2 and 3, 4, 5, we would need the operator

$$Y'_{[F]} = [[12]][[345]]Y_{[F]}. \tag{22d}$$

We now follow the convention that the numbers $1, 2, \cdots, 2j_1 + 2j_2 = N$ will be written in order, that is, the first

row will hold the numbers $1, 2, \cdots, N-n$ while the second row reads $N-n+1,\ N-n+2 \cdots N$. If

$$r_i = N-n+i, \tag{23}$$

then the first row reads $1, 2, \cdots, r_0$ and the second $r_1, r_2, \cdots, r_n = N$. The Young tableau symmetrizer, which corresponds to Eq. (21), is

$$Y_{[N-n,n]} = [[1, \cdots, 2j_1]][[2j_1+1, \cdots, N]] \times\{1, r_1\}\cdots\{n, r_n\}\times[[1, \cdots, r_0]][[r_1 \cdots r_n]]. \tag{24}$$

In this way subscripts belonging to the same j cannot occur in the same column.

Applying the operator in Eq. (24) to an unsymmetrized product or sum of products of n operators of the form s_{i-} will yield either zero or an angular momentum shifting operator. In schematic form then

$$\rho R_{JJ}(J_{1-}J_{2-}) = Y'_{[N-n,n]} \times(\text{product of } n\ s_{i-}\text{'s}) \tag{25a}$$

and

$$R_{JJ}(J_1J_2)\,|j_1j_2; j_1+j_2, j_1+j_2\rangle = \rho\,|j_1j_2, J=j_1+j_2-n, M=J\rangle, \tag{25b}$$

where ρ is a normalization constant. The most convenient product of n s_{i-}'s seems to be

$$P = s_{r_1-}s_{r_2-}\cdots s_{r_n-}\ ; \tag{26}$$

because

$$[[r_1 \cdots r_n]]P = n!\,P \tag{27}$$

and

$$[[1 \cdots r_0]]P = r_0!P, \tag{28}$$

$Y'_{[N-n,n]}P$ reduces to

$$\rho R_{JJ}(J_1, J_2) = Y'_{[N-n,n]}P = n!\,r_0![[1, \cdots, 2j_1]] \times[[2j_1+1, \cdots, N]]\times\{1, r_1\}\cdots\{n, r_n\}P. \tag{29}$$

Straightforward simple manipulation, as shown in the Appendix, yields

$$\rho R_{JJ}(J_1J_2) = r_0!\,(2j_1)!\,(2j_2)! \times\sum_{q=0}^{n}\binom{n}{q}^2\Big/\binom{2j_1}{q}\binom{2j_2}{n-q}J_{1-}J_{2-}\,. \tag{30}$$

Applying $\rho P_{jj}(J_1J_2)$ to $|j_1j_1\rangle\,|j_2j_2\rangle$ and using Eqs. (15), (16a), (20), and (23), we obtain

$$\rho\,|j_1j_2J=j_1+j_2-n, M=J\rangle = (j_1+j_2+J)!\,(2j_1)!\,(2j_2)!\,(j_1-j_2-J)! \times\sum_{q=0}^{n}(-1)^q\binom{n}{q}\Big/\left[\binom{2j_1}{q}\binom{2j_2}{n-q}\right]^{1/2} \times|j_1, j_1-q\rangle\,|j_2, j_2-n+q\rangle. \tag{31}$$

To evaluate ρ, Eq. (15) yields

$$\binom{n}{q}\Big/\left[\binom{2j_1}{q}\binom{2j_2}{n-q}\right]^{1/2} = n!\,/\sqrt{(2j_1)!\,(2j_2)!} \times\left(\frac{(2j_1-q)!\,(2j_2-n+q)!}{q!\,(n-q)!}\right)^{1/2}. \tag{32}$$

Making the identifications

$$m_1 = j_1-q,\ m_2 = j_2-n+q, \tag{33}$$

Eq. (31) now reads as

$$\rho\,|j_1j_2; J, M=J\rangle = (j_1+j_2+J)!\,\sqrt{(2j_1)!\,(2j_2)!} \times[(j_1+j_2-J)!\,]^2\times\sum_{m_1+m_2=J}(-1)^{j_1-m_1} \times\left(\frac{(j_1+m_1)!\,(j_2+m_2)!}{(j_1-m_1)!\,(j_2-m_2)!}\right)^{1/2}|j_1m_1\rangle\,|j_2m_2\rangle \tag{34}$$

and

$$\rho^2 = (j_1+j_2+J)!\,\sqrt{(2j_1)!\,(2j_2)!} \times[(j_1+j_2-J)!\,]^2 \times\sum_{m_1+m_2=M}\frac{(j_1+m_1)!\,(j_2+m_2)!}{(j_1-m_1)!\,(j_2-m_2)!} \tag{35}$$

Hamermesh (Ref. 2, pp. 372 and 373), shows that

$$\sum_{m_1+m_2=M}\frac{(j_1+m_1)!\,(j_2+m_2)!}{(j_1-m_1)!\,(j_2-m_2)!} = (j_1-j_2+J)! \times(j_2-J_1+J)!\binom{J+j_1+j_2+1}{j_1+j_2-J}. \tag{36}$$

Thus,

$$R_{JJ} = \left(\binom{2j_1}{n}\binom{2j_2}{n}\binom{j_1+j_2+J+1}{2J+1}\right)^{-1/2}\sum_q(-1)^q\binom{n}{q} \times J_{1-}^qJ_{2-}^{n-q}/q!\,(n-q)!\,\binom{2j_1}{q}\binom{2j_1}{n-q}. \tag{37}$$

Multiplying R_{JJ} by

$$J_-^{J-M} = (J_{1-}+J_{2-})^{J-M} = \sum_p\binom{J-M}{p}J_{1-}^pJ_{2-}^{J-M-p} \tag{38}$$

and using Eq. (16a), we obtain after some manipulation

$$|j_1j, JM\rangle = \left[\binom{2j_1}{n}\binom{2j_2}{n}(j_1+j_2+J+1)\right]^{1/2} \times\left[\binom{2j_1}{j_1+m_1}\binom{2j_2}{j_2+m_2}\Big/\binom{2j_2}{j_2+M}\right]^{1/2} \times\left[\sum_q(1)^q\binom{n}{q}\binom{j_1-m_1}{q}\binom{j_2-m_2}{n-q}\Big/\binom{2j_1}{q}\binom{2j_2}{n-q}\right] \times|j_1m_1\ j_2m_2\rangle. \tag{39}$$

Within a phase, then, it can be shown that

$$\langle j_1m_1, j_2m_2\,|J=j_1+j_2-n, M\rangle = \left(\binom{2j_1}{n}\binom{2j_2}{n}\binom{2j_1}{j_1+m_1}\binom{2j_2}{j_2+m_2} \Big/\binom{j_1+j_2+J+1}{n}\binom{2J}{J+M}\right)^{1/2}$$

$$\times\sum_q(-1)^q\binom{j_1+j_2-J}{q}$$
$$\times\binom{j_2-j_1+J}{j_2+m_2-q}\binom{j_1-j_2+J}{j_1-m_1-q}. \tag{40}$$

The above formulas are expressed in terms of binomial coefficients instead of factorials for ease of eventual computation.

IV. DISCUSSION

The analysis in this paper was aimed primarily to obtaining the CGC of R_3, which is locally isomorphic to $SU(2)$. From the group-theoretic viewpoint, the most important relations are expressed by Eqs. (5), (6), and (7), yielding the group's structure. Taking advantage of the fact that irreducible representations (IR's) of $SU(2)$ may be represented by two-rowed Young tableaux and that the highest weight state is totally symmetric leads to a straightforward prescription for generating all desired states.

Therefore, the generalization to higher groups is almost immediate. The IR's of the $SU(N)$ group may be represented by $(N+1)$-rowed Young tableaux. Because the highest weight state of a product representation is nondegenerate, we may find the correct symmetrizing operators that will project onto the desired state, i.e., the correct symmetry-breaking shift operators analogous to R_{JJ} of Eq. (1).

Subsequent work is being directed toward finding such operators for higher groups, mainly for obtaining CGC by computer. Future notes will deal with higher groups and the "strategies" imposed by nonlinear Young diagrams and multiple reducibility.

APPENDIX : SYMMETRY OPERATORS

The symmetrizer

$$[[a_r]]\equiv[[a_1,\cdots,a_r]]\equiv\sum_{\gamma=1}^{n!}P_\gamma \tag{A1}$$

contains the $r!$ possible permutations defined on $a_1,\cdots,a_r$. We are operating with $[[a_r]]$ on products and sums of products of the s_{i-}, remembering that $(s_{1-})^2=0$, and that i's belong only to the set

$$a_1, a_2, \cdots, a_r.$$

(It is trivial to extend this to the general case.)

Thus, consider

$$P_n=s_{i_1-}s_{i_2-}\cdots s_{i_n-} \tag{A2}$$

and

$$i_1<i_2<\cdots<i_n. \tag{A2'}$$

There are $\binom{r}{n}$ different ways of choosing a P_n, and the operator $[[a_r]]$ applied to a P_n produces a sum containing all of the possible P_n's the same number of times. Operating with $[[a_r]]$ on the already symmetric sum of all the P_n's reproduces that sum times the factor $r!$. That means

$$[[a_r]]P_n=\left[r!\Big/\binom{r}{n}\right]\times(\text{sum of all } P_n\text{'s}). \tag{A3}$$

If we define

$$J_-=\sum_{i=1}^{r}s_{a_i-},$$

then from Eq. (14)

$$[[a_r]]P_n=\left[\frac{r!}{n!}\Big/\binom{r}{n}\right]J_-^n. \tag{A4}$$

To evaluate the right-hand side of Eq. (29), we define the quantities $p^{i,j\cdots}$, where

$$p^{i,j\cdots}=\prod_{k\neq i,j}s_{k-}. \tag{A5}$$

The action of $\{i,r_i\}$ on P_n is

$$\{i,r_i\}P_n=P_n-s_{i-}p^i,$$

so that

$$\{1,r_1\}\{2,r_2\}\cdots\{n,r_n\}P=P-\sum_{i=1}^{n}s_{i-}p^i$$
$$+\sum_{j>1}s_{i-}s_{i-}p^{ij}\pm\cdots+(-1)^n s_{1-}s_{2-}\cdots s_{n-}. \tag{A6}$$

Using Eq. (A3) applied to the P's

$$[[2j_1+1,\cdots,2j_1+2j_2]]\{1,r_1\}\cdots\{nr_n\}$$
$$=(2j_2)!\sum_q\left\{(-1)^q J_{2-}^{n-q}\hat{J}_{1-}^q\Big/\left[(n-q)!\,q!\binom{n_1}{n-q}\right]\right\}, \tag{A7}$$

where J_{1-} is defined by Eqs. (19a) and (19b) and

$$\hat{J}_{1-}=\sum_{i=1}^{n}s_{i-}. \tag{A8}$$

The last step is to apply Eq. (A3) to the $\hat{J}_{1-}^q$, each of which contain $\binom{n}{q}$ terms so that

$$[[1,\cdots,2j_1]][[2j_1+1,\cdots,2j_1-2j_2]]\{1,r_1\}$$
$$\{2,r_2\}\cdots\{nr_n\}P=\frac{(2j_1)!\,(2j_2)!}{n!}$$
$$\times\sum_q(-1)^q\binom{n}{q}^2\binom{2j_1}{q}\binom{2j_2}{n-q}J_{1-}^{n-q}. \tag{A9}$$

*This research was performed under the auspices of the U.S. Atomic Energy Commission.

[1]D.E. Rutherford, *Substitutional Analysis* (Edinburgh U.P., Edinburgh, 1948).
[2]M. Hamermesh, *Group Theory and Its Application to Physical Problems* (Addison-Wesley, Reading, Mass., 1962).
[3]M.E. Rose, *Elementary Theory of Angular Momentum* (Wiley, New York, 1957).
[4]A. de Shalit and I. Talmi, *Nuclear Shell Theory* (Academic, New York, 1963).

Some solutions of complex Einstein equations

J. F. Plebański*

Centro de Investigación y de Estudios Avanzados del I.P.N., AP 14-740, México 14, D.F., Mexico
(Received 11 April 1975)

Complex V_4's are investigated where $\Gamma_{\dot{A}\dot{B}} = 0$ and therefore *a fortiori* equations $G_{ab} = 0$ are fulfilled. A general theory of spaces of this type is outlined and examples of nontrivial solutions of all degenerate algebraic types are provided.

1. NOTATION, CONVENTIONS AND TERMINOLOGY

The fact that the large family of type D solutions of Einstein–Maxwell equations[1] can be explicitly exhibited as a real cross section of a complex double Kerr–Schild metric and that its generalization derived with Demiański[2] has the same property, has stimulated the author's interest in (i) the complex Riemannian geometry and (ii) the double Kerr–Schild metrics as such. The second subject will be extensively studied in a forthcoming publication by Schild and the present author. This article is basically intended as an outline of some general facts and results concerned with the complex Riemannian geometry as such.

The complex four-dimensional Riemannian space is a pair consisting of a four-dimensional differential analytic manifold M_4 and (with $e^a \in \Lambda^1$, $a = 1,2,3,4$) the metric given by

$$V_4 : ds^2 =: 2e^1e^2 + 2e^3e^4 = g_{ab}e^a e^b \in \Lambda^1 \otimes \Lambda^1. \tag{1.1}$$

The tetradial indices—the forms e^a define a null tetrad—are to be manipulated by

$$(g_{ab}) = \begin{pmatrix} 0&1&0&0 \\ 1&0&0&0 \\ 0&0&0&1 \\ 0&0&1&0 \end{pmatrix} = (g^{ab}). \tag{1.2}$$

The Pauli matrix is a 2×2 matrix with entries in Λ^1 and connects Λ^1 with spinorial objects

$$(g^{A\dot{B}}) := \sqrt{2} \begin{pmatrix} e^4, & e^2 \\ e^1, & -e^3 \end{pmatrix}. \tag{1.3}$$

Thus

$$ds^2 = -\det(g^{A\dot{B}}) = \tfrac{1}{2}\epsilon_{AB}\epsilon_{\dot{C}\dot{D}}g^{A\dot{C}}g^{B\dot{D}}. \tag{1.4}$$

Consider now two independent sets of 2×2 complex matrices with determinant equal to one:

$$\begin{aligned} &l^{A'}{}_{A} \in SL(2,\mathbb{C}) \to \det(l^{A'}{}_{A}) = 1, \\ &\bar{l}^{\dot{B}'}{}_{\dot{B}} \in S\bar{L}(2,\mathbb{C}) \to \det(\bar{l}^{\dot{B}'}{}_{\dot{B}}) = 1. \end{aligned} \tag{1.5}$$

It is now obvious that $e^{a'}$ defined by

$$\sqrt{2} \begin{pmatrix} e^{4'}, & e^{2'} \\ e^{1'}, & -e^{3'} \end{pmatrix} := (l^{A'}{}_{A}\bar{l}^{\dot{B}'}{}_{\dot{B}}g^{A\dot{B}}) \tag{1.6}$$

still gives $2e^{1'}e^{2'} + 2e^{3'}e^{4'} = ds^2$. Thus, we can consider as the tetradial gauge group

$$\mathcal{G} := SL(2,\mathbb{C}) \times S\bar{L}(2,\mathbb{C}). \tag{1.7}$$

[More precisely, one could work with the tetradial gauge imposing on the 2×2 matrices only the condition $\det(l^{A'}{}_{A})\det(\bar{l}^{\dot{B}'}{}_{\dot{B}}) = 1$; the obvious group of the tetradial gauge in V_4, $O(4,\mathbb{C})$, decomposes into the product $O(3,\mathbb{C}) \times O(3,\mathbb{C})$; this fact permits us to identify the tetradial gauge group with $\mathcal{G}$.]

Although we are interested in the complex geometry, we can proceed with the standard conventions of the spinor calculus, transforming respectively undotted and dotted indices by objects from $SL(2,\mathbb{C})$ and $S\bar{L}(2,\mathbb{C})$, forgetting however about the condition $(l^{A'}{}_{A})^* = \bar{l}^{\dot{A}'}{}_{\dot{A}}$ from the case of real V_4 of the signature $(+++-)$, which causes $\mathcal{G}$ to reduce to $SL(2,\mathbb{C})$ homomorphic with $SO(3,1,R)$.

[In the "real" case, the permitted gauge transformations must conserve the condition $(e^2)^* = e^1$; in the complex geometry where this condition is abandoned, matrices $SL(2,\mathbb{C})$ and $S\bar{L}(2,\mathbb{C})$ are independent.]

The equation

$$g^{A\dot{B}} \wedge g^{C\dot{D}} = S^{AC}\epsilon^{\dot{B}\dot{D}} + \epsilon^{AC}S^{\dot{B}\dot{D}} \tag{1.8}$$

defines in Λ^2 the two forms of the spin tensor:

$$\begin{aligned} S^{AB} &:= \tfrac{1}{2}\epsilon_{\dot{R}\dot{S}}g^{A\dot{R}} \wedge g^{B\dot{S}} = \tfrac{1}{2}e^a \wedge e^b S_{ab}{}^{AB}, \\ S^{\dot{A}\dot{B}} &:= \tfrac{1}{2}\epsilon_{RS}g^{R\dot{A}} \wedge g^{S\dot{B}} = \tfrac{1}{2}e^a \wedge e^b S_{ab}{}^{\dot{A}\dot{B}}. \end{aligned} \tag{1.9}$$

Explicitly, we have:

$$\begin{aligned} &S^{11} = 2e^4 \wedge e^2, \quad S^{12} = e^1 \wedge e^2 + e^3 \wedge e^4, \quad S^{22} = 2e^3 \wedge e^1, \\ &S^{\dot{1}\dot{1}} = 2e^4 \wedge e^1, \qquad\qquad S^{\dot{2}\dot{2}} = 2e^3 \wedge e^2. \\ &S^{\dot{1}\dot{2}} = -e^1 \wedge e^2 + e^3 \wedge e^4, \end{aligned} \tag{1.10}$$

The forms S^{AB} and $S^{\dot{A}\dot{B}}$ are respectively self-dual and anti-self-dual under Hodge's star operation:

$$*S^{AB} = S^{AB}, \quad *S^{\dot{A}\dot{B}} = -S^{\dot{A}\dot{B}}. \tag{1.11}$$

[The duality operation $*$ acting on $\omega \in \Lambda^p$,

$$\omega = \frac{1}{p!}\omega_{a_1\dots a_p}e^{a_1} \wedge \circ\cdots\circ \wedge e^{a_p}, \tag{1.12}$$

maps it into

$$\begin{aligned} *\omega = \frac{1}{p!p'!} &\circ \exp\frac{i\pi}{2}(pp' - 2)\epsilon^{a_1\dots a_p}{}_{b_1\dots b_{p'}} \\ &\circ \omega_{a_1\dots a_p}e^{b_1} \wedge \circ\cdots\circ \wedge e^{b_{p'}} \end{aligned} \tag{1.13}$$

where indices of the Levi–Civita symbol $\epsilon_{a_1\dots a_4}$ are manipulated by (g^{ab}) and $p' := 4 - p$; this definition is so constructed that $**\omega = \omega$ for every ω.]

Now, the connection forms $\Gamma_{ab}=\Gamma_{[ab]}\in\Lambda^1$ are defined by the first structure equations[3]

$$de^a=e^b\wedge\Gamma^a{}_b \tag{1.14}$$

and if ∂_a is the inverse tetrad (acting on any scalar T: $(d-e^a\partial_a)T=0$), then the Ricci rotation coefficients ($\Gamma_{abc}e^c=:\Gamma_{ab}$) can be computed from the commutators

$$\partial_b\partial_a-\partial_a\partial_b=(\Gamma^s{}_{ab}-\Gamma^s{}_{ba})\partial_s. \tag{1.15}$$

The objects Γ_{ab} are equivalent to their spinorial images

$$\begin{aligned}&\Gamma_{AB}:=-\tfrac{1}{4}\Gamma_{ab}S^{ab}{}_{AB},\quad \bar\Gamma_{\dot A\dot B}:=-\tfrac{1}{4}\Gamma_{ab}S^{ab}{}_{\dot A\dot B}\\&\Gamma_{ab}=-\tfrac{1}{2}S_{ab}{}^{AB}\Gamma_{AB}-\tfrac{1}{2}S_{ab}{}^{\dot A\dot B}\bar\Gamma_{\dot A\dot B}.\end{aligned} \tag{1.16}$$

The explicit form of these relations is

$$\begin{aligned}(\Gamma_{AB})&=-\frac{1}{2}\begin{pmatrix}2\Gamma_{42}, & \Gamma_{12}+\Gamma_{34}\\ \Gamma_{12}+\Gamma_{34}, & 2\Gamma_{31}\end{pmatrix},\\(\bar\Gamma_{\dot A\dot B})&=-\frac{1}{2}\begin{pmatrix}2\Gamma_{41}, & -\Gamma_{12}+\Gamma_{34}\\ -\Gamma_{12}+\Gamma_{34}, & 2\Gamma_{32}\end{pmatrix}\end{aligned} \tag{1.17}$$

and under the tetradial gauge $\mathcal{G}$ from (1.7) the connections transform according to

$$\begin{aligned}\Gamma^{A'}{}_{B'}&=l^{A'}{}_A l^{-1}{}_{B'}{}^B\Gamma^A{}_B+l^{A'}{}_S\,dl^{-1}{}_{B'}{}^S,\\ \bar\Gamma^{\dot A'}{}_{\dot B'}&=\bar l^{\dot A'}{}_{\dot A}\bar l^{-1}{}_{\dot B'}{}^{\dot B}\bar\Gamma^{\dot A}{}_{\dot B}+\bar l^{\dot A'}{}_{\dot S}\,d\bar l^{-1}{}_{\dot B'}{}^{\dot S}.\end{aligned} \tag{1.18}$$

[This fact will be of crucial importance in our further considerations: (1.18) exhibits explicitly the (irreducible) decomposition of the transformational properties of Γ's which occurs on the level of the spinorial images and which corresponds to the independent factors in the (cross) factors (1.7) for the gauge group.]

Notice that while

$$dg^{A\dot B}=g^{A\dot S}\wedge\bar\Gamma^{\dot B}{}_{\dot S}+g^{S\dot B}\wedge\Gamma^A{}_S, \tag{1.19}$$

then

$$dS^{AB}=-3S^{(AB}\Gamma^{C)}{}_C,\quad dS^{\dot A\dot B}=-3S^{(\dot A\dot B}\bar\Gamma^{\dot C)}{}_{\dot C} \tag{1.20}$$

Consider now the two spinorial curvature forms

$$\begin{aligned}R^A{}_B&:=d\Gamma^A{}_B+\Gamma^A{}_S\wedge\Gamma^S{}_B\\&=-\tfrac{1}{2}C^A{}_{BCD}S^{CD}+\frac{R}{24}S^A{}_B+\tfrac{1}{2}C^A{}_{B\dot C\dot D}S^{\dot C\dot D},\end{aligned} \tag{1.21a}$$

$$\begin{aligned}\bar R^{\dot A}{}_{\dot B}&:=d\bar\Gamma^{\dot A}{}_{\dot B}+\bar\Gamma^{\dot A}{}_{\dot S}\wedge\bar\Gamma^{\dot S}{}_{\dot B}\\&=-\tfrac{1}{2}\bar C^{\dot A}{}_{\dot B\dot C\dot D}S^{\dot C\dot D}+\frac{R}{24}S^{\dot A}{}_{\dot B}+\tfrac{1}{2}C_{CD}{}^{\dot A}{}_{\dot B}S^{CD}.\end{aligned} \tag{1.21b}$$

The equalities in (1.21) correspond to the Cartan structure equations $d\Gamma^a{}_b+\Gamma^a{}_s\wedge\Gamma^s{}_b=\frac{1}{2}R^a_{bcd}e^c\wedge e^d$, with R_{abcd} being the tetrad components of the Riemann tensor. The symbols used in (1.21) are defined as follows: Let $R_{ab}:=R^s{}_{abs}=$ Ricci tensor, $R=R^s_s=$ scalar curvature, $C_{ab}:=R_{ab}-\frac{1}{4}Rg_{ab}=$ Ricci tensor with extracted trace, and let C_{abcd} be the conformal curvature. Then we have

$$\begin{aligned}C_{ABCD}&:=\tfrac{1}{16}S^{ab}{}_{AB}C_{abcd}S^{cd}{}_{CD}=C_{(ABCD)},\\ \bar C_{\dot A\dot B\dot C\dot D}&:=\tfrac{1}{16}S^{ab}{}_{\dot A\dot B}C_{abcd}S^{cd}{}_{\dot C\dot D}=\bar C_{(\dot A\dot B\dot C\dot D)},\\ C_{abcd}&=\tfrac{1}{4}S_{ab}{}^{AB}C_{ABCD}S_{cd}{}^{CD}+\tfrac{1}{4}S_{ab}{}^{\dot A\dot B}\bar C_{\dot A\dot B\dot C\dot D}S_{cd}{}^{\dot C\dot D},\end{aligned} \tag{1.22}$$

and

$$C_{AB\dot C\dot D}=\tfrac{1}{4}g^a{}_{A\dot C}g^b{}_{B\dot D}C_{ab}\leftrightarrow C_{ab}=g_a{}^{A\dot C}g_b{}^{B\dot D}C_{AB\dot C\dot D} \tag{1.23}$$

where of course $g^{A\dot B}=:e^ag_a{}^{A\dot B}$. Thus, C_{ABCD} and $\bar C_{\dot A\dot B\dot C\dot D}$ represent the two (independent in the complex geometry) spinorial images of the conformal curvature; $C_{AB\dot C\dot D}=C_{(AB)\dot C\dot D}=C_{AB(\dot C\dot D)}$ is the spinorial image of C_{ab}.

Following Debney, Kerr and Schild[4] we will also use the abbreviations for the independent components of C_{ABCD} and $\bar C_{\dot A\dot B\dot C\dot D}$:

$$\begin{aligned}&C^{(5)}=2C_{1111},\quad C^{(4)}=2C_{1112},\\&C^{(3)}=2C_{1122},\quad C^{(2)}=2C_{1222},\quad C^{(1)}=2C_{2222}\end{aligned} \tag{1.24}$$

and

$$\begin{aligned}&\bar C^{(5)}=2\bar C_{\dot1\dot1\dot1\dot1},\quad \bar C^{(4)}=2\bar C_{\dot1\dot1\dot1\dot2},\\&\bar C^{(3)}=2\bar C_{\dot1\dot1\dot2\dot2},\quad \bar C^{(2)}=2\bar C_{\dot1\dot2\dot2\dot2},\quad \bar C^{(1)}=2\bar C_{\dot2\dot2\dot2\dot2}.\end{aligned} \tag{1.25}$$

It presents now no difficulty to generalize the Petrov–Penrose classification for the case of a complex V_4. We will say that K^A and $\bar K^{\dot A}$ are the two types of P-spinors (Penrose's spinors—undotted and dotted) iff

$$C_{ABCD}K^AK^BK^CK^D=0,\quad \bar C_{\dot A\dot B\dot C\dot D}\bar K^{\dot A}\bar K^{\dot B}\bar K^{\dot C}\bar K^{\dot D}=0. \tag{1.26}$$

Of course, we are interested in P-spinors of both types which are linearly independent in pairs. Thus, following Penrose,[4] by applying the fundamental theorem of algebra, we conclude that, representing the spinorial images according to

$$C_{ABCD}=\alpha_{(A}\beta_B\gamma_C\delta_{D)},\quad \bar C_{\dot A\dot B\dot C\dot D}=\bar\alpha_{(\dot A}\bar\beta_{\dot B}\bar\gamma_{\dot C}\bar\delta_{\dot D)}, \tag{1.27}$$

we can describe the schemes of possible coincidences for the spinors of both types as the Cartesian product of two Penrose's diagrams:

$$\begin{aligned}&C_{ABCD}\\&[1-1-1-1]\\&\quad[2-2]\qquad[2-1-1]\\&\quad[-]\qquad\quad[4]\qquad[3-1]\\&\bar C_{\dot A\dot B\dot C\dot D}\\&\times[1-1-1-1]\\&\quad[2-2]\qquad[2-1-1]\\&\quad[-]\qquad\quad[4]\qquad[3-1].\end{aligned} \tag{1.28}$$

Thus, it makes sense to talk about the conformal curvature of a complex V_4 as being of the types $[1-1-1-1]\otimes[2-2]$, $[2-2]\otimes[4]$, etc. In symbols of this type, the first place we reserve for C_{ABCD} and the second, for $\bar C_{\dot A\dot B\dot C\dot D}$.[5] Without possibility of confusion, we can also adopt here the terminology from the real V_4's of the signature $(+++-)$, talking about the spaces of the types $G\otimes N$, $D\otimes D$, $N\otimes N$, etc., identifying, as customary, $G\equiv[1-1-1-1]$, $D\equiv[2-2]$, $N\equiv[4]$. Of, course, the space $[-]\otimes[-]$ has vanishing conformal curvature and hence is conformally flat.

It presents no difficulty to generalize the concept of

TABLE I.

Hellish	Earthly	Heavenly
$S^{\dot{A}\dot{B}}$	$g^{\dot{A}\dot{B}}$	S^{AB}
$\bar{\Gamma}^{\dot{A}\dot{B}}$		Γ_{AB}
$\bar{C}^{\dot{A}\dot{B}\dot{C}\dot{D}}$	$C^{\dot{A}\dot{B}\dot{C}\dot{D}}$	C_{ABCD}
	R	

the Debever vectors on the case considered. Let K_μ be a null vector; then it can always be considered as induced by a pair of spinors of both types:

$$K_\mu\, dx^\mu = -\tfrac{1}{2} g^{\dot{A}B} K_A \bar{K}^{\dot{B}}. \tag{1.29}$$

The object

$$K_{\alpha\beta\gamma\delta} \equiv \mathcal{D}(K_\mu) C_{\alpha\beta\gamma\delta} := K^\mu K_{[\alpha} C_{\beta]\mu\nu[\gamma} K_{\delta]} K^\nu \tag{1.30}$$

has all symmetries of the conformal curvature tensor and hence is characterized entirely by its two spinorial images analogous to C_{ABCD} and $\bar{C}^{\dot{A}\dot{B}\dot{C}\dot{D}}$ which can be easily computed as given by

$$\begin{aligned} K_{ABCD} &= \tfrac{1}{16} K_A K_B K_C K_D \cdot \bar{C}^{\dot{P}\dot{Q}\dot{R}\dot{S}} \bar{K}_{\dot{P}} \bar{K}_{\dot{Q}} \bar{K}_{\dot{R}} \bar{K}_{\dot{S}}, \\ K^{\dot{A}\dot{B}\dot{C}\dot{D}} &= \tfrac{1}{16} \bar{K}^{\dot{A}} \bar{K}^{\dot{B}} \bar{K}^{\dot{C}} \bar{K}^{\dot{D}} \cdot C_{PQRS} K^P K^Q K^R K^S. \end{aligned} \tag{1.31}$$

It easily follows that the linear operation $\mathcal{D}(K_\mu)$ defined by (1.30) is nil-potent:

$$\mathcal{D}^2(K_\mu) = 0. \tag{1.32}$$

Moreover, one easily infers that

$$\mathcal{D}(K_\mu) C_{\alpha\beta\gamma\delta} = 0 \tag{1.33}$$

is the necessary and sufficient condition for both K_A and $\bar{K}^{\dot{B}}$ to be generic P-spinors of both types. We will thus call the null vectors defined by (1.29), with both K_A and $\bar{K}^{\dot{B}}$ being generic P-spinors, the generalized DP (Debever–Penrose) vectors. In a general complex V_4 all of them are complex. The number of different (in the sense of the linear independence in pairs) DP vectors depends on the type of V_4; e.g., for the type $G \otimes G$ there exist 16 such vectors, while for the type $N \otimes N$ we have only one DP vector. The spaces of the type [something] $\otimes$ [–] determine only the spinor K_A in DP vector; the spinor $\bar{K}^{\dot{B}}$ can be here selected completely arbitrarily.

Notice that the concept of the generalized DP vector applies *mutatis mutandi* also in the case of a real V_4 of the signature $(+\,+\,+\,-)$ where the P-spinors of the second type became complex conjugates of the P-spinors of the first type. Thus, with the conformal curvature of the type G one can construct precisely 16 such complex vectors; the 4 real vectors among these 16 objects will be the standard DP vectors. Of course, there will be numerous additional relations among the discussed 16 vectors.

Now, Newman and Penrose call such[6] complex V_4's which have in our notation $\bar{C}^{\dot{A}\dot{B}\dot{C}\dot{D}} = 0$ "right-flat heaven" and if $C_{ABCD} = 0$, correspondingly, "left-flat heaven." Following them (in a sense), but insisting on a more contrasting (perhaps one could say—"manichean") terminology, we propose to call all objects which are $S\bar{L}(2,\mathbb{C})$ scalars and are geometric objects with respect to $SL(2,\mathbb{C})$, the "heavenly objects." Parallelly the objects which with respect to $SL(2,\mathbb{C})$ are scalars and are geometric objects with respect to $S\bar{L}(2,\mathbb{C})$, will be the "hellish objects." The objects whose components mix in an irreducible fashion under $SL(2,\mathbb{C})$ and $S\bar{L}(2,\mathbb{C})$ transformations will be then the "earthly objects." An absolute scalar with respect to $\mathcal{G} = SL(2,\mathbb{C}) \times S\bar{L}(2,\mathbb{C})$—in a degenerate sense because it cannot be in heaven and hell simultaneously—we assign to earth.

Therefore, the objects discussed up to now, can be classified according to Table I.

The space[6] where $\bar{C}^{\dot{A}\dot{B}\dot{C}\dot{D}} = 0$ we will call "weak heaven": heaven"; we permit here for C_{ab} and R being $\neq 0$; a complex space V_4 where there exists such a choice for the null tetrad that

$$\bar{\Gamma}^{\dot{A}\dot{B}} = 0 \tag{1.34}$$

we will call the "strong heaven."[7]

With these "hellish" objects vanishing we have *a forteriori* from (1.21b)

$$\bar{C}^{\dot{A}\dot{B}\dot{C}\dot{D}} = 0, \quad C_{AB\dot{C}\dot{D}} = 0, \quad R = 0. \tag{1.35}$$

Therefore, with $G_{ab} := R_{ab} - \tfrac{1}{2} g_{ab} R$ we conclude that the Einstein vacuum equations (in the complex V_4)

$$G_{ab} = 0 \tag{1.36}$$

are automatically fulfilled in the "strong heaven." Of course, "strong heaven" is nontrivial iff $C_{ABCD} \neq 0$; otherwise V_4 is flat.

2. HEAVENLY TETRAD AND FIRST HEAVENLY EQUATION

Assuming (1.34), it follows from (1.20) that

$$dS^{\dot{A}\dot{B}} = 0. \tag{2.1}$$

Thus (in a simply connected region) there exists such $U^{\dot{A}\dot{B}} \in \Lambda^1$ that

$$S^{\dot{A}\dot{B}} = dU^{\dot{A}\dot{B}}. \tag{2.2}$$

Using the explicit form of $S^{\dot{A}\dot{B}}$ from (1.10), we infer now that

$$2e^4 \wedge e^1 = dU^{\dot{1}\dot{1}}, \quad 2e^3 \wedge e^2 = dU^{\dot{2}\dot{2}}, \tag{2.3}$$

so that $dU^{\dot{1}\dot{1}} \wedge dU^{\dot{1}\dot{1}} = 0 = dU^{\dot{2}\dot{2}} \wedge dU^{\dot{2}\dot{2}}$. Thus, by applying the Darboux theorem in our complex V_4, we deduce the existence of such scalars p, q, r, s that

$$\begin{aligned} &2e^4 \wedge e^1 = 2\, dp \wedge dq = 2\, d(p\, dq + d\tau) \\ &2e^3 \wedge e^2 = 2\, dr \wedge ds = 2\, d(r\, ds + d\sigma) \\ &\quad \to dV := e^1 \wedge e^2 \wedge e^3 \wedge e^4 \\ &\qquad = dp \wedge dq \wedge dr \wedge ds. \end{aligned} \tag{2.4}$$

It follows that (i) $\{x^\mu\} = \{pqrs\}$ can be used as the independent coordinates and (ii) that the heavenly tetrad is given in terms of these coordinates by

$$\begin{aligned} e^1 &= A\, dp + B\, dq, \quad e^3 = E\, dr + F\, ds, \\ -e^4 &= C\, dp + D\, dq, \quad e^2 = G\, dr + H\, ds. \end{aligned} \tag{2.5}$$

Entering with it into (2.4) we infer that the structural functions must fulfill two conditions

$$AD - BC = 1, \tag{2.6a}$$

$$EH - FG = 1. \tag{2.6b}$$

We still must assure that $\overset{\bullet\bullet}{S}{}^{12} = -e^1 \wedge e^2 + e^3 \wedge e^4$ be closed which requires

$$\begin{aligned} d(-e^1 \wedge e^2 + e^3 \wedge e^4) &= -d\{(AG - CE)\,dp \wedge dr + (AH - CF)\,dp \wedge ds \\ &\quad + (BG - DE)\,dq \wedge dr + (BH - DF)\,dq \wedge ds\} = 0. \end{aligned} \tag{2.7a}$$

This condition is equivalent to the equations

$$\begin{aligned} &(AG - CE)_q - (BG - DE)_p = 0, \quad (AG - CE)_s - (AH - CF)_r = 0, \\ &(AH - CF)_q - (BH - DF)_p = 0, \quad (BG - DE)_s - (BH - DF)_r = 0. \end{aligned} \tag{2.7b}$$

Last equations can be readily seen to imply the existence of a function $\Omega = \Omega(pqrs)$—called subsequently the first key function—such that

$$\begin{aligned} &AG - CE = \Omega_{pr}, \quad BG - DE = \Omega_{qr} \\ &AH - CF = \Omega_{ps}, \quad BH - DF = \Omega_{qs}. \end{aligned} \tag{2.7c}$$

We then easily see that

$$\overset{\bullet\bullet}{S}{}^{12} = d(\Omega_p\,dp + \Omega_q\,dq) = -d(\Omega_r\,dr + \Omega_s\,ds) \tag{2.8}$$

so that $d\overset{\bullet\bullet}{S}{}^{12} = 0$ is assured.

Considering now (2.7c) as four equations on $EFGH$ with $ABCD$ assumed known, one easily solves them, using (2.4a) and obtaining

$$\begin{aligned} &E = B\Omega_{pr} - A\Omega_{qr}, \quad G = D\Omega_{pr} - C\Omega_{qr}, \\ &F = B\Omega_{ps} - A\Omega_{qs}, \quad H = D\Omega_{ps} - C\Omega_{qs}. \end{aligned} \tag{2.9}$$

This substituted into (2.6b) yields for the key function Ω the first heavenly equation

$$\begin{vmatrix} \Omega_{pr} & \Omega_{ps} \\ \Omega_{qr} & \Omega_{qs} \end{vmatrix} = 1. \tag{2.10}$$

Of course, ABC and D cancel out here because of (2.6a); quite similarly, if we substitude our tetrad into ds^2 with EFG and H of the form (2.9), we obtain

$$\begin{aligned} V_4 : ds^2 &= 2\Omega_{pr}\,dp\,dr + 2\Omega_{ps}\,dp\,ds \\ &\quad 2\Omega_{qr}\,dq\,dr + 2\Omega_{qs}\,dq\,ds. \end{aligned} \tag{2.11}$$

Therefore, we have demonstrated that the most general "strong heaven" V_4 is determined by one key function Ω which fulfills (2.10) and defines $g_{\mu\nu}$ by (2.11). Observe that with coordinates ordered as $\{x^\mu\} = \{pqrs\}$,

$$(g_{\mu\nu}) = \begin{pmatrix} 0 & 0 & \Omega_{pr} & \Omega_{ps} \\ 0 & 0 & \Omega_{qr} & \Omega_{qs} \\ \Omega_{pr} & \Omega_{qr} & 0 & 0 \\ \Omega_{ps} & \Omega_{qs} & 0 & 0 \end{pmatrix},$$

$$(g^{\mu\nu}) = \begin{pmatrix} 0 & 0 & \Omega_{qs} & -\Omega_{qr} \\ 0 & 0 & -\Omega_{ps} & \Omega_{pr} \\ \Omega_{qs} & -\Omega_{ps} & 0 & 0 \\ -\Omega_{qr} & \Omega_{pr} & 0 & 0 \end{pmatrix} \tag{2.12}$$

and

$$\det(g_{\mu\nu}) = 1. \tag{2.13}$$

For the arbitrary scalar ϕ we have thus

$$\begin{aligned} \phi^{;\alpha}_{;\alpha} &= \frac{1}{\sqrt{g}}\partial_\mu \sqrt{g}\, g^{\mu\nu}\partial_\nu \phi \\ &= \partial_p(\Omega_{qs}\phi_r - \Omega_{qr}\phi_s) - \partial_q(\Omega_{ps}\phi_r - \Omega_{pr}\phi_s) \\ &\quad + \partial_r(\Omega_{sq}\phi_p - \Omega_{sp}\phi_q) - \partial_s(\Omega_{rq}\phi_p - \Omega_{rp}\phi_q). \end{aligned} \tag{2.14}$$

It follows that if we interpret Ω as a scalar, then

$$\Omega^{;\alpha}_{;\alpha} = 4. \tag{2.15}$$

Now, it is clear that ABC and D in the heavenly tetrad—with EFG and H interpreted according to (2.9)—just correspond to the residual freedom of $SL(2,\mathbb{C})$ gauge. Thus, not losing generality but only taking a definite choice for this gauge, we can in particular assume

$$A = 1 = D, \quad B = 0 = C\ . \tag{2.16}$$

This gives for the heavenly tetrad

$$\begin{aligned} &e^1 = dp, e^4 = -dq, \qquad \partial_1 = \partial_p, \partial_4 = -\partial_q, \\ &e^2 = \Omega_{pr}\,dr + \Omega_{ps}\,ds, \leftrightarrow \partial_2 = \Omega_{qs}\partial_r - \Omega_{qr}\partial_s, \\ &-e^3 = \Omega_{qr}\,dr + \Omega_{qs}\,ds, \quad \partial_3 = \Omega_{ps}\partial_r - \Omega_{pr}\partial_s. \end{aligned} \tag{2.17}$$

We can now determine from (1.14) the connection forms Γ_{ab}. The result is that: The "hellish" connections indeed duly vanish:

$$\Gamma_{41} = 0, \quad -\Gamma_{12} + \Gamma_{34} = 0, \quad \Gamma_{32} = 0. \tag{2.18}$$

The "heavenly" connections are given by

$$\begin{aligned} &\alpha := \Gamma_{12} = \Gamma_{34} = -K\,dp + L\,dq, \\ &\beta := \Gamma_{42} = -L\,dp + M\,dq, \\ &\gamma := \Gamma_{31} = -N\,dp + K\,dq \end{aligned} \tag{2.19}$$

where

$$\begin{aligned} K &:= \partial_2\Omega_{pp} = \Omega_{ps}\Omega_{qrp} - \Omega_{pr}\Omega_{qsp} \equiv \Omega_{qs}\Omega_{prp} - \Omega_{qr}\Omega_{psp} \\ &= \Omega_{pr}\Omega_{ps}[\ln(\Omega_{pr}/\Omega_{ps})]_q, \\ L &:= -\partial_3\Omega_{qq} = \Omega_{qr}\Omega_{psq} - \Omega_{qs}\Omega_{prq} \equiv \Omega_{pr}\Omega_{qsq} - \Omega_{ps}\Omega_{qrq} \\ &= \Omega_{qr}\Omega_{qs}[\ln(\Omega_{qs}/\Omega_{qr})]_p, \\ M &:= \partial_2\Omega_{qq} = \Omega_{qs}\Omega_{qrq} - \Omega_{qr}\Omega_{qsq} \\ &= -\Omega_{qr}\Omega_{qs}[\ln(\Omega_{qs}/\Omega_{qr})]_q, \\ N &:= -\partial_3\Omega_{pp} = \Omega_{pr}\Omega_{psp} - \Omega_{ps}\Omega_{prp} \\ &= -\Omega_{pr}\Omega_{ps}[\ln(\Omega_{pr}/\Omega_{ps})]_p. \end{aligned} \tag{2.20}$$

[The identities in (2.20) follow from (2.10); formulas with logarithms apply if one assumes $\Omega_{ps}\Omega_{pr} \neq 0$.]

Of course, (2.18) already assures that the V_4 under

study, being "strong heaven" has $G_{ab}=0$ and $\bar{C}_{\dot{A}\dot{B}\dot{C}\dot{D}}=0$. Now, the curvature quantities C_{ABCD}, or equivalently $C^{(a)}$, $a=1,\ldots,5$ can be computed from (1.21a) which specialized on the present case amounts to

$$\left.\begin{matrix} d\beta+2\beta\wedge\alpha \\ d\alpha+\beta\wedge\gamma \\ d\gamma+2\alpha\wedge\gamma \end{matrix}\right\} = \tfrac{1}{2}e^4\wedge e^2 \left\{\begin{matrix} C^{(5)} \\ C^{(4)}+\tfrac{1}{2}(e^1\wedge e^2+e^3\wedge e^4) \\ C^{(3)} \end{matrix}\right. \left\{\begin{matrix} C^{(4)} \\ C^{(3)}+\tfrac{1}{2}e^3\wedge e^1 \\ C^{(2)} \end{matrix}\right. \left\{\begin{matrix} C^{(3)} \\ C^{(2)} \\ C^{(1)} \end{matrix}\right. . \tag{2.21}$$

Using the symbol dV from (2.4), we have thus

$$\left.\begin{matrix} C^{(5)} \\ C^{(4)} \\ C^{(3)} \end{matrix}\right\} dV = -2e^3\wedge e^1\wedge \left[\left\{\begin{matrix} d\beta+2\beta\wedge\alpha \\ d\alpha+\beta\wedge\gamma \\ d\gamma+2\alpha\wedge\gamma \end{matrix}\right.\right. , \tag{2.22}$$

$$\left.\begin{matrix} C^{(3)} \\ C^{(2)} \\ C^{(1)} \end{matrix}\right\} dV = -2e^4\wedge e^2\wedge \left\{\begin{matrix} d\beta+2\beta\wedge\alpha \\ d\alpha+\beta\wedge\gamma \\ d\gamma+2\alpha\wedge\gamma \end{matrix}\right. .$$

Substituting here (2.19) and (2.17), one easily finds that

$$\begin{aligned} &\tfrac{1}{2}C^{(5)}=\partial_2 M, \quad \tfrac{1}{2}C^{(4)}=\partial_2 L, \\ &\tfrac{1}{2}C^{(3)}=\partial_2 K\equiv-\partial_3 L, \\ &\tfrac{1}{2}C^{(2)}=-\partial_3 K, \quad \tfrac{1}{2}C^{(1)}=-\partial_3 N. \end{aligned} \tag{2.23}$$

3. AN ALTERNATIVE FORM OF THE HEAVENLY TETRAD AND THE SECOND HEAVENLY EQUATION

Although the description of the heavenly tetrad in terms of Ω satisfying (2.10) is symmetric and simple enough, we are able to provide a still simpler—and more convenient for many purposes—alternative description of this tetrad. Let

$$x:=\Omega_p, \quad y:=\Omega_q. \tag{3.1}$$

Then Eq. (2.10) takes the form (forgetting for a moment about coordinates p and q)

$$\frac{\partial(x,y)}{\partial(r,s)}=1, \tag{3.2}$$

so that certainly x and y can be used as new coordinates in the place of r and s. The tetrad (2.17) is now given by

$$\begin{aligned} &e^1=dp, \qquad e^2=dx+A\,dp+B\,dq, \\ &-e^4=dq, \quad -e^3=dy+B\,dp+C\,dq, \end{aligned} \tag{3.3}$$

where in terms of Ω we have

$$A=-\Omega_{pp}, \quad B=-\Omega_{pq}, \quad C=-\Omega_{qq}. \tag{3.4}$$

We will now consider the heavenly tetrad (3.3) *prima facie* understanding A, B, and C as the three structural functions given in coordinates $\{x^\mu\}=\{pqxy\}$, The forms $S^{\dot{A}\dot{B}}$ induced by (3.3) must be all closed. The forms $S^{\dot{1}\dot{1}}=2e^4\wedge e^1=2\,dp\wedge dq$ and $S^{\dot{1}\dot{2}}=-e^1\wedge e^2+e^3\wedge e^4=dx\wedge dp+dy\wedge dq$ are closed independently of the shape of A, B, and C. Thus, with

$$\tfrac{1}{2}S^{\dot{2}\dot{2}}=e^3\wedge e^2=dx\wedge dy+(AC-B^2)\,dp\wedge dq +A\,dp\wedge dy+B(dx\wedge dp+dq\wedge dy)+C\,dx\wedge dq \tag{3.5}$$

we obtain from $dS^{\dot{2}\dot{2}}=0$ the conditions

$$\begin{aligned} &A_x+B_y=0, \\ &B_x+C_y=0; \end{aligned} \tag{3.6a}$$

$$\begin{aligned} &(AC-B^2)_x+B_q-C_p=0, \\ &(AC-B^2)_y-A_q+B_p=0. \end{aligned} \tag{3.6b}$$

One easily shows that (3.6a) implies (and is implied by) the existence of a function $\theta=\theta(pqxy)$ such that

$$A=-\theta_{yy}, \quad B=\theta_{xy}, \quad C=-\theta_{xx}. \tag{3.7}$$

But this substituted into (3.6) gives

$$\left.\begin{matrix} \partial_x \\ \partial_y \end{matrix}\right\}(\theta_{xx}\theta_{yy}-(\theta_{xy})^2+\theta_{xp}+\theta_{yq})=0. \tag{3.8}$$

Consequently, $\theta_{xx}\theta_{yy}-\theta^2{}_{xy}+\theta_{xp}+\theta_{yq}=f_p(p,q)$, with $f(p,q)$ being arbitrary. Therefore, introducing $\Theta:=\theta-xf$ we arrive at the conclusion that with:

$$A=-\Theta_{yy}, \quad B=\Theta_{xy}, \quad C=-\Theta_{xx}, \tag{3.9}$$

and the second key function $\Theta=\Theta(pqxy)$ fulfilling the second heavenly equation

$$\Theta_{xx}\Theta_{yy}-(\Theta_{xy})^2+\Theta_{xp}+\Theta_{yq}=0, \tag{3.10}$$

the forms $S^{\dot{A}\dot{B}}$ are all closed. [Indeed, (3.9) used in (3.5) gives

$$\tfrac{1}{2}S^{\dot{2}\dot{2}}=e^3\wedge e^2=dx\wedge dy+d\Theta_y\wedge dp+dq\wedge d\Theta_x.] \tag{3.11}$$

The heavenly metric given in the terms of the present coordinates and the second key function is thus

$$\begin{aligned} V_4:\ ds^2=&\,2\,dp(dx-\Theta_{yy}\,dp+\Theta_{xy}\,dq) \\ &+2\,dq(dy+\Theta_{xy}\,dp-\Theta_{xx}\,dq). \end{aligned} \tag{3.12}$$

For the natural tetrad we have now

$$\begin{aligned} &e^1=dp, \quad -e^4=dq, && \partial_2=\partial_x, \quad -\partial_3=\partial_y, \\ &e^2=dx-\Theta_{yy}\,dp+\Theta_{xy}\,dq, && \leftrightarrow \quad \partial_1=\partial_p+\Theta_{yy}\partial_x-\Theta_{xy}\partial_y, \\ &-e^3=dy+\Theta_{xy}\,dp-\Theta_{xx}\,dq, && -\partial_4=\partial_q-\Theta_{xy}\partial_x+\Theta_{xx}\partial_y. \end{aligned} \tag{3.13}$$

Now, in the present coordinates we have for the metric tensor and its inverse

$$(g_{\mu\nu})=\begin{pmatrix} -2\Theta_{yy} & 2\Theta_{xy} & 1 & 0 \\ 2\Theta_{xy} & -2\Theta_{xx} & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

$$\det(g_{\mu\nu})=1.$$

$$(g^{\mu\nu})=\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 2\Theta_{yy} & -2\Theta_{xy} \\ 0 & 1 & -2\Theta_{xy} & 2\Theta_{xx} \end{pmatrix} \tag{3.14}$$

Knowing $(g^{\mu\nu})$ we find that, for any scalar ϕ,

$$\phi^{;\alpha}_{;\alpha} = \frac{1}{\sqrt{g}}\partial_\mu \sqrt{g} g^{\mu\nu}\partial_\nu \phi$$
$$= 2[(\phi_{xp} + \phi_{yq}) + \partial_x(\Theta_{yy}\phi_x - \Theta_{xy}\phi_y) - \partial_y(\Theta_{xy}\phi_x - \Theta_{xx}\phi_y)]. \quad (3.15)$$

This, specialized for $\phi = \Theta$ (assuming that Θ is a scalar), yields an important conclusion:

$$\Theta^{;\alpha}_{;\alpha} = 2(\Theta_{xx}\Theta_{yy} - \Theta^2_{xy})$$
$$= -2(\Theta_{xp} + \Theta_{yq}). \quad (3.16)$$

The connections Γ_{ab} worked out from the tetrad (3.13) maintain the general form (2.18)–(2.19); we obtain, however, for the functions KLM and N working with the present coordinates a surprisingly simple result:

$$K = \Theta_{xyy},\ L = \Theta_{xxy},\ M = \Theta_{xxx},\ N = \Theta_{yyy}. \quad (3.17)$$

Similarly, the quantities which characterize the conformal curvature are surprisingly simple:

$$\tfrac{1}{2}C^{(5)} = \Theta_{xxxx},\ \tfrac{1}{2}C^{(4)} = \Theta_{xxxy},$$
$$\tfrac{1}{2}C^{(3)} = \Theta_{xxyy}, \quad (3.18)$$
$$\tfrac{1}{2}C^{(2)} = \Theta_{xyyy},\ \tfrac{1}{2}C^{(1)} = \Theta_{yyyy}.$$

4. SOME EXAMPLES

The purpose of this section is to provide some "living specimens" of (nontrivial) "heavenly metrics," i.e., to give some specific solutions either of (2.10) or (3.10) and to examine their structure.

A simple solution of (3.10) has the form of

$$\Theta = \tfrac{1}{2}\int^x dx \int^x dx\, F(x,q) \quad (4.1)$$

where the function of two variables $F(x,q)$ is arbitrary. Thus, the metric:

$$ds^2 = 2\,dp\,dx + 2\,dq(dy - \tfrac{1}{2}F(x,q)\,dq) \quad (4.2)$$

is heavenly. Working with the natural tetrad (3.13), we have connections such that only $M = \frac{1}{2}F_x$ can be nonzero; while, out of all $C^{(a)}$'s, only $C^{(5)} = F_{xx}(x,q)$ can be nonzero. Therefore, with $F_{xx} \neq 0$, this solution of the (complex) Einstein equations is of the type $[4]\otimes[-]$.

As the second example, with $\Lambda = \Lambda(qr)$ arbitrary, consider the function

$$\Omega = p\cdot\int^r dr' \exp[\Lambda(qr')] + s\cdot\int^q dq' \exp[-\Lambda(q'r)] \quad (4.3)$$

which fulfills (2.10). With $\Omega_{ps} = 0$, from (2.20) we have $K = 0 = N$ and therefore, applying (2.23),

$$C^{(1)} = C^{(2)} = C^{(3)} = 0. \quad (4.4)$$

For L and M one finds that

$$L = -\Lambda_q,$$
$$M = p(\Lambda_{qq} + 2\Lambda_q\Lambda_q) - s\exp(-2\Lambda)\Lambda_{qr}, \quad (4.5)$$

and consequently by applying (2.23)

$$\tfrac{1}{2}C^{(4)} = -\exp(-\Lambda)\Lambda_{qr},$$
$$\tfrac{1}{2}C^{(5)} = [p\exp(-6\Lambda)\partial_q \exp(6\Lambda) - s\exp(-2\Lambda)\partial_r]$$
$$\times \exp(-\Lambda)\Lambda_{qr}. \quad (4.6)$$

Therefore, if $\Lambda_{qr} \neq 0$, the heavenly metric

$$V_4:\ ds^2 = 2e^{\Lambda}\,dp\,dr + 2\exp(-\Lambda)\,dq\,ds$$
$$+ 2(pe^{\Lambda}\Lambda_q - s\exp(-\Lambda)\Lambda_r)\,dr\,dq \quad (4.7)$$

is of the type $[3-1]\times[-]$. (If $\Lambda_{qr} = 0$, then $C^{(a)} = 0$ and the space is flat).

It is of interest to observe that using new coordinates

$$u := pe^{\Lambda},\quad v := s\exp(-\Lambda), \quad (4.8)$$

we can rewrite (4.7) in a simpler form

$$V_4:\ ds^2 = 2\,dr(du - u\Lambda_r(rq)\,dr)$$
$$2\,dq(dv + v\Lambda_q(rq)\,dq), \quad (4.9)$$

which exhibits the fact that the space being considered is a special case (with both distinguished null vectors being gradients) of a double Kerr–Schild metric in the sense of Ref. 1. [Note that (4.2) is a single KS metric.]

We will state here some facts which follow as a special case from the general theory of DKS metrics. Let $\{x^\mu\} = \{uvrq\}$ be independent coordinates. Consider then the DKS metric defined by

$$e^1 = dr,\quad e^2 = du + P\,dr,$$
$$e^3 = dq,\quad e^4 = dv + Q\,dq \quad (4.10)$$

with P and Q being arbitrary structural functions. Then we have

$$\Gamma_{42} = 0,\ \Gamma_{41} = -P_v\,dr,\ \Gamma_{12} = P_u\,dr,\ \Gamma_{34} = Q_v\,dq,$$
$$\Gamma_{31} = -(P_q - QP_v)\,dr + (Q_r - PQ_u)\,dq,\ \Gamma_{32} = Q_u\,dq. \quad (4.11)$$

Using (1.21) one can compute the curvature quantities induced by (4.10). The complete list of these quantities is given below:

$$\bar{C}^{(5)} = -2P_{vv},\ \bar{C}^{(4)} = -P_{uv},\ \bar{C}^{(2)} = -Q_{uv},\ \bar{C}^{(1)} = -2Q_{uu}$$
$$\bar{C}^{(3)} = -\tfrac{1}{3}(P_{uu} + Q_{vv}),\ R = -2(P_{uu} + Q_{vv}), \quad (4.12)$$

$$C^{(5)} = C^{(4)} = 0,\ C^{(3)} = -\tfrac{1}{3}(P_{uu} + Q_{vv}),\ C^{(2)} = P_{uq} - Q_{vr},$$
$$-\tfrac{1}{2}C^{(1)} = (\partial_q + Q_v)(P_q - QP_v) + (\partial_r - P_u)(Q_r - PQ_u), \quad (4.13)$$

$$C_{34} = -C_{12} = \tfrac{1}{2}(P_{uu} - Q_{vv}),\ C_{41} = -P_{uv},\ C_{32} = -Q_{uv},$$
$$C_{44} = C_{22} = C_{42} = 0$$
$$C_{33} = 2(Q_r - PQ_u)_u,\ C_{11} = 2(P_q - QP_v)_v,$$
$$-C_{31} = (P_q - QP_v)_v + (Q_r - PQ_u)_v. \quad (4.14)$$

The general result contained in (4.10) up to (4.14) has various interesting special subcases. In particular, if we require that the tetrad (4.10) describe a "weak heaven," i.e.,

$$\bar{C}^{(a)} = 0,\quad a = 1, \ldots, 5, \quad (4.15)$$

then, in agreement with (4.12), the structural functions take the form

$$P = \tfrac{1}{2}au^2 + b_1u + c_1 + f_1v, \quad (4.16)$$
$$Q = -\tfrac{1}{2}av^2 - b_2v - c_2 - f_2u,$$

where $a(qr), \ldots, f_2(qr)$ are all arbitrary functions of q and r only. Besides (4.15) these P and Q imply $R = 0$ and

$$C^{(5)} = C^{(4)} = C^{(3)} = 0,\ C^{(2)} = ua_q + va_r + b_{1q} + b_{2r}, \quad (4.17)$$
$$C^{(1)} = \text{in the form from (4.13)},$$

so that the conformal curvature is of the type $[3-1]$

$\otimes[-]$, while the nontrivial components of C_{ab} are

$$C_{34}=-C_{12}=a,\quad -C_{31}=ua_q+va_r+b_{1q}-b_{2r}+2f_1f_2,$$
$$C_{33}=-2[f_{2r}-(au+b_1)f_2],\quad C_{11}=2[f_{1q}+(av+b_2)f_1]. \tag{4.18}$$

Suppose now that we require additionally that $C_{33}=C_{31}=C_{11}=0$ permitting only for $C_{34}=-C_{12}=a=\text{const}$, to be $\neq 0$. If in fact $a\neq 0$, then the structural functions amount to

$$P=\tfrac{1}{2}au^2-u\Lambda_r(qr)+c_1(qr),$$
$$Q=-\tfrac{1}{2}av^2+v\Lambda_q(qr)-c_2(qr) \tag{4.19}$$

where $a\neq 0$ and $\Lambda(qr)$, $c_1(qr)$, $c_2(qr)$ are arbitrary. The only nontrivial curvature quantities are here given by

$$C_{34}=-C_{12}=a,\quad \tfrac{1}{2}C^{(2)}=-\Lambda_{qr},$$
$$\tfrac{1}{2}C^{(1)}=(u\,e^{-\Lambda}\partial_q e^{\Lambda}-ve^{\Lambda}\partial_r e^{-\Lambda}-2auv)\Lambda_{qr}$$
$$+a(uc_{2r}+vc_{1q})+c_{2rr}-c_{1qq}-\Lambda_r c_{2r}-\Lambda_q c_{1q}. \tag{4.20}$$

In the limit $a\to 0$, $c_1\to 0\to c_2$ and the space in question coincides with (4.9); the curvature quantities

$$\tfrac{1}{2}C^{(2)}=-\Lambda_{qr},$$
$$\tfrac{1}{2}C^{(1)}=(ue^{-\Lambda})\,\partial_q e^{\Lambda}-ve^{\Lambda}\partial_r e^{-\Lambda})\,\Lambda_{qr} \tag{4.21}$$

are considerably more symmetric than the equivalent set in Eq. (4.6).

If one requires in (4.18) more strongly $C_{ab}=0$, then the space becomes a strong heaven. The corresponding structural functions are now given by

$$P=-u\left(\Lambda_r+\frac{\dot F}{F-G}\right)+c_1+\frac{\dot F}{F-G}e^{\Lambda}\cdot v,$$
$$Q=v\left(\Lambda_q+\frac{\dot G}{F-G}\right)-c_2-\frac{\dot G}{F-G}e^{-\Lambda}\cdot u \tag{4.22}$$

where $F=F(r)$, $G=G(q)$, $\Lambda=\Lambda(qr)$, $c_1=c_1(qr)$, and $c_2(qr)$ are arbitrary. The curvature is characterized here by

$$\tfrac{1}{2}C^{(2)}=-\Lambda_{qr} \tag{4.23}$$

and some $C^{(1)}$; even when $\Lambda_{qr}\to 0$, the functions c_1 and c_2 allow that the space can reduce to something nontrivial of the type $N\otimes[-]$. Observe that although the tetrad (4.10) with P and Q from (4.22) produces some "hellish" connections $\bar\Gamma^{\cdot\cdot}_{A\dot B\dot C}$, there exists an $S\bar L(2,\mathbb{C})$ gauge in which these objects do vanish.

Now, as the next simple example we consider the function

$$\Theta=\frac{\beta}{2\alpha(\alpha-1)}x^{\alpha}y^{1-\alpha}, \tag{4.24}$$

so that

$$\Theta_{xx}=\frac{\beta}{2}x^{\alpha-2}y^{1-\alpha},$$

$$\Theta_{xy}=-\frac{\beta}{2}x^{\alpha-1}y^{-\alpha},\quad \Theta_{yy}=\frac{\beta}{2}x^{\alpha}y^{-\alpha-1}, \tag{4.24}$$

which fulfills (3.10) for all values of constants α and β. From (3.18) it follows that we have

$$C^{(a)}=\partial_x\partial_y(-1)^{a+1}x^{\alpha+2-a}y^{-\alpha-3+a}. \tag{4.25}$$

On the other hand, if $K^2\neq 0$, then the equation for the P-spinor K^A, $C_{ABCD}K^AK^BK^CK^D=0$ amounts to

$$\sum_{a=1}^{5}\binom{4}{a-1}C^{(a)}(K^1/K^2)^{5-a}=0. \tag{4.26}$$

Substituting here from (4.25) we find that this equation is equivalent to

$$\partial_x\partial_y(xK^1-yK^2)^4x^{\alpha-3}y^{-\alpha-2}=0, \tag{4.27}$$

or explicitly

$$(xK^1-yK^2)^2\{(\alpha+2)(\alpha+1)(xK^1)^2+(\alpha-2)(\alpha-3)(yK^2)^2$$
$$-2(\alpha+1)(\alpha-2)(xK^1)(yK^2)\}=0. \tag{4.28}$$

It follows that, irrespective of the value of α, the spinor $K^A=(y,x)$ is always (at least) a double P-spinor, and that the factor with quadratic form in (4.27) has vanishing discriminant if and only if $\alpha=-1$ or $\alpha=2$. Therefore, if $\alpha=-1$, 2 the solution is of the type $[2-2]\otimes[-]$ and with $\alpha\neq -1$, 2 it is of the type $[2-1-1]\times[-]$.

Therefore, we have succeeded in this section of producing explicitly examples of heavenly metric of all possible algebraically degenerate types.

Of course, this work and its results are to be considered as a technical step forward within the general goal toward which relativity is striving in recent years, i.e., to produce general techniques which would lead to general solutions of the Einstein equations on a real manifold. The complex solutions and complex geometry, although attracting interest at this moment (see, e.g., Refs. 8–10, ought to be considered only as an intermediate step.

ACKNOWLEDGMENT

I would like to express my gratitude to Dr. J.D. Finley for illuminating discussions and his active participation in deriving the results outlined in Sec. 2. A discussion with Dr. E.T. Newman and Dr. R. Penrose during the "Riddle of Gravity" symposium at Syracuse is also appreciated.

*On leave of absence from University of Warsaw, Warsaw, Poland.

[1]J. Plebañski, Ann. Phys. 90, 196 (1975).
[2]J. Plebañski, and M. Demiański, preprint, November 1973.
[3]The null tetrad formalism used throughout this text, follows notation and convention of the real null tetrad formalism used

in G. Debney, R. Kerr, and A. Schild, J. Math. Phys. 10, 1842 (1969), abandoning, however, of course the assumption that $(e^2)^* = e^1$ and all its implications.

[4]R. Penrose, Ann. Phys. (N.Y.) 10, 171 (1960).

[5]The ordered symbol, however, describes the type of V_4 as referred to the specific orientation of the tetrad considered given, modulo $\mathcal{G}$ transformation only. Any *improper* tetrad transformation $e^{a'} = l^{a'}{}_b e^b$, with $\det(l^{a'}{}_b) = -1$, maps type $[A]\otimes[B]$ into $[B]\otimes[A]$; for that reason, if one thinks about the type of V_4 irrespective of the orientation for the tetrad, we have to interpret $[A]\otimes[B]$ rather as a commutative tensor product.

[6]That is, an *oriented* complex V_4 as per Ref. 5.

[7]Of course, the improper tetrad transofrmation, e.g., $e^1 \to e^2$, $e^2 \to e^1$, $e^3 \to e^3$, $e^4 \to e^4$, would change "strong heaven" into "strong hell" with $\Gamma_{AB} = 0$.

[8]E.T. Newman and J. Winicour, J. Math. Phys. 15, 426 (1974).

[9]R.W. Lind and E.T. Newman, J. Math. Phys. 15, 1103 (1974).

[10]E.T. Newman and J. Winicour, J. Math. Phys. 15, 113 (1974).

Null geodesic surfaces and Goldberg–Sachs theorem in complex Riemannian spaces

J. F. Plebański*

Centro de Investigación y de Estudios Avanzados del IPN, México 14, D.F., Mexico

S. Hacyan

Instituto de Astronomía, Universidad Nacional Autónoma de México, México 20, D.F., Mexico
(Received 29 April 1975)

An extension of the Goldberg-Sachs theorem for the case of a complex V_4 is given with a simple proof. The interpretation of the theorem, however, no longer applies the concept of the geodesic and shearless congruence of null directions; instead, the existence of a geodesic 2-surface (complex), the tangent vectorial space to which (i) contains only null vectors, (ii) is parallelly propagated along the surface, is now essential.

1. INTRODUCTION

This paper follows Ref. 1 where the formalism of the null tetrad (essentially in the same version as that used in Ref. 2), together with its equivalent spinorial form, was adopted for the purposes of complex Riemannian geometry in four dimensions. Thus, we follow here the notation and terminology of Ref. 1. In particular, with $\mathcal{G} = \mathrm{SL}(2,\mathbb{C}) \times \mathrm{S\bar{L}}(2,\mathbb{C})$ being the tetradial gauge group, the objects which are scalars with respect to the second factor in $\mathcal{G}$ (i.e., can possess undotted spinorial indices only) will be occasionally called "heavenly"; while, correspondingly, the objects which are scalars with respect to the first factor in $\mathcal{G}$ (i.e., can be endowed with dotted spinorial indices only) will be called "hellish."

The fundamental "undotted" objects are S^{AB} (the base of self-dual 2-forms), Γ_{AB} (the connection 1-forms Γ_{42}, $\Gamma_{12}+\Gamma_{34}$, Γ_{31}), and C_{ABCD} (the spinorial image of the conformal curvature—in the notation of Ref. 2 these are equivalent to the $C^{(a)}$, $a=1,\ldots,5$). The corresponding "dotted" objects are: $S^{\dot{A}\dot{B}}$ (the base of anti-self-dual 2-forms), $\Gamma_{\dot{A}\dot{B}}$ (the connection 1-forms Γ_{41}, $-\Gamma_{12}+\Gamma_{34}$, Γ_{32}), and $\bar{C}_{\dot{A}\dot{B}\dot{C}\dot{D}}$ (in the notation of Ref. 2 these are equivalent to the $\bar{C}^{(a)}$, $a=1,\ldots,5$). In distinction to the case of a real V_4 of signature $(+++-)$, in a general complex V_4 the fundamental "dotted" objects are not complex conjugates of the corresponding "undotted" objects and, thus, possess an "independent existence" and transformation properties. In particular, it follows—as was noticed in Ref. 1.—that the generalized Petrov–Penrose algebraic classification of the conformal curvature of a complex V_4 is just the cartesian product of two Penrose diagrams[3] which correspond to the spinorial structures of the independent objects C_{ABCD} and $C_{\dot{A}\dot{B}\dot{C}\dot{D}}$.

Now, it is well known in the theory of real V_4's of the signature $(+++-)$ that there exists an interesting relationship between the algebraic degeneration of the conformal curvature and the existence of a geodesic and shearless congruence of (real) null directions. This relationship, known under the name of the Goldberg–Sachs theorem,[4] was found to play an instrumental role in most of the successful developments in the theory of exact solutions recorded in the last decade.

The objective of the present article is to extend the Goldberg–Sachs theorem on the level of complex V_4's with their specific typology of algebraical structures. It turns out that, quite similarly to the case of a real V_4, when working with a complex V_4, by applying the Cartan structure formulas and Bianchi identities, we can demonstrate the existence of a complex extension of the Goldberg–Sachs theorem. From the point of view of the interpretation, we must, however, here deviate significantly from the intuitions formed from the case of a real V_4: The role of the one-dimensional extremal null variety—a null and shearless geodesic—is taken, in the complex case, by the two-dimensional counterpart of the null geodesic.

2. THE COMPLEX VERSION OF GS THEOREM

Consider a complex V_4 given in terms of a null tetrad e^a:

$$V_4 : ds^2 = 2e^1e^2 + 2e^3e^4 \in \Lambda^1 \otimes \Lambda^1. \tag{2.1}$$

Then, with connections $\Gamma_{ab} = \Gamma_{[ab]} \in \Lambda^1$ defined by $de^a = e^b \wedge \Gamma^a{}_b$, the Cartan structure equations separate into the "undotted" and "dotted" sets of equations. The "undotted" set is given by

$$\begin{aligned}\mathcal{A} :&= d\Gamma_{42} + \Gamma_{42} \wedge (\Gamma_{12} + \Gamma_{34}) \\ &= \tfrac{1}{2}C^{(5)} e^4 \wedge e^2 + \tfrac{1}{2}C^{(4)}(e^1 \wedge e^2 + e^3 \wedge e^4) \\ &\quad + [\tfrac{1}{2}C^{(3)} - R/12]e^3 \wedge e^1 - \tfrac{1}{2}C_{44}\, e^4 \wedge e^1 \\ &\quad - C_{42}(-e^1 \wedge e^2 + e^3 \wedge e^4) - \tfrac{1}{2}C_{22}\, e^3 \wedge e^2\end{aligned} \tag{2.2a}$$

$$\begin{aligned}\mathcal{B} :&= d(\Gamma_{12} + \Gamma_{34}) + 2\Gamma_{42} \wedge \Gamma_{31} \\ &= C^{(4)} e^4 \wedge e^2 + [C^{(3)} + R/12](e^1 \wedge e^2 + e^3 \wedge e^4) \\ &\quad + C^{(2)} e^3 \wedge e^1 - C_{41}\, e^4 \wedge e^1 - C_{12}(-e^1 \wedge e^2 + e^3 \wedge e^4) \\ &\quad + C_{32}\, e^3 \wedge e^2\end{aligned} \tag{2.2b}$$

$$\begin{aligned}\mathcal{C} :&= d\Gamma_{31} + (\Gamma_{12} + \Gamma_{34}) \wedge \Gamma_{31} \\ &= [\tfrac{1}{2}C^{(3)} - R/12]e^4 \wedge e^2 + \tfrac{1}{2}C^{(2)}(e^1 \wedge e^2 + e^3 \wedge e^4) \\ &\quad + C^{(1)} e^3 \wedge e^1 - \tfrac{1}{2}C_{11}\, e^4 \wedge e^1 + \tfrac{1}{2}C_{31}(-e^1 \wedge e^2 + e^3 \wedge e^4) \\ &\quad - \tfrac{1}{2}C_{33}\, e^3 \wedge e^2.\end{aligned} \tag{2.2c}$$

These relations involve only "undotted" Γ's and

$C^{(a)}$'s. (Notations used here follows conventions of Ref. 1; e.g., C_{ab} are the null tetrad components of the Ricci tensor with extracted trace, R is the scalar curvature.) The "dotted" set of structure equations arises formally from (2.2) by the following interchange of indices:

$$\begin{matrix}1\to 2\\ 2\to 1\\ 3\to 3\\ 4\to 4\end{matrix} \quad \text{and} \quad C^{(a)}\to \bar{C}^{(a)}. \tag{2.3}$$

{Of course, with real V_4 of signature $(+++-)$, where $e^2=(e^1)^*$, the operation in Eq. (2.3) is just the complex conjugation; in a complex V_4, in general, $\bar{C}^{(a)}\neq[C^{(a)}]^*$.} We can also notice that Eqs. (2.2) are formally invariant with respect to the interchanges:

$$\begin{matrix}1\to 2 & & C^{(5)}\to +C^{(1)}\\ 2\to 1 & & C^{(4)}\to -C^{(2)}\\ 3\to 4 & \text{and} & C^{(3)}\to +C^{(3)}.\\ 4\to 3 & & C^{(2)}\to -C^{(4)}\\ & & C^{(1)}\to +C^{(5)}\end{matrix} \tag{2.4}$$

Lemma A[5]: Suppose now that

$$C_{44}=C_{42}=C_{22}=0, \quad \Gamma_{422}=\Gamma_{424}=0. \tag{2.5}$$

Then, by reading off the $e^4\wedge e^2$ component of (2.2a), one first infers that

$$C^{(5)}=0. \tag{2.6}$$

Now, by applying the operator d on Eqs. (2.2) we obtain the Bianchi identities, which [assuming Eq. (2.5)], as far as $C^{(4)}$ is concerned, state that

$$[-\partial_4+(\Gamma_{124}+\Gamma_{344})+4\Gamma_{421}]C^{(4)}=0, \tag{2.7a}$$

$$[-\partial_2+(\Gamma_{122}+\Gamma_{342})-4\Gamma_{423}]C^{(4)}=0. \tag{2.7b}$$

This we can now use in the general commutator

$$(\partial_4\partial_2-\partial_2\partial_4)F=(\Gamma^s_{24}-\Gamma^s_{42})\,\partial_s F, \tag{2.8}$$

specialized for $F=\ln[C^{(4)}]$, obtaining the integrability condition of the two equations (2.7a, b):

$$\begin{aligned}&\partial_4(\Gamma_{124}+\Gamma_{344}-4\Gamma_{423})-\partial_2(\Gamma_{124}+\Gamma_{344}+4\Gamma_{421})\\&-(\Gamma_{122}+\Gamma_{342}-4\Gamma_{423})(\Gamma_{124}+\Gamma_{412})\\&-(\Gamma_{124}+\Gamma_{344}+4\Gamma_{421})(\Gamma_{324}-\Gamma_{342})=0.\end{aligned} \tag{2.9}$$

On the other hand, we can directly read off from Eq. (2.2) that

$$\begin{aligned}\tfrac{1}{2}C^{(4)}&=-\partial_2\Gamma_{421}+\Gamma_{423}(\Gamma_{412}-\Gamma_{421})+\Gamma_{421}\Gamma_{342},\\ \tfrac{1}{2}C^{(4)}&=-\partial_4\Gamma_{423}+\Gamma_{421}(\Gamma_{423}-\Gamma_{324})+\Gamma_{424}\Gamma_{124},\\ C^{(4)}&=-\partial_4(\Gamma_{122}+\Gamma_{342})-\partial_2(\Gamma_{124}+\Gamma_{344})+\\&\quad-(\Gamma_{122}+\Gamma_{342})(\Gamma_{124}+\Gamma_{412})\\&\quad-(\Gamma_{124}+\Gamma_{344})(\Gamma_{324}-\Gamma_{342}).\end{aligned} \tag{2.10}$$

Multiplying the first two of these relations by four, and adding to the last we obtain

$$5C^{(4)}=\text{left-hand member of (2.9)}]=0, \tag{2.11}$$

which contradicts our silent assumption that $C^{(4)}\neq 0$, which was made when specializing (2.8) for $F=\ln[C^{(4)}]$. Therefore,

$$C^{(4)}=0. \tag{2.12}$$

Thus, we have demonstrated the implication

$$(C_{44}=C_{42}=C_{22}=0) \text{ and } (\Gamma_{422}=0=\Gamma_{424})\to[C^{(5)}=0=C^{(4)}],$$

which is the content of our Lemma A.

Lemma B[5]: Suppose now that

$$C_{ab}=0 \text{ and } C^{(5)}=0=C^{(4)}. \tag{2.14}$$

Then, from the special Bianchi identities, $R=\text{const}=-4\lambda$; i.e., although we assume that the Ricci tensor with extracted trace is vanishing, the cosmological constant may be present. With (2.14) assumed, we can work out from (2.2) that, as the consequence of the Bianchi identities,

$$\begin{aligned}&(a_1)\quad -3\Gamma_{424}C^{(3)}=0,\\&(a_2)\quad [\partial_2+2(\Gamma_{122}+\Gamma_{342})+\Gamma_{423}]C^{(1)}\\&\qquad+[\partial_3+(\Gamma_{123}+\Gamma_{343})-4\Gamma_{312}]C^{(2)}-3\Gamma_{313}C^{(3)}=0;\\&(b_1)\quad -3\Gamma_{422}C^{(3)}=0,\\&(b_2)\quad [-\partial_4+2(\Gamma_{124}+\Gamma_{344})+\Gamma_{421}]C^{(1)}\\&\qquad+[\partial_1+(\Gamma_{121}+\Gamma_{341})+4\Gamma_{314}]C^{(2)}-3\Gamma_{311}C^{(3)}=0;\\&(c_1)\quad [-\partial_4+3\Gamma_{421}]C^{(3)}-2\Gamma_{424}C^{(2)}=0,\\&(c_2)\quad -\Gamma_{422}C^{(1)}+[-\partial_2-(\Gamma_{122}+\Gamma_{342})-2\Gamma_{423}]C^{(2)}\\&\qquad+[-\partial_3+3\Gamma_{312}]C^{(3)}=0;\\&(d_1)\quad [-\partial_2-3\Gamma_{423}]C^{(3)}-2\Gamma_{422}C^{(2)}=0,\\&(d_2)\quad \Gamma_{424}C^{(1)}+[\partial_4+(\Gamma_{124}+\Gamma_{344})-2\Gamma_{421}]C^{(2)}\\&\qquad+[\partial_1-3\Gamma_{314}]C^{(3)}=0.\end{aligned} \tag{2.15}$$

This established, we immediately infer that if only one of the three quantities $C^{(3)}$, $C^{(2)}$, and $C^{(1)}$ is different from zero, then necessarily

$$\Gamma_{424}=0=\Gamma_{422}. \tag{2.16}$$

If, along with $C_{ab}=0$, we also have $C^{(a)}=0$ and $\lambda=0$, then, in the terminology of Ref. 1, V_4 represents a "strong hell" where, by using the freedom of $SL(2,\mathbb{C})$ gauge, we can so select the tetrad that $\Gamma_{AB}=0$ and, in particular, $\Gamma_{42}=0$, so that, consequently, (2.16) applies *a forteriori*.

Therefore, we can now state our Lemma B in the form of an implication:

$$(C_{ab}=0) \text{ and } (C^{(5)}=0=C^{(4)})\to(\Gamma_{422}=0=\Gamma_{424}), \tag{2.17}$$

understanding, in the subcase $C^{(a)}=0$ and $\lambda=0$, the arrow as the possibility of such a choice for the null tetrad that the thesis applies.[6]

Taking now the intersection of our Lemmas A and B, we derive the Goldberg–Sachs theorem as extended on the complex Riemannian space:

Theorem: The existence of a choice for the null tetrad such that $\Gamma_{424}=0=\Gamma_{422}$ is the necessary and sufficient condition for the complex V_4, which fulfills the (complex) Einstein empty space equations $G_{ab}=0$, to have the "undotted" spinorial image of the conformal curvature

tensor algebraically degenerate, with $C^{(5)}=0=C^{(4)}$, and with e^3 defining some (at least) double Debever–Penrose null direction.[7]

Now, an important comment must be made: Our complex extension of the Goldberg–Sachs theorem involves Ricci rotation coefficients Γ_{424} and Γ_{422}, but does not concern Γ_{414} and Γ_{411}, which remain in general independent.

In the case when we take a real cross section of the V_4 under study of the signature $(+++-)$, by postulating additionally $(e^2)^*=e^1$, we have then $(\Gamma_{424}=0=\Gamma_{422})$ $-(\Gamma_{414}=0=\Gamma_{411})$ and, as is well known, conditions $\Gamma_{424}=0=\Gamma_{411}$ have the interpretation of securing that the discussed geodesic congruence has vanishing shear. Thus, in that case, the theorem from the present article reduces to the standard version of the GS theorem for the real geometry: The existence of a geodesic and shearfree congruence of null directions is the necessary and sufficient condition for the conformal curvature of V_4 to be degenerate, when $G_{ab}=\lambda g_{ab}$ is assumed.

But geodesic directions in the sense of Ref. 1 are "terestrial" objects: if $\pounds_v v_\mu=\lambda v_\mu$ is maintained as the definition of a geodesic vector in the complex geometry, then, with null $e^3=v_\mu\,dx^\mu$ selected as one of the members of the null tetrad, the conditions that this direction is a geodesic, $\Gamma_{424}=0=\Gamma_{414}$, involve both the "heavenly" and "hellish" connections. On the other hand, it is obvious that our complex GS theorem (with the assumption of $C_{ab}=0$) involves only the "undotted" quantities

$$[C^{(5)}=0=C^{(4)}]\longleftrightarrow[\Gamma_{422}=0=\Gamma_{424}];\tag{2.18}$$

and has a purely "dotted" version obtained by the application of the transformation (2.3):

$$[\bar{C}^{(5)}=0=\bar{C}^{(4)}]\longleftrightarrow[\Gamma_{411}=0=\Gamma_{414}].\tag{2.19}$$

Still two other variants can be obtained by applying the permutation of indices as in (2.4), and with $C_{ab}=0$ all the time assumed:

$$[C^{(1)}=0=C^{(2)}]\longleftrightarrow[\Gamma_{311}=0=\Gamma_{313}],\tag{2.20}$$

and

$$[\bar{C}^{(1)}=0=\bar{C}^{(2)}]\longleftrightarrow[\Gamma_{322}=0=\Gamma_{323}].\tag{2.21}$$

The same remark applies with respect to our Lemmas A and B: They clearly possess two "undotted" and two "dotted" versions.

These comments suggest that, while looking for the geometric interpretation of the conditions $\Gamma_{424}=0=\Gamma_{422}$ (with Γ_{414} and Γ_{411} left as independent quantities!), it is advisable to consider geometric objects which, from the point of view of the tetradial gauge group, are "heavenly." A natural candidate of this type is the 2-form of the area element of a 2-surface, endowed with the self-duality property. This is the basic idea which has lead to the considerations of the subsequent section.

3. NULL GEODESIC SURFACES

A 2-form which represents an element of area of a 2-surface must be necessarily simple: Thus, the two likely candidates for the "heavenly" self-dual area elements of 2-surfaces are

$$S^{22}=2e^3\wedge e^1,\quad S^{11}=2e^4\wedge e^2.\tag{3.1}$$

Now, quite generally, if in a V_n there is a simple p-form $(n>p)$,

$$\Lambda^p\ni\omega=\Theta_1\wedge\cdots\wedge\Theta_p\neq 0\tag{3.2}$$

(the forms Θ_i are defined for all points of V_n), then ω can be interpreted as the area element of a p-surface, V_p if, with $\Theta_i=\Theta_{i\mu}\,dx^\mu$ and in the local coordinates $\{x^\mu\}$, the equations

$$\frac{\partial x^\mu}{\partial\tau^i}=C_i^j\Theta_j^\mu(x)\tag{3.3}$$

are integrable for $x^\mu=x^\mu(\tau^1\cdots\tau^p)$. [The $p\times p$ matrix (C_j^i) with $\det(C_j^i)=1$ is here arbitrary.] Now, if we define the tangent vectorial space $\Upsilon\subset\Lambda^1$ by the condition $v\in\Upsilon\longleftrightarrow v\wedge\omega=0$, it is well known that the integrability conditions of (3.3) amount to the statement that the set of forms Θ_i is closed under the Lie bracket, i.e., that (see, e.g., Ref. 8)

$$\pounds_{\Theta_i}\Theta_j-\pounds_{\Theta_j}\Theta_i\in\Upsilon.\tag{3.4}$$

Thus, the form S^{22} is an integrable element of surface if

$$e^{1\nu}e^3{}_{\mu;\nu}-e^{3\nu}e^1{}_{\mu;\nu}=\alpha e^1{}_\mu+\beta e^3{}_\mu.\tag{3.5}$$

Since $\Gamma^a{}_{bc}:=-e^a{}_{\mu;\nu}e_b{}^\mu e_c{}^\nu$, one easily finds that this condition is fulfilled if and only if

$$\Gamma_{424}=0=\Gamma_{422},\tag{3.6}$$

i.e., the conditions that S^{22} is an integrable element of surface coincide precisely with the conditions which appear in the thesis of our GS theorem, extended on the case of the complex V_4.

(It can be observed that $S^{11}=2e^4\wedge e^2$ represents an integrable surface element if

$$\Gamma_{313}=0=\Gamma_{311};\tag{3.7}$$

similarly, the "hellish" objects $S^{\dot2\dot2}=2e^3\wedge e^2$ and $S^{\dot1\dot1}=2e^4\wedge e^1$ are integrable 2-surface elements if, in the first case,

$$\Gamma_{414}=0=\Gamma_{411},\tag{3.8}$$

and, in the second case,

$$\Gamma_{323}=0=\Gamma_{322}.)\tag{3.9}$$

Let now $u,v\in\Upsilon$, where $\Lambda^1\supset\Upsilon\ni u\longleftrightarrow u\wedge e^3\wedge e^1=0$, and consider the vector $\pounds_u v$; the condition that this vector is in Υ, i.e.,

$$u,v\in\Upsilon\rightarrow\pounds_u v\in\Upsilon,\tag{3.10}$$

stated explictly, amounts to the following: for every α, β, γ, and δ,

$$(\alpha e^1{}_\nu+\beta e^3{}_\nu)(\gamma e^1{}_\mu+\delta e^3{}_\mu)_{;\nu}\,dx^\mu\wedge e^3\wedge e^1=0,\tag{3.11}$$

or, equivalently,

$$(\alpha\delta\Gamma_{422}+\beta\delta\Gamma_{424})\,e^1\wedge e^2\wedge e^3$$
$$+(\alpha\gamma\Gamma_{422}+\beta\gamma\Gamma_{424})\,e^1\wedge e^3\wedge e^4=0.\tag{3.12}$$

This, of course, again holds if (3.6) is fulfilled.

Summing up, we find that the conditions $\Gamma_{424}=0=\Gamma_{422}$ are necessary and sufficient for the two null congruences e^3 and e^1 to be interpreted as spanning the tangent space Υ to a family of (integrable) 2-surfaces, in such a manner that Υ is parallelly propagated along all these surfaces. Of course, each vector $u\in\Upsilon$ is null, $(u,u)=0$.

Comparing this with the situation to which we are accustomed when working with null geodesics, we immediately see that a perfect analogy applies: In both cases, the two-and, respectively, one-dimensional varieties possess null tangent spaces which are parallelly propagated along the variety. Therefore, the 2-surfaces with the element of area $e^3\wedge e^1$, whose existence is assured if $\Gamma_{424}=0=\Gamma_{422}$, generalize the concept of null geodesic on the level of the two-dimensional varieties. We will call them null geodesic surfaces.

It should be observed, however, that one-dimensional geodesics exhibit the additional property of being an extremal variety. Thus, it remains to examine whether the corresponding generalization also applies in the two-dimensional case. Now, the idea of a two-dimensional variety which generalizes the one-dimensional extremal (geodesic) is well known in differential geometry and amounts to an abstract treatment of the 2-surface, bounded by a given curve, which has minimal area, i.e., of the two-dimensional membrane. Extremal surfaces were already studied by Bompiani[9] as early as 1919; Rayski[10] proposed to apply three-dimensional spacelike surfaces of extremal volume as a possibly useful tool in general relativity.

For our purposes it is useful to outline briefly the general concept of the extremal p-dimensional variety V_p in a Riemannian space V_n ($n>p$; in general V_n can be complex). Following Ref. 9, let, in the local coordinates $\{x^\mu\}$ ($\mu=1,\ldots,n$), the subspace V_p be given as parametrized by $\tau^i=(\tau^1,\ldots,\tau^p)$:

$$V_p: x^\mu=x^\mu(\tau^i). \tag{3.13}$$

Then the induced metric on V_p is

$$a_{ij}:=g_{\mu\nu}\frac{\partial x^\mu}{\partial\tau^i}\frac{\partial x^\nu}{\partial\tau^j},\quad a:=\det(a_{ij}). \tag{3.14}$$

If now V_{p-1} is a given closed subspace bounding a region $R_p\subset V_p$, then the "volume" of R_p is

$$V=\int_{R_p}dV,\quad dV:=\sqrt{a}\,d\tau^1\cdots d\tau^p. \tag{3.15}$$

Of course, V as a functional of x^μ (τ^i) is independent of the choice of parametrization.

If the induced metric is nonsingular, $a\neq 0$, and (a^{ij}) exists, then the Euler–Lagrange equations which follow from $\delta V=0$ amount to[11]

$$a^{ij}\frac{D^2x^\mu}{\partial\tau^i\partial\tau^j}=b^i\frac{\partial x^\mu}{\partial\tau^i}, \tag{3.16}$$

where the arbitrary b^i's are related to the ambiguity of choice for the parametrization; by selecting a preferred (affine) parameterization one can make $b^i=0$. (Of course, the covariant derivative is used here in the sense of V_n and it employs $\left\{{\alpha\atop\beta\gamma}\right\}$ as connections.)

On the other hand, if the tangent space T to V_p which is spanned by the vectors $\partial x^\mu/\partial\tau^i$ is parallelly propagated along V_p, i.e., if there exist such C^k_{ij} that

$$\frac{D^2x^\mu}{\partial\tau^i\partial\tau^j}=C^k_{ij}\frac{\partial x^\mu}{\partial\tau^k}, \tag{3.17}$$

then Eqs. (3.16) are fulfilled *a forteriori* (with some specific b^i's) so that V_p, endowed with the property (3.17), atuomatically has extremal volume.

Although the geodesic null surfaces generated by $e^3\wedge e^1$ (when $\Gamma_{424}=0=\Gamma_{422}$ is assumed) possess a tangent space which is parallelly propagated along these surfaces, we cannot directly apply the construction outlined above claiming that the geodesic null surfaces are extremal. This is so because in the present case $a_{ij}=0\to a=0$ (e^1 and e^3 are null and orthogonal) so that $V\equiv 0$, while (a^{ij}) does not exist and the Euler–Lagrange equations in the form (3.16) lose sense. [It can be observed that essentially the same difficulty is already encountered in the case of the standard one-dimensional geodesic: The action $\int(g_{\mu\nu}\,dx^\mu\,dx^\nu)^{1/2}\equiv 0$ cannot be directly used in order to deduce the equations of null geodesic.]

However, quite generally, if in (3.14) $a_{ij}=0$, and V from (3.15) cannot be used as the action functional, we can conveniently use as the action

$$A:=\frac{1}{2}\int\Lambda^{ij}g_{\mu\nu}\frac{\partial x^\mu}{\partial\tau^i}\frac{\partial x^\nu}{\partial\tau^j}\,d\tau^1\wedge\cdots\wedge d\tau^p, \tag{3.18}$$

with $\lambda^{ij}=\lambda^{ji}$ being Lagrange multipliers. This action yields as the equations of the null extremal variety V_p:

$$g_{\mu\nu}\frac{\partial x^\mu}{\partial\tau^i}\frac{\partial x^\nu}{\partial\tau^j}=0, \tag{3.19a}$$

$$\lambda^{ij}\frac{D^2x^\mu}{\partial\tau^i\partial\tau^j}=b^i\frac{\partial x^\mu}{\partial\tau^i}, \tag{3.19b}$$

where $b^i=-\partial\Lambda^{ij}/\partial\tau^j$ can be affected by the choice of the parametrization. If $p=1$, Eqs. (3.19) with $\Lambda^{11}\neq 0$ describe the null geodesic. For $p=2$, in complex V_4, by differentiating covariantly Eqs. (3.19), one easily infers that all derivatives $D^2x^\mu/\partial\tau^i\partial\tau^j$ are orthogonal to all vectors from the null space spanned by $\partial x^\mu/\partial\tau^1$ and $\partial x^\mu/\partial\tau^2$, and, therefore, are vectors belonging to this null space. Thus, there exist C^k_{ij} such that (3.17) applies and, consequently, (3.19b) reduce to the equations $\lambda^{ij}C^k_{ij}+\partial\Lambda^{ki}/\partial\tau^i=0$ for the uninteresting Lagrange multipliers.

It follows that the action (3.19) yields the geodesic null surface as an extremal variety. The last argument was included in order to exemplify the fact that geodesic null surfaces are extremal varieties, derivable from an action principle: Of course, similarly to the case of one-dimensional null geodesics, there exist many different variational principles which lead to the same geodesic null surfaces.

4. CONCLUSIONS

The union of results of Sec. 3 and 4 permits us to give the following formulation of the Goldberg–Sachs theorem as extended to the case of the complex V_4:

Theorem: The undotted or dotted components of the spinorial image of the conformal curvature in empty

complex V_4 (i.e., $G_{ab}=0$) is algebraically degenerate if and only if there exists a (complex) congruence of geodesic null surfaces with parallelly propagated tangent space which contains the degenerate (generalized) DP vector.

The result stated above exhibits the important role of the (extremal) two-dimensional null varieties in the geometry of complex V_4. Therefore, one can say that, in a sense, "analytic continuation" of general relativity leads to the concept of the (complex) null "string" ≡ null geodesic surface; this concept appears in a natural fashion in the "complex anatomy" of the GS theorem. Whether the complex null strings may be useful for more mundane purposes—in particular, whether they can serve as mathematical models of some physical objects—remains an open question.

ACKNOWLEDGMENT

We appreciate a discussion with J. D. Finley.

*On leave of absence from University of Warsaw, Warsaw, Poland.

[1]J. F. Plebañski, *Some Solutions of Complex Einstein Equation*, to be published in J. Math. Phys. (1975).
[2]G. Debney, R. Kerr and A. Schild, J. Math. Phys. **10**, 1842 (1969).
[3]R. Penrose, Ann. Phys. (N.Y.) **10**, 171 (1960).
[4]J. Goldberg and R. Sachs, Acta Phys. Polon. Suppl. 22:13 (1962).
[5]J. F. Plebañski, *Spinors, Tetrads and Forms*, unpublished monograph of Centro de Investigación y Estudios Avanzados del IPN, México 14, D. F. Chapter VI, Sec. 5.
[6]If $C^{(a)}=0=C_{ab}$, but $\lambda\neq 0$, it is likely that the lemma still holds. This point will be examined elsewhere in a more general context.
[7]It is obvious from lemmas *A* and *B* that if, $G_{ab}=\lambda g_{ab}$, $\lambda=\text{const}\neq 0$, the theorem is valid in all cases of algebraic degeneration, except when "heavenly" $C^{(a)}$'s vanish; although it is likely that the theorem applies in this last case.
[8]K. Yano, *The Theory of Lie Derivatives and its Applications* (North-Holland, Amsterdam, 1956).
[9]E. Bompiani, Comptes Rendus **169**, 840 (1919).
[10]J. Rayski, Acta Phys. Polon. **20**, 509 (1961).
[11]L. P. Eisenhart, *Riemannian Geometry* (Princeton U. P., Princeton, 1928), Sec. 52.

Structural properties of the canonical U(3) Racah functions and the U(3) : U(2) projective functions

J. D. Louck*

Theoretical Division, Los Alamos Scientific Laboratory, University of California, Los Alamos, New Mexico 87545

M. A. Lohe†‡ and L. C. Biedenharn†

Department of Physics, Duke University, Durham, North Carolina 27706
(Received 27 May 1975)

The class of U(3) Racah functions which are identically zero are determined from the canonical splitting of the multiplicity. These results imply the form of a special class of (projective) tensor operators. The function G_q associated with the "stretched" (maximal null space) Wigner operator is generalized and shown to be applicable in determining the denominator for the minimal null space operator.

1. INTRODUCTION AND SUMMARY

The construction of the set of all canonical tensor operators in the unitary group U(3) is a formidable task which has been the subject of intensive effort over the past few years.[1-6] The existence of this set of tensor operators has been proven quite early,[6] but the actual construction—as opposed to the "in principle" constructability used in the existence proof—is by no means straightforward.

A major part of the explicit construction is the effort to achieve some kind of intuitive understanding of the properties possessed by the actual answers; for this purpose—which we feel to be the real essence of the problem—a formal answer expressed, say, as an unstructured multiple sum over complicated components is effectively no answer at all. In pursuing this objective we have been led to discuss first the unit tensor ("Wigner") operators[7] which were shown to be composed from simpler components: the U(3) : U(2) projective operators {also known as "isoscalar" factors in physical applications of unitary symmetry [Gell-Mann's SU(3) structure]}.

For the SU(3) projective operators, whose tensor operator irrep labels are $\langle p00\rangle$ and $\langle qq0\rangle$, completely explicit results have been known for some time[8]; these operators[9] are the "building blocks" of the general SU(3) operator $\langle pq0\rangle$, but (as indicated above) the actual formulas expressing this fact are too complicated to admit of anything but the most superficial interpretation (or use).

The definitive characteristic which permits the complete construction of the $\langle p00\rangle$ and $\langle qq0\rangle$ operators is that their operator patterns [the operator-label-space "Gel'fand-like" patterns—denoted by $\Gamma=(\Gamma_{ij})$—which express the canonical labelling] are (necessarily) multiplicity-free. Expressed differently, in a more suggestive way, the elementary operators have *stretched* operator patterns, i.e., patterns of the form

$$(\diagdown\!\vdots\!\cdot)\ \text{and/or}\ (\cdot\!\vdots\!\diagup),$$

where the dots denote the labels Γ_{ij} and the lines denote "tied" (numerically equal) labels.

It has proved possible to extend to projective operators, having general irrep labels $[pq0]$, but specialized to stretched operator patterns Γ_s, this explicit construction. This result (cf. Ref. 4, I and II) takes the form

$$\begin{bmatrix} & (\Gamma_s) & \\ p & q & 0 \\ & p & 0 \\ & & p \end{bmatrix} = (-1)^{p-\Delta_1} F_R \begin{pmatrix} [\Delta_1 & \Delta_2 & \Delta_3] \\ [p & q & 0] \\ & [p & 0] \end{pmatrix} \Big/ D\begin{pmatrix} [\Delta_1 & \Delta_2 & \Delta_3] \\ [p & q & 0] \end{pmatrix}. \quad (1.1)$$

In this result F_R is a restricted arrow-pattern function whose value is read off directly from the pattern calculus rules,[2,4] and the denominator function has the following form:

$$D\begin{pmatrix} [\Delta_1 & \Delta_2 & \Delta_3] \\ [p & q & 0] \end{pmatrix}([m]) = \left[\frac{\Delta_1!\,\Delta_2!\,\Delta_3}{(p+q)!\,\Delta'([m]+[\Delta])} \prod_{i<j=1}^{3} (\Delta_i+\Delta_j+1)!\right] \times \begin{pmatrix} x_{ij}+\Delta_i \\ \Delta_i+\Delta_j+1 \end{pmatrix}^{1/2} \left(\frac{(p+q)!}{(p-q)!\,G_q(\Delta;x)}\right)^{1/2}, \quad (1.2a)$$

where

$$x_{ij} = p_{i3} - p_{j3} \quad (1.2b)$$

and

$$G_q(\Delta;x) = G_q(\Delta_1\Delta_2\Delta_3; x_1x_2x_3), \quad (1.2c)$$

$$x_i = x_{jk} \quad (i,j,k \text{ cyclic in } 1,2,3). \quad (1.2d)$$

This appears to be a very complicated result, but the structure can be easily understood. The first factor, F_R, is the square root of a product of linear factors; this structure is completely known (and easily written out) from the pattern calculus rules.[2] The second factor, $D(\cdots)$, is the denominator function which plays a basic structural role in the canonical splitting of the multiplicity: *the set of irrep labels* $\{[m]\}$ *on which the denominator function vanishes uniquely characterizes the stretched operator pattern* Γ_s, *and conversely.*

Expressed differently, one denotes the representation spaces determined by the set of irreps $\{[m]\}$ on

which D vanishes as the *null space* of the operator $\langle {}_{p\ q\ 0}^{\Gamma_s} \rangle$. The essential result then becomes: *The null space determined by Eq.* (1.2) *uniquely characterizes* $\langle {}_{p\ q\ 0}^{\Gamma_s} \rangle$.

{Let us note, for completeness, that a special U(2) operator pattern, namely $({}_{p}^{p}{}^{0})$, appears in Eq. (1.1); in the general case [i.e., an arbitrary U(2) operator pattern on the left-hand side of Eq. (1.1)] the function F_R would be multiplied by a polynomial [which for $({}_{p}^{p}{}^{0})$ reduces to 1]. It follows that the denominator function—which alone can introduce singularities into the general projective operator—is indeed characteristic of the set of projective operators, and therefore of the unit tensor operator $\langle {}_{p\ q\ 0}^{\Gamma_s} \rangle$ itself.}

What is the null space determined by Eq. (1.2)? There are two pieces to this null space, which account for the form in which Eq. (1.2a) is written:

(a) There is a set of null space lines determined by the linear factors in the product over $i<j$. (See note added in proof.)

(b) There is a triangular set of null space points determined by the zeroes of the polynomial G_q. (Note that these zeroes of G_q occur at intersections of null space lines so that the net result is a first order null space zero. Note also that when the arguments of the G_q function are in a "constricted position"—see Fig. 1 and Eq. (2.15) below—this triangle of zeroes can develop one or more null space lines.)

This summarizes the current status of the tensor operator construction in U(3); we are now in a position to sketch the results to be presented in the present paper. We will generalize the projective operator in Eq. (1.1) to comprise that class of projective operators for which the form of Eq. (1.1) remains valid, that is, a product of a (known) pattern calculus function and an (unknown) denominator function [a function of U(3) and U(2) invariant operators], to be determined. This is the class of "least complicated" projective operators and a major result of the paper will be to give explicitly one such projective operator *for every operator pattern* Γ.

We will determine, in other words, the explicit denominator function for all tensor operators in SU(3). Accordingly, we will be able to verify, in complete detail, the null space properties of the general U(3) tensor operator and illustrate how this structure accords fully with the intertwining number-null space construction discussed in Ref. 4, Paper II.

This is a rather sizeable task, which, as will be seen, has introduced its own insights and subtleties that deserve discussion. One may better understand these new structures if one proceeds through simpler examples. Accordingly, we will present the work in two parts, of which the first part (the present paper) will confine itself to developing the *form* of a class of least complicated projective operators, showing how the denominator functions enter. We will also present in this part the denominator function for the Wigner operator having the minimal null space [the maximal null space denominator function is given by Eqs. (1.2)]. These minimal null space operators include the SU(3) generators.

The method whereby these minimal null space operators are determined is quite special, and depends upon a corresponding special property of the U(3) Racah functions. Section 3 is concerned with this development, which is found to depend on determining which Racah functions vanish identically (a topic of interest in its own right).

Section 4 applies the Racah function results of Sec. 3 to the construction of a class of least complicated operators.

It is the surprising result of Sec. 5 that the denominator function for the minimal null space operators is actually the same function G_q determined for $\Gamma_1=\Gamma_s$ [the stretched pattern, but for *different* parameters (arguments)]. Thus, for simplicity, the next section, Sec. 2, is devoted to this extension of the parameter domain of the G_q function. This extension will in turn lead to a neat reformulation of some of the pattern calculus rules; these results—albeit purely mnemonic—have an appeal of their own which we hope the reader will share.

2. EXTENSION OF THE DENOMINATOR FUNCTION G_q

A. Resume of properties of G_q

The properties of the denominator function G_q were developed (Ref. 4, I, II) in considerable detail, not so much for intrinsic interest, but as a means toward proving the validity of the useful polynomial form [Eq. (2.2) below]. First let us write the function G_q in the form

$$G_q(\xi;x)=G_q\begin{pmatrix}\xi_1 & \xi_2+x_1 & \xi_3-x_1\\ \xi_2 & \xi_3+x_2 & \xi_1-x_2\\ \xi_3 & \xi_1+x_3 & \xi_2-x_3\end{pmatrix}, \tag{2.1}$$

$$\begin{aligned}G_q(\xi;x)=(-1)^q q!\sum_{(k)}&\binom{\xi_1+\xi_2+\xi_3-k_1-k_2-k_3}{k_4}\\ &\times f_{q,k_1}(\xi_1,\xi_2+x_1,\xi_3-x_1)f_{q,k_2}(\xi_2,\xi_3+x_2,\xi_1-x_2)\\ &\times f_{q,k_3}(\xi_3,\xi_1+x_3,\xi_2-x_3),\end{aligned} \tag{2.2}$$

where the sum is over all nonnegative integers k_1,k_2, k_3,k_4 which add to q, $k_1+k_2+k_3+k_4=q$, and

$$f_{q,k}(xyz)\equiv(q-r)!r!\binom{x}{q-r}\binom{y}{r}\binom{z}{r}. \tag{2.3}$$

In the defining equations above for G_q, we have introduced $\xi_i=\Delta_i$, since the definition does not require that the ξ_i be integral. It is essential, however, that the x_i variables [see Eq. (1.2b)] obey the constraint

$$\sum_{i=1}^{3}x_i=0. \tag{2.4}$$

Let us now list some of the properties of $G_q(\xi,x)$:

(a) $G_q(\xi;x)$ is a polynomial of degree $2q$ in the variables x_i.

(b) $G_q(\xi;x)$ is unchanged by a transposition or any permutation of the rows or columns of the 3×3 array of parameters in Eq. (2.1).

(c) $G_q(\xi;x)$ has trigonal symmetry in the barycentric (Möbius) plane $\{x_i\}$. There is a center of symmetry at the point:

$x_i = (\xi_k - \xi_j)/2$, (ijk) cyclic.

(d) $G_q(\xi; x)$ has zeroes at the lattice points of an equilateral triangle of side q whose vertex and corresponding two sides are on the lines $x_1 = \xi_3$, $x_3 = \xi_2$. The symmetry in (c) shows that there are six such triangles of zeroes.

B. The necessity of extending the definition

We presented in Ref. 4II an example of the remarkable symmetry properties of the G_q function. This example concerned the zeroes of the function $G_3(352; x_1x_2x_3)$ and, as we will see, supplies the motivation for generalizing the definition of G_q.

Consider Fig. 1 (reproduced from Ref. 4II). The remarkable property shown by this figure is that if we remove the two lines of zeroes (shown as — · — lines) *we obtain a new figure* (of six equilateral triangles) *obeying the symmetry* [item (c) above] *characteristic of the G_q function itself.* Since the degree of $G_3(352; x)$ is six, and we remove two linear factors, the resulting degree $2q = 6 - 2 = 4$ checks with the size of the new triangles (side $= 2 = q$). But the result $[G_3(352; x)/(\text{two linear factors})]$ does not fit the definition, Eq. (2.2), for a function G_2 of the type $G_2(\Delta_1'\Delta_2'\Delta_3'; x)$ in which the $x_i = p_{j3} - p_{k3}$ variables are unaltered. In studying the polynomial which remains after removing the linear terms from $G_q(\Delta; x)$, we have been led to introduce what appears, at first glance, to be a more general polynomial. We will show subsequently [cf. Eq. (2.13d)] that this new polynomial is, in fact, of the form

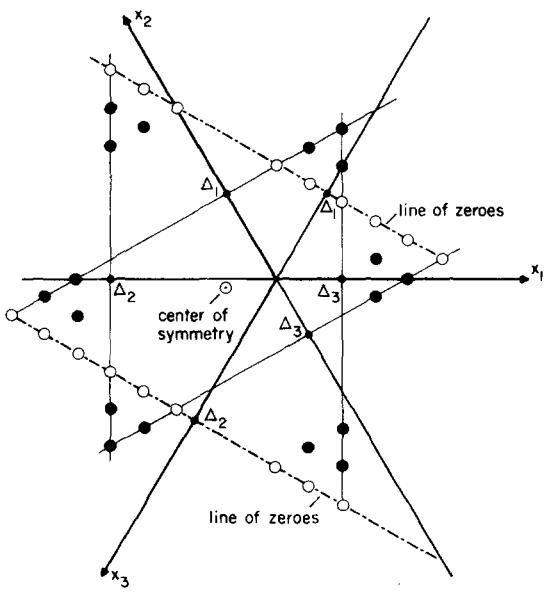


FIG. 1. Zeroes of the polynomial $G_3(352; x_1x_2x_3)$. This polynomial vanishes at each of the six points (three large open circles and three large solid circles) of each of the six equilateral triangles symmetrically placed about the center of symmetry at the point $(-3/2, 1/2, 1)$. The linear factors of the polynomial are $(x_3 + 3)(x_3 - 5)$. Hence, the polynomial also vanishes on the lines $x_3 = -3$ and $x_3 = 5$ (the dash–dot lines). Removing these linear factors from the polynomial leaves a new polynomial which still vanishes at each of three points (the large solid circles) of each of six equilateral triangles which are still symmetrically placed about the center of symmetry.

$G_{q'}(\xi'; x')$ in which we *simultaneously* shift the ξ_i and the x_i. The form we present below (with *unshifted* x_i) has, however, an interesting interpretation in terms of the pattern calculus rules (which we will discuss); this interpretation will, in fact, make the form given below the more useful of the two possible forms.

We now define:

$$G_q(\Xi; x) = G_q\begin{pmatrix} \xi_{11} & \xi_{12}+x_1 & \xi_{13}-x_1 \\ \xi_{21} & \xi_{22}+x_2 & \xi_{23}-x_2 \\ \xi_{31} & \xi_{32}+x_3 & \xi_{33}-x_3 \end{pmatrix}, \tag{2.5}$$

$$\begin{aligned} G_q(\Xi; x) = (-1)^q q! \sum_{(k)} & \binom{z-k_1-k_2-k_3}{k_4} \\ & \times f_{q,k_1}(\xi_{11}, \xi_{12}+x_1, \xi_{13}-x_1) \\ & \times f_{q,k_2}(\xi_{21}, \xi_{22}+x_2, \xi_{23}-x_2) \\ & \times f_{q,k_3}(\xi_{31}, \xi_{32}+x_3, \xi_{33}-x_3), \end{aligned} \tag{2.6}$$

where the sum is over all nonnegative integers k_1, k_2, k_3, k_4 which add to q, $k_1+k_2+k_3+k_4 = q$, and, as before,

$$f_{q,r}(xyz) = (q-r)!\,r!\binom{x}{q-r}\binom{y}{r}\binom{z}{r}. \tag{2.7}$$

Let us call the new, more general, 3×3 array introduced in Eq. (2.5) the "Ξ-array":

$$\Xi = \begin{pmatrix} \xi_{11} & \xi_{12} & \xi_{13} \\ \xi_{21} & \xi_{22} & \xi_{23} \\ \xi_{31} & \xi_{32} & \xi_{33} \end{pmatrix}, \tag{2.8a}$$

$$z = \sum_i \xi_{ij} = \sum_k \xi_{mk}, \quad \textit{for every } (j, m). \tag{2.8b}$$

The essential characteristic of the Ξ-array is that it is a magic square. The constraint, Eq. (2.4), is also imposed in the new definition of G_q.

It is easily seen that the previous definition accords fully with the new definition, as a special case.

Let us now establish several general properties of the polynomials defined by Eqs. (2.5)–(2.8), for a general Ξ-array which satisfies the magic square constraint.

The first general property follows directly from the same considerations used in Ref. 4II, namely that: *The polynomial $G_q(\Xi; x)$ for a general Ξ-array is invariant under transposition and under all permutations of rows, and of columns, of the Ξ-array matrix.*

In order to establish next the triangles of zeroes that characterize the G_q function, we make use of the identity, Eq. (2.3a) of Ref. 4II. (This identity is the Saalschütz identity tailored to the present discussion.) This step reduces Eq. (2.6) to the form (we choose to carry out the summation over k_2, i.e., we combine the binomial factor with f_{q,k_2} and sum out k_2):

$$\begin{aligned} G_q(\Xi; x) = (-1)^q q! \sum_{k_1+k_2+k_3=q} & f_{q,k_1}(\xi_{11}, \xi_{12}+x_1, \xi_{13}-x_1) \\ & \times f_{q,k_2}(\xi_{21}, z-\xi_{23}+x_2+k_2-q, z-\xi_{22}-x_2+k_2-q) \\ & \times f_{q,k_3}(\xi_{31}, \xi_{32}+x_3, \xi_{33}-x_3). \end{aligned} \tag{2.9}$$

(Note that k_2 has been restored in the summation and the magic square condition has been used.)

From this expression now consider the following fac-

tors which appear under the summation:

$$\binom{\xi_{13}-x_1}{k_1}\binom{\xi_{33}-x_3}{k_3}\binom{\xi_{23}-z-x_2+q-1}{k_2}. \tag{2.10}$$

[Note that the last term comes from using $\binom{-\lambda}{k}=(-1)^q \times \binom{\lambda+q-1}{q}$.] Now suppose that $\xi_{13}-x_1$ and $\xi_{33}-x_3$ are nonnegative integers satisfying

$$(\xi_{13}-x_1)+(\xi_{33}-x_3)\leq q-1. \tag{2.11}$$

Then also $\xi_{23}-z-x_2+q-1$ is a nonnegative integer. Thus, under the condition (2.11), we see that in Eq. (2.10) the sum of nonnegative integers in the binomial coefficients is

$$(\xi_{13}-x_1)+(\xi_{33}-x_3)+(\xi_{23}-z-x_2+q-1)=q-1,$$

and, since $k_1+k_2+k=q$, at least one of the factors in (2.10) is always zero whenever (2.11) holds. We have thus proved: *The polynomial $G_q(\Xi;x)$ has value zero whenever $\xi_{13}-x_1$ and $\xi_{33}-x_3$ are nonnegative integers satisfying*

$$(\xi_{13}-x_1)+(\xi_{33}-x_3)\leq q-1.$$

This result establishes that there exists a triangular lattice of zeroes, q zeroes on each side of the triangle, with one vertex of the triangle placed at $(x_1x_3)=(\xi_{13},\xi_{33})$. Using the symmetries established for the Ξ-array, this result extends to define six equilateral triangles of zeroes.

The six triangles on which $G_q(\Xi;x)$ vanishes are depicted in Fig. 2. The diagram is drawn under the assumption that the ξ_{ij} are all nonnegative. The ξ_{ij} situated near the dots designate *distances* from the origin.

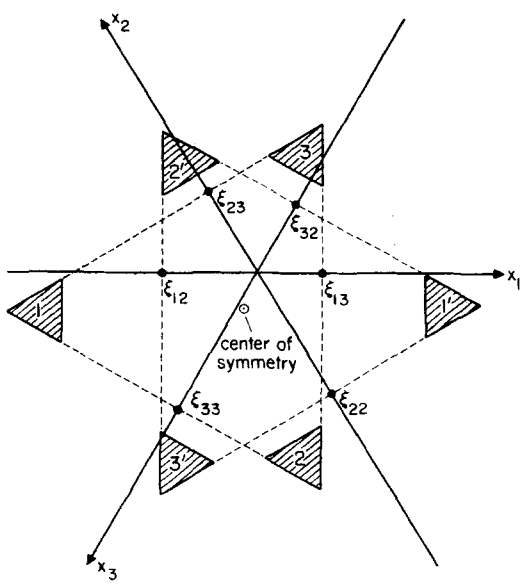


FIG. 2. Zeroes of the polynomial $G_q(\Xi;x)$. Regions 1, 2, and 3 are defined, respectively, by the conditions $(\xi_{33}-x_3)+(\xi_{23}-x_2)\leq q-1$, $(\xi_{13}-x_1)+(\xi_{33}-x_3)\leq q-1$, and $(\xi_{23}-x_2)+(\xi_{13}-x_1)\leq q-1$. Regions 1′, 2′, and 3′ are, respectively, the reflections 1, 2, and 3 through the center of symmetry located at $[(\xi_{13}-\xi_{12})/2,\ (\xi_{23}-\xi_{22})/2,\ (\xi_{33}-\xi_{32}/2]$ and are defined by the condtions $(\xi_{32}+x_3)+(\xi_{22}+x_2)\leq q-1$, $(\xi_{12}+x_1)+(\xi_{32}+x_3)\leq q-1$, and $(\xi_{22}+x_2)+(\xi_{12}+x_1)\leq q-1$. The lattice points within and on the boundaries of these six triangles are zeroes of $G_q(\Xi;x)$.

Let us now replace the Ξ-array by $-\Xi+q-1$, where Ξ is still *defined generally* and the notation $-\Xi+q-1$ symbolizes $-\xi_{ij}+q-1$. This leaves us still with a magic square. Next consider $G_q(-\Xi+q-1;-x)$. Observe that triangle 2 in Fig. 2 now becomes the triangle of points defined by the conditions

$$(\xi_{13}-x_1)+(\xi_{33}-x_3)\geq q-1. \tag{2.12}$$

The triangle of points defined by these conditions is precisely the reflection of triangle 2 of Fig. 2 through the base line nearest the origin. Observe that this same phenomenon happens with respect to each of the six triangles! *The operations $\Xi\rightarrow-\Xi+q-1$, $x\rightarrow-x$ carry the points of the six triangles depicted in Fig. 2 into the points of the six triangles obtained by rotating each of the original triangles about its base line nearest the origin.* (A remarkable feat!)

This "geometric" result will be of considerable importance in the construction of the minimal null space denominator function in Sec. 5.

The polynomial $G_q(\Xi;x)$ appears to be more general than $G_q(\xi;x)$. However, upon putting

$$\begin{pmatrix}\xi_{11} & \xi_{12}+x_1 & \xi_{13}-x_1\\ \xi_{21} & \xi_{22}+x_2 & \xi_{23}-x_2\\ \xi_{31} & \xi_{32}+x_3 & \xi_{33}-x_3\end{pmatrix}=\begin{pmatrix}\xi_1 & \xi_2+x_1' & \xi_3-x_1'\\ \xi_2 & \xi_3+x_2' & \xi_1-x_2'\\ \xi_3 & \xi_1+x_3' & \xi_2-x_3'\end{pmatrix} \tag{2.13a}$$

so that

$$\xi_i=\xi_{i1}, \tag{2.13b}$$

we see that the magic square conditions (2.8) are just sufficient to determine uniquely the shifted x_i' variables [still satisfying $\sum_i x_i'=0$]:

$$\begin{aligned}x_1'&=x_1+\xi_{12}-\xi_{21},\\ x_2'&=x_2+\xi_{22}-\xi_{31},\\ x_3'&=x_3+\xi_{32}-\xi_{11}.\end{aligned} \tag{2.13c}$$

Thus we have proved

$$G_q(\Xi;x)=G_q(\xi;x'). \tag{2.13d}$$

Note that we could also now derive the symmetries and zeroes of $G_q(\Xi;x)$ given by Fig. 2, directly from the symmetries and zeroes of $G_q(\xi;x')$ proved in Ref. 4II. Indeed, we can now write out recursion relations, etc., directly for the $G_q(\Xi;x)$. More important is the following result (cf. Proposition 3 of Ref. 4II): *The set of zeros of the six triangles depicted in Fig. 2 uniquely determine $G_q(\Xi;x)$ up to a multiplicative factor which depends at most on the $\xi_i=\xi_{i1}$.*

C. A reduction formula

We began the discussion of Sec. 2B above by noting—from our motivating example of $G_3(352;x)$—the necessity of generalizing the Ξ-array in the original definition of the G_q function. Let us now complete the discussion of our motivating example by giving the general form of the relationship between the $G_q(\Delta;x)$ and the new $G_q(\Xi;x)$ function.

For this purpose, we introduce the "step functions" λ_i defined in the following manner:

$$\lambda_i = \begin{cases} q - \Delta_i & \text{for } \Delta_i \le q \\ 0 & \text{for } \Delta_i > q. \end{cases} \tag{2.14}$$

The introduction of this new notation might appear rather trivial, while, in fact, it will enable us to unify many results which otherwise would be presented as different formulas or cases. It is worth noting specific instances in which this unification occurs: (1) the reduction formula, Eq. (2.15), below; (2) the multiplicity formula, Eq. (2.17); (3) the intertwining number-null space diagram, Fig. 3 below; (4) a generalization of the pattern calculus rules [cf. Sec. 2D]; (5) the general identification of operator patterns, Eq. (4.4a); and (6) the limit formula given in Appendix A.

Returning now to the discussion of the reduction formula, we prove directly the following relation[10]:

$$G_q\begin{pmatrix} \Delta_1 & \Delta_2 + x_1 & \Delta_3 - x_1 \\ \Delta_2 & \Delta_3 + x_2 & \Delta_1 - x_2 \\ \Delta_3 & \Delta_1 + x_3 & \Delta_2 - x_3 \end{pmatrix}$$
$$= [q]_{\lambda_1+\lambda_2+\lambda_3} \prod_{i=1}^{3} (-1)^{\lambda_i} \frac{[\Delta_i]_{\lambda_j+\lambda_k} [\Delta j + x_i]_{\lambda_i} [\Delta_k - x_i]_{\lambda_i}}{[q - \lambda_j - \lambda_k]_{\lambda_i}}$$
$$\times G_{q-\lambda_1-\lambda_2-\lambda_3}\begin{pmatrix} \Delta_1 + \lambda_1 - \lambda_2 - \lambda_3 & \Delta_2 - \lambda_1 + x_1 & \Delta_3 - \lambda_1 - x_1 \\ \Delta_2 + \lambda_2 - \lambda_3 - \lambda_1 & \Delta_3 - \lambda_2 + x_2 & \Delta_1 - \lambda_2 - x_2 \\ \Delta_3 + \lambda_3 - \lambda_1 - \lambda_2 & \Delta_1 - \lambda_3 + x_3 & \Delta_2 - \lambda_3 - x_3 \end{pmatrix}, \tag{2.15}$$

where ijk are cyclic in 123 and we have introduced the notation

$$[x]_a = x(x-1)\cdots(x-a+1), \quad a = 1, 2, \cdots, \tag{2.16a}$$

$$[x]_0 = 1 \tag{2.16b}$$

for a falling factorial.

We denote the Ξ-array in Eq. (2.15) as a "reduced Ξ-pattern." Note that this reduced Ξ-array automatically obeys the magic square constraint.

D. Interpretation in terms of the pattern calculus

The results obtained in Subsections A–C above lead to a considerable simplification in our previous discussion of the intertwining number–null space diagram.

For example, instead of the complicated (tabular) discussion of the multiplicity of a given Δ-pattern—which led in all to eight separate cases to consider—we may now give a single comprehensive formulation. Using the step functions λ_i, *we may write the following single formula for the multiplicity* $\mathcal{M}$ of the Δ-pattern $\Delta \in [p\ q\ 0]$:

$$\mathcal{M} = q - \lambda_1 - \lambda_2 - \lambda_3 + 1, \tag{2.17}$$

where the λ_i in this result are defined by Eq. (2.14).

We may now also reformulate the "coordinate points" of the intertwining number—null space diagram in the manner shown in Fig. 3 (cf. Ref. 4II). The importance of the intertwining number—null space diagram has been discussed in Refs. 3, 4. The new result we wish to note here is that the polynomial $G_{q-\lambda_1-\lambda_2-\lambda_3}(\Xi; x)$, where Ξ is the reduced pattern given in Eq. (2.15), vanishes on the lattice points of the equilateral triangle $P_1'P_2'P_3'$ (cf. Fig. 3), where

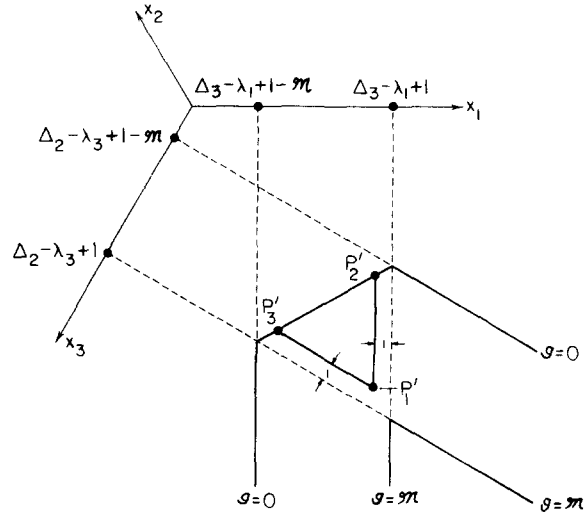


FIG. 3. The intertwining number-null space diagram. The intertwining number $\mathcal{G}$ is defined at each lattice point (points having integral coordinates) of the Möbius plane. At each lattice point in the pie-shaped region bounded by the bent solid line, $\mathcal{G} = \mathcal{M}$, the value of $\mathcal{G}$ is $\mathcal{M}$; at lattice points in the region between the bent solid lines $\mathcal{G} = 0$ and $\mathcal{G} = \mathcal{M}$, the value of $\mathcal{G}$ varies from 1 to $\mathcal{M} - 1$, its value being k on those lattice points on the bent line which is k units away from the $\mathcal{G} = 0$ line. At all other lattice points $\mathcal{G} = 0$. The polynomial $G_{q-\lambda_1-\lambda_2-\lambda_3}(\Xi; x)$, where Ξ is the reduced Ξ-pattern displayed in Eq. (2.15), vanishes on the lattice points of the equilateral triangle $P_1'P_2'P_3'$.

$$P_1' = (\Delta_3 - \lambda_1, -\Delta_2 - \Delta_3 + \lambda_1 + \lambda_3, \Delta_2 - \lambda_3),$$
$$P_2' = (\Delta_3 - \lambda_1, -\Delta_2 - \Delta_3 - 1 + q - \lambda_2, \Delta_2 - q + 1 + \lambda_1 + \lambda_2),$$
$$P_3' = (\Delta_3 - q + 1 + \lambda_2 + \lambda_3, -\Delta_2 - \Delta_3 - 1 + q - \lambda_2, \Delta_2 - \lambda_3). \tag{2.18}$$

These results can be readily verified by a point by point change of notation from the eight cases detailed in Ref. 4II. The eight sets of values which the λ's can assume handles the situation so much better that it holds out hope that the general problem in U(n) can be similarly handled.

Not only does the new formulation simplify previous results, but it also allows a similar unification of some of the pattern calculus rules. We will now show that *the reduced* Ξ*-pattern given in Eq.* (2.15) *may be associated directly with the linear factors that appear in the denominator function.*

The usefulness of this result is best understood from an example. Let us consider the ⟨420⟩ operator (the 27-plet operator given in complete detail by Castilho-Alcarás, *et al.*[11]). In particular, consider the projective operator

$$\begin{bmatrix} & 3 & \\ 4 & & 0 \\ 4 & 2 & 0 \\ 4 & & 0 \\ & 0 & \end{bmatrix}.$$

For this operator the original Ξ-pattern is

$$\begin{pmatrix} 3 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},$$

and the reduced Ξ-pattern is

$$\begin{pmatrix} 2 & 1 & 2 \\ 2 & 1 & 2 \\ 1 & 3 & 1 \end{pmatrix} .$$

[Note that there is no Δ-pattern that will give this reduced Ξ-pattern, unless we also make the corresponding shift of the x_i given by Eq. (2.13c).]

For this operator the shifts are $\Delta=[312]$. The linear factors in the denominator (1.2) are obtained from the pattern-on-pattern rule for totally symmetric operators; this yields 12 linear factors. But the reduction of $G_q(\Delta;x)$ yields two linear factors. These two factors cancel 2 of the 12 factors leaving 10 factors.

Let us now show how these 10 linear factors may be obtained directly from the Ξ-pattern. We write out the Ξ-pattern, including the x-variables:

$$\begin{pmatrix} 2 & 1+x_1 & 2-x_1 \\ 2 & 1+x_2 & 2-x_2 \\ 1 & 3+x_3 & 1-x_3 \end{pmatrix} . \tag{2.19}$$

It is the integer pairs (1,2), (1,2), and (3,1) appearing in columns 2 and 3 which are important in the rule for obtaining the linear factors. We now give the rule. To determine the linear factors depending on $x_1=p_{23}-p_{33}$, we write out two sets of three dots as shown below. With the integer pair (1,2) appearing in row 1 (the row containing x_1), we now associate arrows going between dots 2 and 3 in each of these sets of three dots. In the first set of three dots, the first integer in the pair (1,2) tells us to draw 1 arrow (we write the 1 under dot 2) from dot 2 to dot 3, and to write the second integer 2 above dot 3. For the second set of three dots, we reverse the procedure by drawing 2 arrows from dot 3 to dot 2, writing the 2 under dot 3 and the 1 above dot 2. We thus arrive at the following diagram.

We now use our standard $p(\text{tail})-[p(\text{head})+\delta]$ rule of the pattern calculus, where δ is the integer sitting above the arrow head. Thus, the factors in x_1 are

$$(p_{23}-p_{33}-2)(p_{33}-p_{23}-1)(p_{33}-p_{23}), \tag{2.20a}$$

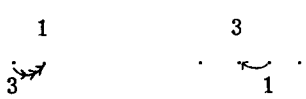
where the first factor $(p_{23}-p_{33}-2)$ comes from the single arrow in the first set of three dots and the factors $(p_{33}-p_{23}-1)(p_{33}-p_{23})$ come from the two arrows in the second set of three dots.

We now repeat this procedure for rows 2 and 3 of the Ξ-pattern. The diagrams for row 2 are

yielding factors

$$(p_{33}-p_{13}-2)(p_{13}-p_{33}-1)(p_{13}-p_{33}). \tag{2.20b}$$

The diagrams for row 3 are



yielding factors

$$(p_{13}-p_{23}-1)(p_{13}-p_{23})(p_{13}-p_{23}+1)(p_{23}-p_{13}-3). \tag{2.20c}$$

The 10 linear factors given by the rules above are just the linear factors obtained directly from Eqs. (1.2) and (2.15).

All linear factor in the U(3) *denominator function* (1.2) *are obtained directly from the reduced* Ξ*-pattern by this rule.* The proof is by direct verification.

Observe how nicely this new rule contains *both* the extremal denominator rule[2] and the pattern-on-pattern rule[4] as special cases.

The denominator function (1.2) now takes the form

$$D\begin{pmatrix} [\Delta_1 & \Delta_2 & \Delta_3] \\ [p & q & 0] \end{pmatrix}([m])$$
$$= \left[\frac{\prod_{i=1}^{3}(\Delta_i-\lambda_j-\lambda_k)!\,(q-\lambda_j-\lambda_k)!}{q!\,(p-q)!\,[(q-\lambda_1-\lambda_2-\lambda_3)!\,]^2}\right]^{1/2}$$
$$\times\left[\frac{|\text{product of linear factors}|}{G_{q-\lambda_1-\lambda_2-\lambda_3}(\Xi;x)}\right]^{1/2}, \tag{2.21}$$

where the product of linear factors is given by the rule described above and Ξ is the reduced pattern (2.15).

In our discussion of $G_q(\Xi;x)$, we have chosen to define the ξ-values as arbitrary (not necessarily integer-valued) so that one may consider the G_q functions as a continuous function of the Δ's extending the meaning beyond the integer valued Δ's that alone are meaningful for U(3). This allows one to consider the six triangles of zeroes in general position in the Möbius plane, and to continuously vary this position. Thus, as the triangles are allowed to continuously approach the center of symmetry, when the triangles get too close we have the phenomenon illustrated by the motivating example, $G_3(352;x)$: Quite suddenly the form of the function changes and *lines of zeroes appear*. The appearance of these lines implies that a restriction has occurred in the reduction of the U(3) direct product—this information being contained implicitly in the denominator function and in G_q itself. *In such a way the purely group theoretic aspects of the tensor operator problem are contained in the remarkable algebraic properties of the associated functions.*

These interpretive concepts will be extended considerably when the full set of denominator functions is obtained in the second part of the present work.

3. A CLASS OF ZERO U(3) RACAH FUNCTIONS

It is our purpose in this section to develop further the relationship between the denominator functions (of projective operators) and the Racah functions, confining our attention to U(3).

A. Zero projective operators[12]

Let us recall that the canonical splitting conditions[6] imply the following zero projective operators:

$$\begin{bmatrix} & (\Gamma_k) & \\ p & q & r \\ & p \quad r & \\ & \gamma & \end{bmatrix} = 0 \tag{3.1a}$$

for all γ which satisfy

$$p \geq \gamma \geq p-k+2. \tag{3.1b}$$

The index k in Eq. (3.1a) takes on the values $k=2,\ldots,\mathcal{M}$, where the $\mathcal{M}$ distinct operator patterns $(\Gamma_1),(\Gamma_2),\ldots,(\Gamma_{\mathcal{M}})$ all have the same Δ-pattern, $[\Delta_1\ \Delta_2\ \Delta_3]$. $\mathcal{M}$ is the multiplicity of this Δ-pattern in $[pqr]$, and it may be given the form:

$$\mathcal{M}=\begin{cases}(q-r+1)-(\lambda_1+\lambda_2+\lambda_3) & \text{for } p\geq\Delta_i\geq r\\ 0 & \text{otherwise,}\end{cases} \tag{3.2}$$

as discussed in Sec. 2.

Equation (3.1a) by no means enumerates all of the projective operators which are zero; this fact is immediately evident from examining the explicit results for the 27-plet operator given by Castilho-Alcarás, *et al.*[11]

The zeroes given by Eq. (3.1) do, however, uniquely determine (up to phases) the U(3) Wigner (and, hence, Racah) coefficients; it follows that all other zeroes are, in principle, derivable. It is useful for our purpose (deriving zero Racah invariant operators) to extend the class of zero operators considerably beyond those defining the canonical splitting, Eq. (3.1).

In particular, we seek to prove that the following projective operators are the *zero operator*:

$$\left[\begin{array}{ccc} & (\Gamma_k) & \\ p & q & r \\ \alpha & & \beta \\ & \gamma & \end{array}\right]=0 \tag{3.3a}$$

for all indices $k=2,3,\ldots,\mathcal{M}$ and all lexical patterns $\binom{\alpha\ \beta}{\gamma}$ which satisfy the inequality

$$\alpha\geq\gamma\geq\beta+(p-k-r+2). \tag{3.3b}$$

If we let (m) and $(m')_{\alpha+\beta,\gamma}$ denote the Gel'fand patterns

$$(m)=\begin{pmatrix} m_{13} & & m_{23} & & m_{33}\\ & m_{12} & & m_{22} & \\ & & m_{11} & & \end{pmatrix} \tag{3.4a}$$

and

$$(m')_{\alpha+\beta,\gamma}=\begin{pmatrix} m_{13}+\Delta_1 & & m_{23}+\Delta_2 & & m_{33}+\Delta_3\\ & m_{12}+\gamma & & m_{22}+\alpha+\beta-\gamma & \\ & & m_{11}+\delta & & \end{pmatrix}, \tag{3.4b}$$

respectively, then the proof of Eq. (3.3) is equivalent to proving that the following Wigner coefficients are zero for *arbitrary lexical labels* $m_{13}\geq m_{12}\geq m_{23}\geq m_{22}\geq m_{33}$:

$$\langle(m')_{\alpha+\beta,\gamma}|\left\langle\begin{array}{ccc} & (\Gamma_k) & \\ p & q & r \\ \alpha & & \beta \\ & \delta & \end{array}\right\rangle|(m)\rangle=0 \tag{3.4c}$$

for all α,β,γ which satisfy Eq. (3.3b) and for all lexical δ.

The starting point of the proof is the canonical U(3) splitting conditions[6] in the form:

$$\langle(m')_{p+r,\gamma}|\left\langle\begin{array}{ccc} & (\Gamma_k) & \\ p & q & r \\ p & & \gamma \\ & \delta & \end{array}\right\rangle|(m)\rangle=0 \tag{3.5a}$$

for $k\geq 2$ and

$$p\geq r\geq p-k+2. \tag{3.5b}$$

We will prove that the more general property expressed by Eqs. (3.4) and (3.3) follows from Eqs. (3.5).

The proof is by induction. We consider that k is specified (fixed), and let N denote an integer such that $0\leq N\leq k-2$. *We assume that Eqs.* (3.4) *have been proved for all* α *and* β *which satisfy*

$$\alpha-\beta\geq p-r-N,$$

and for all γ *which satisfy* Eq. (3.3b) (induction hypothesis). We will prove that this assumption implies that Eqs. (3.4) are true for all α and β which satisfy

$$\alpha-\beta\geq p-r-N-1$$

and for all γ *which satisfy Eq.* (3.3b) induction closes the induction loop on N, and since Eqs. (3.4) are true for $N=0$, they are true for $N=0,1,\ldots,k-2$, that is, for all lexical α,β,γ which satisfy Eqs. (3.3b).

We first prove that the induction hypothesis implies the validity of the following two equations for all α,β which satisfy $\alpha-\beta\geq p-r-N$:

$$\langle(m')_{\alpha+\beta-1,\gamma}|\left[E_{32},\left\langle\begin{array}{ccc} & (\Gamma_k) & \\ p & q & r \\ \alpha & & \beta \\ & \delta & \end{array}\right\rangle\right]|(m)\rangle=0 \tag{3.6a}$$

for

$$\alpha-1\geq\gamma\geq\beta+(p-k-r+2), \tag{3.6b}$$

$$\langle(m')_{\alpha+\beta+1,\gamma}|\left[E_{23},\left\langle\begin{array}{ccc} & (\Gamma_k) & \\ p & q & r \\ \alpha & & \beta \\ & \delta & \end{array}\right\rangle\right]|(m)\rangle=0 \tag{3.7a}$$

for

$$\alpha\geq\gamma\geq\beta+1+(p-k-r+2). \tag{3.7b}$$

To prove these results, consider the expansion of Eq. (3.6a) obtained by writing the commutator in the form

$$E_{32}\langle\rangle-\langle\rangle E_{32}. \tag{3.8}$$

We evaluate the matrix element of the first term by letting $E_{32}^{\dagger}=E_{23}$ act on the final state vector $|(m')_{\alpha+\beta-1,\gamma}\rangle$; we evaluate the matrix element of the second term by letting E_{32} act on $|(m)\rangle$. The result of this calculation is:

$$\begin{aligned}\langle(m')_{\alpha+\beta-1,\gamma}|&\left[E_{32},\left\langle\begin{array}{ccc} & (\Gamma_k) & \\ p & q & r \\ \alpha & & \beta \\ & \delta & \end{array}\right\rangle\right]|(m)\rangle\\ &=(\#)\langle(m')_{\alpha+\beta,\gamma+1}|\left\langle\begin{array}{ccc} & (\Gamma_k) & \\ p & q & r \\ \alpha & & \beta \\ & \delta & \end{array}\right\rangle|(m)\rangle\\ &\quad+(\#)\langle(m')_{\alpha+\beta,\gamma}|\left\langle\begin{array}{ccc} & (\Gamma_k) & \\ p & q & r \\ \alpha & & \beta \\ & \delta & \end{array}\right\rangle|(m)\rangle\end{aligned}$$

$$+(\#)\langle (m')_{\alpha+\beta,\,\gamma+1}| \left\langle \begin{matrix} & (\Gamma_k) & \\ p & q & r \\ \alpha & & \beta \\ & \delta & \end{matrix} \right\rangle |(m)\rangle_{m_{12}-m_{12}-1}$$

$$+(\#)\langle (m')_{\alpha+\beta,\,\gamma}| \left\langle \begin{matrix} & (\Gamma_k) & \\ p & q & r \\ \alpha & & \beta \\ & \delta & \end{matrix} \right\rangle |(m)\rangle_{m_{22}-m_{22}-1}, \tag{3.9}$$

where # denotes numbers which are known matrix elements of E_{23} and E_{32}, the detailed form being unimportant, and where the notation for the last two matrix elements designates that the labels m_{12} and m_{22} are to be replaced by $m_{12}-1$ and $m_{22}-1$, respectively. The right-hand side of Eq. (3.9) is zero for all α, β satisfying $\alpha-\beta \geq p-r-N$ and all γ satisfying Eq. (3.6b), this conclusion being a consequence of the induction hypothesis. Thus, we have proved that the induction hypothesis implies Eqs. (3.6).

Equations (3.7) are similarly implied by the induction hypothesis.

The next step in the proof is to use the tensor operator property

$$\left[E_{32}, \left\langle \begin{matrix} & (\Gamma_k) & \\ p & q & r \\ \alpha & & \beta \\ & \delta & \end{matrix} \right\rangle \right] = (\#) \left\langle \begin{matrix} & (\Gamma_k) & \\ p & q & r \\ \alpha-1 & & \beta \\ & \delta & \end{matrix} \right\rangle + (\#) \left\langle \begin{matrix} & (\Gamma_k) & \\ p & q & r \\ \alpha & & \beta-1 \\ & \delta & \end{matrix} \right\rangle \tag{3.10}$$

in Eq. (3.6a). Substituting this relation into Eq. (3.6a) and noting that the matrix element of the second operator on the right-hand side of Eq. (3.10) is zero by the induction hypothesis, we obtain

$$\langle (m')_{\alpha+\beta-1,\,\gamma}| \left\langle \begin{matrix} & (\Gamma_k) & \\ p & q & r \\ \alpha-1 & & \beta \\ & \delta & \end{matrix} \right\rangle |(m)\rangle = 0 \tag{3.11}$$

for all α, β, γ which satisfy $\alpha-\beta \geq p-r-N$ and $\alpha-1 \geq \gamma \geq \beta+(p-k-r+2)$. Similarly, from Eq. (3.7), we obtain

$$\langle (m')_{\alpha+\beta+1,\,\gamma}| \left\langle \begin{matrix} & (\Gamma_k) & \\ p & q & r \\ \alpha & & \beta+1 \\ & \delta & \end{matrix} \right\rangle |(m)\rangle = 0 \tag{3.12}$$

for all α, β, γ which satisfy $\alpha-\beta \geq p-r-N$ and $\alpha \geq \gamma \geq \beta+1+(p-k-r+2)$.

The two results, Eqs. (3.11) and (3.12), extend the induction hypothesis for N to $N+1$, and complete the proof of the general validity of Eqs. (3.4) for all lexical patterns $\binom{\alpha}{\gamma\ \beta}$ which satisfy Eq. (3.3b). This result, in turn, implies the zero projective operators described by Eqs. (3.3).

Thus, we have shown that the following projective operators are zero:

$$\left[\begin{matrix} & (\Gamma_k) & \\ p & q & r \\ \alpha & & \beta \\ & \gamma & \end{matrix} \right] = 0 \tag{3.13a}$$

for all α, β, and γ which satisfy

$$\alpha \geq \gamma \geq \beta+(p-k-r+2). \tag{3.13b}$$

We may now *Hermitian conjugate* the operators (3.13) and make an appropriate re-identification of labels to prove that

$$\left[\begin{matrix} & (\Gamma_k) & \\ p & q & r \\ \alpha & & \beta \\ & \gamma & \end{matrix} \right] = 0 \tag{3.14a}$$

for all α, β, γ which satisfy

$$\alpha-(p-k-r+2) \geq \gamma \geq \beta. \tag{3.14b}$$

While we have no proof that these two results [Eqs. (3.13) and (3.14)] exhaust *all* cases of zero projective operators, they do account for those cases which are presently known from direct calculation.

B. Application to Racah invariants[13]

The zeroes identified above imply that a class of Racah operators are zero. We shall now prove that the following class of Racah invariant operators are the zero operator[12]:

$$\left\{ \begin{pmatrix} [M]+[\Delta(\Lambda)] \\ (\Gamma_k'') \end{pmatrix} \begin{pmatrix} (\Lambda) \\ [M'] \\ (\Gamma_j') \end{pmatrix} \begin{pmatrix} [M] \\ (\Gamma_i) \end{pmatrix} \right\} = 0 \tag{3.15a}$$

for all indices i, j, k which satisfy

$$i+j-k \geq [M'_{13}-\Delta_1(\Lambda)]-[M'_{33}-\Delta_3(\Lambda)]+2, \tag{3.15b}$$

where the integers i, j, k must also be in the intervals $1 \leq i \leq \mathcal{M}$, $1 \leq j \leq \mathcal{M}'$, $1 \leq k \leq \mathcal{M}''$, where $\mathcal{M}$ is the multiplicity of $\Delta=\Delta(\Gamma_i)$ in $[M]$, etc.

This result is a direct consequence of the coupling law for projective operators and the zeroes established in Eqs. (3.13) and (3.14).

To prove Eqs. (3.15), consider the coupling law for projective operators[3]:

$$\sum_{\gamma',\gamma} \left[\begin{pmatrix} [M]+[\Delta(\Lambda)] \\ (\gamma'') \end{pmatrix} \begin{pmatrix} (\Lambda) \\ [M'] \\ (\gamma') \end{pmatrix} \begin{pmatrix} [M] \\ (\gamma) \end{pmatrix} \right] \left[\begin{matrix} (\Gamma_j') \\ [M'] \\ (\gamma') \end{matrix} \right] \left[\begin{matrix} (\Gamma_i) \\ [M] \\ (\gamma) \end{matrix} \right] = \sum_{k=1}^{\mathcal{M}''} \left\{ \begin{pmatrix} [M]+[\Delta(\Lambda)] \\ (\Gamma_k'') \end{pmatrix} \begin{pmatrix} (\Lambda) \\ [M'] \\ [\Gamma_j'] \end{pmatrix} \begin{pmatrix} [M] \\ (\Gamma_i) \end{pmatrix} \right\} \left[\begin{matrix} (\Gamma_k'') \\ [M]+[\Delta(\Lambda)] \\ (\gamma'') \end{matrix} \right]. \tag{3.16}$$

In this coupling law, we now choose the operator pattern

$$(\gamma'') = \begin{pmatrix} \gamma_{12}'' & \gamma_{22}'' \\ & \gamma_{11}'' \end{pmatrix} \tag{3.17a}$$

such that

$$\gamma_{12}'' \geq \gamma_{11}'' \geq \gamma_{22}'' + M_{13} - M_{33} + M'_{13} - M'_{33} - i - j + 3. \tag{3.17b}$$

This choice of (γ'') has the effect of restricting the summation over (γ) and (γ') on the left-hand side of Eq. (3.16) such that *either*

$$\gamma_{12} \geq \gamma_{11} \geq \gamma_{22} + (M_{13} - M_{33} - i + 2) \tag{3.18a}$$

or

$$\gamma'_{12} \geq \gamma'_{11} \geq \gamma'_{22} + (M'_{13} - M'_{33} - j + 2). \tag{3.18b}$$

or both. [The left-hand sides of the inequalities stated in Eqs. (3.17b) and (3.18) are, of course, just the betweenness conditions, and could have been omitted, since they are always implicit.] Equations (3.18), in turn, imply that the left-hand side of Eq. (3.16) is zero for all (γ'') patterns satisfying Eq. (3.17b), this conclusion following from the fact that one or the other (or both) of the projective operators on the left-hand side is zero (the result just proven above).

The proof that the choice given by Eqs. (3.17) forces the summation patterns (γ) and (γ') into the intervals given by Eqs. (3.18a) and (3.18b), respectively, may be given in the following manner. Because $\gamma_{11} + \gamma'_{11} = \gamma''_{11}$, it follows from Eq. (3.17b) that $\gamma_{11} + \gamma'_{11} \geq \gamma''_{22} + M_{13} - M_{33} + M'_{13} - M'_{33} - i - j + 3$. Therefore, if we put

$$\gamma_{11} = \gamma_{22} + M_{13} - M_{33} - i + 2 + \lambda \tag{3.19a}$$

and

$$\gamma'_{11} = \gamma'_{22} + M'_{13} - M'_{33} - j + 2 + \lambda', \tag{3.19b}$$

then λ and λ' satisfy $\lambda + \lambda' \geq \gamma''_{22} - \gamma_{22} - \gamma'_{22} - 1$. However, in the square-bracket invariant appearing in the left-hand side in Eq. (3.16), it is necessary that $[\gamma''_{12}\ \gamma''_{22}]$ be contained in the U(2) direct product $[\gamma'_{12}\ \gamma'_{22}] \times [\gamma_{12}\ \gamma_{22}]$, and this condition requires $\gamma''_{22} \geq \gamma_{22} + \gamma'_{22}$. Thus, we find $\lambda + \lambda' \geq -1$, which implies that at least one of the integers λ, λ' is nonnegative. Hence, the choice of (γ'') given by Eq. (3.17b) forces at least one of the patterns (γ), (γ') into the range of values where the corresponding projective operator is zero.

We next note from the result for the zero projective operators that for a given lower pattern $\binom{\alpha}{\gamma}{}^{\beta}$ the operators which are zero are those of index k satisfying $k \geq (p - r + 2) - (\gamma - \beta)$. Thus, the summation on the right-hand side of Eq. (3.16) is over $k = 1, 2, \ldots, [M_{13} + \Delta_1(\Lambda)] - [M_{33} + \Delta_3(\Lambda)] + 1 - (\gamma''_{11} - \gamma''_{22})$. If we put

$$\gamma''_{11} - \gamma''_{22} = [M_{13} + \Delta_1(\Lambda)] - [M_{33} + \Delta_3(\Lambda)] - k_0 + 1, \tag{3.20a}$$

then as k_0 runs over the values

$$k_0 = 1, 2, \ldots, i + j - 2 - [M'_{13} - \Delta_1(\Lambda)] - [\Delta_3(\Lambda) - M'_{33}], \tag{3.20b}$$

all lexical values of γ''_{11}, γ''_{22} which satisfy Eq. (3.17b) are enumerated. [For each difference (3.20a) corresponding to the values k_0 given by (3.20b), there exists at least one value of γ''_{12} giving a lexical pattern—note that $k_0 = 1$ yields the largest possible difference $\gamma''_{11} - \gamma''_{22}$ and forces $\gamma''_{12} = M_{13} + \Delta_1(\Lambda)$.]

We have thus proved:

$$\sum_{k=1}^{k_0} \left\{ \begin{pmatrix} [M] + \Delta(\Lambda) \\ (\Gamma''_k) \end{pmatrix} \begin{pmatrix} (\Lambda) \\ [M'] \\ (\Gamma'_j) \end{pmatrix} \begin{pmatrix} [M] \\ (\Gamma_i) \end{pmatrix} \right\} \left[\begin{matrix} (\Gamma''_k) \\ [M] + [\Delta(\Lambda)] \\ (\gamma'') \end{matrix} \right] = 0 \tag{3.21}$$

for all k_0 in the range given by Eq. (3.20b), where the difference $\gamma''_{11} - \gamma''_{22}$ is given by Eq. (3.20a). Let us also note that it is necessary to have

$$i + j \geq [M'_{13} - \Delta_1(\Lambda)] - [\Delta_3(\Lambda) - M'_{33}] + 3 \tag{3.22}$$

in order that there be at least one term in Eq. (3.21).

In Eq. (3.21), we now set $\gamma''_{12} = M_{13} + \Delta_1(\Lambda)$, $\gamma''_{22} = M_{33} + \Delta_3(\Lambda)$, and $\gamma''_{11} = M_{13} + \Delta_1(\Lambda) - k + 1$, noting that the projective functions having these lower labels are nonzero[14] operators for each $k = 1, 2, \ldots, \mathcal{M}''$. We now choose $k_0 = 1, 2, \ldots$, in turn, to establish the result, Eqs. (3.15).

C. Interpretation

Let us summarize. By rather lengthy detailed arguments—which are technically straightforward—we have suceeded in parts A and B, above, in deriving the fact that a certain class of Racah functions has the value zero. The importance of this result for tensor operator structures is that it defines the *boundary for the nonvanishing Racah operators.* {We shall demonstrate subsequently that the Racah operators on the boundary—one step away from the zero operators, i.e., those for which

$$k = i + j - 1 - [M'_{13} - \Delta_1(\Lambda)] + [M'_{33} - \Delta_3(\Lambda)], \tag{3.23}$$

do not vanish identically.} It will be shown that the Racah invariants on the boundary having $\Lambda = (\max)$, are the ratios of denominator functions. Combined with the determination (in the second part of this paper) of all denominator functions, this will lead to a determination of a class of boundary Racah invariants.

It is helpful at this point to recall the analogous situation in regard to the U(2) Racah functions. For the U(2) Racah invariants, the boundary functions *were always monomials*; and these monomials were ratios of U(2) denominator functions (products of linear factors.) [Boundary U(2) Racah function have one or more of the four angular momentum triangles reduced to a line: $j_1 + j_2 = j_3$ is valid numerically and not just vectorially.] In particular, however, the fundamental U(2) Racah coefficient (spin $\frac{1}{2}$) has *only* boundary values; for this case, *knowledge of U(2) denominators is definitive.*

Knowledge of the fundamental Racah invariants suffices to determine (recursively) *all* Racah invariants. Thus—because of the simplicity that fundamental U(2) Racah functions are all on the boundary—the determination of all U(2) Racah functions is particularly simple.

The surprising difficulty underlying the U(3) Racah functions stems from the fact that not all fundamental Racah invariants lie on the boundary, in fact, almost all do not. It is this extra degree of freedom (which, it should be noted, *explicitly denies the possibility of a subgroup constraint*) that accounts for the difficulty inherent to the U(3) problem.

Nevertheless, our construction of the zero Racah invariants for U(3) does lead to important conclusions, as we illustrate in the next section.

4. THE FORM OF CERTAIN BOUNDARY RACAH INVARIANTS

Having discovered a class of zero Racah invariants, we next consider the class of Racah functions having indices i, j, k in the left-hand side of Eq. (3.15b) such that we are one unit off from the zero operator, i.e., those for which the indices i, j, k satisfy

$$k = i + j - 1 - [M'_{13} - \Delta_1(\Lambda)] + [M'_{33} - \Delta_3(\Lambda)]. \tag{4.1a}$$

We call the Racah invariants (3.15a) of this type *boundary Racah invariants.*

In Eq. (3.16), we choose the entries γ''_{11} and γ''_{22} in the (γ'') pattern such that

$$\gamma''_{11} - \gamma''_{22} = M_{13} - M_{33} + M'_{13} - M'_{33} - i - j + 2. \tag{4.1b}$$

In consequence of the zero projective operators (3.3) and the zero Racah operators (3.15), the right-hand side of Eq. (3.16) reduces to the *single term* having indices i, j, and k satisfying Eq. (4.1a). In the left-hand side, the choice of indices (4.1b) forces γ_{11} and γ'_{11} to the values

$$\gamma_{11} = \gamma_{22} + M_{13} - M_{33} - i + 1, \tag{4.1c}$$

$$\gamma'_{11} = \gamma'_{22} + M'_{13} - M'_{33} - j + 1. \tag{4.1d}$$

Thus, we obtain

$$\sum_{\substack{\gamma_{12},\gamma_{22}\\ \gamma'_{12},\gamma'_{22}}} \left[\begin{pmatrix} [M + [\Delta(\Lambda)]] \\ (\gamma'') \end{pmatrix} \begin{pmatrix} (\Lambda) \\ [M'] \\ (\gamma') \end{pmatrix} \begin{pmatrix} [M] \\ (\gamma) \end{pmatrix} \right] \left[\begin{matrix} (\Gamma'_j) \\ [M'] \\ (\gamma') \end{matrix} \right] \left[\begin{matrix} (\Gamma_i) \\ [M] \\ (\gamma) \end{matrix} \right]$$
$$= \left\{ \begin{pmatrix} [M] + [(\Lambda)] \\ (\Gamma''_k) \end{pmatrix} \begin{pmatrix} (\Lambda) \\ [M'] \\ (\Gamma'_j) \end{pmatrix} \begin{pmatrix} [M] \\ (\Gamma_i) \end{pmatrix} \right\} \left[\begin{matrix} (\Gamma''_k) \\ [M] + [\Delta(\Lambda)] \\ (\gamma'') \end{matrix} \right] \tag{4.2}$$

for the specific choice of labels given by Eqs. (4.1a)–(4.1d), where the summation on the left-hand side is over all lexical values such that $\gamma_{12} + \gamma'_{12} = \gamma''_{12}$ and $\gamma_{22} + \gamma'_{22} = \gamma''_{22}$. The choice of labels (4.1b) has the effect of isolating the Racah invariant operator having the *smallest* index k [for given i, j, $[M']$, (Λ)] for which the operator is not zero.

Equation (4.2) is a formula for obtaining particular Racah coefficients when certain projective coefficients are known.

Let us illustrate the use of Eq. (4.2) by specializing it still further. We take $(\Lambda) = (\max)$ and choose $\gamma''_{11} = \gamma''_{12} = M_{13} + M'_{13}$. This forces the left-hand side to the single term in which $\gamma_{11} = \gamma_{12} = M_{13}$ and $\gamma'_{11} = \gamma'_{12} = M'_{13}$. The equation reduces to

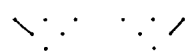
$$\begin{bmatrix} (\max) \\ M'_{13} \quad M'_{23} \quad M'_{33} \\ M'_{13} \quad \gamma'_{22} \\ M'_{13} \end{bmatrix} \begin{pmatrix} M_{13} & M_{23} & M_{33} \\ M_{13} & \gamma_{22} \end{pmatrix}$$
$$\times \begin{bmatrix} (\Gamma'_j) \\ M'_{13} \quad M'_{23} \quad M'_{33} \\ M'_{13} \quad \gamma'_{22} \\ M'_{13} \end{bmatrix} \begin{bmatrix} (\Gamma_i) \\ M_{13} \quad M_{23} \quad M_{33} \\ M_{13} \quad \gamma_{22} \\ M_{13} \end{bmatrix}$$
$$= \left\{ \begin{pmatrix} [M] + [M'] \\ (\Gamma''_{i+j-1}) \end{pmatrix} \begin{pmatrix} (\max) \\ [M'] \\ (\Gamma'_j) \end{pmatrix} \begin{pmatrix} [M] \\ (\Gamma_i) \end{pmatrix} \right\}$$
$$\times \begin{bmatrix} (\Gamma''_{i+j-1}) \\ M_{13} + M'_{13} M_{23} + M'_{23} M_{33} + M'_{33} \\ M_{13} + M'_{13} \qquad \gamma_{22} + \gamma'_{22} \\ M_{13} + M'_{13} \end{bmatrix}, \tag{4.3a}$$

in which

$$\gamma_{22} = M_{33} + i - 1, \quad \gamma'_{22} = M'_{33} + j - 1. \tag{4.3b}$$

It is useful at this point to discuss the explicit patterns to be associated with the null space ordering. We have assigned the label Γ_1 to that operator in a given multiplicity set, having largest null space. Since the null spaces are simply ordered, Γ_2 has the next largest null space, etc., down to $\Gamma_{\mathcal{M}}$, the operator having the minimal null space.

In Ref. 4II we have shown explicitly for the Γ_1 operator that the limit properties uniquely assign to this operator a *stretched pattern,* that is, a pattern tied in one or both of the following ways.



It is the great advantage of the step functions λ_i that these patterns can be given a comprehensive form. Letting $\Delta(\Gamma) = [\Delta_1\ \Delta_2\ \Delta_3]$, we find for the stretched Γ_1 pattern:

$$\begin{pmatrix} p & q & r \\ (\Gamma_1) \end{pmatrix} = \begin{pmatrix} p & q & r \\ \Delta_1 + \Delta_2 - \lambda_3 - r & & \lambda_3 + r \\ & \Delta_1 & \end{pmatrix}, \tag{4.4a}$$

where λ_3 is given by Eq. (2.14).

We extend this ordering now to the full multiplicity set, and assign to the operator labelled by Γ_k the explicit pattern

$$\begin{pmatrix} p & q & r \\ (\Gamma_k) \end{pmatrix} = \begin{pmatrix} p & q & r \\ \Delta_1 + \Delta_2 - \lambda_3 - k - r + 1 & & \lambda_3 + k + r - 1 \\ & \Delta_1 & \end{pmatrix}. \tag{4.4b}$$

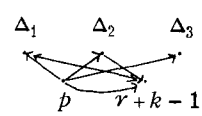
For preciseness, let us note that this step is not obvious, and requires proof. We anticipate this proof (which will be accomplished by examination of the explicitly constructed operators). It is useful to anticipate this result in order to unify the notation.

The structural content of Eq. (4.3) may be summarized by the statement: *Except for normalization, projective operators of the type*

$$\begin{bmatrix} (\Gamma_k) \\ p \quad q \qquad r \\ p \qquad r+k-1 \\ p \end{bmatrix} \tag{4.5}$$

multiply by addition of their corresponding patterns. While Eq. (4.3) is the proof of this result, it could have been anticipated from the properties of the arrow patterns associated with operators of the type (4.5).

To see this, consider the restricted arrow pattern for the operator (4.5).



(4.6)

In this arrow pattern there are $p - \Delta_i$ arrows going from point 1 in row 2 to point i in row 3, $\Delta_i - r - k + 1$ arrows going from point i in row 3 to point 2 in row 2, and $p - r - k + 1$ arrows going from point 1 to point 2 in row 2. (For $k = 1, 2, \ldots, \mathcal{M}$ it is always true that $\Delta_i \geq r + k - 1$, where $\mathcal{M}$ is the multiplicity of $[\Delta]$.)

We next observe that the arrow pattern retains the *same form* (4.6) for arbitrary $[p\ q\ r]$ (the number of

arrows changes, but not the direction). This, in turn, means there are no opposing arrows in the restricted arrow patterns for the two operators on the left-hand side of Eq. (4.3a). *Indeed, we may attribute the additive property of the two operator patterns to the geometrical property of no opposing arrows.*

We can extract still further information from Eq. (4.3a). Assume that the values of the operator (4.5) on arbitrary U(3) irrep labels $[m]=[m_{13}\ m_{23}\ m_{33}]$ and arbitrary U(2) irrep labels $[m']=[m_{12}\ m_{22}]$ are of the form

$$\begin{bmatrix} & (\Gamma_k) & \\ p & q & r \\ & p & r+k-1 \\ & & p \end{bmatrix}\begin{pmatrix}[m]\\[m']\end{pmatrix} = (-1)^{p-\Delta_1}\binom{q-r}{k-1}^{-1/2} \times F_R\begin{pmatrix}[\Delta_1 & \Delta_2 & \Delta_3]\\ [p & q & r]\\ [p & r+k-1]\end{pmatrix}\begin{pmatrix}[m]\\[m']\end{pmatrix}\Bigg/ D\begin{pmatrix}(\Gamma_k)\\ [p & q & r]\end{pmatrix}([m]), \quad (4.7)$$

where F_R denotes the restricted pattern calculus factor described in Ref. 4I, and denominator function

$$D\begin{pmatrix}(\Gamma_k)\\ [p & q & r]\end{pmatrix}([m]) \quad (4.8)$$

is, as yet, unknown, but will be given in the second paper.

We can give a straightforward constructive proof of Eq. (4.7), using Eq. (4.3). First of all, let us note that the projective function occurring in Eq. (4.3a) is a special case of a result given by Eq. (62) of Ref. 2. Its value is

$$\begin{bmatrix} & (\max) & \\ M'_{13} & M'_{23} & M'_{33} \\ & M'_{13} & \gamma'_{22} \\ & M'_{13} & \end{bmatrix}\begin{pmatrix} M_{13} & M_{23} & M_{33}\\ & M_{13} & \gamma_{22}\end{pmatrix} = \left[\frac{\binom{M'_{23}-M'_{33}}{j-1}\binom{M_{23}-M_{33}}{i-1}}{\binom{M_{23}+M'_{23}-M_{33}-M'_{33}}{i+j-1}}\right]^{1/2}. \quad (4.9)$$

We next particularize Eq. (4.3a) to the following case:

$$\begin{bmatrix} & 0 & \\ 0 & & 0 \\ 1 & 0 & -1 \\ & 1 & 0 \\ & & 1 \end{bmatrix}\binom{q-r-1}{i-2}^{1/2}\begin{bmatrix} & & (\Gamma_{i-1}) & & \\ p-1 & & q & & r+1 \\ & p-1 & & r+i-1 & \\ & & p-1 & & \end{bmatrix} = (\text{Racah invariant})\binom{q-r}{i-1}^{1/2}\begin{bmatrix} & (\Gamma_i) & \\ p & q & r \\ & p & r+i-1 \\ & & p \end{bmatrix}. \quad (4.10a)$$

The [10 − 1] projective function in this result is given by[1,3]

$$\begin{bmatrix} & 0 & \\ 0 & & 0 \\ 1 & 0 & -1 \\ & 1 & 0 \\ & & 1 \end{bmatrix} = -F_R\begin{pmatrix}[0 & 0 & 0]\\ [1 & 0 & -1]\\ & [1 & 0]\end{pmatrix}\Bigg/ D\begin{pmatrix} & 0 & \\ 0 & & 0 \\ 1 & 0 & -1\end{pmatrix} \quad (4.10b)$$

where

$$D\begin{pmatrix} & 0 & \\ 0 & & 0 \\ 1 & 0 & -1\end{pmatrix}([m]) = [\tfrac{1}{3}G_1(111;x)]^{1/2}, \quad (4.10c)$$

in which $G_1(111;x)$ is a special case of the functions studied in I and II of Ref. 4:

$$G_1(111;x) = x_1^2+x_2^2+x_3^2-6. \quad (4.10d)$$

In the next step, we iterate Eq. (4.10a) in p, r, and i, multiplying by the appropriate Racah invariant at each step of the iteration. The result is

$$\mathcal{Q}\begin{bmatrix} & (\Gamma_i) & \\ p & q & r \\ & p & r+i-1 \\ & & p \end{bmatrix} = \binom{q-r}{i-1}^{-1/2}\begin{bmatrix} & & 0 & & \\ & 0 & & 0 & \\ 1 & & 0 & & -1 \\ & 1 & & 0 & \\ & & 1 & & \end{bmatrix}^{i-1} \times\begin{bmatrix} & (\Gamma_1) & \\ p-i+1 & q & r+i-1 \\ & p-i+1 & r+i-1 \\ & & p-i+1 \end{bmatrix}, \quad (4.11)$$

where $\mathcal{Q}$ denotes a string of $i-1$ Racah invariants, and the relationship is valid for all $i=1,2,\ldots,\mathcal{N}$ (for $i=1$ both sides are identically the same). Observe that the (Γ_1) operator appearing in the right-hand side of Eq. (4.11) is the "stretched" case studied in I and II of Ref. 4. It has already been proved to have the form (4.7). Using this result and Eq. (4.11), and noticing that there are no opposing arrows in the restricted arrow patterns for the operators appearing in the right-hand side, we immediately establish the validity of the general form (4.7) for all $k=1,2,\ldots,\mathcal{N}$.

The phase in Eq. (4.7) is fixed by the phase of the operators appearing in the right-hand side of Eq. (4.11), since each of the Racah invariants appearing in $\mathcal{Q}$ is of the type

$$\left\{\begin{pmatrix}p & q & r\\ & (\Gamma_i) & \end{pmatrix}\begin{pmatrix} & (\max) & \\ 1 & 0 & -1 \\ 0 & & 0 \\ & 0 & \end{pmatrix}\begin{pmatrix}p-1 & q & r+1\\ & (\Gamma_{i-1}) & \end{pmatrix}\right\},$$

and has phase (+1). The phase of the right-hand side is given by $(-1)^{i-1}(-1)^{p-i+1-\Delta_1}=(-1)^{p-\Delta_1}$.

We may now go one step further and substitute the structural form (4.7) back into Eq. (4.3a) to obtain the following form for the boundary Racah functions having $\Lambda=(\max)$:

$$\left\{\begin{pmatrix}[M]+[M']\\ (\Gamma''_{i+j-1})\end{pmatrix}\begin{pmatrix}(\max)\\ [M']\\ [\Gamma'_j]\end{pmatrix}\begin{pmatrix}[M]\\ (\Gamma_i)\end{pmatrix}\right\}([m]+[\Delta'']) = \frac{D\begin{pmatrix}(\Gamma''_{i+j-1})\\ [M]+[M']\end{pmatrix}([m])}{\left[D\begin{pmatrix}(\Gamma'_j)\\ [M']\end{pmatrix}D\begin{pmatrix}(\Gamma_i)\\ ([M])\end{pmatrix}\right]([m])}. \quad (4.12)$$

Equations (4.7) and (4.12) generalize, in form, the results given in I of Ref. 4 for "stretched" patterns. We emphasize, however, that the typical denominator function appearing in Eqs. (4.7) and (4.12) are as yet unknown—except for the stretched case $k=1$. In the next section, using a special technique, we will determine yet another one of these denominator functions, leaving to the second paper the difficult task of determining the general denominator function.

Equation (4.2) has a number of other interesting specializations, leading to a variety of results, all explicit except for the unknown denominator functions [e.g., choose $\Lambda=(\max)$, $\gamma''_{12}=M_{13}+M'_{13}$ and $\gamma''_{22}=M_{33}+M'_{33}$]. We will not present these results, since the most important problem at hand is the determination of the denominator functions themselves as illustrated by Eqs. (4.7) and (4.12).

5. THE DENOMINATOR FUNCTION FOR THE MINIMAL NULL SPACE TENSOR OPERATORS

In I of Ref. 4, we developed special methods for evaluating the "stretched" (maximal null space) denominator function, which we denote now by the more explicit functional notation

$$D\begin{pmatrix}(\Gamma_1)\\ [M]\end{pmatrix}([m]). \tag{5.1a}$$

{The symbols express the fact that the denominator function D indexed by the stretched operator pattern Γ_1 of the tensor operator with irrep labels $[M]$ has the U(3) Young pattern $[m]$ as its argument.} In this section, we will give a special technique for obtaining the *minimal* null space denominator function, denoted by

$$D\begin{pmatrix}(\Gamma_{/\!/})\\ [M]\end{pmatrix}([m]). \tag{5.1b}$$

[Here the $\Gamma_{/\!/}$ denotes the minimal null space operator pattern $\Gamma_{/\!/}$ associated with the multiplicity set having the common shift $\Delta(\Gamma_{/\!/})=\Delta(\Gamma_1)=(\Delta_1\ \Delta_2\ \Delta_3)$.]

The technique we will use for this determination is one of systematically generalizing special cases which are relatively easy to calculate. A completely different —and far more comprehensive (but difficult)—technique will be used in the second part of the present paper, to obtain the general denominator function, but we feel that alternative techniques are of interest in their own right, as helpful illustrations of the interconnections existing in the subject.

Let us begin then by considering the special case of self-conjugate tensor operators; these have the irrep labels $[2k\ k\ 0]$ in SU(3), or more suggestively in U(3), $[k\ 0\ -k]$. [This is less of a special case than it appears to be at first glance, since, in a very real sense, knowledge of all self-conjugate operators in SU(3) would be tantamount to knowledge of all SU(3) tensor operators—but the necessary constructions are not yet fully explicit.]

Consider then the minimal null space tensor operator, $\Gamma_{/\!/}=(_0{}^0{}_0)$ $(/\!/=k+1)$,

$$\left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ \alpha & & \beta \\ & \gamma & \end{matrix} \right\rangle \tag{5.2}$$

having $\Delta(\Gamma_{k+1})=[0\ 0\ 0]$. For $k=1$, these operators are the U(3) octet of generators (normalized); for $k=2$, they are the 27-plet of operators calculated explicitly in Ref. 11.

Observe that because of the canonical splitting conditions the projective operators obey the result [cf. Eqs. (3.13) and (3.14)]

$$\left[\begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ k & & -k \\ & \gamma & \end{matrix} \right] \neq 0 \ \text{only for } \gamma=0. \tag{5.3}$$

This fact enables us to "promote" the projective operator to a genuine Wigner operator, i.e.,

$$\left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ k & & -k \\ & \gamma & \end{matrix} \right\rangle = \left[\begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ k & & -k \\ & 0 & \end{matrix} \right] \left\langle \begin{matrix} 0 \\ k \\ \gamma \end{matrix} \begin{matrix} \\ -k \\ \end{matrix} \right\rangle, \tag{5.4}$$

since the sum becomes, by (5.3), but a single term.

Furthermore, we have the following very simple coupling law:

$$\left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ 1 & 0 & -1 \\ 1 & & -1 \\ & 1 & \end{matrix} \right\rangle \left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ k & & -k \\ & k & \end{matrix} \right\rangle = \left\{ \left(\begin{matrix} k+1 & 0 & -k-1 \\ & 0\ \ 0 & \\ & 0 & \end{matrix} \right) \left(\begin{matrix} & 1 & \\ 1 & & 0 \\ 1 & 0 & -1 \\ 0 & & 0 \\ & 0 & \end{matrix} \right) \left(\begin{matrix} k & 0 & -k \\ & 0\ \ 0 & \\ & 0 & \end{matrix} \right) \right\} \left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ k+1 & 0 & -k-1 \\ & k+1 & \end{matrix} \right\rangle. \tag{5.5}$$

This result is an immediate consequence of the properties of the Racah invariants for minimal null space patterns developed in Sec. 3. Since the Racah operator appearing in Eq. (5.5) is a U(3) invariant, we may iterate Eq. (5.5) to obtain the formal result:

$$\left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ k & & -k \\ & k & \end{matrix} \right\rangle = [\text{U(3) invariant}] \times \left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ 1 & 0 & -1 \\ 1 & & -1 \\ & 1 & \end{matrix} \right\rangle^{k}, \tag{5.6}$$

where the superscript k denotes the k-fold product of the operator shown. Although the U(3) invariant operator appearing in this result is, at this stage, unknown, Eq. (5.6) clearly does have structural content.

We know from Refs. 1 and 3 that the particular operator appearing in Eq. (5.6) is a normalized generator and has the alternative form:

$$\left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ 1 & 0 & -1 \\ & 1 & -1 \\ & & 1 \end{matrix} \right\rangle = I_2^{-1/2} E_{12}, \tag{5.7}$$

where I_2 is a known second rank invariant, and E_{12} is Weyl's form for this generator in U(3). Combining Eqs. (5.6) and (5.7), we obtain the equivalent form

$$\left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ & k & -k \\ & & k \end{matrix} \right\rangle = F_k^{-1} (E_{12})^k, \tag{5.8}$$

where F_k is U(3) invariant operator, to be determined yet.

One may give a boson operator realization for these (abstract) equations which may be helpful in appreciating intuitively the content of these results. We have the familiar realization:

$$\left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ 1 & 0 & -1 \\ & 1 & -1 \\ & & 1 \end{matrix} \right\rangle \rightarrow I_2^{-1/2} \sum_{i=1}^{3} a_1^i \bar{a}_2^i, \tag{5.7*}$$

which then shows that Eq. (5.8) takes the form:

$$\left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ & k & -k \\ & & k \end{matrix} \right\rangle = [\text{function of U(3) invariants only}] \times \left(\sum_{i=1}^{3} a_1^i \bar{a}_2^i \right)^k. \tag{5.8*}$$

The result expressed in Eq. (5.8) [(5.8*)] is useful in that it allows one to determine the operators

$$\left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ \alpha & & \beta \\ & \gamma & \end{matrix} \right\rangle \tag{5.9}$$

for all labels α, β, γ by using the commutation relations with the generators, subject only to an (unknown) overall normalization. It follows that one method for obtaining the invariant factor F_k appearing in Eq. (5.8) is to generate the operators (5.9) from(5.8) and then to normalize:

$$\sum_{\alpha,\beta,\gamma} \left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ \alpha & & \beta \\ & \gamma & \end{matrix} \right\rangle \left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ \alpha & & \beta \\ & \gamma & \end{matrix} \right\rangle^{\dagger} = I_0, \tag{5.10}$$

where $I_0([m]) = 0$ if $[m]$ belongs to the null space of the operator (5.9), and otherwise $I_0([m]) = 1$. Although straightforward in principle, this procedure is extremely laborious in practice.

There exists a second, more practical, method for obtaining the factor F_k, and hence the desired denominator function. This is a technique used frequently in quantum mechanics prior to the development of angular momentum theory—*the method of multiplet averaging.* Let us apply this method to the operator

$$\left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ & k & -k \\ & & 0 \end{matrix} \right\rangle, \tag{5.11}$$

where, for greatest convenience, we have chosen $\gamma = 0$ so that the operator is *diagonal.* [This point can be demonstrated by the fact that k commutations of $E_{21} = J_-$ with Eq. (5.8), noting that $E_{12} = J_+$, show that the operator (5.11) is a polynomial in $J^2 = J_z(J_z + 1) + J_- J_+$ and $J_z = (E_{11} - E_{22})/2$.]

The multiplet averaging method gives the relation

$$\sum_{(m)} \left\langle \begin{pmatrix} [m] \\ (m) \end{pmatrix} \middle| \left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ & k & -k \\ & & 0 \end{matrix} \right\rangle \middle| \begin{pmatrix} [m] \\ (m) \end{pmatrix} \right\rangle^2 = \frac{\text{dimension of multiplet}}{\text{dimension of operator}} I_0([m]) = [-x_1 x_2 x_3 / 2(k+1)^3] I_0([m]). \tag{5.12}$$

The validity of this relation is established in Appendix B.

[We can now recognize the importance of the realization given by Eq. (5.8/5.8*). Were one to introduce Eq. (5.4) into the multiplet sum, Eq. (5.12), the sum over m_{11} would eliminate the U(2) Wigner operator, leaving a sum (over the projective operator) *which could not be carried out* (with information at hand). In sharp contrast to Eq. (5.4), the result in Eq. (5.8/5.8*) implies *precisely the additional* (m_{12}, m_{22}) *information necessary to carry out the complete multiplet sum.* These points are illustrated explicitly below.]

Substitution of Eq. (5.4) (for $\gamma = 0$) into the multiplet sum formula (5.12) yields

$$\sum_{m_{12}, m_{22}} (m_{12} - m_{22} + 1) \left\langle \begin{pmatrix} m_{13} m_{23} m_{33} \\ m_{12} m_{22} \end{pmatrix} \middle| \left[\begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ & k & -k \\ & & 0 \end{matrix} \right] \middle| \begin{pmatrix} m_{13} m_{23} m_{33} \\ m_{12} m_{22} \end{pmatrix} \right\rangle^2 = [-x_1 x_2 x_3 (2k+1) / 2(k+1)^3] I_0([m]), \tag{5.13a}$$

where we have utilized the result

$$\sum_{m_{11}} \left\langle \begin{pmatrix} m_{12} m_{22} \\ m_{11} \end{pmatrix} \middle| \left\langle \begin{matrix} 0 & \\ k & -k \\ & 0 \end{matrix} \right\rangle \middle| \begin{pmatrix} m_{12} m_{22} \\ m_{11} \end{pmatrix} \right\rangle^2 = (m_{12} - m_{22} + 1)/(2k+1). \tag{5.13b}$$

We now proceed in the following manner: From Eqs. (5.4) and (5.8) (set $\gamma = k$), we have

$$\left[\begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ & k & -k \\ & & 0 \end{matrix} \right] \left\langle \begin{matrix} 0 & \\ k & -k \\ & k \end{matrix} \right\rangle = F_k^{-1} (E_{12})^k. \tag{5.14}$$

Taking the (nonzero) matrix element of this result yields:

$$\left\langle \begin{pmatrix} m_{13}m_{23}m_{33} \\ m_{12}m_{22} \end{pmatrix} \middle| \begin{bmatrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ & k & -k \\ & 0 & \end{bmatrix} \middle| \begin{pmatrix} m_{13}m_{23}m_{33} \\ m_{12}m_{22} \end{pmatrix} \right\rangle = (-1)^k F_k^{-1}([m]) \left[\frac{k!k!(2k+1)}{(m_{12}-m_{22}+1)} \begin{pmatrix} m_{12}-m_{22}+k+1 \\ 2k+1 \end{pmatrix} \right]^{1/2}$$

$$= (-1)^k F_k^{-1}([m]) \left[\frac{k!k!}{(2k)!} \prod_{s=1}^{k} [(p_{12}-p_{22})^2 - s^2] \right]^{1/2}. \tag{5.15}$$

We now utilize this result in Eq. (5.13a) to obtain

$$F_k^{-2}([m]) \sum_{m_{12},m_{22}} \begin{pmatrix} m_{12}-m_{22}+k+1 \\ 2k+1 \end{pmatrix}$$

$$= [-x_1x_2x_3/2(k+1)(k+1)!(k+1)!]I_0([m]), \tag{5.16}$$

where we note that the summation over m_{12}, m_{22} is such that $m_{13} \geq m_{12} \geq m_{23} \geq m_{22} \geq m_{33}$. Noting that

$$\sum_{m_{12},m_{22}} \begin{pmatrix} m_{12}-m_{22}+k+1 \\ 2k+1 \end{pmatrix} = -\sum_{i=1}^{3} \begin{pmatrix} x_i+k+1 \\ 2k+3 \end{pmatrix}, \tag{5.17}$$

we obtain the following expression for $F_k^2([m])$[15]:

$$F_k^2([m]) = \frac{2(k+1)(k+1)!(k+1)!}{x_1x_2x_3} \sum_{i=1}^{3} \begin{pmatrix} x_i+k+1 \\ 2k+3 \end{pmatrix}, \tag{5.18}$$

where we again recall that $x_1 = p_{23} - p_{33} = m_{23} - m_{33} + 1$, $x_2 = p_{33} - p_{13} = m_{33} - m_{13} - 2$, $x_3 = p_{13} - p_{23} = m_{13} - m_{23} + 1$.

It is easily established that x_1, x_2, and x_3 are factors in the sum appearing in Eq. (5.18). Hence, $F_k^2([m])$ *is a polynomial of degree* $2k$.

Moreover, because of the complete symmetry in the x_i, the function F_k^{-2} is hexagonally symmetric in the Möbius plane and has a center of symmetry at the origin.

What are the zeroes of the polynomial defined by Eq. (5.18)? Let us first show that $x_1 = 1$ yields a line containing $2k$ zeroes. Introducing $x_1 = 1$ into Eq. (5.18), and removing $x_2 = -1 - x_3$, yields the form

$$F_k^2([m]) = \frac{2k![(k+1)]^3}{(2k+2)!} \begin{pmatrix} x_3+k+1 \\ k \end{pmatrix} \begin{pmatrix} x_3-1 \\ k \end{pmatrix}. \tag{5.19}$$

It is clear from this form that the polynomial F_k^2 has k zeroes along the line $x_1 = 1$ at $x_3 = 1, 2, \ldots, k$ and also k zeroes along the line $x_1 = 1$ at $x_2 = 1, 2, \ldots, k$.

Similarly, by direct substitution into the definition (5.18), it is easily established that: *The polynomial* $F_k^2([m])$ *vanishes at each lattice point of the equilateral triangle whose* $(x_1x_2x_3)$ *vertices are* $(1, -2, 1)$, $(k, -1-k, 1)$ *and* $(1, -1-k, k)$.

From the symmetry in x_i, it then follows that there are precisely *six* such triangles of zeroes symmetrically placed about the center of symmetry at the origin.

The results we have obtained are definitive. *For it follows from the definition and properties of the denominator function* $G_q(\Xi; x)$, *as discussed in Sec.* 2, *that the properties established for* $F_k^2([m])$ *uniquely determine* $F_k^2([m])$ *to be a multiple of the* G_q *function.* (These defining properties are degree of polynomial, symmetry, and location of triangles of zeroes.) The relative normalization is determined from special cases and the explicit result is found to be

$$F_k^2([m]) = \frac{k+1}{(2k+1)!} G_k \begin{pmatrix} -1 & -1+x_1 & -1-x_1 \\ -1 & -1+x_2 & -1-x_2 \\ -1 & -1+x_3 & -1-x_3 \end{pmatrix}. \tag{5.20}$$

The projective function (5.15) is now given explicitly by

$$\left\langle \begin{pmatrix} m_{13}m_{23}m_{33} \\ m_{12}m_{22} \end{pmatrix} \middle| \begin{bmatrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ & k & -k \\ & 0 & \end{bmatrix} \middle| \begin{pmatrix} m_{13}m_{23}m_{33} \\ m_{12}m_{22} \end{pmatrix} \right\rangle$$

$$= (-1)^k k! \left[\frac{(2k+1)\prod_{s=1}^{k}[(p_{12}-p_{22})^2 - s^2]}{(k+1)G_k} \right]^{1/2}, \tag{5.21}$$

where the G_k in this result has the arguments displayed in Eq. (5.20).

It is worth noting that these results agree in complete detail with similar results given earlier for the octet and 27-plet operators.

It is interesting to observe that Eq. (5.21) has the following structural form:

$$\left\langle \begin{pmatrix} m_{13}m_{23}m_{33} \\ m_{12}m_{22} \end{pmatrix} \middle| \begin{bmatrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ & k & -k \\ & 0 & \end{bmatrix} \middle| \begin{pmatrix} m_{13}m_{23}m_{33} \\ m_{12}m_{22} \end{pmatrix} \right\rangle = (-1)^k D \begin{pmatrix} 0 \\ k \quad -k \end{pmatrix} (m_{12}m_{22}) \Big/ D \begin{pmatrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \end{pmatrix} (m_{13}m_{23}m_{33}). \tag{5.22}$$

The numerator in this result is the U(2) denominator function given by the pattern-on-pattern calculus rules[41]:

$$D\begin{pmatrix} 0 \\ k & -k \end{pmatrix}(m_{12}m_{22}) \equiv D\begin{pmatrix} k \\ 2k & 0 \end{pmatrix}(m_{12}m_{22})$$
$$\equiv D\begin{pmatrix} [k & k] \\ [2k & 0] \end{pmatrix}(m_{12}m_{22}) = \left[\frac{k!k!}{2k!}\prod_{s=1}^{k}[(p_{12}-p_{22})^2 - s^2]\right]^{1/2}, \tag{5.23}$$

and the U(3) denominator function is given by Eq. (5.20)[16]:

$$D\begin{pmatrix} 0 \\ 0 & 0 \\ k & 0 & -k \end{pmatrix}([m]) = \left[\frac{k+1}{(2k+1)!}\, G_k\begin{pmatrix} -1 & -1+x_1 & -1-x_1 \\ -1 & -1+x_2 & -1-x_2 \\ -1 & -1+x_2 & -1-x_3 \end{pmatrix}\right]^{1/2} \tag{5.24}$$

We now also know from Eq. (4.12) the explicit form of the Racah function appearing in Eq. (5.5):

$$\left\{\begin{pmatrix} k+1 & 0 & -k-1 \\ & 0 & 0 \\ & & 0 \end{pmatrix}\begin{pmatrix} & 1 & \\ 1 & & 0 \\ 1 & 0 & -1 \\ & 0 & \\ & & 0 \end{pmatrix}\begin{pmatrix} k & 0 & -k \\ & 0 & 0 \\ & & 0 \end{pmatrix}\right\}([m]) = \frac{D\begin{pmatrix} 0 \\ 0 & 0 \\ k+1 & 0 & -k-1 \end{pmatrix}([m])}{D\begin{pmatrix} 0 \\ 0 & 0 \\ 1 & 0 & -1 \end{pmatrix}([m])\; D\begin{pmatrix} 0 \\ 0 & 0 \\ k & 0 & -k \end{pmatrix}([m])}. \tag{5.25}$$

In view of Eqs. (5.22) and (5.25), it is revealing now to go back and examine the structure of the projective operator coupling and examine the structure of the projective operator coupling law corresponding to Eq. (5.5). It reads

$$\left\{\begin{pmatrix} k+1 & & -k-1 \\ & 0 & \end{pmatrix}\begin{pmatrix} & 1 & \\ 1 & & -1 \\ & 0 & \end{pmatrix}\begin{pmatrix} k & & -k \\ & 0 & \end{pmatrix}\right\}\begin{bmatrix} & 0 & \\ 0 & & 0 \\ 1 & 0 & -1 \\ & 0 & \end{bmatrix}\begin{bmatrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ & 0 & \end{bmatrix}$$
$$= \left\{\begin{pmatrix} k+1 & 0 & -k-1 \\ & 0 & 0 \\ & & 0 \end{pmatrix}\begin{pmatrix} & 1 & \\ 1 & & 0 \\ 1 & 0 & -1 \\ & 0 & \\ & & 0 \end{pmatrix}\begin{pmatrix} k & 0 & -k \\ & 0 & 0 \\ & & 0 \end{pmatrix}\right\}\begin{bmatrix} & 0 & \\ 0 & & 0 \\ k+1 & 0 & -k-1 \\ & 0 & \end{bmatrix}. \tag{5.26}$$

The U(2) Racah function appearing in this result has the values

$$\left\{\begin{pmatrix} k+1 & & -k-1 \\ & 0 & \end{pmatrix}\begin{pmatrix} & 1 & \\ 1 & & -1 \\ & 0 & \end{pmatrix}\begin{pmatrix} k & & -k \\ & 0 & \end{pmatrix}\right\}(m_{12}m_{22}) = \frac{D\begin{pmatrix} 0 \\ k+1 & -k-1 \end{pmatrix}(m_{12}m_{22})}{D\begin{pmatrix} 0 \\ 1 & -1 \end{pmatrix}(m_{12}m_{22})\; D\begin{pmatrix} 0 \\ k & -k \end{pmatrix}(m_{12}m_{22})}. \tag{5.27}$$

Observe how when we substitute Eqs. (5.25) and (5.27) into Eq. (5.26) the denominator functions simply combine with the projective functions (5.22) to form unity, so that the coupling law is trivially satisfied.

There is one further property we wish to note—the limit $m_{33} \to -\infty$ of the Racah functions (5.25). It has been conjectured[3] that a Racah function always limits to a square-bracket function, and this property has held in each instance where it has been tested.[3,4II] The present example also validates this limit property:

$$\lim_{m_{33}\to-\infty}\left\{\begin{pmatrix} k+1 & 0 & -k-1 \\ & 0 & 0 \\ & & 0 \end{pmatrix}\begin{pmatrix} & 1 & \\ 1 & & 0 \\ 1 & 0 & -1 \\ & 0 & \\ & & 0 \end{pmatrix}\begin{pmatrix} k & 0 & -k \\ & 0 & 0 \\ & & 0 \end{pmatrix}\right\}(m_{13}m_{23}m_{33})$$
$$= \left[\begin{pmatrix} k+1 & 0 & -k-1 \\ & 0 & 0 \\ & & 0 \end{pmatrix}\begin{pmatrix} & 1 & \\ 1 & & 0 \\ 1 & 0 & -1 \\ & 0 & \\ & & 0 \end{pmatrix}\begin{pmatrix} k & 0 & -k \\ & 0 & 0 \\ & & 0 \end{pmatrix}\right](m_{13}m_{23}) = [3(k+2)(k+2)/2(2k+2)(2k+3)]^{1/2}. \tag{5.28}$$

This result is proved in Appendix A, where we also demonstrate how this limit property is used to assign the operator pattern $\binom{0}{0}$ to the symbol Γ_{k+1}—a result we have assumed, for convenience, throughout this paper.

While the results given in Eq. (5.21) do indeed suffice to give the desired denominator function of the special class of self-conjugate $\Gamma_{/\!/}$ operators, it would appear that we are very far from the class of *all* $\Gamma_{/\!/}$ operators. If one accepts the idea that Eq. (5.20) establishes the *form* of the desired general answer, then we can immediately

generalize the G_q function that appears here by taking the arguments of the function to be in "general position." Explicitly, this is the assertion that for the general denominator the G_k function in Eq. (5.21) is replaced by

$$G_k \to G_{q-\lambda_1-\lambda_2-\lambda_3}\begin{pmatrix} -\Delta_1-2\lambda_1+q-1 & -\Delta_2-\lambda_2-\lambda_3+q-1-x_1 & -\Delta_3-\lambda_2-\lambda_3+q-1+x_1 \\ -\Delta_2-2\lambda_2+q-1 & -\Delta_3-\lambda_3-\lambda_1+q-1-x_2 & -\Delta_1-\lambda_3-\lambda_1+q-1+x_2 \\ -\Delta_3-2\lambda_3+q-1 & -\Delta_1-\lambda_1-\lambda_2+q-1-x_3 & -\Delta_2-\lambda_1-\lambda_2+q-1+x_3 \end{pmatrix}. \tag{5.29}$$

The fact that Eq. (5.21) has no linear factors appearing in the denominator can be seen to be a property restricted to the special class of operators considered in this equation. For the general $\Gamma_{/\!/}$ operator, we may reason this way. For the maximal null space case, Γ_1, we know in detail *that the linear factors are read off the reduced Ξ-pattern by the rule stated (and discussed) in Sec.* 2.

Next let us recall that it is an abstract general property of the null space diagram that the null space of Γ_t differs[17] from that of Γ_{t+1} by *precisely those six lines which form the boundary of the null space of* Γ_{t+1}. This (abstractly proven) fact enables us to determine the linear factors in Γ_t *from a knowledge of the linear factors in* Γ_1. In particular, we may state now the linear factors (multiplying the G-function) in the general $\Gamma_{/\!/}$ case: *The linear factors for* $\Gamma_{/\!/}$ *are read off the reduced Ξ-pattern (by the rules of Sec* 2) *after each entry in the reduced Ξ-pattern is decreased by the integer* -1. [One easily verifies that, for the special case represented by Eq. (5.20), there are no linear factors.]

We shall not attempt to validate the argument placing G_k in general position, since this properly belongs to the systematic discussion of the general Γ_t denominators (which will be given in the next paper). It has been our object in the present paper to show, by simpler methods, that the functional form $G_q(\Xi; x)$, found for the stretched denominator function Γ_1, also suffices for the denominator function for the minimal null space $\Gamma_{/\!/}$ denominator as well. This is as far as one can go along this line, since the general Γ_t denominator will require generalization of the G_q function itself.

ACKNOWLEDGMENTS

We wish to express our gratitude to Dr. E. de Vries and Dr. A. J. van Zanten for assisting with some of the preliminary calculations of the denominator functions.

APPENDIX A

The purpose of this appendix is to demonstrate how specific operator patterns are assigned to the symbols $\Gamma_1, \Gamma_2, \ldots, \Gamma_{/\!/}$ by the use of limit properties.[18]

This limit procedure has already been used in Ref. 4II [cf. Sec. 4D] to assign generally the "stretched" operator pattern Γ_1 given by Eq. (4.4a). It is useful to recall this procedure to illustrate the usefulness of the λ-notation [cf. Eq. (2.14)] in unifying results. In terms of the λ-notation, Eqs. (5.6a,b) of Ref. 4II are written as a single result:

$$G_q(\Delta; x) \approx \binom{q}{\lambda_3}[\Delta_1]_{\lambda_3}[\Delta_2]_{\lambda_3}[\Delta_3]_{q-\lambda_3}[\Delta_1+\Delta_2-2\lambda_3]_{q-\lambda_3}[x_3-\Delta_2-\Delta_3+q-1]_{\lambda_3}[x_3+\Delta_1]_{\lambda_3}(-m_{33})^{2q-2\lambda_3}, \tag{A1}$$

where $\approx$ symbolizes the asymptotic form of $G_q(\Delta; x)$ for large $-m_{33}$.

Similarly, Eqs. (5.7a,b) and (5.8a,b) are unified to the single result:

$$\lim_{m_{33}\to-\infty}\left\{\left([p \quad q \quad 0]_{(\Gamma_s)}\right)\left([p-q \quad \overset{(\max)}{0} \quad 0]_{(\Gamma')}\right)\left([q \quad q \quad 0]_{(\Gamma'')}\right)\right\}([m]+[\Delta])$$

$$=\left[\begin{pmatrix} p & & q & 0 \\ \Delta_1+\Delta_2-\lambda_3 & & \lambda_3 & \\ & \Delta_1 & & \end{pmatrix}\left([p-q \quad \overset{(\max)}{0} \quad 0]_{(\Gamma')}\right)\left([q \quad q \quad 0]_{(\Gamma'')}\right)\right](m_{13}+\Delta_1,\ m_{23}+\Delta_2)$$

$$=\delta_{q-\lambda_3,\Delta_3''}\left[\frac{(\Delta_1-\lambda_3)!(\Delta_2-\lambda_3)!(p-q-\Delta_3')!p!(q-\lambda_3)!}{(\Delta_1')!(\Delta_2')!(q-\Delta_1'')!(q-\Delta_2'')!(p-\lambda_3)!(\Delta_1+\Delta_2-\lambda_3)!}\right.$$
$$\left.\times\frac{(x_{12}+\Delta_1''-\Delta_2'')(x_{12}+\Delta_1''-\Delta_2-1)!(x_{12}+\Delta_1''-q-1)!(x_{12}+\Delta_1-\lambda_3)!}{(x_{12}+\Delta_1-\Delta_2'')!(x_{12}+q-\Delta_2'')!(x_{12}-\Delta_2+\lambda_3-1)!}\right]^{1/2}. \tag{A2}$$

It is this relation which is used to establish the identification, Eq. (4.4a), of the stretched pattern (now denoted by Γ_1).

Let us now illustrate how this limit property is used to assign the minimal null space operator pattern $\Gamma_{\not\eta} = \Gamma_{k+1}$ to be

$$\begin{pmatrix} k & 0 & -k \\ & (\Gamma_{k+1}) & \end{pmatrix} = \begin{pmatrix} k & 0 & -k \\ & 0 & \\ & 0 & \end{pmatrix}. \tag{A3}$$

If we had not anticipated this answer, the Racah coefficient appearing in the left-hand side of Eq. (5.28) would have been written

$$\left\{ \begin{pmatrix} k+1 & 0 & -k-1 \\ & (\Gamma_{k+2}) & \end{pmatrix} \begin{pmatrix} & 1 & \\ 1 & & 0 \\ 1 & 0 & -1 \\ 0 & & 0 \\ & 0 & \end{pmatrix} \begin{pmatrix} k & 0 & -k \\ & (\Gamma_{k+1}) & \end{pmatrix} \right\}, \tag{A4}$$

in which Γ_{k+1} and Γ_{k+2} appear as symbols specifying minimal null space operators, but as yet unidentified with numerical triangular patterns [the $\binom{0}{0}$ pattern in the middle is already known from Ref. 3 to be the assignment induced by limits].

The square-bracket function appearing in Eq. (5.28) is also known (from Ref. 3), and it has the value indicated in Eq. (5.28). The idea now is to prove directly from Eq. (5.25) [We now replace the Racah operator by the one given by Eq. (A4), having unassigned patterns] that the limit of the Racah function (A4) is indeed the numerical value given in Eq. (5.28).

It is straightforward to determine from Eqs. (5.24), (5.20), and (5.18) that

$$D\begin{pmatrix} (\Gamma_{k+1}) \\ k \quad 0 \quad -k \end{pmatrix}([m]) \approx \left[\frac{(k+1)!(k+1)!}{(2k+1)!}\right]^{1/2} (-m_{33})^k \tag{A5}$$

for large $-m_{33}$ [note that we would be using Γ_{k+1} in the left-hand side of Eq. (5.24) had we not anticipated the assignment (A3)]. By direct substitution of (A5) into the right-hand side of Eq. (5.25), we find that the limit $m_{33} \to -\infty$ of the Racah function (A4) is indeed the value given in Eq. (5.28), i.e., we thus prove:

$$\lim_{m_{33} \to -\infty} \left\{ \begin{pmatrix} k+1 & 0 & -k-1 \\ & (\Gamma_{k+2}) & \end{pmatrix} \begin{pmatrix} & 1 & \\ 1 & & 0 \\ 1 & 0 & -1 \\ 0 & & 0 \\ & 0 & \end{pmatrix} \begin{pmatrix} k & 0 & -k \\ & (\Gamma_{k+1}) & \end{pmatrix} \right\} ([m])$$

$$= \left[\begin{pmatrix} k+1 & 0 & -k-1 \\ & 0 \quad 0 & \\ & 0 & \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ & 0 \quad 0 & \\ & 0 & \end{pmatrix} \begin{pmatrix} k & 0 & -k \\ & 0 \quad 0 & \\ & 0 & \end{pmatrix} \right] (m_{13} m_{23}). \tag{A6}$$

We emphasize that the patterns appearing in the square-bracket invariant are not in question—these patterns are already understood as U(2) patterns from the definition[3] of a square-bracket invariant. We thus induce uniquely the pattern assignment (A3), using the same limit concept which identified the stretched pattern [and which eventually will be used to establish the general assignment, Eq. (4.4b)].

APPENDIX B

The purpose of this appendix is to prove the multiplet averaging formula, Eq. (5.12).

The starting place for the proof is the product law [cf. Eq. (2.43) of Ref. 3] for Wigner operators, the following product being a special application of the general law:

$$\left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ k & & -k \\ & 0 & \end{matrix} \right\rangle \left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ k & & -k \\ & 0 & \end{matrix} \right\rangle = \sum_{(\Gamma)} \sum_{(M)} \sum_{(\Lambda)} \left\langle \begin{pmatrix} [k \quad 0 \quad -k] + [\Delta(\Lambda)] \\ (M) \end{pmatrix} \middle| \left\langle \begin{matrix} & (\Lambda) & \\ k & 0 & -k \\ k & & -k \\ & 0 & \end{matrix} \right\rangle \middle| \left\langle \begin{matrix} k & 0 & -k \\ & k & -k \\ & 0 & \end{matrix} \right\rangle \right\rangle \quad \times$$

$$\times \left\{ \left(\begin{matrix} [k\ 0 & -k] + [\Delta(\Lambda)] \\ & (\Gamma) \end{matrix} \right) \left(\begin{matrix} (\Lambda) \\ k\ 0\ -k \\ 0\ 0 \\ 0 \end{matrix} \right) \left(\begin{matrix} k & 0 & -k \\ 0 & 0 \\ & 0 \end{matrix} \right) \right\} \left\langle \left(\begin{matrix} [k\ 0 & -k] + [\Delta(\Lambda)] \\ & (\Gamma) \end{matrix} \right) \right\rangle, \tag{B1}$$

where we note that for each lexical pattern (Λ) the sum is over those lexical patterns (Γ) such that

$$\Delta \left(\begin{matrix} [k\ 0 & -k] + \Delta(\Lambda) \\ & (\Gamma) \end{matrix} \right) = [0\,0\,0]. \tag{B2}$$

(B2)

We now take the diagonal matrix element of Eq. (B1) between Gel'fand states having labels given by Eq. (3.4a), summing over all lexical patterns

$$(m) = \begin{pmatrix} m_{12} & & m_{22} \\ & m_{11} & \end{pmatrix}. \tag{B3}$$

We next observe that the trace of any tensor operator mapping of the irrep space $[m_{13}\, m_{23}\, m_{33}]$ into itself (any tensor operator having $\Delta = [0\,0\,0]$) is zero unless the operator is the identity mapping, i.e., unless $[k\ 0\ -k] + \Delta(\Lambda) = [0\,0\,0]$ and $(M) = (\Gamma) = \left(\begin{smallmatrix} 0 & \\ & 0 \end{smallmatrix}\right)$ in the summation on the right-hand side of Eq. (B1). For this term, the summation over the U(2) labels (B3) gives the dimension, $D([m])$, of the irrep $[m]$. Thus, we obtain

$$\sum_{(m)} \left\langle \left(\begin{matrix} [m] \\ (m) \end{matrix} \right) \middle| \left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ & k\ -k & \\ & 0 & \end{matrix} \right\rangle \middle| \left(\begin{matrix} [m] \\ (m) \end{matrix} \right) \right\rangle^2 = \left\langle \left(\begin{matrix} 0 & 0 & 0 \\ & 0 & 0 \\ & & 0 \end{matrix} \right) \middle| \left\langle \begin{matrix} & -k & \\ 0 & & -k \\ k & 0 & -k \\ & 0\ 0 & \\ & 0 & \end{matrix} \right\rangle \middle| \left\langle \begin{matrix} k & 0 & -k \\ & k & -k \\ & & 0 \end{matrix} \right\rangle \right\rangle$$

$$\times \left\{ \left(\begin{matrix} 0 & 0 & 0 \\ & 0 & 0 \\ & & 0 \end{matrix} \right) \left(\begin{matrix} & -k & \\ 0 & & -k \\ k & 0 & -k \\ & 0\ 0 & \\ & 0 & \end{matrix} \right) \left(\begin{matrix} k & 0 & -k \\ & 0 & 0 \\ & & 0 \end{matrix} \right) \right\} ([m])\, D([m]). \tag{B4}$$

It remains to evaluate the Wigner and Racah coefficients in Eq. (B4). The Wigner coefficient may be evaluated from the pattern calculus rules, and its value is $[D(k\ 0\ -k)]^{-1/2}$. The Racah coefficient has the value $I_0([m])\,[D(k\ 0\ -k)]^{-1/2}$ (this result is proved below), so that the right-hand side of Eq. (B4) is

$$[D([m])/D(k\ 0\ -k)]\, I_0([m]), \tag{B5}$$

as stated in the multiplet sum rule, Eq. (5.12).

We still must prove that the Racah coefficient occurring in Eq. (B4) has the value $I_0([m])\,[D(k\ 0\ -k)]^{-1/2}$. One method of showing this result is to appeal directly to the definition of Racah operators in terms of Wigner operators, given by Eq. (2.46) of Ref. 3. For the case at hand, we obtain

$$\left\{ \left(\begin{matrix} 0 & 0 & 0 \\ & 0 & 0 \\ & & 0 \end{matrix} \right) \left(\begin{matrix} & -k & \\ 0 & & -k \\ k & 0 & -k \\ & 0\ 0 & \\ & 0 & \end{matrix} \right) \left(\begin{matrix} k & 0 & -k \\ & 0 & 0 \\ & & 0 \end{matrix} \right) \right\} = \sum_{(\bar{M})} \left\langle \left(\begin{matrix} 0 & 0 & 0 \\ & 0 & 0 \\ & & 0 \end{matrix} \right) \middle| \left\langle \begin{matrix} & -k & \\ 0 & & -k \\ k & 0 & -k \\ & (M) & \end{matrix} \right\rangle \middle| \left\langle \begin{matrix} k & 0 & -k \\ & (\bar{M}) & \end{matrix} \right\rangle \right\rangle \left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ & (M) & \end{matrix} \right\rangle \left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ & (\bar{M}) & \end{matrix} \right\rangle. \tag{B6}$$

The right-hand side of this result may now be simplified by using the Hermitian conjugation properties expressed by

$$\left\langle \begin{matrix} & -k & \\ 0 & & -k \\ k & 0 & -k \\ & (M) & \end{matrix} \right\rangle^{\dagger} = (-1)^{\phi(M)} D^{-1/2} \left\langle \begin{matrix} & k & \\ k & & 0 \\ k & 0 & -k \\ & (\bar{M}) & \end{matrix} \right\rangle D^{1/2}, \tag{B7a}$$

$$\left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ & (M) & \end{matrix} \right\rangle^{\dagger} = (-1)^{\phi(M)} \left\langle \begin{matrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \\ & (\bar{M}) & \end{matrix} \right\rangle, \tag{B7b}$$

where $\phi(M) = M_{12} + M_{22} + M_{11}$ and D is the dimension operator. The first result, Eq. (B7a), may be proved by induction, using the proven relation for the fundamental Wigner operators [cf. Eq. (4.7) of Ref. 3]; the second result Eq. (B7b), may be proved from the generator realization of these tensor operators [cf. Eq. (5.8)]. We omit the details of these proofs.

Equation (B7a) is now used to evaluate the Wigner coefficient appearing in Eq. (B6):

$$\left\langle \begin{pmatrix} 0 & 0 & 0 \\ & 0 & 0 \\ & & 0 \end{pmatrix} \middle| \begin{pmatrix} & -k & \\ 0 & & -k \\ k & 0 & -k \end{pmatrix}_{(M)} \middle| \begin{pmatrix} k & 0 & -k \\ & (\bar{M}) & \end{pmatrix} \right\rangle = \frac{(-1)^{\phi(M)}}{[D(k\ 0\ -k)]^{1/2}} \left\langle \begin{pmatrix} k & 0 & -k \\ & (\bar{M}) & \end{pmatrix} \middle| \begin{pmatrix} k & 0 & -k \\ & (\bar{M}) & \end{pmatrix} \right\rangle = (-1)^{\phi(M)} [D(k\ 0\ -k)]^{-1/2}. \quad \text{(B8)}$$

Operating on the state vector

$$\left| \begin{matrix} [m] \\ (m) \end{matrix} \right\rangle$$

with Eq. (B6), using at the same time the results of Eqs. (B7b) and (B8), we obtain

$$\left\{ \begin{pmatrix} 0 & 0 & 0 \\ & 0 & 0 \\ & & 0 \end{pmatrix} \begin{pmatrix} & -k & \\ 0 & & -k \\ k & 0 & -k \end{pmatrix}_{(M)} \begin{pmatrix} k & 0 & -k \\ & 0 & 0 \\ & & 0 \end{pmatrix} \right\} ([m]) = [D(k\ 0\ -k)]^{-1/2} \left(\sum_{(M)} \left\langle \begin{pmatrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \end{pmatrix}_{(M)} \left\langle \begin{pmatrix} & 0 & \\ 0 & & 0 \\ k & 0 & -k \end{pmatrix}_{(M)} \right\rangle^{\dagger} \right) \left| \begin{matrix} [m] \\ (m) \end{matrix} \right\rangle$$

$$= [D(k\ 0\ -k)]^{-1/2} I_0([m]). \quad \text{(B9)}$$

Note added in proof: It is the *norm* of an operator which determines its null space and not its *normalization*, the latter being the significance of a denominator function. However, in the canonical splitting, the norm N_Γ and the denominator D_Γ are related by $N_\Gamma([m]) = D_\Gamma([m])\,[D([m]+[\Delta])/D([m])]^{1/2}$, where D is the dimension operator. Thus, in making statements about the null space, we ignore the factor $[D'([m]+[\Delta])]^{-1/2}$ occurring in Eq. (1.2) and the factor $[D'([m])]^{1/2}$ which occurs as a factor in the product over $i<j$. Since the factor $D'([m]) = D([m])/1!2!\cdots(n-1)!$ does not vanish in the lexical region ($x_1 \geq 1$, $x_2 \leq -2$, $x_3 \geq 1$) of the intertwining number-null space diagram, this statement is technically correct. We have added this remark to avoid any possible confusion on this point.

*Work performed under the auspices of the USERDA.
†Supported in part by the NSF under Grant GP-14116.
‡Supported in part by an Overseas Scholarship from the Royal Commission for the Exhibition of 1851.

[1]Early work on the subject includes: J. J. DeSwart, Rev. Mod. Phys. **35**, 916 (1963); T. A. Brody, M. Moshinsky, and I. Renero, J. Math. Phys. **6**, 1540 (1965); G. E. Baird and L. C. Biedenharn, J. Math. Phys. **6**, 1847 (1964).
[2]L. C. Biedenharn and J. D. Louck, Commun. Math. Phys. **8**, 89 (1968).
[3]J. D. Louck and L. C. Biedenharn, J. Math. Phys. **2**, 2368 (1970).
[4]L. C. Biedenharn, J. D. Louck, E. Chacón, and M. Cifton, J. Math. Phys. **13**, 1957 (1972) (I); L. C. Biedenharn and J. D. Louck, J. Math. Phys. **13**, 1985 (1972) (II); J. D. Louck and L. C. Biedenharn, J. Math. Phys. **14**, 1336 (1973) (III).
[5]J. P. Draayer and Y. Akiyama, J. Math. Phys. **14**, 1904 (1973).
[6]L. C. Biedenharn, A. Giovannini, and J. D. Louck, J. Math. Phys. **8**, 691 (1967).
[7]G. E. Baird and L. C. Biedenharn, J. Math. Phys. **5**, 1730 (1964).
[8]E. Chacón, M. Cifton, and L. C. Biedenharn, J. Math. Phys. **13**, 577 (1972); S. J. Alisauskas, A.-A. A. Jucys, and A. P. Jucys, J. Math. Phys. **13**, 1349 (1972).
[9]These operators are themselves constructed from the two elementary operators. General results for U(n) are known, in complete detail, for all elementary operators[2] and all adjoint operators.[3]
[10]The $[\Delta_j + x_i]_{\lambda_i}[\Delta_k - x_i]_{\lambda_i}$ factors in this result were already noted in Ref. 4II by examining the eight possible cases. The numerical factors may be obtained by several techniques, one being to compare the asymptotic forms of each side of Eq. (2.15) [cf. Appendix A and Eq. (2.13d)].
[11]J. A. Castilho-Alcarás, L. C. Biedenharn, K. T. Hecht, and G. Neely, Ann. Phys. (N.Y.) **60**, 85 (1970).
[12]Draayer and Akiyama (Ref. 5) have noted some of the zeroes which occur in the projective operators and the Racah functions. The fact that all such zeroes are derivable directly from the basic splitting conditions, Eqs. (3.1), was not noted, however.
[13]It has been emphasized by Hecht [K. T. Hecht, J. Math. Phys. **15**, 2148 (1974)] that in applications to physical problems Racah coefficients are often more useful than Wigner coefficients, and one of our goals is to obtain explicit algebraic expressions for all U(3) Racah functions.
[14]This fact may be proved from the construction given in Sec. 4F of Ref. 4I.
[15]The expression $F_R^2([m])$ vanishes (see the discussion which follows) at the same values of (m_{13}, m_{23}, m_{33}) as does $I_0([m])$, so that it is not necessary to keep the factor $I_0([m])$ in Eq. (5.18).
[16]This identification of the denominator function could be off by factors depending on k. However, the subsequent relations Eqs. (5.25)—(5.28), show that the identification made is correct.
[17]We are now using Γ_t, $t = 1, 2, \ldots, \mathcal{M}$, to denote the operator patterns $\Gamma_1, \ldots, \Gamma_{\mathcal{M}}$, since the index k previously used has been given a different meaning in this section.
[18]See Ref. 3 for the notations used in this appendix.

Group theory and propagation of operator averages

C. Quesne*

Physique Théorique et Mathématique, Université Libre de Bruxelles, Brussels, Belgium
(Received 17 June 1975)

The propagation of operator averages, which is the basis of French's spectral distribution method, is reformulated in the framework of group theory. The concept of complementary groups is extensively used. It is shown that the possibility of propagating averages is intimately connected with the absence of state labeling problem. The construction of the propagation operators is examined, and for those cases where it is not trivial, a new way of approach is suggested by establishing a link with recent group theoretical advance in the construction of subgroup invariants in the universal covering algebra of a group. Finally the discussion is illustrated by some examples taken, or not, from current literature.

1. INTRODUCTION

During the last few years, the spectral distribution method mainly developed by French and coworkers[1–3] has proved quite useful in nuclear spectroscopy. Its use has been based on the possibility of propagating the operator averages, defined in given subspaces of the fixed number of particles spaces, from low to high values of the number of particles (or holes). Although it has been obvious from the beginning that this procedure was intimately connected with group theory, the precise relationship has not been clearly understood up to now.

In this paper we shall reformulate the average propagation in the framework of group theory. For this purpose, the concept of complementary groups, introduced previously by Moshinsky and Quesne,[4] turns out to be useful and for this reason is briefly reviewed in Sec. 2.

Section 3 is devoted to the discussion of the average propagation in the light of the concepts just introduced. We show the existence of close connections between some difficulties encountered in the propagation process and classical problems of group theory, such as those coming from the use of noncanonical chains of subgroups. In particular we stress the usefulness of the construction of an integrity basis for the subgroup invariants in the enveloping algebra of a group when there is a state labeling problem in the group to subgroup reduction. This establishes a link between an interesting group theoretical problem, whose solution was formulated recently in general terms,[5] and some physical applications.

Finally, in Sec. 4, we illustrate the general discussion by some examples taken from the current physical literature or corresponding to new averaging processes.

2. NONINVARIANCE GROUPS AND COMPLEMENTARY SUBGROUPS

In the second quantized picture, fermions are represented by creation and annihilation operators, b^+_α and b_α respectively, where α denotes all the quantum numbers specifying the one-particle states. In the case of nucleons in a central field, α stands for $nljm$ or nlm_lm_s for identical nucleons, and $nljmm_t$ or $nlm_lm_sm_t$ for neutrons and protons in the isospin formalism. All n-particle states can be written as linear combinations of the states

$$|\alpha_1\alpha_2\cdots\alpha_n\rangle = b^+_{\alpha_1}b^+_{\alpha_2}\cdots b^+_{\alpha_n}|0\rangle,\quad \alpha_1<\alpha_2<\cdots<\alpha_n, \tag{2.1}$$

where $|0\rangle$ is the vacuum state.

When the number of one-particle states is finite and equal to $\mathcal{N}$ (so that $0\leq n\leq\mathcal{N}$), it is well known[4,6] that the largest noninvariance group that can be built is the group $U(2\mathcal{N})$, the generators of whose Lie algebra $\mathbf{C}^{\alpha'_1\alpha'_2\cdots\alpha'_{n'}}_{\alpha_1\alpha_2\cdots\alpha_n}$ can transform any state (2.1) into any other one,

$$\mathbf{C}^{\alpha'_1\alpha'_2\cdots\alpha'_{n'}}_{\alpha_1\alpha_2\cdots\alpha_n}|\alpha''_1\alpha''_2\cdots\alpha''_{n''}\rangle$$

$$=\delta_{n'n''}\delta_{\alpha_1\alpha''_1}\cdots\delta_{\alpha'_{n'}\alpha''_{n'}}|\alpha_1\alpha_2\cdots\alpha_n\rangle. \tag{2.2}$$

This group was called supergroup by Judd.[6] It is clear from Eq. (2.2) that all n-particle states, $0\leq n\leq\mathcal{N}$, belong to the irreducible representation (IR) [1] of $U(2^{\mathcal{N}})$.

The supergroup contains a smaller noninvariance group $O^*(2\mathcal{N}+1)$, the generators of whose Lie algebra are the one-particle creation and annihilation operators, and their commutators.[7] All n-particle states, $0\leq n\leq\mathcal{N}$, belong to the single IR $[\frac{1}{2}\mathcal{N}]$ of this group. The group $O^*(2\mathcal{N}+1)$ contains all the symmetry groups, the unitary group $U(\mathcal{N})$, and its subgroups.

In general we are concerned with a chain of groups

$$\begin{array}{ccccccc} U(2^{\mathcal{N}}) & \supset & O^*(2\mathcal{N}+1) & \supset & U(\mathcal{N}) & \supset & G, \\ [1] & & [\frac{1}{2}\mathcal{N}] & & [1^n] & & \lambda \end{array} \tag{2.3}$$

below each one of which we give the corresponding IR to which the states belong. The n-particle states can be written as

$$|n\varphi\lambda\mu\rangle = P^+_{n\varphi\lambda\mu}|0\rangle, \tag{2.4}$$

where μ characterizes the row of the IR λ of G, φ distinguishes between the various equivalent IR's λ of G contained in the IR $[1^n]$ of $U(\mathcal{N})$, and $P^+_{n\varphi\lambda\mu}$ is an appropriate linear combination of products of n creation operators b^+_α. The generators of $U(2^{\mathcal{N}})$ can be rewritten in the basis (2.4) as $\mathbf{C}^{n'\varphi'\lambda'\mu'}_{n\varphi\lambda\mu}$ and are characterized by the property

$$\mathbf{C}^{n'\varphi'\lambda'\mu'}_{n\varphi\lambda\mu}|n''\varphi''\lambda''\mu''\rangle = \delta_{n'n''}\delta_{\varphi'\varphi''}\delta_{\lambda'\lambda''}\delta_{\mu'\mu''}|n\varphi\lambda\mu\rangle. \tag{2.5}$$

They are rather complicated linear combinations of the operators

$$\mathbf{D}^{n'\varphi'\lambda'\mu'}_{n\varphi\lambda\mu} = P^+_{n\varphi\lambda\mu}P^{n'\varphi'\lambda'\mu'}, \tag{2.6}$$

which can be inverted to give the latter in terms of the former.[4] Here

$$P^{n\varphi\lambda\mu}=[P^{+}_{n\varphi\lambda\mu}]^{+}. \tag{2.7}$$

The noninvariance groups $U(2\mathcal{N})$ and $O^{+}(2\mathcal{N}+1)$ are especially useful when discussing the concept of complementary groups.[4] Two groups G and G^{C}, whose direct product is a subgroup of a larger group H, are referred to as "complementary" within a definite IR of H, if there is a one-to-one correspondence between all the IR's of G and G^{C} contained in this IR of H. All Casimir operators of one of the subgroups depend linearly on the Casimir operators of the other one in this IR of H.[8] In the following, we shall take for H either noninvariance group, whose IR is unique, so that we shall not mention it anymore and shall speak of "complementary groups within a group H."

In Ref. 4, it is shown that any subgroup G of $U(\mathcal{N})$, including itself, does have a complementary group within the supergroup and that its structure can be obtained easily. We shall call it the trivial complementary group of G and denote it by G^{TC}:

$$U(2\mathcal{N})\supset G\times G^{TC}. \tag{2.8}$$

Its generators are obtained from the generators of $U(2\mathcal{N})$ which are scalar with respect to G, and are thus

$$C^{n'\varphi'\lambda}_{n\varphi\lambda}=\sum_{\mu}\mathbf{C}^{n'\varphi'\lambda\mu}_{n\varphi\lambda\mu}. \tag{2.9}$$

G^{TC} is a direct sum of unitary groups:

$$G^{TC}=\sum_{\lambda}\oplus U(d(\lambda)), \tag{2.10}$$

each of which is associated with a given IR λ of G and has a dimension equal to the number of times $d(\lambda)$ that the IR λ of G is contained in the IR [1] of the supergroup. All the states belonging to the various equivalent IR's λ of G belong to the IR [1] of the group $U(d(\lambda))$ and to the IR's [0] of the groups $U(d(\lambda'))$ with $\lambda'\neq\lambda$.

On the contrary all the subgroups G of $U(\mathcal{N})$ do not have a complementary group within the noninvariance group $O^{+}(2\mathcal{N}+1)$. When a complementary group does exist, it is not trivial at all, and will be denoted by G^{C}:

$$O^{+}(2\mathcal{N}+1)\supset G\times G^{C}. \tag{2.11}$$

The generators of its Lie algebra are obtained from those of $O^{+}(2\mathcal{N}+1)$, which are scalar with respect to G.

Known examples of complementary groups are those associated with unitary,[9] orthogonal,[10] or symplectic[11] subgroups of $U(\mathcal{N})$. The corresponding chains of groups are

$$\begin{array}{llll} O^{+}(2\mathcal{N}+1)\supset & U(q)\times U(r), & \mathcal{N}=qr, & \\ [\tfrac{1}{2}\mathcal{N}] & \lambda \quad \tilde{\lambda} & & \end{array} \tag{2.12}$$

$$\begin{array}{llll} O^{+}(2\mathcal{N}+1)\supset & O^{+}(2q+1)\times O(2r), & \mathcal{N}=(2q+1)r, & \\ [\tfrac{1}{2}\mathcal{N}] & \lambda \quad \lambda'' & & \end{array} \tag{2.13}$$

and

$$\begin{array}{llll} O^{+}(2\mathcal{N}+1)\supset & Sp(2q)\times Sp(2r), & \mathcal{N}=2qr. & \\ [\tfrac{1}{2}\mathcal{N}] & \lambda \quad \lambda' & & \end{array} \tag{2.14}$$

Below each group we give the IR's to which the states belong. For the unitary groups, λ denotes a partition of n $(0\leq n\leq\mathcal{N})$ into q integers not exceeding r, and $\tilde{\lambda}$ its conjugate partition. For the orthogonal and symplectic groups, λ denotes a partition into q integers not exceeding r, λ' is the partition into r integers not exceeding q, whose conjugate partition is such that $\tilde{\lambda}'_i=r-\lambda_{q+1-i}$, and λ'' is the set of r half-integers defined by $\lambda''_i=\lambda'_i+\frac{1}{2}$.

Whenever the complementary group G^{C} does exist, the generators of the Lie algebra of G^{TC} can be expressed as polynomials in the generators of the Lie algebra of G^{C}, i.e., the algebra of G^{TC} is a (finite-dimensional) representation of the universal enveloping algebra[12] of G^{C}, built in terms of fermion creation and annihilation operators. Indeed the generators of G^{TC} carry any state transforming according to a definite IR of G into zero or a state transforming irreducibly in the same way. But all these states belong to the same IR of G^{C}, so that the transformation of any one of them into any other one can always be produced by successive applications of generators of G^{C}. There is no such simple prescription for the generators of G^{TC} when G^{C} does not exist, and in this case their explicit building is generally quite complicated.

A special type of trivial complementary group will play an important role in the following section. It is associated with a subgroup $\mathcal{G}$ of $U(\mathcal{N})$ defined as follows: If G is a subgroup of $SU(\mathcal{N})$, then

$$\mathcal{G}=U(1)\times G, \tag{2.15a}$$

the single generator of U(1) being the number operator N; if G is not a subgroup of $SU(\mathcal{N})$, i.e., if N belongs to its Lie algebra, then

$$\mathcal{G}=G. \tag{2.15b}$$

The IR's of $\mathcal{G}$ are characterized by $\Lambda=(n,\lambda)$. The algebra of

$$\mathcal{G}^{TC}=\sum_{\Lambda}\oplus U(d(\Lambda)) \tag{2.16}$$

is generated by the operators

$$C^{\varphi'\Lambda}_{\varphi\Lambda}=\sum_{\mu}\mathbf{C}^{n\varphi'\lambda\mu}_{n\varphi\lambda\mu}, \tag{2.17}$$

which are linear combinations of the operators

$$D^{\varphi'\Lambda}_{\varphi\Lambda}=\sum_{\mu}\mathbf{D}^{n\varphi'\lambda\mu}_{n\varphi\lambda\mu}. \tag{2.18}$$

It is made of all the scalars with respect to G which leave the number of particles invariant, i.e., the polynomials in the generators of $U(\mathcal{N})$ which are scalar with respect to G. Therefore, the algebra of $\mathcal{G}^{TC}$ is nothing less than a (finite-dimensional) representation of the subalgebra of scalars with respect to G of the universal enveloping algebra of $U(\mathcal{N})$, which representation is built in terms of fermion creation and annihilation operators. It contains itself as a subalgebra a representation of the set of all polynomials in the Casimir operators of $\mathcal{G}$, i.e., polynomials in N [Casimir operator of $U(\mathcal{N})$], and in the Casimir operators of G.

3. PROPAGATION OF OPERATOR AVERAGES

Let us consider the average of a k-body operator,

$$O(k)=\sum_{\substack{\varphi\lambda\mu\\ \varphi'\lambda'\mu'}}\langle k\varphi\lambda\mu|O(k)|k\varphi'\lambda'\mu'\rangle\,\mathbf{D}^{k\varphi'\lambda'\mu'}_{k\varphi\lambda\mu}, \tag{3.1}$$

in a subspace of dimension $\dim(\lambda)$ of the n-particle space, defined as the representation space of an IR λ of a subgroup G of $U(\mathcal{N})$, and specified by additional quantum numbers φ if necessary,

$$\langle O(k)\rangle^{n\varphi\lambda}=[\dim(\lambda)]^{-1}\sum_{\mu}\langle n\varphi\lambda\mu|O(k)|n\varphi\lambda\mu\rangle. \tag{3.2a}$$

Instead of characterizing the subspace in which the average is calculated by some IR's of the chain (2.3) of subgroups of $O^*(2\mathcal{N}+1)$, it will be more useful in the following discussion to specify it by some IR Λ of the subgroup $\mathcal{G}$ of $O^*(2\mathcal{N}+1)$, defined in relation (2.15). As $\dim(\Lambda)=\dim(\lambda)$, relation (3.2a) becomes

$$\langle O(k)\rangle^{\varphi\Lambda}=[\dim(\Lambda)]^{-1}\sum_{\mu}\langle\varphi\Lambda\mu|O(k)|\varphi\Lambda\mu\rangle, \tag{3.2b}$$

where φ distinguishes between the various equivalent IR's Λ of $\mathcal{G}$ when necessary. As it will appear in the examples of Sec. 4, this way of writing the average gives a unified description of all the cases treated up to now.

Introducing expansion (3.1) into relation (3.2b), we get

$$\langle O(k)\rangle^{\varphi\Lambda}=\sum_{\substack{\varphi'\Lambda'\mu'\\ \varphi''\Lambda''\mu''}}\{\langle\varphi'\Lambda'\mu'|O(k)|\varphi''\Lambda''\mu''\rangle \times[\dim(\Lambda)]^{-1}\sum_{\mu}\langle\varphi\Lambda\mu|\mathbf{D}^{\varphi''\Lambda''\mu''}_{\varphi'\Lambda'\mu'}|\varphi\Lambda\mu\rangle\}, \tag{3.3}$$

where the summation over Λ' and Λ'' is restricted to $\Lambda'=(k,\lambda')$, and $\Lambda''=(k,\lambda'')$. Only the part of the operator $\mathbf{D}^{\varphi''\Lambda''\mu''}_{\varphi'\Lambda'\mu'}$ which is scalar with respect to $\mathcal{G}$ can contribute to the sum over μ and its matrix elements are independent of μ, so that

$$[\dim(\Lambda)]^{-1}\sum_{\mu}\langle\varphi\Lambda\mu|\mathbf{D}^{\varphi''\Lambda''\mu''}_{\varphi'\Lambda'\mu'}|\varphi\Lambda\mu\rangle =\delta_{\Lambda'\Lambda''}\delta_{\mu'\mu''}\langle\varphi\Lambda|[\dim(\Lambda')]^{-1}\sum_{\mu'}\mathbf{D}^{\varphi''\Lambda'\mu'}_{\varphi'\Lambda'\mu'}|\varphi\Lambda\rangle. \tag{3.4}$$

Relation (3.3) becomes

$$\langle O(k)\rangle^{\varphi\Lambda}=\sum_{\varphi'\varphi''\Lambda'}\{\langle\varphi\Lambda|D^{\varphi''\Lambda'}_{\varphi'\Lambda'}|\varphi\Lambda\rangle \times[\dim(\Lambda')]^{-1}\sum_{\mu'}\langle\varphi'\Lambda'\mu'|O(k)|\varphi''\Lambda'\mu'\rangle\}, \tag{3.5}$$

where the operator $D^{\varphi''\Lambda'}_{\varphi'\Lambda'}$ is defined by Eq. (2.18).

Following French,[2] we say that the average of $O(k)$ can be propagated from its defining subspaces $\varphi'\Lambda'$, $\Lambda'=(k,\lambda')$, if it can be expressed, for any φ and Λ, as a linear combination of the averages of $O(k)$ in its defining subspaces. From relation (3.5), it is clear that a necessary and sufficient condition for this to happen whatever $O(k)$ may be is that

$$\langle\varphi\Lambda|D^{\varphi''\Lambda'}_{\varphi'\Lambda'}|\varphi\Lambda\rangle =\delta_{\varphi'\varphi''}\langle\varphi\Lambda|D^{\varphi'\Lambda'}_{\varphi'\Lambda'}|\varphi\Lambda\rangle, \tag{3.6}$$

for any $\varphi,\varphi',\varphi''$, Λ and Λ' such that $\Lambda'=(k,\lambda')$, because then

$$\langle O(k)\rangle^{\varphi\Lambda}=\sum_{\varphi'\Lambda'}\langle\varphi\Lambda|D^{\varphi'\Lambda'}_{\varphi'\Lambda'}|\varphi\Lambda\rangle\langle O(k)\rangle^{\varphi'\Lambda'}. \tag{3.7}$$

In this case, it is possible to define an operator

$$\hat{O}(k)=\sum_{\varphi'\Lambda'}\langle O(k)\rangle^{\varphi'\Lambda'}D^{\varphi'\Lambda'}_{\varphi'\Lambda'}, \tag{3.8}$$

which is trace equivalent to $O(k)$, i.e., such that

$$\langle\hat{O}(k)\rangle^{\varphi\Lambda}=\langle O(k)\rangle^{\varphi\Lambda} \tag{3.9}$$

for any φ and Λ. The operators $D^{\varphi'\Lambda'}_{\varphi'\Lambda'}$, of which it is a linear combination, are called propagation operators. The calculation of average (3.9) in any subspace $\varphi\Lambda$ then reduces to the construction of the propagation operators, and the determination of their diagonal matrix elements.

At this point there arise two main questions. The first one is to know in which cases the propagation is possible, i.e., conditions (3.6) are satisfied. The second one is to give prescriptions for building the propagation operators when these do exist. We now proceed to examine both problems successively.

Conditions (3.6) are trivially satisfied if there is no need for additional quantum numbers φ. It is thus obvious that a sufficient condition for the propagation to be possible for any k is that any IR Λ' of $\mathcal{G}$ be contained at most once in the IR $[\frac{1}{2}\mathcal{N}]$ of $O^*(2\mathcal{N}+1)$, or that any IR λ' of G be contained at most once in the IR $[1^n]$ of $U(\mathcal{N})$. If this is not the case, and that additional quantum numbers are necessary for some value of n, the propagation is still possible for a given value of k if any IR $\Lambda'=(k,\lambda')$ of $\mathcal{G}$ is contained at most once in $[\frac{1}{2}\mathcal{N}]$. However, if for a given k and a given λ', the IR $\Lambda'=(k,\lambda')$ is contained at least twice in $[\frac{1}{2}\mathcal{N}]$, conditions (3.6) will generally not be fulfilled because a state $|n\varphi\lambda\mu\rangle$ may have simultaneously parents of the type $|k\varphi'\lambda'\mu'\rangle$ and $|k\varphi''\lambda'\mu'\rangle$ (some simple examples of this property can be easily constructed). We thus see coming out for the first time connections between the propagation procedure and the state labeling problem. Other relations will appear later on.

When the propagation is possible, how can the propagation operators be constructed? The operators $D^{\varphi'\Lambda'}_{\varphi'\Lambda'}$ are linear combinations of the generators $C^{\varphi''\Lambda'}_{\varphi'\Lambda'}$ of the trivial complementary group $\mathcal{G}^{TC}$ of $\mathcal{G}$. It is thus clear that the study of the structure of this group is intimately connected with the construction of the propagation operators. This implies that a procedure has to be found to build the generators of $\mathcal{G}^{TC}$ in a direct way, without constructing explicitly the n-particle states.

When the algebra of $\mathcal{G}^{TC}$ coincides with its subalgebra of polynomials in the Casimir operators of $\mathcal{G}$, the problem is easily solved because the Casimir operators are well known for all classical groups, as well as their eigenvalues in all IR's. Sufficient conditions for this to happen can be found:

(i) $\mathcal{G}$ is the first step of a canonical chain of subgroups of $U(\mathcal{N})$ [therefore $\mathcal{G}=U(\mathcal{N}-1)$], because then the

Casimir operators of the groups of the chain form a complete set of commuting operators[13];

(ii) $\mathcal{G}$ is the direct product of two complementary groups, $\mathcal{G}=\bar{G}\times\bar{G}^{C}$, because then the generators of $\bar{G}^{TC}$ are polynomials in the generators of $\bar{G}^{C}$, and $\mathcal{G}^{TC}$ is the subgroup of $\bar{G}^{TC}$ which is left invariant by the transformations of $\bar{G}^{C}$, so that its generators are polynomials in the Casimir operators of $\bar{G}^{C}$ (or of $\bar{G}$ as the latter can be expressed in terms of the former);

(iii) $\mathcal{G}$ is the direct product of a group $\bar{G}$ and a canonical subgroup $\bar{G}'$ of its complementary group $\bar{G}^{C}$, $\mathcal{G}=\bar{G}\times\bar{G}'$, $\bar{G}'\subset\bar{G}^{C}$, because then the generators of $\mathcal{G}^{TC}$ are those polynomials in the generators of $\bar{G}^{C}$ which are left invariant by the transformations of $\bar{G}'$, and are therefore polynomials in the Casimir operators of $\bar{G}^{C}$ (or $\bar{G}$) and $\bar{G}'$.

The construction of the generators of $\mathcal{G}^{TC}$ is not as trivial when the algebra of polynomials in the Casimir operators of $\mathcal{G}$ is a proper subalgebra of the algebra of $\mathcal{G}^{TC}$. In this case the propagation operators can be easily built only for those k values for which the number of linearly independent polynomials in the Casimir operators of $\mathcal{G}$ is the same as that of defining subspaces (k,λ').

Such a situation may happen even when there is no state labeling problem in the reduction $U(N)\supset G$ for the IR's $[1^n]$ of $U(N)$. In Sec. 4, we give an example of such a case, corresponding to the reduction $U(6)\supset O^{+}(3)$, which occurs for identical fermions in a single shell of angular momentum 5/2.

When there is a state labeling problem in the reduction $U(N)\supset G$, the above-mentioned subalgebra is always proper because the algebra of $\mathcal{G}^{TC}$ contains at least one generator of the type $C^{\varphi''\Lambda'}_{\varphi'\Lambda'}$ $(\varphi'\neq\varphi'')$, which cannot be expressed in terms of Casimir operators of $\mathcal{G}$ as it connects two different IR's of $\mathcal{G}$. In fact it is well known[4] that in this case the algebra of $\mathcal{G}^{TC}$ contains the supplementary operators whose eigenvalues define the additional quantum numbers which complete the classification of n-particle states.

In both cases, the problem of building the generators of $\mathcal{G}^{TC}$ will be solved if we can construct an integrity basis, i.e., a minimal set of operators of $\mathcal{G}^{TC}$ in terms of which all the generators of $\mathcal{G}^{TC}$ can be written as polynomial expressions. This is a difficult problem to tackle, but some definite progress was made recently in its general formulation and its explicit solution in some particular cases.[5] Therefore, there is some hope that it will be possible in a near future to build the propagation operators in their full generality at least for some of those chains of groups for which we are restricted now to low k values.

It is quite easy to generalize the results of the discussion to the case of an operator O of mixed particle rank, not greater than u,

$$O=\sum_{k=0}^{u}O(k). \tag{3.10}$$

We may distinguish two cases:

(i) When any IR Λ' of $\mathcal{G}$ is contained at most once in the IR $[\frac{1}{2}N]$ of $O^{+}(2N+1)$, the average of any operator O can be propagated from its defining subspaces $\Lambda'=(k,\lambda')$, $0\leq k\leq u$, whatever u may be;

(ii) When $\bar{u}$ is the first value of n for which an IR Λ' of $\mathcal{G}$ is contained more than once in the IR $[\frac{1}{2}N]$ of $O^{+}(2N+1)$, the average of any operator O of maximum particle rank $u<\bar{u}$ can be propagated from its defining subspaces $\Lambda'=(k,\lambda')$, $0\leq k\leq u$; however, the average of an operator O for which $u\geq\bar{u}$ cannot be propagated in general. The problems encountered in the explicit construction of the propagation operators remain of course the same as before.

Finally, let us mention that a modification of definition (3.2) is used by several authors when dealing with subgroups G of $U(N)$ for which there is a state labeling problem.[14,15] It consists in taking the average of the operator $O(k)$ over all the equivalent IR's Λ of $\mathcal{G}$ contained in the IR $[\frac{1}{2}N]$ of $O^{+}(2N+1)$,

$$\langle O(k)\rangle^{\Lambda}=[d(\Lambda)\dim(\Lambda)]^{-1}\sum_{\varphi\mu}\langle\varphi\Lambda\mu|O(k)|\varphi\Lambda\mu\rangle. \tag{3.11}$$

Here $d(\Lambda)$ is the number of times the IR Λ of $\mathcal{G}$ is contained in the IR $[\frac{1}{2}N]$ of $O^{+}(2N+1)$ or $[1]$ of $U(2N)$, and is clearly a member of the set of numbers which define the dimensions of the unitary subgroups of $\mathcal{G}^{TC}$. The average (3.11) is thus taken in the representation space of the IR $\Lambda\times[1]$ of the subgroup $\mathcal{G}\times U(d(\Lambda))$ of $U(2N)$. The operators $D^{\varphi''\Lambda'}_{\varphi'\Lambda'}$ which contribute to such an average are therefore linear combinations of the generators of the subgroup $\sum_{\Lambda'\neq\Lambda}\oplus U(d(\Lambda'))$ of $\mathcal{G}^{TC}$. Trivial changes have to be applied to the relations derived above, and all the main conclusions remain valid. Moreover, it should be clear that whenever other scalar operators enter the construction of the propagation operators in addition to the Casimir operators, average (3.11) is more natural and much easier to calculate than (3.2) because in general those operators cannot be simultaneously diagonalized so that their diagonal matrix elements in the subspaces characterized by φ and Λ are difficult to determine.

We proceed now to illustrate the discussion of this section by same examples mostly taken from current applications of the spectral distribution method.

4. DISCUSSION OF SOME PARTICULAR CASES

(i) There is no state labeling problem and the algebra of $\mathcal{G}^{TC}$ is generated by the polynomials in the Casimir operators of $\mathcal{G}$.

Let us consider first the case where $\mathcal{G}=\bar{G}\times\bar{G}^{C}$. Then $\mathcal{G}$ is one of the three direct product groups of relations (2.12), (2.13), and (2.14). Only examples of the first type have been considered up to now because the other two imply averaging over states with different particle numbers. They correspond to r values equal to 1 (scalar averaging[1,2]), 2 (fixed isospin averaging[1]), or 4 (fixed supermultiplet averaging[16]). Any operator can be propagated, and the propagation operators are polynomials in the Casimir operators of $U(r)$, $G_1=N$, G_2, $G_3,\ldots,G_r$. For scalar averaging, they are thus polynomials in the number operator, for fixed isospin averaging polynomials in N and $\mathbf{T}^2$, and for fixed supermultiplet averaging polynomials in N, G_2, G_3, and G_4.

Let us consider next a case where $\mathcal{G}=\bar{G}\times\bar{G}'$, and $\bar{G}'$ is a canonical subgroup of $\bar{G}^C$: fixed seniority averaging for identical nucleons, the seniority being either that in a single subshell, or the multi-shell generalized seniority.[1,17] We have indeed $\mathcal{G}=\mathrm{Sp}(2q)\times\mathrm{U}(1)$, where $U(1)$ is a canonical subgroup of the quasispin group Sp(2), generated by the operator $S_0=\frac{1}{2}(N-q)$. Any operator can be propagated, and the propagation operators are polynomials in N and $\mathbf{S}^2$, the square of the quasispin operator.

(ii) There is no state labeling problem, but the algebra of $\mathcal{G}^{\mathrm{TC}}$ is not generated by the polynomials in the Casimir operators of $\mathcal{G}$ only.

An example of such a case is given by the quite trivial fixed angular momentum averaging for identical fermions in a single shell of angular momentum 5/2, corresponding to the chain of groups

$$\mathrm{U}(2^6)\supset\mathrm{O}^*(13)\supset\mathcal{G}=\mathrm{U}(1)\times\mathrm{O}^*(3). \tag{4.1}$$

The IR's of the last group are (0,0), (1,5/2), (2,0), (2,2), (2,4), (3,3/2), (3,5/2), (3,9/2), (4,0), (4,2), (4,4), (5,5/2), and (6,0), and are contained only once in the IR $[\frac{1}{2}^6]$ of $\mathrm{O}^*(13)$. Moreover, it can be easily verified that the generators of $\mathcal{G}^{\mathrm{TC}}$ cannot be built in terms of polynomials in the Casimir operators N and $\mathbf{J}^2$ only. However, it is sufficient to add to them the square of the quasispin operator $\mathbf{S}^2$ to get an integrity basis of $\mathcal{G}^{\mathrm{TC}}$ in terms of which all the propagation operators can be explicitly constructed.

(iii) There is a state labeling problem. Many examples of such a situation are known. Let us mention the fixed supermultiplet and SU(3) symmetry averaging,[18] the fixed supermultiplet, spin and isospin averaging,[19] and the fixed seniority, isospin, and reduced isospin averaging for both types of nucleons. The corresponding groups are $\mathcal{G}=\mathrm{SU}(3)\times\mathrm{U}(4)$, $\mathcal{G}=\mathrm{U}(q)\times\mathrm{SU}(2)\times\mathrm{SU}(2)$, and $\mathcal{G}=\mathrm{Sp}(2q)\times\mathrm{U}(2)$, and the state labeling problem occurs in the reduction $\mathrm{U}(q)\supset\mathrm{SU}(3)$, $\mathrm{U}(4)\supset\mathrm{SU}(2)\times\mathrm{SU}(2)$, and $\mathrm{Sp}(4)\supset\mathrm{U}(2)$, respectively. Another important case corresponds to the fixed angular momentum averaging for high enough N values.

The usefulness of the construction of an integrity basis of $\mathcal{G}^{\mathrm{TC}}$ can be illustrated by the following example. Let us consider the fixed supermultiplet and orbital angular momentum averaging for the p-shell nuclei. The corresponding group is $\mathcal{G}=\mathrm{O}^*(3)\times\mathrm{U}(4)$, where the rotation group in orbital space is a subgroup of the complementary group U(3) of U(4). The propagation operators for $u=2$ operator averages can be built in terms of the Casimir operators 1, N, N^2, G_2, and $\mathbf{L}^2$. However, for $u=4$ operator averages, the full integrity basis is needed. It was shown recently[5] that it includes the operators N, G_2, G_3, $\mathbf{L}^2$, as well as a third-order operator $X^{(3)}$, and a fourth-order operator $X^{(4)}$. In terms of these, the propagation operators for $u=4$ operator averages can be written as linear combinations of 1, N, N^2, N^3, N^4, G_2, NG_2, N^2G_2, G_3, NG_3, $G_2{}^2$, $\mathbf{L}^2$, $N\mathbf{L}^2$, $N^2\mathbf{L}^2$, $\mathbf{L}^4$, $G_2\mathbf{L}^2$, $X^{(3)}$, $NX^{(3)}$, and $X^{(4)}$.

*Maître de recherches F. N. R. S.

[1] J. B. French, in *Nuclear Structure*, edited by A. Hossain, Harun-ar-Rashid, and M. Islam (North-Holland, Amsterdam, 1967).
[2] J. B. French and K. F. Ratcliff, Phys. Rev. C 3, 94 (1971).
[3] F. S. Chang, J. B. French, and T. H. Thio, Ann. Phys. (N.Y.) 66, 137 (1971).
[4] M. Moshinsky and C. Quesne, J. Math. Phys. **11**, 1631 (1970).
[5] B. R. Judd, W. Miller Jr., J. Patera, and P. Winternitz, J. Math. Phys. 15, 1787 (1974).
[6] B. R. Judd and J. P. Elliott, *Topics in Atomic and Nuclear Theory* (University of Canterbury Press, New Zealand, 1970), p. 21.
[7] B. R. Judd, in *Group Theory and Applications*, edited by E. M. Loebl (Academic, New York, 1968), Vol. 1.
[8] M. C. K. Aguilera-Navarro and V. C. Aguilera-Navarro, J. Math. Phys. 16, 72 (1975).
[9] M. Moshinsky, *Group Theory and the Many Body Problem* (Gordon and Breach, New York, 1968).
[10] K. Helmers, Nucl. Phys. 69, 593 (1965).
[11] K. Helmers, Nucl. Phys. 23, 594 (1961).
[12] N. Jacobson, *Lie Algebras* (Wiley, New York, 1962).
[13] I. M. Gel'fand and M. L. Tseitlin, Dokl. Akad. Nauk SSSR 71, 825 (1950).
[14] K. F. Ratcliff, Phys. Rev. C 3, 117 (1971).
[15] C. Jacquemin, Phys. Lett. 43B, 253 (1973).
[16] J. C. Parikh, Ann. Phys. (N.Y.) 76, 202 (1973).
[17] C. Quesne and S. Spitz, Ann. Phys. (N.Y.) 85, 115 (1974).
[18] J. C. Parikh, Phys. Lett. 41B, 468 (1972).
[19] R. U. Haq and J. C. Parikh, Nucl. Phys. A 220, 349 (1974).

The electromagnetic field on a simplicial net*

Rafael Sorkin†

California Institute of Technology, Pasadena, California 91125
(Received 11 April 1975)

The "Regge calculus" approach is extended to the electromagnetic case. To this end an "affine" tensor formalism and associated exterior calculus are developed. The simplicial approach to linear field equations is illustrated by the two-dimensional scalar wave equation, on which also a discussion of the treacherous character of the continuum limit is based.

I. INTRODUCTION

A previous paper[1] developed the simplicial approach to the purely metrical field ("Regge calculus"). Therein the formalism was brought to a point where full-blown computer calculations ought to be possible, but my own attempts succeeded only in developing programs to calculate all the basic quantities but not in solving efficiently or reliably the basic set of equations for the time evolution problem. (See Sec. IV B of Ref. 1).

This paper will extend the formalism of Ref. 1 to the case of the coupled metrical and electromagnetic thatches ("geometrodynamics"). Since the electromagnetic part of the equations is linear, the new calculational problem is probably not much harder than the old.

In working with tensors defined relative to an n-simplex, it is convenient to use a system of coordinates which reflects the $(n+1)$-fold character of the vertices. Section II elaborates such a formalism, and Sec. III develops the "exterior calculus" in those terms. The key result which allows one to formulate the electromagnetic action is furnished by condition (3) of the theorem, in which Sec. III culminates.

Section IV presents the equations for the electromagnetic thatch and verifies all the formal consequences of these equations that one has become used to in the continuum.

Finally, Secs. V and VI discuss how the simplicial approach works out in a particularly simple situation—a "massless" scalar thatch on a two-dimensional net. It appears that the simplicial approach will agree with the finite difference scheme only "on the average." In particular, the investigation of the continuum limit by Taylor expansion at a point is in general misleading.

The notations and terminology of this paper agree with those of Ref. 1.

II. AFFINE COORDINATES

A. Affine coordinates

By considering a point P in the interior of an n-simplex as the centroid of $n+1$ masses t^j placed at the vertices $v_0 \cdots v_n$ we can express it as an "affine sum"

$$P = \sum_{j=0}^{n} t^j v_j \Big/ \sum_{j=0}^{n} t^j$$

of the vertices v_j. Renormalizing the masses, we can write

$$P = \sum_{j=0}^{n} t^j v_j, \tag{1}$$

in which

$$\sum_{j=0}^{n} t^j = 1 \tag{2}$$

and with all the $t^j > 0$. By relaxing this latter condition we can express any point in the affine space S of the simplex in the form (1), (2).[2] We call $v_0, v_1, \ldots, v_n$ an affine point basis for S.

A vector of an affine space is the "difference" of two points which we write as $Q - P$ or $\overrightarrow{PQ}$. Let V denote the space of all vectors of S. If

$$P = \sum p^j v_j, \quad Q = \sum q^j v_j,$$

then we take for coordinate of $\overrightarrow{PQ}$ the differences $x^j \equiv q^j - p^j$. Then (2) implies

$$\sum_{j=0}^{n} x^j = 0. \tag{3}$$

Another way to explain these coordinates is as components of $\overrightarrow{PQ}$ relative to the (redundant) "barycentric basis" comprising the $n+1$ vectors

$$\mathbf{e}_i = v_i - \frac{1}{n+1} \sum_{k=0}^{n} v_k. \tag{4}$$

A simple computation verfies this:

$$\begin{aligned} \sum_i x^i \mathbf{e}_i &= \sum_i x^i v_i - \Big(\sum_i x_i\Big)\Big(\frac{1}{n+1} \sum_k v_k\Big) \\ &= \sum_i (q^i - p^i) v_i \quad [\text{by (3)}] \\ &= \sum q^i v_i - \sum p^i v_i \\ &= Q - P = \overrightarrow{PQ}. \end{aligned}$$

Corresponding to the basis $(\mathbf{e}_j)$ for V, we introduce for the dual space V^* a basis $(\mathbf{e}^i)$ defined by

$$\langle \mathbf{e}^j, \mathbf{e}_k \rangle = \tilde{\delta}^j_k \equiv \delta^j_k - \frac{n}{n+1} = \begin{cases} \dfrac{1}{n+1} & \text{if } j = k, \\ -\dfrac{1}{n+1} & \text{if } j \neq k. \end{cases} \tag{5}$$

Notice that

$$\sum_k \mathbf{e}_k = 0, \quad \sum_k \mathbf{e}^k = 0, \tag{6}$$

$$\sum_k \mathbf{e}_k \otimes \mathbf{e}^k = \mathbf{1}. \tag{7}$$

Let us check the last relation, for example, by applying its left-hand side to the vector $a=\sum a^i\mathbf{e}_i$. First, however, we point out the lemma:

Lemma: If for any quantities Q_j, $j=0,\ldots,n$, $\sum_j Q_j=0$, then

$$Q_j=\tilde{\delta}^k_j Q_k. \tag{8}$$

Continuing with the check, we have

$$\begin{aligned}\sum_k \mathbf{e}_k\otimes\mathbf{e}^k\cdot a&=\sum_k \mathbf{e}_k\left\langle \mathbf{e}^k,\sum_j a^j\mathbf{e}_j\right\rangle\\&=\sum_{k,j}\mathbf{e}_k a^j\langle\mathbf{e}^k,\mathbf{e}_j\rangle\\&=\sum_j a^j\sum_k\tilde{\delta}^k_j\mathbf{e}_k\\&=\sum_j a^j\mathbf{e}_j\quad[\text{by the lemma and (6)}]\\&=a.\end{aligned}$$ QED

If T is any sort of tensor relative to the vector space V, we define its *affine components* $T^{j\cdots k}_{i\cdots m}$ by contracting it with the relevant product of basis vectors $\mathbf{e}_j,\mathbf{e}^k$. Then (7) guarantees the expansion:

$$T=\sum_{\substack{j\cdots k\\ i\cdots m}} T^{j\cdots k}_{i\cdots m}\mathbf{e}_j\otimes\cdots\otimes\mathbf{e}_k\otimes\mathbf{e}^i\cdots\otimes\mathbf{e}^m, \tag{9}$$

from which follows, with the aid of (6),

$$\sum_j T^{j\cdots k}_{i\cdots m}=0 \tag{10}$$

for any index j, up or down. This last result is the distinguishing feature of affine components and, together with (5) and the lemma, it guarantees that contraction works as usual by summing on the contracted indices.

Finally, we derive the affine components of two "special tensors." The "Kronecker delta tensor" δ has components formed as follows (in a slightly cumbersome notation):

$$\tilde{\delta}^j_k=\delta^\mu_\nu(\mathbf{e}_k)^\nu(\mathbf{e}^j)_\mu=(\mathbf{e}_k)^\nu(\mathbf{e}^j)_\nu=\langle\mathbf{e}_k,\mathbf{e}^j\rangle=\tilde{\delta}^j_k,$$

which shows the consistency of our earlier definition (5).

The other "special" tensor we will need is the epsilon symbol, which strictly speaking is not a tensor but a tensor density and thus defined *a priori* only up to an overall factor. We fix this factor by setting

$$\tilde{\epsilon}^{12\cdots n}=+1,$$

from which it is easy to evaluate the other components using the antisymmetry of $\tilde{\epsilon}^{i\cdots j}$ and the sum rule (10). Thus, for example,

$$\tilde{\epsilon}^{213\cdots n}=-\tilde{\epsilon}^{12\cdots n}=-1$$

and

$$\tilde{\epsilon}^{023\cdots n}+\tilde{\epsilon}^{123\cdots n}+\tilde{\epsilon}^{223\cdots n}+\cdots+\tilde{\epsilon}^{n23\cdots n}=0,$$

$$\tilde{\epsilon}^{023\cdots n}+1+0+\cdots+0=0,\ \text{so that}\ \tilde{\epsilon}^{023\cdots n}=-1.$$

Let $j_0j_1\cdots j_n$ be any permutation of the indices $01\cdots n$. Then

$$\tilde{\epsilon}^{j_1\cdots j_n}=\begin{cases}+1 & \text{if the permutation } j_0j_1\cdots j_n \text{ is even,}\\ -1 & \text{if the permutation } j_0j_1\cdots j_n \text{ is odd.}\end{cases} \tag{11}$$

We note without proof that our definition is equivalent (for the contravariant ϵ) to

$$\epsilon=\overrightarrow{v_0v_1}\wedge\overrightarrow{v_0v_2}\wedge\cdots\wedge\overrightarrow{v_0v_n}.$$

A final subtlety needs mention. Let $n=3$ for definiteness. Then under the usual definitions

$$\epsilon^{\sigma\mu\nu}\epsilon_{\sigma\alpha\beta}=\delta^\mu_\alpha\delta^\nu_\beta-\delta^\mu_\beta\delta^\nu_\alpha.$$

That the analogous formula apply to $\tilde{\epsilon}^{ijk}$ and $\tilde{\epsilon}_{ilm}$ requires, as is easily checked, that $\tilde{\epsilon}^{123}\tilde{\epsilon}_{123}=\frac{1}{4}=1/(n+1)$. Accordingly, we define the covariant ϵ with components of magnitude $(n+1)^{-1}$:

$$\tilde{\epsilon}_{j_1\cdots j_n}=[1/(n+1)]\tilde{\epsilon}^{j_1\cdots j_n}. \tag{12}$$

With these definitions all the expected formulas obtain.

B. The metric tensor of a simplex

As pointed out in Ref. 1, to specify the $n(n+1)/2$ edge lengths of an n-simplex is equivalent to specifying a flat metric for the interior of that simplex. In this section we calculate the affine components $\tilde{g}_{ij}$ of this metric. Let l^2_{ij} be the length squared of the edge joining v_i to v_j. Then, since, plainly, $\overrightarrow{v_iv_j}=\mathbf{e}_j-\mathbf{e}_i$,

$$\begin{aligned}l^2_{ij}&=\langle g,\overrightarrow{v_iv_j}\otimes\overrightarrow{v_iv_j}\rangle\\&=\langle g,(\mathbf{e}_j-\mathbf{e}_i)\otimes(\mathbf{e}_j-\mathbf{e}_i)\rangle\\&=\langle g,\mathbf{e}_j\otimes\mathbf{e}_j\rangle-2\langle g,\mathbf{e}_i\otimes\mathbf{e}_j\rangle+\langle g,\mathbf{e}_i\otimes\mathbf{e}_i\rangle\\&=\tilde{g}_{jj}-2\tilde{g}_{ij}+\tilde{g}_{ii}\\&\equiv A_{ij}.\end{aligned}$$

By forming the combination $\tilde{\delta}^k_i\tilde{\delta}^l_jA_{kl}$, we can, in view of (8) and (10) as applied to $\tilde{\delta}$, recover $\tilde{g}_{ij}$:

$$\tilde{\delta}^k_i\tilde{\delta}^l_jl^2_{kl}=0-2\tilde{g}_{ij}+0,\quad \tilde{g}_{ij}=-\tfrac{1}{2}l^2_{kl}\tilde{\delta}^k_i\tilde{\delta}^l_j, \tag{13}$$

which says that $\tilde{g}_{ij}$ is just $-\frac{1}{2}l^2_{ij}$ "rendered affine" or "projected into the affine tensors."

Now suppose that $\tilde{A}^{jk}$ are the affine components of some tensor. Then according to (13)

$$\begin{aligned}\tilde{A}^{ij}\tilde{g}_{ij}&=-\tfrac{1}{2}(\tilde{A}^{ij}\tilde{\delta}^k_i\tilde{\delta}^l_j)l^2_{kl}\\&=-\tfrac{1}{2}\tilde{A}^{kl}l^2_{kl};\end{aligned}$$

in other words, one has the general

Replacement rule: If $\tilde{g}_{jk}$ occurs with both indices contracted against affine indices, then it can be replaced by $-\frac{1}{2}l^2_{jk}$.

C. "Geometrical" tensors

In this subsection we fix some normalizations and derive a useful expression for the volume of a simplex.

Let the wedge product be defined in the usual way and normalized so that, for instance, the wedge product $a\wedge b\wedge c$ of three vectors consists of six terms each with coefficient ±1. Then we take the product $a\wedge b$ to represent the parallelogram determined by a and b, and $\frac{1}{2}a\wedge b$ the triangle or 2-simplex spanned by them. In general, the normalized product

$$\omega=(1/m!)a_1\wedge\cdots\wedge a_m \tag{14}$$

will represent the m-simplex spanned by vectors $a_1\cdots a_m$. We also introduce the more conveniently normalized contraction

$$\langle\omega\,|\,\phi\rangle \equiv \langle\omega,\phi\rangle/m!, \tag{15}$$

where m is the rank of the forms ω, ϕ. For example, if $m=2$, then

$$\langle\omega\,|\,\phi\rangle = \tfrac{1}{2}\omega^{\mu\nu}\phi_{\mu\nu}.$$

With these definitions the volume represented by any rank m totally antisymmetric tensor is

$$\mathrm{vol}(\omega) = |\langle\omega\,|\,\omega\rangle|^{1/2} \equiv \|\omega\|. \tag{16}$$

(The absolute value is needed because of the indefinite metric, i.e., the volume is defined to be a real number.) Thus, for example, the bone [012] of some 4-simplex σ corresponds to the tensor

$$\begin{aligned}\omega &= (1/2!)\overrightarrow{v_0v_1}\wedge\overrightarrow{v_0v_2}\\ &= (1/2!)(\mathbf{e}_1-\mathbf{e}_0)\wedge(\mathbf{e}_2-\mathbf{e}_0)\\ &= (1/2!)(\mathbf{e}_1\wedge\mathbf{e}_2+\mathbf{e}_2\wedge\mathbf{e}_0+\mathbf{e}_0\wedge\mathbf{e}_1)\\ &= (1/2!)\tilde{\epsilon}^{ij}_{(012)}\mathbf{e}_i\otimes\mathbf{e}_j,\end{aligned}$$

where $\tilde{\epsilon}^{ij}_{(012)}$ is of course formed from the indices 0, 1, 2 in the manner of (11). In general the m subsimplex with vertices $k_0\cdots k_m$ corresponds via (14) to the tensor with affine components:

$$\tilde{\omega}^{j_1\cdots j_m} = (1/m!)\epsilon^{j_1\cdots j_m}_{(k_0\cdots k_m)}. \tag{17}$$

According to (16) the volume V of this simplex is given by

$$\pm V^2 = \langle\omega\,|\,\omega\rangle = (m!)^{-1}\tilde{\omega}^{i\cdots j}\tilde{g}_{ia}\cdots\tilde{g}_{jb}\tilde{\omega}^{a\cdots b}.$$

By the replacement $\tilde{g}^2_{ij} \to -\frac{1}{2}l^2_{ij}$ discovered in the previous subsection we convert this result to one expressed directly in terms of edge lengths:

$$\pm\mathrm{vol}^2 = (-\tfrac{1}{2})^m(1/m!)\tilde{\omega}^{j_1\cdots j_m}\tilde{\omega}^{k_1\cdots k_m}\prod_{a=1}^{m} l^2_{j_ak_a}. \tag{18}$$

To find the volume of any m-simplex of the net, we can work within the m-dimensional affine space spanned by that simplex and (calling its vertices $0\cdots m$) write

$$\tilde{\omega}^{j_1\cdots j_m} = (1/m!)\tilde{\epsilon}^{j_1\cdots j_m}.$$

Then

$$\pm\mathrm{vol}^2 = (-\tfrac{1}{2})^m(m!)^{-3}\tilde{\epsilon}^{j_1\cdots j_m}\tilde{\epsilon}^{k_1\cdots k_m}l^2_{j_1k_1}\cdots l^2_{j_mk_m}. \tag{19}$$

To facilitate numerical evaluation of such expressions, we introduce the concept of a "bordered determinant" which has the form (with A representing any $m\times m$ matrix)

$$B(A) \equiv \begin{vmatrix} 0 & 1 & 1 & \cdots & 1\\ 1 & & & & \\ 1 & & & & \\ \cdot & & A & & \\ \cdot & & & & \\ \cdot & & & & \\ 1 & & & & \end{vmatrix}$$

Then the expression

$$\tilde{\epsilon}^{j_1\cdots j_m}\tilde{\epsilon}^{k_1\cdots k_m}A_{j_1k_1}\cdots A_{j_mk_m}$$

can be evaluated as $-m!B(A)$, the proof being left to the interested reader. Thus we get the expression for volume in terms of edges, as

$$\pm\mathrm{vol}^2 = -(-\tfrac{1}{2})^m(m!)^{-2}\begin{vmatrix} 0 & 1 & 1 & \cdots & 1\\ 1 & & & & \\ 1 & & l^2_{ij} & & \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ 1 & & & & \end{vmatrix}, \tag{20}$$

a result which appears in Ref. 3.

For a triangle we find (setting $m=2$, $x=l^2_{01}$, $y=l^2_{02}$, $z=l^2_{12}$)

$$\begin{aligned}\pm A^2 &= -(-\tfrac{1}{2})^2(2!)^{-2}\begin{vmatrix} 0 & 1 & 1 & 1\\ 1 & 0 & x & y\\ 1 & x & 0 & z\\ 1 & y & z & 0\end{vmatrix}\\ &= -\tfrac{1}{16}[x^2+y^2+z^2-2(xy+yz+zx)].\end{aligned}$$

III. SIMPLICIAL EXTERIOR CALCULUS

Let

Σ_0 = set of all 0-simplexes (vertices) of the net,

Σ_1 = set of all oriented 1-simplexes (legs) of the net,

Σ_2 = set of all oriented 2-simplexes of the net, etc.,

and represent a typical oriented 2-simplex, e.g., as $[xyz]$, where $x, y, z \in \Sigma_0$. Then we define[4]

a 0-*form* (scalar thatch) as a map $\phi\colon \Sigma_0 \to R$,

a 1-*form* (co-vector thatch) as a map $A\colon \Sigma_1 \to R$ such that $A(xy) = -A(yx)$,

a 2-*form* as a map $F\colon \Sigma_2 \to R$ such that $F(xyz) = -F(yxz) = F(yzx)$, etc.

To understand these definitions, one could think of $A(xy)$, e.g., as the line integral $\int_x^y A_\mu\,dx^\mu$ of some field A_μ along the leg $[xy]$. If, then, $F = dA$, then Stokes' theorem becomes

$$\begin{aligned}F(xyz) &\longleftarrow \int_{[xyz]}F_{\mu\nu}\,da^{\mu\nu} = \oint_{[xy]+[yz]+[zx]}A_\mu\,dx^\mu\\ &\longleftarrow A(xy)+A(yz)+A(zx).\end{aligned}$$

Generalizing this relation to arbitrary dimension we define the operator "d" from m- to $(m+1)$-forms as follows:

$$d\omega(k_0k_1\cdots k_m) = \sum_{j=0}^{m}(-1)^j\omega(k_0\cdots\hat{k}_j\cdots k_m), \tag{21}$$

where the "hat" indicates omission. It is easy to check that

$$d\,d\omega = 0: \tag{22}$$

$$\begin{aligned}d\,d\omega(k_0\cdots k_{m+1}) &= \sum_{j=0}^{m+1}(-1)^j\,d\omega(k_0\cdots\hat{k}_j\cdots k_{m+1})\\ &= \sum_{j=0}^{m+1}(-1)^j\sum_{l=0}^{m+1}(-1)^l\,\mathrm{sgn}(l-j)\\ &\quad\times\omega(k_0\cdots\hat{k}_a\cdots\hat{k}_b\cdots k_{m+1})\\ &\qquad[\text{where } a=\min(j,l),\ b=\max(j,l)]\\ &= \sum_{j,l=0}^{m+1}(-1)^{j+l}\,\mathrm{sgn}(l-j)\\ &\quad\times\omega(k_0\cdots\hat{k}_a\cdots\hat{k}_b\cdots k_{m+1})\\ &= 0\end{aligned}$$

since $\mathrm{sgn}(l-j)$ alone in the penultimate expression is antisymmetric in the exchange of j with l. QED

The main theorem of this section is in part a partial converse to (22).

Consider now a particular m-form ω and a particular simplex $\sigma \in \Sigma_4$ and ask whether ω "extends to the interior of σ," i.e., whether there is defined in the space of σ an m-form $\omega(\sigma)$ which agrees with ω on all the m-subsimplexes of σ. By definition $\omega(\sigma)$ "agrees with" ω on $[x_0 \cdots x_m]$ iff

$$\langle \omega(\sigma), [x_0 \cdots x_m] \rangle = \omega(x_0 \cdots x_m). \tag{23}$$

If ω does extend to σ, then we can form a simple expression [analogous to (13)] for the affine components of $\omega(\sigma)$:

From above we know that, calling $\sigma = [01 \cdots n]$,

$$m![k_0 \cdots k_m] = \overrightarrow{k_0 k_1} \wedge \overrightarrow{k_0 k_2} \wedge \cdots \wedge \overrightarrow{k_0 k_m} = (\mathbf{e}_{k_1} - \mathbf{e}_{k_0}) \wedge (\mathbf{e}_{k_2} - \mathbf{e}_{k_0}) \wedge \cdots (\mathbf{e}_{k_m} - \mathbf{e}_{k_0}).$$

In the expansion of the right-hand side, only terms lacking or linear in $\mathbf{e}_{k_0}$ survive since $\mathbf{e}_{k_0} \quad \mathbf{e}_{k_0} = 0$, and one obtains, in a hopefully clear notation,

$$m![k_0 \cdots k_m] = \mathbf{e}_{k_1} \wedge \cdots \wedge \mathbf{e}_{k_m} - \sum_{j=1}^{m} \mathbf{e}_{k_1} \wedge \cdots \wedge \underset{(j)}{\mathbf{e}_{k_0}} \wedge \cdots \wedge \mathbf{e}_{k_m}. \tag{24}$$

Because of (6) we can isolate $\mathbf{e}_{k_1} \wedge \cdots \wedge \mathbf{e}_{k_m}$ by summing on k_0:

$$\begin{aligned} &m! \sum_{k_0=0}^{n} [k_0 \cdots k_m] = (n+1)\mathbf{e}_{k_1} \wedge \cdots \wedge \mathbf{e}_{k_m}, \\ &\mathbf{e}_{k_1} \wedge \cdots \wedge \mathbf{e}_{k_m} = \frac{m!}{1+n} \sum_{k_0=0}^{n} [k_0 k_1 \cdots k_m]. \end{aligned} \tag{25}$$

Applying $\omega(\sigma)$ to both sides, we have

$$\begin{aligned} &m!\,\tilde{\omega}(\sigma)_{k_1 \cdots k_m} = \frac{m!}{1+n} \sum_{k_0=0}^{n} \langle \omega(\sigma), [k_0 \cdots k_m] \rangle, \\ &\tilde{\omega}(\sigma)_{k_1 \cdots k_m} = \frac{1}{1+n} \sum_{k_0=0}^{n} \omega(k_0 k_1 \cdots k_m). \end{aligned} \tag{26}$$

In order to study this condition more closely, we make the definition (relative to the simplex σ)

$$S\omega(k_1 \cdots k_m) = \frac{1}{1+n} \sum_{k_0=0}^{n} \omega(k_0 k_1 \cdots k_n) \tag{27}$$

so that (26) can be expressed in the droll form

$$\tilde{\omega}(\sigma)_{k_1 \cdots k_m} = S\omega(k_1 \cdots k_m). \tag{28}$$

It is easy to see that $S\omega$ is an $(m-1)$-form (on σ) when ω is an m-form, and that

$$S^2 = 0. \tag{29}$$

We can also verify the important relation (relative to σ as always)

$$Sd + dS = 1. \tag{30}$$

Proof:

$$\begin{aligned} dS\omega(k_0 \cdots k_m) &= \sum_{j=0}^{m} (-1)^j S\omega(k_0 \cdots \hat{k}_j \cdots k_m) \\ &= \sum_{j=0}^{m} (-1)^j \frac{1}{1+n} \sum_{l=0}^{n} \omega(l k_0 \cdots \hat{k}_j \cdots k_m) \\ &= \frac{1}{1+n} \sum_{j=0}^{m} \sum_{l=0}^{n} \omega(k_0 \cdots \underset{(j)}{l} \cdots k_m), \\ S\,d\omega(k_0 \cdots k_m) &= \frac{1}{1+n} \sum_{l=0}^{n} d\omega(l k_0 \cdots k_m) \\ &= \frac{1}{1+n} \sum_{l=0}^{n} \omega(k_0 \cdots k_m) \\ &\quad - \sum_{j=0}^{m} (-1)^j \omega(l k_0 \cdots \hat{k}_j \cdots k_m) \\ &= \omega(k_0 \cdots k_m) - \frac{1}{1+n} \sum_{l=0}^{n} \\ &\quad \times \sum_{j=0}^{m} \omega(k_0 \cdots \underset{(j)}{l} \cdots k_m). \end{aligned}$$

Comparing the two expressions completes the proof.

Returning to the question whether ω extends to σ, we note that the formula for $\omega(\sigma)$ given in (28) or (26) will define a form in the space of σ whether or not ω extends to σ. If we call this form ρ, then the condition that ρ agree with ω on σ (which is just that that ω extend to σ) becomes

$$\langle \rho, [k_0 \cdots k_m] \rangle = \omega(k_0 \cdots k_m).$$

But by (24)

$$\begin{aligned} &\langle \rho, [k_0 \cdots k_m] \rangle \\ &= \frac{1}{m!} \langle \rho, \mathbf{e}_{k_1} \wedge \cdots \wedge \mathbf{e}_{k_m} \rangle \\ &\quad - \frac{1}{m!} \sum_{j=1}^{m} \langle \rho, \mathbf{e}_{k_1} \wedge \cdots \underset{(j)}{\mathbf{e}_{k_0}} \wedge \cdots \mathbf{e}_{k_m} \rangle \\ &= \tilde{\rho}_{k_1 \cdots k_m} - \sum_{j=1}^{m} \tilde{\rho}_{k_1 \cdots \underset{(j)}{k_0} \cdots k_m} \\ &= \frac{1}{1+n} \sum_{l=0}^{n} \omega(l k_1 \cdots k_m) - \sum_{j=1}^{m} \frac{1}{1+n} \\ &\quad \times \sum_{l=0}^{n} \omega(l k_1 \cdots \underset{(j)}{k_0} \cdots k_m) \\ &= \frac{1}{1+n} \sum_{l=0}^{n} \omega(l k_1 \cdots k_m) + \frac{1}{1+n} \sum_{j=1}^{m} \sum_{l=0}^{n} \omega(k_0 k_1 \cdots \underset{(j)}{l} \cdots k_m) \\ &= \frac{1}{1+n} \sum_{l=0}^{n} \sum_{j=0}^{m} \omega(k_0 \cdots \underset{(j)}{l} \cdots k_m). \end{aligned}$$

Comparing this with the proof of (30) furnishes the condition for ω to extend to σ in the form

$$dS\omega = \omega. \tag{31}$$

We can now prove the following fundamental theorem which has been the goal of this subsection:

Theorem: Let ω be any form defined on a net including the simplex σ and set $\Omega = S\omega$, as defined in (27). Then the following three conditions are equivalent:

(1) $\omega = d\Omega$,

(2) $d\omega = 0$,

(3) ω extends to σ, the extension being furnished by (26).

Proof: We just saw that we can replace (3) by the condition

(3′) $dS\omega = \omega$;

we already know by (22) that (1) $\Rightarrow$ (2), and (3′) $\Rightarrow$ (1) is

obvious. To complete the circle of implication, we need only (2)⇒(3′) which follows immediately from (30) applied to ω. QED

By the way, Ω becomes unique through the condition $S\Omega=0$, which follows from (29).

IV. THE ELECTROMAGNETIC THATCH

(All components in this section are affine components—but the tilde (~) will usually be omitted.)

A. The source free thatch equations

In this section we assume a net Σ with fixed metric thatch l^2_{ij} and the associated metric tensors $g(\sigma)$ for each $\sigma\in\Sigma_4$.

The "vector potential" A is a 1-form on Σ, as defined in Sec. III, and $F=dA$ is the electromagnetic thatch. By the theorem of the previous chapter F extends in each cell $\sigma\in\Sigma_4$ to a tensor $F_{ij}(\sigma)$ given by (26). Calling $V(\sigma)$ the volume of σ, we take for the action

$$\begin{aligned} S_e &= -\tfrac{1}{2}\sum_{\sigma\in\Sigma_4} V(\sigma)\langle F(\sigma)\,|\,F(\sigma)\rangle_\sigma \\ &= -\tfrac{1}{4}\sum_\sigma V(\sigma)F^{ij}(\sigma)F_{ij}(\sigma) \\ &= -\tfrac{1}{4}\sum_\sigma V(\sigma)g(\sigma)^{ia}g(\sigma)^{jb}F(\sigma)_{ij}F(\sigma)_{ab}. \end{aligned} \tag{32}$$

The thatch equations equate to zero the variation of S with respect to the thatch A:

$$\frac{\partial S}{\partial A(ij)}=0 \quad \text{for all legs } [ij]\in\Sigma_1. \tag{33}$$

Well,

$$\delta S_e=-\tfrac{1}{2}\sum_\sigma V(\sigma)F^{ij}(\sigma)\delta F_{ij}(\sigma).$$

But according to (26)

$$\begin{aligned} F_{ij}(\sigma) &= \tfrac{1}{5}\sum_{k\in\sigma} F(kij) \\ &= \tfrac{1}{5}\sum_k [A(ij)+A(jk)+A(ki)] \\ &= A(ij)+\tfrac{1}{5}\sum_k [A(jk)+A(ki)] \end{aligned} \tag{34}$$

Because F_{ij} are affine components we can also write

$$\begin{aligned} F_{ij}(\sigma) &= \tilde\delta^a_i\tilde\delta^b_j F_{ab}(\sigma) \\ &= \tilde\delta^a_i\tilde\delta^b_j A(ab), \end{aligned} \tag{35}$$

where the remaining terms vanish because of (10) applied to $\tilde\delta^j_k$. Just as for $\tilde g_{jk}$ one discovers from (35) the *replacement rule*

$$\tilde F_{jk}\to A(jk) \tag{36}$$

whenever $\tilde F_{jk}$ occurs with both j and k contracted against affine indices.

This allows us to express δS_e directly in terms of δA:

$$\delta S_e=-\tfrac{1}{2}\sum_\sigma V(\sigma)F^{ij}(\sigma)\delta A(ij) \tag{37}$$

There is thus one thatch equation for each leg of the net:

$$\sum_\sigma V(\sigma)F^{ij}(\sigma)=0, \tag{38}$$

where, of course, σ ranges only over those 4-simplexes for which the expression has sense, i.e., for those of which $[i]$ and $[j]$ are vertices.

We can also express $F^{ij}(\sigma)$, and thereby the thatch equations, directly in terms of A:

$$\begin{aligned} F^{ij}(\sigma) &= g^{ia}(\sigma)g^{jb}(\sigma)F_{ab}=g^{ia}(\sigma)g^{jb}(\sigma)A(ab) \\ F^{ij}(\sigma) &= \tfrac{1}{2}h^{ijab}(\sigma)A(ab), \end{aligned} \tag{39}$$

where

$$h^{ijab}=g^{ia}g^{jb}-g^{ib}g^{ja}.$$

B. The equations with a source—Charge conservation

If there is prescribed a source J, then the action has an additional term

$$S_i=\sum_{[ij]\in\Sigma_1} A(ij)J(ij), \tag{40}$$

in which $J(ij)$ should be considered, not as a 1-form, but rather as a "vector density" or "current." In place of (37) stands (half of)

$$-\sum_\sigma V(\sigma)F^{ij}(\sigma)\delta A(ij)+J(ij)\delta A(ij)$$

so that the thatch equations become

$$\sum_\sigma V(\sigma)F^{ij}(\sigma)=J(ij). \tag{41}$$

The natural interpretation of J regards $J(ij)$ as the charge flowing "along" leg $[ij]$ of the net. It is as if Σ_1 were an electrical network, A the potential drop, and J the current. Then the conservation of charge (like one of Kirchhoff's laws) reads

$$\sum_j J(ij)=0 \tag{42}$$

and follows from (41) because of the rule (10).

We can also cast the conservation law in an "integral" form as opposed to its "local" statement (42): Let $\Omega\subset\Sigma_0$ be all the vertices in some region of the net and form the two expressions

$$\sum_{i\in\Omega}\sum_{k\in\Sigma_0} J(ik) \quad\text{and}\quad \sum_{i,k\in\Omega} J(ik).$$

The first vanishes by the equation of conservation (42) and the second by the antisymmetry of J. Then

$$\begin{aligned} 0 &= \sum_{i\in\Omega}\sum_{k\in\Sigma_0} J(ik) \\ &= \sum_{i\in\Omega}\left(\sum_{k\in\Omega}+\sum_{k\notin\Omega}\right)J(ik) \\ &= 0+\sum_{i\in\Omega}\sum_{k\notin\Omega} J(ik), \\ \sum_{i\in\Omega}\sum_{k\notin\Omega} J(ik) &= 0. \end{aligned} \tag{43}$$

In words: "The total charge leaving the region Ω vanishes."

C. Gauge invariance

As usual, $F=dA$ determines A only up to an addition of the form $d\theta$, for arbitrary 0-form θ. Since A does not occur explicitly in S_e, we are free to require invariance under the "gauge transformation"

$$A\to A+d\theta \tag{44}$$

as long as the interaction term (40) is unaffected. But under (44) S_i acquires an additional term

$$\tfrac{1}{2}\sum_{i,j} d\theta(ij)J(ij)\ ^{5}$$
$$=\tfrac{1}{2}\sum [\theta(j)-\theta(i)]J(ij)$$
$$=-\sum_i \theta(i)\sum_j J(ij)$$

whence gauge invariance requires

$$\sum_j J(ij)=0 \tag{45}$$

since θ is arbitrary. This is exactly the familiar connection between gauge invariance and charge conservation.

Since the gauge freedom of A introduces a free number for each vertex of the net, one can remove this freedom by imposing one condition at each vertex. One which suggests itself is

$$\sum_j A(ij)=0 \quad \text{at all } i\in\Sigma_0. \tag{46}$$

This looks something like the "Lorentz gauge," but it is not, since A is a 1-form rather than a current.

D. Coupling to the metric thatch—The energy-momentum tensor

Equation (38) already includes the effects of an arbitrary background metric. To find the reciprocal influence of the electromagnetic thatch on the metric, we must evaluate

$$T(ij)=-\frac{\partial S_e}{\partial l_{ij}^2}\ . \tag{47}$$

Writing (32) in the form

$$S_e=\sum_\sigma L(\sigma)$$
$$L(\sigma)\equiv -\tfrac{1}{4}V(\sigma)g^{\mu\mu}g^{\nu\nu}F_{\mu\nu}(\sigma)F_{\mu\nu}(\sigma), \tag{48}$$

and, varying the metric $g(\sigma)$ interior to σ, one finds

$$2\delta L=-\tfrac{1}{2}\delta V\langle F,F\rangle - Vg^{\mu\mu}\delta g^{\nu\nu}F_{\mu\nu}F_{\mu\nu}$$
$$=-\tfrac{1}{2}\delta V\langle F,F\rangle + Vg_{\mu\mu}\delta g_{\nu\nu}F^{\mu\nu}F^{\mu\nu}.$$

If we express this in affine components, then δV assumes a simple form which follows readily from the method of Sec. II C:

$$\delta V=\tfrac{1}{2}V\tilde{g}^{ij}\delta\tilde{g}_{ij} \tag{49}$$

whence

$$2\delta L=V\tilde{g}_{ii}\delta\tilde{g}_{jj}\tilde{F}^{ij}\tilde{F}^{ij}-\tfrac{1}{4}V\tilde{F}^{ab}\tilde{F}_{ab}\tilde{g}^{jj}\delta\tilde{g}_{jj}$$
$$=V\delta\tilde{g}_{jj}(\tilde{F}^{ij}\tilde{g}_{ii}\tilde{F}^{ij}-\tfrac{1}{4}\tilde{F}^{ab}\tilde{F}_{ab}\tilde{g}^{jj})$$
$$=V(\sigma)\delta\tilde{g}_{jk}(\sigma)\tilde{T}^{jk}(\sigma), \tag{50}$$

in which

$$\tilde{T}^{j}{}_{k}(\sigma)\equiv\tilde{F}^{ja}(\sigma)\tilde{F}_{ka}(\sigma)-\tfrac{1}{4}\tilde{F}^{ab}(\sigma)\tilde{F}_{ab}(\sigma)\tilde{\delta}^{j}{}_{k} \tag{51}$$

is the well-known formation in terms of $\tilde{F}_{ij},\tilde{g}_{ij}$. Applying the replacement rule of Sec. II B converts (50) into

$$\delta L=-\tfrac{1}{4}\delta l_{jk}^2 V\tilde{T}^{jk}$$

so that, finally,

$$\delta S=\sum_\sigma \delta L(\sigma)$$
$$=-\tfrac{1}{4}\sum_\sigma\sum_{j,k}\delta l_{jk}^2 V(\sigma)\tilde{T}^{jk}(\sigma)$$
$$=-\tfrac{1}{4}\sum_{j,k}\delta l_{jk}^2\sum_\sigma V(\sigma)\tilde{T}^{jk}(\sigma)$$
$$=-\sum_{[j,k]\in\Sigma_1}\delta l_{jk}^2 T_e(jk),$$
$$\text{where}\quad T_e(jk)=\tfrac{1}{2}\sum_\sigma V(\sigma)\tilde{T}^{jk}(\sigma). \tag{52}$$

Thus the thatch equation (4) of Ref. 1 becomes for the present case

$$G(jk)=T_e(jk). \tag{53}$$

V. EXAMPLE: THE WAVE EQUATION IN TWO DIMENSIONS

To develop some feeling for the behavior of simplicial equations, we can study a particularly simple, linear case: the two-dimensional wave equation.

If ϕ is the basic scalar thatch then, in analogy with the continuum theory, we choose for the action

$$S=\sum_{\sigma\in\Sigma_n}L(\sigma)V(\sigma) \tag{54}$$

where

$$L(\sigma)=\langle d\phi(\sigma)\mid d\phi(\sigma)\rangle$$
$$=\tfrac{1}{2}\tilde{g}^{ij}(\sigma)d\tilde{\phi}_i(\sigma)d\tilde{\phi}_j(\sigma). \tag{55}$$

Here, of course, $g(\sigma)$ and $d\phi(\sigma)$ are defined as in Secs. II B and III, respectively.

There is also a replacement rule for $d\tilde{\phi}_j$. Explicitly

$$d\tilde{\phi}_i(\sigma)=\frac{1}{1+n}\sum_{k\in\sigma}d\phi(ki)$$
$$=\frac{1}{1+n}\sum_k \phi(i)-\phi(k),$$
$$d\tilde{\phi}_i(\sigma)=\phi(i)-\langle\phi\rangle_\sigma, \tag{56}$$

where $\langle\phi\rangle_\sigma$ is the average value of ϕ in the simplex ϕ. Then, since $\sum_j\tilde{g}^{ij}=0$,

$$\tilde{g}^{ij}(\sigma)d\tilde{\phi}_j(\sigma)=\sum_{j\in\sigma}\tilde{g}^{ij}(\sigma)\phi(j)$$

and

$$L(\sigma)=\sum_{i,j\in\sigma}\tfrac{1}{2}\tilde{g}^{ij}(\sigma)\phi(i)\phi(j), \tag{57}$$

whence

$$S=\tfrac{1}{2}\sum_{\substack{i,j\in\Sigma_0\\ \sigma\in\Sigma_n}}\tilde{g}^{ij}(\sigma)\phi(i)\phi(j)V(\sigma). \tag{58}$$

[The sum is naturally over only those values which make sense—those for which $i\in\sigma$ and $j\in\sigma$. Equivalently one can define $\tilde{g}^{ij}(\sigma)\equiv(0)$ for all the nonsensical values.]

Varying $\phi(i)$:

$$\frac{\partial S}{\partial\phi(i)}=\sum_{j,\sigma}\tilde{g}^{ij}(\sigma)\phi(j)V(\sigma), \tag{59}$$

the vanishing of which constitutes the thatch equation for vertex i.

So far everything was general. We now specialize to various two-dimensional nets with flat metric. To evaluate $\tilde{g}^{ij}(\sigma)$, the following formula, which can be proved by the methods of Sec. II C, will prove very convenient:

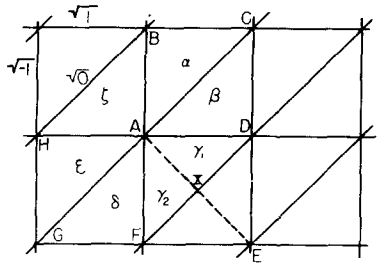


FIG. 1. A rectangular net for a two-dimensional flat space—time. The diagonal lines are lightlike.

$$\langle F(i), F(j)\rangle = [\tilde{g}/(n-1)!\,]\tilde{g}^{ij} = \pm nn!\,V^2\tilde{g}^{ij}. \tag{60}$$

Here everything relates to a particular n-simplex; $F(i)$ is the oriented face opposite to the vertex i, V the volume of the simplex, and

$$n!\,\tilde{g} = \tilde{\epsilon}^{i\cdots j}\tilde{g}_{ia}\cdots\tilde{g}_{jb}\tilde{\epsilon}^{a\cdots b}, \tag{61}$$

of course.

Work first with the net of Fig. 1 (without the dotted line), and consider the equation of vertex A. Because all the cells have the same volume V, Eq. (59) becomes

$$\sum_{\sigma}\sum_{j\in\sigma}\tilde{g}^{ij}(\sigma)\phi(j) = 0$$

or, in view of (60),

$$\sum_{\sigma}\sum_{j\in\sigma}\langle F(i), F(j)\rangle\phi(j) = 0. \tag{62}$$

There are two types of cell in the net, of which α and β are exemplars. For α one finds from (60) (order: $A\ B\ C$)

$$\tilde{g}^{ij}(\alpha) = \begin{matrix} -1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 1 \end{matrix}, \tag{63}$$

and from this $\tilde{g}^{ij}(\beta)$ must be (order: $A\ D\ C$)

$$\tilde{g}^{ij}(\beta) = \begin{matrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{matrix}. \tag{64}$$

The equation of A is then

$$[\tilde{g}^{AA}(\alpha)+\tilde{g}^{AA}(\beta)+\tilde{g}^{AA}(\gamma)+\tilde{g}^{AA}(\delta)+\tilde{g}^{AA}(\epsilon)+\tilde{g}^{AA}(\zeta)]\phi(A)$$
$$+[\tilde{g}^{AB}(\zeta)+\tilde{g}^{AB}(\alpha)]\phi(B)+[\tilde{g}^{AC}(\alpha)+\tilde{g}^{AC}(\beta)]\phi(C)$$
$$+\cdots+[\tilde{g}^{AH}(\epsilon)+\tilde{g}^{AH}(\zeta)]\phi(H) = 0$$

or

$$(-1+1+0-1+1+0)\phi(A)+(1+1)\phi(B)+(0+0)\phi(C)$$
$$+(-1-1)\phi(D)+\cdots+(-1-1)\phi(H) = 0, \tag{65}$$

$$\phi(B)-\phi(D)+\phi(F)-\phi(H) = 0, \quad \phi(B)+\phi(F) = \phi(H)+\phi(D),$$

which is exactly the equation used by the method of finite differences, in place of $\Box^2\phi = 0$ (in two dimensions).

It is remarkable that $\phi(A)$, $\phi(C)$, $\phi(G)$ drop out of the equation completely. It is also odd that the vertices of the net fall into two variationally unrelated subsets, but there seems to be no way to set up a net which avoids this and still has basic equations of the type (65). The net indicated in Fig. 2, for example, relates every point to every other, but through the typical equations

$$\phi(a)+2\phi(h)+\phi(d) = \phi(b)+\phi(c)+\phi(e)+\phi(f), \tag{66}$$

which could be thought of as the sum of the two equations

$$\phi(a)+\phi(h) = \phi(f)+\phi(b) \quad\text{and}\quad \phi(d)+\phi(h) = \phi(c)+\phi(e).$$

The most disconcerting phenomenon implied by (62) is that of the totally unrelated vertex as illustrated by Fig. 3. The subnet pictured, which might be the refinement indicated by the dotted line in Fig. 1, consists of four cells (triangles). According to (58) their contribution to the action is a sum of terms in $\phi(X)\phi(X)$, $\phi(X)\phi(A), \ldots, \phi(A)\phi(F), \cdots$. From (63) and (64) the coefficient of $\phi(X)\cdot\phi(X)$ is

$$\tfrac{1}{8}\sum_{j=1}^{4}\tilde{g}^{XX}(j) = \tfrac{1}{8}(-4+4-4+4) = 0,$$

while that of $\phi(X)\phi(A)$, e.g., is,

$$\tfrac{1}{4}\sum_{j=1}^{2}\tilde{g}^{XA}(j) = \tfrac{1}{4}(2-2) = 0.$$

In other words $\phi(X)$ drops out of the action completely! In fact the expression for S is the same for both nets: The dotted line makes no difference.

Lest all these surprises give the impression that the simplicial approach is especially productive of anomalies, we should add that for any other than the 1—1 ratio of sides, the net of Fig. 1 reproduces exactly the equation of the usual finite difference approximation. And, though we have stuck to flat space—time, the simplectic scheme comes into its own only with a curved background metric—which it handles with no extra trouble.

As a final example we take the two-dimensional potential equation. Using a "square" net with the topology

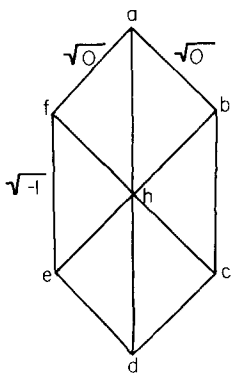


FIG. 2. A simple net in two dimensions.

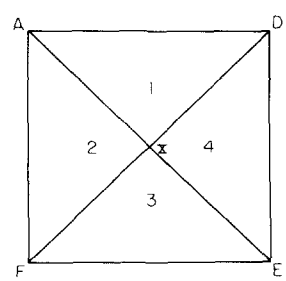


FIG. 3. A refinement of the net of Fig. 1. $\phi(x)$ makes no contribution to the action.

pictured in Fig. 1 one finds for vertex A, e.g., the equation

$$4\phi(A)=\phi(H)+\phi(D)+\phi(B)+\phi(F), \tag{67}$$

which just says that $\phi(A)$ is the average of the neighboring values. And, of course, this is the well-known characteristic feature of a solution of $\nabla^2\phi=0$.

VI. THE CHARACTER OF THE CONTINUUM LIMIT

A. General considerations

All the thatch equations discussed above or in Ref. 1 share this feature: The thatch represents a true field but one of a very simple (possibly singular) type. Thus $A(jk)$ corresponds to a piecewise linear electromagnetic potential, while $l^2(jk)$ defines a piecewise flat manifold. In terms of these fields one defines the action in the usual way; then the thatch equations just assert the stationarity of the action—but only for variations which maintain the correspondence of the field to some thatch. By using ever finer nets one allows for ever more delicate variations of the field so that, in the limit of an infinitely fine net, one expects the solution thatch to correspond exactly to the true field.[6]

On the other hand, as we will see below in particular examples, it is in general *false* that the limit of a particular thatch equation is the correct field equation at that point. In other words, the discretization of an exact continuum solution will not produce a solution of the thatch equations, even in the continuum limit!

To understand this better, remember that $A(jk)$, for example, corresponds to a piecewise linear field. At any given point this can agree with $A_\mu(x)$ only to terms of the first order in dx (precisely those involved in the definition of the action!); it can reflect the second derivatives of A_μ only *on the average* over a small region. Thus one can expect A_μ and even $F_{\mu\nu}$, but not $\partial_\mu F_{\mu\nu}$ to become exact in the continuum limit. The field equations $\partial_\mu F_{\mu\nu}=0$ can become exact only after averaging.

We can arrive at this conclusion again by a somewhat different argument. Let Ω be a region of spacetime and consider for simplicity the scalar thatch $\phi(j)$. In the continuum limit δS must vanish for any smooth variation $\delta\phi(x)$ of the field $\phi(x)$. In particular, it must vanish for the variation $\delta\phi=\text{const}$ within Ω, $\delta\phi=0$ outside. But for such a variation δS is just the sum

$$\delta\phi\sum_{j\in\Omega}\frac{\partial S}{\partial\phi(j)}\,.$$

(The boundary terms are negligible if the net is sufficiently fine.) We conclude that even though the thatch equations $\partial s/\partial\phi(j)$ may fail individually, their sum

$$\sum_{j\in\Omega}\frac{\partial S}{\partial\phi(j)}=0 \tag{68}$$

over any finite region Ω will be valid.

For the thatches A, l^2 the same argument applies except that, in place of $\delta\phi=\text{const}$, one must put a variation of the $l^2(ij)$ [respectively $\delta A(ij)$] which corresponds to $\delta g_{\mu\nu}=\text{const}$ [resp. $\delta A=\text{const}$]. There will be ten [resp. six] linearly independent such variations.

B. Illustration

The simplest example of these considerations is the flat scalar wave equation in two dimensions. We will examine the thatch equation (59) for various nets and show that while the continuum limit of (59) is always a homogeneous second order differential equation, it is sometimes the wrong one. As expected, however, an appropriate sum of these equations (over a "unit cell" of the lattice) always reduces to the correct equation in the continuum limit.

Let us write (59) for the vertex $[0]\in\Sigma_0$ in the form

$$\sum_k \mu(k)\phi(k)=0, \tag{69}$$

where

$$\mu(k)\equiv\sum_\sigma \tilde{g}^{0k}(\sigma)V(\sigma) \tag{70}$$

and k ranges over $\mathfrak{S}_0(\mathfrak{S}_1([0]))$. If we expand $\phi(x)$ about $[0]$ (assuming flat space–time recall), then

$$\phi(j)=\phi(0)+\phi'(0)\cdot\overrightarrow{0j}+\tfrac{1}{2}\phi''(0)\cdot\overrightarrow{0j}\otimes\overrightarrow{0j}+\cdots, \tag{71}$$

and (69) becomes

$$\phi(0)\sum_k\mu(k)+\phi'(0)\cdot\sum_k\mu(k)\overrightarrow{0k}$$
$$+\tfrac{1}{2}\phi''(0)\cdot\sum_k\mu(k)\overrightarrow{0k}\otimes\overrightarrow{0k}+\cdots=0. \tag{72}$$

Since one can prove in general (flat space) that

$$\sum_k\mu(k)=0,\quad\sum_k\mu(k)[k]=0, \tag{73}$$

(72) becomes, to second order in $\overrightarrow{0k}$,

$$\tfrac{1}{2}\phi''(0)\cdot\sum_k\mu(k)[k]\otimes[k]=0. \tag{74}$$

[The notation of the second equation of (73) makes sense because of the first, just as that of (74) in turn makes sense because of (73).] In the continuum limit this has the form

$$a^{\mu\nu}\partial_\mu\partial_\nu\phi(0)=0. \tag{75}$$

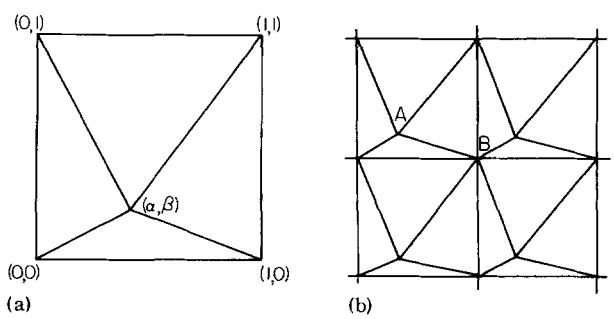
Unfortunately, $a^{\mu\nu}\neq g^{\mu\nu}$ in general.

Consider, for example, the star shown in Fig. 4(a). The corresponding equation (75) works out as

$$-\frac{1}{2}(1+q)\frac{\partial^2\phi}{\partial x^2}-\frac{1}{2}\left(1+\frac{1}{q}\right)\frac{\partial^2\phi}{\partial y^2}=0, \tag{76}$$

in which

$$q=\alpha(1-\alpha)/\beta(1-\beta).$$



(a) (b)

FIG. 4. (a) A star in flat two-dimensional space–time. The vertices are labelled by their rectangular coordinates. (b) A complete net made up of repetitions of this star.

(For convenience we deal with a positive definite metric.) This differs from the correct $\nabla^2\phi = 0$ unless $q = 1$, i.e., unless (α, β) lies on one of the diagonals of the square.

On the other hand, for a vertex such as B in Fig. 4(b), (75) becomes, with the same normalization,

$$-\frac{1}{2}(3-q)\frac{\partial^2\phi}{\partial x^2} - \frac{1}{2}\left(3-\frac{1}{q}\right)\frac{\partial^2\phi}{\partial y^2} = 0. \tag{77}$$

Since on the average there are equal numbers of vertices of types A and B, the average thatch equation is the average of (76) and (77):

$$-\nabla^2\phi = 0,$$

which is the correct continuum equation.

Notice that, in forming the average thatch equation, it was enough to consider one equation for each type of net vertex. In a net where all vertices were equivalent, each individual equation would already be completely typical. This explains why the nets of Part V produced the correct continuum limit without any averaging process.

ACKNOWLEDGMENTS

Jim Gunn was very helpful during the development of these ideas, eapecially those concerning the continuum limit.

*Supported in part by the National Science Foundation [MPS75-01398].

†Present address: Department of Applied Mathematics and Astronomy, University College, Cardiff CF1 1XL, Wales.

[1]R. Sorkin, "The Time Evolution Problem in Regge Calculus," Phys. Rev. D **12**, (1975), to appear.

[2]An "affine space" is just a vector space in which no point is distinguished as the "origin." Its symmetry group thus includes translations as well as the linear maps appropriate to a vector space.

[3]J. A. Wheeler, in *Relativity, Groups and Topology*, edited by C. DeWitt and B. DeWitt (Gordon and Breach, New York, 1964).

[4]k-forms are also called k-cochains in combinatorial topology.

[5]The factor $\frac{1}{2}$ avoids double counting of legs in (40).

[6]Unfortunately, the action is not positive definite (hyperbolic equation); so there will also be some requirement for overall stability. That is not discussed at all here.

Dynamic stability and thermodynamics in kinetic theory and fluid mechanics

Miroslav Grmela

Centre de Recherches Mathématiques, Université de Montréal, Montréal, Québec, Canada H3C 3J7
(Received 24 June 1975)

The study of the relationship between thermodynamic and local dynamic stability that has been developed in the previous papers is further extended. The dynamical systems considered include the multicomponent fluid dynamics and the two component Enskog–Vlasov dynamics.

I. INTRODUCTION

Nonequilibrium statistical mechanics introduces a family of dynamical systems $\mathcal{S}$. Every element $s \in \mathcal{S}$ of this family can be regarded as a sum of the experience obtained from measuring and observing the time development of a class of physical systems C_s. The set of observations and measurements O_s used, i.e., the empirical base of the dynamical system s, has the property $O_s \neq O_{s'}$ if and only if $s \neq s'$. Only results of the measurements O_s can decide whether the time development of a given physical system is described by the dynamical system s, i.e., whether the given physical system is an element of C_s. In order to distinguish the physical systems inside C_s a set of phenomenological quantities ρ_s is introduced by every dynamical system $s \in \mathcal{S}$. Thus, in fact, every $s \in \mathcal{S}$ is again a family of dynamical systems parametrized by ρ_s. ρ_s is an element of a set $\{\rho_s\}$. The map between the elements of $\{\rho_s\}$ and the elements of C_s can be obtained again only as a result of measurements and observations O_s.

Two questions arise naturally. The first question is as to whether there exist some properties common for all $s \in \mathcal{S}$. These properties could be then used to characterize $\mathcal{S}$ in a space of all dynamical systems. Two arguments supporting the existence of such common properties can be mentioned. Experience shows that one can find, for every physical system c, the sets of observations and measurements $O^{(c)}_{s_H}$ and $O^{(c)}_{s_T}$, such that $c \in C_{s_H}$, $c \in C_{s_T}$ and s_H is the Hamiltonian dynamics and s_T is thermodynamics. In other words, the experience seems to indicate that for every physical system c there is $O^{(c)}_{s_H}$ such that the time development measured is the time development derived from a Hamiltonian dynamical system (e.g., classical or quantum mechanics) and also $O^{(c)}_{s_T}$ such that the relations among the time independent properties of the system c measured are the relations derived from thermodynamics. The second question is as to whether and how the dynamical system s_{II} associated with $O_{s_{II}}$ can be derived from s_I associated with O_{s_I}, if $C_{s_{I,II}} = C_{s_I} \cap C_{s_{II}}$ is not empty and O_{s_I} consists of observations and measurements that are more detailed than the ones that are elements of $O_{s_{II}}$. In particular, one is interested in finding a characterization of $C_{s_{I,II}} \subset C_{s_I}$ in terms of some restrictions in $\{\rho_{s_I}\}$ and in finding the map $\{\rho_{s_I}\} \to \{\rho_{s_{II}}\}$ for $c \in C_{s_{I,II}}$.

These two questions have been studied in Refs. 1,2, and the study is continued in this and in the subsequent paper,[3] for the family of dynamical systems consisting of three elements: fluid dynamics (the dynamical system s_{II}), dynamical systems introduced by kinetic theory (the dynamical systems s_I), and thermodynamics. In Sec. 2 of this paper a dynamical system is introduced and some of its properties are studied. Two particular cases of this dynamical system, namely the multicomponent fluid dynamics and the two component Enskog–Vlasov dynamics, are discussed in detail in Secs. 3 and 4. Thermodynamics is obtained as a result of (A) the restricted problem of fixed points, (B) the local dynamic stability of some fixed points, and (C) compatibility of the problems (A), (B). We say that the problems (A), (B) are compatible if there exists a real valued function $\mathbb{V}$ such that the problem (A) is equivalent to the problem of critical points of $\mathbb{V}$ and the problem (B) is equivalent to an investigation of convexity of $\mathbb{V}$. The function $\mathbb{V}$ evaluated at its critical points is the thermodynamic potential. The conditions (in the form of restrictions in the phenomenological quantities) that guarantee the compatibility are found for all cases discussed. The theory naturally unifies and generalizes some results known in nonequilibrium thermodynamics, hydrodynamics, and in the theory of the Boltzmann equation. The results of physical nature come hand-in-hand with the results of mathematical nature (existence of solutions, etc.). The subsequent paper[3] uses the results of this paper to investigate in detail the relationship between the two component Enskog–Vlasov dynamics and the two component fluid dynamics.

2. A GENERAL THEORY

This section supplements and improves the general discussion developed in Ref. 2. The dynamical systems introduced here include as special cases all dynamical systems of fluid mechanics and kinetic theory discussed in Refs. 1,2 and Secs. 3 and 4 of this paper.

Let a state of the physical system considered be fully (with respect to the given set of measurements and observations) described by f. All admissible states (again with respect to the given set of measurements and observations) form a set H. The structure (algebraical, topological, etc.) of H should not be *a priori* postulated. It should appear as a result of the comparison between the properties of the trajectories of the dynamical system and the observed time evolution. It is necessary however to start with some structure. We shall assume that H is a smooth manifold, locally topologically isomorphic to a complete Hilbert space H with the inner product (φ, ψ), $\varphi, \psi \in H$. In all cases discussed in this and the subsequent paper, H will be the L_2 space and

$(\,,\,)$ will denote the standard inner product in the L_2 space. The norms of the tangent spaces $T_f\mathcal{H}$ of $\mathcal{H}$ at f are not necessarily equivalent to the norm in H. A general inner product in $T_f\mathcal{H}$ is $\langle\varphi,\psi\rangle=(\varphi,A_f\Psi)$, where φ, $\psi\in T_f\mathcal{H}$, and A_f is a bounded self-adjoint positive definite linear operator. Information about A_f will be obtained from the discussion of dynamics in $\mathcal{H}$. Moreover, an involution $\mathcal{J}:\mathcal{H}\to\mathcal{H}$ is defined. In order to simplify our notation, we shall also use the same symbol $\mathcal{J}$ for the linear map induced by $\mathcal{J}$ on the tangent spaces of $\mathcal{H}$. The subset of $\mathcal{H}$ that is invariant with respect to $\mathcal{J}$ is denoted $\mathcal{H}^+$, its complement $\mathcal{H}^-$; thus $\mathcal{H}\equiv\{\mathcal{H}^+,\mathcal{H}^-\}$. We shall assume that $\mathcal{J}A_f\mathcal{J}=A_f$ for all $f\in\mathcal{H}$. As reflection of the fact that all dynamical systems introduced by nonequilibrium statistical mechanics are related to Hamiltonian mechanics (see the Introduction), we can consider $\mathcal{J}$ as inherited from the involution $\mathcal{J}_{\rm cm}$ in classical mechanics defined by reversing velocities of all particles composing the system considered. Besides $\mathcal{H}$, we shall need another smooth manifold, H^* that is defined as an image of $\mathcal{H}$ under the transformation $T:\mathcal{H}\to H^*$. The transformation T is a part of the phenomenological quantities ρ. Let T_fT denote the first derivative of T at $f\in\mathcal{H}$. We shall assume that $\mathcal{J}T_fT\mathcal{J}=T_fT$ for all $f\in\mathcal{H}$. Any vector field $R:\mathcal{H}\to T\mathcal{H}$, where $T\mathcal{H}$ denotes the tangent bundle of $\mathcal{H}$, can be split into the even and odd part $R^\pm=\frac{1}{2}(R\pm\mathcal{J}R\mathcal{J})$. We shall assume that the vector field generating the time development of f has the following form:

$$\begin{aligned}\frac{\partial f}{\partial t}&=R^-_{\rho^-}(f)+R^+_{\rho^+}(f)\\&=R^*_{\tilde\rho^-}(f,f^*)+R^*_{\tilde\rho^+}(f^*).\end{aligned}\tag{2.1}$$

The phenomenological quantities $\{\rho\}$ introduced by (2.1) are $\{\rho\}\equiv\{\rho^+,\rho^-\}\equiv\{\tilde\rho^+,\tilde\rho^-,T\}$. For later use, we introduce also $\tilde\rho^{++}$ that denotes the complement of $\tilde\rho^+\cup\tilde\rho^-$ in $\tilde\rho^+$. The Eq. (2.1) defines a family of vector fields parametrized by ρ. Not all of these vector fields define a dynamical system, in particular then, not all of them define a dynamical system introduced by nonequilibrium statistical mechanics. We shall find $\{\rho\}_{ph}\subset\{\rho\}$ such that (2.1) represents local dynamical systems and these dynamical systems are elements of $\mathcal{S}$, i.e., elements of the family of dynamical systems introduced by nonequilibrium statistical mechanics. By the local dynamical system, we mean the dynamical system generated by the linearized vector field at a fixed point of the vector field. The meaning of the statement that a dynamical system is an element of $\mathcal{S}$ will be made precise in the Secs. 2A, 2B, and 2C.

First we shall restrict ourselves to a subset $\{\rho\}_+$ of $\{\rho\}$ such that the following conditions (2.$\mathcal{A}$a) and 2.$\mathcal{A}$b) are satisfied:

The solutions of $R^{*+}_{\tilde\rho^+}(f^{*+})=0$, where $f^{*+}\in\mathcal{H}^{*+}$, consist of a family of submanifolds $\mathcal{H}^*_{0,q}$ of $\mathcal{H}^*$ parametrized by a set of parameters q. $\mathcal{H}^*_{0,q}$ is independent of $\tilde\rho^{++}$. $\mathcal{H}^*_{0,q}$ is not the whole manifold $\mathcal{H}^*$ for any q. (2.$\mathcal{A}$a)

Let P^{*+}_0 denote the vector field $R^{*+}_{\tilde\rho^+}$ linearized at $f^*_0=Tf_0$; f^*_0 is the solution of $R^{*+}_{\tilde\rho^+}f^*=0$ independent of the position coordinates of the system considered. $H^*_0=T_{f_0}\mathcal{H}^*$ is isomorphic to H and H^{*c} denotes the complement of $T_{f_0}\mathcal{H}^*_{0,q}$ in H^*_0. We shall assume that

P^{*+}_0 is a densely defined, closed, self-adjoint, negative definite operator in H^{*c}_0. (2.$\mathcal{A}$b)

We shall see that (2.$\mathcal{A}$a) and (2.$\mathcal{A}$b) are closely related. Their physical meaning will become clear in the Examples at the end of this Sec. and in Secs. 3 and 4.

It is well known from experience that for every physical system there is a set of observations and measurements O_T, such that if O_T is applied to the system prepared before (in most cases the preparation consists in leaving the system unperturbed from an outside influence for a sufficiently long time) then the experience obtained results in the theory called thermodynamics. The state of the prepared system is called an equilibrium state. Thermodynamics postulates the existence of the equilibrium state. Other dynamical theories introduced by nonequilibrium statistical mechanics have to be able to derive the equilibrium state and thermodynamics in terms of their description of states and the time development equations. In our case the structure of thermodynamics must be obtained from (2.1), and expressed through the phenomenological quantities ρ. One can think of several possibilities. For example, thermodynamics might appear as a theory of fixed points of (2.1), supplemented possibly by some restrictions, boundary conditions, etc. The fact that thermodynamics does not include the time dependence does not mean that it must appear only through a study of fixed points. A study of invariant manifolds or a study of ergodicity-type properties of trajectories (as it is widely used in the case of the Hamiltonian dynamics) might replace the study of fixed points.

In the context of (2.1), (2.$\mathcal{A}$a), and (2.$\mathcal{A}$b), concretely then in the context of fluid mechanics and kinetic theory, thermodynamics can be obtained as the result of solutions to the following three problems: (A) restricted problem of fixed points, (B) local dynamical stability of some fixed points, (C) compatibility of (A), (B).

A. Restricted problem of fixed points

The equilibrium state f_{eq} is defined as a solution of the following restricted problem of fixed points

$$\begin{aligned}&f\in\mathcal{H}^+,\\&R^+f^+=0,\\&R^-f^+\big|_{R^+f^+=0}=0.\end{aligned}\tag{2.2}$$

The last equation means $R^-f^+=0$ restricted to the solutions of $R^+f^+=0$. By f_0 we denote the solution of (2.2) that is independent of the position coordinates of the system.

B. Local dynamic stability of f_0

The tangent space $H_0=T_{f_0}\mathcal{H}$ is by assumption a Hilbert space topologically isomorphic to the L_2 space. The inner product in H_0 is denoted by $\langle\varphi,\psi\rangle\equiv(\varphi,A\psi)$, where φ,

$\psi \in H_0$, and A is a bounded self-adjoint positive definite linear operator, i.e., $H_0 \to H_0$, $\mathcal{J}A\mathcal{J} = A$. We shall say that f_0 is locally stable if and only if there exists a unique solution φ_t to the initial value problem

$$\begin{aligned} &\frac{\partial \varphi}{\partial t} = P_0\varphi, \\ &\varphi_{t=0} = \varphi_0 \in D(P_0) \subset H_0, \\ &\|\varphi_0\| \text{ is bounded,} \end{aligned} \tag{2.3}$$

for all $0 \leq t < \infty$ and $\|\varphi_t\| = (\langle \varphi_t, \varphi_t \rangle)^{1/2}$ bounded for all $0 \leq t < \infty$. P_0 denotes the vectorfield R linearized at f_0, $D(P_0)$ denotes the domain of P_0. By using the Hille–Yoshida–Phillips theory,[4] the necessary and sufficient condition for the local dynamic stability of f_0 is: (i) P_0 is densely defined and closed; (ii) both P_0 and its adjoint $P_0^\dagger$ are dissipative, i.e., $\langle \varphi, P_0\varphi \rangle \leq 0$ for all $\varphi \in D(P_0)$, and $\langle \varphi, P_0^\dagger \varphi \rangle \leq 0$ for all $\varphi \in D(P_0^\dagger)$. Equivalently, there exists an operator A satisfying the properties listed above such that $(\varphi, AP_0\varphi) \leq 0$ for all $\varphi \in D(P_0)$, and $(\varphi, (AP_0)^\dagger \varphi) \leq 0$ for all $\varphi \in D(P_0^\dagger)$.

C. Compatibility of the problems 2A and 2B

We shall say that the restricted problem of fixed points (the problem 2A) and the problem of the local dynamic stability (the problem 2B) are compatible if and only if there exists a function $\mathbb{V}: H \times U \to \mathbb{R}$, $U \subset \mathbb{R}^m$, m is a positive integer, such that the problem 2A is equivalent to $\delta\mathbb{V}/\delta f = 0$ and the linear operator $A: H_0 \to H_0$ that solves the problem 2B is equal to $\delta^2\mathbb{V}/\delta F(f)\delta F(f')$, where $F: H \to H$ is an invertible transformation. We shall be able to find $\{\rho\}_{ph} \subset \{\rho\}_* \subset \{\rho\}$, such that the problems 2A and 2B are compatible, and also the functions $\mathbb{V}$ for all the special cases of (2.1) discussed in Secs. 3 and 4. In all these cases the function $\mathbb{V}$ appears to be linear in the variables $\sigma = (\sigma_1, \ldots, \sigma_m) \in U$, i.e.,

$$\mathbb{V}(f;\sigma) = S(f) + \sum_{i=1}^{m} \sigma_i v_i(f), \tag{2.4}$$

where S and v_i, $i = 1, \ldots, m$, are functions $H \to \mathbb{R}$. The function $\mathbb{V}_{th}: U \to \mathbb{R}$ is defined as $\mathbb{V}$ restricted to the solutions of $\delta\mathbb{V}/\delta f = 0$. $\mathbb{V}_{th}$ is interpreted as the thermodynamic potential of the physical system considered. More precisely, $\sigma_1, \ldots, \sigma_m$ are the intensive thermodynamic parameters (thermodynamic fields in the terminology of Griffiths and Wheeler[5]), $\sigma_{m+1} = \sigma_{m+1}(\sigma_1 \cdots \sigma_m) \equiv \mathbb{V}(f_{eq}(\sigma_1, \ldots, \sigma_m), \sigma_1, \ldots, \sigma_m)$, and f_{eq} denotes a solution of $\delta\mathbb{V}/\delta f = 0$ (we shall not discuss in this paper the case when this equation has more solutions, physically, the case of phase transitions). From (2.4), we have: (i) Let s_i, $i = 1, \ldots, m$, denote the thermodynamical conjugates of σ_i, i.e., $s_i = \partial\mathbb{V}_{th}/\partial\sigma_i$, $i = 1, \ldots, m$. From the definition of $\mathbb{V}_{th}$ and from (2.4) one obtains immediately $s_i = v_i(f_{eq})$. This equality provides useful information about the relation between f and the thermodynamic observations and measurements. (ii) Let $\mathbb{V}$ be defined as in (2.4) and let $\mathbb{V}_\sigma: H \to \mathbb{R}$; $(f, \sigma) \overset{\mathbb{V}}{\mapsto} \mathbb{V}(f, \sigma)$ reaches its nondegenerate maximum at f_σ for all $\sigma \in U_0$, where U_0 is an open neighborhood of $\sigma_0 \in U_0$, $0 \in U_0$. Then $\mathbb{V}_{th}: U_0 \to \mathbb{R}$, and is convex at σ_0. Indeed $\mathbb{V}(f_{\sigma_0+\sigma_1}; \sigma_0 + \sigma_1) = \mathbb{V}(f_{\sigma_0+\sigma_1}; \sigma_0) + \mathbb{V}(f_{\sigma_0+\sigma_1}; \sigma_1) \leq \mathbb{V}(f_{\sigma_0}; \sigma_0) + \mathbb{V}(f_{\sigma_1}; \sigma_1)$; σ_0, σ_1, $\sigma_0 + \sigma_1 \in U_0$. In view of the thermodynamic interpretation of $\mathbb{V}_{th}$, we have shown the relationship between the convexity of the function $\mathbb{V}$ with respect to its dependence on f and the convexity of $\mathbb{V}_{th}$ with respect to its dependence on σ, or equivalently the thermodynamic stability condition for $\mathbb{V}_{th}$.

The general considerations developed in this section will be used in Secs. 3 and 4 in the context of fluid mechanics and kinetic theory. An additional interesting illustration of the general considerations of this section can be obtained from the following two examples.

Example 1

The classical Liouville equation describing the classical dynamics of N identical particles can be considered as a special (degenerate) case of (2.1). The assumption (2.$\mathcal{A}$a) is not satisfied since the solution of $R^{*+}f^{*+} = 0$ is the whole manifold H. The manifold H is the linear L_2 space, $H^* \equiv H$ is its dual (i.e., $\mathcal{T}$ is the identity map). The phenomenological quantities P are the Hamiltonians. The involution $\mathcal{J}$ is defined by $f(\mathbf{r}_i, \mathbf{v}_i) \overset{\mathcal{J}}{\to} f(\mathbf{r}_i - \mathbf{v}_i)$, $i = 1, \ldots, N$, where $(\mathbf{r}_i \mathbf{v}_i)$ denotes the position coordinate and velocity of the ith particle respectively, and $H \ni f$: $\mathbb{R}^{3N} \times \mathbb{R}^{3N} \to \mathbb{R}$. By using Stone's theorem (Stone's theorem may be considered as a special case of the Hille–Yoshida–Phillips theorem used in Sec. 2B[4]), we find that any state $f \in D \subset H$, where D is the domain of the Liouville operator, is locally (and in this case also globally) stable. Thus, the function $\mathbb{V}$ can be for example $\mathbb{V} = K(f,f)$, where K is a positive constant, but the restricted problem of fixed points (2.2) is not equivalent to the problem of the critical points of $\mathbb{V}$. It is an easy exercise to see that if in (2.2) $R^+f^+ = 0$, which is trivial in this case, is replaced by an ad hoc restriction $f = f_{th} = N(\beta)\, n(\mathbf{r}_1 \cdots \mathbf{r}_N) \exp[-\frac{1}{2}\beta(v_1^2 + \cdots + v_N^2)]$ and f_{th} are called thermalized functions[6], β is a positive constant, $n(\mathbf{r}_1 \cdots \mathbf{r}_N) > 0$, and $N(\beta)$ is determined by the requirement $\int d^3\mathbf{v}_1 \cdots d^3\mathbf{v}_N f_{th} = n(\mathbf{r}_1 \cdots \mathbf{r}_N)$, then there exists a single function

$$\begin{aligned} \mathbb{V}_L(f;\beta,\alpha) = \int d^3\mathbf{v}_1 \cdots d^3\mathbf{v}_N \int d^3\mathbf{r}_1 \cdots d^3\mathbf{r}_N [&f \ln f \\ &+ \beta V_{pot}(\mathbf{r}_1 \cdots \mathbf{r}_N) f + \tfrac{1}{2}\beta(v_1^2 + \cdots v_N^2) f - \alpha f] \end{aligned}$$

that has the required properties, i.e., $\delta\mathbb{V}_L/\delta f = 0$ is equivalent to (2.2) where $R^+f^+ = 0$ is replaced by $f = f_{th}$ and

$$A \equiv \frac{\delta^2 \mathbb{V}_L}{\delta f \delta f'}\bigg|_{f_0}.$$

The quantity V_{pot} denotes the potential energy of N particles.

Example 2

Let H be n-dimensional manifold, $H^- \equiv 0$, $(f_1, \ldots, f_n) \overset{\mathcal{T}}{\to} (F_1(f), \ldots, F_n(f))$, $R^- \equiv 0$ and R^{*+} satisfies (2.$\mathcal{A}$a) and (2.$\mathcal{A}$b). For example, $R^{*+}f^* \equiv -\operatorname{grad} \mathbb{V}^{*+}$, where $\mathbb{V}^{*+}$ is a real valued function that has one nondegenerate minimum at f_0^*, determined by $F_1(f) = \alpha_1, \ldots, F_n(f) = \alpha_n$. The equilibrium state $f_0 = \mathcal{T}^{-1}f_0^*$ is locally stable if and only if the matrix $B \equiv (B_{ij}) \equiv (\partial F_i(f)/\partial f_j)|_{f_0}$, is an invertible, self-adjoint, positive definite matrix.

[*Proof*: The linearized vector field is $P_0 = R^{*+}B$. By using the Liapunov theorem,[7] or the Hille–Yashida–Phillips theory, the equilibrium state f_0 is locally stable

if and only if the equation $A^{\dagger}R^{*+}B + B^{\dagger}R^{*+}A = KA^{\dagger}R^{*+}A$ (K is a positive constant) or equivalently $R^{*+}(BA^{-1}) + (BA^{-1})^{\dagger} \times R^{*+} = KR^{*+}$ has a unique solution A that is a symmetric positive definite, nonsingular matrix. Clearly the only solution of the above equation is $KA = 2B$]. Thus, we see that the requirement of local stability of f_0 restricts the choice of $\bar{T}$ and brings a natural inner product for H_0.

3. n-COMPONENT FLUID DYNAMICS

The state of a system in the n-component fluid dynamics is completely described by $f \equiv (C_1, \ldots, C_{n-1}, E, N, \mathbf{U})$, where $C_1, \circ\circ\circ, C_{n-1}$, E, and N are functions $\Omega \to \mathbb{R}^+$ of class at least C^2 and $\mathbf{U}$: $\Omega \to T\Omega = \Omega \times \mathbb{R}^3$) also of class at least C^2. Ω is a bounded open subset of $\mathbb{R}^3$ in which the physical system considered is confined; $\mathbb{R}^+$ denotes the positive real line. We shall assume that Ω has a smooth boundary $\partial\Omega$ and the volume of Ω equals to one. Elements of Ω, position vectors, are denoted by $\mathbf{r}$. $T\Omega$ denotes the tangent bundle of Ω. The physical meaning of f, that will appear later from the study of thermodynamics is the following: $C_1, \ldots, C_{n-1}$ are the local concentrations of the components $1, \ldots, n-1$; E is the local inner energy; N is the local density and $\mathbf{U}$ is the local velocity. The involution $\mathcal{J}$ is defined by $(C_1, \ldots, C_{n-1}, E, N, \mathbf{U}) \xrightarrow{\mathcal{J}} (C_1, \ldots, C_{n-1}, E, N, -\mathbf{U})$. Thus, H^+ is composed of all quantities that behave as scalars under an orthogonal transformation of coordinates in Ω; H^- is composed of all vectors. The importance of the involution $\mathcal{J}$ and of the corresponding decomposition of H is thus evident in this case.

We shall introduce the following family of the time development equations for f:

$$\frac{\partial(NC_i)}{\partial t} = -\frac{\partial}{\partial r_\alpha}(U_\alpha NC_i) + \sum_{k=1}^{n} L_{ik}\frac{\partial^2}{\partial r_\alpha \partial r_\alpha} C_k^* + L_{in}\frac{\partial^2}{\partial r_\alpha \partial r_\alpha} E^*,$$

$$\frac{\partial(NE)}{\partial t} = -\frac{\partial}{\partial r_\alpha}(U_\alpha NE) - N^*\frac{\partial U_\alpha}{\partial r_\alpha} + \sum_{k=1}^{n} L_{nk}\frac{\partial^2}{\partial r_\alpha \partial r_\alpha} C_k^* + L_{nn}\frac{\partial^2}{\partial r_\alpha \partial r_\alpha} E^*, \quad (3.1)$$

$$\frac{\partial N}{\partial t} = -\frac{\partial}{\partial r_\alpha}(U_\alpha N),$$

$$\frac{\partial(NU_\alpha)}{\partial t} = -\frac{\partial}{\partial r_\beta}(U_\beta NU_\alpha) - \frac{\partial N^*}{\partial r_\alpha} + \eta\frac{\partial^2 U_\beta^*}{\partial r_\beta \partial r_\beta} + (\tfrac{1}{3}\eta + \eta_v)\frac{\partial^2 U_\beta^*}{\partial r_\beta \partial r_\alpha}.$$

Only terms linear in first- and second-order derivatives with respect to the position coordinates are considered in (3.1). Quadratic terms, etc., in derivatives of $\mathbf{U}$ do not influence our discussion at all. Terms quadratic, etc., in derivatives of f^+ can be included provided the solution to the problem (2.2) [see (3.2)] is not changed. The equations (3.1) are not linear since the transformation $\bar{T}: H \to H^*$, $(C_1, \ldots, C_{n-1}, E, N, \mathbf{U}) \mapsto (C_1^*, \ldots, C_{n-1}, E^*, N^*, \mathbf{U}^*)$ is not linear. Moreover, we assume that $f^*(\mathbf{r})$ depends only on $f(\mathbf{r})$ at the same position vector $\mathbf{r}$. This assumption can be relaxed in expense of complications in the following calculations. The last two terms on the rhs of (3.1) correspond to $R^{*+}(f^*)$, the rest is $R^{*-}(f, f^*)$. The phenomenological quantities $\tilde{P}^{*+}$ are L_{ik}, $i,k = 1, \circ\circ\circ, n$, η, and η_v considered as functions of f^{*+}. There are no other phenomenological quantities (except $\tilde{P}^{*+}$ and $\bar{T}$) introduced by (3.1). The assumption (2.$\mathcal{A}$a) is satisfied if we assume that

$C_i^*|_{\partial\Omega} = \alpha_i^*$, $i = 1, \ldots, n-1$, $E^*|_{\partial\Omega} = \beta^*$, where where $\alpha_1^*, \ldots, \alpha_{n-1}^*, \beta^*$ are constants. Moreover, L_{ik}, $i, k = 1, \ldots, n$, is a nonsingular positive definite matrix for all $f^* \in H^*$ considered. (3.$\mathcal{A}$a)

If we assume moreover, that

$\eta_0 = \eta|_{f_0^*}$ and $(\frac{1}{3}\eta_0 + \eta_{v,0})$ are positive, $(\eta_{v,0} = \eta_v|_{f_0^*})$, and the boundary conditions for the local dynamics are taken as such[8] that P_0^{*+} is self-adjoint,

then also (2.$\mathcal{A}$b) is satisfied. (3.$\mathcal{A}$b)

3A. Restricted problem of fixed points

If (3.$\mathcal{A}$a) and (3.$\mathcal{A}$b) are satisfied, then $R^{*+}f^{*+} = 0$ is equivalent to $\partial\mathbb{V}^{*+}(f^{*+})/\partial f^{*+} = 0$, where

$$\mathbb{V}^{*+} = \sum_{i,k=1}^{n}\sum \frac{\partial}{\partial r_\alpha}(C_i^*)L_{ik}\frac{\partial}{\partial r_\alpha}(C_k^*)$$

(we have used the notation $E^* \equiv C_n^*$). The only critical point of $\mathbb{V}^{*+}$ is $(\partial/\partial r_\alpha)C_i^* = 0$, $i = 1, \ldots, n$. The family of sets $H_{0,q}^{*+}$ is thus $H_{0,q}^{*+} \equiv \{\alpha_1^*, \ldots, \alpha_{n-1}^*, \beta^*, N^*, 0\}$. Indeed $H_{0,q}^{*+}$ is independent of $\tilde{P}^{*+} = \{L_{ik}, \eta, \eta_v, i,k = 1, \ldots, n\}$. The set of parameters q is now $q \equiv \{\alpha_1^* \circ\circ \alpha_{n-1}^*, \beta^*\}$. The equation $R^- f^* = 0$ restricted to $H_{0,q}^{*+}$ is $\partial N^*/\partial r_\alpha$ $\partial N^*/\partial r_\alpha|_{H_{0,q}^{*+}} = 0$. Thus, the problem (2.2) is now

$$\begin{aligned} &U_\alpha = 0, \\ &C_i^* = \alpha_i^*, \quad i = 1, \ldots, n-1, \\ &E^* = \beta^*, \\ &N^*|_{H_{0,q}^{*+}} = P^*, \end{aligned} \quad (3.2)$$

where $\alpha_1^* \circ\circ \alpha_{n-1}^*$, β^*, and P^* are constants. The position coordinate independent solution f_0 of (3.2) is $f_0 \equiv (C_{1,0}, \ldots, C_{n-1,0}, E_0, N_0, 0)$.

3B. Local dynamic stability of f_0

The linear vector field P_0 obtained by linearizing (3.1) at f_0 generates the time evolution of $\phi \in H_0$,

$$\frac{\partial\phi}{\partial t} = P_0\phi, \quad (3.3)$$

where H_0 is the tangent space of H at f_0. One can check that P_0 can be written in the form

$$P_0 = L\mathcal{M},$$

where

$$L = \begin{matrix} \\ n-1 \\ +1 \\ 3 \end{matrix}\overbrace{\begin{matrix} \\ \\ \end{matrix}}^{\begin{matrix} n+1 \end{matrix}}\overbrace{\begin{matrix} \\ \\ \end{matrix}}^{\begin{matrix} 3 \end{matrix}}\begin{Bmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{Bmatrix}, \quad \mathcal{M} = \begin{matrix} \\ n-1 \\ +1 \\ 3 \end{matrix}\overbrace{\begin{matrix} \\ \\ \end{matrix}}^{\begin{matrix} n+1 \end{matrix}}\overbrace{\begin{matrix} \\ \\ \end{matrix}}^{\begin{matrix} 3 \end{matrix}}\begin{Bmatrix} \mathcal{M}_{11} & \mathcal{M}_{12} \\ \mathcal{M}_{21} & \mathcal{M}_{22} \end{Bmatrix},$$

$$\mathcal{L}_{11}=\frac{1}{N_0}\begin{bmatrix} L_0 & 0 \\ & 0 \\ 0\cdots 0 & 0 \end{bmatrix}\frac{\partial^2}{\partial r_\alpha \partial r_\alpha},\quad L_0=\begin{bmatrix} L_{11}\cdots L_{1n} \\ L_{1n}\cdots L_{mn} \end{bmatrix},$$

$$M_{11}=\begin{bmatrix} M_0 & m_1 \\ m_2 & m_3 \end{bmatrix},\quad M_0=\begin{bmatrix} \frac{\partial C_1^*}{\partial C_1} & \cdots\frac{\partial C_1^*}{\partial E} \\ \frac{\partial E^*}{\partial C_1} & \cdots\frac{\partial E^*}{\partial E} \end{bmatrix},$$

$$m_1=\begin{bmatrix} \frac{\partial C_1^*}{\partial N} \\ \frac{\partial N^*}{\partial N} \end{bmatrix},\quad m_2=\begin{bmatrix} \frac{\partial N^*}{\partial C_1}\cdots\frac{\partial N^*}{\partial E} \end{bmatrix},\quad m_3=\frac{\partial N^*}{\partial N},$$

$$\mathcal{L}_{12}=-\frac{1}{N_0}\frac{1}{k_0}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ P^*\frac{\partial}{\partial r_1} & P^*\frac{\partial}{\partial r^2} & P^*\frac{\partial}{\partial r_3} \\ (N_0)^2\frac{\partial}{\partial r_1} & (N_0)^2\frac{\partial}{\partial r_2} & (N_0)^2\frac{\partial}{\partial r_3} \end{bmatrix},$$

$$\mathcal{M}_{12}=-\eta_1\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{\partial}{\partial r_1} & \frac{\partial}{\partial r_2} & \frac{\partial}{\partial r_3} \end{bmatrix},$$

$$\mathcal{L}_{21}=-\frac{1}{N_0}\begin{bmatrix} 0\cdots\frac{\partial}{\partial r_1} \\ 0\cdots\frac{\partial}{\partial r_2} \\ 0\cdots\frac{\partial}{\partial r_3} \end{bmatrix},\quad \mathcal{M}_{21}\equiv 0,$$

$$\mathcal{L}_{22}=\frac{1}{N_0}\eta_0\begin{bmatrix} 1&0&0 \\ 0&1&0 \\ 0&0&1 \end{bmatrix}\frac{\partial^2}{\partial r_\alpha \partial r_\alpha},\quad \mathcal{M}_{22}=k_0\begin{bmatrix} 1&0&0 \\ 0&1&0 \\ 0&0&1 \end{bmatrix}$$

The quantities L, M_0, m_1, m_2, m_3 are understood to be evaluated at $\mathcal{T}f_0$. We have also used the notation $\eta_0=\eta|\mathcal{T}f_0$: $\eta_1=(\frac{1}{3}\eta+\eta_v)|\mathcal{T}f_0$; $k_0=\partial U^*{}_\alpha/\partial U_\alpha|\mathcal{T}f_0$. The operator P_0 is thus

$$P_0=\begin{bmatrix} (P_0)_{11} & (P_0)_{12} \\ (P_0)_{21} & (P_0)_{22} \end{bmatrix}=\begin{bmatrix} \mathcal{L}_{11}\mathcal{M}_{11} & \mathcal{L}_{12}\mathcal{M}_{22} \\ \mathcal{L}_{21}\mathcal{M}_{11} & \mathcal{L}_{21}\mathcal{M}_{12}+\mathcal{L}_{22}\mathcal{M}_{22} \end{bmatrix}.$$

Clearly,

$$P_0^+=\begin{bmatrix} (P_0)_{11} & 0 \\ 0 & (P_0)_{22} \end{bmatrix},\quad P_0^-=\begin{bmatrix} 0 & (P_0)_{12} \\ (P_0)_{21} & 0 \end{bmatrix}.$$

Following the discussion in Sec. 2C, and in view of the the assumptions in (3.$\mathcal{A}$a), and (3.$\mathcal{A}$b), we shall look for an operator A,

$$A=\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}, \tag{3.4}$$

that is invertible, self-adjoint, position definite and AP_0^+ is self-adjoint, dissipative and AP_0^- is skew-adjoint. The operators A_{12} and A_{21} in (3.4) are zeros because of the requirement $\mathcal{J}A\mathcal{J}=A$. For the purpose of the calculations, we introduce

$$A_{11}=\begin{bmatrix} & & a_{1,n+1} \\ & A_0 & \cdot \\ & & \cdot \\ & & a_{n+1,n} \\ a_{n+1,1} & \cdots & a_{n,n+1}\ a_{n+1,n+1} \end{bmatrix}.$$

Using the notation introduced above, we have

$$AP_0=\begin{bmatrix} A_{11}\mathcal{L}_{11}\mathcal{M}_{11} & A_{11}\mathcal{L}_{12}\mathcal{M}_{22} \\ A_{22}\mathcal{L}_{21}\mathcal{M}_{11} & A_{22}\mathcal{L}_{21}\mathcal{M}_{12}+A_{22}\mathcal{L}_{22}\mathcal{M}_{22} \end{bmatrix}.$$

By using basically the same arguments as in Example 2 of Sec. 2, we deduce from the requirement that AP_0^+ is self-adjoint, that $A_0=\tilde{K}M_0$, and that

$$A_{22}=\tilde{Q}k_0\begin{bmatrix} 1&0&0 \\ 0&1&0 \\ 0&0&1 \end{bmatrix},$$

where $\tilde{K}$ and $\tilde{Q}$ are positive constants. From the requirement that AP_0^- is skew-adjoint, we have

$$(A_{11}\mathcal{L}_{12}\mathcal{M}_{22})^\dagger=-A_{22}\mathcal{L}_{21}\mathcal{M}_{11}, \tag{3.6}$$

or explicitly

$$\begin{aligned} \tilde{K}P\tilde{a}_{in}+\tilde{K}(N_0)^2a_{i,n+1} &= \tilde{Q}k_0\frac{\partial N^*}{\partial C_i},\quad i=1,\ldots,n-1, \\ \tilde{K}P\tilde{a}_{nn}+\tilde{K}(N_0)^2a_{n,n+1} &= \tilde{Q}k_0\frac{\partial N^*}{\partial E}, \\ \tilde{K}P\tilde{a}_{n+1,n}+\tilde{K}(N_0)^2a_{n+1,n+1} &= \tilde{Q}k_0\frac{\partial N^*}{\partial N}. \end{aligned} \tag{3.7}$$

Equation (3.7) is a system of $(n+1)$ equations for $(n+1)$ unknowns, $a_{i,n+1}$; $i=1,\ldots,n+1$, that can be readily solved. Finally, we have

$$A=\begin{pmatrix} \frac{\partial C_1^*}{\partial C_1}\cdots\frac{\partial C_1^*}{\partial E} & \frac{1}{N_0^2}\left(Kk_0\frac{\partial N^*}{\partial C_1}-P^*\frac{\partial C^*{}_1}{\partial E}\right) & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \frac{\partial E^*}{\partial C_1}\cdots\frac{\partial E^*}{\partial E} & \frac{1}{N_0^2}\left(Kk_0\frac{\partial N^*}{\partial E}-P^*\frac{\partial E^*}{\partial E}\right) & 0 & 0 & 0 \\ \frac{\partial C_1^*}{\partial N}\cdots\frac{\partial E^*}{\partial N} & \frac{1}{N_0^2}\left(Kk_0\frac{\partial N^*}{\partial N}-P^*\frac{\partial E^*}{\partial N}\right) & 0 & 0 & 0 \\ 0\ \cdots\ 0 & 0 & \tilde{Q}k_0 & 0 & 0 \\ 0\ \cdots\ 0 & 0 & 0 & \tilde{Q}k_0 & 0 \\ 0\ \cdots\ 0 & 0 & 0 & 0 & \tilde{Q}k_0 \end{pmatrix}, \tag{3.8}$$

where $K=\tilde{Q}/\tilde{K}>0$.

4. COMPATIBILITY OF THE PROBLEMS 3A AND 3B

Our problem now is to find $\{P\}$ such that $A_0 > 0$, $A > 0$, $k_0 > 0$ and the problems 3A and 3B are compatible, thus to find explicitly the function $\mathbb{V}$ introduced in 2C. We shall look for $\mathbb{V}$ in the form

$$\mathbb{V}(C_1, \ldots, C_{n-1}, E, N, \mathbf{U}; \alpha_1^*, \ldots, \alpha_{n-1}^*, \beta^*, \gamma^*) = \int_\Omega d^3\mathbf{r}\Big[S(C_1, \ldots, C_{n-1}, E, 1/N) - \sum_{i=1}^{n-1} \alpha_i^* C_i - \beta^* E - \gamma^* \frac{1}{N} + \tfrac{1}{2} k_0 U_\alpha U_\alpha\Big].$$

If the transformation T is defined by

$$\frac{\partial S}{\partial C_i} = C_i^*, \quad \frac{\partial S}{\partial E} = E^*, \quad \frac{\partial S}{\partial (1/N)} = N^* E^*, \quad \frac{\partial S}{\partial \mathbf{U}} = \mathbf{U}^*,$$

then indeed the necessary condition for an extremum of (3.9) is equivalent to (3.2), and (3.8) is equivalent to the second variation of $\mathbb{V}$ provided $\gamma^*/\beta^* = P^*$ and $-\beta^* = K k_0 > 0$, $\tilde{Q} = 1$. Thus $\{P_{ph}\}$ is defined by (3.10), (3.$\mathcal{A}$a), (3.$\mathcal{A}$b), and by the requirement that $\mathbb{V}$ reaches its non-degenerate minimum at f_0.

Now, if we require that $E_{th} = \int_\Omega d^3r\, E(\mathbf{r})$, $V_{th} = \int_\Omega d^3r[1/N(\mathbf{r})]$ and $C_{i,th} = \int_\Omega d^3\mathbf{r}\, C_i(r)$, then $\beta^* = -1/T$. That implies $E^* = -1/\tau$, τ is the local temperature and T is the thermodynamic temperature; $P^* = P$ implies $N^* = p$, where p is the local pressure and P is the thermodynamic pressure; $\alpha_i^* = -(1/T)(\mu_{i,th} - \mu_{n,th})$ implies $C_i^* = -(1/\tau)(\mu_i - \mu_n)$, where μ_i is the local chemical potential of the ith component, and $\mu_{i,th}$ is the thermodynamic chemical potential of the ith component. On the other hand, from the thermodynamic meaning of α_i^*, β^*, P^*, we would obtain the physical meaning (in the relation to the thermodynamic observations and measurements) of $C_i, E, N, \mathbf{U}$.

The time development equations (3.1) together with $C_i^* = -(1/\tau)(\mu_i - \mu_n)$, $E^* = -1/\tau$, $N^* = p$ are the well known equations of fluid mechanics.[9] Our considerations served basically two purposes. (i) We have illustrated the general theory developed in Sec. 2 that applies to other dynamical theories of nonequilibrium statistical mechanics, namely to the kinetic theory discussed in the next section. (ii) The mathematical problem of the existence and uniqueness of the solution to the system of the linear partial differential equations (3.3) has been solved. In fluid mechanics, our discussion can be compared with at least two other, different, approaches to fluid dynamics. The first is the classical standard approach that starts by generalizing thermodynamics to the local thermodynamics,[9] the second is the approach developed by Coleman and Noll.[10] None of these alternative approaches to fluid mechanics has been fully extended to other dynamical theories of non-equilibrium statistical mechanics.

5. TWO COMPONENT ENSKOG-VLASOV DYNAMICS

A state of the system in two component kinetic theory is completely described by $f \equiv (f_1, f_2)$, where f_i: $\Omega \times \mathbb{R}^3 \to \mathbb{R}^+$, $(\mathbf{r}, \mathbf{v}) \to f_i(\mathbf{r}, \mathbf{v})$, $i = 1, 2$, $\mathbb{R}^+$ is the positive real line, $\mathbf{r} \in \Omega$ is the same as in Sec. 3, and $\mathbf{v}$ denotes the velocity. The necessary differentiability of f will be assumed. All admissible f form a smooth manifold $\mathcal{H}$. The involution $\mathcal{J}$ is defined by $(f_1(\mathbf{r}, \mathbf{v}), f_2(\mathbf{r}, \mathbf{v})) \mapsto (f_1(\mathbf{r}, -\mathbf{v}), f_2(\mathbf{r}, -\mathbf{v}))$.

The family of the time development equations for f considered in this Sec. is a straightforward generalization of the Enskog–Vlasov kinetic equation discussed in Ref. 1. The even part R^+ of R consists of the Boltzmann collision operators for hard spheres.[11] By using the well known properties of the Boltzmann collision operators, we easily obtain that the requirements (2.Aa), and (2.Ab) are satisfied, provided

$$f^* = \ln f + F_0^*(f), \tag{4.1}$$

where $F_0^*(f)$ is an arbitrary element of $\mathcal{H}_0^*$ that is given by

$$\mathcal{H}_0^{*+} = \begin{pmatrix} \psi_1(\mathbf{r}) + \tfrac{1}{2}\beta(\mathbf{r}) m_1 v^2 \\ \psi_2(\mathbf{r}) + \tfrac{1}{2}\beta(\mathbf{r}) m_2 v^2 \end{pmatrix}, \tag{4.2}$$

where ψ_i: $\Omega \to \mathbb{R}^+$ and β: $\Omega \to \mathbb{R}^+$ are arbitrary, and m_i is the mass of a particle of the i-th component. $F_0^*(f)$ will be expressed later through the phenomenological quantities introduced in the vector field (see 4.16). The vector field R^{*+} is in full detail

$$R^{*+}(f^*) = \begin{pmatrix} J_{B,1}^*(f_1^*) + J_{B,12}^*(f_1^*, f_2^*) \\ J_{B,2}^*(f_2^*) + J_{B,21}^*(f_2^*, f_1^*) \end{pmatrix}, \tag{4.3}$$

where

$$J_{B,i}^*(f_i^*) = \theta_i^B[n_1(\mathbf{r}), n_2(\mathbf{r})] \int d^3\mathbf{v}_1 \exp(-F_{0,1}^* - F_{0,1,I}^*) \times \int d^2\kappa (g_\alpha \kappa_\alpha)[\exp(f_i^{*\prime} + f_{i,I}^{*\prime}) - \exp(f_i^* + f_{i,I}^*)],$$

where $f_{i,I}^* \equiv f_i^*(\mathbf{r}, \mathbf{v}_I')$; $f_{i,I}^* \equiv f_i^*(r, v_I)$, etc., and $(\mathbf{v}, \mathbf{v}_I)$ and $(\mathbf{v}, \mathbf{v}_I')$ are related by the two parameter family of linear, invertible, volume preserving transformations T_κ: $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3 \times \mathbb{R}^3$, where $\mathbf{v}' = \mathbf{v} + \kappa(g_\alpha \kappa_\alpha)$, $\mathbf{v}_I' = \mathbf{v}_I - \kappa(g_\alpha \kappa_\alpha)$, $g = \mathbf{v}_I - \mathbf{v}$, κ is a unit vector, and θ_i^B are positive real valued functions of $n_i(\mathbf{r}) = \int d^3\mathbf{v} f_i(\mathbf{r}, \mathbf{v})$.

$$J_{B,ij}^*(f_i^*, f_j^*) = \theta_{ij}^B[n_1(\mathbf{r}), n_2(\mathbf{r})] \int d^3 v_j \exp(-F_{0,i}^* - F_{0,j}^*) \times \int d^2\kappa (g_\alpha^{ji} \kappa_\alpha)[\exp(f_i^{*\prime} + f_j^{*\prime}) - \exp(f_i^* + f_i^*)],$$

where $(\mathbf{v}_i', \mathbf{v}_j')$ and $(\mathbf{v}_i, \mathbf{v}_j)$ are related by the two-parameter family of linear, invertible, volume preserving transformations $T_{12,\kappa}$: $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3 \times \mathbb{R}^3$; $\mathbf{v}_1' = \mathbf{v}_1 + N_2 \kappa(g_\alpha^{21} \kappa_\alpha)$; $\mathbf{v}_2' = \mathbf{v}_2 - N_1 \kappa(g_\alpha^{21} \kappa_\alpha)$, $\mathbf{g}^{21} = \mathbf{v}_2 - \mathbf{v}_1$, $N_i = 2m_i/(m_1 + m_2)$, and $\theta_{ij}^B = \theta_{ji}^B$ are positive real valued functions of $n_1(\mathbf{r})$ and $n_2(\mathbf{r})$.

The odd part R^- of the vector field R is given by

$$R^-(f) = \begin{pmatrix} -v_\alpha \partial f_1/\partial r_\alpha + J_{E,1}(f_1) + J_{V,1}(f_1) + J_{E,12}(f_1, f_2) + J_{V,12}(f_1, f_2) \\ -v_\alpha \partial f_2/\partial r_\alpha + J_{E,2}(f_2) + J_{V,2}(f_2) + J_{E,21}(f_2, f_1) + J_{V,21}(f_2, f_1) \end{pmatrix},$$

$$J_{E,i}(f_i) = \int d^2\kappa \int_{(g_\alpha \kappa_\alpha) > 0} d^3\mathbf{v}_I (g_\gamma \kappa_\gamma) \kappa_\alpha \Big[\theta_i^E[n_1(\mathbf{r}), n_2(\mathbf{r})]\Big(f_i' \frac{\partial}{\partial r_\alpha} f_{i,I}' + f_i \frac{\partial}{\partial r_\alpha} f_{i,I}\Big) + \frac{1}{2} \frac{\partial \theta_i^E[n_1(\mathbf{r}), n_2(\mathbf{r})]}{\partial r_\alpha} (f_i' f_{i,I}' + f_i f_{i,I})\Big],$$

$$J_{V,i}(f_i)=\frac{1}{m_i}\frac{\partial f_i}{\partial v_\alpha}\int_\Omega d^3\mathbf{r}_I\int d^3\mathbf{v}_I\frac{\partial}{\partial r_\alpha}[V_i(|\mathbf{r}-\mathbf{r}_I|)]f_i(\mathbf{r}_I,\mathbf{v}_I), \tag{4.4}$$

where $V_i(r)$ is the long range part of the two particle potential of the ith component,

$$J_{E,ij}(f_i,f_j)=\int d^2\kappa\int_{(g^{ji}_\alpha\kappa_\alpha)>0} d^3\mathbf{v}_j(g^{ji}_\gamma\kappa_\gamma)\kappa_\alpha$$
$$\times\ \theta^E_{ij}(n_1(\mathbf{r}),n_2(\mathbf{r}))\ f_i'\frac{\partial}{\partial r_\alpha}f_j'+f_i\frac{\partial}{\partial r_\alpha}f_j$$
$$+\frac{1}{2}\frac{\partial\theta^E_{ij}(n_1(\mathbf{r}),n_2(\mathbf{r}))}{\partial r_\alpha}(f_i'f_j'+f_if_j)\ ,$$
$$J_{V,ij}(f_i,f_j)=\frac{1}{m_i}\frac{\partial f_i}{\partial r_\alpha}\int d^3\mathbf{r}_j\int d^3\mathbf{v}_j\frac{\partial}{\partial r_\alpha}$$
$$\times V_{12}(|\mathbf{r}_i-\mathbf{r}_j|)f_j(\mathbf{r}_j,\mathbf{v}_j),$$

$V_{12}=V_{ij}=V_{ji}$ is the long range two partial potential between the particle of the i-th and j-th component, and θ^E_j and θ^E_{ij} are positive real valued functions of $n_1(\mathbf{r})$ and $n_2(\mathbf{r})$. Thus, the phenomenological quantities $\mathcal{P}$ introduced by the two component Enskog–Vlasov dynamics are

$$\mathcal{P}=\{\theta^B_i,\ \theta^B_{ij},\ \theta^E_i,\ \theta^E_{ij},\ V_i,\ V_{ij},\ m_i,\ i=1,2,\ i\neq j\}$$

4A. RESTRICTED PROBLEM OF FIXED POINTS

The submanifold $\mathcal{H}^{**}_0$ has been obtained already in (4.1). From $R^-f^+|_{\mathcal{H}^{**}_0}=0$, we have

$$\beta(\mathbf{r})=\beta, \tag{4.5}$$

where β is a positive constant. For $n_1(\mathbf{r})$ and $n_2(\mathbf{r})$, we obtain

$$0=-v_\alpha\frac{\partial n_i}{\partial r_\alpha}-2v_\alpha\theta^E_in_i\frac{\partial n_i}{\partial r_\alpha}-v_\alpha\frac{\partial\theta^E_i}{\partial r_\alpha}n_i^2$$
$$-2v_\alpha\theta^E_{ij}n_i\frac{\partial n_j}{\partial r_\alpha}-v_\alpha\frac{\partial\theta^E_{ij}}{\partial r_\alpha}n_in_j$$
$$-\beta v_\alpha n_i\frac{\partial}{\partial r_\alpha}\int d^3\mathbf{r}_IV_i(|r-\mathbf{r}_I|)n_i(\mathbf{r}_I)$$
$$-\beta v_\alpha n_i\frac{\partial}{\partial r_\alpha}\int d^3\mathbf{r}_IV_{12}(|r-\mathbf{r}_I|)n_j(\mathbf{r}_I), \tag{4.6}$$

$i,j=1,2,\ i\neq j$. In order to keep the notation as simple as possible, we have absorbed in the symbols θ^E in (4.6), the multiplicative factor $(2/3)\pi$. One can see easily that (4.6) can be written in the form $v_\alpha\partial/\partial r_\alpha\,\Phi_i(n_1(\mathbf{r}),n_2(\mathbf{r}))=0$, thus in the form $\Phi_i=\tilde{\alpha}_i$, that does not include derivatives with respect to the position coordinates [$\tilde{\alpha}_i$ are constants and Φ_i are functions of $n_1(\mathbf{r})$, $n_2(\mathbf{r})$] if and only if

$$\theta^E_{ij}=\frac{\delta\Theta^E}{\delta n_i\delta n_j}, \tag{4.7}$$

where Θ^E is a real valued function of $n_1(\mathbf{r})$, $n_2(\mathbf{r})$ of class at least C^3 and $\theta^E_{ii}\equiv\theta^E_i$. By using (4.7), one can write (4.6)

$$\ln n_i+\theta^E_in_i+\frac{\delta\Theta^E}{\delta n_i}+n_j\theta^E_{ij}+\beta\Big(\int d^3\mathbf{r}'[V_i(|\mathbf{r}-\mathbf{r}'|)n_i(\mathbf{r}')$$
$$+V_{12}(|\mathbf{r}-\mathbf{r}'|)n_j(\mathbf{r}')]\Big)=\tilde{\alpha}_i,\quad i,j=1,2,\ i\neq j. \tag{4.8}$$

The constants $\tilde{\alpha}_i$ are independent of $n_1(\mathbf{r})$, $n_2(\mathbf{r})$ but they can depend on β. The position coordinates independent solution of (4.8) is thus

$$f_0\equiv\begin{pmatrix}n_{0,1}(\beta m_1/2\pi)^{3/2}\exp(-\frac{1}{2}\beta m_1v^2)\\ n_{0,2}(\beta m_2/2\pi)^{3/2}\exp(-\frac{1}{2}\beta m_2v^2)\end{pmatrix}, \tag{4.9}$$

where $n_{0,1}, n_{0,2}$, β characterizing f_0 in (4.9) are related to α_1, α_2, β introduced in (4.8) by requiring that (4.9) is a solution of (4.8).

4B. LOCAL DYNAMIC STABILITY OF f_0

By linearizing (4.3) and (4.4) at f_0 [taking into account (4.1)], one obtains the linearized vector field P_0 governing the time evolution of $f_0\phi\in H_0$

$$\frac{\partial\phi}{\partial t}=P_0\phi, \tag{4.10}$$

where H_0 is the tangent space of $\mathcal{H}$ at f_0. $P_0=P_0^+ + P_0^-$. From the standard properties of the Boltzmann collision operator,[11] we obtain that $A_1P_0^+$ is a self-adjoint non-positive linear operator, i.e.,

$$A_1P_0^+=\begin{bmatrix}(A_1P_0^+)_{11} & (A_1P_0^+)_{12}\\ (A_1P_0^+)_{21} & (A_1P_0^+)_{22}\end{bmatrix}, \tag{4.11}$$

$$A_1=\begin{bmatrix}n_{0,1}(\beta m_1/2\pi)^{3/2}\exp(-\frac{1}{2}\beta m_1v^2) & 0\\ 0 & n_{0,2}(\beta m_2/2\pi)^{3/2}\exp(-\frac{1}{2}\beta m_2v^2)\end{bmatrix}, \tag{4.12}$$

$$(A_1P_0^+)_{ii}\phi_i=(\beta m_i/2\pi)^{3/2}\exp(-\tfrac{1}{2}\beta m_iv_i^2)[n_{0,i}\theta^B_{0,i}(K_i^{(1)}+K_i^{(2)}-K_i^{(3)}-K_i^{(4)})\phi_i+n_{0,j}\theta^B_{0,ji}(K_i^{(5)}-K_i^{(7)})\phi_i],$$

$$(A_1P_0^+)_{ij}\phi_j=(\beta m_i/2\pi)^{3/2}\exp(-\tfrac{1}{2}\beta m_iv_2)n_{0,j}\theta^B_{0,ij}(K_i^{(6)}-k_i^{(8)})\phi_j,$$

$$K_i^{(1)}\phi_i=\int d^2\kappa\int d^3\mathbf{v}_j(\mathbf{v}_j-\mathbf{v}_i)_\alpha\kappa_\alpha(\beta m_i/2\pi)^{3/2}\exp(-\tfrac{1}{2}\beta m_iv_j^2)\phi_i(\mathbf{v}_i'),$$

$$K_i^{(2)}\phi_i=\int d^2\kappa\int d^3\mathbf{v}_j(\mathbf{v}_j-\mathbf{v}_i)_\alpha\kappa_\alpha(\beta m_i/2\pi)^{3/2}\exp(-\tfrac{1}{2}\beta m_iv_j^2)\phi_i(\mathbf{v}_j'),$$

$$K_i^{(3)}\phi_i=\phi_i(\mathbf{v})\int d^3\kappa\int d^3\mathbf{v}_j(\mathbf{v}_j-\mathbf{v}_i)_\alpha\kappa_\alpha(\beta m_i/2\pi)^{3/2}\exp(-\tfrac{1}{2}\beta m_iv_j^2),$$

$$K_i^{(4)}\phi_i=\int d^2\kappa\int d^3\mathbf{v}_j(\mathbf{v}_j-\mathbf{v}_i)_\alpha\kappa_\alpha(\beta m_i/2\pi)^{3/2}\exp(-\tfrac{1}{2}\beta m_iv_j^2)\phi_i(\mathbf{v}_j),$$

$$K_i^{(5)}\phi_i=\int d^2\kappa\int d^3\mathbf{v}_j(\mathbf{v}_j-\mathbf{v}_i)_\alpha\kappa_\alpha(\beta m_j/2\pi)^{3/2}\exp(-\tfrac{1}{2}\beta m_jv_j^2)\phi_i(\mathbf{v}_i'),$$

$$K_i^{(6)}\phi_j=\int d^2\kappa\int d^3\mathbf{v}_j(\mathbf{v}_j-\mathbf{v}_i)_\alpha\kappa_\alpha(\beta m_j/2\pi)^{3/2}\exp(-\tfrac{1}{2}\beta m_jv_j^2)\phi_j(\mathbf{v}_j'),$$

$$K_i^{(7)}\phi_i = \phi_i(\mathbf{v}_i)\int d^2\kappa \int d^3\mathbf{v}_j(\mathbf{v}_j-\mathbf{v}_i)_\alpha \kappa_\alpha (\beta m_j/2\pi)^{3/2}\exp(-\tfrac{1}{2}\beta m_j v_j^2),$$

$$K_i^{(8)}\phi_j = \int d^2\kappa \int d^3\mathbf{v}_j(\mathbf{v}_j-\mathbf{v}_i)_\alpha \kappa_\alpha (\beta m_j/2\pi)^{3/2}\exp(-\tfrac{1}{2}\beta m_j v_j^2)\phi_j(\mathbf{v}_j),$$

$i,j=1,2$, $i\neq j$. The transformations T_κ relating $(\mathbf{v}_i,\mathbf{v}_j)$ and $(\mathbf{v}_i',\mathbf{v}_j')$ are used in $K^{(1)},\ldots,K_i^{(4)}$, and the transformations $T_{12,\kappa}$ are used in $K_i^{(5)},\ldots,K_i^{(8)}$, $\theta^B_{0,ij}=\theta^B_{ij}|_{f_0}$.

By linearizing R^- at f_0, we obtain

$$P_0^- = \begin{bmatrix} (P_0^-)_{11} & (P_0^-)_{12} \\ (P_0^-)_{21} & (P_0^-)_{22} \end{bmatrix}, \tag{4.13}$$

where $(P_0^-)_{ii}=D+Q_i$, $(P_0^-)_{ij}=Q_{ij}$,

$$Q_i = n_{0,i}[\theta^E_{0,i}(DG_i^{(2)}-G_i^{(2)}D)+\tilde\theta^E_{0,i}DG_i^{(2)} + (3/4\pi)\theta^E_{0,j}D_i^{(1)}+\beta DG_i^{(1)}G_i^{(2)}],$$

$$Q_{ij} = n_{0,j}[\theta^E_{0,ij}(DG_j^{(2)}-G_j^{(2)}D)+\tilde\theta^E_{0,ij}DG_j^{(2)} + (3/4\pi)\theta^E_{0,j}D_j^{(2)}+\beta DG_j^{(1)}G_{12}^{(2)}].$$

We have used the notation

$$D=-v_\alpha\frac{\partial}{\partial r_\alpha},$$

$$G_i^{(1)}\phi_i = \int d^3\mathbf{v}_1\left(\frac{m_i\beta}{2\pi}\right)^{3/2}\exp(-\tfrac{1}{2}\beta m_i v_1^2)\phi_i(\mathbf{r},\mathbf{v}_1),$$

$$G_i^{(2)}\phi_i = \int d^3\mathbf{r}_1 V_i(|r-r_1|)\phi_i(\mathbf{r}_1,\mathbf{v});$$

$$D_i^{(1)}\phi_i = \frac{\partial}{\partial r_\alpha}\int d^2\kappa\int d^3\mathbf{v}_1\left(\frac{m_i\beta}{2\pi}\right)^{3/2} \times\exp(-\tfrac{1}{2}m_i\beta v_1^2)(\mathbf{v}_1-\mathbf{v})_\gamma\kappa_\gamma\kappa_\alpha\phi(\mathbf{r},\mathbf{v}_1'),$$

$$G_{12}^{(2)}\phi_j = \int d^3\mathbf{r}_1 V_{12}(|\mathbf{r}-\mathbf{r}_1|)\phi_j(\mathbf{r}_1,\mathbf{v}),$$

$$D_i^{(2)}\phi_i = \frac{\partial}{\partial r_\alpha}\int d^2\kappa\int d^3\mathbf{v}_1\left(\frac{m_j\beta}{2\pi}\right)^{3/2} \times\exp(-\tfrac{1}{2}m_j\beta v_j^2)(\mathbf{v}_j-\mathbf{v})_\gamma\kappa_\gamma\kappa_\alpha\phi(\mathbf{r},\mathbf{v}_j),$$

$$\theta^E_{0,i}=\theta^E_i\Big|_{f_0},\quad \tilde\theta^E_{0,i}=\frac{\delta\theta_i^E}{\delta n_i}\Big|_{f_0},\quad \theta^E_{0,ij}=\theta^E_{ij}\Big|_{f_0},$$

$$\tilde\theta^E_{0,ij}=\tilde\theta^E_{0,ji}=\left(\frac{\delta\theta^E_{ij}}{\delta n_j}+\frac{\delta\theta^E_i}{\delta n_j}\right),$$

$i,j=1,2$, $i\neq j$. In the operators $D^{(1)}$ and $D^{(2)}$ the transformations T_κ and $T_{12,\kappa}$ are used respectively.

Following the arguments of Sec. 2B, taking into account the properties of the Boltzmann collision operators (4.$\mathcal{A}$a), (4.$\mathcal{A}$b), we shall look for the nonsingular self-adjoint positive definite operator A, $(\mathcal{J}A\mathcal{J}=A)$ such that AP_0^+ is self-adjoint and nonpositive and AP_0^- is skew-adjoint. Since we know already that $A_1P_0^+$ is self-adjoint and nonpositive, we look for $A=A_1A_2$, where

$$A_2 = \begin{bmatrix} (A_2)_{11} & (A_2)_{12} \\ (A_2)_{21} & (A_2)_{22} \end{bmatrix}, \tag{4.14}$$

$(A_2)_{ii}=1+\tilde A_{ii}$ such that $\tilde A_{ii}P_0^+=0$ and $(A_2)_{ij}P_0^+=0$.

From (4.12), one obtains

$$(A_2)_{ii}=1+n_{0,i}[2\theta^E_{0,i}+\tilde\theta^E_{0,i}+\beta G_i^{(1)}]G_i^{(2)},$$

$$(A_2)_{ij}=n_{0,j}[2\theta_{0,ij}+\tilde\theta^E_{0,ij}+\beta G_j^{(1)}]G_{12}^{(2)}.$$

One can check easily that indeed AP_0^+ is self-adjoint and nonpositive and AP_0^- is skew-adjoint.

4C. COMPATIBILITY OF THE PROBLEMS 4A AND 4B

Our problem again is to find a function $\mathbb{V}$ such that Eqs. (4.2), (4.5) and (4.8) are equivalent to the necessary condition for extremum of $\mathbb{V}$, and the operator A obtained in (4.11) and (4.14) equals the second derivative of $\mathbb{V}$ at f_0. It is easy to check that the function $\mathbb{V}$ satisfying these requirements is

$$\mathbb{V}(f_1,f_2;\beta,\alpha_1,\alpha_2) = \int_\Omega d^3\mathbf{r}\int d^3\mathbf{v}\sum_{i,j=1}^{2}\sum (f_i\ln f_i + f_i\frac{\delta\Theta^E}{\delta n_i}+\tfrac{1}{2}\beta m_i v^2 f_i-\alpha_i f_i+\tfrac{1}{2}\beta\int_\Omega d^3\mathbf{r}_1 \times\int d^3\mathbf{v}_1[V_i(|r-r_1|)f_i(\mathbf{r},\mathbf{v})f_i(\mathbf{r}_1,\mathbf{v}_1) + V_{12}(|\mathbf{r}-\mathbf{r}_1|)f_i(\mathbf{r},\mathbf{v})f_j(\mathbf{r}_1,\mathbf{r}_1)], \tag{4.15}$$

$\alpha_i=\tilde\alpha_i+\frac{3}{2}\ln(\beta m_i/2\pi)$. The necessary conditions for $\mathbb{V}$ to reach its nondegenerate minimum at f_0 is equivalent to positive definiteness of A. We would like to point out that the two component Enskog–Vlasov dynamics has the structures introduced in Sec 2 only if θ^E_{ij} are generated from one function, Θ^E. The choice of V_i and V_{12} is restricted by the requirement that $G_i^{(2)}$ and $G_{12}^{(2)}$ are bounded. By comparing (4.15) with (2.4), we have

$$S(f)=\int_\Omega d^3\mathbf{r}\int d^3\mathbf{v}\sum_{i=1}^{2} f_i\ln f_i + f_i\frac{\delta\Theta^E}{\delta n_i}. \tag{4.16}$$

The transformation $\mathcal{T}$ (see 4.1) is now defined as

$$f^*=\frac{\delta S(f)}{\delta f}. \tag{4.17}$$

Thus, the unknown in map $F_0^*(f)$ (4.1) is now obtained in terms of Θ^E. If we assign to β, α_1, α_2 the usual thermodynamic meaning, [i.e., $\beta=1/T$, $\alpha_i=\mu_i/T$, where T is the temperature and μ_i is the chemical potential of the i-th component] then $\mathbb{V}_{\rm th}=\gamma=p\beta$; p denotes the pressure, and the thermodynamic inner energy equals

$$\sum_{i,j=2}^{2}\sum\int_\Omega d^3\mathbf{r}\int d^3\mathbf{v}(\tfrac{1}{2}n_i v^2 f_{0,i}+\tfrac{1}{2}\beta\int_\Omega d^3\mathbf{r}_1 \times\int d^3\mathbf{v}_1[V_i(|\mathbf{r}-\mathbf{r}_1|)f_{0,i}(\mathbf{r},\mathbf{v})f_{0,i}(\mathbf{r}_1,\mathbf{v}_1) + V_{12}(|\mathbf{r}-\mathbf{r}_1|)f_{0,i}(\mathbf{r},\mathbf{v})f_{0,j}(\mathbf{r}_1,\mathbf{v}_1)],$$

and the concentration of the ith component equals

$$\int d^3\mathbf{r}\int d^3\mathbf{v} f_{0,i}(\mathbf{r},\mathbf{v})\left(\sum_{j=1}^{2}\int d^2\mathbf{r}\int d^3\mathbf{v} f_{0,j}(\mathbf{r},\mathbf{v})\right)^{-1}.$$

This section basically served two purposes. (i) We have illustrated the general theory introduced in Sec. 2. (ii) We have proven the existence and uniqueness of the solution to the system of linear integro-differential equations (4.10). Results of this section will be used in the next paper[3] for an investigation of the relationship between two component Enskog–Vlasov dynamics and the two component fluid dynamics.

ACKNOWLEDGMENT

I am indebted to Professor Van Vliet for his support of this research. I have also benefited from discussions with Professor Luis de Sobrino and Professor Donald Dawson.

[1]M. Grmela, J. Math. Phys. 15, 35 (1974).
[2]M. Grmela, Helv. Phys. Acta 47, 667,677 (1974).
[3]M. Grmela, J. Math. Phys. 16, 2450 (1975).
[4]K. Yoshida, *Functional Analysis* (Springer-Verlag, 1968).
[5]R. B. Griffiths and J. C. Wheeler, Phys. Rev. A 2, 1047 (1970).
[6]J. Yvon, *Correlation and Entropy in Classical Statistical Mechanics* (Pergamon, New York, 1969).
[7]P. Lancaster, *Theory of Matrices* (Academic, New York, 1969).
[8]G. Hellwig, *Differential Operators of Mathematical Physics* (Addison-Wesley, Reading, Mass., 1967).
[9]S. R. deGroot and P. Mazur, *Non-equilibrium Thermodynamics* (North-Holland, Amsterdam, 1962). I. Gyarmati, *Non-equilibrium Thermodynamics* (Springer-Verlag, Berlin, 1970).
[10]C. Truesdell, *Rational Thermodynamics* (McGraw-Hill, New York, 1969).
[11]S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-Uniform Gases* (Cambridge U. P., Cambridge, 1952).

Long time behavior of solutions to the linearized two component Enskog–Vlasov kinetic equations

Miroslav Grmela

Centre de Recherches Mathématiques, Université de Montréal, Montréal, Québec, Canada H3C 3J7
(Received 24 June 1975)

Fluid dynamics is obtained from the study of the long time behavior of solutions to the two component Enskog–Vlasov kinetic equations. Both thermodynamic and dynamic phenomenological coefficients of fluid mechanics are expressed in terms of the phenomenological quantities entering the Enskog–Vlasov-type kinetic equations.

1. INTRODUCTION

The general discussion of the relationship between thermodynamic and local dynamic stability developed in Ref. 1, is used in this paper to investigate the relationship between the two component Enskog—Vlasov dynamics and the two component fluid dynamics. The method developed for the one component systems[2–4] is followed closely. Since the problem is well defined from both a physical and mathematical point of view, it can be discussed rigorously. No ad hoc relations among the state variables of kinetic theory, fluid mechanics and thermodynamics are needed. This type of relation appears as a result. A single calculation gives both dynamic and thermodynamic phenomenological coefficients of fluid mechanics, expressed as functions of the phenomenological quantities of kinetic theory. The results for the one component systems[4] reproduced the van der Waals type critical phenomena for dynamic and thermodynamic phenomenological coefficients. The critical phenomena derived from the results of this paper will be discussed in a subsequent paper.

If the Enskog—Vlasov type kinetic equations are reduced to the Enskog equations only, then our results can be compared with the results obtained by using the Chapman—Enskog method.[5,6]

2. DESCRIPTION OF THE METHOD

In order to explain the main idea of the method, we do not need all details of kinetic theory and fluid mechanics. The general description immediately suggests a possibility of applying the method in the context of other dynamical theories of nonequilibrium statistical mechanics. All details of the two component kinetic theory and fluid mechanics can be found in Ref. 1; all details needed in this paper will be introduced in Secs. 3 and 4.

Let $f^{(i)}$ completely describe a state of a class $\mathcal{C}$ of two component physical systems. The complete description means complete with respect to the set of observation and measurements $\mathcal{O}^{(i)}$, the results of which form the empirical basis for the dynamical system indexed by i. The index $i=1$, will denote kinetic theory; $i=2$, fluid mechanics. All admissible states (with respect to $\mathcal{O}^{(i)}$) form a set $\mathcal{H}^{(i)}$. We assume[1] that $\mathcal{H}^{(i)}$ is a smooth infinite-dimensional manifold locally topologically isomorphic to the Hilbert space L_2. It is assumed moreover,[1] that an involution $\mathcal{I}^{(i)}: \mathcal{H}^{(i)} \to \mathcal{H}^{(i)}$ is defined for $i=1,2$. The subset of $\mathcal{H}^{(i)}$ that is invariant with respect to $\mathcal{I}^{(i)}$ is denoted $\mathcal{H}^{(i)+}$, its complement $\mathcal{H}^{(i)-}$.

The experience from the set of observations and measurements $\mathcal{O}^{(i)}$ is summed up in the time evolution equation,

$$\frac{\partial f^{(i)}}{\partial t} = R^{(i)} f^{(i)}, \tag{1}$$

where $R^{(i)}$ is a class of vector fields (i.e., the maps $\mathcal{H}^{(i)} \to T\mathcal{H}^{(i)}$, $T\mathcal{H}^{(i)}$ denote the tangent bundle of $\mathcal{H}^{(i)}$) that generates the time development of $f^{(i)}$ parametrized by a set of phenomenological quantities $\mathcal{P}^{(i)}$. We shall define $R^{(i)\pm} = \frac{1}{2}(R^{(i)} \pm \mathcal{I}^{(i)} R^{(i)} \mathcal{I}^{(i)})$.

The complete solution of (1) consists of the set $\mathcal{S}^{(i)}$ of all trajectories passing through all $f^{(i)} \in \mathcal{H}^{(i)}$ for all $\mathcal{P}^{(i)} \in \mathcal{P}^{(i)}$. Intuitively, the mathematical procedure involved in the derivation of fluid mechanics ($i=2$) from kinetic theory ($i=1$) is the "pattern" recognition in $\mathcal{S}^{(1)}$. The "pattern" in $\mathcal{S}^{(1)}$ will be characterized by $f^{(2)}$, its time development will be governed by $R^{(2)}$. The phenomenological quantities $\mathcal{P}^{(2)}$ will be expressed in terms of $\mathcal{P}^{(1)}$.

We have shown[1] that thermodynamics is obtained by recognizing the pattern of locally stable fixed points in $\mathcal{S}^{(1)}$ or $\mathcal{S}^{(2)}$. Discussing this pattern recognition, a function $\mathbf{V}^{(i)}: \mathcal{H}^{(i)} \times U \to \mathbb{R}$; $\mathbf{V}^{(i)}(f^{(i)}; \sigma_1, \ldots, \sigma_m) = S^{(i)}(f^{(i)}) + \sum_{j=1}^{m} \sigma_j v_j^{(i)}(f^{(i)})$, is naturally introduced.[1] $S^{(i)}: \mathcal{H}^{(i)} \to \mathbb{R}$, $v_j^{(i)}: \mathcal{H}^{(i)} \to \mathbb{R}$, $i=1, 2$; U is an open subset of R^m ($m=3$ for two component systems); elements of U are the intensive thermodynamic parameters (thermodynamic fields in the terminology of Griffiths and Wheeler[7]) characterizing the equilibrium states. The thermodynamically stable fixed points $f_{eq}^{(i)}$ are identified with the dynamically stable fixed points that appear to be exactly the nondegenerate critical points of $\mathbf{V}^{(i)}$. The thermodynamic potential $\sigma_{m+1} = \sigma_{m+1}(\sigma_1, \ldots, \sigma_m)$ is defined as $\sigma_{m+1} = \mathbb{V}^{(i)}(f_{eq}^{(i)}(\sigma_1, \ldots, \sigma_m); \sigma_1, \ldots, \sigma_m)$, where $f_{eq}^{(i)}$ is a critical point of $\mathbf{V}^{(i)}$ (i.e., $\mathbf{V}^{(i)}$ reaches its maximum or minimum at $f^{(i)}$). The thermodynamical conjugate s_j of σ_j is defined as $s_j = \partial \sigma_{m+1} / \partial \sigma_j$, thus, from the definition of σ_{m+1}, $s_j = v_j^{(i)}(f_{eq}^{(i)})$. The identification of $v_j^{(1)}(f^{(1)})$ and $v_j^{(2)}(f^{(2)})$ immediately leads to information on the relationship between $f^{(1)}$ and $f^{(2)}$. All details of this pattern recognition can be found in Ref. 1.

Let $f_0^{(i)}$ be independent of the position coordinates and represent a stable (both thermodynamically and dynamically) equilibrium state. $H_0^{(i)}$ denotes the tangent space

of $\mathcal{H}^{(i)}$ at $f_0^{(i)}$. The time development of the elements $\phi^{(i)}$ of $H_0^{(i)}$ is governed by

$$\frac{\partial \phi^{(i)}}{\partial t} = P_0^{(i)} \phi^{(i)}, \tag{2}$$

where $P_0^{(i)}$ denote the family of the linearized vector fields $R^{(i)}$, evaluated at $f_0^{(i)}$ and parametrized by $\mathcal{P}^{(i)}$ and by the parameters σ_j, $j=1,\ldots,m$, that are needed to specify $f_0^{(i)}$. From (2) and the requirement of the local dynamic stability of $f_0^{(i)}$ ($f_0^{(i)}$ is locally stable if and only if the norm of $\phi^{(i)}$ is bounded for all positive times), the linear operator $A^{(i)}$: $H_0^{(i)} \to H_0^{(i)}$ is obtained. $A^{(i)}$ equals the second derivative of $\mathbb{V}^{(i)}$ with respect to $f^{(i)}$ evaluated at $f_0^{(i)}$. The operator $A^{(i)}$ is bounded, self-adjoint, positive definite, i.e., $\mathcal{J}^{(i)} A^{(i)} \mathcal{J}^{(i)} = A^{(i)}$. Moreover, if $H_0^{(i)}$ is equipped with the inner product $\langle \phi, \psi \rangle^{(i)} = (\phi, A^{(i)} \psi)$, where $(\ ,\ )$ denotes the standard inner products in L_2 space, then $P_0^{(i)+}$ is self-adjoint and nonpositive and $P_0^{(i)-}$ is skew-adjoint.

Let $\mathcal{S}_0^{(i)}$ denote the set of all trajectories of dynamical system (2). The pattern recognition in $\mathcal{S}_0^{(1)}$ leading to the local fluid dynamics is the following. Let $H_0^{(1,2)} \subset H_0^{(1)}$ have the following three properties: (i) $H_0^{(1,2)} \sim H_0^{(2)}$, where $\sim$ means isomorphism such that the norms in $H_0^{(1,2)}$ and $H_0^{(2)}$ define an equivalent topology, but the norms themselves are not necessarily equivalent. (ii) $H_0^{(1,2)}$ is invariant with respect to the time development generated by $P_0^{(1)}$. (iii) $H_0^{(1,2)}$ is asymptotic in the sense that $H_0^{(1,2)}$ characterizes the long time behavior of trajectories; if for example, the part of the spectrum of $P_0^{(1)}$ that is closet to zero consists of the point spectrum, then the corresponding eigenfunctions will be used to define $H_0^{(1,2)}$. The local fluid dynamics is obtained as the local kinetic theory dynamics restricted to $H_0^{(1,2)}$ and represented in $H_0^{(2,1)}$ that is completely isomorphic (i.e., also the norms are equivalent) to $H_0^{(2)}$. The only difficult problem of this pattern recognition process is to find $H_0^{(1,2)}$. Everything else, including the problem of finding the space $H_0^{(2,1)}$, is in principle, a simple problem. It should be pointed out that the space $H_0^{(1,2)}$ with the properties (i), (ii), (iii) introduced above does not necessarily exist. The spectral theorem for the Fourier transform of $P_0^{(1)}$ (Sec. 3) guarantees the existence of $H_0^{(1,2)}$ and also shows a method for finding it.

3. THE LOCAL TWO COMPONENT ENSKOG-VLASOV DYNAMICS

A state of the class $\mathcal{C}$ of two component systems is described in kinetic theory by $f^{(1)} \equiv (f_1, f_2)$, $f_i: \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^+$, $(i=1,2)$; $(\mathbf{r},\mathbf{v}) \mapsto f_i(\mathbf{r},\mathbf{v})$; $\mathbf{r},\mathbf{v}$ denotes the position vector and velocity respectively, $\mathbb{R}^+$ denotes the positive real line. The involution $\mathcal{J}^{(1)}$ is defined by $(f_1(\mathbf{r},\mathbf{v}), f_2(\mathbf{r},\mathbf{v})) \overset{\mathcal{J}^{(1)}}{\mapsto} (f_1(\mathbf{r},-\mathbf{v}), f_2(\mathbf{r},-\mathbf{v}))$. The function $\mathbb{V}^{(1)}$ appears to be[1]

$$\begin{aligned}
&\mathbb{V}^{(1)}(f^{(1)}; \beta, \alpha_1, \alpha_2) \\
&= \int d^2\mathbf{r} \int d^3\mathbf{v} \sum_{j=1}^{2} \sum_{\substack{k=1 \\ j\neq k}}^{2} \Big(f_i \ln f_j + f_j \frac{\delta \Theta}{\delta \mu_j} + \tfrac{1}{2}\beta m_j v^2 f_j - \alpha_j f_j \\
&\quad + \tfrac{1}{2}\beta \int d^3\mathbf{r}_1 \int d^3\mathbf{v}_1 [V_j(|\mathbf{r}-\mathbf{r}_2|) f_j(\mathbf{r},\mathbf{v}) f_j(\mathbf{r}_1,\mathbf{v}_1) \\
&\quad + V_{jk}(|\mathbf{r}-\mathbf{r}_1|) f_j(\mathbf{r},\mathbf{v}) f_k(\mathbf{r}_1,\mathbf{v}_1)] \Big)
\end{aligned} \tag{3}$$

We have used the notation $n_j(\mathbf{r}) = \int d^3\mathbf{v} f_j(\mathbf{r},\mathbf{v})$. The nondegenerate critical point $f_0^{(1)}$ of $\mathbb{V}^{(1)}$ that is, moreover, independent of the position coordinates is $f_0^{(1)} \equiv (f_{0,1}, f_{0,2})$; $f_{0,j} = n_0 \ (\beta m_j/2\pi)^{3/2} \exp(-\frac{1}{2}\beta_j v^2)$; $j=1,2$. The quantities $n_{0,1}$, $n_{0,2}$, β are positive constants. Since $f_0^{(1)}$ is a critical point of $\mathbb{V}^{(1)}$, we obtain relations between $n_{0,1}$, $n_{0,2}$ and α_1, α_2. The thermodynamic potential $\gamma = \gamma(\beta, \alpha_1, \alpha_2)$ is

$$\begin{aligned}
\gamma = \mathbb{V}^{(1)}(f_0^{(1)}(\beta_1,\alpha_1,\alpha_2); \beta_1,\alpha_1,\alpha_2) = \sum_{j,k=1}^{2} \sum \big(&n_{0,j} + (n_{0,j})^2 \theta_{0,j} \\
&+ \tfrac{1}{2}\beta (n_{0,j})^2 V_{0,j} + \tfrac{1}{2}\beta n_{0,j} n_{0,j} V_{0,jk} \big).
\end{aligned} \tag{4}$$

The thermodynamical meaning of β, α_j, γ is the following: $\beta = 1/T$, T is the temperature; $\alpha_j = \beta\mu_j$, μ_j is the chemical potential of the jth component; $\gamma = \beta p$, p is the pressure. By using (3), the thermodynamical conjugates of α_j, β, and γ are expressed in terms of $f^{(1)}$. Thus, the kinetic theory state variable $f^{(1)}$ is related to the thermodynamic state variables. The phenomenological quantities $\mathcal{P}^{(1)}$ introduced by the two component Enskog–Vlasov dynamics are $\mathcal{P}^{(1)} \equiv \{\Theta, V_j, V_{jk} = V_{kj}, m_j,$ and the phenomenological quantities introduced of the hard sphere Boltzmann collision operators; $j,k=1,2,\ j\neq k\}$. Θ is a real valued function of $n_1(\mathbf{r})$ and $n_2(\mathbf{r})$ of class C^3. The quantities θ_j and θ_{jk} appearing in $R^{(1)-}$ [see (6)] are defined by $\theta_j = (\delta^2\Theta/\delta n_j \delta n_j)$; $\theta_{jk} = (\delta^2\Theta/\delta n_j \delta n_k)$; $j,k=1,2$, $j\neq k$. V_{jk} are the long range interaction potentials between the particles of the jth component and the particles of the first and the second component, respectively. The mass of the particles of jth component is denoted by m_j.

The two component Enskog–Vlasov dynamics is linearized at $f_0^{(1)}$. By $\phi^{(1)}$ we denote elements of the tangent space $\tilde{H}_0^{(1)}$ of $\mathcal{H}^{(1)}$ at $f_0^{(1)}$ on which the linearized dynamics [Eq. (2) with $i=1$] is defined. By $\phi(\mathbf{k},\mathbf{v}) \equiv (\phi_1(\mathbf{k},\mathbf{v}),\ \phi_2(\mathbf{k},\mathbf{v}))$, we shall denote the Fourier transform of

$$\phi(\mathbf{r},\mathbf{v}) \equiv (\phi_1(\mathbf{r},\mathbf{v}), \phi_2(\mathbf{r},\mathbf{v})), \text{ i.e.,}$$

$$\phi_j(\mathbf{k},\mathbf{v}) = 1/(2\pi)^{3/2} \int d^3\mathbf{r} \exp(-i\mathbf{k}\mathbf{r}) \phi_j(\mathbf{r},\mathbf{v});$$

$j=1,2$ and i denotes the imaginary unit.

The vector $\mathbf{k}$ is fixed, $\mathbf{k} \equiv (0,0,k)$. In order to simplify the notation we shall write $\phi_j(\mathbf{v})$ instead of $\phi_j(\mathbf{k},\mathbf{v})$. The space $H_0^{(1)}$, on which the linearized Fourier transformed (with $\mathbf{k}$ fixed) two component Enskog–Vlasov dynamics, is now defined as a complex Hilbert space whose elements are $\phi_j: \mathbb{R}^3 \to \mathbb{C}$; $(\mathbf{v}) \mapsto \phi_j(\mathbf{v})$, $j=1,2$. The inner product $\langle \psi, \psi \rangle^{(1)}$ in $H_0^{(1)}$ is defined by $\langle \phi, \psi \rangle^{(1)} = (\phi, A^{(1)}\psi)$, where $(\phi,\psi) \equiv (\phi_1,\psi_1) + (\phi_2,\psi_2)$ denotes the standard inner product in the $\mathbb{C}L_2$ space. Instead of considering $P_0^{(1)}$ in $H_0^{(1)}$, it will be convenient to consider $A^{(1)}P_0^{(1)}$ in the $\mathbb{C}L_2$ space.

The operator $A^{(1)}$ is[1]

$$A^{(1)} = \begin{bmatrix} A_{11}^{(1)} & A_{12}^{(1)} \\ A_{21}^{(1)} & A_{22}^{(1)} \end{bmatrix},$$

where

$$\begin{aligned}
A_{jj}^{(1)} \phi_j = (\beta m_j/2\pi)^{3/2} &\exp(-\tfrac{1}{2}\beta m_j v^2) \\
&\times [n_{0,j} + w_j (n_{0,j})^2 \int d^3\mathbf{v}_I (\beta m_j/2\pi)^{3/2} \\
&\times \exp(-\tfrac{1}{2}\beta m_j v_I^2) \phi_j(\mathbf{v}_I)],
\end{aligned}$$

$$A^{(1)}_{jl}\phi_l = (\beta m_j/2\pi)^{3/2}\exp(-\tfrac{1}{2}\beta m_j v^2)n_{0,j}n_{0,j}w_{jl}\int d^3\mathbf{v}_l(\beta n_l/2\pi)^{3/2}$$
$$\times\exp(-\tfrac{1}{2}\beta m_l v_l^2)\phi_l(\mathbf{v}_l),$$
$$w_j = 2\theta_{0,j} + \tilde{\theta}_{0,j} + \beta V_{0,j},$$
$$w_{jl} = w_{lj} = 2\theta_{0,jl} + \tilde{\theta}_{0,jl} + \beta V_{0,jl},\quad j,l=1,2;\quad j\neq l.$$

The operator $A^{(1)}$ is bounded, self-adjoint and positive definite if and only if $f_0^{(1)}$ is locally dynamically stable. The operator $A^{(1)}P_0^{(1)}$ equals the sum of the odd part $A^{(1)}P_0^{(1)-} = k\beta_1$, and the even part $A^{(1)}P_0^{(1)+} = \beta_0$.

$$A^{(1)}P_0^{(1)-} = \begin{bmatrix} (A^{(1)}P_0^{(1)-})_{11} & (A^{(1)}P_0^{(1)-})_{12} \\ (A^{(1)}P_0^{(1)-})_{21} & (A^{(1)}P_0^{(1)-})_{22} \end{bmatrix}, \tag{6}$$

where:

$$(A^{(1)}P_0^{(1)-})_{jj}\phi_j^{(1)}$$
$$= k\beta_1 = (\beta m_j/2\pi)^{3/2}\exp(-\tfrac{1}{2}\beta m_j v^2)(-ikn_{0j})$$
$$\times[\mathbf{v}_3\phi_j^{(1)} + n_{0,j}(\theta_{0,j} + \tilde{\theta}_{0,j} + \beta V_{0,j})\int d^3\mathbf{v}_I(\beta m_j/2\pi)^{3/2}$$
$$\times\exp(-\tfrac{1}{2}\beta m_j v_I^2)(\mathbf{v}_I + \mathbf{v})_3\phi_j^{(1)}(\mathbf{v}_I) + (3/4\pi)n_{0,j}\theta_{0,j}\int d^2\kappa$$
$$\times\int d^3\mathbf{v}_I(\beta m_j/2\pi)^{3/2}\exp(-\tfrac{1}{2}\beta m_j v_1^2)(\mathbf{v}-\mathbf{v}_I)_\delta\kappa_\delta\kappa_3\phi_j^{(1)}$$
$$\times(\mathbf{v}_I - \kappa(\mathbf{v}_I - \mathbf{v})_\alpha\kappa_\alpha)];$$
$$(A^{(1)}P_0^{(1)-})_{jl}\phi_l^{(1)} = (\beta m_j/2\pi)^{3/2}\exp(-\tfrac{1}{2}\beta m_j v^2)(-ikn_{0,j}n_{0,l})$$
$$\times[(\theta_{0,jl} + \tilde{\theta}_{0,jl} + \beta V_{0,jl})\int d^2\mathbf{v}_l(\beta m_l/2\pi)^{3/2}$$
$$\times\exp(-\tfrac{1}{2}\beta m_l v_l^2)(\mathbf{v}_I + \mathbf{v})_3\phi_l^{(1)}(\mathbf{v}_l) + (3/4\pi)\theta_{0,jl}\int d^2\kappa$$
$$\times\int d^3\mathbf{v}_l(\beta m_l/2\pi)^{3/2}\exp(-\tfrac{1}{2}\beta m_l v_l^2)$$
$$\times(\mathbf{v}-\mathbf{v}_l)_\delta\kappa_\delta\kappa_3\phi_l^{(1)}[\mathbf{v}_l - (2m_l/m_1 + m_2)[\mathbf{v}_l - \mathbf{v})_\alpha\kappa_\alpha]];$$

$j,l=1,2$, $j\neq l$. In (6) we used the notation $\theta_{0,j} = \delta^2\Theta/\delta n_j\delta n_j$, $\tilde{\theta}_{0,j} = \delta^3\Theta/\delta n_j\delta n_j\delta n_j$, $\theta_{0,jl} = \theta_{0,lj} = \delta^2\Theta/\delta n_l\delta n_j$, $\tilde{\theta}_{0,jl} = \tilde{\theta}_{0,lj} = \delta^3\Theta/\delta n_l\delta n_l\delta n_j + \delta^3\Theta/\delta n_j\delta n_j\delta n_l$.

All derivatives are evaluated at $(n_{0,1}, n_{0,2})$; $j,l=1,2$, $j\neq l$; $V_{0,j} = \int d^3\mathbf{r}' V_j(|\mathbf{r}-\mathbf{r}'|)$, $V_{0,jl} = \int d^3\mathbf{r}' V_{jl}(|\mathbf{r}-\mathbf{r}'|)$. The operator $A^{(1)}P_0^{(1)-}$ is a skew-adjoint (in $\mathbb{C}L_2$ space) linear operator.

The operator $A^{(1)}P_0^{(1)+}$ is the symmetric linearized Boltzmann collision operator for the two component hard sphere gas.[1] We shall not need its explicit form,[1] only its properties that can be easily obtained from the well known properties of the Boltzmann collision operator[5] will be used. (i) The nullspace H_0 of $A^{(1)}P_0^{(1)+}$ (i.e., $H_0 = \{\phi\in H_0^{(1)};\ A^{(1)}P_0^{(1)+}\phi = 0\}$) is the six-dimensional Hilbert space spanned by

$$\begin{pmatrix}1\\0\end{pmatrix},\ \begin{pmatrix}0\\1\end{pmatrix},\ \begin{pmatrix}m_1v^2\\m_2v^2\end{pmatrix},\ \begin{pmatrix}m_1\mathbf{v}\\m_2\mathbf{v}\end{pmatrix}.$$

(ii) The operator $A^{(1)}P_0^{(1)+}$, is self-adjoint and negative definite in the complement $H_0^{(1)c}$ of H_0 in $H_0^{(1)}$. (iii) $A^{(1)}P_0^{(1)+} = K - N$, where K is a bounded and compact linear operator.

$$N\phi =$$
$$\begin{bmatrix} n_{0,1}(n_{0,1}\nu_{11}(v) + n_{0,2}\nu_{12}(v)) & 0 \\ 0 & n_{0,2}(n_{0,2}\nu_{22}(v) + n_{0,1}\nu_{21}(v)) \end{bmatrix}\begin{pmatrix}\phi_1\\\phi_2\end{pmatrix}$$

where $\nu_{jl}(v)$ are real valued functions such that $\nu_{jl}(v) \sim v$ for $v\to\infty$.

Now, we can prove for $A^{(1)}P_0^{(1)}$, the familiar spectral theorem.[2,8-10] The residual spectrum of $A^{(1)}P_0^{(1)}$ is empty. The halfplane $\mathrm{Re}\omega > 0$ belongs to the resolvent set; in the region $-\nu_0 \le \mathrm{Re}\omega \le 0$ there can lie only points of the point spectrum,

$$0 < \nu_0 = \min_{\substack{i,v\\ i\neq j}}\{n_{0,i}\nu_{ii}(v) + n_{0,j}\nu_{ij}(v)\}.$$

If k is sufficiently small then the sixfold degenerate eigenvalue $\omega = 0$ of $A^{(1)}P_0^{(1)+} = (A^{(1)}P_0^{(1)})|_{k=0}$ splits into six eigenvalues, ω_i, $i=0,1,\ldots,5$, that are the closest to zero spectral points of $A^{(1)}P_0^{(1)}$. The values of ω_i and the corresponding eigenfunction ψ_i, $i=0,1,\ldots,5$ can be calculated by using the standard perturbation method. The perturbation series is convergent for k sufficiently small.

The proof of this theorem follows the arguments used first by Wing[8] and McLennan.[2] The arguments are based on the theorems of Weyl, Kato and Rellich[9] (see also the proof of Theorem 3 in Ref. 10).

4. THE LOCAL TWO COMPONENT FLUID DYNAMICS

A state of the class $\mathcal{C}$ of two component systems is described in fluid dynamics by $f^{(2)} \equiv (C_1, E, N, \mathbf{U})$, where C_1, E, N are functions $\mathbb{R}^3 \to \mathbb{R}^+$, $(\mathbf{r}) \mapsto C_1(\mathbf{r})$, etc., $\mathbf{U}$: $\mathbb{R}^3 \to \mathbb{R}^3$, $(\mathbf{r}) \mapsto \mathbf{U}(\mathbf{r})$. The involution $\mathcal{J}^{(2)}$ is defined by $(C_1, E, N, \mathbf{U}) \overset{\mathcal{J}^{(2)}}{\mapsto} (C_1, E, N, -\mathbf{U})$. The function $\mathbb{V}^{(2)}$ appears to be[1]

$$\mathbb{V}^{(2)}(f^{(2)};\beta,\gamma,\delta) = \int d^3\mathbf{r}(S^{(2)}(f^{(2)}) + \beta E + \gamma(1/N) + (1/2)\beta U_\alpha U_\alpha + \delta C_1); \tag{7}$$

where β,γ have the same meaning in thermodynamics as the same symbols introduced in the Sec. 3, $\delta = \alpha_2 - \alpha_1$. By using (7), the thermodynamic conjugates of β,γ,δ are expressed in terms of $f^{(2)}$, thus, $f^{(2)}$ is related to the thermodynamic state variables. The position coordinate independent equilibrium state $f_0^{(2)} \equiv (C_{0,1}, E_0, N_0, 0)$ is the nondegenerate critical point of $\mathbb{V}^{(2)}$. Thus, we have the one to one relations between the positive constants $C_{0,1}$, E_0, N_0 and β,γ,δ. In order to simplify the notation, we shall write $f_0^{(2)} \equiv (C_1, E, N, 0)$. The phenomenological quantities $\mathcal{P}^{(2)}$ introduced by the two component fluid mechanics are $\mathcal{P}^{(2)} = \{\mathcal{T}, L, \eta, \eta_v\}$, where $\mathcal{T}$ is a map

$$(C_1(\mathbf{r}),\ E(\mathbf{r}),\ N(\mathbf{r}),\ \mathbf{U}(\mathbf{r}))$$
$$\overset{\mathcal{T}}{\mapsto}(C_1^*(\mathbf{r}),\ E^*(\mathbf{r}),\ N^*(\mathbf{r}),\ \mathbf{U}^*(\mathbf{r})),\quad L = \begin{pmatrix} d & \lambda_{12} \\ \lambda_{12} & \lambda_{22}\end{pmatrix}$$

is a nonsingular positive definite matrix, its entries are functions of $f^{(2)}$: η, η_v also depend on $f^{(2)}$ and $\eta > 0$, $(\tfrac{1}{3}\eta + \eta_v) > 0$. The map $\mathcal{T}$ is related to $S^{(2)}$ by the relations $\partial S^{(2)}/\partial C_1 = C_1^*$, $\partial S^{(2)}/\partial E = E^*$, $\partial S^{(2)}/\partial(1/N) = N^*E^*$.

The two component fluid dynamics is linearized at $f_0^{(2)}$. By $\phi^{(2)}$, we denote an element of the tangent space $H_0^{(2)}$ of $\mathcal{H}^{(2)}$ at $f_0^{(2)}$ on which the linearized and Fourier transform [with fixed $\mathbf{k} \equiv (0,0,k)$] is defined. $H_0^{(2)}$ is a six-dimensional complex Hilbert space. The inner product $\langle\phi,\psi\rangle^{(2)}$ in $H_0^{(2)}$ is defined by $\langle\phi,\psi\rangle^{(2)} = (\phi, A^{(2)}\psi)$, where (ϕ,ψ) denotes the standard inner product in a six-dimensional $\mathbb{C}L_2$ space. The operator $A^{(2)}$ is[1]

$$A^{(2)}=\begin{matrix}\\ 3\{\\ 3\{\end{matrix}\overset{\overset{3}{\frown}\ \overset{3}{\frown}}{\begin{bmatrix} A_{11}^{(2)} & 0 \\ 0 & 1 \end{bmatrix}},\qquad (8)$$

$$A_{11}^{(2)}=\begin{bmatrix} \Delta_{c_1} & \Delta_e & \Delta_n \\ \Delta_e & \beta\tau_e & \beta\tau_n \\ \Delta_n & \beta\tau_n & (1/M^2)(p_n-\gamma\tau_n) \end{bmatrix},$$

$\tau=-1/E^*$, $\Delta=(1/\beta)C_1^*$, $p=N^*$; the indices c_1, e, n mean differentiations with respect to C_1, E, N respectively. All derivatives are evaluated at $f_0^{(2)}$. The operator, $P_0^{(2)+}$ equals

$$P_0^{(2)+}=-\frac{1}{M}k^2\overset{\overset{3}{\frown}\ \overset{3}{\frown}}{\begin{bmatrix} \beta L_{11}A_{11}^{(2)} & 0 \\ 0 & L_{22} \end{bmatrix}}\begin{matrix}\}3\\ \}3\end{matrix},\qquad (9)$$

where

$$L_{11}=\begin{bmatrix} d & \lambda_{12} & 0 \\ \lambda_{12} & \lambda_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix},\quad L_{22}=\begin{bmatrix} \frac{4}{3}\eta+\eta_v & 0 & 0 \\ 0 & \eta & 0 \\ 0 & 0 & \eta \end{bmatrix}.$$

The operator, $P_0^{(2)-}$ equals

$$P_0^{(2)-}=-\frac{ik}{M}\begin{bmatrix} 0&0&0&0&0&0\\ 0&0&0&\gamma/\beta&0&0\\ 0&0&0&M^2&0&0\\ p_{c_1}&p_e&p_n&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0 \end{bmatrix}.\qquad (10)$$

5. SPACE $H_0^{(1,2)}$

In addition to the notation introduced in Sec. 3, we shall use

$$n_{0,1}+n_{0,2}=n,\quad m_1n_{0,1}=M_1,\quad m_2n_{0,2}=M_2,$$
$$M_1+M_2=M,\quad M_1=MC_1,\quad M_2=MC_2,$$
$$N_i=2m_i(m_1+m_2)^{-1},\quad i=1,2,\quad N=2m_1m_2(m_1+m_2)^{-1},$$
$$r_i=n_{0,i}+(n_{0,i})^2w_i,\quad i=1,2,\quad r_{12}=n_{0,1}n_{0,2}w_{12}=r_{21},$$
$$r=\det\begin{bmatrix} r_2 & r_{12} \\ r_{12} & r_2 \end{bmatrix},\quad r_1C_2-r_{12}C_1=R_1,\qquad (11)$$
$$r_{12}C_2-r_2C_1=R_2,\quad C_2R_1-C_1R_2=D,\quad a_1=-MR_2/r,$$
$$a_2=MR_1/r.$$

The nullspace H_0 of $A^{(1)}P_0^{(1)+}$ is the six-dimensional Hilbert space spanned by (see Sec. 3)

$$h_0=\begin{pmatrix} a_1 \\ a_2 \end{pmatrix},\quad h_1=\begin{pmatrix} C_2 \\ -C_1 \end{pmatrix},\quad h_2=\frac{1}{n}\begin{bmatrix} \frac{1}{2}m_1v^2-3/2\beta \\ \frac{1}{2}m_2v^2-3/2\beta \end{bmatrix},$$
$$h_{2+i}=\frac{1}{M}\begin{pmatrix} m_1v_i \\ m_2v_i \end{pmatrix},\quad i=1,2,3.\qquad (12)$$

The vectors h_i, $i=0,1,\ldots,5$, form an orthogonal basis in H_0, i.e., $\langle h_i,h_j\rangle=\Lambda_i\delta_{ij}$, where $\Lambda\equiv[(M^2/r)D,D,3/2\beta^2n,1/\beta M,1/\beta M,1/\beta M]$. Hereafter $\langle\ ,\ \rangle$ means $\langle\ ,\ \rangle^{(1)}$. The spectral theorem (see Sec. 3) for $A^{(1)}P_0^{(1)}$ allows us to find $H_0^{(1,2)}$ by applying the standard perturbation method. The perturbation $k\beta_0=ikA^{(1)}P_0^{(1)-}$ (k is a small parameter) will bring about splitting of the sixfold degenerate eigenvalue $\omega=0$ of $\beta_0=A^{(1)}P_0^{(1)+}$ into six eigenvalues $\omega_0,\omega_1,\ldots,\omega_5$ that are the closest to the zero spectral points $A^{(1)}P_0^{(1)}$. The corresponding eigenfunctions will be $\psi_0,\ldots,\psi_5$. We shall calculate $\omega_0,\ldots,\omega_5$ up to order k^2, and $\psi_0,\ldots,\psi_5$ up to order k. The calculations are simple in principle, so that we shall write only the results.

$$\beta_1h_0=-iM\begin{pmatrix} v_3C_1 \\ v_3C_2 \end{pmatrix},\quad \beta_1h_1=-i\begin{pmatrix} v_3R_1 \\ v_3R_2 \end{pmatrix},$$
$$\beta_1h_2=-i\frac{1}{n}\begin{pmatrix} v_3X_1^{(2,0)}+\frac{1}{2}v_3v^2X_1^{(2,1)} \\ v_3X_2^{(2,0)}+\frac{1}{2}v_3v^2X_2^{(2,1)} \end{pmatrix},$$
$$\beta_1h_3=-i\frac{1}{M}\begin{pmatrix} X_1^{(3,0)}+v_3^2X_1^{(3,1)}+\frac{1}{2}v^2X_1^{(3,2)} \\ X_2^{(3,0)}+v_3^2X_2^{(3,1)}+\frac{1}{2}v^2X_2^{(3,2)} \end{pmatrix},\qquad (13)$$
$$\beta_1h_4=-i\frac{1}{M}\begin{pmatrix} v_3v_2X_1^{(3,1)} \\ v_3v_2X_2^{(3,1)} \end{pmatrix},\quad \beta_1h_5=-i\frac{1}{M}\begin{pmatrix} v_3v_1X_1^{(3,1)} \\ v_3v_1X_2^{(3,1)} \end{pmatrix},$$

where

$$X_i^{(2,0)}=-(1/2\beta)[3n_i+n_i^2\theta_i+N_1N_2n_1n_2\theta_{12}$$
$$+2N_i(1-N_i)m_1n_2\theta_{12}],$$
$$X_i^{(2,1)}=m_i(n_i+\tfrac{3}{2}n_i^2\theta_i+\tfrac{3}{2}N_1N_2n_1n_2\theta_{12}),$$
$$X_i^{(3,0)}=(1/\beta)[n_i^2(\theta_i+\tilde\theta_i+\beta V_i)+n_1n_2(\theta_{12}+\tilde\theta_{12}+\beta V_{12})$$
$$-(1-N_j)n_1n_2\theta_{12}],$$
$$X_i^{(3,1)}=m_i(n_i+\tfrac{2}{5}n_i^2\theta_i+\tfrac{3}{5}N_jn_1n_2\theta_{12}),$$
$$X_i^{(3,2)}=m_i(\tfrac{2}{5}n_i^2\theta_i+\tfrac{2}{5}N_jn_1n_2\theta_{12}),\quad i=1,2,\ j\neq i.$$

We have simplified the notation by writing $n_1,\theta_1,\ldots,$ etc., instead of $n_{0,1},\theta_{0,1},\ldots,$ etc. By using (13) one gets

$$W_{ij}=\langle h_i^*,\beta_1h_j\rangle=W_{ji},$$
$$W_{ij}=0\ \text{ except for }\ W_{03}=-i(1/\beta),$$
$$W_{13}=-i(1/\beta M)(R_1+R_2),$$
$$W_{23}=-i(1/nM\beta^2)(n_1+n_2+n_1^2\theta_1+n_2^2\theta_2+n_1n_2\theta_{12}).$$

The eigenfunctions ψ_i, $i=0,\ldots,5$ that span $H^{(1,2)}$ are

$$\psi_i=\psi_i^{(\mathrm{I})}+k\psi_i^{(\mathrm{II})}+k\psi_i^{(\mathrm{III})}.\qquad (15)$$

$\psi_i^{(\mathrm{I})}$ lie in H_0 and are written in the basis $\{h_i\}$, $i=0,\ldots,5$:

$$-\psi_0^{(\mathrm{I})}=((\Lambda_1\Lambda_2/\Lambda_0)W_{03}^2,\ \Lambda_2W_{13}W_{03},\ \Lambda_1W_{23}W_{03},\ \Lambda\Lambda_1\Lambda_2W_{03},\ 0,\ 0),$$

$$-\psi_1^{(\mathrm{I})}=((\Lambda_1\Lambda_2/\Lambda_0)W_{03}^2,\ \Lambda_2W_{13}W_{03},\ \Lambda_1W_{23}W_{03},\ -\Lambda\Lambda_1\Lambda_2W_{03},\ 0,\ 0),$$

$$\psi_2^{(\mathrm{I})}=((\psi_2^{(\mathrm{I})})_1,\ (\psi_2^{(\mathrm{I})})_2,\ (\psi_2^{(\mathrm{I})})_3,\ 0,\ 0,\ 0),\qquad (15a)$$

$$\psi_3^{(\mathrm{I})} = ((\psi_3^{(\mathrm{I})})_1,\ (\psi_3^{(\mathrm{I})})_2,\ (\psi_3^{(\mathrm{I})})_3,\ 0,\ 0,\ 0),\quad \psi_4^{(\mathrm{I})} = (0,\ 0,\ 0,\ 0,\ 1,\ 0),\quad \psi_5^{(\mathrm{I})} = (0,\ 0,\ 0,\ 0,\ 0,\ 1),$$

where

$$(\psi_i^{(\mathrm{I})})_1 = M_{23} - (\kappa_i - M_{22})\lambda_1 W_{13} W_{03},\ (\psi_i^{(\mathrm{I})})_2 = (\kappa_i - M_{22})(\lambda_0 W_{23}^2 + \lambda_1 W_{03}^2),\ (\psi_i^{(\mathrm{I})})_3 = -M_{23} W_{03}/W_{23} - (\kappa_i - M_{22}) W_{13} W_{23},\ i = 2,3,$$

$\Lambda = -i(W_{03}^2/\lambda_0\lambda_3 + W_{13}^2/\lambda_1\lambda_3 + W_{23}^2/\lambda_2\lambda_3)^{1/2}$ and $M_{i,j}$, κ_i $(i,j=2,3)$, are defined in (16). From (12) and (15a), we have

$$\langle \psi_i^{(\mathrm{I})},\ \psi_j^{(\mathrm{I})} \rangle = \Psi_i \delta_{ij},\quad \Psi_0 = \Psi_1 = \frac{1}{\beta^2} 2\lambda_1\lambda_2 \left(\frac{\lambda_1\lambda_2}{\lambda_0} \frac{1}{\beta^2} + \lambda_2 W_{13}^2 + \lambda_1 W_{23}^2 \right),\ \Psi_2 = M_{23}^2 \tilde{\lambda}_2 + (\kappa_2 - M_{22})^2 \tilde{\lambda}_3,$$

$$\Psi_3 = M_{23}^2 \tilde{\lambda}_2 + (\kappa_3 - M_{22})^2 \tilde{\lambda}_3,\ \Psi_4 = \Psi_5 = \tilde{\lambda}_2 = \lambda_0 + \frac{W_{03}^2}{W_{23}} \lambda_2,\ \tilde{\lambda}_3 = \lambda_0 \lambda_2^2 W_{03}^2 W_{13}^2 + \lambda_1(\lambda_0 W_{23}^2) + \lambda_2 H_0^2 W_{13}^2)^2 + \lambda_2 \lambda_0^2 W_{13}^2 W_{23}^2. \tag{15b}$$

Continuing the calculations of ψ_i, we have

$$\psi_2^{(\mathrm{II})} = \beta_0^{-1} \chi_2, \tag{15c}$$

where

$$\chi_0 = (-\Lambda A - \beta_1)\psi_0,\ \chi_1 = (\Lambda A - \beta_1)\psi_1,$$
$$\chi_2 = -\beta_1 \psi_2,\ \chi_3 = -\beta_1 \psi_3,\ \chi_4 = -\beta_1 \psi_4 = \chi_5, \tag{15d}$$

and $\psi_i^{(\mathrm{II})}$ are orthogonal to H_0. The functions χ_i can be written explicitly by using (13). To obtain $\beta_0^{-1}\chi_i$ one has to solve the system of integral equations involving the Boltzmann collision operators. The term $k\psi_i^{(\mathrm{III})}$ in (15) is an arbitrary vector of H_0. We shall write

$$\psi_i^{(\mathrm{III})} = i \sum_{j=0}^{5} g_j^i \psi_j^{(\mathrm{I})}, \tag{15e}$$

where $g_j^i = 0$ for $i = j$, and for $i = 4,5$ and all j, and $j = 4,5$ and all i. The twelve remaining coefficients g_j^i are unknown and will be specified later in Sec. 8.

The eigenvalues $\omega_0, \ldots, \omega_5$ are

$$\begin{aligned}
&\omega_0 = \omega k^2 - ik\Lambda,\\
&\omega_1 = \omega k^2 + ik\Lambda,\\
&\omega_2 = (\tilde{\lambda}_2 \tilde{\lambda}_3)^{-1}(M_{22}\tilde{\lambda}_3 + M_{33}\tilde{\lambda}_2 + D),\\
&\omega_3 = (\tilde{\lambda}_2 \tilde{\lambda}_3)^{-1}(M_{22}\tilde{\lambda}_3 + M_{33}\tilde{\lambda}_2 - D),\\
&\omega_4 = \omega_5 = (\Psi_4)^{-1} M_{44},
\end{aligned} \tag{16}$$

where

$$\begin{aligned}
&M_{ij} = (\tilde{\chi}_i, \beta_0^{-1}\tilde{\chi}_j) = M_{ji},\quad i,j = 2,3,\\
&M_{ii} = (\chi_i, \beta_0^{-1}\chi_i),\quad i = 4,5,\\
&\tilde{\chi}_i = -\beta_1 \tilde{h}_i,\quad i = 2,3,\\
&\tilde{h}_2 = (1,\ 0,\ -W_{03}/W_{23},\ 0,\ 0,\ 0),\\
&-\tilde{h}_3 = (\lambda_2 W_{03} W_{13},\ -\lambda_0 W_{23}^2 - \lambda_2 W_{03}^2,\ \lambda_0 W_{12} W_{23},\ 0,\ 0,\ 0),\\
&\langle \tilde{h}_i, \tilde{h}_j \rangle = \tilde{\lambda}_i \delta_{ij},\quad i,j = 2,3,\\
&D^2 = M_{22}^2 \tilde{\lambda}_3^2 + M_{33}^2 \tilde{\lambda}_3^2 + 2\tilde{\lambda}_2 \tilde{\lambda}_3 (2M_{23}^2 - M_{22} M_{33}),\\
&\omega = (\Psi_0)^{-1}(\chi_0, \beta_0^{-1}\chi_0) = (\Psi_1)^{-1}(\chi_1, \beta_0^{-1}\chi_1).
\end{aligned}$$

The last equality follows from $\Psi_0 = \Psi_1$ and from $\psi_1 = \mathcal{J}^{(1)}\psi_0 = \mathcal{J}^{(\Lambda)}\psi_0$, where $\mathcal{J}^{(1)}$ is the involution defined in Sec. 3 and $\mathcal{J}^{(\Lambda)}$ is the transformation induced by $\Lambda \to -\Lambda$. In (15), we used $\kappa_2 = \tilde{\lambda}_2\omega_2$; $\kappa_3 = \tilde{\lambda}_2\omega_3$.

6. SPACE $H_0^{(2,1)}$

Now, we look for $H_0^{(2,1)}$ with basis ϕ_i, $i = 0, \ldots, 5$ such that the time development of $a_i = \langle \phi^{(1)}, \phi_i \rangle$, derived from $d\phi^{(1)}/dt = P_0^{(1)}\phi^{(1)}$ restricted to $H_0^{(1,2)}$, will be identical to $\partial\phi^{(2)}/\partial t = P_0^{(2)}\phi^{(2)}$ with $P_0^{(2)}$ in (9), (10). Moreover the inner product in the space spanned by $\{a_i\}$, $i = 0, \ldots, 5$, induced from the inner product in $H_0^{(1)}$ will be identical to the inner product in $H_0^{(2)}$, thus $(A^{(2)})^{-1} \equiv (\langle \phi_i, \phi_j \rangle)$. In order to identify $H_0^{(2)}$ to the space spanned by $\{a_i\}$, $i = 0, \ldots, 5$, we have to find $\{\phi_i\}$, $j = 0, \ldots, 5$, such that the involution $\mathcal{J}^{(1)}$ induces the involution $\mathcal{J}^{(2)}$ (notice also that $H_0^{(2)-}$ is spanned by vectors, $H_0^{(2)+}$ by scalars). By using (13), we easily obtain $\partial\langle h_0\phi^{(2)}\rangle/\partial t = -ikM\langle h_3, \phi^{(1)}\rangle$. These considerations suggest the following choice of $\{\phi_i\}$:

$$\begin{aligned}
&\phi_0 = (K_1, K_2, 0, 0, 0, 0),\\
&\phi_1 = (K_3, K_4, \eta/M, 0, 0, 0),\\
&\phi_2 = (1, 0, 0, 0, 0, 0),\\
&\phi_3 = (0, 0, 0, 1, 0, 0),\\
&\phi_4 = (0, 0, 0, 0, 1, 0),\\
&\phi_5 = (0, 0, 0, 0, 0, 1).
\end{aligned}$$

The functions ϕ in (17) are written in the basis $\{h_i\}$ defined in (12). The four coefficients $K_1, \ldots, K_4$ will be specified later. The matrix $A^{-1} \equiv (\langle \phi_i, \phi_j \rangle)$ is readily obtained

$$A^{-1} = \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & A_{22}^{-1} \end{bmatrix}, \tag{18}$$

where

$$A_{22}^{-1} = \frac{1}{\beta M} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and

$$A_{11}^{-1} = \frac{\lambda_1 M}{r} \begin{bmatrix} K_1 + K_2^2 \dfrac{r}{M} & K_1 K_3 + K_2 K_4 \dfrac{r}{M} & K_1 \\ K_1 K_3 + K_2 K_4 \dfrac{r}{M} & K_3^2 + \dfrac{r}{M}\left(K_4^2 + \dfrac{3}{2\beta^2}\dfrac{n}{M^2\lambda_1}\right) & K_3 \\ K_1 & K_3 & 1 \end{bmatrix}.$$

The inverse A, of A^{-1} that should be identical to $A^{(2)}$ is

$$A = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}, \tag{19}$$

where

$$A_{22}=\beta M\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix},$$

$$A_{11}=\frac{(2\beta^2/3)Mr}{K_1(K_1-1)+K_2^2(r/M)}\times\begin{bmatrix}K_4^2+\frac{3}{2\beta^2}\frac{n}{M^2\lambda_1}&-K_2K_4&(A_{11})_1\\-K_2K_4&K_2^2&K_1K_2K_4-K_2^2K_3\\(A_{11})_1&K_1K_2K_4-K_2^2K_3&(A_{11})_2\end{bmatrix},$$

$$(A_{11})_1=K_2K_3K_4-K_1K_4^2-3n/2\beta^2M^2\lambda_1,$$

$$(A_{11})_2=K_1^2K_4^2+K_2^2K_3^2-2K_1K_2K_3K_4+(3n/2\beta^2M^2\lambda_1)\times[K_1^2+(r/M)K_2^2-(r/M)K_2^2K_4^2].$$

We have to know how to express $\{\phi_i\}$ in terms of $\{\psi_i\}$ and $\{\psi_i\}$ in terms of $\{\phi_i\}$, $i=0,\ldots,5$. Since we know both $\{\phi_i\}$ [see (17)] and $\{\psi_i\}$ [see (15)] we find

$$\phi_i=\sum_{j=0}^{5}\alpha_j^i\psi_j,\qquad \psi_i=\sum_{j=0}^{5}\tilde{\alpha}_j^i\phi_j, \tag{20}$$

where

$$(\alpha_j^i)\equiv\alpha=\alpha^{(\mathrm{I})}+k\alpha^{(\mathrm{II})},\quad(\tilde{\alpha}_j^i)\equiv\tilde{\alpha}^{(\mathrm{I})}+k\tilde{\alpha}^{(\mathrm{II})},$$

$$\alpha^{(\mathrm{II})}=-i\alpha g,\quad\tilde{\alpha}^{(\mathrm{II})}=ig\tilde{\alpha}$$

[we have used the notation $g=(g_j^i)$—see (15e)] and

$$\alpha^{(\mathrm{I})}=\begin{bmatrix}(\alpha^{(\mathrm{I})})_{11}&0\\0&(\alpha^{(\mathrm{I})})_{22}\end{bmatrix},$$

$$(\alpha^{(\mathrm{I})})_{11}=\begin{bmatrix}\alpha_0^0&\alpha_0^1&\alpha_0^2&\alpha_0^3\\\alpha_0^0&\alpha_0^1&\alpha_0^2&-\alpha_0^3\\\alpha_2^0&\alpha_2^1&\alpha_2^2&0\\\alpha_3^0&\alpha_3^1&\alpha_3^2&0\end{bmatrix},\quad\alpha_l^i=(\Psi_l)^{-1}\langle\psi_l^{(\mathrm{I})},\phi_j\rangle.$$

Using (15) and (17), we obtain

$$\alpha_0^0=(\Psi_0)^{-1}\lambda_1\lambda_2(K_1W_{03}^2+K_2W_{03}W_{13}),$$

$$\alpha_0^1=(\Psi_0)^{-1}\lambda_1\lambda_2[K_3W_{03}^2+K_4W_{03}W_{13}+(n/M)W_{03}W_{23}],$$

$$\alpha_0^2=(\Psi_0)^{-1}\lambda_1\lambda_2W_{03}^2,\quad\alpha_0^3=(\Psi_0)^{-1}\lambda_1\lambda_2\lambda_3\Lambda W_{03},\quad\text{etc.},$$

$$\alpha_{22}^{(\mathrm{I})}=\begin{bmatrix}1&0\\0&1\end{bmatrix},$$

$$\tilde{\alpha}^{(\mathrm{I})}=\begin{bmatrix}(\tilde{\alpha}^{(\mathrm{I})})_{11}&0\\0&(\tilde{\alpha}^{(\mathrm{I})})_{22}\end{bmatrix},$$

$$(\tilde{\alpha}^{(\mathrm{I})})_{11}=\begin{bmatrix}\tilde{\alpha}_0^0&\tilde{\alpha}_0^0&\tilde{\alpha}_0^2&\tilde{\alpha}_0^3\\\tilde{\alpha}_1^0&\tilde{\alpha}_1^0&\tilde{\alpha}_1^2&\tilde{\alpha}_1^3\\\tilde{\alpha}_2^0&\tilde{\alpha}_2^0&\tilde{\alpha}_2^2&\tilde{\alpha}_2^3\\\tilde{\alpha}_3^0&-\tilde{\alpha}_3^0&0&0\end{bmatrix}$$

$$(\tilde{\alpha}^{(\mathrm{I})})_{22}=\begin{bmatrix}1&0\\0&1\end{bmatrix}.$$

The quantities $\tilde{\alpha}_j^i$ can be readily obtained since $\tilde{\alpha}^{(\mathrm{I})}$ is clearly the inverse of $\alpha^{(\mathrm{I})}$. We mention only that $\tilde{\alpha}_3^0=-1/2\alpha_0^3$. The two component Enskog—Vlasov dynamics restricted to $H_0^{(1,2)}$ and represented in $H_0^{(2,1)}$ is now

$$\frac{da}{dt}=Ba,$$

or in components

$$\frac{da_i}{dt}=B_{ij}a_i, \tag{21}$$

where

$$B_{ij}=\sum_{l=0}^{5}\overline{(\alpha_l^i\bar{\omega}_l\tilde{\alpha}_j^l)};$$

the bar means the complex conjugation.

7. COMPARISON OF B^- AND $P_0^{(2)-}$

We shall see that the comparison of the odd parts of B [see (21)] and $P_0^{(2)}$ [see (10)] allows us to find K_1, K_2, K_3, K_4 and thus to find $A^{(2)}$ (i.e., the thermodynamic phenomenological coefficient of fluid mechanics) in terms of the phenomenological quantities of the Enskog—Vlasov dynamics. The comparison of B^+ and $P_0^{(2)+}$ that is done in the next section will allow us to find the twelve so far unspecified coefficients g_j^i, and the dynamic phenomenological coefficients of fluid mechanics in terms of the phenomenological quantities of the Enskog—Vlasov dynamics.

By using (20), (16) and (21), we obtain

$$B_{03}^-=(1/\lambda_3)(K_1W_{03}+K_2W_{13}),$$

$$B_{13}^-=(1/\lambda_3)(K_3W_{03}+K_4W_{13}+(n/M)W_{23}),$$

$$B_{23}^-=(1/\lambda_3)W_{03},\quad B_{30}^-=2\Lambda\tilde{\alpha}_0^0\alpha_0^3,\qquad B_{31}^-=2\Lambda\tilde{\alpha}_1^0\alpha_0^3,\quad B_{32}^-=2\Lambda\tilde{\alpha}_2^0\alpha_0^3. \tag{22}$$

All remaining B_{ij}^- (i.e., B_{ij} proportional to k) equal zero. By setting (22) equal to (10), using (4), (8), and (19), one obtains five equations for four unknowns $K_1,\ldots,K_4$. This system of equations has one solution (one equation is redundant which serves also as a check of our calculations). The solution is

$$K_1=-(1/MD)C_1C_2(R_1+R_2),$$

$$K_2=(1/D)C_1C_2,$$

$$K_3=-(1/M^2D)\det\begin{bmatrix}k_1&R_1\\k_2&R_2\end{bmatrix},\qquad K_4=-(1/MD)\det\begin{bmatrix}C_1&k_1\\C_2&k_2\end{bmatrix}, \tag{23}$$

where we used the symbols introduced in (11) and

$$k_1=[(1-N_1)/M]n_1n_2(m_1+m_2)+(\beta/2)[(1+C_2)V_1n_1^2-C_1V_2n_2^2+2C_2V_{12}n_1n_2],$$

$$k_2=[(1-N_2)/M]n_1n_2(m_1+m_2)+(\beta/2)[(1+C_1)V_2n_2^2-C_2V_1n_1^2+2C_1V_{12}n_1n_2].$$

Combining (22) with (17), we have obtained the relation between the state variables of kinetic theory and fluid mechanics. The relation obtained should be consistent with the relation obtained by comparing the corresponding terms (staying by the same thermodynamic fields) in $\mathbb{V}^{(1)}$ [see (3)] and $\mathbb{V}^{(2)}$ [see (7)]. Indeed, the relation obtained from (22) and (17) is identical to the relation obtained from

$$C_i(\mathbf{r})\left[\sum_{j=1}^{2} m_j \int d^3\mathbf{v} f_j(\mathbf{r},\mathbf{v})\right] = m_i \int d^3\mathbf{v} f_i(\mathbf{r},\mathbf{v}),$$

$$E(\mathbf{r}) = \int d^3\mathbf{v} \sum_{i,j=1}^{2} \sum (1/2) m_i v^2 f_i \tag{24}$$

$$+ \int d^3\mathbf{r}_1 \int d^3\mathbf{v}_1 V_{ij}(|\mathbf{r}-\mathbf{r}_1|) f_i f_j \Big].$$

By using (23) and (19), we obtain all thermodynamic parameters entering fluid dynamics (the "stiffness matrix" in Tisza's terminology[11]) expressed in terms of $p^{(1)}$.

8. COMPARISON OF B^+ AND $P_0^{(2)+}$

In this section, we obtain the twelve unknown coefficients g_j^i, the matrix L and η, η_v. By putting B_{i3}^+ and B_{3i}^+, $i=1,2,3$, equal to zero, one obtains

$$g_3^2 = g_2^3 = 0, \quad g_1^0 + g_0^1 = 0, \quad g_0^2 + g_1^2 = 0,$$
$$g_0^2 + g_1^3 = 0, \quad g_2^0 + g_2^1 = 0, \quad g_3^0 + g_3^1 = 0.$$

From (25) and (15), we thus have $\langle \psi_i, \psi_j \rangle = \Psi_i \delta_{ij} + O(k^2)$.

From B_{44}^+ and B_{55}^+, one obtains

$$\eta = -(1/k^2) M_{44}. \tag{26}$$

From B_{33}^+, by using (26), (15), and (9), η_v is obtained. The last problem is to solve the equations

$$b = L_{11} A_{11}^{(2)}, \tag{27}$$

where

$$b = \begin{bmatrix} B_{00}^+ & B_{01}^+ & B_{02}^+ \\ B_{10}^+ & B_{11}^+ & B_{12}^+ \\ B_{20}^+ & B_{21}^+ & B_{22}^+ \end{bmatrix}$$

and the matrices L_{11} and $A_{11}^{(2)}$ are defined in (9). The entries of the matrix b are

$$B_{ij} = D_{ij} + E_{ij}, \tag{28}$$

where

$$D_{ij} = \sum_{l=0}^{5} \alpha^{(1)i}{}_l \tilde{\omega}_l \tilde{\alpha}^{(1)l},$$

$\tilde{\omega}$ denotes the part of ω proportional to k^2, and

$$E_{ji} = 2\Lambda(\alpha^{(1)j}{}_0 \tilde{\alpha}^{(1)0}{}_i 2g_0^1 + \alpha^{(1)j}{}_0 \tilde{\alpha}^{(1)2}{}_i g_2^1 + \alpha^{(1)j}{}_2 \tilde{\alpha}^{(1)0}{}_i g_0^2$$
$$+ \alpha^{(1)j}{}_0 \tilde{\alpha}^{(1)3}{}_i g_3^1 + \alpha^{(1)j}{}_3 \tilde{\alpha}^{(1)0}{}_i g_0^3).$$

The matrix $A_{11}^{(2)}$ is known from Sec. 7, thus

$$L_{11} = b(A_{11}^{(2)})^{-1}. \tag{29}$$

By putting $(L_{11})_{1i}$ and $(L_{11})_{3i}$, $i=1,2,3$, equal to zero [see the definition of L_{11} in (9)], we have five equations for five unknowns g_0^1, g_2^1, g_3^1, g_0^2, g_0^3. The rest will give L.

9. CONCLUSION

The local two component fluid dynamics (8)—(10) has been derived from the local two component Enskog—Vlasov dynamics. The map $p^{(1)} \to p^{(2)}$ is materialized in the formulas (11)—(29) of Secs. 5—8. The map becomes explicit after solving the system of integral equations (15c) and after performing the corresponding simple arithmetic manipulations that are involved in the notation used in Secs. 5—8.

Finally, we want to summarize the motivation for this paper, the results obtained and the possible applications of the results obtained.

(i) A general structure of the non-Hamiltonian dynamics appearing in nonequilibrium statistical mechanics has been suggested in Ref. 12. The structure has arisen from the study of kinetic theory, fluid mechanics,[1] their interrelationship,[4] and their relationship to thermodynamics. The closest more general dynamical systems of nonequilibrium statistical mechanics are the two component kinetic theory and fluid dynamics. Indeed, we can imagine that the system considered is still the one component system, but our approach to the system (through the observations and measurements) is more detailed. The elements composing the system are not indistinguishable as they are in the one component system but they belong to one of the two (distinguishable by our more precise measurements), kinds of elements. We have shown[1] that the two component dynamical systems of kinetic theory and fluid dynamics possess the structure suggested in Ref. 4 (formulated in a better way in Ref. 1), provided some relations among the phenomenological quantities introduced by the dynamical systems are satisfied. For example, in kinetic theory, the phenomenological quantities θ_{ij} [see (6)] must be derivable from one functional Θ. The unity of the structure is well demonstrated by comparing the explicit calculations in Secs. 5—8 and the corresponding calculations for one component systems.[1]

(ii) A very interesting problem is the behavior of the map $p^{(1)} \to p^{(2)}$ near the critical (in the sense of thermodynamics) points. For one component systems, we have reproduced the van der Waals dynamic and static critical phenomena. In the next paper, we shall derive the dynamic and static critical phenomens based on the map $p^{(1)} \to p^{(2)}$ obtained in this paper. We are also now in position to approach the problem of dynamic and static critical phenomena, indirectly, by using the scaling hypothesis of Green.[13] The one component system is considered as the two component system, the scaling parameter is the ratio of the concentrations of this "artificially created" two component system. The map $p^{(1,1)} \to p^{(1,2)}$ from the phenomenological quantities of two component kinetic theory to the phenomenological quantities of one component kinetic theory (obtained by putting together the results of this paper and Ref. 4) will depend on the scaling parameter. The fixed point in the dependence on the scaling parameter defines the critical values of the thermodynamic parameters α, β (so-called scaling hypothesis).

(iii) The experimental values of $p^{(2)}$ can be compared with the values of $p^{(2)}$ obtained from the map $p^{(1)} \to p^{(2)}$ derived in this paper, provided the experimental values for $p^{(1)}$ are known.

ACKNOWLEDGMENTS

I am indebted to Professor Karel Van Vliet for his support of this research. I have also benefited from dis-

cussions with Professor Luis de Sobrino, Professor Donald Dawson, and Professor G. Scharf.

[1]M. Grmela, J. Math. Phys. 16, 2441 (1975).
[2]J. A. McLennan, Phys. Fluids 8, 1580 (1965).
[3]G. Scharf, Helv. Phys. Acta 40, 929 (1967); 42, 5 (1969).
[4]M. Grmela, Helv. Phys. Acta 47, 677 (1974).
[5]S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-Uniform Gases* (Cambridge U. P., Cambridge, 1952), 2nd ed., Chap. 16, results of H. H. Thorne.
[6]L. Barajas, L. S. Garcia—Colin, and E. Pina, J. Stat. Phys. 7, 161 (1973).
[7]R. B. Griffiths and J. C. Wheeler, Phys. Rev. A 2, 1047 (1970).
[8]G. H. Wing, *An Introduction to Transport Theory* (Wiley, New York, 1962).
[9]T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1966).
[10]M. Grmela, J. Math. Phys. 15, 35 (1974).
[11]L. Tisza, *Generalized Thermodynamics* (M.I.T. Press, Cambridge, Mass., 1966).
[12]M. Grmela, Helv. Phys. Acta 47, 667 (1974).
[13]M. S. Green, Phys. Rev. A 8, 1998 (1973).

Solutions for the general cylindrically symmetric stationary dust model

John Charles Zimmerman

Department of Physics and Astronomy, Clemson University, Clemson, South Carolina 29632
(Received 14 August 1975)

For a dust-filled space-time possessing cylindrical symmetry, the field equations form an underdetermined set. As demonstrated by King [Commun. Math. Phys. **38**, 157 (1974)], by carefully selecting a function it is possible to generate solutions which are either well-behaved or are characterized by one of a number of different types of singularity. The three particular choices for the functions we take produce two nonsingular solutions and one with a Weyl singularity.

1. INTRODUCTION

The gravitational field equations for a cylindrical stationary dust model can be derived by using the tetrad approach of Ellis.[1] We pick an orthonormal tetrad field $\{e_a\}$ in which $e_0 = u$, $e_0 \cdot e_0 = -1$, $e_\alpha \cdot e_\beta = \delta_{\alpha\beta}$, and $e_0 \cdot e_\alpha = 0$ ($\alpha, \beta = 1,2,3$ and $a, b = 0,1,2,3$). The commutation relations of e_a, e_b define the quantities γ^c_{ab} by

$$[e_a, e_b] = \gamma^c_{ab} e_c. \tag{1}$$

The nonzero components of the vorticity vector

$$\omega^a = \tfrac{1}{2}\eta^{abcd} u_b u_{c;d}, \tag{2}$$

and the shear tensor

$$\sigma_{ab} = u_{(a;b)} - \tfrac{1}{3}u^c{}_{;c}(g_{ab} + u_a u_b), \tag{3}$$

can be written in terms of the γ^c_{ab} as

$$\omega = \omega^1 = -\tfrac{1}{2}\gamma^0_{23} = A\cosh\lambda, \tag{4}$$

$$\sigma = \sigma_{23} = -\tfrac{1}{2}\gamma^2_{03} = A\sinh\lambda, \tag{5}$$

where A and λ are functions of r. As King[2] has shown, the field equations, which form an underdetermined set, can be written in terms of an unknown function $\alpha(r)$. These give us:

$$\rho = 4A^2 = -2\alpha''/\alpha, \tag{6}$$

$$e^{2\lambda} = \alpha\alpha''/[2(\alpha' - 1)] = A^2\alpha^2/(1 - \alpha'), \tag{7}$$

$$\gamma_1 = \gamma^1_{13} = (1/\alpha)(\alpha' - 1), \tag{8}$$

$$\gamma_2 = 1/\alpha, \tag{9}$$

where

$$\alpha' = \frac{d\alpha}{dr}, \quad \alpha'' = \frac{d^2\alpha}{dr^2}, \tag{10}$$

and ρ is the matter density.

The form of the metric one obtains by this procedure is

$$ds^2 = \exp[2\textstyle\int\gamma_2 dr][-(f^2 - c^2)dt^2 + (b^2 - e^2)d\phi^2 + 2(fe - bc)d\phi\,dt] + dr^2 + \exp[2\textstyle\int\gamma_1 dr]dz^2, \tag{11}$$

where c, f, b, and e are functions of r satisfying:

$$\frac{db}{dr} = 2\sigma e, \quad \frac{de}{dr} = 2\omega b - \gamma_2 e, \tag{12}$$

$$\frac{dc}{dr} = 2\sigma f, \quad \frac{df}{dr} = 2\omega c - \gamma_2 f. \tag{13}$$

Upon specifying $\alpha(r)$, both the cosmological solution and $\alpha(r)$ must satisfy several conditions imposed by the symmetry and the field equations:

(1) $\rho > 0$, $e^{2\lambda} > 0$ for $0 \leq r \leq$ some r_0.

(2) As $r \to 0$, $\alpha(r) = r + O(r^3)$.

(3) ρ, ω, and σ must be even functions of r.

(4) γ_1 and γ_2 must be odd functions of r.

(5) β, the "boost" to a parallel propagated frame, is finite, where

$$\beta = -\int_{r_0}^{r}(\sigma + \omega)\,dr = -\int_{r_0}^{r} A e^{\lambda} dr. \tag{14}$$

(6) The elementary flatness condition (the circumference of an infinitesimal circle about the axis $r = 0$ equals $2\pi r$), which is expressed as

$$\lim_{r\to 0}(1/r)\exp(\textstyle\int\gamma_2 dr)(b^2 - e^2)^{1/2} = 1, \tag{15}$$

must be satisfied.

By carefully selecting $\alpha(r)$, we can arrange it so that the fluid flow lines are incomplete because the vorticity and shear diverge at a finite value of the proper radial coordinate. If ρ is finite in this frame at that same value of r, then we encounter a Weyl singularity.

2. NONSINGULAR SOLUTIONS

Suppose considering the cylindrical nature of the symmetry, that we choose $\alpha(r) = 2J_1(r)$, where $J_1(r)$ is a first-order Bessel function of the first kind. Then, from (6) and (7),

$$\rho = \tfrac{3}{2} - \tfrac{1}{2}J_3(r)[J_1(r)]^{-1}, \tag{16}$$

$$e^{2\lambda} = [\tfrac{3}{2}J_1^2(r) - \tfrac{1}{2}J_3(r)J_1(r)][1 - J_0(r) - J_2(r)]^{-1}. \tag{17}$$

For $0 \leq r \lesssim 3.2$, $\rho > 0$ and $e^{2\lambda} > 0$. Applying l'Hospital's rule yields

$$\lim_{r\to 0} e^{2\lambda} = 3, \tag{18}$$

thereby excluding the existence of a Weyl singularity or an intermediate singularity.

The remaining conditions on the solution are easily checked. Because Bessel functions of order $J_{2n}(r)$

$(n=0,1,2,3,\cdots)$ are even functions and $J_{2n+1}(r)$ are odd functions, Conditions 3 and 4 are fulfilled. The statement of elementary flatness requires some calculation:

$$\alpha(r)=r-\tfrac{1}{8}r^3+O(r^5)=r[1-r^2/8+O(r^4)], \tag{19}$$

$$\int\gamma_2 dr=\int(1/\alpha)dr=\ln(r)+r^2/16+O(r^4), \tag{20}$$

so that

$$\exp(\int\gamma_2 dr)=r\exp[r^2/16+O(r^4)]. \tag{21}$$

Now, from (4) and (5),

$$\omega=(1/\sqrt{2})[1-r^2/96+O(r^4)], \tag{22}$$

$$\sigma=(1/2\sqrt{2})[1-r^2/48+O(r^4)], \tag{23}$$

which when substituted into (12) give

$$\frac{db}{dr}=\frac{1}{\sqrt{2}}e, \tag{24}$$

$$\frac{de}{dr}=\frac{2}{\sqrt{2}}b-\frac{e}{r}, \tag{25}$$

after dropping terms of order r^2 and higher. The solution is:

$$e(r)=(1/\sqrt{2})r+O(r^3), \tag{26}$$

$$b(r)=1+O(r^2); \tag{27}$$

(15) is therefore satisfied.

We have constructed an example of a nonpathological solution for a general cylindrically symmetric stationary dust model. Our selection for $\alpha(r)$ is a simple function of r which appears to be an obvious choice because of the symmetry, i.e., a Bessel function, and it has given us a well behaved model.

As a second choice, let ρ be a constant,[2] or,

$$\rho=-2\alpha''/\alpha=4A^2=2B^2=\text{const}. \tag{28}$$

Then

$$\alpha=(1/B)\sin(Br), \tag{29}$$

so that

$$e^{2\lambda}=\tfrac{1}{2}\sin^2Br/(1-\cos Br). \tag{30}$$

Take the limit as $r\to 0$. Because $e^{2\lambda}$ is an indeterminate form at this limit, we need to apply l'Hospital's rule. Then,

$$\lim_{r\to 0}e^{2\lambda}=1. \tag{31}$$

That this limit is nonzero can also be seen by making a simple expansion of $e^{2\lambda}$ for small r. We get

$$e^{2\lambda}=1-\tfrac{1}{4}B^2r^2-O(r^4). \tag{32}$$

In Ref. 2, this was accidentally given as an example of a solution demonstrating a Weyl singularity. It does, though, possess the pathological behavior of having an infinite number of axes $\alpha=0$, and also for $r\to 0$, ω is constant and σ goes to zero. This can be seen from the definitions of ω and σ and the expansion in (32). These show that for small values of r,

$$\omega=(1/\sqrt{2})B, \tag{33}$$

$$\sigma=-(1/8\sqrt{2})B^3r^2+O(r^4). \tag{34}$$

Here, the solution of (12) is:

$$e(r)=\sqrt{2}r+O(r^5), \tag{35}$$

$$b(r)=2/B+O(r^4). \tag{36}$$

Proceeding as before, the elementary flatness condition is fulfilled.

Although the solution (29) does not give a Weyl singularity the choice $\alpha=C\sin Br$, where C and B are constants satisfying $CB<1$ does give a Weyl singularity with the same density (28).[3]

3. A WEYL SINGULARITY

Let us assume for the form of ρ, an oscillating matter density which under suitable conditions (to be specified later), never vanishes or becomes negative. Set

$$\rho=k-4q\cos 2r, \tag{37}$$

where k and q are constants. Then from (6), $\alpha(r)$ must satisfy

$$\alpha''+\tfrac{1}{2}(k+4q\cos 2r)\alpha=\alpha''+(a-2q\cos 2r)\alpha=0. \tag{38}$$

Here, $a=k/2$ is referred to as the "characteristic number" and can be calculated once q is set. We recognize (38) as Mathieu's equation.[4] This equation has both even and odd solutions. Because of Condition 2, we will pick the family of odd solutions, and in particular, $se_1(r,q)$, a first-order sine-elliptic function. Then,

$$\alpha(r)=se_1(r,q)=\sum_{s=0}^{\infty}B_{2s+1}(q)\sin(2s+1)r, \tag{39}$$

where:

$$\begin{aligned}
&B_1=1,\\
&B_3=-\tfrac{1}{8}q+\tfrac{1}{64}q^2-\tfrac{1}{1536}q^3-\tfrac{11}{36846}q^4+\cdots,\\
&B_5=\tfrac{1}{192}q^2-\tfrac{4}{4608}q^3-\tfrac{1}{24576}q^4+\cdots,\\
&B_7=-\tfrac{1}{9216}q^3+\tfrac{1}{49152}q^4+\cdots,\\
&B_9=\tfrac{1}{737280}q^4+\cdots,
\end{aligned} \tag{40}$$

up to orders of q^4. Now,

$$e^{2\lambda}=\sum_{s=0}^{\infty}B_{2s+1}(q)\sin(2s+1)r\sum_{t=0}^{\infty}B_{2t+1}(q)(2t+1)^2\sin(2t+1)r \times 2^{-1}\Big(1-\sum_{u=0}^{\infty}B_{2u+1}(q)(2u+1)\cos(2u+1)r\Big)^{-1}. \tag{41}$$

To go any further in our study of this solution, we need to choose a value for q. Let $q=0.1$. Then,

$$\begin{aligned}
a&=1-q-\tfrac{1}{8}q^2+\tfrac{1}{64}q^3-\tfrac{1}{1536}q^4-\cdots\\
&\cong 0.8988.
\end{aligned} \tag{42}$$

Substituting this q value into (40) and using these results to investigate (41), we find that, for all r,

$$e^{2\lambda}>0, \tag{43}$$

and

$$\lim_{r\to 0}e^{2\lambda}=0. \tag{44}$$

From (4) and (5), we see that this solution possesses a Weyl singularity. That this singular behavior of ω and σ, and therefore the components of the Weyl tensor,[2] has a logarithmic structure, can be seen by noting that

$$e^{2\lambda}=k_1r^2+k_2r^4+O(r^6), \tag{45}$$

or

$$\lambda = \tfrac{1}{2}\ln(k_1) + \tfrac{1}{2}\ln(r^2), \tag{46}$$

where k_i are constants, and terms of order r^4 and higher have been dropped. Then

$$\omega = k_3 + k_4[\ln(r^2)]^2 + \cdots, \tag{47}$$

and

$$\sigma = k_5 \ln(r^2) + \cdots. \tag{48}$$

For the value $q = 0.1$, the matter density never approaches zero at any multiple of π. However, if $q = 0.32$, then $a \cong 1.3354$ and the dust is clumped into periodic separate groupings. Therefore, to prevent the occurrence of a negative density,

$$0 < q \lesssim 0.33. \tag{49}$$

The first five conditions in Sec. 1 are easily verified. However, because of the logarithmic singularities in ω and σ, it is difficult to investigate the metric functions $b(r)$ and $e(r)$ and to draw any conclusions concerning the elementary flatness condition.

4. CONCLUSION

The underdetermined nature of the gravitational field equations for a cylindrical stationary dust model allows us the freedom to generate a cosmology having either singular or nonsingular behavior. The choice for the unknown function $\alpha(r)$ must satisfy six conditions stemming from the symmetry and the field equations. Using this choice to form the quantity $e^{2\lambda}$, the nature of the solution is determined from

$$\lim_{r \to 0} e^{2\lambda}. \tag{50}$$

For both $\alpha(r) = 2J_1(r)$ and $\alpha(r) = B^{-1}\sin Br$, we find that this limit is finite. This implies a nonsingular solution (though pathological in the latter case). For $\alpha(r) = C\sin Br$ where $0 < BC < 1$, a Weyl singularity results.[3] If $\alpha(r) = se_1(r, q)$, the density oscillates and the limit of (50) vanishes, producing an example of a Weyl singularity. Since the Weyl components diverge as a logarithmic singularity, further investigation of this solution appears difficult due to the singular nature of the equations determining the metric.

ACKNOWLEDGMENT

The author would like to express his sincere thanks to Dr. John R. Ray for his encouragement and suggestions during the course of the present work.

[1] G. F. R. Ellis, J. Math. Phys. 8, 1171 (1967).
[2] A. R. King, Comm. Math. Phys. 38, 157 (1974).
[3] A. R. King (private communication).
[4] N. W. McLachlan, *Theory and Application of Mathieu Functions* (Dover, New York, 1964).

Invariance transformations, invariance group transformations, and invariance groups of the sine-Gordon equations*

Sukeyuki Kumei

Department of Physics, University of the Pacific, Stockton, California 95204
(Received 5 May 1975)

We investigate a structure of continuous invariance transformations connected to the identity transformation. The transformations considered do not necessarily form a group. We clarify the relationship between the infinitesimal invariance transformation and the finite invariance transformation by showing explicitly how the infinitesimal transformations are woven into the finite one. The analysis leads to a new method of finding generators of the invariance group transformation. The results are useful in the study of symmetry properties, or group theoretic structure, of differential equations. We use the results in studying the group properties of the sine-Gordon equation $u_{xt} = \sin u$, and indicate that the equation is invariant under an infinite number of one-parameter groups; the groups obtained are of a more general type than that dealt with by Lie. These findings are used to prove the group theoretic origin of the well-known conservation laws associated with the sine-Gordon equation.

INTRODUCTION

The discovery of the puzzling behavior of nonlinear wave "solitons" in various fields of applied science has triggered extensive study of nonlinear dispersive waves.[1] One of the basic properties of the system which admits a soliton appears to be the possession of an infinite number of conserved quantities. As has been shown by Lax,[2] the existence of such conserved quantities is closely related to the soliton behavior of the waves. In spite of their importance in elucidating the nature of nonlinear waves, it seems that no one as yet has obtained a clear understanding of the origin of such conserved quantities.[3]

It is well known that both in classical and quantum mechanics the conservation law reflects the existence of symmetry in the system. In classical mechanics, Noether's theorem associates one conserved quantity with each invariance group of the action integral. In quantum mechanics, we can associate one conserved quantity Q, which satisfies the equation $[Q, H] + i\partial Q/\partial t = 0$, with each invariance group of the time-dependent Schröedinger equation.[4] From these experiences, it is natural to wonder whether there exists an invariance group associated with each conservation law of nonlinear waves.

In the present and in future communications, we will investigate this question by applying Lie's infinitesimal analysis[5] and its generalization[5d] to the differential equations governing the waves. In this paper, we approach the question by studying the group theoretic aspect of continuous invariance transformations, which has been proved useful in systematically deriving a series of conservation laws.

In Sec. II, we analyze continuous invariance transformations (not necessarily a group transformation) connected to the identity transformation, to clarify the relationship between local and global invariance transformations. The results will be used in Sec. III to elucidate the group theoretic aspect of continuous invariance transformations of differential equations. In Sec. IV, we apply the generalization of Lie's theory to find some invariance groups of the sine-Gordon equation, $u_{xt} = \sin u$. In Sec. V, by using the result of Sec. III, we develop a new method of finding generators of an invariance group of differential equations. The method will be used, with the aid of the Bäcklund transformation, to show that the sine-Gordon equation is invariant under an infinite number of one-parameter groups. In Sec. VI, we investigate a relation between these groups and a series of conservation laws of the sine-Gordon equation.

I. PRELIMINARY

We consider a partial differential equation of the form

$$F(z^i, u, u_j, u_{kl}, \cdots) = 0, \tag{1}$$

where $z^i = (z^1, z^2, \ldots, z^n)$, $u_j = (\partial_1 u, \ldots, \partial_n u)$, etc. Let's suppose that there exists a solution $u = u(z^i, \alpha)$, which depends on a parameter α continuously. Assuming that it is analytic near $\alpha = 0$, we expand the solution in a Taylor series in α,

$$u = \sum_{k=0}^{\infty} \frac{\alpha^k}{k!} u^k(z^i), \quad u^k = \{(\partial_\alpha)^k u\}_{\alpha=0}. \tag{2}$$

Putting this solution into equation (1), we obtain a sequence of partial differential equations which will successively determine a possible form of the u^k. In particular, the first term u^0 must be a solution of equation (1). If the equation is linear, all the u^k's must also satisfy the same equation. In the case of nonlinear differential equations, however, all the equations are different. First, the differential equation for u^1 becomes homogeneous linear and involves the first solution u^0; we then obtain a nonhomogeneous linear equation for the u^k, $k > 1$, which has the same homogeneous part as the u^1; the nonhomogeneous term depends upon the u^0, $u^1, \ldots, u^{k-1}$ and their derivatives. By a deductive argument, we expect that if only the nonhomogeneous solution is taken for $u^2, u^3 \ldots, u^{k-1}$, the nonhomogeneous solution for u^k will have a strong functional dependence on the u^0. We consider the sine—Gordon equation $u_{xt} - \sin u = 0$ as an example. The equation for u^0, u^1, and u^2 is found to be

$$u^0_{xt} - \sin u^0 = 0, \quad u^1_{xt} - u^1 \cos u^0 = 0,$$
$$u^2_{xt} - u^2 \cos u^0 = -(u^1)^2 \sin u^0.$$

It is surprising that we can find many solutions for u^1 and u^2 which can be expressed as simple functions of u^0 and its derivatives; a few examples are

$$u^1 = u^0_x \quad \text{and} \quad u^2 = u^0_{xx},$$

or (3a–d)

$$u^1 = u^0_{xxx} + \tfrac{1}{2}(u^0_x)^3 \quad \text{and}$$
$$u^2 = u^0_{xxxxx} + 3(u^0_x)^2 u^0_{xxx} + \tfrac{9}{4}(u^0_x)^4 u^0_{xx} + 9u^0_x u^0_{xx} u^0_{xxx} + 3(u^0_{xx})^3.$$

The existence of such solutions is directly connected to the origin of the infinite number of conservation laws, and the study of the origin of such solutions will provide a key in understanding the origin of the conservation laws. We ask how a nonhomogeneous solution for u^k will depend on the u^0 if we take only the nonhomogeneous solutions for $u^2, \ldots, u^{k-1}$; this problem requires a careful analysis of invariance transformations.

II. RELATION BETWEEN AN INFINITESIMAL AND A FINITE INVARIANCE TRANSFORMATION

We have considered an example in which one solution u is continuously connected to another solution u^0 through a parameter α. This may be considered as a continuous transformation of u^0 to u; it is a special case of the continuous invariance transformation which is connected to the identity transformation.

We consider a set of transformations of the coordinates of the n-dimensional vector space $R^n(x^1, \ldots, x^n)$ which analytically depends on the parameter α, and becomes the identity transformation for $\alpha = 0$:

$$x^i \to \bar{x}^i = X^i(x^j, \alpha), \quad x^i = X^i(x^j, 0). \tag{4}$$

We also consier an equation $F(x^i) = F(x^1, \ldots, x^m) = 0$ which is defined in the subspace $R^m(x^1, \ldots, x^m)$ of R^n. The equation $F(x^i) = 0$ defines a manifold S, or hypersurface in R^m. We define the invariance transformation in the following way:

> Transformation (4) is a continuous invariance transformation of the equation $F(x^i) = 0$, if the condition $F(X^i(x^j, \alpha)) = 0$ is satisfied for the continuous values of α on the manifold S defined by $F(x^i) = 0$.

Geometrically, this implies that an invariance transformation carries a point on S into another point on S. We first investigate this invariance condition in detail, and will come back to the invariance transformation of the differential equation in the next section.

Under the condition we have imposed on the transformation, we can expand $X^i(x^j, \alpha)$ in a Taylor series in α by:

$$\bar{x}^i = X^i(x^j, \alpha) = x^i + \sum_{k=1}^{\infty} \frac{\alpha^k}{k!} \xi^i_k, \quad \xi^i_k = \{(\partial_\alpha)^k X^i(x^j, \alpha)\}_{\alpha=0}. \tag{5}$$

Defining the differential operator U_k by

$$U_k = \sum_{j=1}^{n} \xi^j_k \partial_j, \tag{6}$$

we can write (5) as

$$\bar{x}^i = \left(1 + \sum_{k=1}^{\infty} \frac{\alpha^k}{k!} U_k\right) x^i. \tag{7}$$

We analyze the effect of this transformation on an analytic function $f(x^i)$ defined in R^n. Expanding $f(\bar{x}^i)$ in a Taylor series in α, we obtain

$$f(\bar{x}^i) = \sum_{k=0}^{\infty} \frac{\alpha^k}{k!} \phi_k(x^i), \quad \phi_k = \{(\partial_\alpha)^k f(\bar{x}^i)\}_{\alpha=0} = A_k f(x^i),$$

where

$$A_0 = 1 \quad \text{and}$$
$$A_k = k! \sum \left\{ \left[\prod_{j=1}^{s} \left([(p_j!)^{q_j} (q_j!)]^{-1} \prod_{l=1}^{q_j} \xi^{r^j_l}_{p_j} \right) \right] \times \prod_{j=1}^{s} \prod_{l=1}^{q_j} \partial_{r^j_l} \right\}, \tag{9}$$

where p_j and q_j are the integers satisfying the conditions

$$\sum_{j=1}^{s} p_j q_j = k, \quad 1 \le p_i < p_j \le k \quad \text{for} \quad i < j, \quad 1 \le q_j.$$

Here, we apply the summation convention with respect to the indices r^j_l. The choice of the sets $(p_1, \ldots, p_s)$ and $(q_1, \ldots, q_s)$ satisfying the conditions is not unique, and the sum in (9) is to be taken with respect to each of such sets. Using the differential operator A_k, we can write the effect of the coordinate transformation on the function $f(x^i)$ as

$$f(\bar{x}^i) = \left(1 + \sum_{k=1}^{\infty} \frac{\alpha^k}{k!} A_k\right) f(x^i) = T(\alpha) f(x^i). \tag{10}$$

We note that $T(\alpha) x^i = x^i + \alpha \xi^i_1 + \frac{1}{2}\alpha^2 \xi^i_2 + \cdots = \bar{x}^i$ recovers the definition we started with.

Now we suppose that the continuous transformation $T(\alpha)$ leaves the equation $F(x^i) = 0$ defined in R^m invariant in the sense defined above. Then, the following statement will be obvious:

> The transformation $T(\alpha)$ is a continuous invariance transformation of the equation $F(x^i) = 0$, if and only if $A_k F(x^i) = 0$ on the manifold S defined by $F(x^i) = 0$. (A)

Although this provides the condition for a transformation to be an invariance transformation, it is very difficult to get any clear view of the structure of the transformation unless a considerable simplification of expression (9) is made; it is crucial to oberve that we can re-express (9) as

$$A_k = k! \sum \frac{(\bar{U}_1)^{q_1}}{(p_1!)^{q_1}(q_1!)} \cdot \cdots \cdot \frac{(\bar{U}_s)^{q_s}}{(p_s!)^{q_s}(q_s!)}, \tag{11}$$

where we take the same rule of summation as for (9). The remarkable feature of this expression is the fact that all the $\bar{U}_k$'s are first-order differential operators. We write down the first four generators in this form:

$$A_1 = \bar{U}_1, \quad A_2 = (\bar{U}_1)^2 + \bar{U}_2, \quad A_3 = (\bar{U}_1)^3 + 3\bar{U}_1\bar{U}_2 + \bar{U}_3,$$
$$A_4 = (\bar{U}_1)^4 + 6(\bar{U}_1)^2\bar{U}_2 + 3(\bar{U}_2)^2 + 4\bar{U}_1\bar{U}_3 + \bar{U}_4, \tag{12a–12d}$$

where

$$\bar{U}_1 = U_1 = \xi^i_1 \partial_i, \quad \bar{U}_2 = U_2 - \xi^i_1 \xi^j_{1,i} \partial_j$$
$$\bar{U}_3 = U_3 + (-3\xi^i_1 \xi^j_{2,i} + 2\xi^i_1 \xi^k_{1,i} \xi^j_{1,k} + 2\xi^i_1 \xi^k_1 \xi^j_{1,ik}) \partial_j,$$
$$\bar{U}_4 = U_4 + (-4\xi^i_1 \xi^j_{3,i} - 3\xi^i_2 \xi^j_{2,i} + 9\xi^i_1 \xi^k_{1,i} \xi^j_{2,k} + 6\xi^i_1 \xi^k_1 \xi^j_{2,ik}$$

$$+3\xi_2^i\xi_{1,i}^k\xi_{1,k}^j+3\xi_2^i\xi_1^k\xi_{1,ik}^j-6\xi_1^i\xi_{1,i}^j\xi_{1,k}^k$$
$$-3\xi_1^i\xi_1^l\xi_{1,il}^k\xi_{1,k}^j$$
$$-12\xi_1^i\xi_1^l\xi_{1,i}^k\xi_{1,ik}^j-3\xi_1^i\xi_1^j\xi_1^k\xi_{1,ikl}^j)\partial_j. \quad \text{(13a)—(13d)}$$

The importance of the decomposition into this form will be recognized if we remember the basic lemma used in the theory of continuous group transformations:

If two first-order differential operators $U_a=\xi_a^i\partial_i$ and $U_b=\xi_b^i\partial_i$, $i=1,2,\ldots,n$, satisfy the conditions $U_jf(x^i)=0$, $j=a,b$, on the manifold defined by $f(x^i)=0$, then we have $U_aU_bf(x^i)=0$ on the same manifold.

Successive applications of the lemma to the invariance condition (A), lead to the conclusion that all the operators $\overline{U}_k$ in the expression (11) must satisfy the condition:

$$\overline{U}_kF(x^i)=0 \text{ on the manifold } S \text{ defined by } F(x^i)=0. \quad \text{(B)}$$

This allows us to draw the following conclusion:

All the A_k's are constructed from first-order differential operators Q, which satisfy the condition $QF(x^i)=0$ on $F(x^i)=0$. (C)

In Lie's theory of group transformations, the operator which satisfies condition (B) is called a generator of the invariance group. We suppose that the largest invariance group of the equation $F(x^i)=0$ is an r-parameter group with generators Q_i. Then, all the operators which satisfy condition (B) can be written as

$$\overline{U}_k=\sum_{i=1}^{r}a_k^iQ_i\,. \quad (14)$$

In particular, if we let $a_k^i=0$ for $k\geq 2$, we obtain $A_k=(\overline{U}_1)^k$ from (11), and the operator $T(\alpha)$ in (10) reduces to an exponential operator,

$$T(\alpha)=\sum_{k=0}^{\infty}\frac{\alpha^k}{k!}(\overline{U}_1)^k=e^{\alpha\overline{U}_1}. \quad (15)$$

Result (C) is significant in studying the structure of invariance transformations because it clarifies the constraints on and arbitrariness of an invariance transformation. The vital fact is that if we have a complete set of generators of the invariance group of the equation $F(x^i)=0$, then any continuous invariance transformation connected to the identity transformation can be constructed from these generators.

Now, the problem is how to find such generators for a given equation $F(x^i)=0$. The basic idea of deriving the generators was established by Lie, and we will illustrate it briefly after the discussion of differential equations.

III. INVARIANCE TRANSFORMATIONS OF DIFFERENTIAL EQUATIONS

We have considered a set of coordinate transformations in R^n which leave the equation $F(x^i)=0$ defined in R^m invariant. We now introduce some functional relations among the coordinates, which are compatible with the equation $F(x^i)=0$; such relations will restrict further the domain of manifold S.

We consider a function $u(x^i)$ defined in the $(k-1)$-dimensional space $R^{k-1}(x^1,\ldots,x^{k-1})$ with $k<m$, and assume the following:

The coordinate x^k is determined by the relation $x^k=u(x^i)$, and the coordinates $x^{k+1},\ldots,x^n$ are determined as the derivatives of $u(x^i)$ with respect to the coordinates $x^1,\ldots,x^{k-1}$. For instance, $x^{k+1}=\partial_1u$, $\ldots$, $x^{2k-1}=\partial_{k-1}u$, $x^{2k}=\partial_1\partial_1u$, $x^{2k+1}=\partial_1\partial_2u$, $\cdots$. (16)

We suppose that R^n is chosen in such a way that if it contains a coordinate corresponding to a jth derivative, then coordinates for all the other jth derivatives also appear in R^n. The condition we have imposed are compatible with the equation $F(x^i)=0$ only if the function $u(x^i)$ is a solution of the equation $F(x^i)=0$, interpreted as a partial differential equation by considering x^i's as derivatives defined by (16). Each solution will define a submanifold $\overline{S}$ of the manifold S, called the solution surface.

Now, we consider a continuous coordinate transformation (4) under which a manifold satisfying condition (16) is always mapped onto a manifold which also satisfies condition (16), with $x^k=u(x^i,\alpha)$. In analyzing such transformations, it is convenient to introduce the following definitions:

Basic coordinates and jth order coordinates

We call the coordinates $x^1,\ldots,x^k$ basic coordinates; and the coordinates corresponding to the jth order derivatives, jth order coordinates. For instance, in (16), $x^{k+1},\ldots,x^{2k-1}$ are first-order, and x^{2k}, x^{2k+1} are second-order.

Basic space and jth extended space

We call the vector space $(x^1,\ldots,x^i)$, jth extended space if it consists of all and only the basic coordinates and a complete set of the first through the jth order coordinates. In particular, we call the 0th extended space $(x^1,\ldots,x^k)$, the basic space.

Basic transformation

We call the transformation of a set of basic coordinates, the basic transformation.

Basic operator and jth extended operator

We call the operator $\overset{j}{Q}=\sum_{i=1}^{l}\xi^i\partial_i$ of the transformation in the jth extended space the jth extended operator. The 0th extended operator $\overset{0}{Q}=\sum_{i=1}^{k}\xi^i\partial_i$, will be called the basic operator.

It is clear that under condition (16), the transformation of the basic coordinates will determine the transformation of the rest of the coordinates. In particular, if a basic operator is given, we can determine all the extended operators. Now, we require that such transformation leaves the equation $F(x^i)=0$ invariant. The geometrical meaning of the invariance transformation is more important; the invariance transformation maps one solution surface $\overline{S}$ to another solution surface $\overline{S}'$ (or onto itself), both of which are on S. A discovery of such a transformation will lead to a new solution of the differential equation. The transformation studied most extensively is the group transformation. Lie considered an invariance group transformation of the form

$$x^i \to \bar{x}^i = e^{\alpha Q} x^i = x^i + \alpha \xi^i(x) + \cdots, \quad i = 1, \ldots, k,$$
$$x = (x^1, \ldots, x^k), \tag{17}$$

in which infinitesimal terms of the basic transformation depend only on the basic coordinates. It is important to note that, under such assumptions, a finite transformation of the coordinate x^p does not involve any coordinate whose order is higher than the order of x^p. This guarantees that the jth extended space is closed under the transformation. The existence of such a closed space enables us to elegantly construct a finite group transformation, via the method of characteristics, from a generator of the group.

Anderson, Kumei, and Wulfman, however, found that there exist invariance groups of time-dependent Schrödinger equations which are not of Lie's form.[4] They generalized Lie's theory by allowing infinitesimal terms ξ^i of the basic coordinates to depend on the coordinates of higher order:

$$x^i \to \bar{x}^i = e^{\alpha Q} x^i = x^i + \alpha \xi^i(x) + \cdots, \quad i = 1, \ldots, k,$$
$$x = (x^1, \ldots, x^l), \quad l \geq k. \tag{18}$$

Here, the order of the coordinates in ξ^i is not restricted, and coordinates of any order may appear.[6] We note, however, that we no longer have any closed finite-dimensional space under such a group transformation.[7] Although this does not cause any problem in finding generators of invariance transformations, we can no longer apply the method of characteristics in finding a finite transformation. This generalization, however, is absolutely necessary to uncover all the invariance groups inherent in the differential equations.[8] Before we show that the sine–Gordon equation admits such invariance groups, we answer the question raised at the end of Sec. I. To put the problem into our present language, we rewrite (1) as

$$F(x^i) = 0 \text{ with } x^i = z^i, \quad i = 1, \ldots, n,$$
$$x^{n+1} = u, \quad x^{n+2} = u_1, \cdots, \tag{1'}$$

and (2) as

$$\bar{x}^i = x^i, \quad i = 1, \ldots, n,$$
$$\bar{x}^{n+1} = X^{n+1}(x^i, \alpha) = x^{n+1} + \alpha \xi_1^{n+1} + \tfrac{1}{2}\alpha^2 \xi_2^{n+1} + \cdots. \tag{2'}$$

From a transformational viewpoint, the statement that $\bar{u}$ is a solution of equation (1) is equivalent to saying that transformation (2′) leaves equation (1′) invariant. For such a transformation, as we have found, we can write $\xi_k^{n+1} = A_k x^{n+1}$ (or $u^k = A_k u^0$ in the old notation). This leads to the conclusion,

> If a differential equation $F(z^i, u, u_j, u_{k1}, \cdots) = 0$ admits a solution $u(z^i, \alpha) = \sum_{k=0}^{\infty} (\alpha^k/k!) u^k(z^i)$ which depends analytically on α near $\alpha = 0$, then u^k is always written as $u^k(z^i) = A_k u^0(z^i)$, where u^0 is a solution of the same equation and the operator A_k is constructed by (11) from the generators Q of invariance group transformations of $F = 0$ by which the independent variables z^i are unchanged. In particular, $u^1 = Qu^0$. Furthermore, if only the inhomogeneous solution is taken for every u^k, $k > 1$, then $u^k = (Q)^k u^0$, and a resulting solution is expressed as $u(z^i) = e^{\alpha Q} u^0(z^i)$. (D)

IV. SOME INVARIANCE GROUP TRANSFORMATIONS OF SINE-GORDON EQUATION

We now go back to the analysis of the solutions (3a)–(3d). According to result (D), these solutions clearly indicate the existence of invariance groups, or symmetries, of the sine–Gordon equation. We will systematically determine generators of the invariance group transformations and will reveal new symmetries of the equation.

We first specialize the general formulation given above to a case in which we have three basic coordinates x^1, x^2, and x^3, and $F(x^i)$ is chosen as

$$F(x^1, x^2, x^3, x^4, x^5, x^6, x^7, x^8) = x^7 - \sin x^3.$$

As stated in (16), we establish the following constraints on the coordinates:

$$x^3 = u(x^1, x^2), \quad x^4 = u_1, \quad x^5 = u_2, \quad x^6 = u_{11},$$
$$x^7 = u_{12}, \quad x^8 = u_{22} \quad x^9 = u_{111}, \quad x^{10} = u_{112},$$
$$x^{11} = u_{122}, \quad x^{12} = u_{222}, \quad x^{13} = u_{1111}, \quad x^{14} = u_{1112},$$
$$x^{15} = u_{1122}, \quad x^{16} = u_{1222}, \quad x^{17} = u_{2222}, \quad x^{18} = u_{11111},$$
$$x^{19} = u_{11112}, \quad x^{20} = u_{11122}, \quad x^{21} = u_{11222}, \quad x^{22} = u_{12222},$$
$$x^{23} = u_{22222}, \tag{19}$$

where subscripts 1 and 2 indicate the derivatives with respect to x^1 and x^2. We now consider a transformation, of the generalized form (18), in which x^1 and x^2 are unchanged and the infinitesimal transformation of x^3 depends on x^3 and the first through the third-order coordinates:

$$\bar{x}^1 = x^1, \quad \bar{x}^2 = x^2,$$
$$\bar{x}^3 = x^3 + \alpha \xi^3(x^3, x^4, x^5, x^6, x^8, x^9, x^{12}). \tag{20}$$

We note that the inclusion of the coordinates x^7, x^{10}, and x^{11} is redundant because we can replace them by the coordinates in ξ^3 after we have introduced the condition $F = 0$. The infinitesimal transformation, induced by (20), of the coordinate corresponding to the derivative $(\partial_1)^m(\partial_2)^n u$ is calculated as[9]

$$\bar{x}^i = x^i + \alpha(\partial_1)^m(\partial_2)^n \xi^3 = x^i + \alpha \xi^i. \tag{21}$$

Here, the partial derivative should be interpreted as

$$(\partial_1)^m(\partial_2)^n \xi^3 = (\partial_1)^m(\partial_2)^n \times \xi^3(u(x^1, x^2), u_1(x^1, x^2), \ldots, u_{222}(x^1, x^2)) \tag{22}$$

For instance

$$\xi^4 = \xi^3_3 x^4 + \xi^3_4 x^6 + \xi^3_5 x^7 + \xi^3_6 x^9 + \xi^3_8 x^{11} + \xi^3_9 x^{13} + \xi^3_{12} x^{16}, \tag{23}$$

where ξ^3_i is the derivative of ξ^3 with respect to the coordinate x^i contained in ξ^3, and should not be confused with the same notation used in Sec. II. As in Sec. II, we write the infinitesimal transformation in the jth extended space as

$$\bar{x}^i = (1 + \alpha \overset{j}{Q}) x^i, \quad \overset{j}{Q} = \sum_{i=1}^{l} \xi^i \partial_i \quad \text{with} \quad \xi^1 = \xi^2 = 0. \tag{24}$$

Now, we assume that the equation $F = 0$ is invariant under the group transformation whose infinitesimal form is given by (20); the condition is

$$\overset{2}{Q} F = \xi^7 - \xi^3 \cos x^3 = 0 \text{ on } x^7 - \sin x^3 = 0, \tag{25}$$

which is the partial differential equation for ξ^3. The equation $\overset{2}{Q}F=0$ will split into a set of partial differential equation because some of the coordinates which appear in the equation are independent from the coordinates in ξ^3 (Appendix). By solving these equations, we find four independent solutions:

$$\xi^3_a=x^4, \quad \xi^3_b=x^9+\tfrac{1}{2}(x^4)^3, \tag{26a}$$

$$\xi^3_c=x^5, \quad \xi^3_d=x^{12}+\tfrac{1}{2}(x^5)^3. \tag{26b}$$

Obviously, the last two solutions can be obtained from the first two, by interchanging the roles of x^1 and x^2. The second extended operators associated with ξ^3_a and ξ^3_b are calculated from (21) and (24) as

$$\overset{2}{Q}_a=x^4\partial_3+x^6\partial_4+x^7\partial_5+x^9\partial_6+x^{10}\partial_7+x^{11}\partial_8, \tag{27a}$$

$$\begin{aligned}\overset{2}{Q}_b=&\{x^9+\tfrac{1}{2}(x^4)^3\}\partial_3+\{x^{13}+\tfrac{3}{2}(x^4)^2x^6\}\partial_4+\{x^{14}\\&+\tfrac{3}{2}(x^4)^2x^7\}\partial_5+\{x^{18}+3x^4(x^6)^2+\tfrac{3}{2}(x^4)^2x^9\}\partial_6\\&+\{x^{19}+3x^4x^6x^7+\tfrac{3}{2}(x^4)^2x^{10}\}\partial_7+\{x^{20}+3x^4(x^7)^2\\&+\tfrac{3}{2}(x^4)^2x^{11}\}\partial_8.\end{aligned} \tag{27b}$$

We can easily check that they satisfy condition (25); thus, the sine-Gordon equation is invariant under the group transformations generated by these operators.

This result explains the origin of the solutions (3a)—(3d); they are obtained from result (D):

$$u^1_a=\overset{0}{Q}_a x^3=x^4 \tag{28a}$$

and

$$u^2_a=(\overset{1}{Q}_a)^2x^3=x^6, \tag{28b}$$

$$u^1_b=\overset{0}{Q}_b x^3=x^9+\tfrac{1}{2}(x^4)^2 \tag{28c}$$

and

$$\begin{aligned}u^2_b=(\overset{3}{Q}_b)^2x^3=&x^{24}+3(x^4)^2x^{13}+\tfrac{9}{4}(x^4)^4x^6\\&+9x^4x^6x^9+3(x^6)^3.\end{aligned} \tag{28d}$$

We note that we need the extended operators $\overset{1}{Q}_a$ and $\overset{3}{Q}_b$ for calculating the second term u^2. As we stated in Sec. III, this is the general character of the generalized transformation (18) and we need the $k(n-1)$th extended operator to calculate the nth term u^n if the basic transformation contains the k-th order coordinate.

V. GENERATING FUNCTION FOR GENERATORS

We have obtained four generators of the invariance group of the sine-Gordon equation by considering a generalized Lie transformation (18). However, if we had assumed a more general form for ξ^3, we might have been able to produce more generators. It is unfortunate that we have no theory which tells us which coordinates we need in ξ^i of (18) to obtain a complete set of generators, hence we must make some assumptions on the form of ξ^i. In practice, it is not possible to retain too many coordinates in ξ^i because the determining equations for ξ^i become too huge to solve. Therefore, it is highly desirable to have another method for producing the generators, which does not require either such assumptions or the construction of the solutions of determining equations. Here, we provide one such method although the completeness of the set of generators obtained is still not assured.

The idea of the method is to reverse the result in Sec. II. We found that the operator Q which satisfies a condition $QF=0$ on $F=0$ is the building block of any invariance transformation connected to the identity transformation. By reversing this, we argue that if we have an invariance transformation connected to the identity transformation, then we can find at least one such operator. More precisely, we proceed in the following way.

Suppose we have an invariance transformation of the equation $F(x^i)=0$,

$$\bar{x}^i=x^i+\sum_{k=1}^{\infty}\frac{\alpha^k}{k!}A_kx^i=x^i+\sum_{k=1}^{\infty}\frac{\alpha^k}{k!}\xi^i_k, \quad i=1,\ldots,n, \tag{29}$$

in which all the ξ^i_k are known. Using the result (11), we can write,

$$\xi^i_1=A_1x^i=\bar{U}_1x^i, \quad \xi^i_2=A_2x^i=\{(\bar{U}_1)^2+\bar{U}_2\}x^i,$$

$$\xi^i_3=A_3x^i=\{(\bar{U}_1)^3+3\bar{U}_1\bar{U}_2+\bar{U}_3\}x^i,\cdots,$$

where all the $\bar{U}_k$ are first-order differential operators. From the first equation, we obtain

$$\bar{U}_1=\sum_{i=1}^{n}\xi^i_1\partial_i \ . \tag{30}$$

Feeding this into the second, we get $\bar{U}_2x^i=\xi^i_2-(\bar{U}_1)^2x^i$, which provides

$$\bar{U}_2=\sum_{i=1}^{n}[\xi^i_2-\{(\bar{U}_1)^2x^i\}]\partial_i \ . \tag{31}$$

Next we substitute these for the $\bar{U}_1$ and $\bar{U}_2$ in the third equation to determine $\bar{U}_3$. Continuing this process we can obtain a series of operators $\bar{U}_k$, all of which satisfy the invariance condition "$QF=0$ on $F=0$." We note that if the starting transformation (29) happens to form a group, then we only get $\bar{U}_1$ and all the others are equivalently zero for the reason discussed in Sec. II. We may consider the starting transformation (29) as a generating function for generators of an invariance group. The upshot of the method is that only algebraic computations are involved in the process and a computer can be used, whereas the construction of the solutions of the determining equations by computer is very difficult. Obviously, this method can be used to find generators of an invariance group of a differential equation if the constraints (16) are taken into account. We apply the method to the sine-Gordon equation to find additional generators.

We start with the well-known Bäcklund transformation of the sine-Gordon equation,[1]

$$\bar{x}^4-x^4=2\alpha\sin\tfrac{1}{2}(\bar{x}^3+x^3), \quad \alpha(\bar{x}^5+x^5)=2\sin\tfrac{1}{2}(\bar{x}^3-x^3), \tag{32}$$

with the convention established in (19). This transformation guarantees that if x^3 is a solution of the sine-Gordon equation then so is $\bar{x}^3$ for a continuous value of α. A principal use of the Bäcklund transformation is to construct a new solution $\bar{x}^3$ from a known solution x^3 by solving a set of first-order differential equations (32). We assume that the new solution $\bar{x}^3$, is an analytic

function of α in the neighborhood of $\alpha=0$, and so are its derivatives. Then, it is clear from (32) that the transformation is connected to the identity transformation; $\bar{x}^3 \to x^3$ as $\alpha \to 0$. The analyticity assumption allows us to expand the solution $\bar{x}^3$ in the Taylor series in α near $\alpha=0$. Such an expansion is found in the paper by Scott *et al.*,[1] and we rewrite their result:

$$\bar{x}^3 = x^3 + \sum_{k=1}^{\infty} \frac{\alpha^k}{k!} \xi_k^3, \tag{33}$$

with:

$$\begin{aligned}
&\xi_1^3 = 2x^4, \quad \xi_2^3 = 4x^6, \quad \xi_3^3 = 12x^9 + 2(x^4)^3, \\
&\xi_4^3 = 48x^{13} + 48(x^4)^2 x^6, \\
&\xi_5^3 = 240x^{18} + 360(x^4)^2 x^9 + 600x^4(x^6)^2 + 18(x^4)^5, \\
&\xi_6^3 = 1440x^{24} + 2880(x^4)^2 x^{13} \\
&\qquad + 12\,960x^4x^6x^9 + 3840(x^6)^3 + 1440(x^4)^4 x^6 \\
&\xi_7^3 = 10\,080x^{31} + 176\,400(x^6)^2 x^9 + 95\,760x^4(x^9)^2 \\
&\qquad + 141\,120x^4x^6x^{13} + 25\,200(x^4)^2 x^{18} \qquad \text{(34a–g)} \\
&\qquad + 63\,000(x^4)^3(x^6)^2 + 18\,900(x^4)^4 x^9 + 450(x^4)^7, \cdots,
\end{aligned}$$

where we have adopted the convention (19) and $x^{31} = u_{1111111}$. In this specific Bäcklund transformation, the coordinates x^1 and x^2 are unchanged, i.e.,

$$\bar{x}^1 = x^1, \quad \bar{x}^2 = x^2 \quad \text{or} \quad \xi_1^i = \xi_2^i = 0 \text{ for } i \geq 1. \tag{35}$$

The transformations (33) and (35) form the basic transformations, and they provide all the necessary information to follow the above prescription to find $\bar{U}_k$. We list the results up to $\bar{U}_7$:

$$\begin{aligned}
&\bar{U}_2 = \bar{U}_4 = \bar{U}_6 = 0, \quad \bar{U}_1 = 2x^4\partial_3, \quad \bar{U}_3 = \{4x^9 + 2(x^4)^3\}\partial_3, \\
&\bar{U}_5 = \{48x^{18} + 120(x^4)^2 x^9 + 120x^4(x^6)^2 + 18(x^4)^5\}\partial_3, \\
&\bar{U}_7 = \{1440x^{31} + 25\,200(x^6)^2 x^9 + 15\,120x^4(x^9)^2 \\
&\qquad + 20\,160x^4x^6x^{13} + 5040(x^4)^2 x^{18} + 12\,600(x^4)^3(x^6)^2 \\
&\qquad + 6300(x^4)^4 x^9 + 450(x^4)^7\}\partial_3. \qquad \text{(36a–e)}
\end{aligned}$$

Here, we have given the operators in the basic form; the operators in the extended form can be obtained from (21) and (24). By continuing this process, we will be able to find an infinite number of operators which satisfy the invariance condition (25). We can associate one invariance group transformation of the sine-Gordon equation with each of these operators.

VI. A SERIES OF CONSERVATION LAWS AND INVARIANCE GROUPS

In this section, we use notation (16), hence x^k represents a solution of the differential equation $F(x^i)=0$.

We consider an equation $F(x^i)=0$ which can be put into a conservation form:

$$\sum_{i=1}^{k-1} \partial_i f^i = 0, \quad f^i = f^i(x^1, \ldots, x^{k-1}, x^k, \ldots, x^l), \tag{37}$$

where the derivatives are to be taken by considering $x^k, \ldots, x^l$ as functions of $x^1, \ldots, x^{k-1}$. The vector $\mathbf{f} = (f^1, f^2, \ldots, f^{k-1})$ establishes a divergent free flux in the space $R^{k-1}(x^1, \ldots, x^{k-1})$ for each solution of the equation. Now, we assume that the equation $F(x^i)=0$ is invariant under an r-parameter group with the property:

$$\bar{x}^i = T_g(\alpha)x^i = x^i \quad \text{for} \quad i = 1, \ldots, k-1 \quad \text{with}$$

$$T_g(\alpha) = \exp\left(\sum_{i=1}^{r} \alpha^i Q_i\right). \tag{38}$$

We suppose that transformation (38) exists if $|\alpha| < \delta$. Here, δ is a positive number. Under such assumptions, $\bar{x}^k$ represents a new solution of the equation and a corresponding flux, $\bar{\mathbf{f}} = (\bar{f}^1, \bar{f}^2, \ldots, \bar{f}^{2k-1})$ is written as

$$\begin{aligned}
\bar{f}^i &= f^i(x^1, \ldots, x^{k-1}, \bar{x}^k, \ldots, \bar{x}^l) \\
&= T_g(\alpha) f^i(x^1, \ldots, x^{k-1}, x^k, \ldots, x^l).
\end{aligned} \tag{39}$$

The implication of the new flux is the same as the old one, except that it is now for the new solution. However, its power series expansion in α tells us something new about the starting solution x^k; because we have assumed that the transformation $T_g(\alpha)$ exists at least for some range of $|\alpha|$, it acts as a generating function of fluxes; each term of the expansion of (39) in $\alpha^1, \ldots, \alpha^r$ also forms a divergent free flux. We state this as follows:

> If a differential equation $F(x^i)=0$ admits an invariance group with property (38), and if a flux $\mathbf{f}$ of the form (37) exists, then, for any polynomial (E) function $G(Q_1, \ldots, Q_r)$ of the generators of the group, the vector $G\mathbf{f}$ forms a divergent free flux.

Here, we see two basic patterns for a series of divergent free fluxes to arise: one associated with a series $Q_i\mathbf{f}$, $i=1, \ldots, r$ and one associated with a series $(Q_i)^n\mathbf{f}$, $n=1, 2, \cdots$. It will be reasonable to say in general, that the former is more fundamental than the latter because the series of the second type can be mechanically constructed if Q_i is known, although the reverse is not possible. One, however, should not think that the fluxes of the second type are trivial.[10]

We now apply this analysis to the sine-Gordon equation, $F = x^7 - \sin x^3$. The equation can be put into the conservation form by multiplying by x^5;

$$\partial_1 f^1 + \partial_2 f^2 = 0 \quad \text{with} \quad \mathbf{f} = (f^1, f^2) = (\tfrac{1}{2}(x^5)^2, \cos x^3),$$

and the generators (36b)—(36e) can be used to derive new fluxes. We list a few of them, (using the notation $\mathbf{f}_i = \frac{1}{i}\bar{U}_i\mathbf{f}$):

$$\begin{aligned}
&\mathbf{f}_1 : f_1^1 = 2x^5x^7, \quad f_1^2 = -2x^4 \sin x^3, \\
&\mathbf{f}_3 : f_3^1 = \{4x^{14} + 6(x^4)^2 x^7\}x^5, \quad f_3^2 = -\{4x^9 + 2(x^4)^3\}\sin x^3, \\
&\mathbf{f}_5 : f_5^1 = 24\{2x^{25} + 10x^4x^7x^9 + 5(x^4)^2x^{14} + 5x^7(x^6)^2 \\
&\qquad + 10x^4x^6x^{10} + \tfrac{15}{4}(x^4)^4x^7\}x^5, \\
&\qquad f_5^2 = -24\{2x^{10} + 5(x^4)^2x^9 + 5x^4(x^6)^2 + \tfrac{3}{4}(x^4)^5\}\sin x^3, \\
&\mathbf{f}_{3,3} : f_{3,3}^1 = (\bar{U}_3)^2 f^1 = 16\{x^5x^{32} + 3(x^4)^2x^5x^{19} + (x^{14})^2 \\
&\qquad + \{3(x^4)^2x^7 + 9x^4x^5x^6\}x^{14} + 6x^4x^5x^7x^{13} \\
&\qquad + \{9x^5(x^6)^2 + 9x^4x^5x^9 + \tfrac{9}{4}(x^4)^4x^5\}x^{10} \qquad \text{(40a–d)} \\
&\qquad + 9x^5x^6x^7x^9 + 9(x^4)^3x^5x^6x^7 + \tfrac{9}{4}(x^4)^4(x^7)^2\}.
\end{aligned}$$

Here, we have listed only the first component for $\mathbf{f}_{3,3}$. Among these fluxes, the first flux, $\mathbf{f}_1$ is trivial because it is the derivative of $\mathbf{f}$ with respect to x^1.[12] We analyze the known results from our viewpoint. Our

results are clearly different from the fluxes given in the paper by Scott *et al.*[11] Their results, however, can be obtained by taking a linear combination of fluxes with the form $(\bar{U}_1)^a(\bar{U}_2)^b\cdots(\bar{U}_r)^p\mathbf{f}$. In fact, by using (11) and (36a)–(36e), we find that $A_3\mathbf{f}$ and $A_5\mathbf{f}$ recover their results. For instance,

$$A_3 f^1 = \{(\bar{U}_1)^3 + 3\bar{U}_1\bar{U}_2 + \bar{U}_3\} f^1 = 6\{2x^5x^{14} + 4x^4x^{10} + (x^4)^2x^5x^7\}$$
$$= 6\{2u_2u_{1112} + 4u_{12}u_{112} + (u_1)^2u_2u_{12}\},$$

where we interpret $\bar{U}_i$ as a generator extended to a necessary order. Now, we ask which fluxes are most basic among these.

Although this question is very important in analyzing the nature of conservation laws in general, the answer depends on the measure one uses. However, as we have indicated above, the hierarchy becomes quite clear within the framework of group theory; we classify fluxes into two categories:

(1) Basic fluxes: $\mathbf{f}$, $Q_i\mathbf{f}$, $i=1,\ldots,r$,

and

(2) Associated fluxes: $(Q_{i_1})^{n_1}(Q_{i_2})^{n_2}\cdots(Q_{i_p})^{n_p}\mathbf{f}$

with $i_k=1,\ldots,r$, and $n_1+n_2+\cdots+n_p>1$,

and we use the basic fluxes to characterize the conservation law associated with a solution. The remarkable feature of the sine-Gordon equation is that it possesses a series of basic fluxes.

SUMMARY

To conclude this paper, we briefly summarize the results obtained in the present study. In Sec. II, we studied a structural aspect of continuous invariance transformations connected to the identity transformation, and we stated the explicit relation between a continuous invariance transformation and a continuous invariance group transformation [(A), (11), (C)]. In Sec. III, we used the result of Sec. II to analyze invariance properties of differential equations and we uncovered the group theoretic structure, inherent in any solution which depends on a continuous parameter [(D)]. In Sec. V, a new method was given for obtaining generator of an invariance group and it was used to find a series of new generators of an invariance group of the sine-Gordon Eq. [(36b)–(36e)]. In Sec. VI, we gave a group theoretic criteria for the existence of a series of conservation laws associated with solutions of a differential equation [(E)], and this was used to provide a group theoretic explanation of a series of conservation laws of the sine-Gordon equation. The results (40a)–(40d) explicitly indicate that there exist conservation laws whose existence is inexplicable within the Lie's framework of group theory, but still can be explained by group theory if the generalized theory (Ref. 5d) is used. In the next papers, we will show that the conservation laws of the Korteweg–deVries equation and the cubic Schröedinger equation are also related to invariance groups of the generalized Lie type.[13]

Note added in proof: The transformation (10) with A_k defined by (11), (12a–d) has been found to be the power series expansion, in α, of the expression

$$T(\alpha) = e^{\alpha\bar{U}_1}\, e^{(21)^{-1}\alpha^2\bar{U}_2}\, e^{(31)^{-1}\alpha^3\bar{U}_3}\, e^{(41)^{-1}\alpha^4\bar{U}_4}\cdots .$$

ACKNOWLEDGMENTS

I wish to express my grateful acknowledgment to Professor Carl E. Wulfman for helpful discussions, his continuing encouragement and many valuable comments on the manuscript. I also thank the Research Cooperation and the Physics Department of U.O.P. for the support of this research.

APPENDIX: DETERMINING EQUATIONS OF GENERATORS

Although our transformation is more general than that of Lie, the basic idea for obtaining the differential equations (determining equations) for ξ^i is the same as Lie's, and for a detailed discussion of the Lie method we refer the reader to the book by Ovsjannikov[5b] or the book by Bluman and Cole.[5c] Using f for ξ^3, the deteter-mining equations for our problem are the following:

$$f_{9,12}=0,$$
$$f_{3,12}x^4 + f_{4,12}x^6 + f_{5,12}x^7 + f_{6,12}x^9 + f_{8,12}x^{11} + f_{12,12}x^{16} = 0,$$
$$f_{3,9}x^5 + f_{4,9}x^7 + f_{5,9}x^8 + f_{6,9}x^{10} + f_{8,9}x^{12} + f_{9,9}x^{14} = 0,$$
$$f_3x^7 + f_4x^{10} + f_5x^{11} + f_6x^{14} + f_8x^{16} + f_9x^{19} + f_{12}x^{22}$$
$$+ \sum_{\substack{i=3,4,5\\6,8,12}} (f_{3,i}x^5 + f_{4,i}x^7 + f_{5,i}x^8 + f_{6,i}x^{10} + f_{8,i}x^{12}$$
$$+ f_{9,i}x^{14}) = 0,$$

with supplementary conditions:

$$x^7 = \sin x^3, \quad x^{10} = x^4\cos x^3, \quad x^{11} = x^5\cos x^3,$$
$$x^{14} = x^6\cos x^3 - (x^4)^2\sin x^3,$$
$$x^{16} = x^8\cos x^3 - (x^5)^2\sin x^3,$$
$$x^{19} = x^9\cos x^3 - 3x^4x^6\sin x^3 - (x^4)^3\cos x^3,$$
$$x^{22} = x^{12}\cos x^3 - 3x^5x^8\sin x^3 - (x^5)^3\cos x^3,$$

where $f_j = \partial_j f$ and $f_{j,k} = \partial_j\partial_k f$.

*This work was supported by a Research Cooperation Grant.

[1]A.C. Scott, F.Y.F. Chu, and D.W. McLaughlin, Proc. IEEE 61, 1444 (1973) (review article).
[2]P.D. Lax, Comm. Pure Appl. Math. 21, 467 (1968).
[3]It has been suspected that some transformation property of the differential equation governing the wave motion is responsible for the existence of a series of conservation laws. In fact, the restricted Bäcklund transformations (R.B.T.) have provided a systematic way of deriving a series of conservation laws. However, the derivation involves a process of power series expansion of a solution with respect to some parameter. Such a method only exemplifies the existence of a series, but does not explain the origin of individual conservation law. On the discussion of R.B.T. in the theory of solitons, we refer to (a) G.L. Lamb, Rev. Mod. Phys. 43, 99 (1971); (b) D.W. McLaughlin and A.C. Scott, J. Math. Phys. 14, 1817 (1973); (c) H.D. Wahlquist and F.B. Estabrook, Phys. Rev. Lett. 31, 1386 (1973). *We add in proof the following papers on the Bäcklund transformations*: G.L. Lamb, Jr., J. Math. Phys. 15, 2157 (1974); M. Wadati, H. Sanuki, and K. Konno, Progr. Theor. Phys. (Kyoto) 53, 419 (1975).
[4]R.L. Anderson, S. Kumei, and C.E. Wulfman, Rev. Mex. Fis. 21, 1,35 (1972); J. Math. Phys. 14, 1527 (1973).

[5]For Lie's work and its later development, we refer the reader to (a) S. Lie, *Transformationgruppen* (Chelsea, New York, 1970), 3 Vols., Reprints of 1888, 1890, and 1893 eds., S. Lie, *Differentialgreichungen* (Chelsea, New York, 1967), reprint of 1891 ed., S. Lie, *Continuierliche Gruppen* (Chelsea, New York, 1967), reprint of 1893 ed. (b) L.V. Ovsjannikov, *Group theory of differential equations* (Siberian Sec. Acad. of Sci., Novosibirsk, USSR, 1962). This book has been translated into English by G.W. Bluman, Department of Mathematics, University of British Columbia (unpublished). L.V. Ovsjannikov, *Some problems arising in group analysis of differential equations* (Proceeding Conference on Symmetry, Similarity and Group Theoretic Methods in Mechanics, edited by P.G. Glockner and M.C. Singh (University of Calgary Press, Canada, 1974). (c) G.W. Bluman and J.D. Cole, J. Math. Mech. 18, 1025 (1969). G.W. Bluman and J.D. Cole, *Similarity Methods for Differential Equation* (Springer, New York, 1974). (d) R.L. Anderson, S. Kumei and C.E. Wulfman, Phys. Rev. Lett. 28, 988 (1972).

[6]We note that the well-known contact transformations of ordinary differential equations, which were extensively studied by Lie, are a realization of the derivative-dependent transformations in which only the first-order derivative appears.

[7]If the equation is an ordinary differential equation, it is always possible to find a closed space.

[8]Several years ago, Professor G.M. Lamb kindly raised the question of the relation between this generalization and the Bäcklund transformation, which depends on first-order derivatives. The basic difference is the fact that the Bäcklund transformation is not a group transformation in general, whereas our generalization allows us to construct a group transformation. We should consider that a Lie type transformation and the Bäcklund transformation are complementary in the sense that neither of them subsumes the other. Lie's infinitesimal approach, however, will be superior in the structural analysis of continuous invariance transformations.

[9]The general formula of the expression of the extended operator will be found in the paper by R.L. Anderson and S. Davison, J. Math. Anal. Appl. 48, 301 (1974).

[10]We define "trivial" flux in the following way. We consider a set $S\{f_1, f_2, \ldots, f_l\}$ which consists of divergence free fluxes, $f_1, \ldots, f_l$ and their derivatives of any order. We note that the derivatives are also divergence free. Now, a flux f is said to be trivial with respect to the set S, if f can be expressed as a linear combination of the members of the set S. In this sense, the flux $Q^n f$ is, in general, nontrivial with respect to the set $S\{f, Qf, Q^2 f, \ldots, Q^{n-1} f\}$. For instance, the flux $f_{3,3}$ of (40) is nontrivial with respect to the set $S\{f, \bar{U}_3^1 f\}$.

[11]Eq. VI. B. 7 in Ref. 1.

[12]This is due to the special character of the operator $\bar{U}_1$; the operation of $\bar{U}_1$ on the variable x^i, $i > 2$, is equivalent to the differentiation of the function u with respect to x^1. For instance, $\bar{U}_1 x^3 = x^4$, $(\bar{U}_1^1)^2 x^3 = x^6$ and $\bar{U}_1^1 x^5 = x^7$ are the transformations $u \to u_1$, $u \to u_{11}$ and $u_2 \to u_{12}$. Because of this property, the fluxes obtained from $(\bar{U}_1)^n f$ are all trivial.

[13]S. Kumei, "Group theoretic studies of conservation laws of nonlinear dispersive waves" (II, III, IV) (submitted for publication).

General techniques for single and coupled quantum anharmonic oscillators*

Francis R. Halpern

Department of Physics, University of California, San Diego, La Jolla, California 92037
(Received 6 March 1975; final revision received 9 May 1975)

Matrix mechanic methods are used to find approximate equations and solutions for quantum anharmonic oscillator problems. A series of hypotheses are introduced that truncate and partially decouple the infinite set of coupled equations that specify the problem in the matrix mechanics formulation. The dependent variables or unknowns in these equations are the matrix elements of the coordinate and momentum operators. The independent variables are the matrix indices and coupling strengths. The equations themselves specify that the off diagonal matrix elements of the Hamiltonian and the commutators expressed in terms of the unknowns vanish and that the diagonal commutator matrix elements vanish except for canonical pairs in which case they are equal to $-i\hbar$. The truncation and decoupling hypotheses offer an orderly procedure for dealing approximately with the vast array of equations of the exact problems. Only the leading behavior of the coordinate and momentum operator matrix elements is found in terms of the matrix indices and coupling parameters. Although general techniques are presented to find the equations, the solutions discussed and the applications are brief extensions of problems that have already been treated.

I. INTRODUCTION

In the earliest formulation of quantum mechanics by Heisenberg is implicitly in terms of matrices.[1] The subsequent clarification by Born and Jordan explicitly formulates the problem in terms of matrices.[2] In von Neumann's codification the matrix mechanics method appears first.[3] Despite this priority there are very few attempts to directly solve quantum mechanical problems by matrix methods. Anharmonic oscillators are treated in both Heisenberg's original paper and that of Born and Jordan.

The matrix formulation of a problem is simple. Given a Hamiltonian H that is a function of n coordinates and n momenta, it is required to find $2n$ matrices $p_1, \ldots, p_n$ and $x_1, \ldots, x_n$ such that $H(p_1, \ldots, p_n, x_1, \ldots, x_n)$ is a diagonal matrix and that the $2n^2 - n$ commutation rules $[p_a, p_b] = 0$, $[x_a, x_b] = 0$, $[p_a, x_b] = -i\hbar\delta_{ab}$ are satisfied. The solution of the problem requires the specification of the quantities $(p_a)_{r_1, r_2, \ldots, r_n; s_1, s_2, \ldots, s_n}$ and $(x_b)_{r_1, \ldots, r_n; s_1, \ldots, s_n}$ the unknowns of the problem. These quantities are functions of the matrix indices $r_1, r_2, \ldots, r_n; s_1, s_2, \ldots, s_n$ and of any parameters that may appear in the Hamiltonian. The equations of the problem are the conditions placed on the matrix elements of the Hamiltonian and the commutators:

$$\begin{aligned}
&\left(H_{r_1, r_2, \ldots, r_n; s_1, s_2, \ldots, s_n}\right) = E_{r_1, r_2, \ldots, r_n}\,\delta_{r_1 s_1}\,\delta_{r_2 s_2} \cdots \delta_{r_n s_n} \\
&[p_a, p_b]_{r_1, r_2, \ldots, r_n; s_1, s_2, \ldots, s_n} = 0 \\
&[x_a, x_b]_{r_1, r_2, \ldots, r_n; s_1, s_2, \ldots, s_n} = 0 \\
&[p_a, p_b]_{r_1, r_2, \ldots, r_n; s_1, s_2, \ldots, s_n} = -i\hbar\delta_{ab}\delta_{r_1 s_1}\delta_{r_2 s_2} \cdots \delta_{r_n s_n}.
\end{aligned} \tag{1}$$

There are many infinities of equations here rather intricately coupled to each other. In order to solve them, the alternative to a very clever or lucky ansatz is a systematic scheme to truncate and decouple these equations.

If the Hamiltonian H has the form

$$H = \tfrac{1}{2}\sum_{a=1}^{m}(p_a^2 + \omega_a^2 x_a^2) + V(x_1, \ldots, x_n)$$

with V a polynomial of degree higher than the second that satisfies the stability requirement $\lim_{r\to\infty} V = +\infty$, where $r = (x_1^2 + \cdots + x_n^2)^{1/2}$, the problem is conventionally called an anharmonic oscillator. For this class of problems experience gained with some simple examples suggests a set of hypotheses that permit the truncation and decoupling of Eqs. (1).[4]

It is these truncated and decoupled equations that shall be developed in the remainder of this paper. Almost all approximate methods of solution of quantum anharmonic oscillator problems rely on some truncation and decoupling scheme to find a manageable set of approximate equations. For example, in perturbation theory the equations are segregated according to the powers of various parameters that occur in the potential V above. It is known that approximations of this type diverge but that they may be asymptotic.[5] Padé series based on these perturbation formulas converge at least for the single oscillator.[6] In other schemes the number of degrees of freedom are kept finite. Approximation of this type may converge but are not rapidly convergent.[7] These approximations also suffer the defect of not revealing the qualitative structure of the solutions even when giving numerically accurate answers.

In the current method the matrix elements of the coordinate and momentum operators are ranked according to their expected significance, and a perturbation procedure in this index is developed. The equations that are generated are difficult but not intractable. It seems as if it will be possible to treat coupled problems. Not only are the numerical values of the energy levels computed but the values of the coordinate and momentum operator matrix elements. Most significantly the qualitative dependence of these quantities are clearly exhibited. For the first time the effects of coupling on these matrix elements is partially shown. In the two degree of freedom oscillator problem, if the degrees of freedom are called x and y, then the x and p_x operators have a y de-

pendence and vice versa. It has not yet been possible to find the matrix element for the x and p_x operators to change the y quantum number, but that aspect of the problem which belongs to the next order of approximation is under investigation and will be reported in a subsequent paper.

In the next section these hypotheses are stated. Section III contains the calculation of matrix elements of the commutators and the Hamiltonian with the help of the hypotheses. In Sec. IV the results of Refs. 4 are slightly extended in the more general language provided here. In Sec. V the difficulties with these methods are reviewed.

II. THE HYPOTHESES ABOUT THE BEHAVIOR OF THE MATRIX ELEMENTS OF THE COORDINATE AND MOMENTUM OPERATORS

The matrix element, either coordinate or momentum, $A_{r_1,\ldots,r_n;s_1,\ldots,s_n}$ is usually regarded as a function of the variables $r_1, r_2, \ldots, r_n, s_1, s_2, \ldots, s_n$. The first hypothesis is that the behavior of A is simpler in terms of the variables $r_1+s_1, r_2+s_2, \ldots, r_n+s_n, r_1-s_1, r_2-s_2, \ldots, r_n-s_n$. Thus the functions ξ and π are introduced according to

$$(x_a)_{r_1,r_2,\ldots,r_n;s_1,s_2,\ldots,s_n} = \xi_a(r_1+s_1,\ldots,r_n+s_n;\ r_1-s_1,\ldots,r_n-s_n),$$

$$(p_b)_{r_1,r_2,\ldots,r_n;s_1,s_2,\ldots,s_n} = i\pi_b(r_1+s_1,\ldots,r_n+s_n;\ r_1-s_1,\ldots,r_n-s_n).$$

The second hypothesis is that the ξ and π are slowly varying functions of their first n arguments. This is already suggested by $\sqrt{r+s}$ behavior of the single harmonic oscillator. This hypothesis is a decoupling hypothesis since in an equation that couples $\xi(s)$ and $\xi(s+1)$ the difference can be ignored and the two terms set equal to lowest order. The third hypothesis is that the ξ and π are rapidly decreasing functions of their second set of variables. Again for the single harmonic oscillator this is true. The functions vanish if $|r-s|>1$. This hypothesis is a truncation hypothesis. In an order of approximation only those unknowns for which all $|r-s|$ are less than a specified number are included. This reduces to a finite number both the number of unknowns and the number of equations since infinitely many equations involve only unknowns for which $|r-s|$ exceed the cutoff value for the order of approximation.

III. CALCULATION OF COMMUTATORS AND HAMILTONIANS

The commutators are treated first since they are applicable to all problems. The commutator of two operators A and B is carried out. Wherever it does not obscure the discussion the notations $\mathbf{r}, \mathbf{s}, \mathbf{t}$ are substituted for $r_1, \ldots, r_n, s_1, \ldots, s_n, t_1, \ldots, t_n$, etc.:

$$\begin{aligned}[A,B]_{\mathbf{r},\mathbf{s}} &= \sum_{\mathbf{t}} (A_{\mathbf{rt}}B_{\mathbf{ts}} - B_{\mathbf{rt}}A_{\mathbf{ts}}) \\ &= \sum_{\mathbf{t}} [\alpha(\mathbf{r}+\mathbf{t};\ \mathbf{r}-\mathbf{t})\,\beta(\mathbf{t}+\mathbf{s};\mathbf{t}-\mathbf{s}) \\ &\quad - \beta(\mathbf{r}+\mathbf{t};\mathbf{r}-\mathbf{t})\,\alpha(\mathbf{t}+\mathbf{s};\mathbf{t}-\mathbf{s})] \\ &= \sum_{\mathbf{t}} [\alpha(\mathbf{r}+\mathbf{s}+\mathbf{t}-\mathbf{s};\mathbf{r}-\mathbf{t})\beta(\mathbf{r}+\mathbf{s}+\mathbf{t}-\mathbf{r};\mathbf{t}-\mathbf{s}) \\ &\quad - \beta(\mathbf{r}+\mathbf{s}+\mathbf{t}-\mathbf{s};\mathbf{r}-\mathbf{t})\alpha(\mathbf{r}+\mathbf{s}+\mathbf{t}-\mathbf{r};\mathbf{t}-\mathbf{s})].\end{aligned}$$

Now let $\mathbf{r}-\mathbf{t}=\mathbf{w}$ and $\mathbf{t}-\mathbf{s}=\mathbf{v}$ so that $\mathbf{w}+\mathbf{v}=\mathbf{r}-\mathbf{s}$.

Now by hypothesis III $\mathbf{v}$ and $\mathbf{w}$ must be small or either ξ or π will be small. Thus it is possible to expand ξ and π about $\mathbf{r}+\mathbf{s}$ this gives

$$\begin{aligned}[A,B]_{\mathbf{r},\mathbf{s}} = \sum_{\mathbf{w},\mathbf{v}} \delta_{\mathbf{w}+\mathbf{v},\mathbf{r}-\mathbf{s}} [&\alpha(\mathbf{r}+\mathbf{s},\mathbf{w})\,\beta(\mathbf{r}+\mathbf{s},\mathbf{v}) \\ &- \beta(\mathbf{r}+\mathbf{s},\mathbf{w})\,\alpha(\mathbf{r}+\mathbf{s},\mathbf{v}) \\ &+ \mathbf{v}\cdot\partial\,\alpha(\mathbf{r}+\mathbf{s},\mathbf{w})\,\beta(\mathbf{r}+\mathbf{s},\mathbf{v}) \\ &- \alpha(\mathbf{r}+\mathbf{s},\mathbf{w})\mathbf{w}\cdot\partial\,\beta(\mathbf{r}+\mathbf{s},\mathbf{v}) \\ &- \mathbf{v}\cdot\partial\,\beta(\mathbf{r}+\mathbf{s};\mathbf{w})\,\alpha(\mathbf{r}+\mathbf{s};\mathbf{v}) \\ &+ \beta(\mathbf{r}+\mathbf{s},\mathbf{w})\,\mathbf{w}\cdot\partial\,\alpha(\mathbf{r}+\mathbf{s},\mathbf{v})] \\ &+ \text{higher order terms in } \mathbf{v} \text{ and } \mathbf{w}.\end{aligned}$$

The zero order term vanishes by interchanging $\mathbf{v}$ and $\mathbf{w}$ in either but not both terms in the bracket. Again interchanging $\mathbf{v}$ and $\mathbf{w}$ in the first order term permits the more compact expression

$$\begin{aligned}[A,B]_{\mathbf{r},\mathbf{s}} = 2\sum_{\mathbf{w},\mathbf{v}} \delta_{\mathbf{w}+\mathbf{v},\mathbf{r}-\mathbf{s}} [&\mathbf{v}\cdot\partial\,\alpha(\mathbf{r}+\mathbf{s};\mathbf{w})\,\beta(\mathbf{r}+\mathbf{s};\mathbf{v}) \\ &- \alpha(\mathbf{r}+\mathbf{s};\mathbf{w})\,\mathbf{w}\cdot\partial\,\beta(\mathbf{r}+\mathbf{s};\mathbf{v})].\end{aligned}$$

The second order term in $\mathbf{v}$ and $\mathbf{w}$ vanishes for symmetry reasons so that the expression is correct up to terms of order three which will be neglected. The diagonal commutators are of particular interest:

$$\begin{aligned}[A,B]_{\mathbf{r},\mathbf{r}} &= 2\sum_{\mathbf{v}} \mathbf{v}\cdot\partial\ \alpha(\mathbf{r}+\mathbf{s};-\mathbf{v})\,\beta(\mathbf{r}+\mathbf{s};\mathbf{v}) \\ &\quad + \alpha(\mathbf{r}+\mathbf{s};-\mathbf{v})\,\mathbf{v}\cdot\partial\ \beta(\mathbf{r}+\mathbf{s};\mathbf{v}) \\ &= 2\sum_{\mathbf{v}} \mathbf{v}\cdot\partial\,[\alpha(\mathbf{r}+\mathbf{s};-\mathbf{v})\,\beta(\mathbf{r}+\mathbf{s};\mathbf{v})].\end{aligned}$$

If the operators A and B are actually coordinate or momentum operators, there are additional symmetries that permit further simplification. The coordinate matrices are assumed to be real and symmetric so that $\xi(\mathbf{r};-\mathbf{w})=\xi(\mathbf{r};\mathbf{w})$. The momentum matrices are assumed to be pure imaginary antisymmetric so that $\pi(\mathbf{r};-\mathbf{w}) = -\pi(\mathbf{r};\mathbf{w})$. These symmetries are consequences of time reversal invariance. Applied to the commutators these symmetries yield

$$[p_a,p_b]_{\mathbf{r},\mathbf{r}} = -2i^2\sum_{\mathbf{v}} \mathbf{v}\cdot\partial[\pi_a(\mathbf{r}+\mathbf{r},\mathbf{v})\,\pi_b(\mathbf{r}+\mathbf{r},\mathbf{v})]=0,$$

$$[x_a,x_b]_{\mathbf{r},\mathbf{r}} = 2\sum_{\mathbf{v}} \mathbf{v}\cdot\partial[\xi_a(\mathbf{r}+\mathbf{r},\mathbf{v})\,\xi_b(\mathbf{r}+\mathbf{r},\mathbf{v})]=0.$$

These two expressions vanish because they are odd in $\mathbf{v}$. There is nothing surprising about these results since the diagonal matrix elements of the commutator of two symmetric or two antisymmetric matrices vanish. It is reassuring that these properties have not been lost in the preceding manipulation. The commutator of a momentum and a coordinate are given by

$$[p_a,x_b]_{\mathbf{r},\mathbf{r}} = -2i\sum_{\mathbf{v}} \mathbf{v}\cdot\partial[\pi_a(\mathbf{r}+\mathbf{r},\mathbf{v})\,\xi_b(\mathbf{r}+\mathbf{r},\mathbf{v})]. \qquad (2)$$

This does not vanish since it is even in $\mathbf{v}$ since π is an odd and ξ an even function of $\mathbf{v}$. There is no particular

improvement possible in the nondiagonal commutator matrix elements. Thus the commutator equations become

$$\sum_{\mathbf{v}} \mathbf{v}\cdot\partial[\pi_a(\mathbf{r}+\mathbf{r},\mathbf{v})\,\xi_b(\mathbf{r}+\mathbf{r},\mathbf{v})] = -\tfrac{1}{2}\hbar\,\delta_{ab} \quad \text{(diagonal)},$$

$$\sum_{\mathbf{w},\mathbf{v}} \delta_{\mathbf{w}+\mathbf{v},\mathbf{r}-\mathbf{s}}[\mathbf{v}\cdot\partial\alpha(\mathbf{r}+\mathbf{s},\mathbf{w})\,\beta(\mathbf{r}+\mathbf{s},\mathbf{v}) - \alpha(\mathbf{r}+\mathbf{s},\mathbf{w})\mathbf{w}\cdot\partial\beta(\mathbf{r}+\mathbf{s},\mathbf{v})] = 0 \quad \text{(off-diagonal)}, \tag{3}$$

where α and β run over all momentum and coordinate operators. The truncation hypothesis will be appealed to select a finite set of these equations. Since the sums are over $\mathbf{v}$ and $\mathbf{w}$, it is not difficult to isolate sums and terms with $\mathbf{v}$ and $\mathbf{w}$ bounded.

There is, however, a serious difficulty which requires additional rather more ad hoc assumptions to resolve it. The commutator equations are not completely decoupled. The differentiation which is really more properly a finite difference operation couples adjacent values of the first argument, for example, $\mathbf{v}\cdot\partial\,\alpha(\mathbf{r}+\mathbf{s},\mathbf{w}) = v_1[\alpha(\mathbf{r}_1+\mathbf{s}_1+1;\mathbf{w}) - \alpha(r_1+s_1;w)]+\cdots$, where irrelevent terms have been neglected. The difference between $\alpha(r_1+s_1+1,\mathbf{w})$ and $\alpha(r_1+s_1,\mathbf{w})$ cannot be neglected since that is the whole effect so that the differentiation is a coupling. For the diagonal commutator matrix elements this can be overcome because the differential equations can be integrated but for the off-diagonal commutator matrix equations even in the one body problem the equations are nonintegrable and the coupling must be removed by another less satisfactory method. By examining the equations, Hamiltonian, and commutator and gaining some experience in dealing with oscillator problems, it has been possible to propose solutions with free constants. These can be substituted in the various equations and the differentiations can be explicitly carried out. The resulting equations for the constants are decoupled. This leaves something to be desired in rigor and generality but does permit decoupling. The method is illustrated in the examples.

Next consider the matrix element of the Hamiltonian. Start with the quadratic terms, the harmonic parts

$$\begin{aligned}(p_a^2)_{\mathbf{rs}} &= \sum_{\mathbf{t}} (p_a)_{\mathbf{rt}}\,(p_a)_{\mathbf{ts}}\\ &= -\sum_{\mathbf{t}} \pi_a(\mathbf{r}+\mathbf{t};\mathbf{r}-\mathbf{t})\,\pi_a(\mathbf{t}+\mathbf{s};\mathbf{t}-\mathbf{s})\\ &= -\sum_{\mathbf{t}} \pi_a(\mathbf{r}+\mathbf{s}+\mathbf{t}-\mathbf{s};\mathbf{r}-\mathbf{t})\,\pi_a(\mathbf{r}+\mathbf{s}+t-r);\ \mathbf{t}-\mathbf{s})\\ &= -\sum_{\mathbf{w},\mathbf{v}} \delta_{\mathbf{w}+\mathbf{v};\mathbf{r}-\mathbf{s}}\,\pi_a(\mathbf{r}+\mathbf{s}+\mathbf{v};\mathbf{w})\,\pi_a(\mathbf{r}+\mathbf{s}-\mathbf{w};\mathbf{v})\\ &= -\sum \delta_{\mathbf{w}+\mathbf{v};\mathbf{r}-\mathbf{s}}[\pi_a(\mathbf{r}+\mathbf{s};\mathbf{w})\,\pi_a(\mathbf{r}+\mathbf{s};\mathbf{v})\\ &\quad +\mathbf{v}\cdot\partial\,\pi_a(\mathbf{r}+\mathbf{s};\mathbf{w})\,\pi_a(\mathbf{r}+\mathbf{s};\mathbf{v})\\ &\quad -\pi_a(\mathbf{r}+\mathbf{s};\mathbf{w})\mathbf{w}\cdot\partial\,\pi_a(\mathbf{r}+\mathbf{s};\mathbf{v})] + O(\mathbf{v}^2,\mathbf{w}\cdot\mathbf{v},\mathbf{w}^2).\end{aligned}$$

The linear term in $\mathbf{w}$ and $\mathbf{v}$ vanishes by the permutation symmetry between $\mathbf{w}$ and $\mathbf{v}$. This gives

$$(p_a^2)_{\mathbf{rs}} = -\sum_{\mathbf{w},\mathbf{v}} \delta_{\mathbf{w}+\mathbf{v},\mathbf{r}-\mathbf{s}}\,\pi_a(\mathbf{r}+\mathbf{s};\mathbf{w})\,\pi_a(\mathbf{r}+\mathbf{s},\mathbf{v}) + O(\mathbf{v}^2,\mathbf{w}\cdot\mathbf{v},\mathbf{w}^2). \tag{4}$$

No special properties of the functions π were used to derive this result so it follows that

$$(x_b)^2_{\mathbf{rs}} = \sum_{\mathbf{w},\mathbf{v}} \delta_{\mathbf{w}+\mathbf{v};\mathbf{r}-\mathbf{s}}\,\xi_b(\mathbf{r}+\mathbf{s};\mathbf{w})\,\xi_b(\mathbf{r}+\mathbf{s};\mathbf{v}) + O(\mathbf{w}^2,\mathbf{w}\cdot\mathbf{v},\mathbf{v}^2). \tag{5}$$

As a sample of how the anharmonic terms can be treated ($\sum A_{abcd}x_a x_b x_c x_d)_{\mathbf{rs}}$ is considered. The coefficient A_{abcd} is assumed to be invariant under the permutation group on (a,b,c,d):

$$\begin{aligned}&\Big(\sum_{abcd} A_{abcd}x_a x_b x_c x_d\Big)_{\mathbf{rs}}\\ &= \sum_{abcd\mathbf{tuv}} A_{abcd}(x_a)_{\mathbf{rt}}\,(x_b)_{\mathbf{tu}}\,(x_c)_{\mathbf{uv}}\,(x_d)_{\mathbf{vs}}\\ &= \sum A_{abcd}\,\xi_a(\mathbf{r}+\mathbf{t};\mathbf{r}-\mathbf{t})\,\xi_b(\mathbf{t}+\mathbf{u};\mathbf{t}-\mathbf{u})\\ &\quad\times\xi_c(\mathbf{u}+\mathbf{v};\mathbf{u}-\mathbf{v})\,\xi_d(\mathbf{v}+\mathbf{s};\mathbf{v}-\mathbf{s})\\ &= \sum A_{abcd}\,\xi_a(\mathbf{r}+\mathbf{s}+\mathbf{t}-\mathbf{s};\mathbf{r}-\mathbf{t})\\ &\quad\times\xi_b(\mathbf{r}+\mathbf{s}+\mathbf{t}+\mathbf{u}-\mathbf{r}-\mathbf{s};\mathbf{t}-\mathbf{u})\\ &\quad\times\xi_c(\mathbf{r}+\mathbf{s}+\mathbf{u}+\mathbf{v}-\mathbf{r}-\mathbf{s};\mathbf{u}-\mathbf{v})\,\xi_d(\mathbf{r}+\mathbf{s}+\mathbf{v}-\mathbf{r};\mathbf{v}-\mathbf{s}).\end{aligned}$$

Now let $\mathbf{r}-\mathbf{t}=\mathbf{w}_1$, $\mathbf{t}-\mathbf{u}=\mathbf{w}_2$, $\mathbf{u}-\mathbf{v}=\mathbf{w}_3$, $\mathbf{v}-\mathbf{s}=\mathbf{w}_4$ so that $\mathbf{w}_1+\mathbf{w}_2+\mathbf{w}_3+\mathbf{w}_4=\mathbf{r}-\mathbf{s}$:

$$\begin{aligned}&\Big(\sum_{abcd} A_{abcd}x_a x_b x_c x_d\Big)_{\mathbf{rs}}\\ &= \sum_{\substack{abcd\\ \mathbf{w}_1\mathbf{w}_2\mathbf{w}_3\mathbf{w}_4}} A_{abcd}\,\delta_{\mathbf{w}_1+\mathbf{w}_2+\mathbf{w}_3+\mathbf{w}_4,\mathbf{r}-\mathbf{s}}\,\xi_a(\mathbf{r}+\mathbf{s}+\mathbf{w}_2+\mathbf{w}_3+\mathbf{w}_4;\mathbf{w}_1)\\ &\quad\times\xi_b(\mathbf{r}+\mathbf{s}-\mathbf{w}_1+\mathbf{w}_3+\mathbf{w}_4;\mathbf{w}_2)\,\xi_c(\mathbf{r}+\mathbf{s}-\mathbf{w}_1-\mathbf{w}_2+\mathbf{w}_4;\mathbf{w}_3)\\ &\quad\times\xi_d(\mathbf{r}+\mathbf{s}-\mathbf{w}_1-\mathbf{w}_2-\mathbf{w}_3;\mathbf{w}_4).\\ &= \sum A_{abcd}\,\delta_{\mathbf{w}_1+\mathbf{w}_2+\mathbf{w}_3+\mathbf{w}_4,\mathbf{r}-\mathbf{s}}\{\xi_a(\mathbf{r}+\mathbf{s};\mathbf{w}_1)\,\xi_b(\mathbf{r}+\mathbf{s};\mathbf{w}_2)\\ &\quad\times\xi_c(\mathbf{r}+\mathbf{s};\mathbf{w}_3)\,\xi_d(\mathbf{r}+\mathbf{s};\mathbf{w}_4)+[(\mathbf{w}_2+\mathbf{w}_3+\mathbf{w}_4)\cdot\partial\,\xi_a(\mathbf{r}+\mathbf{s};\mathbf{w}_1)]\\ &\quad\times\xi_b(\mathbf{r}+\mathbf{s};\mathbf{w}_2)\,\xi_c(\mathbf{r}+\mathbf{s};\mathbf{w}_3)\,\xi_d(\mathbf{r}+\mathbf{s};\mathbf{w}_4)+\xi_a(\mathbf{r}+\mathbf{s};\mathbf{w}_1)\\ &\quad\times[(-\mathbf{w}_1+\mathbf{w}_3+\mathbf{w}_4)\cdot\partial\,\xi_b(\mathbf{r}+\mathbf{s};\mathbf{w}_2)]\,\xi_c(\mathbf{r}+\mathbf{s};\mathbf{w}_3)\,\xi_d(\mathbf{r}+\mathbf{s};\mathbf{w}_4)\\ &\quad+\xi_a(\mathbf{r}+\mathbf{s};\mathbf{w}_1)\,\xi_b(\mathbf{r}+\mathbf{s};\mathbf{w}_2)[(-\mathbf{w}_1-\mathbf{w}_2+\mathbf{w}_4)\cdot\partial\,\xi_c(\mathbf{r}+\mathbf{s};\mathbf{w}_3)]\\ &\quad\times\xi_d(\mathbf{r}+\mathbf{s};\mathbf{w}_4)+\xi_a(\mathbf{r}+\mathbf{s};\mathbf{w}_1)\,\xi_b(\mathbf{r}+\mathbf{s};\mathbf{w}_2)\,\xi_c(\mathbf{r}+\mathbf{s};\mathbf{w}_3)\\ &\quad\times[(-\mathbf{w}_1-\mathbf{w}_2-\mathbf{w}_3)\cdot\partial\,\xi_d(\mathbf{r}+\mathbf{s};\mathbf{w}_4)]\}+O(\mathbf{w}_i^2,\mathbf{w}_i\cdot\mathbf{w}_j).\end{aligned}$$

Now it will be shown that the linear term vanishes because of symmetry so that

$$\begin{aligned}&\Big(\sum_{abcd} A_{abcd}x_a x_b x_c x_d\Big)_{\mathbf{rs}}\\ &= \sum_{\substack{a,b,c,d\\ \mathbf{w}_1,\mathbf{w}_2,\mathbf{w}_3,\mathbf{w}_4}} A_{abcd}\,\delta_{\mathbf{w}_1+\mathbf{w}_2+\mathbf{w}_3+\mathbf{w}_4;\mathbf{r}-\mathbf{s}}\,\xi_a(\mathbf{r}+\mathbf{s};\mathbf{w}_1)\,\xi_b(\mathbf{r}+\mathbf{s};\mathbf{w}_2)\\ &\quad\times\xi_c(\mathbf{r}+\mathbf{s};\mathbf{w}_3)\,\xi_d(\mathbf{r}+\mathbf{s};\mathbf{w}_4).\end{aligned} \tag{6}$$

The leading part of any anharmonic term can be written down at sight following this example. Since the first argument of all the ξ functions is $\mathbf{r}+\mathbf{s}$, an immense amount of decoupling has taken place.

To show that the linear term vanishes, note that A_{abcd} and $\delta_{\mathbf{w}_1+\mathbf{w}_2+\mathbf{w}_3+\mathbf{w}_4+\mathbf{r}-\mathbf{s}}$ are invariant under the permutations

of a, b, c, d and w_1, w_2, w_3, w_4 respectively. The term in question is of the form $A_{abcd}\delta_{w_1+w_2+w_3+w_4+r-s}F(a,b,c,d;w_1,w_2,w_3,w_4)$. Let $P_{a\,b\,c\,d}$ and P_{w_1,w_2,w_3,w_4} be the sum of the 4! permutation operators in the permutation groups on $a\ b\ c\ d$ and w_1, w_2, w_3, w_4 respectively. Since a, b, c, d and w_1, w_2, w_3, w_4 are dummy indices of summation

$$\sum A_{abcd}\delta_{w_1+w_2+w_3+w_4;r-s}\ F(a,b,c,d;w_1,w_2,w_3,w_4)$$

$$=(4!)^{-2}\sum P_{abcd}\,P_{w_1,w_2,w_3,w_4}\,\delta_{w_1+w_2+w_3+w_4;r-s}A_{abcd}$$

$$\times F(a,b,c,d;w_1,w_2,w_3,w_4)$$

and, since A and δ are invariant,

$$=(4!)^{-2}\sum \delta_{w_1+w_2+w_3+w_4;r-s}\,A_{abcd}P_{w_1w_2w_3w_4}P_{abcd}$$

$$\times F(a,b,c,d;w_1,w_2,w_3,w_4).$$

The permutations applied to F will now be shown to vanish. Consider the terms in F of the form $[(w_2+w_3+w_4)\circ\partial\,\xi_a(w_1)]\,\xi_b(w_2)\,\xi_c(w_3)\,\xi_d(w_4)$. There are six permutations which leave it invariant. For example, the simultaneous interchange of (w_2w_3) and (bc). Obviously with each member of the permutation group on $b\ c\ d$ there is a member of the permutation group on $w_2w_3w_4$ such that the product leaves the expression above unaltered. Thus the coefficient of $\partial\,\xi_a(w_1)\xi_b(w_2)\xi_c(w_3)\xi_d(w_4)$ from this part is $(6w_2+6w_3+6w_4)$. If $a\ b$ and w_1w_2 are interchanged in the part

$$\xi_a(w_1)[(-w_1+w_3+w_4)\circ\partial\xi_b(w_2)\xi_c(w_3)\xi_d(w_4)],$$

the expression $(-w_2+w_3+w_4)\circ\partial\xi_a(w_1)\xi_b(w_2)\xi_c(w_3)\xi_d(w_4)$ results. Again there are six permutations that leave the term $\partial\xi_a(w_1)\xi_b(w_2)\xi_c(w_3)\xi_d(w_4)$ unaltered. This time because of the minus sign the coefficient becomes $2(w_2+w_3+w_4)$. Interchanging $(a\ c)$ $(w_1\ w_2)$ in the term $\xi_a(w_1)\xi_b(w_2)(-w_1-w_2-w_4)\circ\partial\xi_c(w_3)\xi_d(w_4)$ and carrying out the six permutations gives the result $-2(w_2+w_3+w_4)\circ\partial\xi_a(w_1)\xi_b(w_2)\xi_c(w_3)\xi_d(w_4)$. Finally the interchange $(a\ d)$ $(w_1\ w_4)$ applied to the expression $\xi_a(w_1)\xi_b(w_2)\xi_c(w_3)\,(-w_1-w_2-w_3)\circ\partial\xi_d(w_4)$ followed by the six permutations gives the result $-6(w_2+w_3+w_4)\cdot\partial\xi_a(w_1)\times\xi_b(w_2)\xi_c(w_3)\xi_d(w_4)$. The sum of these contributions is zero. All other terms can be similarly examined, and the result is that they all vanish so the linear term vanishes.

It is reassuring to examine the combinatorics and see that nothing has been left out. If a term such as $w_2\circ\partial\xi_a(w_1)\xi_b(w_2)\xi_c(w_3)\xi_d(w_4)$ is identified as a monomial the original expression has 12 monomials. After all the permutations are performed there will be $24\circ24\circ12$ monomials not all distinct. There are 24 ways of associating $a\ b\ c\ d$ with w_1, w_2, w_3, w_4; there are four choices of where the differentiation can be applied and finally three choices of which w can be dotted into the gradient. Thus there are $24\cdot4\circ3$ distinct monomials. Each monomial must occur $24\circ24\circ12/24\cdot4\circ3=24$ times. For those treated above this is true 12 times with a plus sign and 12 times with a minus to give a net zero.

Thus we have found equations for commutators (2) and (3) and the quadratic (4) and (5) and the quartic terms (6) that occur in the Hamiltonian. Terms of higher than fourth degree can be treated similarly. These techniques permit any oscillator problem to be reduced to equations of this type. Algebraic equations for the Hamiltonian and first order nonlinear partial differential equations for the commutators are obtained. In the next section several examples are considered.

IV. EXAMPLES

The first example is an oscillator with an harmonic and an anharmonic term. The Hamiltonian H is given by

$$H=p^2+x^2+\lambda x^4.$$

The purely anharmonic oscillator p^2+x^4 has been treated earlier. For this problem the expression for the commutator C and Hamiltonian H become

$$\tfrac{1}{2}C_{rr}=i\sum_v v\,\partial[\pi(2r,v)\xi(2r,v)]=-\tfrac{1}{2}\hbar i,$$

$$C_{rs}=\sum_{w,v}\delta_{w+v;r-s}[v\partial\pi(r+s;w)\,\xi(r+s;v)$$

$$-\pi(r+s;w)w\partial\xi(r+s;v)],$$

$$H_{rs}=\sum_{w,v}\delta_{w+v;r-s}[-\pi(r+s;w)\,\pi(r+s,v)$$

$$+\xi(r+s;w)\,\xi(r+s;v)]$$

$$+\lambda\sum_{w_1,w_2,w_3,w_4}\delta_{w_1+w_2+w_3+w_4,r-s}\,\xi(r+s,w_1)$$

$$\times\xi(r+s,w_2)\,\xi(r+s,w_2)\,\xi(r+s,w_4).$$

The successive approximations are given by the sequence of equations for the nth approximation:

$$H_{s,s+2}=H_{s,s+4}=\cdots=H_{ss+2n}=0,$$

$$C_{ss}=-i\hbar,$$

$$C_{ss+2}=C_{ss+4}=\cdots=C_{ss+2n-2}=0.$$

The lowest order equations are explicitly

$$-\pi^2(r+s,1)+\xi^2(r+s,1)+4\lambda\xi^4(r+s,1)=0,$$

$$2\,\partial[(\pi(2r,1)\,\xi(2r,1)]=-\hbar/2,$$

where $s=r+2$. There are two unknown functions $\xi(2r,1)$ and $\pi(2r,1)$. The difference between $\xi(2r+2,1)$ and $\xi(2r,1)$ is negligible at this order of approximation. The truncation hypothesis suggests a set of two equations. There really are an infinite set of pairs of equations but they are decoupled. Terms such as

$$x_{r,r+1}x_{r+1,r+2}x_{r+2,r+3}x_{r+3,r+2}$$

$$=\xi(2r+1,1)\xi(2r+3,1)\xi(2r+5,1)\xi(2r+5,-1),$$

that occur in the quartic term are set equal to $\xi^4(2r+2,1)$. In this way the intricate couplings are broken. The derivatives in commutator also couple various terms. The diagonal commutator is integrable so that this coupling is also dissolved. The integration of the commutator gives

$$\pi(2r,1)\xi(2r,1)=-\tfrac{1}{2}\hbar r.$$

If this equation is used to eliminate π from the Hamiltonian equation, it gives

$$-\hbar^2r^2/4\xi^2+\xi^2+4\lambda\xi^4=0$$

or

$$16\lambda\xi^6 + 4\xi^4 - \hbar^2 r^2 = 0.$$

To simplify this equation, let $\xi = (\hbar r/2)^{1/2}\eta$ so that the equation becomes

$$2a\eta^6 + \eta^4 - 1 = 0,$$

where $a = \lambda\hbar r$. The solution for ξ^2 is given by

$$\xi^2 = (\tfrac{1}{2}\hbar r)^{1/2}\,((1/2\lambda\hbar r)^{1/3} - \tfrac{1}{3}$$
$$+ \{2(\lambda\hbar r)^2 - \tfrac{1}{27} + 2\lambda\hbar r[(\lambda\hbar r)^2 - \tfrac{1}{27}]^{1/2}\}^{1/3}$$
$$+ \{2(\lambda\hbar r)^2 - \tfrac{1}{27} - 2\lambda\hbar r[(\lambda\hbar r)^2 - \tfrac{1}{27}]^{1/2}\}^{1/3})$$

For $(\lambda\hbar r)^2 > \frac{1}{27}$ all the quantities are real while for $(\lambda\hbar r)^2 < \frac{1}{27}$ the radicals are to be taken as complex conjugates. There are two significant features of the solution that are preserved in higher orders. First ξ^2 is a function of the variables $\hbar r$ and $\lambda\hbar r$. Secondly the dependence of ξ^2 on $\hbar r$ is very simple while the dependence on $\lambda\hbar r$ is intricate but weak.

The $\hbar r$ dependence of ξ^2 is the same as for a harmonic oscillator just $(\hbar r/2)$. The $\lambda\hbar$ dependence is monotonic starting at 1 for $\lambda\hbar r = 0$ and becoming asymptotically equal to $(2\lambda\hbar r)^{-1/6}$ for large values of $\lambda\hbar r$. It is easy to find expansion for large and small values of $\lambda\hbar r$. These are

$$\left(\frac{2}{\hbar r}\right)^{1/2}\xi = \begin{cases} 1 - \frac{1}{4}(2\lambda\hbar r) + \frac{9}{32}(2\lambda\hbar r)^2 - \frac{55}{128}(2\lambda\hbar r)^3 + \cdots, & 2\lambda\hbar r \text{ is small,} \\ (2\lambda\hbar r)^{-1/6}[1 - \frac{1}{6}(2\lambda\hbar r)^{-2/3} + \frac{1}{24}(2\lambda\hbar r)^{-4/3} - \frac{1}{48}(2\lambda\hbar r)^{-6/3} + \cdots], & 2\lambda\hbar r \text{ is large.} \end{cases}$$

For intermediate values rational functions such as

$$\left(\frac{2}{\hbar r}\right)^{1/2}\xi \approx \frac{1}{1 + (2\lambda\hbar r)^{1/6}},$$

$$\left(\frac{2}{\hbar r}\right)^{1/2}\xi \approx \frac{1 + (2\lambda\hbar r)^{1/6}}{1 + (2\lambda\hbar r)^{1/6} + (2\lambda\hbar r)^{2/6}},$$

$$\left(\frac{2}{\hbar r}\right)^{1/2}\xi \approx \frac{1 + (2\lambda\hbar r)^{1/6} + (2\lambda\hbar r)^{2/6}}{1 + (2\lambda\hbar r)^{1/6} + (2\lambda\hbar r)^{2/6} + (2\lambda\hbar r)^{3/6}},$$

which agree with the first, one, two, and three terms of the series expansions in $(2\lambda\hbar r)^{1/6}$ for both large and small values of $(2\lambda\hbar r)^{1/6}$ can be used to interpolate with very good accuracy. These rational interpolating functions are hard to construct when the degree of the numerator is six or larger.

For the next higher approximation the exact solution cannot be given and only the asymptotic series and the rational interpolating are available.

Now consider the second approximation to this problem. The functions $\xi(r,1)$, $\xi(r,3)$, $\pi(r,1)$, and $\pi(r,3)$ are considered and the equations are $C_{ss} = -i\hbar$, $C_{ss+2} = H_{ss+2} = H_{ss+4} = 0$. The diagonal commutator equation C_{ss} is

$$\sum v\,\partial[\pi(2r;v)\,\xi(2r;v)] = -\tfrac{1}{2}\hbar$$

The summation index v takes on the values ± 1 and ± 3. Since the argument of the sum is even in v, it follows that

$$\partial[\pi(2r;1)\,\xi(2r;1) + 3\pi(2r,3)\,\xi(2r;3)] = -\tfrac{1}{4}\hbar$$

or that

$$\pi(2r;1)\,\xi(2r;1) + 3\pi(2r,3)\,\xi(2r;3) = -\tfrac{1}{2}\hbar r$$

The Hamiltonian equations $H_{ss+2} = 0$ and $H_{ss+4} = 0$ are

$$H_{ss+2} =$$
$$-\pi^2(2r+2;1) + 2\pi(2r+2;1)\pi(2r+2;3) + \xi^2(2r+2;1)$$
$$+ 2\xi(2r+2;1)\,\xi(2r+2;3) + \lambda[4\xi^4(2r+2;1)$$
$$+ 12\xi^3(2r+2;1)\,\xi(2r+2;3) + 12\xi^2(2r+2;1)\,\xi^2(2r+2;3)$$
$$+ 12\xi(2r+2;1)\,\xi^3(2r+2;3)] = 0,$$
$$H_{ss+4} =$$
$$-\pi(2r+4;1)\,\pi(2r+4;3) + \xi(2r+4;1)\,\xi(2r+4;3)$$
$$+ \lambda[\xi^4(2r+4;1) + 12\xi^3(2r+4;1)\,\xi(2r+4;3)$$
$$+ 6\xi^2(2r+4;1)\,\xi^2(2r+4;3) + 12\xi(2r+4;1)\,\xi^3(2r+4;3)]$$
$$= 0.$$

The commutator equation $C_{ss+2} = 0$ is given by

$$C_{ss+2} = -\partial\pi(2r+2;3)\,\xi(2r+2;1) + \partial\pi(2r+2;1)\,\xi(2r+2;1)$$
$$- 3\partial\pi(2r+2;1)\,\xi(2r+2;3) - 3\pi(2r+2;3)\,\partial\xi(2r+2;1)$$
$$- \pi(2r+2;1)\partial\xi(2r+2;1) - \pi(2r+2;1)\partial\xi(2r+2;3) = 0$$

This equation unlike the diagonal commutator equation is not completely integrable. A general integral of the form

$$F[\xi(2r+2;1),\ \xi(2r+2;3),\ \pi(2r+2;1),\ \pi(2r+2;3),\ r] = C$$

does not exist. If such an integral existed, then $dF = \mu C_{ss+2}$, where μ is a function, an integrating factor for $C_{ss+2} = 0$. A little calculation shows, however, that there is no function μ such that the integrability conditions

$$\frac{\partial^2 F}{\partial x_1 \partial x_2} = \frac{\partial^2 F}{\partial x_2 \partial x_1}$$

are satisfied with x_1 and x_2 standing for any of the variables $\xi(2r+2;1)$, $\xi(2r+2;3)$, $\pi(2r+2;1)$, and $\pi(24+2;3)$. The set of equations still has a solution, but it depends on special integrals of $C_{ss+2} = 0$. Since this equation couples different values of r, the coupling is not yet broken. In principle the three other algebraic equations could be solved in terms of one of the variables and then C_{ss+2} could be written as an ordinary differential equation in this variable. Because of the complicated character of the three algebraic equations, this does not appear to be a practical technique.

The solution to the equations of the lowest order of approximation does, however, give suggestions about how a solution of this set of equations may be achieved. The functions ξ and π can be written as $(\hbar r/2)^{1/2}$ times function of $(2\lambda\hbar r)^{1/6}$ and these functions will have a weak

dependence on $(2\lambda\hbar r)^{1/6}$. The power series expansions in $(2\lambda\hbar r)^{1/6}$ and $(2\lambda\hbar r)^{-1/6}$ can be found and the rational interpolating function can also be constructed.

As an ansatz take

$$\xi(2r;1)=\sqrt{\hbar r/2}\ a[(2\lambda\hbar r)^{1/6}],$$
$$\xi(2r;3)=\sqrt{\hbar r/2}\ b[(2\lambda\hbar r)^{1/6}],$$
$$\pi(2r;1)=\sqrt{\hbar r/2}\ c[(2\lambda\hbar r)^{1/6}],$$
$$\pi(2r;3)=\sqrt{\hbar r/2}\ d[(2\lambda\hbar r)^{1/6}].$$

The equations that a,b,c,d must satisfy are

$$ac+3bd=1,$$
$$-12ad-12bc+(2\lambda\hbar r)^{1/6}$$
$$\times(-ad'+ac'-3bc'-3da'-ca'-cb')=0,$$
$$-c^2+2cd+a^2+2ab+(2\lambda\hbar r)$$
$$(a^4+3a^3+3a^2b^2+3ab^3)=0,$$
$$-4cd+4ab+(2\lambda\hbar r)(a^4+12a^3b+6a^2b^2+12ab^3)=0.$$

The power series expansions for large and small values of $(2\lambda\hbar r)$ are

$$a=\begin{cases}1-\frac{3}{8}(2\lambda\hbar r)+\frac{75}{128}(2\lambda\hbar r)^2-\frac{3471}{4096}(2\lambda\hbar r)^2+\cdots,\\ (2\lambda\hbar r)^{-1/6}[0.9196+0.007\,756\,(2\lambda\hbar r)^{-4/6}\\ -0.007\,054\,(2\lambda\hbar r)^{-8/6}+\cdots],\end{cases}$$

$$b=\begin{cases}\frac{3}{64}(2\lambda\hbar r)^2+\frac{627}{4096}(2\lambda\hbar r)^2+\cdots,\\ (2\lambda\hbar r)^{-1/6}[0.0131+0.050\,96\,(2\lambda\hbar r)^{-4/6}\\ +0.000\,1602\,(2\lambda\hbar r)^{-8/6}+\cdots],\end{cases}$$

$$c=\begin{cases}-1-\frac{3}{8}(2\lambda\hbar r)+\frac{57}{128}(2\lambda\hbar r)^2-\frac{4911}{4090}(2\lambda\hbar r)^3+\cdots,\\ (2\lambda\hbar r)^{1/6}[-1.0793+0.038\,11\,(2\lambda\hbar r)^{-4/8}\\ +0.008\,459\,(2\lambda\hbar r)^{-8/6}+\cdots],\end{cases}$$

$$d=\begin{cases}-\frac{1}{4}(2\lambda\hbar r)+\frac{54}{128}(2\lambda\hbar r)^2-\frac{2331}{4090}(2\lambda\hbar r)^3+\cdots,\\ (2\lambda\hbar r)^{1/6}[-0.1943+0.077\,15\,(2\lambda\hbar r)^{4/6}\\ +0.001\,101\,(2\lambda\hbar r)^{-8/6}+\cdots].\end{cases}$$

The lowest order equations for the pair of oscillators described by the Hamiltonian

$$H=p_1^2+p_2^2+x_1^2+x_2^2+Ax_1^4+2Bx_1^2x_2^2+Cx_2^4$$

were previously developed. They are

$$16A\xi_1^6(1,0;1,0)+16B\xi_1^4(1,0;1,0)\xi_2^2(0,1;0,1)+4\xi_1^4(1,0;1,0)-n^2=0,$$
$$16B\xi_1^2(1,0;1,0)\xi_2^4(0,1;0,1)+16C\xi_2^6(0,1;0,1)+4\xi_2^4(0,1;0,1)-m^2=0.$$

The substitutions

$$\xi_1(1,0;1,0)=\sqrt{n/2}\ \eta_1,$$
$$\xi_2(0,1;0,1)=\sqrt{m/2}\ \eta_2$$

extract the harmonic dependence and give the equations

$$(2An)\eta_1^6+2Bm\,\eta_1^4\eta_2^2+\eta_1^4-1=0,\qquad(7)$$
$$2Bn\,\eta_1^2\eta_2^4+(2Cm)\eta_2^6+\eta_2^4-1=0.$$

For positive values of B, η_1 and η_2 are slowly varying functions of m and n with $0\le\eta_1,\ \eta_2\le1$. Power series expansions can be developed in the variables $(2An)^{\pm1}$, $(2Bm)^{\pm1}$, $(2Bn)^{\pm1}$, $(2Cm)^{\pm}$ depending on the size of these variables. Rational interpolating functions can also be constructed for intermediate values of the variables. For negative values of B the situation is more complicated since η_1 and η_2 may become infinite.

In the more complicated approximation in which $\xi_1(\mathbf{a},1,0)$, $\xi_1(\mathbf{a},0,1)$, $\xi_2(0,1,1,0)$, and $\xi_2(0,1,0,1)$ and the corresponding π's are different from zero the equations

$$\partial_1[\pi_1(2r;1,0)\xi_1(2r;1,0)]+\partial_2[\pi_1(2r;0,1)\xi_1(2r;0,1)]=-\hbar/4,$$
$$\partial_1[\pi_1(2r;1,0)\xi_2(2r;1,0)]+\partial_2[\pi_1(2r;0,1)\xi_2(2r;0,1)]=0,$$
$$\partial_1[\pi_2(2r;1,0)\xi_2(2r;1,0)]+\partial_2[\pi_2(2r;0,1)\xi_1(2r;0,1)]=0,$$
$$\partial_1[\pi_2(2r;1,0)\xi_2(2r;1,0)]+\partial_2[\pi_2(2r;0,1)\xi_2(2r;0,1)]=-\hbar/4,$$

$$-\pi_1^2(1,0)+\pi_2^2(1,0)+\xi_1^2(1,0)+\xi_2^2(1,0)+4A\xi_1^4(1,0)$$
$$+12A\xi_1^2(1,0)\xi_1^2(0,1)+8B\xi_1^2(1,0)\xi_2^2(1,0)$$
$$+16B\xi_1(1,0)\xi_1(0,1)\xi_2(1,0)\xi_2(0,1)+4B\xi_1^2(1,0)\xi_2^2(0,1)$$
$$+4B\xi_1^2(0,1)\xi_2^2(1,0)+4C\xi_2^4(0,1)+12\xi_2^2(0,1)\xi_2^2(1,0)=0,$$
$$-2\pi_1(1,0)\pi_1(0,1)+2\pi_2(1,0)\pi_2(0,1)+2\xi_1(1,0)\xi_1(0,1)$$
$$+2\xi_2(1,0)\xi_2(0,1)+12A\xi_1^3(1,0)\xi_1(0,1)+12A\xi_1(1,0)\xi_1^3(0,1)$$
$$+12B\xi_1(1,0)\xi_1(0,1)\xi_2^2(1,0)+12B\xi_1(1,0)\xi_1(0,1)\xi_2^2(0,1)$$
$$+12B\xi_1^2(0,1)\xi_2(1,0)\xi_2(0,1)+12C\xi_2^3(1,0)\xi_2(0,1)$$
$$+12C\xi_2(1,0)\xi_2^3(1,0)=0,$$
$$-\pi_1^2(0,1)+\pi_2^2(0,1)+\xi_1^2(0,1)+\xi_2^2(0,1)+4A\xi_1^4(0,1)$$
$$+12A\xi_1^2(1,0)\xi_1^2(0,1)+8B\xi_1^2(0,1)\xi_2^2(0,1)$$
$$+16B\xi_1(1,0)\xi_1(0,1)\xi_2(1,0)\xi_2(0,1)+4B\xi_1^2(1,0)\xi_2^2(0,1)$$
$$+4B\xi_1^2(0,1)\xi_2^2(1,0)+4C\xi_2^4(0,1)+12C\xi_2^2(0,1)\xi_2^2(1,0)$$
$$=0,$$
$$-2\pi_1(1,0)\pi_1(0,1)-2\pi_2(1,0)\pi_2(0,1)+2\xi_1(1,0)\xi_1(0,1)$$
$$+2\xi_2(1,0)\xi_2(0,1)+12A\xi_1^3(1,0)\xi_1(0,1)$$
$$+12A\xi_1(1,0)\xi_1^3(0,1)+12B\xi_1(1,0)\xi_1(0,1)\xi_2^2(1,0)$$
$$+12B\xi_1(1,0)\xi_1(0,1)\xi_2^2(0,1)+12B\xi_1^2(1,0)\xi_2(1,0)\xi_2(0,1)$$
$$+12B\xi_1^2(0,1)\xi_2(1,0)\xi_2(0,1)+12C\xi_2^3(1,0)\xi_2(0,1)$$
$$+12C\xi_2(1,0)\xi_2^3(0,1)=0.$$

These equations are easily solved by setting

$$\pi_1(0,1)=\pi_2(1,0)=\xi_1(0,1)=\xi_2(1,0)=0.$$

The equations then reduce to those of the previous order of approximation. The next order is much more complicated and will be the subject of a subsequent calculation.

V. CONCLUSIONS

The intention of the present methods is to reduce the problems of anharmonic oscillators to sets of algebraic equations. Ideally the general features of the solutions of these equations would be evident from the equations and exact numerical solutions would be feasible with computers. Successive approximation would yield solutions of increasing accuracy.

For the single oscillator the problem is almost in this state. The equations can be written down and the qualitative behavior of the solutions is self-evident and quantitative solutions are accessible. There are improvements that still can be made. The approximations are severe for small values of the quantum numbers, and the ground state and first few excited states should be treated more carefully. The possibility exists of linearizing the second and higher order approximation rather than attempting to solve the nonlinear equations in each order. This would accomplish a great deal in the direction of decoupling the nondiagonal commutator equations.

In the cases of two or more coupled oscillators the new phenomena of the commutators of the type

$$[p_a, p_b], \ [x_a, x_b], \ \text{and} \ [p_a, x_b]$$

for $a \neq b$ occur. The relative numerical size of these matrix elements is still unknown. This information is vital in developing a systematic procedure to go beyond the lowest order reported here. In order to study these numbers, a good solution of (7) is required. This seems to be the main obstacle at this time to bring the solution of the problem of several oscillators to the same level as that of a single oscillator.

[1]W. Heisenberg, Z. Physik 33, 879 (1925).
[2]M. Born and P. Jordan, Z. Physik 34, 858 (1925).
[3]J. von Newmann, *Mathematische Grundlagen der Quantemmechanik* (Springer-Verlag, Berlin, 1932), p. 6.
[4]F. Halpern, J. Math. Phys. 15, 733 (1974); F. Halpern and T.W. Yonkman, J. Math. Phys. 15, 1718 (1974).
[5]C. Bender and T.T. Wu, Phys. Rev. 184, 1231 (1969).
[6]J.J. Loeffel, A. Martin, B. Simon, and A.S. Wightman, Phys. Lett. B 30, 656 (1969).
[7]S.N. Biswas, K. Datta, R.P. Saxena, R.K. Srivastava, and V.S. Varma, Phys. Rev. D 4, 3617 (1971).

Separation of variables in the Hamilton–Jacobi, Schrödinger, and related equations. II. Partial separation*

Peter Havas†

Department of Physics, Temple University, Philadelphia, Pennsylvania 19122
(Received 24 July 1975)

Sufficient conditions are given for coordinate systems in which the Hamilton–Jacobi equation and the Schrödinger and related equations are partially separable in n dimensions. For the first equation, the solution is assumed to be a sum of $\nu+\tau$ functions of a single variable, and of μ $[0\leq\mu\leq(n-\nu-\tau)/2]$ groups of other variables; for the other equations, products of such functions are assumed. These assumptions lead to $\nu+\tau$ completely separated differential equations, of which ν are linear in the separation constants, and μ partially separated equations depending on the remaining $n-\nu-\tau$ variables. The general forms of the various metric tensors g^{kl} of the Riemannian spaces V_n as well as of the allowed potentials V corresponding to the different possible types of such equations are determined; they are identical for the Hamilton–Jacobi equation and for the other equations studied, except that for the latter some of the metrics are further restricted by a condition on their determinants. The results are established by methods similar to those used in Paper I of this series for complete separation, and include the results obtained there as special cases. In the course of determining the allowed forms of the g^{kl} it is also established that there exist $\nu+\tau+\mu$ independent first integrals linear and quadratic in the momenta for the dynamical systems described by the Hamilton–Jacobi or the Schrödinger equation. The ν linear ones are homogeneous, and the $\tau+\mu$ quadratic ones correspond to homogeneous quadratic integrals of the geodesics of the V_n. These results imply the existence of ν Killing vectors and of $\tau+\mu$ Killing tensors of rank two for the V_n. Further polynomial integrals can be constructed; those integrals of degree r which are independent of the original $\nu+\tau+\mu$ integrals each correspond to an independent Killing tensor of rank r.

I. INTRODUCTION

In spite of the importance of the method of separation of variables for the solution of partial differential equations (PDE's), there are very few systematic attempts in the literature to determine the coordinate systems which permit the solution of a particular PDE by this method. The only PDE for which all coordinate systems, i.e., all forms of the metric tensor of a Riemannian space V_n, as well as all forms of the potential V have been determined which allow complete separation is the Hamilton–Jacobi (H–J) equation[1,2]; the results were reviewed and elaborated on in a recent paper[3] (in the following referred to as I). This paper also gave sufficient conditions for coordinate systems in which the Schrödinger, Helmholtz, and Laplace equation are completely separable. It was shown that in n dimensions there are $n+1$ types of such systems, n of which are in general nonorthogonal. These systems correspond to those for which the H–J equation is separable, but except for the "essentially geodesic case" of Levi-Civita[4] (for which V must vanish) they are more restricted due to a condition on the determinant of the metric.

If studies of conditions for complete separation are scarce, those of conditions where separation is only partial are almost nonexistent. Of main concern to this paper is the work of Paul Stäckel on the H–J equation[5] generalizing this author's results for complete separation in orthogonal coordinate systems. There are also a few recent studies of the Schrödinger, Laplace, Helmholtz, and related equations,[6–8] which also contain some references to earlier literature.

In Sec. II of this paper we develop sufficient conditions for coordinate systems which allow partial separation (in a sense to be specified later) of the H–J equation, which include the conditions of Burgatti[1] and Dall'Acqua[2] for complete separation as special cases. Corresponding conditions for the Schrödinger, Helmholtz, and Laplace equation, which include the conditions given in I for complete separation as special cases, are found in Sec. IV. The methods used are closely analogous to those used in I; the main complication is notational. As in the case of complete separation, it is found that the Schrödinger and related equations allow the same metric tensors and potentials as the H–J equation, apart from a condition on the determinant of the metric and certain differentiability conditions.

In the course of determining the allowed forms of the metric tensor it is also established that there exist a number of independent first integrals linear and quadratic in the momenta for the dynamical system described by the H–J equation. In Sec. III these integrals are determined, further polynomial integrals are constructed via Poisson's theorem, and these integrals are related to the homogeneous polynomial integrals of the geodesics of the V_n, as well as to the Killing tensors admitted by the V_n. Analogous results are established in Sec. V for the polynomial constants of the motion of the Schrödinger equation. The results of this paper are briefly summarized and discussed in Sec. VI.

II. THE HAMILTON–JACOBI EQUATION

The H–J equation (after separation of the time-dependent part) in a V_n with coordinates q_i $(i=1\cdots n)$ and metric g_{kl} is

$$H(q_1\cdots q_n, p_1\cdots p_n)=E,$$

$$H\equiv T(q_1\cdots q_n, p_1\cdots p_n)+V(q_1\cdots q_n), \tag{1}$$

where

$$T \equiv \tfrac{1}{2}\sum_{k,l=1}^{n} g^{kl}p_k p_l,$$

$$\sum_{m=1}^{n} g_{km}g^{lm} = \delta_k{}^l, \tag{2}$$

$$p_i \equiv \frac{\partial W(q_1 \cdots q_n)}{\partial q_i}, \quad i = 1 \cdots n.$$

In the following it will be understood that the range of the Latin indices is from 1 to n, but for our purposes it will be preferable always to indicate summation explicitly rather than to use a summation convention; similarly, since in this paper we are concerned with special coordinate systems rather than tensorial relations valid in all coordinate systems, no tensorial properties will be implied by the position of any index, except for the co- and contravariant metric tensors g_{kl} and g^{kl}.

Complete separation of Eq. (1) requires

$$W(q_1 \cdots q_n) = \sum_{i=1}^{n} W_i(q_i). \tag{3}$$

We shall be concerned with partially separated solutions of the form

$$W(q_1 \cdots q_n) = \sum_{\alpha=1}^{\nu} W_\alpha(q_\alpha) + \sum_{\rho=\nu+1}^{\nu+\tau} W_\rho(q_\rho) + \sum_{\chi=1}^{\mu} W_\chi(q_{\chi 1} \cdots q_{\chi h_\chi}). \tag{4}$$

Here we have divided the variables and the corresponding indices into three groups

$$q_\alpha, \quad \alpha = 1 \cdots \nu, \tag{5a}$$

$$q_\rho, \quad \rho = \nu+1 \cdots \nu+\tau, \tag{5b}$$

$$q_{\chi\psi}, \quad \chi = 1 \cdots \mu, \quad \psi = 1 \cdots h_\chi,$$

$$0 \leq \mu \leq \tfrac{1}{2}(n-\nu-\tau), \quad h_\chi \geq 2, \tag{5c}$$

$$\sum_{\chi=1}^{\mu} h_\chi = n-\nu-\tau,$$

to be called variables and indices of the first, second, and third kind, respectively. Those of the first and second kind are the same as were used in I (following Refs. 4, 1, and 2); those of the third kind were first introduced by Stäckel. It should be noted that the variables of the third kind are each characterized by two numbers, the first one denoting the particular group of variables, and the second the variables within the group; to avoid confusion, these will be bracketed when they appear in the metric tensor or in the p's.

In the following, letters of the Greek alphabet up to and including ν will be used for indices of the first kind, those between ν and τ for the second kind, and those beyond τ for the third kind, except that the letter μ is reserved for the number of groups of variables of the third kind, and the index η can range over variables of the second kind and the index denoting the particular group of variables of the third kind. Latin indices will be used if the entire range $1 \cdots n$ is covered. The range of any summation will be indicated by $\sum^{\mathrm{I}}$ for variables of the first kind, by $\sum^{\mathrm{II}}$ for variables of the second kind, by $\sum^{\mathrm{III}'}$ for those variables of the third kind which denote the groups of variables, and by $\sum^{\mathrm{III}''}$ for the ones denoting the variables within the group.

Any one or two kinds of variables may be absent. If the first kind is absent, we define $\nu = 0$. This is the case discussed by Stäckel.[5] There, no explicit distinction was made between variables of the second and third kind, but in the presence of variables of the first kind this distinction is essential; unlike Stäckel, we therefore do not allow h_χ to take the value 1.

If the variables of the second kind are absent, we define $\tau = 0$. If those of the third kind are absent, we recover the case of complete separation discussed in I. Clearly we will obtain results which are not contained in I only if $n \geq 3$.

The division of the coordinates given by (5a–c) is fully characterized by the numbers $\nu, \tau, h_1 \cdots h_\mu$. The corresponding metrics and partially separated equations will be referred to as of type $T_{\nu,\tau,h_1 \cdots h_\mu}$; they are generalizations of those of type T_ν discussed in I.

If both the first and second kind of variables are absent, no variable is completely separated, and if in addition χ takes on only the value 1 (and thus $h_1 = n$), there is no separation at all.

Corresponding to the coordinates of the first kind, and of those of the second and third kind combined, we introduce two sets of arbitrary continuous functions, one set

$$\varphi_{\alpha\beta}(q_\alpha), \quad \alpha, \beta = 1 \cdots \nu, \tag{6a}$$

where each function depends on a single variable of the first kind only, and with determinant ϕ_{I}, and another

$$\varphi_{\rho\eta}(q_\rho), \quad \rho = \nu+1 \cdots \nu+\tau, \quad \eta = 1 \cdots \tau+\mu, \tag{6b}$$

$$\varphi_{\chi\eta}(q_{\chi 1} \cdots q_{\chi h_\chi}), \quad \chi = 1 \cdots \mu, \quad \eta = 1 \cdots \tau+\mu,$$

depending on the variables of the second and third kind as indicated, and with determinant ϕ_{II}. ϕ_{I} and ϕ_{II} are assumed not to vanish. The cofactors of $\varphi_{\alpha\beta}$, $\varphi_{\rho\eta}$, and $\varphi_{\chi\eta}$ in their respective determinants are denoted by $\phi_{\alpha\beta}$, $\phi_{\rho\eta}$, and $\phi_{\chi\eta}$, respectively; note that they do *not* depend on q_α, q_ρ, and the $q_{\chi\psi}$ respectively. If one or both of the determinants contain only a single element φ_{ii}, we will define $\phi_{ii} \equiv 1$. We will frequently use the well-known relations

$$\sum_{\alpha}^{\mathrm{I}} \varphi_{\alpha\beta}\phi_{\alpha\kappa} = \sum_{\alpha}^{\mathrm{I}} \varphi_{\beta\alpha}\phi_{\kappa\alpha} = \phi_{\mathrm{I}}\delta_\beta{}^\kappa, \tag{7a}$$

$$\sum_{\rho}^{\mathrm{II}} \varphi_{\rho\eta}\phi_{\rho\eta'} = \sum_{\chi}^{\mathrm{III}'} \varphi_{\chi\eta}\phi_{\chi\eta'} = \phi_{\mathrm{II}}\delta_\eta{}^{\eta'},$$

$$\sum_{\eta}^{\mathrm{II+III}'} \varphi_{\rho\eta}\phi_{\sigma\eta} = \phi_{\mathrm{II}}\delta_\rho{}^\sigma, \quad \sum_{\eta}^{\mathrm{II+III}'} \varphi_{\chi\eta}\phi_{\chi'\eta} = \phi_{\mathrm{II}}\delta_\chi{}^{\chi'}, \tag{7b}$$

without explicitly referring to them.

We also introduce six other sets of functions

$$f_\rho{}^\kappa(q_\rho), \quad F_\rho{}^{\kappa\lambda}(q_\rho) = F_\rho{}^{\lambda\kappa}, \quad u_\rho(q_\rho),$$

$$\rho = \nu+1 \cdots \nu+\tau, \quad \kappa, \lambda = 1 \cdots \nu, \tag{8a}$$

and

$$A_\chi{}^{\varphi\psi}(q_{\chi 1}\cdots q_{\chi h_\chi}) = A_\chi{}^{\psi\varphi},$$

$$G_\chi{}^{\kappa\lambda}(q_{\chi 1}\cdots q_{\chi h_\chi}) = G_\chi{}^{\lambda\kappa}, \quad u_\chi(q_{\chi 1}\cdots q_{\chi h_\chi}),$$

$$\chi = 1\cdots\mu, \quad \varphi, \psi = 1\ldots h_\chi, \quad \kappa, \lambda = 1\cdots\nu, \tag{8b}$$

all of which except the u_ρ's and u_χ's must be continuous, and three sets of arbitrary constants ("separation constants")

$$c_\alpha, \quad \alpha = 1\ldots\nu, \tag{9a}$$

$$c_\rho, \quad \rho = \nu+1\cdots\nu+\tau, \tag{9b}$$

$$c_\chi, \quad \chi = 1\cdots\mu. \tag{9c}$$

For any χ, the determinant of $A_\chi{}^{\varphi\psi}$ is assumed not to vanish. We now consider the three sets of differential equations

$$\frac{dW_\alpha}{dq_\alpha} = \sum_\beta{}^{\mathrm{I}}\varphi_{\alpha\beta}(q_\alpha)c_\beta, \quad \alpha = 1\cdots\nu, \tag{10a}$$

$$\frac{dW_\rho}{dq_\rho} = \sum_\beta{}^{\mathrm{I}} f_\rho{}^\beta(q_\rho)c_\beta \pm \sum_{\kappa,\lambda}{}^{\mathrm{I}} F_\rho{}^{\kappa\lambda}(q_\rho)c_\kappa c_\lambda + \sum_\eta{}^{\mathrm{II+III'}}\varphi_{\rho\eta}(q_\rho)c_\eta - 2u_\rho(q_\rho)^{1/2}, \quad \rho = \nu+1\cdots\nu+\tau, \tag{10b}$$

$$\sum_{\varphi,\psi}{}^{\mathrm{III''}} A_\chi{}^{\varphi\psi}\frac{\partial W_\chi}{\partial q_{\chi\varphi}}\frac{\partial W_\chi}{\partial q_{\chi\psi}} = \sum_{\kappa,\lambda}{}^{\mathrm{I}} G_\chi{}^{\kappa\lambda}c_\kappa c_\lambda + \sum_\eta{}^{\mathrm{II+III'}}\varphi_{\chi\eta}c_\eta - 2u_\chi, \quad \chi = 1\cdots\mu, \tag{10c}$$

and shall derive the forms of g^{kl} and V following from these equations by eliminating the separation constants from them and comparing the resulting equation with Eqs. (1), (2).

Multiplying the set (10a) by $\phi_{\alpha\kappa}/\phi_{\mathrm{I}}$ and summing over α, we obtain

$$\sum_\alpha{}^{\mathrm{I}}\frac{\phi_{\alpha\kappa}}{\phi_{\mathrm{I}}}\frac{dW_\alpha}{dq_\alpha} = c_\kappa. \tag{11a}$$

Substituting this into (10b), rearranging terms, and squaring, we get

$$\left(\frac{dW_\rho}{dq_\rho} - \sum_{\alpha,\beta}{}^{\mathrm{I}}\frac{f_\rho{}^\beta\phi_{\alpha\beta}}{\phi_{\mathrm{I}}}\frac{dW_\alpha}{dq_\alpha}\right)^2 = \sum_{\alpha,\beta,\kappa,\lambda}{}^{\mathrm{I}}\frac{F_\rho{}^{\kappa\lambda}\phi_{\alpha\kappa}\phi_{\beta\lambda}}{\phi_{\mathrm{I}}{}^2}\frac{dW_\alpha}{dq_\alpha}\frac{dW_\beta}{dq_\beta} + \sum_\eta{}^{\mathrm{II+III'}}\varphi_{\rho\eta}c_\eta - 2u_\rho. \tag{11b}$$

Similarly, substituting (11a) into (10c), we get

$$\sum_{\varphi,\psi}{}^{\mathrm{III''}} A_\chi{}^{\varphi\psi}\frac{\partial W_\chi}{\partial q_{\chi\varphi}}\frac{\partial W_\chi}{\partial q_{\chi\psi}} = \sum_{\kappa,\lambda}{}^{\mathrm{I}} G_\chi{}^{\kappa\lambda}\frac{\phi_{\alpha\kappa}\phi_{\beta\lambda}}{\phi_{\mathrm{I}}{}^2}\frac{dW_\alpha}{dq_\alpha}\frac{dW_\beta}{dq_\beta} - 2u_\chi. \tag{11c}$$

Squaring (11a) and changing dummy indices, we obtain

$$\sum_{\alpha,\beta}{}^{\mathrm{I}}\frac{\phi_{\alpha\gamma}\phi_{\beta\gamma}}{\phi_{\mathrm{I}}{}^2}\frac{dW_\alpha}{dq_\alpha}\frac{dW_\beta}{dq_\beta} = c_\gamma{}^2. \tag{12a}$$

We now multiply (11b) by $\phi_{\rho\eta}/\phi_{\mathrm{II}}$ and sum over ρ; similarly we multiply (11c) by $\phi_{\chi\eta}/\phi_{\mathrm{II}}$ and sum over χ. Adding the resulting equations and rearranging, we get

$$\sum_\rho{}^{\mathrm{II}}\frac{\phi_{\rho\eta}}{\phi_{\mathrm{I}}}\left(\frac{dW_\rho}{dq_\rho} - \sum_{\alpha,\beta}{}^{\mathrm{I}}\frac{f_\rho{}^\beta\phi_{\alpha\beta}}{\phi_{\mathrm{I}}}\frac{dW_\alpha}{dq_\alpha}\right)^2 + \sum_\chi{}^{\mathrm{III'}}\sum_{\varphi,\psi}{}^{\mathrm{III''}}\frac{\phi_{\chi\eta}}{\phi_{\mathrm{II}}}A_\chi{}^{\varphi\psi}\frac{\partial W_\chi}{\partial q_{\chi\varphi}}\frac{\partial W_\chi}{\partial q_{\chi\psi}} - \sum_{\alpha,\beta,\kappa,\lambda}{}^{\mathrm{I}}\left[\sum_\rho{}^{\mathrm{II}}F_\rho{}^{\kappa\lambda}\phi_{\rho\eta} + \sum_\chi{}^{\mathrm{III'}}G_\chi{}^{\kappa\lambda}\phi_{\chi\eta}\right] \times\frac{\phi_{\alpha\kappa}\phi_{\beta\lambda}}{\phi_{\mathrm{I}}{}^2\phi_{\mathrm{II}}}\frac{dW_\alpha}{dq_\alpha}\frac{dW_\beta}{dq_\beta} + \sum_\rho{}^{\mathrm{II}}\frac{2u_\rho\phi_{\rho\eta}}{\phi_{\mathrm{II}}} + \sum_\chi{}^{\mathrm{III'}}\frac{2u_\chi\phi_{\chi\eta}}{\phi_{\mathrm{II}}} = c_\eta. \tag{12b}$$

We now sum Eqs. (12a) and (12b) over γ and η, respectively, add the resulting equations, and put

$$\sum_\gamma{}^{\mathrm{I}}c_\gamma{}^2 + \sum_\eta{}^{\mathrm{II+III'}}c_\eta = 2E. \tag{13}$$

Then we obtain

$$\sum_{\alpha,\beta,\gamma}{}^{\mathrm{I}}\frac{\phi_{\alpha\gamma}\phi_{\beta\gamma}}{\phi_{\mathrm{I}}{}^2}\frac{dW_\alpha}{dq_\alpha}\frac{dW_\beta}{dq_\beta} + \sum_\rho{}^{\mathrm{II}}\sum_\eta{}^{\mathrm{II+III'}}\frac{\phi_{\rho\eta}}{\phi_{\mathrm{II}}}\left(\frac{dW_\rho}{dq_\rho} - \sum_{\alpha,\beta}{}^{\mathrm{I}}\frac{f_\rho{}^\beta\phi_{\alpha\beta}}{\phi_{\mathrm{I}}}\frac{dW_\alpha}{dq_\alpha}\right)^2 + \sum_\chi{}^{\mathrm{III'}}\sum_\eta{}^{\mathrm{II+III'}}\sum_{\varphi,\psi}{}^{\mathrm{III''}}\frac{\phi_{\chi\eta}}{\phi_{\mathrm{II}}}A_\chi{}^{\varphi\psi}\frac{\partial W_\chi}{\partial q_{\chi\varphi}}\frac{\partial W_\chi}{\partial q_{\chi\psi}} - \sum_\eta{}^{\mathrm{II+III'}}\sum_{\alpha,\beta,\kappa,\lambda}{}^{\mathrm{I}}\left[\sum_\rho{}^{\mathrm{II}}F_\rho{}^{\kappa\lambda}\phi_{\rho\eta} + \sum_\chi{}^{\mathrm{III'}}G_\chi{}^{\kappa\lambda}\phi_{\chi\eta}\right] \times\frac{\phi_{\alpha\kappa}\phi_{\beta\lambda}}{\phi_{\mathrm{I}}{}^2\phi_{\mathrm{II}}}\frac{dW_\alpha}{dq_\alpha}\frac{dW_\beta}{dq_\beta} + 2\sum_\eta{}^{\mathrm{II+III'}}\left[\sum_\rho{}^{\mathrm{II}}\frac{u_\rho\phi_{\rho\eta}}{\phi_{\mathrm{II}}} + \sum_\chi{}^{\mathrm{III'}}\frac{u_\chi\phi_{\chi\eta}}{\phi_{\mathrm{II}}}\right] = 2E, \tag{14}$$

which is indeed of the form (1) with (2).

We shall first consider the case where the variables of the second kind are absent, so that $\tau = 0$, i.e., the type $T_{\nu,0,h_1\cdots h_\mu}$. Then Eq. (14) reduces to

$$\sum_{\alpha,\beta}{}^{\mathrm{I}}\left[\sum_\gamma{}^{\mathrm{I}}\frac{\phi_{\alpha\gamma}\phi_{\beta\gamma}}{\phi_{\mathrm{I}}{}^2} - \sum_{\chi,\eta}{}^{\mathrm{III'}}\sum_{\kappa,\lambda}{}^{\mathrm{I}}G_\chi{}^{\kappa\lambda}\phi_{\chi\eta}\frac{\phi_{\alpha\kappa}\phi_{\beta\lambda}}{\phi_{\mathrm{I}}{}^2\phi_{\mathrm{II}}}\right]\frac{dW_\alpha}{dq_\alpha}\frac{dW_\beta}{dq_\beta} + \sum_{\chi,\eta}{}^{\mathrm{III'}}\sum_{\varphi,\psi}{}^{\mathrm{III''}}\frac{\phi_{\chi\eta}}{\phi_{\mathrm{II}}}A_\chi{}^{\varphi\psi}\frac{\partial W_\chi}{\partial q_{\chi\varphi}}\frac{\partial W_\chi}{\partial q_{\chi\psi}} + 2\sum_{\chi,\eta}{}^{\mathrm{III'}}\frac{u_\chi\phi_{\chi\eta}}{\phi_{\mathrm{II}}} = 2E. \tag{15}$$

The form of g^{kl} and of V follows from comparison with Eqs. (1) and (2). Since the metric tensor must be continuous, all arbitrary functions entering it must be continuous. Furthermore, since g must not be allowed to vanish, it follows from the form of $g^{\{\chi\varphi\}\{\chi\psi\}}$ that for each χ we must require $\det A_\chi{}^{\varphi\psi} \neq 0$. The form of the W_α follows from integration of Eq. (10a); because of Eq. (2), constants of integration can be omitted.

Summarizing our results we thus have

Theorem I: The Hamilton–Jacobi equation (1), (2) can be solved by partial or complete separation of variables in any V_n ($n \geq 1$) whose contravariant metric ten-

sor is of the form

$$g^{\alpha\beta}=\sum_{\kappa,\lambda=1}^{\nu}\left[\delta_{\kappa\lambda}-\sum_{\chi,\eta=1}^{\mu}\frac{G_\chi^{\kappa\lambda}\phi_{\chi\eta}}{\phi_{II}}\right]\frac{\phi_{\alpha\kappa}\phi_{\beta\lambda}}{\phi_I^2},$$

$$g^{\alpha\eta}=0,$$

$$g^{(\chi\varphi)(\chi'\psi)}=0,\quad \chi\neq\chi',$$

$$g^{(\chi\varphi)(\chi\psi)}=\left(\sum_{\eta=1}^{\mu}\frac{\phi_{\chi\eta}}{\phi_{II}}\right)A_\chi^{\varphi\psi},$$

$$\alpha,\beta=1\cdots\nu,\qquad 0<\nu\leq n\quad(\nu\neq n-1),$$

$$\chi,\chi',\eta=1\cdots\mu,\qquad 0\leq\mu<n-\nu,$$

$$\varphi,\psi=1\cdots h_\chi,\qquad h_\chi\geq 2,$$

$$\sum_{\chi=1}^{\mu}h_\chi=n-\nu,$$

where $G_\chi^{\kappa\lambda}(q_{\chi1}\cdots q_{\chi h_\chi})=G_\chi^{\lambda\kappa}$ and $A_\chi^{\varphi\psi}(q_{\chi1}\cdots q_{\chi h_\chi})=A_\chi^{\psi\varphi}$ (with $\det A_\chi^{\psi\varphi}\neq 0$) are arbitrary continuous real functions of groups of h_χ variables each, ϕ_I and ϕ_{II} are the determinants ($\neq 0$) of two sets of arbitrary continuous real functions $\varphi_{\alpha\beta}(q_\alpha)$ $(\alpha,\beta=1\cdots\nu)$ and $\varphi_{\chi\eta}(q_{\chi1}\cdots q_{\chi h_\chi})$ $(\chi,\eta=1\cdots\mu)$ respectively, and $\phi_{\alpha\beta}$ and $\phi_{\chi\eta}$ are the cofactors of $\varphi_{\alpha\beta}$ and $\varphi_{\chi\eta}$ in these determinants (with $\phi_{ii}\equiv 1$ if one or both determinants consist of a single element φ_{ii} only); the potential energy V must be of the form

$$V(q_{11}\cdots q_{1h_1},\cdots q_{\mu1}\cdots q_{\mu h_\mu})=\sum_{\chi,\eta=1}^{\mu}\frac{u_\chi\phi_{\chi\eta}}{\phi_{II}},$$

where the $u_\chi(q_{\chi1}\cdots q_{\chi h_\chi})$ are arbitrary real functions of groups of h_χ variables each. The solution is of the form

$$W=\sum_{\alpha,\beta=1}^{\nu}c_\beta\int\varphi_{\alpha\beta}(q_\alpha)\,dq_\alpha+\sum_{\chi=1}^{\mu}W_\chi(q_{\chi1}\cdots q_{\chi h_\chi}),$$

where the W_χ are solutions of

$$\sum_{\varphi,\psi=1}^{\mu}A_\chi^{\varphi\psi}\frac{\partial W_\chi}{\partial q_{\chi\varphi}}\frac{\partial W_\chi}{\partial q_{\chi\psi}}=\sum_{\kappa,\lambda=1}^{\nu}G_\chi^{\kappa\lambda}c_\kappa c_\lambda+\sum_{\eta=1}^{\mu}\varphi_{\chi\eta}c_\eta-2u_\chi,$$

and the c_i $(i=1\cdots n)$ are arbitrary real constants, subject to the condition

$$\sum_{\gamma=1}^{\nu}c_\gamma^{\,2}+\sum_{\eta=1}^{\mu}c_\eta=2E.$$

If $\nu=n$, the separation is complete; if $\nu\neq n$, n must be ≥ 3.

Clearly, if $\nu=n$ $(n\geq 1)$, we only have variables of the first kind, and $g^{\alpha\beta}$ reduces to the form (20) of I, from which we can obtain the form (21) of I for $g_{\alpha\beta}$, in agreement with Theorem I of that paper.

Now we consider the case where the variables of the first kind are absent, so that $\nu=0$, i.e., the type $T_{0,\tau,h_1\cdots h_\mu}$. Then Eqs. (13) and (14) reduce to

$$\sum_{\eta}^{II+III'}c_\eta=2E \tag{16}$$

and

$$\sum_{\rho}^{II}\sum_{\eta}^{II+III'}\frac{\phi_{\rho\eta}}{\phi_{II}}\left(\frac{dW_\rho}{dq_\rho}\right)^2+\sum_{\chi}^{III'}\sum_{\eta}^{II+III'}\sum_{\varphi,\psi}^{III''}\frac{\phi_{\chi\eta}}{\phi_{II}}A_\chi^{\varphi\psi}\frac{\partial W_\chi}{\partial q_{\chi\varphi}}\frac{\partial W_\chi}{\partial q_{\chi\psi}}+2\sum_{\eta}^{II+III'}\left[\sum_{\rho}^{II}\frac{u_\rho\phi_{\rho\eta}}{\phi_{II}}+\sum_{\chi}^{III'}\frac{u_\chi\phi_{\chi\eta}}{\phi_{II}}\right]=2E. \tag{17}$$

The form of g^{kl} and of V follows as before. The form of the W_ρ follows from integration of Eq. (10b); the constants of integration can be omitted as before. Thus we have

Theorem II: The Hamilton–Jacobi equation (1), (2) can be solved by partial or complete separation of variables in any V_n $(n\geq 1)$ whose contravariant metric tensor is of the form

$$g^{\rho\rho}=\sum_{\eta=1}^{\mu}\frac{\phi_{\rho\eta}}{\phi_{II}},$$

$$g^{\rho\sigma}=0,\quad \rho\neq\sigma,$$

$$g^{(\chi\varphi)(\chi'\psi)}=0,\quad \chi\neq\chi',$$

$$g^{(\chi\varphi)(\chi\psi)}=\left(\sum_{\eta=1}^{\mu}\frac{\phi_{\chi\eta}}{\phi_{II}}\right)A_\chi^{\varphi\psi},$$

$$\rho,\sigma=1\cdots\tau,\quad 0<\tau\leq n\quad(\tau\neq n-1),$$

$$\left.\begin{aligned}&\chi,\chi'=1\cdots\mu\\&\eta=1\cdots\tau+\mu\end{aligned}\right\},\quad 0\leq\mu<n-\tau,$$

$$\varphi,\psi=1\cdots h_\chi,\quad h_\chi\geq 2,$$

$$\sum_{\chi=1}^{\mu}h_\chi=n-\tau,$$

where $A_\chi^{\varphi\psi}$ is defined as in Theorem I, ϕ_{II} is the determinant ($\neq 0$) of a set of arbitrary continuous real functions $\varphi_{\rho\eta}(q_\rho)$, $\varphi_{\chi\eta}(q_{\chi1}\cdots q_{\chi h_\chi})$ $(\rho=1\cdots\tau,\ \chi=1\cdots\mu,$ $\eta=1\cdots\tau+\mu)$, $\phi_{\rho\eta}$ and $\phi_{\chi\eta}$ are the cofactors of $\varphi_{\rho\eta}$ and $\varphi_{\chi\eta}$ in this determinant; the potential energy V must be of the form

$$V(q_1\cdots q_\tau,\ q_{11}\cdots q_{1h_1}\cdots q_{\mu1}\cdots q_{\mu h_\mu})=\sum_{\eta=1}^{\tau+\mu}\left[\sum_{\rho=1}^{\tau}\frac{u_\rho\phi_{\rho\eta}}{\phi_{II}}+\sum_{\chi=1}^{\mu}\frac{u_\chi\phi_{\chi\eta}}{\phi_{II}}\right],$$

where the $u_\rho(q_\rho)$ are arbitrary real functions of a single variable each, and the u_χ are defined as in Theorem I. The solution is of the form

$$W=\pm\sum_{\rho=1}^{\tau}\int\left(\sum_{\eta=1}^{\tau+\mu}\varphi_{\rho\eta}(q_\rho)c_\eta-2u_\rho(q_\rho)\right)^{1/2}dq_\rho+\sum_{\chi=1}^{\mu}W_\chi(q_{\chi1}\cdots q_{\chi h_\chi}),$$

where the W_χ are solutions of

$$\sum_{\varphi,\psi=1}^{h_\chi}A_\chi^{\varphi\psi}\frac{\partial W_\chi}{\partial q_{\chi\varphi}}\frac{\partial W_\chi}{\partial q_{\chi\psi}}=\sum_{\eta=1}^{\tau+\mu}\varphi_{\chi\eta}c_\eta-2u_\chi,$$

and the c_η $(\eta=1\cdots\tau+\mu)$ are arbitrary real constants, subject to the condition

$$\sum_{\eta=1}^{\tau+\mu}c_\eta=2E.$$

If $\tau=n$, the separation is complete; if $\tau\neq n$, n must be ≥ 3.

Apart from notation, this theorem agrees with the results of Stäckel.[5] If $\tau=n$ $(n\geq 1)$, we only have variables of the second kind, and the theorem reduces to a special case of Theorem II of I.

We also note that

$$g^{(\chi\varphi)(\chi\psi)}=0 \quad \text{for} \quad \varphi\neq\psi$$
$$\text{if} \quad A_\chi{}^{\varphi\psi}=A_\chi{}^{\psi\varphi}=0 \quad \text{for} \quad \varphi\neq\psi, \tag{18}$$

and thus for this special choice of arbitrary functions Theorem II yields orthogonal coordinate systems. This also holds for Theorem I if in addition to the simple choice (18) for the $A_\chi{}^{\varphi\psi}$'s the other arbitrary functions are such that

$$\sum_{\kappa,\lambda}^{\mathrm{I}}\left[\delta_{\kappa\lambda}-\sum_{\chi,\eta}^{\mathrm{III}'}\frac{G_\chi{}^{\kappa\lambda}\phi_{\chi\eta}}{\phi_{\mathrm{II}}}\right]\frac{\phi_{\alpha\kappa}\phi_{\beta\lambda}}{\phi_{\mathrm{I}}{}^2}=0, \tag{19}$$

a condition which is satisfied only for very special relations between the various functions.

Now we consider the general case. Proceeding as in the two special cases above, we readily obtain from Eq. (14)

Theorem III: The Hamilton–Jacobi equation (1), (2) can be solved by partial or complete separation of variables in any V_n $(n\geq 1)$ whose contravariant metric tensor is of the form

$$g^{\alpha\beta}=\sum_{\kappa,\lambda=1}^{\nu}\left[\delta_{\kappa\lambda}-\sum_{\eta=1}^{\tau+\mu}\left(\sum_{\rho=\nu+1}^{\nu+\tau}\frac{F_\rho{}^{\kappa\lambda}\phi_{\rho\eta}}{\phi_{\mathrm{II}}}+\sum_{\chi=1}^{\mu}\frac{G_\chi{}^{\kappa\lambda}\phi_{\chi\eta}}{\phi_{\mathrm{II}}}\right)\right]\times\frac{\phi_{\alpha\kappa}\phi_{\beta\lambda}}{\phi_{\mathrm{I}}{}^2}+\delta^{\alpha\beta}\sum_{\rho=\nu+1}^{\nu+\tau}\frac{\phi_{\rho\eta}}{\phi_{\mathrm{II}}}\left(\sum_{\gamma=1}^{\nu}\frac{f_\rho{}^\gamma\phi_{\alpha\gamma}}{\phi_{\mathrm{I}}}\right)^2,$$

$$g^{\alpha\rho}=-2\sum_{\beta=1}^{\nu}\sum_{\eta=1}^{\tau+\mu}\frac{f_\rho{}^\beta\phi_{\alpha\beta}\phi_{\rho\eta}}{\phi_{\mathrm{I}}\phi_{\mathrm{II}}},$$

$$g^{\alpha(\chi\varphi)}=0,$$

$$g^{\rho\sigma}=0, \quad \rho\neq\sigma,$$

$$g^{\rho\rho}=\sum_{\eta=1}^{\tau+\mu}\frac{\phi_{\rho\eta}}{\phi_{\mathrm{II}}},$$

$$g^{\rho(\chi\varphi)}=0,$$

$$g^{(\chi\varphi)(\chi'\psi)}=0, \quad \chi\neq\chi',$$

$$g^{(\chi\varphi)(\chi\psi)}=\left(\sum_{\eta=1}^{\tau+\mu}\frac{\phi_{\chi\eta}}{\phi_{\mathrm{II}}}\right)A_\chi{}^{\varphi\psi},$$

$$\alpha,\beta=1\cdots\nu,$$

$$\rho,\sigma=\nu+1\cdots\nu+\tau,$$

$$\left.\begin{aligned}&\chi,\chi'=1\cdots\mu\\&\eta=1\cdots\tau+\mu\end{aligned}\right\}, \quad 0\leq\mu<n-\nu-\tau,$$

$$\varphi,\psi=1\cdots h_\chi, \quad h_\chi\geq 2,$$

$$\sum_{\chi=1}^{\mu}h_\chi=n-\nu-\tau,$$

where $G_\chi{}^{\kappa\lambda}=G_\chi{}^{\lambda\kappa}$ and $A_\chi{}^{\varphi\psi}=A_\chi{}^{\psi\varphi}$ are defined as in Theorem I, $F_\rho{}^{\kappa\lambda}(q_\rho)=F_\rho{}^{\lambda\kappa}$ and $f_\rho{}^\alpha(q_\rho)$ are arbitrary continuous real functions of a single variable each, ϕ_{I} and ϕ_{II} are the determinants $(\neq 0)$ found from a set of arbitrary continuous real functions $\varphi_{\alpha\beta}(q_\alpha)(\alpha,\beta=1\cdots\nu)$ and a similar set of functions $\varphi_{\rho\eta}(q_\rho)$ $(\rho=\nu+1\cdots\nu+\tau)$ and $\varphi_{\chi\eta}(q_{\chi1}\cdots q_{\chi h_\chi})(\chi=1\cdots\mu,\ \eta=1\cdots\tau+\mu)$ respectively, and $\phi_{\alpha\beta}$, $\phi_{\rho\eta}$, and $\phi_{\chi\eta}$ are the cofactors of $\varphi_{\alpha\beta}$, $\varphi_{\rho\eta}$, and $\varphi_{\chi\eta}$ in these determinants (with $\phi_{ii}=1$ if one or both determinants consists of a single element φ_{ii} only). The potential energy V must be of the form

$$V(q_{\nu+1}\cdots q_\tau,q_{11}\cdots q_{1h_1}\cdots q_{\mu1}\cdots q_{\mu h_\mu})=\sum_{\eta=1}^{\tau+\mu}\left[\sum_{\rho=\nu+1}^{\tau}\frac{u_\rho\phi_{\rho\eta}}{\phi_{\mathrm{II}}}+\sum_{\chi=1}^{\mu}\frac{u_\chi\phi_{\chi\eta}}{\phi_{\mathrm{II}}}\right],$$

where the $u_\rho(q_\rho)$ and $u_\chi(q_{\chi1}\cdots q_{\chi h_\chi})$ are defined as in Theorems II and I, respectively. The solution is given by

$$W=\sum_{\alpha,\beta=1}^{\nu}\int\varphi_{\alpha\beta}(q_\alpha)c_\beta\,dq_\alpha+\sum_{\rho=\nu+1}^{\tau}\int\left[\sum_{\beta=1}^{\nu}f_\rho{}^\beta(q_\rho)c_\beta\pm\left(\sum_{\kappa,\lambda=1}^{\nu}F_\rho{}^{\kappa\lambda}(q_\rho)c_\kappa c_\lambda+\sum_{\eta=1}^{\tau+\mu}\varphi_{\rho\eta}(q_\rho)c_\eta-2u_\rho(q_\rho)\right)^{1/2}\right]dq_\rho+\sum_{\chi=1}^{\mu}W_\chi(q_{\chi1}\cdots q_{\chi h_\chi}),$$

where W_χ is a solution of

$$\sum_{\varphi,\psi=1}^{h_\chi}A_\chi{}^{\varphi\psi}\frac{\partial W_\chi}{\partial q_{\chi\varphi}}\frac{\partial W_\chi}{\partial q_{\chi\psi}}=\sum_{\kappa,\lambda=1}^{\nu}G_\chi{}^{\kappa\lambda}c_\kappa c_\lambda+\sum_{\eta=1}^{\tau+\mu}\varphi_{\chi\eta}c_\eta-2u_\chi,$$

and the c_i $(i=1\cdots m)$ are arbitrary constants, subject to the condition

$$\sum_{\gamma=1}^{\nu}c_\gamma{}^2+\sum_{\eta=1}^{\tau+\mu}c_\eta=2E.$$

If $\tau=n$, the separation is complete; if $\tau\neq n$, n must be ≥ 3.

Theorem III clearly contains Theorems I and II as special cases. Furthermore, if $\nu+\tau=n$ $(n\geq 1)$, it reduces to Theorem II of I.

III. POLYNOMIAL FIRST INTEGRALS FOR CLASSICAL DYNAMICAL SYSTEMS

Theorem III represents the solution of the problem of partial separation of variables for solutions of the H–J equation (1), (2) of the form (4). If the variables of the third kind are absent, it would also provide the complete solution of the integration of the dynamical system described by the H–J equation or equivalently by the canonical equations

$$\frac{dq_i}{dt}=\frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt}=-\frac{\partial H}{\partial q_i}, \tag{20}$$
$$i=1\cdots n,$$

where H is the function given by (1), (2), but now the momenta p_i are variables on the same footing as the coordinates q_i rather than related to them through $W(q_1\cdots q_n)$. This follows from the Hamilton–Jacobi theorem,[9,10] which states that if a complete integral $W(q_1\cdots q_n,c_1\cdots c_n)$ of the H–J equation depending on n arbitrary constants c_i is known, then the integrals of Eqs. (20), depending on $2n$ arbitrary constants b_i, c_i, are given by

$$p_i=\frac{\partial W}{\partial q_i}, \quad b_i=-\frac{\partial W}{\partial c_i}, \tag{21}$$
$$i=1\cdots n.$$

In the presence of n-ν-τ variables of the third kind ($\nu+\tau\neq n$) the solution of the H—J equation given by Theorem III depends on only $\nu+\tau+\mu$ arbitrary constants. While this is not sufficient for the complete integration of Eqs. (20), it would be possible to obtain $n+\nu+\tau+\mu$ independent integrals if the dependence of the solution on the c_i's were known explicitly.[11] However, in general this is not the case, since the part $\sum^{III'} W_\chi$ of the solution is determined by the PDE's (10c), which are not separable by assumption.

However, in proving Theorem III we have obtained $\nu+\tau+\mu$ independent first integrals of Eqs. (20) explicitly; these are given by Eqs. (11a) and (12b), where the derivatives of the W_i's with respect to the coordinates q_k are to be replaced by the momenta p_k according to Eq. (2). The integrals (11a) are linear in the momenta, and thus their squares (12a) are integrals quadratic in them. Furthermore, the sum of the quadratic integrals (12a) and (12b) yields the energy integral (14); however, obviously (12a) and (14) are not independent of the original $\nu+\mu$ integrals.

Since from Eqs. (20) and (2)

$$\frac{dq_k}{dt}=\sum_{l=1}^{n} g^{kl}p_l, \quad p_l=\sum_k g_{lk}\frac{dq_k}{dt}, \tag{22}$$

any integral which is a polynomial of mth degree in the momenta is a polynomial of the same degree in the velocities, and conversely.

Thus we have

Theorem IV: If the Hamilton—Jacobi equation (1), (2) allows partial or complete separation of variables according to Theorem III, Hamilton's equations (20) possess ν independent integrals linear in the momenta

$$I_\gamma\equiv\sum_{\alpha=1}^{\nu}\frac{\phi_{\alpha\gamma}}{\phi_I}p_\alpha=c_\gamma, \quad \gamma=1\cdots\nu, \tag{A}$$

and $\tau+\mu$ independent integrals quadratic in the momenta

$$\begin{aligned} I_\eta\equiv&\sum_{\rho=\nu+1}^{\nu+\tau}\frac{\phi_{\rho\eta}}{\phi_{II}}\left(p_\rho-\sum_{\alpha,\beta=1}^{\nu}\frac{f_\rho^{\ \beta}\phi_{\alpha\beta}}{\phi_I}p_\alpha\right)^2\\ &+\sum_{\chi=1}^{\mu}\sum_{\varphi,\psi=1}^{h_\chi}\frac{\phi_{\chi\eta}}{\phi_{II}}A_\chi^{\ \varphi\psi}p_{(\chi\varphi)}p_{(\chi\psi)}\\ &-\sum_{\alpha,\beta,\kappa,\lambda=1}^{\nu}\left[\sum_{\rho=\nu+1}^{\nu+\tau}F_\rho^{\kappa\lambda}\phi_{\rho\eta}+\sum_{\chi=1}^{\mu}G_\chi^{\ \kappa\lambda}\phi_{\chi\eta}\right]\\ &\times\frac{\phi_{\alpha\kappa}\phi_{\beta\lambda}}{\phi_I^{\,2}\phi_{II}}p_\alpha p_\beta+\sum_{\rho=\nu+1}^{\nu+\tau}\frac{2u_\rho\phi_{\rho\eta}}{\phi_{II}}+\sum_{\chi=1}^{\mu}\frac{2u_\chi\phi_{\chi\eta}}{\phi_{II}}=c_\eta, \end{aligned} \tag{B}$$

$$\eta=1\cdots\tau+\mu,$$

where all functions and constants are defined as in Theorem III, and the constants are subject to the conditions stated there. In the case of complete separation, $\nu+\tau+\mu$ equals n. From the integrals (A) there follow ν quadratic integrals

$$I_\gamma^{\,2}=\sum_{\alpha,\beta=1}^{\nu}\frac{\phi_{\alpha\gamma}\phi_{\beta\gamma}}{\phi_I^{\,2}}p_\alpha p_\beta=c_\gamma^{\,2}, \tag{C}$$

and the integrals (B) and (C) imply the energy integral

$$H(q_1\cdots q_n,p_1\cdots p_n)\equiv\tfrac{1}{2}\sum_{k,l=1}^{n}g^{kl}p_kp_l+V=E. \tag{D}$$

In the absence of variables of the first kind, Theorem IV reduces to the results found by Stäckel[5] for quadratic integrals, and if in addition the $A_\chi^{\ \varphi\psi}$'s are subject to condition (18), it further reduces to his theorem on orthogonal quadratic integrals, which generalized an earlier result by Di Pirro.[12]

However, this does not necessarily exhaust the integrals linear or quadratic in the momenta. By Poisson's theorem,[9,10] the Poisson bracket

$$[I,I']\equiv\sum_{i=1}^{n}\left(\frac{\partial I}{\partial q_i}\frac{\partial I'}{\partial p_i}-\frac{\partial I}{\partial p_i}\frac{\partial I'}{\partial q_i}\right) \tag{23}$$

for any two integrals is again an integral. This follows from the relation

$$[I,H]=0 \tag{24}$$

valid for any integral, and Jacobi's identity

$$[A,[B,C]]+[B,[C,A]]+[C,[A,B]]=0, \tag{25}$$

valid for any three functions of the p_i and q_i; application to I, I', and H and use of Eq. (24) immediately yields

$$[[I,I'],H]=0 \tag{26}$$

as required. The integral thus obtained is not necessarily new, however, since $[I,I']$ might vanish, be identically equal to a constant, or be a function of other integrals. In particular, in the case of orthogonal coordinate systems the $\nu+\tau+\mu$ independent integrals (A) or (C) and (B) do not have any variables in common and thus all Poisson brackets formed from them vanish; in the terminology of Lie, they form a function group and are in involution.[9,13]

But if not all coordinates are orthogonal, this is not necessarily the case. Any new integrals obtained by Poisson's theorem (which from the form of the I_γ's and I_η's will again be polynomials in the momenta) can be used as members of new Poisson brackets to construct other polynomial integrals; this process may be continued step by step until all further Poisson brackets vanish, are constants, or yield only combinations of known integrals. Unfortunately, this process has to be carried out separately for any particular H. But as the total number of independent integrals equals $2n$, and each step except the last one must yield at least one new integral, this requires at most $2n-(\nu+\tau+\mu)$ steps. Since each step can increase the degree of the highest polynomial by at most one, and the initial set (A), (B) was at most quadratic in the momenta, this process can not lead to any independent polynomial integral of degree higher than $2n-(\nu+\tau+\mu)+2$, and at most to one polynomial of this degree. The procedure is straightforward but tedious, and no new general features become apparent. Therefore, we shall only give the results for the linear and quadratic integrals following from the integrals (A) and (B) of Theorem IV; cubic integrals can be obtained by forming the Poisson brackets of two different integrals (B), and all the new integrals obtained can be used to form further integrals of first to fourth degree in the next step, first to fifth in the succeeding one, and so on.

Thus we have

Theorem V: Under the conditions for which Theorem IV holds, further integrals of the dynamical system described by the Hamiltonian H may be obtained from the ν independent integrals (A) and the $\tau+\mu$ independent integrals (B), linear and quadratic in the momenta, respectively, by forming their Poisson brackets. This yields $\frac{1}{2}\nu(\nu-1)+\nu(\tau+\mu)$ integrals

$$I_{\gamma\delta}\equiv[I_\gamma, I_\delta]$$
$$=\sum_{\alpha,\beta=1}^{\nu}\left\{\frac{\phi_{\beta\delta}}{\phi_{\mathrm{I}}}\frac{\partial}{\partial q_\beta}\left(\frac{\phi_{\alpha\gamma}}{\phi_{\mathrm{I}}}\right)-\frac{\phi_{\beta\gamma}}{\phi_{\mathrm{I}}}\frac{\partial}{\partial q_\beta}\left(\frac{\phi_{\alpha\delta}}{\phi_{\mathrm{I}}}\right)\right\}p_\alpha=c_{\gamma\delta},$$
$$\gamma,\delta=1\cdots\nu,\quad \gamma\neq\delta, \tag{E}$$
$$I_{\gamma\eta}\equiv[I_\gamma, I_\eta]$$
$$=\sum_{\alpha,\beta,\gamma',\delta,\epsilon=1}^{\nu}\sum_{\rho=\nu+1}^{\nu+\tau}\frac{2\phi_{\rho\eta}f_\rho^{\ \epsilon}}{\phi_{\mathrm{I}}\phi_{\mathrm{II}}}$$
$$\times\left(\phi_{\delta\gamma}\frac{\partial}{\partial q_\delta}\left(\frac{\phi_{\alpha\epsilon}}{\phi_{\mathrm{I}}}\right)-\phi_{\delta\epsilon}\frac{\partial}{\partial q_\delta}\left(\frac{\phi_{\alpha\gamma}}{\phi_{\mathrm{I}}}\right)\right)\left(p_\rho-\frac{f_\rho^{\ \gamma'}\phi_{\beta\gamma'}}{\phi_{\mathrm{I}}}p_\beta\right)p_\alpha$$
$$-\sum_{\alpha,\beta,\delta,\kappa,\lambda=1}^{\nu}\left(\sum_{\rho=\nu+1}^{\nu+\tau}F_\rho^{\ \kappa\lambda}\phi_{\rho\eta}+\sum_{\chi=1}^{\mu}G_\chi^{\ \kappa\lambda}\phi_{\chi\eta}\right)\frac{1}{\phi_{\mathrm{II}}}$$
$$\times\left(\frac{\phi_{\delta\gamma}}{\phi_{\mathrm{I}}}\frac{\partial}{\partial q_\delta}\left(\frac{\phi_{\alpha\kappa}\phi_{\beta\lambda}}{\phi_{\mathrm{I}}^{\,2}}\right)-\frac{2\phi_{\beta\kappa}\phi_{\delta\lambda}}{\phi_{\mathrm{I}}^{\,2}}\frac{\partial}{\partial q_\delta}\left(\frac{\phi_{\alpha\gamma}}{\phi_{\mathrm{I}}}\right)\right)p_\alpha p_\beta,$$
$$\gamma=1\cdots\nu,\quad \eta=1\cdots\tau+\mu, \tag{F}$$

which are polynomials of first and second degree in the p_i's, respectively. This set of integrals may contain members which vanish identically, are identically equal to a constant, or are functionally dependent on the integrals (A) and (B); for the subset which is independent, the constants $c_{\gamma\delta}$ and $c_{\gamma\eta}$ can assume arbitrary real values. This subset and the original set (A), (B) can be used to construct further integrals by the same method, yielding polynomials of first to third degree. This process can be continued until no further independent integrals result. The number of independent polynomial integrals obtained by this method is at most equal to $2n-(\nu+\tau+\mu)$; no such integral can be of degree greater than $2n-(\nu+\tau+\mu)+2$, and there can be at most one integral of this degree.

We are not concerned here with the general problem of integration for dynamical systems described by Eqs. (20), but only note that the results of Theorem IV and V are of a form suggesting direct application of Lie's method.[9,13]

We also note that the ν linear integrals (A) are homogeneous, as are all other integrals linear in the momenta which may be obtained by successive application of Poisson's theorem to (A) and the resultant integrals. These integrals, by (22), correspond to integrals linear and homogeneous in the velocities for the geodesics of the V_n under consideration. But the existence of such integrals is the necessary and sufficient condition for the V_n to admit a group of motions G_ν, and thus for the existence of ν Killing vectors, and conversely.[14]

It should be noted, however, that the association of ν integrals and G_ν refers to *linearly* independent linear integrals. While a geodesic of a V_n can admit at most $2n$ functionally independent integrals of any type, it can possess up to $\frac{1}{2}n(n+1)$ linearly independent linear ones (for spaces of constant curvature), which for $n>3$ exceeds the number of independent integrals.[15] Thus our method does not necessarily establish the largest G_ν of the V_n under consideration. Conversely, the knowledge of a G_ν does not establish the possibility of separability of the H–J equation. The existence of a G_ν is an invariant property of the V_n, while separability is a property of particular coordinate systems, and also depends on the form of the potential energy; e.g., all S_n admit a $\frac{1}{2}n(n+1)$ – parameter group of motions, while there exists an infinity of coordinate systems and potential energies in each of these spaces which does not permit any separation at all.

While all linear integrals obtained by our method are homogeneous, the quadratic integrals (B) and all polynomial integrals of degree 3 or higher obtained by the method of Theorem V are homogeneous in general only if all u_ρ's and u_η's vanish and thus $V=0$. But then Eqs. (1), (2), or (20) determine the geodesics of the V_n. For these geodesics to admit k linearly independent homogeneous polynomial integrals of degree r, the V_n must admit k independent solutions of Killing's equation[14,16–19]

$$K_{(ab\cdots l;m)}=0,\quad a,b\cdots=1\cdots n, \tag{27}$$

where the "Killing tensor" $K_{a\ldots l}$ is symmetric and of rank r, the semicolon denotes covariant differentiation, and the brackets denote symmetrization. The linear integrals and the associated Killing vectors considered above correspond to $r=1$ (and frequently, especially in the older literature, the term "Killing's equation" is reserved for this case); the case $r=2$ has recently been the subject of several investigations[20]

One solution of Eq. (27) for $r=2$ which always exists is the metric tensor g_{ab} itself, since all its covariant derivatives vanish. Furthermore, obviously solutions for any $r>1$ can always be constructed from solutions of lower rank by forming their tensor product and symmetrizing. These solutions correspond to homogeneous polynomial integrals of degree r which are products of such integrals of lower degree, and are not considered as independent solutions of Eqs. (27).

If $V\neq 0$, not all polynomial integrals are homogeneous, as noted above. However, the existence of *any* polynomial integral of Eq. (1), (2), or (20) of highest degree r requires that the geodesics possess a homogeneous polynomial integral of degree r, and thus Eqs. (27) must still hold[16,17]; in addition, we must have[16]

$$\sum_{\alpha=1}^{n}K_{ab\ldots l}g^{am}\frac{\partial V}{\partial q_m}=0. \tag{28}$$

Of course, the existence of any polynomial integrals of degree r obtained by the method of Theorem V assures that the conditions (27) and (28) are satisfied for a Killing tensor of rank r. We are not concerned here with the explicit form of the Killing tensors of the V_n's described by the metrics of Theorem III.

From the above, we have

Theorem VI: The V_n with g^{kl} given by Theorem III admits at least ν Killings vectors, and thus a group of motions with ν parameters. Corresponding to each

polynomial integral of the dynamical system described by the Hamiltonian H of Theorem IV, of the form given in Theorems IV and V, there exists a homogeneous polynomial integral of the geodesics of the V_n of the same form except for all u_ρ's and u_χ's being equal to zero. The number of independent Killing tensors of rank r admitted by the V_n is at least equal to the number of such independent homogeneous integrals of degree r; in particular, there exist at least $\tau+\mu$ Killing tensors of rank two.

IV. THE SCHRÖDINGER AND RELATED EQUATIONS

We now consider the time-independent Schrödinger equation associated with the Hamiltonian (1) (using units such that $\hbar=1$)

$$H\psi\equiv T(q_1\cdots q_n,p_1\cdots p_n)\psi+V(q_1\cdots q_n)\psi=E\psi,$$
$$p_k\equiv\frac{1}{i}\frac{\partial}{\partial q_k}, \tag{29}$$

where

$$T\psi=-\tfrac{1}{2}\sum_{k,l=1}^{n}\frac{1}{g^{1/2}}\frac{\partial}{\partial q_k}\left(g^{1/2}g^{kl}\frac{\partial\psi}{\partial q_l}\right), \tag{30}$$
$$g\equiv|\det g_{kl}|,$$

which follows from the time-dependent Schrödinger equation by separating off the time. If $V=0$, it reduces to the Helmholtz equation (which contains the Laplace equation as a special case), which in like manner follows from the wave equation, the damped wave equation, and the diffusion or heat conduction equation, as discussed in I, where it was also shown that for an n-dimensional metric tensor of signature $n-2$ it includes the $(n-1)$-dimensional wave equation.

We shall be concerned with the problem of finding coordinate systems and potentials which allow partially separated solutions of Eqs. (29), (30) of the form

$$\psi(q_1\cdots q_n)=\prod_{\alpha=1}^{\nu}\psi_\alpha(q_\alpha)\prod_{\rho=\nu+1}^{\nu+\tau}\psi_\rho(q_\rho)\prod_{\chi=1}^{\mu}\psi_\chi(q_{\chi1}\cdots q_{\chi h_\chi}), \tag{31}$$

with the same division of coordinates (5a–c) as for the H–J equation. We also introduce the same sets of functions (6a, b), (8a, b) and of constants (9a–c) as before. However, because of the way they enter the subsequent calculations, these functions (except for the u_ρ's and u_η's) now must be required to be of class C^1.

At this stage of I, we introduced two sets of differential equations, one of which contained two differential operators $O_\rho{}^{(1)}$ and $O_\rho{}^{(2)}$, which were determined subsequently. To simplify our calculations, we immediately make use of the forms (49) and (41) of these operators obtained in I and consider the three sets of differential equations

$$\frac{1}{i}\frac{d\psi_\alpha}{dq_\alpha}=\sum_\beta{}^{\mathrm{I}}\varphi_{\alpha\beta}(q_\alpha)c_\beta\psi_\alpha,\quad\alpha=1\cdots\nu, \tag{32a}$$

$$-\frac{d}{dq_\rho}\left(f_\rho(q_\rho)\frac{d\psi_\rho}{dq_\rho}\right)$$
$$=\sum_\beta{}^{\mathrm{I}}\frac{1}{i}\left(\frac{d}{dq_\rho}[f_\rho{}^\beta(q_\rho)c_\beta\psi_\rho]+f_\rho{}^\beta c_\beta\frac{d\psi_\rho}{dq_\rho}\right)$$
$$+\left(\sum_{\kappa,\lambda}{}^{\mathrm{I}}F_\rho{}^{\kappa\lambda}(q_\rho)c_\kappa c_\lambda+\sum_\eta{}^{\mathrm{II+III'}}\varphi_{\rho\eta}(q_\rho)c_\eta-2u_\rho(q_\rho)\right)\psi_\rho,$$
$$\rho=\nu+1\cdots\nu+\tau, \tag{32b}$$

$$-\sum_{\varphi,\psi}{}^{\mathrm{III''}}\frac{\partial}{\partial q_{\chi\varphi}}\left(A_\chi{}^{\varphi\psi}\frac{\partial\psi_\chi}{\partial q_{\chi\psi}}\right)$$
$$=\left(\sum_{\kappa,\lambda}{}^{\mathrm{I}}G_\chi{}^{\kappa\lambda}c_\kappa c_\lambda+\sum_\eta{}^{\mathrm{II+III'}}\varphi_{\chi\eta}-2u_\chi\right)\psi_\chi,$$
$$\chi=1\cdots\mu, \tag{32c}$$

where the f_ρ's are a set of functions of a single variable (of the second kind) each, which also must be required to be of class C^1. We shall proceed to determine the conditions under which the system of differential equations (32) is equivalent to the PDE (29), (30). The procedure used is analogous to the one developed in I for the case of complete separation.

Multiplication of the set (32a) by ψ/ψ_α yields

$$\frac{\partial\psi}{\partial q_\alpha}=i\sum_\beta{}^{\mathrm{I}}\varphi_{\alpha\beta}(q_\alpha)c_\beta\psi. \tag{33}$$

Multiplying this by $\phi_{\alpha\kappa}/\phi_{\mathrm{I}}$ and summing over α, we obtain

$$\sum_\alpha{}^{\mathrm{I}}\frac{\phi_{\alpha\kappa}}{\phi_{\mathrm{I}}}\frac{\partial\psi}{\partial q_\alpha}=ic_\kappa\psi. \tag{34}$$

Differentiating this with respect to q_ρ gives

$$\sum_\alpha{}^{\mathrm{I}}\frac{\phi_{\alpha\kappa}}{\phi_{\mathrm{I}}}\frac{\partial^2\psi}{\partial q_\alpha\partial q_\rho}=ic_\kappa\frac{\partial\psi}{\partial q_\rho}. \tag{35}$$

Differentiating Eq. (34) with respect to q_γ instead and using Eq. (35) in the resulting expression, we get

$$\sum_\alpha{}^{\mathrm{I}}\frac{\partial}{\partial q_\gamma}\left(\frac{\phi_{\alpha\kappa}}{\phi_{\mathrm{I}}}\frac{\partial\psi}{\partial q_\alpha}\right)=-c_\kappa\sum_\beta{}^{\mathrm{I}}\varphi_{\gamma\beta}(q_\gamma)c_\beta\psi. \tag{36}$$

Multiplying this by $-\phi_{\gamma\lambda}/\phi_{\mathrm{I}}$ and summing over γ, we obtain

$$-\sum_{\alpha,\gamma}{}^{\mathrm{I}}\frac{\phi_{\gamma\lambda}}{\phi_{\mathrm{I}}}\frac{\partial}{\partial q_\gamma}\left(\frac{\phi_{\alpha\kappa}}{\phi_{\mathrm{I}}}\frac{\partial\psi}{\partial q_\alpha}\right)=c_\kappa c_\lambda\psi. \tag{37}$$

Putting $\kappa=\lambda$ in (37), we get, changing dummy indices,

$$-\sum_{\alpha,\beta}{}^{\mathrm{I}}\frac{1}{\phi_{\mathrm{I}}}\frac{\partial}{\partial q_\alpha}\left(\frac{\phi_{\alpha\gamma}\phi_{\beta\gamma}}{\phi_{\mathrm{I}}}\frac{\partial\psi}{\partial q_\beta}\right)=c_\gamma{}^2\psi. \tag{38}$$

We now multiply the sets (32b) and (32c) by ψ/ψ_ρ and ψ/ψ_χ, respectively, and eliminate the c's from the resulting expressions by means of Eqs. (34)–(37). Changing dummy indices, we get

$$-\frac{\partial}{\partial q_\rho}\left(f_\rho\frac{\partial\psi}{\partial q_\rho}\right)$$
$$=-\sum_{\alpha,\beta}{}^{\mathrm{I}}\left(\frac{\partial}{\partial q_\rho}\left[\frac{f_\rho{}^\alpha\phi_{\beta\alpha}}{\phi_{\mathrm{I}}}\frac{\partial\psi}{\partial q_\beta}\right]-\frac{f_\rho{}^\alpha\phi_{\beta\alpha}}{\phi_{\mathrm{I}}}\frac{\partial^2\psi}{\partial q_\beta\partial q_\rho}\right)$$
$$-\sum_{\beta,\gamma,\kappa,\lambda}{}^{\mathrm{I}}F_\rho{}^{\kappa\lambda}\frac{\phi_{\gamma\lambda}}{\phi_{\mathrm{I}}}\frac{\partial}{\partial q_\gamma}\left(\frac{\phi_{\beta\kappa}}{\phi_{\mathrm{I}}}\frac{\partial\psi}{\partial q_\beta}\right)$$

$$+\sum_{\eta}^{\mathrm{II}+\mathrm{III}'} \varphi_{\rho\eta}\, c_\eta\, \psi - 2u_\rho\psi \tag{39}$$

and

$$-\sum_{\varphi,\psi}^{\mathrm{III}''} \frac{\partial}{\partial q_{\chi\varphi}} \left(A_\chi^{\varphi\psi} \frac{\partial\psi}{\partial q_{\chi\psi}}\right)$$

$$= -\sum_{\beta,\gamma,\kappa,\lambda}^{\mathrm{I}} G_\chi^{\kappa\lambda} \frac{\phi_{\gamma\lambda}}{\phi_{\mathrm{I}}} \frac{\partial}{\partial q_\gamma} \left(\frac{\phi_{\beta\kappa}}{\phi_{\mathrm{I}}} \frac{\partial\psi}{\partial q_\beta}\right)$$

$$+\sum_{\eta}^{\mathrm{II}+\mathrm{III}'} \varphi_{\chi\eta}\, c_\eta\, \psi - 2u_\chi\psi. \tag{40}$$

Now we multiply Eqs. (34) and (39) by $\phi_{\rho\eta}/\phi_{\mathrm{II}}$ and $\phi_{\chi\eta}/\phi_{\mathrm{II}}$ and sum over ρ and χ, respectively, add the resulting equations, change dummy indices, and rearrange terms to obtain

$$-\sum_{\rho}^{\mathrm{II}} \frac{\phi_{\rho\eta}}{\phi_{\mathrm{II}}} \frac{\partial}{\partial q_\rho} \left(f_\rho \frac{\partial\psi}{\partial q_\rho}\right)$$

$$-\sum_{\chi}^{\mathrm{III}'} \sum_{\varphi,\psi}^{\mathrm{III}''} \frac{\phi_{\chi\eta}}{\phi_{\mathrm{II}}} \frac{\partial}{\partial q_{\chi\varphi}} \left(A_\chi^{\varphi\psi} \frac{\partial\psi}{\partial q_{\chi\psi}}\right)$$

$$-\sum_{\rho}^{\mathrm{II}} \sum_{\alpha,\beta}^{\mathrm{I}} \frac{\phi_{\rho\eta}}{\phi_{\mathrm{II}}} \left(\frac{\partial}{\partial q_\rho} \left[\frac{f_\rho{}^\alpha \phi_{\beta\alpha}}{\phi_{\mathrm{I}}} \frac{\partial\psi}{\partial q_\beta}\right] + \frac{f_\rho{}^\alpha \phi_{\beta\alpha}}{\phi_{\mathrm{I}}} \frac{\partial^2\psi}{\partial q_\beta\, \partial q_\rho}\right)$$

$$+\sum_{\alpha,\beta,\kappa,\lambda}^{\mathrm{I}} \left(\sum_{\rho}^{\mathrm{II}} F_\rho{}^{\kappa\lambda} \phi_{\rho\eta} + \sum_{\chi}^{\mathrm{III}'} G_\chi^{\kappa\lambda}\, \phi_{\chi\eta}\right)$$

$$\times \frac{\phi_{\alpha\lambda}}{\phi_{\mathrm{I}}\phi_{\mathrm{II}}} \frac{\partial}{\partial q_\alpha} \left(\frac{\phi_{\beta\kappa}}{\phi_{\mathrm{I}}} \frac{\partial\psi}{\partial q_\beta}\right)$$

$$+2\sum_{\rho}^{\mathrm{II}} \frac{u_\rho \phi_{\rho\eta}}{\phi_{\mathrm{II}}} \psi + 2\sum_{\chi}^{\mathrm{III}'} \frac{u_\chi \phi_{\chi\eta}}{\phi_{\mathrm{II}}} \psi = c_\eta \psi. \tag{41}$$

Summing Eqs. (38) and (41) over γ and η, respectively, adding the resulting equations, and using the definition (13), we get

$$-\sum_{\alpha,\beta,\gamma}^{\mathrm{I}} \frac{1}{\phi_{\mathrm{I}}} \frac{\partial}{\partial q_\alpha} \left(\frac{\phi_{\alpha\gamma}\, \phi_{\beta\gamma}}{\phi_{\mathrm{I}}} \frac{\partial\psi}{\partial q_\beta}\right)$$

$$-\sum_{\eta}^{\mathrm{II}+\mathrm{III}'} \sum_{\rho}^{\mathrm{II}} \frac{\phi_{\rho\eta}}{\phi_{\mathrm{II}}} \frac{\partial}{\partial q_\rho} \left(f_\rho \frac{\partial\psi}{\partial q_\rho}\right)$$

$$-\sum_{\eta}^{\mathrm{II}+\mathrm{III}'} \sum_{\chi}^{\mathrm{III}'} \sum_{\varphi,\psi}^{\mathrm{III}''} \frac{\phi_{\chi\eta}}{\phi_{\mathrm{II}}} \frac{\partial}{\partial q_{\chi\varphi}} \left(A_\chi^{\varphi\psi} \frac{\partial\psi}{\partial q_{\chi\psi}}\right)$$

$$+\sum_{\eta}^{\mathrm{II}+\mathrm{III}'} \sum_{\rho}^{\mathrm{II}} \sum_{\alpha,\beta}^{\mathrm{I}} \frac{\phi_{\rho\eta}}{\phi_{\mathrm{II}}}$$

$$\times \left(\frac{\partial}{\partial q_\rho} \left[\frac{f_\rho{}^\alpha\, \phi_{\beta\alpha}}{\phi_{\mathrm{I}}} \frac{\partial\psi}{\partial q_\beta}\right] + \frac{f_\rho{}^\alpha \phi_{\beta\alpha}}{\phi_{\mathrm{I}}} \frac{\partial^2\psi}{\partial q_\beta\, \partial q_\rho}\right)$$

$$+\sum_{\eta}^{\mathrm{II}+\mathrm{III}'} \sum_{\alpha,\beta,\kappa,\lambda}^{\mathrm{I}} \left(\sum_{\rho}^{\mathrm{II}} F_\rho{}^{\kappa\lambda} \phi_{\rho\eta} + \sum_{\chi}^{\mathrm{III}'} G_\chi^{\kappa\lambda}\, \phi_{\chi\eta}\right)$$

$$\times \frac{\phi_{\alpha\lambda}}{\phi_{\mathrm{I}}\phi_{\mathrm{II}}} \frac{\partial}{\partial q_\alpha} \left(\frac{\phi_{\beta\kappa}}{\phi_{\mathrm{I}}} \frac{\partial\psi}{\partial q_\beta}\right)$$

$$+2\sum_{\eta}^{\mathrm{II}+\mathrm{III}'} \left(\sum_{\rho}^{\mathrm{II}} \frac{u_\rho \phi_{\rho\eta}}{\phi_{\mathrm{II}}} + \sum_{\chi}^{\mathrm{III}'} \frac{u_\chi\, \phi_{\chi\eta}}{\phi_{\mathrm{II}}}\right) \psi = 2E\psi. \tag{42}$$

This is indeed of the form of Eq. (29) with (30).

We immediately consider the general case. It is apparent that the potential energy has the same form as for the H–J equation as given in Theorem III. Furthermore, we must have

$$g^{1/2} g^{\rho\rho} = F(q_1 \cdots q_\nu,\, q_{11} \cdots q_{\mu h_\mu})\, f_\rho \sum_{\eta}^{\mathrm{II}+\mathrm{III}'} \phi_{\rho\eta},$$

$$g^{1/2} = \phi_{\mathrm{II}} F(q_1 \cdots q_\nu,\, q_{11} \cdots q_{\mu h_\mu}), \tag{43}$$

where F cannot depend on any of the variables of the second kind, since it must be independent of the value of ρ.

Now, as in I, we can introduce new sets of functions

$$\varphi'_{\rho\eta}(q_\rho) \equiv \frac{\varphi_{\rho\eta}(q_\rho)}{f_\rho(q_\rho)}\,, \quad f'_\rho{}^\alpha(q_\rho) \equiv \frac{f_\rho{}^\alpha(q_\rho)}{f_\rho(q_\rho)}\,,$$

$$F'_\rho{}^{\kappa\lambda}(q_\rho) \equiv \frac{F_\rho{}^{\kappa\lambda}(q_\rho)}{f_\rho(q_\rho)}\,, \quad u'_\rho(q_\rho) \equiv \frac{u_\rho(q_\rho)}{f_\rho(q_\rho)}\,, \tag{44}$$

from which we get

$$\phi'_{\rho\eta} = \frac{f_\rho\, \phi_{\rho\eta}}{P}\,, \quad \phi'_{\chi\eta} = \frac{\phi_{\chi\eta}}{P}\,, \quad \phi'_{\mathrm{II}} = \frac{\phi_{\mathrm{II}}}{P}\,,$$

$$P \equiv \prod_{\sigma=\nu+1}^{\nu+\tau} f_\rho(q_\rho), \tag{45}$$

where $\phi'_{\rho\eta}$ and $\phi'_{\chi\eta}$ are the cofactors of $\varphi'_{\rho\eta}$ and $\varphi_{\chi\eta}$, respectively (which again do not depend on the q_ρ and the $q_{\chi\psi}$, respectively), and ϕ'_{II} is the new determinant of the $\varphi'_{\rho\eta}$ and $\varphi_{\chi\eta}$. Then we get from Eqs. (43) and (45)

$$g^{1/2} = \phi'_{\mathrm{II}} F(q_1 \cdots q_\nu,\, q_{11} \cdots q_{\mu h_\mu}) P. \tag{46}$$

From Eqs. (30) and (34) we must have

$$g^{1/2} g^{(\chi\varphi)(\chi\psi)} = A_\chi^{\varphi\psi}\, \phi'_{\chi\eta}\, G(q_1 \cdots q_{\nu+\tau}) P,$$

$$g^{1/2} = \phi'_{\mathrm{II}}\, G(q_1 \cdots q_{\nu+\tau}) P, \tag{47}$$

where G cannot depend on any of the variables of the third kind since $g^{1/2}$ must be independent of χ.

Comparing (46) and (47), we see that F and G must be the same function of the first kind of variables alone, i.e.,

$$G(q_1 \cdots q_{\nu+\tau}) = G(q_1 \cdots q_\nu) = F(q_1 \cdots q_\nu). \tag{48}$$

We now rewrite Eq. (42), taking into account the dependence of the various ϕ_{kl} on their argument as well as the definitions and equations (44)–(48). Since the original functions (6) and (8) were arbitrary, we can use the new quantities (44) and (45) *without* the prime without loss of generality; however, this redefinition must be taken into account in Eq. (32b). Thus, Eq. (42) becomes

$$-\sum_{\alpha,\beta,\gamma}^{\mathrm{I}} \frac{1}{\phi_{\mathrm{I}}} \frac{\partial}{\partial q_\alpha} \left(\frac{\phi_{\alpha\gamma}\, \phi_{\beta\gamma}}{\phi_{\mathrm{I}}} \frac{\partial\psi}{\partial q_\beta}\right)$$

$$-\sum_{\eta}^{\mathrm{II}+\mathrm{III}'} \sum_{\rho}^{\mathrm{II}} \frac{\phi_{\rho\eta}}{\phi_{\mathrm{II}}} \frac{1}{f_\rho} \frac{\partial}{\partial q_\rho} \left(f_\rho \frac{\partial\psi}{\partial q_\rho}\right)$$

$$-\sum_{\eta}^{\mathrm{II}+\mathrm{III}'} \sum_{\chi}^{\mathrm{III}'} \sum_{\varphi,\psi}^{\mathrm{III}''} \frac{\phi_{\chi\eta}}{\phi_{\mathrm{II}}} \frac{\partial}{\partial q_{\chi\varphi}} \left(A_\chi^{\varphi\psi} \frac{\partial\psi}{\partial q_{\chi\psi}}\right)$$

$$+\sum_{\eta}^{\mathrm{II}+\mathrm{III}'} \sum_{\rho}^{\mathrm{II}} \sum_{\alpha,\beta}^{\mathrm{I}} \frac{\phi_{\rho\eta}}{\phi_{\mathrm{II}}}$$

$$\times\left(\frac{\partial}{\partial q_\rho}\left[\frac{f_\rho f_\rho{}^\alpha \phi_{\beta\alpha}}{\phi_{\rm I}}\frac{\partial\psi}{\partial q_\beta}\right]+\frac{f_\rho f_\rho{}^\alpha \phi_{\beta\alpha}}{\phi_{\rm I}}\frac{\partial^2\psi}{\partial q_\beta \partial q_\rho}\right)$$

$$+\sum_\eta^{\rm II+III'}\sum_{\alpha,\beta,\kappa,\lambda}^{\rm I}\left(\sum_\rho^{\rm II}F_\rho{}^{\kappa\lambda}\phi_{\rho\eta}+\sum_\chi^{\rm III'}G_\chi{}^{\kappa\lambda}\phi_{\chi\eta}\right)$$

$$\times\frac{\phi_{\alpha\lambda}}{\phi_{\rm I}\phi_{\rm II}}\frac{\partial}{\partial q_\alpha}\left(\frac{\phi_{\beta\kappa}}{\phi_{\rm I}}\frac{\partial\psi}{\partial q_\beta}\right)$$

$$+2\sum_\eta^{\rm II+III'}\left(\sum_\rho^{\rm II}\frac{u_\rho\phi_{\rho\eta}}{\phi_{\rm II}}+\sum_\chi^{\rm III'}\frac{u_\chi\phi_{\chi\eta}}{\phi_{\rm II}}\right)\psi=2E. \quad (49)$$

Comparing the first and third sum with (30), we see that $F(q_1\cdots q_\nu)$ must be proportional to $\phi_{\rm II}$; the constant of proportionality can be taken as one without loss of generality. Therefore, we have the condition

$$g^{1/2}=\phi_{\rm I}\phi_{\rm II}\prod_{\sigma=\nu+1}^{\nu+\tau}f_\sigma(q_\sigma), \quad (50)$$

just as in I, Eq. (47). The various components of g^{kl} can then be identified by comparison with (30), and are found to be of the form given in Theorem III. No condition beyond Eq. (50) is required, and thus the presence of variables of the third kind does not impose any new restrictions in comparison with the allowed metrics for the H–J equation.

Equations (32a) can be readily integrated. It remains to study the reality conditions. As in the case of the H–J equation, for the metric to be real, all φ_{ij} should be real. To obtain real solutions for (32a), all constants c_i must be real. But we can also construct real solutions by taking a set of equations (32a) with complex separation constants. These will result in ψ's which are complex; however, we can also start from the complex conjugate of our original set (32a) to obtain solutions ψ^* complex conjugate to the solutions ψ, and because of the linearity of our equations their sum will then be a real solution of Eqs. (29), (30). Furthermore, unlike the case of the nonlinear H–J equation, solutions with different values of the separation constants can be linearly superposed, but the values of these constants may be subject to restrictions due to boundary conditions.

Thus we have

Theorem VII: The Schrödinger, Helmholtz, and Laplace equations can be solved by partial or complete separation of variables in any V_n $(n>1)$ whose metric is of the form given in Theorem III and whose determinant is of the form

$$g=\phi_{\rm I}{}^2\phi_{\rm II}{}^2\prod_{\sigma=\nu+1}^{\nu+\tau}f_\sigma^2(q_\sigma). \quad (A)$$

where the $f_\sigma(q_\sigma)$ are functions of class C^1 of a single variable each; for the Schrödinger equation the potential energy must be of the form given in Theorem III. The particular solutions obtained by this method are of the form

$$\psi=A\exp\left(i\sum_{\alpha,\beta=1}^{\nu}c_\beta\int\varphi_{\alpha\beta}(q_\alpha)\,dq_\alpha\right)$$

$$\times\prod_{\rho=\nu+1}^{\nu+\tau}\psi_\rho(q_\rho)\prod_{\chi=1}^{\mu}\psi_\chi(q_{\chi1}\cdots q_{\chi h_\chi})$$

(+c.c. if a real solution is desired),

where the ψ_ρ are solutions of

$$\frac{1}{f_\rho}\frac{d}{dq_\rho}\left(f_\rho\frac{d\psi_\rho}{dq_\rho}\right)$$

$$=-\frac{1}{i}\sum_{\beta=1}^{\nu}\left(\frac{1}{f_\rho}\frac{d}{dq_\rho}\left(f_\rho f_\rho{}^\beta c_\beta\psi_\rho\right)+f_\rho{}^\beta c_\beta\frac{d\psi_\rho}{dq_\rho}\right)$$

$$-\left(\sum_{\kappa,\lambda=1}^{\nu}F_\rho{}^{\kappa\lambda}c_\kappa c_\lambda+\sum_{\eta=1}^{\tau+\mu}\varphi_{\rho\eta}c_\eta-2u_\rho\right)\psi_\rho,$$

$$\rho=\nu+1\cdots\nu+\tau, \quad (B)$$

and the ψ_χ are solutions of

$$\sum_{\varphi,\psi=1}^{h_\chi}\frac{\partial}{\partial q_{\chi\varphi}}\left(A_\chi{}^{\varphi\psi}\frac{\partial\psi_\chi}{\partial q_{\chi\psi}}\right)$$

$$=-\left(\sum_{\kappa,\lambda=1}^{\nu}G_\chi{}^{\kappa\lambda}c_\kappa c_\lambda+\sum_{\eta=1}^{\tau+\mu}\varphi_{\chi\eta}c_\eta-2u_\chi\right)\psi_\chi,$$

$$\chi=1\cdots\mu. \quad (C)$$

The various functions (except the u_ρ's and u_η's) are of class C^1, and otherwise all are defined as in Theorem III and A and the c_i's are arbitrary complex constants, subject to the condition

$$\sum_{\gamma=1}^{\nu}c_\gamma^2+\sum_{\eta=1}^{\tau+\mu}c_\eta=\begin{cases}2E & \text{(Schrödinger or Helmholtz equation)}\\ 0 & \text{(Laplace equation)}\end{cases}$$

and restrictions due to boundary conditions on ψ. The solutions for different values of the c_i's satisfying these conditions can be linearly superposed.

V. POLYNOMIAL CONSTANTS OF THE MOTION OF QUANTUM DYNAMICAL SYSTEMS

Theorem VII represents the solution of the problem of partial separation of variables for solutions of the form (31) of the Schrödinger equation (29), (30) and related equations. The Schrödinger equation is the quantum mechanical analogue of the classical H–J equation for a system of mass points, whereas the Helmholtz, Laplace, and other equations discussed in Sec. IV do not necessarily describe the behavior of a mechanical system; even when they do, this description, except in special cases, is classical rather than quantum. Nevertheless, it will be convenient in the following to use the language of quantum mechanics appropriate for the Schrödinger equation to describe mathematical results common to all equations considered in Sec. IV.

In Sec. III we discussed a number of first integrals of Hamilton's equations of motion (20) obtained in Sec. II in the course of determining the metrics for which the H–J equation is separable. In complete analogy, we have obtained in Sec. IV a number of constants of the motion of the system described by the Schrödinger equation (29), (30), as will be shown below. In discussing these results, we shall for convenience reserve the term "integral" for the classical case of Secs. II and III, and the term "constant of the motion" for the quantum mechanical case under consideration; although frequently these terms are used interchangeably in both cases in

the literature, here such a use might be somewhat confusing.

The integrals considered in Sec. III were polynomials in the canonical momenta p_i (or equivalently in the velocities). The constants of the motion considered here involve operators which are polynomials in the momentum operators $p_i = -i\partial/\partial q_i$, but since they also involve functions of the coordinates, care has to be taken to use the proper order of the various factors, as usual in quantum mechanics. For purposes of comparison with the results of Sec. III, the constants of the motion will be written in terms of p_i rather than explicitly in terms of differential operators.

The integrals of Sec. III could take on arbitrary constant values, but for any solution describing a particular motion of the dynamical system they assume a single particular set of values; because of the nonlinearity of the classical equations (1), (2), or (20), solutions characterized by two different sets of constants cannot be superposed. Such solutions can be superposed for the linear equation (29), (30), however, and thus a state described by the wavefunction ψ associated with a particular value of the energy E may consist of a sum of terms each characterized by a particular set of constants, a phenomenon called "degeneracy." The interpretation of such a state can be found in any book on quantum mechanics and will not be discussed here.

In I, we considered the time-dependent H—J and Schrödinger equations, and then separated off a time-dependent factor to arrive at time-independent PDE's. In this paper, for simplicity, we immediately restricted ourselves to the time-independent equations (1), (2) and (29), (30) instead of recapitulating this procedure. However, it should be kept in mind that in these equations E arose as a separation constant, and thus its value is not preassigned. In particular, the solution Ψ of the time-dependent Schrödinger equation can consist of a superposition of solutions ψ of the time-independent equation (29), (30) belonging to different values of E, each multiplied by the appropriate time factors. A state described by Ψ may involve a finite or even infinite number of separation constants, and thus "constants of the motion"; again, the interpretation of such a state is well known.

In Sec. IV we obtained $\nu+\tau+\mu$ independent constants of the motion explicity; these are given by Eqs. (34) and (41), linear and quadratic in the momentum operators, respectively. Applying the operator of Eq. (34) twice, we obtained the quadratic constant of the motion (38), and summing all quadratic constants of the motion (38) and (41) yielded Eq. (42), which expresses the fact that the Hamiltonian H is a constant of the motion by the original equation (29), (30), giving the energy E. Obviously, (38) and (42) are not independent of the original constants of the motion. It should also be noted that the definitions and equations (44) and (45) must be inserted into the integrals (34), (38), (41), and (42), with the primes omitted, to be consistent with our earlier procedure.

Thus we have

Theorem VIII: If the Schrödinger equation (29), (30) allows partial or complete separation of variables according to Theorem VII, there exist ν independent constants of the motion, linear in the momentum operators $p_i = -i\partial/\partial q_i$,

$$I_\gamma\psi \equiv \sum_{\alpha=1}^{\nu} \frac{\phi_{\alpha\gamma}}{\phi_{\mathrm{I}}} p_\alpha\psi = c_\gamma\psi, \quad \gamma = 1 \cdots \nu, \tag{A}$$

and $\tau+\mu$ independent constants of the motion quadratic in the momentum operators

$$\begin{aligned} I_\eta\psi \equiv & \sum_{\rho=\nu+1}^{\nu+\tau} \frac{\phi_{\rho\eta}}{\phi_{\mathrm{II}}} \frac{1}{f_\rho} p_\rho(f_\rho p_\rho\psi) + \sum_{\rho=\nu+1}^{\nu+\tau} \sum_{\alpha,\beta=1}^{\nu} \frac{\phi_{\rho\eta}}{\phi_{\mathrm{II}}} \frac{1}{f_\rho} \\ & \times\left(p_\rho\left[\frac{f_\rho f_\rho{}^\alpha \phi_{\beta\alpha}}{\phi_{\mathrm{I}}} p_\beta\psi\right] + \frac{f_\beta f_\rho{}^\alpha \phi_{\beta\alpha}}{\phi_{\mathrm{I}}} p_\rho p_\beta\psi\right) \\ & + \sum_{\chi=1}^{\mu} \sum_{\varphi,\psi=1}^{h_\chi} \frac{\phi_{\chi\eta}}{\phi_{\mathrm{II}}} p_{\{\chi\varphi\}}(A_\chi{}^{\varphi\psi} p_{\{\chi\psi\}}) \\ & + \sum_{\alpha,\beta,\kappa,\lambda=1}^{\nu} \left(\sum_{\rho=\nu+1}^{\nu+\tau} F_\rho{}^{\kappa\lambda}\phi_{\rho\eta} + \sum_{\chi=1}^{\mu} G_\chi{}^{\kappa\lambda}\phi_{\chi\eta}\right) \\ & \times \frac{\phi_{\alpha\lambda}}{\phi_{\mathrm{I}}\phi_{\mathrm{II}}} p_\alpha\left(\frac{\phi_{\beta\kappa}}{\phi_{\mathrm{I}}} p_\beta\psi\right) \\ & + 2\sum_{\rho=\nu+1}^{\nu+\tau} \frac{u_\rho\phi_{\rho\eta}}{\phi_{\mathrm{II}}}\psi + 2\sum_{\chi=1}^{\mu} \frac{u_\chi\phi_{\chi\eta}}{\phi_{\mathrm{II}}}\psi = c_\eta\psi, \end{aligned} \tag{B}$$

where all functions and constants are defined as in Theorem VII, and the constants are subject to the conditions stated there. In the case of complete separation, $\nu+\tau+\mu$ equals n. From the integrals (A) then follow ν quadratic constants of the motion

$$I_\gamma{}^2 = \sum_{\alpha,\beta=1}^{\nu} \frac{1}{\phi_{\mathrm{I}}} p_\alpha\left(\frac{\phi_{\alpha\gamma}\phi_{\beta\gamma}}{\phi_{\mathrm{I}}} p_\beta\psi\right) = c_\gamma{}^2\psi. \tag{C}$$

The constants of the motion (B) and (C) imply the original differential equation

$$H\psi \equiv \tfrac{1}{2} \sum_{k,l=1}^{n} \frac{1}{g^{1/2}} p_k(g^{1/2} g^{kl} p_l\psi) + V\psi = E\psi, \tag{D}$$

providing an additional constant of the motion quadratic in the momentum operators.

However, this does not necessarily exhaust the constants of the motion linear or quadratic in the momenta. We can follow the arguments given in Sec. III, using the quantum mechanical Poisson brackets for two operators I and I', defined as

$$[I, I'] = -i(II' - I'I) \tag{51}$$

instead of the classical Poisson brackets (23). Then Jacobi's identity (25) still holds, and Eq. (24) is valid for any constant of the motion. Thus a quantum mechanical "Poisson theorem" follows from Eq. (26), and allows us to obtain additional constants of the motion from pairs of known ones. Again, the constants of the motion thus obtained might vanish identically, be identically equal to a constant, or be a function of other constants of the motion. In particular, for complete separation, just as in the classical case, the Poisson brackets of all independent constants of the motion (A) or (C) and (B) vanish; then, by (51), the operators commute.

As is well known, if the quantum brackets (51) vanish,

the physical quantities represented by the operators are simultaneously measurable. Thus all constants of the motion individually can be measured simultaneously with the energy; but this does not imply that any two of these constants of the motion also can always be measured simultaneously. For this, in general they would have to commute; the corresponding classical quantities (which of course can always be measured simultaneously) must be in involution. However, this condition holds only if the two noncommuting operators are required to possess the same complete set of eigenfunctions; but, for states for which their commutator has zero eigenvalue, simultaneous measurement is possible.[21] Since our considerations on separability of the Schrödinger equation apply only to particular states, the possible appearance of two or more operators which (though commuting with H) do not commute with each other does not pose any difficulty in interpretation.

The existence of two constants of the motion which do not commute is a necessary, though not always sufficient, condition for degeneracy of the energy of the system.[22] Thus Theorem VII either provides us with a set of $\nu+\tau+\mu$ commuting operators, or with a condition necessary for degeneracy.

To have only nondegenerate states requires coordinate systems based on complete sets of normally commuting independent dynamical variables. However, the number of operators in such a set is not a characteristic of the dynamical system under investigation, but depends on the choice of operators, unlike its classical counterpart of Sec. III, where the full specification of a solution always requires $2n$ numbers, i.e., $2n$ integrals which are functions of the dynamical variables.[23] Thus, while it was possible to conclude in the classical case (Theorem V) that the construction of polynomial integrals via Poisson's theorem must terminate after at most $2n-(\nu+\tau+\mu)$ steps and could not lead to a polynomial of degree higher than $2n-(\nu+\tau+\mu)+2$, no analogous conclusion can be drawn in the present case of construction of new polynomial constants of the motion via the quantum mechanical Poisson theorem.

As in the classical case discussed in Sec. III, the procedure used is straightforward, but tedious, and therefore we again will give only the linear and quadratic constants of the motion. Comments analogous to those of Sec. III apply to the construction of further constants of the motion.

To bring out the analogies and differences of the classical and quantum mechanical expressions, it is best to carry out all differentiations implied by the p_i's except those which involve the wavefunction directly. In particular, this shows that the classical and quantum mechanical $I_{\gamma\delta}$ are identical.

Thus we have

Theorem IX: Under the condition for which Theorem VIII holds, further constants of the motion of the dynamical system described by Hamiltonian H may be obtained from the ν independent constants of the motion (A) and the $\tau+\mu$ independent constants of the motion (B), linear and quadratic in the momentum operators, respective, by forming their quantum mechanical Poisson brackets. This yields $\frac{1}{2}\nu(\nu-1)+\nu(\tau+\mu)$ constants of the motion

$$
\begin{aligned}
I_{\gamma\delta}\psi &\equiv [I_\gamma, I_\delta]\psi \\
&= \sum_{\alpha,\beta=1}^{\nu} \left\{ \frac{\phi_{\beta\delta}}{\phi_{\mathrm{I}}} \frac{\partial}{\partial q_\beta}\left(\frac{\phi_{\alpha\gamma}}{\phi_{\mathrm{I}}}\right) \right. \\
&\quad \left. - \frac{\phi_{\beta\gamma}}{\phi_{\mathrm{I}}} \frac{\partial}{\partial q_\beta}\left(\frac{\phi_{\alpha\delta}}{\phi_{\mathrm{I}}}\right) \right\} p_\alpha \psi = c_{\gamma\delta}\psi, \\
&\quad \gamma,\delta = 1 \cdots \nu, \quad \gamma \neq \delta, \qquad \text{(E)}
\end{aligned}
$$

$$
\begin{aligned}
I_{\gamma\eta}\psi &\equiv [I_\gamma, I_\eta]\psi \\
&= \sum_{\alpha,\delta,\epsilon=1}^{\nu} \sum_{\rho=\nu+1}^{\nu+\tau} \frac{\phi_{\rho\eta}}{\phi_{\mathrm{I}}\phi_{\mathrm{II}}} \\
&\quad \times \left(\phi_{\delta\gamma} \frac{\partial}{\partial q_\delta}\left(\frac{\phi_{\alpha\epsilon}}{\phi_{\mathrm{I}}}\right) - \phi_{\delta\epsilon} \frac{\partial}{\partial q_\delta}\left(\frac{\phi_{\alpha\gamma}}{\phi_{\mathrm{I}}}\right) \right) \\
&\quad \times \left\{ f_\rho{}^{\epsilon} p_\alpha p_\alpha - \frac{1}{i}\frac{1}{f_\rho}\frac{\partial}{\partial q_\rho}\left(f_\rho f_\rho{}^{\epsilon}\right) p_\alpha \right\} \psi \\
&\quad - \sum_{\alpha,\beta,\delta,\kappa,\lambda=1}^{\nu} \left(\sum_{\rho=\nu+1}^{\nu+\tau} F_\rho{}^{\kappa\lambda}\phi_{\rho\eta} + \sum_{\chi=1}^{\mu} G_\chi{}^{\kappa\lambda}\phi_{\chi\eta} \right) \frac{1}{\phi_{\mathrm{I}}\phi_{\mathrm{II}}} \\
&\quad \times \left\{ 2\left(\phi_{\delta\gamma}\frac{\partial}{\partial q_\delta}\left(\frac{\phi_{\alpha\lambda}}{\phi_{\mathrm{I}}}\right) - \phi_{\delta\lambda}\frac{\partial}{\partial q_\delta}\left(\frac{\phi_{\alpha\gamma}}{\phi_{\mathrm{I}}}\right) \right) \frac{\phi_{\beta\kappa}}{\phi_{\mathrm{I}}} p_\alpha p_\beta \right. \\
&\quad + \frac{1}{i}\left(\phi_{\alpha\gamma}\frac{\partial}{\partial q_\alpha}\left(\frac{\phi_{\delta\lambda}}{\phi_{\mathrm{I}}}\right)\frac{\partial}{\partial q_\delta}\left(\frac{\phi_{\beta\kappa}}{\phi_{\mathrm{I}}}\right) \right. \\
&\quad \left. - \phi_{\delta\lambda}\frac{\partial}{\partial q_\alpha}\left(\frac{\phi_{\beta\gamma}}{\phi_{\mathrm{I}}}\right)\frac{\partial}{\partial q_\delta}\left(\frac{\phi_{\alpha\kappa}}{\phi_{\mathrm{I}}}\right) \right) p_\beta \\
&\quad - \frac{1}{i}\left(\phi_{\alpha\kappa}\frac{\partial^2}{\partial q_\alpha \partial q_\delta}\left(\frac{\phi_{\beta\gamma}}{\phi_{\mathrm{I}}}\right) \right. \\
&\quad \left. \left. - \phi_{\alpha\gamma}\frac{\partial^2}{\partial q_\alpha \partial q_\delta}\left(\frac{\phi_{\beta\kappa}}{\phi_{\mathrm{I}}}\right) \right) \frac{\phi_{\delta\lambda}}{\phi_{\mathrm{I}}} p_\beta \right\} \psi = c_{\gamma\eta}\psi, \qquad \text{(F)} \\
&\quad \gamma = 1 \cdots \nu, \quad \eta = 1 \cdots \tau+\mu,
\end{aligned}
$$

which are polynomials of first and second degree in the p_i's, respectively. This set of constants of the motion may contain members which vanish identically, are identically equal to a constant, or are functionally dependent on the constants of the motion (A) and (B); for the subset which is independent, the constants $c_{\gamma\delta}$ and $c_{\gamma\eta}$ can assume arbitrary values, apart from restrictions imposed by the boundary conditions. This subset and the original set (A), (B) can be used to construct further constants of the motion by the same method, yielding polynomials of first to third degree. This process can be continued until no further independent constants of the motion result.

The V_n of Theorems VII–IX has a metric tensor of the same form as in Theorems III–VI; the additional condition (A) of Theorem VII on the determinant of the metric has no effect on the number or form of the polynomial first integrals of the classical dynamical system corresponding to the system described by the quantum mechanical Hamiltonian (D) of Theorem VIII, or on the considerations leading to Theorem VI. Thus we have

Theorem X: The V_n with g^{kl} given by Theorem VII admits at least ν Killing vectors, and thus a group of

motions with ν parameters. Its geodesics admit the same homogeneous polynomials as those of the V_n of Theorem III. The number of independent Killing tensors of rank r admitted by the V_n is at least equal to the number of such independent homogeneous integrals of degree r; in particular, there exist at least $\tau+\mu$ Killing vectors of rank two.

VI. DISCUSSION

This paper dealt with the problem of partial separability of the (nonlinear) Hamilton–Jacobi equation, and of the (linear) Schrödinger, Helmholtz, and related equations. The solutions were assumed to be a sum of the form (4) in the first case, and a product of the form (31) in the second one, both involving terms which depended on μ groups of $n-\nu-\tau$ unseparated "variables of the third kind," and otherwise only $\nu+\tau$ terms depending on a single variable each (ν of the first and τ of the second kind). These are not the only possible forms of solutions in which the variables are only partially separated; in particular, this paper is not concerned with the question of "R-separability," where solutions are allowed to depend on the same variable both through terms involving this variable alone and through a term involving all variables.[6–8]

The methods used here to establish forms of the metric tensor for which solutions of the form (4) or (31) exist are an extension of those used in I to establish such terms for completely separated solutions. All the forms found in I are special cases of those found here, valid in the absence of variables of the third kind.

Theorems I–III contain forms of the metric tensor for which the H–J equation is partially separable in n-dimensional space, Theorem III being the most general, including the other theorems as special cases. Theorem VII asserts that the Schrödinger and related equations are partially separable for the same forms of the metric tensor as given in Theorem III, subject only to the additional condition (A) on the determinant, apart from additional differentiability conditions on some of the arbitrary functions entering the metric tensor. A similar condition was found in I, Theorem IV, for the case of complete separation, the only difference being that the factor ϕ_{II}^2 now depends on variables of the third as well as the second kind.

Whether the conditions imposed on the metrics in Theorems III and VII are not only sufficient, but also necessary, to ensure the possibility of separability of the equations considered, has so far not been established (except for the case of complete separability of the H–J equation[2]). Just as in I, the difficulty is due to the fact that separation of variables in general involves *all* separation constants in *each* separated ODE, some of them bilinearly; it remains to be shown that Eqs. (11) and (32) represent indeed the most general form of the separated ODE's for solutions of the form (4) or (31), or that still more complicated forms are possible.

The metrics of type T_n considered in I all correspond to flat space; for those of the other types this is not necessarily the case. For most phyiscal applications we do have to impose the requirement of flatness, however, which may further restrict the allowed forms of the metric. For ordinary two- and three-dimensional space, these forms will be given elsewhere.[24] The results for complete separation were briefly discussed in I; for partial separation of the form (4) or (31) additional results are of course obtained only for three dimensions.

Apart from the type T_0 considered in I, all types considered here correspond in general to nonorthogonal coordinate systems except if the arbitrary functions satisfy a number of special conditions which follow immediately from Theorem III; all remarks on such systems made in Sec. IV of I still apply. In particular, Eqs. (B) of Theorem VII still are self-adjoint as thev stand even for nonorthogonal systems, provided that all $f_\rho{}^\beta$ vanish. If they do not, the equations can still be made self-adjoint by a suitable integrating factor. Thus we can always obtain equations of the Sturm–Liouville type for the variables of the second and third kind.

In the course of determining the allowed forms of the tensors it is also established that there exist $\nu+\tau+\mu$ independent first integrals linear and quadratic in the momenta (or velocities), respectively, for the dynamical system described by the H–J equation, given in Theorem IV; these integrals may be used to construct further polynomial integrals (Theorem V). Similarly, there exist $\nu+\tau+\mu$ such independent constants of the motion for the Schrödinger equation (Theorem VIII), which may be used to construct further polynomial constants of the motion (Theorem IX). The quadratic integrals (F) of Theorem V are homogeneous, while the corresponding quadratic constants of the motion (F) of Theorem IX are not. If $V\neq 0$, in general none of the integrals or constants of the motion of degree 3 or higher obtained by the methods of Theorems V or IX are homogeneous. These and related points will be discussed in detail elsewhere.

The class of V_n's of Theorem III and that of the same form, but further restricted by condition (A), of Theorem VII, both admit at least ν Killing vectors and at least $\tau+\mu$ Killing tensors of rank two. They both admit the same homogeneous polynomial integrals for their geodesics, and the number of independent Killing tensors of rank r is at least equal to the number of such independent homogeneous integrals of degree r (Theorems VII and X).

*Research supported in part by the National Science Foundation and through the grant of a study leave by Temple University.
†Part of this work was carried out in 1969 at the Department of Mathematics, Birkbeck College, University of London, whose hospitality is gratefully acknowledged.

[1] P. Burgatti, Rend. Acad. Lincei 20 (1), 108 (1911).
[2] F. A. Dall'Acqua, Rend. Circ. Mat. Palermo **33**, 341 (1912).
[3] P. Havas, J. Math. Phys. **16**, 1461 (1975).
[4] T. Levi-Civita, Math. Ann. **59**, 383 (1904).
[5] P. Stäckel, Ann. Mat. **25**, 55 (1897).
[6] N. Levison, B. Bogert, and R. M. Redheffer, J. Appl. Math. 7, 241 (1949).
[7] P. Moon and E. Spencer, J. Franklin Inst. 253, 585 (1952).
[8] E. G. Kalnins and W. Miller Jr., J. Math. Phys. 15, 1728 (1974); C. P. Boyer, E. G. Kalnins and W. Miller Jr., J. Math. Phys. **16**, 499 (1975).

[9]G. Prange, in *Encyclopädie der mathematischen Wissenschaften* (Teubner, Leipzig, 1935), Vol. IV/2, Art. 12 and 13, Sec. 17.
[10]L. A. Pars, *Analytical Dynamics* (Heinemann, London, 1965).
[11]R. Lehmann-Filhès, Astron. Nachr. 165, 209 (1904).
[12]G. A. Di Pirro, Ann. Mat. 24, 315 (1896).
[13]F. Engel and K. Faber, *Die Liesche Theorie der partiellen Differentialgleichungen erster Ordnung* (Teubner, Leipzig and Berlin, 1932).
[14]L. P. Eisenhart, *Riemannian Geometry* (Princeton University, Princeton, 1949), Secs. 39 and 70.
[15]Another possibility of associating the existence of integrals of a dynamical system with invariance under a group is through an action integral and Noether's theorem; there, too, the number of integrals obtained from invariance considerations equals the number of group parameters, but may exceed the number of functionally independent integrals, as discussed in P. Havas, Acta Phys. Austr. 38, 145 (1973).
[16]M. M. G. Ricci and T. Levi-Civita, Math. Ann. 54, 125 (1901) [English translation and adaptation by R. Hermann, *Ricci and Levi-Civita's Tensor Analysis Paper* (Math. Sci. Press, Brookline, Mass., 1975)].
[17]J. E. Wright, *Invariants of Quadratic Differential Forms* (Cambridge University, Cambridge, 1908).
[18]M. Walker and R. Penrose, Comm. Math. Phys. 18, 265 (1970).
[19]P. Sommers, J. Math. Phys. 14, 787 (1973).
[20]R. P. Hughston and P. Sommers, Comm. Math. Phys. 32, 147 (1973); I. Hauser and R. J. Malhiot, J. Math. Phys. 15, 816 (1974); 16, 150 (1975), and references given there.
[21]Compare, e.g., H. A. Kramers, "Die Grundlagen der Quantentheorie," in *Hand-und Jahrbuch der Chemischen Physik*, edited by A. Eucken and K. L. Wolf (Akademische Verlagsgesellschaft m. b. H., Leipzig, 1938), Vol. I, p. 145.
[22]E. P. Kemble, *The Fundamental Principles of Quantum Mechanics* (McGraw-Hill, New York, 1937), p. 311.
[23]Ref. 22, p. 287.
[24]P. Havas (to be published).

Similarity solution for viscous and thermal gas flow in a cone*

N. Liron

Department of Applied Mathematics, The Weizmann Institute of Science, Rehovot, Israel

H. E. Wilhelm

Department of Electrical Engineering, Colorado State University, Fort Collins, Colorado 80523
(Received 26 June 1975)

The nonlinear, partial differential equations describing the compressible flow of a viscous and heat conducting gas in a cone are reduced to two coupled, ordinary, nonlinear differential equations by means of a self-similar transformation. These are solved numerically for the velocity and temperature distributions in the cone. It is shown that for given flow numbers R, P, and M, laminar flows exist only up to a critical cone angle θ_0.

INTRODUCTION

The nonlinear partial differential equations of gasdynamics and magnetogasdynamics can be reduced to ordinary nonlinear differential equations in the case of radial flow between inclined, plane walls.[1] It is shown herein that also the nonlinear partial differential equations describing the radial flow of a compressible gas in a cylindrical cone can be reduced to ordinary differential equations, if the viscosity (μ) and heat conductivity (λ) depend on temperature $\bar{T}$ like $\mu \sim \bar{T}^{1/2}$ and $\lambda \sim \bar{T}^{1/2}$. This temperature dependence of the transport coefficients is exhibited by many monotonic gases the atoms of which interact like rigid, elastic spheres. According to kinetic theory, the viscosity and heat conductivity of a rigid sphere gas are given by[2]

$$\mu = (5\sqrt{\pi}/16)Q^{-1}(mk\bar{T})^{1/2}, \quad \lambda = (75\sqrt{\pi}/64)Q^{-1}k(m^{-1}k\bar{T})^{1/2},$$

where k is the Boltzmann constant, m is the mass, and $Q = \pi d^2$ is the transport cross section of the atoms (d = interaction diameter). The rigid sphere gas has a relatively large Prandtl number,

$$P = c_p\mu/\lambda = (5k/2m)\mu/\lambda = 2/3.$$

This means that the thermal energy transport has a noticeable influence on the momentum transport in the flow, i.e., the form of the velocity distribution.

BOUNDARY-VALUE PROBLEM

For the analysis of gas flow in a cone, a spherical coordinate system $(\bar{r}, \theta, \phi)$ is introduced the origin $(\bar{r}=0)$ and polar axis $(\theta=0)$ of which coincide with the apex and the axis of rotational symmetry of the cone, respectively (Fig. 1). The walls of the conical duct are defined by the plane $[\theta=\theta_0,\ \bar{r}_1 \le \bar{r} \le \bar{r}_2]$, where $0 < \theta_0 < \pi$ and $\bar{r}_1$ and $\bar{r}_2$ are the radii at which the gas is let in and removed in the experiment $(0 < r_1 < r_2 < \infty)$. All flow fields are functions of $\bar{r}$ and θ in steady state for reasons of symmetry. The pressure, $\bar{p}(\bar{r}, \theta)$, density $\bar{\rho}(\bar{r}, \theta)$, temperature, $\bar{T}(\bar{r}, \theta)$, and velocity, $\bar{\mathbf{v}} = \{\bar{u}(\bar{r}, \theta), 0, 0\}$, fields of the flow are normalized with respect to their reference values at a fixed point, $\bar{r}_1 \le \bar{r}_0 \le \bar{r}_2$, $\hat{\theta} < \theta$. The nondimensional flow fields are then given by

$$p(r, \theta) = \bar{p}(\bar{r}, \theta)/\bar{p}_0, \quad \rho(r, \theta) = \bar{\rho}(\bar{r}, \theta)/\bar{\rho}_0, \quad T = \bar{T}(\bar{r}, \theta)/\bar{T}_0,$$
$$u(r, \theta) = \bar{u}(\bar{r}, \theta)/\bar{u}_0, \quad r = \bar{r}/\bar{r}_0,$$

where

$$\bar{p}_0 = \bar{p}(\bar{r}_0, \hat{\theta}) > 0, \quad \bar{\rho}_0 = \bar{\rho}(\bar{r}_0, \hat{\theta}) > 0, \quad \bar{T}_0 = \bar{T}(\bar{r}_0, \hat{\theta}) > 0,$$
$$\bar{u}_0 = \bar{u}(\bar{r}_0, \hat{\theta}) \neq 0, \quad \bar{p}_0 = R\bar{\rho}_0\bar{T}_0$$

and $R = k/m$ is the ideal gas constant. In applications it is convenient to use as reference point $\bar{r}_0 = \bar{r}_1$, $\hat{\theta} = 0$.

The nondimensional velocity $[u = u(r, \theta)]$, temperature $[T = T(r, \theta)]$, density $[\rho = \rho(r, \theta)]$, and pressure $[p = p(r, \theta)]$ fields of the rigid sphere gas flow in the cone are described by the nonlinear boundary-value problem[3]:

$$\rho u\frac{\partial u}{\partial r} = -\frac{1}{\gamma M^2}\frac{\partial p}{\partial r} - \frac{2}{3R}\frac{\partial}{\partial r}(\tilde{\mu}\nabla\cdot\mathbf{v}) + \frac{2}{R}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\tilde{\mu}\frac{\partial u}{\partial r}\right) - \frac{2\tilde{\mu}}{r^2}u + \frac{1}{2r\sin\theta}\frac{\partial}{\partial\theta}\left(\frac{\sin\theta}{r}\tilde{\mu}\frac{\partial u}{\partial\theta}\right)\right], \tag{1}$$

$$0 = -\frac{1}{\gamma M^2}\frac{1}{r}\frac{\partial p}{\partial\theta} - \frac{2}{3R}\frac{\partial}{r\partial\theta}(\tilde{\mu}\nabla\cdot\mathbf{v}) + \frac{1}{R}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r\tilde{\mu}\frac{\partial u}{\partial\theta}\right) + \frac{2}{r\sin\theta}\frac{\partial}{\partial\theta}\left(\frac{\sin\theta}{r}\tilde{\mu}u\right) + \frac{\tilde{\mu}}{r^2}\frac{\partial u}{\partial\theta} - \frac{2\tilde{\mu}\cot\theta}{r^2}u\right], \tag{2}$$

$$\frac{\partial}{\partial r}(\rho u) + \frac{2}{r}\rho u = 0, \tag{3}$$

$$\rho u\frac{\partial T}{\partial r} = \frac{\gamma}{PR}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\tilde{\lambda}\frac{\partial T}{\partial r}\right) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\left(\frac{\sin\theta}{r}\tilde{\lambda}\frac{\partial T}{\partial\theta}\right)\right] - (\gamma-1)p\nabla\cdot\mathbf{v} - \frac{2\gamma(\gamma-1)M^2}{3R}\tilde{\mu}(\nabla\cdot\mathbf{v})^2 + \frac{2\gamma(\gamma-1)M^2}{R}\tilde{\mu}\left[\left(\frac{\partial u}{\partial r}\right)^2 + 2\left(\frac{u}{r}\right)^2 + \frac{1}{2}\left(\frac{1}{r}\frac{\partial u}{\partial\theta}\right)^2\right], \tag{4}$$

$$p = \rho T, \tag{5}$$

$$\nabla\cdot\mathbf{v} \equiv r^{-2}\partial(r^2u)/\partial r, \tag{6}$$

where

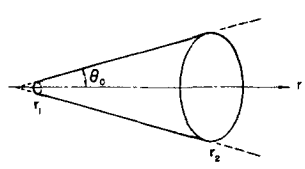


FIG. 1. Geometry of conical gas flow.

$$u(r, \theta = \theta_0) = 0, \tag{7}$$

$$T(r, \theta = \theta_0) = \hat{T}(r), \tag{8}$$

and

$$\tilde{\mu} \equiv \mu(\bar{T})/\mu(\bar{T}_0) = T^{1/2}, \quad \tilde{\lambda} = \lambda(\bar{T})/\lambda(\bar{T}_0) = T^{1/2} \tag{9}$$

are the nondimensional (T-dependent) viscosity and heat conductivity of the rigid gas sphere, respectively. The Reynolds (R), Prandtl (P), and Mach (M) numbers and the adiabatic coefficient (γ) are defined by

$$R = \bar{\rho}_0 \bar{u}_0 \bar{r}_0/\mu_0, \quad P = \mu_0 c_p/\lambda_0, \quad M = \bar{u}_0/(\gamma \bar{p}_0/\bar{\rho}_0)^{1/2}, \quad \gamma = c_p/c_v,$$
$$\mu_0 \equiv (5\sqrt{\pi}/16) Q^{-1} (m k \bar{T}_0)^{1/2},$$
$$\lambda_0 \equiv (75\sqrt{\pi}/64) Q^{-1} k (m^{-1} k \bar{T}_0)^{1/2}. \tag{10}$$

Instead of defining (self-similar) inlet ($r = r_1$) and outlet ($r = r_2$) boundary conditions for the radial velocity field $u(r, \theta)$, the improper boundary condition,

$$\Phi = 2\pi \int_0^{\theta_0} \rho(r, \theta) u(r, \theta) r^2 \sin\theta \, d\theta, \tag{11}$$

is introduced which specifies that the flux rate through the cone is independent of radius r, $\Phi = \text{const}$. Due to the normalization, the flow fields have also to satisfy the identities

$$p(r_0, \hat{\theta}) \equiv 1, \quad \rho(r_0, \hat{\theta}) \equiv 1, \quad T(r_0, \hat{\theta}) \equiv 1, \quad u(r_0, \hat{\theta}) \equiv 1. \tag{12}$$

Equations (1)—(5) represent the gasdynamic equations in spherical coordinates[3] for radial, ϕ-independent flow. Equations (7) and (8) specify that the gas does not slip and approaches a temperature distribution $\hat{T}(r)$ at the inner surface $\theta = \theta_0$ of the cone, respectively.

SIMILARITY TRANSFORMATION

The nonlinear partial differential Eqs. (1)—(4) with variable transport coefficients $\mu(T)$ and $\lambda(T)$ should be reducible to ordinary nonlinear differential equations by means of the similarity ansatz for the nondimensional flow fields,

$$u(r, \theta) = f(\theta)/r, \tag{13}$$

$$\rho(r, \theta) = g(\theta)/r, \tag{14}$$

$$p(r, \theta) = h(\theta)/r^3, \tag{15}$$

$$T(r, \theta) = (1/r^2) h(\theta)/g(\theta) \equiv \phi(\theta)/r^2, \tag{16}$$

with

$$\tilde{\mu}(r, \theta) = \tilde{\lambda}(r, \theta) = T^{1/2} = \phi^{1/2}(\theta)/r \equiv \psi(\theta)/r, \tag{17}$$

where

$$\phi(\theta) \equiv h(\theta)/g(\theta), \quad \psi(\theta) \equiv \phi^{1/2}(\theta). \tag{18}$$

This transformation amounts to a separation of the independent variables r and θ in the flow fields. The dependence r^{-1} in Eqs. (13)—(14) is suggested by the continuity Eq. (3) or the invariance of the flow rate Φ in Eq. (11). The potential energy density $\sim p(r, \theta)$ in Eq. (15) is expected to decrease like r^{-3} in a spherical coordinate system. The dependence r^{-2} in Eq. (16) follows from the state Eq. (5) in conjunction with Eqs. (14)—(15).

Indeed, by substituting Eqs. (13)—(18) into Eqs. (1), (2), and (4), one obtains coupled, ordinary differential equations for the fields $f(\theta)$, $g(\theta)$, and $h(\theta)$ [$\phi \equiv h/g$, $\psi \equiv \phi^{1/2}$]:

$$-gf^2 = (3/\gamma M^2) h + (2/R)\psi f + (2/R)[-\psi f + \tfrac{1}{2}(\psi f')' + \tfrac{1}{2}\cot\theta \, \psi f'], \tag{19}$$

$$h' = \tfrac{4}{3}(\gamma M^2/R)(\psi f)', \tag{20}$$

$$-2fg\psi^2 = (2\gamma/PR)[2\psi^3 + \psi^2\psi' \cot\theta + \psi^2\psi'' + 2\psi\psi'^2] - (\gamma - 1) hf - [2\gamma(\gamma - 1) M^2/3R]\psi f^2 + [2\gamma(\gamma - 1) M^2/R] \times [3f^2\psi + \tfrac{1}{2}\psi f'^2]. \tag{21}$$

Equation (20) is readily integrated to $h = (4\gamma M^2/3R)\psi f + h_0$. In view of $f(\theta_0) = 0$ [Eq. (7)] it is $h_0 \equiv h(\theta_0) = 1 - (4\gamma M^2/3R)$ since $h(\hat{\theta}) = \psi(\hat{\theta}) = f(\hat{\theta}) = 1$. Thus, Eq. (20) gives

$$h = 1 - \tfrac{4}{3}(\gamma M^2/R)(1 - \psi f). \tag{22}$$

Upon elimination of $g \equiv h/\psi^2$ [Eq. (18)] and h [Eq. (22)] in Eqs. (19) and (21), one arrives at the following nonlinear boundary-value problem for the functions $f(\theta) \lesseqgtr 0$ and $\psi(\theta) > 0$:

$$f'' + [\cot\theta + \psi'/\psi] f' + 4f + R(1 - 4\gamma M^2/3R) f^2/\psi^3 + \tfrac{4}{3}\gamma M^2 f^3/\psi^2 + (3R/\gamma M^2)(1 - 4\gamma M^2/3R)/\psi = 0, \tag{23}$$

$$\psi'' + [\cot\theta + 2\psi'/\psi]\psi' + 2\psi + \tfrac{2}{3}(3\gamma - 1) M^2 P f^2/\psi + \tfrac{1}{2}(\gamma - 1) M^2 P f'^2/\psi + [(3 - \gamma)/2\gamma] PR(1 - 4\gamma M^2/3R) f/\psi^2 = 0, \tag{24}$$

where

$$f(\theta = \theta_0) = 0, \tag{25}$$

$$\psi(\theta = \theta_0) = \hat{\psi} > 0, \tag{26}$$

and

$$\Phi = 2\pi \int_0^{\theta_0} f\psi^{-2} [1 - (4\gamma M^2/3R)(1 - \psi f)] \sin\theta \, d\theta, \tag{27}$$

$$f(\hat{\theta}) = 1, \quad \psi(\hat{\theta}) = 1, \tag{28}$$

in accordance with Eqs. (7)—(8) and (11)—(12). It should be noted that the boundary value in Eq. (26) is $\hat{\psi} = \hat{\phi}^{1/2} = r\sqrt{\hat{T}(r)}$ by Eqs. (8) and (16). This restricts the thermal boundary-value to self-similar distributions $\hat{T}(r) \sim r^{-2}$. Other types of thermal boundary conditions can be considered, however; e.g.[1], at a thermally insulated wall it is $\partial T(r, \theta = \theta_0)/\partial\theta = 0$, i.e., $\phi'(\theta = \theta_0) = 0$.

NUMERICAL RESULTS

For computer calculations, it is suitable to rewrite the boundary value problem in Eqs. (23)—(26) and (28) in the form

$$\frac{d^2 f}{dx^2} + \frac{df}{dx}\left(\theta_0 \cot\theta_0 x + \psi^{-1}\frac{d\psi}{dx}\right) + \theta_0^2 R\left(\frac{4}{R} f + (1 - s)\frac{f^2}{\psi^3} + s\frac{f^3}{\psi^2} + \frac{4}{Rs}(1 - s)\frac{1}{\psi}\right) = 0, \tag{29}$$

$$\frac{d^2\psi}{dx^2} + \frac{d\psi}{dx}\left(\theta_0 \cot\theta_0 x + 2\psi^{-1}\frac{d\psi}{dx}\right) + \theta_0^2 R\left\{\frac{2}{R}\psi + P\left[\frac{1}{2}\left(3 - \frac{1}{\gamma}\right) s\frac{f^2}{\psi} + \left(1 - \frac{1}{\gamma}\right)\frac{3s}{8}\left(\frac{df}{dx}\right)^2 \psi^{-1} + \frac{3 - \gamma}{2\gamma}(1 - s)\frac{f}{\psi^2}\right]\right\} = 0, \tag{30}$$

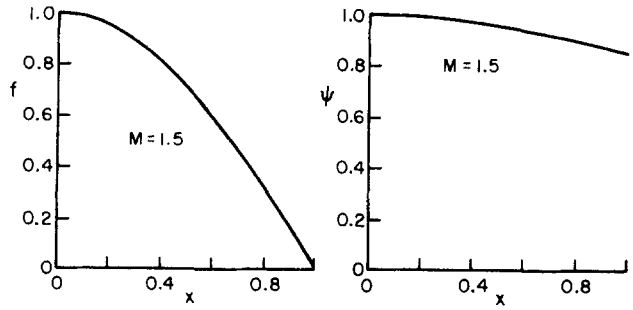


FIG. 2. $f(x)$ and $\psi(x)$ for $M=1.5$, $\theta_0^2R=2.621$.

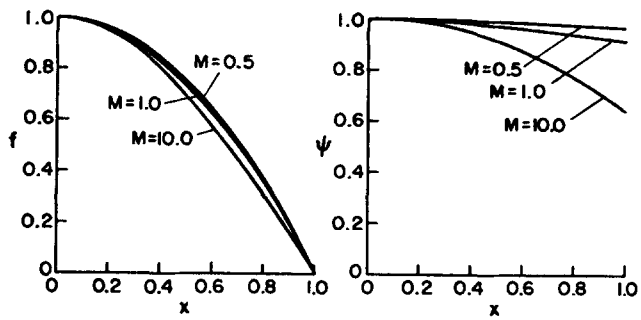


FIG. 3. $f(x)$ and $\psi(x)$ for (i) $M=0.5$, $\theta_0^2R=0.506$, (ii) $M=1.0$, $\theta_0^2R=1.585$, and (iii) $M=10.0$, $\theta_0^2R=5.272$.

where

$$f(x=1)=0, \quad f(x=0)=1 \quad (\hat{\theta}=0), \tag{31}$$

$$\psi(x=1)=\hat{\psi}, \quad \psi(x=0)=1 \quad (\hat{\theta}=0), \tag{32}$$

and

$$df(x=0)/dx=0, \quad d\psi(x=0)/dx=0, \tag{33}$$

for reasons of symmetry. The normalized independent variable x and the parameter s are

$$x\equiv\theta/\theta_0, \quad 0\leq x\leq 1, \tag{34}$$

$$s\equiv\tfrac{4}{3}\gamma M^2/R. \tag{35}$$

For the case of large Reynolds numbers, which is of the greatest practical importance, it is $R\gg 1$ and $\theta_0^2 \ll 1$ such that $R\theta^2=O[1]$ and $s=O[R^{-1}]$ for $M^2\ll R$. Furthermore, it is $\theta_0\cot\theta_0 x\approx\theta_0/\theta_0 x=x^{-1}$ for $\theta_0\ll 1$. Accordingly, Eqs. (29) and (30) may be reduced to

$$f''+f'\left(\frac{1}{x}+\frac{\psi'}{\psi}\right)+(\theta_0^2R)\left(\frac{f^2}{\psi^2}+\frac{3}{\gamma M^2}\right)\frac{1}{\psi}=0, \tag{29*}$$

$$\psi''+\psi'\left(\frac{1}{x}+2\frac{\psi'}{\psi}\right)+(\theta_0^2R)P\left(\frac{3-\gamma}{2\gamma}\right)\frac{f}{\psi^2}=0. \tag{30*}$$

For the monatomic rigid sphere gas under consideration, $\gamma=5/3$ and $P=2/3$ are fixed numbers. Hence, if M' is also prescribed, the solutions $f(x)$ and $\psi(x)$ are the same for different R and θ_0 as long as the combination $\alpha=\theta_0^2R$ remains unchanged [Eqs. (29)*–(30)*]. $\alpha=\theta_0^2R$ represents an eigenvalue of the boundary-value problem.

The conclusions for large Reynolds numbers are confirmed by the numerical solutions of the original Eqs. (29)–(33). Examples of velocity distributions $f(x)$ and (square root) temperature distributions $\psi(x)$ are shown for $\gamma=5/3$, $P=2/3$, and various M and $\alpha=\theta_0^2R$ in Figs. 2 and 3.

In Figs. 2 and 3, the cone angles are, e.g., $\theta_0\times 10^3=0.711$, 1.259, 1.619, 2.296 for $M=0.5$, 1.0, 1.5, 10.0 if $R=10^6$. The numerical results indicate that (laminar) flows exist with increasing Reynolds number R only for decreasing cone angles θ_0. As the Mach number M increases, solutions exist for larger cone angles θ_0. The profiles $f(x)$ [$\psi(x)$] become less [more] steep as M increases. For $\gamma=5/3$ and $P=2/3$, the eigenvalues $\alpha=\theta_0^2R$ are given in dependence of M in Table I, as well as θ_0 for the typical $R=10^6$.

Thus, we have obtained a new exact solution of the compressible gasdynamic equations with viscous and thermal dissipation.

APPENDIX

In this appendix we show that a flow of the type described above (self-similar) must necessarily be a pure outflow, nonzero at $\theta=0$, and symmetric. This establishes conditions (31)–(33) plus $f\geq 0$.

A term $(\cot\theta)f'$ appears in Eq. (23), and similarly a term $(\cot\theta)\psi'$ appears in Eq. (24). Since all other terms are finite at $\theta=0$, and since we obviously demand continuity of the equations through $\theta=0$, we must have

$$f'(0)=\psi'(0)=0. \tag{A1}$$

Equation (A1) already establishes the symmetry of $f(\theta)$, $\psi(\theta)$ around $\theta=0$, since the transformation $\theta\to-\theta$ leaves Eqs. (23), (24) unchanged, and Eq. (A1) assures also the same initial conditions. Moreover, we may write

$$f'(\theta)\approx a(\theta)\tan\theta, \quad \theta\ll 1, \tag{A2}$$

with

$$a=a(\theta)\neq 0. \tag{A3}$$

To show condition (A3), assume $a(0)=0$. Then

$$a(\theta)\approx\theta^{\epsilon}, \quad \epsilon>0 \text{ for } \theta\ll 1, \tag{A4}$$

and thus

$$f''(\theta)\approx\theta^{\epsilon-1}\tan\theta+\theta^{\epsilon}\cos^{-2}\theta, \quad \theta\ll 1, \tag{A5}$$

and

$$f''(0)=0. \tag{A6}$$

Either $f(0)=0$ implying $\frac{4}{3}\gamma M^2/R=1$, from Eq. (23), and therefore, from Eq. (22) $h(0)=0$—in other words the (absolute) pressure is zero, which is physically unacceptable—or $f(0)\neq 0$, and we can then take $\hat{\theta}=0$ and

TABLE I. $\alpha=\theta_0^2R$ for $\gamma=5/3$, $P=2/3$, and various M. Associated θ_0 for $R=10^6$.

M	0.5	1.0	1.5	2.0	4.0	6.0	10.0
α	0.506	1.585	2.621	3.393	4.713	5.072	5.272
$\theta_0\times 10^3$	0.711	1.259	1.619	1.842	2.171	2.252	2.296

then $f(0)=1$. Equation (23) then yields (substituting $\theta=0$) the condition

$$1+(3/\gamma M^2)(1-\tfrac{4}{3}\gamma M^2/R)=0. \tag{A7}$$

But $h(\theta_0)\geq 0$ implying

$$1-\tfrac{4}{3}\gamma M^2/R\geq 0, \tag{A8}$$

so that condition (A7) is impossible. This establishes condition (A3).

It follows that

$$\lim_{\theta\to 0} f''(\theta)=\lim_{\theta\to 0} f'(\theta)\cot\theta=a\neq 0. \tag{A9}$$

Similarly one shows

$$\lim_{\theta\to 0} \psi''(\theta)=\lim_{\theta\to 0} \psi'(\theta)\cot\theta=b\neq 0. \tag{A10}$$

We now show that the flow must be a pure outflow, i.e., $u\geq 0$ throughout. Assume $u<0$ at some point, and let a negative minimum of u be at $\hat{\theta}$, i.e., $u_0=u(\hat{\theta})<0$, and thus $R<0$. $f(\hat{\theta})$ is a (local) maximum of $f(\theta)$. If $\hat{\theta}\neq 0$, then by Eq. (23)

$$f''(\hat{\theta})=-R(1+3/\gamma M^2)>0, \tag{A11}$$

but for a maximum $f''(\hat{\theta})\leq 0$, a contradiction. If $\hat{\theta}=0$, then by Eq. (A9), $f''(0)=a$ and thus $a<0$. From Eqs. (23) and (A9)

$$[f''(\theta)+f'(\theta)\cot\theta]\big|_{\theta=0}=2a=-R(1+3/\gamma M^2)>0, \tag{A12}$$

again a contradiction. Thus $u\geq 0$ throughout, and $f\geq 0$.

If $f(0)=0$, then it must be a local minimum. By Eqs. (A9), (23)

$$[f''(\theta)+f'(\theta)\cot\theta]\big|_{\theta=0}=2a$$
$$=-(3R/\gamma M^2)(1-\tfrac{4}{3}\gamma M^2/R)\psi^{-1}<0, \tag{A13}$$

since we have already established $R>0$. But this implies a local maximum and $f\equiv 0$. We may thus assume $f(0)\neq 0$ and take $\hat{\theta}=0$. This establishes conditions (31)–(33).

Note added in proof: An other similarity transformation for compressible cone flow has been found recently by van der Werff[4] for r-dependent viscosity, $\mu\propto r^{-1}$, and heat conductivity, $\lambda\propto r^{-1}$.

*Supported in part by the U.S. Office of Naval Research.

[1]H.E. Wilhelm, J. Math. Phys. 14, 1930 (1973); Phys. Fluids 17, 360 (1974).
[2]S. Chapman and T.G. Cowling, *Theory of Nonuniform Gases* (Cambridge U.P., Cambridge, 1970).
[3]B.B. Bird, W.E. Stewart, and E.N. Lightfoot, *Transport Phenomena* (Wiley, New York, 1960).
[4]T.J. van der Werff, Phys. Fluids 18, 384 (1975).

SU(4) Clebsch–Gordan coefficients*

Veronika Rabl, George Campbell, Jr., and Kameshwar C. Wali

Physics Department, Syracuse University, Syracuse, New York 13210
(Received 29 July 1975)

We give tables of Clebsch–Gordan coefficients for the products of SU(4) representations $15 \otimes 15$ and $20 \otimes 15$, decomposed with respect to SU(3).

1. INTRODUCTION

The recent discoveries[1] of extremely narrow states suggest the existence of a new quantum number[2] in hadronic interactions. The simplest way to incorporate this new quantum number is to extend the SU(3) symmetry of hadronic interactions to that of SU(4). The latter will certainly be a more approximate symmetry than the former. Nevertheless, if the new quantum number exists, SU(4) is expected to play an important role both in the spectroscopy of hadrons and in unified renormalizable gauge theories.

The purpose of this paper is to present tables of Clebsch–Gordan coefficients of SU(4) decomposed with respect to SU(3). To the best of our knowledge this information does not exist in the literature, although SU(4) symmetry has been in physics for a long time.[3] Since its main use has been in nuclear physics, the available information concerning its decomposition has been mainly with respect to SU(2). For particle theory, it is apparent that the decomposition with respect to SU(3) will be more useful.

The great deal of qualitative and semiquantitative success of the quark model in SU(3) suggests a similar model within the framework of SU(4) symmetry. From this point of view, the most useful representations are those that can be obtained from the direct product of $4 \otimes 4^*$ and $(4 \otimes 4 \otimes 4)$. Of particular importance for us in this paper are the 15-dimensional adjoint representation which will contain the mesons and the 20-dimensional representation contained in $4 \otimes 4 \otimes 4$ which will accomodate the $\frac{1}{2}^+$ baryons.

In the next section, we review and summarize a few results concerning the irreducible representations of SU(4) necessary for our purpose. We also discuss the SU(3) content of some of the simple irreducible representations and their physical identification. Section 3 is devoted to the specification of the conventions made in the tables of Clebsch–Gordan coefficients given subsequently.

2. TENSOR REPRESENTATIONS OF SU(4)

Lee has discussed[4] most of the essential properties of SU(n) necessary for applications in particle physics. We shall specialize his results to SU(4). Let the fundamental representation be denoted by

$$Q = \begin{bmatrix} u \\ d \\ s \\ c \end{bmatrix}, \tag{2.1}$$

where u, d and s form an SU(3) triplet and c, the fourth quark, is an SU(3) singlet carrying the additional quantum number, charm $C = +1$. The transformation law for Q is given by

$$Q \to Q' = UQ, \tag{2.2}$$

where U is a 4×4 unitary unimodular transformation. The corresponding contragradient representation Q^* transforms as

$$Q^* \to Q^{*\prime} = U^* Q^*. \tag{2.3}$$

Starting from the fundamental representation (2.1), we can generate the higher-dimensional representations by taking tensor products. Tensors that correspond to the standard Young tableaux of a given pattern then form the basis of an irreducible representation.[4] For SU(4), Young tableaux have at most three rows and, if we enumerate all Young tableaux with this restriction, we obtain a complete set of irreducible tensors. The dimensionality n of the representation is given by the formula

$$n = \frac{(\lambda_1 - \lambda_2 + 1)(\lambda_1 - \lambda_3 + 2)(\lambda_1 - \lambda_4 + 3)(\lambda_2 - \lambda_3 + 1)}{12} \times (\lambda_2 - \lambda_4 + 2)(\lambda_3 - \lambda_4 + 1)\,, \tag{2.4}$$

where λ_i are nonnegative integers which specify the number of "boxes" in the ith row. We have given in Table I some of the Young tableaux, the dimensionality, and the corresponding tensors that describe the irreducible representation.

The 15-dimensional adjoint representation which is spanned by the tensors $T_{\alpha\beta,\gamma,\delta}$ can also be given in terms

TABLE I.

Young tableaux	Dimensionality	Tensor
	4	T_{α}
	15	$T_{\alpha\beta,\gamma,\delta}$
	20	$T_{\alpha\beta\gamma}$
	20	$T_{\alpha\beta,\gamma}$
	20	$T_{\alpha\beta,\gamma\delta}$
	45	$T_{\alpha\beta\gamma,\delta}$
	84	$T_{\alpha\beta\gamma\delta,\,\epsilon\sigma,\,\mu\omega}$

of the components of a mixed tensor M^a_α, where

$$M^a_\alpha = \epsilon^{\alpha\beta\gamma\delta} T_{\beta\alpha,\gamma,\delta}, \tag{2.5}$$

with the trace condition

$$M^a_a = 0. \tag{2.6}$$

The mesons which belong to the 15-dimensional adjoint representation can be written in the matrix form

$$[M^a_\alpha] = \begin{bmatrix} \pi^0/\sqrt{2} + \eta/\sqrt{6} + \eta_c/\sqrt{12} & \pi^+ & K^+ & \bar{D}^0 \\ \pi^- & -\pi^0/\sqrt{2} + \eta/\sqrt{6} + \eta_c/\sqrt{12} & K^0 & D^- \\ K^- & \bar{K}^0 & -2\eta/\sqrt{6} + \eta_c/\sqrt{12} & F^- \\ D^0 & D^+ & F^+ & -3\eta_c/\sqrt{12} \end{bmatrix} \tag{2.7}$$

which is the obvious generalization of the 3×3 matrix representation for the meson octet in SU(3). The additional mesons[5] $(\bar{D}^0, D^-F^-)$, (D^+, D^0, F^+), and η_c transform as 3, 3*, and 1 representations of SU(3) respectively, so that the SU(3) content of the 15-dimensional representation is (8, 3, 3*, 1).

The quark model suggests that the $\frac{1}{2}^+$ baryons belong to the 20-dimensional representation contained in $4 \otimes 4 \otimes 4$. This representation is described by the tensors $T_{\alpha\beta,\gamma}$ or alternately by the tensors $B^{a,b}_\alpha$, where

$$B^{a,b}_\alpha = \epsilon^{abc\gamma} T_{\beta\alpha,\gamma}, \tag{2.8}$$

with the null trace conditions

$$B^{\alpha,a}_\alpha = B^{a,\alpha}_\alpha = 0. \tag{2.9}$$

Its SU(3) content can easily be found to be (8, 6, 3*, 3). All the new baryons that appear have a nonzero value of the additional quantum number, charm. The new particles in both 15 and 20 are listed in Table II.

3. EXPLANATION OF THE TABLES OF CLEBSCH-GORDAN COEFFICIENTS

As stated in the Introduction, the main purpose of this paper is to decompose SU(4) with respect to SU(3). A state in a given irreducible representation is completely specified by

$$\phi(\{R, \mu,\}I, I_z Y, C), \tag{3.1}$$

where R and μ label the SU(4) and $SU(3) \subset SU(4)$ representations, respectively. I, I_z are the isotopic spin quantum numbers and Y, C stand for the hypercharge and charm, respectively. We shall denote I, I_z, Y, and C collectively by the symbol ν, so that the abbreviated notation for the state will read

$$\phi(\{R, \mu\}, \nu). \tag{3.2}$$

The transformation law for these states can then be written as

$$\phi'(\{R, \mu\}, \nu) = \sum_{\mu'} \sum_{\nu'} \phi(\{R, \mu'\}, \nu') D^{(R,\mu')}_{\nu',\nu}. \tag{3.3}$$

For given R and μ, $D^{(R,\mu)}_{\nu',\nu}$ is a μ-dimensional unitary matrix which describes the corresponding irreducible representation of SU(3). Considering product states, we can easily see that SU(4) Celbsch–Gordan coefficients (CG)

$$\begin{pmatrix} R_1 & R_2 & R_\gamma \\ \mu_1\nu_1 & \mu_2\nu_2 & \mu_{\gamma'}\nu \end{pmatrix} \tag{3.4}$$

can be written as

$$\begin{pmatrix} R_1 & R_2 & R_\gamma \\ \mu_1\nu_1 & \mu_2\nu_2 & \mu_{\gamma'}\nu \end{pmatrix} = \left(\begin{matrix} R_1 & R_2 \\ \mu_1 & \mu_2 \end{matrix}\middle| \begin{matrix} R_\gamma \\ \mu_{\gamma'} \end{matrix}\right) \begin{pmatrix} \mu_1 & \mu_2 & \mu_{\gamma'} \\ \nu_1 & \nu_2 & \nu \end{pmatrix}, \tag{3.4'}$$

where $\left(\begin{smallmatrix} R_1 & R_2 \\ \mu_1 & \mu_2 \end{smallmatrix}\middle| \begin{smallmatrix} R_\gamma \\ \mu_{\gamma'} \end{smallmatrix}\right)$ can be regarded as SU(3) singlet factors.[6] $\left(\begin{smallmatrix} \mu_1 & \mu_2 & \mu_{\gamma'} \\ \nu_1 & \nu_2 & \nu \end{smallmatrix}\right)$ is an SU(3) CG given by

$$\begin{pmatrix} \mu_1 & \mu_2 & \mu_{\gamma'} \\ \nu_1 & \nu_2 & \nu \end{pmatrix} = \left(\begin{matrix} \mu_1 & \mu_2 \\ I_1 Y_1 & I_2 Y_2 \end{matrix}\middle| \begin{matrix} \mu_{\gamma'} \\ IY \end{matrix}\right) C^{I_1 I_2 I}_{I_{1z} I_{2z} I_z}, \tag{3.5}$$

where $\left(\begin{smallmatrix} \mu_1 & \mu_2 \\ I_1 Y_1 & I_1 Y_2 \end{smallmatrix}\middle| \begin{smallmatrix} \mu_{\gamma'} \\ IY \end{smallmatrix}\right)$ are the SU(2) singlet or isoscalar factors and $C^{I_1 I_2 I}_{I_{1z} I_{2z} I_z}$ are the usual SU(2) isospin CG's. Our tables give the SU(3) singlet factors $\left(\begin{smallmatrix} R_1 & R_2 \\ \mu_1 & \mu_2 \end{smallmatrix}\middle| \begin{smallmatrix} R_\gamma \\ \mu_{\gamma'} \end{smallmatrix}\right)$.

TABLE IIA. Quantum numbers of new mesons in the 15-dimensional SU(4) representation.

SU(3)	Label	Isospin		Hypercharge	Charm
		I	I_z		
	$\bar{D}^0$		$+1/2$		
{3}		1/2		0	−1
	D^-		$-1/2$		
	F^-	0	0	−1	
	D^+		$+1/2$		
{3*}		1/2		0	+1
	D^0		$-1/2$		
	F^+	0	0	1	
{1}	η_c	0	0	0	0

TABLE IIB. Quantum numbers of new baryons in the 20-dimensional SU(4) representation.

SU(3)	Label	Isospin		Hypercharge	Charm
		I	I_z		
	C_1^{++}		1		
		1		1	
	C_1^+		0		
{6}	C_1^0		−1		+1
	S^+		1/2	0	
		1/2			
	S^0		−1/2		
	T^0	0	0	−1	
	A^+		1/2		
{3*}		1/2		0	+1
	A^0		−1/2		
	C_0^+	0	0	1	
	X_u^{++}		1/2		
{3}		1/2		1	+2
	X_d^+		−1/2		
	X_s^+	0	0	0	

In computing the singlet factors, as well as fixing the phases, it is useful to consider the various SU(2) subgroups of SU(4). For this purpose it is convenient to write the infinitesimal generators in terms of the non-Hermitian matrices E_{ij}, which satisfy the commutation relations

$$[E_{ij}, E_{kl}] = \delta_{jk}E_{il} - \delta_{il}E_{kj},$$

$$\text{for } i,j,k,l = 1,2,3,4 \text{ and } E_{ij}{}^{\dagger} = E_{ji}. \tag{3.6}$$

In the fundamental representation the E matrices are very simple. They are the so-called matrix units, e_{ij}, where e_{ij} is the matrix in which the element belonging to the ith row and the jth column is unity and all the other elements are zero. We can define the raising and lowering operators of the different subgroups as follows:

$$\begin{aligned} &E_{12} = I_+, \quad E_{13} = V_+, \\ &E_{23} = U_+, \quad E_{14} = L_+, \\ &E_{34} = K_+, \quad E_{24} = M_+, \\ &\text{and } I_- = E_{12}{}^{\dagger} = E_{21}, \text{ etc.} \end{aligned} \tag{3.7}$$

The operators (3.7) generate, in addition to the well-known I-spin, U-spin, and V-spin subgroups of SU(3) $\subset$ SU(4), three additional SU(2) subgroups which we have called K, L, and M:

$$\begin{aligned} &[K_+, K_-] = 2K_3 = E_{33} - E_{44}, \\ &[L_+, L_-] = 2L_3 = E_{11} - E_{44}, \\ &[M_+, M_-] = 2M_3 = E_{22} - E_{44}. \end{aligned} \tag{3.8}$$

The six diagonal operators I_3, U_3, K_3, V_3, L_3, and M_3 obviously cannot all be independent. They can be expressed in terms of the physical quantum numbers such as baryon number, hypercharge, and charm.[7] Thus,

$$\begin{aligned} U_3 &= \tfrac{1}{2}(-I_3 + \tfrac{3}{2}Y - C/2), \\ V_3 &= \tfrac{1}{2}(I_3 + \tfrac{3}{2}Y - C/2), \\ K_3 &= \tfrac{1}{2}(B - Y - C), \\ L_3 &= \tfrac{1}{2}(I_3 + \tfrac{1}{2}Y + B - \tfrac{3}{2}C), \\ M_3 &= \tfrac{1}{2}(-I_3 + \tfrac{1}{2}Y + B - \tfrac{3}{2}C). \end{aligned} \tag{3.9}$$

Phase conventions

We recall that in SU(3) the relative phases of the matrix elements of any two out of the three sets $I_\pm$, $U_\pm$, $V_\pm$ can be chosen arbitrarily. In addition to the usual convention that the matrix elements of $I_\pm$ be positive, deSwart[8] chose to make the matrix elements of $V_\pm$ positive in order to fix the relative phases of the different SU(2) multiplets contained in an irreducible representation of SU(3). Likewise, in SU(4) we can choose the phases of three out of the six sets of shift operators. However, as discussed by Baird and Biedenharn,[9] the most convenient and general phase convention for any SU(n) is to define the matrix elements of $E_{i,i+1}$, $i = 1, \ldots, n-1$, to be positive. For SU(4) this leads us to choose $I_\pm$, $U_\pm$, and $K_\pm$ to have positive matrix elements. Since this choice also affects SU(3) the resulting isoscalar factors to be used in conjunction with our SU(4) tables have been provided. Our phase conventions for SU(4) CG can then be summarized as follows:

TABLE III. SU(3) content of representations arising in the products $15 \otimes 15$ and $20 \otimes 15$.

Label	Young tableau	SU(3) content	Charm
4*		{3*}	1
		{1}	0
15*		{3*}	1
		{8}	0
		{1}	0
		{3}	−1
20		{3}	2
		{6}	1
		{3*}	1
		{8}	0
20′		{1}	3
		{3}	2
		{6}	1
		{10}	0
20″		{6}	1
		{8}	0
		{6*}	−1
36*		{6*}	2
		{15*}	1
		{3*}	1
		{8}	0
		{1}	0
		{3}	−1
45		{3}	2
		{6}	1
		{3*}	1
		{10}	0
		{8}	0
		{15}	−1
45*		{15*}	1
		{10*}	0
		{8}	0
		{6*}	−1
		{3}	−1
		{3*}	−2
60*		{15}	2
		{15*}	1
		{6}	1
		{10*}	0
		{8}	0
		{6*}	−1
84		{6*}	2
		{15*}	1
		{3*}	1
		{27}	0
		{8}	0
		{1}	0
		{15}	−1
		{3}	−1
		{6}	−2
140		{8}	3
		{15}	2
		{6*}	2
		{3}	2
		{24}	1
		{15*}	1
		{6}	1
		{3*}	1
		{27}	0
		{10}	0
		{8}	0
		{15}	−1

TABLE IV. SU(3) singlet factors for $15 \otimes 15$ of SU(4). CG series $15 \otimes 15 = 84 \oplus 45 \oplus 45^* \oplus 20'' \oplus 15_D \oplus 15_F \oplus 1$.

$\{27\}$ $C=0$

	84 +
$\{8\}\otimes\{8\}$	1

$\{10\}$ $C=0$

	45 −
$\{8\}\otimes\{8\}$	1

$\{10^*\}$ $C=0$

	45* −
$\{8\}\otimes\{8\}$	1

$\{15\}$ $C=-1$

	45 −	84 +
$\{8\}\otimes\{3\}$	$1/\sqrt{2}$	$1/\sqrt{2}$
$\{3\}\otimes\{8\}$	$-1/\sqrt{2}$	$1/\sqrt{2}$

$\{15^*\}$ $C=1$

	45* −	84 +
$\{8\}\otimes\{3^*\}$	$1/\sqrt{2}$	$1/\sqrt{2}$
$\{3^*\}\otimes\{8\}$	$-1/\sqrt{2}$	$1/\sqrt{2}$

$\{8\}$ $C=0$

	15_D +	15_F −	$20''$ +	45 −	45* −	84 +
$(\{8\}\otimes\{8\}_D$	$\sqrt{5}/3$	0	$\sqrt{5}/2\sqrt{3}$	0	0	$-1/6$
$(\{8\}\otimes\{8\})_F$	0	$\sqrt{3}/2$	0	$1/2\sqrt{2}$	$-1/2\sqrt{2}$	0
$\{3^*\}\times\{3\}$	$1/\sqrt{6}$	$-1/2\sqrt{2}$	$-1/2\sqrt{2}$	$\sqrt{3}/4$	$-\sqrt{3}/4$	$\sqrt{5}/2\sqrt{6}$
$\{3\}\otimes\{3^*\}$	$1/\sqrt{6}$	$1/2\sqrt{2}$	$-1/2\sqrt{2}$	$-\sqrt{3}/4$	$\sqrt{3}/4$	$\sqrt{5}/2\sqrt{6}$
$\{8\}\otimes\{1\}$	$-1/3\sqrt{2}$	0	$1/\sqrt{6}$	$1/2$	$1/2$	$\sqrt{5}/3\sqrt{2}$
$\{1\}\otimes\{8\}$	$-1/3\sqrt{2}$	0	$1/\sqrt{6}$	$-1/2$	$-1/2$	$\sqrt{5}/3\sqrt{2}$

$\{6\}$ $C=1$

	$20''$ +	45 −
$\{8\}\otimes\{3^*\}$	$1/\sqrt{2}$	$1/\sqrt{2}$
$\{3^*\}\otimes\{8\}$	$-1/\sqrt{2}$	$1/\sqrt{2}$

$\{6^*\}$ $C=-1$

	$20''$ +	45^* −
$\{8\}\otimes\{3\}$	$1/\sqrt{2}$	$1/\sqrt{2}$
$\{3\}\otimes\{8\}$	$-1/\sqrt{2}$	$1/\sqrt{2}$

$\{6\}$ $C=-2$

	84 +
$\{3\}\otimes\{3\}$	1

$\{6^*\}$ $C=2$

	84 +
$\{3^*\}\otimes\{3^*\}$	1

$\{3\}$ $C=-1$

	15_D +	15_F −	45^* −	84 +
$\{8\}\otimes\{3\}$	$2/3$	$1/\sqrt{3}$	$-1/\sqrt{6}$	$-1/3\sqrt{2}$
$\{3\}\otimes\{8\}$	$-2/3$	$1/\sqrt{3}$	$-1/\sqrt{6}$	$1/3\sqrt{2}$
$\{3\}\otimes\{1\}$	$1/3\sqrt{2}$	$1/\sqrt{6}$	$1/\sqrt{3}$	$2/3$
$\{1\}\otimes\{3\}$	$1/3\sqrt{2}$	$-1/\sqrt{6}$	$-1/\sqrt{3}$	$2/3$

TABLE IV. (*Continued*)

$\{3\}$ $C=2$

	45 −
$\{3^*\}\otimes\{3^*\}$	1

$\{3^*\}$ $C=1$

	15_D +	15_F −	45 −	84 +
$\{8\}\otimes\{3^*\}$	$-2/3$	$1/\sqrt{3}$	$1/\sqrt{6}$	$1/3\sqrt{2}$
$\{3^*\}\otimes\{8\}$	$2/3$	$1/\sqrt{3}$	$1/\sqrt{6}$	$-1/3\sqrt{2}$
$\{3^*\}\otimes\{1\}$	$1/3\sqrt{2}$	$-1/\sqrt{6}$	$1/\sqrt{3}$	$2/3$
$\{1\}\otimes\{3^*\}$	$1/3\sqrt{2}$	$1/\sqrt{6}$	$-1/\sqrt{3}$	$2/3$

$\{3^*\}$ $C=-2$

	45* −
$\{3\}\otimes\{3\}$	1

$\{1\}$ $C=0$

	1 +	15_D +	15_F −	84 +
$(\{8\}\otimes\{8\})_D$	$2\sqrt{2}/\sqrt{15}$	$2/3$	0	$-1/3\sqrt{5}$
$\{3^*\}\otimes\{3\}$	$-1/\sqrt{5}$	$1/\sqrt{6}$	$1/\sqrt{2}$	$-\sqrt{2/15}$
$\{3\}\otimes\{3^*\}$	$1/\sqrt{5}$	$-1/\sqrt{6}$	$1/\sqrt{2}$	$\sqrt{2/15}$
$\{1\}\otimes\{1\}$	$-1/\sqrt{15}$	$\sqrt{2}/3$	0	$4\sqrt{2}/3\sqrt{5}$

TABLE V. SU(3) singlet factors for $20\otimes 15$ of SU(4). CG series $20\otimes 15=140\oplus 60^*\oplus 36^*\oplus 20'\oplus 20_1\oplus 20_2\oplus 4^*$.

$\{27\}$ $C=0$

	140 +
$\{8\}\otimes\{8\}$	1

$\{24\}$ $C=1$

	140 +
$\{6\}\otimes\{8\}$	1

$\{15\}$ $C=2$

	60* −	140 +
$\{6\}\otimes\{3^*\}$	$1/\sqrt{2}$	$1/\sqrt{2}$
$\{3\}\otimes\{8\}$	$-1/\sqrt{2}$	$1/\sqrt{2}$

$\{15\}$ $C=-1$

	140 +
$\{8\}\otimes\{3\}$	1

TABLE V. (*Continued*)

{15*} $C=1$

	36* +	60* −	140 +
{8}⊗{3*}	$\sqrt{3}/2\sqrt{2}$	1/2	$\sqrt{3}/2\sqrt{2}$
{6}⊗{8}	$-3/4$	$\sqrt{3}/2\sqrt{2}$	1/4
{3*}⊗{8}	$-1/4$	$-\sqrt{3}/2\sqrt{2}$	3/4

{10} $C=0$

	20′ −	140 +
{8}⊗{8}	$\sqrt{2/3}$	$1/\sqrt{3}$
{6}⊗{3}	$-1/\sqrt{3}$	$\sqrt{2/3}$

{10*} $C=0$

	60* −
{8}⊗{8}	1

{8} $C=3$

	140 +
{3}⊗{3*}	1

{8} $C=0$

	20_1 +	20_2 −	36* +	60* −	140 +
$(\{8\}\otimes\{8\})_D$	$\sqrt{65}/4\sqrt{6}$	0	$-\sqrt{5}/4\sqrt{2}$	$\sqrt{5}/4\sqrt{2}$	$-1/4\sqrt{6}$
$(\{8\}\otimes\{8\})_F$	$-5/4\sqrt{78}$	$2\sqrt{2}/\sqrt{13}$	$-3/4\sqrt{2}$	$-1/4\sqrt{2}$	$\sqrt{5}/4\sqrt{6}$
{8}⊗{1}	$1/2\sqrt{39}$	$1/\sqrt{13}$	1/2	1/2	$\sqrt{5}/2\sqrt{3}$
{6}⊗{3}	$-17/4\sqrt{78}$	$-\sqrt{2/13}$	$-3/4\sqrt{2}$	$3/4\sqrt{2}$	$\sqrt{5}/4\sqrt{6}$
{3*}⊗{3}	$3\sqrt{3}/4\sqrt{26}$	$-\sqrt{2/13}$	$-1/4\sqrt{2}$	$-3/4\sqrt{2}$	$\sqrt{15}/4\sqrt{2}$

{6} $C=1$

	20_1 +	20_2 −	20′ −	60* −	140 +
{8}⊗{3*}	$17/6\sqrt{26}$	$2\sqrt{2}/\sqrt{39}$	$\sqrt{2}/3$	$1/2\sqrt{2}$	$\sqrt{5}/6$
{6}⊗{8}	$-5\sqrt{5}/12\sqrt{39}$	$4\sqrt{5}/3\sqrt{13}$	$-\sqrt{5}/3\sqrt{3}$	$-\sqrt{5}/4\sqrt{3}$	$1/6\sqrt{6}$
{6}⊗{1}	$-7/3\sqrt{39}$	$-1/3\sqrt{13}$	$-2/3\sqrt{3}$	$1/\sqrt{3}$	$\sqrt{10}/3\sqrt{3}$
{3}⊗{3}	$11/6\sqrt{13}$	$-2/\sqrt{39}$	$-1/3$	$-1/2$	$\sqrt{5}/3\sqrt{2}$
{3*}⊗{8}	$-\sqrt{13}/4\sqrt{3}$	0	$1/\sqrt{3}$	$-\sqrt{3}/4$	$\sqrt{5}/2\sqrt{6}$

{3} $C=2$

	20_1 +	20_2 −	20′ −	140 +
{6}⊗{3*}	$-11/3\sqrt{26}$	$2\sqrt{2}/\sqrt{39}$	$-\sqrt{2}/3$	$1/3\sqrt{2}$
{3}⊗{8}	$17/3\sqrt{78}$	$4\sqrt{2}/3\sqrt{13}$	$-2\sqrt{2}/3\sqrt{3}$	$-1/3\sqrt{6}$
{3}⊗{1}	$4/3\sqrt{39}$	$-5/3\sqrt{13}$	$-2/3\sqrt{3}$	$4/3\sqrt{3}$
{3*}⊗{3*}	$1/\sqrt{39}$	$2/\sqrt{13}$	$1/\sqrt{3}$	$1/\sqrt{3}$

{6*} $C=2$

	36* +	140 +
{3}⊗{8}	$-\sqrt{3}/2$	1/2
{3*}⊗{3*}	1/2	$\sqrt{3}/2$

{6*} $C=-1$

	60* −
{8}⊗{3}	1

TABLE V. (*Continued*)

{3} $C=-1$

	36* +
{8}⊗{3}	−1

{3*} $C=1$

	4* +	20_1 +	20_2 −	36* +	140 +
{8}⊗{3*}	$2/\sqrt{15}$	$-3/2\sqrt{13}$	$4/\sqrt{39}$	$1/2\sqrt{10}$	$1/2\sqrt{2}$
{6}⊗{8}	$-\sqrt{2/5}$	$-\sqrt{13}/2\sqrt{6}$	0	$-\sqrt{3}/2\sqrt{20}$	$1/4\sqrt{3}$
{3}⊗{3}	$1/\sqrt{5}$	$-1/\sqrt{39}$	$-2/\sqrt{13}$	$-\sqrt{3/10}$	$1/\sqrt{6}$
{3*}⊗{8}	$-\sqrt{2}/3\sqrt{5}$	$7/2\sqrt{78}$	$4\sqrt{2}/3\sqrt{13}$	$-11/4\sqrt{15}$	$1/4\sqrt{3}$
{3*}⊗{1}	$-2/3\sqrt{5}$	$2/\sqrt{39}$	$-1/3\sqrt{13}$	$\sqrt{2/15}$	$\sqrt{2/3}$

{1} $C=0$

	4* +	36* +
{8}×{8}	$-2/\sqrt{5}$	$-1/\sqrt{5}$
{3*}⊗{3}	$1/\sqrt{5}$	$-2/\sqrt{5}$

{1} $C=3$

	20′ −
{3}⊗{3*}	−1

TABLE VI. Isoscalar factors for $\{8\}\otimes\{8\}$. CG series $\{8\}\otimes\{8\}=\{27\}\oplus\{10\}\oplus\{10^*\}\oplus\{8_D\}\oplus\{8_F\}\oplus\{1\}$.

$Y=-1$ $I=1/2$

	27 +	8_D +	8_F −	10 −
$\Xi\pi$	$1/2\sqrt{5}$	$3/2\sqrt{5}$	1/2	−1/2
$\Sigma\bar{K}$	$-1/2\sqrt{5}$	$-3/2\sqrt{5}$	1/2	−1/2
$\Xi\eta$	$3/2\sqrt{5}$	$-1/2\sqrt{5}$	1/2	1/2
$\Lambda\bar{K}$	$3/2\sqrt{5}$	$-1/2\sqrt{5}$	−1/2	−1/2

$Y=-2$ $I=0$

	10 −
$\Xi\bar{K}$	−1

$Y=-2$ $I=1$

	27 +
$\Xi\bar{K}$	1

TABLE VI. (*Continued*)

$Y=-1$ $I=3/2$

	27 +	10* −
$\Xi\pi$	$1/\sqrt{2}$	$-1/\sqrt{2}$
$\Sigma\bar{K}$	$1/\sqrt{2}$	$1/\sqrt{2}$

$Y=0$ $I=0$

	27 +	8_D +	8_F −	1 +
$N\bar{K}$	$-\sqrt{3}/2\sqrt{5}$	$-1/\sqrt{10}$	$1/\sqrt{2}$	$-1/2$
ΞK	$\sqrt{3}/2\sqrt{5}$	$1/\sqrt{10}$	$1/\sqrt{2}$	$1/2$
$\Sigma\pi$	$-1/2\sqrt{10}$	$-\sqrt{3/5}$	0	$\sqrt{3}/2\sqrt{2}$
$\Lambda\eta$	$3\sqrt{3}/2\sqrt{10}$	$-1/\sqrt{5}$	0	$-1/2\sqrt{2}$

$Y=0$ $I=1$

	27 +	8_D +	8_F −	10 −	10* −
$N\bar{K}$	$1/\sqrt{5}$	$-\sqrt{3/10}$	$-1/\sqrt{6}$	$-1/\sqrt{6}$	$1/\sqrt{6}$
ΞK	$1/\sqrt{5}$	$-\sqrt{3/10}$	$1/\sqrt{6}$	$1/\sqrt{6}$	$-1/\sqrt{6}$
$\Sigma\pi$	0	0	$\sqrt{2/3}$	$-1/\sqrt{6}$	$1/\sqrt{6}$
$\Sigma\eta$	$\sqrt{3/10}$	$1/\sqrt{5}$	0	$1/2$	$1/2$
$\Lambda\pi$	$\sqrt{3/10}$	$1/\sqrt{5}$	0	$-1/2$	$-1/2$

$Y=0$ $I=2$

	27 +
$\Sigma\pi$	1

$Y=1$ $I=1/2$

	27 +	8_D +	8_F −	10* −
$N\pi$	$-1/2\sqrt{5}$	$-3/2\sqrt{5}$	$1/2$	$1/2$
ΣK	$1/2\sqrt{5}$	$3/2\sqrt{5}$	$1/2$	$1/2$
$N\eta$	$3/2\sqrt{5}$	$-1/2\sqrt{5}$	$-1/2$	$1/2$
ΛK	$3/2\sqrt{5}$	$-1/2\sqrt{5}$	$1/2$	$-1/2$

$Y=1$ $I=3/2$

	27 +	10 −
$N\pi$	$1/\sqrt{2}$	$-1/\sqrt{2}$
ΣK	$1/\sqrt{2}$	$1/\sqrt{2}$

TABLE VI. (*Continued*)

$Y=2$ $I=0$

	10* −
NK	1

$Y=2$ $I=1$

	27 +
NK	1

TABLE VII. Isoscalar factors for $\{8\}\otimes\{3\}$. CG series $\{8\}\otimes\{3\}=\{15\}+\{6^*\}+\{3\}$.

$I=1$ $Y=-1$ $C=-1$

	6* −	15 +
πF^-	$1/\sqrt{2}$	$1/\sqrt{2}$
$\bar{K}\bar{D}$	$-1/\sqrt{2}$	$1/\sqrt{2}$

$I=1/2$ $Y=0$ $C=-1$

	3 −	6* −	15 +
$\pi\bar{D}$	3/4	$\sqrt{3/8}$	1/4
KF^-	$-\sqrt{3/8}$	1/2	$\sqrt{3/8}$
$\eta\bar{D}$	1/4	$-\sqrt{3/8}$	3/4

$I=0$ $Y=-1$ $C=-1$

	3 −	15 +
$\bar{K}\bar{D}$	$\sqrt{3}/2$	1/2
ηF^-	−1/2	$\sqrt{3}/2$

$I=0$ $Y=1$ $C=-1$

	6* −
$K\bar{D}$	1

$I=1$ $Y=1$ $C=-1$

	15 +
$K\bar{D}$	1

$I=3/2$ $Y=0$ $C=-1$

	15 +
$\pi\bar{D}$	1

$I=1/2$ $Y=-2$ $C=-1$

	15 +
$\bar{K}F^-$	1

TABLE VIII. Isoscalar factors for $\{8\}\otimes\{3^*\}$. CG series $\{8\}\otimes\{3^*\}=\{15^*\}+\{6\}+\{3^*\}$.

$I=1$ $Y=1$ $C=1$

	6 −	15* +
πF^+	$1/\sqrt{2}$	$1/\sqrt{2}$
KD	$-1/\sqrt{2}$	$1/\sqrt{2}$

$I=1/2$ $Y=0$ $C=1$

	3* −	6 −	15* +
πD	3/4	$-\sqrt{3/8}$	−1/4
$\bar{K}F^+$	$\sqrt{3/8}$	1/2	$\sqrt{3/8}$
ηD	−1/4	$-\sqrt{3/8}$	3/4

$I=0$ $Y=1$ $C=1$

	3* −	15* +
KD	$\sqrt{3}/2$	−1/2
ηF^+	1/2	$\sqrt{3}/2$

TABLE XIII. (*Continued*)

$I=0$ $Y=-1$ $C=1$

	6 −
$\bar{K}D$	−1

$I=1$ $Y=-1$ $C=1$

	15* +
$\bar{K}D$	1

$I=3/2$ $Y=0$ $C=1$

	15* +
πD	1

$I=1/2$ $Y=2$ $C=1$

	15* +
KF^*	1

TABLE IX. Isoscalar factors for $\{6\}\otimes\{8\}$. CG series $\{6\}\otimes\{8\}=\{24\}\oplus\{15^*\}\oplus\{6\}\oplus\{3^*\}$.

$I=2$ $Y=1$ $C=1$

	24 +
$C_1\pi$	1

$I=3/2$ $Y=2$ $C=1$

	24 +
C_1K	1

$I=1/2$ $Y=2$ $C=1$

	15* −
C_1K	1

$I=1/2$ $Y=-2$ $C=1$

	24 +
$T\bar{K}$	1

$I=3/2$ $Y=0$ $C=1$

	24 +	15* −
$S\pi$	$\sqrt{2/3}$	$-\sqrt{1/3}$
$C_1\bar{K}$	$\sqrt{1/3}$	$\sqrt{2/3}$

$I=1$ $Y=-1$ $C=1$

	24 +	15* −
$S\bar{K}$	$\sqrt{2/3}$	$\sqrt{1/3}$
$T\pi$	$\sqrt{1/3}$	$-\sqrt{2/3}$

$I=1$ $Y=1$ $C=1$

	24 +	15* −	6 −
$C_1\pi$	$-1/\sqrt{15}$	$1/\sqrt{3}$	$\sqrt{3/5}$
$C_1\eta$	$\sqrt{2/5}$	$1/\sqrt{2}$	$-1/\sqrt{10}$
SK	$2\sqrt{2/15}$	$-1/\sqrt{6}$	$\sqrt{3/10}$

$I=1/2$ $Y=0$ $C=1$

	24 +	15* −	6 −	3* +
$C_1\bar{K}$	$-\sqrt{2/15}$	$-1/2\sqrt{6}$	$3/2\sqrt{5}$	$\sqrt{3}/2\sqrt{2}$
$S\pi$	$-1/\sqrt{15}$	$5/4\sqrt{3}$	$3/2\sqrt{10}$	$-\sqrt{3}/4$
$S\eta$	$\sqrt{3/5}$	$\sqrt{3}/4$	$1/2\sqrt{10}$	$\sqrt{3}/4$
TK	$1/\sqrt{5}$	$-1/2$	$\sqrt{3/10}$	$-1/2$

$I=0$ $Y=-1$ $C=1$

	24 +	6 −
$S\bar{K}$	$-\sqrt{2/5}$	$\sqrt{3/5}$
$T\eta$	$\sqrt{3/5}$	$\sqrt{2/5}$

$I=0$ $Y=1$ $C=1$

	15* −	3* +
$C_1\pi$	$-1/2$	$\sqrt{3}/2$
SK	$\sqrt{3}/2$	$1/2$

TABLE X. Isoscalar factors for $\{6\}\otimes\{3\}$. CG series $\{6\}\otimes\{3\}=\{10\}\oplus\{8\}$.

$Y=-2$ $I=0$ $C=0$

	10 +
TF^-	1

$Y=-1$ $I=1/2$ $C=0$

	10 +	8 −
SF^-	$\sqrt{2/3}$	$\sqrt{1/3}$
$T\bar{D}$	$\sqrt{1/3}$	$-\sqrt{2/3}$

$Y=0$ $I=0$ $C=0$

	8 −
$S\bar{D}$	1

$Y=0$ $I=1$ $C=0$

	10 +	8 −
C_1F^-	$\sqrt{1/3}$	$\sqrt{2/3}$
$S\bar{D}$	$\sqrt{2/3}$	$-\sqrt{1/3}$

$Y=1$ $I=1/2$ $C=0$

	8 −
$C_1\bar{D}$	1

$Y=1$ $I=3/2$ $C=0$

	10 +
$C_1\bar{D}$	1

TABLE XI. Isoscalar factors for $\{6\}\otimes\{3^*\}$. CG series $\{6\}\otimes\{3^*\}=\{15\}\oplus\{3\}$.

$Y=-1$ $I=1/2$ $C=2$

	15 +
TD	1

$Y=0$ $I=0$ $C=2$

	15 +	3 −
SD	$-1/\sqrt{2}$	$1/\sqrt{2}$
TF^+	$1/\sqrt{2}$	$1/\sqrt{2}$

$Y=0$ $I=1$ $C=2$

	15 +
SD	1

$Y=1$ $I=1/2$ $C=2$

	15 +	3 −
C_1D	$-1/2$	$\sqrt{3}/2$
SF^+	$\sqrt{3}/2$	$1/2$

$Y=1$ $I=3/2$ $C=2$

	15 +
C_1D	1

$Y=2$ $I=1$ $C=2$

	15 +
C_1F^+	1

TABLE XII. Isoscalar factors for $\{3\}\otimes\{3\}$. CG series $\{3\}\otimes\{3\}=\{6\}\oplus\{3^*\}$.

$I=1/2\ \ Y=-1\ \ C=-2$

	3^* −	6 +
$\bar{D}F^-$	$1/\sqrt{2}$	$1/\sqrt{2}$
$F^-\bar{D}$	$-1/\sqrt{2}$	$1/\sqrt{2}$

$I=1\ \ Y=0\ \ C=-2$

	6 +
$\bar{D}\bar{D}$	1

$I=0\ \ Y=-2\ \ C=-2$

	6 +
F^-F^-	1

$I=0\ \ Y=0\ \ C=-2$

	3^* −
$\bar{D}\bar{D}$	1

TABLE XIII. Isoscalar factors for $\{3\}\otimes\{3^*\}$. CG series $\{3\}\otimes\{3^*\}=\{8\}\oplus\{1\}$.

$I=1\ \ Y=0\ \ C=0$

	8 +
$\bar{D}D$	1

$I=1/2\ \ Y=1\ \ C=0$

	8 +
$\bar{D}F^+$	1

$I=1/2\ \ Y=-1\ \ C=0$

	8 +
F^-D	1

$I=0\ \ Y=0\ \ C=0$

	1 −	8 +
$\bar{D}D$	$\sqrt{2/3}$	$-1/\sqrt{3}$
F^-F^+	$1/\sqrt{3}$	$\sqrt{2/3}$

TABLE XIV. Isoscalar factors for $\{3^*\}\otimes\{3^*\}$. CG series $\{3^*\}\otimes\{3^*\}=\{6^*\}\oplus\{3\}$.

$I=1/2\ \ Y=1\ \ C=2$

	3 −	6^* +
DF^+	$1/\sqrt{2}$	$1/\sqrt{2}$
F^+D	$-1/\sqrt{2}$	$1/\sqrt{2}$

$I=0\ \ Y=0\ \ C=2$

	3 −
DD	-1

$I=0\ \ Y=2\ \ C=2$

	6^* +
F^+F^+	1

$I=1\ \ Y=0\ \ C=2$

	6^* +
DD	1

(a) For each R_γ we consider the highest state $(\mu_\gamma,\nu)_H$, i.e., the state with highest I_z in the highest-dimensional $\mu_{\gamma'}$, and we choose the coefficient

$$\begin{pmatrix} R_1 & R_2 & R_\gamma \\ \mu_1\nu_1 & \mu_2\nu_2 & (\mu_\gamma,\nu)_H \end{pmatrix}$$

to be positive for the highest $(\mu_1 I_1)$ that occurs. If this is not sufficient, we shall in addition require the highest $(\mu_2 I_2)$. The only ambiguity occurs if $(\mu_\gamma,\nu)_H$ is an SU(3) octet; to resolve it, we shall define $\{8_D\}$ to be a higher representation than $\{8_F\}$.

(b) The relative phases between different SU(3) representations in a given R_γ are obtained by requiring the matrix elements of $K_\pm$ to be positive.

The plus or minus sign underneath each R_γ in the tables denotes the phase factor $\eta_1=\pm 1$ associated with

the symmetry property of the CG:

$$\begin{pmatrix} R_1 & R_2 & R_\gamma \\ \mu_1\nu_1 & \mu_2\nu_2 & \mu_\gamma,\nu \end{pmatrix} = \eta_1 \begin{pmatrix} R_2 & R_1 & R_\gamma \\ \mu_2\nu_2 & \mu_1\nu_1 & \mu_\gamma,\nu \end{pmatrix}. \tag{3.10}$$

For the SU(3) singlet factors this implies

$$\left(\begin{matrix} R_1 & R_2 \\ \mu_1 & \mu_2 \end{matrix}\middle| \begin{matrix} R_\gamma \\ \mu_\gamma, \end{matrix}\right) = \eta_1\xi_1 \left(\begin{matrix} R_2 & R_1 \\ \mu_2 & \mu_1 \end{matrix}\middle| \begin{matrix} R_\gamma \\ \mu_\gamma, \end{matrix}\right), \tag{3.11}$$

where ξ_1 is the usual phase defined for the SU(3) CG by[7]

$$\begin{pmatrix} \mu_1 & \mu_2 & \mu_\gamma, \\ \nu_1 & \nu_2 & \nu \end{pmatrix} = \xi_1 \begin{pmatrix} \mu_2 & \mu_1 & \mu_\gamma, \\ \nu_2 & \nu_1 & \nu \end{pmatrix}. \tag{3.12}$$

ACKNOWLEDGMENTS

Two of the authors (V.R. and K.C.W.) would like to thank Professor B.W. Lee for his hospitality at FERMILAB, where this work was initiated. We would also like to thank Professor L.C. Biedenharn for a discussion concerning phase conventions in SU(n).

*Work supported in part by the U.S. Atomic Energy Commission (ERDA).

[1]J.J. Aubert *et al.*, Phys. Rev. Lett. **33**, 1404 (1974); J.-E. Augustin *et al.*, *ibid.* 33, 1406 (1974); C. Bacci *et al.*, *ibid.* 33, 1408 (1974).

[2]Amati *et al.*, Nuovo Cimento 34, 1732 (1964); J.D. Bjorken and S. Glashow, Phys. Rev. Lett. **11**, 255 (1964); Z. Maki and Y. Ohnuki, Prog. Theor. Phys. 32, 144 (1964); Y. Hara, Phys. Rev. 34, B701 (1964).

[3]E. Wigner, Phys. Rev. **51**, 106 (1937).

[4]B.W. Lee, *Brandeis University Summer Institute in Theoretical Physics*, 1965, edited by M. Chrétien and S. Deser (Gordon and Breach, New York, 1966); see also H.J. Lipkin, *Lie Groups for Pedestrians* (North-Holland, Amsterdam, 1966).

[5]M.K. Gaillard, B.W. Lee, and J.L. Rosner, Rev. Mod. Phys. **47**, 277 (1975).

[6]These correspond to the $SU(4)$: $SU(3)$ reduced Wigner coefficients of J.D. Louck and L.C. Biedenharn, J. Math. Phys. 4, 2368 (1970). We call them SU(3) singlet factors in analogy with isoscalar factors in SU(3).

[7]Here we choose to keep the standard definition $Y=B+S$ corresponding to a hypercharge assignment $Y=\frac{1}{3}$ for the charmed quark. If the hypercharge of the charmed quark is $-\frac{2}{3}$, Y both here and in the tables has to be replaced by $Y'=B+S-C$.

[8]J.J. deSwart, Rev. Mod. Phys. 35, 916 (1963).

[9]G.E. Baird and L.C. Biedenharn, J. Math. Phys. 5, 1723 (1964).

Lie theory and separation of variables. 8. Semisubgroup coordinates for $\Psi_{tt} - \Delta_2 \Psi = 0$

E. G. Kalnins

Centre de Recherches Mathématiques, Université de Montréal, Montréal 101, P.Q., Canada

Willard Miller, Jr.

School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455
(Received 22 January 1975)

We classify and study all coordinate systems which permit R-separation of variables for the wave equation in three space–time variables and such that at least one of the variables corresponds to a one-parameter symmetry group of the wave equation. We discuss 33 such systems and relate them to orbits of commuting operators in the enveloping algebra of the conformal group $SO(3,2)$.

I. INTRODUCTION

This paper is one of a series[1-7] devoted to uncovering the relationships between the symmetry group of a linear second order partial differential equation and the coordinate systems in which variables separate for that equation. Here, we study the wave equation

$$(\partial_{00} - \partial_{11} - \partial_{22})\Psi(x) = 0 \qquad (*)$$

in three space–time variables. The symmetry group of this equation is locally isomorphic to the ten-parameter group $SO(3,2)$. In Paper 9 of this series we will derive explicitly the possible orthogonal coordinate systems with respect to which variables separate or R-separate in (*). (More precisely we shall list all coordinate systems obtained from confocal cyclides and their limits.[8]) We will show that each such system corresponds to a two-dimensional (commuting) subspace of the space of second order symmetric elements in the enveloping algebra of $so(3,2)$. Here, the elements of $so(3,2)$ are first order differential operators which are symmetries of (*). If the commuting operators Q, S form a basis for such a subspace, then the separated solutions Ψ of (*) associated with this coordinate system are characterized by the eigenvalue equations $Q\Psi = \lambda\Psi$, $S\Psi = \mu\Psi$, where the eigenvalues λ, μ are the separation constants. The group $SO(3,2)$ acts on the enveloping algebra of $so(3,2)$ via the adjoint representation and preserves the rank of operators in the enveloping algebra. In particular the infinitesimal symmetries of (*) generate the identity component of $SO(3,2)$ and the symmetry I such that $I\Psi(x) = \Psi(-x)$ lies in the component not connected with the identity. Under this action the two-dimensional commuting subspaces of symmetric second order elements are decomposed into $SO(3,2)$-orbits. We regard coordinate systems attached to subspaces on the same orbit as equivalent, i.e., one such system can be obtained from any other one by an $SO(3,2)$ transformation.

Much of this paper is an introduction to the problem of separation of variables for (*). Most of the detailed calculations will be presented in Paper 9 of this series and subsequent publications. In Sec. 1 we compute the symmetry algebra of (*) in two different bases and by taking a Fourier transform we construct the well-known Hilbert space H_+ of positive energy solutions of (*). On H_+ the symmetry operators of $so(3,2)$ exponentiate to yield a unitary irreducible representation of a covering group $\widetilde{SO(3,2)}$ of the identity component in $SO(3,2)$. In Sec. 2 we determine explicitly the action of $\widetilde{SO(3,2)}$ on H_+. Most of the results of this section appear to be new.

The remainder of the paper is devoted to separation of variables. If a separable coordinate system corresponds to a subspace where there exist operators $Q = A^2$, $S = B^2$ with $[A,B] = 0$ and $A, B \in so(3,2)$, we call such coordinates *subgroup coordinates*. In this case one can diagonalize the first-order operators A, B. These systems are the best-known and easiest to find. More generally, if there exist operators Q, S such that $Q = A^2$, $[A,S] = 0$, and $A \in so(3,2)$, we call these coordinates *semisubgroup coordinates*. Here, one can diagonalize the first order operator A. If there exists no pair Q, S such that Q is a square of some $A \in so(3,2)$, we call the coordinates *nonsubgroup*. Nonsubgroup coordinates are the most intractible of all separable coordinates and appear the least frequently in applications.

A given $A \in so(3,2)$ may correspond to several (or to no) semisubgroup systems. Indeed, if Ψ satisfies both (*) and $A\Psi = i\lambda\Psi$, then, since A is a symmetry of (*), we can use standard Lie theory and introduce new variables y_0, y_1, y_2 such that $A = \partial_{y_0} + f(y)$ (where f may be zero) and $\Psi(y) = r(y)\exp(i\lambda y_0)\Phi_\lambda(y_1, y_2)$, where r is a fixed function satisfying $\partial_{y_0} r + fr = 0$. Then (*) reduces to a second order partial differential equation (†) for Φ_λ in the two variables y_1, y_2. The possible semisubgroup systems A^2, S thus correspond to the possible coordinate systems such that the reduced equation (†) separates. In particular S corresponds to a second order symmetry of the reduced equation.

In Secs. 3–7 we examine the possible semisubgroup systems. The systems are of seven types corresponding to seven choices for A. Using the notation for elements of $so(3,2)$ introduced in Sec. 1, we find that these types are:

1]. $A = \Gamma_{45}$. Then (†) becomes the eigenvalue equation for the Laplace operator on the sphere S_2. We find two coordinate systems.[4,9]

2]. $A=P_0$ and (†) is the reduced wave equation (4.1). We find four coordinate systems.

3]. $A=P_2$ and (†) is the Klein–Gordon equation (4.5). We find 11 coordinate systems.[3,10]

4]. $A=D$ and (†) is the eigenvalue equation (4.9) for the Laplace operator on a hyperboloid. We find nine coordinate systems.[4]

5]. $A=P_0+P_1$ and (†) is the free particle Schrödinger equation (5.1). We find four coordinate systems.[5]

6]. $A=M_{12}$ and (†) is the Euler–Poisson–Darboux (EPD) equation (6.1). We find nine coordinate systems.

7]. $A=\frac{1}{2}(\Gamma_{23}-\Gamma_{45})$ and (†) is Eq. (7.1). The problem of separation of variables for coordinates of this type is currently under study. There are at least three coordinate systems.

Eliminating duplicate coordinate systems we obtain a total of 33 distinct semisubgroup systems at this writing, 27 of which have already been discussed in Refs. 1–7. The systems of types 6] and 7] are related to unitary representations of the universal covering group of $SL(2,R)$ which belong to the discrete series. They will be discussed in detail in future papers.

At this writing we have determined all $A\in so(3,2)$ such that a separable coordinate system corresponds to some commuting pair A^2,S and such that S belongs to the enveloping algebra of the symmetry algebra of the reduced equation (†) associated with A^2. However, there are some omissions on our list, due to the fact that diagonalization of A does not uniquely determine the variable y_0 which is split off to obtain (†). The systems we have omitted correspond to nonorthogonal coordinates and are such that S is not expressible in terms of the symmetry algebra of (†). These systems prove rather intractible from the group theoretical viewpoint. The proofs of the above remarks follow from the results of Sec. 8. In Paper 9 and later publications we will settle these questions by using other techniques to explicitly list all systems (orthogonal or not) such that (*) R-separates.

In only a few representative cases do we explicitly list the coordinate systems. For 27 systems the expressions are given in Refs. 1–7 and 10. In Paper 9 we will derive explicitly all orthogonal systems allowing R-separation in (*) and obtain the corresponding semisubgroup systems as special cases.

Finally, in Sec. 8 we classify the orbits in $so(3,2)$ under the adjoint action of $SO(3,2)$ to see why not every $A\in so(3,2)$ corresponds to a semisubgroup system of the form A^2,S.

The special functions appearing in this paper are all defined as in the Bateman Project.[11]

1. $SO(3,2)$ AND THE WAVE EQUATION

We are concerned with the wave equation

$$(\partial_{00}-\partial_{11}-\partial_{22})\psi(x)=0,\quad x=(x_0,x_1,x_2). \tag{1.1}$$

As usual[5] we define the symmetry algebra of (1.1) to be the set of all linear differential operators

$$L=\sum_{\alpha=0}^{2}a_\alpha(x)\partial_\alpha+b(x)$$

such that $L\psi$ is a (local) solution of (1.1) whenever ψ is a (local) solution.

It is well known that the possible operators L form a ten-dimensional Lie algebra, isomorphic to $so(3,2)$, where the commutator is the usual Lie bracket.[12] As a convenient basis for this model of $so(3,2)$ we choose the momentum operators

$$P_\alpha=\partial_\alpha,\quad \alpha=0,1,2, \tag{1.2}$$

the generators of homogeneous Lorentz transformations

$$M_{12}=x_1\partial_2-x_2\partial_1,\quad M_{01}=x_0\partial_1+x_1\partial_0,$$
$$M_{02}=x_0\partial_2+x_2\partial_0, \tag{1.3}$$

the generator of dilatations

$$D=-(\tfrac{1}{2}+x_0\partial_0+x_1\partial_1+x_2\partial_2), \tag{1.4}$$

and the generators of special conformal transformations

$$K_0=-x_0+(x\circ x-2x_0^2)\partial_0-2x_0x_1\partial_1-2x_0x_2\partial_2,$$
$$K_1=x_1+(x\cdot x+2x_1^2)\partial_1+2x_1x_0\partial_0+2x_1x_2\partial_2, \tag{1.5}$$
$$K_2=x_2+(x\cdot x+2x_2^2)\partial_2+2x_2x_0\partial_0+2x_2x_1\partial_1,$$

where

$$x\cdot y=x_0y_0-x_1y_1-x_2y_2=x_0y_0-\mathbf{x}\cdot\mathbf{y}. \tag{1.6}$$

The commutation relations follow from (1.27) and (1.28) which will be derived later.

These symmetry operators can be exponentiated to obtain a local Lie transformation group of symmetries of (1.1).[12,13] In particular, the momentum and Lorentz operators generate the Poincaré group of symmetries,

$$\psi(x)\to\psi(\Lambda^{-1}(x-a)),\quad a=(a_0,a_1,a_2),\quad \Lambda\in SO(1,2), \tag{1.7}$$

the dilatation operator generates

$$(\exp\lambda D)\psi(x)=\exp(-\lambda/2)\psi[\exp(-\lambda)x],\quad \lambda\in R, \tag{1.8}$$

and the K_α generate the special conformal transformations

$$\exp(a_0K_0+a_1K_1+a_2K_2)\psi(x)$$
$$=[1+2x\circ a+(a\cdot a)(x\circ x)]^{-1/2}$$
$$\times\Psi\left(\frac{x+a(x\circ x)}{1+2x\circ a+(a\circ a)(x\circ x)}\right). \tag{1.9}$$

In addition we shall consider the inversion, space reflection, and time reflection operators,

$$R\psi(x)=(1/\sqrt{-x\cdot x})\psi(-x/x\circ x),$$
$$S\psi(x)=\psi(x_0,-x_1,x_2), \tag{1.10}$$
$$T\psi(x)=\psi(-x_0,x_1,x_2),$$

which are not generated by the local Lie symmetries (1.2)–(1.5).

As is well-known,[12,14] by formally taking the Fourier transform in the variables x_α we can express a solution $\psi(x)$ of (1.1) in the form

$$\psi(x)=\frac{1}{4\pi}\int_{-\infty}^{\infty}\int[\exp(ik\cdot x)f(k_1,k_2)$$

$$+\exp(i\hat{k}\cdot x)\tilde{f}(k_1,k_2)]\,d\mu(\mathbf{k}), \tag{1.11}$$

where $k_0=+(k_1^2+k_2^2)^{1/2}$, $\hat{k}=(-k_0,k_1,k_2)$, and $d\mu(\mathbf{k})=dk_1\,dk_2/k_0$.

Let $H=H_+\oplus H_-$ be the space of all ordered pairs of complex-valued functions $\{f(k_1,k_2),\tilde{f}(k_1,k_2)\}=\mathbf{F}(k_1,k_2)$ defined on R_2 such that

$$\int\int(|f|^2+|\tilde{f}|^2)\,d\mu(\mathbf{k})<\infty \tag{1.12}$$

(Lebesgue integral), and consider the indefinite inner product on H given by

$$\langle\mathbf{F},\mathbf{G}\rangle=\int\int(f\bar{g}-\tilde{f}\bar{\tilde{g}})\,d\mu(\mathbf{k}). \tag{1.13}$$

Then, as is well-known,[12,14] the functions ψ, Φ related to F, G by (1.11) satisfy

$$\langle\Psi,\Phi\rangle\equiv\langle\mathbf{F},\mathbf{G}\rangle=2i\int_{x_0=t}\int(\psi(x)\partial_0\bar{\Phi}(x)-[\partial_0\psi(x)]\bar{\Phi}(x))dx_1\,dx_2 \tag{1.14}$$

independent of t. More precisely (1.14) can be derived from (1.13) by first considering the dense subspace of H consisting of C^∞ functions with compact support bounded away from (0,0) and then passing to the limit. For $\mathbf{F}\in H$ the corresponding $\psi(x)$ is a solution of (1.1) in the sense of distribution theory; it may not be true that Ψ is two times continuously differentiable in each variable.

The operators (1.2)–(1.5) acting on solutions of (1.1) induce corresponding operators on H under which H_+ and H_- are separately invariant. Indeed with repeated integrations by parts we can establish that the action of these operators on H_+ is

$$P_0=ik_0,\quad P_j=-ik_j,\quad j=1,2, \tag{1.15}$$

$$M_{12}=k_1\partial_{k_2}-k_2\partial_{k_1},\quad M_{01}=k_0\partial_{k_1},\quad M_{02}=k_0\partial_{k_2}, \tag{1.16}$$

$$D=\tfrac{1}{2}+k_1\partial_{k_1}+k_2\partial_{k_2}, \tag{1.17}$$

$$K_0=ik_0(\partial_{k_1k_1}+\partial_{k_2k_2}), \tag{1.18}$$

$$K_1=i(k_1\partial_{k_1k_1}-k_1\partial_{k_2k_2}+2k_2\partial_{k_1k_1}+\partial_{k_1}),$$

$$K_2=i(-k_2\partial_{k_1k_1}+k_2\partial_{k_2k_2}+2k_1\partial_{k_1k_2}+\partial_{k_2}).$$

The action on H_- is the same except that k_0 is replaced by $-k_0$ in each of (1.15)–(1.18). Moreover, it is straightforward to verify that these operators are skew-Hermitian on H_+ and H_- separately.

The induced operators S and T on H are

$$S\mathbf{F}(k_1,k_2)=\mathbf{F}(-k_1,k_2)=(f(-k_1,k_2),\tilde{f}(-k_1,k_2)),$$

$$T\mathbf{F}(k_1,k_2)=(\tilde{f}(k_1,k_2),f(k_1,k_2)). \tag{1.19}$$

Thus, H_+ and H_- are invariant under S, but these spaces are interchanged by T. In view of this interchange property of T we will henceforth limit ourselves to consideration of elements in the Hilbert space H_+, or what amounts to the same thing, the positive energy solutions

$$\Psi(x)=\frac{1}{4\pi}\int_{-\infty}^{\infty}\int\exp(ik\cdot x)f(k_1,k_2)\,d\mu(\mathbf{k}). \tag{1.20}$$

The inner product on H_+ is

$$\langle f,g\rangle=\int_{-\infty}^{\infty}\int f(k_1,k_2)\bar{g}(k_1,k_2)\,d\mu(\mathbf{k}) \tag{1.21}$$

and

$$\langle\Psi,\Phi\rangle\equiv\langle f,g\rangle=4i\int_{x_0=t}\int\Psi(x)\partial_0\bar{\Phi}(x)\,dx_1\,dx_2$$

$$=-4i\int_{x_0=t}\int\bar{\Phi}(x)\partial_0\Psi(x)\,dx_1\,dx_2. \tag{1.22}$$

Furthermore, if Ψ is given by (1.20), we have

$$f(k_1,k_2)=\frac{k_0}{\pi}\int_{-\infty}^{\infty}\int\Psi(x)\exp(-ik\cdot x)\,dx_1\,dx_2. \tag{1.23}$$

By employing arguments analogous to those in Ref. 12, one can show that H_+ is invariant under R and

$$Rf(\mathbf{k})=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int\cos\sqrt{2l\cdot k}\,f(\mathbf{l})\,d\mu(\mathbf{l}),\quad f\in H_+,$$

$$R^2=E, \tag{1.24}$$

where E is the identity operator on H_+. Clearly, R extends to a unitary self-adjoint operator on H_+ with eigenvalues ±1. It follows from the configuration-space realization of our operators that

$$RK_jR^{-1}=P_j,\quad j=1,2,\quad RK_0R^{-1}=-P_0,\quad RDR^{-1}=-D,$$

$$RM_{\alpha\beta}R^{-1}=M_{\alpha\beta},\quad R=R^{-1}. \tag{1.25}$$

At this point it is convenient to introduce another basis for our Lie algebra of symmetry operators which clearly displays the isomorphism between this algebra and $so(3,2)$. We define $so(3,2)$ as the ten-dimensional Lie algebra of 5×5 matrices A such that $AG+GA^t=0$, where 0 is the zero matrix and

$$G=\begin{pmatrix}1&&&&0\\&1&&&\\&&1&&\\&&&-1&\\0&&&&-1\end{pmatrix}.$$

Let $\mathcal{E}_{ij}$ be the 5×5 matrix with a 1 in row i, column j and zeros elsewhere. Then the matrices

$$\Gamma_{ab}=\mathcal{E}_{ab}-\mathcal{E}_{ba}=-\Gamma_{ba},\quad a\neq b,$$

$$\Gamma_{aB}=\mathcal{E}_{aB}+\mathcal{E}_{Ba}=\Gamma_{Ba},\quad 1\leq a,b\leq3, \tag{1.26}$$

$$\Gamma_{AB}=-\mathcal{E}_{AB}+\mathcal{E}_{BA}=-\Gamma_{BA},\quad 4\leq A,B\leq5,$$

form a basis for $so(3,2)$ with commutation relations

$$[\Gamma_{ab},\Gamma_{cd}]=\delta_{bc}\Gamma_{ad}+\delta_{ad}\Gamma_{bc}+\delta_{ca}\Gamma_{db}+\delta_{db}\Gamma_{ca},$$

$$[\Gamma_{aB},\Gamma_{cd}]=-\delta_{ad}\Gamma_{cB}+\delta_{ac}\Gamma_{dB},$$

$$[\Gamma_{Ab},\Gamma_{45}]=\delta_{A5}\Gamma_{4b}-\delta_{A4}\Gamma_{5b}, \tag{1.27}$$

$$[\Gamma_{aB},\Gamma_{cD}]=\delta_{BD}\Gamma_{ac}-\delta_{ac}\Gamma_{BD}.$$

This Γ-basis is related to our other basis via

$$P_0=\Gamma_{14}+\Gamma_{45},\quad P_1=\Gamma_{12}+\Gamma_{25},\quad P_2=\Gamma_{13}+\Gamma_{35},$$

$$K_0=\Gamma_{14}-\Gamma_{45},\quad K_1=\Gamma_{12}-\Gamma_{25},\quad K_2=\Gamma_{13}-\Gamma_{35},\qquad (1.28)$$
$$M_{12}=\Gamma_{23},\quad M_{01}=\Gamma_{42},\quad M_{02}=\Gamma_{43},\quad D=\Gamma_{15}.$$

Furthermore, we can set $R=-G$.

For our model of $so(3,2)$ we have

$$P_0^2-P_1^2-P_2^2=K_0^2-K_1^2-K_2^2=0, \qquad (1.29)$$

where the result for the K-operators follows from (1.25). Furthermore, direct computations yield

$$\begin{aligned}&\Gamma_{12}^2+\Gamma_{13}^2+\Gamma_{23}^2=\Gamma_{45}^2+\tfrac{1}{4},\\ &M_{12}^2-M_{01}^2-M_{02}^2=-D^2+\tfrac{1}{4},\\ &\Gamma_{45}^2-\Gamma_{41}^2-\Gamma_{51}^2=\Gamma_{23}^2+\tfrac{1}{4}.\end{aligned} \qquad (1.30)$$

If $\{\Psi_\alpha(x)\}$ is an orthonormal (ON) basis for the Hilbert space of positive energy solution of (1.1), then (in the sense of distributions)

$$\sum_\alpha \overline{\Psi_\alpha(x)}\Psi_\alpha(x')=\Delta_+(x-x')=\frac{1}{16\pi^2}\int_{-\infty}^{\infty}\exp[ik\cdot(x'-x)]\,d\mu(\mathbf{k}), \qquad (1.31)$$

where the distribution Δ_+ defined by (1.31) has the explicit expression

$$\Delta_+(x)=\begin{cases}\dfrac{2\pi i}{(t^2-r^2)^{1/2}}, & t>r,\\ -\dfrac{2\pi i}{(t^2-r^2)^{1/2}}, & t<-r,\quad r=(x_1^2+x_2^2)^{1/2},\\ \dfrac{2\pi}{(r^2-t^2)^{1/2}}, & -r<t<r.\end{cases} \qquad (1.32)$$

The computation of (1.32) is carried out in analogy with the corresponding result for four-dimensional space-time.[15] It follows immediately that

$$\Psi(x)=\langle\Psi,\Delta_+(x'-x)\rangle, \qquad (1.33)$$

where the integration is carried out over $\mathbf{x}'$.

2. THE ACTION OF THE CONFORMAL GROUP

It is well known that the representation of $so(3,2)$ on $\mathcal{H}_+$ induced by the operators (1.15)–(1.18) exponentiates to a global irreducible unitary representation of a covering group $\check{S}O(3,2)$ of the identity component of $SO(3,2)$.[12] The maximal connected compact subgroup of $\check{S}O(3,2)$ is $SO(3)\times SO(2)$, where $SO(3)$ is generated by Γ_{12}, Γ_{13}, Γ_{23}, and $SO(2)$ by Γ_{45}. We will determine the explicit action of this subgroup on $\mathcal{H}_+$ as well as the action of several other interesting subgroups of $\check{S}O(3,2)$.

The operators M_{01}, M_{02}, M_{12} generate a subgroup of $\check{S}O(3,2)$ isomorphic to $SO(2,1)$. The action of this subgroup on $\mathcal{H}_+$ is determined by

$$\begin{aligned}&(\exp\theta M_{12})f(\mathbf{k})=f(k_1\cos\theta-k_2\sin\theta,k_1\sin\theta+k_2\cos\theta),\\ &(\exp aM_{01})f(\mathbf{k})=f(k_1(a),k_2),\\ &k_1(a)=[e^a(k_1+k_0)^2-e^{-a}k_2^2]/2(k_1+k_0),\quad f\in\mathcal{H}_+.\end{aligned} \qquad (2.1)$$

(The result for M_{02} follows easily from that for M_{01}.)

The P_a generate a translation subgroup of $\check{S}O(3,2)$:

$$(\exp\textstyle\sum a_\alpha P_\alpha)f(\mathbf{k})=\exp(ia\cdot k)f(\mathbf{k}). \qquad (2.2)$$

Unitary operators of the form $\exp\sum a_\alpha K_\alpha$ are somewhat more difficult to compute explicitly. However, the subgroup $SO(2,1)$, (2.1), transforms the vector a under the adjoint action and there are only three distinct cases to consider: (1) $a=(a_0,0;0)$, $a_0\neq 0$, timelike; (2) $a=(0,a_1,0)$, $a_1\neq 0$, spacelike; (3) $a=(a_1,a_1,0)$, lightlike.

We start with the timelike case. Note that the quantities $f_{l_1l_2}$,

$$\begin{aligned}&f_{l_1l_2}(\mathbf{k})=\delta(k_1-l_1)\delta(k_2-l_2)k_0,\quad -\infty<l_j<\infty,\\ &P_jf_{l_1l_2}=-l_jf_{l_1l_2},\quad j=1,2,\quad P_0f_{l_1l_2}=il_0f_{l_1l_2},\end{aligned} \qquad (2.3)$$

form a basis for $\mathcal{H}_+$ of generalized eigenvectors of the commuting operators P_α. The orthogonality relation is

$$\begin{aligned}&\langle f_{l_1l_2},f_{s_1s_2}\rangle=\delta(l_1-s_1)\delta(l_2-s_2)l_0,\\ &l_0=(l_1^2+l_2^2)^{1/2}.\end{aligned} \qquad (2.4)$$

Thus, the quantities $g_{l_1l_2}=Rf_{l_1l_2}$,

$$g_{l_1l_2}(k)=(1/2\pi)\cos\sqrt{2l\cdot k}, \qquad (2.5)$$

form a basis for $\mathcal{H}_+$ of generalized eigenvectors of the commuting operators K_α:

$$\begin{aligned}&K_jg_{l_1l_2}=-il_jg_{l_1l_2},\quad K_0g_{l_1l_2}=-il_0g_{l_1l_2},\\ &\langle g_{l_1l_2},g_{s_1s_2}\rangle=\delta(l_1-s_1)\delta(l_2-s_2)l_0.\end{aligned} \qquad (2.6)$$

(Here we are using the fact that R is unitary.)

The unitary operator $\exp aK_0$ takes the form

$$(\exp aK_0)f(\mathbf{s})=\int_{-\infty}^{\infty}G(a,\mathbf{l},\mathbf{s})f(\mathbf{l})\,d\mu(\mathbf{l}),\quad f\in\mathcal{H}_+, \qquad (2.7)$$

where

$$\begin{aligned}G(a,\mathbf{l},\mathbf{s})&=\langle\exp(aK_0)f_{\mathbf{l}},f_{\mathbf{s}}\rangle=\langle R\exp(-aP_0)Rf_{\mathbf{l}},f_{\mathbf{s}}\rangle\\ &=\langle\exp(-aP_0)g_{\mathbf{l}},g_{\mathbf{s}}\rangle\\ &=\frac{1}{4\pi^2}\iint\exp(-iak_0)\cos\sqrt{2\mathbf{l}\cdot k}\cos\sqrt{2\mathbf{s}\cdot k}\,d\mu(\mathbf{k}).\end{aligned} \qquad (2.8)$$

We can evaluate this integral by expanding the cosines,

$$\cos\sqrt{2l\cdot k}=\sum_{n=-\infty}^{\infty}\exp[in(\theta-\varphi)]J_{2n}[2(l_0k_0)^{1/2}],$$

$$l_1+il_2=l_0\exp(i\varphi),\quad k_1+ik_2=k_0\exp(i\theta),$$

and integrating term-by-term. The result is

$$G(a,\mathbf{l},\mathbf{s})=-(i/2\pi a)\exp[i(s_0+t_0)/a]\times\cos[(1/a)\sqrt{2(s_0l_0+s_1l_1+s_2l_2)}]. \qquad (2.9)$$

To compute the action of $\exp aK_1$, we need a basis of generalized eigenvectors of the commuting operators M_{02} and P_1. The basis is

$$h_{\lambda\mu}(\mathbf{k})=(1/\sqrt{2\pi})\delta(S-\mu)\exp(i\lambda T),\quad -\infty<\mu,\lambda<\infty, \qquad (2.10)$$

$$S=k_1,\quad T=\ln(k_0+k_2),\quad d\mu(\mathbf{k})=dS\,dT.$$

Indeed,

$$\begin{aligned}&M_{02}h_{\lambda\mu}=i\lambda h_{\lambda\mu},\quad P_1h_{\lambda\mu}=-i\mu h_{\lambda\mu},\\ &\langle h_{\lambda\mu},h_{\lambda'\mu'}\rangle=\delta(\mu-\mu')\delta(\lambda-\lambda').\end{aligned} \qquad (2.11)$$

A straightforward computation, using the unitarity of R, shows that the $Rh_{\lambda\mu}$ are corresponding eigenfunctions

of M_{02} and K_1 satisfying the same orthogonality relations as the $h_{\lambda\mu}$. Here,

$$M_{02}(Rh_{\lambda\mu}) = i\lambda(Rh_{\lambda\mu}), \quad K_1(Rh_{\lambda\mu}) = -i\mu Rh_{\lambda\mu},$$

$$Rh_{\lambda\mu}(\mathbf{k}) = \frac{\exp(i\lambda T)}{(2\pi)^{3/2}2^{1/2}} \begin{cases} 4(\mu/S)^{i\lambda}\cosh(\lambda\pi)K_{2i\lambda}(2\sqrt{S\mu}), & S\mu > 0, \\ 2(-\mu/S)^{i\lambda}[K_{2i\lambda}(2\exp(-\pi i/2)\sqrt{-S\mu}) & \\ +K_{2i\lambda}(2\exp(\pi i/2)\sqrt{-S\mu})], & S\mu < 0, \end{cases} \tag{2.12}$$

where $K_\lambda(z)$ is a MacDonald function. The unitary operator $\exp aK_1$ assumes the form

$$(\exp aK_1)f(\mathbf{s}) = \int_{-\infty}^{\infty}\int H(a,\mathbf{l},\mathbf{s})f(\mathbf{l})\,d\mu(\mathbf{l}), \tag{2.13}$$

where

$$H(a,\mathbf{l},\mathbf{s}) = \langle \exp(aK_1)f_{\mathbf{l}}, f_{\mathbf{s}}\rangle = \int_{-\infty}^{\infty}\int \langle f_{\mathbf{l}}, \exp(-aK_1)Rh_{\mu\lambda}\rangle$$

$$\langle Rh_{\mu\lambda}, f_{\mathbf{s}}\rangle\, d\mu\, d\lambda = \int_{-\infty}^{\infty}\int \exp(-i\mu a)\overline{Rh}_{\mu\lambda}(\mathbf{l})Rh_{\mu\lambda}(\mathbf{s})\,d\mu\,d\lambda$$

$$= \frac{1}{8\pi|a|}\exp[-i(s_1+l_1)/a]$$

$$\times\cos\left(\frac{s_1(l_2+l_0)-l_1(s_2+s_0)}{a(s_2+s_0)^{1/2}(l_2+l_0)^{1/2}}\right), \tag{2.14}$$

as follows from a tedious computation.

The computation of $\exp[a(K_0+K_1)]$ can most conveniently be carried out in Sec. 5, where we relate this operator to the free-particle Schrödinger equation.

The dilatation operator D generates the one-parameter group $\exp aD$,

$$(\exp aD)f(\mathbf{k}) = \exp(a/2)f(e^a\mathbf{k}). \tag{2.15}$$

We can now easily exponentiate the compact generator $\Gamma_{45} = \frac{1}{2}(P_0 - K_0)$. Indeed, the operators P_0, D, and K_0 generate an $SL(2,R)$ subgroup of $\tilde{SO}(3,2)$. It is easy to verify the relation

$$\exp 2\theta\Gamma_{45} = \exp(\tan\theta\, P_0)\exp(-\sin\theta\cos\theta K_0)$$

$$\times\exp(-2\ln\cos\theta\, D) \tag{2.16}$$

on $SL(2,R)$, and, evaluating the right-hand side of this expression, we find

$$(\exp 2\theta\Gamma_{45})f(\mathbf{k}) = \frac{i\csc\theta}{2\pi}\exp[-i(k_0+l_0)\cot\theta]$$

$$\times\cos[\csc\theta\sqrt{2(k_0l_0+k_1l_1+k_2l_2)}]f(\mathbf{l})\,d\mu(\mathbf{l}),$$

$$\theta \neq n\pi. \tag{2.17}$$

Similarly, the operators P_1, D, and K_1 generate an $SL(2,R)$ subgroup of $\tilde{SO}(3,2)$ and one can verify the relation

$$\exp 2\theta\Gamma_{12} = \exp(\tan\theta\, P_1)\exp(\sin\theta\cos\theta K_1)\exp(-2\ln\cos\theta\, D),$$

$$2\Gamma_{12} = K_1 + P_1,$$

or

$$(\exp 2\theta\Gamma_{12})f(\mathbf{k}) = \frac{\exp\{-ik_1[(\sin^2\theta+1)/\sin\theta\cos\theta]\}}{8\pi|\sin\theta|}$$

$$\times\iint \exp(-il_1\cot\theta)$$

$$\times\cos\left[\frac{k_1(l_2+l_0)-l_1(k_2+k_0)}{\sin\theta\,(k_2+k_0)^{1/2}(l_2+l_0)^{1/2}}\right]f(\mathbf{l})\,d\mu(\mathbf{l}),$$

$$\theta \neq n\pi/2. \tag{2.18}$$

The operators (2.18) together with the operators $\exp\theta M_{12}$, (2.1), determine the action of the $SO(3)$ subgroup.

3. THE LAPLACE OPERATOR

On restriction of our irreducible representation of $\tilde{SO}(3,2)$ to the compact subgroup $SO(3)$ this representation decomposes into a direct sum of irreducible representations D_l of $SO(3)$, $\dim D_l = 2l+1$. We will determine a convenient basis for $\mathcal{H}_+$, which exhibits this decomposition. This is a basis of eigenvectors of the commuting operators Γ_{45} and Γ_{23}:

$$\Gamma_{45}f = i\lambda f, \quad \Gamma_{23}f = imf, \quad -i\Gamma_{45} = \tfrac{1}{2}k_0(-\partial_{k_1k_1} - \partial_{k_2k_2} + 1). \tag{3.1}$$

By setting $k_1 = k_0\cos\theta$, $k_2 = k_0\sin\theta$, it is easy to show that the ON basis of eigenvectors is

$$f^{(l,m)}(\mathbf{k}) = [(l-m)!/\pi(l+m)!\,]^{1/2}(2k_0)^m\exp(-k_0)$$

$$\times L_{l-m}^{(2m)}(2k_0)\exp(im\theta), \tag{3.2}$$

$$\lambda = l + \tfrac{1}{2}, \quad l = 0, 1, 2, \cdots, \quad m = -l, -l+1, \ldots, l.$$

Here, $L_n^{(\alpha)}(z)$ is a Laguerre polynomial.

From this result and the first of Eqs. (1.30) we see that the $\{f^{(l,m)}: m = l, l-1, \ldots, -l\}$ for fixed l form an ON basis for the representation D_l of $SO(3)$. Furthermore, on restriction to $SO(3)$ our representation decomposes as

$$\sum_{l=0}^{\infty}\oplus D_l.$$

From the known recurrence relations for Laguerre polynomials we find

$$\Gamma_{15}f^{(l,m)} = \tfrac{1}{2}\sqrt{(l-m+1)(l+m+1)}\,f^{(l+1,m)}$$

$$-\tfrac{1}{2}\sqrt{(l-m)(l+m)}\,f^{(l-1,m)},$$

$$\Gamma_{42}f^{(l,m)} = -\tfrac{1}{4}\sqrt{(l+m+2)(l+m+1)}\,f^{(l+1,m+1)} \tag{3.3}$$

$$+\tfrac{1}{4}\sqrt{(l-m)(l-m-1)}\,f^{(l-1,m+1)} + \tfrac{1}{4}\sqrt{(l+m)(l+m-1)}$$

$$\times f^{(l-1,m-1)} - \tfrac{1}{4}\sqrt{(l-m+1)(l-m+2)}\,f^{(l+1,m-1)}.$$

Using (3.1), (3.3) and taking commutators, we can compute the action of any $\Gamma_{\alpha\beta}$ on this basis.

Note the close connection between the eigenvalue equation $\Gamma_{45}f = i\lambda f$ and the quantum Kepler problem in two-dimensional space:

$$Hg = \mu g, \quad H = -\partial_{xx} - \partial_{yy} + e/r, \quad r = (x^2+y^2)^{1/2},$$

$$\iint |g|^2\,dx\,dy < \infty. \tag{3.4}$$

The two eigenvalue equations can be identified provided we set $k_1 = x\sqrt{-\mu}$, $k_2 = y\sqrt{-\mu}$, $\mu = -e^2/4\lambda^2$. The eigenvalue problems are defined on Hilbert spaces with different inner products, but from the Virial theorem[16] we see that if the energy eigenvalue μ belongs to the point spectrum of H and g is a corresponding eigenvector, then g also has finite norm in $\mathcal{H}_+$. Conversely, if f is an eigenfunction of Γ_{45}, then $\iint |f|^2\,dx\,dy < \infty$ and f cor-

responds to an energy eigenvalue μ in the point spectrum of H. Since the eigenvalues λ of Γ_{45} are $\lambda = l + \frac{1}{2}$, $l = 0, 1, 2, \cdots$, it follows that the point eigenvalues of H are $\mu_l = -e^2/4(l+\frac{1}{2})^2$. Although this is a satisfying explanation of the point spectrum of H, it sheds no light on the continuous spectrum of H since Γ_{45} has only point spectrum.

Using (1.20), we can compute the corresponding ON basis of positive energy solutions $\psi^{(l,m)}(x)$ of (1.1):

$$\psi^{(l,m)}(x) = 1/4\pi \int\int_{-\infty}^{\infty} \exp(ik\cdot x) f^{(l,m)}(\mathbf{k})\, d\mu(\mathbf{k})$$
$$= [(l-m)!/4\pi(l+m)!]^{1/2} \exp[im(\alpha - \pi/2)] \quad (3.5)$$
$$\times \int_0^{\infty} \exp[(ix_0 - 1)k_0](2k_0)^m J_m(k_0 r) L_{l-m}^{(2m)}(2k_0)\, dk_0.$$

$x_1 = r\cos\alpha$, $x_2 = r\sin\alpha$.

In terms of the coordinates

$$x_0 = \frac{\sin\Psi}{\cos\sigma - \cos\Psi}, \quad x_1 = \frac{\sin\sigma\cos\alpha}{\cos\sigma - \cos\Psi},$$
$$x_2 = \frac{\sin\sigma\sin\alpha}{\cos\sigma - \cos\Psi}, \quad (3.6)$$

variables R-separate in (1.1) and (3.5) to give

$$\Psi^{(l,m)} = i\sqrt{(l-m)!/8\pi(l+m)!}\sqrt{\cos\sigma - \cos\Psi}$$
$$\times \exp[-i\Psi(l+\tfrac{1}{2})] \exp[im(\alpha - \pi/2)] P_l^m(\cos\sigma)$$
$$= [(-i)^{m-1}/\sqrt{4l+2}]\sqrt{\cos\sigma - \cos\Psi}\, \exp[-i\Psi(l+\tfrac{1}{2})] Y_l^m(\sigma, \alpha), \quad (3.7)$$

where P_l^m is an associated Legendre function and Y_l^m is a spherical harmonic. (We can always parametrize so that $\cos\sigma - \cos\Psi > 0$.) Indeed, in our three-variable model we find

$$\Gamma_{45} = -\partial_\psi + \tfrac{1}{2}\sin\psi/(\cos\sigma - \cos\psi), \quad \Gamma_{23} = \partial_\alpha. \quad (3.8)$$

Thus $\psi^{(l,m)}(x) = \sqrt{\cos\sigma - \cos\psi}\exp[-i\psi(l+\frac{1}{2})]\exp(im\alpha)g(\sigma)$, and, substituting into (1.1), we see that variables R-separate and $g(\sigma)$ is a linear combination of $P_l^m(\cos\sigma)$ and $Q_l^m(\cos\sigma)$. Evaluating the integral (3.5) for special values of the parameters, e.g., $\sigma = 0, \pi$, we establish (3.7).

For future use we point out that there is another model of our irreducible representation of $\tilde{SO}(3,2)$ in which the eigenfunctions of Γ_{45} and Γ_{23} take an especially simple form. The representation space is the Bargmann Hilbert space $\mathcal{J}_2$ consisting of all entire functions $h(z_1, z_2)$ such that[17]

$$\int_{C\times C} |h|^2\, d\xi(\mathbf{z}) < \infty, \quad d\xi(\mathbf{z}) = (\exp[-(|z_1|^2 + |z_2|^2)]/\pi^2)$$
$$\times dx_1\, dx_2\, dy_1\, dy_2,$$
$$z_j = x_j + iy_j, \quad j = 1, 2. \quad (3.9)$$

The inner product is

$$\langle f, h\rangle = \int_{C\times C} \bar{f}h\, d\xi(\mathbf{z}).$$

The carrier space for our representation is not $\mathcal{J}_2$ but the subspace $\mathcal{J}_2^+$ consisting of all $h \in \mathcal{J}_2$ such that $h(-z_1, -z_2) = h(z_1, z_2)$. The functions

$$f^{(l,m)}(z_1, z_2) = z_1^{l+m} z_2^{l-m}/\sqrt{(l+m)!\,(l-m)!},$$
$$l = 0, 1, 2, \cdots, \quad m = l, l-1, \ldots, -l, \quad (3.10)$$

form an ON basis for $\mathcal{J}_2^+$. Setting

$$\Gamma_{45} = (i/2)(z_1\partial_{z_1} + z_2\partial_{z_2} + 1), \quad \Gamma_{15} = \tfrac{1}{2}(z_1 z_2 - \partial_{z_1 z_2}),$$
$$\Gamma_{23} = (i/2)(z_1\partial_{z_1} - z_2\partial_{z_2}), \quad \Gamma_{42} = \tfrac{1}{4}(\partial_{z_1 z_1} + \partial_{z_2 z_2} - z_1^2 - z_2^2) \quad (3.11)$$

and comparing with expressions (3.3), we see that we have a new model of our representation of $\tilde{SO}(3,2)$ in which the functions $f^{(l,m)}(\mathbf{k})$ can be identified with the functions (3.10). The explicit unitary mapping U from $\mathcal{H}^+$ to $\mathcal{J}_2^+$ which commutes with the group action is

$$Uf(z_1, z_2) = \int\int_{R_2} U(\mathbf{k}, \mathbf{z}) f(\mathbf{k})\, d\mu(\mathbf{k}), \quad f \in \mathcal{H}^+, \quad (3.12)$$

where

$$U(\mathbf{k}, \mathbf{z}) = \sum_{l,m} \bar{f}^{(l,m)}(\mathbf{k}) f^{(l,m)}(\mathbf{z})$$
$$= [\exp(k_0 + z_1 + z_2)/\sqrt{\pi}]\cosh[\sqrt{2k_0}$$
$$\times (z_1 \exp(-i\theta/2) - z_2 \exp(i\theta/2))], \quad (3.13)$$

$k_1 = k_0\cos\theta$, $k_2 = k_0\sin\theta$.

For convenience we will list the pairs of commuting second order elements in the Lie algebra of $\tilde{SO}(3,2)$ which correspond to each coordinate system we discuss. Thus the system (3.6), (3.7) corresponds to the operators

1) Γ_{45}^2, Γ_{23}^2.

There is also a Lamé-type coordinate system related to the $SO(3)$ subgroup and determined by

2) Γ_{45}^2, $\Gamma_{12}^2 + a^2\Gamma_{13}^2$, $a \neq 0$,

which we shall not treat here. The relationships between 1) and 2) are discussed in Refs. 4 and 9. These systems correspond to the eigenvalue equation for the Laplace operator on the sphere S_2.

4. DIAGONALIZATION OF P_0, P_2, AND D

We next look for those coordinate systems permitting separation of variables in (1.1) such that the corresponding basis functions Ψ are eigenfunctions of P_0: $P_0\Psi = i\lambda\Psi$. For such systems we have $\Psi(x) = \exp(i\lambda x_0) \times \Phi(x_1, x_2)$, where

$$(\partial_{11} + \partial_{22} + \lambda^2)\Phi = 0. \quad (4.1)$$

Thus the equation for the eigenfunctions reduces to the Helmholtz equation (4.1). Now P_0 commutes with every element in the Euclidean subalgebra $\mathcal{E}(2)$ generated by P_1, P_2, and M_{12}. The Euclidean group in the plane $E(2)$ with Lie algebra $\mathcal{E}(2)$ is the symmetry group of (4.1). It is known that this equation separates in exactly four coordinate systems, each system corresponding to a symmetric second- order element in the enveloping algebra of $\mathcal{E}(2)$. See Refs. 1, 6, and 18 for discussion of these matters as well as listings of the coordinates and the related eigenfunctions. The pairs of commuting second order symmetric operators associated with these systems are

3) P_0^2, P_1^2, Cartesian,

4) P_0^2, M_{12}^2, polar,

5) $P_0^2,\ M_{12}P_2+P_2M_{12}$, parabolic cylinder,

6) $P_0^2,\ M_{12}^2+P_2^2$, elliptic.

On $\mathcal{H}^+$ the requirement $P_0 f = i\lambda f$ implies $f(\mathbf{k}) = \delta(k_0-\lambda)g_\lambda(\theta)$, where $\lambda>0$, $k_1=k_0\cos\theta$, $k_2=k_0\sin\theta$. The search for the functions g_λ reduces to a study of the Hilbert space $L_2[0,2\pi]$ on which $E(2)$ acts via

$$P_1=-i\lambda\cos\theta,\quad P_2=-i\lambda\sin\theta,\quad M_{12}=\partial_\theta. \tag{4.2}$$

As is well known, these operators determine a unitary irreducible representation of $E(2)$ on $L_2[0,2\pi]$. Once the eigenfunctions $g_{\lambda\mu}(\theta)$ of the second operator in 3)–6) have been determined, the corresponding separable solutions $\Psi_{\lambda\mu}$ of (1.1) can be obtained from the relation

$$\Psi_{\lambda\mu}(x)=\frac{\exp(i\lambda x_0)}{4\pi}\int_0^{2\pi}\exp[-i\lambda(x_1\cos\theta+x_2\sin\theta)]g_{\lambda\mu}(\theta)\,d\theta. \tag{4.3}$$

The eigenfunctions $g_{\lambda\mu}$ and the corresponding integrals (4.3) have been worked out in Refs. 1 and 6. For future reference we give the basis 4):

$$P_0\Psi_{\lambda n}=i\lambda\Psi_{\lambda n},\quad M_{12}\Psi_{\lambda n}=in\Psi_{\lambda n},\quad n=0,\pm1,\pm2,\cdots,$$
$$f_{\lambda n}(k_0,\theta)=\delta(k_0-\lambda)\exp(in\theta)/\sqrt{2\pi},\quad \langle f_{\lambda n},f_{\lambda' n'}\rangle=\delta(\lambda-\lambda')\delta_{nn'}, \tag{4.4}$$
$$\Psi_{\lambda n}(x_0,r,\varphi)=[(-i)^n/\sqrt{8\pi}\,]\exp[i(\lambda x_0+n\varphi)]J_n(\lambda r),$$
$$x_1=r\cos\varphi,\quad x_2=r\sin\varphi.$$

Now we search for coordinate systems allowing separation of variables in (1.1) such that the basis functions Ψ are eigenfunctions of P_2: $P_2\Psi=-i\gamma\Psi$. Here we have $\Psi(x)=\exp(-i\gamma x_2)\Phi(x_0,x_1)$, where

$$(\partial_{00}-\partial_{11}+\gamma^2)\Phi=0. \tag{4.5}$$

The operator P_2 commutes with the pseudo-Euclidean subalgebra $\mathcal{E}(1,1)$ generated by P_0, P_1, and M_{01} and, indeed, the pseudo-Euclidean group $E(1,1)$ is the symmetry group of (4.5). This equation separates in nine coordinate systems associated with nine symmetric second order operators in the enveloping algebra of $\mathcal{E}(1,1)$. Details on the coordinates and basis functions are given in Refs. 1, 3, and 10. The pairs of commuting operators associated with the corresponding solutions of (1.1) are

3)′ $P_2^2,\ P_0P_1$,

7) $P_2^2,\ M_{01}^2$,

8) $P_2^2,\ M_{01}P_0+P_0M_{01}$,

9) $P_2^2,\ M_{01}^2-(P_0+P_1)^2$,

10) $P_2^2,\ M_{01}^2+(P_0+P_1)^2$,

11) $P_2^2,\ M_{01}(P_0-P_1)+(P_0-P_1)M_{01}-(P_0+P_1)^2$,

12) $P_2^2,\ M_{01}^2-P_0P_2$,

13) $P_2^2,\ M_{01}^2+P_1^2$,

14) $P_2^2,\ M_{01}^2-P_1^2$.

The case 3)′ is equivalent to 3).

On $\mathcal{H}^+$ the requirement $P_2 f=-i\gamma f$ implies $f(\mathbf{k})=\delta(k_2-\gamma)g_\gamma(\xi)$, where $-\infty<\gamma<\infty$, $k_1=|k_2|\sinh\xi$, $k_0=|k_2|\cosh\xi$. The search for eigenfunctions reduces to a study of the Hilbert space $L_2(R)$ on which $E(1,1)$ acts via

$$P_0=i|\gamma|\cosh\xi,\quad P_1=-i|\gamma|\sinh\xi,\quad M_{01}=\partial_\xi. \tag{4.6}$$

These operators define a unitary irreducible representation of $E(1,1)$ on $L_2(R)$. After the eigenfunctions $g_{\gamma\mu}(\xi)$ of the second operator in 7)–14) have been determined, the corresponding separable solutions $\Psi_{\gamma\mu}$ of (1.1) follow from

$$\Psi_{\gamma\mu}(x)=\frac{\exp(-i\gamma x_2)}{4\pi}\int_{-\infty}^{\infty}\exp[i|\gamma|(x_0\cosh\xi-x_1\sinh\xi)]\times g_{\gamma\mu}(\xi)\,d\xi. \tag{4.7}$$

The eigenfunctions $g_{\gamma\mu}$ and the integrals (4.7) are computed in Ref. 3 for cases 7)–12) and various overlaps between these bases have been determined. Here we give only the basis 7):

$$P_2\Psi_{\gamma\mu}=-i\gamma\Psi,\quad M_{01}\Psi_{\gamma\mu}=i\mu\Psi_{\gamma\mu},\quad -\infty<\gamma,\mu<\infty,$$
$$f_{\gamma\mu}(k_2,\xi)=\delta(k_2-\gamma)\exp(i\mu\xi)/\sqrt{2\pi},$$
$$\langle f_{\gamma\mu},f_{\gamma'\mu'}\rangle=\delta(\gamma-\gamma')\delta(\mu-\mu'), \tag{4.8}$$
$$\Psi_{\gamma\mu}(x_2,\rho,\theta)=(1/\sqrt{8\pi^4})\exp[i(\mu\theta-\gamma x_2)]K_{i\mu}(i|\gamma|\rho),$$
$$x_0=\rho\cosh\theta,\quad x_1=\rho\sinh\theta.$$

Next we look for coordinate systems yielding separation of variables in (1.1) such that the basis functions Ψ are eigenfunctions of D: $D\Psi=-i\nu\Psi$. In this case we have $\Psi(x)=\rho^{i\nu-1/2}\Phi(s_0,s_1,s_2)$, where

$$x_\alpha=\rho s_\alpha,\quad \rho\geq0,\quad s_0^2-s_1^2-s_2^2=\epsilon$$

and $\epsilon=+1,-1$, or 0 depending on whether $x\cdot x>0$, <0 or $=0$. It follows from the second of Eqs. (1.30) that

$$(M_{12}^2-M_{01}^2-M_{02}^2)\Phi(s)=(\nu^2+\tfrac{1}{4})\Phi(s). \tag{4.9}$$

Now D commutes with the subalgebra $so(2,1)$ generated by M_{12}, M_{01}, and M_{02}, and in fact $SO(2,1)$ is the symmetry group of (4.9). This equation separates in nine coordinate systems associated with nine symmetric second order operators in the enveloping algebra of $so(2,1)$. The details for the case $\epsilon=+1$ are worked out in Refs. 4 and 10. The pairs of commuting operators associated with separated solutions of (1.1) are

15) $D^2,\ M_{12}^2$, spherical,

16) $D^2,\ M_{01}^2$, equidistant,

7)′ $D^2,\ (M_{12}-M_{02})^2$, horocyclic,

17) $D^2,\ M_{12}^2+a^2M_{01}^2$, elliptic,

18) $D^2,\ M_{01}^2-a^2M_{12}^2$, hyperbolic $(0<a<1)$

19) $D^2,\ -M_{12}M_{02}-M_{02}M_{12}+aM_{01}^2$, semihyperbolic $(0<a<\infty)$

20) $D^2,\ aM_{01}^2+M_{02}^2+M_{12}^2-M_{02}M_{12}-M_{12}M_{02}$, elliptic-parabolic $(0<a)$

21) $D^2,\ -aM_{01}^2+M_{02}^2+M_{12}^2-M_{02}M_{12}-M_{12}M_{02}$, hyperbolic-parabolic $(0<a)$

22) D^2, $M_{02}M_{01}+M_{01}M_{02}-M_{01}M_{12}-M_{12}M_{01}$, semicircular-parabolic.

[System 7)′ is equivalent to 7).]

On $\mathcal{H}^+$ the requirement $Df=-i\nu f$ implies $f(\mathbf{k})=k_0^{-i\nu-1/2}h_\nu(\theta)$, where $-\infty<\gamma<\infty$, $k_1=k_0\cos\theta$, $k_2=k_0\sin\theta$. The eigenfunction problem thus reduces to a study of the Hilbert space $L_2[0,2\pi]$ on which $SO(2,1)$ acts via

$$M_{12}=\partial_\theta,\quad M_{01}=-\sin\theta\partial_\theta-(i\nu+\tfrac{1}{2})\cos\theta,$$
$$M_{01}=\cos\theta\partial_\theta-(i\nu+\tfrac{1}{2})\sin\theta. \tag{4.10}$$

These operators define a unitary irreducible representation of $SO(2,1)$ which is single-valued and belongs to the principal series: $l=-\frac{1}{2}+i|\nu|$. Once the eigenfunctions $h_{\nu\alpha}(\theta)$ of the second operator in 15)–22) have been determined, the corresponding separable solutions $\Psi_{\nu\alpha}$ of (1.1) can be obtained from

$$\Psi_{\nu\alpha}(x)=\frac{\rho^{i\nu-1/2}}{4\pi}\Gamma(\tfrac{1}{2}-i\nu)\int_0^{2\pi}\exp[\pm i\pi(\tfrac{1}{2}-i\nu)/2]\,|s_0-s_1\cos\theta-s_2\sin\theta|^{i\nu-1/2}h_{\nu\alpha}(\theta)\,d\theta, \tag{4.11}$$

where the plus sign occurs when $s_0-s_1\cos\theta-s_2\sin\theta>0$ and the minus sign occurs when this expression is <0. We list explicitly the basis 17):

$$D\Psi_{\nu m}=-i\nu\Psi_{\nu m},\quad M_{12}\Psi_{\nu m}=im\Psi_{\nu m},\quad -\infty<\nu<\infty,\quad m=0,\pm1,\cdots,$$
$$f_{\nu m}(k_0,\theta)=k_0^{-i\nu-1/2}\exp(im\theta)/2\pi,\quad \langle f_{\nu m},f_{\nu'm'}\rangle=\delta(\nu-\nu')\delta_{mm'},$$
$$\Psi_{\nu m}(x)=\frac{\rho^{i\nu-1/2}}{4\pi}\Gamma(-i\nu-m+\tfrac{1}{2})\exp[i\pi(\tfrac{1}{2}-i\nu)/2]\times P^m_{i\nu-1/2}(\cosh a)\exp(im\varphi), \tag{4.12}$$
$$x_0=\rho\cosh a,\quad x_1=\rho\sinh a\cos\varphi,\quad x_2=\rho\sinh a\sin\varphi.$$

Here, the expression for $\Psi_{\nu m}(x)$ is the one valid for $x\cdot x>0$, $x_0>0$. There is a similar result for the case $x\cdot x>0$, $x_0<0$, but the expression for the case $x\cdot x<0$ is somewhat more complicated:

$$\Psi_{\nu m}(x)=\frac{\rho^{i\nu-1/2}}{\sqrt{8\pi}}\left(\frac{\exp(-i\pi/4-\nu\pi/2)+(-1)^m\exp(i\pi/4+\nu\pi/2)}{\exp(-\pi\nu)+\exp(\pi\nu)}\right)\times(\cosh a)^{-1/2}P^{-i\nu}_{-m-1/2}(\tanh a)\exp(im\varphi), \tag{4.13}$$
$$x_0=\rho\sinh a,\quad x_1=\rho\cosh a\cos\varphi,$$
$$x_2=\rho\cosh a\sin\varphi,\quad a>0.$$

For $a<0$, $\Psi_{\nu m}(x)$ is equal to expression (4.13) multiplied by $(-1)^m$ and a replaced with $-a$. Finally, for $x\cdot x=0$, $x_0>0$ we have

$$\Psi_{\nu m}(x)=\frac{\exp[i\pi(\frac{1}{2}-i\nu)/2]}{4\pi\sqrt{2\pi}}\Gamma(-i\nu-m+\tfrac{1}{2})\times\left[\exp[-(\tfrac{1}{2}+i\nu)a]\frac{\Gamma(-i\nu)}{\Gamma(\frac{1}{2}-m-i\nu)}+\exp[(-\tfrac{1}{2}+i\nu)a]\frac{\Gamma(i\nu)}{\Gamma(\frac{1}{2}-m+i\nu)}\right]\exp(im\varphi) \tag{4.14}$$
$$x_0=e^a,\quad x_1=e^a\cos\varphi,\quad x_2=e^a\sin\varphi.$$

5. THE SCHRÖDINGER EQUATION

Of special interest are the coordinate systems permitting separation of variables in (1.1) such that the basis functions Ψ are eigenfunctions of P_0+P_1: $(P_0+P_1)\Psi=i\beta\Psi$. For this case we have $\Psi(x)=\exp(is\beta)\Phi(t,x_2)$, where $2s=x_0+x_1$, $2t=x_1-x_0$. The function Φ satisfies the free particle Schrödinger equation

$$(i\beta\partial_t+\partial_{x_2x_2})\Phi(t,x_2)=0. \tag{5.1}$$

This equation admits as symmetries the operators

$$P=P_2,\quad K_{-2}=-P_0+P_1,\quad \mathcal{E}=P_0+P_1,\quad \beta=\tfrac{1}{2}(M_{02}-M_{12}),$$
$$D=-D-M_{01},\quad K_2=-\tfrac{1}{4}(K_0+K_1), \tag{5.2}$$

which all commute with $P_0+P_1=\mathcal{E}$. These operators form a basis for the six-dimensional Schrödinger algebra of (5.1). Indeed the script notation and the commutation relations for the basis agree with that found in Ref. 5, where equation (5.1) was analyzed. The pairs of commuting operators associated with separable systems for (5.1) are (deleting squares as in Ref. 5):

3)″ P_0+P_1, P_2, free particle,

23) P_0+P_1, $P_0-P_1-\frac{1}{4}K_0-\frac{1}{4}K_1$, oscillator,

24) P_0+P_1, $P_0-P_1+aM_{12}-aM_{02}$, $a\neq0$, linear potential,

25) P_0+P_1, $D+M_{01}$, repulsive oscillator.

[The so-called free-particle coordinates 3)″ are the same as 3).] On $\mathcal{H}^+$ the requirement $(P_0+P_1)f=i\beta f$ implies $f(\mathbf{k})=u\delta(u-\beta)l_\beta(v)$, where $\beta>0$, $u=k_0-k_1$, $v=k_2$. Thus, the search for the l_β reduces to a study of the Hilbert space $L_2(R)$ on which the Schrödinger group acts via

$$\mathcal{E}=i\beta,\quad P=-iv,\quad \beta=\tfrac{1}{2}\beta\partial_v,\quad D=-\tfrac{1}{2}-v\partial_v,$$
$$K_{-2}=-iv^2/\beta,\quad K_2=-\tfrac{1}{4}\beta\partial_{vv}. \tag{5.3}$$

As shown in Ref. 5, these operators determine an irreducible unitary representation of the Schrödinger group on $L_2(R)$. Once the eigenfunctions $l_{\beta n}(v)$ of the second operators in 23)–25) have been determined, the corresponding separable solutions $\Psi_{\beta n}$ of (1.1) can be computed from

$$\Psi_{\beta n}(x)=\frac{\exp(i\beta s)}{4\pi}\int_{-\infty}^{\infty}\exp[-i(v^2t/\beta+vx_2)]l_{\beta n}(v)\,dv. \tag{5.4}$$

For reference we list the basis 23):

$$(P_0+P_1)\Psi_{\beta n}=i\beta\Psi_{\beta n},\quad (P_0-P_1-\tfrac{1}{4}K_0-\tfrac{1}{4}K_1)\Psi_{\beta n}=i(n+\tfrac{1}{2})\Psi_{\beta n},$$
$$n=0,1,2,\cdots,\quad f_{\beta n}(u,v)=\frac{u\delta(u-\beta)}{(n!\sqrt{2\pi}\,2^n)^{1/2}}\exp(-v^2/\beta)H_n\times\left(v\left(\frac{2}{\beta}\right)^{1/2}\right), \tag{5.5}$$
$$\langle f_{\beta n},f_{\beta'n'}\rangle=\beta\delta(\beta-\beta')\delta_{nn'},$$
$$\Psi_{\beta n}(s,t,x_2)=\frac{\exp(i\beta s)(-i)^n}{4(n!\,2^n\pi\sqrt{2\pi})^{1/2}}\left(\frac{\beta}{1+it}\right)^{1/2}\left(\frac{1-it}{it+1}\right)^{n/2}\times\exp\left(\frac{-x_2^2\beta}{4(1+it)}\right)H_n\left(x_0\left(\frac{\beta}{2(1+t^2)}\right)^{1/2}\right).$$

Using the u,v coordinates we can easily compute $\exp[a(K_0+K_1)]$. Indeed from the expression for K_2 in (5.3) and formula (3.8) of Ref. 5 we find

$$\{\exp[a(K_0+K_1)f]\}(k_1,k_2)$$

$$= \frac{1}{[4\pi i a(k_0-k_1)]^{1/2}} \int_{-\infty}^{\infty} \exp\left(\frac{-(k_2-w)^2}{4ia(k_0-k_1)}\right) \times f\left(\frac{w^2-(k_0-k_1)^2}{2(k_0-k_1)}, w\right) dw, \quad f \in \mathcal{H}^+. \tag{5.6}$$

6. THE EPD EQUATION

Next we look for coordinate systems yielding separation of variables for (1.1) such that the basis functions Ψ are eigenfunctions of M_{12}: $M_{12}\Psi = im\Psi$. We have $\Psi(x) = \exp(im\varphi)\Phi(x_0, r)$, where

$$x_1 = r\cos\varphi, \quad x_2 = r\sin\varphi,$$

and Φ satisfies the Euler–Poisson–Darboux equation

$$\left(\partial_{00} - \partial_{rr} - \frac{1}{r}\partial_r + \frac{m^2}{r^2}\right)\Phi = 0 \tag{6.1}$$

or

$$(\Gamma_{45}^2 - \Gamma_{41}^2 - \Gamma_{51}^2)\Phi = (\Gamma_{23}^2 + \tfrac{1}{4})\Phi = -(m+\tfrac{1}{2})(m-\tfrac{1}{2})\Phi \tag{6.2}$$

from the last of expressions (1.30). The symmetry group of (6.1) is $SL(2,R)$ and is generated by the symmetry operators $\Gamma_{45}, \Gamma_{41}, \Gamma_{51}$.

The coordinate systems in which (1.1) separates via (6.1) are characterized by the following pairs of commuting operators [$\Gamma_{23} = M_{12}$, $\Gamma_{15} = D$, $\Gamma_{45} = \frac{1}{2}(P_0 - K_0)$, $\Gamma_{14} = \frac{1}{2}(P_0 + K_0)$]:

$$\begin{aligned}
&1)' \quad \Gamma_{23}^2, \ \Gamma_{45}^2,\\
&4)' \quad \Gamma_{23}^2, \ (\Gamma_{45}+\Gamma_{14})^2,\\
&17)' \quad \Gamma_{23}^2, \ \Gamma_{15}^2,\\
&26) \quad \Gamma_{23}^2, \ 2\Gamma_{14}^2 + \Gamma_{45}\Gamma_{14} + \Gamma_{14}\Gamma_{45},\\
&27) \quad \Gamma_{23}^2, \ 2\Gamma_{45}^2 + \Gamma_{45}\Gamma_{14} + \Gamma_{14}\Gamma_{45},\\
&28) \quad \Gamma_{23}^2, \ \Gamma_{14}^2 + a(\Gamma_{45}\Gamma_{15} + \Gamma_{15}\Gamma_{45}), \quad a \neq 0,\\
&29) \quad \Gamma_{23}^2, \ \Gamma_{45}^2 + a\Gamma_{15}^2, \quad a \neq 0,\\
&30) \quad \Gamma_{23}^2, \ a\Gamma_{14}^2 + \Gamma_{15}^2, \quad a \neq 0,\\
&31) \quad \Gamma_{23}^2, \ (\Gamma_{14}+\Gamma_{45})\Gamma_{15} + \Gamma_{15}(\Gamma_{14}+\Gamma_{45}).
\end{aligned} \tag{6.3}$$

These statements will be proved and the corresponding coordinate systems derived in another publication.

On $\mathcal{H}^+$ the requirement $M_{12}f = imf$ implies $f(\mathbf{k}) = \exp(im\theta)j_m(k_0)$ where $m = 0, \pm 1, \pm 2, \cdots$, $k_1 = k_0\cos\theta$, $k_2 = k_0\sin\theta$. The eigenfunction problem reduces to a study of the Hilbert space $L_2[0,\infty]$ on which $SL(2,R)$ acts via

$$\begin{aligned}
\Gamma_{45} &= \frac{i}{2}k_0\left(-\partial_{k_0k_0} - \frac{1}{k_0}\partial_{k_0} + \frac{m^2}{k_0^2} + 1\right),\\
\Gamma_{14} &= \frac{i}{2}k_0\left(\partial_{k_0k_0} + \frac{1}{k_0}\partial_{k_0} - \frac{m^2}{k_0^2} + 1\right),\\
\Gamma_{15} &= \tfrac{1}{2} + k_0\partial_{k_0}.
\end{aligned} \tag{6.4}$$

This action is irreducible and unitary equivalent to a single-valued representation of $SL(2,R)$, not $SO(2,1)$, from the negative discrete series $D^-_{|m|-1/2}$, as can be seen from (6.2) and (3.2). Indeed, the eigenvalues of Γ_{45} in this model are $i(n+\frac{1}{2})$, $n = |m|, |m|+1, |m|+2, \cdots$. This model of D_l^- has been studied by a number of authors, e.g., Refs. 19, 20, but the connection with separation of variables in the EPD equation has not been pointed out before. In another publication we will use this and other models of the discrete series D_l^- to study the spectra of the operators (6.3) and derive special function expansions related to the EPD equation.

7. DIAGONALIZATION OF $\Gamma_{23} - \Gamma_{45}$

Finally we look for separable solutions of (1.1) such that the basis functions χ are eigenfunctions of $L = \frac{1}{2}(\Gamma_{23} - \Gamma_{45})$: $L\chi = i\kappa\chi$. Introducing the coordinates (3.6) and setting $\chi(x) = \sqrt{\cos\sigma - \cos\Psi}\,\Phi$, we find

$$\left(\frac{1}{\sin\sigma}\partial_\sigma(\sin\sigma\partial_\sigma) - \partial_{\Psi\Psi} - \tfrac{1}{4} + \frac{1}{\sin^2\sigma}\partial_{\alpha\alpha}\right)\Phi = 0. \tag{7.1}$$

Now we choose as independent coordinates σ, β, ρ where $\beta = \alpha + \Psi$, $\rho = \alpha - \Psi$. In terms of these coordinates, the induced action of L on Φ is $L = \partial_\beta$ and the solutions of $L\chi = i\kappa\chi$ are $\chi(x) = \sqrt{\cos\sigma - \cos\Psi}\,\exp(i\kappa\beta)\Theta$, where

$$[\sin^2\sigma\partial_{\sigma\sigma} + \cos\sigma\sin\sigma\partial_\sigma + (\tfrac{1}{4} - \kappa^2)\cos^2\sigma - 2i\kappa(\sin^2\sigma + 1)\partial\rho + \cos^2\sigma\partial_{\rho\rho} - \tfrac{1}{4}]\Theta(\sigma,\rho) = 0. \tag{7.2}$$

The symmetry algebra of (7.2) is $sl(2,R)$ with basis

$$A = \tfrac{1}{2}(\Gamma_{23} + \Gamma_{45}), \quad B = \tfrac{1}{2}(\Gamma_{24} + \Gamma_{35}), \quad C = \tfrac{1}{2}(\Gamma_{25} - \Gamma_{34}) \tag{7.3}$$

and commutation relations

$$[A,B] = C, \quad [C,A] = B, \quad [C,B] = A. \tag{7.4}$$

Moreover, it is straightforward to verify the identity

$$A^2 - B^2 - C^2 = L^2 + \tfrac{1}{4}. \tag{7.5}$$

In terms of the coordinates σ, β, ρ we also have $A = \partial_\rho$ as the action of A on the solutions of (7.2). From (3.1), (3.2) we see that the spectrum of L is given by $\kappa = \frac{1}{2}(s+\frac{1}{2})$, $s = 0, 1, 2, \cdots$, and for fixed κ the eigenvalues μ of $-iA$, $A\Theta = i\mu\Theta$, are $\mu = l + \frac{1}{4} - s/2$, $l = 0, 1, 2, \cdots$. In terms of (7.5), Eq. (7.2) becomes

$$(A^2 - B^2 - C^2)\Theta = -(-\kappa - \tfrac{1}{2})(-\kappa + \tfrac{1}{2})\Theta \tag{7.6}$$

so that the solutions of (7.2) form the basis space for a model of the representation $D^-_{-\kappa-1/2}$, (negative discrete series), of the twofold covering group of $SL(2,R)$.

The problem of separation of variables for (7.2) has not been settled yet. We will investigate this problem in a future paper to see if there are for this symmetry algebra exactly nine separable systems corresponding to the nine orbits of second-order operators just as found in Secs. 4 and 6. At the moment we know definitely only the subgroup systems:

$$\begin{aligned}
&1)' \quad L^2, \ A^2,\\
&32) \quad L^2, \ B^2,\\
&33) \quad L^2, \ (A+B)^2.
\end{aligned}$$

Except for 1)' our $\mathcal{H}_+$-model is not very convenient for systems of this type. Thus, in later publications we shall employ the model (3.11) and other models of the discrete series, such as those found in Refs. 19 and 21, to study this case. This conclude our list of semisubgroup coordinate systems in which variables separate in (1.1).

8. CONCLUDING REMARKS

In the previous sections we have classified all separable coordinate systems for the wave equation such that the defining operators take the form A^2, S where $A \in so(3,2)$, S is a second order symmetric element in the enveloping algebra of $so(3,2)$ and $[A,S]=0$. As explained in the Introduction, $SO(3,2)$ acts on the pair via the adjoint representation to generate an orbit of $SO(3,2)$—equivalent pairs of commuting second order operators. Coordinate systems corresponding to equivalent pairs of operators are considered as equivalent.

By choosing appropriate examples it is easy to show that there are elements A in $so(3,2)$ for which there is no S such that the pair A^2, S corresponds to a separable coordinate system. To see why this is the case, we classify the orbits in $so(3,2)$ under the adjoint action of $SO(3,2)$. This classification has been obtained in principle by Zassenhaus,[22] but we present the results here in a much more explicit form. Indeed, we list the possible eigenvalues of a 5×5 matrix $A \in so(3,2)$ and for each choice of eigenvalues we list a canonical form $\Gamma \in so(3,2)$ such that $\Gamma = TAT^{-1}$ for some $T \in SO(3,2)$, i.e., we list an element Γ on each $SO(3,2)$ orbit in $so(3,2)$. From the relation $A^t = -GAG$ it follows easily that $\det(A-\lambda E) = -\det(A+\lambda E)$, where $\lambda = \alpha + i\beta \in \mathbb{C}$ and E is the 5×5 identity matrix. Thus, $\lambda = 0$ is always an eigenvalue of A and, if $\lambda \neq 0$ is an eigenvalue, then so are $-\lambda$ and $\bar{\lambda}$. We use the notation $\lambda(n)$, $n=2,3,4,5$, to signify that λ corresponds to a generalized eigenvector x of rank n, i.e., n is the smallest integer m such that $(A-\lambda E)^m x = 0$.

	Possible Eigenvalue	*Canonical Form* $\Gamma \in so(3,2)$
1.	$0, \pm\lambda, \pm\bar{\lambda}$ $\lambda = \alpha + i\beta$, $\alpha, \beta \neq 0$	$\alpha(\Gamma_{24}+\Gamma_{35}) + \beta(\Gamma_{23}-\Gamma_{45})$
2.	$0, \pm\alpha, \pm\beta$ $\beta \neq 0$	$\alpha\Gamma_{24} + \beta\Gamma_{35}$
3.	$0, \pm\alpha, \pm i\beta$ $\alpha, \beta \neq 0$	$\alpha\Gamma_{14} + \beta\Gamma_{23}$
4.	$0, \pm i\alpha, \pm i\beta$ $\alpha, \beta \neq 0$	$\alpha\Gamma_{23} - \beta\Gamma_{45}$
5.	$0, \alpha(2), -\alpha(2)$	$P_0 + P_1 + \alpha(D + M_{01})$
6.	$0, i\alpha(2), -i\alpha(2)$ $\alpha \neq 0$	$2\alpha L + A + B$ [see (7.3)]
7.	$0, 0, 0, \pm i\beta$ $\beta \neq 0$	$\beta\Gamma_{23}$ or $\beta\Gamma_{45}$
8.	$O(3), i\alpha, -i\alpha$ $\alpha \neq 0$	$1/\sqrt{2}(\Gamma_{35}+\Gamma_{45}) + \alpha\Gamma_{12}$
9.	$O(3), \pm\alpha$ $\alpha \neq 0$	$1/\sqrt{2}(\Gamma_{13}+\Gamma_{15}) + \alpha\Gamma_{24}$
10.	$O(3), 0, 0$	P_0 or P_1
11.	$O(5)$	$1/\sqrt{2}(\Gamma_{13}+\Gamma_{15})$ $+\frac{1}{2}(\Gamma_{25}-\Gamma_{23}+\Gamma_{34}+\Gamma_{45})$

Suppose $A = \Gamma$ takes the form 1. It is straightforward to show that the operators $2B = \Gamma_{24}+\Gamma_{35}$ and $2L = \Gamma_{23}-\Gamma_{45}$ commute with each other and that the only nontrivial elements of $so(3,2)$ commuting with A are linear combinations of B and L. Thus the separable coordinate system associated with a pair A^2, S and of the type discussed in the last paragraphs of the Introduction is equivalent to the system associated B^2 and L^2, i.e., the system 32). Similar remarks hold for cases 2—9.

The operator of case 11 lies on the same orbit as the linear potential operator 24): $P_0 - P_1 + aM_{12} - aM_{02}$. This latter operator commutes only with a linear combination of itself and $P_0 + P_1$.

It is clear that we can use our $\mathcal{H}_+$-model to compute the spectra corresponding to pairs A^2, S and to compute overlaps between basis functions corresponding to distinct pairs, in analogy with the procedures developed in Refs. 1—7. We will present these results in forthcoming papers.

[1] W. Miller, Jr., SIAM J. Math. Anal. **5**, 626 (1974).
[2] W. Miller, Jr., SIAM J. Math. Anal. **5**, 822 (1974).
[3] E.G. Kalnins and W. Miller, Jr., J. Math. Phys. **15**, 1025 (1974).
[4] E.G. Kalnins and W. Miller, Jr., J. Math. Phys. **15**, 1263 (1974).
[5] E.G. Kalnins and W. Miller, Jr., J. Math. Phys. **15**, 1728 (1974).
[6] C. Boyer, E.G. Kalnins, and W. Miller, Jr., J. Math. Phys. **16**, 499 (1975).
[7] C. Boyer, E.G. Kalnins, and W. Miller, Jr., J. Math. Phys. **16**, 512 (1975).
[8] M. Bôcher, *Die Reihenentwickelungen der Potentialtheorie* (Leipzig, 1894).
[9] J. Patera and P. Winternitz, J. Math. Phys. **14**, 1130 (1973).
[10] E.G. Kalnins, SIAM J. Math. Anal. **6**, 340 (1975).
[11] A. Erdélyi *et al.*, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vols. I and II.
[12] H.A. Kastrup, Phys. Rev. **140**, B183 (1965); G. Mack and I.T. Todorov, J. Math. Phys. **10**, 2078 (1969); L. Gross, J. Math. Phys. **5**, 687 (1964).
[13] W. Miller, Jr., *Symmetry Groups and Their Applications* (Academic, New York, 1972).
[14] S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Row, Peterson, Evanston, Ill. 1961), Chap. 3.
[15] P.A.M. Dirac, Proc. Camb. Phil. Soc. **30**, 100 (1934).
[16] A.S. Davydov, *Quantum Mechanics* (Addison-Wesley, Reading, Mass., 1965), p. 51.
[17] V. Bargmann, Comm. Pure Appl. Math. **14**, 187 (1961).
[18] P. Winternitz and I. Fris, Yad Fiz. **1**, 889 (1965) [Sov. J. Nucl. Phys. **1**, 636 (1965)].
[19] P. Sally, *Analytic Continuation of the Irreducible Unitary Representations of the Universal Covering Group of SL(2,R)*, AMS Mem., No. 69 (Am. Math. Soc., Providence, R.I., 1967).
[20] W. Montgomery and L. O'Raifeartaigh, J. Math. Phys. **15**, 380 (1974).
[21] I.M. Gel'fand, M.I. Graev, and N. Ya. Vilenkin, *Generalized Functions, Vol. 5* (Academic, New York, 1966).
[22] H. Zassenhaus, Can. Math. Bull. **1**, 31, 101, 183 (1958).

Cylindrically symmetric solutions of a scalar–tensor theory of gravitation

T. Singh

Department of Mathematics, University of Gorakhpur, Gorakhpur 273001, India
(Received 3 December 1974; revised manuscript 2 April 1975)

The cylindrically symmetric solutions for the Einstein–Rosen metric of a scalar–tensor theory proposed by Dunn have been obtained. A method has been given by which one can obtain, under certain conditions, solutions of this scalar–tensor theory from known solutions of the empty space field equations of Einstein's theory of gravitation. It is also found that one of the solutions of the scalar–tensor theory is nonsingular in the sense of Bonnor. Further some special solutions are obtained which reduce to the well-known solution of Levi-Civita and a time dependent solution obtained by Misra and Radhakrishna.

1. INTRODUCTION

Motivated by the ideas of Mach, Brans and Dicke[1] introduced an alternative theory of gravitation involving a scalar function as well as the metric tensor. This scalar–tensor theory is not purely geometrical, however, as the scalar field is introduced in a rather ad hoc manner into the Riemannian manifold.

Several attempts have been made to cast a scalar–tensor theory of gravitation in a wider geometrical context. Brans and Dicke observed in their work the formal connection between their theory and that of Jordan[2] which uses a five-dimensional manifold. Ross[3] has constructed a scalar–tensor theory of gravitation using the Weyl formulation of Riemannian geometry, and Sen and Dunn[4] have introduced a scalar–tensor theory modeled on a modification of Riemannian geometry suggested by Lyra.

Recently Dunn[5,6] has introduced a geometry which differs from the usual Riemannian geometry in that its linear connection has nonvanishing torsion defined in terms of a scalar function. In this way both the metric tensor and the scalar field have a well-defined geometric meaning in the spirit of general relativity. Based on these considerations Dunn[6] has formulated a scalar–tensor theory of gravitation whose field equations are identical in the vacuum case to those given by Dicke[7] in an alternative presentation of the Brans–Dicke theory. Further, Dunn[6] has also found the static spherically symmetric solution of these field equations and it has been found that, with a proper choice of parameter, this theory agrees with experimental results in the three classical tests of red shift, light deflection, and perihelion advance.

In the present paper we have investigated cylindrically symmetric solutions of the field equations of the scalar–tensor theory proposed by Dunn[6] for the Einstein–Rosen metric[8,9] and have given a method by which one can obtain, under certain conditions, solutions of the scalar–tensor theory from known solutions of Einstein's theory of gravitation. It is also found that one of the solutions of the scalar–tensor theory is nonsingular in the sense of Bonnor.[10] Further, some special solutions have also been obtained, two of which reduce in Riemannian space–time to the well-known solution of Levi-Civita[11] and a time dependent solution obtained by Misra and Radhakrishna.[12]

2. FIELD EQUATIONS AND THEIR SOLUTION

This section deals with the general solution to the vacuum field equations of the scalar–tensor theory of gravitation formulated by Dunn[6] in the cylindrically symmetric case with Einstein–Rosen metric. In the regions of space–time with zero charge and mass densities the field equations of this scalar–tensor theory are

$$R_{ij} - \tfrac{1}{2}g_{ij}R = (6k^2/\lambda^2)(\lambda_{,i}\lambda_{,j} - \tfrac{1}{2}g_{ij}\lambda_{,s}\lambda^{,s}), \tag{2.1}$$

$$(\partial/\partial x^s)(\lambda_{,j}\sqrt{-g}\,g^{js}) - (\lambda_{,s}\lambda_{,i}/\lambda)g^{si}\sqrt{-g} = 0, \tag{2.2}$$

where R_{ij} is the Ricci tensor and R the curvature scalar of the metric g_{ij}, λ is scalar field, and k is a constant. If $k=0$ or λ is a constant, the connection of the space–time is metric-preserving and torsion-free, i.e., we have the Riemannian geometry.

The cylindrically symmetric Einstein–Rosen metric is

$$ds^2 = \exp(\eta-\delta)(dt^2 - dr^2) - r^2e^{-\delta}d\phi^2 - e^{\delta}dz^2, \tag{2.3}$$

where δ and η are functions of r and t only and r, ϕ, z, t correspond respectively to x^1, x^2, x^3, x^4 coordinates. The cylindrical symmetry assumed obviously implies that $\lambda_{,2}=\lambda_{,3}=0$, i.e., λ is a function of r and t only.

With the metric (2.3) the field equations (2.1) and (2.2) reduce to the following equations

$$\delta_{11} - \delta_{44} + \delta_1/r = 0, \tag{2.4}$$

$$\eta_1 = \tfrac{1}{2}r(\delta_1^2 + \delta_4^2) - 6k^2r(h_1^2 + h_4^2), \tag{2.5}$$

$$\eta_4 = r(\delta_1\delta_4) - 12k^2r(h_1h_4), \tag{2.6}$$

$$\eta_{11} - \eta_{44} + \tfrac{1}{2}(\delta_1^2 - \delta_4^2) = 6k^2(h_1^2 - h_4^2), \tag{2.7}$$

$$h_{11} - h_{44} + h_1/r = 0, \tag{2.8}$$

where we have put

$$\lambda = e^h. \tag{2.9}$$

The lower suffixes 1 and 4 after an unknown function denote partial differentiation with respect to r and t, respectively. Equation (2.7) can be obtained from Eqs. (2.4), (2.5), (2.6), and (2.8), and so we shall consider the solution of these four equations only. Also the integrability condition for (2.5) and (2.6), viz., $\eta_{14}=\eta_{41}$ is satisfied by virtue of (2.4) and (2.8).

Equations (2.4) and (2.8) are Euclidean wave equations in cylindrical coordinates whose solutions can be

obtained by well-known methods.[13] Equation (2.4) determines δ while (2.8) determines the scalar function h, both representing waves. As is evident from (2.5) and (2.6), η consists of two parts one depending on δ and the other depending on h. This shows that in the scalar–tensor theory[6] the space–time metric (2.3) depends on the tensor field g_{ij} and the scalar field λ both.

There are infinite number of possible combinations of δ and h that can be used to obtain a solution of the scalar–tensor theory. However, if we restrict ourselves to the case when δ and h are functionally related then (2.4) and (2.8) yield

$$\delta = ah + b, \tag{2.10}$$

where a and b are arbitrary constants. In view of (2.10) Eq. (2.4) reduces to (2.8) which determines the form of scalar function h. The use of (2.10) in (2.4) and (2.5) gives

$$\eta_1 = \tfrac{1}{2} r K^2 (h_1^2 + h_4^2), \tag{2.11}$$

$$\eta_4 = r K^2 (h_1 h_4), \tag{2.12}$$

where

$$K^2 = a^2 - 12k^2. \tag{2.13}$$

Now making the substitution

$$h = \psi / K, \quad \text{i.e.,} \quad \lambda = \exp(\psi / K) \tag{2.14}$$

in (2.8), (2.11), and (2.12), we have

$$\psi_{11} - \psi_{44} + \psi_1 / r = 0, \tag{2.15}$$

$$\eta_1 = \tfrac{1}{2} r (\psi_1^2 + \psi_4^2), \tag{2.16}$$

$$\eta_4 = r(\psi_1 \psi_4). \tag{2.17}$$

The integrability condition for (2.16) and (2.17) is satisfied by virtue of (2.15). Hence, whenever ψ is known from (2.15), h and λ are determined from (2.14), η from (2.16) and (2.17), and δ is determined from (2.10) and (2.14), which can be written in the form

$$\delta = [a/(a^2 - 12k^2)^{1/2}]\, \psi + b. \tag{2.18}$$

Equation (2.15) is known to have a solution of the form

$$[AJ_0(\omega r) + BY_0(\omega r)] \cos(\omega t + \epsilon),$$

where A, B, ω, ϵ are constants and $J_0(\omega r)$ and $Y_0(\omega r)$ are Bessel's functions of order zero and of the first and second kind, respectively. Since $Y_0(\omega r) \to -\infty$ as $r \to 0$ we take $B = 0$ and a typical solution of (2.15) is of the form

$$\psi = AJ_0(\omega r) \cos(\omega t + \epsilon). \tag{2.19}$$

The general solution of (2.15) can obviously be obtained by superposing terms of the form (2.19) and from that the solution corresponding to any desired situation can be obtained by a proper choice of the arbitrary constants.

Consider now the space–time metric

$$ds^2 = \exp(\eta - \psi)(dt^2 - dr^2) - r^2 e^{-\psi} d\phi^2 - e^{\psi} dz^2, \tag{2.20}$$

where ψ and η are functions of r and t only. The empty space field equations of Einstein's theory corresponding to (2.20) reduce to (2.15), (2.16), and (2.17) only. Hence of the five equations (2.14)–(2.18) taken along with the space–time metric (2.3) which represent an empty space solution of the scalar–tensor theory, three equations, viz., (2.15), (2.16), and (2.17) taken along with the metric (2.20) represent an empty space–time in Einstein's theory of gravitation. Thus, we have established the following result:

"For every solution of (2.15), (2.16), and (2.17) corresponding to the metric (2.20) which represents an empty space–time in Einstein's theory of gravitation, we have a solution given by (2.3), (2.14), and (2.18), ψ and η remaining the same, which represents an empty space–time in the scalar–tensor theory.[6]"

3. A NONSINGULAR SOLUTION OF THE SCALAR-TENSOR THEORY

Einstein and Rosen[8] and Rosen[9] have obtained solutions of the wave equations of the type (2.15) which lead to particular cases of the metric (2.20) corresponding to progressive or stationary gravitational waves. These solutions contain singularity along the axis of z, presumably representing the source of the waves. Later Bonnor[10] obtained a nonsingular solution of (2.15) by adapting the procedure applied by Synge.[14] Bonnor[10] has shown that Eqs. (2.15)–(2.17) have nonsingular solution given by

$$\psi = 2\sqrt{2}\, c \{[u + (u^2 + u^2)^{1/2}] / (u^2 + v^2)\}^{1/2}, \tag{3.1}$$

$$\eta = -\{[2c^2 r^2 (u^2 - v^2)] / (u^2 + v^2)^2\} + (c^2/l^2)\{[(r^2 - t^2 - l^2) / (u^2 + v^2)^{1/2}] + 1\}, \tag{3.2}$$

where

$$u = r^2 - t^2 + l^2, \quad v = 2lt, \tag{3.3}$$

c and l being arbitrary constants. Hence, corresponding to a nonsingular solution of Einstein's empty space field equations given by (2.20), (3.1), (3.2), and (3.3), we have a solution of the scalar–tensor theory given by (2.3), (2.14), (2.18), (3.1), (3.2), and (3.3), which is also nonsingular in the sense of Bonnor.[10]

4. SOME SPECIAL SOLUTIONS

In this section we shall obtain some special solutions of the scalar–tensor theory.[6] We consider the field equations (2.4), (2.5), (2.6), and (2.8) along with the relation (2.9) and the space–time metric (2.3) and making certain assumptions we obtain the corresponding exact solutions.

Case (i)

δ and h are functions of r only. Then from (2.4) and (2.8) we have

$$\delta = -qh \tag{4.1}$$

and

$$h = \log(r / r_0), \quad \text{i.e.,} \quad \lambda = r / r_0, \tag{4.2}$$

where q and r_0 are constants. These give

$$\delta = -q \log(r / r_0). \tag{4.3}$$

The use of (4.2) and (4.3) in (2.5) and (2.6) yields

$$\eta = (\tfrac{1}{2} q^2 - 6k^2) \log(r / r_0). \tag{4.4}$$

Thus (4.2), (4.3), and (4.4) with metric (2.3) constitute the solution of the scalar–tensor theory. This solution is also valid for $k=0$ if we restrict the possible values of q such that $q \neq 0$. In this case the geometry of the space–time is Riemannian and the space–time metric becomes

$$ds^2 = \left(\frac{r}{r_0}\right)^{\frac{q^2+2q}{2}}(dt^2 - dr^2) - \left(\frac{r}{r_0}\right)^q r^2 d\phi^2 - \left(\frac{r}{r_0}\right)^{-q} dz^2. \tag{4.5}$$

This is the well-known Levi–Civita solution[11] corresponding to a line mass placed along the z axis.

Case (ii)

δ and h are functions of t only. In this case the solution of the field equations is

$$h = k_1 t, \quad \text{i.e.,} \quad \lambda = k_1^0 e^t, \tag{4.6}$$

$$\delta = k_2 t \tag{4.7}$$

and

$$\eta = \tfrac{1}{2}(k_2^2 - 6k^2 k_1^2) r^2 t^2; \tag{4.8}$$

where k_1, k_1^0, and k_2 are arbitrary constants.

Case (iii)

δ is a function of t only and h is a function of r only. In this case the solution of the field equations is

$$\delta = -mt, \tag{4.9}$$

$$h = k_3 \log r, \tag{4.10}$$

and

$$\eta = m^2 r^2/4 - 6k^2 k_3^2 \log r, \tag{4.11}$$

where m and k_3 are arbitrary constants.

If we take $k=0$, the geometry of the space–time is Riemannian and the corresponding space–time metric is

$$ds^2 = \exp(m^2 r^2/4 + mt)(dt^2 - dr^2) - r^2 \exp(mt)\, d\phi^2 - \exp(-mt)\, dz^2.$$

This is the metric mentioned by Misra and Radhakrishna.[12] This is a nonstatic solution of Einstein's empty space field equations and is of type II according to Pirani's criterion.[15]

ACKNOWLEDGMENTS

The author is thankful to Professor K. B. Lal, Professor and Head, Department of Mathematics, Gorakhpur University, Gorakhpur, India, for constant encouragement and keen interest in this work. Sincere thanks are also due to the referee for pointing out to the author a recent paper by Rao *et al.*[16] of which he was unaware at the time of writing this paper.

[1] C. Brans and R. H. Dicke, Phys. Rev. **124**, 925 (1961).
[2] P. Jordan, *Schwerkraft und Weltall* (Vieweg and Sohn, Braunschweig, 1955).
[3] D. K. Ross, Phys. Rev. D **5**, 284 (1972).
[4] D. K. Sen and K. A. Dunn, J. Math. Phys. **12**, 578 (1971).
[5] K. A. Dunn, "A Geometric Model for Scalar–Tensor Theories of Gravitation" (preprint).
[6] K. A. Dunn, J. Math. Phys. **15**, 2229 (1974).
[7] R. H. Dicke, Phys. Rev. **125**, 2163 (1962).
[8] A. Einstein and N. Rosen, J. Franklin Inst. **223**, 43 (1937).
[9] N. Rosen, Bull. Res. Counc. Israel **3**, 328 (1954).
[10] W. B. Bonnor, J. Math. Mech. **6**, 203 (1957).
[11] T. Levi-Civita, C.R. Acad. Lincei **28** *i*, 101 (1919).
[12] M. Misra and L. Radhakrishna, Proc. Nat. Inst. Sci. (India) A**28**, 632 (1962).
[13] C. A. Coulson, *Waves* (Oliver and Boyd, London, 1955), p. 14.
[14] J. L. Synge, *Relativity—The Special Theory* (North-Holland, Amsterdam, 1956), Chap. IX, pp. 358–61.
[15] F. A. E. Pirani, Phys. Rev. **105**, 1089 (1957).
[16] J. R. Rao, R. N. Tiwari, and K. S. Bhamra, Ann. Phys. (N.Y.) **87**, 470 (1974).

The structure of groups with index-3 subgroups and Landau's second theorem

L. V. Meisel and D. M. Gray

Physical Science Division, Benet Weapons Laboratory, Watervliet Arsenal, Watervliet, New York 12189

E. Brown

Rensselaer Polytechnic Institute, Troy, New York 12181
and Physical Science Division, Benet Weapons Laboratory, Watervliet Arsenal, Watervliet, New York 12189
(Received 24 February 1974; revised manuscript received 7 May 1975)

For any group G_0 which contains an index-3 subgroup G, it is shown that either (a) G is an invariant subgroup or (b) G contains an index-2 subgroup G_A, where G_A is an invariant subgroup of G_0. For case (a), G and its cosets give rise to three operators which span a stable three-dimensional subspace of the group algebra which further reduces to three one-dimensional stable subspaces. For case (b), G_A and its cosets give rise to six operators which span a six-dimensional stable subspace of the group algebra which reduces to two one-dimensional and two two-dimensional irreducible stable subspaces of the group algebra. The irreducible representations and the corresponding basis elements of the group algebra are given for both cases. Landau's second theorem pertaining to second order phase transitions is proven.

I. INTRODUCTION

In this paper we discuss some features of groups containing a subgroup consisting of one third of the group elements. We derive the structure of such groups and give explicit irreducible representations which are characteristic of them. Interest in such groups springs from their relevance to the theory of second order phase transitions. In their classic text on statistical physics, Landau and Lifshitz[1] make the statement: "It appears that the following theorem is also true: No second-order phase transition can exist for transitions involving the decrease by a factor three of the number of symmetry elements (owing to the existence of third-order terms in the expansion of the thermodynamic potential)." In a recent review article, Cracknell[2] conjectured that this theorem probably could not be proven in the general case.

Thus, we were motivated to a general study of such groups and to a proof of Landau's theorem. We have obtained two distinct proofs, one of which is based on group representation theory which is presented here. The other proof will appear separately.[3]

II. STRUCTURE AND REPRESENTATIONS

We shall denote sets of group elements by capital letters and members of sets by corresponding lower case letters, e.g., $h \in H$ or $h_A \in H_A$. The order of any such set will be denoted as an absolute value, e.g., the order of set H is denoted $|H|$. We shall be discussing a group G_0/G. In the second case, there exists a subgroup with $|G| = N$ and the distinct left cosets $H = h_1 G$ and $K = k_1 G$. In an obvious notation $G_0 = G + H + K$.

As is shown in the Appendix there are two situations that can occur. In one case G is an invariant subgroup, so that G, H, and K are the elements of the factor group G_0/G. In the second case, there exists a subgroup of G, namely G_A, where $|G_A| = |G|/2$, which is an invariant subgroup of G_0. For this case the factor group G_0/G_A is of order six. The sets G, H, and K each divide in half respectively into sets G_A, G_B, H_A, H_B, K_A, and K_B which are the elements of this factor group. The group multiplication table for this case is shown in Table I.

For this second case, we introduce the operators $\tilde{G}_A$, $\tilde{G}_B$, $\tilde{H}_A$, $\tilde{H}_B$, $\tilde{K}_A$, and $\tilde{K}_B$. The first is defined as $|G_A|^{-1}\sum g_A$, with corresponding definitions for the rest. Under multiplication these operators form a group isomorphic to G_0/G_A, which has three classes, viz., $\tilde{G}_A$, $(\tilde{H}_A, \tilde{K}_A)$, $(\tilde{G}_B, \tilde{H}_B, \tilde{K}_B)$. Regarded as elements of the group algebra,[4] the six operators are bases for a six-dimensional subspace, which reduces like the regular representation, into two one-dimensional and two two-dimensional irreducible subspaces.

We introduce the operators $\tilde{G}_+$, $\tilde{G}_-$, $\tilde{H}_+$, etc., in which $\tilde{G}_\pm = \tilde{G}_A \pm \tilde{G}_B$, and correspondingly for the others. We express the operators that span the irreducible subspaces Γ_{1+}, Γ_{1-}, Γ_2, and Γ_2' in terms of these operators as follows:

$$\Gamma_{1+}: \tilde{A}_+ = \tilde{G}_+ + \tilde{H}_+ + \tilde{K}_+;$$

$$\Gamma_{1-}: \tilde{A}_- = \tilde{G}_- + \tilde{H}_- + \tilde{K}_-;$$

$$\Gamma_2: \tilde{E}_1 = \sqrt{1/6}(2\tilde{G}_+ - \tilde{H}_+ - \tilde{K}_+),$$
$$\tilde{E}_2 = \sqrt{1/2}(\tilde{H}_+ - \tilde{K}_+);$$

$$\Gamma_2': \tilde{E}_1' = \sqrt{1/2}(\tilde{H}_- - \tilde{K}_-),$$
$$\tilde{E}_2' = \sqrt{1/6}(2\tilde{G}_- - \tilde{H}_- - \tilde{K}_-).$$

TABLE I. Multiplication table for the set operators. The table gives the product operation of an element (set) on the left and an element (set) at the top, e.g., $K_A H_B = G_B$.

	G_A	G_B	H_A	H_B	K_A	K_B
G_A	G_A	G_B	H_A	H_B	K_A	K_B
G_B	G_B	G_A	K_B	K_A	H_B	H_A
H_A	H_A	H_B	K_A	K_B	G_A	G_B
H_B	H_B	H_A	G_B	G_A	K_B	K_A
K_A	K_A	K_B	G_A	G_B	H_A	H_B
K_B	K_B	K_A	H_B	H_A	G_B	G_A

TABLE IIa. Character table for the irreducible representations of the set operator group for the case that $|G_B| = 0$ [$\omega = \exp(2\pi i/3)$].

	G_A	H_A	K_A
Γ_0	1	1	1
Γ	1	ω^*	ω
Γ^*	1	ω	ω^*

TABLE IIb. Character table for the irreducible representations of the set operator group for the case that $|G_B| = |G_A|$. Read C_{H_A} as the class of H_A, etc.

	C_{G_A}	$2C_{H_A}$	$3C_{G_B}$
Γ_{1+}	1	1	1
Γ_{1-}	1	1	−1
Γ_2 or Γ_2'	2	−1	0

The case in which G is an invariant subgroup corresponds to the second case except that G_B, H_B, and K_B are absent. In this case there are three one-dimensional representations of the factor group. This can be thought of as a limit of the second case in which $\tilde{G}_A = \tilde{G}_+ = \tilde{G}_- = \tilde{G}$, etc. It is of interest to note that the two-dimensional subspace reduces into the one-dimensional subspaces:

$$\Gamma: \tilde{E}_1 + i\tilde{E}_2 = \sqrt{2/3}(\tilde{G}_A + \omega\tilde{H}_A + \omega^*\tilde{K}_A)$$

and

$$\Gamma^*: \tilde{E}_1 - i\tilde{E}_2 = \sqrt{2/3}(\tilde{G}_A + \omega^*\tilde{H}_A + \omega\tilde{K}_A),$$

where $\omega = \exp(2\pi i/3)$. The characters of the irreducible representations are given in Table II.

III. PROOF OF LANDAU'S SECOND THEOREM

In Landau's theory[1] of second order phase transitions the density function ρ is expressed as $\rho = \rho_0 + \delta\rho$, where ρ_0 has the symmetry of the large group G_0. The covering group G of the lower symmetry phase is the subgroup of G_0, which leaves $\delta\rho$ (and thus ρ) invariant. Furthermore, it was shown that $\delta\rho$ can be expressed as a linear combination of basis functions for a single irreducible representation other than the identical representation of G_0 (i.e., $\delta\rho = c\phi$, where ϕ is a basis function for an irreducible representation of G_0 other than the identical representation).

From an investigation of the transformation properties of ϕ it is a straightforward matter to see that ϕ and its partners $\phi_H = h\phi$ and $\phi_K = k\phi$ are basis functions for an irreducible representation of the factor group G_0/G in the first case and G_0/G_A in the second case. (N.B.: the projection operator for the identical representation must give zero on ϕ, i.e., $(\tilde{G}_+ + \tilde{H}_+ + \tilde{K}_+)\phi = C(\phi + \phi_H + \phi_K) = 0$, where $C = 1$ in the first case and 2 in the other so that ϕ may have at most one independent partner.) As shown by Landau, a continuous transition cannot occur if one can construct a third order polynomial, in terms of ϕ and its partner functions, which is invariant under the operations of G_0. Such a third order invariant can always be constructed if the symmetrized cube of the appropriate irreducible representation contains the identical representation of G_0.[5] Using the formula[6]

$$\chi_S^3(R) = (1/6)[2\chi(R^3) + 3\chi(R^2)\chi(R) + \chi(R)^3]$$

for the character of the symmetrized cube, and denoting the symmetrized cube of representation μ by $[\mu]^3$, one can show that $[\Gamma]^3 = [\Gamma^*]^3 = \Gamma_0$, the identical representation of G_0; and $[\Gamma_2]^2 = \Gamma_{1+} + \Gamma_{1-} + \Gamma_2$, where Γ_{1+} is the identical representation of G_0. We need not consider Γ_{1-} since ϕ cannot be a basis function for this representation. Thus, the identical representation is contained in the symmetrized cube of each representation for which ϕ can be a basis function, proving the theorem.

APPENDIX

We want to show that: (1) When G is not invariant, it always contains an invariant subgroup G_A, where $|G_A| = |G|/2$, and (2) the factor group G/G_A has the multiplication table given in Table I. We give an outline of the proof below. It is useful to consult the table as we establish its properties.

When G is not invariant, the left cosets G, $H = hG$, and $K = kG$ will not coincide with the right cosets G, $H' = Gh$, and $K' = Gk$. Without loss in generality we assume $H \neq H'$. Clearly these cosets have at least one element in common. Thus, there exists a subset of G, which we call G_A, such that $hG_A = G_A h$. It is straightforward to show that G_A is a subgroup. We denote the complex consisting of the remaining elements of G by G_B. We examine the effect of G_A on the cosets $\{G,H,K\}$ and find $G_A\{G,H,K\} = \{G,H,K\}$. Similarly, $G_B\{G,H,K\} = \{G,K,H\}$. (These results follow directly from the rearrangement theorem.) Furthermore, it can be shown that H and K divide into complexes H_A, H_B, K_A, and K_B respectively, such that $H_A\{H,K\} = \{K,G\}$, $H_B\{H,K\} = \{G,K\}$, $K_A\{H,K\} = \{G,H\}$, and $K_B\{H,K\} = \{H,G\}$. This clearly exhausts all possibilities since G plays the role of a right identity. The multiplication properties of these complexes can be deduced directly from the above. For example, $H_A K_B = G_B$. In this way one constructs the multiplication table given in Table I. From the fact that $G_B G_B = G_A$ we conclude that $|G_B| \leq |G_A|$ since there must be at least $|G_B|$ terms in any product involving G_B from the rearrangement theorem. Using $G_B G_A = G_B$, we similarly conclude that $|G_A| \leq |G_B|$. Thus $|G_A| = |G_B|$. By similar arguments we see that there are $|G_A|$ elements in each complex, each of which is clearly a coset of G_A, which is seen to be an invariant subgroup.

[1] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon, London, 1958).
[2] A. P. Cracknell, Advan. Phys. **23**, 673 (1974).
[3] E. Brown and L. V. Meisel, submitted to Phys. Rev.
[4] A review of group algebra concepts may be found in Per-Olov Löwdin, Rev. Mod. Phys. **39**, 259 (1967).
[5] The pertinent representations of G_0 are those that correspond to the factorization given here. These representations are unfaithful in that the same representative matrix is used for each element in a given coset. This matrix is the same one faithful in that the same representative matrix is used for each element in a given coset. This matrixis the same one that is used in the factor group for the element corresponding to that coset. Thus there is a correspondence in the character of these representations of G_0 with those of the factor group.
[6] G. Ya Lyubarskii, *The Application of Group Theory in Physics* (Pergamon, London, 1960).

Local and covariant quantizations of linearized Einstein's equations. II

L. Bracci

Istituto di Fisica dell'Università, Pisa, Italy

F. Strocchi

INFN, Sezione di Pisa and Scuola Normale Superiore, Pisa, Italy
(Received 26 December 1974)

The local and covariant quantizations of the linearized Einstein's equations, in which the space-time translations may not commute with the metric operator, are analyzed. A general characterization of the subspace H' of physical vectors is given in terms of the two point function, independently of the Fock space representation of the gravitational potential $h_{\mu\nu}$. Under very general conditions all the H' are found to coincide with that of the Gupta formulation, thus showing that the physical content is the same as in the Gupta case. The analog of the Landau gauge in quantum gravity is defined and explicitely constructed through a Fock space representation of $h_{\mu\nu}$. Spinless gravitons can occur only in a quantization which exhibits unconventional features.

1. INTRODUCTION

A detailed analysis of the possible local and covariant quantizations of the linearized Einstein's equations (LEE) was given in a previous paper,[1] hereafter referred as I, under very general conditions. A technical assumption underlying the above classification was that the space—time translations were described by unitary operators in the Fock space in which the gravitational potential

$$h(f) \equiv \int h_{\mu\nu}(x) f^{\mu\nu}(x)\, d^4x$$

was defined as a local and covariant operator. The motivation for the above assumption was that it allowed a standard spectral representation[2] of the space—time translations and an easy exploitation of the spectral condition.[3] In the present paper we want to extend our analysis to the case in which nonunitary representations of space—time translations are allowed. This possibility cannot be ruled out by general physical arguments and in fact a careful investigation on the meaning of physical symmetry in theories with indefinite metric (a feature forced by locality and covariance) legitimates the use of operators which are η-unitary but not unitary.[4]

Our analysis will be performed along the lines of Wightman quantum field theory.[3] In this framework, a local and covariant quantization of the gravitational potential, or a gauge for $h_{\mu\nu}(x)$, is specified by a Hilbert space H, a sesquilinear form $\langle \cdot, \cdot \rangle$ and a distinguished subspace H', such that $\langle \cdot, \cdot \rangle \geq 0$ on H' and the linearized Einstein's equations hold as mean values on H'.[5] The arbitrariness in the choice of the gauge is therefore reduced to the definition of $h_{\mu\nu}(x)$ as an operator valued distribution in a Hilbert space H, and to the selection of the subspace H', whose elements are candidates to describe physical states. The Fock space realization of $h_{\mu\nu}(x)$ is determined by the form of the two point function $\langle \psi_0, h_{\mu\nu}(x) h_{\rho\sigma}(y) \psi_0 \rangle$, via a generalized reconstruction theorem. Therefore, a preliminary step in the classification of the possible gauges is the determination of the most general form of the two point function (Sec. 2). In Sec. 3 we show that under very general conditions the subspace H' is uniquely determined and we give its characterization in general. A particularly interesting feature is that the subspaces H' of all the possible gauges coincide with the subspace H' of the Gupta gauge[1,7] and therefore all the theories are equivalent as long as one sticks to H'. As a consequence, the spin content of the theory is the same as that of the Gupta gauge, i.e., only the helicities ± 2 occur in the physical space H'/H''. Thus, in contrast to statements appeared in the literature,[8] a local and covariant quantization of LEE does not admit physical gravitons of zero helicity, without exhibiting some pathological feature.

In Sec. 4 we give a Fock space representation of a particular gauge which is the analog of the Landau gauge in quantum electrodynamics.

In Sec. 5 we discuss the case of theories which deviate from the above characterization. They all have some pathological feature:

(i) the vacuum is not a cyclic vector, for the elements of H', with respect to the polynomials in the smeared fields $R_{\alpha\beta\gamma\delta}(f)$;

(ii) the mean values $\langle \phi, \Box h_{\mu\nu}(f) \psi \rangle$ do not vanish, in general, for $\phi, \psi \in H'$.

2. LOCAL AND COVARIANT QUANTIZATION OF LINEARIZED EINSTEIN'S EQUATIONS. DEFINITIONS AND BASIC ASSUMPTIONS

In this paper we will follow Wightman approach[3] to quantum field theory as close as possible, even if not all the Wightman axioms can be satisfied. A detailed discussion and motivation of the following basic definitions may be found in Ref. 1.

Definition 1: A local and covariant quantization of the linearized Einstein's equations or a gauge for $h_{\mu\nu}(x)$ is specified by:

(a) An operator valued tempered distribution $h_{\mu\nu}(x)$, in a Hilbert space H.

(b) A representation U of the Poincaré group in $\mathcal{H}$.

(c) A nondegenerate sesquilinear form $\langle \cdot, \cdot \rangle$ on $\mathcal{H}$ with respect to which the representation U is unitary.

(d) a distinguished subspace $\mathcal{H}' \subset \mathcal{H}$ such that

(i) the restriction of the sesquilinear form $\langle \cdot, \cdot \rangle$ to $\mathcal{H}'$ is bounded and nonnegative

$$\langle \psi, \psi \rangle \geq 0 \quad \text{for } \psi \in \mathcal{H}',$$

(ii) the Riemann operators $R_{\mu\nu\rho\sigma}(f) = \int R_{\mu\nu\rho\sigma}(x) f(x) d^4x$, $f \in \mathcal{S}$, defined by

$$R_{\mu\nu\rho\sigma} = \tfrac{1}{2}(g_{\alpha\mu} g_{\beta\sigma} \partial_\nu \partial_\rho + g_{\alpha\nu} g_{\beta\rho} \partial_\mu \partial_\sigma - g_{\alpha\mu} g_{\beta\rho} \partial_\nu \partial_\sigma - g_{\alpha\nu} g_{\beta\sigma} \partial_\mu \partial_\rho) h^{\alpha\beta}, \tag{1}$$

leave $\mathcal{H}'$ invariant and the Einstein's equations hold as mean values in $\mathcal{H}'$

$$\langle \phi, g^{\mu\sigma} R_{\mu\nu\rho\sigma}(f) \psi \rangle = 0 \tag{2}$$

for all $\phi, \psi \in \mathcal{H}'$, with ψ in the domain of $R_{\mu\nu\rho\sigma}(f)$;

(iii) the representation U leaves $\mathcal{H}'$ invariant; there exists a unique vector ψ_0, the vacuum vector, which is invariant under the representation U, lies in $\mathcal{H}'$, and is a cyclic vector for the operators $h_{\mu\nu}(f)$.

(e) The Fourier transform of $\langle \psi_0, h_{\mu\nu}(x) h_{\rho\sigma}(y) \psi_0 \rangle$ has support contained in the future light cone $\bar{V}_+$.

(f) The operators $h_{\mu\nu}(x)$ obey local commutativity

$$[h_{\mu\nu}(x), h_{\rho\sigma}(y)] = 0 \quad \text{for } (x-y)^2 < 0 \tag{3}$$

and transform covariantly under U

$$U(a, \Lambda) h_{\mu\nu}(x) U(a, \Lambda)^{-1} = \Lambda_\mu^{-1\rho} \Lambda_\nu^{-1\sigma} h_{\rho\sigma}(\Lambda x + a). \tag{4}$$

(g) The physical states are elements of the quotient space $\mathcal{H}_{\text{phys}} \equiv \mathcal{H}'/\mathcal{H}''$, on which the representation $U(a, \Lambda)$ becomes unitary.

As it is well known, the LEE in vacuum essentially describe a free field theory. Thus, we may assume that all the Wightman functions are determined by the two point function. The classification of all the possible quantizations of LEE can therefore be done in terms of the two point function as discussed in Sec. 3.

Theorem 1: In a local covariant quantization of the linearized Einstein's equations in vacuum the most general form of the two point function is the following:

$$\begin{aligned} W_{\mu\nu\rho\sigma}(x-y) &= \langle \psi_0, h_{\mu\nu}(x) h_{\rho\sigma}(y) \psi_0 \rangle \\ &= -ic(g_{\mu\rho} g_{\nu\sigma} + g_{\nu\rho} g_{\mu\sigma} - g_{\mu\nu} g_{\rho\sigma})[D^+(x-y) + a(x-y)^2] \\ &+ (g_{\mu\nu} \partial_\rho \partial_\sigma + g_{\rho\sigma} \partial_\mu \partial_\nu) F(x-y) \\ &+ (g_{\mu\rho} \partial_\nu \partial_\sigma + g_{\nu\rho} \partial_\mu \partial_\sigma + g_{\mu\sigma} \partial_\nu \partial_\rho + g_{\nu\sigma} \partial_\mu \partial_\rho) G_1(x-y) \\ &+ \partial_\mu \partial_\nu \partial_\rho \partial_\sigma G_2(x-y) \end{aligned} \tag{5}$$

where

$$D^+(x-y) = \frac{i}{2(2\pi)^3} \int \exp[-ik(x-y)] \frac{d^3k}{|\mathbf{k}|}$$

and F, G_1, G_2 are Lorentz invariant distributions.

Proof: By covariance and spectral condition $W_{\mu\nu\rho\sigma}$ is of the form[9]

$$\begin{aligned} W_{\mu\nu\rho\sigma}(x) &= (g_{\mu\rho} g_{\nu\sigma} + g_{\nu\rho} g_{\mu\sigma}) F_1 + g_{\mu\nu} g_{\rho\sigma} F_2 \\ &+ g_{\mu\nu} \partial_\rho \partial_\sigma F_3 + g_{\rho\sigma} \partial_\mu \partial_\nu F_4 \\ &+ (g_{\mu\rho} \partial_\nu \partial_\sigma + g_{\nu\rho} \partial_\mu \partial_\sigma + g_{\mu\sigma} \partial_\nu \partial_\rho + g_{\nu\sigma} \partial_\mu \partial_\rho) F_5 \\ &+ \partial_\mu \partial_\nu \partial_\rho \partial_\sigma F_6 . \end{aligned}$$

On the other hand, by condition (d, ii) one must have

$$\langle \psi_0, R_{\mu\nu}(x) R_{\rho\sigma}(y) \psi_0 \rangle = 0.$$

This implies

$$\Box F_1 = \text{const}, \quad \partial_\mu \partial_\nu \partial_\rho \partial_\sigma (F_1 + F_2) = 0,$$

i.e.,

$$F_1 = -icD^+ + ax^2 + b, \quad F_2 = +icD^+ + a'x^2 + b'$$

(a, b, c, a', b' constants, $c > 0$). With the conventional normalization of $R_{\mu\nu\rho\sigma}$ one has $c = 1$. This convention will be adopted in the following, unless otherwise stated. By a proper redefinition of the gauge functions, the terms $a'x^2 + b'$ and b can be eliminated. Finally, weak local commutativity implies[10]

$$W_{\mu\nu\rho\sigma}(x) = W_{\rho\sigma\mu\nu}(x),$$

i.e.,

$$F_3(x) = F_4(x) \equiv F(x).$$

Remark 1: One can easily show that the constant a vanishes if any of the following properties holds:

(i) the cluster property holds in $\mathcal{H}'$, i.e.,

$$\lim_{\lambda \to \infty} \langle \psi_0, h(f) h(g_{\lambda v}) \psi_0 \rangle = 0,$$

where $f_{\mu\nu}$ and $g_{\mu\nu}$ are test functions of compact support such that $h(f)\psi_0 \in \mathcal{H}'$, λ is positive, v is any spacelike vector and $g_{\lambda v} \equiv g(x - \lambda v)$. [Clearly, if $h(g)\psi_0 \in \mathcal{H}'$, also $h(g_{\lambda v})\psi_0 \in \mathcal{H}'$].

(ii) The metric operator η is bounded and it has a bounded inverse.

Proof: In case (i) we choose $f, g \in \mathcal{D}$ and such that $\partial^\mu f_{\mu\nu} = 0$, $\partial^\mu g_{\mu\nu} =$ [this implies $\tilde{f}_{\mu\nu}(0) = \tilde{g}_{\mu\nu}(0) = 0$] and

$$F \equiv 2 \frac{\partial}{\partial k^\rho} \tilde{f}_{\mu\nu} \frac{\partial}{\partial k_\rho} \tilde{g}^{\mu\nu} - \frac{\partial}{\partial k^\rho} \tilde{f}_\alpha{}^\alpha \frac{\partial}{\partial k_\rho} \tilde{g}_\beta{}^\beta \Big|_{k=0} \neq 0.$$

Thus,

$$\lim_{\lambda \to \infty} \langle \psi_0, h(f) h(g_{\lambda v}) \psi_0 \rangle = -iacF$$

Condition (i) then implies $a = 0$. In case (ii) the proof runs parallel to that of Ref. 5, Proposition 2.4. One defines a linear functional

$$F_g: \quad F_g(\psi_0) = 0, \quad F_g(h(f)\psi_0) = -2a \left(\frac{\partial}{\partial k_\mu} \frac{\partial}{\partial k^\mu} \left[\tilde{f}_{\rho\sigma}(k) \tilde{g}^{\rho\sigma}(k) \right] \right)_{k=0}$$

and one shows that it is bounded in $\mathcal{H}$. Thus, there exists a vector ϕ_g such that

$$F_g(\psi) = (\phi_g, \psi) = \langle \eta^{-1} \phi_g, \psi \rangle.$$

The function $\tilde{g}_{\mu\nu}(k)$ is chosen in such a way that $\tilde{g}_{\mu\nu}(0) = 0$ and $(\partial \tilde{g}_{\mu\nu}/\partial k^\alpha)(0) = 0$. Therefore,

$$U(a, 1) \eta^{-1} \phi_g = \eta^{-1} \phi_g.$$

By the uniqueness of the vacuum $\eta^{-1}\phi_g = \lambda\psi_0$ and by Lorentz invariance one concludes $a=0$. In the following we will consider the case $a=0$.

If the space–time translations commute with the metric operator (i.e., they are not only η-unitary but also unitary), the Fourier transform of the invariant distributions F, G_1, G_2 are measures. In this case, the Fock space representation of $h_{\mu\nu}(f^{\mu\nu})$ is of the form given by Eq. 10 of Ref. 1, apart from special gauge transformations.[6] In the following the Fourier transform of F, G_1, G_2 are not assumed to be measure (e.g., Sec. 4) and in this respect the present analysis extends that of Ref. 1.

3. CHARACTERIZATION OF H'

In most of the following we will be concerned with theories where the following condition is satisfied:

$$\langle\phi, \Box h_{\mu\nu}(f)\psi\rangle = 0, \quad \forall\, \phi, \psi \in H'. \tag{6}$$

They cover all the gauges known to us and in our opinion, the failing of condition (6) leads to some pathological features, as discussed in Sec. 5.

In this section we will show that a characterization of the maximal subspace H' having the properties listed in Definition 1 may be given in general, without explicitly using the Fock space representation of $h(f)$.

Theorem 2: The state $h(f^1)\cdots h(f^n)\psi_0$ belongs to H' iff the test functions $f^i_{\mu\nu} = f^i_{\nu\mu}$, $i = 1\cdots n$, satisfy the following condition: Either

$$(\alpha)\quad \partial^\mu f^i_{\mu\nu}(x) = 0 \tag{7}$$

or

(β) there exists a $b^i_\nu(x) \in \mathcal{S}(\mathbb{R}^4)$ such that

$$\partial^\mu(f^i_{\mu\nu} - \partial_\mu b^i_\nu - \partial_\nu b^i_\mu + g_{\mu\nu}\partial^\rho b^i_\rho) = 0. \tag{8}$$

Furthermore, if $h(f^1)\cdots h(f^n)\psi_0 \in H'$ and at least one of the test function $f^i_{\mu\nu}$ is of the form

$$f^i_{\mu\nu} = \partial_\mu b^i_\nu + \partial_\nu b^i_\mu - g_{\mu\nu}\partial^\rho b^i_\rho, \tag{9}$$

then the vector has zero norm, i.e., it belongs to H''.

Proof: It suffices to characterize the one particle subspace $H'_1 \subset H'$ since by standard technique one can get the whole H'. H'_1 can be easily characterized by using the two point function: If a vector $h(f)\psi_0$ belongs to H'_1, the following equation must hold as a consequence of Eq. (2):

$$\begin{aligned} 0 &= \langle\psi_0, R_{\mu\nu}(x)h(f)\psi_0\rangle \\ &= \int \exp(-ikx)\, d^4k\{\delta(k^2)\theta(k_0)[2k_\nu k^\lambda \tilde{f}_{\lambda\mu} + 2k_\mu k^\lambda \tilde{f}_{\lambda\nu}] \\ &\quad + \tilde{F}(k)k_\rho k_\sigma \tilde{f}^{\rho\sigma}[k^2 g_{\mu\nu} + 2k_\mu k_\nu]\}. \end{aligned} \tag{10}$$

This implies that $k^\lambda \tilde{f}_{\lambda\mu}$ must be of the form

$$k^\lambda \tilde{f}_{\lambda\mu} = k_\mu \tilde{g} + t_\mu \tag{11}$$

where $\tilde{g}$, $t_\mu \in \mathcal{S}(\mathbb{R}^4)$ and $t_\mu(k) = 0$ on the mantle C_+ of the future light cone. By standard arguments[11,12] one concludes that t_μ may be written in the form

$$t_\mu(k) = k^2 \tilde{u}_\mu(k), \quad \tilde{u}_\mu \in \mathcal{S}. \tag{12}$$

On the other hand, Eq. (6) implies $\Box\, \partial_\mu\partial_\nu\partial_\rho\partial_\sigma F = 0$ and, as a consequence, Eq. (10) requires $\tilde{g}(k) = 0$ on C_+, i.e.,

$$\tilde{g}(k) = k^2\tilde{G}(k), \quad \tilde{G} \in \mathcal{S}.$$

In conclusion Eq. (8) holds with $b_\nu = u_\nu + k_\nu G$.

To show that $h(f)\psi_0$, with $f_{\mu\nu}$ of the form (9), belongs to H'' it suffices to note that by Eq. (2) and Eq. (6) the following "supplementary" condition holds in H'

$$\langle\psi, [\partial^\mu h_{\mu\nu}(x) - \tfrac{1}{2}\partial_\nu h^\rho{}_\rho(x)]\phi\rangle = 0, \tag{13}$$

$\forall\, \psi, \phi \in H'$. Furthermore, for $f_{\mu\nu}$ of the form (9), $h_{\mu\nu}(f^{\mu\nu}) = 2(\partial^\mu h_{\mu\nu} - \frac{1}{2}\partial_\nu h^\rho{}_\rho)(b^\nu)$. This shows that $h(f)\psi_0 \in H''$.

It is easy to see that the subspace H' defined by (α), (β) does indeed have the properties listed in Definition 1 and therefore it is maximal. The characterization of H', as given by Theorem 2 remains unchanged if the constant a of Eq. (5) is not zero.

The above characterization shows that the subspaces H' of all the possible gauges, in which Eq. (6) holds, are essentially identical with the subspace H' of the Gupta gauge. In fact, the indefinite scalar product between vectors of H', induced by the two point function, is the same as in the Gupta formulation for all the two point functions characterized by Theorem 1 and independently of the Fock space representation of $h_{\mu\nu}(f)$.

Corollary 1: Let $(h_{\mu\nu}, H, \langle\cdot,\cdot\rangle, H')$ be a local and covariant gauge in which Eq. (6) holds. Then one can define a linear mapping $V: H' \to H'_G$ (G stands for Gupta) such that $\forall\, \psi, \phi \in H'$, and $\psi_G = V\psi$, $\phi_G = V\phi$,

$$\langle\psi, \phi\rangle = \langle\psi_G, \phi_G\rangle_G.$$

Proof: It is sufficient to prove the statement for vectors of the form $\psi_f = h(f)\psi_0$. Then, by Theorem 2, any test function f, such that $h(f)\psi_0 \in H'$, may be decomposed in the form $f_{\mu\nu} = f^T_{\mu\nu} + f^0_{\mu\nu}$, such that $\partial^\mu f^T_{\mu\nu} = 0$ and $\psi_{f^0} \in H''$. Then if $h^G_{\mu\nu}$ denotes the Fock space representation of $h_{\mu\nu}$ in the Gupta gauge,

$$\langle\psi_f, \psi_g\rangle = \langle h^G(f^T)\psi_0, h^G(g^T)\psi_0\rangle_G.$$

An easy consequence of the above Corollary is that the physical content of the theory is the same for all the local and covariant quantizations discussed above. For a detailed discussion we refer to Ref. 1 where the physical content of the Gupta gauge is analyzed.

In particular we have

Corollary 2: Any local and covariant quantization of the linearized Einstein's equations, in which Eq. (6) holds, gives rise to a quotient space $H_{\text{phys}} = H'/H''$ in which only states with helicity ± 2 can occur.

4. THE ANALOG OF THE LANDAU GAUGE

The relevance of the Landau gauge in the discussion of physical problems in quantum electrodynamics,[13] suggests that an analogous gauge may be of interest for quantum gravitation. One of the important features of the Landau gauge in quantum electrodynamics is that the corresponding propagator selects only transverse photons since the supplementary condition $\partial^\mu A_\mu = 0$ is satisfied as an operator equation. In the local and covariant quantizations characterized in the previous sections the supplementary condition is

$$\langle\partial^\mu h_{\mu\nu}(x) - \tfrac{1}{2}\partial_\nu h^\mu{}_\mu(x)\rangle_{H'} = 0$$

and therefore the analog of the Landau gauge is characterized by the requirement that the equation

$$\partial^\mu h_{\mu\nu}(x) - \tfrac{1}{2}\partial_\nu h^\mu{}_\mu(x) = 0 \tag{14}$$

holds as an operator equation. One might hope to define a local and covariant gauge in which each term of Eq. (14) vanishes:

$$\partial^\mu h_{\mu\nu}(x) = 0, \quad h^\mu{}_\mu(x) = 0, \tag{15}$$

just as in the massive case.[14]

Proposition 1: In a local and covariant gauge for $h_{\mu\nu}(f)$, Eqs. (15) imply that the two point function of $R_{\alpha\beta\gamma\delta}$ vanishes, i.e., the theory is trivial.

Proof: In terms of the two point function (5), Eqs. (15) give

$$(g_{\nu\sigma}\partial_\mu + g_{\mu\sigma}\partial_\nu)(cD^+ + \Box G_1) + g_{\mu\nu}\partial_\sigma(\Box F - cD^+)$$
$$+ \partial_\mu\partial_\nu\partial_\sigma(F + \Box G_2 + 2G_1) = 0, \tag{16}$$
$$g_{\mu\nu}(\Box F - 2cD^+) + \partial_\mu\partial_\nu(4F + 4G_1 + \Box G_2) = 0, \tag{16'}$$

respectively. It is not difficult to prove that the above equations are inconsistent unless $c = 0$. In this case the two point function of $R_{\alpha\beta\gamma\delta}$ vanishes.

Definition 2: By an *analog of the Landau gauge* in quantum gravitation we mean a local and covariant gauge in which Eq. (14) holds as an operator equation. and the cluster property holds in H' [property (ii) of Remark 1].

Proposition 2: For the gauges characterized in Definition 2 the two point function takes the form

$$W_{\mu\nu\rho\sigma}(x) = (g_{\mu\rho}g_{\nu\sigma} + g_{\nu\rho}g_{\mu\sigma} - g_{\mu\nu}g_{\rho\sigma})D^+$$
$$- (g_{\mu\rho}\partial_\nu\partial_\sigma + g_{\nu\rho}\partial_\mu\partial_\sigma + g_{\mu\sigma}\partial_\nu\partial_\rho + g_{\nu\sigma}\partial_\mu\partial_\rho)$$
$$\times(\Box^{-1}D^+ + ND^+) + M(g_{\mu\nu}\partial_\rho\partial_\sigma + g_{\rho\sigma}\partial_\mu\partial_\nu)D^+$$
$$+ 2M\partial_\mu\partial_\nu\partial_\rho\partial_\sigma(\Box^{-1}D^+ + LD^+) \tag{17}$$

where L, M, N are constants, $\Box^{-1}$ is defined as in Ref. 5, and the constant c of Theorem 1 has been chosen equal to one.

Proof: In terms of the two point function (5), Eq. (14) implies

$$(g_{\nu\sigma}\partial_\mu + g_{\mu\sigma}\partial_\nu)(D^+ + \Box G_1) + \tfrac{1}{2}g_{\mu\nu}\partial_\sigma\Box F + \partial_\mu\partial_\nu\partial_\sigma(\tfrac{1}{2}\Box G_2 - F) = 0. \tag{18}$$

The general solution of Eq. (18) is of the following form:

$$F = MD^+ + ax^4 + bx^2 + c_0,$$
$$G_1 = -\Box^{-1}D^+ - ND^+ + \tfrac{1}{2}ax^4 + b_1x^2 + c_1,$$
$$G_2 = 2M\Box^{-1}D^+ + LD^+ - \frac{a}{12}x^6 + b_2x^4 + c_2x^2 + c_3.$$

The constants c_0, c_1, c_2, c_3, can be chosen equal to zero since they do not contribute to the two point function. On the other hand, the terms proportional to a, b, b_1, b_2 must vanish by an argument similar to that of Remark 1.

The Fock space representation of the gravitational potential, characterized by the two point function (17), can be easily obtained in terms of the representation of the potential $h^0_{\mu\nu}$, corresponding to $L = M = N = 0$. In this case the indefinite scalar product between the states $\psi_f \equiv h^0(f)\psi_0$, $\psi_t \equiv h^0(t)\psi_0$, ($f_{\mu\nu} = f_{\nu\mu} \in \mathcal{S}$, $t_{\mu\nu} = t_{\nu\mu} \in \mathcal{S}$) is given by

$$\langle\psi_f, \psi_t\rangle = \pi\int\theta(k_0)\delta''(k^2)[\bar{F}^{\mu\nu}T_{\mu\nu} - \tfrac{1}{2}\bar{F}_\lambda{}^\lambda T_\sigma{}^\sigma]\,d^4k, \tag{19}$$

where

$$F_{\mu\nu}(k) \equiv k^2(f_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}f_\lambda{}^\lambda) - k_\mu k^\lambda f_{\lambda\nu} - k_\nu k^\lambda f_{\lambda\mu}, \tag{20}$$

and similarly for $T_{\mu\nu}$. According to the definition[15] of $\theta\delta''$

$$\int\theta(k_0)\delta''(k^2)f(k)\,d^4k \equiv \int\left[\left(\frac{1}{2k_0}\frac{\partial}{\partial k_0}\right)^2\frac{f}{2k_0}\right]_{k_0=|\mathbf{k}|}d^3k.$$

Equation (19) is well defined, since $\bar{F}_{\mu\nu}T^{\mu\nu} - \tfrac{1}{2}\bar{F}_\lambda{}^\lambda T_\sigma{}^\sigma$ is sufficiently well behaved at $\mathbf{k} = 0$.

The above product (19) can also be written in the form

$$\langle\psi_f\psi_t\rangle = (\psi_f, \eta\psi_t) = N\int\frac{d^3k}{k_0}\frac{1}{k_0^4}\bar{\psi}_f\psi_t,$$

where ψ_f denotes the multicomponent vector

$$\psi_f \equiv \sqrt{\pi}\begin{pmatrix} F_{\mu\nu} \\ k_0(\partial/\partial k_0)F_{\mu\nu}\end{pmatrix} \tag{21}$$

(and similarly for ψ_t), N is a suitable normalization constant, and the metric η is represented by the matrix

$$\eta = \frac{1}{8N}\begin{pmatrix}\eta_{11} & \eta_{12} \\ \eta_{21} & \eta_{22}\end{pmatrix}$$

with

$$\eta_{11} = \tfrac{3}{2}(g_{\mu\sigma}g_{\nu\rho} + g_{\mu\rho}g_{\nu\sigma} - g_{\mu\nu}g_{\rho\sigma}) + 4g_{0\mu}g_{0\nu}g_{0\rho}g_{0\sigma}$$
$$- (g_{0\mu}g_{0\nu}g_{\rho\sigma} + g_{0\rho}g_{0\sigma}g_{\mu\nu}),$$
$$\eta_{12} = \eta_{21} = -\tfrac{3}{2}(g_{\mu\rho}g_{\nu\sigma} + g_{\nu\rho}g_{\mu\sigma} - g_{\mu\nu}g_{\rho\sigma}) - 2g_{0\mu}g_{0\nu}g_{\rho\sigma}$$
$$+ (g_{0\mu}g_{0\rho}g_{\nu\sigma} + g_{0\nu}g_{0\rho}g_{\mu\sigma}$$
$$+ g_{0\mu}g_{0\sigma}g_{\nu\rho} + g_{0\nu}g_{0\sigma}g_{\mu\rho}),$$
$$\eta_{22} = g_{\mu\rho}g_{\nu\sigma} + g_{\nu\rho}g_{\mu\sigma} - g_{\mu\nu}g_{\rho\sigma}.$$

The closure of the linear manifold spanned by vectors of the form (21) [$f_{\mu\nu} \in \mathcal{S}(R^4)$] with respect to the scalar product

$$(\psi_f, \psi_t) = \pi\int_{C_+}\frac{d^3k}{k_0^5}\sum_{\mu,\nu}\left(\bar{F}_{\mu\nu}T_{\mu\nu} + k_0{}^2\frac{\partial}{\partial k_0}\bar{F}_{\mu\nu}\frac{\partial}{\partial k_0}T_{\mu\nu}\right)$$

is the one particle Hilbert space $H^{(1)}$. The Hilbert space H is obtained by the standard construction $H = \oplus_{n=1}^{\infty}H^{(n)}$. The Fock space representation of the field is

$$h^0(f) = h^0_+(f) + h^0_-(f),$$
$$(h^0_+(f)\psi)^{(n)}_{\mu_1\nu_1\cdots\mu_n\nu_n} = \sqrt{n+1}\int\theta(k_0)\delta''(k^2)$$
$$\times F^{\mu\nu}(k)\psi^{(n+1)}_{\mu\nu\mu_1\nu_1\cdots\mu_n\nu_n}(kk_1\cdots k_n)\,d^4k,$$
$$(h^0_-(f)\psi)^n_{\mu_1\nu_1\cdots\mu_n\nu_n} = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}[F_{\mu\nu}(-k_j) - \tfrac{1}{2}g_{\mu\nu}F_\lambda{}^\lambda(-k_j)]$$
$$\times\psi^{(n-1)}_{\mu_1\nu_1\cdots\widehat{\mu_j\nu_j}\cdots\mu_n\nu_n}(k_1\cdots\hat{k}_j\cdots k_n).$$

It is not difficult to prove that

$$\partial_\mu h^{\mu\nu} - \tfrac{1}{2}\partial^\nu h_\lambda{}^\lambda = 0.$$

A very simple argument shows that η is not degenerate. Indeed, if $\eta\psi$ is such that $(\eta\psi, \eta\psi) = 0$, it must be of the form

$$\eta\psi \equiv \eta \begin{pmatrix} N_{\mu\nu} \\ M_{\mu\nu} \end{pmatrix} = \begin{pmatrix} k^2 A_{\mu\nu} \\ k^2 B_{\mu\nu} \end{pmatrix}.$$

Solving for $N_{\mu\nu}$ and $M_{\mu\nu}$, one easily finds $N_{\mu\nu} = k^2 N'_{\mu\nu}$, $M_{\mu\nu} = k^2 M'_{\mu\nu}$, i.e., $(\psi, \psi) = 0$.

The operator η is obviously bounded in $\mathcal{H}^{(1)}$. By a suitable choice of the normalization constant N, the norm of η can be made less than 1. As a consequence one can extend the definition of η to the whole space by putting

$$(\eta\psi)^{(n)} = \eta^{\otimes n}\psi^{(n)}, \quad n = 1, 2, \cdots,$$

thereby obtaining a bounded operator.

Finally, the representation of h, corresponding to nonvanishing L, M, N, is obtained by a special gauge transformation[5,6] on $h^0(f)$.

5. GRAVITONS WITH ZERO HELICITY

In this Section we will discuss possible deviations from the characterization given in Sec. 3, namely theories in which Eq. (6) does not hold.

It is important to remark that if the space–time translations commute with the metric operator the Fourier transform of F, G_1, and G_2 are measures and therefore by the Gårding decomposition of Lorentz invariant measures[16] the Fock space obtained by a generalized reconstruction theorem splits into a direct integral over the mass. Thus, quite generally, one may restrict to the case in which $\tilde{F}$, $\tilde{G}_1$, and $\tilde{G}_2$ have support contained in $\{k^2 = 0\}$. In that case $\tilde{F}$, $\tilde{G}_1$, and $\tilde{G}_2$ must be proportional to $\delta(k^2)$ and therefore Eq. (6) trivially holds.

Equation (6) holds also in any local and covariant quantization in which $\Box R_{\lambda\mu\nu\rho}\mathcal{H}' = 0$ and $D_0 \equiv \{$set of vectors obtained by applying polynomials in the smeared fields $R_{\lambda\mu\nu\rho}(f^{\lambda\mu\nu\rho})$ to the vacuum$\}$ is dense in $\mathcal{H}' \ominus \mathcal{H}''$. This latter property is guaranteed by the requirement that in the physical space $\mathcal{H}_{\text{phys}} = \mathcal{H}'/\mathcal{H}''$ the theory can be formulated in terms of the observable fields $R_{\mu\nu\rho\sigma}$.

In conclusion the plausibility of Eq. (6) seems out of discussion and in any case Eq. (6) holds in all the gauges known to us.[17]

It is possible however to exhibit a local and covariant quantization in which Eq. (6) fails. For example, a two point function (5) with

$$F(x) = 0, \quad G_1(x) = 0,$$

$$G_2(x) = \lambda\Box^{-2}D^+(x), \quad \lambda > 0,$$

defines a theory of this kind.

The vectors $h(f)\psi_0$ belongs to $\mathcal{H}'_{(1)}$ if either $\partial^\mu f_{\mu\nu} = 0$ or there exist a function a_μ such that $\partial^\mu(f_{\mu\nu} - g_{\mu\nu}\partial^\rho a_\rho + \partial_\mu a_\nu + \partial_\nu a_\mu) = 0$. It is easy to check that all the properties listed in Definition 1 are fulfilled. Moreover, the vectors of the form $h_{\mu\nu}(-g_{\mu\nu}\partial^\rho a_\rho + \partial_\mu a_\nu + \partial_\nu a_\mu)\psi_0$ have nonvanishing norm and zero helicity.

Similar unconventional gauges can be defined also for quantum electrodynamics, and yield states of zero helicity in the physical space $\mathcal{H}_{\text{phys}}$. Thus the existence of these pathological gauges is not peculiar to quantum gravity.

[1] L. Bracci and F. Strocchi, J. Math. Phys. **13**, 1151 (1972).
[2] M.A. Naimark, *Normed Rings* (Noordhoff, Groningen, 1959).
[3] R. Streater and A.S. Wightman, *PCT, Spin and Statistics and All That* (Benjamin, New York, 1964).
[4] L. Bracci, G. Morchio, and F. Strocchi, Commun. Math. Phys. (to be published).
[5] For a general definition of gauge see F. Strocchi and A.S. Wightman, J. Math. Phys. (to be published). For the quantization of the gravitational potential see Ref. 1 and Ref. 6.
[6] L. Davis, Princeton, senior thesis, 1974.
[7] S.N. Gupta, Proc. Phys. Soc. (Lond. A **65**, 161 (1952).
[8] J. Bertrand, "Mass Zero Spin Two Particles in Quantum Field Theory," talk at the Meeting on Group Theory, Marseille, June 1974.
[9] F. Strocchi, Phys. Rev. **166**, 1302 (1968).
[10] R. Jost, Helv. Phys. Acta **30**, 409 (1957).
[11] L. Gårding and L. Lions, Nuovo Cimento Suppl. **14** (1959), Lemma 8.4.
[12] L. Schwartz, *Théorie des distributions* (Hermann, Paris, 1966), especially, p. 123.
[13] L.D. Landau and I.M. Khalatnikov, Sov. Phys. JETP **2**, 69 (1956).
[14] V.I. Ogievetsky and I.V. Polubarinov, Ann. Phys. (N.Y.) **35**, 167 (1965).
[15] V.S. Vladimirov, *Les functions de plusieurs variables complexes* (Dunad, Paris, 1967), p. 284; I.M. Gel'fand and G.E. Shilov, *Generalized Functions* (Academic, New York, 1964).
[16] L. Gårding and J. Lions, Ref. 11, Sec. 8.
[17] In particular, Eq. (6) holds in the gauges analyzed in Ref. 8.