The triad-interaction representation of homogeneous turbulence

Jon Lee

Aerospace Research Laboratories, Wright-Patterson AFB, Ohio 45433
(Received 24 May 1974; revised manuscript received 22 November 1974)

The triad-interaction representation has been presented for the 2D and 3D homogeneous flows. This has several advantages over the usual Fourier-amplitude representation: (i) The incompressibility is built into the equation as in the vorticity equation. (ii) For a given wave vector, the number of dynamic equations is one less than that of the Fourier-amplitude equations. (iii) In the inviscid limit, energy and enstrophy are conserved in 2D, whereas the 3D flow conserves energy and helicity. (iv) The entire family of triad interactions is categorized into two classes in 2D and four classes in 3D, according to the geometry of triad wave vectors. Lastly, (v) the necessary conditions for isotropy in 3D emerge as the reflexional, rotational, and spherical symmetries in the wave vector space, whereas polar symmetry is only the requirement in 2D. The triad-interaction representation has proved very useful in the investigation of isolating constants of motion and the statistical theory of nonisotropic turbulence.

1. THE EQUATIONS OF EDDY MOTION

By Fourier analyzing the homogeneous velocity field, the Navier–Stokes equations carry over to the Fourier-amplitude equations which are, unfortunately, an infinite set of ordinary differential equations. Since the incompressibility prohibits longitudinal motion in the Fourier space, the Fourier-amplitude equations may be projected into a plane perpendicular to the wave vector. We can thus decompose the quadratic nonlinearity into the three-Fourier-mode interactions, called hereafter the triad interactions (TI), which extend over infinitely coupled triad wave vectors. As early as in 1958, Kraichnan[1] discussed the symbolic equations of eddy motion consisting entirely of such TI's. Although some recent turbulence[2–4] work make use of the representation in terms of TI's, the TI structure of the Navier–Stokes equations has not yet been exhibited in detail. The purpose of this paper is therefore to give an explicit TI representation for the 2D and 3D homogeneous flows, and hence categorize the nonlinear interactions into various geometric classes. Although we shall not discuss any application here, the TI representation lends itself very naturally to the investigation of isolating constants of motion for the homogeneous turbulence[5] and the statistical theory of nonisotropic turbulence.[6]

Using the expansion

$$U_i(\mathbf{x}, t) = \sum_{\mathbf{k}} U_i(\mathbf{k}, t) \exp(i\mathbf{k} \cdot \mathbf{x}), \qquad (1.1)$$

where

$$\mathbf{k} = (2\pi/L) \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} \quad (n_1, n_2, n_3 = \pm 1, \pm 2, \cdots),$$

the Fourier-amplitude equations for the incompressible Navier–Stokes dynamics become

$$(\partial/\partial t + \nu k^2)U_i(\mathbf{k}, t) = -ik_m P_{ij}(\mathbf{k}) \sum_{\mathbf{k}=\mathbf{p}+\mathbf{q}} U_j(\mathbf{p}, t)U_m(\mathbf{q}, t), \qquad (1.2)$$

where ν is the kinematic viscosity and $P_{ij}(\mathbf{k}) = \delta_{ij} - k_i k_j / k^2$. The reality requirement states $U_i(\mathbf{k}) = U_i^*(-\mathbf{k})$. Since the incompressibility $k_i U_i(\mathbf{k}) = 0$ restricts the motion to a plane perpendicular to $\mathbf{k}$, we may project the velocity field into a subspace spanned[2] by polarization vectors $\epsilon^\mu(\mathbf{k})$ ($\mu = 1, 2$)

$$U_i(\mathbf{k}, t) = \sum_{\mu=1,2} \epsilon_i^\mu(\mathbf{k}) u^\mu(\mathbf{k}, t), \qquad (1.3)$$

where

$$\mathbf{k} \cdot \epsilon^\mu(\mathbf{k}) = 0,$$

$$\epsilon^\mu(\mathbf{k}) \cdot \epsilon^\lambda(\mathbf{k}) = \delta_{\mu\lambda},$$

$$\sum_{\mu=1,2} \epsilon_i^\mu(\mathbf{k}) \epsilon_j^\mu(\mathbf{k}) = P_{ij}(\mathbf{k}). \qquad (1.4)$$

The first of (1.4) is the orthogonality of $\mathbf{k}$ with $\epsilon^\mu(\mathbf{k})$, the second is the orthonormality, and the last is the identity of the orthonormal vectors ($\epsilon^1(\mathbf{k})$, $\epsilon^2(\mathbf{k})$, $\mathbf{k}/k$). The 2D flow has only one polarization vector normal to $\mathbf{k}$. The choice of $\epsilon^\mu(\mathbf{k})$ is, however, not unique in 3D, for any orthonormal vectors will do as long as they are both perpendicular to $\mathbf{k}$.

The introduction of (1.3) into (1.2) gives the TI representation that we are seeking:

$$(\partial/\partial t + \nu k^2)u^\mu(\mathbf{k}, t) = -i \sum_{\rho,\lambda} \sum_{\mathbf{k}=\mathbf{p}+\mathbf{q}} \phi_{\mathbf{k}|\mathbf{p},\mathbf{q}}^{\mu|\rho,\lambda} u^\rho(\mathbf{p}, t)u^\lambda(\mathbf{q}, t), \qquad (1.5)$$

where the coupling coefficient is

$$\phi_{\mathbf{k}|\mathbf{p},\mathbf{q}}^{\mu|\rho,\lambda} = (\mathbf{k} \cdot \epsilon^\lambda(\mathbf{q}))(\epsilon^\mu(\mathbf{k}) \cdot \epsilon^\rho(\mathbf{p})). \qquad (1.6)$$

Because of the first of (1.4), the incompressibility is embodied in (1.5) as in the vorticity equation. In 2D, (1.5) is a single scalar equation for $u^1(\mathbf{k})$, identical to the vorticity equation. In the 3D case, there are two equations for $u^\mu(\mathbf{k})$ in contrast to the three equations for $U_i(\mathbf{k})$ in (1.2). This reduction in number of equations is due to the vorticity vector $\omega(\mathbf{k})$ having two independent components, since $\mathbf{k} \cdot \omega(\mathbf{k}) = 0$.

To facilitate the enumeration of TI's, we restrict the wave vector domain to the positive range and split the convolution sum into two parts

$$(\partial/\partial t + \nu k^2)u^\mu(\mathbf{k}, t)$$
$$= -i \sum_{\rho,\lambda} \left(\sum_{\mathbf{p}>0}^{\mathbf{p}<\mathbf{k}} \phi_{\mathbf{k}|\mathbf{p},(\mathbf{k}-\mathbf{p})}^{\mu|\rho,\lambda} u^\rho{}^*(\mathbf{p}, t)u^\lambda{}^*(\mathbf{k}-\mathbf{p}, t) \right.$$
$$\left. + \sum_{\mathbf{p}>0} \bar{\phi}_{\mathbf{k}|\mathbf{p},(\mathbf{k}+\mathbf{p})}^{\mu|\rho,\lambda} u^\rho{}^*(\mathbf{p}, t)u^\lambda{}^*(\mathbf{k}+\mathbf{p}, t) \right), \qquad (1.7)$$

where

$$\bar{\phi}_{\mathbf{k}|\mathbf{p},\mathbf{q}}^{\mu|\rho,\lambda} = \phi_{\mathbf{k}|\mathbf{p},\mathbf{q}}^{\mu|\rho,\lambda} + \phi_{\mathbf{k}|\mathbf{q},\mathbf{p}}^{\mu|\lambda,\rho},$$

is the symmetrized coupling coefficient. We have reversed the directions of p and (k − p) in the first sum of (1.7) and the direction of (k + p) in the second sum. Hence, the triad wave vector now has the head-to-tail configuration.

Several constants of motion are known in the inviscid limit. First, the energy conservation is

$$\frac{1}{2}\frac{d}{dt}\sum_{\mathbf{k}}\sum_{\mu}|u^{\mu}(\mathbf{k})|^{2}=0. \tag{1.8}$$

Second, the enstrophy is also conserved in 2D:

$$\frac{1}{2}\frac{d}{dt}\sum_{\mathbf{k}}k^{2}|u^{1}(\mathbf{k})|^{2}=0, \tag{1.9}$$

which follows from $\omega_{3}(\mathbf{k})=i(k_{1}\epsilon_{2}^{1}-k_{2}\epsilon_{1}^{1})u^{1}(\mathbf{k})$. Lastly, the 3D flow conserves the helicity which is defined by the scalar product of $U_{i}(\mathbf{x},t)$ and the vorticity $\omega_{i}(\mathbf{x},t)$:

$$\frac{d}{dt}\sum_{\mathbf{k}}\sum_{\mu\neq\lambda}i\,\epsilon_{ijm}k_{j}\epsilon_{m}^{\mu}(\mathbf{k})\,\epsilon_{i}^{\lambda}(\mathbf{k})\,u^{\mu}(\mathbf{k})\,u^{\lambda *}(\mathbf{k})=0, \tag{1.10}$$

where ϵ_{ijm} is the unit alternating tensor ($=0$ if i,j,m are not distinct; $=\pm 1$ when i,j,m are cyclic and acyclic, respectively). According to Moffatt,[7] helicity measures the degree of knottedness of the vortex lines and is a particular consequence of Kelvin's circulation theorem. We shall see that energy, enstrophy, and helicity are conserved in the TI representation.

2. TWO-DIMENSIONAL FLOW

We may drop the polarization vector indices ($\mu,\rho,\lambda=1$) in 2D and also the viscous term for the notational compactness

$$\dot{u}(\mathbf{k})=-i\left(\sum_{p>0}^{\mathbf{k}/\mathbf{k}}\phi_{\mathbf{k}|\mathbf{p},(\mathbf{k}-\mathbf{p})}u^{*}(\mathbf{p})u^{*}(\mathbf{k}-\mathbf{p})\right.$$

$$\left.+\sum_{p>0}\bar{\phi}_{\mathbf{k}|\mathbf{p},(\mathbf{k}+\mathbf{p})}u^{*}(\mathbf{p})\,u^{*}(\mathbf{k}+\mathbf{p})\right), \tag{2.1}$$

where the dot denotes $\partial/\partial t$. By choosing $L=2\pi$, the wave vector space consists of plane lattice points $\mathbf{k}=\binom{n_{1}}{n_{2}}$ ($n_{1},n_{2}=$ positive integers). For a given $\mathbf{k}$, the first sum of (2.1) gives two types of terms. The first type is pairs that can be combined under the symmetrized coupling coefficient, and the second type is a single term corresponding to $p=k/2$. When the triad wave vectors collapse to a single vector, we find from (1.6) that $\phi_{\mathbf{k}|\mathbf{p},\mathbf{q}}=0$. So, discounting such degenerate triad wave vectors, (2.1) for the 3×3 lattice may be put in the vector form

$$\begin{pmatrix}\dot{u}\binom{1}{1}\\\dot{u}\binom{1}{2}\\\dot{u}\binom{2}{1}\\\dot{u}\binom{2}{3}\\\dot{u}\binom{3}{3}\\\dot{u}\binom{3}{2}\end{pmatrix}=-i\left\{\begin{pmatrix}\bar{\phi}_{1\,\substack{12\\1\,23}}\\\bar{\phi}_{1\,\substack{21\\2\,31}}\\0\\\bar{\phi}_{2\,\substack{11\\3\,12}}\\0\\0\end{pmatrix}+\begin{pmatrix}\bar{\phi}_{1\,\substack{23\\1\,12}}\\0\\\bar{\phi}_{2\,\substack{31\\1\,21}}\\0\\0\\\bar{\phi}_{3\,\substack{12\\2\,11}}\end{pmatrix}+\begin{pmatrix}0\\\bar{\phi}_{1\,\substack{23\\2\,13}}\\\bar{\phi}_{2\,\substack{31\\1\,32}}\\0\\\bar{\phi}_{3\,\substack{12\\3\,21}}\\0\end{pmatrix}\right\},$$

$$\tag{2.2}$$

where $\bar{\phi}_{\mathbf{k}|\mathbf{p}\mathbf{q}}$ stands for $\bar{\phi}_{\mathbf{k}|\mathbf{p},\mathbf{q}}u^{*}(\mathbf{p})u^{*}(\mathbf{q})$. Note that each vector has no more than three nonzero elements, and this, of course, is independent of the size of wave vector lattice under consideration. (Although $\bar{\phi}_{3\,\substack{12\\3\,21}}=0$, we have retained it in the last vector for symmetric representation.)

A. The fundamental triad interaction

The structure of TI can be exhibited by stripping all the zeros in any vector of (2.2):

$$\dot{u}(\mathbf{K})=-i\,\bar{\phi}_{\mathbf{K}|\mathbf{P},\mathbf{Q}}\,u^{*}(\mathbf{P})u^{*}(\mathbf{Q}),$$

$$\dot{u}(\mathbf{P})=-i\,\bar{\phi}_{\mathbf{P}|\mathbf{Q},\mathbf{K}}\,u^{*}(\mathbf{Q})u^{*}(\mathbf{K}), \tag{2.3}$$

$$\dot{u}(\mathbf{Q})=-i\,\bar{\phi}_{\mathbf{Q}|\mathbf{K},\mathbf{P}}\,u^{*}(\mathbf{K})u^{*}(\mathbf{P}),$$

where $\mathbf{K}+\mathbf{P}+\mathbf{Q}=0$. The coupling coefficients obey

$$\bar{\phi}_{\mathbf{K}|\mathbf{P},\mathbf{Q}}+\bar{\phi}_{\mathbf{P}|\mathbf{Q},\mathbf{K}}+\bar{\phi}_{\mathbf{Q}|\mathbf{K},\mathbf{P}}=0, \tag{2.4}$$

for terms like $(\epsilon(\mathbf{K})\cdot\mathbf{P})$ and $(\epsilon(\mathbf{K})\cdot\mathbf{Q})$ add up to zero on account of $\mathbf{K}+\mathbf{P}+\mathbf{Q}=0$. The detailed energy conservation of (2.3) directly follows from (2.4), and hence this implies the overall conservation (1.8) because energy being conserved by each and every TI. The coupling coefficients obey another constraint,

$$K^{2}\bar{\phi}_{\mathbf{K}|\mathbf{P},\mathbf{Q}}+P^{2}\bar{\phi}_{\mathbf{P}|\mathbf{Q},\mathbf{K}}+Q^{2}\,\bar{\phi}_{\mathbf{Q}|\mathbf{K},\mathbf{P}}=0, \tag{2.5}$$

from which follows the enstrophy conservation (1.9). To verify (2.5), we write out the left-hand side (lhs) by using $(\epsilon(\mathbf{K})\cdot\mathbf{P})=-(\epsilon(\mathbf{K})\cdot\mathbf{Q})$ and similar relations involving $\epsilon(\mathbf{P})$ and $\epsilon(\mathbf{Q})$:

$$\text{lhs}=(KPQ)^{-1}[K(\epsilon(\mathbf{K})\cdot\mathbf{P})\,PQ(\epsilon(\mathbf{P})\cdot\epsilon(\mathbf{Q}))\,(P^{2}-Q^{2})$$

$$+P(\epsilon(\mathbf{P})\cdot\mathbf{Q})\,QK(\epsilon(\mathbf{Q})\cdot\epsilon(\mathbf{K}))\,(Q^{2}-K^{2})$$

$$+Q(\epsilon(\mathbf{Q})\cdot\mathbf{K})\,KP\,(\epsilon(\mathbf{K})\cdot\epsilon(\mathbf{P}))\,(K^{2}-P^{2})]. \tag{2.6}$$

Since $K(\epsilon(\mathbf{K})\cdot\mathbf{P})=P(\epsilon(\mathbf{P})\cdot\mathbf{Q})=Q(\epsilon(\mathbf{Q})\cdot\mathbf{K})$, we can show by using the cosine laws that the curly bracket vanishes identically.

B. The coupling coefficients

Let us express $\mathbf{K}$ and the polarization vector in polar coordinates

$$\mathbf{K}=K\begin{pmatrix}\cos\theta_{k}\\\sin\theta_{k}\end{pmatrix}\quad\text{and}\quad\epsilon(\mathbf{K})=\begin{pmatrix}\sin\theta_{k}\\-\cos\theta_{k}\end{pmatrix}, \tag{2.7}$$

where $\cos\theta_{k}=K_{1}/K$ and $\sin\theta_{k}=K_{2}/K$. Similar expressions exist for $\mathbf{P}$ and $\mathbf{Q}$, and their polarization vectors. Note that $\epsilon(\mathbf{K})$ is reflexional about $\mathbf{K}$, i.e., $\epsilon(\mathbf{K})=-\epsilon(-\mathbf{K})$. In the first sum of (2.1), we have reversed the directions of both p and (k − p), but only the direction of (k + p) in the second sum. By tracing the derivation, we see that the first two elements in each vector of (2.2) were generated by the second sum, whereas the last element was originated from the first sum. Let us now denote by K, P, and Q the wave vectors in the order that they appear in column vectors. Then, considering the wave vector reversal and the reflexionality of ϵ, the coupling coefficients have the representation

$$\bar{\phi}_{\mathbf{K}|\mathbf{P},\mathbf{Q}}=-(\mathbf{K}\cdot\epsilon(\mathbf{P}))(\epsilon(\mathbf{K})\cdot\epsilon(\mathbf{Q}))-(\mathbf{K}\cdot\epsilon(\mathbf{Q}))(\epsilon(\mathbf{K})\cdot\epsilon(\mathbf{P})),$$

$$\bar{\phi}_{\mathbf{P}|\mathbf{Q},\mathbf{K}}=-(\mathbf{P}\cdot\epsilon(\mathbf{Q}))(\epsilon(\mathbf{P})\cdot\epsilon(\mathbf{K}))-(\mathbf{P}\cdot\epsilon(\mathbf{K}))(\epsilon(\mathbf{P})\cdot\epsilon(\mathbf{Q})),$$
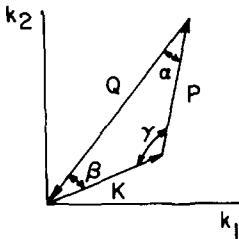
FIG. 1. Plane triad wave vector.

$$\bar{\phi}_{Q|K,P} = (Q \cdot \epsilon(K))(\epsilon(Q) \cdot \epsilon(P)) + (Q \cdot \epsilon(P))(\epsilon(Q) \cdot \epsilon(K)).$$
(2.8)

First, express (2.8) in polar coordinates, and then simplify it further by denoting the internal angles opposite to K, P, Q by $\alpha = \theta_p - \theta_q$, $\beta = \theta_q - \theta_k$, $\gamma = \pi + \theta_k - \theta_p$, respectively (Fig. 1):

$$\bar{\phi}_{K|P,Q} = K \sin(\beta - \gamma),$$
$$\bar{\phi}_{P|Q,K} = P \sin(\gamma - \alpha),$$
$$\bar{\phi}_{Q|K,P} = Q \sin(\alpha - \beta).$$
(2.9)

This shows that the $\bar{\phi}$'s are determined by the shape of triangle formed by K, P, and Q, and not by its orientation in the 2D wave vector space. Note that (2.9) agrees with Kraichnan's for the interacting shear-wave system.[8]

C. Classification of the triad interactions

According to the geometry of triad wave vectors, we can categorize the TI's in 2D into 2 classes. The first

two vectors in (2.2) have triad wave vectors which form triangles symmetric about the diagonal. So, we shall say they are of the symmetric pair class. For such a pair, $\bar{\phi}$'s are the same except for sign

$$\left(\bar{\phi}_{\substack{1|12 \\ 1|23}}, \bar{\phi}_{\substack{1|21 \\ 2|31}}, \bar{\phi}_{\substack{2|11 \\ 3|12}}\right) = -\left(\bar{\phi}_{\substack{1|23 \\ 1|12}}, \bar{\phi}_{\substack{2|31 \\ 1|21}}, \bar{\phi}_{\substack{3|12 \\ 2|11}}\right).$$
(2.10)

Next, the last vector in (2.2) has the triad wave vector forming an isosceles triangle with the base $Q = \binom{3}{3}$, hence it is said belonging to the isosceles class. For such a TI,

$$\bar{\phi}_{Q|K,P} = 0,$$
(2.11)

and, of course,

$$\bar{\phi}_{K|P,Q} = -\bar{\phi}_{P|Q,K}.$$

Let us denote by $(K, P, Q)_m$ the column vector with the nonzero elements $\bar{\phi}_{K|P,Q} u^*(P) u^*(Q)$, $\bar{\phi}_{P|Q,K} u^*(Q) u^*(K)$, $\bar{\phi}_{Q|K,P} u^*(K) u^*(P)$, where

$m = 1$ for the isosceles class,

$m = 2$ for the symmetric pair class.

Since it is superfluous to enumerate both of the symmetric pair, we shall adopt the convention that a TI of the symmetric class be represented by either one of the pair. Then, the subscript m actually refers to multiplicity. Using this compact notation, the TI's in the 5×5 lattice may be summarized as

$$\dot{u}_2 = -i\left\{\left[\begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix}_2 + \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}_1\right]\right.$$

$$\underbrace{}_{\text{third-order}}$$

$$+ \begin{pmatrix} 1 & 1 & 2 \\ 1 & 3 & 4 \end{pmatrix}_2 + \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \end{pmatrix}_2 + \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 4 \end{pmatrix}_2 + \begin{pmatrix} 2 & 1 & 3 \\ 1 & 3 & 4 \end{pmatrix}_2 + \begin{pmatrix} 2 & 2 & 4 \\ 1 & 3 & 4 \end{pmatrix}_2 + \begin{pmatrix} 1 & 3 & 4 \\ 3 & 1 & 4 \end{pmatrix}_1$$

$$\underbrace{}_{\text{fourth-order}}$$

$$+\left[\begin{pmatrix} 1 & 1 & 2 \\ 1 & 4 & 5 \end{pmatrix}_2 + \begin{pmatrix} 1 & 1 & 2 \\ 2 & 3 & 5 \end{pmatrix}_2 + \begin{pmatrix} 1 & 2 & 3 \\ 1 & 4 & 5 \end{pmatrix}_2 + \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 5 \end{pmatrix}_2 + \begin{pmatrix} 2 & 1 & 3 \\ 2 & 3 & 5 \end{pmatrix}_2 + \begin{pmatrix} 2 & 1 & 3 \\ 1 & 4 & 5 \end{pmatrix}_2 + \begin{pmatrix} 1 & 3 & 4 \\ 1 & 4 & 5 \end{pmatrix}_2 + \begin{pmatrix} 1 & 3 & 4 \\ 2 & 3 & 5 \end{pmatrix}_2 + \begin{pmatrix} 2 & 2 & 4 \\ 2 & 3 & 5 \end{pmatrix}_2 + \begin{pmatrix} 2 & 2 & 4 \\ 1 & 4 & 5 \end{pmatrix}_2\right.$$

$$+ \begin{pmatrix} 3 & 1 & 4 \\ 2 & 3 & 5 \end{pmatrix}_2 + \begin{pmatrix} 3 & 1 & 4 \\ 1 & 4 & 5 \end{pmatrix}_2 + \begin{pmatrix} 2 & 3 & 5 \\ 1 & 4 & 5 \end{pmatrix}_2 + \begin{pmatrix} 3 & 2 & 5 \\ 1 & 4 & 5 \end{pmatrix}_2 + \begin{pmatrix} 2 & 3 & 5 \\ 3 & 2 & 5 \end{pmatrix}_1 + \begin{pmatrix} 1 & 4 & 5 \\ 4 & 1 & 5 \end{pmatrix}_1\left.\right]\right\},$$
(2.12)

$$\underbrace{}_{\text{fifth-order}}$$

where u_2 is the column vector $(u\binom{1}{1}, u\binom{1}{2}, \ldots)$. Note that the first square bracket recapitulates (2.2). The second bracket represents the 11 TI's which involve the wave vectors in the fourth-order lattice shell, and the third bracket the 30 TI's of the fifth-order lattice shell. Due to the proliferation of TI's, algebraically handling TI terms becomes very tedious as the lattice order increases. It is however, more efficient to generate the TI's for a given lattice system by a diagrammatic procedure (the detail of which will not be presented here).

3. THREE-DIMENSIONAL FLOW

We can enumerate (1.7) in 3D lattice in much the same way as in 2D. Even for the smallest $3 \times 3 \times 3$ lattice, the equation comparable to (2.2) involves vectors of the dimension 14, so writing them down here in the usual notation is impractical. Let us therefore compress the vector by deleting all the zeros in it. Ignoring degenerate triad wave vectors, the equations of eddy motion for the $3 \times 3 \times 3$ lattice may be put in the compressed vector form

$$\dot{u}_3^\mu = -i \sum_{p,\lambda} \left\{ \begin{pmatrix} \overline{\phi}_1 |^{12}_{\substack{1|23\\1|12}} \\ \overline{\phi}_1 |^{21}_{\substack{2|31\\1|21}} \\ \overline{\phi}_2 |^{11}_{\substack{3|12\\2|11}} \end{pmatrix} + \begin{pmatrix} \overline{\phi}_1 |^{23}_{\substack{1|12\\1|12}} \\ \overline{\phi}_2 |^{31}_{\substack{1|21\\1|21}} \\ \overline{\phi}_3 |^{12}_{\substack{2|11\\2|11}} \end{pmatrix} + \begin{pmatrix} \overline{\phi}_1 |^{12}_{\substack{1|23\\1|23}} \\ \overline{\phi}_1 |^{21}_{\substack{2|31\\2|31}} \\ \overline{\phi}_2 |^{11}_{\substack{3|12\\3|12}} \end{pmatrix} + \begin{pmatrix} \overline{\phi}_1 |^{23}_{\substack{1|12\\1|23}} \\ \overline{\phi}_2 |^{31}_{\substack{1|21\\2|31}} \\ \overline{\phi}_3 |^{12}_{\substack{2|11\\3|12}} \end{pmatrix} + \begin{pmatrix} \overline{\phi}_1 |^{12}_{\substack{2|13\\1|23}} \\ \overline{\phi}_1 |^{21}_{\substack{1|32\\2|31}} \\ \overline{\phi}_2 |^{11}_{\substack{3|21\\3|12}} \end{pmatrix} + \begin{pmatrix} \overline{\phi}_2 |^{13}_{\substack{1|12\\1|23}} \\ \overline{\phi}_1 |^{32}_{\substack{1|21\\2|31}} \\ \overline{\phi}_3 |^{21}_{\substack{2|11\\3|12}} \end{pmatrix} \right.$$

$$\left. + \begin{pmatrix} \overline{\phi}_1 |^{23}_{\substack{2|13\\1|23}} \\ \overline{\phi}_2 |^{31}_{\substack{1|32\\2|31}} \\ \overline{\phi}_3 |^{12}_{\substack{3|21\\3|12}} \end{pmatrix} + \begin{pmatrix} \overline{\phi}_2 |^{13}_{\substack{1|23\\1|23}} \\ \overline{\phi}_1 |^{32}_{\substack{2|31\\2|31}} \\ \overline{\phi}_3 |^{21}_{\substack{3|12\\3|12}} \end{pmatrix} + \begin{pmatrix} \overline{\phi}_1 |^{23}_{\substack{2|13\\1|12}} \\ \overline{\phi}_2 |^{31}_{\substack{1|32\\1|21}} \\ \overline{\phi}_3 |^{12}_{\substack{3|21\\2|11}} \end{pmatrix} + \begin{pmatrix} \overline{\phi}_1 |^{12}_{\substack{1|12\\1|23}} \\ \overline{\phi}_1 |^{21}_{\substack{1|21\\2|31}} \\ \overline{\phi}_2 |^{11}_{\substack{2|11\\3|12}} \end{pmatrix} + \begin{pmatrix} \overline{\phi}_1 |^{23}_{\substack{1|23\\1|12}} \\ \overline{\phi}_2 |^{31}_{\substack{1|21\\1|21}} \\ \overline{\phi}_3 |^{12}_{\substack{2|11\\2|11}} \end{pmatrix} + \begin{pmatrix} \overline{\phi}_2 |^{13}_{\substack{2|13\\1|23}} \\ \overline{\phi}_1 |^{32}_{\substack{1|32\\2|31}} \\ \overline{\phi}_3 |^{21}_{\substack{3|12\\3|12}} \end{pmatrix} \right\} . \tag{3.1}$$

Here $\overline{\phi}_{k|pq}$ stands for $\phi^{\mu|\rho,\lambda}_{k|p,q} u^\rho*(p) u^{\lambda}*(q)$. Since u_3^μ is the column vector

$$\left(u^\mu \begin{pmatrix} 1\\1\\1 \end{pmatrix}, u^\mu \begin{pmatrix} 1\\2\\1 \end{pmatrix}, u^\mu \begin{pmatrix} 2\\1\\1 \end{pmatrix}, u^\mu \begin{pmatrix} 2\\2\\1 \end{pmatrix}, u^\mu \begin{pmatrix} 1\\1\\2 \end{pmatrix}, u^\mu \begin{pmatrix} 1\\2\\2 \end{pmatrix}, u^\mu \begin{pmatrix} 2\\1\\2 \end{pmatrix}, u^\mu \begin{pmatrix} 2\\3\\2 \end{pmatrix}, u^\mu \begin{pmatrix} 3\\2\\2 \end{pmatrix}, u^\mu \begin{pmatrix} 3\\3\\2 \end{pmatrix}, u^\mu \begin{pmatrix} 2\\2\\3 \end{pmatrix}, u^\mu \begin{pmatrix} 2\\3\\3 \end{pmatrix}, u^\mu \begin{pmatrix} 3\\2\\3 \end{pmatrix}, u^\mu \begin{pmatrix} 3\\3\\3 \end{pmatrix} \right)$$

the first vector in reality represents

$$\left(\overline{\phi}_1 |^{12}_{\substack{1|23\\1|12}}, \ \overline{\phi}_1 |^{21}_{\substack{2|31\\1|21}}, \ 0, \ 0, \ 0, \ 0, \ 0, \ \overline{\phi}_2 |^{11}_{\substack{3|12\\2|11}}, \ 0, \ 0, \ 0, \ 0, \ 0, \ 0 \right),$$

and others are interpreted likewise.

A. The coupling coefficients

We express $(K, \epsilon^1(K), \epsilon^2(K))$ by using the spherical coordinates similar to Euler's angles. First of all, we have

$$K = K \begin{pmatrix} \cos\theta_k \cos\eta_k \\ \sin\theta_k \cos\eta_k \\ \sin\eta_k \end{pmatrix}, \tag{3.2}$$

where $\cos\theta_k = K_1/K'$, $\sin\theta_k = K_2/K'$, $\cos\eta_k = K'/K$, and $\sin\eta_k = K_3/K$ $(K' = \sqrt{K_1^2 + K_2^2})$. Now choose a unit column vector $e = (\cos\theta_k, \sin\theta_k, 0)$. We then fix $\epsilon^1(K)$ by $e \times K / |e \times K|$ (Fig. 2),

$$\epsilon^1(K) = \begin{pmatrix} \sin\theta_k \\ -\cos\theta_k \\ 0 \end{pmatrix}_e, \tag{3.3a}$$
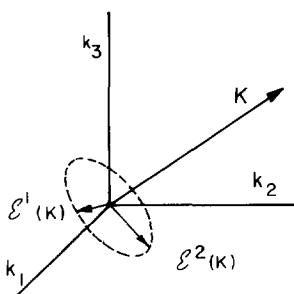
FIG. 2. Polarization vectors.

and $\epsilon^2(K)$ by $K \times \epsilon^1(K) / |K \times \epsilon^1(K)|$,

$$\epsilon^2(K) = \begin{pmatrix} \cos\theta_k \sin\eta_k \\ \sin\theta_k \sin\eta_k \\ -\cos\eta_k \end{pmatrix}_e. \tag{3.3b}$$

Note that $\epsilon^1(K)$ is reflexional, but $\epsilon^2(K)$ is not:

$$\epsilon^1(K) = -\epsilon^1(-K), \quad \epsilon^2(K) = \epsilon^2(-K). \tag{3.4}$$

We indicate by the subscript e that the representation (3.3) does in general depend on the choice of e. It is, however, evident from the cross product that a unit vector in the plane containing the K_3 axis and K will uniquely determine the polarization vectors up to a sign. In other words, an e in the azimuthal angle range $(\pi/2 - \eta_k, \ \pi/2]$ will give the same $\epsilon^\mu(K)$ as (3.3), whereas in the range $[0, \ \pi/2 - \eta_k)$ the polarization vectors are just the reflexion of (3.3) about K.

Considering the wave vector reversal and (3.4), the coupling coefficients have the following representation:

$$\phi^{\mu|\rho,\lambda}_{K|P,Q} = \mp [(K \cdot \epsilon^\rho(P))(\epsilon^\mu(K) \cdot \epsilon^\lambda(Q)) + (K \cdot \epsilon^\lambda(Q))(\epsilon^\mu(K) \cdot \epsilon^\rho(P))],$$

$$\phi^{\rho|\lambda,\mu}_{P|Q,K} = \mp [(P \cdot \epsilon^\lambda(Q))(\epsilon^\rho(P) \cdot \epsilon^\mu(K)) + (P \cdot \epsilon^\mu(K))(\epsilon^\rho(P) \cdot \epsilon^\lambda(Q))],$$

$$\phi^{\lambda|\mu,\rho}_{Q|K,P} = \pm [(Q \cdot \epsilon^\mu(K))(\epsilon^\lambda(Q) \cdot \epsilon^\rho(P)) + (Q \cdot \epsilon^\rho(P))(\epsilon^\lambda(Q) \cdot \epsilon^\mu(K))]. \tag{3.5}$$

Here we take the upper and lower signs for $\lambda = 1$ and 2, respectively. As a typical numerical example, we present here the values of $\overline{\phi}$'s for the first vector of (3.1):

$$(\bar{\phi}^{1|1,1}_{K|P,Q}, \quad \bar{\phi}^{1|1,1}_{P|Q,K}, \quad \bar{\phi}^{1|1,1}_{Q|K,P}) \quad = (-8, \quad 11, \quad -3)/\sqrt{13 \cdot 5 \cdot 2},$$

$$(\bar{\phi}^{1|1,2}_{K|P,Q}, \quad \bar{\phi}^{1|2,1}_{P|Q,K}, \quad \bar{\phi}^{2|1,1}_{Q|K,P}) \quad = (-11, \quad 7, \quad 4)/\sqrt{17 \cdot 13 \cdot 5 \cdot 2},$$

$$(\bar{\phi}^{1|2,1}_{K|P,Q}, \quad \bar{\phi}^{2|1,1}_{P|Q,K}, \quad \bar{\phi}^{1|1,2}_{Q|K,P}) \quad = (11, \quad -2, \quad -9)/\sqrt{6 \cdot 13 \cdot 5 \cdot 2},$$

$$(\bar{\phi}^{1|2,2}_{K|P,Q}, \quad \bar{\phi}^{2|2,1}_{P|Q,K}, \quad \bar{\phi}^{2|1,2}_{Q|K,P}) \quad = (7, \quad -84, \quad 77)/\sqrt{17 \cdot 6 \cdot 13 \cdot 5 \cdot 2},$$

$$(\bar{\phi}^{2|1,1}_{K|P,Q}, \quad \bar{\phi}^{1|1,2}_{P|Q,K}, \quad \bar{\phi}^{1|2,1}_{Q|K,P}) \quad = (-2, \quad -7, \quad 9)/\sqrt{3 \cdot 13 \cdot 5 \cdot 2},$$

$$(\bar{\phi}^{2|1,2}_{K|P,Q}, \quad \bar{\phi}^{1|2,2}_{P|Q,K}, \quad \bar{\phi}^{2|2,1}_{Q|K,P}) \quad = (33, \quad 5, \quad -38)/\sqrt{17 \cdot 3 \cdot 13 \cdot 5 \cdot 2},$$

$$(\bar{\phi}^{2|2,1}_{K|P,Q}, \quad \bar{\phi}^{2|1,2}_{P|Q,K}, \quad \bar{\phi}^{1|2,2}_{Q|K,P}) \quad = (-11, \quad 14, \quad -3)/\sqrt{6 \cdot 3 \cdot 13 \cdot 5 \cdot 2},$$

$$(\bar{\phi}^{2|2,2}_{K|P,Q}, \quad \bar{\phi}^{2|2,2}_{P|Q,K}, \quad \bar{\phi}^{2|2,2}_{Q|K,P})$$

$$= (-111, \quad 120, \quad -9)/\sqrt{17 \cdot 6 \cdot 3 \cdot 13 \cdot 5 \cdot 2}. \tag{3.6}$$

B. The fundamental triad interaction

Each vector in (3.1) represents the fundamental TI

$$\dot{u}^\mu(\mathbf{K}) = -i \sum_{\rho,\lambda} \bar{\phi}^{\mu|\rho,\lambda}_{K|P,Q} \, u^\rho{}^*(\mathbf{P}) u^\lambda{}^*(\mathbf{Q}),$$

$$\dot{u}^\mu(\mathbf{P}) = -i \sum_{\rho,\lambda} \bar{\phi}^{\mu|\rho,\lambda}_{P|Q,K} \, u^\rho{}^*(\mathbf{Q}) u^\lambda{}^*(\mathbf{K}), \tag{3.7}$$

$$\dot{u}^\mu(\mathbf{Q}) = -i \sum_{\rho,\lambda} \bar{\phi}^{\mu|\rho,\lambda}_{Q|K,P} \, u^\rho{}^*(\mathbf{K}) u^\lambda{}^*(\mathbf{P}),$$

over $\mathbf{K} + \mathbf{P} + \mathbf{Q} = 0$. Writing it out in detail, we have

$$
\begin{pmatrix} \dot{u}^1(\mathbf{K}) \\ \dot{u}^1(\mathbf{P}) \\ \dot{u}^1(\mathbf{Q}) \\ \dot{u}^2(\mathbf{K}) \\ \dot{u}^2(\mathbf{P}) \\ \dot{u}^2(\mathbf{Q}) \end{pmatrix}
= -i \left\{
\begin{pmatrix} \bar{\phi}^{1|11}_{K|PQ} \\ \bar{\phi}^{1|11}_{P|QK} \\ \bar{\phi}^{1|11}_{Q|KP} \\ 0 \\ 0 \\ 0 \end{pmatrix}
+ \begin{pmatrix} \bar{\phi}^{1|12}_{K|PQ} \\ \bar{\phi}^{1|21}_{P|QK} \\ 0 \\ 0 \\ 0 \\ \bar{\phi}^{2|11}_{Q|KP} \end{pmatrix}
+ \begin{pmatrix} \bar{\phi}^{1|21}_{K|PQ} \\ 0 \\ \bar{\phi}^{1|12}_{Q|KP} \\ 0 \\ \bar{\phi}^{2|11}_{P|QK} \\ 0 \end{pmatrix}
+ \begin{pmatrix} \bar{\phi}^{1|22}_{K|PQ} \\ 0 \\ 0 \\ 0 \\ \bar{\phi}^{2|21}_{P|QK} \\ \bar{\phi}^{2|12}_{Q|KP} \end{pmatrix}
\right.
$$

$$
\left.
+ \begin{pmatrix} 0 \\ \bar{\phi}^{1|12}_{P|QK} \\ \bar{\phi}^{1|21}_{Q|KP} \\ \bar{\phi}^{2|11}_{K|PQ} \\ 0 \\ 0 \end{pmatrix}
+ \begin{pmatrix} 0 \\ \bar{\phi}^{1|22}_{P|QK} \\ 0 \\ \bar{\phi}^{2|12}_{K|PQ} \\ 0 \\ \bar{\phi}^{2|21}_{Q|KP} \end{pmatrix}
+ \begin{pmatrix} 0 \\ 0 \\ \bar{\phi}^{1|22}_{Q|KP} \\ \bar{\phi}^{2|21}_{K|PQ} \\ \bar{\phi}^{2|12}_{P|QK} \\ 0 \end{pmatrix}
+ \begin{pmatrix} 0 \\ 0 \\ 0 \\ \bar{\phi}^{2|22}_{K|PQ} \\ \bar{\phi}^{2|22}_{P|QK} \\ \bar{\phi}^{2|22}_{Q|KP} \end{pmatrix}
\right\}. \tag{3.8}
$$

The energy conservation follows because the vector $(u^1{}^*(\mathbf{K}), \ u^1{}^*(\mathbf{P}), \ \ldots, \ u^2{}^*(\mathbf{Q}))$ is orthogonal to each and every vector in (3.8), since

$$\bar{\phi}^{\mu|\rho,\lambda}_{K|P,Q} + \bar{\phi}^{\rho|\lambda,\mu}_{P|Q,K} + \bar{\phi}^{\lambda|\mu,\rho}_{Q|K,P} = 0, \tag{3.9}$$

for $\mathbf{K} + \mathbf{P} + \mathbf{Q} = 0$. The overall conservation (1.8) is then a direct consequence of this.

In the spherical representation, the helicity becomes

$$H = \sum_{\mathbf{k}} ik(u^1(\mathbf{k}) u^2{}^*(\mathbf{k}) - u^2(\mathbf{k}) u^1{}^*(\mathbf{k})). \tag{3.10}$$

For the fundamental TI, the helicity conservation (1.10) therefore demands

$$\sum_{\mu,\rho,\lambda=1,2} \left((-1)^\mu \, K \, \bar{\phi}^{\mu|\rho,\lambda}_{K|P,Q} + (-1)^{3-\rho} \, P \, \bar{\phi}^{3-\rho|\lambda,3-\mu}_{P|Q,K} \right.$$

$$\left. + (-1)^{3-\lambda} \, Q \, \bar{\phi}^{3-\lambda|3-\mu,\rho}_{Q|K,P} \right) u^{3-\mu}{}^*(\mathbf{K}) u^\rho{}^*(\mathbf{P}) u^\lambda{}^*(\mathbf{Q})$$

$$+ \text{complex conjugate} = 0. \tag{3.11}$$

Since we do not have a concise analytical demonstration that the sums of three $\bar{\phi}$'s in the parenthesis are identically zero, it is more direct to verify (3.11) by using the typical numerical values for $\bar{\phi}$'s given in (3.6).

At this point in the discussion, we wish to point out the essential difference in the 2D and 3D flow dynamics. In 2D, since there is one polarization vector for a given $\mathbf{k}$, the only possible *quadratic* constant of motion is $|u^1(\mathbf{k})|^2$. Hence, (1.8) and (1.9) may be combined into one constraint:

$$\frac{1}{2} \frac{d}{dt} \sum_{\mathbf{k}} (c_1 + c_2 k^2) |u^1(\mathbf{k})|^2 = 0 \quad (c_1, \ c_2 = \text{const}) . \tag{3.12}$$

Therefore, the energy and enstrophy conservations are invariance restrictions imposed simultaneously on each and every TI. In contradistinction, the *quadratic* constants of motion in 3D are expressed by components of the Hermitian matrix

$$\begin{pmatrix} |u^1(\mathbf{k})|^2 & u^1(\mathbf{k}) u^2{}^*(\mathbf{k}) \\ u^2(\mathbf{k}) u^1{}^*(\mathbf{k}) & |u^2(\mathbf{k})|^2 \end{pmatrix} .$$

Note that (1.8) is invariance of the trace summed over all $\mathbf{k}$, whereas (1.10) states invariance of k times the imaginary part of the off-diagonal, also summed over all $\mathbf{k}$. In the 3D flow, therefore, the energy and helicity conservations are invariance restrictions imposed separately on two groups of TI's.

C. Classification of the triad interactions

The TI's of 3D flow can be divided into four classes, two of which are quite similar to the 2D flow case. First of all, the triad wave vectors of the first two vectors in (3.1) form triangles which are symmetric about the plane containing the k_3 axis and the diagonal of the $k_1 - k_2$ plane, called hereafter the diagonal plane. Similarly, the next six vectors can also be paired off into symmetric pairs. Hence, the first eight vectors will be said belonging to the symmetric pair class. As in the 2D case, the symmetric pair has $\bar{\phi}$'s which are the same except for sign; for example,

$$\left(\bar{\phi}_{1|12}^{\mu|\rho\lambda}, \ \bar{\phi}_{1|21}^{\rho|\lambda\mu}, \ \bar{\phi}_{2|11}^{\lambda|\mu\rho}\right) = (-1)^{\mu+\rho+\lambda}\left(\bar{\phi}_{1|23}^{\mu|\rho\lambda}, \ \bar{\phi}_{2|31}^{\rho|\lambda\mu}, \ \bar{\phi}_{3|12}^{\lambda|\mu\rho}\right).$$

(3.13)

Secondly, the ninth vector in (3.1) has the triad wave vector forming an isosceles triangle, so it belongs to the isosceles class. Similar to (2.11), certain $\bar{\phi}$'s for the base $\mathbf{Q} = \binom{3}{3}$ are identically zero

$$\bar{\phi}_{\mathbf{Q}|\mathbf{K},\mathbf{P}}^{1|\mu,\rho} = 0. \tag{3.14}$$

The TI's mentioned above are essentially the same kinds that we have already encountered in the 2D flow, except that symmetry is now with respect to the diagona plane. In 3D, however, there are TI's whose triad wave vectors lie in either the diagonal plane or the plane containing the k_3 axis and an off-diagonal of the $k_1 - k_2$

plane, called the off-diagonal plane. The last three vectors in (3.1) have triad wave vectors in the diagonal plane, hence they will be said belonging to the diagonal plane class. Although these three classes exhaust the TI's in (3.1), there is another class, called the off-diagonal plane class, having the triad wave vector in an off-diagonal plane. The TI's of such a class, however, appear in symmetric pairs. We shall see them first in the $4 \times 4 \times 4$ lattice. Quite a few $\bar{\phi}$'s vanish identically for the TI of both diagonal and off-diagonal plane classes

$$\bar{\phi}_{\mathbf{K}|\mathbf{P},\mathbf{Q}}^{2|1,1} = \bar{\phi}_{\mathbf{P}|\mathbf{Q},\mathbf{K}}^{2|1,1} = \bar{\phi}_{\mathbf{Q}|\mathbf{K},\mathbf{P}}^{2|1,1} = 0,$$

$$\bar{\phi}_{\mathbf{K}|\mathbf{P},\mathbf{Q}}^{\mu|\rho,\lambda} = \bar{\phi}_{\mathbf{P}|\mathbf{Q},\mathbf{K}}^{\rho|\lambda,\mu} = \bar{\phi}_{\mathbf{Q}|\mathbf{K},\mathbf{P}}^{\lambda|\mu,\rho} = 0,$$

(3.15)

where $(\mu,\rho,\lambda) = (1,1,1)$, $(1,2,2)$, $(2,1,2)$, and $(2,2,1)$.

Again, we denote by $(\mathbf{K},\mathbf{P},\mathbf{Q})_m$ the column vector with the nonzero elements

$$\bar{\phi}_{\mathbf{K}|\mathbf{P},\mathbf{Q}}^{\mu|\rho,\lambda} u^\rho*(\mathbf{P})u^{\lambda}*(\mathbf{Q}), \ \bar{\phi}_{\mathbf{P}|\mathbf{Q},\mathbf{K}}^{\mu|\rho,\lambda} u^\rho*(\mathbf{Q})u^{\lambda}*(\mathbf{K}), \ \bar{\phi}_{\mathbf{Q}|\mathbf{K},\mathbf{P}}^{\mu|\rho,\lambda} u^\rho*(\mathbf{K})u^{\lambda}*(\mathbf{P}),$$

where

$m = 1$ for the isosceles class,

$m = 1'$ for the diagonal plane class,

$m = 2$ for the symmetric pair class,

$m = 2'$ for the off-diagonal plane class.

Using this supercompact notation, (3.1) becomes

$$\dot{u}_3^\mu = -i \sum_{\rho,\lambda} \left\{ \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 1 & 1 & 2 \end{pmatrix}_2 + \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}_2 + \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 1 & 2 & 3 \end{pmatrix}_2 + \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 1 & 2 & 3 \end{pmatrix}_2 + \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix}_1 + \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix}_{1'} + \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 1 & 2 \end{pmatrix}_{1'} + \begin{pmatrix} 2 & 1 & 3 \\ 2 & 1 & 3 \\ 1 & 2 & 3 \end{pmatrix}_{1'} \right\}.$$

(3.16)

Since the subscript m (ignoring any primes) refers to multiplicity, we see that (3.16) represents 12 TI's. After a tedious enumeration, the TI's involving wave vectors in the fourth-order lattice shell can be summarized as follows:

$$\left[\begin{pmatrix} 1 & 1 & 2 \\ 1 & 3 & 4 \\ 1 & 1 & 2 \end{pmatrix}_2\right] + \left[\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 1 & 1 & 2 \end{pmatrix}_2 + \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 4 \\ 1 & 1 & 2 \end{pmatrix}_2 + \begin{pmatrix} 2 & 1 & 3 \\ 1 & 3 & 4 \\ 1 & 1 & 2 \end{pmatrix}_2\right] + \left[\begin{pmatrix} 1 & 3 & 4 \\ 2 & 2 & 4 \\ 1 & 1 & 2 \end{pmatrix}_2 + \begin{pmatrix} 3 & 1 & 4 \\ 1 & 3 & 4 \\ 1 & 1 & 2 \end{pmatrix}_1 + \begin{pmatrix} 1 & 3 & 4 \\ 1 & 3 & 4 \\ 1 & 1 & 2 \end{pmatrix}_{1'}\right]$$

$$+ \left[\begin{pmatrix} 1 & 1 & 2 \\ 1 & 3 & 4 \\ 1 & 2 & 3 \end{pmatrix}_2 + \begin{pmatrix} 1 & 1 & 2 \\ 1 & 3 & 4 \\ 2 & 1 & 3 \end{pmatrix}_2 + \begin{pmatrix} 1 & 1 & 2 \\ 2 & 2 & 4 \\ 1 & 2 & 3 \end{pmatrix}_{2'}\right]$$

$$+ \left[\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 1 & 2 & 3 \end{pmatrix}_2 + \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 2 & 1 & 3 \end{pmatrix}_2 + \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 4 \\ 1 & 2 & 3 \end{pmatrix}_2 + \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 4 \\ 2 & 1 & 3 \end{pmatrix}_2 + \begin{pmatrix} 2 & 1 & 3 \\ 1 & 3 & 4 \\ 1 & 2 & 3 \end{pmatrix}_2 + \begin{pmatrix} 2 & 1 & 3 \\ 1 & 3 & 4 \\ 2 & 1 & 3 \end{pmatrix}_2\right]$$

$$+ \left[\begin{pmatrix} 1 & 3 & 4 \\ 2 & 2 & 4 \\ 1 & 2 & 3 \end{pmatrix}_2 + \begin{pmatrix} 1 & 3 & 4 \\ 2 & 2 & 4 \\ 2 & 1 & 3 \end{pmatrix}_2 + \begin{pmatrix} 3 & 1 & 4 \\ 1 & 3 & 4 \\ 1 & 2 & 3 \end{pmatrix}_2 + \begin{pmatrix} 1 & 3 & 4 \\ 1 & 3 & 4 \\ 1 & 2 & 3 \end{pmatrix}_{1'} + \begin{pmatrix} 1 & 3 & 4 \\ 1 & 3 & 4 \\ 2 & 1 & 3 \end{pmatrix}_{1'} + \begin{pmatrix} 2 & 2 & 4 \\ 2 & 2 & 4 \\ 1 & 2 & 3 \end{pmatrix}_{1'}\right]$$

$$
+\left[\begin{pmatrix}1&1&2\\1&1&2\\1&3&4\end{pmatrix}_{1'}\right]+\left[\begin{pmatrix}1&1&2\\1&2&3\\1&3&4\end{pmatrix}_{2}+\begin{pmatrix}1&1&2\\1&2&3\\2&2&4\end{pmatrix}_{2}+\begin{pmatrix}1&1&2\\1&2&3\\3&1&4\end{pmatrix}_{2}\right]
$$

$$
+\left[\begin{pmatrix}1&2&3\\2&1&3\\1&3&4\end{pmatrix}_{2}+\begin{pmatrix}1&2&3\\2&1&3\\2&2&4\end{pmatrix}_{1}+\begin{pmatrix}1&2&3\\1&2&3\\3&1&4\end{pmatrix}_{1'}+\begin{pmatrix}1&2&3\\1&2&3\\1&3&4\end{pmatrix}_{1'}+\begin{pmatrix}1&2&3\\1&2&3\\2&2&4\end{pmatrix}_{1'}\right]
$$

$$
+\left[\begin{pmatrix}1&1&2\\1&3&4\\1&3&4\end{pmatrix}_{2}+\begin{pmatrix}1&1&2\\1&3&4\\2&2&4\end{pmatrix}_{2}+\begin{pmatrix}1&1&2\\1&3&4\\3&1&4\end{pmatrix}_{2}+\begin{pmatrix}1&1&2\\2&2&4\\1&3&4\end{pmatrix}_{2'}\right]
$$

$$
+\left[\begin{pmatrix}1&2&3\\1&3&4\\1&3&4\end{pmatrix}_{2}+\begin{pmatrix}1&2&3\\1&3&4\\2&2&4\end{pmatrix}_{2}+\begin{pmatrix}1&2&3\\2&2&4\\1&3&4\end{pmatrix}_{2}+\begin{pmatrix}1&2&3\\2&2&4\\2&2&4\end{pmatrix}_{2}+\begin{pmatrix}2&1&3\\1&3&4\\1&3&4\end{pmatrix}_{2}+\begin{pmatrix}2&1&3\\1&3&4\\2&2&4\end{pmatrix}_{2}+\begin{pmatrix}2&1&3\\1&3&4\\3&1&4\end{pmatrix}_{2}\right.
$$

$$
\left.+\begin{pmatrix}1&2&3\\1&3&4\\3&1&4\end{pmatrix}_{2}+\begin{pmatrix}1&2&3\\2&2&4\\3&1&4\end{pmatrix}_{2}\right]+\left[\begin{pmatrix}1&3&4\\3&1&4\\1&3&4\end{pmatrix}_{2}+\begin{pmatrix}2&2&4\\1&3&4\\3&1&4\end{pmatrix}_{2}+\begin{pmatrix}1&3&4\\2&2&4\\1&3&4\end{pmatrix}_{2}+\begin{pmatrix}1&3&4\\2&2&4\\2&2&4\end{pmatrix}_{2}+\begin{pmatrix}1&3&4\\3&1&4\\2&2&4\end{pmatrix}_{1}\right.
$$

$$
\left.+\begin{pmatrix}1&3&4\\1&3&4\\2&2&4\end{pmatrix}_{1'}+\begin{pmatrix}2&2&4\\2&2&4\\1&3&4\end{pmatrix}_{1'}+\begin{pmatrix}1&3&4\\1&3&4\\3&1&4\end{pmatrix}_{1'}\right].
$$

(3.17)

The introduction of the above into (3.16) would complete the equations of eddy motion for the $4\times4\times4$ lattice. As in the 2D case, we can devise a diagrammatic procedure for generating more efficiently the TI's of a given cubic lattice.

4. STATISTICAL SYMMETRIES FOR ISOTROPY

The TI representation is particularly instrumental in bringing out the symmetry conditions necessary for isotropy, thereby allowing us to quantify such non-isotropic effects as the departures from rotational and reflexional invariances. We begin with the velocity correlation tensor

$$R_{ij}(\mathbf{x},\mathbf{r})=(2\pi/L)^{p}\langle U_{i}(\mathbf{x})U_{j}(\mathbf{x}+\mathbf{r})\rangle, \tag{4.1}$$

where p is the dimension, and $\langle\ \rangle$ denotes ensemble average. Its Fourier representation is then

$$R_{ij}(\mathbf{r})=(2\pi/L)^{p}\sum_{\mathbf{k}}\langle U_{i}{}^{*}(\mathbf{k})U_{j}(\mathbf{k})\rangle\exp(i\mathbf{k}\cdot\mathbf{r}). \tag{4.2}$$

Since $(2\pi/L)^{p}\Sigma_{\mathbf{k}}\to\int d\mathbf{k}$ in the limit as $L\to\infty$, we may identify $\langle U_{i}^{*}(\mathbf{k})U_{j}(\mathbf{k})\rangle$ with the spectral tensor

$$\Phi_{ij}(k)=\langle U_{i}^{*}(\mathbf{k})U_{j}(\mathbf{k})\rangle. \tag{4.3}$$

The spectral tensor shows the Hermitian symmetry[9] $\Phi_{ij}(\mathbf{k})=\Phi_{ji}(-\mathbf{k})=\Phi_{ji}^{*}(\mathbf{k})$. Further, it has the isotropic representation

$$\Phi_{ij}(\mathbf{k})=P_{ij}(\mathbf{k})E_{2}(k)/\pi k, \tag{4.4}$$

where $E_{2}(k)$ is the energy spectrum in 2D, and

$$\Phi_{ij}(\mathbf{k})=P_{ij}(\mathbf{k})E_{3}(k)/4\pi k^{2}, \tag{4.5}$$

where $E_{3}(k)$ is the energy spectrum in 3D.

Returning to (4.3), we now express the spectral tensor in terms of (1.3). First, we have in 2D

$$\Phi_{ij}(\mathbf{k})=P_{ij}(\mathbf{k})\langle|u(\mathbf{k})|^{2}\rangle. \tag{4.6}$$

In view of (4.4), isotropy demands $\langle|u(\mathbf{k})|^{2}\rangle$ be polar symmetric in the 2D wave vector space.

$$\langle|u(\mathbf{k})|^{2}\rangle\to E_{2}(k)/\pi k. \tag{4.7}$$

Second, the spectral tensor in 3D becomes

$$\Phi_{ij}(\mathbf{k})=\sum_{\mu}\epsilon_{i}^{\mu}(\mathbf{k})\epsilon_{j}^{\mu}(\mathbf{k})\langle|u^{\mu}(\mathbf{k})|^{2}\rangle+\sum_{\mu\neq\lambda}\epsilon_{i}^{\mu}(\mathbf{k})\epsilon_{j}^{\lambda}(\mathbf{k})\langle u^{\mu*}(\mathbf{k})u^{\lambda}(\mathbf{k})\rangle.$$

(4.8)

Upon comparing with (4.5), we therefore find the conditions necessary for isotropy:

(i) $\langle u^{\mu*}(\mathbf{k})u^{\lambda}(\mathbf{k})\rangle=0$ $(\mu\neq\lambda)$

(ii) $\langle|u^{\mu}(\mathbf{k})|^{2}\rangle$ are independent of μ,   (4.9)

(iii) $\langle|u^{\mu}(\mathbf{k})|^{2}\rangle\to E_{3}(k)/4\pi k^{2}$.

Note that the condition (i) imposes the reflexional symmetry, (ii) guarantees the rotational symmetry, and (iii) demands the spherical symmetry of energy spectrum functions in the 3D wave vector space.

Suppose that (ii) and (iii) are maintained, but the reflexional symmetry (i) is relaxed. Then, the spectral tensor would have the additional term which changes sign with the reflexion of the axes

$$\Phi_{ij}(\mathbf{k})=P_{ij}(\mathbf{k})\frac{E_{3}(k)}{4\pi k^{2}}+\frac{iF(k)}{8\pi k^{4}}\epsilon_{ijm}k_{m}. \tag{4.10}$$

Moffatt[10] calls this the spectral tensor for "pseudo-isotropic" turbulence. He has pointed out that the real function $F(k)$, measuring the lack of reflexional symmetry, can be related to the helicity. Similar to (4.1), define the helicity correlation

$$H(\mathbf{x},\mathbf{r})=(2\pi/L)^{3}\langle\omega_{i}(\mathbf{x})U_{i}(\mathbf{x}+\mathbf{r})\rangle, \tag{4.11}$$

which in the Fourier space becomes

$$H(\mathbf{r})=-(2\pi/L)^{3}\sum_{\mathbf{k}}ik_{j}\epsilon_{ijm}\Phi_{im}(\mathbf{k})\exp(i\mathbf{k}\cdot\mathbf{r}). \tag{4.12}$$

Introducing (4.10) into (4.12) and replacing $(2\pi/L)^{3}\Sigma_{\mathbf{k}}\to\int d\mathbf{k}$, we obtain the spectral form

$$H(0) = \int d\mathbf{k} \, \frac{F(k)}{4\pi k^2}. \tag{4.13}$$

Under the pseudo-isotropy assumption, $\langle u^{\mu*}(\mathbf{k})u^{\lambda}(\mathbf{k})\rangle$ $(\mu \neq \lambda)$ are also the scalar function of k. We can then relate $F(k)$ to the helicity (3.10):

$$F(k)/4\pi k^2 \sim ik(\langle u^1(\mathbf{k})\,u^{2*}(\mathbf{k})\rangle - \langle u^2(\mathbf{k})u^{1*}(\mathbf{k})\rangle). \tag{4.14}$$

5. PROSPECTUS FOR FUTURE WORK

The TI representation has been presented for both the 2D and 3D homogeneous flows. This appears to be more powerful than the usual Fourier-amplitude representation: First of all, the incompressibility is built into the equation as in the vorticity equation. Secondly, for a given wave vector the number of dynamic equations is one less than that of the Fourier-amplitude equations. Thirdly, in the inviscid limit, the 2D flow conserves energy and enstrophy, whereas the 3D flow obeys the energy and helicity conservations. Fourthly, we can categorize the entire family of TI's into two classes in 2D and four classes in 3D, according to the geometry of triad wave vectors. Lastly, the isotropy conditions in 3D are the reflexional, rotational, and spherical symmetries in the wave vector space. On the other hand, polar symmetry is all that we need for the 2D isotropy.

Attempts have already been made to apply the TI representation to turbulence problems. We shall mention here a few of them which are of the immediate concern:

(i) The companion paper[5] provides the proof that energy and enstrophy are the isolating constants of motion in 2D, whereas the 3D flow has energy and helicity as the isolating constants of motion. The existence of the quadratic constants of motion justifies the canonical equilibrium distribution.

(ii) The TI representation is well suited for the formulation of nonisotropic turbulence[6] because one can quantify nonisotropy in terms of departures from the three symmetry requirements.

(iii) We would expect that this representation is also useful for the simulation work, for it involves one less dynamic equation than the Fourier-amplitude representation. Besides, by selecting certain classes of TI's, one might simplify the turbulence dynamics in a systematic fashion without gross misrepresentation. This is purely a conjecture at present, and we wish to explore this in another report.

[1]R.H. Kraichnan, Phys. Rev. 109, 1407 (1958).
[2]R. Balescu and A. Senatorski, Ann. Phys. 58, 587 (1970).
[3]R.H. Kraichnan, J. Fluid Mech. 56, 287 (1972).
[4]J. Lee, J. Math. Phys. 15, 1571 (1974).
[5]J. Lee, J. Math. Phys. 16, 1367 (1975).
[6]J. Lee, "Intrinsic Nonisotropy of the Homogeneous Turbulence with No Mean Flow" (to be published).
[7]H.K. Moffatt, J. Fluid Mech. 35, 117 (1969).
[8]R.H. Kraichnan, Phys. Fluids 6, 1603 (1963).
[9]G.K. Batchelor, *Theory of Homogeneous Turbulence* (Cambridge U.P., Cambridge, 1956), Chaps. II and III.
[10]H.K. Moffatt, J. Fluid Mech. 41, 435 (1970).

Isolating constants of motion for the homogeneous turbulence of two and three dimensions

Jon Lee

Aerospace Research Laboratories, Wright-Patterson AFB, Ohio 45433
(Received 24 May 1974; revised manuscript received 22 November 1974)

For the inviscid eddy motion in a finite-dimensioned Fourier space, it is stated that energy and enstrophy are the isolating constants of motion for the 2D homogeneous turbulence. In contrast, the 3D isotropic turbulence has energy as the only constant of motion. If we relax the reflexional invariance, however, helicity emerges as another invariant; hence energy and helicity are said to be the isolating constants of motion for the helical turbulence. Although these are the key assumptions in the construction of equilibrium distributions, they have heretofore been accepted, without proof, as a natural property of the Navier–Stokes dynamics. This paper provides the proof. We have shown here that quadratic constants of motion for the individual triad-interactions collapse to energy–enstrophy in 2D, but to energy and helicity in 3D.

1. PROBLEM STATEMENT

Following the statistical mechanical argument, statistical equilibrium states have been deduced for both the 2D and 3D homogeneous, inviscid turbulent flows. In 2D, the equipartition of energy—enstrophy among the Fourier modes follows directly from the canonical distribution which is a function of energy and enstrophy.[1] If we suppose that total energy is the only isolating constant of motion for the 3D isotropic turbulence, then energy equipartition is the immediate consequence[2–4] and hence the equilibrium energy spectrum of the form k^2. Now, as Moffatt[5] has pointed out, helicity emerges as another constant of motion when we relax the reflexional invariance of isotropic turbulence. Kraichnan[6] has therefore modified the k^2 energy spectrum by incorporating both energy and helicity into the equilibrium distribution.

As one might infer from the above, the existence of statistical equilibrium states is crucially hinged on two things: One is the knowledge of isolating constants of motion and the other is the Gaussian property of equilibrium distribution, i. e., the canonical distribution. Note that the canonical distribution can lead to equipartition states only when constants of motion are of the quadratic form. In the classical statistical mechanics,[7,8] energy equipartition is deduced for an isolated mechanical system which has the Maxwell–Boltzmann distribution as equilibrium solution of the Liouville equation. If there is another constant of motion besides energy, the trajectories will be restricted to the intersection of the energy surface and a hypersurface specified by that constant of motion. Then, energy equipartition is not possible because the trajectory cannot cover every extension of the energy surface, i. e., the metric decomposability.

In this paper, we shall prove that energy and enstrophy are the quadratic isolating constants of motion in 2D, whereas the 3D homogeneous turbulence has energy and helicity as the constants of motion. In the so-called triad-interaction (TI) representation,[9] we have decomposed the quadratic nonlinearity of homogeneous turbulence into the TI's extending over infinitely coupled triad wave vectors. Since each TI has two constants of motion (Sec. 2), the main point of investigation is the

fate of isolating constants of motion for the individual TI's as they are coupled according to the Navier—Stokes dynamics. In 2D, the constants of motion of each TI correspond to the individual energy and enstrophy respectively. Then, coupling the TI's of the 2D Navier—Stokes equations, we find that quadratic constants of motion for the individual TI's collapse to total energy and enstrophy (Sec. 3). The situation is, however, quite different in 3D. There the energy and helicity conservations are observed by two groups of TI's. Strict isotropy admits only one group of TI's. Upon coupling these TI's, it is found that all but one constant of motion become destroyed, thereby leaving total energy as the only isolating constant of motion for the 3D isotropic turbulence (Sec. 4A). When the reflexional invariance is relaxed, the other group of TI's also comes into play. Coupling them, however, constants of motion for the individual TI's collapse to helicity (Sec. 4B). Hence, total energy and helicity survive as the quadratic isolating constants of motion for the 3D homogeneous turbulence.

2. THE TRIAD-INTERACTION SYSTEM

In the companion paper[9] (called hereafter I), we have presented the detailed categorization of TI's which constitute the nonlinear dynamics of both 2D and 3D flows. It therefore behooves us to investigate here the constant of motion of a single TI system (I. 2. 3) written as

$$\dot{u}(\mathbf{k}_1) = - i\phi_1 u^*(\mathbf{k}_2)u^*(\mathbf{k}_3),$$
$$\dot{u}(\mathbf{k}_2) = - i\phi_2 u^*(\mathbf{k}_3)u^*(\mathbf{k}_1), \qquad (2.1)$$
$$\dot{u}(\mathbf{k}_3) = - i\phi_3 u^*(\mathbf{k}_1)u^*(\mathbf{k}_2),$$

where $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0$, and $\phi_1 = \bar{\phi}_{\mathbf{k}_1 | \mathbf{k}_2, \mathbf{k}_3}$, $\phi_2 = \bar{\phi}_{\mathbf{k}_2 | \mathbf{k}_3, \mathbf{k}_1}$, $\phi_3 = \bar{\phi}_{\mathbf{k}_3 | \mathbf{k}_1, \mathbf{k}_2}$. This is identical to the 3-complex-mode system numerically investigated by Kraichnan.[10] The energy—enstrophy conservation (I. 3. 12) is assured by

$$a_1 \phi_1 + a_2 \phi_2 + a_3 \phi_3 = 0, \qquad (2.2)$$

where $a_n = c_1 + c_2 k_n^2$ ($c_1, c_2 = $ const). Now decompose (2. 1) into the real and imaginary parts by $u(\mathbf{k}_n) = A_n + iB_n$:

$$\begin{pmatrix} \dot{A}_1 \\ \dot{A}_2 \\ \dot{A}_3 \\ \dot{B}_1 \\ \dot{B}_2 \\ \dot{B}_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \phi_1 B_2 B_3 \\ \phi_2 B_3 B_1 \\ \phi_3 B_1 B_3 \end{pmatrix} - \begin{pmatrix} \phi_1 A_2 B_3 \\ \phi_2 B_3 A_1 \\ 0 \\ 0 \\ 0 \\ \phi_3 A_1 A_2 \end{pmatrix}$$

$$- \begin{pmatrix} \phi_1 B_2 A_3 \\ 0 \\ \phi_3 A_1 B_2 \\ 0 \\ \phi_2 A_3 A_1 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ \phi_2 A_3 B_1 \\ \phi_3 B_1 A_2 \\ \phi_1 A_2 A_3 \\ 0 \\ 0 \end{pmatrix}. \qquad (2.3)$$

This conserves energy—enstrophy because $(a_1 A_1, a_2 A_2, \ldots, a_3 B_3)$ is orthogonal to each and every vector in (2.3).

Prior to investigating (2.3) in its entirety, insight may be gained from a simpler 3-mode system represented by any vector, e.g.,

$$\dot{B}_1 = \phi_1 B_2 B_3,$$
$$\dot{B}_2 = \phi_2 B_3 B_1, \qquad (2.4)$$
$$\dot{B}_3 = \phi_3 B_1 B_2,$$

with the constraints

$$\phi_1 + \phi_2 + \phi_3 = 0, \qquad (2.5)$$

$$k_1^2 \phi_1 + k_2^2 \phi_2 + k_3^2 \phi_3 = 0. \qquad (2.6)$$

We notice, however, that (2.4) is the classical Euler's problem[11] of a rigid body moving with one point fixed under no external forces. Since the energy

$$\tfrac{1}{2}(B_1^2 + B_2^2 + B_3^2) = e, \qquad (2.7)$$

is the constant of motion, the trajectory must lie on the energy sphere. But we also find that

$$\tfrac{1}{2}(B_1^2 - B_2^2 \phi_1/\phi_2) = \text{const},$$
$$\tfrac{1}{2}(B_1^2 - B_3^2 \phi_1/\phi_3) = \text{const}. \qquad (2.8)$$

Clearly, the linear combination of (2.8) is also a constant of motion:

$$\tfrac{1}{2}[(\beta+\gamma)B_1^2 - \beta(\phi_1/\phi_2)B_2^2 - \gamma(\phi_1/\phi_3)B_3^2] = \text{const}, \qquad (2.9)$$

where β and γ are the scaling constants. The trajectories are therefore restricted to the intersection of the energy sphere and the ellipsoid (2.9), i.e., the polhode of Poinsot's geometric solution. With the choice of $\beta = -(\phi_2/\phi_1)k_2^2$ and $\gamma = -(\phi_3/\phi_1)k_3^2$, (2.9) reduces to the enstrophy

$$\tfrac{1}{2}(k_1^2 B_1^2 + k_2^2 B_2^2 + k_3^2 B_3^2) = \omega, \qquad (2.10)$$

provided (2.6) is obeyed. Of course, all this could have been deduced alternately from the analytical solution in terms of the Jacobian elliptic functions.

Returning to the complete system, we now look for constants of motion of the quadratic form

$$\sum_{i=1}^{3} (\alpha_i A_i^2 + \alpha_{i+3} B_i^2) = \text{const}. \qquad (2.11)$$

Here, α_i must satisfy the equations imposed by (2.3):

$$\alpha_1 \phi_1 + \alpha_2 \phi_2 + \alpha_6 \phi_3 = 0,$$
$$\alpha_1 \phi_1 + \alpha_5 \phi_2 + \alpha_3 \phi_3 = 0,$$
$$\alpha_4 \phi_1 + \alpha_2 \phi_2 + \alpha_3 \phi_3 = 0, \qquad (2.12)$$
$$\alpha_4 \phi_1 + \alpha_5 \phi_2 + \alpha_6 \phi_3 = 0.$$

Without circumlocution, we consider the case $\alpha_1 = \alpha_4$, $\alpha_2 = \alpha_5$, $\alpha_3 = \alpha_6$. Then, (2.11) and (2.12) reduce respectively to

$$\alpha_1 |u(k_1)|^2 + \alpha_2 |u(k_2)|^2 + \alpha_3 |u(k_3)|^2 = \text{const}, \qquad (2.13)$$

$$\alpha_1 \phi_1 + \alpha_2 \phi_2 + \alpha_3 \phi_3 = 0. \qquad (2.14)$$

Posed in this way, the constants of motion are solutions of (2.14). Obviously, the solution vectors $\alpha^1 = (1, -\phi_1/\phi_2, 0)$ and $\alpha^2 = (1, 0, -\phi_1/\phi_3)$ correspond to (2.8). Also the linear combination

$$\alpha = ((\beta+\gamma), -\beta\phi_1/\phi_2, -\gamma\phi_1/\phi_3), \qquad (2.15)$$

is a solution corresponding to (2.9). In view of (2.5), (2.15) reduces to $\alpha^e = (1, 1, 1)$ with $\beta = -\phi_2/\phi_1$ and $\gamma = -\phi_3/\phi_1$. It further becomes $\alpha^\omega = (k_1^2, k_2^2, k_3^2)$ with $\beta = -k_2^2 \phi_2/\phi_1$ and $\gamma = -k_3^2 \phi_3/\phi_1$, when (2.6) is obeyed. In order words, if (2.5) and (2.6) are imposed, the solution (2.15) can be spanned alternately by α^e and α^ω. Geometrically speaking, α is in a plane perpendicular to the vector (ϕ_1, ϕ_2, ϕ_3). Let us call it the α plane. Since the nonlinear interactions are built up of TI's, we can determine the isolating constant of motion of a turbulent flow by the intersection of the α planes.

3. TWO-DIMENSIONAL FLOW

The TI's of the 2D flow have been grouped in (I.2.12) up to the fifth-order lattice shell. Since each TI has two constants of motion, we shall show here that coupling TI's à la (I.2.12) does not generate extraneous constants of motion. In other words, the intersection of α planes is a hypersurface spanned by α^e and α^ω in an appropriate Euclidean space. Denote the wave vectors in plane lattice by $k_1 = \binom{1}{1}$, $k_2 = \binom{1}{2}$, $k_3 = \binom{2}{2}$, $k_4 = \binom{2}{1}$, $k_5 = \binom{1}{3}$, $k_6 = \binom{2}{3}$, $k_7 = \binom{3}{3}$, $k_8 = \binom{3}{2}$, etc.

A. The 3 × 3 lattice system

For constants of motion of the quadratic form

$$\sum_{\substack{i=1 \\ \neq 3, 5}}^{8} \alpha_i |u(k_i)|^2 = \text{const}, \qquad (3.1)$$

the three vectors in (I.2.2) impose the following conditions:

$$\alpha_1 \bar{\phi}_{k_1 | k_2, k_6} + \alpha_2 \bar{\phi}_{k_2 | k_6, k_1} + \alpha_6 \bar{\phi}_{k_6 | k_1, k_2} = 0,$$
$$\alpha_1 \bar{\phi}_{k_1 | k_4, k_8} + \alpha_4 \bar{\phi}_{k_4 | k_8, k_1} + \alpha_8 \bar{\phi}_{k_8 | k_1, k_4} = 0, \qquad (3.2)$$
$$\alpha_2 \bar{\phi}_{k_2 | k_4, k_7} + \alpha_4 \bar{\phi}_{k_4 | k_7, k_2} + \alpha_7 \bar{\phi}_{k_7 | k_2, k_4} = 0.$$

Since the first two are redundant because of (I.2.10), we may consider only one of them, say

$$\alpha_1 \phi_{11} + \alpha_2 \phi_{12} + \alpha_6 \phi_{13} = 0, \qquad (3.3)$$

where $\phi_{11} = \bar{\phi}_{\mathbf{k}_1 \mid \mathbf{k}_2, \mathbf{k}_6}$, $\phi_{12} = \bar{\phi}_{\mathbf{k}_2 \mid \mathbf{k}_6, \mathbf{k}_1}$, $\phi_{13} = \bar{\phi}_{\mathbf{k}_6 \mid \mathbf{k}_1, \mathbf{k}_2}$, together with

$$\alpha_2 = \alpha_4 \quad \text{and} \quad \alpha_6 = \alpha_8. \tag{3.4}$$

Furthermore, in view of (I.2.11), the last of (3.2) gives rise to $\alpha_2 = \alpha_4$ and $\alpha_7 =$ arbitrary. In this way, the constant of motion for the 3×3 lattice is determined by a single equation (3.3) identical to (2.14), hence total energy and enstrophy are the only isolating constants of motion.

B. The 4 × 4 lattice system

Consider constants of motion

$$\sum_{\substack{i=1 \\ \neq 7,10}}^{15} \alpha_i \mid u(\mathbf{k}_i) \mid^2 = \text{const}. \tag{3.5}$$

After eliminating redundant relations, we obtain from the first two square brackets of (I.2.12)

$$
\begin{aligned}
&\alpha_1 \bar{\phi}_{\mathbf{k}_1 \mid \mathbf{k}_2, \mathbf{k}_6} + \alpha_2 \bar{\phi}_{\mathbf{k}_2 \mid \mathbf{k}_6, \mathbf{k}_1} + \alpha_6 \bar{\phi}_{\mathbf{k}_6 \mid \mathbf{k}_1, \mathbf{k}_2} = 0, \\
&\alpha_1 \bar{\phi}_{\mathbf{k}_1 \mid \mathbf{k}_5, \mathbf{k}_{11}} + \alpha_5 \bar{\phi}_{\mathbf{k}_5 \mid \mathbf{k}_{11}, \mathbf{k}_1} + \alpha_{11} \bar{\phi}_{\mathbf{k}_{11} \mid \mathbf{k}_1, \mathbf{k}_5} = 0, \\
&\alpha_1 \bar{\phi}_{\mathbf{k}_1 \mid \mathbf{k}_6, \mathbf{k}_{12}} + \alpha_6 \bar{\phi}_{\mathbf{k}_6 \mid \mathbf{k}_{12}, \mathbf{k}_1} + \alpha_{12} \bar{\phi}_{\mathbf{k}_{12} \mid \mathbf{k}_1, \mathbf{k}_6} = 0, \\
&\alpha_2 \bar{\phi}_{\mathbf{k}_2 \mid \mathbf{k}_3, \mathbf{k}_{12}} + \alpha_3 \bar{\phi}_{\mathbf{k}_3 \mid \mathbf{k}_{12}, \mathbf{k}_2} + \alpha_{12} \bar{\phi}_{\mathbf{k}_{12} \mid \mathbf{k}_2, \mathbf{k}_3} = 0, \\
&\alpha_2 \bar{\phi}_{\mathbf{k}_4 \mid \mathbf{k}_5, \mathbf{k}_{12}} + \alpha_5 \bar{\phi}_{\mathbf{k}_5 \mid \mathbf{k}_{12}, \mathbf{k}_4} + \alpha_{12} \bar{\phi}_{\mathbf{k}_{12} \mid \mathbf{k}_4, \mathbf{k}_5} = 0, \\
&\alpha_2 \bar{\phi}_{\mathbf{k}_4 \mid \mathbf{k}_6, \mathbf{k}_{13}} + \alpha_6 \bar{\phi}_{\mathbf{k}_6 \mid \mathbf{k}_{13}, \mathbf{k}_4} + \alpha_{13} \bar{\phi}_{\mathbf{k}_{13} \mid \mathbf{k}_4, \mathbf{k}_6} = 0,
\end{aligned} \tag{3.6}
$$

with $\alpha_2 = \alpha_4$, $\alpha_5 = \alpha_9$, $\alpha_6 = \alpha_8$, $\alpha_{11} = \alpha_{15}$, and $\alpha_{12} = \alpha_{14}$. Abbreviating the coupling coefficients by ϕ_{ij}, we put it in the matrix form

$$
\begin{pmatrix}
\phi_{11} & \phi_{12} & 0 & 0 & \phi_{13} & 0 & 0 & 0 \\
\phi_{21} & 0 & 0 & \phi_{22} & 0 & \phi_{23} & 0 & 0 \\
\phi_{31} & 0 & 0 & 0 & \phi_{32} & 0 & \phi_{33} & 0 \\
0 & \phi_{41} & \phi_{42} & 0 & 0 & 0 & \phi_{43} & 0 \\
0 & \phi_{51} & 0 & \phi_{52} & 0 & 0 & \phi_{53} & 0 \\
0 & \phi_{61} & 0 & 0 & \phi_{62} & 0 & 0 & \phi_{63}
\end{pmatrix} \alpha = 0, \tag{3.7}
$$

where the column vector $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_5, \alpha_6, \alpha_{11}, \alpha_{12}, \alpha_{13})$. Since the rank of coefficient matrix is 6, the solution can be spanned by two vectors of the dimension 8:

$$\alpha = \beta \alpha^1 + \gamma \alpha^2, \tag{3.8}$$

where

$$\alpha^1 = (1, -\phi_{11}/\phi_{12}, (\phi_{11}\phi_{41}/\phi_{12} + \phi_{43}\phi_{31}/\phi_{33})/\phi_{42}, p_1/\phi_{52}, 0,$$
$$- (\phi_{21} + \phi_{22}p_1/\phi_{52})/\phi_{23}, -\phi_{31}/\phi_{33}, \phi_{11}\phi_{61}/\phi_{12}\phi_{63}),$$
$$\alpha^2 = (1, 0, \phi_{43}p_2/\phi_{42}\phi_{33}, \phi_{53}p_2/\phi_{52}\phi_{33}, -\phi_{11}/\phi_{13},$$
$$- (\phi_{21} + \phi_{53}\phi_{22}p_2/\phi_{52}\phi_{33})/\phi_{23}, -p_2/\phi_{33}, \phi_{11}\phi_{62}/\phi_{13}\phi_{63}).$$

Here, $p_1 = \phi_{11}\phi_{51}/\phi_{12} + \phi_{31}\phi_{53}/\phi_{33}$ and $p_2 = \phi_{31} - \phi_{11}\phi_{32}/\phi_{13}$. Using the energy constraint

$$\phi_{i1} + \phi_{i2} + \phi_{i3} = 0, \quad (i = 1, \ldots, 6), \tag{3.9}$$

we can reduce (3.8) to $\alpha^e = (1,1,1,1,1,1,1,1)$ with $\beta = -\phi_{12}/\phi_{11}$ and $\gamma = -\phi_{13}/\phi_{11}$. Also, invoking the enstrophy constraint

$$k_1^2\phi_{11} + k_2^2\phi_{12} + k_6^2\phi_{13} = 0,$$
$$k_1^2\phi_{21} + k_5^2\phi_{22} + k_{11}^2\phi_{23} = 0,$$
$$k_1^2\phi_{31} + k_6^2\phi_{32} + k_{12}^2\phi_{33} = 0,$$
$$k_2^2\phi_{41} + k_3^2\phi_{42} + k_{12}^2\phi_{43} = 0, \tag{3.10}$$
$$k_2^2\phi_{51} + k_5^2\phi_{52} + k_{12}^2\phi_{53} = 0,$$
$$k_2^2\phi_{61} + k_6^2\phi_{62} + k_{13}^2\phi_{63} = 0,$$

(3.8) reduces to $\alpha^\omega = (k_1^2, k_2^2, k_3^2, k_5^2, k_6^2, k_{11}^2, k_{12}^2, k_{13}^2)$ with $\beta = -k_2^2\phi_{12}/\phi_{11}$ and $\gamma = -k_6^2\phi_{13}/\phi_{11}$. We have thus shown that α^e and α^ω span the solution in the 8D Euclidean space. That is, coupling six irreducible TI's does not generate extraneous constants of motion, hence this lattice system has no isolating constant of motion other than total energy and enstrophy.

C. The higher-order lattice system

For the 5×5 lattice system, the analysis becomes very tedious due to the proliferation of TI's; however, the topological structure of constants of motion remains the same. Considering all the TI's in (I.2.12), we now obtain a system of 20 equations but only for 14 unknowns [note that both (3.3) and (3.7) are underdetermined systems]. For this over-determined system, we find that the rank of coefficient matrix is 12 and hence α^e and α^ω again span the solution in the 14D Euclidean space. For a lattice of the higher order, the rank of coefficient matrix cannot be less than the number of unknowns by more than 2 because the system is always over-determined. This therefore assures that the intersection of the α planes is spanned by α^e and α^ω. Consequently, the 2D flow has no isolating constant of motion other than total energy and enstrophy.

4. THREE-DIMENSIONAL FLOW

Consider the Hermitian matrix

$$
\begin{pmatrix}
\mid u^1(\mathbf{k}_n) \mid^2 & u^1(\mathbf{k}_n)u^{2*}(\mathbf{k}_n) \\
u^2(\mathbf{k}_n)u^{1*}(\mathbf{k}_n) & \mid u^2(\mathbf{k}_n) \mid^2
\end{pmatrix}. \tag{4.1}
$$

As pointed out in Sec. 3B of I, the trace is energy

$$\mid u^1(\mathbf{k}_n) \mid^2 + \mid u^2(\mathbf{k}_n) \mid^2, \tag{4.2}$$

and helicity is given by the off-diagonals

$$ik(u^1(\mathbf{k}_n)u^{2*}(\mathbf{k}_n) - u^2(\mathbf{k}_n)u^{1*}(\mathbf{k}_n)). \tag{4.3}$$

Since the energy and helicity conservations are mutually exclusive constraints on the two separate groups of TI's, we can investigate constants of motion for isotropic turbulence independent of those of helical turbulence.

Before examining the $3 \times 3 \times 3$ lattice (I.3.16), we shall establish constants of motion of the fundamental TI under the various triad wave vector configurations. Labeling $\mathbf{K}$, $\mathbf{P}$, and $\mathbf{Q}$ by $\mathbf{k}_1$, $\mathbf{k}_2$, and $\mathbf{k}_3$, respectively, we rewrite (I.3.8) by abbreviating the coupling coefficients by ϕ_{ij}.

$$
\begin{pmatrix} \dot{u}^1(\mathbf{k}_1) \\ \dot{u}^1(\mathbf{k}_2) \\ \dot{u}^1(\mathbf{k}_3) \\ \dot{u}^2(\mathbf{k}_1) \\ \dot{u}^2(\mathbf{k}_2) \\ \dot{u}^2(\mathbf{k}_3) \end{pmatrix} = -i \left\{ \begin{pmatrix} \phi_{11} u^{1*}(\mathbf{k}_2) u^{1*}(\mathbf{k}_3) \\ \phi_{12} u^{1*}(\mathbf{k}_3) u^{1*}(\mathbf{k}_1) \\ \phi_{13} u^{1*}(\mathbf{k}_1) u^{1*}(\mathbf{k}_2) \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \phi_{21} u^{1*}(\mathbf{k}_2) u^{2*}(\mathbf{k}_3) \\ \phi_{22} u^{2*}(\mathbf{k}_3) u^{1*}(\mathbf{k}_1) \\ 0 \\ 0 \\ 0 \\ \phi_{23} u^{1*}(\mathbf{k}_1) u^{1*}(\mathbf{k}_2) \end{pmatrix} \right.
$$

$$
+ \begin{pmatrix} \phi_{31} u^{2*}(\mathbf{k}_2) u^{1*}(\mathbf{k}_3) \\ 0 \\ \phi_{33} u^{1*}(\mathbf{k}_1) u^{2*}(\mathbf{k}_2) \\ 0 \\ \phi_{32} u^{1*}(\mathbf{k}_3) u^{1*}(\mathbf{k}_1) \\ 0 \end{pmatrix} + \begin{pmatrix} \phi_{41} u^{2*}(\mathbf{k}_2) u^{2*}(\mathbf{k}_3) \\ 0 \\ 0 \\ 0 \\ \phi_{42} u^{2*}(\mathbf{k}_3) u^{1*}(\mathbf{k}_1) \\ \phi_{43} u^{1*}(\mathbf{k}_1) u^{2*}(\mathbf{k}_2) \end{pmatrix} + \begin{pmatrix} 0 \\ \phi_{52} u^{1*}(\mathbf{k}_3) u^{2*}(\mathbf{k}_1) \\ \phi_{53} u^{2*}(\mathbf{k}_1) u^{1*}(\mathbf{k}_2) \\ \phi_{51} u^{1*}(\mathbf{k}_2) u^{1*}(\mathbf{k}_3) \\ 0 \\ 0 \end{pmatrix}
$$

$$
+ \begin{pmatrix} 0 \\ \phi_{62} u^{2*}(\mathbf{k}_3) u^{2*}(\mathbf{k}_1) \\ 0 \\ \phi_{61} u^{1*}(\mathbf{k}_2) u^{2*}(\mathbf{k}_3) \\ 0 \\ \phi_{63} u^{2*}(\mathbf{k}_1) u^{1*}(\mathbf{k}_2) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \phi_{73} u^{2*}(\mathbf{k}_1) u^{2*}(\mathbf{k}_2) \\ \phi_{71} u^{2*}(\mathbf{k}_2) u^{1*}(\mathbf{k}_3) \\ \phi_{72} u^{1*}(\mathbf{k}_3) u^{2*}(\mathbf{k}_1) \\ 0 \end{pmatrix} + \left. \begin{pmatrix} 0 \\ 0 \\ 0 \\ \phi_{81} u^{2*}(\mathbf{k}_2) u^{2*}(\mathbf{k}_3) \\ \phi_{82} u^{2*}(\mathbf{k}_3) u^{2*}(\mathbf{k}_1) \\ \phi_{83} u^{2*}(\mathbf{k}_1) u^{2*}(\mathbf{k}_2) \end{pmatrix} \right\}.
$$

(4.4)

The ϕ_{ij} obey the energy constraint (I.3.9),

$$\phi_{i1} + \phi_{i2} + \phi_{i3} = 0, \quad (i = 1, \ldots, 8),$$ (4.5)

and the helicity constraint (I.3.11),

$$-k_1\phi_{11} + k_2\phi_{72} + k_3\phi_{63} = 0,$$
$$-k_1\phi_{21} + k_2\phi_{82} - k_3\phi_{53} = 0,$$
$$-k_1\phi_{31} - k_2\phi_{52} + k_3\phi_{83} = 0,$$
$$-k_1\phi_{41} - k_2\phi_{62} - k_3\phi_{73} = 0,$$
$$k_1\phi_{51} + k_2\phi_{32} + k_3\phi_{23} = 0,$$ (4.6)
$$k_1\phi_{61} + k_2\phi_{42} - k_3\phi_{13} = 0,$$
$$k_1\phi_{71} - k_2\phi_{12} + k_3\phi_{43} = 0,$$
$$k_1\phi_{81} - k_2\phi_{22} - k_3\phi_{33} = 0.$$

A. Isotropic turbulence

Consider the constant of motion of the energy form (4.2),

$$\sum_{i=1}^{3} \sum_{\mu=1,2} , \quad \alpha_i^\mu \left| u^\mu(\mathbf{k}_i) \right|^2 = \text{const.}$$ (4.7)

The eight vectors in (4.4) impose that

$$
\begin{pmatrix} \phi_{11} & \phi_{12} & \phi_{13} & 0 & 0 & 0 \\ \phi_{21} & \phi_{22} & 0 & 0 & 0 & \phi_{23} \\ \phi_{31} & 0 & \phi_{33} & 0 & \phi_{32} & 0 \\ \phi_{41} & 0 & 0 & 0 & \phi_{42} & \phi_{43} \\ 0 & \phi_{52} & \phi_{53} & \phi_{51} & 0 & 0 \\ 0 & \phi_{62} & 0 & \phi_{61} & 0 & \phi_{63} \\ 0 & 0 & \phi_{73} & \phi_{71} & \phi_{72} & 0 \\ 0 & 0 & 0 & \phi_{81} & \phi_{82} & \phi_{83} \end{pmatrix} \alpha = 0,
$$ (4.8)

where the column vector $\alpha = (\alpha_1^1, \alpha_2^1, \alpha_3^1, \alpha_1^2, \alpha_2^2, \alpha_3^2)$.

First of all, for the symmetric pair class $(m = 2)$ (Sec. 3C of I), ϕ_{ij} are typically nonzero. One may reduce (4.8) to a form

$$
\begin{pmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \\ \psi_{31} & \psi_{32} \\ \psi_{41} & \psi_{42} \end{pmatrix} \begin{pmatrix} \alpha_2^1 \\ \alpha_3^1 \end{pmatrix} = 0,
$$ (4.9)

where α_1^1 and α_i^2 are expressed in terms of α_2^1 and α_3^1. Without explicitly presenting here the expressions for ψ_{ij}, we simply state that the rank of (ψ_{ij}) is 1. This in turn implies that the coefficient matrix of (4.8) has the rank 5, hence the solution is $\beta\alpha^e$. In other words, the intersection of eight α planes is simply the vector α^e itself for the TI of $m = 2$. Secondly, for the isosceles class $(m = 1)$, (I.3.14) reduces (4.8) to

$$\begin{pmatrix} (\phi_{21}+\phi_{22}) & \phi_{23} \\ (\phi_{41}+\phi_{42}) & \phi_{43} \\ (\phi_{61}+\phi_{62}) & \phi_{63} \\ (\phi_{81}+\phi_{82}) & \phi_{83} \end{pmatrix} \begin{pmatrix} \alpha_1^1 \\ \alpha_3^1 \end{pmatrix} = 0, \qquad (4.10)$$

where $\alpha_1^1 = \alpha_2^1 = \alpha_1^2 = \alpha_2^2$ and $\alpha_3^1 =$ arbitrary. Because of (4.5), the coefficient matrix of (4.10) has the rank 1, hence the solution is again $\beta\alpha^e$ for the TI of $m=1$. Lastly, the situation is quite different for the diagonal and off-diagonal plane classes ($m=1'$ and $2'$). There, in view of (I.3.15), we find that (4.8) reduces to a single equation

$$\alpha_1^2 \phi_{81} + \alpha_2^2 \phi_{82} + \alpha_3^2 \phi_{83} = 0, \qquad (4.11)$$

with $\alpha_1^1 = \alpha_2^1 = \alpha_3^1$. Therefore, the TI's of $m=1'$ and $2'$ have two constants of motion, whereas the TI's of $m=1$ and 2 had none other than energy.

Having separately examined TI's of the different classes, we are now in a position to collectively investigate constants of motion for the $3\times3\times3$ lattice. For the notational compactness, let us relabel the wave vectors in cubic lattice

$$\mathbf{k}_1 = \begin{pmatrix}1\\1\\1\end{pmatrix}, \quad \mathbf{k}_2 = \begin{pmatrix}1\\2\\1\end{pmatrix}, \quad \mathbf{k}_3 = \begin{pmatrix}2\\1\\1\end{pmatrix}, \quad \mathbf{k}_4 = \begin{pmatrix}2\\2\\1\end{pmatrix}, \quad \mathbf{k}_5 = \begin{pmatrix}1\\1\\2\end{pmatrix},$$

$$\mathbf{k}_6 = \begin{pmatrix}1\\2\\2\end{pmatrix}, \quad \mathbf{k}_7 = \begin{pmatrix}2\\1\\2\end{pmatrix}, \quad \mathbf{k}_8 = \begin{pmatrix}2\\3\\2\end{pmatrix}, \quad \mathbf{k}_9 = \begin{pmatrix}3\\2\\2\end{pmatrix}, \quad \mathbf{k}_{10} = \begin{pmatrix}3\\3\\2\end{pmatrix},$$

$$\mathbf{k}_{11} = \begin{pmatrix}2\\2\\3\end{pmatrix}, \quad \mathbf{k}_{12} = \begin{pmatrix}2\\3\\3\end{pmatrix}, \quad \mathbf{k}_{13} = \begin{pmatrix}3\\2\\3\end{pmatrix}, \quad \mathbf{k}_{14} = \begin{pmatrix}3\\3\\3\end{pmatrix}.$$

Then, for the constant of motion

$$\sum_{i=1}^{14} \sum_{\mu=1,2} \alpha_i^\mu \, |u^\mu(\mathbf{k}_i)|^2 = \text{const.} \qquad (4.12)$$

the equations that α_i^μ must satisfy are ($\mu, \rho, \lambda = 1, 2$)

$$\alpha_1^\mu \overline{\phi}_{\mathbf{k}_1|\mathbf{k}_2, \mathbf{k}_8}^{\mu|\rho,\lambda} + \alpha_2^\mu \overline{\phi}_{\mathbf{k}_2|\mathbf{k}_8, \mathbf{k}_1}^{\mu|\rho,\lambda} + \alpha_8^\mu \overline{\phi}_{\mathbf{k}_8|\mathbf{k}_1, \mathbf{k}_2}^{\mu|\rho,\lambda} = 0,$$

$$\alpha_1^\mu \overline{\phi}_{\mathbf{k}_1|\mathbf{k}_6, \mathbf{k}_{12}}^{\mu|\rho,\lambda} + \alpha_6^\mu \overline{\phi}_{\mathbf{k}_6|\mathbf{k}_{12}, \mathbf{k}_1}^{\mu|\rho,\lambda} + \alpha_{12}^\mu \overline{\phi}_{\mathbf{k}_{12}|\mathbf{k}_1, \mathbf{k}_6}^{\mu|\rho,\lambda} = 0,$$

$$\alpha_2^\mu \overline{\phi}_{\mathbf{k}_2|\mathbf{k}_5, \mathbf{k}_{12}}^{\mu|\rho,\lambda} + \alpha_5^\mu \overline{\phi}_{\mathbf{k}_5|\mathbf{k}_{12}, \mathbf{k}_2}^{\mu|\rho,\lambda} + \alpha_{12}^\mu \overline{\phi}_{\mathbf{k}_{12}|\mathbf{k}_2, \mathbf{k}_5}^{\mu|\rho,\lambda} = 0,$$

$$\alpha_2^\mu \overline{\phi}_{\mathbf{k}_2|\mathbf{k}_7, \mathbf{k}_{14}}^{\mu|\rho,\lambda} + \alpha_7^\mu \overline{\phi}_{\mathbf{k}_7|\mathbf{k}_{14}, \mathbf{k}_2}^{\mu|\rho,\lambda} + \alpha_{14}^\mu \overline{\phi}_{\mathbf{k}_{14}|\mathbf{k}_2, \mathbf{k}_7}^{\mu|\rho,\lambda} = 0, \qquad (4.13)$$

$$\alpha_2^\mu \overline{\phi}_{\mathbf{k}_2|\mathbf{k}_3, \mathbf{k}_{10}}^{\mu|\rho,\lambda} + \alpha_3^\mu \overline{\phi}_{\mathbf{k}_3|\mathbf{k}_{10}, \mathbf{k}_2}^{\mu|\rho,\lambda} + \alpha_{10}^\mu \overline{\phi}_{\mathbf{k}_{10}|\mathbf{k}_2, \mathbf{k}_3}^{\mu|\rho,\lambda} = 0,$$

$$\alpha_1^2 \overline{\phi}_{\mathbf{k}_1|\mathbf{k}_5, \mathbf{k}_{11}}^{2|2,2} + \alpha_5^2 \overline{\phi}_{\mathbf{k}_5|\mathbf{k}_{11}, \mathbf{k}_1}^{2|2,2} + \alpha_{11}^2 \overline{\phi}_{\mathbf{k}_{11}|\mathbf{k}_1, \mathbf{k}_5}^{2|2,2} = 0,$$

$$\alpha_1^2 \overline{\phi}_{\mathbf{k}_1|\mathbf{k}_4, \mathbf{k}_{10}}^{2|2,2} + \alpha_4^2 \overline{\phi}_{\mathbf{k}_4|\mathbf{k}_{10}, \mathbf{k}_1}^{2|2,2} + \alpha_{10}^2 \overline{\phi}_{\mathbf{k}_{10}|\mathbf{k}_1, \mathbf{k}_4}^{2|2,2} = 0,$$

$$\alpha_4^2 \overline{\phi}_{\mathbf{k}_4|\mathbf{k}_5, \mathbf{k}_{14}}^{2|2,2} + \alpha_5^2 \overline{\phi}_{\mathbf{k}_5|\mathbf{k}_{14}, \mathbf{k}_4}^{2|2,2} + \alpha_{14}^2 \overline{\phi}_{\mathbf{k}_{14}|\mathbf{k}_4, \mathbf{k}_5}^{2|2,2} = 0,$$

with $\alpha_2^1 = \alpha_2^3 = \alpha_1^2 = \alpha_3^2$, $\alpha_6^\mu = \alpha_7^\mu$, $\alpha_8^\mu = \alpha_9^\mu$, $\alpha_{12}^\mu = \alpha_{13}^\mu$, $\alpha_1^1 = \alpha_5^1 = \alpha_{11}^1$, $\alpha_1^1 = \alpha_4^1 = \alpha_{10}^1$, and $\alpha_4^1 = \alpha_5^1 = \alpha_{14}^1$. Equation (4.13) can be traced to TI's in the order that they appear in (I.3.16).

Since the first four of (4.13) refer to the TI of $m=2$, each has the full representation (4.8). The fifth referring to $m=1$ has the reduced form (4.10), whereas the last three all referring to $m=1'$, are of the form (4.11). Summing up, (4.13) represents 39 equations for the 24 unknowns, some of which are, however, constrained by the equality relations. Writing it in the matrix form, we can show that the rank of coefficient matrix is 23, and hence the solution is $\beta\alpha^e$. That is, the intersection of 39

α planes is along the vector α^e in the 24D Euclidean space. For a finite cubic lattice of the higher order, the number of TI's always exceeds that of the unknowns [see (I.3.17)], so that the rank of coefficient matrix cannot be less than the number of unknowns by more than 1. This therefore assures that α^e is the intersection of α planes for the TI's of a finite-dimensional lattice system. Hence, total energy survives as the only isolating constant of motion for the isotropic turbulence.

B. Helical turbulence

If the reflexional invariance is relaxed, the off-diagonals of (4.1) are no longer zero. Hence, helical turbulence will admit constants of motion of the helicity form (4.3)

$$\sum_{i=1}^{3} i(\alpha_i^1 u^1(\mathbf{k}_i) u^{2*}(\mathbf{k}_i) - \alpha_i^2 u^2(\mathbf{k}_i) u^{1*}(\mathbf{k}_i)) = \text{const.} \qquad (4.14)$$

In analogy to (4.8), we have from (4.4)

$$\begin{pmatrix} \phi_{11} & 0 & 0 & 0 & -\phi_{72} & -\phi_{63} \\ \phi_{21} & 0 & \phi_{53} & 0 & -\phi_{82} & 0 \\ \phi_{31} & \phi_{52} & 0 & 0 & 0 & -\phi_{83} \\ \phi_{41} & \phi_{62} & \phi_{73} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\phi_{51} & -\phi_{32} & -\phi_{23} \\ 0 & 0 & \phi_{13} & -\phi_{61} & -\phi_{42} & 0 \\ 0 & \phi_{12} & 0 & -\phi_{71} & 0 & -\phi_{43} \\ 0 & \phi_{22} & \phi_{33} & -\phi_{81} & 0 & 0 \end{pmatrix} \alpha = 0. \qquad (4.15)$$

For the symmetric pair class, we can reduce (4.15) to a form similar to (4.9).

$$\begin{pmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \\ \psi_{31} & \psi_{32} \\ \psi_{41} & \psi_{42} \end{pmatrix} \begin{pmatrix} \alpha_1^1 \\ \alpha_2^1 \end{pmatrix} = 0, \qquad (4.16)$$

where α_3^1 and α_i^2 are expressed in terms of α_1^1 and α_2^1. Since the rank of (ψ_{ij}) turns out to be 1, the TI of $m=2$ has the solution $\beta\alpha^h$, where $\alpha^h = (k_1, k_2, k_3, k_1, k_2, k_3)$. Next, for the isosceles class, (I.3.14) reduces (4.15) to

$$\begin{pmatrix} (\phi_{11}-\phi_{72}) & -\phi_{63} \\ (\phi_{31}+\phi_{52}) & -\phi_{83} \\ -(\phi_{32}+\phi_{51}) & -\phi_{23} \\ (\phi_{12}-\phi_{71}) & -\phi_{43} \end{pmatrix} \begin{pmatrix} \alpha_1^1 \\ \alpha_3^1 \end{pmatrix} = 0, \qquad (4.17)$$

with $\alpha_1^1 = \alpha_2^1 = \alpha_1^2 = \alpha_2^2$ and $\alpha_3^1 =$ arbitrary. By using the numerical ϕ_{ij} values for a typical TI of $m=1$, we can check that the coefficient matrix of (4.17) has the rank 1. Hence, the intersection of eight α planes is again the vector α^h.

Finally, for the diagonal and off-diagonal plane classes, (I.3.15) reduces (4.15) to

$$\begin{pmatrix} \phi_{21} & 0 & \phi_{53} & 0 & -\phi_{82} & 0 \\ \phi_{31} & \phi_{52} & 0 & 0 & 0 & -\phi_{83} \\ 0 & \phi_{22} & \phi_{33} & -\phi_{81} & 0 & 0 \end{pmatrix} \alpha = 0. \qquad (4.18)$$

The rank of coefficient matrix is 3, hence the solution is

$$\alpha = \beta\alpha^1 + \gamma\alpha^2 + \delta\alpha^3, \tag{4.19}$$

where

$$\alpha^1 = (1, 0, 0, 0, \phi_{21}/\phi_{82}, \phi_{31}/\phi_{83}),$$

$$\alpha^2 = (0, 1, 0, -\phi_{21}/\phi_{81}, 0, \phi_{52}/\phi_{83}),$$

$$\alpha^3 = (0, 0, 1, -\phi_{31}/\phi_{81}, -\phi_{52}/\phi_{82}, 0).$$

Since (4.19) would reduce to α^h with the choice of $\beta = k_1, \gamma = k_2, \delta = k_3$, it is concluded that the TI's of $m = 1'$ and $2'$ have two isolating constants of motion besides helicity, while the TI's of $m = 1$ and 2 had none other than helicity.

Following the discussion of Sec. 4A when we collectively examine the 12 TI's of the $3 \times 3 \times 3$ lattice, the intersection of α planes turns out to be along the vector α^h in the 24D Euclidean space. This is because there are still more equations than the unknowns, but the rank of coefficient matrix is just one less than the number of unknowns. Consequently, for a cubic lattice of the arbitrary order, helicity remains as the only constant of motion for the TI's related to the off-diagonals of (4.1), just as those related to the diagonals have none other than energy.

5. EQUILIBRIUM DISTRIBUTIONS

Total energy and enstrophy are the isolating constants of motion for the 2D homogeneous turbulence in a finite-dimensional Fourier space. Consider the canonical distribution $\mathcal{J}_2$ which is a function of the energy and enstrophy of N Fourier modes

$$\mathcal{J}_2 = \prod_{n=1}^{N} [a(k_n)/\pi] \exp[-a(k_n)|u(k_n)|^2]. \tag{5.1}$$

where $a(k_n) = c_1 + c_2 k_n^2$. Following the classical statistical mechanics,[12] one can readily show that (5.1) is the equilibrium solution of the Liouville equation for the 2D inviscid motion of N Fourier modes. Here the conservation property (2.2) plays the essential role. We may, perhaps, justify the Gaussian form for the inviscid eddy distribution (5.1) by invoking the *central limit* theorem,[13] although the actual turbulent distribution is far from being Gaussian. The modal energies averaged over $\mathcal{J}_2$ are

$$\langle |u(k_n)|^2 \rangle = 1/a(k_n), \tag{5.2}$$

implying energy—enstrophy equipartition. This then leads[1] to the equilibrium energy spectrum (I.4.7)

$$E_2(k) \sim \frac{k}{c_1 + c_2 k^2}, \tag{5.3}$$

for the 2D isotropic turbulence.

Since the equilibrium distribution of the 3D homogeneous turbulence must be a function of energy and helicity, we write the canonical distribution $\mathcal{J}_3$ in the form (see the Appendix)

$$\mathcal{J}_3 = \prod_{n=1}^{N} \frac{\sqrt{\det D_n}}{\pi^2} \exp[-(\mathbf{u}, Q\mathbf{u})]. \tag{5.4}$$

Here,

$$\mathbf{u} = \begin{pmatrix} u^1(\mathbf{k}_n) \\ u^2(\mathbf{k}_n) \end{pmatrix},$$

$$Q = \begin{pmatrix} c_3 & -ic_4 k_n \\ ic_4 k_n & c_3 \end{pmatrix}, \quad D_n = \begin{pmatrix} c_3 + c_4 k_n & 0 \\ 0 & c_3 - c_4 k_n \end{pmatrix},$$

and $(,)$ is the scalar product. Taking average of (4.1) over $\mathcal{J}_3$, we get

$$\begin{pmatrix} \langle |u^1(\mathbf{k}_n)|^2 \rangle & \langle u^1(\mathbf{k}_n)u^{2*}(\mathbf{k}_n) \rangle \\ \langle u^2(\mathbf{k}_n)u^{1*}(\mathbf{k}_n) \rangle & \langle |u^2(\mathbf{k}_n)|^2 \rangle \end{pmatrix} = \begin{pmatrix} c_3/C & -ic_4 k_n/C \\ ic_4 k_n/C & c_3/C \end{pmatrix},$$

$$\tag{5.5}$$

where $C = c_3^2 - c_4^2 k_n^2$. The equilibrium energy spectrum (I.4.9) is given by the diagonals

$$E_3(k) \sim \frac{c_3 k^2}{c_3^2 - c_4^2 k^2}, \tag{5.6}$$

and the off-diagonals give the helicity spectrum (I.4.14),

$$F(k) \sim \frac{c_4 k^4}{c_3^2 - c_4^2 k^2}, \tag{5.7}$$

agreeing with Kraichnan's[6] results. Kraichnan has noted that the equilibrium dynamics in 3D are devoid of the positive/negative temperature states of the 2D flow.[14,15] This is because energy and enstrophy are the invariants simultaneously constraining each and every TI in 2D, hence they can compete with each other for dynamic dominance. In the 3D flow, however, the energy and helicity conservations are mutually exclusive restrictions on the diagonal and off-diagonal elements of (4.1). They will therefore affect dynamics of the two distinct groups of TI's. When we impose the reflexional invariance, the cross correlation $\langle u^\mu(\mathbf{k}_n)u^{\lambda*}(\mathbf{k}_n) \rangle$ ($\mu \neq \lambda$) vanishes identically, so that the covariance matrix (5.5) reduces to a diagonal form with the same elements. Thus the energy spectrum goes over smoothly to the energy equipartition state in the limit of zero helicity $(c_4 \rightarrow 0)$.

APPENDIX: DERIVATION OF THE CANONICAL DISTRIBUTION (5.4)

In view of (4.2) and (4.3), we write $\mathcal{J}_3$ as a function of the energy and helicity of N Fourier modes

$$\mathcal{J}_3 = \prod_{n=1}^{N} K_n \exp[-c_3(|u^1(\mathbf{k}_n)|^2 + |u^2(\mathbf{k}_n)|^2)$$
$$- ic_4 k_n (u^1(\mathbf{k}_n)u^{2*}(\mathbf{k}_n) - u^2(\mathbf{k}_n)u^{1*}(\mathbf{k}_n))], \quad (c_3, c_4 = \text{const}).$$

The immediate task is the evaluation of the normalization constant K_n. To this end, consider the integral

$$I = \int \exp[-(\mathbf{u}, Q\mathbf{u})]\, d\mathbf{u}.$$

Since Q is Hermitian, the quadratic form can be diagonalized by $\mathbf{u} = P\mathbf{x}$ $[P = 2^{-1/2}(\begin{smallmatrix} 1 & 1 \\ i & i \end{smallmatrix})]$:

$$I = \int \exp[-(\mathbf{x}, D_n\mathbf{x})]\, d\mathbf{x}.$$

We can now normalize the quadratic by

$$\mathbf{x} = D_n^{-1/2}\mathbf{z} \quad \left[D_n^{-1/2} = \begin{pmatrix} (c_3 + c_4 k_n)^{-1/2} & 0 \\ 0 & (c_3 - c_4 k_n)^{-1/2} \end{pmatrix}\right].$$

Whence,

$$K_n^{-1} = (\det D_n)^{-1/2} \int \exp[-(\mathbf{z}, \mathbf{z})]\, d\mathbf{z} = \pi^2 (\det D_n)^{-1/2}.$$

To get the last equality, we have used the manipulations of complex Gaussian distribution.[16]

ACKNOWLEDGMENT

[1]R.H. Kraichnan, Phys. Fluids 10, 1417 (1967).
[2]J.M. Burgers, Proc. Acad. Sci., Amsterdam 32, 414 (1929); 36, 620 (1933); and Verhandel. Kon. Nederl. Akad. Wentenshap., Amsterdam (Afd. Natuurk) 17, 1 (1939).
[3]T.D. Lee, Quart. Appl. Math. 10, 69 (1952).
[4]E. Hopf, J. Rat. Mech. Anal. 1, 87 (1952); E. Hopf and E. Titt, 2, 587 (1953).
[5]H.K. Moffatt, J. Fluid Mech. 35, 117 (1969).
[6]R.H. Kraichnan, J. Fluid Mech. 59, 745 (1973).
[7]R.C. Tolman, The Principles of Statistical Mechanics (Oxford U.P., Oxford, 1938).
[8]D. ter Haar, Elements of Statistical Mechanics (Rinehart, New York, 1954).
[9]J. Lee, J. Math. Phys. 16, 1359 (1975).
[10]R.H. Kraichnan, Phys. Fluids 6, 1603 (1963).
[11]H. Lamb, Higher Mechanics (Cambridge U.P., Cambridge, 1943).
[12]P. Ehrenfest and T. Ehrenfest, The Conceptual Foundations of the Statistical Approaches in Mechanics (Cornell U.P., Ithaca, N.Y., 1959).
[13]A.I. Khinchin, Mathematical Foundations of Statistical Mechanics (Dover, New York, 1949).
[14]I. Cook and J.B. Taylor, Phys. Rev. Lett. 28, 82 (1972).
[15]D. Montgomery, Phys. Lett. A 39, 7 (1972).
[16]N.R. Goodman, Ann. Math. Stat. 34, 152 (1963).

An invariance theory for second-order variational problems

J. D. Logan and J. S. Blakeslee

Department of Mathematics, Kansas State University, Manhattan, Kansas 66502
(Received 24 January 1975)

This paper investigates the invariance properties of second-order variational problems when the configuration space is subjected to an r-parameter local Lie group of transformations. In particular, the recent results of Hanno Rund on first-order problems are extended to the higher order case: A new set of fundamental invariance identities are derived for single and multiple integral problems, and new proofs of the Zermelo conditions and Noether's theorem are presented. The results are applied to a variational problem whose second-order Lagrangian depends upon a scalar field in Minkowski space, and some conformal identities are obtained.

1. INTRODUCTION

By a second-order variational problem we mean a variational problem whose Lagrange function depends on derivatives up to the second order. In this paper we investigate the invariance properties of such problems when the configuration space is subjected to an r-parameter local Lie group of transformations. The significance of second-order problems, particularly in elasticity and relativity, is well known.

The invariance theory for first-order problems, i.e., problems in which the Lagrangian depends only on first derivatives, is well documented. In particular, we point out the now classical paper of Noether[1] in which she shows that first integrals, or conservation laws, can be obtained directly and explicitly from the invariance properties of the variational integral, be it of any given order. More recently, Rund[2] has derived a new set of fundamental invariance identities for first-order problems which provide simple access to the Noether theorem and have interesting consequences of their own. Some of these consequences have been explored in a series of articles by Logan.[3,4,5]

The goal of this communication is to extend Rund's results to second-order problems and to investigate the consequences thereof. It will be shown, for example, that Noether's theorem in the higher derivative case follows simply from the second-order invariance identities, thereby providing a new, simpler proof of the Noether theorem which avoids completely the so-called "fundamental variational formula" which expresses the total variation of J when both the dependent and independent variables undergo "infinitesimal" variations. To a certain degree, this formula is complicated and difficult to prove rigorously (see Anderson[6] for a derivation in the higher derivative case). In addition to leading to Noether's theorem, the invariance identities that we shall obtain will provide a method for determining groups of transformations under which the fundamental integral is invariant—and therefore Noether's theorem may be applied to determine first integrals of the Euler equations, which in the present case are fourth-order ordinary differential equations (in general nonlinear). Interpreted still differently, the invariance identities, which will involve the Lagrangian L and the generators of the group of transformations, can be used to characterize all Lagrangians which are invariant under a given group of transformations. We also shall have

some brief remarks and comparisons to make concerning canonical variables for second-order problems, as well as give a new derivation of the Zermelo conditions for these problems. Finally, we shall apply the results to obtain a set of invariance identities for multiple integral problems which are invariant under the so-called special conformal group in the case that the field functions are conformal scalars.

2. INVARIANCE CRITERIA

To carry out the above program we shall study the multiple integral problem and deduce the results for single integrals as a special case. We consider, then, a variational integral whose Lagrangian depends upon m independent coordinates, n "field" functions, and first and second derivatives of these field functions. Notationally, we have

$$J = \int_{G_m} L(t^1, \ldots, t^m; x^1(t), \ldots, x^n(t); \ldots, \dot{x}^k_\alpha(t), \ldots; \ldots,$$

$$\ddot{x}^k_{\alpha\beta}(t), \cdots) \, dt^1 \cdots dt^m \qquad (2.1)$$

where G_m is a simply-connected bounded region in R^m, $t = (t^1, \ldots, t^m)$, and

$$\dot{x}^k_\alpha = \frac{\partial x^k}{\partial t^\alpha}, \quad \ddot{x}^k_{\alpha\beta} = \frac{\partial^2 x^k}{\partial t^\beta \partial t^\alpha}.$$

The "dots" are over the latter quantities to remind us that the quantities are derivatives. Here, and in the sequel, small case italic indices j, k, l range over $1, \ldots, n$, while lower case Greek indices α, $\beta, \ldots$ range over $1, \ldots, m$. The summation convention will be used freely, with one additional notational convenience; in the Lagrangian we assume the second derivatives $\ddot{x}_{\alpha\beta}$ occur only when $\alpha \leqslant \beta$. Therefore, in the summation convention, when terms of the form $L_{\ddot{x}_{\alpha\beta}}$ (or $\partial^2 L / \partial \ddot{x}_{\alpha\beta}$) are summed on either index α or β, we assume that it is carried out in such a way that $\alpha \leqslant \beta$. We also assume that L is of class C^3 in each of its arguments and that the $x^k(t)$ are of class $C^4(G_m)$.

We now consider the r-parameter family of transformations

$$\bar{t}^\alpha = \phi^\alpha(t, \epsilon^1, \ldots, \epsilon^r),$$

$$\qquad (2.2)$$

$$\bar{x}^k = \psi^k(t, x; \epsilon^1, \ldots, \epsilon^r)$$

on (t, x)-space where $t = (t^1, \ldots, t^m)$ and $x = (x^1, \ldots, x^n)$. We assume that ϕ^α and ψ^k are of class C^∞ in each of

their arguments. It is further assumed that to the values $\epsilon^1 = \epsilon^2 = \cdots = \epsilon^r = 0$ of the parameters correspond the identity transformation $\bar{t} = t$, $\bar{x} = x$. Also, we suppose that the transformation $\bar{t}^\alpha = \phi^\alpha(t, \epsilon^1, \ldots, \epsilon^r)$ can be inverted to obtain $t^\alpha = T^\alpha(\bar{t}, \epsilon^1, \ldots, \epsilon^r)$. Then, any m-dimensional hypersurface $x = x(t)$, $t \in G_m$, in (t, x)-space gets mapped via (2.2) into an r-parameter family of hypersurfaces $\bar{x} = \bar{x}(\bar{t})$ in $(\bar{t}, \bar{x})$-space. This family is given by

$$\bar{x}^k = \psi^k(T(\bar{t}, \epsilon), x(T(\bar{t}, \epsilon)), \epsilon) = \bar{x}^k(\bar{t})$$

where we have denoted $\epsilon = (\epsilon^1, \ldots, \epsilon^r)$, $T(\bar{t}, \epsilon) = (T^1(\bar{t}, \epsilon), \ldots, T^m(\bar{t}, \epsilon))$, and $x = (x^1, \ldots, x^n)$. It is important to note that, according to definition,

$$x^k(\phi(t, \epsilon)) = \psi^k(t, x(t), \epsilon) \tag{2.3}$$

where $\phi = (\phi^1, \ldots, \phi^m)$. At this point we remark that we have required that the $\bar{t}$ transformation in (2.2) not depend on x in order to insure the invertibility of the transformation for every $x(t)$.

Since $\epsilon = 0$ gives the identity transformation, it follows from expanding the right-hand sides of (2.2) in a Taylor series about $\epsilon = 0$ that

$$\bar{t}^\alpha = t^\alpha + \epsilon^s \tau_s^\alpha(t) + o(\epsilon),$$

$$\bar{x}^k = x^k + \epsilon^s \xi_s^k(t, x) + o(\epsilon), \tag{2.4}$$

where $o(\epsilon)$ are terms for which $o(\epsilon)/\|\epsilon\| \to 0$ as $\|\epsilon\| \to 0$. The quantities τ_s^α and ξ_s^k are given by

$$\tau_s^\alpha = \left(\frac{\partial \phi^\alpha}{\partial \epsilon^s}\right)_0 \quad \xi_s^k(t, x) = \left(\frac{\partial \psi^k}{\partial t^\alpha}\right)_0 \tag{2.5}$$

where $(\ )_0$ denotes the fact that the quantity is to be evaluated at $\epsilon = 0$. The quantities τ_s^α and ξ_s^k are commonly called the generators of the infinitesimal transformation given by (2.4). Here, and subsequently, the index s will range over $s = 1, \ldots, r$.

We are now in position to define what is understood by saying that J defined by (2.1) is invariant under the r-parameter family of transformations (2.2). Essentially, we require that $J(\bar{x}(\bar{t})) - J(x(t))$ is $o(\epsilon)$ plus terms which are divergences and linear in the ϵ^α. More precisely:

Definition 1: The fundamental integral defined by (2.1) is divergence-invariant under the r-parameter family of transformations (2.2) if, and only if, there exist mr functions $F_s^\alpha = F_s^\alpha(t, x)$ of class C^1 such that

$$\int_{\bar{G}_m} L\left(\bar{t}, \ldots, \bar{t}^m; \ldots, \bar{x}^k(\bar{t}), \ldots, ; \ldots, \frac{\partial \bar{x}^k}{\partial \bar{t}^\alpha}(\bar{t}), \ldots, ; \ldots, \right.$$

$$\left. \frac{\partial^2 \bar{x}}{\partial \bar{t}^\beta \partial \bar{t}^\alpha}(\bar{t}), \cdots \right) d\bar{t}^1 \cdots d\bar{t}^m$$

$$- \int_{G_m} L(t^1, \ldots, t^m; \ldots, x^k(t), \ldots, ; \ldots, \dot{x}_\alpha^k(t), \ldots, ; \tag{2.6}$$

$$\ldots, \ddot{x}_{\alpha\beta}^k(t), \cdots) dt^1 \cdots dt^m$$

$$= o(\epsilon) + \epsilon^s \int_{G_m} \frac{\partial}{\partial t^\alpha} F_s^\alpha(t, x(t)) dt^1 \cdots dt^m$$

for every region G_m, for every ϵ, and for every $x(t) = (x^1(t), \ldots, x^n(t))$ of class C^4.

3. THE INVARIANCE IDENTITIES

We now state and prove the following fundamental result which relates the Lagrangian and its derivatives to the generators of the infinitesimal family of transformations (2.4).

Theorem 1: If the fundamental integral (2.1) is divergence-invariant under the r-parameter family of transformations (2.2), then it is necessary that the Lagrangian satisfy the r identity relations

$$L_{t^\alpha} \tau_s^\alpha + L_{x^k} \xi_s^k + L_{\dot{x}_\alpha^k}\left(\frac{d\xi_s^k}{dt^\alpha} - \dot{x}_\beta^k \frac{d\tau_s^\beta}{dt^\alpha}\right)$$

$$+ L_{\ddot{x}_{\alpha\beta}^k}\left(\frac{d^2\xi_s^k}{dt^\alpha dt^\beta} - \ddot{x}_{\gamma\beta}^k \frac{d\tau_s^\gamma}{dt^\alpha} - \ddot{x}_{\alpha\gamma}^k \frac{d\tau_s^\gamma}{dt^\beta} - \dot{x}_\gamma^k \frac{d^2\tau_s^\gamma}{dt^\alpha dt^\beta}\right) \tag{3.1}$$

$$+ L \frac{d\tau_s^\alpha}{dt^\alpha} - \frac{d}{dt^\alpha} F_s^\alpha = 0$$

where $s = 1, \ldots, r$, and the subscripts on L denote partial differentiation.

The proof of this theorem is similar to the proof of the first-order case given by Rund[2]; only a few additional difficulties are encountered. The basic idea is to differentiate (2.6) with respect to ϵ^s and afterward set $\epsilon = 0$. To this end, we obtain

$$\frac{\partial}{\partial \epsilon^s}\left(\int_{\bar{G}_m} L\left(\bar{t}^1, \ldots, \bar{t}^m; \ldots, \bar{x}^k(\bar{t}), \ldots, \frac{\partial \bar{x}^k}{\partial \bar{t}^\alpha}(\bar{t}), \ldots; \right.\right.$$

$$\left.\left. \cdots, \frac{\partial^2 \bar{x}}{\partial \bar{t}^\alpha \partial \bar{t}^\beta}(\bar{t}), \cdots\right) d\bar{t}^1 \cdots d\bar{t}^m\right)_0$$

$$= \int_{G_m} \frac{d}{dt^\alpha} F_s^\alpha dt^1 \cdots dt^m.$$

We now change variables in the integral on the left according to

$$\bar{t}^\alpha = \phi^\alpha(t, \epsilon). \tag{3.2}$$

Hence, we obtain

$$\frac{\partial}{\partial \epsilon^s}\left(\int_{G_m} L\left(t^1, \ldots, t^m; \ldots, \bar{x}^k(\bar{t}), \ldots; \ldots, \frac{\partial \bar{x}^k}{\partial \bar{t}^\alpha}, \ldots; \right.\right.$$

$$\left.\left. \cdots, \frac{\partial^2 \bar{x}^k}{\partial \bar{t}^\beta \partial \bar{t}^\alpha}, \cdots\right) \det\left(\frac{\partial \bar{t}^\alpha}{\partial t^\beta}\right) dt^1 \cdots dt^m\right)_0$$

$$= \int_{G_m} \frac{d}{dt^\alpha} F_s^\alpha dt^1 \cdots dt^m$$

where it is implicitly understood that $\bar{t}^\alpha$ in the integral on the left is given by (3.2). By differentiating under the integral sign and noting that the resulting expression is valid for all regions of integration G_m, we conclude that

$$\frac{\partial}{\partial \epsilon^s}\left(L\left(\bar{t}^1, \ldots, \bar{t}^m; \ldots, \bar{x}^k(\bar{t}), \ldots; \ldots, \frac{\partial \bar{x}^k}{\partial \bar{t}^\alpha}, \ldots; \right.\right.$$

$$\left.\left. \cdots, \frac{\partial^2 \bar{x}^k}{\partial \bar{t}^\beta \partial \bar{t}^\alpha}, \cdots\right)\right)_0 \tag{3.3}$$

$$+ L\left(\frac{\partial}{\partial \epsilon^s}\det\left(\frac{\partial \bar{t}^\alpha}{\partial t^\beta}\right)\right)_0 = \frac{dF_s^\alpha}{dt^\alpha}$$

where we also have used the fact that

$$\det\left(\frac{\partial \bar{t}^\alpha}{\partial t^\beta}\right)_0 = 1.$$

An easy calculation (see Rund[2]) shows that

$$\left(\frac{\partial}{\partial \epsilon^s}\det\left(\frac{\partial \bar{t}^\alpha}{\partial t^\beta}\right)\right)_0 = \frac{d\tau_s^\alpha}{dt^\alpha}. \tag{3.4}$$

Therefore, upon expanding $\partial L/\partial \epsilon^s$ in (3.3), we obtain

$$L_{t^\alpha}\tau_s^\alpha + L_{x^k}\xi_s^k + L_{\bar{x}_\alpha^k}\left(\frac{\partial}{\partial \epsilon^s}\frac{\partial \bar{x}^k}{\partial \bar{t}^\alpha}\right)_0 + L_{\bar{x}_{\alpha\beta}^k}\left(\frac{\partial}{\partial \epsilon^s}\frac{\partial^2 \bar{x}^k}{\partial \bar{t}^\beta \partial \bar{t}^\alpha}\right)_0$$

$$\tag{3.5}$$

$$+ L\frac{d\tau_s^\alpha}{dt^\alpha} - \frac{dF_s^\alpha}{dt^\alpha} = 0 \quad (s=1,\ldots,r).$$

To compute the remaining two terms which are evaluated at $\epsilon = 0$ in (3.5), we differentiate both sides of (2.3) with respect to t^β to obtain

$$\frac{\partial \bar{x}^k}{\partial \bar{t}^\alpha}\frac{\partial \bar{t}^\alpha}{\partial t^\beta} = \frac{\partial \bar{x}^k}{\partial t^\beta}. \tag{3.6}$$

Next, we differentiate (3.6) with respect to ϵ^s to get, after simplification and evaluation at $\epsilon = 0$,

$$\left(\frac{\partial}{\partial \epsilon^s}\frac{\partial \bar{x}^k}{\partial \bar{t}^\beta}\right)_0 = \frac{d\xi_s^k}{dt^\beta} - \dot{x}_\alpha^k\frac{d\tau_s^\alpha}{dt^\beta}. \tag{3.7}$$

To obtain (3.7) we also noted that

$$\left(\frac{\partial}{\partial \epsilon^s}\frac{\partial \bar{x}^k}{\partial t^\beta}\right)_0 = \frac{d\xi_s^k}{dt^\beta}, \quad \left(\frac{\partial}{\partial \epsilon^s}\frac{\partial \bar{t}^\alpha}{\partial t^\beta}\right)_0 = \frac{d\tau_s^\alpha}{dt^\beta}$$

which follow easily from (2.4).

Now, we differentiate (3.6) with respect to t^γ to obtain

$$\frac{\partial^2 \bar{x}^k}{\partial \bar{t}^\delta \partial \bar{t}^\alpha}\frac{\partial \bar{t}^\delta}{\partial t^\gamma}\frac{\partial \bar{t}^\alpha}{\partial t^\beta} + \frac{\partial \bar{x}^k}{\partial \bar{t}^\alpha}\frac{\partial^2 \bar{t}^\alpha}{\partial t^\gamma \partial t^\beta} = \frac{\partial^2 \bar{x}^k}{\partial t^\gamma \partial t^\beta}.$$

Differentiation of this expression with respect to ϵ^s yields

$$\left(\frac{\partial}{\partial \epsilon^s}\frac{\partial^2 \bar{x}^k}{\partial t^\gamma \partial t^\beta}\right)_0 = \frac{d^2 \xi_s^k}{dt^\gamma dt^\beta} - \ddot{x}_{\delta3}^k\frac{d\tau_s^\delta}{dt^\gamma} - \ddot{x}_{\gamma\alpha}^k\frac{d\tau_s^\alpha}{dt^\beta} - \dot{x}_\alpha^k\frac{d^2\tau_s^\alpha}{dt^\gamma dt^\beta} \tag{3.8}$$

where we have again noted from (2.4) that

$$\left(\frac{\partial}{\partial \epsilon^s}\frac{\partial^2 \bar{t}^\alpha}{\partial t^\gamma \partial t^\beta}\right)_0 = \frac{d^2\tau_s^\alpha}{dt^\gamma dt^\beta}, \quad \left(\frac{\partial}{\partial \epsilon^s}\frac{\partial^2 \bar{x}^k}{\partial t^\gamma \partial t^\beta}\right)_0 = \frac{d^2\xi_s^k}{dt^\gamma dt^\beta}.$$

Upon substitution of (3.7) and (3.8) into (3.5), we obtain the fundamental invariance identities given by (3.1) and therefore complete the proof of Theorem 1.

4. THE SINGLE INTEGRAL CASE

We now specialize the results of the preceeding sections to singular integral

$$J = \int_{t_0}^{t_1} L(t,x,\dot{x},\ddot{x})\,dt \tag{4.1}$$

where $x = (x^1,\ldots,x^n)$, $\dot{x} = (\dot{x}^1,\ldots,\dot{x}^n)$, and $\ddot{x} = (\ddot{x}^1,\ldots,\ddot{x}^n)$. This specialization can be realized by taking $m = 1$ in the preceeding discussion. Then, if (4.1) is invariant up to a divergence under the r-parameter family of transformations

$$\bar{t} = \phi(t,\epsilon^1,\ldots,\epsilon^r),$$

$$\tag{4.2}$$

$$\bar{x}^k = \psi^k(t,x,\epsilon^1,\ldots,\epsilon^r),$$

then it necessarily follows that

$$\frac{\partial L}{\partial t}\tau_s + \frac{\partial L}{\partial x^k}\xi_s^k + \frac{\partial L}{\partial \dot{x}^k}\left(\frac{d\xi_s^k}{dt} - \dot{x}^k\frac{d\tau_s}{dt}\right) \tag{4.3}$$

$$+ \frac{\partial L}{\partial \ddot{x}^k}\left(\frac{d^2\xi_s^k}{dt^2} - 2\ddot{x}^k\frac{d\tau_s}{dt} - \dot{x}^k\frac{d^2\tau_s}{dt^2}\right) + L\frac{d\tau_s}{dt} - \frac{dF_s}{dt} = 0$$

for $s = 1,\ldots,r$, where

$$\tau_s(t) = \left(\frac{\partial \phi}{\partial \epsilon^s}\right)_0, \quad \xi_s^k(t,x) = \left(\frac{\partial \psi^k}{\partial \epsilon^s}\right)_0, \tag{4.4}$$

Equations (4.3) are the fundamental invariance identities for single integral problems.

We will now show that the classical Noether identities for single integrals follows easily from the identities (4.3) (for multiple integrals the proof is the same). Before proceeding, we recall that the Euler equations corresponding to the functional (4.1) are given by

$$E_k(L) \equiv \frac{\partial L}{\partial x^k} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}^k} + \frac{d^2}{dt^2}\frac{\partial L}{\partial \ddot{x}^k} = 0, \quad k = 1,\ldots,n. \tag{4.5}$$

Essentially, the Noether identities state that r linear combinations of the Euler expressions $E_k(L)$ are divergences (or exact derivatives). More precisely:

Theorem (Noether): if the fundamental integral (4.1) is divergence-invariant under the r-parameter family of transformations (4.2), then the following r identities hold true:

$$-E_k(L)(\xi_s^k - \dot{x}^k\tau_s) = \frac{d}{dt}\left[L\tau_s + \left(\frac{\partial L}{\partial \dot{x}^k} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{x}^k}\right)(\xi_s^k - \dot{x}^k\tau_s)\right.$$

$$\left. + \frac{\partial L}{\partial \ddot{x}}\frac{d}{dt}(\xi_s^k - \dot{x}^k\tau_s) - F_s\right] \tag{4.6}$$

for $s = 1,\ldots,r$, where the τ_s and ξ_s^k are given explicitly by (4.4). We note that (4.6) agrees with the expressions obtained by Anderson.[6]

In order to prove the Theorem, the Noether identities involve at least fourth-order derivatives of the functions $x^k(t)$ [through the expressions $d^2/dt^2(\partial L/\partial \ddot{x}^k)$] whereas the invariance identities involve at most second order derivatives of the $x^k(t)$. This distinction is quite important and it points the way to obtaining (4.6) from (4.3). First, we note that the following simple identities are valid:

$$\frac{\partial L}{\partial t}\tau_s = \frac{dL}{dt}\tau_s - \dot{x}^k\frac{\partial L}{\partial x^k}\tau_s - \ddot{x}^k\frac{\partial L}{\partial \dot{x}^k}\tau_s - \dddot{x}^k\frac{\partial L}{\partial \ddot{x}^k}\tau_s,$$

$$-\dot{x}^k\frac{\partial L}{\partial \dot{x}^k}\frac{d\tau_s}{dt} = -\frac{d}{dt}\left(\dot{x}^k\frac{\partial L}{\partial \dot{x}^k}\tau_s\right) + \ddot{x}^k\frac{\partial L}{\partial \dot{x}^k}\tau_s + \dot{x}^k\frac{d}{dt}\frac{\partial L}{\partial \dot{x}^k}\tau_s,$$

$$-\dot{x}^k\frac{\partial L}{\partial \ddot{x}^k}\frac{d^2\tau_s}{dt^2} = -\frac{d}{dt}\left(\dot{x}^k\frac{\partial L}{\partial \ddot{x}^k}\frac{d\tau_s}{dt}\right) + \ddot{x}^k\frac{\partial L}{\partial \ddot{x}^k}\frac{d\tau_s}{dt}$$

$$+ \frac{d}{dt}\left(\ddot{x}^k\frac{d}{dt}\frac{\partial L}{\partial \ddot{x}^k}\tau_s\right) - \dddot{x}^k\frac{d}{dt}\frac{\partial L}{\partial \ddot{x}^k}\tau_s - \dot{x}^k\frac{d^2}{dt^2}\frac{\partial L}{\partial \ddot{x}^k}\tau_s,$$

$$\frac{\partial L}{\partial \ddot{x}^k}\frac{d\xi_s^k}{dt} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^k}\xi_s^k\right) - \frac{d}{dt}\frac{\partial L}{\partial \ddot{x}^k}\xi_s^k,$$

$$\frac{\partial L}{\partial \ddot{x}^k}\frac{d^2\xi_s^k}{dt^2} = \frac{d}{dt}\frac{\partial L}{\partial \ddot{x}^k}\left(\frac{d\xi_s^k}{dt}\right) + \frac{d^2}{dt^2}\frac{\partial L}{\partial \ddot{x}^k}\xi_s^k - \frac{d}{dt}\left(\frac{d}{dt}\frac{\partial L}{\partial \ddot{x}^k}\xi_s^k\right).$$

When these expressions are substituted into (4.3) and simplifications are made, the Noether identities (4.6) follow. Therefore, we have a proof of the Noether identities which essentially involves only differentiation and not a discussion or development of the total variation of the fundamental integral.

As an application of (4.3), we consider the case in a so-called parameter invariant integral of the type (4.1), in which it is assumed that it is invariant under arbitrary transformations

$$\bar{t} = \phi(t,\epsilon)$$

which leave the x^k unaffected. When this situation occurs, the quantities ξ_s^k vanish identically, and (4.3) reduces to

$$\frac{\partial L}{\partial t}\tau_s - \dot{x}^k\frac{\partial L}{\partial \dot{x}^k}\frac{d\tau_s}{dt} - 2\ddot{x}^k\frac{\partial L}{\partial \ddot{x}^k}\frac{d\tau_s}{dt^2} - \dot{x}^k\frac{\partial L}{\partial \ddot{x}^k}\frac{d^2\tau_s}{dt^2} + L\frac{d\tau_s}{dt} = 0$$

or

$$\frac{\partial L}{\partial t}\tau_s + \left(L - \dot{x}^k\frac{\partial L}{\partial \dot{x}^k} - 2\ddot{x}^k\frac{\partial L}{\partial \ddot{x}^k}\right)\frac{d\tau_s}{dt} - \dot{x}^k\frac{\partial L}{\partial \ddot{x}^k}\frac{d^2\tau_s}{dt^2} = 0.$$

Due to the arbitrary nature of the transformation (5.4), we obtain the set of conditions

$$\frac{\partial L}{\partial t} = 0, \quad L - \dot{x}^k\frac{\partial L}{\partial \dot{x}^k} - 2\ddot{x}^k\frac{\partial L}{\partial \ddot{x}^k} = 0, \quad \dot{x}^k\frac{\partial L}{\partial \ddot{x}^k} = 0.$$

These are conditions that the Lagrangian L and its derivatives must satisfy if the variational integral is to be parameter invariant. These are the classical *Zermelo conditions* for second-order problems, now derived in a new way from the fundamental invariance identities. See Rund[7] for an *ab initio* derivation of these conditions.

For future reference, we further note that if $x(t) = (x^1(t),\ldots,x^n(t))$ is an extremum, i.e., it satisfies the governing equations (4.5), then the invariance of (4.1) under the r-parameter family of transformations (4.2) leads to the r equations

$$L\tau_s + \left(\frac{\partial L}{\partial \dot{x}^k} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{x}^k}\right)(\xi_s^k - \dot{x}^k\tau_s)$$
$$+ \frac{\partial L}{\partial \ddot{x}^k}\frac{d}{dt}(\xi_s^k - \dot{x}^k\tau_s) - F_s = \text{const} \qquad (4.7)$$

which are first integrals of the governing differential equations (4.5). In different words, equations (4.7) represent conservation laws since the left-hand-side is constant whenever $x(t)$ is an extremal. The latter fact is generally what is taken to be Noether's theorem.

From a practical point of view, the fundamental invariance identities (4.3) can be used to find first integrals of a system of fourth-order ordinary differential equations. Simply, the procedure is to determine, if possible, the corresponding variational problem for the given system, solve the invariance identities (3.1) for τ_s and ξ_s^k to determine an infinitesimal group of trans-

formations under which the variational problem is invariant, and finally apply Noether's theorem to determine first integrals. We remark that for certain fourth-order problems a suitable Lagrangian may not be able to be found.

To be more precise, if we rewrite the identities (4.3) by expanding the total derivatives of τ_s and ξ_s^k, then for certain Lagrangians the identities become polynomials in the $\dot{x}^k$ and the products $\dot{x}^k\dot{x}^l$. Since these directional arguments are arbitrary in (4.3), we may equate to zero the coefficients in the resulting polynomial and consequently the identities (4.3) are transformed into a system of second order partial differential equations in which the generators τ_s and ξ_s^k are regarded as the unknowns. This resulting system of partial differential equations can be regarded as a general set of Killing equations (see Logan[5] for a complete discussion of this matter for first-order Lagrangians.) Hence, in theory, the above method provides a group-theoretic method for determining first integrals of the governing Euler equation.

5. CHARACTERIZATION OF LAGRANGIANS: CONFORMAL IDENTITIES

Having shown that the fundamental invariance identities give rise to both a simple proof of Noether's theorem and to a method for determining first integrals or conservation laws, we now discuss another important interpretation of these fundamental identities, namely, how they can be used to characterize classes of Lagrangians which possess given invariance properties under a known group of transformations. In particular, Eqs. (3.1) [or Eqs. (4.3) in the single integral case] can be regarded as a system of r first-order quasilinear partial differential equations in the unknown Lagrange function L; the generators τ_s^α and ξ_s^k are defined by the invariance transformation according to (2.5) [or (4.4)] and the F_s^α are the divergence terms which serve to determine the type of invariance ($F_s^\alpha = 0$ for absolute invariance). For first-order Lagrangians, this problem has been discussed in detail in Logan[4]. For example, there it was shown that if the fundamental integral (defined on four-dimensional space—time) is absolutely invariant under the fifteen-parameter special conformal group, where the Lagrangian depends on a scalar function or on the components of a covariant vector field and first derivatives of those components, then it is possible to completely characterize those Lagrange functions possessing the given invariance criteria.

In this article we shall derive some conformal identities for second-order problems in which the Lagrange function depends upon a scalar field and derivatives up to the second order. The fundamental integral takes the form

$$J = \int_D L(t^1,\ldots,t^4, x(t^1,\ldots,t^4); \dot{x}_1(t^1,\ldots,t^4),$$
$$\ldots, \dot{x}_4(t^1,\ldots,t^4);\ldots,\ddot{x}_{\alpha\beta}(t^1,\ldots,t^4),\cdots) dt^1\cdots dt^4$$
$$(5.1)$$

where D is a cylinder in space—time. Following the usual convention, we adopt the spacetime coordinates

1377   J. Math. Phys., Vol. 16, No. 7, July 1975

J.D. Logan and J.S. Blakeslee   1377

$t^1 = ict$, t^2, t^3, t^4 of special relativity, where $i^2 = -1$ and c is the velocity of light *in vacuo*. Therefore, the Minkowski metric $g_{\alpha\beta}$ is given by $g_{\alpha\beta} = 0$ if $\alpha \neq \beta$, and $g_{\alpha\alpha} = 1$, $\alpha = 1, \ldots, 4$. In Minkowski space, the special conformal group can be written explicitly as follows:

(i) translations (four parameters)

$$\bar{t}^\alpha = t^\alpha + \epsilon^\alpha \tag{5.2}$$

(ii) Rotations (six parameters)

$$\bar{t}^\alpha = t^\alpha + \omega^{\alpha\beta}t^\beta, \quad \omega^{\alpha\beta} = -\omega^{\beta\alpha} \tag{5.3}$$

(iii) Dilation (one parameter)

$$\bar{t}^\alpha = t^\alpha + \gamma t^\alpha \tag{5.4}$$

(iv) Inversions (four parameters)

$$\bar{t}^\alpha = t^\alpha + (2t^\alpha t_\lambda - t^\nu t_\nu \delta_\lambda^\alpha)\eta^\lambda. \tag{5.5}$$

In (5.2) through (5.5), the parameters are ϵ^α, $\omega^{\alpha\beta}$, γ, and η^λ. In (5.5) δ_β^α is the Kronecker delta. Above and in the sequel lower case Greek letters will range over $1, \ldots, 4$; we note that the index s of preceeding sections has now been replaced by a lower case Greek index.

The scalar field is determined by the single field function $x(t) = x(t^1, \ldots, t^4)$ which is invariant under (5.2)–(5.5), i.e., $\bar{x}(\bar{t}) = x(t)$. Consequently, the generators ξ_λ are given by

$$\xi_\lambda = 0$$

in the case of each separate transformation. Therefore, it follows from Theorem 1 that if the fundamental integral (5.1) is to be absolutely invariant under the conformal transformations (5.2)–(5.5), then

$$\frac{\partial L}{\partial t^\alpha}\tau_s^\alpha - \frac{\partial L}{\partial \dot{x}_\alpha}\dot{x}_\beta \frac{d\tau_s^\beta}{dt^\alpha}$$

$$+ \frac{\partial L}{\partial \ddot{x}_{\alpha\beta}}\left(-\ddot{x}_{\gamma\beta}\frac{d\tau_s^\gamma}{dt^\alpha} - \ddot{x}_{\alpha\gamma}\frac{d\tau_s^\gamma}{dt^\beta} - \dot{x}_\gamma \frac{d^2\tau_s^\gamma}{dt^\alpha dt^\beta}\right) \tag{5.6}$$

$$+ L\frac{d\tau_s^\alpha}{dt^\alpha} = 0$$

where the τ_s^α are the generators of (5.2)–(5.5).

In the case of translations, $\tau_s^k = \delta_s^k$. Substitution into (5.6) yields the four identities

$$L_{t^\alpha} = 0, \quad \alpha = 1, \ldots, 4, \tag{5.7}$$

which state that the Lagrangian cannot depend explicitly upon the spacetime coordinates $t^1, \ldots, t^4$.

Under the dilation, $\tau^\alpha = t^\alpha$ and so

$$\frac{\partial\tau^\alpha}{\partial t^\beta} = \delta_\beta^\alpha, \quad \frac{\partial^2\tau^\alpha}{\partial t^\beta \partial t^\gamma} = 0$$

Substitution of these quantities into (5.6) yields, after simplification,

$$\frac{\partial L}{\partial \dot{x}_\alpha}\dot{x}_\alpha + 2\frac{\partial L}{\partial \ddot{x}_{\alpha\beta}}\ddot{x}_{\alpha\beta} = 4L. \tag{5.8}$$

Under rotations, $\tau_{(\mu\nu)}^\alpha = (\delta_\mu^\alpha t^\nu - \delta_\nu^\alpha t^\mu)$ for $(\mu, \nu) \in S = \{(1,2), (1,3), (1,4), (2,3), (2,4), (3,4)\}$. Consequently,

$$\frac{d\tau_{(\mu\nu)}^\alpha}{dt^\beta} = (\delta_\mu^\alpha \delta_\beta^\nu - \delta_\nu^\alpha \delta_\beta^\mu)$$

and

$$\frac{d^2\tau_{(\mu\nu)}^\alpha}{dt^\beta dt^\gamma} = 0, \quad \frac{d\tau_{(\mu\nu)}^\alpha}{dt^\alpha} = 0.$$

In this case, the invariance identities (5.6) become

$$M_{\mu\nu} \equiv -\dot{x}_\nu L_{\dot{x}_\mu} + \dot{x}_\mu L_{\dot{x}_\nu} - \ddot{x}_{\mu\beta} L_{\ddot{x}_{\nu\beta}} + \ddot{x}_{\nu\beta}L_{\ddot{x}_{\mu\beta}} - \ddot{x}_{\beta\mu}L_{\ddot{x}_{\beta\nu}}$$

$$+ \ddot{x}_{\beta\nu}L_{\ddot{x}_{\beta\mu}} = 0 \tag{5.9}$$

where $(\mu\nu) \in S$.

Under the inversions (5.5), we have $t_\lambda^\alpha = 2t^\alpha t_\lambda - t^\nu t_\nu \delta_\lambda^\alpha$ and so the following relations hold:

$$\frac{d\tau_\lambda^\alpha}{dt^\beta} = 2(t^\alpha \delta_\beta^\lambda + t^\lambda \delta_\beta^\alpha - t^\beta \delta_\lambda^\alpha), \quad \frac{d\tau_\lambda^\alpha}{dt^\alpha} = 8t^\lambda,$$

$$\frac{d^2\tau_\lambda^\alpha}{dt^\lambda dt^\beta} = 2(\delta_\beta^\lambda \delta_\gamma^\alpha + \delta_\beta^\alpha \delta_\gamma^\lambda - \delta_\lambda^\alpha \delta_\gamma^\beta).$$

After substitution of these quantities into (5.6) and considerable simplification and rearrangement of indices along with use of (5.7), we obtain the four identities

$$(4L - L_{\dot{x}_\alpha \dot{x}_\alpha} - 2L_{\ddot{x}_{\alpha\beta}}\ddot{x}_{\alpha\beta})t^\lambda + M_{\alpha\lambda}t^\alpha$$

$$+ L_{\dot{x}_\alpha\alpha}\dot{x}_\lambda - L_{\ddot{x}_{\alpha\lambda}}\dot{x}_\alpha - L_{\ddot{x}_{\lambda\beta}}\dot{x}_\beta = 0.$$

In conjunction with (5.8) and (5.9), these identities reduce to

$$L_{\ddot{x}_{\alpha\alpha}}\dot{x}_\lambda = L_{\ddot{x}_{\alpha\lambda}}\dot{x}_\alpha + L_{\ddot{x}_{\lambda\beta}}\dot{x}_\beta. \tag{5.10}$$

We summarize our results in the following theorem:

Theorem 3: Let J be given by (5.1) where $x(t)$ is a scalar function. Then, a necessary condition for J to be absolutely invariant under the special conformal group (5.2)–(5.5) is that the Lagrangian L satisfy the following fifteen identites:

(i) $L_{t^\alpha} = 0$, $\alpha = 1, \ldots, 4$,

(ii) $\dot{x}_\alpha L_{\dot{x}_\alpha} + 2\ddot{x}_{\alpha\beta}L_{\ddot{x}_{\alpha\beta}} = 4L$,

(iii) $M_{\mu\nu} = 0$, $(\mu, \nu) \in S$,

(iv) $L_{\ddot{x}_{\alpha\alpha}}\dot{x}_\lambda = L_{\ddot{x}_{\alpha\lambda}}\dot{x}_\alpha + L_{\ddot{x}_{\lambda\beta}}\dot{x}_\beta$, $\lambda = 1, \ldots, 4$,

where $M_{\mu\nu}$ is given by (5.9).

Conditions (i)—(iv) of Theorem 3 are conformal identities that may be used as a convenient test for conformal invariance of a given Lagrange function which depends on a scalar field. In the special case that L does not depend upon first derivatives of the field, but only on the second, condition (ii) becomes

$$\ddot{x}_{\alpha\beta} L_{\ddot{x}_{\alpha\beta}} = 2L.$$ (5.11)

By Euler's theorem on homogeneous functions, Eq.

(5.11) implies that L must be homogeneous of degree 2 in the second derivatives.

[1]E. Noether, Nachrichten Ges. der Wissenschaften Gottingen, 1918 Math-Phys. Kl., 235 (1918).
[2]H. Rund, Utilitas Mathematica 2, 205 (1972).
[3]J. D. Logan, J. Math. Anal. Appl. 42, 191 (1973).
[4]J. D. Logan, J. Math. Anal. Appl. 48, 618 (1974).
[5]J. D. Logan, Utilitas Mathematica (to be published).
[6]D. Anderson, J. Phys. A 6, 299 (1973).
[7]H. Rund, Ann. Math. Pura Appl. 55, 77 (1961).

Simplified calculations for radiation reaction forces

R. Tabensky and D. Villarroel

Departamento de Física, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Santiago, Chile
(Received 20 December 1974)

The Lorentz–Dirac equation of motion for the electron is derived by a new method which makes tedious power series expansions unnecessary.

1. INTRODUCTION

The Lorentz–Dirac equation of motion[1] for the electron has been treated extensively in the literature.[2] This diversity is mainly due to the fact that Maxwell equations do not determine the radiation reaction forces without ambiguity, essentially because of the divergent electron self-energy. However, the general view seems to be that Lorentz–Dirac equation describes correctly the electron dynamics. The standard derivations of these equation are rather long and tedious in spite of its simplicity.

We intend to show that radiation reaction calculations can be done very simply if appropriate geometrical objects are used in Minkowski space. A new cutoff prescription in dealing with the electron infinite self-energy based on light cones is introduced. We follow the approach of Rohrlich[3] and Teiltelboim,[4] where the total four-momentum P_μ of the electromagnetic field is the main object under study. The splitting of P_μ into its bound and radiation parts will emerge naturally from our calculation. The Lorentz–Dirac equation of motion is obtained by equating the time derivative of P_μ to the driving external force.

We now describe the way the calculation is done leaving the details for the next section: The energy–momentum tensor constructed from Lienard–Wiechert retarded potentials is[3,5]

$$T_{\mu\nu} = (e^2/4\pi)[- \kappa^{-2}((k\dot v)^2 + \dot v^2)k_\mu k_\nu + \kappa^{-3}(2(k\dot v)k_\mu k_\nu$$
$$- (k\dot v)(k_\mu v_\nu + k_\nu v_\mu)$$
$$+ k_\mu \dot v_\nu + k_\nu \dot v_\mu) + \kappa^{-4}(- k_\mu k_\nu + k_\mu v_\nu + k_\nu v_\mu - \tfrac{1}{2}\eta_{\mu\nu})].$$

$$(1.1)$$

The following notation has been used: the x^μ are the Minkowski coordinates, $x^\mu = z^\mu(\tau)$ is the electron world line (EWL) parametrized by its proper time τ, $v^\mu = dz^\mu/d\tau = \dot z^\mu$, $\kappa = v_\mu(x^\mu - z^\mu(\tau))$, where τ is the retarded proper time of event x^μ, and $k^\mu = \kappa^{-1}(x^\mu - z^\mu(\tau))$ a light vector. The signature of the Minkowski tensor is -2. The velocity of light is chosen to be 1.

Following the standard procedure of field theory, we define the total four-momentum of the electromagnetic field by

$$P^\mu = \int_\sigma T^{\mu\nu} d\sigma_\nu.$$

$$(1.2)$$

In Teitelboim's work[4] σ is a spacelike hyperplane that cuts the EWL orthogonally. We shall assume σ to be an arbitrary spacelike hyperplane. Such a restriction is unimportant, since the value of P^μ is independent of the detailed shape of σ outside some finite neighborhood of

the EWL, as Gauss' integral theorem applied to $T^{\mu\nu}{}_{,\nu} = 0$ shows. As it is well known, P^μ as defined by Eq. (1.2) is divergent; therefore, some cutoff prescription is necessary (a detailed study of the dependence of P^μ on the cutoff is given in Ref. 6). We shall use the following one: Let σ cut the EWL at proper time $\bar\tau$, pick any $\tau < \bar\tau$, and draw the future light cone C emerging from $z^\mu(\tau)$, as shown in Fig. 1. The portion of σ within C is a three-sphere as seen by an observer with four-velocity u^μ orthogonal to σ. We denote this sphere by Sp. Integral (1.2) is performed on the domain σ-Sp only, and thus it is finite. The three-dimensional picture of Sp is that of a sphere of light emited continuously from τ to $\bar\tau$. Within Sp, when $\tau \to \bar\tau$, it is assumed that the contribution to the total four-momentum is $m_0 v^\mu$, where m_0 is the bare mass, which, when added to the electromagnetic mass, gives the observed electron mass m. We therefore write

$$P^\mu = m_0 v^\mu + \int_{\sigma\text{-Sp}} T^{\mu\nu} d\sigma_\nu.$$

$$(1.3)$$

This last integral splits naturally into two pieces, P^μ_B and P^μ_R, which require separate computation. In order
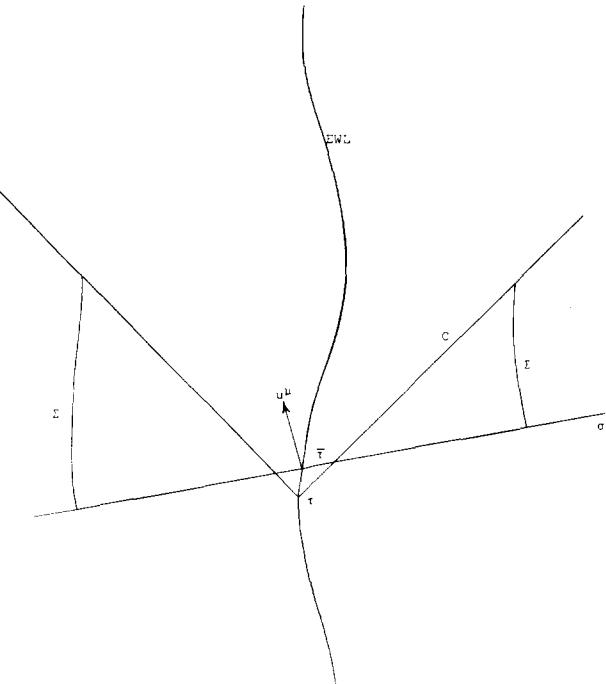

FIG. 1. The world diagram describing various hypersurfaces used in evaluating the electron four-momentum: σ is a spacelike hyperplane cutting the electron world line (EWL) at proper time $\bar\tau$; Σ is a time like tube surrounding the EWL; C is the future light cone with vertex at $z^\mu(\tau)$.

to exhibit this splitting we surround the EWL by a time-like tube Σ (see Fig. 1) which eventually tends to spatial infinity. In applying Gauss' integral theorem on $T^{\mu\nu}{}_{,\nu}$ in the region bounded by Σ, σ and C we get

$$P^\mu = m_0 v^\mu(\bar\tau) + P^\mu_B + P^\mu_R, \tag{1.4}$$

where

$$P^\mu_B = \int_C T^{\mu\nu} dC_\nu \tag{1.5}$$

and

$$P^\mu_R = - \int_\Sigma T^{\mu\nu} d\Sigma_\nu. \tag{1.6}$$

The integral P^μ_B over the light cone C is performed up to its intersection with σ. The second integral P^μ_R is performed over any timelike tube Σ that tends to spatial infinity, from the infinite past up to its intersection with C.

P^μ_B is identified with the four-momentum bound to the electron since it depends on the kinematical variables of the EWL at $\tau = \bar\tau$ only,[4] in the limit $\tau \to \bar\tau$. On the other hand, P^μ_R depends on the entire electron history up to τ, and is calculated out of values of $T^{\mu\nu}$ at spatial infinity. It is identified with the total radiated four-momentum up to proper time τ.

We apply this method to the electron angular momentum obtaining expressions for the bound and radiated parts.[7] Similar calculations are done for the scalar field.

2. VECTOR FIELD

In calculating the bound four-momentum we need the light-cone volume element given by[8]

$$dC^\mu = \kappa^2 \, d\kappa \, d\Omega k^\mu, \tag{2.1}$$

where $d\Omega$ is the solid angle seen by an observer at rest with the electron. Equation (1.1) yields the following relation:

$$T^{\mu\nu} k_\nu = (e^2/8\pi)\kappa^{-4} k^\mu, \tag{2.2}$$

and therefore

$$P^\mu_B = \frac{e^2}{8\pi} \int_{\kappa=\kappa_0}^{\infty} d\kappa \, d\Omega \kappa^{-2} k^\mu = \frac{e^2}{8\pi} \int d\Omega \kappa_0^{-1} k^\mu, \tag{2.3}$$

where κ_0 denotes the values of κ at the intersection of C and σ, which is obtained from Fig. 1 by noting that $\{\kappa_0 k^\mu - [z^\mu(\bar\tau) - z^\mu(\tau)]\}u_\mu = 0$, that is $\kappa_0^{-1} = uk/u(\bar z - z)$. P^μ_B is now given by

$$P^\mu_B = \frac{(e^2/8\pi)}{u(\bar z - z)} \int d\Omega (uk) k^\mu. \tag{2.4}$$

This expression is easily integrated to give

$$P^\mu_B = \frac{2}{3} e^2 \frac{(uv)v^\mu - \frac{1}{4} u^\mu}{u[z(\bar\tau) - z(\tau)]}. \tag{2.5}$$

The leading divergent term of P^μ_B when $\tau \to \bar\tau$ is

$$P^\mu_B = \frac{2}{3} e^2 \epsilon^{-1} \{ v^\mu(\bar\tau) - u^\mu/4[uv(\bar\tau)] \}, \tag{2.6}$$

where $\epsilon = \bar\tau - \tau$. This self-energy divergence cannot be incorporated to m_0 in Eq. (1.4) unless $u^\mu = v^\mu(\bar\tau)$, that is, unless σ cuts the EWL orthogonally.[9] In order to renormalize the divergence, we assume $u^\mu = v^\mu(\bar\tau)$. In

this case we get[4]

$$P^\mu_B = \frac{e^2}{2\epsilon} v^\mu(\bar\tau) - \frac{2}{3} e^2 \dot v^\mu(\bar\tau) + O(\epsilon). \tag{2.7}$$

We now consider the radiated four-momentum P^μ_R defined by Eq. (1.6). The exact shape of the tube Σ is not important as can be seen from $T^{\mu\nu}{}_{,\nu} = 0$. We choose a tube used by Bhabha,[10] defined by the equation $\kappa = \text{const}$, its volume element is given by[10]

$$d\Sigma_\mu = \{[1 - \kappa(k\dot v)]k_\mu - v_\mu\} \kappa^2 \, d\tau \, d\Omega. \tag{2.8}$$

From here and Eq. (1.1) we get

$$T_{\mu\nu} d\Sigma^\nu = (e^2/4\pi)[(k\dot v)^2 + \dot v^2]k_\mu \, d\tau \, d\Omega, \tag{2.9}$$

in the limit $\kappa \to \infty$. A simple computation of angular integrals leads to

$$P^\mu_R = -\frac{2}{3} e^2 \int_{-\infty}^\tau \dot v^2 v^\mu \, d\tau, \tag{2.10}$$

which is the well-known Larmor radiation formula. The convergence of this integral requires that the acceleration vanishes at $\tau = -\infty$.

The Lorentz—Dirac equation is obtained by combining Eqs. (1.4), (2.7), and (2.10) to get

$$m\dot v_\mu - \frac{2}{3} e^2 (\ddot v_\mu + \dot v^2 v_\mu) = F^{ex}_\mu, \tag{2.11}$$

where F^{ex}_μ is the external driving force.

The calculation of the bound and radiated angular momentum is equally simple. The total angular-momentum density is as usual given by

$$M^{\lambda\mu\nu} = x^\lambda T^{\mu\nu} - x^\mu T^{\lambda\nu}. \tag{2.12}$$

The angular momentum $M^{\lambda\mu}$ of the electron field is given by

$$M^{\lambda\mu} = \int_\sigma M^{\lambda\mu\nu} d\sigma_\nu. \tag{2.13}$$

where, as for P^μ, σ is the hyperplane that cuts orthogonally the EWL at $z(\bar\tau)$. The angular momentum (2.13) is split in its bound $(M^{\lambda\mu}_B)$ and radiated $(M^{\lambda\mu}_R)$ parts, in exactly the same way as the four-momentum. The following two formulas are needed:

$$M^{\lambda\mu\nu} dC_\nu = z^{[\lambda}k^{\mu]}\kappa^{-2} \, d\kappa \, d\Omega \tag{2.14}$$

and

$$M^{\lambda\mu\nu} d\Sigma_\nu = 2\{-(k\dot v)k^{[\lambda}v^{\mu]} + k^{[\lambda}\dot v^{\mu]} + [(k\dot v)^2 + \dot v^2]k^{[\lambda}z^{\mu]}\}, \tag{2.15}$$

where the limit $\kappa \to \infty$ has been taken in the second equation. The notation $a^{[\mu}b^{\nu]} = \frac{1}{2}(a^\mu b^\nu - a^\nu b^\mu)$ has been used.

From Eq. (2.14) we see that

$$M^{\lambda\mu}_B = z^\lambda P^\mu_B - z^\mu P^\lambda_B, \tag{2.16}$$

which is the result one expects for material particles. From Eq. (2.15), after angular integrations are performed, we get

$$M^{\lambda\mu}_R = -\frac{4}{3} e^2 \int_{-\infty}^\tau \dot v^2 z^{[\lambda}v^{\mu]} \, d\tau + \frac{4}{3} e^2 \int_{-\infty}^\tau v^{[\lambda}\dot v^{\mu]} d\tau. \tag{2.17}$$

This result was obtained by López and Villarroel.[7]

3. SCALAR FIELD

In this section we apply the same technique to the

massless scalar field of a point source. The field equation to be solved is

$$\Box \Phi = 4\pi g \int_{-\infty}^{+\infty} \delta^4(x - z(\tau)) \, d\tau. \tag{3.1}$$

The retarded solution to this equation is known to be $\Phi = g\kappa^{-1}$. The energy-momentum tensor is given by

$$4\pi T_{\mu\nu} = \Phi_{,\mu}\Phi_{,\nu} - \tfrac{1}{2}\eta_{\mu\nu}\Phi_{,\alpha}\Phi_{,}{}^{\alpha}, \tag{3.2}$$

which gives

$$T_{\mu\nu} = \frac{g^2}{4\pi} \left[\kappa^{-2}(k\mathring{v})^2 k_\mu k_\nu + \kappa^{-3}(-2k_\mu k_\nu + v_\mu k_\nu + v_\nu k_\mu - \eta_{\mu\nu}) \right.$$

$$\left. + \kappa^{-4}(k_\mu k_\nu - v_\mu k_\nu - v_\nu k_\mu + v_\mu v_\nu + \tfrac{1}{2}\eta_{\mu\nu}) \right], \tag{3.3}$$

when $\Phi = g\kappa^{-1}$. From this expression we obtain

$$T^{\mu\nu} dC_\nu = (v^\mu - \tfrac{1}{2}k^\mu)\kappa^{-2} \, d\kappa \, d\Omega \tag{3.4}$$

and

$$T^{\mu\nu} d\Sigma_\nu = (k\mathring{v})^2 k^\mu \, d\tau \, d\Omega, \quad \text{as } \kappa \to \infty, \tag{3.5}$$

which, after angular integrations are performed, gives

$$P^\mu_B = (g^2/2\epsilon)v^\mu - (g^2/3)\mathring{v}^\mu + O(\epsilon) \tag{3.6}$$

and

$$P^\mu_R = -\tfrac{1}{3}g^2 \int_{-\infty}^{\tau} \mathring{v}^2 v_\mu \, d\tau. \tag{3.7}$$

The equation of motion that follows from (3.6) and (3.7), after mass renormalization, is

$$m\mathring{v}_\mu - \tfrac{1}{3}g^2(\mathring{\mathring{v}}_\mu + \mathring{v}^2 v_\mu) = F^{ex}_\mu. \tag{3.8}$$

The bound angular momentum is given by the same expression (2.16) as for the electron. The radiated angular momentum that results is obtained from (2.17) by replacing $\tfrac{2}{3}e^2$ by $\tfrac{1}{3}g^2$.

[1]P.A.M. Dirac, Proc. Roy. Soc. London A 167, 148 (1938).
[2]F. Rohrlich, in *The Physicist's Conception of Nature*, edited by T. Mehra (Reidel, Dordrecht, Holland, 1973), pp. 331—69.
[3]F. Rohrlich, *Classical Charged Particles* (Addison-Wesley, Reading, Mass., 1965).
[4]C. Teitelboim, Phys. Rev. D 1, 1572 (1970).
[5]A.O. Barut, *Electrodynamics and Classical Theory of Fields and Particles* (Macmillan, New York, 1964).
[6]R. Tabensky, "Electrodynamics and the electron equation of motion," to be published.
[7]C.A. López and D. Villarroel, "Bound and emitted angular momenta of a classical accelerated point charge," to be published in Phys. Rev. D.
[8]J.L. Synge, Ann. Mat. 84, 33 (1970).
[9]R. Tabensky and D. Villarroel, "Classical electron mass renormalization," to be published.
[10]H.J. Bhabha, Proc. Roy. Soc. London A 172, 384 (1939).

1382   J. Math. Phys., Vol. 16, No. 7, July 1975

R. Tabensky and D. Villarroel   1382

Analysis of the dispersion of low frequency uniaxial waves in heterogeneous periodic elastic media

George N. Balanis*

Applied Theory, Inc., 1010 Westwood Blvd., Los Angeles, California 90024

The dispersion of harmonic uniaxial waves in heterogeneous periodic elastic media is investigated. The frequency dependence of the wave phase velocity is obtained in the form of a power series valid for small frequencies and arbitrary spatial variations of the heterogeneities. The dominant dispersion term is always negative and proportional to the square of the frequency. Near the static limit of zero frequency the dispersion is of the normal type—the group velocity, which is also a quadratic decreasing function of the frequency, decreases faster than the phase velocity.

INTRODUCTION

Recently, the effects of heterogeneities on wave propagation in composite materials have acquired technical interest. In bilaminate materials the dispersion effects have been examined by Sun, Achenbach and Herrmann,[1] Lee and Yang,[2] Balanis[3] and Peck and Gurtman,[4] experimental investigations have been carried out by Lundergan and Drumheller,[5] models have been proposed by Hegemier and Nayfeh[6] and Barker.[7]

In the following sections a general method, similar to one described by Friedrichs,[8] is applied to spatially heterogeneous periodic elastic media to determine the dispersion effects on uniaxial propagation. The analysis is also applicable to other media similar to elastic media such as dielectrics and transmission lines.

STATEMENT OF PROBLEM

Let the time harmonic dependence be $\exp(-i\omega t)$ where ω is the applied frequency (a real number). Let x, $u(x, \omega)$, $\sigma(x, \omega)$ represent the Lagrangian particle position, harmonic particle velocity and harmonic particle compressive stress. Finally, let the reference density $\rho(x)$, inverse constraint modulus $m(x)$, and reference sound speed $c(x)$, $c(x) = 1/\sqrt{\rho(x)m(x)}$, be positive, piecewise-continuous periodic functions of x with period L. It is well known that the longitudinal wave motion is governed by the equations

$$\frac{\partial}{\partial x}\begin{bmatrix} u(x, \omega) \\ \sigma(x, \omega) \end{bmatrix} = i\omega \begin{bmatrix} 0 & m(x) \\ \rho(x) & 0 \end{bmatrix}\begin{bmatrix} u(x, \omega) \\ \sigma(x, \omega) \end{bmatrix}, \tag{1}$$

whose general solution has the Floquet type form

$$\begin{bmatrix} u(x, \omega) \\ \sigma(x, \omega) \end{bmatrix} = \begin{bmatrix} f(x, \omega) \\ g(x, \omega) \end{bmatrix}\exp\{i[\omega x/v_p(\omega)]\}$$

$$+ \begin{bmatrix} h(x, \omega) \\ s(x, \omega) \end{bmatrix}\exp\{-i[\omega x/v_p(w)]\} \tag{2}$$

where f, g, h, s are periodic functions of x with period L. The effects of the density and constraint modulus heterogeneities on the frequency dependence of the phase velocity $v_p(\omega)$ are the main interest of this paper.

ANALYSIS

Equation (1) is linear. Accordingly, the following Friedrich's type representation is possible:

$$\begin{bmatrix} u(x, \omega) \\ \sigma(x, \omega) \end{bmatrix} = B(x, \omega)\begin{bmatrix} u(0, \omega) \\ \sigma(0, \omega) \end{bmatrix} \tag{3}$$

where $B(x, \omega)$ is a square matrix. The differential equation satisfied by this matrix can be obtained by substituting (3) in (1)

$$\frac{\partial B}{\partial x}(x, \omega) = i\omega\begin{bmatrix} 0 & m(x) \\ \rho(x) & 0 \end{bmatrix}B(x, \omega) \tag{4}$$

and

$$B(0, \omega) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{5}$$

The velocity and stress at the left boundary of the first cell of the periodic medium are related to the velocity and stress at the right cell boundary by Eq. (3) evaluated at $x = L$. Since all cells are the same it follows that the fields at any lattice point are connected to the boundary fields at $x = 0$ by the relation

$$\begin{bmatrix} u(nL, \omega) \\ \sigma(nL, \omega) \end{bmatrix} = B^n(L, \omega)\begin{bmatrix} u(0, \omega) \\ \sigma(0, \omega) \end{bmatrix} \tag{6}$$

where $n = 0, 1, 2, 3, \cdots$. The wave phase velocity, which is related to the eigenvalues of the matrix $B(L, \omega)$, enters the analysis when the boundary field vector is decomposed to the eigenvectors of $B(L, \omega)$. If $\exp\{\pm i[\omega/v_p(\omega)]L\}$ are the matrix eigenvalues and $\begin{bmatrix} 1 \\ Z_{\pm}(\omega) \end{bmatrix}$ the corresponding eigenvectors, then, at the frequencies at which $Z_+(\omega) \neq Z_-(\omega)$, we find

$$\begin{bmatrix} u(0, \omega) \\ \sigma(0, \omega) \end{bmatrix} = a(\omega)\begin{bmatrix} 1 \\ Z_+(\omega) \end{bmatrix} + b(\omega)\begin{bmatrix} 1 \\ Z_-(\omega) \end{bmatrix}$$

which, when substituted in (6), gives the lattice waves in a form comparable to (2)

$$\begin{bmatrix} u(nL,\,\omega) \\ \sigma(nL,\,\omega) \end{bmatrix} = a(\omega)\begin{bmatrix} 1 \\ Z_+(\omega) \end{bmatrix}\exp\{inL[\omega/v_p(\omega)]\}$$
$$+\,b(\omega)\begin{bmatrix} 1 \\ Z_-(\omega) \end{bmatrix}\exp\{-inL[\omega/v_p(\omega)]\} \qquad (7)$$

where

$$a(\omega) = \frac{\sigma(0,\,\omega) - Z_-(\omega)u(0,\,\omega)}{Z_+(\omega) - Z_-(\omega)},$$
$$b(\omega) = \frac{Z_+(\omega)u(0,\,\omega) - \sigma(0,\,\omega)}{Z_+(\omega) - Z_-(\omega)}.$$

The theorem below and the discussion which follows it summarize the low frequencies dispersion effects caused by heterogeneities. The proof of the theorem follows from Eqs. (4) and (5) and is given in the Appendix. The elements of $B(L,\,\omega)$ are indicated by $b_{11}(\omega)$, $b_{12}(\omega)$, $b_{21}(\omega)$ and $b_{22}(\omega)$.

THEOREM

The wave phase velocity has the following properties:

(a) It depends on the trace of $B(L,\,\omega)$ according to the dispersion relation

$$\cos\frac{\omega L}{v_p(\omega)} = \frac{b_{11}(\omega) + b_{22}(\omega)}{2}, \qquad (8)$$

where $b_{11}(\omega)$ and $b_{22}(\omega)$ are real even functions of ω.

(b) It is real for all ω in the region $-\bar{c}/\sqrt{3}L < \omega < \bar{c}/\sqrt{3}L$.

(c) It has a low frequency power series expansion which involves only even powers of the frequency

$$v_p(\omega) = \bar{c}(1 - r_2\omega^2 - r_4\omega^4 - \cdots) \qquad (9)$$

where

$$\bar{c} = \left[\left(\frac{1}{L}\int_0^L \rho(x)\,dx\right)^{1/2}\left(\frac{1}{L}\int_0^L m(x)\,dx\right)^{1/2}\right]^{-1} \qquad (10)$$

$$r_2 = \frac{1}{2}\left(\frac{\bar{c}}{L}\right)^2\left[\frac{2}{4!}\left(\frac{L}{\bar{c}}\right)^4\right.$$
$$-\int_0^L m(x)\int_0^x \rho(v)\int_0^y m(z)\int_0^z \rho(w)\,dw\,dz\,dy\,dx$$
$$\left.-\int_0^L \rho(x)\int_0^x m(v)\int_0^y \rho(z)\int_0^z m(w)\,dw\,dz\,dy\,dx\right] \qquad (11)$$

and

$$\left|r_4\omega^4 + r_6\omega^6 + \cdots\right| = O\left[\left(\frac{\omega L}{\bar{c}}\right)^4\right] \text{ as } \left|\frac{\omega L}{\bar{c}}\right| \to 0.$$

(d) The coefficient r_2 is always nonnegative; positive if $\rho(x)c(x)$ depends on position and zero if $\rho(x)c(x)$ is constant.

THEOREM DISCUSSION

The speed $\bar{c}$, which enters in the above theorem, has physical meaning. Let $\epsilon(x,\,\omega)$ be the strain, $\epsilon(x,\,\omega) = (\partial\zeta/\partial x)(x,\,\omega)$ where $\zeta(x,\,\omega)$ is the harmonic particle

displacement. If the medium is in static equilibrium, which is the case when $\omega = 0$, the stress is independent of position. Let $\zeta_{eq}(x)$, σ_{eq} be the equilibrium particle displacement and equilibrium stress. By integrating Hooke's Law

$$\sigma(x,\,\omega) = -\epsilon(x,\,\omega)/m(x)$$

over a unit cell, we find

$$\sigma_{eq} = -\left(\frac{1}{L}\int_0^L m(x)\,dx\right)^{-1}\epsilon_{eq} \qquad (12)$$

where ϵ_{eq} is the macroscopic strain for the cell, i.e.,

$$\epsilon_{eq} = \frac{\zeta_{eq}(L) - \zeta_{eq}(0)}{L}.$$

When the average density $\bar{\rho}$, $\bar{\rho} = (1/L)\int_0^L \rho(x)\,dx$, is introduced in Eq. (12), we obtain

$$\sigma_{eq} = -\bar{\rho}\,\bar{c}^2\,\epsilon_{eq}$$

which implies that $\bar{c}$ may be interpreted as the macroscopic static sound speed of each unit cell.

Equation (8) represents the dispersion relation for the phase velocity of any periodic medium. Many analyses for particular periodic media have shown that their dispersion relations exhibit the phenomenon of pass and stop bands. Band structure is also evident in (8). The frequency ranges for which

$$\left|\frac{b_{11}(\omega) + b_{22}(\omega)}{2}\right| \leq 1$$

are the pass bands where the phase velocity is real. The stop bands, where the phase velocity is complex, occur when

$$\left|\frac{b_{11}(\omega) + b_{22}(\omega)}{2}\right| > 1.$$

The theorem asserts that the low frequency range

$$|\omega| < \frac{1}{\sqrt{3}}\frac{\bar{c}}{L}$$

is always within a pass band. It is interesting to note that as the thickness of the cell goes to zero $\bar{c}/L \to \infty$ and the pass band increases in size indefinitely. However, this result is not surprising since as $L \to 0$ the cell tends to look more and more like a homogenous cell which, as is well known, has no stop bands.

In contrast to the lossy dispersion encountered in propagation in viscoelastic materials, the dispersion in heterogeneous elastic media is effected without loss of energy. The following argument illuminates the reason the phase velocity power series expansion involves only even powers of the frequency. The transient velocity lattice waves that propagate in the $+x$ direction are

$$U(nL,\,t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} a(\omega)\exp\{inL[\omega/v_p(\omega)]\}\exp(-i\omega t)\,d\omega.$$

In order that these waves be real functions

$$v_p^*(\omega) = v_p(-\omega) \qquad (13)$$

where * means complex conjugate. An assumed phase

velocity expansion which contains both even and odd power terms,

$$v_p(\omega) = \overline{c}(1 - v_1\omega - v_2\omega^2 - v_3\omega^3 - \cdots),$$

must, in accord with (13), have purely imaginary odd power coefficients, in contrast to the even power coefficients which should be real. However, for lossless propagation the power series expansion of the phase $nL\omega/v_p(\omega)$ must be real. Thus, the purely imaginary coefficients $v_1, v_3, v_5, \cdots$, which are indicative of dissipation, are identically equal to zero.

The theorem asserts that in the low frequency regime the phase velocity is always quadratic and attains its maximum value $\overline{c}$ at the static limit of zero frequency. The low frequency group velocity $v_g(\omega)$,

$$v_g(\omega) = [(\partial/\partial\omega)\omega/v_p(\omega)]^{-1}, \quad v_g(\omega) = \overline{c}\left[1 - 3v_2\omega^2 + O\left(\frac{\omega L}{\overline{c}}\right)^4\right]$$

has similar properties. Thus, in the low frequency regime, spatial heterogeneities in the medium impedance affect the wave motion by letting the lower frequencies travel faster. The motion of the higher frequencies, which see the detail of the heterogeneous material structure, is impeded by more reflections. Accordingly, in low frequency propagation of transient pulses the higher frequency portion of the pulse should be behind the main low frequency part which travels with the higher speeds. Solutions of pulse propagation demonstrate this effect of heterogeneities in elastic media.[3] In contrast to other types of dispersion, elastic heterogeneities generate the normal type of dispersion in the low frequency regime—the wave phase velocity is always larger than the group velocity

$$v_p(\omega) - v_g(\omega) = \overline{c}\left[2v_2\omega^2 + O\left(\frac{\omega L}{\overline{c}}\right)^4\right].$$

ACKNOWLEDGMENTS

APPENDIX: THEOREM PROOF

(a) The eigenvalues of the matrix $B(L, \omega)$ depend on its determinant and its trace. The eigenvalue product is equal to the determinant; the sum is equal to the trace. It is easily deduced from (4) that the determinant is independent of position,

$$\frac{\partial}{\partial x}\det[B(x, \omega)] = 0.$$

From (5) follows that

$$\det[B(L, \omega)] = 1$$

and thus each eigenvalue is the inverse of the other. The trace of $B(L, \omega)$ is the sum at $x = L$ of the matrix elements $b_{11}(x, \omega)$ and $b_{22}(x, \omega)$. Their properties can be found from matrix Eqs. (4) and (5) whose explicit form

is

$$\frac{\partial b_{11}}{\partial x}(x, \omega) = i\omega m(x)b_{21}(x, \omega), \quad (A1)$$

$$\frac{\partial b_{21}}{\partial x}(x, \omega) = i\omega\rho(x)b_{11}(x, \omega), \quad (A2)$$

$$b_{11}(0, \omega) = 1, \quad (A3)$$

$$b_{21}(0, \omega) = 0 \quad (A4)$$

and

$$\frac{\partial}{\partial x}b_{12}(x, \omega) = i\omega m(x)b_{22}(x, \omega), \quad (A5)$$

$$\frac{\partial b_{22}}{\partial x}(x, \omega) = i\omega\rho(x)b_{12}(x, \omega), \quad (A6)$$

$$b_{22}(0, \omega) = 1, \quad (A7)$$

$$b_{12}(0, \omega) = 0. \quad (A8)$$

The uniqueness of the solution of the above equations indicates that

$$b_{11}(x, \omega) = b_{11}^*(x, \omega) = b_{11}(x, -\omega),$$

$$b_{22}(x, \omega) = b_{22}^*(x, \omega) = b_{22}(x, -\omega),$$

$$b_{12}(x, \omega) = -b_{12}^*(x, \omega) = -b_{12}(x, -\omega),$$

$$b_{21}(x, \omega) = -b_{21}^*(x, \omega) = -b_{21}(x, -\omega).$$

Therefore, the elements b_{11} and b_{22} are real even functions of ω. Their sum at $x = L$, which is connected to the eigenvalues by the relation

$$\exp\{i[\omega L/v_p(\omega)]\} + \exp\{-i[\omega L/v_p(\omega)]\} = b_{11}(\omega) + b_{22}(\omega),$$

results in (8).

(b) The low frequency behavior of the elements of $B(L, \omega)$ determines the low frequency behavior of the phase velocity. When ω is small an iterative procedure applied to Eqs. (A1)—(A8) shows that at $x = L$ the elements of $B(L, \omega)$ possess the following power series expansions:

$$b_{11}(\omega) = 1 + (i\omega)^2\overline{A}_2 + (i\omega)^4\overline{A}_4 + \cdots,$$

$$b_{12}(\omega) = i\omega\overline{B}_1 + (i\omega)^3\overline{B}_3 + (i\omega)^5\overline{B}_5 + \cdots,$$

$$b_{21}(\omega) = i\omega\overline{C}_1 + (i\omega)^3\overline{C}_3 + (i\omega)^5\overline{C}_5 + \cdots,$$

$$b_{22}(\omega) = 1 + (i\omega)^2\overline{D}_2 + (i\omega)^4\overline{D}_4 + \cdots.$$

where

$$\overline{A}_2 = \int_0^L m(y_1)\int_0^{y_1}\rho(y_2)\,dy_2\,dy_1$$

.
.
.

$$\overline{A}_k = \int_0^L m(y_1)\int_0^{y_1}\rho(y_2)\cdots\int_0^{y_{k-2}}m(y_{k-1})$$

$$\times \int_0^{y_{k-1}}\rho(y_k)\,dy_k\,dy_{k-1}\cdots dy_2\,dy_1$$

.
.
.

$$\overline{B}_1 = \int_0^L m(y_1)\,dy_1$$

.
.
.

George N. Balanis   1385

$$\bar{B}_l = \int_0^L m(y_1) \int_0^{y_1} \rho(y_2) \int_0^{y_2} m(y_3)$$

$$\cdots \int_0^{y_{l-2}} \rho(y_{l-1}) \int_0^{y_{l-1}} m(y_l) \, dy_l \, dy_{l-1} \cdots dy_3 \, dy_2 \, dy_1$$

.
.
.

$$\bar{C}_1 = \int_0^L \rho(y_1) \, dy_1$$

.
.
.

$$\bar{C}_l = \int_0^L \rho(y_1) \int_0^{y_1} m(y_2) \int_0^{y_2} \rho(y_3) \cdots \int_0^{y_{l-2}} m(y_{l-1})$$

$$\times \int_0^{y_{l-1}} \rho(y_l) \, dy_l \, dy_{l-1} \cdots dy_3 \, dy_2 \, dy_1$$

.
.
.

$$\bar{D}_2 = \int_0^L \rho(y_1) \int_0^{y_1} m(y_2) \, dy_2 \, dy_1$$

。
。
。

$$\bar{D}_k = \int_0^L \rho(y_1) \int_0^{y_1} m(y_2) \cdots \int_0^{y_{k-2}} \rho(y_{k-1})$$

$$\times \int_0^{y_{k-1}} m(y_k) \, dy_k \, dy_{k-1} \cdots dy_2 \, dy_1$$

.
.
.

and $l = 1, 3, 5, \cdots$ and $k = 2, 4, 6, \cdots$. When the cell width is substituted as the upper limit of all integrations in the equations above we find

$$\bar{A}_k \leq \left(\frac{L}{\bar{c}}\right)^k, \tag{A9}$$

$$\bar{D}_k \leq \left(\frac{L}{\bar{c}}\right)^k. \tag{A10}$$

Now, the absolute value of the right-hand side of (8) is given by

$$\left|\frac{b_{11}(\omega) + b_{22}(\omega)}{2}\right| = \left|1 - \frac{1}{2}\left(\frac{\omega L}{\bar{c}}\right)^2 + \frac{\omega^4}{2}(\bar{A}_4 + \bar{D}_4)\right.$$

$$\left. - \frac{\omega^6}{2}(\bar{A}_6 + \bar{D}_6) + \cdots\right| \tag{A11}$$

and is bounded from above by

$$\left|1 - \frac{1}{2}\frac{\omega L}{\bar{c}}^2\right| + \frac{1}{2}[\omega^4(\bar{A}_4 + \bar{D}_4) + \omega^6(\bar{A}_6 + \bar{D}_6) + \cdots]$$

which, in turn, is similarly bounded by

$$1 - \frac{1}{2}\left(\frac{\omega L}{\bar{c}}\right)^2 + \left(\frac{\omega L}{\bar{c}}\right)^4 \frac{1}{1 - (\omega L/\bar{c})^2}$$

when ω is in the low frequency range $|\omega L/\bar{c}| < 1/\sqrt{3}$. However, it is easy to show that in this frequency range the last upper bound is less than 1. Thus,

$$\left|\frac{b_{11}(\omega) + b_{22}(\omega)}{2}\right| < 1$$

and the phase velocity is real when $|\omega L/\bar{c}| < 1/\sqrt{3}$.

(c) Let $\xi(\omega)$ be defined as

$$\xi(\omega) = \left[1 - \left(\frac{b_{11}(w) + b_{22}(\omega)}{2}\right)^2\right]^{1/2}. \tag{A12}$$

The dependence of this parameter to small values of the frequency can be obtained from (A11). When this equation is squared we find

$$\left(\frac{b_{11} + b_{22}}{2}\right)^2 = 1 - \epsilon^2 + \epsilon^4\left[\frac{1}{4} + \left(\frac{\bar{c}}{L}\right)^4(\bar{A}_4 + \bar{D}_4)\right] + h_1 \tag{A13}$$

where the small parameter ϵ is given by

$$\epsilon = \omega L/\bar{c}.$$

The remainder h_1, which is an even function of ω, is of order higher than ϵ^4

$$h_1 = O(\epsilon^6).$$

From (A12) and (A13) follows the low frequency expansion for ξ

$$\xi = \epsilon\left\{1 - \frac{1}{2}\epsilon^2\left[\frac{1}{4} + \left(\frac{\bar{c}}{L}\right)^4(\bar{A}_4 + \bar{D}_4)\right] + O(\epsilon^4)\right\}. \tag{A14}$$

Now, the phase velocity is related to ξ by the following equation obtained from (8) and (A12)

$$v_p = \bar{c}[\epsilon/\sin^{-1}(\xi)]. \tag{A15}$$

The phase velocity expansion, whose dominant terms are noted in (9)—(11), results from the power series expansion of the right-hand side of (A15) for small values of ϵ. The useful inverse sine series is given by

$$\sin^{-1}(\xi) = \xi + \frac{1}{6}\xi^3 + \frac{1}{2}\frac{3}{4}\frac{1}{5}\xi^5 + \cdots.$$

(d) For the purpose of showing that v_2 is nonnegative we introduce the transformation $x' = x'(x)$,

$$x' = \int_0^x m(y) \, dy,$$

which is continuous and one-to-one. Let d and q stand for

$$d = \int_0^L m(y) \, dy, \quad q = 2\left(\frac{L}{\bar{c}}\right)^2 v_2.$$

The ratio of the density to the constraint modulus is noted by $n(x')$ and is equal to the square of the medium impedance

$$n(x') = \frac{\rho(x(x'))}{m(x(x'))}.$$

Then, the quantity q, which has the same sign as v_2, is related to $n(x')$ as follows:

$$q = \frac{2}{4!} d^2\left(\int_0^d n(x') \, dx'\right)^2 - \int_0^d \int_0^{x'} n(v')$$

$$\times \int_0^{y'} \int_0^{z'} n(w') \, dw' \, dz' \, dy' \, dx' - \int_0^d n(x') \int_0^{x'} \int_0^{y'} n(z')$$

$$\times \int_0^{z'} dw' \, dz' \, dy' \, dz'. \tag{A16}$$

The function $n(x')$ is piecewise continuous and admits the Fourier series expansion

$$\eta(x') = \bar{\eta} + \sum_{p=1}^{\infty} \left(a_p \cos\frac{2\pi p}{d} x' + b_p \sin\frac{2\pi p}{d} x' \right) \qquad \text{(A17)}$$

where the Fourier coefficients are

$$\bar{\eta} = \frac{1}{d} \int_0^d \eta(x')\, dx'$$

and

$$a_p = \frac{2}{d} \int_0^d \eta(x') \cos\left(\frac{2\pi}{d} p x'\right) dx',$$

$$b_p = \frac{2}{d} \int_0^d \eta(x') \sin\left(\frac{2\pi}{d} p x'\right) dx'.$$

When the expansion (A17) is substituted for η in Eq. (A16), and the rather lengthy but straightforward algebra is carried out, we find that q is connected to the Fourier coefficients by the following simple relation

$$q = \frac{d^4}{8\pi^2} \sum_{p=1}^{\infty} \frac{a_p^2 + b_p^2}{p^2}. \qquad \text{(A18)}$$

Clearly, q and, therefore, v_2 are always positive if the impedance varies with position, and zero if the impedance is constant.

[1]C. T. Sun, J. D. Achenbach, and G. Herrmann, J. Appl. Mech. 35, 467 (1968).
[2]E. H. Lee, and W. H. Yang, SIAM J. Appl. Math. 25, 492 (1973).
[3]G. N. Balanis, J. Appl. Mech. 40, 815 (1973).
[4]J. C. Peck and G. A. Gurtman, J. Appl. Mech. 36, 479 (1969).
[5]C. D. Lundergan and D. S. Drumheller, J. Appl. Phys. 42, 669 (1971).
[6]G. A. Hegemier and A. H. Nayfeh, J. Appl. Mech. 40, 503 (1973).
[7]L. M. Barker, J. Compos. Mater. 5, 140 (1971).
[8]K. O. Friedrichs, *Advanced Ordinary Differential Equations* (Gordon and Breach, New York, 1965), pp. 103—108.

1387   J. Math. Phys., Vol. 16, No. 7, July 1975

George N. Balanis   1387

The influence of linear damping on nonlinearly coupled positive and negative energy waves

V. Fuchs

Direction Sciences de Base, Institut de Recherche de l'Hydro-Québec, Varennes, Québec, Canada
(Received 26 December 1974)

The linearly damped response to the nonlinear resonant mixing of two monochromatic coherent waves, involving modes of different energy sign, is shown to be always explosively unstable. Degeneration theory, modified to encompass explosively unstable solutions, is then applied to distinguish regions of negligible and strong damping, where the equations can be solved analytically. Effective damping, characterized by a damping rate ν, much higher than the (normalized) initial excitation U_0, of the source waves, increases the explosion time by a factor of ν/U_0.

I. INTRODUCTION

A wave is said to be explosively unstable if its amplitude A diverges in a finite time t_{ex}. For three monochromatic, coherent waves a behavior of this type takes place as a result of their nonlinear resonant interaction if the wave with the highest frequency has energy of opposite sign to the waves with lower frequencies,[1] a situation which can arise in an anisotropic plasma.[2,3]

At perfect frequency and wavenumber matching, the corresponding equations,[1,3] including the effect of linear damping on the wave amplitudes but neglecting its possible contribution[4] to the wave-phases, are to lowest nonlinear order

$$\frac{dA_i}{dt} + \nu_i A_i = M_i A_j A_k,\tag{1}$$

$$(i, j, k) = (1, 2, 3) \text{ cycl.}$$

The initial conditions $A_{i0} \equiv A_i(0)$ satisfy $A_{i0} \neq 0$ for at least two waves. All quantities in Eqs. (1) are real, the ν_i are nonnegative, and the coupling coefficients M_i, characteristic of the plasma and the respective wave-mode, are positive.

The system of Eqs. (1) is integrable by quadratures if and only if $\nu_1 = \nu_2 = \nu_3 \equiv \nu$, a property which generally follows from the existence of two independent invariants for this case, $A_1^2 - A_2^2 = C_{12}$ and $A_1^2 - A_3^2 = C_{13}$. The actual integration proceeds as in the conservative case,[1] after having applied the transformations introduced by Armstrong *et al.*,[5]

$$\bar{A}_i = A_i e^{\nu t}, \quad \bar{t} = \nu^{-1}(1 - e^{-\nu t}),\tag{2}$$

which reduce (1) to

$$\frac{d\bar{A}_i}{d\bar{t}} = M_i \bar{A}_j \bar{A}_k.\tag{3}$$

When the ν_i are not all the same, which is the physically more realistic case, the solutions of (1) have to be studied by approximate methods. Wilhelmsson[6] proceeds along lines suggested by the transformations (2), which are generalized in keeping with the requirement that the development of each wave be initially controlled by its respective damping rate ν_i, but depends asymptotically on a symmetrical form of the ν_i. The approximate solutions thus constructed are equal to the exact solutions for $\nu_1 = \nu_2 = \nu_3$, but for appreciably large differences in the damping rates their validity is question-

able. For example, the solutions predict stability whenever $3/\sum \nu_i$ is larger than the conservative explosion time. This result is in contradiction with a fact pointed out by Jungwirth[7] and proved quite generally in Sec. III below, namely that when only one wave is damped the interaction remains unstable for any finite value of ν.

Jungwirth's study[7] is more analytic in nature, attempting to analyze the effect of damping when the initial conditions are equal for all three waves. The discussion concentrates principally on the physically more significant cases of highly asymmetric damping conditions, with only one or two of the waves subject to damping. In particular, in the case of only one damped wave, it is shown that damping can be effective only during a limited interval of time, a fact which figures prominently in the present analysis.

In further studies of a more general nature, Wilhelmsson *et al.*[4] find the necessary conditions for instability, and Wang[8] presents sufficient conditions for stability.

In the present study we treat the inherently unstable situation in which only one wave is subject to damping. We propose to study this case by extending degeneration theory,[9] applied previously[10] to strongly damped, stable, positive energy waves, to unstable situations. We will then be able to define regions of strong and weak damping, respectively, where analytical approximations are available.

In Sec. II we present, for further reference, conservative solutions valid for arbitrary initial conditions.

In Sec. III we discuss system (1) in the phase space of a generating function for the wave amplitudes, and outline, from the point of view of degeneration theory, the difference between exponential and explosive instabilities.

In Sec. IV we develop the degeneration technique in the vicinity of the explosion time.

II. CONSERVATIVE SOLUTIONS

In this section we present the solutions of system (1), in the absence of damping, for arbitrary initial conditions, and discuss the explosion time as a function of the initial conditions.

The substitution[4]

$$u_i = (M_j M_k)^{1/2} A_i, \quad (i, j, k) = (1, 2, 3) \text{ cycl} \tag{4}$$

transforms (1) into the more convenient form

$$u_i' + \nu_i u_i = u_j u_k, \tag{5}$$

where the prime denotes differentiation with respect to time t.

Let us now consider the case of all $\nu_i = 0$. The symmetry of Eqs. (5) allows us to assume, without loss of generality, that the initial conditions satisfy the inequalities

$$u_{30} \leqslant u_{20} \leqslant u_{10}. \tag{6}$$

The solutions for other initial conditions can be obtained by the appropriate permutation of indices in the solutions corresponding to case (6). Let us first consider the case

(i) $u_{30} < u_{20} < u_{10}$

Denote

$$a^2 = u_{10}^2 - u_{30}^2, \quad b^2 = u_{20}^2 - u_{30}^2, \tag{7}$$

$$k^2 = (a^2 - b^2)/a^2. \tag{8}$$

The solution u_3 is the Jacobian elliptic function[11]

$$u_3(t) = a\, cs[-at + tn^{-1}(a/u_{30}, k), k], \tag{9}$$

with k as the modulus. The solutions $u_{1,2}$ follow from the invariants

$$u_1^2 - u_3^2 = a^2, \quad u_2^2 - u_3^2 = b^2. \tag{10}$$

The functions u_i diverge at

$$t_{ex} = (1/a) tn^{-1}(a/u_{30}, k), \tag{11}$$

where $tn^{-1} \in (0, K(k))$, and $K(k)$ is the complete elliptic integral of the first kind.

Of particular interest is the special case $u_{30} = 0$, which yields

$$u_3 = u_{20} tn(u_{10} t, k) \tag{12}$$

with

$$t_{ex} = \frac{1}{u_{10}} K(k) = \begin{cases} \dfrac{\pi}{2u_{20}}, & u_{10} \to u_{20} & (13a) \\[3mm] \dfrac{1}{u_{10}} \ln \dfrac{4u_{10}}{u_{20}}, & u_{10} \gg u_{20}. & (13b) \end{cases}$$

It is now interesting to follow the degradation of the solutions from Jacobian elliptic functions through transcendental elementary-to-elementary functions, as we allow coalescence of the initial conditions.

(ii) $u_{30} = u_{20} < u_{10}$

$$u_3 = \frac{a}{\sinh[-at + \sinh^{-1}(a/u_{20})]} \tag{14}$$

$$t_{ex} = \frac{1}{a} \sinh^{-1} \frac{a}{u_{20}} = \begin{cases} \dfrac{1}{u_{20}}, & u_{10} \to u_{20} & (15a) \\[3mm] \dfrac{1}{u_{10}} \ln \dfrac{2u_{10}}{u_{20}}, & u_{10} \gg u_{20}. & (15b) \end{cases}$$

(iii) $u_{30} < u_{20} = u_{10}$

$$u_3 = \frac{a}{\tan[-at + \tan^{-1}(a/u_{30})]}. \tag{16}$$

(iv) $u_{30} = u_{20} = u_{10}$

$$u = \frac{u_0}{-u_0 t + 1}. \tag{17}$$

An indication of how the initial conditions affect the explosion time follows, by way of comparison, from (13) and (15). First, as a function of the strongest excitation u_{10}, t_{ex} drastically changes its functional form as u_{10} increases, resulting in a considerable reduction of t_{ex}. Second, as a function of the weakest excitation, u_{30}, t_{ex} changes at most by a factor of $\pi/2$.

In the subsequent analysis, therefore, we shall always make the simplifying assumption $u_{30} = 0$.

III. EFFECT OF DAMPING: BASIC CONSIDERATIONS

Now, let the wave u_3 be subject to damping for initial conditions $u_{30} = 0$, $u_{10} u_{20} \neq 0$. Physically, this corresponds to the generation of a damped response u_3 by nonlinear mixing of the waves u_1 and u_2.

Representing this case are Eqs. (5) in the form

$$U' = u u_3, \tag{18a}$$

$$u' = U u_3, \tag{18b}$$

$$u_3' + \nu u_3 = u U, \tag{18c}$$

with the initial conditions $u_{30} = 0$,

$$U_0 = \max(u_{10}, u_{20}) \text{ and } u_0 = \min(u_{10}, u_{20}). \tag{19}$$

Thus $U_0 \geqslant u_0$, and the single invariant of system (18) is

$$U^2 - u^2 = U_0^2 - u_0^2 = C^2. \tag{20}$$

Let us assume $C \neq 0$ and make use of (20) to convert (18) into a second order system for a single generating function. The function ψ defined by

$$U = C \cosh\psi, \tag{21a}$$

$$u = C \sinh\psi \tag{21b}$$

satisfies (20) identically. Further, with

$$u_3 = \psi' \tag{22}$$

Eqs. (18a, b) are satisfied identically, while Eq. (18c) becomes

$$\psi'' + \nu\psi' - \tfrac{1}{2}C^2 \sinh2\psi = 0,$$
$$\tag{23}$$
$$\psi_0 = \cosh^{-1} \frac{U_0}{C}, \quad \psi_0' = 0.$$

Equation (23) has just one singular point in phase space (ψ, ψ'), namely on unstable saddle point at $(0, 0)$. In the presence of such a singular point, the linear damping term cannot remove the instability of the solution,[9] which thus remains explosive for any finite rate of damping. This property of the solution excludes the straightforward application of degeneration theory to (23), in contrast to the case of stable positive energy waves.[10] As a matter of fact, degeneration in the explosively unstable case is not even similar to the degeneration of exponentially unstable systems. It is quite instructive to follow the procedure in the latter case.

Let us, therefore, consider the linear analogy of (23), that is,

$$y'' + \nu y' - C^2 y = 0,$$

$$y(0) = y_0 > 0, \quad y'(0) = 0. \tag{24}$$

The degenerate equation associated with (24) is

$$\nu z' - C^2 z = 0,$$

$$z(0) = y_0. \tag{25}$$

Its solution, $z = y_0 \exp(C^2 t/\nu)$, satisfies the relation

$$\frac{\nu z'}{z''} = \nu^2/C^2 \tag{26}$$

for all $t \in \langle 0, \infty \rangle$. Therefore, if $\nu^2 \gg C^2$, Eq. (24) is degenerescent for all $t > 0$, and $y \approx z$. Indeed, as is easily verified, when $4C^2/\nu^2 \ll 1$, the solution of (24) becomes

$$y \to y_0 [1 - O(C^2/\nu^2)] \exp(C^2 t/\nu). \tag{27}$$

Now, let us turn to the degenerate equation

$$\nu \phi' - \tfrac{1}{2} C^2 \sinh 2\phi = 0, \tag{28}$$

$$\phi_0 = \psi_0$$

associated with (23). Its solution

$$\phi(t) = \tanh^{-1} Q(t) \tag{29}$$

$$Q(t) = \frac{u_0}{U_0} \exp(C^2 t/\nu) \in \langle u_0/U_0, 1 \rangle \tag{30}$$

diverges as $Q \to 1$, that is, at

$$t_{\text{ex}}^{(\phi)} = \frac{\nu}{C^2} \ln \frac{U_0}{u_0}. \tag{31}$$

The u-representation corresponding to ϕ is, from (21) and (22),

$$U^{(\phi)2} = C^2 \frac{1}{1 - Q^2}, \tag{32a}$$

$$u^{(\phi)2} = C^2 \frac{Q^2}{1 - Q^2}, \tag{32b}$$

$$u_3^{(\phi)} = \frac{C^2}{\nu} \frac{Q}{1 - Q^2}. \tag{32c}$$

In the special case $C = 0$, for which the ψ-representation does not exist, we obtain, upon taking the limit $u_0 \to U_0$,

$$t_{\text{ex}}^{(\phi)} = \nu/2U_0^2 \tag{33}$$

$$u^{(\phi)2} = U^{(\phi)2} = U_0^2/(1 - 2U_0^2 t/\nu) \tag{34a}$$

$$u_3^{(\phi)} = U^{(\phi)2}/\nu. \tag{34b}$$

We now define the degree of degeneracy $D(t)$,

$$D(t) = \frac{\nu \phi'}{\phi''} = \frac{\nu^2}{C^2} \frac{1 - Q^2}{1 + Q^2}. \tag{35}$$

First, it is easy to show that if

$$\frac{\nu^2}{u_0^2 + U_0^2} < 1, \tag{36}$$

then $D(t) < 1$ for all $t \in \langle 0, t_{\text{ex}}^{(\phi)} \rangle$ and $u_0/U_0 \in (0, 1)$. Thus, condition (36) is sufficient to assure that Eq. (23) be never degenerescent. In this (weakly damped) regime,

the solution ψ of Eq. (23) is determined by the first and third terms in (23) and the effect of damping can be neglected altogether.

Let therefore $\nu \gg U_0$. Then there can exist an interval of time in which $D(t) \gg 1$, but, since ϕ approaches a singularity as $t \to t_{\text{ex}}^{(\phi)}$, ϕ'' eventually dominates over ϕ' and degeneration breaks down. In the case of strong damping, therefore, Eq. (23) can never be degenerescent on the entire interval $\langle 0, t_{\text{ex}}^{(\psi)} \rangle$, where $t_{\text{ex}}^{(\psi)}$ is the explosion time of the solution ψ of Eq. (23).

IV. DEGENERATION NEAR SINGULARITY

We will now proceed to show that, as damping becomes effective $(\nu \gg U_0)$, the solution ψ is well-approximated by ϕ on the major part of $\langle 0, t_{\text{ex}}^{(\psi)} \rangle$. The procedure is to demonstrate that both $t_{\text{ex}}^{(\phi)}$ and $t_{\text{ex}}^{(\psi)}$ fall within the limits of an interval whose extent is small in comparison with $t_{\text{ex}}^{(\phi)}$, and that Eq. (23) is degenerescent for almost all times up to $t_{\text{ex}}^{(\phi)}$.

To this effect let us prove the following two statements. Let

$$\nu/U_0 \gg 1 \tag{37}$$

Then:

(a) The explosion time $t_{\text{ex}}^{(\psi)}$ of the solution ψ of Eq. (23), and the explosion time $t_{\text{ex}}^{(\phi)}$, given by (31), lie within the interval

$$\langle t_1, t_1 + 2/U_0 \rangle. \tag{38}$$

Here, U_0 is given by (19) and t_1 is determined by the function

$$t(A) \equiv t_A = \frac{\eta}{\nu} \ln \frac{U_0}{u_0} \frac{\eta - A}{\eta + A} \tag{39}$$

(with $\eta = \nu^2/C^2$), the inverse of $A = D(t)$, where $D(t)$ is the degree of degeneracy (35).

(b) The times $t_{\text{ex}}^{(\phi)}$, $t_{\text{ex}}^{(\psi)}$ and t_1 satisfy the inequalities

$$|t_{\text{ex}}^{(\phi)} - t_{\text{ex}}^{(\psi)}|/t_{\text{ex}}^{(\phi)} < 4U_0/\nu, \tag{40a}$$

$$(t_{\text{ex}}^{(\phi)} - t_1)/t_{\text{ex}}^{(\phi)} < 4U_0^2/\nu^2. \tag{40b}$$

We start with (a). By definition, ϕ approximates ψ very well at t_A when $A \gg 1$, whereas around t_1 degeneracy breaks down, and the solutions ϕ and ψ diverges. For $t > t_1$ the solution ψ approaches its region of singularity, characterized by the dominance of ψ'' over $\nu \psi'$. In this interval the damping term can therefore be neglected and the solution is well-approximated by the function (9) with the appropriate initial conditions given by the amplitudes (32) evaluated at $t = t_1$. Let us denote by $\psi^{(s)}$ the solution thus obtained. The quality of the approximation $\psi^{(s)}$ is essentially determined by the deviation of the initial conditions from the actual values of $\psi(t_1)$ and $\psi'(t_1)$. In the region $\langle 0, t_1 \rangle$, where ψ'' represents a small correction term in Eq. (23), we easily find $\phi > \psi$ and $\phi' > \psi'$, since ψ is convex, implying $\psi'' > 0$. Thus $\psi(t_1)$ and $\psi'(t_1)$ lie inside the intervals $(\phi(t_A), \phi(t_1))$ and $(\phi'(t_A), \phi'(t_1))$, respectively, where A is large enough to guarantee a high degree of degeneracy at t_A. For the sake of definiteness, let us take $A = \nu/U_0$.

1390   J. Math. Phys., Vol. 16, No. 7, July 1975

V. Fuchs   1390

We will now assess the effect which this margin of error in initial conditions will have on the estimated explosion time $t_{ex}^{(\psi)} \approx t_1 + t_{ex}^{(s)}$. Shifting the initial conditions from t_1 to t_A along the solutions (32) gives

$$u_1^{(s)}(0) \equiv U^{(\phi)}(t_A) = C[(\eta + A)/2A]^{1/2}, \tag{41a}$$

$$u_2^{(s)}(0) \equiv u^{(\phi)}(t_A) = C[(\eta - A)/2A]^{1/2} \tag{41b}$$

$$u_3^{(s)}(0) \equiv u_3^{(\phi)}(t_A) = \frac{\nu}{2A\eta}(\eta^2 - A^2)^{1/2} \tag{41c}$$

Obviously $u_{10}^{(s)} > u_{20}^{(s)}$, but also $u_{20}^{(s)} > u_{30}^{(s)}$ when $A \geqslant 1$. Therefore, the solution $u_3^{(s)}(t)$ is given by the solution (9) of case (i) with

$$a^{(s)2} = [C^2(\eta + A)/4A^2\eta](2A\eta - \eta + A), \tag{42a}$$

$$b^{(s)2} = [C^2(\eta - A)/4A^2\eta](2A\eta - \eta - A) \tag{42b}$$

$$k^{(s)2} = [4A^2\eta/(\eta + A)(2A\eta - \eta + A)]. \tag{43}$$

Since (37) implies $\eta \gg A$, it follows that

$$k^{(s)} \approx (2A/\eta)^{1/2} \ll 1 \tag{44}$$

and, consequently, $K(k^{(s)}) \to \pi/2$, $cs \to cot$, $tn \to tan$, so that the Jacobian elliptic solution becomes

$$u_3^{(s)} = a^{(s)} \cot\left(- a^{(s)}t + \tan^{-1}\frac{a^{(s)}}{u_{30}^{(s)}}\right). \tag{45}$$

The corresponding explosion time $t_{ex}^{(s)}$ is

$$t_{ex}^{(s)} = \frac{1}{a^{(s)}}\tan^{-1}\frac{a^{(s)}}{u_{30}^{(s)}} \approx \frac{\sqrt{2A}}{\nu}\tan^{-1}\sqrt{2A}, \tag{46}$$

so that

$$t_{ex}^{(s)} < 2A/\nu \tag{47}$$

Further, making use again of $\eta \gg A$, we obtain, from (39),

$$t_A \approx \frac{\nu}{C^2}\ln\frac{U_0}{u_0} - \frac{2A}{\nu} = t_{ex}^{(\phi)} - \frac{2A}{\nu}. \tag{48}$$

The total estimated explosion time, taking t_1 as the time separating the degenerate and singular regions, is

$$t_{ex}^{(\psi)} \approx t_1 + t_{ex}^{(s)}. \tag{49}$$

Thus, since $A = \nu/U_0$,

$$t_1 < t_{ex}^{(\psi)} < t_1 + 2/U_0. \tag{50}$$

Further, from (48),

$$t_{ex}^{(\phi)} = t_1 + 2/\nu < t_1 + 2/U_0, \tag{51}$$

which completes the proof of statement (a).

Let us now turn to (b). Obviously,

$$|t_{ex}^{(\psi)} - t_{ex}^{(\phi)}| < 2/U_0. \tag{52}$$

Thus

$$\rho = \frac{t_{ex}^{(\phi)}}{|t_{ex}^{(\phi)} - t_{ex}^{(\psi)}|} > \frac{U_0\nu}{2C^2}\ln\frac{U_0}{u_0}. \tag{53}$$

Since $C^2 = U_0^2 - u_0^2$, we may write

$$\rho > \frac{\nu}{2U_0}\frac{1}{1 - x^2}\ln\frac{1}{x}, \tag{54}$$

with $x = u_0/U_0 \in (0, 1)$, so that

$$\rho > \nu/4U_0. \tag{55}$$

Further,

$$\sigma = \frac{t_{ex}^{(\phi)}}{t_{ex}^{(\phi)} - t_1} = \frac{\nu^2}{2C^2}\ln\frac{U_0}{u_0}, \tag{56}$$

so that

$$\sigma = \frac{\nu^2}{2U_0^2}\frac{1}{1 - x^2}\ln\frac{1}{x} > \frac{\nu^2}{4U_0^2}, \tag{57}$$

which completes the proof of (b).

V. CONCLUSION

When only one of three waves participating in the resonant interaction between positive and negative energy modes is damped, the interaction remains explosively unstable for any finite damping rate, ν. The system of nonlinear differential equations (18), describing this nonconservative interaction, is characterized by the existence of only one invariant (20) and therefore cannot be integrated by quadratures. However, the system can be transformed into a second order differential equation (23) for a generating function, defined by (21), allowing the application of degeneration theory, generalized to encompass explosively unstable situations. This procedure enables us to ascertain that strong damping substantially increases the explosion time.

In more definite terms, when the rate of damping is small,

$$\nu^2/(u_0^2 + U_0^2) < 1, \tag{58}$$

Eq. (23) is never degenerescent, the effect of damping can be neglected, and the explosion time is approximated by the conservative expression (13).

On the other hand, when damping is effective,

$$\nu/U_0 \gg 1, \tag{59}$$

it follows from (40a) and (40b) that Eq. (23) is degenerescent on the interval $\langle 0, t_d \rangle$, where

$$t_d = t_{ex}^{(\psi)}[1 - O(U_0/\nu)], \tag{60}$$

while

$$t_{ex}^{(\psi)} = t_{ex}^{(\phi)}[1 \pm O(U_0/\nu)]. \tag{61}$$

Therefore, the solution ψ of Eq. (23) can be approximated by the solution (29) of the degenerate equation (28) with an acceptable margin of error.

By way of comparison, we infer from (13a, b) and (33), (31), respectively, that effective damping increases the explosion time by a factor of ν/U_0.

A few words concerning the initial conditions should be added. The degenerate approximations (32a, b) for U and u satisfy the proper initial conditions U_0 and u_0, respectively, whereas $u_3^{(\phi)}$ does not. As shown in Ref.

10, the connection between the initial condition $u_3(0) = 0$ and the degenerate expression (32c) is realized by the multiplication factor $1 - \exp(-\nu t)$. On the time scale of the explosion time $t_{ex}^{(\phi)}$ this transition, characterized by the relaxation time $1/\nu$ is very fast, as can be seen from (56), and thus has a negligible effect on the evolution of the solution.

As a final observation, we wish to point out that the procedure of degeneration presented herein can be applied, with appropriate modifications, to any dissipative second order system with a negative nonlinear control term leading to a saddle point singularity in phase space (Y, Y'), provided that we specify $Y \neq 0$ as an initial condition.

ACKNOWLEDGMENTS

[1] B. Coppi, M.N. Rosenbluth, and R.N. Sudan, Ann. Phys. 55, 207 (1969).
[2] M.N. Rosenbluth, B. Coppi, and R.N. Sudan, Ann. Phys. 55, 248 (1969).
[3] J. Fukai, S. Krishan, and E.G. Harris, Phys. Fluids 13, 3031 (1970).
[4] H. Wilhelmsson, L. Stenflo, and F. Engelmann, J. Math. Phys. 11, 1738 (1970).
[5] J.A. Armstrong, N. Bloembergen, J. Ducuing, and P.S. Pershan, Phys. Rev. 127, 1918 (1962).
[6] H. Wilhelmsson, Physica Scripta 2, 113 (1970).
[7] K. Jungwirth, Czech. J. Phys. B22, 756 (1972).
[8] P.K.C. Wang, J. Math. Phys. 13, 943 (1972).
[9] N. Minorsky, Nonlinear Oscillations (Van Nostrand, Princeton, 1962), Chap. 26, 1.
[10] V. Fuchs and G. Beaudry, J. Math. Phys. 16, 616 (1975).
[11] P.F. Byrd and M.D. Friedman, Handbook of Elliptic Integrals for Engineers and Scientists (Springer, New York, 1971).

1392   J. Math. Phys., Vol. 16, No. 7, July 1975

V. Fuchs   1392

The Hamiltonian $H = (- 1/2)\,d^2/dx^2 + x^2/2 + \lambda/x^2$ reobserved*

L. Lathouwers†

Dienst Voor Teoretische en Wiskundige Natuurkunde, Rijksuniversitair Centrum, Antwerp, Belgium
(Received 11 February 1975)

The Schrödinger problem for the title Hamiltonian is considered as a perturbed one-dimensional harmonic oscillator. Exact bound state solutions can be derived from a classical differential equation in the theory of Laguerre polynomials. These solutions are valid and analytically dependent on λ only in a limited range of the perturbation strength. Within this region the oscillator Hamiltonian restricted to odd and even parity subspaces is unitary equivalent to H restricted over the respective perturbed subspaces. It is shown that due to the singular nature of the perturbation the allowed λ range is narrowed if side conditions are imposed to make the wavefunctions "physically interpretable."

1. EXACT SOLUTION OF THE PERTURBED PROBLEM IN $L^2(-\infty, +\infty)$

The differential operator H of the eigenvalue problem

$$H\psi(x) = E\psi(x), \tag{1}$$

where $\psi(x)$ belongs to $L^2(-\infty, +\infty)$, can conveniently be written in the form

$$H = H_0 + \lambda V,$$
$$H_0 = -\tfrac{1}{2}d^2/dx^2 + x^2/2, \quad V = 1/x^2. \tag{2}$$

The unperturbed Hamiltonian H_0 is the harmonic oscillator operator while V is a singular perturbation the strength of which is measured by the parameter λ. For convenience I have set $\hbar = 1$ and considered a particle of unit mass. The unperturbed eigenvalue problem, obtained from (1) in the limit $\lambda \to 0$, has the well-known solutions

$$E_n^0 = n + \tfrac{1}{2}, \quad \psi_n^0(x) = (2^n n! \sqrt{\pi})^{-1/2} e^{-x^2/2} H_n(x), \tag{3}$$

where n is a nonnegative integer and $H_n(x)$ are the Hermite polynomials. The solutions of (1) can be obtained directly if one identifies the problem with the following differential equation occurring in the theory of orthogonal polynomials (Ref. 1):

$$-\frac{1}{2}\frac{d^2\varphi_n^{(\alpha)}(x)}{dx^2}$$
$$- \left(2n + \alpha + 1 - \frac{x^2}{2} + \frac{1-4\alpha^2}{8x^2}\right)\varphi_n^{(\alpha)}(x) = 0, \tag{4}$$

which, when $\alpha > -1$, has solutions in $L^2(0, +\infty)$ given by

$$\varphi_n^{(\alpha)}(x) = x^{\alpha+1/2} e^{-x^2/2} L_n^{(\alpha)}(x^2), \tag{5}$$

where the $L_n^{(\alpha)}$ are the standard generalized Laguerre polynomials. Identifying the differential operator in (4) with H, one finds an expression for the perturbation strength

$$\lambda = \tfrac{1}{2}\alpha^2 - \tfrac{1}{8}. \tag{6}$$

Since λ depends only on α^2, the eigenvalues and eigenfunctions will occur in pairs associated with $+\alpha$ or $-\alpha$, where α is given by

$$\alpha = +\tfrac{1}{2}\sqrt{1 + 8\lambda}. \tag{7}$$

From (4) one sees that the eigenvalues are

$$E_n^{(\pm\alpha)} = 2n \pm \alpha + 1 \tag{8a}$$

with corresponding eigenfunctions, which are either odd or even since H is parity invariant, given by

$$\psi_n^{(\pm\alpha)}(x) = [n!/\Gamma(n \pm \alpha + 1)]^{1/2}$$
$$\times \begin{cases} \binom{+}{+}\varphi_n^{(\pm\alpha)}(|x|), & x \ge 0, \\ \binom{-}{+}\varphi_n^{(\pm\alpha)}(|x|), & x < 0. \end{cases} \tag{8b}$$

Such a pair of solutions exists only if α belongs to the half open interval $[0, +1[$ since for $\alpha = +1$ the eigenfunction $\psi_n^{(-\alpha)}(x)$ is no longer square integrable. The following orthogonality relations hold:

$$\int_{-\infty}^{+\infty} \bar{\psi}_n^{(\epsilon\,\alpha)}(x)\psi_{n'}^{(\epsilon'\,\alpha)}(x)\,dx = \delta_{nn'}\delta_{\epsilon\epsilon'}, \tag{9}$$

where ϵ and ϵ' may equal ± 1. For $\epsilon = \epsilon'$ this follows from the orthogonality relations of Laguerre polynomials while for $\epsilon \ne \epsilon'$ (9) holds for parity reasons, as can be verified directly from (5).

The function $\alpha = \alpha(\lambda)$ has a branch point at $\lambda = \tfrac{1}{8}$. Introducing a cut from $-\tfrac{1}{8}$ to $-\infty$ in the complex λ plane, an analytic continuation of α can be realized on a double Riemann sheet. Taking the limit $\lambda \to 0$ along a certain path on this surface, we have that

$$\lim_{\lambda \to 0} E_n^{(-\alpha)} = 2n + \tfrac{1}{2} = E_{2n}^0,$$
$$\lim_{\lambda \to 0} E_n^{(+\alpha)} = 2n + \tfrac{3}{2} = E_{2n+1}^0, \tag{10}$$

i.e., the eigenvalues $E_n^{(-\alpha)}$ and $E_n^{(+\alpha)}$ are the perturbed energies corresponding to the unperturbed even and odd labelled levels respectively. A similar result holds for the eigenfunctions which are analytical functions of λ on the Riemann sheet as can be seen from (5). It follows that

$$\lim_{\lambda \to 0} \psi_n^{(-\alpha)}(x) = (-)^n \psi_{2n}^0(x),$$
$$\lim_{\lambda \to 0} \psi_n^{(+\alpha)}(x) = (-)^n \psi_{2n+1}^0(x), \tag{11}$$

which is easily checked using the relations between Hermite and Laguerre polynomials:

$$H_{2n}(x) = (-)^n 2^{2n} n! \, L_n^{(-1/2)}(x^2),$$

$$H_{2n+1}(x) = (-)^n 2^{2n+1} n! \, x L_n^{(+1/2)}(x^2). \tag{12}$$

According to (8a) the perturbed energy spectrum of (1) consists of two sets of equidistant levels $E_n^{(\pm\alpha)}$ which can be obtained from the unperturbed spectrum by uniformly shifting the even (odd) parity levels over a distance $\Delta = -\alpha + \frac{1}{2}$ $(-\Delta)$. The corresponding nonidentical twin eigenfunctions $\psi_n^{(-\alpha)}(x)$ and $\psi_n^{(+\alpha)}(x)$ exist only for α in the half open interval $[0, +1[$. From (6) one can then immediately derive the allowed range for the perturbation strength. It follows that λ has to be situated in the half open interval $[-\frac{1}{8}, +\frac{3}{8}[$. These bounds have a very precise mathematical meaning, in fact, the interval $[-\frac{1}{8}, +\frac{3}{8}[$ is nothing but the neighborhood of $\lambda = 0$ within which the solutions depend analytically upon the perturbation strength. The upper limit $\lambda = +\frac{3}{8}$, obtained for $\alpha = +1$, gives the maximum value of λ for which the solutions of (1) can be represented by (5). Indeed the wavefunction $\psi_n^{(-\alpha)}(x)$ ceases to be square integrable at $\alpha = +1$. Therefore, the form of these wavefunctions must change in a singular way at $\lambda = +\frac{3}{8}$. This is also reflected in the fact that the limit of the ground state is zero $(\lim_{\lambda \to 0} E_0^{(-\alpha)} = 0)$ which is clearly inconsistent since the Hamiltonian is positive definite for $\lambda = +\frac{3}{8}$. At the lower limit for the perturbation strength $\lambda = -\frac{1}{8}$ the perturbation is negative definite. It can be shown from a theorem by Kato (Ref. 2) (as has been mentioned also in Ref. 3) that this number is the minimum value for which H is still bounded from below. Physically this is associated with the possibility for the particle to "fall" to the center, i.e., the occurrence of infinite negative energies (Ref. 4). Here the point $\lambda = -\frac{1}{8}$ showed up in a mathematical way as a branch point beyond which the solution of (1) (eigenfunctions and eigenvalues) is no longer analytically dependent on λ. These bounds for the perturbation strength can also be obtained from the solutions of (1) near the origin. One finds that the characteristic exponents (which describe the singularities of the solutions near $x = 0$) are given by $(1 \pm \sqrt{1+8\lambda})/2$. It is a well-known fact (Ref. 5) that second order differential operators are no longer bounded from below if the characteristic exponents are imaginary, i.e., if $\lambda < -\frac{1}{8}$. Furthermore, if $\lambda \geqslant +\frac{3}{8}$, the solution corresponding to the minus sign is not square integrable at the origin. In fact, $\lambda = +\frac{3}{8}$ is the transition point from limit circle to limit point case in the Weyl classification of second order differential equations (Ref. 6). From a theorem by Titchmarsh (Ref. 7) one can show that for $\lambda \in [-\frac{1}{8}, +\frac{3}{8}[$ the differential operator is of the limit circle type at $x = 0$ (i.e., there exist two linearly independent solutions square integrable at $x = 0$) while for $\lambda \geqslant +\frac{3}{8}$ one is in the limit point case (i.e., there is a unique solution square integrable at $x = 0$).

In Refs. 3 and 8 the solutions $\psi_n^{(-\alpha)}(x)$ were considered to be improper eigenfunctions, in spite of their square integrability, due to their behavior at the origin (see, further, Sec. III). In order to get a complete set of eigenfunctions, one considered as "physically acceptable solutions" on the real line the functions $\varphi_n^{(+\alpha)}(x)$ with the prescription

$$\varphi_n^{(+\alpha)}(-x) = \pm \, \varphi_n^{(+\alpha)}(x), \quad x \geqslant 0. \tag{13}$$

The resulting eigenfunctions are $\psi_n^{(+\alpha)}(x)$ and a new function $\phi_n^{(+\alpha)}(x)$, which is energywise degenerate with $\psi_n^{(+\alpha)}(x)$. However, this degeneracy is not removed in the limit $\lambda \to 0$ while the unperturbed even parity levels do not show up at all. Therefore, one is in fact treating another eigenvalue problem, corresponding to an operator defined by the integral kernel

$$H(x, x') = \sum_{n=0}^{\infty} E_n^{(+\alpha)} (\psi_n^{(+\alpha)}(x)\bar{\psi}_n^{(+\alpha)}(x')$$

$$+ \phi_n^{(+\alpha)}(x)\bar{\phi}_n^{(+\alpha)}(x')) \tag{14}$$

rather than the eigenvalue problem (1). Observe that the limit of (14) as $\lambda \to 0$ is only "half an harmonic oscillator" (odd parity levels only) with twofold degenerate levels (Ref. 3).

II. UNITARY EQUIVALENCE

The question of unitary equivalence which was raised in Ref. 3 can now be treated as follows. Let $H_0^{(+)}$ be the restriction of H_0 to the even parity subspace $^{(+)}L^2(-\infty, +\infty)$ spanned by the unperturbed eigenstates $\{\psi_{2n}^0(x)\}$. Similarly we introduce $H^{(-\alpha)}$ as the restriction of H over the subspace generated by $\{\psi_n^{(-\alpha)}(x)\}$. The unitary mapping $U(-\alpha)$ which maps the above subspaces onto each other immediately establishes a unitary equivalence between $H^{(-\alpha)} - \Delta I^{(-\alpha)}$ and $H_0^{(+)}$ $(I^{(-\alpha)}$ is the unit operator in the space spanned by $\{\psi_n^{(-\alpha)}(x)\}$). Indeed it is easily checked that

$$[U(-\alpha)H_0^{(+)}U^*(-\alpha) - (H^{(-\alpha)} - \Delta I^{(-\alpha)})]\psi_n^{(-\alpha)}(x) = 0, \tag{15}$$

where Δ is the level shift. In the limit $\lambda \to 0$, $U(-\alpha)$ equal $I^{(-1/2)}$ and (15) becomes a trivial identity. In the same way one can prove that $H_0^{(-)}$ is unitary equivalent to $H^{(+\alpha)} - \Delta I^{(+\alpha)}$ (with obvious notations for the operator). The existance of the properties of the restricted Hamiltonians demonstrated above is not very surprising. Indeed, the perturbation causes a uniform shift Δ $(-\Delta)$ of even (odd) parity levels. The unitary equivalence of $H^{(-\alpha)}$ and $H_0^{(+)}$ $(H^{(+\alpha)}$ and $H_0^{(-)})$ is nothing but the operator analog of this special relation between the unperturbed and perturbed spectrum.

III. CONSEQUENCES OF PROBABILISTIC INTERPRETATIONS OF NONRELATIVISTIC QUANTUM MECHANICS

The original formulations of nonrelativistic quantum mechanics (QM) given by Heisenberg, Born, Jordan (matrix mechanics) and Schrödinger (wave mechanics) were shown to be mathematically equivalent by the latter one (Ref. 9). This was accomplished essentially by proving that the sequential Hilbert space is isometric to the space of square integrable functions. A little later Dirac (Ref. 10) managed to formulate QM in terms of elements of an abstract Hilbert space. In fact, he showed that any realization of the abstract properties of Hilbert space in different mathematical forms gives a possible formulation of QM. Therefore, as is, for instance stated by von Neumann (Ref. 11), the only restriction on the wavefunction in the Schrödinger picture is its single-valuedness and its

square integrability. The integral $\int_a^b |\psi(x)|^2 dx$ is then to be interpreted as the probability of finding the particle in the interval $[a, b]$. However, some authors (see, e.g., Refs. 8 and 12) consider a wavefunction physically interpretable only if $\psi(x)$ and $\psi(x)d\psi(x)/dx$ are continuous. These conditions assure that the point probability density $|\psi(x)|^2$ and the probability density current $\mathrm{Re}[\bar{\psi}(x)(\nabla/i)\psi(x)]$ vary continuously. In the case of singular potentials these conditions turn out to be too severe since one might lose physically observed states (e.g., the ground state of the hydrogen atom).

An interesting illustration of the implications of continuity conditions can now be given for the eigenvalue problem (1). The solutions as derived in Sec. I correspond to the most general point of view, namely the solution in $L^2(-\infty, +\infty)$ (notice that all spectra are entirely discrete such that no difficulties occur due to continuous parts in the spectra). If one restricts oneself to continuous wavefunctions, it follows from (5) that α has to be smaller than $+\frac{1}{2}$, which gives an allowed range of $[-\frac{1}{8}, 0[$ for the perturbation strength. This means that continuous wavefunctions exist only for not too strong negative definite perturbations. If in addition one requires that $\psi(x)d\psi(x)/dx$ is continuous, one finds that $\alpha = 0$ such that $\lambda = -\frac{1}{8}$ would give the only possible perturbed Hamiltonian which would be isolated from the unperturbed $\lambda = 0$ problem. As an alternative one may choose not to consider the $\psi_n^{(-\alpha)}(x)$ solutions at all (Refs. 3, 8) but then, as was shown in Sec. I, one is no longer considering the same eigenvalue problem.

The final justification for the conditions one imposes on the wavefunctions is, of course, the agreement of theory and experiment. The analysis given above merely shows that in the case of singular interactions too severe restrictions may have drastic consequences.

ACKNOWLEDGMENTS

*Work supported by the Interuniversitair Instituut voor Kernwetenschappen, Belgium and the Swedish Institute, Sweden.

†Presently at the Department of Quantum Chemistry, Uppsala University, P.O. Box 518, S-751 20, Uppsala, Sweden.

[1]M. Abromowitz and I. Stegun, *Handbook of Mathematical Functions* (National Bureau of Standards, Washington, D. C., 1964), p. 781.

[2]T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1966), p. 346, Theorem 4.2.

[3]F. Zirilli, J. Math. Phys. 15, 1202 (1974).

[4]L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Pergamon, New York, 1958), Sec. 35.

[5]F. Rellich, Math. Ann. 122, 343 (1951).

[6]H. Weyl, Math. Ann. 68, 220 (1910).

[7]E. C. Titchmarsh, *Eigenfunction Expansions* (Oxford U. P., Oxford, 1962), Sec. 5.15.

[8]F. Calogero, J. Math. Phys. 10, 2191 (1969).

[9]E. Schrödinger, Ann. Phys. 79, 361 (1925).

[10]P. A. M. Dirac, Proc. Roy. Soc. Lond. A 109, 642 (1925).

[11]J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton U. P., Princeton, N.J., 1955), p. 29, note 34.

[12]A. Messiah, *Mechanique Quantique* (Dunod, Paris, 1959), p. 65.

On convergent iterations and their applications to Hartree–Fock equations

A. Kelemen

Institut für Theoretische Physik, Frankfurt/Main, West Germany
(Received 9 September 1974)

A method for solving a restricted class of nonlinear equations is presented and applied in detail to solution of the Hartree–Fock (HF) equations.

1. INTRODUCTION

In a previously published note[1] a method has been presented which guarantees convergent iterations yielding a solution of the Hartree—Fock (HF) equations. Two aspects of this procedure, which falls within the framework of the steepest descent methods, are, however, not particularly satisfactory. First, in order to solve Eq. (23) of Ref. 1, we had to replace the differentials by finite differences which leads to difficulties in a computational application. Second, the method was designed exclusively for finding a solution of the HF equations.[2] These two restrictions will be removed in this paper.

The subsequent section is devoted to the reconstruction of the method of convergent iterations in a much more general manner than in Ref. 1. This is similar to the results of Ref. 3 where not only was the treatment of orthogonality constraints included but, moreover, the numerical solution of the HF equations was carried out by means of the developed methods. The steepest descent algorithm has furthermore been extended to the multiconfiguration interaction self-consistent field procedure. The results have been given in Ref. 4.

In the remaining sections we apply the developed theory to the HF equations. Section 3 is devoted to casting the method presented in Ref. 1 into its most practical form. Here the potential is modified in such a manner that the degeneracy of the HF Hamiltonian is removed. In Sec. 4, on the other hand, it is not assumed that the HF one-particle wavefunctions form an orthonormal set so that the direct work with certain determinants and their minors is unavoidable. Details of related investigations may be found in Ref. 5. Various aspects of the theory connected with the inclusion of orthonormality conditions are treated in Sec. 5. Compared with procedures used in Ref. 1 the methods developed in these sections appear to be more efficient.

2. METHOD OF CONVERGENT ITERATIONS

Let $p_1, \ldots, p_N$ be a set of unknown quantities, with respect to which the set of equations

$$F_\alpha(p_1, \ldots, p_N) = 0, \quad \alpha = 1, \ldots, N \tag{1}$$

is to be solved. For the latter functions we shall assume that there exists such a sufficiently smooth function $W(p_1, \cdots, p_N)$ with the property $(\partial_\alpha \equiv \partial/\partial p_\alpha)$

$$F_\alpha(p_1, \ldots, p_N) = \partial_\alpha W(p_1, \ldots, p_N), \tag{2}$$

where

$$W_0 \leqslant W(p_1, \ldots, p_N) \tag{3}$$

holds for any $p_1, \ldots, p_N$. Furthermore, we introduce a continuous parameter s and postulate the following equations $(dp(s)/ds = \dot{p})$:

$$\dot{p}_\alpha = -\partial_\alpha W. \tag{4}$$

Now

$$\dot{W} = \sum_\alpha \partial_\alpha W \dot{p}_\alpha = -\sum_\alpha (\partial_\alpha W)^2 \leqslant 0, \tag{5}$$

so that $W(s)$ is a descending function of s. Making use of the inequalities (3) and (5), we obtain

$$\lim_{s \to \infty} \dot{W}(s) = \lim_{s \to \infty} \partial_\alpha W(s) = 0, \tag{6}$$

so that, in view of (4),

$$\lim_{s \to \infty} p_\alpha(s) = \bar{p}_\alpha, \tag{7}$$

$\bar{p}_\alpha$ being the solution of (1). It is obvious that Eqs. (4) are not the most general ones. Consider, e.g., such a situation where W has to be supplemented by equations of constraint

$$K_{\beta'}(p_1, \ldots, p_N) = G_\beta(p_1, \ldots, p_N) = G_\beta^*(p_1, \ldots, p_N) = 0,$$
$$\beta' = 1, \ldots, M', \beta = 1, \ldots, M, \tag{8}$$

in order to arrive at (1). (The asterisk refers to the complex conjugate.) We have divided the constraints into the real K_β and complex G_β groups.

In this case, Eq. (4) has to be generalized and we propose the following form $(\partial_\beta^* \equiv \partial/\partial p_\beta^*)$:

$$\dot{p}_\alpha = -\left(\partial_\alpha^* W - \sum_\beta \mu_\beta \partial_\alpha^* K_\beta - \sum_\beta \lambda_\beta \partial_\alpha^* G_\beta - \sum_\beta \bar{\lambda}_\beta \partial_\alpha^* G_\beta^*\right), \tag{9}$$

$$\dot{p}_\alpha^* = -\left(\partial_\alpha W - \sum_\beta \mu_\beta \partial_\alpha K_\beta - \sum_\beta \lambda_\beta \partial_\alpha G_\beta - \sum_\beta \bar{\lambda}_\beta \partial_\alpha G_\beta^*\right). \tag{10}$$

Our aim is to show that these equations lead also to the conditions (7). In the first step, the μ_β, λ_β, and $\bar{\lambda}_\beta$ have to be defined. To do this we shall require that (8) be fulfilled for each value of s. Thus, e.g.,

$$0 = \dot{G}_\beta = \sum_\alpha \partial_\alpha G_\beta \dot{p}_\alpha + \sum_\alpha \partial_\alpha^* G_\beta \dot{p}_\alpha^*$$

$$= -\sum_\alpha \partial_\alpha G_\beta \left(\partial_\alpha^* W - \sum_\gamma \mu_\gamma \partial_\alpha^* K_\gamma - \sum_\gamma \lambda_\gamma \partial_\alpha^* G_\gamma - \sum_\gamma \bar{\lambda}_\gamma \partial_\alpha^* G_\gamma^*\right)$$

$$- \sum_\alpha \partial_\alpha^* G_\beta \left(\partial_\alpha W - \sum_\gamma \mu_\gamma \partial_\alpha K_\gamma - \sum_\gamma \lambda_\gamma \partial_\alpha G_\gamma - \sum_\gamma \bar{\lambda}_\gamma \partial_\alpha G_\gamma^*\right) \tag{11}$$

and the corresponding equations for $\dot{G}_\beta^*$ and $\dot{K}_\beta$. Consequently, from

$$\sum_\gamma \mu_\gamma \left\{\sum_\alpha \partial_\alpha K_\beta \partial_\alpha^* K_\gamma + \sum_\alpha \partial_\alpha^* K_\beta \partial_\alpha K_\gamma\right\} + \sum_\gamma \lambda_\gamma \left\{\sum_\alpha \partial_\alpha K_\beta \partial_\alpha^* G_\gamma\right.$$

$$+\sum_{\alpha}\partial_{\alpha}^{*}K_{\beta}\partial_{\alpha}G_{\gamma}\Big\} +\sum_{\gamma}\overline{\lambda}_{\gamma}\Big\{\sum_{\alpha}\partial_{\alpha}K_{\beta}\partial_{\alpha}^{*}G_{\gamma}^{*}+\sum_{\alpha}\partial_{\alpha}^{*}K_{\beta}\partial_{\alpha}G_{\gamma}^{*}\Big\}$$

$$=\sum_{\alpha}\partial_{\alpha}K_{\beta}\partial_{\alpha}^{*}W+\sum_{\alpha}\partial_{\alpha}^{*}K_{\beta}\partial_{\alpha}W, \tag{12}$$

$$\sum_{\gamma}\mu_{\gamma}\Big\{\sum_{\alpha}\partial_{\alpha}G_{\beta}\partial_{\alpha}^{*}K_{\gamma}+\sum_{\alpha}\partial_{\alpha}^{*}G_{\beta}\partial_{\alpha}K_{\gamma}\Big\}+\sum_{\gamma}\lambda_{\gamma}\Big\{\sum_{\alpha}\partial_{\alpha}G_{\beta}\partial_{\alpha}^{*}G_{\gamma}$$

$$+\sum_{\alpha}\partial_{\alpha}^{*}G_{\beta}\partial_{\alpha}G_{\gamma}\Big\}+\sum_{\gamma}\overline{\lambda}_{\gamma}\Big\{\sum_{\alpha}\partial_{\alpha}G_{\beta}\partial_{\alpha}^{*}G_{\gamma}^{*}+\sum_{\alpha}\partial_{\alpha}^{*}G_{\beta}\partial_{\alpha}G_{\gamma}^{*}\Big\}$$

$$=\sum_{\alpha}\partial_{\alpha}G_{\beta}\partial_{\alpha}^{*}W+\sum_{\alpha}\partial_{\alpha}^{*}G_{\beta}\partial_{\alpha}W, \tag{13a}$$

$$\sum_{\gamma}\mu_{\gamma}\Big\{\sum_{\alpha}\partial_{\alpha}G_{\beta}^{*}\partial_{\alpha}^{*}K_{\gamma}+\sum_{\alpha}\partial_{\alpha}^{*}G_{\beta}^{*}\partial_{\alpha}K_{\gamma}\Big\}+\sum_{\gamma}\lambda_{\gamma}\Big\{\sum_{\alpha}\partial_{\alpha}G_{\beta}^{*}\partial_{\alpha}^{*}G_{\gamma}$$

$$+\sum_{\alpha}\partial_{\alpha}^{*}G_{\beta}^{*}\partial_{\alpha}G_{\gamma}\Big\}+\sum_{\gamma}\overline{\lambda}_{\gamma}\Big\{\sum_{\alpha}\partial_{\alpha}G_{\beta}^{*}\partial_{\alpha}^{*}G_{\gamma}^{*}+\sum_{\alpha}\partial_{\alpha}^{*}G_{\beta}^{*}\partial_{\alpha}G_{\gamma}^{*}\Big\}$$

$$=\sum_{\alpha}\partial_{\alpha}G_{\beta}^{*}\partial_{\alpha}^{*}W+\sum_{\alpha}\partial_{\alpha}^{*}G_{\beta}^{*}\partial_{\alpha}W, \tag{13b}$$

we first see that

$$\overline{\lambda}_{\gamma}=\lambda_{\gamma}^{*}. \tag{14}$$

Now Eqs. (12) and (13) can be solved with respect to μ and λ, whereby the required relation between μ's, λ's, and p's is established. In analogy with (5) we calculate

$$\dot{W}=\sum_{\alpha}\partial_{\alpha}W\dot{p}_{\alpha}+\sum_{\alpha}\partial_{\alpha}^{*}W\dot{p}_{\alpha}^{*}$$

$$=-2\sum_{\alpha}\Big\{\partial_{\alpha}W-\sum_{\beta}\mu_{\beta}\partial_{\alpha}K_{\beta}-\sum_{\beta}\lambda_{\beta}\partial_{\alpha}G_{\beta}-\sum_{\beta}\lambda_{\beta}^{*}\partial_{\alpha}G_{\beta}^{*}\Big\}$$

$$\times\Big\{\partial_{\alpha}^{*}W-\sum_{\beta}\mu_{\beta}\partial_{\alpha}^{*}K_{\beta}-\sum_{\beta}\lambda_{\beta}\partial_{\alpha}^{*}G_{\beta}-\sum_{\beta}\lambda_{\beta}^{*}\partial_{\alpha}^{*}G_{\beta}^{*}\Big\}\le 0. \tag{15}$$

In deriving (15) we have used Eqs. (12) and (13). Instead of (5) we thus obtain

$$\lim_{s\to\infty}\dot{W}(s)=\lim_{s\to\infty}\Big\{\partial_{\alpha}W-\sum_{\beta}\mu_{\beta}\partial_{\alpha}K_{\beta}-\sum_{\beta}\lambda_{\beta}\partial_{\alpha}G_{\beta}-\sum_{\beta}\lambda_{\beta}^{*}\partial_{\alpha}G_{\beta}^{*}\Big\}=0 \tag{16}$$

and Eq. (7). The latter statement requires some explanation. Since Eqs. (9) and (10) are linear differential equations of the first order, the knowledge of, say, $p_1(0),\ldots,p_N(0)$ is sufficient for finding $p_1(s),\ldots,p_N(s)$. Further, if $p_1(0),\ldots,p_N(0)$ fulfill (8), then, by virtue of (11), $p_1(s),\ldots,p_N(s)$ also fulfill (8). Accordingly, the last equality in (16) may be looked upon as a variational formulation of (1) where we have utilized the Lagrange multiplier technique. That is why Eq. (7) holds also in this case. At this point we return to Eqs. (9) and (10) in order to show that the modification

$$\dot{p}_{\alpha}=-f(s)\Big(\partial_{\alpha}^{*}W-\sum_{\beta}\lambda_{\beta}\partial_{\alpha}^{*}G_{\beta}-\sum_{\beta}\overline{\lambda}_{\beta}\partial_{\alpha}^{*}G_{\beta}^{*}-\sum_{\beta}\mu_{\beta}\partial_{\alpha}^{*}K_{\beta}\Big), \tag{17}$$

$$\dot{p}_{\alpha}^{*}=-f(s)\Big(\partial_{\alpha}W-\sum_{\beta}\lambda_{\beta}\partial_{\alpha}G_{\beta}-\sum_{\beta}\overline{\lambda}_{\beta}\partial_{\alpha}G_{\beta}^{*}-\sum_{\beta}\mu_{\beta}\partial_{\alpha}K_{\beta}\Big) \tag{18}$$

does not affect the inequality (15). In (17) and (18) $f(s)$ represents an arbitrary but positive function of s. However, if we let

$$s=g(s') \tag{19}$$

where $g(s')$ is given by the solution of

$$\frac{dg(s')}{ds'}=\frac{1}{f[g(s')]}, \tag{20}$$

then $p_{\alpha}'(s')$ defined as

$$p_{\alpha}'(s')=p_{\alpha}(g(s')) \tag{21}$$

is a solution of Eqs. (9) and (10). Consequently, if $f(s)$ has no zeros for $s>0$, we obtain from (7)

$$\overline{p}_{\alpha}=\lim_{s\to\infty}p_{\alpha}(s)=\lim_{s'\to\infty}p_{\alpha}'(s'). \tag{22}$$

From this we see that the function $f(s)$ plays no role in solving (1) and we need no longer consider it. One of possible ways of solving (4) or (9) and (10) consists in approximating $p_{\alpha}(s)$ by the first two terms of the Taylor expansion in s,

$$p_{\alpha}(s)\to p_{\alpha}+s\overline{p}_{\alpha}, \tag{23}$$

where p_{α} is arbitrary and $\overline{p}_{\alpha}$ is given by the right-hand side of (4). Inserting (23) into W, we can minimize W with respect to s so that

$$\sum_{\alpha}\overline{p}_{\alpha}\partial_{\alpha}W]_{p_{\alpha}(s)=p_{\alpha}+s\overline{p}_{\alpha}}=0, \tag{24}$$

the left-hand side being a rational function of s if W is a rational function of p_{α}. In the latter case the solving of (24) can be converted to finding the zeros $\overline{s}$ of a certain polynominal for which mathematical methods have already been developed. If for $\overline{s}$ W becomes minimal, we replace $p_{\alpha}+\overline{s}\overline{p}_{\alpha}$ by p_{α}^{new} and repeat the whole procedure. In the following sections we shall apply in detail the method described here to the HF equations in order to construct a theory of convergent iterations.

3. CONVERGENT ITERATIONS AND THE HF THEORY WITH A MODIFIED TWO-BODY INTERACTION

Consider the quantity

$$W(\varphi_1,\ldots,\varphi_N)=\sum_{\alpha}\frac{\langle\varphi_{\alpha}|T|\varphi_{\alpha}\rangle}{\langle\varphi_{\alpha}|\varphi_{\alpha}\rangle}+\sum_{\alpha<\beta}\frac{\{\varphi_{\alpha}\varphi_{\beta}|V|\varphi_{\alpha}\varphi_{\beta}\}}{\langle\varphi_{\alpha}|\varphi_{\alpha}\rangle\langle\varphi_{\beta}|\varphi_{\beta}\rangle}, \tag{25}$$

where φ_{α}, $\alpha=1,\ldots,N$, are one-particle wavefunctions and $\{\varphi_{\alpha}\varphi_{\beta}|V|\varphi_{\alpha}\varphi_{\beta}\}$ is an antisymmetrized matrix element of V. The variation of (25) consists of the variation of the kinetic part

$$\sum_{\alpha}\langle\delta\varphi_{\alpha}|\frac{1}{\langle\varphi_{\alpha}|\varphi_{\alpha}\rangle}\Big(T-\frac{\langle\varphi_{\alpha}|T|\varphi_{\alpha}\rangle}{\langle\varphi_{\alpha}|\varphi_{\alpha}\rangle}\Big)|\varphi_{\alpha}\rangle+\text{c.c.} \tag{26}$$

and of the potential part

$$\sum_{\alpha}\langle\delta\varphi_{\alpha}|\sum_{\beta}\frac{1}{\langle\varphi_{\alpha}|\varphi_{\alpha}\rangle\langle\varphi_{\beta}|\varphi_{\beta}\rangle}\Big((\varphi_{\beta}|V|\varphi_{\beta})|\varphi_{\alpha}\rangle-(\varphi_{\beta}|V|\varphi_{\alpha})|\varphi_{\beta}\rangle$$

$$-\{\varphi_{\alpha}\varphi_{\beta}|V|\varphi_{\alpha}\varphi_{\beta}\}\frac{1}{\langle\varphi_{\alpha}|\varphi_{\alpha}\rangle}|\varphi_{\alpha}\rangle\Big)+\text{c.c.} \tag{27}$$

From these equations we thus obtain

$$T|\varphi_{\alpha}\rangle+\sum_{\beta}\frac{(\varphi_{\beta}|V|\varphi_{\beta})}{\langle\varphi_{\beta}|\varphi_{\beta}\rangle}|\varphi_{\alpha}\rangle-\sum_{\beta}\frac{(\varphi_{\beta}|V|\varphi_{\alpha})}{\langle\varphi_{\beta}|\varphi_{\beta}\rangle}|\varphi_{\beta}\rangle=\epsilon_{\alpha}|\varphi_{\alpha}\rangle \tag{28}$$

with

$$\epsilon_{\alpha}=\frac{\langle\varphi_{\alpha}|T|\varphi_{\alpha}\rangle}{\langle\varphi_{\alpha}|\varphi_{\alpha}\rangle}+\sum_{\beta}\frac{\{\varphi_{\alpha}\varphi_{\beta}|V|\varphi_{\alpha}\varphi_{\beta}\}}{\langle\varphi_{\alpha}|\varphi_{\alpha}\rangle\langle\varphi_{\beta}|\varphi_{\beta}\rangle}. \tag{29}$$

A standard calculation yields, however, that

$$(\epsilon_{\alpha}-\epsilon_{\gamma})\langle\varphi_{\gamma}|\varphi_{\alpha}\rangle=0 \tag{30}$$

which shows clearly that (28) and (29) are not the HF
equations: If $\alpha \neq \gamma$ and $\epsilon_\alpha = \epsilon_\gamma$, then $\langle \varphi_\gamma | \varphi_\alpha \rangle = 0$ need not
hold. The orthogonality of $\varphi_\alpha, \varphi_\gamma (\alpha \neq \gamma)$ may be enforced
by the replacement

$$V \to V + \delta V. \tag{31}$$

Here it is explicitly assumed that the new potential re-
moves the degeneracy in ϵ. We now make use of (26),
(27), and (4) to obtain

$$\frac{d}{ds} | \varphi_\alpha \rangle = - T | \varphi_\alpha \rangle - \sum_\beta [(\varphi_\beta | V | \varphi_\beta) | \varphi_\alpha \rangle$$

$$- (\varphi_\beta | V | \varphi_\alpha) | \varphi_\beta)] \frac{1}{\langle \varphi_\beta | \varphi_\beta \rangle} + \left(\frac{\langle \varphi_\alpha | T | \varphi_\alpha \rangle}{\langle \varphi_\alpha | \varphi_\alpha \rangle} \right.$$

$$\left. + \sum_\beta \frac{\{\varphi_\alpha \varphi_\beta | V | \varphi_\alpha \varphi_\beta\}}{\langle \varphi_\alpha | \varphi_\alpha \rangle \langle \varphi_\beta | \varphi_\beta \rangle} \right) | \varphi_\alpha \rangle. \tag{32}$$

For a numerical application of (32), $| \varphi_\alpha \rangle$ will be ex-
panded in a complete set $| \chi_i \rangle$

$$| \varphi_\alpha \rangle = \sum_i (\alpha)_i | \chi_i \rangle, \tag{33}$$

so that (32) becomes

$$(\overset{\circ}{\alpha})_i = - \sum_j T_{ij} (\alpha)_j - \sum_\beta \sum_{j k l} \bar{V}_{ij,kl} (\beta)_j^* (\alpha)_k (\beta)_l \frac{1}{\sum_m (\beta)_m^* (\beta)_m}$$

$$+ \left(\frac{\sum_{jk} T_{jk} (\alpha)_j^* (\alpha)_k}{\sum_j (\alpha)_j^* (\alpha)_j} + \sum_\beta \frac{\sum_{jklm} \bar{V}_{jk,lm} (\alpha)_j^* (\beta)_k^* (\alpha)_l (\beta)_m}{\sum_j (\alpha)_j^* (\alpha)_j \sum_k (\beta)_k^* (\beta)_k} \right) (\alpha)_i \tag{34}$$

where

$$\bar{V}_{ij,kl} = V_{ij,kl} - V_{ij,lk}. \tag{35}$$

The next step consists in performing the replacement
(23)

$$(\alpha)_i (s) \to (\bar{\alpha})_i = (\alpha)_i + (\bar{\alpha})_i s, \tag{36}$$

where $(\bar{\alpha})_i$ is given by the right-hand side of (34). The
optimal value of s is found by inserting (36) into Eq. (25)
and finding the first derivative of W with respect to s.
The result is given by

$$R_1(s) + R_2(s) = 0, \tag{37}$$

where $R_1(s)$ is the kinetic part contribution

$$R_1(s) = \sum_\alpha \frac{\sum_{ij} T_{ij} [(\bar{\alpha})_i^* (\bar{\alpha})_j + (\bar{\alpha})_i^* (\bar{\alpha})_j]}{\sum_i (\bar{\alpha})_i^* (\bar{\alpha})_i}$$

$$- \sum_\alpha \sum_{ij} T_{ij} (\bar{\alpha})_i^* (\bar{\alpha})_j \frac{\sum_k [(\bar{\alpha})_k^* (\bar{\alpha})_k + (\bar{\alpha})_k^* (\bar{\alpha})_k]}{[\sum_i (\bar{\alpha})_i^* (\bar{\alpha})_i]^2} \tag{38}$$

and $R_2(s)$ designates the potential part contribution

$$R_2(s) = \sum_{\alpha \beta} \frac{\sum_{ijkl} \bar{V}_{ij,kl} [(\bar{\alpha})_i^* (\bar{\beta})_j^* (\bar{\alpha})_k (\bar{\beta})_l + (\bar{\alpha})_i^* (\bar{\beta})_j^* (\bar{\alpha})_k (\bar{\beta})_l]}{\sum_i (\bar{\alpha})_i^* (\bar{\alpha})_i \sum_j (\bar{\beta})_j^* (\bar{\beta})_j}$$

$$- \sum_{\alpha \beta} \sum_{ijkl} \bar{V}_{ij,kl} (\bar{\alpha})_i^* (\bar{\beta})_j^* (\bar{\alpha})_k (\bar{\beta})_l \frac{\sum_m [(\bar{\alpha})_m^* (\bar{\alpha})_m + (\bar{\alpha})_m^* (\bar{\alpha})_m]}{[\sum_i (\bar{\alpha})_i^* (\bar{\alpha})_i]^2 \sum_j (\bar{\beta})_j^* (\bar{\beta})_j}. \tag{39}$$

The conversion of (37) in a polynomial is straightforward
and we obtain

$$P(s) \equiv \prod_\alpha \left[\sum_i (\bar{\alpha})_i^* (\bar{\alpha})_i \right]^2 (R_1(s) + R_2(s)) = 0. \tag{40}$$

Writing

$$P(s) = \sum_{i=1}^{4N-1} a_i s^{i-1} \tag{41}$$

we must then evaluate a_i. However, let $s_1, \ldots, s_{4N-1}$ be
nonnegative and nonequal numbers; then we obtain[6]

$$a_j = \sum_{r=0}^{4N-j-1} S_{j+r}^{(4N-1)} (s_1, \ldots, s_{4N-1}) (-1)^r \sum_{i=1}^{4N-1} (-1)^{i+j}$$

$$\times \frac{1}{\prod_{k=1}^{i-1} (s_i - s_k) \prod_{k=i+1}^{4N-1} (s_k - s_i)} s_i^r P(s_i), \tag{42}$$

where $S_p^{(n)}(s_1, \ldots, s_n)$ are the generalized Stirling
numbers of the first kind defined by

$$S_p^{(n+1)} (s_1, \ldots, s_{n+1}) = S_{p-1}^{(n)} (s_1, \ldots, s_n) + s_{n+1} S_p^{(n)} (s_1, \ldots, s_n),$$

$$n \geq 0, \tag{43}$$

with

$$S_{-1}^{(n)} (s_1, \ldots, s_n) = S_{n+1}^{(n)} (s_1, \ldots, s_n) = S_0^{(0)} - 1 = 0. \tag{44}$$

Since a_j are now known, we can apply the standard
numerical methods to find that zero $\bar{s}$ of $P(s)$ for which
$W(s)$ becomes minimal. Replacing

$$(\alpha)_i \to (\alpha)_i + (\bar{\alpha})_i \bar{s} (\equiv (\alpha)_i^{\text{new}}), \tag{45}$$

we repeat the whole procedure beginning with the right-
hand side of (34) which is now the new $(\bar{\alpha})_i$. It is obvious
that the new $(\alpha)_i$ approximate the HF solutions better
than the old ones. An infinite repetition of these itera-
tions yields the HF solutions.

4. CONVERGENT ITERATIONS AND THE HF THEORY WITHOUT THE MODIFIED INTERACTION

The work of the previous section has been based on
(31) which enforced the orthogonality of φ's. We shall
now show that (31) need not be taken into account if,
instead of Eq. (25),

$$W = \langle \Psi | H | \Psi \rangle / \langle \Psi | \Psi \rangle, \tag{46}$$

where $| \Psi \rangle$ is a determinant of unnormalized and non-
orthogonal one-particle wavefunctions multiplied be the
factor $1/\sqrt{N!}$, which in this case has nothing to do with
the normalization. As it will be seen in the following,
we shall be able to convert the problem of solving Eq.
(4) into the problem of repeatedly finding zeros of
another polynomial. The structure of this procedure is,
however, now much more complicated and accordingly
restricted to only small number of particles. For
evaluation of (46) we need the kinetic part

$$\langle \Psi | T | \Psi \rangle = \sum_{\alpha \beta} \langle \varphi_\alpha | T | \varphi_\beta \rangle M_{\alpha,\beta}^{(1)} (-1)^{\alpha+\beta} \tag{47}$$

and the potential part

$$\langle \Psi | V | \Psi \rangle = \sum_{\alpha < \beta} \sum_{\gamma < \delta} \{\varphi_\alpha \varphi_\beta | V | \varphi_\gamma \varphi_\delta\} M_{\alpha\beta,\gamma\delta}^{(2)} (-1)^{\alpha+\beta+\gamma+\delta}, \tag{48}$$

where $M_{\alpha,\beta}^{(1)}$ designates the minor, assigned to the α-row
and β-column, of the first rank of the matrix $\langle \varphi_\alpha | \varphi_\beta \rangle$,
whereas $M_{\alpha\beta,\gamma\delta}^{(2)}$ assigned to the α, β rows and γ, δ
columns is a minor of the second rank of the same
matrix. Similarly, one can introduce the minor of the
third rank. The following relations are of importance
for our future work:

$$M_{\alpha,\beta}^{(1)} = \sum_\delta (-1)^{\gamma+\delta} S[\gamma \alpha] S[\delta \beta] M_{\alpha\gamma,\beta\delta}^{(2)} \langle \varphi_\gamma | \varphi_\delta \rangle, \tag{49}$$

$$\delta M^{(1)}_{\alpha,\beta} = \sum_{\gamma,\delta} (-1)^{\gamma+\delta} S[\gamma\alpha] S[\delta\beta] M^{(2)}_{\alpha\gamma,\beta\delta} \delta \langle \varphi_\gamma \mid \varphi_\delta \rangle \qquad (50)$$

and

$$M_{\alpha\beta,\gamma\delta} = \sum_{\xi} (-1)^{\epsilon+\xi} S[\epsilon\alpha] S[\epsilon\beta] S[\xi\gamma] S[\xi\delta] M^{(3)}_{\alpha\beta\epsilon,\gamma\delta\xi} \langle \varphi_\epsilon \mid \varphi_\xi \rangle, \qquad (51)$$

$$\delta M^{(2)}_{\alpha\beta,\gamma\delta} = \sum_{\epsilon,\xi} (-1)^{\epsilon+\xi} S[\epsilon\alpha] S[\epsilon\beta] S[\xi\gamma] S[\xi\delta] M^{(3)}_{\alpha\beta\epsilon,\gamma\delta\xi} \delta \langle \varphi_\epsilon \mid \varphi_\xi \rangle, \qquad (52)$$

where

$$S[\alpha\beta] = \begin{cases} 1 & \alpha < \beta \\ 0 & \alpha = \beta \\ -1 & \alpha > \beta. \end{cases} \qquad (53)$$

We also introduce the minor $M^{(0)}$ of the zeroth rank which is simply the determinant of the matrix $\langle \varphi_\alpha \mid \varphi_\beta \rangle$, so that

$$M^{(0)} = \sum_{\beta} (-1)^{\alpha+\beta} M^{(1)}_{\alpha,\beta} \langle \varphi_\alpha \mid \varphi_\beta \rangle, \qquad (54)$$

$$\delta M^{(0)} = \sum_{\alpha,\beta} (-1)^{\alpha+\beta} M^{(1)}_{\alpha,\beta} \delta \langle \varphi_\alpha \mid \varphi_\beta \rangle. \qquad (55)$$

According to Eq. (4) we need the expressions

$$(\partial/\partial \langle \varphi_\alpha \mid) \langle \Psi \mid T \mid \Psi \rangle$$
$$= \sum_{\beta} [T \mid \varphi_\beta \rangle M^{(1)}_{\alpha,\beta} (-1)^{\alpha+\beta}$$
$$+ \sum_{\beta\gamma\delta} \langle \varphi_\gamma \mid T \mid \varphi_\beta \rangle (-1)^{\alpha+\beta+\gamma+\delta} S[\gamma\alpha] S[\beta\delta] M^{(2)}_{\alpha\gamma,\beta\delta} \mid \varphi_\delta \rangle] \qquad (56)$$

and

$$(\partial/\partial \langle \varphi_\alpha \mid) \langle \Psi \mid V \mid \Psi \rangle$$
$$= \sum_{\beta\gamma\delta} (\varphi_\beta \mid V \mid \varphi_\gamma) \mid \varphi_\delta \rangle M^{(2)}_{\alpha\beta,\gamma\delta} (-1)^{\alpha+\beta+\gamma+\delta} S[\alpha\beta] S[\gamma\delta]$$
$$+ \sum_{\beta<\gamma} \sum_{\delta<\epsilon} \sum_{\xi} (-1)^{\alpha+\beta+\gamma+\delta+\epsilon+\xi} \{\varphi_\beta\varphi_\gamma \mid V \mid \varphi_\delta\varphi_\epsilon\}$$
$$\times S[\alpha \mid \beta\gamma] S[\xi \mid \delta\epsilon] M^{(3)}_{\alpha\beta\gamma,\delta\epsilon\xi} \mid \varphi_\xi \rangle, \qquad (57)$$

where, for $\beta < \gamma$,

$$S[\alpha \mid \beta\gamma] = \begin{cases} 1 & \alpha < \beta \\ -1 & \beta < \alpha < \gamma \equiv S[\alpha\beta] S[\alpha\gamma], \\ 1 & \gamma < \alpha \end{cases} \qquad (58)$$

so that we may write

$$\frac{d}{ds} \mid \varphi_\alpha \rangle = -\{R(56) + R(57)\}/M^{(0)}$$
$$+ \{R(47) + R(48)\} R(54)/(M^{(0)})^2, \qquad (59)$$

$R(56)$ denoting the right-hand side of (56), etc. We employ now Eq. (33) and obtain

$$(\overset{\circ}{\alpha})_i = (\tilde{\alpha})_i \qquad (60)$$

with

$$(\tilde{\alpha})_i = -\left\{ \sum_{\beta} \sum_{j} T_{ij}(\beta)_j M^{(1)}_{\alpha,\beta} (-1)^{\alpha+\beta} + \sum_{\beta\gamma\delta} \sum_{jk} T_{jk}(\beta)^*_j (\gamma)_k M^{(2)}_{\alpha\beta,\gamma\delta} \right.$$
$$\times S[\beta\alpha] S[\gamma\delta] (-1)^{\alpha+\beta+\gamma+\delta} + \sum_{\beta} \sum_{\gamma<\delta} \sum_{jkl} \overline{V}_{ij,kl}(\beta)^*_j$$
$$\times (\delta)^*_k (\gamma)_l M^{(2)}_{\alpha\beta,\gamma\delta} S[\alpha\beta] (-1)^{\alpha+\beta+\gamma+\delta} + \sum_{\beta<\gamma} \sum_{\delta<\epsilon} \sum_{\xi} (-1)^{\alpha+\beta+\gamma+\delta+\epsilon+\xi}$$
$$\left. \times \sum_{klmn} \overline{V}_{kl,mn}(\beta)^*_k (\gamma)^*_l (\delta)_m (\epsilon)_n M^{(3)}_{\alpha\beta\gamma,\delta\epsilon\xi} S[\alpha \mid \beta\gamma] S[\xi \mid \delta\epsilon] (\xi)_i \right\}$$
$$\Big/ M^{(0)} + \left\{ \sum_{\beta\gamma} \sum_{kl} T_{kl}(\beta)^*_k (\gamma)_l M^{(1)}_{\beta,\gamma} (-1)^{\beta+\gamma} \right.$$
$$\left. + \sum_{\beta<\gamma} \sum_{\delta<\epsilon} \sum_{klmn} \overline{V}_{kl,mn}(\beta)^*_k (\gamma)^*_l (\delta)_m (\epsilon)_n M^{(2)}_{\beta\gamma,\delta\epsilon} (-1)^{\beta+\gamma+\delta+\epsilon} \right\}$$
$$\times \left[\sum_{\beta} (-1)^{\alpha+\beta} M^{(1)}_{\alpha,\beta}(\beta)_i \right] \Big/ (M^{(0)})^2. \qquad (61)$$

In analogy with (36) we introduce the new quantity $(\tilde{\alpha})_i$ where $(\tilde{\alpha})_i$ is defined by Eq. (61). Thus W becomes a function of s of the type

$$W(s) = \left(\sum_{i=0}^{2N} a_i s^i \right) \Big/ \left(\sum_{i=0}^{2N} b_i s^i \right), \qquad (62)$$

where the polynomial in the numerator of (62) is given by

$$\sum_{\beta\gamma} \sum_{kl} T_{kl}(\beta)^*_k (\tilde{\gamma})_l M^{(1)}_{\beta,\gamma} (-1)^{\beta+\gamma}$$
$$+ \sum_{\beta<\gamma} \sum_{\delta<\epsilon} \sum_{klmn} \overline{V}_{kl,mn}(\beta)^*_k (\tilde{\gamma})^*_l (\tilde{\delta})_m (\tilde{\epsilon})_n M^{(2)}_{\beta\gamma,\delta\epsilon} (-1)^{\beta+\gamma+\delta+\epsilon}, \qquad (63)$$

whereas the polynomial in the denominator of (62) is simply the determinant of a matrix with the matrix elements

$$\sum_i (\tilde{\alpha})^*_i (\tilde{\beta})_i \qquad (64)$$

which thus also depends on s. This statements holds also for the minors occurring in (63). It must not be forgotten that the coefficients a_i and b_i in (62) may be calculated in the same way as it was proposed in (42). The optimal value of s is found by minimizing $W(s)$, which, in view of Eq. (62), leads to the problem of finding the zeros of the polynomial

$$\sum_{k=0}^{4N-1} \left\{ \sum_{i=1}^{k+1} (a_i b_{k+1-i} - a_{k+1-i} b_i) i \right\} s^k = 0. \qquad (65)$$

We choose now that zero $\bar{s}$ of (65) at which $W(s)$ has its global minimum so that the new $(\alpha)_i$ is given by $(\alpha)_i + \bar{s}(\tilde{\alpha})_i$ by means of which the new $(\tilde{\alpha})_i$ [see Eq. (61)] are found and the whole procedure is repeated. The limiting value of $(\alpha)_i$ is obviously identical with the corresponding $(\alpha)^{HF}_i$.

The advantages and disadvantages of the method developed in this section can be briefly summarized as follows. On the one hand, we have avoided (31) so that the direct work with V was made possible. Further, the new values of $(\alpha)_i$ originating from each iteration step have been obtained by solving an algebraic equation (65), which as we have already stated, is a mathematically well-defined procedure. The main disadvantage, however, consists in the necessity of repeated calculation of the minors $M^{(0)}$, $M^{(1)}$, $M^{(2)}$, and $M^{(3)}$ so that this procedure is restricted only to a small number of particles.

5. CONVERGENT ITERATIONS AND THE HF THEORY WITH SCHMIDT ORTHOGONALIZATION PROCEDURE

Until now we have exclusively used Eq. (4). But it is obvious that if we put

$$\sum_{\alpha} \langle \varphi_\alpha \mid T \mid \varphi_\alpha \rangle + \sum_{\alpha<\beta} \{\varphi_\alpha\varphi_\beta \mid V \mid \varphi_\alpha\varphi_\beta\} \qquad (66)$$

with the constraints

$$\langle \varphi_\alpha \mid \varphi_\beta \rangle = \delta_{\alpha\beta}, \quad \alpha, \beta = 1, \ldots, N, \qquad (67)$$

we also arrive at the HF equations where Eqs. (9) and (10) may now be employed. Identifying $\langle \varphi_\alpha \mid \varphi_\beta \rangle$, $\alpha < \beta$,

with $G_{\alpha\beta}$ and $\langle\varphi_\alpha|\varphi_\beta\rangle$, $\alpha>\beta$, with $G^*_{\alpha\beta}$ and $\langle\varphi_\alpha|\varphi_\alpha\rangle$ with K_α, we can solve Eqs. (12) and (13) with respect to λ's and μ's and obtain ($\lambda_{\alpha\alpha}=\mu_\alpha$)

$$\lambda_{\alpha\beta}=\langle\varphi_\beta|T|\varphi_\alpha\rangle+\sum_\gamma\{\varphi_\beta\varphi_\gamma|V|\varphi_\alpha\varphi_\gamma\}. \tag{68}$$

This inserted into Eq. (9) yields

$$\frac{d}{ds}|\varphi_\alpha\rangle=-T|\varphi_\alpha\rangle-\sum_\beta[(\varphi_\beta|V|\varphi_\beta)|\varphi_\alpha\rangle-(\varphi_\beta|V|\varphi_\alpha)|\varphi_\beta\rangle]$$

$$+\sum_\nu(\langle\varphi_\nu|T|\varphi_\alpha\rangle+\sum_\gamma\{\varphi_\nu\varphi_\gamma|V|\varphi_\alpha\varphi_\gamma\})|\varphi_\nu\rangle, \tag{69}$$

or using Eq. (33).

$$(\dot\alpha)_i=-\sum_j T_{ij}(\alpha)_j-\sum_\beta\sum_{jkl}\bar V_{ij,kl}(\beta)^*_j(\alpha)_k(\beta)_l$$

$$+\sum_\nu\left(\sum_{jk}(\nu)^*_j T_{jk}(\alpha)_k+\sum_\gamma\sum_{klmn}\bar V_{kl,mn}(\nu)^*_k(\gamma)^*_l(\alpha)_m(\gamma)_n\right)(\nu)_i. \tag{70}$$

Designating the right-hand side of this equation as $(\bar\alpha)_i$, we can again define the new quantity $(\tilde\alpha)_i$ according to the prescription (36). However, $(\tilde\alpha)_i$ do not fulfill the constraints (67) and we shall enforce them by applying the Schmidt orthogonalization procedure

$$[\alpha]_i=(\tilde{\tilde\alpha})_i\Big/\left\{\sum_j(\tilde{\tilde\alpha})_j(\tilde{\tilde\alpha})_j\right\}^{1/2}, \quad \alpha\ge 1, \tag{71}$$

where

$$(\tilde{\tilde\alpha})_i=-\sum_{\beta=1}^{\alpha-1}\sum_j[\beta]^*_j(\tilde\alpha)_j[\beta]_i+(\tilde\alpha)_i, \quad \alpha>1, \tag{72}$$

and

$$(\tilde 1)_i=(\tilde 1)_i. \tag{73}$$

However, we must prove that

$$[\beta]_i-(\beta)_i]_{s=0}=[\dot\beta]_i-(\dot\beta)_i]_{s=0}=0 \tag{74}$$

is satisfied for the new coefficients, because in this case $\mathring W(0)<0$ and $\tilde W(s)$ has a minimum for $s>0$. Here $\tilde W(s)$ stands for $W([1]_1,\cdots,[N]_n)$. The first equality in (74) is trivial. In order to prove the second one, we note that

$$\widetilde{(\alpha+1)}_i]_{s=0}=-\sum_{\beta=1}^\alpha\sum_j\{(\bar\beta)^*_j\widetilde{(\alpha+1)}_j(\tilde\beta)_i+(\tilde\beta)^*_j\widetilde{(\alpha+1)}_j(\bar\beta)_i$$

$$+(\bar\beta)^*_j\widetilde{(\alpha+1)}_j(\bar\beta)_i\}+\widetilde{(\alpha+1)}_i]_{s=0}=\widetilde{(\alpha+1)}_i]_{s=0}, \tag{75}$$

where we have assumed that $(\tilde{\dot\alpha})_i=(\tilde\alpha)_i]_{s=0}$ for $\alpha<\beta$. Taking Eq. (73) into account, we see that $(\tilde{\dot\gamma})_i]_{s=0}=(\bar\alpha)_i]_{s=0}$ holds by induction. Thus, by Eq. (71), we obtain

$$[\dot\alpha]_i=(\bar\alpha)_i-\tfrac{1}{2}(\alpha)_i\sum_j[(\tilde{\tilde\alpha})^*_j(\tilde\alpha)_j+(\tilde\alpha)^*_j(\tilde{\tilde\alpha})_j]]_{s=0}=(\bar\alpha)_i, \tag{76}$$

which proves that the second equality in (74) is also valid.

We are interested in finding the minimum of $\tilde W(s)$. For this we need $\mathring W$,

$$\mathring W=\sum_\alpha\sum_{ij}T_{ij}([\dot\alpha]^*_i[\alpha]_j+[\alpha]^*_i[\dot\alpha]_j)$$

$$+\sum_{\alpha\beta}\sum_{ijkl}\bar V_{ij,kl}([\dot\alpha]^*_i[\beta]^*_j[\alpha]_k[\beta]_l+[\alpha]^*_i[\beta]^*_j[\dot\alpha]_k[\beta]_l), \tag{77}$$

where, due to (71), $[\dot\alpha]_i$ is given by

$$[\dot\alpha]_i=(\tilde{\dot{\tilde\alpha}})_i\Big/\left\{\sum_j(\tilde{\tilde\alpha})^*_j(\tilde{\tilde\alpha})_j\right\}^{1/2}$$

$$-\tfrac{1}{2}(\tilde{\tilde\alpha})_i\sum_j[(\tilde{\dot{\tilde\alpha}})^*_j(\tilde{\tilde\alpha})_j+(\tilde{\tilde\alpha})^*_j(\tilde{\dot{\tilde\alpha}})_j]\Big/\left\{\sum_j(\tilde{\tilde\alpha})_j{}^*(\tilde{\tilde\alpha})_j\right\}^{3/2}, \tag{78}$$

$(\tilde{\dot{\tilde\alpha}})_i$ being generated by the recurrence formula [see Eq. (72)]

$$\alpha>1; \ (\tilde{\dot{\tilde\alpha}})_i=-\sum_{\beta=1}^{\alpha-1}\sum_j\left\{[\dot\beta]^*_j(\bar\alpha)_i[\beta]_i+[\beta]^*_j(\bar\alpha)_j[\beta]_i\right.$$

$$\left.+[\beta]^*_j(\bar\alpha)_j[\dot\beta]_i\right\}+(\bar\alpha)_i \tag{79}$$

with

$$(\tilde{\dot{\tilde 1}})_i=(\tilde 1)_i. \tag{80}$$

The function of W in the procedure developed in this section can be best seen if the iteration algorithm is presented in detail. We start with an arbitrary set of numbers $(\alpha)_i$ with the constraints

$$\sum_i(\alpha)^*_i(\beta)_i=\delta_{\alpha\beta}, \quad \alpha,\beta=1,\dots,N, \tag{81}$$

which are a trivial consequence of Eq. (67). Then the right-hand side of (70) [denoted by $(\bar\alpha)_i$] is evaluated for each α and i. We, further, choose an arbitrary $s>0$ and calculate $(\tilde\alpha)_i=(\alpha)_i+s(\bar\alpha)_i$. Using the relations (71), (73), and (78)—(80), we obtain $[\alpha]_i$ and $[\dot\alpha]_i$ so that $\mathring W$ can also be evaluated. If the latter quantity is negative, another value of s has to be tried until $\mathring W$ becomes positive. Now, let $\mathring W$ be positive, then there exists $\bar s$ such that $0<\bar s<s$ and for which $\tilde W(\bar s)=0$ (there exist sufficiently efficient methods in mathematics for finding this $\bar s$). The last check consists in proving the inequality $\tilde W(\bar s)<\tilde W(0)$ which states that we have found the local minimum of $\tilde W(s)$, the existence of which is guaranteed from our previous considerations. The new $(\alpha)_i$ are defined by the transformation

$$[\alpha]_i]_{s=\bar s}\to(\alpha)_i, \tag{82}$$

where these $(\alpha)_i$ obviously fulfill (81) and one can use them as initial values of another iteration step. Since this algorithm guarantees the convergence of the iterations we obtain $(\alpha)_i\to(\alpha)^{HF}_i$.

6. SUMMARY

The theory of convergent iterations developed in the second section has been applied in detail to the HF equations and we shall briefly compare the various versions presented in Secs. 3, 4, and 5:

(i) Sec. 3: The theory takes on a particularly simple form, however, the two-body interaction has to be modified in order to enforce the orthogonality of φ's. The starting values of a new iteration step are obtained by means of zeros of the polynomial (40).

(ii) Sec. 4: The latter property holds also in this case where the corresponding polynomial is given by Eq. (65).

No. modification of the two-body interaction need be done, but now a number of minors have to be evaluated so that the procedure is limited to a small number of particles.

(iii) Sec. 5: This version of the method makes direct use of the constraints (67) so that the introduction of minors is completely avoided. Since $[\alpha]_i$ are no longer rational functions of s the problem of minimizing W cannot be converted to the problem of finding zeros of a polynomial. Thus a standard but lengthy algorithm (71)—

(73) and (77)—(81) has to be applied.

[1]A. Kelemen, Nucl. Phys. A 220, 391 (1974).
[2]G. Ripka, Lect. Theor. Phys. VIIIC, 237 (1965); G. Ripka, *Advances in Nuclear Physics*, edited by M. Baranger and E. Vogt (Plenum, New York, 1968), Vol. 1, p. 183.
[3]R. McWeeny, Proc. Roy. Soc. (Lond.) A 235, 496 (1956).
[4]R. McWeeny and B.T. Sutcliffe, *Methods of Molecular Quantum Mechanics* (Academic, New York, 1969), p. 145 ff.
[5]P.O. Löwdin, Phys. Rev. 97, 1474 (1955).
[6]A. Kelemen (unpublished).

1401   J. Math. Phys., Vol. 16, No. 7, July 1975

A. Kelemen   1401

Construction of a meromorphic many-channel *p*-wave *S* matrix

Joseph R. Cox and Hugo R. Garcia*

Department of Physics, Florida Atlantic University, Boca Raton, Florida 33432
(Received 26 February 1975)

The simplest possible nonrelativistic many-channel *p*-wave *S* matrix which is meromorphic on its energy Riemann surface, as well as the underlying potential matrix, is explicitly constructed by means of the many-channel Marchenko equations. The results suggest that, contrary to the case in which no coupling between channels is present, such an *S* matrix necessarily increases in complexity with increasing angular momentum.

1. INTRODUCTION

The many-channel problem[1] is an approximation to the nonrelativistic quantum mechanical inelastic scattering problem. The approximation consists of assuming that the target cannot be broken up by the oncoming projectile. Thus, all possible initial and final states (i. e. , channels) of the system consist of the projectile plus the target in the ground state or in any one of a finite number of discrete excited states. (A slight modification of the approximation also provides a crude model for rearrangement processes where the channel reduced mass undergoes a change.) Although a fairly satisfactory understanding of the scattering certainly requires a more sophisticated (and complicated) approach which takes into account the continuum states of the target, such as the one based on the Faddeev equations,[2] the many-channel approximation is, nevertheless, often quite reasonable when the energy is far below the target breakup threshold, and it has furnished considerable insight into the many-particle scattering processes in atomic, nuclear, and particle physics. (See, for example, Refs. 1, 3, and 4.)

Our main purpose in this paper is to construct explicitly the simplest *p*-wave many-channel *S* matrix (for spinless particles) which is meromorphic on its energy Riemann surface, and also to construct the corresponding (time reversal invariant) potential matrix.

Sections 2—4 are introductory in nature. In Sec. 2 the coupled radial Schrödinger equations, the Jost function, and the *S* matrix for arbitrary angular momentum are introduced. Section 3 consists of a brief discussion of the many-channel Marchenko equations which provide a means of solving the inverse problem at fixed angular momentum. Section 4 reviews the Fredholm determinant and its relationship to the *S* matrix, bound states, and resonances of the many-channel system, and briefly discusses the Bargmann potentials. Finally, Secs. 5 and 6 are devoted to the construction, by means of the Marchenko formalism, of a meromorphic *p*-wave *S* matrix together with the associated potential matrix which describes the interparticle forces.

The methods employed in this paper have close counterparts, in the uncoupling limit, to those recently employed in several other areas of physics. These areas include the construction of soliton solutions of the (nonlinear) steady-state Korteweg—de Vries equation in plasma physics,[5,6] and the construction of coherent optical pulse profiles which exhibit lossless propagation during transmission through a resonant atomic medium.[7] In addition, coupled equations related to the two-channel problem have been obtained in connection with the problem of diffraction of an elastic wave by a spherical area included in an infinite homogeneous medium.[8]

2. PRELIMINARIES

The basic radial integral equations of the conventional many-channel inelastic scattering theory[1] may be expressed in matrix notation as

$$\psi_l(K, r) = u_l(K, r) + \int_0^\infty dr' G_l(K; r, r') V(r') \psi_l(K, r').$$
(2.1)

Here $l = 0, 1, \cdots$, $\psi_l(K, r)$ is an $n \times n$ matrix, where n is the number of channels. K is the $n \times n$ diagonal matrix of the channel wavenumbers $k_1, \ldots, k_n$:

$$K_{ij} = k_i \delta_{ij}, \quad i, j = 1, \ldots, n.$$
(2.2)

Conservation of energy $\hbar^2 k_1^2 / 2\mu$ of the system is expressed by

$$k_1^2 = k_j^2 + \Delta_j^2, \quad j = 1, \ldots, n,$$
(2.3)

where μ is the (common) channel reduced mass and $\hbar^2 \Delta_j^2 / 2\mu$, $\Delta_j > 0$, is the threshold energy of the jth channel. The Green's function $G_l(K; r, r')$ in (2.1) is the diagonal $n \times n$ matrix

$$G_l(K; r, r') = (-)^{l+1} K^{-1} u_l(Kr_<) w_l(Kr_>),$$
(2.4)

where u_l and w_l are spherical Bessel functions:

$$u_l(z) = (\pi z / 2)^{1/2} J_{l+1/2}(z),$$

$$w_l(z) = i \exp(i\pi l)(\pi z / 2)^{1/2} H_{l+1/2}(z).$$
(2.5)

$V(r)$ in (2.1) is the $n \times n$ symmetric potential matrix; it will be restricted here to depend only on r and to be "well-behaved"—i. e. , all elements are not as singular as r^{-2} as $r \to 0$, all diagonal elements vanish more rapidly than r^{-2} as $r \to \infty$, and all off-diagonal elements vanish more rapidly than any inverse power of r as $r \to \infty$.

Equation (2.1) implies that $\psi_l(K, r)$ satisfies the coupled radial Schrödinger equations

$$-\psi_l''(K, r) + \left(V(r) + \frac{l(l+1)}{r^2} \right) \psi_l(K, r) = K^2 \psi_l(K, r).$$
(2.6)

We define the $n \times n$ matrix $F_l(K, r)$ as that solution of (2.6) which obeys the boundary condition

$$\lim_{r \to \infty} \exp(-iKr) F_l(K, r) = 1. \tag{2.7}$$

Then the Jost matrix[1] $\mathfrak{L}_l(K)$ is given by

$$\tilde{\mathfrak{L}}_l(K) = (-iK)^l [(2l-1)!!]^{-1} \lim_{r \to 0} r^l F_l(K, r), \tag{2.8}$$

where the tilde denotes matrix transpose, and the S matrix is given by the open-channel submatrix of $S_l(K)$, where

$$S_l(K) = K^{-1/2} \mathfrak{L}_l(-K) \mathfrak{L}_l^{-1}(K) K^{1/2}. \tag{2.9}$$

Since there must be conservation of current in the open channels, this open-channel submatrix is unitary. Another important property of $S_l(K)$ is that at all energies it is symmetric:

$$S_l(K) = \tilde{S}_l(K). \tag{2.10}$$

3. THE MANY-CHANNEL MARCHENKO EQUATIONS

The one-channel inverse scattering problem at fixed angular momentum, i.e., the problem of constructing the potential from scattering (and bound state) data in one partial wave, has been solved long ago.[9] Of several alternative solutions the one given by Marchenko[10] is the most useful for our purposes here. The generalization to many coupled channels of the Marchenko solution has been obtained by Cox[11,12] and some results are as follows.

Suppose that for a given l there are N "all channels closed" bound states of the system with energies $E_j = \hbar^2 (k_1^{(j)})^2 / 2\mu$, $j = 1, \ldots, N$, and that the n-channel matrix $S_l(K)$ is given for all positive energies. Then one constructs the matrix function

$$G_l(r, r') = (1/2\pi) \int_{-\infty}^{+\infty} dk_1 k_1 w_l(Kr) K^{-1/2} [S_l(K) - 1] K^{-1/2} w_l(Kr')$$

$$- \sum_{j=1}^{N} w_l(iK^{(j)}r) A_j W_l(iK^{(j)}r'), \tag{3.1}$$

where the A_j are certain real symmetric matrices, $w_l(Kr)$ is given by (2.5), and the integral in (3.1) is to be done on the real k_1 axis for which $\text{Im} k_2, \ldots, k_n > 0$. $G_l(r, r')$ serves as the kernel of the matrix integral equation

$$A_l(r, r') = G_l(r, r') + \int_r^{\infty} dr'' A_l(r, r'') G_l(r'', r'), \quad r' > r. \tag{3.2}$$

A knowledge of $A_l(r, r')$ then furnishes the potential matrix according to

$$V(r) = -2 \frac{d}{dr} A_l(r, r) \tag{3.3}$$

and the solution $F_l(K, r)$ [see Eqs. (2.6), (2.7)] of the coupled Schrödinger equations according to

$$F_l(K, r) = (-i)^l [w_l(Kr) + \int_r^{\infty} dr' A_l(r, r') w_l(Kr')]. \tag{3.4}$$

It will prove useful for our purposes to note that one can then use (3.4) to reconstruct, via Eqs. (2.8) and (2.9), the original matrix $S_l(K)$ in (3.1), thus verifying the self-consistency of the whole procedure.

Two remarks should be made here concerning Eq. (3.1). First, direct use of (3.1)–(3.3) in order to

analyze scattering data is not possible, because only the open-channel submatrix is accessible to experimental determination. Although it is possible in principle to continue analytically all elements of this submatrix to energies where all channels are closed and thus find $S_l(K)$, no general method for doing this is known to us. Second, the many-channel Marchenko equations are of use only if the integral transform in (3.1) exists, and this imposes stringent requirements on the matrix $S_l(K)$ for $l > 0$. For example, in the one-channel case $\omega_l(Kr) \sim k_1^{-l}$ as $k_1 \to 0$, so the integrand in (3.1) diverges as $k_1 \to 0$ unless $[S_l(K) - 1]$ vanishes ther at least as fast as k_1^{2l}. For this reason the one-channel Marchenko equations are not ordinarily used for $l > 0$. Instead, another formulation of the inverse problem, the Gel'fand—Levitan equations[9] is used. However, even though the Gel'fand—Levitan equations have been generalized to the coupled-channel case,[13] and there is no divergent integral problem there for $l > 0$, the input is not at all accessible to experiment, nor, contrary to the case of the coupled channel Marchenko equations, is it likely that any exact closed form solutions with $S_l(K)$ meromorphic exist even for $l = 0$.[13] Thus, in any investigation of closed-form solutions to coupled channel problems for $l > 0$, the only recourse, to our knowledge, is to use the Marchenko equations with due attention to the difficulties outlined above.

4. PREDHOLM DETERMINANT AND MEROMORPHIC S MATRICES

The matrix integral equation (2.1) is of the Fredholm type and thus has associated with it a Fredholm determinant $\Delta_l(K) = \Delta_l(k_1, \ldots, k_n)$. It is identical with the determinant of the Jost matrix of Sec. 2[1]:

$$\Delta_l(K) = \det \mathfrak{L}_l(K). \tag{4.1}$$

It is this single function $\Delta_l(K)$ which most conveniently characterizes the scattering, bound states, and resonances of the coupled-channel system for a given l.

Suppose in what follows that $\Delta_l(K)$ is known in analytic form. Then all elements of the open-channel S matrix can be constructed from it according to the rules[1]

$$(S_l)_{aa} = \Delta_l(k_1, \ldots, -k_a, \ldots)/\Delta_l(K), \tag{4.2}$$

$$[(S_l)_{ab}]^2 = (S_l)_{aa}(S_l)_{bb} - \Delta_l(k_1, \ldots, -k_a, -k_b, \ldots)/\Delta_l(K), \tag{4.3}$$

and hence the cross sections can be computed.

Associated with $\Delta_l(K)$ [or any other function of K such as $S_l(K)$] is a k_1 Riemann surface consisting of 2^n half-planes, each distinguished from the others by its particular combination of signs of the imaginary parts of all the n k's, and having branch points at $k_1 = \pm \Delta_2, \ldots, \pm \Delta_n$ which for the positive sign correspond to the threshold energies of the second through nth channel, respectively. [One is led to these conclusions by eliminating $k_2, \ldots, k_n$ in favor of k_1 from $\Delta_l(K)$ by means of (2.3) and considering k_1 a complex number.]

The following properties of $\Delta_l(K)$ on this k_1 Riemann surface are well known whenever the potential matrix has finite first and second absolute moments.[1] $\Delta_l(K) \to 1$ as $|k_1| \to \infty$ in the half-plane $\text{Im} k_1, \ldots, k_n > 0$, and it

1403   J. Math. Phys., Vol. 16, No. 7, July 1975

J.R. Cox and H.R. Garcia   1403

is also a regular analytic function there. Furthermore, it can have no zeros in this half-plane except possibly on the imaginary axis. Zeros on the imaginary axis give rise to conventional "all channels closed" bound states of the system. (Zeros which lie on the positive real k_1 axis below the highest channel threshold give rise to bound states "embedded in the continuum.") Finally, all zeros and poles of $\Delta_l(K)$ off any imaginary axis must occur in pairs symmetric about the imaginary axis.

Thus we see that for the purposes of model-building it is natural to consider those systems where $\Delta_l(K)$ for a given l is a meromorphic function of k_1 on the appropriate Riemann surface. (By this we mean that the only singularities of Δ_l on such a surface are poles.) If exact solutions to the coupled Schrödinger equations were known for these systems, then one could construct potential matrices which give rise to the desired number and energies of bound states and resonances simply by making the appropriate choice of poles and zeros of $\Delta_l(K)$.

For the one-channel case, corresponding to a two-particle elastic process via a central potential, such a $\Delta_l(K)$ would be simply a rational function of $k = k_1$,

$$\Delta_l(k) = \prod_i \frac{k - \alpha_i}{k - \beta_i} , \tag{4.4}$$

where, as follows from the discussion above, the number of zeros equals the number of poles, $\text{Im}\beta_i < 0$, and those α_i's in the upper half plane are on the imaginary axis and correspond to bound states. All those potentials $V(r)$ which give rise to $\Delta_l(k)$'s of the form (4.4) are known exactly in closed form for all l in terms of the α_i's and β_i's for the appropriate l. [1] Such potentials are known in the literature as Bargmann potentials.

A generalization of the Bargmann potentials to the case of many coupled s-wave ($l = 0$) channels has also been obtained. [14] The corresponding generalization of (4.4) is of the form

$$\Delta_0(K) = \det \left| 1 + i \sum_j N_j (K + ib^{(j)})^{-1} \right| , \tag{4.5}$$

where the N_j's are certain interrelated $n \times n$ matrices, and the $b^{(j)}$'s are diagonal matrices of order n whose elements $b_p^{(j)}\delta_{pq}$ are connected by

$$(b_1^{(j)})^2 = (b_p^{(j)})^2 - \Delta_p^2, \quad p = 1, \ldots, n, \tag{4.6}$$

where Δ_p is the same as in (2.3).

It turns out that the simplest allowable version of (4.5) with coupling present is of the form

$$\Delta_0(K) = \det[(K + iA)(K + ib)^{-1}], \tag{4.7}$$

where A is a real symmetric $n \times n$ matrix. For $n = 1$ this reduces to a rational function of k with one zero and one pole. However, for $n > 1$ the pole and zero structure of (4.7) is much more complicated. [14]

The main purpose of the following sections is to begin to generalize the Bargmann potentials to the case of many coupled p-wave ($l = 1$) channels by using the many-channel Marchenko equations in a similar manner as was done earlier for the s-wave case. Although we will only arrive at the analog of (4.7) for the coupled p-wave

case, the analysis will point the way for obtaining the analog of (4.5).

5. SEARCH FOR A "ONE POLE" p-WAVE S MATRIX

Assume there are no p-wave bound states. Then the p-wave version of (3.1) reads

$$G(r, r') = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk_1 k_1 w(Kr) K^{-1/2} [S(K) - 1] K^{-1/2} w(Kr'), \tag{5.1}$$

where we have suppressed the subscripts $l = 1$ on $w(Kr)$ and $S(K)$, and where, according to (2.5),

$$w(Kr) = [i - (Kr)^{-1}] \exp(iKr). \tag{5.2}$$

We are of course ignorant of the exact analytical form of $S(K)$ in (5.1). However, recall from Sec. 3 that the k_1 integration is to be done on the real k_1 axis for which $\text{Im}k_2, \ldots, k_n > 0$. If $S(K)$ is a meromorphic function of k_1, as we assume here, then we can evaluate (5.1) by contour integration. The appropriate closed contour c consists of a large semicircle in the half-plane $\text{Im}K > 0$ plus the real k_1 axis with small semicircular detours to avoid the threshold branch points and a possible simple pole at $k_1 = 0$. According to the discussion of Sec. 4, the simplest (nontrivial) meromorphic S matrix then appears to be the one which has only one pole within c. We will call this S matrix a "one pole" S matrix even though in general it would have additional poles in other half-planes. This section is devoted to an attempt to construct a one-pole p-wave S matrix when there are no p-wave bound states. Although the attempt does not prove successful when coupling is present, it is instructive in that it leads unexpectedly to a new s-wave S matrix and also furnishes valuable insight into the correlation between the poles of the S matrix and the "angular momentum" pole of $w(Kr)$ which will be useful in the sucessful construction of a "two-pole" coupled channel S matrix in Sec. 6.

For a one-pole S matrix we expect that (5.1) yields

$$G(r, r') = -w(ibr)Cw(ibr') + (rr')^{-1}D, \tag{5.3}$$

where the pole position is $K = ib$ [as in (4.7)], and C and D are real symmetric $n \times n$ matrices. This conjecture is a natural generalization of the one-pole version of $G(r, r')$ in the one-channel case. In this case, according to the discussion of Sec. 4, we would have

$$\Delta_1(k) = \frac{k + ia}{k + ib} , \quad a > 0, \quad b > 0,$$

and consequently, according to (4.2),

$$S_1(k) = \frac{k - ia}{k + ia} \frac{k + ib}{k - ib} . \tag{5.4}$$

Insertion of (5.4) in (5.1) and interpretation of the integral in the principal value sense then easily leads to an expression of the form of (5.3).

Equation (5.3) may be written in matrix notation as

$$G(r, r') = A(r)B(r'), \tag{5.5}$$

where

$$A(r) = (A_1 A_2), \quad B(r) = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \tag{5.6}$$

and where

$$A_1 = -iw(ibr), \quad A_2 = r^{-1},$$

$$B_1 = -iCw(ibr), \quad B_2 = Dr^{-1}. \tag{5.7}$$

[Note that both $A(r)$ and $B(r)$ are rectangular matrices, since each of their two "elements" in (5.6) is itself an $n \times n$ matrix.] Equation (5.5) is a degenerate kernel in the matrix sense for the (Fredholm) Eq. (3.2) and can thus be solved exactly. The solution is

$$A(r, r') = A(r)Q^{-1}(r)B(r'), \tag{5.8}$$

where

$$Q(r) = 1 - \int_r^\infty dr' B(r')A(r'). \tag{5.9}$$

Insertion of (5.8) into (3.4) and evaluation of the integrals yields

$$F(K, r) = -i[w(Kr) + A(r)Q^{-1}(r)P(r)], \tag{5.10}$$

where

$$p(r) = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}, \tag{5.11}$$

and where

$$p_1 = -C[(K+ib)^{-1} + (Kbr)^{-1}]\exp[i(K+ib)r],$$

$$p_2 = -D(Kr)^{-1}\exp(iKr). \tag{5.12}$$

According to (5.6), (5.7), and (5.9) the $2n \times 2n$ matrix $Q(r)$ is

$$Q(r) = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}, \tag{5.13}$$

where

$$Q_{11} = 1 - Cb^{-1}[\tfrac{1}{2} + (br)^{-1}]\exp(-2br), \tag{5.14a}$$

$$Q_{12} = -C(br)^{-1}\exp(-br), \tag{5.14b}$$

$$Q_{21} = -D(br)^{-1}\exp(-br), \tag{5.14c}$$

$$Q_{22} = 1 - Dr^{-1}. \tag{5.14d}$$

Although (5.10) formally should be a solution of the p-wave version of (2.6), as it stands it is of little use in computing the S matrix because of the necessity of finding the inverse of $Q(r)$ which in the presence of coupling $(n \geq 2)$ is at least a 4×4 matrix. We now show how (5.10) may be cast into a form which does permit computation of the S matrix for arbitrary n.

Let $\theta(r)$ be the inverse of $Q(r)$, and partition $\theta(r)$ into four $n \times n$ blocks in a manner analogous to (5.13):

$$\theta(r) = \begin{pmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{pmatrix}. \tag{5.15}$$

Then the statement $Q(r)\theta(r) = 1$ may be expressed by the four matrix equations

$$Q_{ab}\theta_{bc} = 1\delta_{ac}, \quad a, c = 1, 2, \tag{5.16}$$

where δ_{ab} is the Kronecker delta and where summation from 1 to 2 over repeated indices is implied. Similarly, Eq. (5.10) becomes

$$F(K, r) = -i[w(K, r) + A_a \theta_{ab} p_b], \tag{5.17}$$

which, upon use of (5.12), also may be written

$$F(K, r) = -i\{w(Kr) - A_a \theta_{a1} C[(K+ib)^{-1} + (Kbr)^{-1}]$$

$$\times \exp[(iK - b)r] - A_a \theta_{a2} D(Kr)^{-1}\exp(iKr)\}. \tag{5.18}$$

Evaluating (5.18) at $K = ib$ and then using (5.14a) and (5.14b), as well as (5.7) and (5.16), we obtain

$$F(ib, r) = A_a \theta_{a1}. \tag{5.19}$$

Next, we define the matrix $Z_\zeta(K, r)$ by

$$Z_\zeta(K, r) = [(K - \zeta)^{-1} - i(K\zeta r)^{-1}]\exp[i(K - \zeta)r], \tag{5.20}$$

where ζ is an $n \times n$ diagonal matrix with elements $\zeta_{ij} = \zeta_i \delta_{ij}$ such that

$$\zeta_1^2 = \zeta_j^2 + \Delta_j^2, \quad j = 1, \ldots, n. \tag{5.21}$$

[See (2.3).] Also, we define the projector P_i to be that $n \times n$ matrix which has all elements equal to zero except the ith element on the diagonal, which equals one. Then it follows from (5.18) and (5.20) that

$$W_i = [r^{-1} + A_a \theta_{a1} C(br)^{-1}\exp(-br) + A_a \theta_{a2} Dr^{-1}]P_i D(Kr)^{-1}$$

$$\times \exp(iKr), \tag{5.22}$$

where

$$W_i = \lim_{P_i \zeta \to 0} F(\zeta, r)\zeta D\zeta Z(K, r)P_i. \tag{5.23}$$

Using (5.14b) and (5.14d), as well as (5.16), (5.22) reduces to

$$W_i = A_a \theta_{a2} P_i D(Kr)^{-1}\exp(iKr), \tag{5.24}$$

which, upon summing over i from 1 to n, yields

$$W = A_a \theta_{a2} D(Kr)^{-1}\exp(iKr), \tag{5.25}$$

where

$$W = \sum_i W_i. \tag{5.26}$$

Finally, insertion of (5.19) and (5.25) in (5.18) yields, upon use of (5.23) and (5.26),

$$F(K, r) = i\ -w(Kr) + F(ib, r)CZ_{ib}(K, r)$$

$$+ \sum_i \lim_{P_i \zeta \to 0} [F(\zeta, r)\zeta D\zeta Z_\zeta(K, r)P_i]. \tag{5.27}$$

Equation (5.27), which reveals explicitly the contributions due to the "pole" and threshold branch points, lends itself to analysis much more readily than does (5.10).

The third term on the right side of (5.27) (i.e., the summation term) actually depends on K only through a factor K^{-1} on the right, as may easily be seen from (5.20). Thus, by setting $K = ib$ in (5.27) and using the resulting equation to eliminate the summation term from (5.27), we obtain

$$F(K, r) = \{(1 - ibK^{-1}) + iF(ib, r)[C(K+ib)^{-1} - (2K)^{-1})$$

$$\times \exp(-br + bK^{-1}\exp(br)]\}\exp(iKr). \tag{5.28}$$

On the other hand, setting $K = \zeta$ in (5.27) and using the resulting equation together with (5.27) with $K = ib$ to eliminate the summation term, we obtain

$$F(ib, r) = [b(D - r) - 1]\{[b\exp(br) - \tfrac{1}{2}C\exp(-br)](D - r)$$

$$+ Cb^{-1}\exp(-br)\}^{-1}. \tag{5.29}$$

If conjecture (5.3) is correct, then (5.28) must solve the p-wave version of (2.6). [Note that (5.28) certainly satisfies (2.7).] However, if we use (2.8) and (5.28), but not (5.29), to calculate the Jost matrix, we find that, in the presence of coupling, it does not satisfy (2.10) and hence does not produce a symmetric p-wave S matrix. This is surprising, because, as demonstrated in Appendix I, (5.3) is of a form which guarantees a symmetric potential matrix, which in turn "must" produce a symmetric S matrix. We are thus led to inquire whether (5.28) is indeed a solution of (2.6) for some well-behaved potential matrix. The potential matrix may be calculated explicitly. First, we apply the technique used to go from (5.10) to (5.27) to (5.8). The result is

$$A(r, r') = i\left(- F(ib, r)Cw(ibr')\right.$$
$$\left. + \sum_i \lim_{\rho_i \zeta \to 0} [F(\zeta, r)\zeta D\zeta w(\zeta r')\rho_i]\right). \qquad (5.30)$$

Then, setting $K = \zeta$ in (5.28) and using the resulting equation to evaluate the sum in (5.30), we obtain

$$V(r) = - 2/r^2 - 2\frac{d}{dr}\{F(ib, r)[\tfrac{1}{2}C \exp(- br) + b \exp(br)]\}, \qquad (5.31)$$

where we have also used (3.3). Using (5.28), (5.29), and (5.31) we can check by direct substitution whether (2.6) is satisfied. This is done in Appendix II. What we find there is that (5.28), (5.29) solve the p-wave version of (2.6) with potential (5.31) only if D is diagonal. However, according to (2.8), $F(ib, r)$ must have, for any well-behaved potential matrix, an r^{-1} singularity at $r = 0$, and (5.29) can have this property only if C and D are related by

$$C = 2bDb(Db - 2)^{-1}, \qquad (5.32)$$

which implies in turn that if D is diagonal, then so is C. Unfortunately, if C and D are both diagonal, $V(r)$ is diagonal and there is no coupling. We conclude, therefore, that (5.28) does not, for any choice of C and D, furnish a solution of the p-wave version of (2.6) with a well-behaved, nondiagonal potential matrix.

It can be shown that, as expected, in the uncoupled, one-channel limit, where C, D, and b are numbers connected by (5.32), (5.31) is well behaved at $r = 0$ and is in fact the well-known Bargmann potential which produces the S matrix given by (5.4), where there a is a function of C and D.[1] In this uncoupled case the singular first term on the right side of (5.31) is precisely cancelled by a part of the second term. In the presence of any coupling, however, this part of the second term becomes nondiagonal, the cancellation fails, and $V(r)$ becomes singular at $r = 0$ precisely as $- 2r^{-2}$. However, this precisely cancels the p-wave barrier term in (2.6). Consequently, we are led to the following conclusion. Equations (5.28) and (5.29) with D diagonal and C nondiagonal constitute a solution of the s-wave ($l = 0$) version of (2.6) with the (nonsingular) potential matrix

$$V(r) = - 2\frac{d}{dr}\{F(ib, r)[\tfrac{1}{2}C \exp(- br) + b \exp(br)]\}. \qquad (5.33)$$

The corresponding s-wave Jost matrix may be easily

calculated from (5.28) and the s-wave version of (2.8). We obtain

$$\mathcal{L}_0(K) = (K^2 + iR_1K + R_2)[K(K + ib)]^{-1}, \qquad (5.34)$$

where

$$R_1 = F(ib, 0)(\tfrac{1}{2}C + b), \quad R_2 = [b + F(ib, 0)(\tfrac{1}{2}C - b)]b, \qquad (5.35)$$

and where $F(ib, 0)$ is given in terms of C and D by (5.29). The S matrix is then obtained via Eq. (2.9) or Eqs. (4.1)—(4.3). It is

$$S_0(K) = (K + ib)K^{3/2}(K^2 + iR_1K + R_2)^{-1}$$
$$\times (K^2 - iR_1K + R_2)K^{-3/2}(K + ib)^{-1}, \qquad (5.36)$$

and, as expected, it can be shown to be symmetric.

It is interesting to note that, according to (4.6), we can never have, for any j, $b^{(j)} = 0$ in (4.5). Consequently, the determinant of (5.34) is never of the form (4.5), and thus, to our knowledge, (5.33) and (5.36) constitute a new exact solution to the s-wave coupled channel scattering problem. However, the presence of the factor $K^{-3/2}$ in (5.36) produces infinite cross sections at thresholds for all production reactions, and hence (5.36) is rather pathological.

6. THE "TWO-POLE" p-WAVE S MATRIX

Let us recall the nature of the difficulty encountered in the search for a one-pole solution in Sec. 5. The matrix D had to be diagonal in order to have a solution to the coupled p-wave Schrödinger equations, and at the same time C and D had to be related by (5.32) which in turn implied that C was also diagonal and hence that there could be no coupling. All this suggests to us that (5.3) provides insufficient "degrees of freedom" for a p-wave solution, and so should be replaced by the "two-pole" expression

$$G(r, r') = - w(ibr)Cw(ibr') - w(iar)Ew(iar') + (rr')^{-1}D, \qquad (6.1)$$

where now both C and E are real, symmetric, nondiagonal, $n \times n$ matrices, D is $n \times n$, real and diagonal, and a, as well as b, is diagonal and has its diagonal elements connected by a relation of the form (4.6). We will now show that (6.1) does produce an exact, trouble-free, p-wave solution in the presence of coupling and that this solution is self-consistent in the sense that it gives rise to an S matrix which, when inserted in (5.1), leads again to (6.1).

The method of analysis used on (5.3) is also easily applied to (6.1) with the result that, as expected, the generalization of (5.27) is

$$F(K, r) = i\left(- w(K, r) + F(ia, r)EZ_{ia}(K, r) + F(ib, r)CZ_{ib}(K, r)\right.$$
$$\left. + \sum_i \lim_{\rho_i \zeta \to 0} [F(\zeta, r)\zeta D\zeta Z_\zeta(K, r)\rho_i]\right). \qquad (6.2)$$

Using (6.2) together with (6.2) evaluated at first $K = ib$ and then at $K = ia$ to thus eliminate in two ways the summation term in (6.2), we obtain, respectively,

$$F(K, r) = \{(1 - ibK^{-1}) + iF(ia, r)E[(K + ia)^{-1} - bK^{-1}(a + b)^{-1}]$$
$$\times \exp(- ar) + iF(ib, r)[C((K + ib)^{-1} - (2K)^{-1})$$
$$\times \exp(- br) + bK^{-1} \exp(br)]\} \exp(iKr) \qquad (6.3)$$

and

$$F(K, r) = \{(1 - iaK^{-1}) + iF(ia, r)[E((K + ia)^{-1} - (2K)^{-1})$$

$$\times \exp(-ar) + aK^{-1} \exp(ar)] + iF(ib, r)C[(K + ib)^{-1}$$

$$- aK^{-1}(a + b)^{-1}] \exp(-br)\} \exp(iKr). \qquad (6.4)$$

Finally, setting $K = \zeta$ in (6.3) and using the resulting equation to evaluate the summation term in (6.2), we obtain one additional expression for $F(K, r)$:

$$F(K, r) = i\{(1 - bD)r^{-1} + F(ia, r)E[(K + ia)^{-1} + (Kar)^{-1}$$

$$- bD(a + b)^{-1}(Kr)^{-1}] \exp(-ar) + F(ib, r)$$

$$\times [C((K + ib)^{-1} + (Kbr)^{-1} - D(2Kr)^{-1}) \exp(-br)$$

$$+ bD(Kr)^{-1} \exp(br)]\} \exp(iKr). \qquad (6.5)$$

From (6.5) and (2.8) we see that the condition for $F(K, r)$ to have a r^{-1} singularity at $r = 0$ is

$$a^{-1}\tilde{\mathscr{L}}(ia)E[a^{-1} - bD(a + b)^{-1}] + b^{-1}\tilde{\mathscr{L}}(ib)[\tfrac{1}{2}C(2b^{-1} - D) + bD] = 0. \qquad (6.6)$$

Equation (6.6) sheds new light on the difficulty encountered in Sec. 5. If $E = 0$ [and we assume $\tilde{\mathscr{L}}(ib)$ to be nonsingular], then condition (6.6) reduces to (5.32) which cannot be satisfied for diagonal D and nondiagonal C. Equation (6.6) thus reveals the necessity, in the presence of coupling, of having at least two nondiagonal matrices (C and E) in addition to the diagonal matrix D.

Comparing (6.3) and (6.4), we see that

$$F(ia, r)[- \tfrac{1}{2}E(a + b)^{-1} \exp(-ar) + a(a + b)^{-1} \exp(ar)]$$

$$- F(ib, r)[\tfrac{1}{2}C(a + b)^{-1} \exp(-br) + b(a - b)^{-1} \exp(br)] = 1, \qquad (6.7)$$

from which it follows, according to (2.8), that

$$a^{-1}\tilde{\mathscr{L}}(ia)[(a - \tfrac{1}{2}E)(a + b)^{-1}$$

$$= b^{-1}\tilde{\mathscr{L}}(ib)[\tfrac{1}{2}C(a + b)^{-1} + b(a - b)^{-1}]. \qquad (6.8)$$

If we now use (2.8) to compute the Jost matrix from (6.3), (6.4), and (6.6), we obtain three alternative expressions of the form

$$\tilde{\mathscr{L}}(K) = K[\tilde{\alpha}K^2 + i\tilde{P}K + \tilde{Q}][K(K + ia)(K + ib)]^{-1}, \qquad (6.9)$$

with three alternative expressions for the real matrices $\tilde{\alpha}$, $\tilde{P}$, and $\tilde{Q}$. Comparing these expressions and making use of (6.8), we conclude that

$$\tilde{\alpha} = a^{-1}\tilde{\mathscr{L}}(ia)Ea(a + b)^{-1} + b^{-1}\tilde{\mathscr{L}}(ib)[\tfrac{1}{2}C + b] = 1, \qquad (6.10)$$

$$\tilde{P} = a + b - a^{-1}\tilde{\mathscr{L}}(ia)Ea - b^{-1}\tilde{\mathscr{L}}(ib)Cb, \qquad (6.11)$$

and

$$\tilde{Q} = [a^{-1}\tilde{\mathscr{L}}(ia)E + b^{-1}\tilde{\mathscr{L}}(ib)C - 1]ab \qquad (6.12a)$$

$$= [a^{-1}\tilde{\mathscr{L}}(ia)Ea(a + b)^{-1} + b^{-1}\tilde{\mathscr{L}}(ib)(\tfrac{1}{2}C - b)ab^{-1}]b^2. \qquad (6.12b)$$

From (6.9) and (6.10) we then have

$$\mathscr{L}(K) = K[K^2 + i\tilde{P}K + \tilde{Q}][K(K + ia)(K + ib)]^{-1}. \qquad (6.13)$$

Next, we wish to establish the relationship between P and Q in (6.13). Using (6.12b) we can write (6.6) as

$$\tilde{Q}D = a^{-1}\tilde{\mathscr{L}}(ia)Eb + b^{-1}\tilde{\mathscr{L}}(ib)Ca. \qquad (6.14)$$

On the other hand, using (6.12a) and (6.11) we see that the right side of (6.14) is simply $Q(a^{-1} + b^{-1}) + P$. Thus, assuming Q is nonsingular, we conclude that

$$D = a^{-1} + b^{-1} + \tilde{Q}^{-1}\tilde{P}. \qquad (6.15)$$

Since a, b, and D are diagonal, we conclude from (6.15) that

$$PQ^{-1} = d^{-1}, \qquad (6.16)$$

where d is the diagonal matrix

$$d^{-1} = D - a^{-1}b^{-1}, \qquad (6.17)$$

and also that

$$\tilde{P}Q = \tilde{Q}P. \qquad (6.18)$$

Although it is not easily seen from (6.11), we will find later that, also,

$$P = \tilde{P}. \qquad (6.19)$$

It is interesting that (6.16) and (6.19) also follow simply from the requirement of time reversal invariance. To see this, we ask how P and Q must be related in (6.13) in order to satisfy (2.10). The answer is that we must require that the quantity $\tilde{P}[1 + K^{-2}Q]$ be symmetric for all energies, which, upon use of (2.3), implies (6.19) and also that

$$P_i Q^{-1} P P_i = P_i P Q^{-1}, \qquad (6.20)$$

where P_i is the same projector as was used in (6.2). Summing over i in (6.20) we immediately infer (6.16), i.e., that PQ^{-1} must be diagonal.

Finally, using (6.16) and (6.19) to eliminate $\tilde{Q}$ in (6.13), we obtain

$$\tilde{\mathscr{L}}(K) = K[K^2 + P(iK + d)][K(K + ia)(K + ib)]^{-1}. \qquad (6.21)$$

The Fredholm determinant is then, from (4.1) and (6.21),

$$\Delta(K) = Y(K)[\det(K + ia)(K + ib)]^{-1}, \qquad (6.22)$$

where

$$Y(K) = \det[K^2 + P(iK + d)]. \qquad (6.23)$$

In accordance with the assumption of no p-wave bound states, P and d are assumed to be chosen in such a way that

$$Y(K) \neq 0, \quad \text{Im}k_i > 0, \quad i = 1, \ldots, n. \qquad (6.24)$$

[That such a choice is possible is easily seen by examining (6.23) in the uncoupling limit.]

For the case of two channels, for example, (6.22) becomes

$$\Delta(k_1, k_2) = Y(k_1, k_2)\left(\prod_{i=1}^{2} (k_i + ia_i)(k_i + ib_i)\right)^{-1}, \qquad (6.25)$$

where

$$Y(k_1, k_2) = k_1^2 k_2^2 + (ik_1 + d_1)(ik_2 + d_2)\det P + k_1^2 P_2(ik_2 + d_2)$$

$$+ k_2^2 P_1(ik_1 + d_1), \qquad (6.26)$$

and where d_i and P_i are, respectively, the diagonal elements of d and P. The S matrix is then, according to (6.25), (4.2), and (4.3),

$$S_{11} = \frac{Y(-k_1, k_2)(k_1 + ia)(k_1 + ib_1)}{Y(k_1, k_2)(k_1 - ia_1)(k_1 - ib_1)}, \qquad (6.27a)$$

$$S_{22} = \frac{Y(k_1, -k_2)(k_2 + ia_2)(k_2 + ib_2)}{Y(k_1, k_2)(k_2 - ia_2)(k_2 - ib_2)}, \qquad (6.27b)$$

$$S_{12} = S_{21} = \frac{-2i(k_1 k_2)^{3/2} P_3 (k_2 - ia_2)(k_2 - ib_2)}{Y(k_1, k_2)(k_1 - ia_1)(k_1 - ib_1)}, \qquad (6.27c)$$

where P_3 is the off-diagonal element of P.

The generalization of (5.30) to the "two-pole" case can be shown to be, as expected,

$$A(r, r') = i\left(- F(ib, r)Cw(ibr') - F(ia, r)Ew(iar') \right.$$

$$\left. + \sum_i \lim_{p_i \zeta \to 0} [F(\zeta, r)\zeta D\zeta w(\zeta r')p_i] \right). \qquad (6.28)$$

Using (6.3) to evaluate the sum in (6.28) and then substituting into (3.3), we obtain the potential matrix

$$V(r) = -2\frac{d}{dr}\left\{ - bDr^{-1} + F(ib, r)[r^{-1}(\tfrac{1}{2}C(2b^{-1} - D)\exp(-br) \right.$$

$$+ bD\exp(br)) + C\exp(-br)] + F(ia, r)$$

$$\times [r^{-1}E(a^{-1} - bD(a + b)^{-1} + r)\exp(-ar)]\}. \qquad (6.29)$$

From (6.6) and (6.10) we see that the expression in curly brackets in (6.29) is not singular at $r = 0$ (the terms proportional to r^{-1} and r^{-2} each vanish) and consequently that the potential matrix is now well behaved at $r = 0$ in the sense of the discussion of Sec. 2. Also, it is possible to show, using (6.29) and (2.7), that if $D \neq 0$, then the large-r behavior of $V(r)$ is $4Dr^{-3}$, and if $D = 0$ [which corresponds, according to (6.17), to one particular choice of d for a given a and b], then the behavior is that of a decreasing exponential, i.e., larger of $2\exp(-br)(bC + Cb)\exp(-br)$ and $2\exp(-ar)(aE + Ea)$ $\times \exp(-ar)$.

We want next to show explicitly that the two-pole S matrix we have constructed above, when inserted in (5.1) really does produce (6.1) as it must for self-consistency. This furnishes an indirect check that (6.5) and (6.29) satisfy the coupled p-wave equations (2.6). A direct proof of this fact can be constructed along similar lines as the proof given in Appendix II, but we will not give it here.

According to (2.9) and (6.13), (5.1) now reads

$$G(r, r') = \frac{1}{2\pi}P\int_{-\infty}^{+\infty} dk_1 k_1 w(Kr)A(K)w(Kr'), \qquad (6.30)$$

where P denotes principal value,

$$A(K) = K^{-1/2}[S(K) - 1]K^{-1/2}, \qquad (6.31)$$

and where, also,

$$A(K) = -2iK^{-1}[K^2(P - a - b) - (a + b)Q - abP]$$

$$\times [(K^2 + iKP + Q)(K - ia)(K - ib)]^{-1}. \qquad (6.32)$$

The integral (6.30) is to be evaluated by contour integration over the contour c described at the beginning of Sec. 5. First, let us compute the residues of the simple poles at $K = ia$, $K = ib$, within c, of the integrand of (6.30). Denoting this integrand by $\bar{A}(K)$, writing

$$\bar{A}(K) = G_b(K)(K + ib)^{-1} = G_a(K)(K + ia)^{-1}, \qquad (6.33)$$

and using (2.3) and (4.6), we see that

$$\operatorname*{Res}_{k_1 = ia_1} \bar{A}(K) = a_1^{-1}G_a(ia)a, \quad \operatorname*{Res}_{k_1 = ib_1} \bar{A}(K) = b_1^{-1}G_b(ib)b. \qquad (6.34)$$

Hence, from (6.30) and (6.32)—(6.34), we have

$$2\pi i \operatorname*{Res}_{k_1 = ia_1} \bar{A}(K) = -w(ibr)Cw(ibr'),$$

$$\qquad (6.35)$$

$$2\pi i \operatorname*{Res}_{k_1 = ib_1} \bar{A}(K) = -w(iar)Ew(iar'),$$

where we have defined C and E by

$$C = 2(a + b)(b - P - b^{-1}Q)(b + P - b^{-1}Q)^{-1}(b - a)^{-1}b \quad (6.36)$$

and

$$E = 2(a + b)(a - P - a^{-1}Q)(a + P - a^{-1}Q)^{-1}(a - b)^{-1}a. \quad (6.37)$$

[Since $C = \tilde{C}$ and $E = \tilde{E}$, it follows from (6.16) and either (6.36) or (6.37) that $P = \tilde{P}$.] Equations (6.35) furnish the first two terms on the right side of (6.1). As we will now see, the third term is due to the threshold branch points and the angular momentum pole at the origin.

The residue of the integrand of (6.30) at the simple pole at $k_1 = 0$, $k_j = i\Delta_j$, $j = 2, \ldots, n$, is found to be, upon use of (6.16),

$$\pi i \operatorname*{Res}_{k_1 = 0} \bar{A}(K) = -(rr')^{-1}[(a_1 + b_1)d_1 + a_1 b_1]\lim_{k_1 \to 0}p_i[(K^2 P^{-1} + iK + d)$$

$$\times (K - ia)(K - ib)]^{-1}. \qquad (6.38)$$

Further evaluation of (6.38) depends on the easily proved fact that if M is any nonsingular matrix with elements M_{ij} such that $p_j M = M_{ij} p_j$, then

$$p_j M^{-1} = (M_{jj})^{-1}p_j. \qquad (6.39)$$

Applying (6.39) to (6.38), we find that

$$\pi i \operatorname*{Res}_{k_i = 0} \bar{A}(K) = -(a_1^{-1} + b_1^{-1} + d_1^{-1})(rr')^{-1}p_1. \qquad (6.40)$$

Again using (6.39) it is straightforward to show that the contributions to the integral in (6.30) from the threshold branch points at $k_1 = +\Delta_i$ and $k_1 = -\Delta_i$ are equal and that their sum B_i is simply

$$B_i = -(a_i^{-1} + b_i^{-1} + d_i^{-1})(rr')^{-1}p_i, \quad i = 2, \ldots, n. \qquad (6.41)$$

Finally, adding together (6.40) and (6.41), we obtain

$$-(a^{-1} + b^{-1} + d^{-1})(rr')^{-1},$$

which, upon use of (6.17), indeed yields the third term on the right side of (6.1). (The contribution of the large semicircle in c is, of course, zero.)

7. CONCLUSION

In conclusion, we have found that the simplest symmetric p-wave meromorphic many-channel S matrix is produced by the Fredholm determinant (6.22) and not by the simpler one-pole expression (4.7) which would be the case in the uncoupling limit. We have also found explicitly the underlying potential matrix. It is interesting to note that if it were possible to factor the expression $K^2 + P(iK + d)$ in (6.21) in such a way as to have a (diagonal) factor $(K + ia)$ or $(K + ib)$ on the right, then (6.22) would indeed simplify to an expression of the form (4.7). However, this must be impossible for it is easily verified that the resulting Jost matrix does not then satisfy (2.10) and hence the condition for time reversal invariance is violated. Our results suggests also that in the presence of coupling the algebraic structure of even the simplest possible meromorphic S

matrix becomes progressively more complicated with increasing angular momentum. Finally, we point out that instead of using C, D, and E as matrix parameters in the two-pole solution, it would be simpler to use d and P as parameters. If this is done, then C, D, and E are given by (6.16), (6.17), and (6.37). Also, Eqs. (6.3) and (6.4) with $K = ia$ and $K = ib$, respectively, then furnish the matrices $F(ia, r)$ and $F(ib, r)$ which are needed in the computation of, for example, $V(r)$ in (6.29).

APPENDIX A: PROOF THAT $\widetilde{G}(r,r') = \widetilde{G}(r',r)$ IMPLIES $V(r) = \widetilde{V}(r)$

Assume that

$$\widetilde{G}(r, r') = G(r', r). \tag{A1}$$

Then (3.2) with $r' = r$, together with (3.2), implies that

$$A(r, r) = G(r, r) + \int_r^\infty dr'' A(r, r'') \widetilde{G}(r, r'')$$
$$- \int_r^\infty dr'' [A(r, r'') - G(r, r'')] \widetilde{A}(r, r')$$
$$+ \int_r^\infty dr' [A(r, r') - G(r, r')] \widetilde{A}(r, r'),$$

which, upon use of (3.2) again, becomes

$$A(r, r) = G(r, r) + \int_r^\infty dr' [A(r, r') \widetilde{G}(r, r') + G(r, r') \widetilde{A}(r, r')$$
$$- A(r, r') \widetilde{A}(r, r')] + \int_r^\infty dr' \int_r^\infty dr''$$
$$\times A(r, r'') G(r'', r') \widetilde{A}(r, r'). \tag{A2}$$

Equations (A1) and (A2) imply

$$A(r, r) = \widetilde{A}(r, r). \tag{A3}$$

Hence, upon use of (3.3) and (A3) we conclude that $V(r) = \widetilde{V}(r)$.

APPENDIX B: VERIFICATION THAT $F(K,r)$ OBTAINED OBTAINED IN SEC. 5 SOLVES THE COUPLED SCHRÖDINGER EQUATIONS FOR DIAGONAL D

The p-wave version of (2.6) for $F(K, r)$ is

$$F''(K, r) = [V(r) + 2r^{-2} - K^2] F(K, r). \tag{B1}$$

Substituting (5.28) and (5.31) in (B1) and using the fact that $K^2 + b^2$ is, according to (2.3) and (4.6), a multiple of the unit matrix, we easily find that we indeed have a solution, provided only that $F(ib, r)$ is a solution of (B1) with $K = ib$.

Writing $F(ib, r)$ as given by (5.29) as

$$F(ib, r) = NM^{-1}, \tag{B2}$$

where

$$N = [b(D - r) - 1],$$
$$M = [(b \exp(br) - \tfrac{1}{2}C \exp(-br))(D - r) + Cb^{-1} \exp(-br)]^{-1}, \tag{B3}$$

and defining

$$T = \tfrac{1}{2}C \exp(-br) + b \exp(br), \tag{B4}$$

we find, upon substitution of (B2) into (B1) with $K = ib$ and subsequent use of (5.31), (B3), and (B4), that the condition for solution is

$$M = T'B^{-2}N + TbN^{-1}b^{-2}N. \tag{B5}$$

However, from (B3) and (B4) it follows that

$$M = T'b^{-2}N + Tb^{-1}. \tag{B6}$$

Comparing (B5) and (B6) we see that in order to have a solution it is necessary and sufficient that N and b^{-2} commute, which, according to (B3), means that D must be diagonal.

*Present address: Pratt and Whitney Aircraft, West Palm Beach, Florida.
[1] R.G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966).
[2] L.D. Faddeev, *Mathematical Aspects of the Three-Body Problems in Quantum Scattering Theory*, Publ. 69 (Steklov Math. Inst., Leningrad) (transl. by Davey, New York, 1965).
[3] A.N. Kamal and H.J. Kreuzer, Phys. Rev. 2, 2033 (1970).
[4] R.J. Moore, Phys. Rev. 4, 3755 (1971).
[5] N.J. Zabusky and M.D. Kruskal, Phys. Rev. Lett. 15, 240 (1965).
[6] P.J. Caudrey et al., Phys. Rev. Lett. 30, 237 (1973).
[7] G.L. Lamb, Phys. Rev. Lett. 31, 196 (1973).
[8] P.C. Sabatier, C.R. Hebd. Seances Acad. Sci. B (France), 278, 545 (1974).
[9] I.M. Gel'fand and B.M. Levitan, Dokl. Acad. Nauk. SSSR 77, 557 (1951).
[10] L.D. Faddeev, J. Math. Phys. 4, 72 (1963).
[11] J.R. Cox, thesis, Indiana University, 1962.
[12] J.R. Cox, unpublished notes.
[13] J.R. Cox, Ann. Phys. (N.Y.) 39, 216 (1966).
[14] J.R. Cox, J. Math. Phys. 5, 1065 (1964).

On angular momentum and channel coupling for a meromorphic many-channel S matrix

Joseph R. Cox

Department of Physics, Florida Atlantic University, Boca Raton, Florida 33432
(Received 26 February 1975)

We consider a class of possible (i.e., not known *a priori* to be unitary) nonrelativistic many-channel S matrices meromorphic on their energy Riemann surfaces whose general form is suggested by the inverse scattering problem. These possible S matrices are associated with n (spinless) coupled channels of the same angular momentem l, and they reduce, in the one-channel limit, to the quotient of two polynomials of degree $2m$ in the wavenumber k. It is shown that the assumptions which imply unitarity for real energies of the open-channel submetrix also imply that no channels can be coupled when $m \le l$.

1. INTRODUCTION

In the well-known (spinless) many-channel model of nonrelativistic inelastic scattering of a particle by a target haivng a finite number n of discrete excited states as described by Newton,[1] the S matrix S_+ for orbital angular momentum l is, under the assumption that the interaction forces are sufficiently well behaved, the open-channel submatrix of the $n \times n$ matrix

$$S = K^{-1/2} \mathcal{L}(-K) \mathcal{L}^{-1}(K) K^{1/2}. \tag{1.1}$$

In (1.1) K is the diagonal $n \times n$ matrix of the channel wavenumbers $k_1, \ldots, k_n$; conservation of energy $\hbar^2 k_1^2/2\mu$ of the system requires that

$$k_1^2 = k_j^2 + \Delta_j^2, \quad j = 1, \ldots, n, \tag{1.2}$$

where μ is the (common) channel reduced mass and $\hbar^2\Delta_j^2/2\mu$, $\Delta_j > 0$, is the threshold energy of the jth channel, i.e., it is the energy difference between the ground state and the $(j-1)$th excited state of the target. $\mathcal{L}(K)$ in (1.1) is the $n \times n$ Jost matrix. Time-reversal invariance is assumed; thus, S is symmetric. Hence, using (1.1), we have

$$\left[\tilde{\mathcal{L}}(K) K^{-1} \mathcal{L}(-K) \right]_a = 0, \tag{1.3}$$

where the tilde denotes matrix transpose and where the notation $[M]_a$ denotes the antisymmetric part of M, i.e.,

$$[M]_a = 1/2(M - \tilde{M}). \tag{1.4}$$

Properties of $\mathcal{L}(K)$ in addition to (1.3) include

$$\mathcal{L}(-K^*) = \mathcal{L}^*(K), \tag{1.5}$$

where the asterisk denotes complex conjugate, and also

$$\mathcal{L}'_{\alpha\beta}(k_1, k_2, \ldots, k_\gamma, \ldots, k_n)$$
$$= \mathcal{L}'_{\alpha\beta}(k_1, k_2, \ldots, -k_\gamma, \ldots, k_n), \quad \gamma \ne \alpha, \tag{1.6}$$

where $\mathcal{L}'_{\alpha\beta}$ is related to the matrix elements $\mathcal{L}_{\alpha\beta}$ of $\mathcal{L}(K)$ by

$$\mathcal{L}'_{\alpha\beta} = (k_\alpha/k_\beta)^l \mathcal{L}_{\alpha\beta}. \tag{1.7}$$

As Newton[1] has shown, Eqs. (1.3), (1.5), and (1.6) imply that for real energies S_+ is unitary.

In this paper we are concerned with the specific Jost matrix

$$\tilde{\mathcal{L}}(K) = K^l N(K) K^{-l} Q^{-1}(K), \tag{1.8}$$

where $Q(K)$ is the diagonal matrix

$$Q(K) = \prod_{j=1}^{m} (K + ib^{(j)}) \tag{1.9}$$

and where

$$N(K) = \sum_{i=0}^{m} A_i K^i. \tag{1.10}$$

Each $b^{(j)}$ in (1.9) is a diagonal matrix of order n whose elements $b_p^{(j)}\delta_{pq}$ all have positive real parts and are connected by

$$(b_1^{(j)})^2 = (b_p^{(j)})^2 - \Delta_p^2, \quad p = 1, \ldots, n, \tag{1.11}$$

where Δ_p is the same as in (1.2). The matrices A_i in (1.10) are $n \times n$, not necessarily diagonal, and are independent of K. They may, however, depend on the nonnegative integer l and on m, which is assumed here to be a positive integer. In addition,

$$A_m = 1. \tag{1.12}$$

Finally, the A_i's are such that det $N(K)$ has no zeros when Im k_1, k_2, $\ldots$, $k_n > 0$ unless all k_j's are purely imaginary.

Equation (1.8) for $n > 1$ arises naturally from attempts to generalize the Bargmann potentials[1] to the many-channel case via a heuristic procedure which utilizes a generalization[2] of the Marchenko[3] solution of the one-channel inverse scattering problem at fixed angular momentum. The starting point of this procedure is an assumption about the number and positions of the poles of S in the region Im k_1, k_2, $\ldots$, $k_n > 0$ and the number of bound states. The inverse problem formalism then dictates the underlying solution to the coupled radial Schrödinger equations and hence the Jost matrix and S. Particular coupled closed form solutions which have been obtained in this manner for the many-channel problem $(n > 1)$ include the cases $l = 0$, $m \ge 1$,[2] and $l = 1$, $m = 2$.[4] However, when $l > 0$, certain difficulties appear. One of these, the one to which we address ourselves here, is that some initial assumptions lead only to a diagonal S, i.e., to a situation where all channels are uncoupled. (An example of this is the set of assumptions which lead to (1.8) with $l = 1$, $m = 1$.[4]) It would be very helpful to have a means of identifying such cases from the beginning.

It seems natural to suppose that this difficulty is closely related to the unitarity (and symmetry) of the S matrix. For this reason, we initiate a study of the following question. Suppose (1.8) is given and that the

A_i's (and $b^{(j)}$'s are unknowns. What restrictions are imposed on them by the requirement that S_+ be unitary? In this paper we consider the cases where

$$m \le l. \tag{1.13}$$

It is our main purpose to show that if (1.8) and (1.13) are assumed, then the requirements (1.3), (1.5), and (1.6), which together imply the unitarity of S_+, also imply that S_+ is diagonal. Thus, no finite value of m yields coupled channels for arbitrarily large angular momentum l, or, put another way, the polynomials in K given by (1.9) and (1.10) must be at least of degree $l+1$ in order for any channels to be coupled.

2. UNITARITY OF S_+

It is readily verified, with the aid of (1.7), that (1.8) satisfies (1.6); indeed, $\mathcal{L}^1_{\alpha\beta}$ here is a function of k_α only. Also, Eq. (1.8) when inserted in (1.5) implies

$$(-)^{i+m} A_i^* = A_i \tag{2.1}$$

as well as

$$\{b_j^*\} = \{b_j\}. \tag{2.2}$$

[Equation (2.2) indicates that all poles of $\mathcal{L}_{\alpha\beta}(K)$ occur either on the imaginary axis or in pairs symmetric about the imaginary axis in the k_α plane.] Thus, in order to ensure the unitarity of S_+, we need, in addition to (2.1) and (2.2), only require that the A_i's in (1.8) be such that (1.3) is satisfied. We turn now to this, our main task in this section.

Inserting (1.8) in (1.3) and making use of the fact that, due to (1.2) and (1.11), $K^2 + (b^{(j)})^2$ is a multiple of the unit matrix, (1.3) reduces to

$$\lfloor N(K) K^{-\lambda} N(-K) \rfloor_a = 0, \tag{2.3}$$

where we have defined

$$\lambda = 2l + 1. \tag{2.4}$$

Inserting (1.10) in (2.3), we obtain

$$\sum_{a=1}^{\substack{m \\ m-1}} \sum_{\gamma=0}^{\substack{m-1 \\ m}} [A_\alpha K^{\alpha+\gamma-\lambda} A_\gamma]_a = 0, \tag{2.5}$$

where the sum over α is over odd integers only, the sum over γ is over even integers only, and the first and second upper limits on the sums are for m odd and m even, respectively. (In the remainder of this paper, unless otherwise indicated, all summations are to be computed over odd or even integers only, according as to whether the limits on the summations are odd or even.) Let P_i be the $n \times n$ matrix consisting of all zeros except that the ith entry on the diagonal is unity. Then, in light of (1.2), we can write

$$K = \sum_{i \in r} k_i P_i, \qquad r = \{1, 2, \dots, n\}. \tag{2.6}$$

Letting $\gamma = \beta - \alpha$ in (2.5), interchanging the order of summation, and using (2.6), we obtain

$$\sum_{i \in c} \sum_{\beta=1}^{2m-1} g_{i\beta} k_i^{\beta-\lambda} = 0, \tag{2.7}$$

where we have defined

$$g_{i\beta} = \begin{cases} \sum_{\alpha=1}^{\beta} \Omega_{\alpha, \beta-\alpha}, & \beta \in \left\{1, 3, \dots, \begin{matrix} m-2 \\ m-1 \end{matrix}\right\}, \quad (m > 1) \\[20pt] \sum_{\substack{\alpha = \beta-(m-1) \\ \alpha = \beta-m}}^{\substack{m-1 \\ m}} \Omega_{\alpha, \beta-\alpha}, & \beta \in \left\{\begin{matrix} m, m+2, \dots, \\ m+1, m+3, \dots, \end{matrix} \ 2m-1\right\}, \end{cases} \tag{2.8}$$

and also

$$\Omega_{\alpha, \beta} = [A_\alpha P_i \tilde{A}_\beta]_a. \tag{2.9}$$

[The upper and lower members in the sets in (2.8) are for m odd and m even, respectively.] From (1.13) and (2.4) it follows that (2.7) contains only negative integer powers of k_i^2. Furthermore, as is easily seen with the aid of (1.2), the functions $k_i^{\beta-\lambda}$ are all linearly independent functions of k_i. Hence, under the assumption (1.13), (2.7) implies

$$g_{i\beta} = 0, \quad i \in r, \quad \beta \in \{1, 3, \dots, 2m-1\}. \tag{2.10}$$

Using the fact that, according to (2.9), $\Omega_{\alpha,\beta} = -\Omega_{\beta,\alpha}$ and making some simple changes of variables in the summations in (2.8), Eq. (2.10) can be expressed as

$$\sum_{\substack{\alpha=1 \\ \alpha=0}}^{\substack{\beta-1 \\ \beta-1}} \Omega_{\alpha, \beta-\alpha} = 0, \quad \beta \in \left\{\begin{matrix} m-2 \\ 1, 3, \dots, m-1 \end{matrix}\right\}, \quad (m > 1), \tag{2.11a}$$

$$\sum_{\alpha=\beta+1}^{m} \Omega_{\alpha, \beta+m-\alpha} = 0, \beta \in \left\{\begin{matrix} 0, 2, \dots, \\ 1, 3, \dots, \end{matrix} \ m-1 \right\}. \tag{2.11b}$$

We now work out in detail the implications of (2.11b). To this end, let N be an integer such that $0 \le N < m$, and assume that

$$\sum_{\alpha=\beta+1}^{m} \Omega_{\alpha, \beta+m-\alpha} = 0, \quad \beta \in \{m-1, m-3, \dots, N\}. \tag{2.12}$$

Equation (2.12), together with Eqs. (1.12) and (2.1), implies that there exists a set of $1/2(m-N+1)$ real diagonal $n \times n$ matrices $D_{m-1}, D_{m-3}, \dots, D_N$ such that

$$A_{m-1} = i D_{m-1} \tag{2.13a}$$

and

$$A_\beta = i \left(D_\beta + \sum_{\alpha=\beta+1}^{m-2} A_\alpha D_{m+\beta-\alpha} \right),$$

$$\beta \in \{m-3, m-5, \dots, N\}, \quad (m \ge 3). \tag{2.13b}$$

The proof is by induction on N. First, assume in (2.12) that $N = m-1$. Then, using (1.12) and (2.9), we obtain

$$P_i \tilde{A}_{m-1} = A_{m-1} P_i,$$

which, upon application of the fact that

$$P_i P_j = P_i \delta_{ij}, \tag{2.14}$$

leads to the conclusion that A_{m-1} is diagonal. Furthermore, according to (2.1), A_{m-1} is purely imaginary. Hence, (2.13a) follows. Second, assume that (2.12) [together with (1.12) and (2.1)] implies (2.13) (when $N \ge 2$) and also assume 2.13 with $\beta = N-2$, i.e.,

$$\sum_{\alpha=N-1}^{m} \Omega_{\alpha, m+N-\alpha-2} = 0. \tag{2.15}$$

To complete the proof, we show that (2.13b) with β $=N-2$, i.e.,

$$A_{N-2} = i\left(D_{N-2} + \sum_{\alpha=N-1}^{m-2} A_\alpha D_{m+N-\alpha-2}\right), \qquad (2.16)$$

follows. By using (2.9), (2.15) may be written

$$[A_{N-2} p_i]_a + \Lambda = 0, \qquad (2.17)$$

where we have used (1.12) and have defined

$$\Lambda = \sum_{\alpha=N-1}^{m-2} [A_{m+N-\alpha-2} p_i \tilde{A}_\alpha]_a. \qquad (2.18)$$

Inserting (2.13) in (2.18), we obtain

$$-i\Lambda = \left[p_i \sum_{\alpha=N-1}^{m-2} D_{m+N-2-\alpha}\right]_a + i\Lambda', \qquad (2.19)$$

where

$$\Lambda' = \sum_{\alpha=N+1}^{m-2} \sum_{\alpha'=M+N-1-\alpha}^{m-2} [A_{\alpha'} D_{2m+N-2-\alpha-\alpha'} p_i \tilde{A}_\alpha]_a. \qquad (2.20)$$

Interchanging the order of summation in one of the two terms in brackets in (2.20), we find that

$$\Lambda' = 0. \qquad (2.21)$$

[All subsequent double sums in this paper which are asserted vanish, vanish for similar reasons as (2.20).] Thus, from (2.21), (2.19), (2.18), and (2.17) we have

$$\left[\left(A_{N-2} - i\sum_{\alpha=N-1}^{m-2} A_\alpha D_{m+N-2-\alpha}\right)p_i\right]_a = 0,$$

from which we infer, upon use of (2.14) and (2.1), Eq. (2.16), where D_{N-2} is some real diagonal matrix.

Setting $N=0$ if m is odd and $N=1$ if m is even in (2.13), we have the statement that there exists a set of real diagonal $n \times n$ matrices D_β such that

$$A_{m-1} = iD_{m-1}, \qquad (2.22a)$$

$$A_\beta = i\left(D_\beta + \sum_{\alpha=\beta+1}^{m-2} A_\alpha D_{m+\beta-\alpha}\right),$$

$$\beta \in \left\{\begin{matrix} 0, 2, \cdots, \\ 1, 3, \cdots, \end{matrix} m-3\right\}, \quad (m \geq 3). \qquad (2.22b)$$

We now work out the implications of (2.22) and (2.11a). Let N' be an odd integer such that $1 \leq N' \leq \frac{m-2}{m-1}$, and assume that

$$\sum_{\substack{\alpha=1 \\ \alpha=0}}^{\beta-1} \Omega_{\alpha,\beta-\alpha} = 0, \quad \beta \in \left\{\begin{matrix} m-2, m-4, \cdots, \\ m-1, m-3, \cdots, \end{matrix} N'\right\}. \qquad (2.23)$$

Equation (2.23), together with Eqs. (2.22), (2.1), and (1.12), implies that, according to whether m is odd or even, there exists a set of $1/2(m-N')$ or $1/2(m-N'-1)$ real diagonal $n \times n$ matrices $D_{-\beta}$ such that

$$D_{-\beta} + \sum_{\substack{\alpha=N+1 \\ \alpha=0}}^{m-2} A_\alpha D_{m-\alpha-\beta} = 0,$$

$$\beta \in \left\{\begin{matrix} 2, 4, \cdots, \\ 1, 3, \cdots, \end{matrix} m-N'\right\}, \quad (m > 1). \qquad (2.24)$$

The proof is by induction on N'. First, assume in (2.23) that $N' = \frac{m-2}{m-1}$. Then, according to (2.9), (2.23) becomes

$$\sum_{\substack{\alpha=1 \\ \alpha=0}}^{\frac{\beta-1}{m-1-\alpha}} [A_\alpha p_i \tilde{A}_{m-2-\alpha}]_a = 0. \qquad (2.25)$$

Inserting (2.22) in (2.25), we obtain

$$\sum_{\substack{\alpha=1 \\ \alpha=0}}^{\frac{m-2}{m-1-\alpha}} [A_\alpha D_{m-2-\alpha} p_i]_a = 0, \qquad (2.26)$$

where we have made use of the fact that

$$\sum_{\substack{\alpha=1 \\ \alpha=2}}^{\frac{m-2}{m-\alpha}} \sum_{\substack{\alpha'=m-\alpha-1 \\ \alpha'=m-\alpha}}^{\frac{m-2}{m-\alpha}} [A_\alpha p_i D_{2m-2-\alpha-\alpha'} \tilde{A}_{\alpha'}]_a = 0.$$

Equation (2.26) implies, upon use of (2.14) and (2.1), that for some real diagonal $n \times n$ matrix D_{-2},

$$D_{-2} + \sum_{\substack{\alpha=1 \\ \alpha=0}}^{\frac{m-2}{m-\alpha-1}} A_\alpha D_{m-\alpha-2} = 0, \qquad (2.27)$$

which is (2.24) with $\beta = \frac{2}{1}$. Second, in order to complete the proof of (2.24), we must show that (for $N' > 3$) (2.22) together with (2.24), when inserted in (2.23) with $\beta = N' - 2$, i.e.,

$$\sum_{\substack{\alpha=1 \\ \alpha=0}}^{\frac{N'-2}{N'-3}} \Omega_{\alpha, N'-2-\alpha} = 0, \qquad (2.28)$$

implies, with the aid of (2.1), that there exists a real diagonal matrix $D_{N'-2-m}$ such that

$$D_{N'-2-m} + \sum_{\substack{\alpha=1 \\ \alpha=0}}^{\frac{m-2}{m}} A_\alpha D_{N'-2-\alpha} = 0, \qquad (2.29)$$

i.e., Eq. (2.24) with $\beta = m - (N'-2)$. By using (2.9), (2.28) may be written

$$\sum_{\substack{\alpha=1 \\ \alpha=0}}^{\frac{N'-2}{N'-3}} [A_\alpha p_i \tilde{A}_{N-2-\alpha}]_a = 0. \qquad (2.30)$$

Substituting (2.22) in (2.30), we obtain

$$\sum_{\substack{\alpha=1 \\ \alpha=0}}^{\frac{N'-2}{N'-3}} [A_\alpha D_{N'-2-\alpha} p_i]_a + \psi = 0, \qquad (2.31)$$

where

$$\psi = \sum_{\substack{\alpha=1 \\ \alpha=0}}^{\frac{N'-2}{N'-3}} \sum_{\substack{\alpha'=N' \\ \alpha'=N'-1}}^{\frac{m-2}{m-2}} \psi_{\alpha\alpha'}, \qquad (2.32)$$

$$\psi_{\alpha\alpha'} = [A_\alpha D_{m+N'-2-\alpha-\alpha'} p_i \tilde{A}_{\alpha'}]_a \qquad (2.33)$$

and where we have made use of the fact that

$$\sum_{\substack{\alpha=1 \\ \alpha=2}}^{\frac{N'-2}{N'-3}} \sum_{\substack{\alpha'=N'-1-\alpha}}^{\frac{N'-2}{N'-3}} \psi_{\alpha\alpha'} = 0.$$

Next, setting $-\beta = -\alpha'-2 + N'$ in (2.24) and using the resulting equation to eliminate from (2.32) the sum over α', we find that (2.32) reduces to

$$\psi = \sum_{\substack{\alpha = N' \\ \alpha = N'-1}}^{m-2} [A_\alpha D_{N'-2-\alpha} P_i]_a, \qquad (2.34)$$

where we have used the fact that

$$\sum_{\substack{\alpha = N' \\ \alpha = N'-1}}^{m-2} \sum_{\substack{\alpha' = N' \\ \alpha' = N'-1}}^{m-2} \psi_{\alpha \alpha'} = 0.$$

Finally, combining (2.31) and (2.34), we have

$$\sum_{\substack{\alpha=1 \\ \alpha=0}}^{m-2} [A_\alpha D_{N'-2-\alpha} P_i]_a = 0,$$

which implies, upon use of (2.14) and (2.1), Eq. (2.29), where $D_{N'-2-m}$ is some real diagonal matrix.

Setting $N' = 1$ in (2.25), we have the statement that, in addition to (2.22), there exists a set of real diagonal $n \times n$ matrices $D_{-\beta}$ such that

$$D_{-\beta} + \sum_{\substack{\alpha=1 \\ \alpha=0}}^{m-2} A_\alpha D_{m-\alpha-\beta} = 0, \quad \beta \in \left\{ \begin{matrix} 2, & 4, & \ldots, \\ 1, & 3, & \ldots, \end{matrix} \, m-1 \right\}, \quad (m>1). \qquad (2.35)$$

Thus, Eqs. (2.11), together with (2.1), imply (2.22) and (2.35). [(2.13) and (2.24), which are more general than (2.22) and (2.35), will be utilized in a subsequent paper.]

Equations (2.22) and (2.35) furnish, in a recursive fashion, expressions for the m real diagonal matrices D_α in terms of the A_α's. Indeed, setting $\beta = m-3$ in (2.22b) and using (2.22a), then setting $\beta = m-5$ in (2.22b) and using the previous equations, etc., we obtain

$$iD_{m-1} = -A_{m-1},$$

$$iD_{m-3} = A_{m-3} - A_{m-2} A_{m-1},$$

$$iD_{m-5} = A_{m-5} - A_{m-4} A_{m-1} - A_{m-2} A_{m-3} + A_{m-2}^2 A_{m-1}.$$

The generalization is, clearly,

$$iD_{m-p} = A_{m-p} + \sum_{\sigma_p'} (-)^{h+1} A_{m-p_1} A_{m-p_2} \cdots A_{m-p_h},$$

$$p \in \left\{ 1, 3, \ldots, \begin{matrix} m \\ m-1 \end{matrix} \right\}, \qquad (2.36)$$

where σ_1' is the empty set and σ_p' for $p > 1$ is the set of all ordered decompositions of p into positive integer summands (i.e., $p = p_1 + p_2 + \cdots + p_h$) such that $p_1, p_2, \ldots, p_{h-1}$ are even integers and p_h is an odd integer. [Evidently the right side of (2.36) contains altogether $2^{1/2(p-1)}$ terms.] Next, setting $\beta = \frac{2}{1}$ in (2.35) and using (2.36) as well as the fact that placing an even number q on the left of an ordered list of summands of p (from σ_p') produces an element of the set σ_{q+p}', we obtain

$$iD_{m-p} = \sum_{\sigma_p} (-)^{h+1} A_{m-p_1} A_{m-p_2} \cdots A_{m-p_h}, \quad p = \begin{matrix} m+2 \\ m+1 \end{matrix}, \quad (2.37)$$

where σ_p is that subset of σ_p' which consists of summands no larger than m, i.e., $p_i \le m$. Setting $\beta = \frac{4}{3}$ in (2.35) and using (2.36) and (2.37), we obtain (2.37) except that now $p = \begin{matrix} m+4 \\ m+3 \end{matrix}$. Continuing in this manner, we observe that (2.37) is valid for all p such that

$$p \in \left\{ \begin{matrix} m+2, & m=4, \ldots, \\ m+1, & m+3, \ldots, \end{matrix} \, 2m-1 \right\}. \qquad (2.38)$$

Finally, defining

$$\tau(m-p) = \begin{cases} 1, & p \le m, \\ 0, & p > m, \end{cases}$$

and noting that σ_p' may be replaced by σ_p in (2.36), we can combine (2.36) and (2.37) with (2.38) into a single expression and so obtain

$$iD_{m-p} = \tau(m-p)A_{m-p} + \sum_{\sigma_p} (-)^{h+1} A_{m-p_1} A_{m-p_2} \cdots A_{m-p_h},$$

$$p \in \{1, 3, \ldots, 2m-1\}. \qquad (2.39)$$

[It is not difficult to verify directly that (2.39) solves (2.22) and (2.35).] Thus, when (1.13) holds, the A's must be such that the m expressions given by the right side of (2.39) are diagonal (and purely imaginary).

3. STRUCTURE OF S_+ WHEN $m \le l$

Let us write (2.35) as a matrix equation

$$d + aD = 0, \qquad (3.1)$$

where we have defined

$$a = (A_1 A_3 \cdots A_{m-2}), \qquad (3.2)$$

$$d = (D_{-2} D_{-4} \cdots D_{-(m-1)}), \qquad (3.3)$$

and

$$D = \begin{pmatrix} D_{m-3} D_{m-5} D_{m-7} \cdots \\ {m-1} {m-3} {m-5} \\ D_{m-5} D_{m-7} \cdots \cdots \\ {m-3} {m-5} \\ D_{m-7} \cdots \cdots \cdots \\ {m-5} \\ \cdots \cdots \cdots D_{-(m-3)} \end{pmatrix}. \qquad (3.4)$$

A necessary and sufficient condition for (3.1) to have a solution is[5]

$$d(D^+ D - 1) = 0, \qquad (3.5)$$

where D^+ is the (Moore–Penrose) generalized inverse of D, in which case the general solution is

$$a = e + c[1 - DD^+], \qquad (3.6)$$

where we have defined

$$e = -dD^+ \qquad (3.7)$$

and where

$$c = (C_1 C_3 \cdots C_{m-2}) \qquad (3.8)$$

is a real but otherwise arbitrary matrix partioned, in analogy with (3.2), into $n \times n$ blocks $C_1, \ldots, C_{m-2}$. Let us similarly partition e in (3.7):

$$e = (E_1 E_3 \cdots E_{m-2}). \qquad (3.9)$$

Next, note that, as indicated explicitly in (3.4), D consists entirely of diagonal $n \times n$ blocks, and thus, so does D^+. (This latter statement may be easily verified by noting that D is real and symmetric and hence, according to a result of Decell[6] based on the application of the Cayley–Hamilton theorem to D^2, D^+ is expressible as a polynomial in D with numerical coefficients.) Thus, from (3.3) and (3.7) it follows that the $n \times n$

blocks $E_1 \ldots E_{m-2}$ in (3.9) are each diagonal. Also, from (3.5) and (3.7) we see that Eq. (3.1) with a replaced by e is valid, or, reverting to the notation of (2.35),

$$D_{-\beta} + \sum_{\substack{\alpha=1\\ \alpha=0}}^{m-2} E_\alpha D_{m-\alpha-\beta} = 0, \quad \beta \in \begin{Bmatrix} 2, & 4, & \ldots, \\ 1, & 3, & \ldots, \end{Bmatrix} m-1 \Bigg\}, \quad m \geqslant 3. \tag{3.10}$$

In analogy with (2.22) we then define the diagonal $n \times n$ matrices $E_0, E_2, \ldots, E_{m-1}$ in terms of the $n \times n$ blocks in (3.9) by

$$E_{m-1} = A_{m-1}, \tag{3.11a}$$

$$E_\beta = i\left(D_\beta + \sum_{\alpha=\beta+1}^{m-2} E_\alpha D_{m+\beta-\alpha}\right),$$

$$\beta \in \begin{Bmatrix} 0, & 2, & \ldots, \\ 1, & 3, & \ldots, \end{Bmatrix} m-3 \Bigg\}, \quad m \geqslant 3. \tag{3.11b}$$

(In other words, the E's satisfy precisely the same equations as the A's; however, the E's are diagonal whereas the A's, due to c in (3.6), in general, are not diagonal.) Finally, in analogy with (1.10) and (1.12) we define the *diagonal* matrix polynomial in K,

$$M(K) = \sum_{i=0}^{m} E_i K^i, \quad E_m = 1, \tag{3.12}$$

where the sum extends over both odd and even integers.

We now show that

$$W = 0, \tag{3.13}$$

where W is defined as

$$W = N(K)M(-K) - N(-K)M(K). \tag{3.14}$$

Equation (3.13) is established by a procedure which closely parallels that immediately following (2.3). Substituting (1.10) and (3.12) in (3.14), we obtain

$$W = \sum_{\alpha=1}^{m} \sum_{\beta=\alpha}^{m+\alpha-1} \hat{\Omega}_{\alpha, \beta-\alpha} K^\beta, \tag{3.15}$$

where

$$\hat{\Omega}_{\alpha, \beta} = A_\alpha E_\beta - A_\beta E_\alpha. \tag{3.16}$$

Interchanging the order of summation in (3.15) and defining, in a manner analogous to (2.8),

$$\hat{g}_\beta = \begin{cases} \displaystyle\sum_{\alpha=1}^{\beta} \hat{\Omega}_{\alpha, \beta-\alpha}, & \beta \in \begin{Bmatrix} 1, & 3, & \ldots, \\ \end{Bmatrix} \begin{matrix} m-2 \\ m-1 \end{matrix}\Bigg\}, \\[20pt] \displaystyle\sum_{\substack{\alpha=\beta-(m-1)\\ \alpha=\beta-m}}^{m} \hat{\Omega}_{\alpha, \beta-\alpha}, & \beta \in \begin{Bmatrix} m, & m+2, & \ldots, \\ m+1, & m+3, & \ldots, \end{Bmatrix} 2m-1\Bigg\}, \end{cases} \tag{3.17}$$

we obtain

$$W = \sum_{\beta=1}^{2m-1} \hat{g}_\beta K^\beta. \tag{3.18}$$

Our aim is now to show that

$$\hat{g}_\beta = 0, \quad \beta \in \{1, 3, \ldots, 2m-1\}, \tag{3.19}$$

and hence, by (3.18), $W = 0$. Equation (3.17) can be rewritten, in a manner analogous to that used to obtain the left sides of (2.11),

$$\hat{g}_\beta = \pm \sum_{\substack{\alpha=1\\ \alpha=0}}^{\beta-1} \hat{\Omega}_{\alpha, \beta-\alpha}, \quad \beta \in \begin{Bmatrix} 1, & 3, & \ldots, \\ \end{Bmatrix} \begin{matrix} m-2 \\ m-1 \end{matrix}\Bigg\}, \tag{3.20a}$$

$$\hat{g}_{\beta+m} = \pm \sum_{\alpha=\beta+1}^{m} \hat{\Omega}_{\alpha, \beta+m-\alpha}, \quad \beta \in \begin{Bmatrix} 0, & 2, & \ldots, \\ 1, & 3, & \ldots, \end{Bmatrix} m-1\Bigg\}, \tag{3.20b}$$

where the plus and minus signs refer to the cases m odd and m even, respectively. Let us consider (3.20b) first. From (1.12), (3.12), (3.11a), and (3.16), we see that (3.20b) may be written [the term in (3.21) containing the summation is to be omitted if $\beta = m-1$ or $m-3$]

$$\pm \hat{g}_{\beta+m} = (E_\beta - A_\beta) + i(A_{\beta+1} - E_{\beta+1})D_{m-1} + \sum_{\alpha=\beta+3}^{m-2} \hat{\Omega}_{\beta+m-\alpha}, \tag{3.21}$$

which, upon insertion of (2.22b) and (3.11b), readily reduces to

$$\hat{g}_{\beta+m} = 0, \quad \beta \in \begin{Bmatrix} 0, & 2, & \ldots, \\ 1, & 3, & \ldots, \end{Bmatrix} m-1\Bigg\}, \tag{3.22}$$

where we have made use of the fact that

$$\sum_{\alpha=\beta+3}^{m-2} \sum_{\alpha'=\beta+m-\alpha+1}^{m-2} \hat{\Omega}_{\alpha\alpha'} D_{2m+\beta-\alpha-\alpha'} = 0.$$

Now consider (3.20a). First, let $\beta = m-1$ (m even). Then (3.20a) becomes

$$\hat{g}_{m-1} = \pm \sum_{\alpha=0}^{m-2} \hat{\Omega}_{\alpha, m-\alpha-1},$$

which, upon insertion of (2.22b) and (3.11b) and subsequent use of (2.35) and (3.10) with $\beta = 1$, yields

$$\hat{g}_{m-1} = 0, \tag{3.23}$$

where we have used the fact that

$$\sum_{\alpha=2}^{m-2} \sum_{\alpha'=m-\alpha}^{m-2} \hat{\Omega}_{\alpha\alpha'} D_{2m-\alpha-\alpha'-1} = 0.$$

Returning to (3.20a) and assuming $\beta \neq m-1$, we have, upon use of (2.22b) and (3.11b), as well as a subsequent interchange in order of integration,

$$\mp i\hat{g}_\beta \sum_{\substack{\alpha=1\\ \alpha=0}}^{\beta-1} (A_\alpha - E_\alpha)D_{\beta-\alpha} + \sum_{\substack{\alpha'=\beta+2\\ \alpha'=\beta+1}}^{m-2} \sum_{\substack{\alpha=1\\ \alpha=0}}^{\beta-1} \hat{\Omega}_{\alpha\alpha'} D_{m+\beta-\alpha-\alpha'},$$

$$\beta \in \begin{Bmatrix} 1, & 3, & \ldots, \\ \end{Bmatrix} \begin{matrix} m-2 \\ m-3 \end{matrix}\Bigg\}, \tag{3.24}$$

where we have used the fact that

$$\sum_{\substack{\alpha'=1\\ \alpha'=0}}^{\beta-1} \sum_{\alpha=\beta-\alpha'+1}^{\beta-1} \hat{\Omega}_{\alpha\alpha'} D_{m+\beta-\alpha-\alpha'} = 0.$$

Finally, using (2.35) and (3.10) to eliminate the second summation in the last term in (3.24) and making use of the fact that $\hat{\Omega}_{\alpha, \beta} = -\hat{\Omega}_{\beta, \alpha}$, we obtain, after a second use of (2.35) and (3.10) with $-\beta$ replaced by $\beta-m$,

$$\hat{g}_\beta = 0, \quad \beta \in \begin{Bmatrix} 1, & 3, & \ldots, \\ \end{Bmatrix} \begin{matrix} m-2 \\ m-3 \end{matrix}\Bigg\}. \tag{3.25}$$

Equations (3.25), (3.23), and (3.22) are equivalent to (3.19) which in turn implies, via (3.18), (3.13).

In order to demonstrate that under assumption (1.13) S is diagonal, all that remains is to insert (1.8) in (1.1)

and make use of (3.14) and (3.13) to eliminate $N(-K)$ and thus obtain

$$S = K^{-(l+1/2)}Q(-K)^{-1}M^{-1}(K)M(-K)Q(K)K^{(l+1/2)}. \quad (3.26)$$

[Since K, $Q(K)$, and $M(K)$ are diagonal, so are S and S_+.]

4. SUMMARY AND CONCLUSION

Suppose the A's in (1.10) are given, and they satisfy (2.1). Then the conditions (1.3), (1.5), (1.7) which together imply that for real energies S_+ is unitary also imply, under assumption (1.13), that there exist m real diagonal matrices D_α which are related to the A's by (2.22) and (2.35) and which can be computed from (2.39). However, (2.35) has a solution (for some of the A's) only under assumption (3.5). When it has a solution, there exists a diagonal polynomial $M(K)$ such that $N(K)M(-K)$ is an even function of K. This has the consequence that $p(K)$ is factorable into the factors K^l, a

function of K^2 (not necessarily diagonal), and a diagonal function of K, and hence, by (1.1), leads always to a diagonal S and S_+. As an alternative to taking the A's as given, one could take as given the m real diagonal matrices D_α, subject to (3.5). Then, (3.2) and (3.6)—(3.9) furnish the A's. Of course, the conclusion that S_+ must be diagonal for $m \leq l$ is unchanged. The coupled cases of (1.8) with $m > l$ will be discussed in a future work.

[1] R.G. Newton, *Scattering Theory of Waves and Particles*, (McGraw-Hill, New York, 1966).
[2] J.R. Cox, J. Math. Phys. 5, 1065 (1964).
[3] V.A. Marchenko, Dokl. Akad. Nauk. SSSR 104, 695 (1955).
[4] H.R. Garcia, M.S. thesis, Florida Atlantic Univ., 1973 (unpublished).
[5] T.L. Boullion and P.L. Odell, *Generalized Inverse Matrices* (Wiley-Interscience, New York, 1971).
[6] H.P. Decell, Jr., SIAM Rev. 7, 526 (1965).

1415   J. Math. Phys., Vol. 16, No. 7, July 1975

Joseph R. Cox   1415

On the projective unitary representations of connected Lie groups

Jose F. Cariñena and Mariano Santander*

Departamento de Física Teórica, Facultad de Ciencias, Universidad de Valladolid y G.I.F.T., Spain
(Received 3 September 1974)

We show here the possibility of finding a unique Lie group $\bar{G}$ associated with each connected Lie group G such that every projective unitary representation of G can be lifted to a unitary representation of $\bar{G}$, that is to say, all PUR of G can be found from the UR of only one group $\bar{G}$. This method is applied to the research of PUR of the Galilei group and compared with the preceding ones.

1. INTRODUCTION

This paper is devoted to the study of the (continuous) projective unitary representations (hereafter PUR) of a connected Lie group. We are not going to analyze the fundamental role played by the PUR of the symmetry group of a quantum theory in the Hilbert space formulation, because it is well known. We only attempt to study the relations between the PUR of a connected Lie group G and the (continuous) unitary representations (UR) of other Lie groups related with G. Schur[1] has shown that any PUR of a finite group G can be lifted to a UR of another group $\bar{G}$, univocally associated with G. In this paper we show that the same result holds when G is a connected Lie group.

The plan of this work is intended as follows: Section 2 is devoted to a brief review of Bargmann's method used up to today to find the PUR of a connected Lie group. In Sec. 3 a new method of finding the PUR of a connected Lie group from the UR of another connected Lie group $\bar{G}$ is developed. We call this group $\bar{G}$ related with G the "projective covering group" of G. In order to prevent some misunderstandings, two appendices containing the definitions which we are using are added at the end of this paper. In Sec. 4 we consider the application of the new method to two fundamental groups in physics: the connected Poincaré P and Galilei G groups. In the first case the projective covering group and the universal covering group coincide and every PUR of P can be lifted to a UR of $P* \approx (T_4 \odot SL(2, \mathbb{C}))$ as is well known from the Wigner's work.[2] But not all the PUR of G can be lifted to UR of $G*$, and this fact is also known.[3-5] The study of the PUR of G can be found, e.g., in the review of Lévy-Léblond.[6] The last section is devoted to clarify the relation between the method proposed by us and the ones used by Lévy-Léblond[6] and Varadarajan,[7] which are based on Bargmann's[3] and Mackey's[8] works when they are applied to the Galilei group.

2. THE REDUCTION TO UNITARY REPRESENTATIONS

Because of the definition (see Appendix A, Def. A1) of $PU(H)$ as the factor group $U(H)/U(1)$, it follows that $1 \to U(1) \to U(H) \to PU(H) \to 1$ is a central extension of $PU(H)$ by $U(1)$. Bargmann's method of making up all the PUR, τ, of a connected Lie group G can be stated in the following recipe which can be justified by means of the

lifting homomorphism $\Lambda_{G,\tau}$, although neither $U(H)$ nor $PU(H)$ are itself Lie groups:

(i) Determine the universal covering group $G*$ of G. By $p: G* \to G$ we will denote the covering homomorphism.

(ii) Determine all (equivalence classes of) central extensions of $G*$ by $U(1)$.

(iii) Determine all UR of the middle group corresponding to each extension $y: 1 \to U(1) \to Y \xrightarrow{m} G* \to 1$ and select those which are reduced on $U(1)$ to the identify map. Every one gives rise to a PUR of $G*$.

(iv) Among these representations, pick out only those where the subgroup $m^{-1}(\text{Ker} p)$ of Y is mapped on $U(1)$. This condition suffices to assure that we will obtain not only a PUR of $G*$, but one of G itself.

The preceding method can be pictured in the following commutative diagram:

$$
\begin{array}{ccccccc}
y: 1 \to & U(1) \to & Y & \to & G* \to & 1 & \\
 & \| & | & \tau^* & {}^{\searrow G} & & \\
 & \| & \downarrow & & {}^{\nearrow}_{\tau} & & \\
1 \to & U(1) \to & U(H) \to & PU(H) \to & 1 & &
\end{array}
\qquad (2.1)
$$

We must remark that the use of the universal covering group $G*$ is simply of technical character. In fact we can also obtain all PUR of G from certain UR [step (iii)] of the middle group of each extension $X \in \text{Ext}_0$ $(G, U(1))$.

$$
\begin{array}{ccccccc}
X: 1 \to & U(1) \to & X & \to & G & \to 1 & \\
 & \| & | & & | \tau & & \\
1 \to & U(1) \to & U(H) \to & PU(H) \to & 1 & &
\end{array}
\qquad (2.2)
$$

Even this procedure has the advantage of making steps (i) and (iv) unnecessary, $\text{Ext}_0(G, U(1))$ is generally, and roughly speaking, greater than $\text{Ext}_0(G*, U(1))$, and then we have much more extensions to find the PUR of G.

We can summarize all the previous results as follows: Any PUR of a connected Lie group G can be lifted to an UR of some central extension of $G*$ (or G) by $U(1)$. But, generally there exist several (inequivalent) extensions, and therefore this solution to the problem cannot be considered as a canonical one.

3. THE "PROJECTIVE COVERING GROUP" OF A CONNECTED LIE GROUP

Now we shall develop a new method for studying the PUR of a connected Lie group G by making use of a theorem by Hochschild, which analyzes the possibility of a choice of a connected Lie group M and an epimorphism $\omega: M \to G$ in such a way that any extension of G by A (connected Abelian Lie group) is lifted by $\Lambda_{M, \omega}$ to the trivial extension.

Theorem 3.1 (*Hochschild*[9]): Let A, G, be connected Lie groups A being Abelian. Then, there exists a connected and simply connected Lie group M and a continuous epimorphism $\omega: M \to G$ such that Ker $\Lambda_{M, \omega}$ $= \mathrm{Ext}_0(G, A)$.

This theorem is really a particular case of a result due to Hochschild,[10] and its proof is not reproduced here. We will only show the characterization of the Lie group M by means of its Lie algebra. Let LG, LA be the Lie algebras corresponding to G and A respectively, and let $H_0^2(LG, LA)$ be the second cohomology group of LG and LA relative to the trivial action of LG on LA. It is well known that $H^2_0(LG, LA)$ is also a real finite-dimensional vector space. Let n be the dimension of $H^2_0(LG, LA)$ and let $\xi_1, \ldots, \xi_n$ be a set of representatives in $Z^2_0(LG, LA)$ for a basis of $H^2_0(LG, LA)$. We define a Lie algebra structure LM as follows: The subjacent vector space of LM is a direct sum.

$$LM = LA^{\dim H_0^2(LG, LA)} \oplus LG$$

of LG and n copies of LA. The product law (commutation relations) in LM is given by

$$[(a_1, \ldots, a_n, g), (b_1, \ldots, b_n, h)]$$

$$= (\xi_1(g, h), \ldots, \xi_n(g, h), [g, h])$$

where

$$a_1, \ldots, a_n, b_1, \ldots, b_n \in LA \quad \text{and} \quad g, h \in LG.$$

Now let M be the only connected and simply connected Lie group with Lie algebra LM. The homomorphism π: $LM \to LG$ defined by $\pi: (a_1 \cdots a_n, g) \to g$ induces an epimorphism $\omega: M \to G$ such that $\dot\omega = \pi$. This group M and the "covering" homomorphism ω satisfy the announced properties in the Theorem 3.1.

The application of this theorem to the problem which we are considering is an easy matter: Every PUR, τ, of the connected Lie group G can be lifted to a UR, T, of the middle group of some central extension $X \in \mathrm{Ext}_0(G, U(1))$ [see diagram (2.2)]. Both groups G and $U(1)$ in the extension X verify the conditions of Theorem 3.1 which can be used. Hence, there exists a Lie group $\overline{G}$ and a continuous epimorphism $p: \overline{G} \to G$ such that Ker $\Lambda_{\overline{G}, p} = \mathrm{Ext}_0(G, U(1))$. This fact can be pictured by the following commutative diagram.

$$\begin{array}{ccccc}
1 \longrightarrow U(1) \longrightarrow & U(1) \otimes \overline{G} \longrightarrow & \overline{G} \longrightarrow 1 \\
\| & \downarrow & \downarrow p \\
X: 1 \longrightarrow U(1) \longrightarrow & X \longrightarrow & G \longrightarrow 1 \\
\| & \downarrow T & \downarrow \tau \\
1 \longrightarrow U(1) \longrightarrow & U(H) & PU(H) \longrightarrow 1
\end{array}$$

Therefore, each PUR, τ, of G induces a PUR, $\tau \circ p$, of $\overline{G}$, and this last PUR can always be lifted to a UR of $U(1) \otimes \overline{G}$ which on $U(1)$ is reduced to the identity, that is to say, we obtain a UR of $\overline{G}$. Conversely, let R be a UR of $\overline{G}$ and let $1 \otimes R$ be the UR of $U(1) \otimes \overline{G}$ defined by $1 \otimes R$: $(\lambda, g) \to \lambda R(g)$. Clearly $1 \otimes R$ induces a PUR, $\overline{\tau}$, of $\overline{G}$:

$$\begin{array}{ccccc}
1 \longrightarrow U(1) \longrightarrow & U(1) \otimes \overline{G} \longrightarrow & G \longrightarrow 1 \\
\| & \downarrow {\scriptstyle I \otimes R} & R \quad \downarrow {\scriptstyle \tau} \searrow G \\
1 \longrightarrow U(1) \longrightarrow & U(H) \longrightarrow & PU(H) \longrightarrow 1
\end{array}$$

But with $\overline{\tau}$ is associated a PUR τ of G such that $\overline{\tau} = \tau \circ p$ if and only if Ker $\overline{\tau} \subset$ Ker$p$.[11] This means that R induces a PUR of G if and only if R (Ker$p) \subset U(1)$.

We shall call $\overline{G}$ the projective covering group of G, due to the fundamental role played by $\overline{G}$ in this problem. The name is justified, because, in the simplest cases, in which $H^2_0(LG, \mathbb{R}) = 0$, then $\overline{G}$ coincides with the universal covering group.

We must remember that the subjacent vectorial space of the Lie algebra $L\overline{G}$ of $\overline{G}$ is a direct sum $\mathbb{R}^{\dim H_0^2(LG, \mathbb{R})} \oplus LG$ and the commutation relations[12] are $[(\theta_1, \ldots, \theta_n, g), (\xi_1, \ldots, \xi_n, h)] = (\xi_1(g, h), \ldots, \xi_n(g, h), [g, h])$, where $\theta_1, \ldots, \theta_n, \xi_1, \ldots, \xi_n \in \mathbb{R}$ $g, h \in LG$, and $\xi_1, \ldots, \xi_n$ are a set of representatives in $Z_0^2(LG, \mathbb{R})$ for a basis of $H_0^2(LG, \mathbb{R})$. Then the group $\overline{G}$ is defined as being the only connected and simply connected Lie group with Lie algebra $L\overline{G}$. Furthermore, the covering homomorphism $p: \overline{G} \to G$ is induced by the projection $L\overline{G} \to LG$ defined by $(\theta_1, \ldots, \theta_n, g) \to g$.

Then, each connected Lie group G admits a well-defined "projective covering group" $\overline{G}$, and we can summarize all the preceding results in the following theorem.

Theorem 3.2[13]: Let G be any connected Lie group and $\overline{G}$ the corresponding "projective covering group," and let $p: \overline{G} \to G$ be the covering homomorphism. Then, each PUR, τ, of G can be lifted to a UR, R of $\overline{G}$. Conversely, each UR, R, of $\overline{G}$ mapping into $U(1)$ the kernel of p induces a PUR of G

This theorem can be visualized by the following commutative diagram.

$$\begin{array}{ccccc}
1 \longrightarrow \mathrm{Ker}p \longrightarrow & \overline{G} \xrightarrow{\; p \;} & G \longrightarrow 1 \\
\downarrow & \downarrow R & \downarrow \tau \\
1 \longrightarrow U(1) \longrightarrow & U(H) \longrightarrow & PU(H) \longrightarrow 1
\end{array} \qquad (3.1)$$

Following the method here developed, the determination of all PUR of G consists of:

(a) Determine the group $\overline{G}$ and the homomorphism $p: \overline{G} \to G$.

(b) Pick out, among all UR of $\overline{G}$, those mapping Ker$p$ in $U(1)$. Each one gives rise to a PUR of G, according to diagram (3.1). All PUR of G are obtained in this way.

The procedure we are proposing is, over all, conceptually easier than the proceding ones. In fact, only a group $\bar{G}$ is necessary to find by means of its UR, all the PUR of G. Then, we must find only the UR of $\bar{G}$ and afterwards, pick out the suitable ones following (b). When $H^2_0(LG, \mathbb{R}) = 0$, it is obviously $\bar{G} = G^*$ and then every PUR of G can be lifted to a UR of G^*, and conversely, each UR of G^* mapping Ker$p$ in $U(1)$ gives rise to a PUR of G. This result, which is only a corollary of theorem 3.2, is explicitly contained in Bargmann's paper and is well known.

4. APPLICATIONS

The simplest case corresponds to a connected Lie group G such that $H^2_0(LG, \mathbb{R}) = 0$ [or, equivalently, $H^2_0(G^*, \mathbb{R}) = 0$ or also $H^2_0(G^*, U(1)) = 0$]. This is the case for the tridimensional rotation group SO(3), the Poincaré group P (identity component only), etc.[14] As indicated at the end of Sec. 3, Theorem 3.2 leads in such cases to the result that any PUR of G can be lifted to a UR of G^* because then $\bar{G} = G^*$. This result is well known, and we do not comment on it here.

A second example, which we will study more closely, is the Galilei group $\mathcal{G}$, where the simplest relation $H^2_0(L\mathcal{G}, \mathbb{R}) = 0$ does not hold. Let $\mathcal{G}$ be the connected Galilei group, that is to say, the group of all linear transformations of $\mathbb{R}^4$ of the form

$$(b, \mathbf{a}, \mathbf{v}, R): (\mathbf{r}, t) \to (R\mathbf{r} + \mathbf{v}t + \mathbf{a}, t + b), \qquad (4.1)$$

where R is a 3×3 real orthogonal matrix with det$R = +1$. Our notations on $\mathcal{G}$ agree with those of the work of Lévy-Léblond.[5]

The first step to the determination of all PUR of $\mathcal{G}$ is the research of the group $\bar{\mathcal{G}}$. It is well known that $H^2_0(L\mathcal{G}, \mathbb{R})$ has dimension one[15] and a representative in $Z^2_0(L\mathcal{G}, \mathbb{R})$ for a basis of $H^2_0(L\mathcal{G}, \mathbb{R})$ is the "cocycle" Ξ give by $\Xi(K_i, P_j) = \delta_{ij}M$. the values of Ξ being zero for another pairs of generators (we use the standard choice of the basis[16] in $L\mathcal{G}$). Here M is any fixed element of $\mathbb{R}$, which we take as the new generator of the algebra $L\bar{\mathcal{G}}$. The commutation relations in $L\bar{\mathcal{G}}$ are given by

$$[J_i, J_j] = \epsilon_{ijk}J_k, \qquad [K_i, H] = P_i,$$
$$[J_i, K_j] = \epsilon_{ijk}K_k, \qquad [K_i, P_j] = \delta_{ij}M,$$
$$[J_i, P_j] = \epsilon_{ijk}P_k, \qquad (4.2)$$

and all other commutators vanish.

The transition from the algebra $L\bar{\mathcal{G}}$ to the group $\bar{\mathcal{G}}$ is done by "integrating" the former commutation relations. The set of the group $\bar{\mathcal{G}}$ is $\mathbb{R} \times \mathcal{G}^*$, and the group law is given by

$$(\theta, g)(\zeta, h) = (\theta + \zeta + \omega(g, h), gh), \qquad (4.3)$$

where $\theta, \zeta \in \mathbb{R}$ and $g, h \in \mathcal{G}^*$ and where $\omega: \mathcal{G}^* \times \mathcal{G}^* \to \mathbb{R}$ is the factor system of the universal covering group of the Galilei group given by

$$\omega(g, g') = \tfrac{1}{2}b\mathbf{v}'^2 + \mathbf{v} \cdot \tilde{R}\mathbf{a}', \qquad (4.4)$$

where $g \in \mathcal{G}^*$ has been taken in the form $(b, \mathbf{a}, \mathbf{v}, \tilde{R})$, the only difference with an element of $\mathcal{G}$ being the replacement of $R \in$ SO(3) by $\tilde{R} \in SU(2)$ [the group $SU(2)$ acts on

$\mathbb{R}^3$ via the covering homomorphism $SU(2) \to$ SO(3)]. The relation (4.3) shows that $\bar{\mathcal{G}}$ is a central, not trivial, extension of $\mathcal{G}^*$ by R.

Fortunately, this group $\bar{\mathcal{G}}$ is itself a semidirect product and its irreducible UR can be easily characterized by means of Mackey's theorem.[17,18] In fact, the elements $(0; b, \mathbf{a}, 0, 1)$ form an Abelian normal subgroup N of $\bar{\mathcal{G}}$; the group K of homogeneous Galilei transformations $(0; 0, 0, \mathbf{v}, \tilde{R})$ acts on N as follows:

$$(\mathbf{v}, \tilde{R}): (\theta; b, \mathbf{a}) \to (\theta + \tfrac{1}{2}b\mathbf{v}^2 + \mathbf{v} \cdot \tilde{R}\mathbf{a}, b, \tilde{R}\mathbf{a} + b\mathbf{v}). \quad (4.5)$$

This law gives the semidirect structure of $\bar{\mathcal{G}}$. The group N being isomorphic to $\mathbb{R}^5$, its dual group $\hat{N}$ is also isomorphic to $\mathbb{R}^5$; let us denote its elements by $(m, E, \mathbf{p})$. The group K acts on N as follows:

$$(\mathbf{v}, \tilde{R}): (m, E, \mathbf{p}) \to (m, E + \mathbf{v} \cdot \tilde{R}\mathbf{p} + \tfrac{1}{2}m\mathbf{v}^2, \tilde{R}\mathbf{p} + m\mathbf{v}). \quad (4.6)$$

The orbits under this action can be easily found. The function $(m, E, \mathbf{p}) \to 2mE - \mathbf{p}^2$ is invariant under the action (4.6). Take $\rho = 2mE - \mathbf{p}^2$. The orbits are given by

$$Z_{m, \rho} = \{(m, E, \mathbf{p}) \mid 2mE - \mathbf{p}^2 = \rho\}, \quad m, \rho \in \mathbb{R}, \quad m \neq 0$$
$$Z_{0, \rho} = \{(0, E, \mathbf{p}) \mid -\mathbf{p}^2 = \rho\}, \quad \rho \in \mathbb{R}, \quad \rho < 0, \qquad (4.7)$$
$$Z_{0, 0, E} = \{(0, E, 0)\}, \quad E \in \mathbb{R}.$$

Looking at this orbit structure of $\hat{N}$, it is easy to conclude that $\bar{\mathcal{G}}$ is a regular semidirect product, and, then, all its irreducible UR are obtained from the Mackey method. As $\bar{\mathcal{G}}$ is of "type I,"[18] then all UR are direct sums (integral) of irreducible UR. The little groups of each orbit (4.7) are easily calculated and finally, all irreducible UR of $\bar{\mathcal{G}}$ are classified in five "types" and labeled in the following form:

$$
\begin{aligned}
&\text{I:} && [m, \rho, j], && m, \rho \in \mathbb{R}, && 2j \in Z, && j \geq 0, && m \neq 0; \\
&\text{II:} && [\rho, \xi], && \rho \in \mathbb{R}, && 2\xi \in Z, && \rho < 0; \\
&\text{III:} && [\rho, r, \epsilon], && \rho, r \in \mathbb{R}, && \epsilon = \pm 1, && \rho < 0, && r > 0; \\
&\text{IV:} && [E, j], && E \in \mathbb{R}, && 2j \in Z, && j \geq 0; \\
&\text{V:} && [E, r, \xi], && E, r \in \mathbb{R}, && 2\xi \in Z, && r > 0.
\end{aligned}
\qquad (4.8)
$$

According to Theorem 3.2, it turns out that the irreducible PUR of $\mathcal{G}$ are induced by those (4.8) mapping onto $U(1)$ the kernel of the covering homorphism. But Ker$p$ is contained in the center of $\bar{\mathcal{G}}$ and Schur's lemma shows that any irreducible UR of $\bar{\mathcal{G}}$ applies Ker$p$ on $U(1)$. This last step (b) introduces no further restrictions. Then, the irreducible PUR of $\mathcal{G}$ are also given by (4.8).

One point remains to be studied: the possible equivalence (in a projective sense) of these representations. It turns out that PUR of class II—V with different labels are projectively inequivalent and that the PUR of class I with the same $[m, j]$ but different ρ's are projectively equivalent.[19]

5. DISCUSSION: (GALILEI GROUP)

Although original works on PUR of the Galilei group are neither by Varadarajan[4] nor Lévy-Léblond,[5] we will compare the methods given in these references with the

one we are proposing in order to show clearly and sharply the fundamental difference between both.

The method used by Varadarajan[7] is essentially the one given in (2.1), where the identity automorphism of $U(1)$ has been replaced by the complex conjugation $\lambda \to \lambda^* = \lambda^{-1}$. The difference is only of a technical character and leads to the consequence that from all UR of the middle group of each extension $\chi \in \text{Ext}_0(\mathcal{G}^*, U(1))$ we must choose only those reducing on $U(1)$ to $\lambda \to \lambda^{-1}$. The remainder of the procedure is formally analogous to the one given in Sec. 4, although m is not a real number but an entier [because $\widehat{U(1)} = Z$]. The UR mapping λ on λ^{-1} are precisely the ones corresponding to the orbits with $m = -1$. Although the (real) label m seems not to appear in this method, it is, really, parametrizing the set $\text{Ext}_0(\mathcal{G}^*, U(1))$. Then as in the case of Sec. 4 the last step (iv) introduces no further restrictions for irreducible UR. Attention is called to the fact that this method involves first a "multiplicative" step (ii) (we must consider the UR of many groups) and, after, another "restrictive" step (iii) (select only some UR of each group) which are unnecessary in the method given in the former section.

The method given by Lévy-Léblond is similar to Varadarajan's, but he considers the extensions of $\mathcal{G}^*$ by $\mathbb{R}$. The set $H_0^2(\mathcal{G}^*, \mathbb{R})$ is one dimensional and it is parametrized by $m \in \mathbb{R}$. A peculiar fact is that there is an infinite set of inequivalent extensions (called by Lévy-Léblond[5] $\mathcal{G}_m$), but its middle groups are only two (up to an isomorphism): the direct product $\mathbb{R} \otimes \mathcal{G}^*$ corresponding to $m = 0$ and another group called the "extended Poincaré group," and denoted by $\widetilde{\mathcal{G}}$, for $m \neq 0$ $\mathcal{G}_m = \widetilde{\mathcal{G}}$. From the middle group of each extension labelled by $m \subset \mathbb{R}$ we obtain many more "candidates" to PUR of $\mathcal{G}$ than the true ones. A condition similar to (iii) of Sec. 2 is necessary here, and because of the use of $\mathbb{R}$ and not of $U(1)$ in the extensions, this condition must be somewhat different to (iii) and it is also translated by selecting those UR of each middle group whose corresponding character of the phase subgroup $\mathbb{R}$ is $m = 1$. One must continue as in Varadarajan's method and the absence of m (when it is compared with the method of Sec. 4) is due to the multiplicity of groups (labelled by $m \in \mathbb{R}$) which are needed.

The most peculiar feature is that the "projective covering group" $\widetilde{\mathcal{G}}$ of $\mathcal{G}$ is isomorphic to the "extended Galilei group" ($\widetilde{\mathcal{G}}$ in the Lévy-Léblond's notation). All PUR of $\mathcal{G}$ can be found from the UR of $\widetilde{\mathcal{G}}$ (only one group ! ! !) and therefore of $\widetilde{\mathcal{G}}$, while not only $\widetilde{\mathcal{G}}$ but a set $\widetilde{\mathcal{G}}_m$ with $m \neq 0$, $m \in \mathbb{R}$, and also $\widetilde{\mathcal{G}}_0 \approx \mathbb{R} \otimes \mathcal{G}^*$ are needed for the construction of all PUR of $\mathcal{G}$ in his method. The apparent contradiction is easily clarified by considering that in Lévy-Léblond's method (and also in Bargmann's) a condition (iii) or similar is necessary to assure the existence of a induced projective representation. This condition says: "Pick out from all UR of the middle group of each extension only those mapping the phase subgroup [$\mathbb{R}$ or $U(1)$] in a prescribed manner on $U(\mathcal{H})$." This condition is the one compelling us to take $m = 1$ (η in the Lévy-Léblond paper) and not the fact that "nature chooses a unique and universal value."[20] In the method developed by us no restrictive conditions

appear and the two "multiplicative" and "restrictive" steps are unnecessary, so that this method must be considered simpler than ones preceding.

APPENDIX A: PROJECTIVE UNITARY REPRESENTATIONS

Let $U(\mathcal{H})$ be the unitary group of the Hilbert space $\mathcal{H}$ endowed with the usual topology. The group $U(1)$ is embedded in $U(\mathcal{H})$ as an invariant subgroup.

Definition A-1: We call projective unitary group $PU(\mathcal{H})$ to the factor group $U(\mathcal{H})/U(1)$. It can be endowed with the quotient topology relative to the canonical surjection.

Now, let G be a connected Lie group.

Definition A-2: A unitary representation of G is a continuous homomorphism $G \to U(\mathcal{H})$.

Definition A-3: A projective unitary representation of G is a continuous homomorphism $G \to PU(\mathcal{H})$.

We must remark that each quantum theory having G as symetry group determines an homomorphism of G in the group $\text{Aut}(\mathcal{H})$ of automorphisms of the lattice $\mathcal{L}(\mathcal{H})$ of all subspaces of a Hilbert space $\mathcal{H}$. According to a well-known theorem of Wigner's, if $\dim \mathcal{H} \geq 2$, $\text{Aut}(\mathcal{H})$ is isomorphic to the projective semiunitary group of $\mathcal{H}$, $P\Gamma U(\mathcal{H})$, but if G is a connected Lie group, the image of G under a continuous mapping lies in the connected component containing the identity, and so the elimination from the beginning of the antiunitary ray transformations is not a very restrictive condition when we are considering only connected symetry groups.

A possible relation between PUR and UR can be obtained by lifting: If E is a connected Lie group, p: $E \to G$ a continuous homomorphism, τ a PUR of G and T a UR of E, we will say that T is a lifting of τ if and only if the following diagram is commutative.

$$
\begin{array}{ccccccc}
1 & \longrightarrow & \text{Ker}\, p & \longrightarrow & E & \longrightarrow & G & \longrightarrow & 1 \\
& & \downarrow & & \downarrow T & & \downarrow \tau & & \\
1 & \longrightarrow & U(1) & \longrightarrow & U(\mathcal{H}) & \longrightarrow & PU(\mathcal{H}) & \longrightarrow & 1
\end{array}
$$

We shall also say that τ can be lifted to a UR of E. This definition of lifting of a PUR to a UR is closely related to the lifting of extensions to be defined in the next appendix.

For PUR, the concepts of irreducibility and equivalence are defined in the usual way [there no exists a G-invariant proper projective subspace of $\mathcal{H}$ (projective space associated with $\mathcal{H}$), and there exists an element of $PU(\mathcal{H})$ intertwining the two representations]. We only must call attention to the fact that, while any lifting of an irreducible PUR is also irreducible as UR, two liftings of the same (or equivalent) PUR can be inequivalent. Roughly speaking, irreducibility is not changed by passing from UR to PUR, while equivalences are somewhat weakened.

APPENDIX B: GROUP EXTENSIONS AND LIFTING OF EXTENSIONS

As in the first appendix we will only give some defini-

tions and results. The reader interested in it may find systematic treatments in Refs. 21, 22.

Let G and A be Lie groups, A being Abelian.

Definition B-1: We call central extension of G by A an exact sequence of Lie groups:

$$\xi: 1 \to A \xrightarrow{i} E \xrightarrow{p} G \to 1,$$

where i and p denote continuous homomorphisms and $i(A)$ is contained in the center of E.

We remark that we do not call central extension of G by A to the group E itself because this terminology may be misleading. We call E the "middle group" of the extension ξ. Now, we can define an equivalence relation in the set of all central extensions of G by A and the factor set, denoted $\mathrm{Ext}_0(G, A)$ may be endowed with an Abelian group structure by means of the Baer composition law.[22] Then, this group can be seen to be isomorphic to the second cohomology group $H_0^2(G, A)$ corresponding to the trivial action of G on A.

It is also well known that $H_0^2(G, A)$ is a contravariant functor in the first factor G, that is to say, that every homomorphism $\rho: G' \to G$ defines an homomorphism $\Lambda_{G',\rho}: H_0^2(G, A) \to H_0^2(G', A)$, and therefore another homomorphism named by the same symbol $\Lambda_{G',\rho}: \mathrm{Ext}_0(G, A) \to \mathrm{Ext}_0(G', A)$. The extension $\xi' \in \mathrm{Ext}_0(G', A)$ associated with $\xi \in \mathrm{Ext}_0(G, A)$ is called the lifted extension of ξ by means of ρ. The explicit way to building up this new extension can be found in the Hochschild paper.[23] The two extensions are related by the following commutative diagram,

$$\Lambda_{G',\rho}(\xi): 1 \to A \to F \to G' \to 1$$
$$\parallel \quad \downarrow \quad \downarrow\rho$$
$$\xi: 1 \to A \to E \to G \to 1$$

For Lie algebras analogous definitions can be established. So, we introduce the concepts of Lie algebra extension, and the corresponding equivalence relation. The set $\mathrm{Ext}_0(LG, LA)$ of (equivalence classes of) central extensions of LG by LA may also be endowed with an

Abelian group structure, and so it is isomorphic to $H_0^2(LG, LA)$.[24]

The groups $Z_0^2(LG, LA)$ and $B_0^2(LG, LA)$ of "cocyles" and "coboundaries" are related to $H_0^2(LG, LA)$ as in the group case in such a way that $H_0^2(LG, LA) = Z_0^2(LG, LA)/B_0^2(LG, LA)$.

*In part from the Tesina de Licenciatura (Ref. 13).

[1] I. Schur, J. Reine Angew. Math. 127, 20 (1904); 132, 85 (1907).
[2] E. P. Wigner, Ann. Math. 40, 149 (1939).
[3] V. Bargmann, Ann. Math. 59, 1 (1954).
[4] V. S. Varadarajan, *Geometry of Quantum Theory Vol. II* (Van Nostrand, Princeton, N. J., 1970).
[5] J. M. Lévy-Léblond, "Galilei group and Galilean invariance," in *Group Theory and Its Applications*, Vol. II, edited by E. M. Loebl (Academic, New York, 1971).
[6] See Ref. 5, p. 251.
[7] See Ref. 4, p. 111, Theor. 10.16 and Chap. XII, Sec. 8.
[8] G. W. Mackey, Acta Math. 99, 265 (1958).
[9] G. Hochschild, Ann. Math. 54, 96 (1951).
[10] See Ref. 9, p. 101, Theor. 2.1.
[11] L. Michel, *Invariance in Quantum Mechanics and Group Extensions*, Istambul Lectures (Gordon and Breach, New York, 1964), p. 157, Lemma 3.
[12] We must remark that in Ref. 10 there is an errata concerning these commutation relations.
[13] M. Santander, Tesina de Licenciatura: *Extensiones de Grupos y Simetrías en Mecánica Cuántica*, Universidad de Madrid, 1974.
[14] See, e.g., K. R. Parthasarathy, *Multipliers on Locally Compact Groups*, Lecture Notes in Math., Vol. 93 (Springer, New York, 1969).
[15] See Ref. 5, p. 239.
[16] See Ref. 5, p. 232.
[17] D. J. Simms, *Lie Groups and Quantum Mechanics*, Lecture Notes in Math., Vol. 52 (Springer, New York, 1968).
[18] G. W. Mackey, Bull. Am. Math. Soc. 69, 628 (1963).
[19] See Ref. 5, p. 252.
[20] See Ref. 5, p. 273.
[21] S. MacLane, *Homology* (Springer, New York, 1963).
[22] L. Michel, *Relativistic Invariance and Internal Symmetries*, Brandeis Lectures 1965, (Gordon and Breach, New York, 1966); see also Ref. 11.
[23] See Ref. 9, Sec. 1.
[24] See, e.g., Ref. 9, Sec. 2; also Ref. 4, Chap. X, Sec. 4.

Neutron transport and diffusion in inhomogeneous media. I*

Edward W. Larsen

Courant Institute of Mathematical Sciences, New York University, New York, New York 10012
(Received 18 November 1974)

The asymptotic solution of the neutron transport equation is obtained for large near-critical domains D which possess a cellular, nearly periodic structure. A typical mean free path in D is taken to be of the same order of magnitude as a cell diameter, and these are taken to be small (of order ϵ) compared to a typical diameter of D. The solution is asymptotic with respect to the small parameter ϵ. It is a product of two functions, one determined by a detailed cell calculation and the other obtained as the solution of a time dependent diffusion equation. The diffusion equation contains precursor (delayed neutron) densities, equations for which are derived. The coefficients in the diffusion equation, which are determined using the results of the cell calculation, differ from those now used in engineering applications. The initial condition for the diffusion equation is derived, and the problem of determining the boundary condition is discussed.

INTRODUCTION

Recently, the solution of general neutron transport problems has been obtained for domains D which are large compared to a mean free path, and in which the material properties vary slowly with respect to a mean free path and all secondary neutrons are "prompt."[1-5] The solutions are asymptotic with respect to a small dimensionless parameter ϵ, defined as the ratio of a typical mean free path to a typical dimension of D. For near-critical domains, a diffusion equation with initial and boundary conditions is derived which describes the neutron density ψ several mean free paths away from the boundaries of D, and several collisions after the initial time.

In most reactor cores, however, inhomogeneities in space do not occur "slowly" with respect to a mean free path, and not all secondary neutrons are prompt. A reactor core typically consists of a large periodic array of identical hexagonal or rectangular cells, each cell being only a few mean free paths thick; within each cell the material properties undergo large discontinuities, for instance in passing from a fuel rod to the moderator. In addition, the ratio of delayed to prompt neutrons is small (typically about 1%), but for near-critical domains these delayed neutrons can have a large effect on the neutron density.[6] For such cases, delayed neutrons cannot be realistically ignored.

To account for these effects we shall consider in this paper a near-critical domain with delayed neutrons and nearly periodic, rapidly varying spatial inhomogeneities, and we shall generalize the asymptotic results derived earlier to this case. In particular, we take D to be a large domain consisting of comparatively small, nearly identical cells, arranged periodically. We require a typical mean free path in D to be the same order of magnitude as a cell diameter, and these to be small (of order ϵ) compared to a typical diameter of D. Also we require the ratio of delayed to prompt secondary neutrons to be small (of order ϵ^2).

The main result of this paper is that beyond several mean free paths away from the boundary of D and several collison times after the initial time $t' = 0$, the solution ψ of the neutron transport equation is given by an asymptotic "interior" solution ψ^i of the form

$$\psi^i = A_0(\mathbf{r}, \epsilon t')\phi(\mathbf{r}, \epsilon^{-1}\mathbf{r}, \mathbf{v}, \epsilon t') + O(\epsilon). \tag{A}$$

In (A), ϕ is the solution of the single cell problem described in Sec. 3 and A_0 satisfies the diffusion equation (3.16). If, to leading order, the cells comprising D are identical and their material properties are independent of time, then ϕ is independent of $\mathbf{r}$ and of $\epsilon t'$ and is periodic in $\epsilon^{-1}\mathbf{r}$ across each cell. For this case, (A) shows that the velocity dependence and the periodic fast spatial dependence of ψ^i are contained in ϕ, and are separated from and modulated by the slow time and slow spatial dependence contained in A_0. (This specific case will be described in detail in Sec. 1.) If the material properties of D have no time and fast spatial variation and if there are no delayed neutrons, then the problem for ϕ and the diffusion equation for A_0 simplify to those described in Ref. 3.)

It is hoped that the methods in this paper, which as in Ref. 3 are easily specialized to multigroup or one-speed neutron transport, will be useful in solving realistic reactor core problems. Diffusion equations similar to ours have been proposed as approximate governing equations for reactor cores,[6] and are in common use in many computer codes. However, our equations are derived directly from the neutron transport equation, our method systematically yields the correction terms of order ϵ and higher, and our diffusion coefficients are different from those previously proposed. In addition, our method of derivation takes into account the rapid spatial variation in a cell and gives directly the effective diffusion coefficients, "homogenized" across a cell. In other methods, the effective homogenized coefficients are arrived at heuristically.[7,8] We shall numerically compare our diffusion coefficients to others in common use in a second paper.[9]

The plan of this paper is as follows. Section 1 is a detailed summary of the results of this paper for the case of a domain D whose material properties are, to $O(\epsilon)$, periodic in space and independent of time. These results are then derived for a more general domain D in Secs. 2 through 7.

In Sec. 2 we completely formulate the general neutron transport problem under consideration and state certain assumptions we find necessary to impose on the cell

structure. In Sec. 3 we derive the interior solution ψ^i and the Eqs. (3.16), (3.22) governing A_0. In Sec. 4 we give some results about the coefficients in the diffusion equation. In Sec. 5 we derive the initial conditions for A_0 and the initial layer ψ^0 which describes the transition from the initial condition to the interior or diffusion solution described by ψ^i. ψ^0 decays exponentially with time after the initial time, the decay rate being fast of order ϵ^{-1}. In Sec. 6 we discuss boundary conditions for A_0. In Sec. 7 we construct solutions of the asymptotic equations (3.16), (3.22), and we discuss the critical problem for D.

1. SUMMARY OF RESULTS

In this section we shall summarize our results for a problem of practical interest. We consider a domain D in which (a) the ratio of a typical mean free path to a cell diameter is of order one, (b) the ratio of a typical mean free path to a typical dimension of D is small of order ϵ, (c) the material properties of D are periodic in space and independent of time except for a pertur-bation term of order ϵ^2, and (d) the interior sources and delayed neutron terms are small of order ϵ. Then the neutron transport equation may be written in scaled form as

$$\left(\frac{1}{\epsilon}\frac{\partial}{\partial \tau} + \mathbf{v}\cdot\nabla - \frac{1}{\epsilon}L_0 - \epsilon L_2\right)\tilde{\psi} = \epsilon\sum_i \lambda_i \tilde{C}_i + \epsilon S, \qquad (1.1)$$

$$\frac{1}{\epsilon}\frac{\partial}{\partial \tau}\tilde{C}_i = \epsilon\int \theta_i \tilde{\psi}\, d\mathbf{v}' - \epsilon\lambda_i \tilde{C}_i, \qquad (1.2)$$

where L_0 and L_2 are "prompt" collision operators of the form

$$L_n\tilde{\psi} = \int v'\sigma_{S_n}\tilde{\psi}\, d\mathbf{v}' - v\sigma_{T_n}\tilde{\psi}, \quad n = 0, 2. \qquad (1.3)$$

In (1.1)—(1.3), the space variable $\mathbf{r}$ is stretched so that a typical dimension of D is of order 1. We introduce the fast space variable $\mathbf{r}' = \epsilon^{-1}\mathbf{r}$, in terms of which a typical cell diameter and mean free path are of order one; we also introduce the slow time variable $t = \epsilon^2\tau$. Then in (1.3) we express σ_{S0} and σ_{T0} as functions of velocity and as periodic functions of $\mathbf{r}'$ across a cell, but we require σ_{S0} and σ_{T0} to be independent of t, τ, and $\mathbf{r}$. (Thus L_0 depends parametrically upon $\mathbf{r}'$.) Also in (1.3), we express σ_{S2} and σ_{T2} as functions of $\mathbf{r}$, t, velo-city, and as periodic functions of $\mathbf{r}'$ across a cell, but we require σ_{S2} and σ_{T2} to be independent of τ. (Thus L_2 depends parametrically upon $\mathbf{r}'$, $\mathbf{r}$, and t.)

Next, we require that the equation

$$0 = T\phi(\mathbf{r}', \mathbf{v}) \equiv (\mathbf{v}\cdot\nabla' - L_0)(\mathbf{r}', \mathbf{v}) \qquad (1.4)$$

have a unique normalized solution ϕ which is positive and periodic across a cell. [In (1.4), ∇' acts on $\mathbf{r}'$.]

Under the above conditions (and in fact under more general conditions described in Sec. 2), we derive the following result in Sec. 3: As $\epsilon \to 0$, the solution $\tilde{\psi}$ of (1.1), (1.2) tends, away from the boundaries of D and the initial time, to an "interior" solution ψ^i of the form

$$\psi^i = A_0(\mathbf{r}, t)\phi(\epsilon^{-1}\mathbf{r}, \mathbf{v}) + O(\epsilon), \qquad (1.5)$$

where A_0 satisfies the equations

$$\frac{\partial A_0}{\partial t} = \nabla\cdot/\eta_0\cdot\nabla A_0 + M_2 A_0 + \sum_i \lambda_i Q_i + S_0, \qquad (1.6)$$

$$\frac{\partial Q_i}{\partial t} = \beta_i A_0 - \lambda_i Q_i. \qquad (1.7)$$

In (1.6) and (1.7), $/\eta_0$ and M_2 are defined by

$$/\eta_0 = \int_C \int \mathbf{v}\phi^* T^{-1}\phi\mathbf{v}\, d\mathbf{v}\, d\mathbf{r}', \qquad (1.8)$$

$$M_2(\mathbf{r}, t) = \int_C \int \phi^* L_2\phi\, d\mathbf{v}\, d\mathbf{r}', \qquad (1.9)$$

where C denotes a cell and Q_i, S_0, and β_i are certain moments of $\tilde{C}_i$, S, and θ_i which are described in Sec. 3.

In Eq. (1.8), T^{-1} is the pseudo-inverse of T and ϕ^* is the adjoint eigenfunction corresponding to (1.4).

Since ϕ^* and $T^{-1}\phi\mathbf{v}$ are functions only of $\mathbf{r}'$ and $\mathbf{v}$, $/\eta_0$ is a constant matrix. However, since L_2 depends parametrically upon $\mathbf{r}$ and t, M_2 is in general a function of $\mathbf{r}$ and t.

For rectangular or hexagonal cylindrical cells pos-sessing certain symmetries, we show in Sec. 4 that $/\eta_0$ is a diagonal matrix, provided the coordinate axes are correctly oriented.

The initial condition for A_0 is derived in Sec. 5, along with the initial layer solution of the transport equation, which describes the transition from the initial state to the "diffusion state" described by ψ^i.

The problem to determine the boundary conditions for A_0 is discussed in Sec. 6. If D is surrounded by a vacu-um or by a perfectly absorbing region, then it is asymp-totically correct to take $A_0 = 0$ on the boundary of D.

If the domain D consists of a nearly homogeneous mat-erial, then all quantities in the above equations are in-dependent of the fast space variable $\mathbf{r}'$. However, we may still allow L_2 to depend parametrically upon $\mathbf{r}$ and t, to account for a small, slowly varying (in space and time) perturbation from exact homogeneity.

For this case, the operator T in (1.4) reduces to $T = -L_0$, and we show in Sec. 4 that the eigenfunction ϕ must have the form

$$\phi = \phi(v), \qquad (1.10)$$

Then ψ^i is given by (1.5) with ϕ of the form (1.10), and again Eqs. (1.6)—(1.9) apply; we also show in Sec. 4 that $/\eta_0$ is proportional to the identity matrix:

$$/\eta_0 = M_0\mathcal{I}.$$

Thus the second derivative operator in (1.6) reduces to the usual Laplacian operator.

The above paragraphs describe our results for the given special domain D. In the remaining sections of this paper, we shall derive the above results for more general domains D. In these domains condition (c), given in the first paragraph of this section, is replaced by a weaker and more general condition which is described in Sec. 2.

2. FORMULATION

We consider the neutron transport equation, expressed in terms of scaled quantities by

$$\left(\frac{1}{\epsilon}\frac{\partial}{\partial \tau}+\mathbf{v}\cdot\nabla-\frac{1}{\epsilon}L\right)\tilde{\psi}(\mathbf{r},\mathbf{v},\tau,\epsilon)$$

$$=\sum_i \epsilon\lambda_i\tilde{C}_i(\mathbf{r},\mathbf{v},\tau,\epsilon)+\epsilon S(\mathbf{r},\epsilon^{-1}\mathbf{r},\mathbf{v},\epsilon^2\tau). \qquad (2.1)$$

Here ϵS is the source and $\epsilon^{-1}L$ is the prompt collision operator, defined by

$$L\tilde{\psi}(\mathbf{r},\mathbf{v},\tau,\epsilon)=\int v'\sigma_S(\mathbf{r},\epsilon^{-1}\mathbf{r},\mathbf{v}'\to\mathbf{v},\epsilon^2\tau,\epsilon)\tilde{\psi}(\mathbf{r},\mathbf{v}',\tau,\epsilon)\,dv'$$

$$-v\sigma_T(\mathbf{r},\epsilon^{-1}\mathbf{r},v,\epsilon^2\tau,\epsilon)\tilde{\psi}(\mathbf{r},\mathbf{v},\tau\epsilon). \qquad (2.2)$$

$\tilde{C}_i$ is the precursor density of species i. For stationary fuel reactors, $\tilde{C}_i$ satisfies the o.d.e.

$$\frac{1}{\epsilon}\frac{\partial \tilde{C}_i}{\partial \tau}(\mathbf{r},\mathbf{v},\tau,\epsilon)=\epsilon\int \theta_i(\mathbf{r},\epsilon^{-1}\mathbf{r},\mathbf{v}'\to\mathbf{v},\epsilon^2\tau)\tilde{\psi}(\mathbf{r},\mathbf{v}',\tau,\epsilon)\,dv'$$

$$-\epsilon\lambda_i\tilde{C}_i(\mathbf{r},\mathbf{v},\tau,\epsilon). \qquad (2.3)$$

For moving (fluid) fuel reactors, $\tilde{C}_i$ will satisfy a more general equation than (2.3). Our method can be used to treat such a case, but we shall not do so here.

In the above equations, the scattering and total cross sections are written as $\epsilon^{-1}\sigma_S$ and $\epsilon^{-1}\sigma_T$ to indicate that the mean free path is small of order ϵ. The time variable τ is scaled so that the time to traverse one mean free path is of order unity. τ is defined in terms of unscaled time t' by $\tau=\epsilon^{-1}t'$. The decay constants and source are written as $\epsilon\lambda_i$ and ϵQ to indicate that they are small.

In addition, the scaled source S and cross sections σ_S, σ_T, θ_i are written as functions of $\epsilon^{-1}\mathbf{r}$ and $\epsilon^2\tau$ since these functions have rapid spatial variations on the order of a mean free path, and slow time variation with respect to τ. To indicate that σ_S and σ_T have Taylor expansions in their last variable, we write them as

$$\sigma_S(\mathbf{r},\epsilon^{-1}\mathbf{r},\mathbf{v}'\to\mathbf{v}',\epsilon^2\tau,\epsilon)=\sum_{n=0}^{\infty}\epsilon^n\sigma_{Sn}(\mathbf{r},\epsilon^{-1}\mathbf{r},\mathbf{v}'\to\mathbf{v},\epsilon^2\tau), \qquad (2.4)$$

$$\sigma_T(\mathbf{r},\epsilon^{-1}\mathbf{r},v,\epsilon^2\tau,\epsilon)=\sum_{n=0}^{\infty}\epsilon^n\sigma_{Tn}(\mathbf{r},\epsilon^{-1}\mathbf{r},v,\epsilon^2\tau). \qquad (2.5)$$

We require the cells which comprise D to possess certain symmetries. In particular, we require each cell to have a center (say $\mathbf{r}=0$) about which σ_S and σ_T are to order ϵ symmetric functions of position. We also require σ_S and σ_T to be rotationally symmetric in $\mathbf{v}$ and $\mathbf{v}'$, to order ϵ. Then, with $\mathbf{r}'=\epsilon^{-1}\mathbf{r}$ and $t=\epsilon^2\tau$,

$$\sigma_{Sn}(\epsilon\mathbf{r}',\mathbf{r}',\mathbf{v}'\to\mathbf{v},t)=\sigma_{Sn}(-\epsilon\mathbf{r}',-\mathbf{r}',-\mathbf{v}'\to-\mathbf{v},t), n=0,1, \qquad (2.6)$$

and

$$\sigma_{Tn}(\epsilon\mathbf{r}',\mathbf{r}',v,t)=\sigma_{Tn}(-\epsilon\mathbf{r}',-\mathbf{r}',v,t),\quad n=0,1. \qquad (2.7)$$

Also, we require σ_S and σ_T to be to leading order periodic across each cell. We meet this condition by requiring σ_{S0} and σ_{T0} in (2.6) and (2.7) to be periodic in

$\mathbf{r}'$ across each cell for $\epsilon=0$. Then if $\mathbf{r}'=0$ is the center of a given cell C and $\mathbf{r}'=\mathbf{a}$ is the center of any other cell, we have

$$\sigma_{S0}(0,\mathbf{r}',\mathbf{v}'\to\mathbf{v},t)=\sigma_{S0}(0,\mathbf{r}'+\mathbf{a},\mathbf{v}'\to\mathbf{v},t),\quad \mathbf{r}'\in C, \qquad (2.8)$$

and

$$\sigma_{T0}(0,\mathbf{r}',v,t)=\sigma_{T0}(0,\mathbf{r}'+\mathbf{a},v,t),\qquad \mathbf{r}'\in C. \qquad (2.9)$$

Hereafter, any function of $\mathbf{r}'$ satisfying this condition will be called *periodic*.

To solve (2.1)—(2.5), we define the fast position variable $\mathbf{r}'$ by $\mathbf{r}'=\epsilon^{-1}\mathbf{r}$, and we express $\tilde{\psi}$ and $\tilde{C}_i$ as functions of $\mathbf{r}$ and $\mathbf{r}'$:

$$\tilde{\psi}(\mathbf{r},\mathbf{v},\tau,\epsilon)=\psi(\mathbf{r},\mathbf{r}',v,\tau,\epsilon),$$

$$\tilde{C}_i(\mathbf{r},\mathbf{v},\tau,\epsilon)=C_i(\mathbf{r},\mathbf{r}',\mathbf{v},\tau,\epsilon).$$

Then $\mathbf{v}\cdot\nabla\tilde{\psi}=(\mathbf{v}\cdot\nabla+\epsilon^{-1}\mathbf{v}\cdot\nabla')\psi$, where ∇ and ∇' on the right side of this equation act on $\mathbf{r}$ and $\mathbf{r}'$, respectively. If we define L_n to be the operator in (2.2) with σ_S and σ_T replaced respectively by σ_{Sn} and σ_{Tn}, and we replace $\epsilon^{-1}\mathbf{r}$ by $\mathbf{r}'$ in (2.1)—(2.5), then the equations for ψ and C_i become

$$\left(\epsilon^{-1}\frac{\partial}{\partial \tau}+\mathbf{v}\cdot\nabla+\epsilon\mathbf{v}\cdot\nabla'-\sum_{n=0}^{\infty}\epsilon^{n-1}L_n\right)\psi$$

$$=\epsilon\left(\sum_i\lambda_iC_i+S\right) \qquad (2.10)$$

and

$$\frac{\partial C_i}{\partial t}=\epsilon^2\left(\int \theta_i\psi\,d\mathbf{v}'-\lambda_iC_i\right). \qquad (2.11)$$

For the purposes of analysis, we shall regard $\mathbf{r}$ and $\mathbf{r}'$ as independent variables. To do this, it is necessary to extend the known functions σ_{Sn}, σ_{Tn}, S, and θ_i to be defined for $\mathbf{r}'\ne\epsilon^{-1}\mathbf{r}$. We choose these extensions in such a way that for each $\mathbf{r}$ they are periodic in $\mathbf{r}'$. If the original functions σ_S, σ_T, S, and θ_i are exactly periodic, then σ_{Sn}, σ_{Tn}, S, and θ_i are independent of $\mathbf{r}$ and the above extensions need not be constructed.

3. THE INTERIOR SOLUTION

To determine a solution ψ^i of (2.10) which is appreciable throughout D, we introduce the slow time variable $t=\epsilon^2\tau$ in (2.2)—(2.5) and we assume that ψ and C_i possess asymptotic power series expansions in ϵ:

$$\psi^i\sim\sum_{n=0}^{\infty}\epsilon^n\psi_n(\mathbf{r},\mathbf{r}',\mathbf{v},t), \qquad (3.1)$$

$$C_i\sim\sum_{n=0}^{\infty}\epsilon^nC_{i,n}(\mathbf{r},\mathbf{r}',\mathbf{v},t). \qquad (3.2)$$

Introducing (3.1) and (3.2) into (2.10) and (2.11), using $\partial/\partial\tau=\epsilon^2\partial/\partial t$, and equating the coefficients of different powers of ϵ, we obtain the sequence of equations

$$T\psi_n=\sum_{j=0}^{n-1}L_{n-j}\psi_j-\mathbf{v}\cdot\nabla\psi_{n-1}-\frac{\partial\psi_{n-2}}{\partial t}+\sum_i\lambda_iC_{i,n-2}+\delta_{n2}S, \qquad (3.3)$$

$$\frac{\partial C_{i,n}}{\partial t} = \int \theta_i \psi_n \, dv' - \lambda_i C_{i,n}, \tag{3.4}$$

where

$$T = \mathbf{v} \cdot \nabla' - L_0, \tag{3.5}$$

and where $\psi_{-1} = \psi_{-2} = C_{i,-1} = C_{i,-2} = 0$. Equation (3.3) with $n = 0$ is $T\psi_0 = 0$, which has the general solution

$$\psi_0(\mathbf{r}, \mathbf{r}', \mathbf{v}, t) = A_0(\mathbf{r}, t)\phi(\mathbf{r}, \mathbf{r}', \mathbf{v}, t). \tag{3.6}$$

Here ϕ satisfies $T\phi = 0$, i.e.,

$$0 = \mathbf{v} \cdot \nabla'\phi(\mathbf{r}, \mathbf{r}', \mathbf{v}, t) + v\sigma_{T0}(\mathbf{r}, \mathbf{r}', v, t)\phi(\mathbf{r}, \mathbf{r}', \mathbf{v}, t)$$

$$- \int v'\sigma_{S0}(\mathbf{r}, \mathbf{r}', \mathbf{v}' \to \mathbf{v}, t)\phi(\mathbf{r}, \mathbf{r}', \mathbf{v}', t) \, dv'. \tag{3.7}$$

In this equation $\mathbf{r}$ and t occur as parameters and for fixed values of $\mathbf{r}$ and t, the scaled cross sections σ_{T0} and σ_{S0} are periodic in $\mathbf{r}'$. We require that the solution ϕ of (3.7) be unique (up to a multiplicative constant), positive, and periodic in $\mathbf{r}'$.

These are conditions on σ_{T0} and σ_{S0}. They mean physically that if the domain D were extended periodically to fill R^3, then R^3 would be exactly critical with respect to σ_{T0} and σ_{S0}. This is because $\phi(\mathbf{r}, \mathbf{r}', \mathbf{v}, t)$, extended periodically in $\mathbf{r}'$, would be a positive, bounded solution of the sourceless transport equation (3.7) on R^3 (in which t, we recall, appears only parametrically). Our problem for ψ^i is thus a perturbation of this infinite space problem in the following sense: Our domain D is large but not infinite, our cross sections are almost but not exactly periodic, and our problem is not stationary but depends slowly on time. To leading order, the perturbation about the infinite space problem is accounted for by the undetermined function A_0 in (3.6).

To make ϕ unique, we impose the normalization

$$1 = \left[\int_C d\mathbf{r}' \right]^{-1} \int_C \int \phi(\mathbf{r}, \mathbf{r}', \mathbf{v}, t) \, d\mathbf{v}' \, d\mathbf{r}', \tag{3.8}$$

where C is a cell.

Next we consider Eq. (3.3) with $n = 1$:

$$T\psi_1 = -A_0(L_1\phi) - \mathbf{v} \cdot \nabla(A_0\phi). \tag{3.9}$$

To solve this equation, we consider the Banach space X of functions $f(\mathbf{r}', \mathbf{v})$, periodic in $\mathbf{r}'$, with norm

$$\|f\| = \left[\int_C d\mathbf{r}' \right]^{-1} \int_C \int |f(\mathbf{r}', \mathbf{v})| \, d\mathbf{v} \, d\mathbf{r}'.$$

In this space the operator $T = \mathbf{v} \cdot \nabla' - L_0$ has, by the above constraints on σ_{T0} and σ_{S0}, the eigenvalue $\lambda = 0$ with geometrical multiplicity one. Thus the conjugate operator $T^* = -\mathbf{v} \cdot \nabla' - L_0^*$ has the eigenvalue $\lambda = 0$ with geometrical multiplicity one and with eigenfunction ϕ^*. We assume that $\lambda = 0$ is an isolated point eigenvalue of T, and we normalize ϕ^* by

$$1 = \int_C \int \phi^*\phi \, d\mathbf{v} \, d\mathbf{r}'. \tag{3.10}$$

Then by the alternative theorem,[10] Eq. (3.9) has a solution in X (i.e., periodic in $\mathbf{r}'$) if and only if the solvability condition

$$0 = A_0 \int_C \int \phi^* L_1\phi^* \, d\mathbf{v} \, d\mathbf{r}' + \int_C \int \phi^* \mathbf{v} \cdot \nabla A_0\phi \, d\mathbf{v} \, d\mathbf{r}' \tag{3.11}$$

is satisfied. Because of (2.6) and (2.7),

$$\phi(\mathbf{r}, \mathbf{r}', \mathbf{v}, t) = \phi(\mathbf{r}, -\mathbf{r}', -\mathbf{v}, t) \tag{3.12}$$

(where $\mathbf{r}' = 0$ is the center of a cell), and this relation holds also for ϕ^*. Thus the second integral in (3.11) is zero, and the solvability condition (3.11) for nonzero A_0 reduces to

$$0 = \int_C \int \phi^* L_1\phi \, d\mathbf{v} \, d\mathbf{r}'. \tag{3.13}$$

This is a condition on L_1 which is met, for instance, by $L_1 = 0$.

The general solution of (3.9) may now be written

$$\psi_1 = A_1\phi - A_0 T^{-1}L_1\phi - T^{-1}\mathbf{v} \cdot \nabla A_0\phi. \tag{3.14}$$

Here T^{-1} is the unique pseudo-inverse of T satisfying

$$0 = \int_C \int \phi^* T^{-1}f \, d\mathbf{v} \, d\mathbf{r}'$$

for any function $f(\mathbf{r}', \mathbf{v}) \in X$ such that

$$0 = \int_C \int \phi^* f \, d\mathbf{v} \, d\mathbf{r}'.$$

To proceed, we consider Eq. (3.3) with $n = 2$:

$$T\psi_2 = A_0 L_2\phi + (L_1 - \mathbf{v} \cdot \nabla)(A_1\phi - T^{-1}\mathbf{v} \cdot \nabla A_0\phi - A_0 T^{-1}L_1\phi)$$

$$- \frac{\partial}{\partial t}(A_0\phi) + \sum_i \lambda_i C_{i,0} + S. \tag{3.15}$$

The solvability condition for this equation is, using (3.10) and rearranging,

$$\frac{\partial A_0}{\partial t} = \nabla \cdot (\not{n}_0 \cdot \nabla A_0) + \mathbf{M}_1 \cdot \nabla A_0 + M_2 A_0 + \sum_i \lambda_i Q_i + S_0, \tag{3.16}$$

where

$$\not{n}_0(\mathbf{r}, t) = \int_C \int [\mathbf{v}\phi^* T^{-1}\phi\mathbf{v}] \, d\mathbf{v} \, d\mathbf{r}', \tag{3.17}$$

$$\mathbf{M}_1(\mathbf{r}, t) = \int_C \int [\mathbf{v}\phi^* T^{-1}(\mathbf{v} \cdot \nabla\phi) - (\mathbf{v} \cdot \nabla\phi^*)T^{-1}\mathbf{v}\phi] \, d\mathbf{v} \, d\mathbf{r}', \tag{3.18}$$

$$M_2(\mathbf{r}, t) = \int_C \int \left[\phi^* \mathbf{v} \cdot \nabla T^{-1}\mathbf{v} \cdot \nabla\phi + \phi^* L_2\phi \right.$$

$$\left. - \phi^* L_1 T^{-1}L_1\phi - \phi^* \frac{\partial\phi}{\partial t} \right] d\mathbf{v} \, d\mathbf{r}', \tag{3.19}$$

$$Q_i(\mathbf{r}, t) = \int_C \int \phi^* C_{i,0} \, d\mathbf{v} \, d\mathbf{r}', \tag{3.20}$$

$$S_0(\mathbf{r}, t) = \int_C \int \phi^* S \, d\mathbf{v} \, d\mathbf{r}'. \tag{3.21}$$

The equation for Q_i, obtained from (3.20) and (3.4), is

$$\frac{\partial Q_i}{\partial t} = \beta_i A_0 - \lambda_i Q_i, \tag{3.22}$$

where

$$\beta_i(\mathbf{r}, t) = \int_C \int \phi^*(\mathbf{r}, \mathbf{r}', \mathbf{v}, t)$$

$$\times \int \theta_i(\mathbf{r}, \mathbf{r}', \mathbf{v}' \to \mathbf{v}, t)\phi(\mathbf{r}, \mathbf{r}', \mathbf{v}', t) \, d\mathbf{v}' \, d\mathbf{v} \, d\mathbf{r}'. \tag{3.23}$$

In deriving (3.16) we used the results

$$0 = \int_C \int \mathbf{v}\phi^* T^{-1}L_1\phi \, dv \, d\mathbf{r}',$$

$$0 = \int_C \int \mathbf{v}\phi (T^*)^{-1} L_1^* \phi^* \, dv \, d\mathbf{r}',$$

$$0 = \int_C \int (\mathbf{v}\cdot\nabla\phi)(T^*)^{-1}L_1^*\phi^* \, dv \, d\mathbf{r}',$$

$$0 = \int_C \int \phi^*\mathbf{v}\cdot\nabla T^{-1}L_1\phi \, dv \, d\mathbf{r}'.$$

The first of these holds because $L_1\phi$ satisfies (3.12), so by symmetry $T^{-1}L_1\phi$ along with ϕ^* satisfies (3.12). The other integrals are zero for the same reason.

Equation (3.16) is a diffusion equation for A_0 containing the precursor (delayed neutron) densities Q_i, which satisfy Eqs. (3.22). If (3.16) is satisfied, then ψ_2 can be determined from (3.15). A diffusion equation for A_1 can now be determined as the solvability condition for equation (3.3) with $n = 3$. This equation will contain new precursor densities, and equations for these can be obtained from (3.4) just as was done above for Q_i. This procedure can be continued recursively to generate equations for any A_n, $n \ge 0$, along with the appropriate precursor densities. However, we shall not do so here.

4. THE COEFFICIENTS OF THE DIFFUSION EQUATION

In this section we shall give some results about the form of the coefficients $\mathcal{M}_0$, $\mathbf{M}_1$, and M_2 in Eq. (3.16) for certain cases of physical or computational interest. In the first three cases considered, we make certain assumptions about the dependence of the material properties of D on the variables t, $\mathbf{r}$, and $\mathbf{r}'$ to various orders. In the fourth through sixth cases, we assume that the cells are long cylinders with material properties which are invariant with respect to certain reflections and rotations. In the seventh and final case, we assume that the cells are thin slabs.

1. The material properties of D are independent of t to O(1)

First, let us consider a domain D in which the changes in material properties with respect to t are of order ϵ. Then σ_{T0}, σ_{S0}, ϕ, ϕ^*, $\mathcal{M}_0$, and $\mathbf{M}_1$ are independent of t. The coefficient M_2, which is related to the "local" reactivity, is in general time dependent due to its dependence on L_1 and L_2. We note that θ_i could have been expanded as a power series in ϵ, and if so the first term in its expansion would occur in (3.23). If this were time independent, then β_i would be time independent also.

This case of small changes in time of the material properties is of computational interest because it has general validity, and for it the problem for ϕ becomes one in which t does not even occur parametrically.

2. The material properties of D are independent of r and t to O(1) and to O(ε²)

Next we consider a domain D in which the slow time and spatial variation in σ_S and σ_T is of order ϵ. Then σ_{S0} and σ_{T0} are independent of $\mathbf{r}$ and t. Hence ϕ and ϕ^* are independent of $\mathbf{r}$ and t, $\mathcal{M}_0$ is a constant, $\mathbf{M}_1 = 0$, and the first term in the integrand of M_2 is zero. If also σ_{Sn} and σ_{Tn}, $n = 1, 2$, are independent of $\mathbf{r}$ and t

then M_2 is a constant. This last case occurs for example if σ_S and σ_T are exactly periodic in space and independent of time.

3. The material properties of D are independent of r' to O(1) and to O(ε²)

Now we consider a domain D in which the fast spatial variation in σ_S and σ_T is of order ϵ. For this case σ_{S0} and σ_{T0} are independent of $\mathbf{r}'$ and so ϕ and ϕ^* are independent of $\mathbf{r}'$. It is shown in Ref. 3 that $\phi = \phi(\mathbf{r}, v, t)$ and $\phi^* = \phi^*(\mathbf{r}, v, t)$, where $\mathbf{v} = v\Omega$. Also, $(T^{-1}\phi\mathbf{v})(\mathbf{r}, \mathbf{v}, t) = \Omega(T_1^{-1}\phi v)(\mathbf{r}, v, t)$ where T_1 is the operator

$$(T_1 f)(\mathbf{r}, v, t) = v\sigma_{T0}(\mathbf{r}, v, t)f(\mathbf{r}, v, t)$$
$$- \int v'\sigma_{S0}^1(\mathbf{r}', v' \to v, t)f(\mathbf{r}, v', t)(v')^2 \, dv'.$$

In this equation σ_{S0}^1 is the $n = 1$ coefficient in the Legendre polynomial expansion of σ_{S0}:

$$\sigma_{S0}(\mathbf{r}, \mathbf{v}' \to \mathbf{v}, t) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi}\sigma_{S0}^n(\mathbf{r}, v' \to v, t)P_n(\Omega\cdot\Omega').$$

Equation (3.17) now yields

$$\mathcal{M}_0(\mathbf{r}, t) = \mathcal{Y}\left(\frac{4\pi}{3}\int_C d\mathbf{r}'\right)\int v\phi^*(\mathbf{r}, v, t)(T_1^{-1}\phi v)(\mathbf{r}, v, t)v^2 \, dv. \quad (4.1)$$

Thus if σ_{S0} and σ_{T0} are independent of $\mathbf{r}'$, then $\mathcal{M}_0$ is proportional to the identity matrix.

If all quantities through order ϵ^2 are independent of $\mathbf{r}'$, which occurs if no fast variation is present, then $\mathcal{M}_0$ is given by (4.1) and the integrations over C in (3.16)–(3.23) are superfluous. Also, if σ_{T0} and σ_{S0} are independent of t, then our results agree with those stated in Ref. 3 in different notation.

4. D consists of long cylindrical cells

Next we consider a three-dimensional domain D consisting of long slender cells. This configuration occurs in many reactor cores. If the cells are oriented in the direction of the coordinate axis of r_3, then σ_S and σ_T are functions of r_1' and r_2', but not of r_3'. Thus the solution ϕ of (3.5) is independent of r_3' and is an even function of v_3. For this case, the 3×3 matrix $\mathcal{M}_0$ reduces to a $(2\times2)\times(1\times1)$ matrix in the following sense: If $\mathbf{e}_j$ is a coordinate vector in the direction corresponding to r_j, then $M_{jk} \equiv \mathbf{e}_j \cdot \mathcal{M}_0 \cdot \mathbf{e}_k = 0$ for $j = 3$ and $k = 1, 2$, and conversely.

To see this, we observe from (3.17) that

$$M_{jk} = \int_C \int v_j\phi^* T^{-1}\phi v_k \, dv \, d\mathbf{r}'.$$

If $j = 3$ and $k = 1$ or 2, then $T^{-1}\phi v_k$ is even in v_3 and the integrand is an odd function of v_3. Thus the integral is zero as claimed. If $j = 1$ or 2 and $k = 3$, then

$$M_{jk} = \int_C \int v_k\phi (T^*)^{-1}\phi^* v_j \, dv \, d\mathbf{r}',$$

and this integral is zero for the same reason.

5. The material properties of a long cylindrical cell are symmetric across a plane parallel to the cell

Next, if D consists of long slender cells whose physi-

cal properties are symmetric across the r_1, r_3-plane, then $\mathcal{M}_0$ is diagonal. (To prove this, we need only show $M_{12} = M_{21} = 0$.) This follows from the fact that a rotation of angle π about the r_1 axis must leave $\mathcal{M}_0$ invariant. Thus $\mathcal{M}_0 = R^{-1} \cdot \mathcal{M}_0 \cdot R$, where

$$R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

and it follows immediately that $M_{12} = M_{21} = 0$ as claimed.

6. The material properties of a long cylindrical cell are symmetric across a plane parallel to the cell and invariant with respect to a rotation of the cell by angle $\theta \neq \pi$

If, in addition to the symmetry described in the above paragraph, a cell is also symmetric upon rotation about the r_3-axis by an angle $\theta \neq \pi$, then $M_{11} = M_{22}$. This can occur in hexagonal cells (for which $\theta = \pi/3$) or square cells (for which $\theta = \pi/2$). Once again, $\mathcal{M}_0$ must be invariant with respect to the rotation matrix

$$R = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and the result $M_{11} = M_{22}$ immediately follows.

7. D consists of adjacent thin slabs

As a final special case, we consider a domain consisting of thin one-dimensional slabs. The above considerations apply here, and the matrix $\mathcal{M}_0$ is diagonal if one of the principal directions (say e_3) is taken to be normal to the slab boundaries. Also, $M_{11} = M_{22}$.

In general geometries, we see from (3.14) that the determination of $\mathcal{M}_0$ depends on the determination of $\eta = T^{-1} v\phi$, where η is periodic in r'. The equation for η is $T\eta = v\phi$, and the general solution of this equation may be written

$$\eta = r'\phi + \xi,$$

where $T\xi = 0$. The periodicity conditions on η lead to appropriate conditions for ξ, but we shall not consider these here. We wish only to point out that the problem for ξ may in some instances be simpler to solve than the problem for η.

5. INITIAL CONDITION FOR THE DIFFUSION EQUATION

To determine the initial conditions for A_0 and Q_i, we must undertake an initial layer analysis. In particular, we assume that several mean free paths away from the boundary of D the angular density ψ may be written as

$$\psi = \psi^i + \psi^0. \tag{5.1}$$

We have already determined ψ^i in Sec. 3, and we take ψ^0 to be an initial layer which decays in time after the initial time $\tau = 0$.

To determine ψ^0 we introduce the following expansion:

$$\psi^0 \sim \sum_{n=0}^{\infty} \epsilon^n \hat{\psi}_n(r, r', v, \tau). \tag{5.2}$$

We require ψ^0 to satisfy (2.10), (2.11) with $S = 0$, since S is accounted for by ψ^i. Then, introducing the series (5.2) into (2.10), expanding σ_{Sn}, σ_{Tn}, and θ_i [as defined in (2.4), (2.5), and (2.3)] in powers of $\epsilon^2 \tau$, and equating the coefficients of different powers of ϵ, we obtain as the equation for $\hat{\psi}_0$

$$\frac{\partial \hat{\psi}_0}{\partial \tau} + T_0 \hat{\psi}_0 = 0. \tag{5.3}$$

Here $T_0 = v \cdot \nabla - L_0$, where L_0 is defined by

$$L_0 f(r, r', v) = \int v' \sigma_{S0}(r, r', v' \rightarrow v, 0) f(r, r', v') \, dv'$$
$$- v \sigma_{T0}(r, r', v, 0) f(r, r', v).$$

Upon multiplying (5.3) by $\phi_0^* = \phi^*(r, r', v, 0)$ and integrating, we obtain

$$0 = \frac{\partial}{\partial \tau} \int\int_C \phi_0^* \hat{\psi}_0 \, dv \, dr'.$$

Thus in order that $\hat{\psi}_0 \rightarrow 0$ as $\tau \rightarrow \infty$, it is necessary that

$$0 = \int_C \int \phi^*(r, r', v, 0) \hat{\psi}_0(r, r', v, 0) \, dv \, dr'. \tag{5.4}$$

(This condition is also sufficient if the spectrum of T, other than the simple eigenvalue $\lambda = 0$, lies in the half space $\mathrm{Re}\,\lambda > \gamma > 0$. We assume that this is the case.)

Setting $\tau = 0$ and $\epsilon = 0$ in (5.1) and using $\psi^i = A_0 \phi + 0(\epsilon)$, we obtain, upon eliminating $\hat{\psi}_0$ from (5.4),

$$A_0(r, 0) = \int_C \int \phi^*(r, r', v, 0) \psi(r, r', v, 0) \, dv \, dr'. \tag{5.5}$$

This is the initial condition for A_0. The initial conditions for Q_i are obtained from the prescribed initial conditions for C_i and equation (3.20). The initial condition for $\hat{\psi}_0$ is

$$\hat{\psi}_0(r, r', v, 0) = \psi(r, r', v, 0) - A_0(r, 0)\phi(r, r', v, 0). \tag{5.6}$$

The initial layer $\hat{\psi}_0$ is now the solution of (5.3), (5.6) and may be written in terms of τ as

$$\hat{\psi}_0(r, r', v, \tau) = \frac{1}{\pi i} \int_{-\gamma - i\infty}^{-\gamma + i\infty} \exp(\lambda\tau)(\lambda I + T)^{-1}(I - P)$$
$$\times \psi(r, r', v, 0) \, d\lambda,$$

where P is the projection

$$Pf = \phi \int_C \int \phi^* f \, dv \, dr'.$$

Thus $\hat{\psi}_0$ decays as $\tau \rightarrow \infty$, the decay rate being of order one with respect to τ. But τ is a scaled time variable, given in terms of unscaled time t' by $\tau = \epsilon^{-1} t'$. Therefore, the decay rate of $\hat{\psi}_0$ is fast, of order ϵ^{-1}, with respect to unscaled time.

6. BOUNDARY CONDITIONS FOR THE DIFFUSION EQUATION

To obtain the boundary condition for the diffusion equation, it is necessary to perform a boundary layer analysis near ∂D by introducing stretched spatial coordinates, as was done in Ref. 3. To do this one chooses a point $r_0 \in \partial D$ and defines $r'' = \epsilon^{-1}(r - r_0)$. Then, near r_0, ψ^b is treated as a function of r'' rather than r or r'. ψ^b will asymptotically satisfy Eqs. (2.10), (2.11) with $S = 0$ (since S is accounted for by ψ^i) and will be required to decay with distance from ∂D. Furthermore,

if the boundary conditions are slowly varying in time, then ψ^b will depend parametrically upon t rather than τ. Then for large values of τ (i.e., after several collisions) the neutron density near ∂D will be

$$\psi \sim \psi^i + \psi^b. \tag{6.1}$$

If the incoming neutron density is prescribed, then the boundary condition for A_0 is determined as in Ref. 3 by requiring $\psi \sim \psi^i + \psi^b$ to satisfy this condition and ψ^b to decay with distance from ∂D. If the incoming neutron density is not prescribed but instead D is surrounded by an exterior reflecting region D_{ex}, then we must construct the asymptotic solution ψ_{ex} in D_{ex} and equate ψ_{ex} with ψ on ∂D. This yields equations for A_0 on ∂D.

If D_{ex} is either a perfect absorber or a subcritical domain in which the neutron density is appreciable only within a few mean free paths of ∂D, then ψ_{ex} is either zero or a boundary layer, respectively. For these cases the continuity condition across ∂D is met to leading order by setting $\psi_{ex} = \psi^b = 0$ on ∂D. Furthermore, since ∂D is the boundary of many small cells, it has the form

$$\partial D = \{\xi + \epsilon r(\epsilon^{-1}\xi) \mid \xi \in \Gamma\}. \tag{6.2}$$

Here Γ is a smooth surface close to ∂D. With $O(\epsilon)$ error, we may take the boundary condition on A_0 to be

$$0 = A_0(\mathbf{r}, t), \quad \mathbf{r} \in \Gamma. \tag{6.3}$$

If D_{ex} is a near critical domain, then to leading order ψ_{ex} is determined by the solution of a diffusion equation, and it is inappropriate to set $\psi_{ex} = 0$ on ∂D. For this case, as with the case in which the incoming density is prescribed, the full boundary layer analysis must be undertaken.

In Ref. 3, the boundary layer problem near ∂D reduced to a constant coefficient half space problem for two reasons: The material properties of D had slow spatial variation and the boundary ∂D was smooth. In this paper neither of these assumptions holds, and the boundary layer problem appropriate to the present case is a full three-dimensional problem with variable coefficients. Since this problem cannot in general be treated analytically, we shall not consider it further here.

7. SOLUTION OF THE DIFFUSION EQUATIONS, AND CRITICALITY

Let us take the coefficients in Eqs. (3.16), (3.22) to be independent of t. Also, let D be surrounded by a perfect absorber or subcritical domain, so that the appropriate boundary condition for A_0 is given by (6.3).

Then with initial conditions prescribed for A_0 and Q_i, Eqs. (3.16), (3.22) can be solved using the Laplace transform method. Letting $\hat{A}_0$, $\hat{Q}_i$, and $\hat{S}$ be the Laplace transforms of A_0, Q_i, and S and denoting the diffusion operator in (3.16) by $K = \nabla \cdot /\hbar_0 \cdot \nabla + \mathbf{M}_1 \cdot \nabla + M_2$, we obtain

$$s\hat{A}_0(\mathbf{r}, s) - A_0(\mathbf{r}, 0) = K\hat{A}_0(\mathbf{r}, s) + \sum_i \lambda_i \hat{Q}_i(\mathbf{r}, s) + \hat{S}_0(\mathbf{r}, s)$$

and

$$s\hat{Q}_i(\mathbf{r}, s) - Q_i(\mathbf{r}, 0) = \beta_i \hat{A}_0(\mathbf{r}, s) - \lambda_i \hat{Q}_i(\mathbf{r}, s).$$

Eliminating $\hat{Q}_i$ yields

$$\hat{A}_0(\mathbf{r}, s) = [g(s)I - K]^{-1}\left(A_0(\mathbf{r}, 0) + \hat{S}_0(\mathbf{r}, s)\right.$$
$$\left. + \sum_i \frac{\lambda_i}{s + \lambda_i}Q_i(\mathbf{r}, 0)\right),$$

where

$$g(s) = s - \sum_i \frac{\lambda_i \beta_i}{s + \lambda_i},$$

and where $\hat{A}_0(\mathbf{r}, s) = 0$ for $\mathbf{r} \in \Gamma$.

Using the inverse transform, we obtain

$$A_0(\mathbf{r}, t) = \frac{1}{\pi i}\int_{\alpha - i\infty}^{\alpha + i\infty} \exp(st)\hat{A}_0(\mathbf{r}, s)\, ds,$$

where α lies to the right of the singularities of the integrand. These singularities are those s for which $g(s)$ lies in the spectrum of K.

Let α_0 denote the singularity of $\hat{A}_0$ with greatest real part. Then if $\text{Re}\,\alpha_0 < 0$, D is subcritical; if $\text{Re}\,\alpha_0 = 0$, then D is critical; and if $\text{Re}\,\alpha_0 > 0$, then D is supercritical. This is because solutions of (3.16), (3.22) with $S_0 = 0$ will decay, approach a constant value, or grow as $t \to \infty$ in these respective cases.

The above considerations apply only if the coefficients $/\hbar_0$, $\mathbf{M}_1$, $\mathbf{M}_2$, and β_i of (3.16), (3.22) are independent of t. If these coefficients cannot be regarded as independent of time, then the above considerations do not apply. Such instances occur in reactor dynamics, where the physical properties of D are taken to be time and temperature dependent. For this case the above coefficients are indirectly functions of A_0, and so the problem is no longer even linear. If the dependence of the coefficients on A_0 is made explicit, then (3.16) and (3.22) can be used to study stability near an equilibrium point. However, we shall not consider this topic here.

ACKNOWLEDGMENTS

[1]B.J. Matkowsky, "On the Diffusion Approximation to Transport Theory," 1970 (unpublished).
[2]G.J. Habetler and B.J. Matkowsky, "On Boundary Value Problems in Diffusion Theory as Limits of Boundary Value Problems in Transport Theory," 1971 (unpublished).
[3]E.W. Larsen and J.B. Keller, J. Math. Phys. 15, 75 (1974).
[4]E.W. Larsen, J. Math. Phys. 15, 229 (1974).
[5]G.C. Papanicolaou, "Asymptotic Analysis of Transport Processes," Bull. AMS, March, 1975 (to be published).
[6]D.L. Hetrick, *Dynamics of Nuclear Reactors* (University of Chicago, 1971).
[7]G.I. Bell and S. Glasstone, *Nuclear Reactor Theory* (Van Nostrand, New York, 1970), p. 167.
[8]J.R. Lamarsh, *Introduction to Nuclear Reactor Theory* (Addison-Wesley, Reading, Mass., 1966).
[9]E.W. Larsen, "Neutron Transport and Diffusion in Inhomogeneous Media. II" (unpublished).
[10]F. Riesz and B. Sz-Nagy, *Functional Analysis* (Ungar, New York, 1970).

Bäcklund transformation solutions of the Toda lattice equation*

Hsing-Hen Chen and Chuan-Sheng Liu

Institute for Advanced Study, Princeton, New Jersey 08540
(Received 11 February 1975)

Bäcklund transformation, superposition formula, and multisoliton solutions are constructed for the Toda lattice difference–differential equation with some discussion of its generalizations.

I. INTRODUCTION

To understand many nonlinear, nonergodic phenomena such as Fermi—Pasta—Ulam recurrence, Toda[1] proposed a one-dimensional nonlinear lattice model with neighboring particles interacting through an exponential potential function, $V(r) = a[(1/b)\exp(-br) + r]$, $ab > 0$, which admits exact analytic solution. It is therefore an exactly soluble many-body system. By varying the parameters a and b, this model can represent a wide spectrum of interactions, from the harmonic limit (ab finite, $a \to \infty$, $b \to 0$) to the hard sphere limit (ab finite, $a \to 0$, $b \to \infty$). Therefore, many useful results can be obtained by solving this model. The Hamiltonian function for the dynamics of Toda lattice with unit mass and normalized parameters is

$$H = \sum_n P_n^2 - \{\exp[-(Q_n - Q_{n-1})] - 1\}, \tag{1}$$

where P_n is the canonical momentum of the nth lattice particle and Q_n is the displacement from its equilibrium.

The corresponding equation of motion is

$$\dot{Q}_n = P_n, \quad \dot{P}_n = -\{\exp[-(Q_{n+1} - Q_n)] - \exp[-(Q_n - Q_{n-1})]\}. \tag{2}$$

It is a remarkable discovery by Toda[1] that this nonlinear difference evolution equation allows analytic stable solitary wave (soliton) solutions. He obtained the single-soliton solution and the two-soliton solution. Hirota[2] later generalized it intuitively to N solitons. In the meantime, Gardner[3] *et al.* discovered the multisolitary wave solutions to the Korteweg—de Vries equation and developed a general analytic scheme, the so-called inverse scattering problem method, to solve the KdV equation exactly. On the other hand, there exists another analytic approach, the Bäcklund transformation,[4] to solve a set of nonlinear equations. The usefulness of this latter method was recognized by Lamb,[5] Wahlquist and Estabrook,[6] and was recently demonstrated to be equivalent to the inverse scattering problem[7] for the nonlinear differential equations encompassed in the schemes of Ablowitz *et al.*[8] An inverse scattering problem for the Toda lattice equation was first found and solved by Flaschka.[9] Using the discrete version of inverse problems developed by Kac and Case,[10] he obtained general multisoliton solutions identical to the Hirota's intuitive result. In this paper we will demonstrate the Bäcklund transformation approach of solving the Toda equation, thus providing the first case where a nonlinear evolution difference equation can be solved by a discrete form of Bäcklund transformation. This transformation is derived from an inverse scattering problem equivalent to the Flaschka's. Superposition formula is

then derived from this Bäcklund transformation, and multisoliton solutions are constructed. Because of its simplicity, the Bäcklund transformation method is very appealing in constructing the multisoliton solutions. It also has the advantage to enable one to construct "periodic wave + soliton" solutions, which can not be obtained from the inverse problem.

In the last section, we shall also demonstrate ways to find generalized Toda equations that admit the same inverse problem.

II. SCATTERING PROBLEM AND BÄCKLUND TRANSFORMATION

Let us consider an inverse scattering problem represented by the two coupled wave equations:

$$2a_n v_{1,n} = (4a_n^2 + 2\lambda b_{n+1} - \lambda^2) v_{2,n+1} + \lambda v_{1,n+1},$$
$$2a_n v_{2,n} = (\lambda - 2b_{n+1}) v_{2,n+1} - v_{1,n+1}, \tag{3}$$

$$\dot{v}_{1,n} = A_n v_{1,n} + B_n v_{2,n},$$
$$\dot{v}_{2,n} = C_n v_{1,n} - A_n v_{2,n}, \tag{4}$$

where

$$A_n = -b_n, \quad B_n = \lambda^2 - 2\lambda b_n - 4a_n^2, \quad C_n = 1, \tag{5}$$

$v_{1,n}$ and $v_{2,n}$ are two wavefunctions; a_n and b_n are time dependent potentials with eigenvalue λ independent of time. These equations are integrable if and only if

$$\dot{a}_n = a_n(b_{n+1} - b_n), \quad \dot{b}_n = 2(a_n^2 - a_{n-1}^2). \tag{6}$$

This set is equivalent to the Toda equation (2) by the following identification:

$$2a_n = \exp[-(Q_n - Q_{n-1})/2] \quad \text{and} \quad b_n = -P_{n-1}/2. \tag{7}$$

On the other hand, of letting

$$4a_n^2 = 1 + \dot{w}_n, \quad 2b_n = w_n - w_{n-1}, \tag{8}$$

the Toda equation can also be cast to a convenient form:

$$\ddot{w}_n/(1 + \dot{w}_n) = w_{n+1} + w_{n-1} - 2w_n. \tag{9}$$

To derive the Bäcklund transformation, we first define $u \equiv v_1/v_2$; Eqs. (2) and (3) then reduce to two Riccati equations,

$$(\lambda + u_n)(\lambda - 2b_{n+1} - u_{n+1}) = 4a_n^2,$$
$$(\lambda + u_n)(\lambda - 2b_n - u_n) = 4a_n^2 + \dot{u}_n. \tag{10}$$

These equations represent a mutual transformation, generalized Miura transformation (GMT) between the following two nonlinear equations. One of them is the Toda equation (9) satisfied by w_n, another one satisfied by u_n is the following:

$$u_n = \dot{U}_n, \quad \lambda \ddot{U}_n/(\lambda + \dot{U}_n) = \exp[-(U_n - U_{n-1})]$$
$$- \exp[-(U_{n+1} - U_n)]. \tag{11}$$

It is also obvious that Eq. (9) is invariant under the index transformation, $m \to 2n - m$. Therefore, a similar inverse scattering problem exists with the corresponding index changes, that is,

$$2\tilde{a}_n \tilde{v}_{1,n} = (4\tilde{a}_n^2 + 2\lambda \tilde{b}_{n-1} - \lambda^2)\tilde{v}_{2,n-1} + \lambda \tilde{v}_{1,n-1},$$
$$2\tilde{a}_n \tilde{v}_{2,n} = (\lambda - 2\tilde{b}_{n-1})\tilde{v}_{2,n-1} - \tilde{v}_{1,n-1}, \tag{12}$$

$$\dot{\tilde{v}}_{1,n} = \tilde{A}_n \tilde{v}_{1,n} + \tilde{B}_n \tilde{v}_{2,n},$$
$$\dot{\tilde{v}}_{2,n} = \tilde{C}_n \tilde{v}_{1,n} - \tilde{A}_n \tilde{v}_{2,n}, \tag{13}$$

where

$$\tilde{A}_n = -\tilde{b}_n, \quad \tilde{B}_n = \lambda^2 - 2\lambda \tilde{b}_n - 4\tilde{a}_n^2, \quad \tilde{C}_n = 1.$$

The integrability conditions for these equations are

$$\dot{\tilde{a}}_n = \tilde{a}_n(\tilde{b}_{n-1} - \tilde{b}_n) \quad \text{and} \quad \dot{\tilde{b}}_n = 2(\tilde{a}_n^2 - \tilde{a}_{n+1}^2). \tag{14}$$

The identifications $4\tilde{a}_n^2 = 1 + \dot{w}_n$ and $2\tilde{b}_n = w_n - w_{n+1}$ reduce it again to the Toda equation (9). Define similarly $\tilde{u}_n \equiv \tilde{v}_{1n}/\tilde{v}_{2n}$; we get two Riccati equations

$$(\lambda + \tilde{u}_n)(\lambda - 2\tilde{b}_{n-1} - \tilde{u}_{n-1}) = 4\tilde{a}_n^2,$$
$$(\lambda + \tilde{u}_n)(\lambda - 2\tilde{b}_n - \tilde{u}_n) = 4\tilde{a}_n^2 + \dot{\tilde{u}}_n. \tag{15}$$

After elimination of $\tilde{a}_n$ and $\tilde{b}_n$, we found $\tilde{u}_n$ satisfying

$$\tilde{u}_n = \dot{\tilde{U}}_n, \quad \lambda \ddot{\tilde{U}}_n/(\lambda + \dot{\tilde{U}}_n)$$
$$= \exp[-(\tilde{U}_n - \tilde{U}_{n+1})] - \exp[-(\tilde{U}_{n-1} - \tilde{U}_n)]. \tag{16}$$

Now, suppose (λ, U_n) constitutes a solution to Eq. (11); then Eq. (10) transforms it to w_n, a solution of Toda equation (9). However, $(-\lambda, -U_n)$ would then be a solution to Eq. (16), and correspondingly, Eq. (15) transforms it to w'_n, another solution of Toda equation (9). Therefore, a reversible route is found for the transformations between two solutions w_n and w'_n to the same Toda equation (9). We can illustrate this transformation route as the following:

$$
\begin{array}{ccc}
(\lambda, U_n) & \longleftrightarrow & (-\lambda, -U_n) = (\lambda', \tilde{U}_n') \\
\updownarrow \text{GMT} & & \updownarrow \text{GMT} \\
w_n \longleftarrow & \text{BT} & \longrightarrow w'_n
\end{array}
\tag{17}
$$

Rewrite Eqs. (10) and (15) into

$$(\lambda + u_n)(\lambda - 2b_{n+1} - u_{n+1}) = 4a_{n+1}^2, \tag{18a}$$
$$(\lambda + u_{n+1})(\lambda - 2b_{n+1} - u_{n+1}) = 4a_{n+1}^2 + \dot{u}_{n+1} \tag{18b}$$

and

$$(\lambda + u_{n+1})(\lambda + 2b'_n - u_n) = 4\tilde{a}_{n+1}'^2, \tag{19a}$$
$$(\lambda + u_n)(\lambda + 2b'_n - u_n) = 4\tilde{a}_n'^2 - \dot{u}_n \tag{19b}$$

Using Eq. (8) to change a_n's and b_n's into w_n's and doing some algebra, (19b) + (19b) − (18a) − (18a), we get

$$\dot{u}_{n+1} - \dot{u}_n = (\dot{w}'_{n+1} - \dot{w}_{n+1}) - (\dot{w}'_n - \dot{w}_n) \tag{20}$$
$$+ (u_{n+1} - u_n)[u_n - u_{n+1} + (w'_{n+1} - w_{n+1}) - (w'_n - w_n)].$$

Therefore, $u_n = w'_n - w_n$. Substituting it into either Eq. (18) or (19), we get finally the Bäcklund transformation:

$$(\lambda + w'_n - w_n)(\lambda + w_n - w'_{n+1}) = 1 + \dot{w}_n,$$
$$(\lambda + w'_n - w_n)(\lambda + w_{n-1} - w'_n) = 1 + \dot{w}'_n. \tag{21}$$

This transformation enables us to construct nontrivial solutions from a known one. For example, let $w_n = 0$, a trivial solution; we get from (21) then

$$w'_n = \pm \sinh\varphi \tanh(n\varphi \pm \sinh\varphi t + \alpha), \quad \lambda = \pm \cosh\varphi,$$
or $$\tag{22}$$
$$\dot{w}'_n = \sinh^2\varphi \, \text{sech}^2(n\varphi \pm \sinh\varphi t + \alpha).$$

The latter is a single soliton with amplitude $\sinh^2\varphi$ and travelling with speed $(\sinh\varphi)/\varphi$ in both directions. Substituting this single soliton solution back into the Bäcklund transformation, we can find the two-soliton solution and so on. To do so, it seems that we have to solve a first order nonlinear evolution difference—differential equation, which is not trivial. However, a remarkably nice consequence of the Bäcklund transformation is the possibility of deriving a nonlinear superposition formula for the solutions. This formula enables one to avoid integration quadratures and makes possible the construction of multisoliton solutions by algebraic means only.

The superposition formula is derived from the permutability assumption of the Bäcklund transformation,[4] that is, a solution w_n^3 generated from, say w_n^0, by applying twice the Bäcklund transformation, first with parameter λ_1, then λ_2, is identical to the one obtained with the reversed order. Then, a result consistent to the permutability assumption is obtained as the superposition formula. For Toda lattice, we found the following superposition formula:

$$(\lambda_1 - \lambda_2)[(w_n^3 + w_n^0) - (w_n^1 + w_n^2) + (\lambda_1 + \lambda_2)]$$
$$= (w_n^3 - w_n^0)(w_n^1 - w_n^2). \tag{23}$$

We demonstrate in the following the construction of two-soliton solutions. Starting with vacuum $w_n^0 = 0$, we have, from Eq. (22), single-soliton solutions:

$$w_n^1 = \begin{matrix} \pm \sinh\varphi_1 \tanh(n\varphi_1 \pm \sinh\varphi_1 t + \alpha), \\ \mp \sinh\varphi_1 \coth(n\varphi_1 \pm \sinh\varphi_1 t + \alpha); \end{matrix} \quad \lambda_1 = \pm \cosh\varphi_1$$

and

$$w_n^2 = \begin{matrix} \pm \sinh\varphi_2 \tanh(n\varphi_2 \pm \sinh\varphi_2 t + \alpha_2), \\ \mp \sinh\varphi_2 \coth(n\varphi_2 \pm \sinh\varphi_2 t + \alpha_2); \end{matrix} \quad \lambda_2 = \pm \cosh\varphi_2.$$

From these solutions, both regular and singular two soliton solutions can be obtained. To get regular solutions of physical interest, we follow the rule of Wahlquist and Estabrook,[6] choose regular w_n^1 and singular w_n^2. Two kinds of regular two-soliton solutions exist, with (i) $\lambda_1 \lambda_2 > 0$ and (ii) $\lambda_1 \lambda_2 < 0$, corresponding to the collision of two solitons travelling in the same direction and in the opposite directions respectively. We consider first the case (i). It is convenient to look into $\dot{w}_n^3$ instead of w_n^3 itself. From Eq. (23), we have

$$\dot{w}_n^3 = \frac{t_d}{dt}\left(\frac{\pm(\lambda_1 - \lambda_2)[\sinh\varphi_1 \tanh\xi_1 \mp (\lambda_1 + \lambda_2) - \sinh\varphi_2 \tanh\xi_2]}{(\lambda_1 - \lambda_2) \mp [\sinh\varphi_1 \tanh\xi_1 + \sinh\varphi_2 \tanh\xi_2]}\right), \tag{24}$$

where

$$\xi_1 = n\varphi_1 \pm \sinh\varphi_1 t + \alpha_1 \quad \text{and} \quad \xi_2 = n\varphi_2 \pm \sinh\varphi_2 + \alpha_2.$$

It can be seen easily that as $\xi_1 \to \pm\infty$, the two-soliton solution approaches

$$\dot{w}_n^3 \to \sinh^2\varphi_2 \, \mathrm{sech}^2(\xi_2 \pm \phi_2^\pm), \tag{25}$$

the second soliton solution with a phase shift,

$$2\phi_2^+ = \mp \varphi_2 \pm \ln\left(\frac{\sinh[(\varphi_1 - \varphi_2)/2]}{\sinh[(\varphi_1 + \varphi_2)/2]}\right),$$

$$2\phi_2^- = \pm \varphi_2 \pm \ln\left(\frac{\sinh[(\varphi_1 - \varphi_2)/2]}{\sinh[(\varphi_1 + \varphi_2)/2]}\right),$$

or

$$\Delta\phi_2 \equiv \phi_2^+ + \phi_2^- = \pm \ln\frac{\sinh[(\varphi_1 - \varphi_2)/2]}{\sinh[(\varphi_1 + \varphi_2)/2]} \tag{26}$$

is the phase change experienced by the second soliton induced by the collision. Similarly, the phase change of the first soliton is

$$\Delta\phi_1 = \pm \ln\left(\frac{\sinh[(\varphi_1 + \varphi_2)/2]}{\sinh[(\varphi_1 - \varphi_2)/2]}\right) = -\Delta\phi_2. \tag{27}$$

Therefore, it is clear that Toda solitons reemerge after collision with the same identities but a phase shift, $\Delta\phi$.

The result for case (ii) $(\lambda_1\lambda_2 < 0)$ is similar; we have, instead of (26) and (27), the following phase shift:

$$\Delta\phi_2 = -\Delta\phi_1 = \pm \ln\left(\frac{\cosh[(\varphi_1 - \varphi_2)/2]}{\cosh[(\varphi_1 + \varphi_2)/2]}\right). \tag{28}$$

These results are identical to those obtained by Toda before. However, Eq. (23) enables us to go beyond the two-soliton solutions. Multisoliton solutions are obtained only by iterating the procedure outlined above.

III. CONCLUDING REMARKS

1. Equation (11) generated from (10) is a new member of nonlinear evolution difference–differential equation that admits soliton solutions. The relation $u_n = w_n' - w_n$ relates its solutions to the solutions of Toda equation. We have therefore, as a by-product to our solutions of Toda equation, found a branch of solitonlike solutions to this new equation.

2. The Bäcklund transformation, Eq. (21), is not always nontrivial. If we let $w_n' = w_{n-1/2}$, then the two transformation equations become identical $(\lambda + w_{n-1/2} - w_n)(\lambda + w_n - w_{n+1/2}) = 1 + \dot{w}_n$. It can be shown that solutions to this equation are not solitons.

3. The inverse problem (3) and (4) can be more general, that is, Eq. (3) and (4) does not uniquely determine the functional form of A, B, and C. The integrability conditions are actually

$$(\dot{a}_n/a_n + A_n - A_{n+1})\lambda - B_n - (4a_n^2 + 2\lambda b_{n+1} - \lambda^2)C_{n+1} = 0,$$

$$(A_{n+1} + A_n + \dot{a}_n/a_n)(4a_n^2 + 2\lambda b_{n+1} - \lambda^2) + B_n(\lambda - 2b_{n+1})$$
$$- 8a_n\dot{a}_n - 2\lambda\dot{b}_{n+1} - \lambda B_{n+1} = 0,$$

$$(A_{n+1} + A_n - \dot{a}_n/a_n) + C_n\lambda - (\lambda - 2b_{n+1})C_{n+1} = 0,$$

$$(\lambda - 2b_{n+1})(\dot{a}_n/a_n - A_n + A_{n+1}) + C_n(4a_n^2 + 2\lambda b_{n+1} - \lambda^2)$$
$$+ B_{n+1} + 2\dot{b}_{n+1} = 0. \tag{29}$$

Finite series expansions of A, B, C to different orders of λ leads to different equations of interest. The Toda lattice is but the simplest example. The next generalization would be

$$\dot{a}_n/a_n = b_{n+1}^2 - b_n^2 + a_{n+1}^2 - a_{n-1}^2,$$
$$\dot{b}_n = 2a_n^2(b_{n+1} + b_n) - 2a_{n-1}^2(b_n + b_{n-1}) \tag{30}$$

with the identification

$$A_n = -b_n\lambda + (\dot{a}_n/a_n - b_{n+1}^2 + a_n^2 - a_{n+1}^2),$$

$$B_n = \lambda^3 - b_n\lambda^2 + 2(\dot{a}_n/a_n - b_{n+1}^2 - a_n^2 - a_{n+1}^2)\lambda - 4a_n^2 b_{n+1},$$

$$C_n = \lambda + b_n. \tag{31}$$

Bäcklund transformations and solutions to these generalized Toda equations can be found along the same line as this paper presents. They are similar to the case of Toda equations.

[1]M. Toda, J. Phys. Soc. Japan 22, 431 (1967); 23, 501 (1967); M. Toda and M. Wadati, J. Phys. Soc. Japan 34, 18 (1973).
[2]R. Hirota, J. Phys. Soc. Japan 35, 286 (1973).
[3]C. Gardner, J. Greene, M. Kruskal, and R. Miura, Phys. Rev. Lett. 19, 1095 (1967); P. Lax, Comm. Phys. Appl. Math. 21, 916 (1968).
[4]L. Eisenhart, A Treatise on the Differential Geometry of Curves and Surfaces (Dover, New York, 1960).
[5]G. Lamb, Jr., Rev. Mod. Phys. 43, 99 (1971).
[6]H. Wahlquist and F. Estabrook, Phys. Rev. Lett. 31, 1386 (1973).
[7]H. Chen, Phys. Rev. Lett. 33, 925 (1974).
[8]M. Ablowitz, P. Kaup, A. Newell, and H. Segur, Phys. Rev. Lett. 31, 125 (1973).
[9]H. Flaschka, Prog. Theor. Phys. 51, 703 (1974).
[10]K. Case and M. Kac, J. Math. Phys. 14, 594 (1973).

Harmonic mappings of Riemannian manifolds and stationary vacuum space-times with whole cylinder symmetry

Ahmet Eriş and Yavuz Nutku

Department of Physics, Middle East Technical University, Ankara, Turkey
(Received 16 October 1974)

We consider stationary, cylindrically-symmetric gravitational fields in the framework of harmonic mappings of Riemannian manifolds. In this approach the emphasis is on a correspondence between the solution of the Einstein field equations and the geodesics in an appropriate Riemannian configuration space. Using Hamilton–Jacobi techniques, we obtain the geodesics and construct the resulting space-time geometries. We find that the light cone structure of the configuration space delineates the distinct exterior fields of Lewis and van Stockum which together form the most general solution with whole cylinder symmetry.

1. INTRODUCTION

In a previous paper[1] it was shown that the classical field theories of physics may be formulated in the framework of Eells and Sampson's theory of deformations.[2] There it was found that the field equations of these theories were the conditions for extremizing an invariant functional, "energy", associated with the mapping from the space–time manifold M into another manifold M'. In general M' had to be a principal fiber bundle over M with a structure group which is identified with the gauge group of the physical theory. However, for a restricted class of gravitational fields, namely stationary axisymmetric exterior solutions, it suffices for M' to be a specific Riemannian manifold, and the Einstein field equations can then be formulated as extremal submanifolds of M' according to Eells and Sampson's theory of the harmonic mappings of Riemannian manifolds. The geometry of M' turns out to be that of a hyperboloid as it was first discovered by Matzner and Misner[3] and later elaborated on by Matzner[4] and Neugebauer and Kramer.[5] In this paper we shall specialize further to whole cylinder symmetry in order to reduce the discussion of these exterior fields to a study of the geodesics of M'. The specialization to static spherically symmetric gravitational fields where the Einstein field equations can be formulated as the geodesics of a flat configuration space is discussed in the Appendix. The advantage of this formulation lies in the existence of a very powerful technique for obtaining the geodesics, namely the Hamilton–Jacobi theory. We shall show that on M' there exists a complete set of first integrals of motion and the Hamilton–Jacobi equation is solvable by a separation of variables in a coordinate system adopted to the Killing directions. This enables us to obtain Hamilton's principal function and thereby construct the most general solution of the Einstein field equations compatible with the assumed symmetries. We find that, depending on the range of values taken on by the constant of motion in M', the corresponding space–time solutions fall into three distinct classes. The crucial constant is the "energy" first integral and when it is positive, negative, or zero, that is, for "time-like", "space"-like or "null" geodesics on the hyperboloid we have the vacuum solutions of Lewis[6] and van Stockum.[7] Thus the transition from any one of these solutions into another requires the crossing of the null cone in M'. Finally we examine the relation-

ship between the constants of the motion on M' and the parameters entering into the metric for space–time.

2. HARMONIC MAPPINGS OF RIEMANNIAN MANIFOLDS

In the theory of harmonic mappings of Riemannian manifolds, following Eells and Sampson[2] we start by considering two given Riemannian[8] manifolds M and M' with the metrics

$$ds^2 = g_{ab}\, dx^a\, dx^b \qquad (a, b = 1, 2, \ldots, n), \tag{1}$$

$$ds'^2 = g'_{AB}\, dy^A\, dy^B \qquad (A, B = 1, 2, \ldots, m) \tag{2}$$

respectively and a map $f : M \to M'$. We are interested in constructing invariant functionals associated with this mapping. For this purpose we define for each point P in M an inner product $\langle\ ,\ \rangle_P$ on the space of 2-covariant tensors of the tangent space to M at P. If α and β are two such tensors, then

$$\langle \alpha, \beta \rangle_P \equiv \alpha_{ab}\, \beta_{cd}\, g^{ac}\, g^{bd}, \tag{3}$$

where g is the Riemannian metric tensor on M. A particular example of an invariant functional is the "energy"

$$E(f) = \int_M \langle g, f^* g' \rangle_P\, {}^* 1, \tag{4}$$

where we consider the integral over the volume of M of the inner product of g with the induced metric $f^* g'$. The invariant volume element on M is denoted by ${}^* 1$. In terms of the local coordinates the energy functional is given by

$$E(f) = \int_M g'_{AB}\, \frac{\partial f^A}{\partial x^a}\, \frac{\partial f^B}{\partial x^b}\, g^{ab}\, \sqrt{|g|}\, d^n x \tag{5}$$

and those maps for which the first variation of $E(f)$ vanishes are called harmonic maps. The necessary and sufficient conditions for a map to be harmonic are given by the Euler equations

$$\Delta f^A + \Gamma'^A_{BC}\, \frac{\partial f^B}{\partial x^a}\, \frac{\partial f^C}{\partial x^b}\, g^{ab} = 0, \tag{6}$$

where Γ'^A_{BC} are the coefficients of the Riemannian connection on M' and Δ denotes the covariant Laplace operator on M.

The physical interest of this theory lies in the fact that many familiar equations of physics are derivable from a variational principle where we identify a suitable specialization of the energy functional as the action. In

particular if we consider the problem of stationary, axisymmetric gravitational fields which are described by the line element[9]

$$ds^2 = -\rho e^{2\psi}(dt - \omega d\phi)^2 + \rho^{-1} e^{2(\gamma-\psi)}[(d\rho)^2 + (dz)^2]$$
$$+ \rho e^{-2\psi}(d\phi)^2 \qquad (7)$$

where ψ, γ, and ω are functions of ρ, z only, we find that the Einstein field equations are derivable from the variational principle $\delta I = 0$,

$$I = 2\pi \int [(\nabla\psi)^2 - \tfrac{1}{4}e^{4\psi}(\nabla\omega)^2]\rho d\rho\, dz, \qquad (8)$$

where ∇ is the flat space gradient operator in the coordinates ρ, z. We note that once the solution for ψ and ω is obtained the remaining metric coefficient γ can be obtained by quadratures. The integrand in Eq. (8) is up to a divergence Einstein's Lagrangian density $\sqrt{-g}\,R$ for the metric (7). Therefore, if we consider two Riemannian manifolds with the metrics

$$ds^2 = (d\rho)^2 + (dz)^2 + \rho^2(d\phi)^2, \qquad (9)$$

$$ds'^2 = (d\psi)^2 - \tfrac{1}{4}e^{4\psi}(d\omega)^2 \qquad (10)$$

and a map $f:M \to M'$, the requirement that it be a harmonic map reduces to the condition that the Einstein field equations for (7) be satisfied since the energy functional formed by using (9) and (10) is the same as the action in Eq. (8). Hence M' with the metric (10) is the configuration space for this class of gravitational fields. Note that M is effectively two-dimensional since the maps we shall consider will be independent of ϕ.

In order to construct the harmonic maps and obtain solutions of the Einstein field equations we need to solve the Euler equations (6). The metric in Eq. (10) describes a space of constant curvature. Since M' is not flat, Eqs. (6) are a set of coupled nonlinear partial differential equations. However, we note that if we specialize to whole cylinder symmetry, thus suppressing the z dependence of the metric coefficients in Eq. (7), Eqs. (6) reduce to ordinary differential equations which describe the geodesics of M'. In terms of harmonic mappings of Riemannian manifolds such a specialization results in M becoming one-dimensional while the geometry of M' is unchanged. The problem of obtaining new solutions of the Einstein field equations is now vastly simplified since we require only a knowledge of the geodesics of M'. For this purpose we shall turn to Hamilton—Jacobi theory which is given in the next section.

3. THE HAMILTON-JACOBI THEORY OF GEODESICS AND THE GENERAL EXTERIOR SOLUTION

We shall now investigate the geodesics of M' with the metric (10). For this purpose we write down the Hamilton—Jacobi equation

$$\frac{\partial S}{\partial \rho} + H\left(y^A, \frac{\partial S}{\partial y^A}\right) = 0 \qquad (11)$$

where $y^A = \{\psi, \omega\}$ are the generalized coordinates and we are parametrizing the geodesics by ρ, which will have the same meaning as the space—time coordinate in Eq. (7). The Hamiltonian in Eq. (11) is simply the kinetic energy

$$H = \frac{1}{\sqrt{g}}\, g'^{AB}\, \frac{\partial S}{\partial y^A}\, \frac{\partial S}{\partial y^B} \qquad (12)$$

and the Hamilton—Jacobi equation is given by

$$\rho\frac{\partial S}{\partial \rho} + \left(\frac{\partial S}{\partial \psi}\right)^2 - 4e^{-4\psi}\left(\frac{\partial S}{\partial \omega}\right)^2 = 0. \qquad (13)$$

We require the complete solution of this equation, which must contain two nontrivial arbitrary constants. Equation (13) is readily seen to be separable, and we find the general solution

$$S = -\beta \ln\rho + \alpha\omega \pm \tfrac{1}{2}(4\alpha^2 e^{-4\psi} + \beta)^{1/2}$$
$$\pm \begin{cases} \tfrac{1}{2}\beta^{1/2}\ln\{e^{2\psi}[(4\alpha^2 e^{-4\psi}+\beta)^{1/2} - \beta^{1/2}]\}, & \beta > 0, \\ -\tfrac{1}{2}(-\beta)^{1/2}\tan^{-1}[(4\alpha^2 e^{-4\psi}+\beta)^{1/2}(-\beta)^{-1/2}]\}, & \beta < 0, \end{cases}$$
$$\qquad (14)$$

of the Hamilton—Jacobi equation. The various constants appearing in S are expressed in terms of the momenta as

$$\alpha = P_\omega, \quad \beta = P_\psi^2 - 4e^{-4\psi}P_\omega^2, \qquad (15)$$

and the affine parameter for the geodesics is given by $\ln\rho$. This completes the characterization of the geodesics in M'.

In the Hamilton—Jacobi theory, once the principal function S is constructed, the solution for the generalized coordinates is obtained from S by differentiations. Thus we shall find for the geodesics in M' the functional dependence of the generalized coordinates on the parameter ρ. By virtue of the present formulation, these solutions when substituted into Eq. (7) give us a space—time metric which satisfies the Einstein field equations. But we first note that, as the expression (14) for S suggests, the three cases $\beta > 0$, $\beta < 0$, and $\beta = 0$ must be clearly delineated since they will give rise to different exterior solutions.

In presenting the solutions we are going to avoid labelling the coordinates according to the cylindrical system foreseen in Eq. (7) because the nature of the final solution does not in general warrant the interpretation. For $\beta > 0$, that is, for ω-like geodesics, there are two types of solutions distinguished by the sign of P_ω. If $P_\omega < 0$, the space—time metric becomes

$$ds^2 = -[(x^1)^{1-c} - \nu^2(x^1)^{1+c}](dx^0)^2 + 2\nu(x^1)^{1+c}(dx^0\, dx^3)$$
$$+ (x^1)^{-(1-c^2)/2}[(dx^1)^2 + (dx^2)^2] + (x^1)^{1+c}(dx^3)^2, \qquad (16)$$

where

$$c = 4\beta^{1/2} > 0, \qquad (17)$$

ν is an arbitrary constant, and the range of x^1 is restricted to lie between the values

$$0 < x^1 < |\nu|^{-1/c}. \qquad (18)$$

where as for $P_\psi > 0$ we find

$$ds'^2 = -[\nu^2(x^1)^{1+c} - (x^1)^{1-c}](dx^0)^2 + 2\nu(x^1)^{1+c}(dx^0\, dx^3)$$
$$+ (x^1)^{-(1-c^2)/2}[(dx^1)^2 + (dx^2)^2] - (x^1)^{1+c}(dx^3)^2, \qquad (19)$$

where the definition of the constant c is the same as in Eq. (17), but the range of x^1 is now given by

$$|\nu|^{-1/c} < x^1. \qquad (20)$$

These solutions were first obtained by Lewis,[6] and we shall refer to them as $L\omega_-$ and $L\omega_+$ solutions respectively. If in Eqs. (16) and (19) we take $c = 1$, the result is flat space—time, where $L\omega_-$ is written in a uniformly rotating coordinate system with angular velocity ν,

while for $L\omega_+$ the rotation is a hyperbolic one. Except for this case we must restrict the range of x^1 to the limits in Eqs. (18) and (20). These limiting values correspond to apparent singularities of the metric, but it is not immediately clear that such a restriction is necessary. The fact that stationary observers in, for example, the $L\omega_-$ geometry have their angular velocities bounded by the limits

$$\Omega = -\nu \pm (x^1)^{-c} \tag{21}$$

suggests that $x^1 = 0$ might be a horizon and $x^1 = |\nu|^{-1/c}$ an infinite red-shift surface. However, the curvature invariants become unbounded at both of these points so that they are real singularities and the inequality (18) must hold. Similarly the limits excluded by Eq. (20) are also curvature singularities.

We now turn to the case $\beta < 0$ which correspond to ψ-like goedesics. For this case the space—time metric becomes

$$ds^2 = - x^1 \cos u (dx^0)^2 + 2x^1 \sin u \, dx^0 \, dx^3$$
$$+ (x^1)^{-(1+c^2)/2}[(dx^1)^2 + (dx^2)^2] + x^1 \cos u \, (dx^3)^2, \tag{22}$$

where

$$u = - c \ln(x^1/a), \, c = 4(-\beta)^{1/2} > 0, \tag{23}$$

and a is an arbitrary positive constant. The range of x^1 must now be restricted by

$$a \, e^{\pi(2n-1/2)/c} < x^1 < a \, e^{\pi(2n+1/2)/c}, \tag{24}$$

where both of these excluded limits once again correspond to real singularities. The periodic nature of this solution enables us to choose any n which is an integer. This solution is also due to Lewis,[6] and we shall refer to it as $L\psi$ solution.

The space—time metric corresponding to the case of "null" geodesics where $\beta = 0$ is given by

$$ds^2 = - x^1(1 + a \ln x^1)(dx^0)^2 + 2x^1(dx^0 \, dx^3)$$
$$+ (x^1)^{-1/2}[(dx^1)^2 + (dx^2)^2], \tag{25}$$

where a is an arbitrary constant of integration and, depending on the choice of sign for a, we must restrict x^1 as follows:

$$a < 0, \quad 0 < x^1 < e^{-1/a};$$
$$a > 0, \quad x^1 > e^{-1/a}. \tag{26}$$

All the excluded limiting values for x^1 are curvature singularities. This solution which is due to van Stockum[7] will be denoted by V.

4. DISCUSSION OF THE RESULTS

Theorems on the solution of the partial differential equation of Hamilton and Jacobi enable us to conclude that the solutions obtained in the previous section exhaust the class of stationary exterior solutions of the Einstein field equations with whole cylinder symmetry. We had found that according to the values of the constants of motion on M' they could be divided into the categories $L\omega_-$, $L\omega_+$, $L\psi$, and V. In order to present them in unified scheme we shall turn to an examination

of the phase space for M'. First of all the dynamics[10] of stationary, cylindrically-symmetric gravitational fields takes place on the phase space T^*M' which is the cotangent bundle of M'. This is implicit in our procedure of starting with the Einstein field equations written in the form of geodesic equations of M' and first casting them into Hamiltonian form. In our discussion the central role played by the symplectic structure of T^*M' is manifest even in the process of integration of these equations since it is based on the Hamilton—Jacobi theory. Finally the solutions themselves can be characterized by identifying the appropriate region in the phase space. For purposes of distinguishing between the four classes of solutions it suffices to consider a two dimensional submanifold spanned by the momenta P_ψ, P_ω, and in Fig. 1 we can see an illustration of this situation. The most crucial constant which decides the class of the exterior metric is the "energy" constant β. In particular $\beta = 1/16$ leads to flat space—time, and, as we can see from Fig. 1, the locus of such points are the two hyperbolas which intersect the P_ψ axis at $\pm 1/4$. The $L\omega_-$ solutions are hyperbolae which lie within the "past" null cone of this figure. Similarly $L\omega_+$ solutions are hyperbolas contained in the "future" null cone. The $L\psi$ solutions are represented by the hyperbolas which lie in the "elsewhere" region, that is, they intersect the P_ω axis and there is no distinction between the two branches in this case. Finally the null geodesics which are the asymptotes of all these hyperbolas coorespond to the class V solutions.

CONCLUSION

The virtue of formulating the Einstein field equations in terms of the harmonic mappings of Riemannian manifolds lies in the prominence it gives to the configuration space M'. We have seen that this prominence is justly deserved because of the simplicity it brings to the problems of the formulation as well as the solution of the stationary vacuum gravitational fields with whole cylinder symmetry.

ACKNOWLEDGMENTS
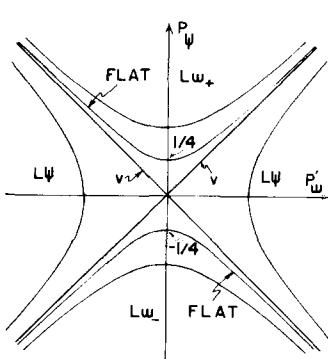
FIG. 1. A submanifold of the phase space for stationary, cylindrically-symmetric gravitational fields. In this figure we display the regions corresponding to the different classes of solutions which together form the most general solution of the Einstein field equations. Particular examples of the indicated class of solutions are the hyperbolas and the asymptotes in this figure. Note that we have $P'_\omega = 2 \, e^{-2\psi} P_\omega$ for the horizontal axis here.

and to Professor F. Öktem for a critical reading of the manuscript. We are also indebted to Professor J. Stachel for helpful suggestions, in particular for kindly pointing out to us the work of Lewis and van Stockum.

APPENDIX

The formalism of Sec. 2 includes as a special case the formulation of spherically symmetric gravitational fields in the framework of harmonic mappings of Riemannian manifolds. However, some simplifications which ensue as a result of this specialization, in particular the reduction of the discussion of spherical fields to geodesic motion in configuration space, are not immediate in the way the problem has been set up for axially symmetric systems. For this purpose we need to start with a new form of the space–time metric, taking into account all the available symmetries. It is well known that the general line element for static spherically symmetric fields can be written in the form

$$ds^2 = -B^{-2} dt^2 + A^{-2} dr^2 + r^2 (d\theta^2 + \sin^2\theta \, d\phi^2), \tag{A1}$$

where A, B are functions of r only. Looking at the Einstein Lagrangian which is obtained in this case,[11] we observe that the solutions of the Einstein field equations can be formulated as the harmonic maps $f: M \to M'$, where

$$ds^2 = dr^2 + r^2 (d\theta^2 + \sin^2\theta \, d\phi^2), \tag{A2}$$

$$ds'^2 = \frac{1}{4AB} \left(\frac{dA^2}{A^2} + 2 \frac{d\,A dB}{AB} \right) \tag{A3}$$

are the metrics for M, M' respectively. In this case the geometry of M' is flat and the transformation

$$A = (\alpha + \beta)^{-2}, \quad B = (\alpha + \beta)/(\alpha - \beta), \tag{A4}$$

brings Eq. (A3) to the form

$$ds'^2 = d\alpha^2 - d\beta^2. \tag{A5}$$

As we have restricted our considerations to harmonic maps which depend only on r, we are interested in the geodesics of this two-dimensional Minkowski space with the affine parameter given by r^{-1}. The discussion of the geodesics is simplified by noting that we have the free-dom of performing Lorentz transformations on M'. The coefficients of the space–time metric (A4) contain α, β only through the retarded and advanced combinations so that they are left invariant under these transformations. (For null geodesics on M' the resulting space–time geometry is therefore singular.) Finally the discrete symmetry $\alpha \to \beta$, $\beta \to \alpha$ of the metric coefficients shows that there is no distinction between α- or β-like geodesics as far as the resulting space–time geometry is concerned. Thus it will suffice to consider a single geodesic which will be a straight line parallel (say) to the β axis, and its equation can be written in the form

$$\alpha = 1, \quad \beta = \tfrac{1}{2} mr^{-1} \tag{A6}$$

where $\tfrac{1}{2}m$ is an arbitrary constant which measures the velocity of the particle in M'. If Eqs. (A6) are substituted into the space–time metric (A1), we find the Schwarzschild solution where m is the Schwarzschild mass. The null cone at the origin of M' distinguishes between the Schwarzschild and Kantowski–Sachs regions of the solution. The α-axis is the asymptotically flat region of the Schwarzschild geometry and crossing the null cone in M' corresponds to passage through the Schwarzschild horizon.

[1]Y. Nutku, Ann. Inst. H. Poincaré A21, 175 (1974).
[2]J. Eells, Jr., and J.H. Sampson, Am. J. Math. 86, 109 (1964); Ann. Inst. Fourier 14, 61 (1964).
[3]R.A. Matzner and C.W. Misner, Phys. Rev. 154, 5, 1229 (1967).
[4]R.A. Matzner (unpublished).
[5]G. Neugebauer and D. Kramer, Ann. Physik 24, 62 (1969).
[6]T. Lewis, Proc. Roy. Soc. (London) A 136, 176 (1932).
[7]W.J. van Stockum, Proc. Roy. Soc. (Edinburgh) 57, 135 (1937).
[8]The adjective Riemannian should really be pseudo-Riemannian since the metrics of interest here need not be positive definite. However, we shall always drop the prefix "pseudo".
[9]R.A. Matzner and Y. Nutku, Astrophys. J. 167, 149 (1971).
[10]In this connection we are using the word "dynamics" loosely for "evolution" in a spacelike direction.
[11]F.A. Kaempffer, Phys. Rev. 165, 1420 (1968).

Orthogonal polynomials. II*

K. M. Case

The Rockefeller University, New York, New York 10021
(Received 20 January 1975)

A class of orthogonal polynomials defined by a weight function of compact support is considered. These are known to satisfy three-term recursion relations. It is shown, under rather weak restrictions, that the traces of powers of the Jacobi matrices formed from the coefficients in the recursion relations are simply related to Fourier coefficients of the logarithm of the weight function.

I. INTRODUCTION

Recently[1] we presented a treatment of a class of orthogonal polynomials from the viewpoint of scattering theory. In particular, a set of sum rules involving the coefficients of the three term recursion relations which these polynomials satisfy emerged rather directly. Experts in the inverse scattering method of solving nonlinear evolution equations will no doubt note the close similarity to the "trace formulas" which occur in that method.[2] Here we show the detailed connection: First, the relation throws considerable light on the nature of the sum rules. Second, this leads to a direct and systematic method to write down all the sum rules.

In Sec. II we briefly summarize the pertinent material from Ref. 1. A derivation of the relevant trace formulas is then presented in Sec. III. While elegant and short, this derivation (at least in the form given) seems to demand very strong conditions. However, in the Appendix it is shown by a more cumbersome approach that the formulas are actually valid under very much weaker restrictions. Later sections are denoted to the explicit form of the first several sum rules. As an application we then write these out for the case of the Legendre polynomials.

II. REVIEW

Here we give a brief summary of the essentials of the previous work (hereafter denoted as I).

Consider polynomials which are orthogonal with respect to a weight function $\rho(\lambda)$ with support that is a finite continuous stretch of the real axis plus, perhaps, a finite number of real discrete points outside of the continuum region, i.e.,

$$d\rho(\lambda) = \rho'(\lambda)\, d\lambda, \quad \lambda_1 \le \lambda \le \lambda_2, \tag{II.1}$$

and $d\rho(\lambda) = \sum_i \rho_i \delta(\lambda - \lambda_i)\, d\lambda$, λ not as above.

It is well known[3] that these polynomials $\psi(\lambda, n)$ satisfy a three term recursion relation of the form

$$a(n+1)\psi(\lambda, n+1) + b(n)\psi(\lambda, n) + a(n)\psi(\lambda, n-1) = \lambda\psi(\lambda, n) \tag{II.2}$$

subject to the initial conditions

$$\psi(\lambda, -1) = 0, \quad \psi(\lambda, 0) = C = 1/\left[\int_{-\infty}^{\infty} d\rho(\lambda)\right]^{1/2}. \tag{II.3}$$

Here the limits $a(\infty)$, $b(\infty)$ exist and the continuum region is

$$b(\infty) - 2a(\infty) \le \lambda \le b(\infty) + 2a(\infty). \tag{II.4}$$

It proves convenient to introduce auxiliary functions $\psi_\pm(z, n)$ which satisfy Eq. (II.2), but instead of the boundary conditions of Eq. (II.4) they satisfy

$$\lim_{n \to \infty} |\psi_\pm(z, n) - z^{\pm n}| = 0, \quad |z| \le 1, \quad |z| \ge 1 \tag{II.5}$$

where

$$\lambda = b(\infty) + a(\infty)(z + z^{-1}). \tag{II.6}$$

Particularly important is the function $f_+(z)$ defined through Eq. (II.2) as

$$f_+(z) = \lambda\psi_+(\lambda, 0) - b(0)\psi_+(\lambda, 0) - a(1)\psi_+(\lambda, 1). \tag{II.7}$$

(This will be called the Jost function.)

Basic properties are

(a) $f_+(z)$ is analytic within the unit circle except for a simple pole at $z = 0$;

(b) the zeros z_i within the unit circle correspond to the discrete jump points;

(c) the boundary value of f_+ as the unit circle is approached from within determines the continuous part of the spectral function. Indeed,

$$\rho' = a(\infty) \sin\theta / \pi C^2 |f_+|^2, \tag{II.8}$$

where

$$z = \exp(i\theta).$$

(d) Conversely, the continuous part of the spectral function plus the z_i (which are real and $|z_i| < 1$) corresponding to the discrete λ_i determine $f_+(z)$. We have the explicit formula

$$f_+(z) = \frac{\Pi_+(z_i - z)\Pi_-(z - z_i)}{z\Pi_i(1 - z_i z)}$$

$$\times \exp\frac{1}{4\pi}\int_{-\pi}^{\pi} \left(\ln\left|\frac{a(\infty)\sin\theta'}{\pi C^2 \rho'}\right|\right)\left(\frac{\exp(i\theta') + z}{\exp(i\theta') - z}\right) d\theta'. \tag{II.9}$$

[Here $\Pi_\pm$ means the product over positive (negative) z_i.]

Some simplification is achieved if we introduce

$$\alpha(n) = a(n)/a(\infty), \quad \beta(n) = \frac{b(n) - b(\infty)}{a(\infty)},$$

$$u = z + z^{-1} = \frac{\lambda - b(\infty)}{a(\infty)}. \tag{II.10}$$

Then Eq. (II.2) becomes

$$(L\psi)(u, n) = u\psi(u, n), \tag{II.11}$$

with

$$(L\psi)(u, n) = \alpha(n + 1)\psi(u, n + 1) + \beta(n)\psi(u, n) + \alpha(n)\psi(u, n - 1).$$
$$\text{(II. 12)}$$

[Effectively the general problem has been reduced to the case $a(\infty) = 1$, $b(\infty) = 0$.]

Let us also introduce a comparison operator L_0 defined by

$$(L_0\psi^{(0)})(u, n) = \psi^{(0)}(u, n + 1) + \psi^0(u, n - 1). \qquad \text{(II. 13)}$$

We denote the solutions of

$$L_0\psi^{(0)} = u\psi^{(0)}$$

subject to the initial conditions of Eq. (II. 3) by $\psi^{(0)}$ and the solutions corresponding to the conditions of Eq. (II. 5) by $\psi_\pm^{(0)}$.

III. THE TRACE FORMULAS

Our program is to relate the traces

$$\mathrm{Tr}(L^p - L_0^p)$$

to various integrals involving the spectral function. The derivation given here may be regarded as heuristic. The argument seems to demand very stringent conditions on the approach of $\alpha(n)$ and $\beta(n)$ to their asymptotic values. However, in the Appendix it will be shown that the formulas do indeed hold under much weaker conditions.

Our program (which closely follows Flaschka[2]) is as follows: We first demonstrate the relation

$$\ln\frac{zf_+(z)}{C_{-1}} = \mathrm{Tr}\sum_{m=1}^{\infty}\frac{1}{mu^m}[L^m - L_0{}^m]. \qquad \text{(III. 1)}$$

(Here C_{-1} is the residue of f_+ at $z = 0$.) Then express f_+ in terms of ρ' by means of Eq. (II. 9). Expanding both sides of Eq. (II. 1) in a Taylor series about $z = 0$ and equating coefficients of the different powers of z then gives the desired formulas.

To prove Eq. (II. 1) let us first find the Green's function corresponding to the operator L, i.e.,

$$(L - u1)G = 1. \qquad \text{(III. 2)}$$

A straightforward construction[4] yields

$$G(u, n;m) = -\frac{\psi(u, n)\psi_+(u, m)a(\infty)}{Cf_+}, \quad n \leq m$$

$$= -\frac{\psi_+(u, n)\psi(u, m)a(\infty)}{Cf_+}, \quad n \geq m. \qquad \text{(III. 3)}$$

Hence

$$\mathrm{Tr}\left(\frac{1}{L - u1} - \frac{1}{L_0 - u1}\right) = -\frac{a(\infty)}{C}\sum_{n=0}^{\infty}\left(\frac{\psi(u, n)\psi_+(u, n)}{f_+}\right.$$

$$\left. -\frac{\psi^0(u, n)\psi_+^0(u, n)}{f_+^0}\right). \qquad \text{(III. 4)}$$

To evaluate the sums let us differentiate with respect to z the Eq. (I. 11) for ψ_+. [$\dot{\psi}_+ = (d/dz)\psi_+$]. We obtain

$$L\dot{\psi}_+ = \frac{du}{dz}\psi_+ + u\dot{\psi}_+. \qquad \text{(III. 5)}$$

Applying the usual operations leading to Green's identity to this and the equation

$$L\psi = u\psi, \qquad \text{(III. 6)}$$

yields the relation

$$\frac{du}{dz}\sum_{n=0}^{N}\psi(u, n)\psi_+(u, n)$$

$$= a(N + 1)[\psi(u, N)\dot{\psi}_+(u, N + 1) - \dot{\psi}_+(u, N)\psi(u, N + 1)]$$

$$+ C\dot{f}_+/a(\infty). \qquad \text{(III. 7)}$$

(Here we have used the boundary conditions on ψ and the definition of f_+.)

Now we want to show that in the limit $N \to \infty$ the difference between Eq. (II. 7) and the same expression with superscripts zero give no contribution from the upper limit. To show this we need the asymptotic behavior of $\psi(u, n)$ and $\psi^0(u, n)$. For $|z| = 1$ this is readily done. Thus in this case the $\psi_\pm$ are both well defined and are independent solutions of our equation. Therefore, we have the representation

$$\psi(u, n) = \frac{C}{z - z^{-1}}[f_-(z)\psi_+(u, n) - f_+(z)\psi_-(u, n)], \qquad \text{(III. 8)}$$

and then the asymptotic behavior

$$\psi(u, n) \to \frac{C}{z - z^{-1}}[f_-(z)z^n - f_+(z)z^{-n}]. \qquad \text{(III. 9)}$$

However, we would like to use this representation well *within* the unit circle. (Indeed to $z = 0$.) Suppose ψ_- can be analytically continued within the unit circle. Then Eqs. (II. 8), (II. 9) will hold there. This can be done at least when $\alpha(n)$, $\beta(n)$ approach their asymptotic values faster than any exponential. Therefore, at least in this case our following derivation is valid. For weaker convergence it can only be regarded as heuristic. In the Appendix, however, we demonstrate that the formulas are valid even when the convergence is as slow as $1/n^2$.

Inserting the expansion of Eq. (II. 9) in Eq. (III. 7), doing the same for quantities with zero subscripts, subtracting and passing to the limit $N \to \infty$ yields

$$\mathrm{Tr}\left[\frac{1}{L - u1} - \frac{1}{L_0 - u1}\right] = -\frac{dz}{du}\left(\frac{\dot{f}_+}{f_+} - \frac{\dot{f}_+^0}{f_+^0}\right)$$

$$\equiv -\frac{d}{du}\ln\{f_+/f_+^0\}. \qquad \text{(III. 10)}$$

If now the left hand of the equation is expanded in powers of $1/u$ and integrated from u to infinity, we obtain

$$-\sum_{m=1}^{\infty}\frac{1}{mu^m}\mathrm{Tr}(L^m - L_0^m).$$

Integrating the right-hand side yields

$$\lim_{z_0 \to 0}\ln\frac{f_+(z)f_+^0(z_0)}{f_+^0(z)f_+(z_0)}.$$

But $f_+^0(z) = 1/z$, and $f_+(z_0) \to C_{-1}/z_0$.

Therefore, we obtain

$$\ln\frac{zf_+(z)}{C_{-1}} = -\sum_{m=1}^{\infty}\frac{1}{mu^m}\mathrm{Tr}(L^m - L_0^m). \qquad \text{(III. 11)}$$

Let us now expand both side in powers of z. Thus writing

$$\ln\frac{zf_+(z)}{C_{-1}} = \sum_{n=1}^{\infty}\gamma_n z^n \qquad \text{(III. 12)}$$

and

$$\frac{1}{u^m} = \sum_m \sigma_{mn} z^n \qquad \text{(III. 13)}$$

and equating coefficients of z^n in Eq. (III. 11), we obtain the relations

$$\gamma_n = \sum_{m=1}^{n} \frac{\sigma_{mn}}{m} \text{Tr}(L_0^m - L^m). \qquad \text{(III. 14)}$$

Some remarks are in order.

(1) The σ_{mn} are universal and simple. Thus,

$$\sigma_{mn} = 0, \quad m > n,$$

$$\sigma_{mn} = 0, \quad \text{unless } n - m \text{ is even, and otherwise,}$$

$$\sigma_{mn} = \begin{pmatrix} \frac{1}{2}(n+m-1) \\ m-1 \end{pmatrix} - 1 \quad (n-m)/2 \qquad \text{(III. 15)}$$

(2) The γ_n are simply expressed in terms of integrals of the spectral function. Thus from Eq. (I. 9) we see that

$$C_{-1} = \prod_i |z_i| \exp\frac{1}{4\pi} \int_{-\tau}^{\tau} \ln\left|\frac{a(\infty)\sin\theta'}{\pi C^2 \rho'}\right| d\theta'. \qquad \text{(III. 16)}$$

From this it follows that

$$\ln\frac{z f_*(z)}{C_{-1}} = \sum_i \ln(1 - z/z_i) - \sum_i \ln(1 - z_i z)$$

$$+ \frac{z}{2\pi} \int_{-\tau}^{\tau} \ln\left|\frac{a(\infty)\sin\theta'}{\pi C^2 \rho'}\right| \frac{\exp(-i\theta')\,d\theta'}{1 - z\exp(-i\theta')}. \qquad \text{(III. 17)}$$

Finally, expanding in powers of z and equating coefficients, we conclude that

$$\gamma_n = \frac{1}{\pi} \int_0^{\tau} \cos n\theta' \ln\left|\frac{a(\infty)\sin\theta'}{\pi C^2 \rho'}\right| d\theta'$$

$$+ \sum_i \frac{[(z_i)^n - (z_i)^{-n}]}{n}. \qquad \text{(III. 18)}$$

IV. SOME EXAMPLES

It is instructive to look in detail at the structure of these sum rules.

(1) $n = 1$

$$\gamma_1 = \sigma_{11} \text{Tr}(L_0 - L) \qquad \text{(IV. 1)}$$

But $\sigma_{11} = 1$ and $\text{Tr}(L_0 - L) = -\sum_{n=0}^{\infty} \beta(n)$, ie.,

$$\gamma_1 = -\sum_{n=0}^{\infty} \beta(n) \qquad \text{(IV. 2)}$$

(2) $n = 2$

$$\gamma_2 = \sigma_{12} \text{Tr}(L_0 - L) + \frac{\sigma_{22}}{2} \text{Tr}(L_0^2 - L^2).$$

Since $\sigma_{12} = 0$, $\sigma_{22} = 1$, and

$$\text{Tr}(L_0^2 - L^2) = -2\sum_{n=1}^{\infty} [\alpha^2(n) - 1] - \sum_{n=0}^{\infty} \beta^2(n).$$

Thus,

$$\gamma_2 = -\sum_{n=1}^{\infty} [\alpha^2(n) - 1] - \frac{1}{2}\sum_{n=0}^{\infty} \beta^2(n) \qquad \text{(IV. 3)}$$

(3) $n = 3$

$$\sigma_{13} = -1, \quad \sigma_{23} = 0, \quad \sigma_{33} = 1$$

$$\therefore \gamma_3 = -\text{Tr}(L_0 - L) + \tfrac{1}{3}\text{Tr}(L_0^3 - L^3),$$

or using Eq. (IV. 1) and the explicit form of the traces:

$$\gamma_3 + \gamma_1 = -\sum_{n=0}^{\infty} \alpha^2(n+1)[\beta(n) + \beta(n+1)] + \frac{\beta^3(n)}{3} \qquad \text{(IV. 4)}$$

(4) $n = 4$

$$\sigma_{14} = \sigma_{34} = 0, \quad \sigma_{24} = -2, \quad \sigma_{44} = 1$$

$$\therefore \gamma_4 + 2\gamma_2 = \tfrac{1}{4}\text{Tr}(L_0^4 - L^4)$$

$$= \frac{1}{4}\sum_{n=0}^{\infty}\{b^4(n) + 2(\alpha^4(n) - 1)$$

$$+ 4[(\alpha^2(n)\alpha^2(n+1) - 1) + \alpha^2(n)\beta^2(n)$$

$$+ \alpha^2(n+1)\beta^2(n) + \alpha^2(n+1)\beta(n)\beta(n+1)]\}. \qquad \text{(IV. 5)}$$

(5) We note that in the identities with n odd (even) only traces of odd (even) powers occur. In particular, in the important special case that $\beta(n) \equiv 0$ half of our identities are the trivial $0 = 0$. Therefore, as our last example, we consider $\beta(n) \equiv 0$ and treat the case $n = 6$. Then

$$\sigma_{16} = \sigma_{36} = \sigma_{56} = 0,$$

$$\sigma_{66} = 1, \quad \sigma_{26} = 3, \quad \sigma_{46} = -4$$

and the identity is

$$\gamma_6 + 4\gamma_4 + 5\gamma_2 = \tfrac{1}{6}\text{Tr}(L_0^6 - L^6)$$

$$= -\frac{1}{3}\sum_{n=0}^{\infty} [\alpha^2(n+1) - 1]$$

$$- \sum_{n=0}^{\infty}\{\alpha^2(n+1)[\alpha^2(n+1)\alpha^2(n+2) + \alpha^4(n+2)$$

$$+ \alpha^2(n+2)\alpha^2(n+3)] - 3\}. \qquad \text{(IV. 6)}$$

V. THE CASE OF LEGENDRE POLYNOMIALS

We consider this as the simplest application. There are no z_i. $\rho' \equiv 1$, $a(\infty) = 1/2 = C^2$. Thus,

$$\gamma_n = \frac{1}{\pi} \int_0^{\tau} \cos n\theta' \ln\left(\frac{\sin\theta'}{\pi}\right) d\theta'$$

$$= \begin{cases} -1/n, & n \text{ even,} \\ 0, & n \text{ odd.} \end{cases} \qquad \text{(V. 1)}$$

The relation of Eq. (IV. 3) is then

$$\frac{1}{2} = \sum_{n=1}^{\infty} [\alpha^2(n) - 1], \qquad \text{(V. 2)}$$

which was previously given in Ref. I.

The relations of Eq. (IV. 5) and (IV. 6) are,

respectively,

$$5 = \sum_{n=1}^{\infty} \{2[\alpha^4(n) - 1] + 4[\alpha^2(n)\alpha^2(n+1) - 1]\} \quad \text{(V.3)}$$

and

$$11 = \sum_{n=1}^{\infty} [\alpha^6(n+1) - 1] + 3\sum_{n=0}^{\infty} \{\alpha^2(n+1)[\alpha^2(n+1)\alpha^2(n+2)$$
$$+ \alpha^4(n+2) + \alpha^2(n+2)\alpha^2(n+3)] - 3\}. \quad \text{(V.4)}$$

If we insert the known form for the coefficients

$$\alpha^2(n) = n^2/(n^2 - 1/4),$$

we obtain relations which can be checked by direct, if extremely tedious, calculation.

VI. CONCLUSION

It has been shown that under stringent convergence conditions for the $a(n)$, $b(n)$ there are extremely simple relations between the $\mathrm{Tr}(L_0^m - L^m)$ and the Fourier coefficients of the logarithm of the spectral function. These are summarized in Eqs. (III.14), (III.15), and (III.18). The extreme simplicity of the results suggest that they are true more generally. In the Appendix it is shown that the relations indeed hold even if the convergence of $a(n)$, $b(n)$ to their asymptotic values is as slow as $1/n^2$.

APPENDIX

In the main text we have found very simple sum rules. These were proved under restrictions on the $\alpha(n)$, $\beta(n)$. The rules very simplicity suggests they may hold more generally. Here we sketch a proof of this by direct computation.

The essence of the rules found earlier is that there is a relation between the coefficients of the power series expansion of $\ln(zf_+(z)/C_{-1})$ in the vicinity of the origin and the various traces of powers of L and L_0. However, previously[1] we found that $zf^+(z)/C_{-1} \to 1$ as $z \to 0$ and it is analytic in the vicinity even if $\alpha(n) \to_{n \to \infty} 1$ and $\beta(n) \to_{n \to \infty} 0$ as slowly as $1/n^2$. Hence there is a power series development. We will show the coefficients are readily computed and indeed are the traces that we expect.

To find, systematically, the power series of $\psi_+(z, n)$ [and hence that of $f_+(z)$] it is convenient to introduce new variables. Thus suppose we write

$$\alpha(n) = 1/\sqrt{g(n-1)g(n)}. \quad \text{(A1)}$$

[We will never actually need the $g(n)$ explicitly but will require that $\lim_{n \to \infty} g(n) = 1$.] Then introducing $\phi_+(z, n)$ by

$$\phi_+(z, n) = \psi_+(z, n)/g(n) \quad \text{(A2)}$$

we find that the three term recursion relation is

$$\phi_+(z, n-1) = [z + z^{-1} - \beta(n)]g(n)\phi_+(z, n) - \phi_+(z, n+1), \quad \text{(A3)}$$

and $f_+(z)$ is given by

$$f_+(z) = \frac{a(\infty)}{\sqrt{g(0)}}\{[z + z^{-1} - \beta(0)]g(0)\phi_+(z, 0) - \phi_+(z, 1)\}. \quad \text{(A4)}$$

Since ϕ_+ satisfies the boundary condition $\lim_{n \to \infty} |\phi^+(z) - z^n| = 0$, we know that

$$\phi_+(z, n) = \sum_{m=0}^{\infty} A^{(n)}(m)z^{n+m} \quad \text{(A5)}$$

and then

$$f_+(z) = \frac{a(\infty)}{\sqrt{g(0)}}\left(g(0)\sum_{m=0}^{\infty} A^{(0)}(m)[z^{m+1} + z^{m-1} - \beta(0)]\right.$$
$$\left. - \sum_{m=0}^{\infty} A^{(1)}(m)z^{m+1}\right). \quad \text{(A6)}$$

If we expand $f_+(z)$ in the Laurent series

$$f_+(z) = \sum_{n=-1}^{\infty} C_n z^n \quad \text{(A7)}$$

and equate coefficients of the powers of z in Eq. (6), we obtain

$$C_n = \frac{a(\infty)}{\sqrt{g(0)}}[g(0)A^{(0)}(n-1) + g(0)A^{(0)}(n+1)$$
$$- g(0)\beta(0)A^{(0)}(n) - A^{(1)}(n-1)]. \quad \text{(A8)}$$

In particular,

$$C_{-1} = a(\infty)\sqrt{g(0)}A^{(0)}(0). \quad \text{(A9)}$$

Similarly, recursion relations for the $A^{(n)}(m)$ are obtained by inserting the expansion of Eq. (A5) into Eq. (A3). We obtain

$$A^{(n-1)}(m) = g(n)A^{(n)}(m-2) + g(n)A^{(n)}(m)$$
$$- \beta(n)g(n)A^{(n)}(m-1) - A^{(n+1)}(m-2). \quad \text{(A10)}$$

Note that this is ideally suited for an interative solution. If we know $A^{(n)}(m')$ for $m' < m$ we can then find $A^{(n)}(m)$.

To start the process consider $m = 0$. Then since $A^{(n)}(m) = 0$, $m < 0$, the equation is

$$A^{(n-1)}(0) = g(n)A^{(n)}(0), \quad \text{(A11)}$$

which yields

$$A^{(n)}(0) = \prod_{i=n+1}^{\infty} g(i). \quad \text{(A12)}$$

As a special case we have

$$A^{(0)}(0) = \prod_{i=1}^{\infty} g(i). \quad \text{(A13)}$$

Then Eq. (9) tells us that

$$C_{-1} = a(\infty)\sqrt{g(0)}\prod_{i=1}^{\infty} g(i), \quad \text{(A14)}$$

which on using Eq. (1) becomes

$$C_{-1} = a(\infty)\prod_{i=1}^{\infty} \frac{1}{\alpha(i)} = a(\infty)\prod_{i=1}^{\infty} a(\infty)/a(i). \quad \text{(A15)}$$

Some remarks:

(i) The explicit form for $g(i)$ is not needed.

(ii) The result agrees with one reported earlier.[1]

(iii) Our present approach has given us one additional relation.

Thus comparing (A15) with Eq. (III.16) we see that

$$a(\infty) \prod_{i=1}^{\infty} a(\infty)/a(i) = \prod_i |z_i| \exp \frac{1}{2\pi} \int_0^H \ln \frac{a(\infty)\sin\theta'}{C^2\rho'} \, d\theta'.$$

(A16)

To obtain the higher C_m it is convenient to decompose $A^{(n)}(m)$ in the form

$$A^{(n)}(m) = A^{(n)}(0)B^{(n)}(m),$$

(A17)

[where therefore $B^{(n)}(0) = 1$]. The recursion relation Eq. (10) becomes

$$B^{(n-1)}(m) - B^{(n)}(m) = K^{(n)}(m),$$

(A18)

where

$$K^{(n)}(m) = B^{(n)}(m-2) - B^{(n+1)}(m-2)$$
$$+ f(n)B^{(n+1)}(m-2) - \beta(n)B^{(n)}(m-1),$$

(A19)

with

$$f(n) = 1 - \frac{1}{g(n)g(n+1)} \equiv 1 - \alpha^2(n+1).$$

[Notice that again we no longer need the explicit form of $g(i)$—they completely disappear from the problem.] Now if $B^{(n)}(m')$ are known for $m' < m$ the Eq. (18) gives us $B^{(n)}(m)$. Indeed,

$$B^{(n)}(m) = \sum_{n'=n+1}^{\infty} K^{(n')}(m).$$

(A20)

Using Eq. (18), we can also slightly simplify Eq. (19) to

$$K^{(n)}(m) = K^{(n+1)}(m-2) + f(n)B^{(n+1)}(m-2) - \beta(n)B^{(n)}(m-1).$$

(A21)

With the initial conditions $B^{(n)}(0) = 1$, $K^{(n)}(0) = 0$, $B^{(n)}(m) = K^{(n)}(m) = 0$, for $m < 0$ we can now proceed to compute the higher $B^{(n)}(m)$ and $K^{(n)}(m)$ iteratively.

Before doing so we give the expression for C_m. These are

$$C_m/C_{-1} \equiv d_{m+1}$$

with

$$d_m = B^{(0)}(M-2) + B^{(0)}(m) - \frac{B^{(1)}(m-2)}{g(0)g(1)} - \beta(0)B^{(0)}(m-1),$$

(A22)

which using Eq. (19) and Eq. (20) yield

$$d_m = B^{(0)}(m) + K^{(0)}(m)$$

$$\equiv \sum_{n'=0}^{\infty} K^{(n')}(m).$$

(A23)

[The significance of the d_m is that the quantity $zf_*(z)/C_{-1}$, whose logarithm we eventually will count, has the Taylor series

$$\frac{zf_*(z)}{C_{-1}} = 1 + \sum_{m=1}^{\infty} d_m z^m. \,]$$

(A24)

Now let us proceed to calculate the d_m. From Eq. (21)

$$K^{(n)}(1) = -\beta(n)B^{(n)}(0) = -\beta(n)$$

and then

$$B^{(n)}(1) = -\sum_{n'=n+1}^{\infty} \beta(n'),$$

(A25)

while

$$d_1 = -\sum_{n=0}^{\infty} \beta(n) = \text{Tr}(L_0 - L).$$

(A26)

Similarly, for $m = 2$, we obtain from Eq. (21)

$$K^{(n)}(2) = f(n)B^{(n+1)}(0) - \beta(n)B^{(n)}(1)$$
$$= f(n) + \beta(n) \sum_{n'=n+1}^{\infty} \beta(n')$$

and so

$$B^{(n)}(2) = \sum_{n'=n+1}^{\infty} f(n') + \sum_{n'=n+1}^{\infty} \beta(n') \sum_{n''=n'+1}^{\infty} \beta(n'')$$

(A27)

and

$$d_2 = \sum_{n=0}^{\infty} f(n) + \sum_{n=0}^{\infty} \beta(n) \sum_{n'=n+1}^{\infty} \beta(n').$$

(A28)

But

$$\sum_{n=0}^{\infty} \beta(n) \sum_{n'=n+1}^{\infty} \beta(n') = -\frac{1}{2}\sum_{n=0}^{\infty} \beta^2(n) + \frac{1}{2}\left(\sum_{n=0}^{\infty} \beta(n)\right)^2.$$

$$\therefore \quad d_2 = \sum_{n=0}^{\infty} [1 - \alpha^2(n+1)] + \frac{1}{2}\sum_{n=0}^{\infty} \beta^2(n) + \frac{1}{2}\left(\sum_{n=0}^{\infty} B(n)\right)^2$$

$$= \frac{1}{2}\text{Tr}(L_0^2 - L^2) + \frac{1}{2}[\text{Tr}(L_0 - L)]^2.$$

(A29)

As a final example to illustrate what happens for higher m we consider the case $\beta(n) = 0$ and m up to 4.

Note that if $\beta(n) \equiv 0$, $K^{(n)}(m) = B^{(n)}(m) = d_m = 0$, m odd. Therefore, we need only d_2 and d_4. From Eqs. (27) and (28)

$$K^{(n)}(2) = f(n), \quad B^{(n)}(2) = \sum_{n'=n+1}^{\infty} f(n'),$$

$$d_2 = \sum_{n=0}^{\infty} f(n).$$

(A30)

Then from Eq. (21)

$$K^{(n)}(4) = f(n+1) + f(n) \sum_{n'=n+2}^{\infty} f(n')$$

and thus

$$d_4 = \sum_{n=0}^{\infty} f(n+1) + \sum_{n=0}^{\infty} f(n) \sum_{n'=n+2}^{\infty} f(n').$$

(A31)

The $(n+1)$ and $(n+2)$ occurring here appear a little strange—but are just what is needed for d_4 to be expressible as traces. Thus,

$$\sum_{n=0}^{\infty} f(n) \sum_{n'=n+2}^{\infty} f(n') = \sum_{n=0}^{\infty} f(n) \sum_{n'=n+1}^{\infty} f(n') - \sum_{n=0}^{\infty} f(n)f(n+1),$$

and since

$$\sum_{n=0}^{\infty} f(n) \sum_{n'=n+1}^{\infty} f(n') = \frac{1}{2}\left[\sum_{n=0}^{\infty} f(n)\right]^2 - \frac{1}{2}\sum_{n=0}^{\infty} f^2(n),$$

we have

$$d_4 = \sum_{n=0}^{\infty} f(n+1) + \frac{1}{2}\left[\sum_{n=0}^{\infty} f(n)\right]^2 - \frac{1}{2}\sum_{n=0}^{\infty} f^2(n) - \sum_{n=0}^{\infty} f(n)f(n+1).$$

Finally, remembering that $f(n) = 1 - \alpha^2(n+1)$, we obtain

$$d_4 = 2\sum_{n=0}^{\infty}\left[\alpha^2(n+1) - 1\right] + \frac{1}{4}\sum_{n=0}^{\infty}\{2[1 - \alpha^4(n+1)]$$

$$+ 4[1 - \alpha^2(n+1)\alpha^2(n+2)]\} + \frac{1}{2}\left(\sum_{n=0}^{\infty}\left[\alpha^2(n+1) - 1\right]\right)^2$$

or

$$d_4 = -\operatorname{Tr}(L_0^2 - L^2) + \tfrac{1}{4}\operatorname{Tr}(L_0^4 - L^4) + \tfrac{1}{2}[\tfrac{1}{2}\operatorname{Tr}(L_0^2 - L^2)]^2. \quad (A32)$$

Let us now show the expressions for the d_m obtained here just give those in the main text for γ_m. We have

$$\ln(zf_+(z)/C_i) = \ln\left(1 + \sum_{m=1}^{\infty} d_m z^m\right)$$

$$= \sum_{n=1}^{\infty} \gamma_n z^n. \quad (A33)$$

Expanding the logarithm in powers of z and equating coefficients give the γ_n in terms of d_m. The first few are

$$\gamma_1 = d_1,$$

which by Eq. (26) is

$$\gamma_1 = \operatorname{Tr}(L_0 - L)$$

$$\gamma_2 = d_2 - \frac{d_1^2}{2}, \quad (A34)$$

which using Eqs. (29) and (26) is

$$\gamma_2 = \tfrac{1}{2}\operatorname{Tr}(L_0^2 - L^2). \quad (A35)$$

Finally we note that for $\beta(n) \equiv 0$,

$$\gamma_n = 0 \quad (A36)$$

and

$$\gamma_4 = d_4 - \frac{d_2^2}{2}$$

$$= -\tfrac{2}{2}\operatorname{Tr}(L_0^2 - L^2) + \tfrac{1}{4}\operatorname{Tr}(L_0^4 - L^4). \quad (A37)$$

It is clear that the relations we have here are just the relations

$$\gamma_n = \sum_{m=1}^{n} \frac{\sigma_{mn}}{m}\operatorname{Tr}(L_0^m - L^m) \quad (A38)$$

of the main text.

One probably should emphasize the great simplification which arises from taking the expansion of the logarithm of $zf_+(z)/C_{-1}$ instead of the function itself.

(i) While both the d_m and γ_n are expressible in terms of $\operatorname{Tr}(L_0^r - L^r)$, the γ_n are *linear* in the traces and only involve even (odd) r for n even (odd).

(ii) The expression for the γ_n in terms of integrals of $\ln\rho'$ is particularly simple.

[1]K. M. Case, J. Math. Phys. 15, 2166 (1974).
[2]V. E. Zakharov and L. D. Faddeev, Funk. Anal. i evo Proloz 5, 18 (1971); H. Flaschka, Prog. Theor. Phys. 51, 703 (1974).
[3]See, for example, N. I. Akhiezer, *The Classical Moment Problem* (Oliver and Boyd, Edinburgh, 1965).
[4]See Appendix, K. M. Case and M. Kac, J. Math. Phys. 14, 594 (1973).

Analytic T matrices for Coulomb plus rational separable potentials

H. van Haeringen and R. van Wageningen

Natuurkundig Laboratorium der Vrije Universiteit, Amsterdam, The Netherlands
(Received 23 January 1975)

The $l = 0$ partial wave projected Coulomb off-shell T matrix $T_{c,l = 0}$ in momentum representation is obtained in closed form. Problems existing in the literature concerning the half- and on-shell behavior of T_c and $T_{c,l}$ are discussed and clarified by means of explicit formulas. The remaining derivations in this paper are based on $T_{c,l = 0}$. We consider the class of N-term separable potentials where the form factors are rational functions of p^2 (in momentum representation). We prove that the $l = 0$ T matrix corresponding to the Coulomb potential plus any such so-called rational separable potential has a very simple form, namely, it can be written in terms of rational functions and the (simple) hypergeometric function with parameters $(1, i\gamma; 1 + i\gamma)$, where γ is the well-known Coulomb parameter. Explicit analytic formulas are derived for a number of simple members of the class, the Yamaguchi potential being one of them. In this particular case the expressions of Zachary and of Bajzer are reproduced who used a method based on the O_4 symmetry.

1. INTRODUCTION

The nonrelativistic few-body problem is considerably simplified by the use of nonlocal separable potentials in place of local potentials. This simplification is justified by the observation that short-range local potentials can be approximated by finite-rank separable potentials in a mathematically well-defined sense. In view of the importance of charged particles in few-body systems, the interest in studying potentials consisting of the sum of a short-range finite-rank potential and the Coulomb potential is not surprising.[1-7] Because of the long-range difficulties involved, often a screened Coulomb potential is proposed with a very large screening parameter.[1,8-11] For the pure Coulomb problem, wave functions and Green's functions in coordinate representation are known in closed form.[12-15] Several equivalent[16] momentum representation expressions (in three-dimensional space, $\mathbf{p} \in R^3$) are known for the Green's function and for the off-shell T matrix.[16-22] For the Coulomb plus Yamaguchi[23] potential the off-shell T matrix in momentum representation is known in closed form[5,6] for $l = 0$. This latter is derived using the O_4 group-theoretic approach first discussed at some length by Fock.[24]

As is well known, in conventional (short-range) potential scattering theory the physical scattering amplitude can be obtained by taking the on-shell limit of the off-shell T matrix. This is no longer true when the potential has a long range such as the Coulomb potential; in fact, the on-shell limit is not defined in this case. Because this trouble just comes from the behavior of the potential at large distance, the same fact holds when an arbitrary short-range potential is added to the Coulomb potential. Nevertheless, it is generally expected that one also can extract, in such a situation, all relevant physical information from the off-shell T matrix. Therefore, it seems to be very interesting to have explicit formulas for off-shell T matrices which describe the Coulomb interaction plus a rather general short-range interaction exactly. Once we have gotten the correct relation between the off-shell T matrices and the physical amplitudes, we could obtain information about the short-range interaction by comparing the

numerical values resulting from the theory and the experiment respectively. We have been able to establish such a relation, and to give a satisfactory justification for it. This will be reported in detail in a subsequent paper.[25]

In the present paper we obtain the $l = 0$ partial wave projected off-shell Coulomb T matrix in momentum representation $T_{c, l=0}$ analytically, starting from a known expression for the complete T_c in three-dimensional $\mathbf{p}$ space (unless stated otherwise we work in momentum representation). Two equivalent explicit formulas for $T_{c, l=0}$ are given at the end of Sec. 3. Concerning the half- and on-shell behavior of T_c and of $T_{c, l}$, some confusion has grown in the literature. As we said above, the difficulties arise exclusively from the long range of the Coulomb potential and remain unaltered when an arbitrary short-range potential is added. These difficulties are discussed in detail in Sec. 4 and clarified with the help of explicit formulas derived in Sec. 3.

In section 5 we give the general formulas to obtain the T matrix corresponding to an arbitrary N-term separable potential plus the Coulomb potential. In Secs. 6 and 7 we derive analytic expressions for a number of off-shell T matrices corresponding to Coulomb plus N-term separable potentials (for $l = 0$ only), starting from one of the explicit expressions for $T_{c, l=0}$ derived in Sec. 3. In particular, for the case of Coulomb plus the (one-term separable) Yamaguchi potential we find complete agreement with the formula of Zachary[5] and of Bajzer.[6] It should be noted, however, that we apply only special function theory. Nowhere in this paper do we use the O_4 group-theoretic approach of Zachary and Bajzer.

In Sec. 8 we define the very large class of rational separable potentials. We reveal the general structure of the explicit formulas for the off-shell T matrices corresponding to any such potential plus the Coulomb potential. It appears that the hypergeometric function $_2F_1(1, i\gamma; 1 + i\gamma; \cdot)$, where γ is the well-known Coulomb parameter,[12] is the only nonrational function involved. Moreover, the method to obtain any such T matrix in

exact closed form follows immediately from the derivations given in Secs. 6–8.

Finally, in Sec. 9 the results and implications are summarized and discussed.

2. NOTATION

Dirac notation is used throughout. Our normalization is determined by the use of delta functions as basis states in momentum representation,

$$\langle \mathbf{p} | \mathbf{k} \rangle = \delta(\mathbf{p} - \mathbf{k}); \tag{1}$$

in the lth partial wave space

$$\langle p | kl \rangle = k^{-2}\delta(p - k). \tag{2}$$

This implies the following basis states in coordinate representation

$$\langle \mathbf{r} | \mathbf{k} \rangle = (2\pi)^{-3/2} e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{3}$$

$$\langle r | kl \rangle = (2/\pi)^{1/2} i^l j_l(kr). \tag{4}$$

Here the variable k stands for the square root of the energy,

$$E = k^2. \tag{5}$$

Functions of k will be defined in the upper-half of the complex k-plane, i.e.,

$$\text{Im} k > 0. \tag{6}$$

Generalized functions (distributions) like (1) and (2) may be defined as limits of functions of k where the appropriate limit

$$\text{Im} k \to 0 \tag{7}$$

is taken, cf., e.g., Ref. 12. The partial wave projection of a rotationally invariant operator A is defined by

$$\langle x' | A_l | x \rangle \equiv \int d\hat{\mathbf{x}} \, P_l(\hat{\mathbf{x}}' \cdot \hat{\mathbf{x}}) \langle \mathbf{x}' | A | \mathbf{x} \rangle, \tag{8}$$

where x stands for r or p. This implies

$$\langle \mathbf{x}' | A | \mathbf{x} \rangle = \sum_{l=0}^{\infty} \frac{2l+1}{4\pi} P_l(\hat{\mathbf{x}}' \cdot \hat{\mathbf{x}}) \langle x' | A_l | x \rangle. \tag{9}$$

In Eq. (9) the infinite sum converges generally not pointwise, but only in the sense of distributions[12,26,27]; compare Sec. 4.

3. THE PARTIAL WAVE PROJECTED COULOMB T OPERATOR $T_{c,l}$

In this section an integral representation for $T_{c,l}$ is derived from one of the known expressions for the complete T_c. At the end of the section $T_{c,l=0}$ is explicitly obtained; the derivations of Secs. 6–8 are based on this expression for $T_{c,l=0}$.

Our starting point is the complete Coulomb T matrix T_c. There are several equivalent[16] expressions known for this function.[16-21] In particular, Schwinger's[19] derivation, based on the O_4 symmetry exhibited by the Coulomb potential, is elegant. Chen and Chen[16] give a fairly complete survey of the work done on two-body Coulomb amplitudes. In place of T_c the resolvent G_c is given by many authors, but in momentum representation the explicit expressions for G_c and T_c are related in a trivial way because one has

$$G_c(E) = G_0(E) + G_0(E) T_c(E) G_0(E). \tag{10}$$

Below we mention some useful expressions for T_c. We define the Coulomb potential

$$V_c(r) \equiv 2k\gamma/r. \tag{11}$$

Here the notation of Messiah[12] is followed. Units are such that $\hbar = 2m = 1$ where m is the reduced mass. It is important to note that the parameter γ is energy-dependent, for (recall $E \equiv k^2$, $\text{Im} k > 0$)

$$k\gamma = \text{real const.} \tag{12}$$

Defining

$$\mathbf{q} \equiv \mathbf{p}' - \mathbf{p}, \quad q \equiv |\mathbf{q}|,$$

we get from (11)

$$\langle \mathbf{p}' | V_c | \mathbf{p} \rangle = \frac{k\gamma}{\pi^2 q^2}. \tag{13}$$

Define further

$$x^2 \equiv 1 + (p^2 - k^2)(p'^2 - k^2)/k^2 q^2,$$

$$y \equiv (x + 1)/(x - 1),$$

then T_c can be expressed as

$$\langle \mathbf{p}' | T_c(k^2) | \mathbf{p} \rangle = \frac{k\gamma}{\pi^2 q^2} + \frac{4ik^3\gamma^2}{\pi^2(p^2 - k^2)(p'^2 - k^2)}$$

$$\times \int_0^1 dl \, l^{i\gamma}(1 - ly)^{-1}(1 - l/y)^{-1}. \tag{14}$$

The first term of the rhs is just the potential term; see Eq. (13). Using the integral representation[28]

$$_2F_1(1, i\gamma; 1 + i\gamma; y) = i\gamma \int_0^1 dl \, l^{i\gamma-1}(1 - ly)^{-1} \tag{15}$$

of this h.f. (hypergeometric function) we get from (14)

$$\langle \mathbf{p}' | T_c(k^2) | \mathbf{p} \rangle = \frac{k\gamma}{\pi^2 q^2}\left(1 + \frac{I(y)}{x}\right), \tag{16}$$

with

$$I(y) = {}_2F_1(1, i\gamma; 1 + i\gamma; y) - {}_2F_1(1, i\gamma; 1 + i\gamma; 1/y). \tag{17}$$

It is clear that $I(y)/x$ is invariant for the transformation $x \to -x$. Thus the sign of the real part of x may be chosen freely. Choose then

$$\text{Re} x < 0,$$

this implies

$$|y| < 1.$$

Using well-known properties of the h.f.[28] [cf., Eq. (32)] and writing it in its infinite series representation, we get the result

$$I(y) = 1 - (-y)^{i\gamma}\Gamma(1 + i\gamma)\Gamma(1 - i\gamma) + 2\gamma^2\sum_{n=1}^{\infty}\frac{y^n}{n^2 + \gamma^2}. \tag{18}$$

We can rewrite Eq. (14) in a different form,

$$\langle \mathbf{p}' | T_c(k^2) | \mathbf{p} \rangle = \frac{k\gamma}{\pi^2 q^2} + \frac{4k\gamma^2}{i\pi^2 q^2}\int_0^1 \frac{dl \, l^{i\gamma}}{4l - (1 - l)^2(x^2 - 1)}. \tag{19}$$

This expression is more convenient for the derivation of

$T_{c,l}$. According to (8), multiplying Eq. (19) by the Legendre polynomial $P_l(\hat{\mathbf{p}}' \cdot \hat{\mathbf{p}})$ and integrating over the angles yields $T_{c,l}$. Now we have the well-known relation between P_l and the Legendre function of the second kind Q_l:

$$Q_l(z) = \frac{1}{2} \int_{-1}^{1} d\alpha\, P_l(\alpha)(z-\alpha)^{-1}. \tag{20}$$

Equation (20) immediately yields the partial wave projected potential matrix elements from Eq. (13):

$$\langle p' | V_{c,l} | p \rangle = \frac{2k\gamma}{\pi p p'}\, Q_l\!\left(\frac{p^2+p'^2}{2pp'}\right). \tag{21}$$

Applying Eqs. (8) and (20) to Eq. (19) gives, after interchanging the order of integration (this is permitted, compare Sec. 4B),

$$\langle p' | T_{c,l}(k^2) | p \rangle = \frac{2k\gamma}{\pi p p'} \left(Q_l\!\left(\frac{p^2+p'^2}{2pp'}\right) - i\gamma \int_0^1 dt\, t^{i\gamma-1} Q_l(z) \right), \tag{22}$$

with

$$z = \frac{1}{2pp'}\left(p^2 + p'^2 - \frac{(1-t)^2}{t} \cdot \frac{(p^2-k^2)(p'^2-k^2)}{4k^2} \right). \tag{23}$$

Difficulties may occur in the integrals of the form $\int dt\, t^{i\gamma} \cdots$ in the neighborhood of $t = 0$. These are discussed in Sec. 4B. In effect we may assume that the exponent $i\gamma$ contains a small positive real part. We simplify Eq. (22) by means of integration by parts. The integrated term just cancels the first term in the rhs of Eq. (22); thus we get

$$\langle p' | T_{c,l}(k^2) | p \rangle = \frac{2k\gamma}{\pi p p'} \int_0^1 dt\, t^{i\gamma} \frac{d}{dt}\, Q_l(z). \tag{24}$$

Whether this formula is useful for a derivation of $T_{c,l}$ in closed form remains to be seen. In this paper we use only $T_{c,l=0}$ for further derivations. This function follows rather easily from Eq. (24). Because of the simple relation

$$\frac{d}{dz}\, Q_0(z) = (1-z^2)^{-1}, \tag{25}$$

the integral in Eq. (24) can (in case $l = 0$) be written in terms of h.f.'s. Rearranging and collecting terms yields the final result

$$\langle p' | T_{c,l=0}(k^2) | p \rangle = \frac{ik}{\pi p p'}\, D, \tag{26}$$

$$D = {}_2F_1(1,\ i\gamma;\ 1+i\gamma;\ a'a) - {}_2F_1(1,\ i\gamma;\ 1+i\gamma;\ a/a') \tag{27}$$

$$+ {}_2F_1(1,\ i\gamma;\ 1+i\gamma;\ 1/(a'a)) - {}_2F_1(1,\ i\gamma;\ 1+i\gamma;\ a'/a),$$

with

$$a \equiv (p-k)/(p+k), \tag{28}$$

$$a' \equiv (p'-k)/(p'+k). \tag{29}$$

A different expression for D is convenient if one wishes to study its half-shell behavior, i.e., $p' \to k$ thus $a' \to 0$. This expression reads

$$D \equiv \Gamma(1+i\gamma)\Gamma(1-i\gamma)[(-a'a)^{i\gamma} - (-a'/a)^{i\gamma}]$$

$$+ {}_2F_1(1,\ i\gamma;\ 1+i\gamma;\ a'a) - {}_2F_1(1,\ -i\gamma;\ 1-i\gamma;\ a'a)$$

$$- {}_2F_1(1,\ i\gamma;\ 1+i\gamma;\ a'/a) + {}_2F_1(1,\ -i\gamma;\ 1-i\gamma;\ a'/a). \tag{30}$$

Equation (30) can be obtained from Eq. (27) as follows. Utilizing the relation[29,30]

$$_2F_1(\lambda,\mu;\nu;z) = \frac{\Gamma(\nu)\Gamma(\mu-\lambda)}{\Gamma(\mu)\Gamma(\nu-\lambda)}(-z)^{-\lambda}$$

$$\times\ {}_2F_1(\lambda, 1-\nu+\lambda; 1-\mu+\lambda; 1/z)$$

$$+ \frac{\Gamma(\nu)\Gamma(\lambda-\mu)}{\Gamma(\lambda)\Gamma(\nu-\mu)}(-z)^{-\mu}$$

$$\times\ {}_2F_1(\mu, 1-\nu+\mu; 1-\lambda+\mu; 1/z), \tag{31}$$

we derive

$${}_2F_1(1, i\gamma; 1+i\gamma; z) + {}_2F_1(1, -i\gamma; 1-i\gamma; 1/z)$$

$$= 1 + \Gamma(1+i\gamma)\Gamma(1-i\gamma)(-z)^{-i\gamma}. \tag{32}$$

It is well known that ${}_2F_1$ has a branch point at infinity.[29] In Eq. (27) it is then clear that two of the four h.f.'s have to be transformed; this is done with the help of Eq. (32), and Eq. (30) is obtained at once. In Eq. (30), the singular behavior of D in the neighborhood of $a' = 0$ is written out in its most transparent form.

Finally, it should be noted that the off-shell $l = 0$ T matrix for the Hulthén potential is known in closed form.[31] Because the Hulthén potential can be seen as a screened Coulomb potential, one could in principle obtain $T_{c,l=0}$ from the analytic formula of Bahethi and Fuda, by letting the screening parameter go to infinity. This formula contains hypergeometric functions ${}_4F_3$ and ${}_3F_2$, however, so it seems that our (more direct) approach is easier.

4. HALF- AND ON-SHELL BEHAVIOR OF THE COULOMB T MATRICES

In this section we discuss the half-shell behavior $p' \to k$ and the on-shell behavior $p' \to k$, $p \to k$ of the complete T_c and of all partial wave projected $T_{c,l}$, i.e., we consider

$$\langle \mathbf{p}' | T_c(k^2) | \mathbf{p} \rangle$$

and

$$\langle p' | T_{c,l}(k^2) | p \rangle, \qquad l = 0,\ 1, \cdots,$$

for $p' \to k$ and for $p' \to k$, $p \to k$. Obviously, k has to be (almost) real, for p' and p are real by definition. Therefore, we take, in this section only, k exactly real (and positive) and we replace k by $k + i\epsilon$ with $\epsilon > 0$ [recall Eq. (6)], whenever it is important to distinguish between k and $k + i\epsilon$.

A. The functions T_c and $T_{c,l}$

We first concentrate on Eq. (19) where the complete T_c is given. The integral in the rhs exists on the condition that

Re$i\gamma > -1$. (33)

This restriction can be removed by replacing the integral along the real axis with a contour integral[19,20]

$$\int_0^1 dt\, t^{i\gamma}(\cdots) \;\Rightarrow\; \frac{1}{e^{-2\pi\gamma}-1}\int_C dt\, t^{i\gamma}(\cdots).$$ (34)

The contour C begins at $t = 1 + i\delta$ where the phase of t is zero, moves to the origin, circles it once and continues to $t = 1 - i\delta$ with $\delta \downarrow 0$. Equation (19) and (34) together give the most general expression for T_c.[32] Now we take p' equal to k in Eq. (19). Then x^2 becomes equal to one, and we see that the resulting integral in the rhs of Eq. (19)

$$\int_0^1 dt\, t^{i\gamma-1}$$ (35)

is divergent, because $i\gamma$ is imaginary. In this case (34) is of no help. The half-shell limit of T_c simply is not defined. In Eqs. (16), (17) this fact becomes manifest; $p' \to k$ implies $y \to 0$ (or $y \to \infty$) and it is well known that the h.f. has a branch point at infinity. In Eq. (18) the singularity is made explicit:

$$\lim_{y\to 0} y^{i\gamma}$$ (36)

is not defined. $T_{c,l}$ contains the same singularity. Taking $p' = k$ in Eqs. (22), (23) we get a divergent integral like (35). In the case $l = 0$ we also have the more explicit formula Eq. (30); the limit $p' \to k$ here again becomes a limit like (36).

Everything we have said about $p' \to k$ applies a fortiori to the on-shell limit $p' \to k$, $p \to k$.

Finally, we remark that for $p' \to p \neq k$ there appear singularities, too. From Eq. (21) it follows that $V_{c,l}$ has a logarithmic singularity, and in Eqs. (26), (27) it can be seen that the same holds for $T_{c,l=0}$, because this h.f. has a logarithmic singularity for the argument becoming equal to one.[28–30] Concerning these singularities we found a discrepancy in the literature (see Ref. 17 of Ref. 33 and compare Ref. 34). This discrepancy seems to have been caused mainly by the mistake of Nutt.[22] See the discussion following (47). These singularities are less interesting so we will not discuss them further in this paper.

B. The distributions T_c and $T_{c,l}$

In Sec. 4A we considered T_c and $T_{c,l}$ as ordinary functions. Now we want to discuss a different approach: We consider T_c and $T_{c,l}$ as distributions (also called generalized functions). Distributions are often tacitly applied by physicists. Such applications are not restricted to the well-known Dirac delta function/distribution. For example, the Coulomb potential in momentum representation can only be obtained with the help of distributions. Straightforward derivation of Eq. (13) from Eq. (11) goes wrong because the Fourier transform of $V_c(r)$ yields a divergent integral. As is (implicitly or explicitly) done in all textbooks, one multiplies by $e^{-\epsilon r}$ and takes the limit $\epsilon \downarrow 0$ after the integration has been carried out; this procedure easily yields Eq. (13) for $\langle p'|V_c|p\rangle$.

Because there are several types of distributions

(see Refs. 26, 27 and references quoted there), we ought to specify the particular one we are going to apply. The distribution appropriate to our purpose is in fact the same as studied by Herbst.[35] According to Herbst the pure Coulomb S matrix, usually written

$$\langle k\,|S_c|k'\rangle = \frac{\gamma}{2\pi i k}\, e^{2i\sigma_0(k)}\delta(k'^2-k^2)\left(\frac{1-\hat{k}'\cdot\hat{k}}{2}\right)^{-1-i\gamma},$$ (37)

is undefined as it stands because it is not an integrable function. Furthermore, any extension is unique only up to a distribution with support at $\hat{k}' = \hat{k}$. Herbst proves that there is at most one unitary extension (in the sense of distributions) of (37) to all k' and k. This extension is just given by the substitution

$$(1-\hat{k}'\cdot\hat{k})^{-1-i\gamma} \;\to\; \lim_{\eta\downarrow 0}(1-\hat{k}'\cdot\hat{k})^{\eta-1-i\gamma}.$$ (38)

This means that S_c is a distribution, defined by

$$\langle k\,|S_c|f\rangle = \lim_{\eta\downarrow 0}\frac{\gamma}{2\pi i k}\, e^{2i\sigma_0(k)}\int_{R^3} dk'\delta(k'^2 - k^2)$$ (39)

$$\times\left(\frac{1-\hat{k}'\cdot\hat{k}}{2}\right)^{\eta-1-i\gamma}\langle k'|f\rangle,$$

where $\langle k'|f\rangle$ is a continuously differentiable and square integrable function. The so-called test functions f belonging to this type of distribution are dense in $L^2(R^3)$.

Analogous to Herbst we define the substitution

$$(\cdots)^{i\gamma} \;\longrightarrow\; \lim_{\eta\downarrow 0}(\cdots)^{\eta+i\gamma}.$$ (40)

We apply (40) in every situation where $i\gamma$ appears in the exponent. It is important to note that this has nothing to do with the fact that $i\gamma$ itself contains a small real part when k contains a small imaginary part according to (12). Remark that there should be, in this respect, no distinction between a repulsive force and an attractive force. Indeed, there is no distinction because (40) is not essentially modified when γ is replaced by $-\gamma$. The limits $\epsilon \downarrow 0$ (ϵ occurring in $k + i\epsilon$) and $\eta \downarrow 0$ are taken independently, the latter after the former. There is an important difference between (38) and (40) as we use it for T_c and $T_{c,l}$. In (38) the singularity is worse than a pole while the singularities in T_c and $T_{c,l}$ are much weaker than poles; in effect they are quite harmless. Therefore, the corresponding test functions are certainly dense in the Hilbert space of square integrable functions.

Applying (40) to Eqs. (16), (18) [or to Eq. (19)] we see that the half- and on-shell T_c are exactly zero. From Eq. (30) it is seen that the same fact holds for $T_{c,l=0}$. This is so because

$$\lim_{\eta\downarrow 0}\lim_{p\to k}\lim_{\epsilon\downarrow 0}\left(\frac{p-k-i\epsilon}{p+k+i\epsilon}\right)^{\eta+i\gamma} = 0.$$ (41)

Obviously, the second and the third limit may be interchanged. Finally, in Eq. (23) z becomes equal to $(p^2 + p'^2)/(2pp')$ when $p' \to k$ or when $p' \to k$, $p \to k$, and in Eq. (22) the Legendre function Q_l then becomes a constant with respect to the integration variable t, the integral becomes elementary, and the half- and on-shell $T_{c,l}$ are exactly zero for all l.

In all other (nonsingular) points convention (40) makes

no difference whatsoever. In conclusion we may say that T_c and $T_{c,\,l}$ are only altered in the on- and half-shell points, where they were not defined before, and in these points they become exactly zero.

Convention (40) is of considerable value because it simplifies the evaluation of our integrals, while we still know what we are doing. Interchanging the order of integrations, as for example has been done to derive Eq. (22), is now justified with the help of (40). Because all occurring integrals are absolutely convergent, the theorems of Fubini and Tonelli may be applied[36] and thus the correct evaluation is easily performed.

C. The operators T_c and $T_{c,\,l}$

In this section we consider T_c and $T_{c,\,l}$ as operators (defined in the corresponding Hilbert spaces of square integrable functions). They may be defined in terms of the resolvents

$$G_0(E) \equiv (E - H_0)^{-1}, \tag{42}$$

$$G_c(E) \equiv (E - H_0 - V_c)^{-1}, \tag{43}$$

where H_0 is the kinetic energy operator. These definitions are

$$G_c = G_0 + G_0 T_c G_0, \qquad G_{c,\,l} = G_{0,\,l} + G_{0,\,l} T_{c,\,l} G_{0,\,l}, \tag{44}$$

where the energy-dependence has been suppressed. Alternatively, they can be defined by the Lippmann—Schwinger (L—S) equations

$$T_c(E) = V_c + V_c G_0(E) T_c(E), \tag{45a}$$

$$T_{c,\,l}(E) = V_{c,\,l} + V_{c,\,l} G_{0,\,l}(E) T_{c,\,l}(E). \tag{45b}$$

It is essential that E is not in the spectrum of the Hamilton operators H_0 and $H_0 + V_c$, for in that case that resolvents are not defined. For E not real positive, the kernels of the L—S equations are of Hilbert—Schmidt class if the potential satisfies certain conditions.[13] This

guarantees that the L—S equations have unique solutions for these potentials. Even for the Coulomb potential, the kernel $V_{c,\,l} G_{0,\,l}(E)$ is of Hilbert—Schmidt class (provided E is not positive), and therefore the solution $T_{c,\,l}(E)$ of equation (45b) is unique.

Writing the operators in momentum representation gives the matrix elements $\langle \mathbf{p}' | T_c(E) | \mathbf{p} \rangle$ and $\langle p' | T_{c,\,l}(E) | p \rangle$. They can be considered as functions of the momenta and of the energy. If the potential has short range (in a certain well defined sense), it is known that these functions are meromorphic in the complex E-plane, cut along the positive real axis. Besides bound-state poles (cf. Ref. 32) other poles can occur. The cut $0 < E < \infty$ is called the unitarity cut. The limits at the upper and lower rim of the unitarity cut exist if the potential has short range. As we have seen, the limits $k \to p$ and $k \to p'$ [cf., Eq. (5)] of the Coulomb T matrices do not exist. These additional singularities, characteristic for the Coulomb potential (and probably for other long-range potentials), are of the type $(p-k)^{i\gamma}$, $(p'-k)^{i\gamma}$. These singularities are clearly integrable. So if we consider $T_c | f \rangle$, where $| f \rangle$ is a Hilbert space vector,

$$\langle \mathbf{p}' | T_c(k^2) | f \rangle = \int_{R^3} d\mathbf{p}\ \langle \mathbf{p}' | T_c(k^2) | \mathbf{p} \rangle \langle \mathbf{p} | f \rangle, \tag{46}$$

we get no trouble at all in this integral because $\langle \mathbf{p} | f \rangle$, being square integrable, is locally integrable.[36] The value of $\langle \mathbf{p}' | T_c(k^2) | \mathbf{p} \rangle$ at $p = k$ is here unimportant.

Strictly speaking, the half- and on-shell points make sense only for the matrix elements, not for the operators. So we may say equally well that the half- and on-shell values of T_c and of $T_{c,\,l}$, considered as operators, are irrelevant.

D. Summary and discussion

We can summarize the discussion of Secs. 4A—4C as follows. When T_c and $T_{c,\,l}$ are considered as

$$\left\{ \begin{array}{l} \text{functions} \\ \text{distributions} \\ \text{operators} \end{array} \right\} \text{ their half- and on-shell values are } \left\{ \begin{array}{l} \text{not defined} \\ \text{exactly zero} \\ \text{irrelevant} \end{array} \right\}. \tag{47}$$

In this connection it is useful to remark that in the literature considerable confusion has arisen about the half- and on-shell Coulomb T matrices. West[37] finds exactly zero for the on-shell T_c. He adds a physical interpretation: $\langle \mathbf{p}' | T_c | \mathbf{p} \rangle$ may be interpreted as the probability to scatter a free state denoted by $| \mathbf{p} \rangle$ into the free state $\langle \mathbf{p}' |$. This probability should indeed be zero because in practice one never has a free state when the pure Coulomb force acts. This is due to its long range. Nutt[22] obtained zero, too, but his derivation is wrong as was shown, e.g., by Nuttall and Stagat.[38] Shastry and Rajagopal,[33] following Nutt, also found zero for the on-shell partial wave projected $T_{c,\,l}$. Most authors obtain the branch-point singularities (of T_c only; to our knowledge $T_{c,\,l}$ is not known in explicit form). The result (47) applies as well when an arbitrary short-range *separable* potential is added to the Coulomb po-

tential. The proof of this statement follows easily from Eqs. (48)—(58) given in the next section. We conjecture that (47) will hold also for the Coulomb potential plus an arbitrary (suitably defined) short-range potential. This can be expected because in the Fourier transform a neighborhood of the point $p = k$ is related to a neighborhood of the point $r = \infty$ (in coordinate representation). Furthermore, only the tail of the potential is important for the behavior of all relevant functions at large distances. Explicit examples of such T matrices corresponding to Coulomb-like potentials are given in the following sections.

We may understand (47) as the formal solution of the T matrix problems. However, concerning the *physical* interpretation more has to be said. This will be done in a forthcoming paper.[25] Compare also the work of Taylor[39,40] and of Marquez.[41]

H. van Haeringen and R. van Wageningen   1445

5. THE TOTAL T OPERATOR

Let the potential V be the sum of two potentials V_c and V_s which are not specified for the time being,

$$V = V_c + V_s, \tag{48}$$

In the Gell-Mann—Goldberger two-potential formalism[42] it follows that the corresponding total T operator can be written[6,7]

$$T = T_c + T_{cs}, \tag{49}$$

where T_c is the T operator corresponding to V_c [cf., Eqs. (44), (45)] and T_{cs} is given by

$$T_{cs}(E) = [1 + T_c(E)G_0(E)] \, t_{cs}(E)[1 + G_0(E)T_c(E)], \tag{50}$$

where the operator $t_{cs}(E)$ satisfies the equation

$$t_{cs}(E) = V_s + V_s G_c(E) t_{cs}(E), \tag{51}$$

The resolvents G_0 and G_c have been introduced already in Eqs. (42), (43). The orbital angular momentum projections of the operators occurring in Eqs. (48)—(51) satisfy these same equations.

Now we specify the operators V_c and V_s. Let V_c be the Coulomb potential [given by Eq. (11) or (13)] and let V_s be an operator of finite rank N (also called an N-term separable potential), which works only in a subspace corresponding to one particular value of the orbital angular momentum l,

$$V_s = -\sum_{i=1}^{N} |g_i\rangle \lambda_i \langle g_i| . \tag{52}$$

In Eq. (52) the projection operator projecting onto the lth partial wave space and the corresponding subscripts l are suppressed. The same is done in the rest of this section. The λ_i are real numbers in order that V_s be Hermitian. Time reversal invariance requires that $\langle p|g_i\rangle$ and $\langle r|g_i\rangle$ are real functions of p and r, respectively. It is well known that the T operator T_s corresponding to the separable potential V_s is separable itself and may be written

$$T_s(E) = -\sum_{i,j=1}^{N} |g_i\rangle \tau_{ij}(E) \langle g_j| . \tag{53}$$

Here the $N \times N$ matrix τ is defined via its inverse:

$$(\tau^{-1}(E))_{ij} \equiv (\Lambda^{-1})_{ij} + \langle g_i | G_0(E) | g_j \rangle, \tag{54}$$

where Λ is a diagonal matrix with elements $\lambda_i \delta_{ij}$. Now $T_s(E)$ is the unique solution of the equation

$$T_s(E) = V_s + V_s G_0(E) T_s(E). \tag{55}$$

Also Eq. (51) has a unique solution. The uniqueness follows almost trivially from the fact that V_s is of finite rank. Because of the resemblance of Eq. (51) to Eq. (55), we can write down the solution of Eq. (51) at once: All we have to do is replace G_0 by G_c in Eqs. (53), (54). Thus in terms of the Coulomb-modified (energy-dependent) form factors $|g_i^c(E)\rangle$ defined by

$$|g_i^c(E)\rangle \equiv (1 + T_c(E)G_0(E)) |g_i\rangle, \tag{56}$$

we obtain our final result for T_{cs} [cf., Eq. (50)],

$$T_{cs}(E) = -\sum_{i,j=1}^{N} |g_i^c(E)\rangle \tau_{ij}^c(E)\langle g_j^c(E^*)| , \tag{57}$$

with

$$(\tau^c(E))_{ij}^{-1} \equiv (\Lambda^{-1})_{ij} + \langle g_i | G_c(E) | g_j \rangle . \tag{58}$$

Equation (44) and definition (56) imply the useful identity

$$G_0(E) |g_i^c(E)\rangle \equiv G_c(E) |g_i\rangle. \tag{59}$$

It is clear from Eqs. (49), (56)—(58) that the total T matrix has been obtained in closed form, once we have derived explicit expressions for $\langle p|g_i^c\rangle$ and for $\langle g_i | G_c | g_j\rangle$. This is what we are going to do in the next sections, for some simple form factors $|g_i\rangle$ in $l = 0$ partial wave space.

6. THE COULOMB MODIFIED FORM FACTORS $\langle p|g^c\rangle$

From now on we consider only the case $l = 0$ and we suppress the subscript l throughout. The form factors $\langle p|g\rangle$ are supposed to be rational functions of p^2. The simplest rational function that is physically acceptable as form factor is the Yamaguchi[23] one

$$\langle p | g_\beta \rangle \equiv \sqrt{2/\pi} \; \frac{1}{p^2 + \beta^2} . \tag{60}$$

Differentiation with respect to the parameter β yields (apart from a trivial factor)

$$\langle p | g_{\beta\beta} \rangle \equiv \sqrt{2/\pi} \; \frac{1}{(p^2 + \beta^2)^2} . \tag{61}$$

In coordinate representation one has

$$\langle r | g_\beta \rangle = e - \beta r/r, \tag{62}$$

$$\langle r | g_{\beta\beta} \rangle = e - \beta r/(2\beta). \tag{63}$$

We obtain the Coulomb-modified form factor

$$\langle p | g_\beta^c(k^2) \rangle = \langle p | (1 + T_c(k^2)G_0(k^2)) | g_\beta \rangle \tag{64}$$

in closed form with the help of Eqs. (26), (27), and (15). The result is a double integral of the type

$$\int_0^\infty dp' \int_0^1 dt \cdots \frac{1}{p'}$$

$$\times \left(\frac{1}{1 - ta'a} - \frac{1}{1 - ta'/a} + \frac{1}{1 - t/(a'a)} - \frac{1}{1 - ta/a'} \right), \tag{65}$$

where as before $a = (p - k)/(p + k)$, $a' = (p' - k)/(p' + k)$, [Eqs. (28), (29)]. The four terms of the integrand in Eq. (65) are taken two by two; the integrand then appears to be a function of $(p')^2$. Interchanging the order of the integrations (this is permitted according to Sec. 4B), one easily performs the integration $\int dp'$. The remaining integral $\int dt$ is again an h.f.. The actual derivation is given at the end of this section. In advance we give here the final result

$$\langle p | g_\beta^c(k^2) \rangle = \frac{\sqrt{2/\pi}}{\beta^2 + p^2} - \frac{\sqrt{2/\pi}}{\beta^2 + k^2} \frac{k}{p} [F_{i\gamma}(Ba) - F_{i\gamma}(B/a)], \tag{66}$$

where

$$B \equiv (\beta + ik)/(\beta - ik). \tag{67}$$

In addition, we introduced here the notation

$$F_{i\gamma}(z) \equiv {}_2F_1(1, \; i\gamma; \; 1 + i\gamma; \; z) \tag{68}$$

which we shall use from now on in this paper. Equation (66) can be converted into [e.g. with the help of Eq. (85)]

$$\langle p \,|\, g^c_\beta(k^2)\rangle = \frac{1}{2p(\beta - ik)} \frac{\sqrt{2/\pi}}{1 + i\gamma}$$

$$\times \left(\frac{p-k}{\beta - ip} \,_2F_1(1, \ i\gamma; \ 2 + i\gamma; \ Ba)\right.$$

$$\left. + \frac{p+k}{\beta + ip} \,_2F_1(1, \ i\gamma; \ 2 + i\gamma; \ B/a)\right).$$

(69)

In Eq. (69) a result of Zachary[5] and of Bajzer[6] is reproduced. [Note the misprint in Eq. (49) of Ref. 6; the argument $(Ba)^{-1}$ of the h.f. mentioned there should be B/a (in our notation)]. The h.f.'s occurring in Eq. (69) differ from the h.f. in (68) with regard to their third parameter. More generally, we meet in this game several different h.f.'s which all have one common property, i.e., they are of the type

$$_2F_1(\lambda, \ \mu - 1 + i\gamma; \ \nu + \mu - 1 + i\gamma; \ z), \quad \lambda, \ \mu, \ \nu = 1, \ 2, \ \cdots.$$

(70)

All these functions can be expressed in any one representative of the class (70) plus rational functions of z. This can be proven with the help of their integral representations.[28] We chose already $_2F_1(1, \ i\gamma; \ 1 + i\gamma; \ z)$ as a simple and convenient representative [see (68)]. This has the additional advantage that it facilitates the comparison of different expressions for the same object, such as in Eqs. (66) and (69). Therefore, $F_{i\gamma}$ will be our standard hypergeometric function throughout this paper.

The derivative of the standard h.f. $F_{i\gamma}$ can also be written in terms of $F_{i\gamma}$ and rational functions. The following equality is easily established:

$$\frac{d}{dz} F_{i\gamma}(z) = \frac{i\gamma}{z}\left(\frac{1}{1-z} - F_{i\gamma}(z)\right).$$

(71)

Utilizing (71) we obtain from (66) by differentiation with respect to β [cf. (61)]

$$\langle p \,|\, g^c_{\beta\beta}(k^2)\rangle = \sqrt{2/\pi}\left(\frac{1}{(\beta^2 + p^2)^2} + \frac{k\gamma/\beta}{(\beta^2 + p^2)(\beta^2 + k^2)}\right.$$

$$\left. - \frac{1 + k\gamma/\beta}{(\beta^2 + k^2)^2} \frac{k}{p}[F_{i\gamma}(Ba) - F_{i\gamma}(B/a)]\right).$$

(72)

The rest of this section is devoted to the derivation of Eq. (66). The reader who is not interested in the details of the proof can immediately go on to Sec. 7. We calculate the integral in the rhs of the following equation:

$$\langle p \,|\, T_c(k^2)G_0(k^2)\,|\, g_\beta\rangle = -\sqrt{2/\pi} \int_0^\infty dp' p'^2 \frac{\langle p \,|\, T_c(k^2)\,|\, p'\rangle}{(p'^2 + \beta^2)(p'^2 - k^2)}.$$

(73)

Substitute for $\langle p \,|\, T_c \,|\, p'\rangle$ Eqs. (26), (27), utilize the integral representation of the h.f. Eq. (15), and interchange the order of the two integrations, then one obtains

$$\langle p \,|\, T_c(k^2)G_0(k^2)\,|\, g_\beta\rangle = \sqrt{2/\pi} \ \frac{k\gamma}{\pi p} \int_0^1 dt \, t^{i\gamma - 1}$$

$$\times \int_0^\infty \frac{dp' p'}{(p'^2 + \beta^2)(p'^2 - k^2)}$$

$$\times \left(\frac{1}{1 - ta(p'-k)/(p'+k)} - \frac{1}{1 - ta(p'+k)/(p'-k)}\right)$$

$$+ (\text{idem, with } p \to -p \text{ thus } a \to 1/a).$$

(74)

The integral $\int dp'$ in Eq. (74) can be simplified utilizing the variable x defined by

$$x \equiv k(1 + ta)/(1 - ta),$$

(75)

where $\text{Im}\, x > 0$ holds because $\text{Im}\, k > 0$. This integral then appears to be an elementary one; we get

$$\int_0^\infty dp' \cdots \text{ in } (74) = \frac{-4kta}{(1 - ta)^2} \int_0^\infty \frac{dp' p'^2}{(p'^2 + \beta^2)(p'^2 - k^2)(p'^2 - x^2)}$$

$$= \frac{-4kta}{(1 - ta)^2} \frac{i\pi/2}{(\beta - ik)(\beta - ix)(k + x)}.$$

(76)

From (75) we have

$$\beta - ix = (\beta - ik)(1 - tBa)/(1 - ta),$$

$$k + x = 2k/(1 - ta),$$

(77)

so that the factors $(1 - ta)$ occurring in the integral $\int dt \cdots$ drop out, and this integral appears to be again an h.f. $_2F_1$ according to

$$\int_0^1 dt \frac{t^{i\gamma}}{1 - tBa} = \frac{1}{1 + i\gamma} \,_2F_1(1, \ 1 + i\gamma; \ 2 + i\gamma; \ Ba).$$

(78)

This h.f. is reduced to the standard $F_{i\gamma}$ by means of

$$F_{i\gamma}(z) \equiv \,_2F_1(1, i\gamma; \ 1 + i\gamma; \ z) = 1 + \frac{i\gamma z}{1 + i\gamma}$$

$$\times \,_2F_1(1, \ 1 + i\gamma; \ 2 + i\gamma, \ z).$$

(79)

Performing all this in Eq. (74), we get the final result

$$\langle p \,|\, T_c(k^2)G_0(k^2)\,|\, g_\beta\rangle = \frac{-k\sqrt{2/\pi}}{p(\beta - ik)^2} \cdot \frac{i\gamma}{1 + i\gamma}$$

(80)

$$\times [a \,_2F_1(1, 1 + i\gamma; 2 + i\gamma; \ Ba) - 1/a \,_2F_1(1, 1 + i\gamma; 2 + i\gamma; \ B/a)]$$

$$= -\frac{\sqrt{2/\pi}}{\beta^2 + k^2} \frac{k}{p}[F_{i\gamma}(Ba) - F_{i\gamma}(B/a)].$$

(81)

Equation (81) proves Eq. (66).

7. THE IN-PRODUCTS $\langle g|G_c|g\rangle$

In this section we obtain analytic expressions for the in-products $\langle g_i \,|\, G_c \,|\, g_j\rangle$ occurring in the T_{cs}-formula, cf. Eqs. (57), (58). The operators V_s and T_{cs} are fixed once we have fixed the N form factors $|g_i\rangle$, $i = 1, \ldots, N$ (plus λ_i and the energy, of course). We suppose that these form factors are of the type (60) and/or (61); their parameters β are possibly different. Because "cross products" $\langle g_i \,|\, G_c \,|\, g_j\rangle$ with $i \neq j$ occur, we have to calculate three different expressions, namely, $\langle g_\alpha \,|\, G_c \,|\, g_\beta\rangle$, $\langle g_{\alpha\alpha} \,|\, G_c \,|\, g_\beta\rangle$, and $\langle g_{\alpha\alpha} \,|\, G_c \,|\, g_{\beta\beta}\rangle$ where α and β are independent parameters. From these three, all possible combinations can be made. Obviously this is sufficient to obtain T_{cs} in closed form [in combination with Eqs. (66) and (72)].

The derivation of $\langle g_\alpha \,|\, G_c \,|\, g_\beta\rangle$ is given at the end of this section. It has to be noted that one could start with Eq. (66), because from Eq. (59) one has the identity

$$\langle g_i \,|\, G_c(k^2) \,|\, g_j\rangle \equiv \langle g_i \,|\, G_0(k^2) \,|\, g^c_j(k^2)\rangle.$$

(82)

We obtain

$$\langle g_\alpha | G_c(k^2) | g_\beta \rangle = \frac{-_2F_1(1, i\gamma; 2+i\gamma; AB)}{(1+i\gamma)(\alpha+\beta)(\alpha-ik)(\beta-ik)},$$ (83)

with

$$A \equiv (\alpha + ik)/(\alpha - ik), \quad B \equiv (\beta+ik)/(\beta-ik).$$ (84)

Taking $\alpha = \beta$ in Eq. (83) we reproduce the result of Zachary.[5] Utilizing the relation

$$\frac{1}{1+i\gamma}\ _2F_1(1,\ i\gamma;\ 2+i\gamma;\ z) = 1/z + (1-1/z)F_{i\gamma}(z),$$ (85)

Eq. (83) can be converted into an equation where $F_{i\gamma}$ is the only nonelementary function,

$$\langle g_\alpha | G_c(k^2) | g_\beta \rangle = \frac{-1}{(\alpha+\beta)(\alpha+ik)(\beta+ik)}$$

$$- \frac{2ik}{(\alpha^2+k^2)(\beta^2+k^2)}\ F_{i\gamma}(AB).$$ (86)

Differentiation of Eq. (86) with respect to α, β and utilizing Eq. (71) yields the final results

$$\langle g_{\alpha\alpha} | G_c(k^2) | g_\beta \rangle = \frac{1}{2\alpha(\alpha+\beta)^2(\alpha^2+k^2)}$$

$$- \frac{\alpha+k\gamma}{\alpha(\alpha+\beta)(\alpha+ik)(\beta+ik)(\alpha^2+k^2)} - \frac{2ik(\alpha+k\gamma)}{\alpha(\alpha^2+k^2)^2(\beta^2+k^2)}\ F_{i\gamma}(AB)$$ (87)

and

$$\langle g_{\alpha\alpha} | G_c(k^2) | g_{\beta\beta} \rangle = \frac{\alpha^2 + \alpha\beta + \beta^2 + k^2 + k\gamma(\alpha+\beta)}{2\alpha\beta(\alpha+\beta)^3(\alpha^2+k^2)(\beta^2+k^2)}$$

$$- \frac{(\alpha+k\gamma)(\beta+k\gamma)}{\alpha\beta(\alpha+\beta)(\alpha+ik)(\beta+ik)(\alpha^2+k^2)(\beta^2+k^2)}$$

$$- \frac{2ik(\alpha+k\gamma)(\beta+k\gamma)}{\alpha\beta(\alpha^2+k^2)^2(\beta^2+k^2)^2}\ F_{i\gamma}(AB).$$ (88)

Now we give the proof of Eq. (83), or alternatively of Eq. (86). According to Eq. (81), we may write

$$\langle g_\alpha | G_0(k^2)T_c(k^2)G_0(k^2) | g_\beta \rangle = \int_0^\infty dp\ \frac{p^2}{(p^2+\alpha^2)(p^2-k^2)}$$

$$\times \frac{2k}{\pi p}\ \frac{i\gamma}{\beta^2+k^2} \int_0^1 dt\ t^{i\gamma-1}\left(\frac{1}{1-tBa} - \frac{1}{1-tB/a}\right).$$ (89)

Define the variable y,

$$y \equiv k\ \frac{1+tB}{1-tB},$$ (90)

interchange the order of the two integrations (cf. Sec. 4B) and recall from Eq. (28) that

$$a - 1/a = -4pk/(p^2 - k^2),$$ (91)

$$a + 1/a = 2(p^2 + k^2)/(p^2 - k^2),$$ (92)

then one arrives at the following formula which shows a strong resemblance to Eq. (76):

$$\langle g_\alpha | G_0(k^2)T_c(k^2)G_0(k^2) | g_\beta \rangle = \frac{-8k^2B}{\pi(\beta^2+k^2)}$$

$$\times i\gamma \int_0^1 \frac{dt\ t^{i\gamma}}{(1-tB)^2} \int_0^\infty \frac{dp p^2}{(p^2+\alpha^2)(p^2-k^2)(p^2-y^2)}.$$ (93)

So this derivation runs along the same lines as the derivation of Eq. (66) in Sec. 6. The integration $\int dp$ gives in analogy to Eq. (76) a factor $(\alpha - iy)$ that can be written

$$\alpha - iy = (\alpha - ik)(1-tAB)/(1-tB).$$ (94)

The factors $(1-tB)$ cancel and one obtains

$$\langle g_\alpha | G_0(k^2)T_c(k^2)G_0(k^2) | g_\beta \rangle = \frac{2k\gamma}{(\alpha-ik)^2(\beta-ik)^2} \int_0^1 dt\ \frac{t^{i\gamma}}{1-tAB},$$ (95)

With the help of Eqs. (78), (79), and (82), Eqs. (83) and (86) follow easily from Eq. (95).

8. T MATRIX FOR COULOMB PLUS ARBITRARY $l = 0$ RATIONAL SEPARABLE POTENTIAL

We call a potential of the form

$$V_s = \sum_{i=1}^N | g_i \rangle \lambda_i \langle g_i |,$$

where the form factors $| g_i \rangle$ in momentum representation, $\langle p | g_i \rangle$,

1. are real rational functions of p^2,

2. are zero at infinity, (96)

3. and have no poles on the positive axis $0 \le p^2 < \infty$,

a *rational separable potential*. Any such form factor can be written (for some $\nu = 1, 2, \cdots$)

$$\langle p | g_{\{\beta\}} \rangle \equiv \frac{P(p^2)}{(p^2+\beta_1^2)(p^2+\beta_2^2)\cdots(p^2+\beta_\nu^2)},$$ (97)

where P is some (real) polynomial of degree smaller than ν. Its realness follows from the fact that the form factor is real. The denominator of the rhs of definition (97) is real, too, but the β's may be complex. They can and will always be chosen such that they satisfy

$$\mathrm{Re}\,\beta_n > 0, \quad n = 1, \ldots, \nu.$$ (98)

These parameters β_n need not all be different. The notation in definition (97) is such that the symbol $\{\beta\}$ stands for $(\beta_1, \ldots, \beta_\nu)$. In (61) we had the particular case $\nu = 2$, we took $\beta_1 = \beta_2 = \beta$ and the polynomial P was taken constant there. Our potential is built up from N form factors $| g_i \rangle$, $i = 1, \ldots, N$. Each form factor has the form (97), but possibly with a different set of β's, e.g.,

$$\beta_{i,1}, \beta_{i,2}, \cdots, \beta_{i,\nu_i} \text{ for each } i = 1, \ldots, N.$$

This more complete notation is cumbersome and would not clarify our formulas so we will not use it. We suppress the index i when the deal with the general form factors (97).

Recalling the discussion following (65), it becomes clear that we need in effect integrals of the type

$$\int_0^\infty dp P(p^2) \prod_{n=1}^s (p^2 + \beta_n^2)^{-1}.$$ (99)

Quite a lot of such integrals are known in closed form.[43,44] Moreover, the general structure can easily be found. In this way one can derive the desired formulas for the $\langle p | g^c \rangle$'s and for the $\langle g | G_c | g \rangle$'s. This will be done at

the end of this section. We give the final results here in advance:

$$\langle p | g^c_{\{B\}}(k^2) \rangle = R_0(p^2) + \sum_{n=1}^{v} R_n(p^2) \frac{1}{p}[F_{i\gamma}(B_n a) - F_{i\gamma}(B_n/a)]$$

(100)

and

$$\langle g_{\{\alpha\}} | G_c(k^2) | g_{\{B\}} \rangle = R_{00} + \sum_{m=1}^{\mu} \sum_{n=1}^{v} R_{mn} F_{i\gamma}(A_m B_n),$$

(101)

where as before $a = (p - k)/(p + k)$, and in analogy to Eq. (84) [cf. Eq. (120)]

$$A_m \equiv (\alpha_m + ik)/(\alpha_m - ik), \quad m = 1, 2, \ldots, \mu,$$

(102)

$$B_n \equiv (\beta_n + ik)/(\beta_n - ik), \quad n = 1, 2, \ldots, v.$$

(103)

R_0 and R_n are rational functions of k, γ, β_1, $\ldots$, β_v, p^2. R_{00} and R_{mn} are rational functions of k, γ, β_1, $\ldots$, β_v, α_1, $\ldots$, α_μ.

The T matrix corresponding to a rational separable potential plus the Coulomb potential is given by Eqs. (49), (57), (58) and (100), (101). The important point is that, apart from rational functions, only the h.f. $F_{i\gamma}$ shows up in the final expressions for T. The arguments occurring in $F_{i\gamma}$ can easily be found. They are exclusively defined by the *poles* of all form factors involved (and by k and p of course), as can be seen from the Eqs. (26), (27), (97), (100)–(103), and (120).

The remaining part of this section is entirely devoted to the derivation of Eqs. (100) and (101). The proof of these equations is in fact a generalization of the proofs given in Secs. 6 and 7, respectively.

We start with the integral

$$^{(s)}I \equiv \int_0^\infty dp\, p^2 \prod_{n=1}^{s} (p^2 + \beta_n^2)^{-1}, \quad s = 2, 3, 4, \cdots,$$

(104)

where

$$\text{Re}\,\beta_n > 0 \quad \text{for } n = 1, 2, \ldots, s.$$

(105)

The following result holds[44]

$$^{(s)}I = \frac{\pi}{2} P_s(\beta_1, \ldots, \beta_s) \prod_{m<n=1}^{s} (\beta_n + \beta_m)^{-1}.$$

(106)

Here P_s is a symmetric homogeneous polynomial in the variables $(\beta_1 \ldots, \beta_s)$. Its degree is $\frac{1}{2}(s-2)(s-3)$, but the highest power of any one of the β's is $s-3$ if $s>2$, and for $s = 2$ this power is zero. Explicitly we have for the first three polynomials

$$P_2 = P_3 = 1,$$

(107)

$$P_4 = \beta_1 + \beta_2 + \beta_3 + \beta_4.$$

(108)

With the generalized Yamaguchi form factor

$$\langle p | g_{\{B\}} \rangle = \prod_{n=1}^{v} (p^2 + \beta_n^2)^{-1},$$

(109)

which is a special case of (97), we get in the same way as in Sec. 6

$$\langle p | T_c(k^2) G_0(k^2) | g_{\{B\}} \rangle = -\frac{4k^2\gamma}{\pi p} \int_0^1 \frac{dt\, t^{i\gamma} a}{(1-ta)^2}$$

$$\times \int_0^\infty \frac{dp'\, p'^2}{(p'^2 - k^2)(p'^2 - x^2)(p'^2 + \beta_1^2) \cdots (p'^2 + \beta_v^2)}$$

$$+ (\text{idem, with } p \to -p \text{ thus } a \to 1/a).$$

(110)

As before the variable x is given by $x/k = (1+ta)/(1-ta)$, where $\text{Im}\,x > 0$ holds because of $\text{Im}\,k > 0$. Application of Eq. (106) gives for the integral $\int dp'$ in the rhs of Eq. (110) the expression

$$\frac{i\pi}{4k}$$

$$\times \frac{(1 - ta)P_{v+2}(-ix, -ik, \beta_1, \ldots, \beta_v)}{(\beta_1 - ix) \cdots (\beta_v - ix)(\beta_1 - ik) \cdots (\beta_v - ik)(\beta_1 + \beta_2) \cdots (\beta_{v-1} + \beta_v)}$$

(111)

where the polynomial P_{v+2} contains at most terms with x^{v-1}. It contains no terms with higher powers of x. In analogy to Eq. (77) we have

$$\beta_n - ix = (\beta_n - ik)(1 - taB_n)/(1-ta).$$

(112)

Multiply numerator and denominator of expression (111) by the factor $(1-ta)^v$. The polynomial P_{v+2} then becomes a polynomial in the variables $(ta, k, \beta_1, \ldots, \beta_v)$, and the numerator contains further a factor $(1-ta)^2$. The denominator becomes a product of the factors $(1-taB_n)$ and of the factors $(\beta_n - ik)^2(\beta_n + \beta_m)$ with $m < n = 1, 2, \ldots, v$.

In the rhs of Eq. (110) we see a factor $(1-ta)^2$ in the denominator, so this factor cancels. Equation (110) then simplifies considerably; we get

$$\langle p | T_c(k^2) G_0(k^2) | g_{\{B\}} \rangle$$

$$= \frac{1}{(\beta_1 - ik)^2 \cdots (\beta_v - ik)^2 (\beta_1 + \beta_2) \cdots (\beta_{v-1} + \beta_v)}$$

$$\times \frac{k\gamma}{ip} \int_0^1 dt\, \frac{at^{i\gamma} P''_{v+2}(ta, k, \beta_1, \ldots, \beta_v)}{(1 - taB_1) \cdots (1 - taB_v)}$$

$$+ (\text{idem, with } p \to -p \text{ thus } a \to 1/a).$$

(113)

For $v = 1$ the polynomial P''_{v+2} is just 1, and Eq. (113) reduces to Eq. (80), apart from a factor $\sqrt{2/\pi}$.

The denominator of the integrand in Eq. (113) can be written in partial fractions. Then we get a sum of v terms as follows:

$$\prod_{n=1}^{v} (1 - taB_n)^{-1} = \sum_{n=1}^{v} P'_n/(1 - taB_n),$$

(114)

where the P'_n are certain polynomials of $(ta, B_1, \ldots, B_v)$. Now we know[28]

$$\int_0^1 dt\, t^{i\gamma-1} \frac{t^\lambda}{1-tz} = \frac{1}{\lambda + i\gamma}\, _2F_1(1, \lambda + i\gamma; \lambda + 1 + i\gamma; z),$$

$$\lambda = 0, 1, \cdots.$$

(115)

Furthermore, iteration of Eq. (79) yields

$$F_{i\gamma}(z) = 1 + \frac{i\gamma z}{1 + i\gamma} + \cdots + \frac{i\gamma z^{\lambda-1}}{\lambda - 1 + i\gamma} + \frac{i\gamma z^\lambda}{\lambda + i\gamma} F_{\lambda+i\gamma}(z);$$

(116)

recall from Eq. (68) the notation

$$F_{\lambda+i\gamma}(z) \equiv\, _2F_1(1, \lambda + i\gamma; \lambda + 1 + i\gamma; z).$$

(68′)

This implies that the integral

$$\int_0^1 dt\, t^{i\gamma-1}(ta)^\lambda/(1-taB_n) \tag{117}$$

can be written in terms of $F_{i\gamma}(aB_n)$ plus rational functions, for $\lambda = 0,\ 1,\ \cdots$. The same holds then if the factor $(ta)^\lambda$ in the integrand of (117) is replaced by an arbitrary polynomial of ta, like we have from Eqs. (113), (114). Applying all this to Eq. (113) we get the result

$$\langle p\,|\,T_c(k^2)G_0(k^2)\,|\,g_{(\beta)}\rangle = r_0 + \sum_{n=1}^{\upsilon} r_n \frac{1}{p}$$
$$\times \left[F_{i\gamma}(B_n a) - F_{i\gamma}(B_n/a)\right], \tag{118}$$

where r_0 and the r_n are rational functions of $(p^2,\ k,\ \gamma,\ \beta_1, \ldots, \beta_\upsilon)$. They are functions of p^2 (and not only of p) because $\langle p\,|\,T_c G_0\,|\,g\rangle$ is *even* in p.

Now that we have evaluated this expression where the form factor is given by Eq. (109), it is only one step further to prove Eq. (118) for a form factor that is an arbitrary rational function of p^2 [see definition (97)]. Write any term $p^{2\lambda}$ occurring in the polynomial in (97) as

$$p^{2\lambda} = p^{2\lambda-2}(p^2 + \beta_1^2 - \beta_1^2). \tag{119}$$

The term containing the factor $(p^2 + \beta_1^2)$ is then reduced by dividing that factor out. Iteration of this procedure reduces the exponent to zero, so any rational form factor can be reduced to a sum of form factors of the type given in Eq. (109). Therefore, Eq. (118) holds for an arbitrary rational form factor, and Eq. (100) has been proven.

The proof of Eq. (101) which we shall give now is not essentially different in structure from the one just given for Eq. (100). With the form factor

$$\langle p\,|\,g_{(\alpha)}\rangle = \prod_{m=1}^{\mu} (p^2 + \alpha_m^2)^{-1}, \tag{120}$$

we have [cf. Eqs. (118) and (15)]

$$\langle g_{(\alpha)}\,|\,G_0(k^2)T_c(k^2)G_0(k^2)\,|\,g_{(\beta)}\rangle = r_{00} - \sum_{n=1}^{\upsilon}\int_0^1 dt \int_0^\infty dp$$
$$\times \frac{p^2 r_n(p^2)}{(p^2 - k^2)(p^2 + \alpha_1^2)\cdots(p^2 + \alpha_\mu^2)}\frac{i\gamma t^{i\gamma-1}}{p}$$
$$\times \left(\frac{1}{1-tB_n a} - \frac{1}{1-tB_n/a}\right) \tag{121}$$

Here r_{00} is a rational function of $(k,\ \gamma,\ \alpha_m,\ \beta_n)$ with $m = 1, \ldots, \mu$; $n = 1, \ldots, \upsilon$. As in (90) define

$$y_n \equiv k(1 + tB_n)/(1 - tB_n), \tag{122}$$

in terms of which Eq. (121) reads

$$\langle g_{(\alpha)}\,|\,G_0(k^2)T_c(k^2)G_0(k^2)\,|\,g_{(\beta)}\rangle$$
$$= r_{00} + 4ik\gamma \sum_{n=1}^{\upsilon}\int_0^1 dt\, t^{i\gamma-1}\frac{tB_n}{(1-tB_n)^2}$$
$$\times \int_0^\infty \frac{dp\, p^2 r_n(p^2)}{(p^2 - k^2)(p^2 - y_n^2)(p^2 + \alpha_1^2)\cdots(p^2 + \alpha_\mu^2)}. \tag{123}$$

Utilizing

$$\alpha_m - iy_n = (\alpha_m - ik)(1 - tA_m B_n)/(1 - tB_n), \tag{124}$$

and recalling the discussion following Eq. (110), one

finds that all factors $(1 - tB_n)$ cancel. The denominator of the integrand in the integral $\int dt$ then contains only factors $(1 - tA_m B_n)$ and other factors that are constant with respect to the variable of integration t. The product of these factors can be written as a sum by means of partial fractions just as in Eq. (114),

$$\prod_{m=1}^{\mu} \frac{1}{1 - tA_m B_n} = \sum_{m=1}^{\mu} \frac{P'_m}{1 - tA_m B_n}, \tag{125}$$

where the P'_m are polynomials in $(tB_n,\ A_1, \ldots, A_\mu)$. Thus we get again integrals of the type (117). All these integrals are then transformed into rational functions of $(k,\ \gamma,\ \alpha_1, \ldots, \alpha_\mu,\ \beta_1, \ldots, \beta_\upsilon)$ and the h.f.'s $F_{i\gamma}(A_m B_n)$. Then the same result is valid if one takes arbitrary rational form factors $\langle g_{(\alpha)}\,|$ instead of the form factors of Eq. (120). Finally, for G_c the same holds as for $G_0 T_c G_0$ according to Eq. (10). This completes the proof of Eq. (101).

9. SUMMARY AND DISCUSSION

We discussed and resolved the half- and on-shell problems which occur in the T matrices (in complete three-dimensional space and in the partial wave spaces $l = 0,\ 1, \cdots$, respectively) of any Coulomb-like potential $V = V_c + V_s$ where V_s is an arbitrary short-range separable potential (Sec. 4). In Secs. 5–7 we obtained exact explicit formulas for off-shell T matrices corresponding to several Coulomb-plus-rational-separable potentials (for $l = 0$ only). To this end, we first derived In Sec. 3 the pure Coulomb $T_{c,l=0}$ from a known expression for the complete T_c in three-dimensional space. Equations (26), (27), and (30) give two equivalent explicit formulas for $T_{c,l=0}$. Finally, in Sec. 8 we derived a general formula for the off-shell T matrix corresponding to any member of the very large class of potentials consisting of the Coulomb potential plus an arbitrary rational separable potential, as defined in (96). Apart from rational functions, the final expressions [Eqs. (100) and (101)] contain only the h.f. $_2F_1(1,\ i\gamma;\ 1 + i\gamma;\cdot)$ where γ is the well-known Coulomb parameter.[12]

As for the physical relevance of the obtained formulas, we like to discuss some important points.

In scattering experiments with charged particles one often deals with the combination of the Coulomb interaction plus some short-range interaction. For example, nonrelativistic models for proton—proton scattering have a short-range potential built in to account for the strong interaction, the Coulomb potential taking care of the charges. The strong interaction here is almost completely unknown at very short distance (< 1 fm), so one looks for a phenomenological fit. The model potential may then be (partly) local or nonlocal (e.g. separable); often one takes the sum of a local potential (for the longer distances) and a nonlocal separable potential (for the short distances). Moreover, as we said already in the introduction, any local short-range potential can be approximated by an N-term separable potential in a mathematically well-defined sense. This explains and justifies our interest in a general kind of separable potential.

Due to the long range of the Coulomb potential, diffi-

culties occur in the theoretical description of the scattering. As we said before, the physical scattering amplitude is (in conventional short-range potential scattering theory) just the on-shell T matrix, apart from a simple factor. This cannot be the case in long-range potential scattering theory, because the on-shell limit is not defined then, as is known (compare Sec. 4). We repeat that this trouble is unaffected by any short-range potential; it is only the Coulomb potential that generates the problems.

As a way out of these problems it has been proposed to screen the Coulomb potential; eventually this screening is then removed by letting the screening radius go to infinity.[1,8-11,45] Since some time this approach can be interpreted in a rather satisfactory way[40,46] at least for two-particle scattering. In practice, of course, one always has screening at some very large distance. But even then it could be questionable whether one is allowed to apply conventional short-range scattering theory; the range of the potential is (very) large, though not infinite. This is a delicate question. Moreover, the screening approach is hampered by the fact that the scattering formulas (T matrices and the like) are not known analytically in general, and they are much more difficult to obtain. Screening can also be applied as long as it appears useful in singular expressions (cf. Refs. 7, 47, 48). This is implicitly done in the application of the on-shell Gell-Mann—Goldberger[42] two-potential formula to Coulomb-like potentials (cf. Ref. 6). The long-range troubles have often caused inaccuracies and inconsistencies in the past. In Sec. 4 we mentioned a few examples of such inconsistencies we found in the literature. Here we give some more. The Coulomb scattering state is often written (e.g. Refs. 17, 7)

$$| \mathbf{k} \pm \rangle_c = (1 + G_c(k^2 \pm i\epsilon)V_c)|\mathbf{k}\rangle , \qquad (126)$$

where G_c is the Coulomb Green's function. Although the formula analogous to (126) for a short-range potential is correct, (126) is at least inaccurate because the limit of $G_c(k^2 \pm i\epsilon)$ for $\epsilon \to 0$ is not clearly defined. Related to this we have the equation (see, e.g., Refs. 1, 21, 49)

$$\langle \mathbf{p}' | T_c | \mathbf{p} \rangle = \langle \mathbf{p}' | V_c | \mathbf{p} + \rangle_c. \qquad (127)$$

This equation is known to be correct when V has short range, but it is inaccurate, to say the least, when V is the Coulomb potential because the half-shell limit of T_c which is implicitly understood in Eq. (127) is not defined. In fact, we claim that the T_c given by Eq. (127) is *not* the conventional T_c defined by the (off-shell-) Eq. (10),

$$G_c(E) = G_0(E) + G_0(E)T_c(E)G_0(E).$$

To summarize one has:

(i) Exact analytic formulas without application of screening, but in such a situation no satisfactory physical interpretation is known.

(ii) A rather satisfactory interpretation when screening is applied, but in general no analytic formulas are known then. There might be problems also in this case because the range is (very) large though not infinite.

This whole problem can be attacked in a rigorous

manner to obtain a correct physical interpretation without the application of screening. All formulations then have to be set in an appropriate distribution-theoretic framework. Our approach is in a certain sense akin to the work of Herbst,[35] Taylor,[40] and Prugovečki and Zorbas.[50] We shall discuss this problem in a subsequent paper.[25] In this connection we note that Sec. 4 of the present paper solves only the *formal* difficulties as we found them in the literature. This is just a first step in the process to obtain the complete solution.

With the approach just mentioned we can in principle derive the desired physical quantities from the T matrix formulas obtained in the present paper, for example, amplitudes and effective range parameters. In advance we report here already that we found a means of obtaining the scattering length a_{cs} and the effective range r_{cs} in closed form for *all* our potentials. For the simplest one of our potentials, namely, Coulomb plus Yamaguchi, these two parameters have been obtained (i) by Harrington[1] in first order perturbation, (ii) by Ali and Rahman and Husain[51] purely numerically. It appears that the first order approximation of our result agrees exactly with the formulas derived by Harrington. The numerical results of Ref. 51 also agree with our closed formulas. More details about this subject will be reported.[25]

ACKNOWLEDGMENTS

[1]D. R. Harrington, Phys. Rev. **139**, B691 (1965).
[2]D. R. Harrington, Phys. Rev. **147**, 685 (1966).
[3]V. A. Alessandrini, H. Fanchiotti, and C. A. Garcia, Phys. Rev. **170**, 935 (1968).
[4]S. Adya, Phys. Rev. **177**, 1406 (1969).
[5]W. W. Zachary, J. Math. Phys. **12**, 1379 (1971); **14**, 2018 (1973).
[6]Ž. Bajzer, Nuovo Cimento A **22**, 300 (1974).
[7]G. Cattapan and G. Pisent, V. Vanzani, Lett. Nuovo Cimento **11**, 650 (1974).
[8]L. D. Faddeev, in *Three-Body Problem in Nuclear and Particle Physics*, edited by J. S. C. McKee and P. M. Rolph (North-Holland, Amsterdam, 1970).
[9]W. F. Ford, J. Math. Phys. **7**, 626 (1966).
[10]J. D. Dollard, J. Math. Phys. **9**, 620 (1968).
[11]J. Zorbas, J. Phys. A **7**, 1557 (1974).
[12]A. Messiah, *Quantum Mechanics* (North-Holland, Amsterdam, 1961).
[13]R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966).
[14]L. Hostler, J. Math. Phys. **5**, 591 (1964).
[15]McA. H. Hull and G. Breit, in *Encyclopedia of Physics*, edited by S. Flügge (Springer, Berlin, 1959), Vol. XLI/1.
[16]J. C. Y. Chen and A. C. Chen, in *Advances in Atomic and Molecular Physics*, edited by D. R. Bates and I. Esterman (Academic, New York, 1972), Vol. 8.
[17]S. Okubo and D. Feldman, Phys. Rev. **117**, 292 (1960).

1451   J. Math. Phys., Vol. 16, No. 7, July 1975

H. van Haeringen and R. van Wageningen   1451

[18]R. A. Mapleton, J. Math. Phys. **2**, 482 (1961).
[19]J. Schwinger, J. Math. Phys. **5**, 1606 (1964).
[20]L. Hostler, J. Math. Phys. **5**, 1235 (1964).
[21]W. F. Ford, Phys. Rev. **133**, B1616 (1964).
[22]G. L. Nutt, J. Math. Phys. **9**, 796 (1968).
[23]Y. Yamaguchi, Phys. Rev. **95**, 1628 (1954).
[24]V. Fock, Z. Physik **98**, 145 (1935).
[25]H. van Haeringen (to be published).
[26]L. Schwartz, *Théorie des distributions*, nouvelle édition (Hermann, Paris, 1966).
[27]F. Constantinescu, *Distributionen und ihre Anwendung in der Physik* (Teubner, Stuttgart, 1974).
[28]W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Springer, Berlin, 1966).
[29]Bateman Manuscript Project, *Higher Transcendental Functions*, edited by A. Erdélyi (McGraw-Hill, New York, 1953).
[30]M. Abramowitz and I. A. Stegun, editors, *Handbook of Mathematical Functions* (Dover, New York, 1965).
[31]O. P. Bahethi and M. G. Fuda, J. Math. Phys. **12**, 2076 (1971).
[32]For an attractive Coulomb force inequality (33) is not satisfied at energies between the ground state and zero energy, but with the help of (34) Eq. (19) still gives T_c correctly. The bound states are given by $i\gamma = -n$, $n = 1, 2, \cdots$. From (34) and Eq. (19) it can be seen that T_c has poles at the bound-state energies.
[33]C. S. Shastry and A. K. Rajagopal, Phys. Rev. A **2**, 781 (1970).
[34]J. S. Ball, J. C. Y. Chen, and D. Y. Wong, Phys. Rev. **173**, 202 (1968).
[35]I. W. Herbst, Commun. Math. Phys. **35**, 181 (1974).
[36]E. Prugovecki, *Quantum Mechanics in Hilbert Space* (Academic, New York, 1971).
[37]G. B. West, J. Math. Phys. **8**, 942 (1967).
[38]J. Nuttall and R. W. Stagat, Phys. Rev. A **3**, 1355 (1971).
[39]J. R. Taylor, *Scattering Theory* (Wiley, New York, 1972).
[40]J. R. Taylor, Nuovo Cimento B **23**, 313 (1974).
[41]L. Marquez, Am. J. Phys. **40**, 1420 (1972).
[42]M. Gell-Mann and M. L. Goldberger, Phys. Rev. **91**, 398 (1953).
[43]I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Academic, New York, 1965).
[44]H. van Haeringen and L. P. Kok, "Some Definite Integrals," Internal Report, 1972.
[45]J. T. Holdeman and R. M. Thaler, Phys. Rev. **139**, B1186 (1965).
[46]M. D. Semon and J. R. Taylor, Nuovo Cimento A **26**, 48 (1975).
[47]J. Hamilton, I. Øverbö, and B. Tromborg, Nucl. Phys. B **60**, 443 (1973).
[48]M. J. Roberts, J. Phys. B **7**, L397 (1974).
[49]M. J. Roberts, J. Phys. B **3**, L115 (1970).
[50]E. Prugovecki and J. Zorbas, J. Math. Phys. **14**, 1398 (1974).
[51]S. Ali, M. Rahman, and D. Husain, Phys. Rev. D **6**, 1178 (1972).

Is something missing in the Boltzmann entropy?*

B. Forte and C. C. A. Sastri

Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1
(Received 13 February 1975)

A representation theorem for entropy functionals on the set of probability densities on the space R_n is proved.
The important feature of the theorem is that the representation contains, in addition to the Boltzmann term,
the continuous analog of the Hartley entropy as well as a term that, in statistical mechanics, corresponds to
the chemical potential and is usually introduced ad hoc into the expression for entropy.

1. INTRODUCTION

The purpose of this paper is to prove a representation theorem for entropy functionals in statistical mechanics. This theorem relies heavily on the main result of Ref. 1.

There does not exist at present a completely satisfactory characterization of entropy in the continuous case. Generally, one of two practices has been followed: Either entropy is characterized in the discrete case and its extension to the continuous case conjectured, or it is characterized in the continuous case but only after assuming that either the entropy functional or some related functional has an integral representation. [2,3] Also, the functional may depend on the choice of units, or it may not include terms which are used in statistical mechanics but introduced ad hoc. In Ref. 4, where the discrete case was treated, a set of "natural" conditions were found which characterize the Shannon and Hartley entropies. A representation for the corresponding functional in the continuous case was found in Ref. 1. This representation has the advantage that it contains, in addition to the usual Boltzmann term, a term that depends on the dimension of the space, which in statistical mechanics is the same (except for a factor) as the number of particles in the system, and hence can be interpreted as the term corresponding to chemical reactions. However, it does not contain the continuous analog of the Hartley entropy. We believe this to be unfortunate, since it is natural that the expression for entropy be affected by any information about the set over which the distribution function is positive—or its complement, the set over which the function vanishes. We overcome this difficulty in the present paper. We do this by imposing certain intuitively natural conditions on the unknown entropy functional and by showing that any functional H_n (on the set of probability densities on R_n) which has those properties necessarily has the form

$$H_n(f) = -\alpha \int_{R_n} f \log f \, d\mu + bn + c \log\mu(A_f) \qquad (1)$$

where f is any probability density s.t. $f \log f$ is integrable, and $A_f := \{x \in R_n \mid f(x) > 0\}$; μ is the Lebesgue measure and α, b, c are real numbers, with $\alpha \geq 0$, $c \geq 0$. The last two terms on the rhs of (1) are the ones missing from the usual Boltzmann entropy.

We shall consider in a later publication the more realistic problem in which the requirement that the exact number of particles in the system—or equivalently, the exact dimension of the space—be known is dropped. In other words, we shall consider the grand canonical ensemble.

2. PRELIMINARIES

The notation used here is essentially the same as in Ref. 1. Let N denote the set of all positive integers and let B_n, $n \in N$, denote the σ-algebra of Borel sets in the Euclidean space R_n. Let μ be the Lebesgue measure on R_n and $V(R_n, B_n, \mu)$ the set of all probability densities in $L_1(R_n, B_n, \mu)$ s.t. $f \log f \in L_1(R_n, B_n, \mu)$. We shall denote the set of all simple probability densities in $V(R_n, B_n, \mu)$ by $\bar{V}(R_n, B_n, \mu)$ and the set of all nonnegative simple functions in $L_1(R_n, B_n, \mu)$ by $S(R_n, B_n, \mu)$. Whenever it is necessary to indicate the dependence on the dimension of the space, we shall do so by attaching a subscript to the appropriate quantities.

For each $f \in V$, let $A_f = \{(x_1, x_2, \ldots, x_n) \in R_n \mid f(x_1, x_2, \ldots, x_n) > 0\}$. Consider the generalized entropy functional H_n on V defined by (1). (The condition that $\alpha \geq 0$ is well known; the condition that $c \geq 0$ cannot be relaxed. See the Appendix for proof.) It has the following properties:

(1) Let $f \in V$. For any positive integer $l < n$, let f', f'' be the marginal densities of f defined by

$$f'(x_1, x_2, \ldots, x_l)$$
$$= \int_{R_{n-l}} f(x_1, x_2, \ldots, x_l, x_{l+1}, \ldots, x_n) \, dx_{l+1} \cdots dx_n, \qquad (2)$$
$$f''(x_{l+1}, x_{l+2}, \ldots, x_n)$$
$$= \int_{R_l} f(x_1, x_2, \ldots, x_l, x_{l+1}, \ldots, x_n) \, dx_1 \, dx_2 \cdots dx_l.$$

Then

$$H_n(f) \leq H_l(f') + H_{n-l}(f'') \quad \text{(subadditivity).}$$

(2) If, for some $l \in N$, $l < n$, and for every $x = (x_1, x_2, \ldots, x_n) \in R_n$

$$f(x) = f'(x_1, x_2, \ldots, x_l) f''(x_{l+1}, x_{l+2}, \ldots, x_n),$$

where f' and f'' are defined as in Eq. (2), then

$$H_n(f) = H_l(f') + H_{n-l}(f'') \quad \text{(additivity).}$$

(3) Let T be an invertible, measure-preserving transformation of R_n into itself and let U_T be the isometry on $L_1(R_n, B_n, \mu)$ induced by T:

$$U_T f(x) = f(Tx) \quad \forall x \in R_n.$$

Then

$$H_n(f) = H_n(U_T f).$$

(4) Let $\{S_i\}$ be a sequence of functions in $S(R_n, B_n, \mu)$ such that $s_i \uparrow f$ a.e. If $f \in V(R_n, B_n, \mu)$, then

$$H_n(s_i/\|s_i\|_1) \to H_n(f) \quad \text{as } i \to \infty.$$

Here, $\| \;\|_1$ denotes the L_1-norm.

Proof: Since properties (1)—(4) were proved in Ref. 1 for the functional

$$-\alpha \int f \log f \, d\mu + bn,$$

it is sufficient to show that the term $\log \mu(A_f)$ has these properties. Again, since properties (2), (3), and (4) are obviously satisfied, we shall check only property (1).

Let $l < n$, $l \in N$, and let f', f'' denote the marginal densities of f defined by Eq. (2). Let A_1, A_2 be the projections of A_f on R_l and R_{n-l} respectively. If $E \subset R_l$, $F \subset R_{n-l}$ denote the sets on which f', f'' are positive respectively, it suffices to show that

$$\mu_n(A_f) \leq \mu_l(E)\mu_{n-l}(F).$$

Divide A_1 into disjoint sets: $A_1 = A_{1,0} \cup A'_{1,0}$, where

$A_{1,0} =$ subset of A_1 on which $f' = 0$,

$A'_{1,0} =$ subset of A_1 on which $f' > 0$.

Clearly

$$A'_{1,0} \subset E \subset A_1.$$

Now either

$$\mu_l(A_{1,0}) = 0 \quad \text{or} \quad \mu_l(A_{1,0}) > 0.$$

If $\mu_l(A_{1,0}) = 0$,

$$\mu_l(A_1) = \mu_l(A'_{1,0}) = \mu_l(E).$$

If $\mu_l(A_{1,0}) > 0$, there are two possibilities:

(i) $\mu_n[(A_{1,0} \times R_{n-l}) \cap A_f] = 0$,

(ii) $\mu_n[(A_{1,0} \times R_{n-l}) \cap A_f] > 0$.

If (i) holds, then obviously $\mu_n(A_f) = \mu_n[(A'_{1,0} \times R_{n-l}) \cap A_f]$. Since $A'_{1,0} \subset E$, this means that except for a set of measure zero, $A_f \subset E \times R_{n-l}$. Now suppose (ii) holds. Equation (2) implies that, for each $x \in R_l$ s. t. $f'(x) = 0$, $f = 0$ a. e. with respect to μ_{n-l} on $\{x\} \times R_{n-l}$. This means that $f = 0$ a. e. with respect to μ_n on $A_{1,0} \times R_{n-l}$. Thus there exists a subset of A_f of positive measure, namely $(A_{1,0} \times R_{n-l}) \cap A_f$, on which $f = 0$. Contradiction. Hence (ii) cannot be true.

Thus either

$$\mu_l(A_{1,0}) = 0 \quad \text{or} \quad \mu_n[(A_{1,0} \times R_{n-l}) \cap A_f] = 0.$$

Similarly, either

$$\mu_{n-l}(A_{2,0}) = 0 \quad \text{or} \quad \mu_n[(R_l \times A_{2,0}) \cap A_f] = 0.$$

($A_{2,0}$ is defined the same way as $A_{1,0}$.)

Observe that

$$\mu_l(A_{1,0}) = 0 \Rightarrow \mu_n[(A_{1,0} \times R_{n-l}) \cap A_f] = 0;$$

similarly for the other.

If $\mu_l(A_{1,0}) = 0 = \mu_{n-l}(A_{2,0})$, $\mu_l(A_1) = \mu_l(E)$, and $\mu_{n-l}(A_2) = \mu_{n-l}(F)$, and the result follows. If $\mu_n[(A_{1,0} \times R_{n-l}) \cap A_f] = 0 = \mu_n[(R_l \times A_{2,0}) \cap A_f]$, then, except for sets of measure zero,

$$A_f \subset E \times R_{n-l} \quad \text{and} \quad A_f \subset R_l \times F,$$

i. e., except for a set of measure zero,

$$A_f \subset E \times F.$$

Hence the result.

3. REPRESENTATION THEOREM

We are now ready to state and prove the main result of this paper.

Theorem: If Φ_n is a functional on $V(R_n, B_n, \mu)$ which has properties (1)—(4), it has the representation (1); in other words,

$$\Phi_n(f) = H_n(f) \quad \forall f \in V.$$

In order to prove this theorem, we shall need the following two lemmas, for the proof of which we refer the reader to Ref. 1.

Lemma 1: Property (3) implies that there exist functions

$$g_n^m: \Gamma_m \times R_m^+ \to R_1, \quad m, n \in N,$$

$$\Gamma_m = \{(p_1, p_2, \ldots, p_m), \; p_i > 0 \text{ for } 1 \leq i \leq m, \; \sum_{i=1}^{m} p_i = 1\},$$

such that for every

$$s_m \equiv \sum_{i=1}^{m} \frac{p_i}{\mu(A_i)} \Big|_{A_i} \tag{3}$$

in $\overline{V}(R_n, B_n, \mu)$, with $A_i \cap A_j = \phi$ for $i \neq j$, $i, j = 1, 2, \ldots, m$, $p_i > 0$, and with $|_{A_i}$ denoting the characteristic function of A_i,

$$\Phi_n(s_m) = g_n^{(m)}(p_1, p_2, \ldots, p_m, \mu(A_1), \mu(A_2), \ldots, \mu(A_m)).$$

Lemma 2: Properties (1)—(4) imply that if $s_1 = |_A / \mu(A)$, $A \in B_n$, then

$$\Phi_n(s_1) = a \log \mu(A) + bn, \quad n \in N, \; a \geq 0, \text{ and } b \in R_1.$$

Remark: In Ref. 1, Lemma 2 is proved for a functional having properties (1)—(3) plus a property slightly more restrictive than (4). The result is true, however, even if the last property is replaced by property (4).

It can be shown[1] that if $\bar{m}$ is any positive number and $p, q, r \in [0, 1]$, $(p_3, p_4, \ldots, p_m) \in \Gamma_{m-2}$,

$$g_n^{(2)}(q, 1-q, \bar{m}, \bar{m})$$

$$- g_n^{(2)}(pr + (1-p)q, p(1-r) + (1-p)(1-q), \bar{m}, \bar{m})$$

$$\overset{m \text{ times}}{\leq g_n^{(m)}(pq, p(1-q), (1-p)p_3, \ldots, (1-p)p_m, \overbrace{\bar{m}, \bar{m}, \ldots, \bar{m}})}$$

$$\overset{m \text{ times}}{- g_n^{(m)}(pr, p(1-r), (1-p)p_3, \ldots, (1-p)p_m, \overbrace{\bar{m}, \bar{m}, \ldots, \bar{m}})}$$

$$\leq g_n^{(2)}((1-p)r + pq, (1-p)(1-r) + p(1-q), \bar{m}, \bar{m})$$

$$- g_n^{(2)}(r, 1-r, \bar{m}, \bar{m}).$$

These inequalities and the symmetry of $g_n^{(m)}(p_1, p_2, \ldots, p_m, \bar{m}, \bar{m}, \ldots, \bar{m})$—guaranteed by property (3)—can be used to apply Lemma 5 of Ref. 4 to $g_n^{(m)}$ and obtain (for $m \geq 2$)

$$g_n^{(m)}(p_1, p_2, \ldots, p_m, \bar{m}, \bar{m}, \ldots, m)$$

$$= -\alpha_n(m, \bar{m}) \sum_{i=1}^{m} p_i \log p_i + \beta_n(m, \bar{m}), \tag{4}$$

where $\alpha_n(m, \bar{m}) \geq 0$ and $\beta_n(m, \bar{m})$ is a real number. Now if $r, s, u, v \in N$, $\bar{m}, \bar{\bar{m}} > 0$, and if $(p_1, p_2, \ldots, p_u) \in \Gamma_u$, $(q_1, q_2, \ldots, q_v) \in \Gamma_v$, the additivity of Φ_n gives

$$g_{r+s}^{(uv)}(p_1 q_1, p_1 q_2, \ldots, p_1 q_v, p_2 q_1, \ldots, p_2 q_v, \ldots, p_u q_v,$$

$$\overbrace{\overline{\overline{m\bar{m}}, m\bar{m}, \ldots, m\bar{m}}}^{u \ v \ \text{times}})$$

$$= g_r^{(u)}(p_1, p_2, \ldots, p_u, \overbrace{\bar{m}, \bar{m}, \ldots, \bar{m}}^{u \ \text{times}})$$

$$+ g_s^{(v)}(q_1, q_2, \ldots, q_v, \overbrace{\bar{\bar{m}}, \bar{\bar{m}}, \ldots, \bar{\bar{m}}}^{v \ \text{times}}).$$

Since this is true for every $(p_1, p_2, \ldots, p_u) \in \Gamma_u$ and every $(q_1, q_2, \ldots, q_v) \in \Gamma_v$, it implies that

$$\alpha_{r+s}(uv, \bar{m}\bar{\bar{m}}) = \alpha_r(u, \bar{m}) = \alpha_s(v, \bar{\bar{m}}),$$

i. e., that $\alpha_r(u, \bar{m}) = \alpha$, a constant, for every $r \geq 1$, $u \geq 2$. Moreover,

$$\beta_{r+s}(uv, \bar{m}\bar{\bar{m}}) = \beta_r(u, \bar{m}) + \beta_s(v, \bar{\bar{m}}).$$

Now if $A = U_{i=1}^m A_i$, where the A_i^s are mutually disjoint sets in R_n, and if $\mu(A_i) = \bar{m}$, $p_i = 1/m$ for each i,

$$\sum_{i=1}^m \frac{p_i}{\mu(A_i)}\bigg|_{A_i} = \frac{1_A}{m\bar{m}}.$$

Hence Lemmas, 1, 2 and Eq. (4) give

$$-\alpha \log(1/m) + \beta_n(m, \bar{m}) = a \log(m\bar{m}) + bn$$

or

$$\beta_n(m, \bar{m}) = a \log \bar{m} + (a - \alpha) \log m + bn.$$

Hence

$$g_n^{(m)}(p_1, p_2, \ldots, p_m, \overbrace{\bar{m}, \bar{m}, \ldots, \bar{m}}^{m \ \text{times}})$$

$$= -\alpha \sum_{i=1}^m p_i \log p_i + a \log \bar{m} + (a - \alpha) \log m + bn$$

$$= -\alpha \sum_{i=1}^m p_i \log(p_i/\bar{m}) + bn + c \log(m\bar{m}),$$

where $c = a - \alpha$. As yet there are no restrictions on c but, as we shall show in the Appendix, subadditivity [property (1)] requires that $c \geq 0$.

Now consider $g_n^{(m)}(p_1, p_2, \ldots, p_m, h_1 \bar{m}, h_2 \bar{m}, \ldots, h_m \bar{m})$, $h_i \in N$, $1 \leq i \leq m$. By partitioning for each i the set of measure $h_i \bar{m}$ into h_i mutually disjoint sets each of which is of measure $\bar{m}$, it is easy to see that

$$g_n^{(m)}(p_1, p_2, \ldots, p_m, h_1 \bar{m}, \ldots, h_m \bar{m})$$

$$= -\alpha \sum_{i=1}^m p_i \log(p_i/h_i \bar{m})$$

$$+ c \log[(h_1 + h_2 + \cdots + h_m) \bar{m}] + bn.$$

Let $s_m \in \bar{V}(R_n, B_n, \mu)$ be defined as in (3). Then there exists a sequence of m-vectors

$$(h_{1r} \bar{m}_r, h_{2r} \bar{m}_r, \ldots, h_{mr} \bar{m}_r), \quad r = 1, 2, \cdots$$

with $\bar{m}_r \downarrow 0$ and $h_{ir} \in N$ for all i and r such that

$$h_{ir} \bar{m}_r \uparrow \mu(A_i) \quad \text{as } r \to \infty.$$

Also, for each r, there exists a set $A_{ir} \subset A_i$ such that $\mu(A_{ir}) = h_{ir} \bar{m}_r$.

The sequence of simple functions

$$s_{mr} = \sum_{i=1}^m \frac{p_i 1_{A_{ir}}}{\mu(A_i)}, \quad r = 1, 2, \ldots,$$

is bounded above by s_m and converges to it monotonically. Hence, by property (4),

$$\Phi_n(s_{mr}/\|s_{mr}\|_1) \to \Phi_n(s_m).$$

However, if we set

$$q_{ir} \equiv \left(\frac{p_i h_{ir} \bar{m}_r}{\mu(A_i)}\right) \bigg/ \sum_{i=1}^m \left(\frac{p_i h_{ir} \bar{m}_r}{\mu(A_i)}\right),$$

then

$$\frac{s_{mr}}{\|s_{m-r}\|_1} = \sum_{i=1}^m \frac{q_{ir} 1_{A_{ir}}}{h_{ir} \bar{m}_r},$$

so that

$$\Phi_n\left(\frac{s_{mr}}{\|s_{mr}\|_1}\right) = -\alpha \sum_{i=1}^m q_{ir} \log\left(\frac{q_{ir}}{h_{ir} \bar{m}_r}\right)$$

$$+ c \log[(h_{1r} + h_{2r} + \cdots + h_{mr}) \bar{m}_r] + bn.$$

As $r \to \infty$, $q_{ir} \to p_i$ and $h_{ir} \bar{m}_r \uparrow \mu(A_i)$. Hence

$$\Phi_n(s_m) = -\alpha \sum_{i=1}^m p_i \log\left(\frac{p_i}{\mu(A_i)}\right) + c \log\left(\sum_{i=1}^m \mu(A_i)\right) + bn.$$

Finally, let $f \in V(R_n, B_n, \mu)$.

For each x let

$$f_r(x) = \begin{cases} (i-1)/2^r & \text{if } (i-1)/2^r \leq f(x) < i/2^r \quad (i = 1, 2, \ldots, 2^r r) \\ r & \text{if } f(x) \geq r. \end{cases}$$

Then $f_r \uparrow f$ everywhere, and since f is integrable, each f_r is integrable.

If

$$A_{ir} \equiv \{x \in R_n \mid (i-1)/2^r \leq f(x) < i/2^r\}$$

and

$$B_r \equiv \{x \in R_n \mid f(x) \geq r\},$$

$$f_r = \sum_{i=1}^{2^r r} \frac{(i-1)}{2^r} 1_{A_{ir}} + r 1_{B_r} \quad \text{for each } r.$$

The integrability of f_r implies that all the sets B_r and A_{ir} except A_{1r} are of finite measure and that $\mu(A_{1r}) = \infty$. Further, the function $f_r/\|f_r\|_1$ can be put in the form (3) with

$$p_{ir} = [(i-1)/2^r \|f_r\|_1] \mu(A_{ir}) \quad (i = 2, \ldots, 2^r r)$$

$$p_r = (r/\|f_r\|_1) \mu(B_r),$$

so that

$$\sum_{i=2}^{2^r r} p_{ir} + p_r = 1.$$

Hence

$$\Phi_n\left(\frac{f_r}{\|f_r\|_1}\right) = -\alpha \sum_{i=2}^{r 2^r} \frac{(i-1) \mu(A_{ir})}{2^r \|f_r\|_1} \log\left(\frac{i-1}{2^r \|f_r\|_1}\right)$$

$$- \alpha \frac{r}{\|f_r\|_1} \mu(B_r) \log\left(\frac{r}{\|f_r\|_1}\right)$$

1455   J. Math. Phys., Vol. 16, No. 7, July 1975

B. Forte and C.C.A. Sastri   1455
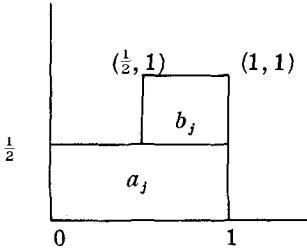
FIG. 1

$$+ c \log \left(\sum_{i=2}^{2^r} \mu(A_{ir}) + \mu(B_r) \right) + bn$$

$$= H_n \left(\frac{f_r}{\| f_r \|_1} \right).$$

But we have already proved that $H_n(f_r/\| f_r \|_1) \to H_n(f)$.

By property (4),

$$\Phi_n(f_r/\| f_r \|_1) \to \Phi_n(f).$$

Hence

$$\Phi_n(f) = H_n(f),$$

and the theorem is proved.

APPENDIX

We shall now show, by constructing an example, that subadditivity requires that $c \geq 0$, given that $a \geq 0$. Let $n = 2$. $\forall j \in N$, let f_j be defined as shown in Fig. 1 with $a_j, b_j > 0$.

Then $\mu(A_{f_j}) = \frac{3}{4}$ and $A_{1j} = A_{2j} = [0, 1]$. Normalization of f_j gives $2a_j + b_j = 4$. The marginal densities f_j', f_j'' of f_j and hence the functionals $H_2(f_j)$, $H_1(f_j')$, $H_1(f_j'')$ can easily be evaluated. Some computation shows that in order for subadditivity to hold, we must have

$$a \left[\frac{a_j}{4} \log \left(\frac{a_j}{2} \right) + \frac{(a_j + b_j)}{4} \log \frac{(a_j + b_j)}{2} - \frac{b_j}{4} \log 2 \right]$$

$$+ c \log \tfrac{3}{4} \leq 0. \qquad (5)$$

For each j, let

$$a_j = (2^j - 1)/2^{j-1}, \quad b_j = 1/2^{j-2}.$$

Then after a small computation (5) reduces to

$$a \left(\frac{1}{2} \log \frac{2^{2j} - 1}{2^{2j}} + \frac{1}{2^{j+1}} \log \frac{2^j + 1}{2^{j+2} - 2^2} \right) + c \log \tfrac{3}{4} \leq 0,$$

i. e. ,

$$\left(c + \frac{a}{2^{j-2}} \right) \log \frac{3}{4} + a \left[\frac{1}{2} \log \frac{2^{2j} - 1}{2^{2j}} \right.$$

$$\left. + \frac{1}{2^{j+1}} \log \left(\frac{2^j + 1}{2^j - 1} \cdot \frac{2^{14}}{3^8} \right) \right] \leq 0.$$

We claim that the quantity in square brackets on the lhs is nonnegative. Assuming for the moment that it is true and using the fact that $a \geq 0$, we must, in order for subadditivity to hold, have

$$c + a/2^{j-2} \geq 0.$$

Since it would be true for each j, it would imply that $c \geq 0$. Now it just remains to show that

$$\frac{1}{2} \log \frac{2^{2j} - 1}{2^{2j}} + \frac{1}{2^{j+1}} \log \left(\frac{2^j + 1}{2^j - 1} \cdot \frac{2^{14}}{3^8} \right) \geq 0.$$

Let

$$1/2^j = x. \quad \text{As } j \uparrow \infty, \; x \downarrow 0.$$

We have to show that

$$\log(1 - x^2) + x \log \frac{2^{14}}{3^8} + x \log \left(\frac{x + 1}{1 - x} \right) \geq 0.$$

The inequality is clearly satisfied when $x = 0$. It suffices then to show that

$$f(x) = \log(1 - x^2) + x \log \frac{2^{14}}{3^8} + x \log \left(\frac{x + 1}{1 - x} \right)$$

is an increasing function as x increases from 0. For this, it is sufficient to show that $f'(x) \geq 0$:

$$f'(x) = \frac{-2x}{1 - x^2} + \log \frac{2^{14}}{3^8} + \log \left(\frac{x + 1}{1 - x} \right) + \frac{x}{x + 1} + \frac{x}{1 - x}$$

$$= \log \frac{2^{14}}{3^8} + \log \left(\frac{x + 1}{1 - x} \right),$$

which is clearly nonnegative.

[1] B. Forte (1974), "Additive and Subadditive Entropy Functionals", a University of Waterloo preprint; to appear in the Proceedings of the Conference on Measures of Information and Their Applications, held in Aug. 1974 in Bombay.
[2] L. L. Campbell, Information and Control 21, 329 (1972).
[3] S. Ikeda, Ann. Inst. Stat. Math. 13, 259 (1962).
[4] J. Aczel, B. Forte and C.T.Ng, Advan. Appl. Prob. 6, 131 (1974).

Evaluation of a class of lattice sums in arbitrary dimensions*

A. N. Chaba

Universidade Federal da Paraíba, Departamento de Física, João Pessoa, Paraíba, Brazil

R. K. Pathria

Department of Physics, University of Waterloo, Waterloo, Ontario, Canada
(Received 4 March 1975)

Starting from the Poisson summation formula in m dimensions, a class of lattice sums is evaluated analytically. The resulting formulas are applicable to the electronic-structure studies of crystalline solids, the analysis of stability of quantized vortex arrays in extreme type-II superconductors and in rotating superfluid helium, and the investigation of Bose–Einstein condensation in finite systems.

I. INTRODUCTION

Recent work by Glasser[1-3] and Zucker[4,5] has revived interest in the analytic evaluation of m-dimensional sums

$$\sum_{\{l_i\}=-\infty}^{\infty}{}' f(l_1, l_2, \ldots, l_m)(l_1^2 + l_2^2 + \cdots + l_m^2)^{-s}, \qquad (1)$$

where $\sum'$ excludes the term with $l_1 = l_2 = \cdots = 0$. While Glasser has concentrated mainly on two-dimensional sums Zucker has investigated sums in 4, 6, and 8 dimensions as well. However, the nature of their techniques is such that it forestalls any comparable degree of success with sums in an odd number of dimensions, including the case $m = 3$.

In our recent work on Bose–Einstein condensation in finite systems[6,7] we have encountered similar sums, with

$$f(l_1, l_2, \ldots, l_m)$$
$$= \exp[-a(l_1^2 + l_2^2 + \cdots + l_m^2)] \quad (a > 0). \qquad (2)$$

These sums have also appeared previously in the work of Fetter, Hohenberg, and Pincus[8] on the stability of an array of quantized vortices in type-II superconductors and in rotating superfluid helium. Furthermore, the behavior of these sums for small values of a turns out to be relevant to the work of Harris and Monkhorst[9] on Madelung sums appearing in the electronic-structure studies of crystalline solids. In the two-dimensional case, in particular, these sums may find direct applications in the general area of surface science. It, therefore, appeared of interest to investigate these sums analytically. The main results of this investigation are being reported here.

II. EVALUATION OF SUMS

We start with the Poisson summation formula

$$\sum_{l=-\infty}^{\infty} \exp(-al^2) = \left(\frac{\pi}{a}\right)^{1/2} \sum_{l=-\infty}^{\infty} \exp(-\pi^2 l^2/a), \qquad (3)$$

which holds for all $a > 0$. The m-dimensional version of (3) gives

$$\sum_{R_m}{}' \left[\exp(-aR^2) - \left(\frac{\pi}{a}\right)^{m/2} \exp(-\pi^2 R^2/a)\right] = \left(\frac{\pi}{a}\right)^{m/2} - 1, \qquad (4)$$

where $R^2 = l_1^2 + l_2^2 + \cdots + l_m^2$. Integrating (4) with respect to a, we obtain the identity

$$\sum_{R_m}{}' \left[\frac{1}{R^2} \exp(-aR^2) + \frac{\pi^{m/2}}{a^{(m-2)/2}} E_{(4-m)/2}\left(\frac{\pi^2 R^2}{a}\right)\right]$$

$$= \begin{cases} \dfrac{2}{m-2}\,\pi^{m/2} a^{(2-m)/2} + a + C_m & (m = 1, 3, 4) \qquad (5a) \\[2mm] -\pi \ln a + a + C_2 & (m = 2), \qquad (5b) \end{cases}$$

where the exponential integral is given by

$$E_n(z) = \int_1^{\infty} t^{-n} \exp(-zt)\,dt \quad \left(n = \frac{4-m}{2} \geq 0\right). \qquad (6)$$

For $m > 4$, the function $E_n(z)$ in (5) has to be replaced by

$$\alpha_n(z) = \int_1^{\infty} t^n \exp(-zt)\,dt \quad \left(n = \frac{m-4}{2} > 0\right); \qquad (7)$$

the right-hand side remains the same as in (5a). We observe that, for all m, the sum in (5) is minimum when $a = \pi$; it follows that

$$C_m > m\pi/(2-m) \quad (m \neq 2) \qquad (8a)$$

and

$$C_2 > \pi \ln \pi - \pi. \qquad (8b)$$

For $m = 4$, we obtain the simple result

$$\sum_{R_4}{}' \left(\frac{1}{R^2}\,[\exp(-aR^2) + \exp(-\pi^2 R^2/a)]\right) = \frac{\pi^2}{a} + a + C_4, \qquad (9)$$

which is symmetric with respect to the interchange $a \leftrightarrow (\pi^2/a)$. For $a = \pi$, it reduces to

$$\sum_{R_4}{}' \frac{1}{R^2} \exp(-\pi R^2) = \pi + \tfrac{1}{2}C_4, \qquad (10)$$

from which the constant C_4 can be readily evaluated. We obtain $C_4 = -5.545178$.

We shall now consider more practical cases.

(i) $m = 2$. In this case our identity takes the form

$$\sum_{R_2}{}' \left[\frac{1}{R^2} \exp(-aR^2) + \pi E_1\left(\frac{\pi^2 R^2}{a}\right)\right] = -\pi \ln a + a + C_2. \qquad (11)$$

Noting that

$$C_2 = \lim_{a \to 0} \left(\sum_{R_2}{}' \frac{1}{R^2} \exp(-aR^2) - \pi \ln(1/a)\right), \qquad (12)$$

we obtain, following Glasser's method,[1]

$$C_2 = \pi \left(\gamma - \ln \frac{[\Gamma(\tfrac{1}{4})]^4}{4\pi^3} \right) = 0.771605. \qquad (13)$$

Using (11) we have been able to simplify considerably the evaluation of the planar sum $\sum' K_0(\mu\sqrt{l_1^2+l_2^2})$ encountered by Fetter et al.,[8] with the result

$$\sum_{l_i=-\infty}^{\infty}{}' K_0(\mu\sqrt{l_1^2+l_2^2}) = \frac{2\pi}{\mu^2} + \frac{1}{2}\ln\left(\frac{\mu^2}{4\pi}\right) + \left(\gamma - \ln\frac{[\Gamma(\tfrac{1}{4})]^2}{2\pi}\right)$$

$$- \frac{\mu^2}{2\pi} \sum_{l_i=-\infty}^{\infty}{}' (l_1^2+l_2^2)^{-1}[\mu^2 + 4\pi^2(l_1^2+l_2^2)]^{-1}, \qquad (14)$$

valid for all $\mu > 0$. The same sum also showed up in our recent analysis of Bose–Einstein condensation in a two-dimensional system of finite size.[7] The reduction effected in (14) then enabled us to discuss successfully the circumstances under which a condensate component could appear in the system.

For $a \ll \pi$, Eq. (11) gives

$$\sum_{R_2}{}' \frac{1}{R^2} \exp(-aR^2) = -\pi\ln a + C_2 + a - O[a\exp(-\pi^2/a)]. \qquad (15)$$

Successive integrations with respect to a now yield

$$\sum_{R_2}{}' \frac{1}{R^4} \exp(-aR^2)$$
$$\simeq 4\zeta(2)\,\beta(2) + \pi a\ln a - (\pi + C_2)a - \tfrac{1}{2}a^2, \qquad (16)$$

$$\sum_{R_2}{}' \frac{1}{R^6} \exp(-aR^2)$$
$$\simeq 4\zeta(3)\,\beta(3) - 4\zeta(2)\,\beta(2)a - \frac{\pi}{2}a^2\ln a + \left(\frac{3\pi}{4} + \frac{C_2}{2}\right)a^2 + \tfrac{1}{6}a^3, \qquad (17)$$

and so on. The constants of integration in these formulas were obtained from the following result due to Hardy[10]:

$$\sum_{l_i=-\infty}^{\infty}{}' (l_1^2+l_2^2)^{-s} = 4\zeta(s)\,\beta(s) \quad (s>1), \qquad (18)$$

where

$$\zeta(s) = \sum_{l=0}^{\infty} (l+1)^{-s}, \quad \beta(s) = \sum_{l=0}^{\infty} (-1)^l (2l+1)^{-s}.$$

It is important to note that the error term in formulas (15)–(17) is such that, though valid for small a, they are useful for sizable values of a as well. For instance, even for $a=1$, we obtain an accuracy of about 1 part in 30 000 from formula (15), about 1 part in 250 000 from formula (16), and about 1 part in 750 000 from formula (17).

Combining (15) with the corresponding one-dimensional result,

$$\sum_{l=-\infty}^{\infty}{}' \frac{1}{l^2} \exp(-al^2) = C_1 - 2\pi^{1/2}a^{1/2} + a - O[a^{3/2}\exp(-\pi^2/a)], \qquad (19)$$

where

$$C_1 = \pi^2/3 = 3.289868, \qquad (20)$$

we get

$$\sum_{l_i=1}^{\infty} \frac{1}{(l_1^2+l_2^2)} \exp[-a(l_1^2+l_2^2)]$$

$$= -\frac{\pi}{4}\ln a + \left(\frac{C_2}{4} - \frac{C_1}{2}\right) + \pi^{1/2}a^{1/2}$$
$$- \tfrac{1}{4}a - O[a\exp(-\pi^2/a)]. \qquad (21)$$

The last result provides an improvement over the one reported earlier by Glasser.[1]

For $a \gg \pi$, we obtain from (11), writing $(\pi^2/a)=b \ll \pi$,

$$\sum_{R_2}{}' E_1(bR^2) = \frac{\pi}{b} + \ln b + \left(\frac{C_2}{\pi} - 2\ln\pi\right) - O[\exp(-\pi^2/b)]. \qquad (22)$$

The main term here corresponds to replacing the summation over $\mathbf{R}$ by an integration while the remaining terms arise from the discreteness of the sum. Again, the nature of the error term in this formula is such that it can be useful even when b is not too small.

(ii) $m = 3$. In this case we have

$$\sum_{R_3}{}' \left[\frac{1}{R^2}\exp(-aR^2) + \frac{\pi^{1/2}}{R}\,\Gamma\left(\frac{1}{2},\frac{\pi^2R^2}{a}\right)\right] = \frac{2\pi^{3/2}}{a^{1/2}} + a + C_3; \qquad (23)$$

clearly,

$$C_3 = \lim_{a\to 0}\left(\sum_{R_3}{}' \frac{1}{R^2}\exp(-aR^2) - \frac{2\pi^{3/2}}{a^{1/2}}\right). \qquad (24)$$

First of all, we note that by using (23) we can express the sum

$$S_1(y) = \sum_{q}{}' \frac{\exp[-2y(q_1^2+q_2^2+q_3^2)^{1/2}]}{y(q_1^2+q_2^2+q_3^2)^{1/2}}, \qquad (25)$$

which appears in the study of Bose–Einstein condensation in a three-dimensional system of finite size,[6] in a form that will allow us to carry out a rigorous discussion of the phenomenon of condensation in terms of a "collapse of the lattice points of the thermogeometric space of the given system towards its origin." Details of this work will be reported elsewhere.

For $a \ll \pi$, we obtain from (23)

$$\sum_{R_3}{}' \frac{1}{R^2}\exp(-aR^2) = \frac{2\pi^{3/2}}{a^{1/2}} + C_3 + a - O[a^{1/2}\exp(-\pi^2/a)]. \qquad (26)$$

Evaluating this sum for a suitable value of a, not necessarily too small, we get

$$C_3 = -8.913633. \qquad (27)$$

This is in perfect agreement with the value reported by Harris and Monkhorst[9] for their constant C which was defined as

$$C = \lim\left(\frac{8\pi^3}{a_1a_2a_3}\sum_{k_3}{}' \frac{1}{k^2} - \int \frac{1}{k^2}\,d^3k\right), \qquad (28)$$

where

$$\mathbf{k} = 2\pi\left(\frac{l_1}{a_1}, \frac{l_2}{a_2}, \frac{l_3}{a_3}\right)$$

and the limit in (28) implies that the included region of the k-space is extended to infinity. In the work of Harris and Monkhorst, C is a measure of the electric potential at a given lattice point when a unit positive charge is placed at each of the other lattice points while a balancing negative charge is distributed uniformly throughout the space. Expressed in units of $(2\pi/a_1)$, C is a function of the ratios a_2/a_1 and a_3/a_1. When $a_1=a_2=a_3$, C is precisely equal to our C_3; cf. (24) and (28).

The foregoing observation suggests that we generalize our results to the anisotropic case where

$$R = (r_1^2 l_1^2 + r_2^2 l_2^2 + r_3^2 l_3^2)^{1/2}. \qquad (29)$$

To do this we simply modify the Poisson summation formula (3) to

$$\sum_{l_i=-\infty}^{\infty} \exp(-a r_i^2 l_i^2) = \frac{1}{r_i} \left(\frac{\pi}{a}\right)^{1/2} \sum_{l_i=-\infty}^{\infty} \exp(-\pi^2 l_i^2 / r_i^2 a)$$

$$(i = 1, 2, 3), \qquad (3')$$

with the result

$$\sum_{R_3}{}' \frac{1}{R^2} \exp(-aR^2)$$

$$\simeq \frac{2\pi^{3/2}}{(r_1 r_2 r_3)a^{1/2}} + C_3(r_1, r_2, r_3) + a \left(C_3 > \frac{-3\pi}{(r_1 r_2 r_3)^{2/3}}\right).$$

$$(26')$$

The correspondence between our constant of integration C_3 and the lattice-structure constant C of Harris and Monkhorst now turns out to be

$$C_{HM}\left(\frac{a_2}{a_1}, \frac{a_3}{a_1}\right) \equiv \frac{a_2}{a_3} C_3\left(\frac{a_2}{a_1}, 1, \frac{a_2}{a_3}\right). \qquad (30)$$

However, the evaluation of C_3 with the help of formula (26') is considerably simpler than the evaluation of C_{HM} by the method of Harris and Monkhorst.

In passing we note that, by combining (15), (19), and (26), we obtain

$$\sum_{l_i=1}^{\infty} \frac{1}{(l_1^2 + l_2^2 + l_3^2)} \exp[-a(l_1^2 + l_2^2 + l_3^2)]$$

$$= \frac{1}{4} \frac{\pi^{3/2}}{a^{1/2}} + \frac{3\pi}{8} \ln a + \tfrac{1}{8}(C_3 - 3C_2 + 3C_1) - \tfrac{3}{4}\pi^{1/2} a^{1/2}$$

$$+ \tfrac{1}{8}a - O[a^{1/2} \exp(-\pi^2/a)]. \qquad (31)$$

Moreover, by successive integrations of (26) with respect to a, we can obtain three-dimensional formulae similar to (16) and (17). The constants of integration appearing there would involve triple sums such as

$$\sum_{\{l_i\}}{}' (l_1^2 + l_2^2 + l_3^2)^{-s},$$

which can be evaluated by using a technique due to Mackenzie.[11,12]

For $a \gg \pi$, we obtain from (23), writing $(\pi^2/a) = b \ll \pi$,

$$\sum_{R_3}{}' \frac{1}{R} \Gamma\left(\tfrac{1}{2}, bR^2\right) = \frac{\pi^{3/2}}{b} + \frac{C_3}{\pi^{1/2}} + 2b^{1/2} - O[\exp(-\pi^2/b)]. \qquad (32)$$

Again, the main term of this result corresponds to replacing the summation over **R** by an integration while the remaining terms arise from the discreteness of the sum. Moreover, this formula is useful even when b is not too small.

III. SOME FURTHER RESULTS

Multiplying (5) by $a^{(m-6)/2}$ and integrating with respect to a, we obtain

$$\sum_{R_m}{}' \frac{1}{R^{m-2}} \left[\Gamma\left(\frac{m-4}{2}, aR^2\right) - \left(\frac{\pi}{a}\right)^{(4-m)/2} \Gamma\left(\frac{m-4}{2}, \frac{\pi^2 R^2}{a}\right)\right]$$

$$= \begin{cases} \dfrac{2\pi^{m/2}}{(m-2)a} - \dfrac{2}{(m-2)a^{(2-m)/2}} - \dfrac{2C_m}{(m-4)a^{(4-m)/2}} + D_m \\ \qquad\qquad\qquad\qquad (m \neq 2, 4) & (33a) \\[8pt] -\dfrac{\pi}{a}\ln a - \dfrac{\pi - C_2}{a} - \ln a + D_2 \quad (m = 2) & (33b) \\[8pt] \dfrac{\pi^2}{a} - a - C_4 \ln a + D_4 \qquad (m = 4), & (33c) \end{cases}$$

where

$$\Gamma\left(\frac{m-4}{2}, z\right) = \begin{cases} z^{(m-4)/2} E_{(6-m)/2}(z) & (m \leq 6) \\ \exp(-z) & (m = 6) \\ z^{(m-4)/2} \alpha_{(m-6)/2}(z) & (m \geq 6). \end{cases}$$

For $a = \pi$, the left-hand side of (33) vanishes identically; therefore,

$$D_2 = 2\ln\pi + 1 - (C_2/\pi), \quad D_4 = C_4\ln\pi,$$

and

$$D_m = \frac{2C_m}{(m-4)\pi^{(4-m)/2}} \quad (m \neq 2, 4).$$

We note that, in view of the relationship

$$\Gamma(-n, z) = \frac{1}{n}\left(\frac{e^{-z}}{z^n} - \Gamma(1-n, z)\right) \quad (n \neq 0), \qquad (34)$$

the identities (33) can be expressed as linear combinations of the identities (5) and the ones obtained from (5) on replacing a by π^2/a; in fact,

$$f_{(33)}(a) = \frac{2}{(4-m)}\left(\frac{f_{(5)}(a)}{a^{(4-m)/2}} - \frac{f_{(5)}(\pi^2/a)}{\pi^{(4-m)/2}}\right) \quad (m \neq 4), \qquad (35)$$

where $f_{(5)}$ and $f_{(33)}$ denote the summands appearing in Eqs. (5) and (33), respectively. Consequently, though useful in some respects, they do not represent anything new, except in the case $m = 4$ when we do obtain a new result, viz.

$$\sum_{R_4}{}' \frac{1}{R^2}\left[E_1(aR^2) - E_1\left(\frac{\pi^2 R^2}{a}\right)\right] = \frac{\pi^2}{a} - a + C_4\ln\left(\frac{\pi}{a}\right). \qquad (36)$$

For $a \ll \pi$, it gives

$$\sum_{R_4}{}' \frac{1}{R^2} E_1(aR^2) = \frac{\pi^2}{a} + C_4\ln\left(\frac{\pi}{a}\right) - a + O[a\exp(-\pi^2/a)], \qquad (37)$$

which is consistent with the fact that, in the limit $a \to 0$,

$$\sum_{R_4}{}' \to \int_0^\infty \frac{1}{R^2} E_1(aR^2)(2\pi^2 R^3 \, dR) = \frac{\pi^2}{a}. \qquad (38)$$

Before concluding we wish to remark that certain other classes of lattice sums can be similarly evaluated if, instead of the conventional Poisson summation formulas (3) and (4), one employes

$$\sum_{l=-\infty}^{\infty} (-1)^l \exp(-al^2) = \left(\frac{\pi}{a}\right)^{1/2} \sum_{l=-\infty}^{\infty} \exp[-\pi^2(l + \tfrac{1}{2})^2/a] \qquad (39)$$

and the corresponding m-dimensional version of this formula. This would lead to Madelung sums for a lattice in which the charges placed at the various lattice points alternate in sign. Moreover, depending on the nature of

the problem at hand, one may as well use a suitable combination of the formulas (3) and (39). Finally, for treating lattices with a lower degree of symmetry, one may start with a more generalized form of the Poisson summation formula, such as

$$\sum_{l_{1,2}=-\infty}^{\infty} \exp[-a(l_1^2 + l_1 l_2 + l_2^2)]$$

$$= \left(\frac{2\pi}{\sqrt{3}a}\right) \sum_{l_{1,2}=-\infty}^{\infty} \exp[-4\pi^2(l_1^2 + l_1 l_2 + l_2^2)/3a], \qquad (40)$$

which is appropriate for a triangular lattice. More generally, one may employ the formula

$$\sum_{\{l_i\}} \exp\left[-a \sum_{i,j=1}^{m} r_{ij} l_i l_j\right]$$

$$= \left(\frac{\pi}{a}\right)^{m/2} \sum_{\{l_i\}} \exp\left[-\frac{\pi^2}{a} \sum_{i,j=1}^{m} r_{ij}^{-1} l_i l_j\right], \qquad (41)$$

where (r_{ij}^{-1}) is the inverse of the matrix (r_{ij}) and, without loss of generality, the determinant $|r_{ij}|$ has been taken to be equal to unity. Generalizations resulting from the use of formulas (39)—(41) will not be pursued here.

ACKNOWLEDGMENTS

*Work supported in part by the National Research Council of Canada.
[1]M.L. Glasser, J. Math. Phys. **14**, 409 (1973). Equation (23) of this reference is in error; it should read as Eq. (21) of the present paper.
[2]M.L. Glasser, J. Math. Phys. **14**, 701 (1973).
[3]M.L. Glasser, J. Math. Phys. **15**, 188 (1974).
[4]I.J. Zucker, J. Math. Phys. **15**, 187 (1974).
[5]I.J. Zucker, J. Phys. A **7**, 1568 (1974); also I.J. Zucker and M.M. Robertson, J. Phys. A (to be published).
[6]S. Greenspoon and R.K. Pathria, Phys. Rev. A **9**, 2103 (1974).
[7]A.N. Chaba and R.K. Pathria, to be published.
[8]A.L. Fetter, P.C. Hohenberg and P. Pincus, Phys. Rev. **147**, 140 (1966).
[9]F.E. Harris and H.J. Monkhorst, Phys. Rev. B **2**, 4400 (1970).
[10]G.H. Hardy, Mess. Math. **49**, 85 (1919).
[11]J.K. Mackenzie, Proc. Phys. Soc. (Lond.) A **63**, 1370 (1950).
[12]B.M.E. van der Hoff and G.C. Benson, Can. J. Phys. **31**, 1087 (1953).

Separation of variables in the Hamilton–Jacobi, Schrödinger, and related equations. I. Complete separation*

Peter Havas†

Department of Physics, Temple University, Philadelphia, Pennsylvania 19122

It was established by Levi-Civita that in n dimensions there exist $n+1$ types of coordinate systems in which the Hamilton–Jacobi equation is separable, n of which are in general nonorthogonal; the form of the separated equations was given by Burgatti and Dall'Acqua. In this paper first the general forms of the $n+1$ types of metric tensors of the corresponding Riemannian spaces V_n are determined. Then, sufficient conditions are given for coordinate systems in which the Schrödinger, Helmholtz, and Laplace equation are separable. It is shown that there again exist $n+1$ types of such systems, whose metric tensors are of the same form as those of the Hamilton–Jacobi equation. However, except for the "essentially geodesic case" of Levi-Civita they are further restricted by a condition on the determinant of the metric; this condition is a generalization of that found by Robertson for orthogonal systems in the case of the Schrödinger equation.

I. INTRODUCTION

The method of separation of variables is probably the most common one for attempting to find solutions of partial differential equations (PDE's), and is mentioned at least in passing in every textbook dealing with such equations. It is therefore rather surprising that the literature contains very few systematic attempts to determine the coordinate systems which permit the solution of a particular PDE by this method.

Almost all equations studied can be associated with a quadratic form

$$T(q_1 \cdots q_n, \dot{q}_1 \cdots \dot{q}_n) \equiv \tfrac{1}{2} \sum_{k,l=1}^{n} g_{kl}(q_1 \cdots q_n) \dot{q}_k \dot{q}_l,$$

$$\dot{q}_i \equiv \frac{dq_i}{dt}, \quad g_{kl} = g_{lk}, \tag{1}$$

where the q_i can be considered as coordinates, t as the time, and the $\tfrac{1}{2}n(n+1)$ functions g_{kl} as the metric tensor of a Riemannian space V_n. In the following it will be understood that the range of Latin indices is from 1 to n, but for our purposes it will be preferable always to indicate summations 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}.

If we are dealing with a problem of classical mechanics, T is interpreted as the kinetic energy and the metric is positive definite. For a conservative system we then have a potential energy V, a Lagrangian

$$L(q_1 \cdots q_n, \dot{q}_1 \cdots \dot{q}_n) \equiv T(q_1 \cdots q_n, \dot{q}_2 \cdots \dot{q}_n) - V(q_1 \cdots q_n), \tag{2}$$

and a Hamiltonian

$$H(q_1 \cdots q_n, p_1 \cdots p_n) \equiv T(q_1 \cdots q_n, p_1 \cdots p_n) + V(q_1 \cdots q_n),$$

$$T(q_1 \cdots q_n, p_1 \cdots p_n) \equiv \tfrac{1}{2} \sum_{k,l=1}^{n} g^{kl} p_k p_l, \tag{3}$$

$$\sum_{m=1}^{n} g_{km} g^{lm} = \delta_k{}^l,$$

where $\delta_k{}^l$ is the Kronecker delta. We too shall restrict ourselves to PDE's which can be associated with (1) and (3), but without necessarily implying a connection with mechanics or a positive definite metric. The problem of separability then reduces to the determination of all coordinate systems, i.e., of all metric tensors g_{kl}, for which solutions exist which can be obtained by separating the PDE associated with (1) into a number of ordinary differential equations (ODE's) and of the form of all potential functions V which allow such a separation.

The most famous PDE associable with (1) is Laplace's, which, as is well known, is separable in Euclidean 3-space E_3 in 11 orthogonal coordinate systems (the ellipsoidal coordinates and their degenerate forms).[1] However, from the point of view of separation of variables it is neither the simplest one nor the one studied most systematically. Surprisingly, from this point of view the simplest PDE, and the only one for which the problem of separation of variables has been solved completely, is a nonlinear one. The system (3) can be associated with the Hamilton–Jacobi (H–J) equation

$$\frac{\partial S}{\partial t} + H(q_1 \cdots q_n, p_1 \cdots p_n) = 0,$$

$$p_i \equiv \frac{\partial S}{\partial q_i}, \quad i = 1 \cdots n. \tag{4}$$

If we assume

$$S(t, q_1 \cdots q_n) = -Et + W(q_1 \cdots q_n), \tag{5}$$

where E is a constant, Eq. (4) with (3) reduces to

$$\tfrac{1}{2} \sum_{k,l=1}^{n} g^{kl} \frac{\partial W}{\partial q_k} \frac{\partial W}{\partial q_l} + V = E, \tag{6}$$

which is a nonlinear first order PDE. Following earlier studies by Liouville and Jacobi, in 1891 Paul Stäckel, investigating the question of finding a complete solution of this equation of the form

$$W(q_1 \cdots q_n) = \sum_{i=1}^{n} W_i(q_i), \tag{7}$$

with each W_i depending on a single variable only, gave

a celebrated solution for orthogonal coordinate systems,[2] which is frequently quoted and has found its way into some textbooks.[3-5] Oddly enough, other work by Stäckel himself,[6] which gave solutions for nonorthogonal systems in two dimensions, and its further developments have been completely ignored in the recent literature, in spite of the existence of a very detailed review article.[7] Generalizing Stäckel's results, Levi-Civita[8] established the necessary and sufficient conditions to be satisfied by H and the corresponding metric if Eq. (6) was to be separable, and deduced the existence of $n + 1$ distinct types of such metrics. Later, Burgatti[9] gave forms of the separated ODE's sufficient to lead to these $n + 1$ types, and subsequently Dall' Acqua[10] proved that these forms were also necessary.

Although the general form of the separated equations has been known for some time, the general form of the metrics following from these has not been given before. In Sec. II we shall establish this form, after briefly summarizing those results of Levi-Civita, Burgatti, and Dall' Acqua which are needed for this purpose. In Sec. III it is shown that the same metrics allow separation of the Schrödinger, Helmholtz, and Laplace equation, subject only to a condition on the determinant of the metric. The results are briefly discussed in Sec. IV.

II. THE HAMILTON–JACOBI EQUATION

Following Levi-Civita and Dall' Acqua, we divide the variables (and their indices) into two groups

$$q_\alpha, \quad \alpha = 1 \cdots \nu, \tag{8a}$$

$$q_\rho, \quad \rho = \nu + 1 \cdots n \tag{8b}$$

(defining $\nu = 0$ if the first group is absent), to be called variables (and indices) of the first and the second kind, respectively. The division into these two groups arises from the role of the variables in the conditions on H given by Levi-Civita, but there is no need to discuss this point here.

In the following, letters of the Greek alphabet up to and including ν will be used for indices of the first kind, and after ν for those of the second kind. Latin indices will be used if the entire range $1 \cdots n$ is covered. To simplify the notation, we shall indicate the range of any summation only by $\sum^I$ for variables of the first kind, by $\sum^{II}$ for those of the second kind, and by $\sum$ for those of both kinds. We assume $n > 1$, as otherwise the problem of separation does not exist.

Corresponding to the coordinates of the first and second kind, respectively, we introduce two sets of arbitrary continuous functions, each depending on a single variable only,

$$\varphi_{\alpha\beta}(q_\alpha), \quad \alpha, \beta = 1 \cdots \nu, \tag{9a}$$

$$\varphi_{\rho\sigma}(q_\rho), \quad \rho, \sigma = \nu + 1 \cdots n, \tag{9b}$$

with nonvanishing determinants ϕ_I and ϕ_{II}, respectively. The cofactors of $\varphi_{\alpha\beta}$ and $\varphi_{\rho\sigma}$ in their respective determinants are denoted by $\phi_{\alpha\beta}$ and $\phi_{\rho\sigma}$, respectively; note that they do *not* depend on q_α and q_ρ, 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

$$\sum_\alpha^I \varphi_{\alpha\beta}\phi_{\alpha\kappa} = \sum_\alpha^I \varphi_{\beta\alpha}\phi_{\kappa\alpha} = \phi_I \delta_\beta{}^\kappa, \tag{10a}$$

$$\sum_\rho^{II} \varphi_{\rho\sigma}\phi_{\rho\tau} = \sum_\rho^{II} \varphi_{\sigma\rho}\phi_{\tau\rho} = \phi_{II} \delta_\rho{}^\tau, \tag{10b}$$

without explicitly referring to them.

We also introduce three sets of arbitrary functions

$$f_\rho{}^\kappa(q_\rho), \quad F_\rho{}^{\kappa\lambda}(q_\rho) = F_\rho{}^{\lambda\kappa}, \quad u_\rho(q_\rho)$$

$$\kappa, \lambda = 1 \cdots \nu, \quad \rho = \nu + 1 \cdots n, \tag{11}$$

each depending on a single variable only and where the $f_\rho{}^K$'s and $F_\rho{}^{K\lambda}$'s must be continuous, and two sets of arbitrary constants ("separation constants")

$$c_\alpha, \quad \alpha = 1 \cdots \nu, \tag{12a}$$

$$c_\rho, \quad \rho = \nu + 1 \cdots n. \tag{12b}$$

The separated ODE's then are

$$\frac{dW_\alpha}{dq_\alpha} = \sum_\beta^I \varphi_{\alpha\beta}(q_\alpha) c_\beta, \quad \alpha = 1 \cdots \nu, \tag{13a}$$

$$\frac{dW_\rho}{dq_\rho} = \sum_\beta^I f_\rho{}^\beta(q_\rho) c_\beta$$

$$\pm \left(\sum_{\kappa,\lambda}^I F_\rho{}^{\kappa\lambda}(q_\rho) c_\kappa c_\lambda \right.$$

$$\left. + \sum_\sigma^{II} \varphi_{\rho\sigma}(q_\rho) c_\sigma - 2u_\rho(q_\rho) \right)^{1/2}, \quad \rho = \nu + 1 \cdots n. \tag{13b}$$

Corresponding to the $n + 1$ possible values of ν, there are thus $n + 1$ different types of equations T_ν. These equations were first obtained by Burgatti,[9] who showed that they are sufficient to allow solutions of the form (7); the proof that they are also necessary was given by Dall' Acqua.[10] The type T_0 is that discovered by Stäckel[2] and T_n is that found (though not solved) by Levi-Civita.[8] In the latter case, $V = 0$ (apart from a trivial additive constant which here and in the other cases is assumed to be absorbed in E); the mechanical problem is that of free particles, and the trajectories following from (1) can be considered as a geodesic in V_n. This type was called "essentially geodesic" by Levi-Civita to distinguish it from the other types which are separable even in the presence of potentials of the form (3); as shown by Dall'Acqua, there are no cases in which the presence of a potential is necessary to achieve separation.

We shall now derive the forms of g^{kl} and V following from (13) by eliminating the separation constants from Eqs. (13) and comparing the resulting equation with Eq. (6).

Multiplying the set (13a) by $\phi_{\alpha\kappa}/\phi_I$ and summing over α, we obtain

$$\sum_\alpha^I \frac{\phi_{\alpha\kappa}}{\phi_I} \frac{dW_\alpha}{dq_\alpha} = c_\kappa. \tag{14a}$$

Substituting this into (13b), rearranging terms, and squaring, we get

$$\left(\frac{dW_\rho}{dq_\rho} - \sum_{\alpha,\beta}^{\mathrm{I}} \frac{f_\rho{}^\beta \phi_{\alpha\beta}}{\phi_{\mathrm{I}}} \frac{dW_\alpha}{dq_\alpha}\right)^2$$

$$= \sum_{\alpha,\beta,\kappa,\lambda}^{\mathrm{I}} \frac{F_\rho{}^{\kappa\lambda} \phi_{\alpha\kappa} \phi_{\beta\lambda}}{\phi_{\mathrm{I}}^2} \frac{dW_\alpha}{dq_\alpha} \frac{dW_\beta}{dq_\beta}$$

$$+ \sum_\sigma^{\mathrm{II}} \varphi_{\rho\sigma}(q_\rho) c_\sigma - 2u_\rho(q_\rho). \tag{14b}$$

Squaring (14a) and changing dummy indices, we obtain

$$\sum_{\alpha,\beta}^{\mathrm{I}} \frac{\phi_{\alpha\gamma}\phi_{\beta\gamma}}{\phi_{\mathrm{I}}^2} \frac{dW_\alpha}{dq_\alpha} \frac{dW_\beta}{dq_\beta} = c_\gamma^2. \tag{15a}$$

Multiplying (14b) by $\phi_{\rho\tau}/\phi_{\mathrm{II}}$, summing over ρ, and rearranging, we get

$$\sum_\rho^{\mathrm{II}} \frac{\phi_{\rho\tau}}{\phi_{\mathrm{II}}} \left(\frac{dW_\rho}{dq_\rho} - \sum_{\alpha,\beta}^{\mathrm{I}} \frac{f_\rho{}^\beta \phi_{\alpha\beta}}{\phi_{\mathrm{I}}} \frac{dW_\alpha}{dq_\alpha}\right)^2$$

$$- \sum_\rho^{\mathrm{II}} \sum_{\alpha,\beta,\kappa,\lambda}^{\mathrm{I}} \frac{F_\rho{}^{\kappa\lambda} \phi_{\alpha\kappa} \phi_{\beta\lambda} \phi_{\rho\tau}}{\phi_{\mathrm{I}}^2 \phi_{\mathrm{II}}} \frac{dW_\alpha}{dq_\alpha} \frac{dW_\beta}{dq_\beta}$$

$$+ \sum_\rho^{\mathrm{II}} \frac{2u_\rho \phi_{\rho\tau}}{\phi_{\mathrm{II}}} = c_\tau. \tag{15b}$$

We now sum Eqs. (15a) and (15b) over γ and τ, respectively, add the resulting equations, and put

$$\sum_\gamma^{\mathrm{I}} c_\gamma^2 + \sum_\tau^{\mathrm{II}} c_\tau = 2E. \tag{16}$$

Then we have

$$\sum_{\alpha,\beta,\gamma}^{\mathrm{I}} \frac{\phi_{\alpha\gamma}\phi_{\beta\gamma}}{\phi_{\mathrm{I}}^2} \frac{dW_\alpha}{dq_\alpha} \frac{dW_\beta}{dq_\beta}$$

$$+ \sum_{\rho,\tau}^{\mathrm{II}} \frac{\phi_{\rho\tau}}{\phi_{\mathrm{II}}} \left(\frac{dW_\rho}{dq_\rho} - \sum_{\alpha,\beta}^{\mathrm{I}} \frac{f_\rho{}^\beta \phi_{\alpha\beta}}{\phi_{\mathrm{I}}} \frac{dW_\alpha}{dq_\alpha}\right)^2$$

$$- \sum_{\rho,\tau}^{\mathrm{II}} \sum_{\alpha,\beta,\kappa,\lambda}^{\mathrm{I}} \frac{F_\rho{}^{\kappa\lambda} \phi_{\alpha\kappa}\phi_{\beta\lambda}\phi_{\rho\tau}}{\phi_{\mathrm{I}}^2 \phi_{\mathrm{II}}} \frac{dW_\alpha}{dq_\alpha} \frac{dW_\beta}{dq_\beta}$$

$$+ \sum_{\rho,\tau}^{\mathrm{II}} \frac{2u_\rho\phi_{\rho\tau}}{\phi_{\mathrm{II}}} = 2E. \tag{17}$$

We first consider the case $\nu = n$, i.e., the type T_n. Then Eqs. (16) and (17) reduce to

$$\sum_\gamma c_\gamma^2 = 2E \tag{18}$$

and

$$\sum_{\alpha,\beta,\gamma} \frac{\phi_{\alpha\gamma}\phi_{\beta\gamma}}{\phi_{\mathrm{I}}^2} \frac{dW_\alpha}{dq_\alpha} \frac{dW_\beta}{dq_\beta} = 2E. \tag{19}$$

Comparison with (3) shows that we must have $V = 0$ and

$$g^{\alpha\beta} = \sum_\gamma \frac{\phi_{\alpha\gamma}\phi_{\beta\gamma}}{\phi_{\mathrm{I}}^2}. \tag{20}$$

This is equivalent to

$$g_{\alpha\beta} = \sum_\gamma \varphi_{\alpha\gamma}(q_\alpha) \varphi_{\beta\gamma}(q_\beta) \tag{21}$$

[as can be readily verified from Eqs. (10a) and the

definition of g^{kl} given in Eq. (3)], a result found explicitly for $n = 3$ by a different method by Dall'Acqua.[11] Furthermore, it was shown earlier by Levi-Civita[8] that for any solution of the essentially geodesic case the Riemann—Christoffel curvature tensor vanishes. Therefore, all metrics (21) are those of flat (Euclidean or pseudo-Euclidean) space S_n. As also shown by Levi-Civita, in the Euclidean case the metrics are all related to their Cartesian form (with Cartesian coordinates y_α) by transformations of the form

$$y_\alpha = \sum_\beta Q_{\beta\alpha}(q_\beta), \qquad Q_{\beta\alpha} = \int \varphi_{\beta\alpha} \, dq_\beta. \tag{22}$$

The extension of his argument to the pseudo-Euclidean case is immediate.

The form of W follows from integration of Eq. (13a); because of Eq. (6), constants of integration can be omitted.

Summarizing our results we thus have

Theorem I: The Hamilton—Jacobi equation (6) (with $V = 0$) can be solved by the method of separation of variables in any S_n ($n > 1$) which has a metric of the form

$$g_{kl} = \sum_{j=1}^n \varphi_{kj}(q_k) \varphi_{lj}(q_l),$$

where the $\varphi_{ij}(q_i)$ are n^2 arbitrary continuous real functions, of a single variable each, with nonvanishing determinant. The solution is given by

$$W = \sum_{i,j=1}^n c_j \int \varphi_{ij}(q_i) \, dq_i,$$

where the c_j are arbitrary real constants, subject only to the condition

$$\sum_{j=1}^n c_j^2 = 2E.$$

Now we consider the general case. The form of g^{kl} follows immediately from a comparison of Eqs. (17) and (6), and that of W from integration of Eqs. (13); integration constants can again be omitted. Furthermore, since the metric tensor must be continuous, all arbitrary functions entering it must be required to be continuous. Thus we have

Theorem II: The Hamilton—Jacobi equation (6) can be solved by the method of separation of variables in any V_n ($n > 1$) whose contravariant metric tensor is of the form

$$g^{\alpha\beta} = \sum_{\gamma=1}^\nu \frac{\phi_{\alpha\gamma}\phi_{\beta\gamma}}{\phi_{\mathrm{I}}^2} - \sum_{\kappa,\lambda=1}^\nu \sum_{\sigma,\tau=\nu+1}^n \frac{(F_\sigma{}^{\kappa\lambda} - f_\sigma{}^\kappa f_\sigma{}^\lambda)\phi_{\alpha\kappa}\phi_{\beta\lambda}\phi_{\sigma\tau}}{\phi_{\mathrm{I}}^2 \phi_{\mathrm{II}}},$$

$$g^{\alpha\rho} = -2 \sum_{\gamma=1}^\nu \frac{f_\rho{}^\gamma \phi_{\alpha\gamma}}{\phi_{\mathrm{I}} \phi_{\mathrm{II}}} \sum_{\tau=\nu+1}^n \phi_{\rho\tau},$$

$$g^{\rho\rho} = \frac{1}{\phi_{\mathrm{II}}} \sum_{\tau=\nu+1}^n \phi_{\rho\tau},$$

$$g^{\rho\sigma} = 0, \qquad \rho \neq \sigma,$$

$$\alpha, \beta = 1 \cdots \nu, \qquad 0 \leq \nu < n,$$

$$\rho, \sigma = \nu + 1 \cdots n,$$

where $F_\rho{}^{\kappa\lambda}(q_\rho) = F_\rho{}^{\lambda\kappa}$ and $f_\rho{}^\gamma(q_\rho)$ are arbitrary continuous real functions of a single variable 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 $\varphi_{\rho\sigma}(q_\rho)$ $(\rho, \sigma = \nu + 1 \cdots n)$ respectively, and $\phi_{\alpha\beta}$ and $\phi_{\rho\sigma}$ are the cofactors of $\varphi_{\alpha\beta}$ and $\varphi_{\rho\sigma}$ in these determinants (with $\phi_{ij} \equiv 1$ if one or both determinants consist of a single element φ_{ii} only); the potential energy V must be of the form

$$V(q_{\nu+1} \cdots q_n) = \sum_{\rho, \tau = \nu+1}^{n} \frac{u_\rho(q_\rho)\phi_{\rho\tau}}{\phi_{\text{II}}},$$

where the $u_\rho(q_\rho)$ are arbitrary real functions of a single variable each. The solution is given by

$$W = \sum_{\alpha, \beta = 1}^{\nu} c_\beta \int \varphi_{\alpha\beta}(q_\alpha)\, dq_\alpha$$

$$+ \sum_{\rho = \nu+1}^{n} \int \left[\sum_{\beta=1}^{\nu} f_\rho^{\beta}(q_\rho) c_\beta \right.$$

$$\pm \left(\sum_{\kappa, \lambda = 1}^{\nu} F_\rho^{\kappa\lambda}(q_\rho)\, c_\kappa c_\lambda \right.$$

$$\left. \left. + \sum_{\sigma = \nu+1}^{n} \varphi_{\rho\sigma}(q_\sigma) c_\sigma - 2u_\rho(q_\rho) \right)^{1/2} \right] dq_\rho,$$

where the c_i $(i = 1 \cdots n)$ are arbitrary real constants, subject to the condition

$$\sum_{\gamma=1}^{\nu} c_\gamma^2 + \sum_{\tau=\nu+1}^{n} c_\tau = 2E.$$

III. THE SCHRÖDINGER AND RELATED EQUATIONS

In addition to the H—J equations (4), there exists another way to associate a PDE with the Hamiltonian (3), the Schrödinger equation

$$i\hbar \frac{\partial \Psi}{\partial t} + H(q_1 \cdots q_n, p_1 \cdots p_n)\Psi = 0,$$

$$p_i \equiv \frac{\hbar}{i} \frac{\partial}{\partial q_i}, \tag{23}$$

where $\hbar$ is Planck's constant divided by 2π. In the following we shall use units such that $\hbar = 1$. As is well known from the consideration of the Laplacian, the part of $H\Psi$ involving the p_i's must be taken as

$$T(q_1 \cdots q_n, p_1 \cdots p_n)\Psi = -\frac{1}{2} \sum_{k,l=1}^{n} \frac{1}{g^{1/2}} \frac{\partial}{\partial q_k} \left(g^{1/2} g^{kl} \frac{\partial \Psi}{\partial q_l} \right),$$

$$g \equiv |\det g_{kl}|, \tag{24}$$

(using Beltrami's second differential parameter, whereas the H—J equation involves the first one). If we assume

$$\Psi(t, q_1 \cdots q_n) = \exp(iEt)\, \psi(q_1 \cdots q_n), \tag{25}$$

where E is a constant, Eq. (23) reduces to

$$H(q_1 \cdots q_n, p_1 \cdots p_n)\psi = E\psi, \tag{26}$$

or explicitly, using (3) and (24),

$$-\frac{1}{2} \sum_{k,l=1}^{n} \frac{1}{g^{1/2}} \frac{\partial}{\partial q_k} \left(g^{1/2} g^{kl} \frac{\partial \psi}{\partial q_l} \right)$$

$$+ V(q_1 \cdots q_n)\psi = E\psi. \tag{27}$$

Obviously the Helmholtz and Laplace equations are special cases of this PDE, the time-independent Schrödinger equation, although they are not associated with mechanical systems described by Hamiltonian functions. Furthermore, regardless of the physical origin of Eq. (27), we shall in the following *not* restrict ourselves to spaces with a positive definite metric. Thus, e.g., Eq. (27), with $V = E = 0$, and an n-dimensional metric tensor of signature $n - 2$, includes the "$(n-1)$-dimensional" wave equation, i.e., the wave equation for $n - 1$ spatial dimensions. Another way of including this equation is by the usual procedure of separating off a time-dependent factor; then we obtain the Helmholtz equation, i.e., Eq. (27) with $V = 0$ and a positive definite metric. Similarly, the Helmholtz equation results from separating off a time-dependent factor in the diffusion or heta conduction equation, or in the damped wave equation.

We shall be concerned with the problem of finding coordinate systems and potentials which allow solutions of Eq. (27) of the form

$$\psi(q_1 \cdots q_n) = \prod_{i=1}^{n} \psi_i(q_i) \tag{28}$$

of products of functions ψ_i of a single variable, which can be found by solving an ODE. In general, the coordinate systems for which such a separation is possible will allow also a corresponding separation for the other PDE's mentioned; exception will be noted later.

Equation (6) was of first order and nonlinear, whereas Eq. (29) is linear, but of second order. In spite of this difference, Eq. (27) can be separated for spaces with metrics of the same form as those found for Eq. (6), with only one additional condition. Separation may well be possible for other forms of the metrics, but cannot be obtained by the method used here, which is patterned after that used for Eq. (6). It should also be noted that because of the linearity of Eq. (27), solutions obtained by separation of variables can be superimposed linearly, in contrast to the corresponding solutions of the nonlinear Eq. (6).

We divide the indices into two groups as before. Instead of Eqs. (13) we consider

$$\frac{1}{i} \frac{d\psi_\alpha}{dq_\alpha} = \sum_\beta{}^{\text{I}} \varphi_{\alpha\beta}(q_\alpha) c_\beta \psi_\alpha, \tag{29a}$$

$$O_\rho^{(2)}[\psi_\rho] = \sum_\beta{}^{\text{I}} \frac{1}{i} \left(O_\rho^{(1)}[f_\rho^{\beta}(q_\rho)c_\beta\psi_\rho] + f_\rho^{\beta} c_\beta \frac{d\psi_\rho}{dq_\rho} \right)$$

$$+ \left(\sum_{\kappa,\lambda}{}^{\text{I}} F_\rho^{\kappa\lambda}(q_\rho)c_\kappa c_\lambda + \sum_\sigma{}^{\text{II}} \varphi_{\rho\sigma}(q_\rho)c_\sigma - 2u_\rho(q_\rho) \right)\psi_\rho, \tag{29b}$$

where φ_{ij}, $F_\rho^{\kappa\lambda}$, f_ρ^{β}, u_ρ, and c_i have the same significance as before, and $O_\rho^{(1)}$ and $O_\rho^{(2)}$ are operators involving functions of q_ρ and derivatives with respect to q_ρ only (of first and second order, respectively), whose form will be determined later. Because of the way they enter the ODE (29a), the φ_{ij}'s and f_ρ^{β}'s now must be required to be of class C^1.

We shall now determine the conditions under which the system of ODE's (29) is equivalent to the PDE (27). In analogy with the procedure used for the H—J equation,

this is achieved by eliminating the separation constants from Eqs. (29) and comparing the resultant equation with Eq. (27).

Multiplication of the set (29a) by ψ/ψ_α yields

$$\frac{\partial \psi}{\partial q_\alpha} = i \sum_\beta^{\mathrm{I}} \varphi_{\alpha\beta}(q_\alpha) c_\beta \psi. \tag{30}$$

Multiplying this by $\phi_{\alpha\kappa}/\phi_{\mathrm{I}}$ and summing over α, we obtain

$$\sum_\alpha^{\mathrm{I}} \frac{\phi_{\alpha\kappa}}{\phi_{\mathrm{I}}} \frac{\partial \psi}{\partial q_\alpha} = i c_\kappa \psi. \tag{31}$$

Differentiating this with respect to q_ρ gives

$$\sum_\alpha^{\mathrm{I}} \frac{\phi_{\alpha\kappa}}{\phi_{\mathrm{I}}} \frac{\partial^2 \psi}{\partial q_\alpha \partial q_\rho} = i c_\kappa \frac{\partial \psi}{\partial q_\rho} . \tag{32}$$

Differentiating Eq. (31) with respect to q_γ instead, and using Eq. (30) in the resulting expression, we get

$$\sum_\alpha^{\mathrm{I}} \frac{\partial}{\partial q_\gamma} \left(\frac{\phi_{\alpha\kappa}}{\phi_{\mathrm{I}}} \frac{\partial \psi}{\partial q_\alpha} \right) = - c_\kappa \sum_\beta^{\mathrm{I}} \varphi_{\gamma\beta}(q_\gamma) c_\beta \psi. \tag{33}$$

Multiplying this by $- \phi_{\gamma\lambda}/\phi_{\mathrm{I}}$ and summing over γ, we obtain

$$- \sum_{\alpha,\gamma}^{\mathrm{I}} \frac{\phi_{\gamma\lambda}}{\phi_{\mathrm{I}}} \frac{\partial}{\partial q_\gamma} \left(\frac{\phi_{\alpha\kappa}}{\phi_{\mathrm{I}}} \frac{\partial \psi}{\partial q_\alpha} \right) = c_\kappa c_\lambda \psi. \tag{34}$$

Putting $\kappa = \lambda$ in (34), we get, changing dummy indices,

$$- \sum_{\alpha,\beta}^{\mathrm{I}} \frac{1}{\phi_{\mathrm{I}}} \frac{\partial}{\partial q_\alpha} \left(\frac{\phi_{\alpha\gamma}\phi_{\beta\gamma}}{\phi_{\mathrm{I}}} \frac{\partial \psi}{\partial q_\beta} \right) = c_\gamma^2 \psi. \tag{35}$$

We now multiply the set (29b) by ψ/ψ_ρ and eliminate the c_κ's everywhere except the last sum in the resulting expression by means of Eqs. (31)—(34). Changing dummy indices, we get

$$O_\rho^{(2)}[\psi] = - \sum_{\beta,\gamma}^{\mathrm{I}} \left(O_\rho^{(1)} \left[\frac{f_\rho{}^\beta \phi_{\gamma\beta}}{\phi_{\mathrm{I}}} \frac{\partial \psi}{\partial q_\gamma} \right] \right.$$
$$+ f_\rho{}^\beta \frac{\phi_{\gamma\beta}}{\phi_{\mathrm{I}}} \frac{\partial^2 \psi}{\partial q_\gamma \partial q_\rho} \Big)$$
$$- \sum_{\beta,\gamma,\kappa,\lambda}^{\mathrm{I}} F_\rho{}^{\kappa\lambda} \frac{\phi_{\gamma\lambda}}{\phi_{\mathrm{I}}} \frac{\partial}{\partial q_\gamma} \left(\frac{\phi_{\beta\kappa}}{\phi_{\mathrm{I}}} \frac{\partial \psi}{\partial q_\beta} \right)$$
$$+ \sum_\sigma^{\mathrm{II}} \varphi_{\rho\sigma} c_\sigma \psi - 2 u_\rho \psi. \tag{36}$$

Multiplying this by $\phi_{\rho\tau}/\phi_{\mathrm{II}}$, summing over ρ, and rearranging, we obtain

$$\sum_\rho^{\mathrm{II}} \frac{\phi_{\rho\tau}}{\phi_{\mathrm{II}}} O_\rho^{(2)}[\psi] + \sum_\rho^{\mathrm{II}} \sum_{\beta,\gamma}^{\mathrm{I}} \frac{\phi_{\rho\tau}}{\phi_{\mathrm{II}}}$$
$$\times \left(O_\rho^{(1)} \left[\frac{f_\rho{}^\beta \phi_{\gamma\beta}}{\phi_{\mathrm{I}}} \frac{\partial \psi}{\partial q_\gamma} \right] + f_\rho{}^\beta \frac{\phi_{\gamma\beta}}{\phi_{\mathrm{I}}} \frac{\partial^2 \psi}{\partial q_\gamma \partial q_\rho} \right)$$
$$+ \sum_\rho^{\mathrm{II}} \sum_{\beta,\gamma,\kappa,\lambda}^{\mathrm{I}} \frac{\phi_{\rho\tau}}{\phi_{\mathrm{II}}} \frac{F_\rho{}^{\kappa\lambda}\phi_{\tau\lambda}}{\phi_{\mathrm{I}}} \frac{\partial}{\partial q_\gamma} \left(\frac{\phi_{\beta\kappa}}{\phi_{\mathrm{I}}} \frac{\partial \psi}{\partial q_\beta} \right)$$
$$+ 2 \sum_\rho^{\mathrm{II}} \frac{\phi_{\rho\tau}}{\phi_{\mathrm{II}}} u_\rho \psi = c_\tau \psi. \tag{37}$$

We now add Eqs. (35) and (37), summing over γ and τ, respectively, change dummy indices, and use the definition (16) to obtain

$$- \sum_{\alpha,\beta,\gamma}^{\mathrm{I}} \frac{1}{\phi_{\mathrm{I}}} \frac{\partial}{\partial q_\alpha} \left(\frac{\phi_{\alpha\gamma}\phi_{\beta\gamma}}{\phi_{\mathrm{I}}} \frac{\partial \psi}{\partial q_\beta} \right)$$
$$+ \sum_{\rho,\tau}^{\mathrm{II}} \frac{\phi_{\rho\tau}}{\phi_{\mathrm{II}}} O_\rho^{(2)}[\psi] + \sum_{\rho,\tau}^{\mathrm{II}} \sum_{\alpha,\beta}^{\mathrm{I}} \frac{\phi_{\rho\tau}}{\phi_{\mathrm{II}}}$$
$$\times \left(O_\rho^{(1)} \left[\frac{f_\rho{}^\alpha \phi_{\beta\alpha}}{\phi_{\mathrm{I}}} \frac{\partial \psi}{\partial q_\beta} \right] + \frac{f_\rho{}^\alpha \phi_{\beta\alpha}}{\phi_{\mathrm{I}}} \frac{\partial^2 \psi}{\partial q_\beta \partial q_\rho} \right)$$
$$+ \sum_{\rho,\tau}^{\mathrm{II}} \sum_{\alpha,\beta,\kappa,\lambda}^{\mathrm{I}} \frac{F_\rho{}^{\kappa\lambda} \phi_{\alpha\lambda}\phi_{\rho\tau}}{\phi_{\mathrm{I}}\phi_{\mathrm{II}}} \frac{\partial}{\partial q_\alpha} \left(\frac{\phi_{\beta\kappa}}{\phi_{\mathrm{I}}} \frac{\partial \psi}{\partial q_\beta} \right)$$
$$+ 2 \sum_{\rho,\tau}^{\mathrm{II}} \frac{\phi_{\rho\tau}}{\phi_{\mathrm{II}}} u_\rho \psi = 2E\psi. \tag{38}$$

We again first consider the case T_n. Then Eqs. (16) and (38) reduce to (18) and

$$- \sum_{\alpha,\beta,\gamma}^{\mathrm{I}} \frac{1}{\phi_{\mathrm{I}}} \frac{\partial}{\partial q_\alpha} \left(\frac{\phi_{\alpha\gamma}\phi_{\beta\gamma}}{\phi_{\mathrm{I}}} \frac{\partial \psi}{\partial q_\beta} \right) = 2E\psi. \tag{39}$$

Comparison with Eq. (27) shows that we again have the form (20) for $g^{\alpha\beta}$ and thus of (21) for $g_{\alpha\beta}$, but that the additional condition

$$g^{1/2} = K \phi_{\mathrm{I}} \tag{40}$$

must hold, where K is a constant. Taking determinants on both sides of (21), we get $g = \phi_{\mathrm{I}}^2$. Thus Eq. (40) is automatically satisfied (with $K = 1$), and no additional restriction is imposed compared to the Hamilton—Jacobi case.

We note that Eq. (39) does not contain a term corresponding to V (and thus for positive E is just the Helmholtz equation). It can easily be seen that we cannot obtain such a term by inclusion of additional functions on the rhs of Eqs. (29a), as these would lead to terms in Eq. (38) which have no counterpart in (27).

Equations (29a) 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 (29a), all constants c_i must be real; but then from (18) E cannot be zero (except for the trivial case of a constant ψ). Therefore, our method does not provide us directly with nontrivial real solutions of Laplace's equation. On the other hand, we can construct real solutions by taking a set of Eqs. (29a) with complex separation constants to satisfy Eq. (18). These will result in ψ's which are complex; however, we can also start from the complex conjugate of our original set (29a) 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 Eq. (27). Furthermore, in the case of the H—J equation the separation constants were unrestricted, apart from Eq. (18); now, however, restrictions may arise through boundary conditions on ψ.

Thus we have

Theorem III: The Schrödinger equation (with $V = 0$), the Helmholtz, and the Laplace equation can be solved by the method of separation of variables in any S_n ($n > 1$) which has a metric of the form given in Theorem I. The particular solutions obtained by this method are of the form

$$\psi = A \exp i \sum_{i,j=1}^{n} c_j \int \varphi_{ij}(q_i)\, dq_i$$

(+ c. c. if a *real* solution is desired),

where the φ_{ij} are arbitrary functions of class C^1 and A and the c_j are arbitrary complex constants, subject only to the condition

$$\sum_{j=1}^{n} c_j^2 = \begin{cases} 2E & \text{(Schrödinger or Helmholtz equation)} \\ 0 & \text{(Laplace's equation).} \end{cases}$$

and to restrictions due to boundary conditions on ψ. The solutions for different values of the c_j's satisfying these conditions can be linearly superimposed.

Now we consider the general case. To determine the possible forms of g^{kl}, we first consider the terms involving $g^{\rho\sigma}$. These are all contained in the first sum of Eq. (37). By comparison with Eq. (27) we must have

$$O_\rho^{(2)}[\psi_\rho] = -\frac{d}{dq_\rho}\left(f_\rho(q_\rho)\frac{d\psi_\rho}{dq_\rho}\right), \tag{41}$$

where $f_\rho(q_\rho)$ is a function of class C^1 to be discussed later and the sign is chosen for convenience of comparison with the H–J equation. We also get $g^{\rho\sigma} = 0$ ($\rho \neq \sigma$) as in the case of the H–J equation. Furthermore, we must have

$$g^{1/2}g^{\rho\rho} = F(q_1 \cdots q_\nu)\, f_\rho(q_\rho)\sum_\tau^{\mathrm{II}}\phi_{\rho\tau},$$
$$g^{1/2} = \phi_{\mathrm{II}}\, F(q_1 \cdots q_\nu), \tag{42}$$

where F cannot depend on any of the variables of the second kind, since it must be independent of the value of ρ, and remains to be further specified.

Now we can introduce new sets of functions

$$\varphi'_{\rho\sigma}(q_\rho) \equiv \frac{\varphi_{\rho\sigma}(q_\rho)}{f_\rho(q_\rho)}, \quad f_\rho'^{\alpha}(q_\rho) \equiv \frac{f_\rho^{\alpha}(q_\rho)}{f_\rho(q_\rho)},$$
$$F_\rho'^{\kappa\lambda}(q_\rho) \equiv \frac{F_\rho^{\kappa\lambda}(q_\rho)}{f_\rho(q_\rho)}, \quad u_\rho'(q_\rho) \equiv \frac{u_\rho(q_\rho)}{f_\rho(q_\rho)}, \tag{43}$$

from which we get

$$\phi'_{\rho\sigma} = \frac{f_\rho \phi_{\rho\sigma}}{P}, \quad \phi'_{\mathrm{II}} = \frac{\phi_{\mathrm{II}}}{P},$$
$$P \equiv \prod_{\sigma=\nu+1}^{n} f_\sigma(q_\sigma), \tag{44}$$

where $\phi'_{\rho\sigma}$ is the cofactor of $\varphi_{\rho\sigma}'$ (which again does not depend on q_ρ) and ϕ'_{II} is the new determinant of the $\varphi_{\rho\sigma}'$'s. Then we get from Eqs. (42) and (44)

$$g^{1/2} = \phi'_{\mathrm{II}}\, F(q_1 \cdots q_\nu) P. \tag{45}$$

We now rewrite Eq. (38), taking into account the dependence of the various ϕ_{ij}'s on their argument as well as the equations and definitions (41), (43), and (44). Since the original functions (9) and (11) were arbitrary,

we can use the new quantities (43) and (44) *without* the prime without loss of generality; however, this definition must be taken into account in Eq. (29b). Thus, Eq. (37) becomes

$$-\sum_{\alpha,\beta,\gamma}^{\mathrm{I}}\frac{1}{\phi_{\mathrm{I}}}\frac{\partial}{\partial q_\alpha}\left(\frac{\phi_{\alpha\gamma}\phi_{\beta\gamma}}{\phi_{\mathrm{II}}}\frac{\partial\psi}{\partial q_\beta}\right)$$

$$-\sum_{\rho,\tau}^{\mathrm{II}}\frac{\phi_{\rho\tau}}{\phi_{\mathrm{II}}}\frac{1}{f_\rho}\frac{\partial}{\partial q_\rho}\left(f_\rho\frac{\partial\psi}{\partial q_\rho}\right)$$

$$+\sum_{\rho,\tau}^{\mathrm{II}}\sum_{\alpha,\beta}^{\mathrm{I}}\frac{\phi_{\rho\tau}}{\phi_{\mathrm{II}}}\frac{1}{f_\rho} \tag{46}$$

$$\times\left(O_\rho^{(1)}\left[\frac{f_\rho f_\rho^{\alpha}\phi_{\beta\alpha}}{\phi_{\mathrm{I}}}\frac{\partial\psi}{\partial q_\beta}\right]+\frac{f_\rho f_\rho^{\alpha}\phi_{\beta\alpha}}{\phi_{\mathrm{I}}}\frac{\partial^2\psi}{\partial q_\beta\partial q_\rho}\right)$$

$$+\sum_{\rho,\tau}^{\mathrm{II}}\sum_{\alpha,\beta,\kappa,\lambda}^{\mathrm{I}}\frac{F_\rho^{\kappa\lambda}\phi_{\alpha\lambda}\phi_{\rho\tau}}{\phi_{\mathrm{I}}\phi_{\mathrm{II}}}\frac{\partial}{\partial q_\alpha}\left(\frac{\phi_{\beta\kappa}}{\phi_{\mathrm{I}}}\frac{\partial\psi}{\partial q_\beta}\right)$$

$$+2\sum_{\rho,\tau}^{\mathrm{II}}\frac{\phi_{\rho\tau}}{\phi_{\mathrm{II}}}u_\rho\psi = 2E\psi.$$

Comparing the first and third sum with Eq. (27), and using Eq. (45), we see that $F(q_1 \cdots q_\nu)$ must be proportional to ϕ_{I}; because of the presence of the factor P the constant of proportionality can be put equal to one without loss of generality. Therefore, we have the condition

$$g^{1/2} = \phi_{\mathrm{I}}\,\phi_{\mathrm{II}}\prod_{\sigma=\nu+1}^{n} f_\sigma(q_\sigma), \tag{47}$$

i.e., only those metrics are allowed for which f_σ's can be found such that (47) holds.

Comparison of the second sum in Eq. (46) with Eq. (27) shows that it should correspond to

$$\sum_\rho^{\mathrm{II}}\sum_\beta^{\mathrm{I}}\frac{1}{g^{1/2}}\left[\frac{\partial}{\partial q_\rho}\left(g^{1/2}g^{\rho\beta}\frac{\partial\psi}{\partial q_\beta}\right)+\frac{\partial}{\partial q_\beta}\left(g^{1/2}g^{\rho\beta}\frac{\partial\psi}{\partial q_\rho}\right)\right].$$
$$\tag{48}$$

This can be achieved by taking

$$O_\rho^{(1)}[f_\rho^{\beta}c_\beta\psi_\rho] = \frac{d}{dq_\rho}(f_\rho^{\beta}c_\beta\psi_\rho), \tag{49}$$

since the factor of $\partial^2\psi/\partial q_\rho\partial q_\beta$ in that sum is independent of q_β. No further conditions are introduced, and the various components of the metric can be identified as being of the form given in Theorem II.

The solutions of Eq. (29) will in general be complex and remarks similar to those made for Theorem III apply. Thus we have [taking account of Eq. (47) and of the effect of redefinitions (43) (with the prime omitted) on Eq. (29b)]:

Theorem IV: The Schrödinger, Helmholtz, and Laplace equations can be solved by the method of separation of variables in any V_n ($n > 1$) whose metric is of the form given in Theorem II and whose determinant is of the form

$$g = \phi_{\mathrm{I}}^2\,\phi_{\mathrm{II}}^2\prod_{\sigma=\nu+1}^{n} f_\sigma^2(q_\sigma), \tag{A}$$

where the $f_\sigma(q_\sigma)$ are functions of class C^1 of a single variable each; for the Schrödinger equation the po-

tential energy must be of the form given in Theorem II. The particular solutions obtained by this method are of the form

$$\psi = A \exp\left(i \sum_{\alpha,\beta=1}^{\nu} c_\beta \int \varphi_{\alpha\beta}(q_\alpha)\, dq_\alpha\right) \prod_{\rho=\nu+1}^{n} \psi_\rho(q_\rho)$$

(+ c.c. if a real solution is desired),

where the ψ_ρ are solutions of

$$\frac{1}{f_\rho} \frac{d}{dq_\rho}\left(f_\rho \frac{d\psi_\rho}{dq_\rho}\right)$$

$$= -\frac{1}{i} \sum_{\beta=1}^{\nu}\left(\frac{1}{f_\rho}\frac{d}{dq_\rho}(f_\rho f_\rho^{\ \beta}c_\beta\psi_\rho) + f_\rho^{\ \beta}c_\beta\frac{d\psi_\rho}{dq_\rho}\right)$$

$$-\left(\sum_{\kappa,\lambda=1}^{\nu} F_\rho^{\ \kappa\lambda}c_\kappa c_\lambda + \sum_{\sigma=\nu+1}^{n}\varphi_{\rho\sigma}c_\sigma - 2u_\rho\right)\psi_\rho, \qquad (B)$$

where the various functions are defined as in Theorem II except that the φ_{ij}'s and $f_\rho^{\ \beta}$'s must now be of class C^1, and A and the c_j's are arbitrary complex constants, subject to the condition

$$\sum_{\gamma=1}^{\nu} c_\gamma^2 + \sum_{\tau=\nu+1}^{n} c_\tau = \begin{cases} 2E & \text{(Schrödinger or Helmholtz equation)} \\ 0 & \text{(Laplace's equation)} \end{cases}$$

and to restrictions due to boundary conditions on ψ. The solutions for different values of the c_j's satisfying these conditions can be linearly superposed.

IV. DISCUSSION

Theorems I and II contain explicitly the general form of the metric tensor for all $n+1$ types for which the Hamilton—Jacobi equation is separable in n-dimensional space. Theorem III asserts that the Schrödinger, Helmholtz, and Laplace equations are separable for the same form of the metric of type T_n as given in Theorem I, and Theorem IV asserts the same in the general case for the general form of the metric given in Theorem II, 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. Whether the conditions imposed on the metrics in Theorems III and IV are not only sufficient, but also necessary, to ensure the possibility of separability for the Schrödinger and related equations, has so far not been established. The difficulty is due to the fact that (contrary to what might have been expected from the familiar elementary examples) separation of variables in general involves *all* separation constants in *each* separated ODE, some of them bilinearly; it remains to be shown that Eqs. (29) represent indeed the most general form of the separated ODE's, or that still more complicated forms are possible.

The metrics of type T_n all correspond to flat space; for those of the other types this is not necessarily the case. For most physical application 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 Euclidean space, these forms will be given elsewhere.[12] Rather remarkably it turns out that the requirement of flatness implies that condition (A) is satisfied automatically in all cases, in agreement with the results for the particular case of orthogonal coordinate systems obtained by Eisenhart[13]; whether this result is true in n dimensions

is currently under investigation. Since condition (A) was the only distinction between the forms of the metric for the H—J and the other equations in Euclidean space, the forms of the metric which allow separation for the Schrödinger, Helmholtz, and Laplace equation in two and three dimensions are identical with those found for the H—J equation by Weinacht.[14]

For any n, the type T_0 implies orthogonal coordinate systems; in this case the conditions following from Theorem IV reduce to those obtained previously by Robertson[15] in the case of the Schrödinger equation on the assumption of orthogonal systems (apart from a trivial mistake in his redefinition of u_ρ). All other types do not require the coordinate systems to be orthogonal, disproving an assertion frequently made in the literature that separation is possible only for orthogonal coordinate systems (for a typical statement see Ref. 16). However, they do contain some subclasses of orthogonal coordinate systems. In particular, if for type T_n (Theorems I and III) we require

$$\sum_j \varphi_{kj}(q_k)\varphi_{lj}(q_l) = 0, \quad k \neq l, \qquad (50)$$

we obtain an orthogonal coordinate system, with

$$g_{kk} = \sum_j \varphi^2_{\ kj}(q_k) \equiv \varphi_k^{\ 2}(q_k) \equiv 1/g^{kk}. \qquad (51)$$

But this metric could also be described within type T_0 (satisfying the Robertson conditions and Theorem IV) by defining φ_{kk} as the same function of q_k as $\varphi_k^{\ 2}$, while all other φ_{kj} vanish, i.e.,

$$\varphi_{kk}(q_k) \equiv \varphi_k^{\ 2}, \qquad \varphi_{kj} = 0, \quad k \neq j, \qquad (52)$$

and thus the metric (51) can also be taken as

$$g^{kk} = \frac{1}{g_{kk}} = \frac{1}{\phi_{\mathbb{I}}}\sum_j \phi_{kj},$$

$$g = \prod_{k=1}^{n}\varphi_{kk} = \phi_{\mathbb{I}} = \phi_{\mathbb{I}}^{\ 2}\prod_{k=1}^{n} f_k^{\ 2},$$

$$f_k = \varphi_{kk}^{\ -1/2}. \qquad (53)$$

This shows that the metrics of the $n+1$ types T_ν are not mutually exclusive, a fact which has been noted in the case of the H—J equation for $n=3$ by Weinacht.[14] Furthermore, the choice (52), (53) allows the introduction of a nontrivial potential, whereas the choice (51) does not; therefore, the metric described by either expression is "essentially geodesic" only by the particular choice (51) of variables of the first kind.

For all types T_ν ($\nu \neq n$), Eqs. (B) correspond to non-orthogonal systems, except if the arbitrary functions satisfy a number of special conditions which follow immediately from Theorem II. On the other hand, these equations are self-adjoint as they stand even for non-orthogonal 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.[17] Thus in all cases we can obtain equations of the Sturm—Liouville type for the variables of the second kind.

In a recent paper, Carter[18] has obtained metrics of a four-dimensional space with signature 2 which satisfy the Einstein—Maxwell equations, possess a two-param-

eter Abelian isometry group, and for which the four-dimensional H—J and Schrödinger equations are separable, although some of the coordinate systems are nonorthogonal. Insofar as they include electromagnetic vector potentials in addition to the scalar one, the results go beyond those of this paper; however, the potentials are introduced differently, so that even in the absence of vector potentials the equations considered take the form (6) or (27) only in special cases. Apart from this, they are special cases of our results, and all metrics obtained are included in those of Theorems I—IV. The results will be discussed in more detail elsewhere.

Recently, a group theoretical method for the description of separation of variables was introduced by Winternitz et al.[19] and applied to various PDE's by Miller et al.[20] Their investigations are less general than those of this paper in restricting themselves to flat space and to only one or two spatial dimensions, though including many more detailed studies of particular coordinate systems and the associated solutions. In the case of the (one- and two-dimensional) Schrödinger equation (23) they are more general in also considering separations other than those obtained here by the reduction (25) to the time-independent equation (26). The results will be compared in detail to those obtained for flat space by the method of this paper (which, as noted above, are identical to those of Ref. 14) in Ref. 12.

This paper only dealt with those metrics which allow complete separation of variables. A subsequent paper[21] will treat metrics which allow only partial separation of variables for the equations considered here. It will also contain a discussion of the first integrals or constants of the motion which follow for dynamical systems from (complete or partial) separability of the H—J or Schrödinger equation.

[1] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Vol. 7, Chap. 5.
[2] P. Stäckel, *Uber die Integration der Hamilton—Jacobi'schen Differentialgleichung mittels Separation der Variabeln* (Habilitationsschrift, Halle a.S, 1891); Math. Ann. **42**, 548 (1893).
[3] C. L. Charlier, *Die Mechanik des Himmels* (Veit & Comp., Leipzig, 1902), Vol. I, Chap II, Sec. 1.
[4] L. P. Eisenhart, *Riemannian Geometry* (Princeton U. P., Princeton, N.J., 1949), 2nd ed., Sec. 39 and Appendix 13.
[5] L.A. Pars, *A Treatise on Analytical Dynamics* (Heinemann, London, 1965), Chap. 18.
[6] P. Stäckel, Math. Ann. **35**, 91 (1890).
[7] G. Prange, in *Encyclopädie der mathematischen Wissenschaften* (Teubner, Leipzig, 1935), Vol. IV/2, Art. 12 and 13, Sec. 19.
[8] T. Levi-Civita, Math. Ann. **59**, 383 (1904).
[9] P. Burgatti, Rend. Acc. Lincei 20 (1), 108 (1911).
[10] F.A. Dall'Acqua, Rend. Circ. Matem. Palermo 33, 341 (1912).
[11] F.A. Dall'Acqua, Math. Ann. **66**, 398 (1908).
[12] P. Havas (to be published shortly).
[13] L. P. Eisenhart, Ann. Math. **35**, 284 (1934).
[14] J. Weinacht, Math. Ann. **91**, 279 (1923).
[15] H. P. Robertson, Math. Ann. **98**, 749 (1927).
[16] P. Moon and D. E. Spencer, J. Franklin Inst. **253**, 585 (1952).
[17] See, e.g., E. L. Ince, *Ordinary Differential Equations* (Longmans, Green, London, 1927), p. 215.
[18] B. Carter, Comm. Math. Phys. 10, 280 (1968).
[19] P. Winternitz and I. Fris, Yad. Fiz. 1, 889 (1965) [Sov. J. Nucl. Phys. 1, 636 (1965)]; A.A. Makarov, J.A. Smorodinsky, Kh. Valiev, and P. Winternitz, Nuovo Cimento A **52**, 1061 (1965); P. Winternitz, I. Lukac, and Y. Smorodinskii, Yad. Fiz. 7, 192 (1968) [Sov. J. Nucl. Phys. 7, 139 (1968)]; J. Patera and P. Winternitz, J. Math. Phys. 14, 1130 (1973).
[20] W. Miller, Jr., SIAM J. Math. Anal. 5, 626, 822 (1974); E.G. Kalnins and W. Miller, Jr., J. Math. Phys. 15, 1025, 1263, 1728 (1974); C.P. Boyer, E.G. Kalnins, and W. Miller, Jr., J. Math. Phys. 16, 499 (1975).
[21] P. Havas (to be published shortly).

Integrals of motion and resonances in a dipole magnetic field

G. Contopoulos

Goddard Space Flight Center, Greenbelt, Maryland

L. Vlahos

University of Thessaloniki, Greece
(Received 26 June 1974; final revision received 21 November 1974)

A method is developed for deriving a third integral of motion, besides the Hamiltonian and the angular momentum, of a charged particle, in a dipole magnetic field. This method is particularly useful in resonance cases, where the usual adiabatic invariants are not applicable. First the Hamiltonian is reduced to a "regular" form, i.e., its lowest order terms are written as $H_2 = \Phi_{12} + \Phi_{22}$, where $\Phi_{12} = (1/2)(a^2 + p_a^2)$, $\Phi_{22} = (1/2)(\omega^2 b^2 + p_b^2)$. Then the third integral can be constructed step by step as a series; in every resonance case a different form of the integral is derived. In the nonresonant cases the Hamiltonian is written in a normal form $H = H^*(\Phi_1, \Phi_2)$, where Φ_1, Φ_2 are canonical variables introduced by using von Zeipel's method. The nonresonant orbits are quasiperiodic with frequencies $\omega_1 = \partial H^*/\partial \Phi_1$, $\omega_2 = \omega \partial H^*/\partial \Phi_2$ and rotation number Rot $= \omega_1/\omega_2$. As an example the location of a particular resonance is found. The comparison with the numerical integrations is satisfactory.

I. INTRODUCTION

The problem of the motion of charged particles in the field of a magnetic dipole has attracted much interest, because of its application to the case of the Earth's magnetic field. It was studied extensively by Störmer[1]; thus it is known as the "Störmer Problem." Some fifty years later this problem came again into focus, after the discovery of the Van Allen belts.

The motion of a charged particle of mass m and charge q in cylindrical coordinates ρ, φ, z can be described by the Hamiltonian

$$H = \frac{1}{2m}\left[p_\rho^2 + p_z^2 + \left(\frac{p_\varphi}{\rho} - \frac{qM\rho}{r^3} \right)^2 \right],\qquad(1)$$

where p_ρ, p_φ, p_z are the canonical momenta, M is the magnetic moment of the dipole, and

$$\rho = r\cos\lambda, \quad z = r\sin\lambda.\qquad(2)$$

The motion of the particle is a combination of a rotation around the "guiding center," with the cyclotron frequency

$$\omega_c = qB/m,\qquad(3)$$

where B is the magnetic field, and a motion of the guiding center along a field line.

The Hamiltonian (1) is time-independent and axisymmetric. Thus we have two integrals of motion, the energy

$$H = \tfrac{1}{2}mv^2,\qquad(4)$$

where v is the velocity of the particle, and the φ-component of the canonical momentum p_φ. If we write p_φ in the form

$$p_\varphi = qM/r_0,\qquad(5)$$

and use now dimensionless variables ρ', φ', z', and a new time t', defined by the relations

$$\rho' = \rho/r_0, \quad \varphi' = \varphi, \quad z' = z/r_0, \quad t' = qMt/mr_0^3,\qquad(6)$$

the Hamiltonian (1) becomes

$$H = \tfrac{1}{2}(p_\rho^2 + p_z^2) + \frac{1}{2}\left(\frac{1}{\rho} - \frac{\rho}{r^3} \right)^2,\qquad(7)$$

where the primes have been suppressed for notational convenience.

Thus the Störmer problem is reduced to a two-dimensional problem with potential

$$V(\rho, z) = \frac{1}{2}\left(\frac{1}{\rho} - \frac{\rho}{r^3} \right)^2.\qquad(8)$$

From the geometry of the equipotential curves it is known that a particle can be trapped only if its energy is less than 1/32. Alfvén[2] proved that apart from H and p_φ, there is an approximate third integral of motion, namely the magnetic moment of the particle

$$\mu = v_\perp^2/2\omega_c,\qquad(9)$$

where $v_\perp$ is the component of velocity perpendicular to the field line. The magnetic moment is an adiabatic invariant if the variations of the field in time and space are small, i.e., if

$$\frac{T}{B}\frac{\partial B}{\partial t} \quad \text{and} \quad \frac{a_c}{B}\frac{\partial B_i}{\partial x_j}\qquad(10)$$

are small, where T and a_c are the period and radius of the cyclotron motion.

A more accurate integral was given by Dragt[3] by using a method of canonical perturbations. Dragt introduced new coordinates, a and b, which are connected with r and λ by the relations

$$a + 1 = r\cos^{-2}\lambda\qquad(11)$$

and

$$b = (\sin\lambda)/r^2.\qquad(12)$$

(In the present paper we write $a + 1$ instead of Dragt's a, in order to consider a as a small quantity). In terms of these coordinates the Hamiltonian (7) becomes

$$H = \frac{1}{2}\left[\left(\frac{p_a}{h_a} \right)^2 + \left(\frac{p_b}{h_b} \right)^2 \right] + \frac{a^2}{2(a+1)^4\cos^6\lambda},\qquad(13)$$

where
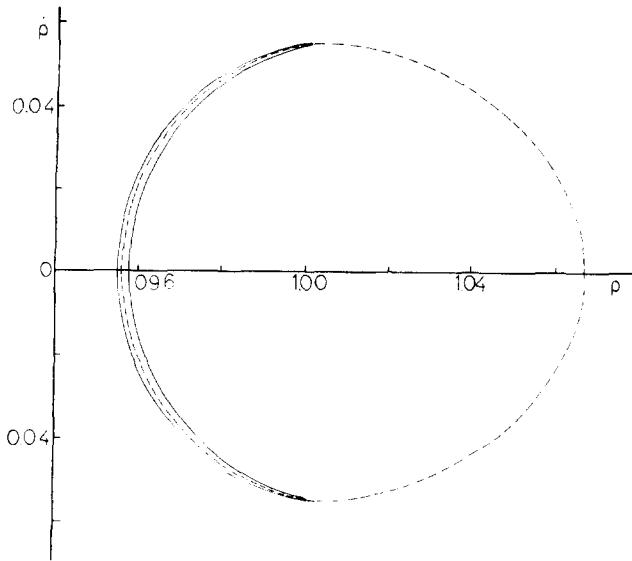
FIG. 1. Empirical invariant curve near resonance (——) compared with the nonresonant theoretical prediction (– – – –).

$$h_a^2 = \frac{\cos^6\lambda}{1 + 3\sin^2\lambda} \, , \tag{14}$$

$$h_b^2 = \frac{(a + 1)^6 \cos^{12}\lambda}{1 + 3\sin^2\lambda} \, , \tag{15}$$

and p_a, p_b are the components of the momentum along a and b.

In this coordinate system the rapid oscillations around the magnetic field lines are described by the quantities a and p_a, and the slow drift of the guiding center on the magnetic field line by b and p_b. Then the new integral is approximately[3]

$$\mu^* = \mu - \frac{1}{2\omega_c} \left(\frac{v_{\parallel}^2}{R}\right) h_a a, \tag{16}$$

where R is the radius of curvature of the magnetic lines and $v_{\parallel}$ is the component of the particle's velocity parallel to them.

If a mechanical system has two degrees of freedom and there exist two isolating integrals of motion (in our case these integrals are H and μ^*), the orbits in the four-dimensional phase-space lie on two-dimensional surfaces, called integral surfaces. These are sections of the energy supersurfaces by the supersurfaces representing the new integral.

An intersection of such a surface by the plane $z = 0$ (or $b = 0$) is called an "invariant curve."

The points of intersection of an orbit by the plane $z = 0$ lie on such an invariant curve. In particular, the periodic orbits are represented by one or more invariant points. Using the integrals (13) and (16), Dragt determined invariant curves and found that they are closed curves, like ellipses, around the point ($\rho = 1$). Comparing, further, the theoretical invariant curves with the invariant curves computed empirically, by numerical integration of orbits, he observed good

agreement between them in most of the cases.

But in the cases where the initial conditions are near those of resonant periodic orbits, there is disagreement between theory and the numerical integrations. Instead of closed curves around the point ($\rho = 1$) one finds empirically elongated curves like crescents (Fig. 1). Such invariant curves are called islands.

The existence of islands cannot be explained either by Alfvén's theory or by the theory of canonical perturbations. This disagreement appears because μ (or μ^*) is not an adiabatic invariant if there is a resonance or a near resonance between the frequencies of the cyclotron motion and the motion of the guiding center. The breakdown of the usual adiabatic invariants near resonances was well known to the first authors that used them,[4] but only recently has it attracted special attention.[5-7]

In the present paper we consider resonance phenomena in the dipole field using a perturbation technique developed in stellar dynamics.[8-10] This method consists of finding, step by step, a new integral of motion, besides the Hamiltonian, which is called a "third" integral. In order to apply this method, we must have the Hamiltonian as a power series in the variables with the lowest order terms in the form

$$H_2 = \tfrac{1}{2}(\omega_a^2 a^2 + p_a^2 + \omega_b^2 b^2 + p_b^2), \tag{17}$$

where the lowest order frequencies, ω_a and ω_b, are different from zero. We call such a form of the Hamiltonian "regular."

If ω_a/ω_b is sufficiently near a rational number n/m, we find resonance phenomena (m islands).

In the case of the dipole field we have a particular difficulty, because if we expand the Hamiltonian (13) (Sec. II) its lowest order terms are not of the regular form (17). Namely the term containing b^2 is missing. Thus we can say that one lowest order frequency is zero, and we cannot consider any resonances between the two frequencies.

In order to avoid this difficulty and bring the Hamiltonian (13) to a "regular" form, we "construct" in Sec. II a term $\tfrac{1}{2}\omega^2 b^2$. This method does not define ω uniquely. The ambiguity is resolved after we reduce the Hamiltonian to a normal form in Sec. III. The normal form of the Hamiltonian also gives the approximate positions of the various islands, corresponding to any given resonance. A numerical application in Sec. III gives satisfactory results.

II. THE "REGULAR" FORM OF THE HAMILTONIAN

If we consider a, b, p_a, p_b as small, we can expand $\sin^2\lambda$ and $\cos^{-6}\lambda$ using the relation

$$(a + 1)^4 b^2 = \sin^2\lambda (1 - \sin^2\lambda)^{-4} \tag{18}$$

and find

$$\sin^2\lambda = b^2 + 4ab^2 + 6a^2 b^2 - 4b^4 + O_5, \tag{19}$$

where O_k means terms of kth degree in the variables a, b, p_a, p_b, and

$$\cos^{-6}\lambda = 1 + 3b^2 + 12ab^2 + 18a^2b^2 - 6b^4 + O_5. \tag{20}$$

Then we calculate

$$h_a^{-2} = 1 + 6b^2 + 24ab^2 + 36a^2b^2 - 9b^4 + O_5, \tag{21}$$

$$h_b^{-2} = 1 - 6a + 21a^2 + 9b^2 - 18ab^2 - 56a^3 + 126a^4 + 27a^2b^2$$
$$+ 3b^4 + O_5, \tag{22}$$

and

$$2V = a^2 - 4a^3 + 10a^4 + 3a^2b^2 - 20a^5 + 35a^6 - 6a^2b^4 + O_7. \tag{23}$$

The relation (13) then takes the form

$$H = H_2 + H_3 + H_4 + H_5 + H_6 + O_7, \tag{24}$$

where

$$H_2 = \tfrac{1}{2}(a^2 + p_a^2) + \tfrac{1}{2}p_b^2, \tag{25}$$

$$H_3 = \tfrac{1}{2}(-4a^3 - 6ap_b^2), \tag{26}$$

$$H_4 = \tfrac{1}{2}(10a^4 + 3a^2b^2 + 21a^2p_b^2 + 6p_a^2b^2 + 9b^2p_b^2), \tag{27}$$

$$H_5 = \tfrac{1}{2}(-20a^5 - 56a^3p_b^2 + 24ap_a^2b^2 - 18ap_b^2b^2), \tag{28}$$

and

$$H_6 = \tfrac{1}{2}(35a^6 + 126a^4p_b^2 + 36a^2p_a^2b^2 - 6a^2b^4 + 27a^2b^2p_b^2$$
$$- 9p_a^2b^4). \tag{29}$$

We notice that H_2 does not contain b^2. Therefore one lowest order frequency is zero. In order to bring the Hamiltonian to a regular form, we must "construct" a term $\tfrac{1}{2}\omega^2b^2$. This we do as follows. By solving Eq. (24) for p_b^2 we find

$$p_b^2 = 2h - a^2 - p_a^2 - 2H_3 - 2H_4 - \cdots, \tag{30}$$

where h is the numerical value of the Hamiltonian (24). Thus the term $\tfrac{1}{2}b^2p_b^2$ of H_4 can be written in the form

$$\tfrac{9}{2}b^2p_b^2 = c\,\tfrac{9}{2}b^2(2h - a^2 - p_a^2 - 2H_3 - 2H_4 - \cdots)$$
$$+ (1-c)\tfrac{9}{2}b^2p_b^2, \tag{31}$$

where c is a constant. Thus, if we set $\omega^2 = 18hc$, we find a term of the form $\tfrac{1}{2}\omega^2b^2$. At this point the value of c is unspecified. However, as we will see in Sec. III, only the choice $c = \tfrac{1}{2}$ gives results that become more and more accurate by including higher order terms.

If we set $c = \tfrac{1}{2}$, the Hamiltonian takes the form

$$\overline{H} = \overline{H}_2 + \overline{H}_3 + \overline{H}_4 + \overline{H}_5 + \overline{H}_6 + O_7, \tag{32}$$

where

$$\overline{H}_2 = \Phi_{12} + \Phi_{22}, \tag{33}$$

with

$$\Phi_{12} = \tfrac{1}{2}(a^2 + p_a^2), \quad \Phi_{22} = \tfrac{1}{2}(\omega^2b^2 + p_b^2), \tag{34}$$

and

$$\omega^2 = 9h, \tag{35}$$

while

$$\overline{H}_3 = \tfrac{1}{2}(-4a^3 - 6ap_b^2), \tag{36}$$

$$\overline{H}_4 = \tfrac{1}{2}(21a^2p_b^2 + 10a^4 + \tfrac{3}{2}b^2p_a^2 - \tfrac{3}{2}b^2a^2 + \tfrac{9}{2}b^2p_b^2), \tag{37}$$

$$\overline{H}_5 = \tfrac{1}{2}(24ap_a^2b^2 - 56a^3p_b^2 - 20a^5 + 9ab^2p_b^2 + 18a^3b^2), \tag{38}$$

and

$$\overline{H}_6 = \tfrac{1}{2}(36a^2p_a^2b^2 + 126a^4p_b^2 + 35a^6 - 45a^4b^2 - 36p_a^2b^4$$
$$- \tfrac{39}{2}a^2b^4 - \tfrac{135}{2}a^2b^2p_b^2 - \tfrac{7}{2}b^4p_b^2). \tag{39}$$

We know that $h < \tfrac{1}{32}$ for trapped orbits; therefore $\omega < 3\sqrt{2}/8$.

We prove now that the equations of motion obtained from the "old" Hamiltonian [Eq. (24)] and from the "new" Hamiltonian [Eq. (32)] are equivalent.

If the "old" Hamiltonian is written in the form

$$H \equiv \tfrac{1}{2}p_b^2 + H'(a, p_a, b, p_b) + \tfrac{9}{4}b^2p_b^2 = h, \tag{40}$$

then the equations of motion are

$$\frac{da}{dt} = \frac{\partial H'}{\partial p_a}, \quad \frac{db}{dt} = p_b + \frac{\partial H'}{\partial p_b} + \tfrac{9}{2}b^2p_b,$$
$$\frac{dp_a}{dt} = -\frac{\partial H'}{\partial a}, \quad \frac{dp_b}{dt} = -\frac{\partial H'}{\partial b} - \tfrac{9}{2}bp_b^2. \tag{41}$$

The "new" Hamiltonian is written

$$\overline{H} \equiv \tfrac{1}{2}p_b^2 + H' + \tfrac{9}{4}b^2(2h - 2H' - \tfrac{9}{2}b^2p_b^2) = h, \tag{42}$$

and the corresponding equations of motion are

$$\frac{da}{dt} = \frac{\partial H'}{\partial p_a}(1 - \tfrac{9}{2}b^2), \quad \frac{db}{dt} = \left(p_b + \frac{\partial H'}{\partial p_b} + \tfrac{9}{2}b^2p_b\right)(1 - \tfrac{9}{2}b^2),$$

$$\frac{dp_a}{dt} = -\frac{\partial H'}{\partial a}(1 - \tfrac{9}{2}b^2), \tag{43}$$

$$\frac{dp_b}{dt} = -\frac{\partial H'}{\partial b}(1 - \tfrac{9}{2}b^2) - \tfrac{9}{2}b(2h - 2H' - 9b^2p_b^2).$$

From Eq. (42) we find that

$$(2h - 2H' - \tfrac{9}{2}b^2p_b^2)(1 - \tfrac{9}{2}b^2) = p_b^2(1 - \tfrac{9}{2}b^2). \tag{44}$$

Therefore, if $1 - \tfrac{9}{2}b^2 \neq 0$, we find again Eq. (40), and the numerical value of the second member of the last Eq. (43) is

$$-\left(\frac{\partial H'}{\partial b} + \tfrac{9}{2}bp_b^2\right)(1 - \tfrac{9}{2}b^2).$$

If we define now a new time t' by the relation

$$dt' = dt(1 - \tfrac{9}{2}b^2), \tag{45}$$

(assuming $1 - \tfrac{9}{2}b^2 > 0$), we find the same equations of motion (41) with t replaced by t'. Thus the orbits are the same, and the only difference is that they are described at a different rate.

The restriction $1 - \tfrac{9}{2}b^2 > 0$ means that $|b| < \sqrt{2}/3$. As $|b|$ increases the truncated formulas (19), (20), $\cdots$

1471   J. Math. Phys., Vol. 16, No. 7, July 1975

G. Contopoulos and L. Vlahos   1471

TABLE I. The values of $\sin^2\lambda$, z, and ρ for various a and b.

	$a = 0$				$a = -0.05$				$a = 0.05$			
b	0	0.1	0.2	0.3	0	0.1	0.2	0.3	0	0.1	0.2	0.3
$\sin^2\lambda$	0	0.0096	0.0347	0.0679	0	0.0079	0.0290	0.0578	0	0.0116	0.0411	0.0788
z	0	0.0972	0.1799	0.2429	0	0.0837	0.1570	0.2152	0	0.1118	0.2041	0.2715
ρ	1	0.9856	0.9484	0.8999	0.95	0.9388	0.9090	0.8689	1.05	1.0318	0.9859	0.9284
$\sin^2\lambda$ Eq. (21)	0	0.0096	0.0336	0.0576	0	0.0078	0.0262	0.0419	0	0.0118	0.0422	0.0770

deteriorate. As an example we give in Table I the values of $\sin^2\lambda$ for various values of a and b, and compare them with the values of the truncated series (19).

We notice that the difference between the accurate value of $\sin^2\lambda$ and that of Eq. (19) is less than 2% for $b = 0.1$, and less than 10% for $b = 0.2$, but it reaches 30% in the case $b = 0.3$ ($a = -0.5$). Therefore, our method can be applied if $|b|$ remains small; hence $|z|$ is also small. It does not apply at all for orbits approaching the origin.

Using the "regular" form of the Hamiltonian (32) we calculate the "third integral" in the form of a series in the variables. The algebraic form of this integral is given by a computer program[10] either far from resonances, or in exact resonances and in near resonance cases.

This integral may be used in finding the invariant curves and in particular the islands corresponding to a given resonance.[9] Numerical applications will be given in a future paper.

III. REDUCTION OF THE HAMILTONIAN TO A NORMAL FORM

If we perform the canonical change of variables $(a, b, p_a, p_b) \rightarrow (\theta_1, \theta_2, \Phi_{12}, \Phi_{22})$ through the equations

$$a = (2\Phi_{12})^{1/2}\sin\theta_1, \quad p_a = (2\Phi_{12})^{1/2}\cos\theta_1,$$

$$b = (2\Phi_{22})^{1/2}/\omega]\sin\omega\theta_2, \quad p_b = (2\Phi_{22})^{1/2}\cos\omega\theta_2, \tag{46}$$

we find the Hamiltonian (32) as a function of $\theta_1, \theta_2, \Phi_{12}, \Phi_{22}$. This is a series of trigonometric terms of the form $\frac{\sin}{\cos}(m\theta_1 \pm n\omega\theta_2)$ with coefficients polynomials in $\Phi_i^{1/2}$. In particular $\bar{H}_2$ has the form (33), while $\bar{H}_3, \bar{H}_4, \bar{H}_5, \bar{H}_6$ are easily found from Eqs. (36)—(39).

We will try now to perform a canonical change of variables such that $\bar{H}$ is brought to a normal form, i.e., it is expressed as a function of the new moments Φ_1, Φ_2 only:

$$\bar{H} = H^*(\Phi_1, \Phi_2). \tag{47}$$

Then we can write $\Phi_1 = \frac{1}{2}(a'^2 + p_a'^2)$, $\Phi_2 = \frac{1}{2}(\omega^2 b'^2 + p_b'^2)$ and the orbits in the new variables are quasiperiodic, i.e., they are harmonic oscillations along the a' and b' axes with frequencies

$$\omega_1 = \partial H^*/\partial\Phi_1, \quad \omega_2 = \omega\,\partial H^*/\partial\Phi_2. \tag{48}$$

The ratio of these two frequencies is called the "rotation number"

$$\text{Rot} = \omega_1/\omega_2. \tag{49}$$

This is the double of the rotation number as defined by Dragt.[3]

In Celestial Mechanics the usual procedure to eliminate the trigohometric terms from a Hamiltonian and bring it to a normal form is von Zeipel's method.[11]

We introduce new variables Θ_i, Φ_i through the equations

$$\Phi_{i2} = \frac{\partial S}{\partial\theta_i} = \Phi_i + \frac{\partial S_3}{\partial\theta_i} + \frac{\partial S_4}{\partial\theta_i} + \cdots,$$

$$\Theta_i = \frac{\partial S}{\partial\Phi_{12}} = \theta_i + \frac{\partial S_3}{\partial\Phi_i} + \frac{\partial S_4}{\partial\Phi_i} + \cdots, \tag{50}$$

where

$$S = \theta_1\Phi_1 + \theta_2\Phi_2 + S_3 + S_4 + \cdots \tag{51}$$

is a generating function, and S_k is of degree k in $\Phi_i^{1/2}$.

If we denote the derivatives at the point ($\Phi_{12} = \Phi_i$, $\theta_i = \Theta_i$) by parentheses, we write, in second approximation in $\Phi_i^{1/2}$,

$$\Phi_{i2} = \Phi_i + \left(\frac{\partial S_3}{\partial\theta_i}\right) + \left(\frac{\partial S_4}{\partial\theta_i}\right) - \left(\frac{\partial^2 S_3}{\partial\theta_i\partial\theta_1}\right)\left(\frac{\partial S_3}{\partial\Phi_1}\right)$$
$$- \left(\frac{\partial^2 S_3}{\partial\theta_i\partial\theta_2}\right)\left(\frac{\partial S_3}{\partial\Phi_2}\right), \tag{52}$$

$$\theta_i = \Theta_i - \left(\frac{\partial S_3}{\partial\Phi_i}\right) - \left(\frac{\partial S_4}{\partial\Phi_i}\right) + \left(\frac{\partial^2 S_3}{\partial\Phi_i\partial\theta_1}\right)\left(\frac{\partial S_3}{\partial\Phi_1}\right)$$
$$- \left(\frac{\partial^2 S_3}{\partial\Phi_i\partial\theta_2}\right)\left(\frac{\partial S_3}{\partial\Phi_2}\right), \tag{53}$$

and similar expressions in higher approximations. Thus

$$\bar{H} = \Phi_1 + \Phi_2 + \left(\frac{\partial S_3}{\partial\theta_1}\right) + \left(\frac{\partial S_3}{\partial\theta_2}\right) + \left(\frac{\partial S_4}{\partial\theta_1}\right) + \left(\frac{\partial S_4}{\partial\theta_2}\right)$$
$$- \left(\frac{\partial^2 S_3}{\partial\theta_1^2}\right)\left(\frac{\partial S_3}{\partial\Phi_1}\right) - \left(\frac{\partial^2 S_3}{\partial\theta_1\partial\theta_2}\right)\left(\frac{\partial S_3}{\partial\Phi_2}\right) - \left(\frac{\partial^2 S_3}{\partial\theta_2\partial\theta_1}\right)\left(\frac{\partial S_3}{\partial\Phi_1}\right)$$
$$- \left(\frac{\partial^2 S_3}{\partial\theta_2^2}\right)\left(\frac{\partial S_3}{\partial\Phi_2}\right) + \left(\frac{\partial\bar{H}_3}{\partial\Phi_{12}}\right)\left(\frac{\partial S_3}{\partial\theta_1}\right) + \left(\frac{\partial\bar{H}_3}{\partial\Phi_{22}}\right)\left(\frac{\partial S_3}{\partial\theta_2}\right)$$
$$- \left(\frac{\partial\bar{H}}{\partial\theta_1}\right)\left(\frac{\partial S_3}{\partial\Phi_1}\right) - \left(\frac{\partial\bar{H}}{\partial\theta_2}\right)\left(\frac{\partial S_3}{\partial\Phi_2}\right) + H_3' + H_4' + \cdots = H^*, \tag{54}$$

where H^* is a function of Φ_i only, and

$$H_k' = \bar{H}_k(\Theta_i, \Phi_i). \tag{55}$$

If we equate terms of the same order in the variables Φ_i, we find equations of the form

$$\left(\frac{\partial S_k}{\partial\theta_1}\right) + \left(\frac{\partial S_k}{\partial\theta_2}\right) = H_k^* - G_k - H_k', \tag{56}$$

where G_k contains all the terms of the first member of Eq. (54) of order k in $\Phi_i^{1/2}$, that depend only in S_i, $\bar{H}_i$ with $i \leq k - 1$.

1472   J. Math. Phys., Vol. 16, No. 7, July 1975

G. Contopoulos and L. Vlahos   1472

The function H_k^* is calculated by setting the mean value of the right-hand side of Eq. (56) with respect to the angles equal to zero. Then Eq. (56) is solved easily, because its right-hand side contains only trigonometric terms.

After several operations we find

$$H_3^* = 0, \tag{57}$$

$$S_3 = 2(2\Phi_1)^{3/2}[(\cos^3\theta_1)/3 - \cos\theta_1] - \tfrac{3}{2}(2\Phi_1)^{1/2}(2\Phi_2)$$

$$\times\{\cos\theta_1 + [1/(1-4\omega^2)](\cos\theta_1\cos2\omega\theta_2$$

$$+ 2\omega\sin\theta_1\sin2\omega\theta_2)\}, \tag{58}$$

$$H_4^* = -\tfrac{15}{2}\Phi_1^2 - [12 - 9/2(1-4\omega^2)]\Phi_1\Phi_2$$

$$- \tfrac{9}{4}[2 + 1/(1-4\omega^2) - 1/2\omega^2]\Phi_2^2, \tag{59}$$

and

$$S_4 = (2\Phi_1)^2\{-\tfrac{25}{16}\sin2\theta_1 + \tfrac{13}{32}\sin4\theta_1 - \tfrac{1}{16}\sin6\theta_1\}$$

$$+ (2\Phi_1)(2\Phi_2)\left[\sin2\theta_1\left(-\frac{3}{16\omega^2} - \frac{27}{16} + \frac{9\omega^2}{4(1-4\omega^2)}\right)\right.$$

$$+ \frac{\sin2\omega\theta_2}{\omega}\left(\frac{15}{16} + \frac{9\omega^2}{4(1-4\omega^2)}\right) + \frac{3}{32}\left(\frac{1}{\omega^2} + \frac{12\omega^2-9}{1-4\omega^2}\right)$$

$$\times\left(\frac{\sin(2\theta_1+2\omega\theta_2)}{1+\omega} + \frac{\sin(2\theta_1-2\omega\theta_2)}{1-\omega}\right) + \tfrac{3}{8}\sin4\theta_1$$

$$+ \frac{3}{16}\frac{\sin(4\theta_1+2\omega\theta_2)}{(1+2\omega)} + \frac{3}{16}\frac{\sin(4\theta_1-2\omega\theta_2)}{(1-2\omega)}$$

$$- \frac{9\omega}{16}\frac{\sin(2\theta_1+4\omega\theta_2)}{(1+2\omega)^2} + \frac{9\omega}{16}\frac{\sin(2\theta_2-4\omega\theta_2)}{(1-2\omega)^2}$$

$$- \frac{9\omega}{8}\frac{\sin4\omega\theta_2}{(1-4\omega^2)}\right] + (2\Phi_2)^2\left[\sin2\theta_1\left(-\frac{9}{16} - \frac{9}{32(1-4\omega^2)}\right)\right.$$

$$+ \frac{\sin2\omega\theta_2}{\omega}\left(\frac{9}{8} + \frac{9\omega^2}{4(1-4\omega^2)}\right) - \frac{9}{16}\frac{\sin(2\theta_1+2\omega\theta_2)}{(1+2\omega)}$$

$$- \frac{9}{16}\frac{\sin(2\theta_1-2\omega\theta_2)}{(1-2\omega)} - \frac{9}{64}\frac{\sin(2\theta_1+4\omega\theta_2)}{(1+2\omega)^2}$$

$$- \frac{9}{64}\frac{\sin(2\theta_2-4\omega\theta_2)}{(1-2\omega)^2} + \frac{9\sin4\omega\theta_2}{64\omega}\left(\frac{1}{2\omega^2} + \frac{1}{1-4\omega^2}\right)\right]. \tag{60}$$

Thus we find, in this approximation,

$$H^* = \Phi_1 + \Phi_2 - \tfrac{15}{2}\Phi_1^2 - [12 - 9/2(1-4\omega^2)]\Phi_1\Phi_2$$

$$- \tfrac{9}{4}[2 + 1/(1-4\omega^2) - 1/2\omega^2]\Phi_2^2 = h, \tag{61}$$

and

$$\omega_1 = \partial H^*/\partial\Phi_1 = 1 - 15\Phi_1 - [12 - 9/2(1-4\omega^2)]\Phi_2, \tag{62}$$

$$\omega_2/\omega = \partial H^*/\partial\Phi_2$$

$$= 1 - [12 - 9/2(1-4\omega^2)]\Phi_1 - \tfrac{9}{2}[2 + 1/(1-4\omega^2)$$

$$- 1/2\omega^2]\Phi_2. \tag{63}$$

We notice that $\omega_1 = 1 + O(h)$, because $\Phi_1 + \Phi_2 = O(h)$. However, if $\Phi_2 = O(h)$, the last term of Eq. (63) is of $O(1)$ because $\omega^2 = O(h)$. Therefore, ω_2/ω is not near 1 and $\text{Rot} = \omega_1/\omega_2$ is not near $1/\omega$.

This is a serious disadvantage because we have no reason to believe that the higher order terms will be small compared to those included in Eqs. (62)–(63), and, as we will see, they are not.

However, if we assume Φ_2 to be of $O(h^2)$ the last term of Eq. (63) is of $O(h)$ and Rot is near $1/\omega$. Therefore, the above theory can give approximately the rotation number if $\Phi_2 = O(h^2)$, which means that $p_b^2(z=0)$ is of $O(h^2)$, or that the maximum b^2 is of $O(h)$; therefore the orbits are near the equator.

At this point we can check that only the choice $\omega^2 = 9h$ [Eqs. (35)] can ensure that Rot is near $1/\omega$. In fact, if we replace $\tfrac{9}{2}b^2p_b^2$ by the expression (31), we find that H^* becomes

$$\Phi_1 + \Phi_2 - \tfrac{15}{2}\Phi_1^2 - [12 - 9/2(1-4\omega^2) - (9/2\omega^2)(1-2c)]\Phi_1\Phi_2$$

$$- \tfrac{9}{4}[2 + 1/(1-4\omega^2) - (1-c)/\omega^2]\Phi_2^2, \tag{64}$$

and

$$\omega_1 = 1 - 15\Phi_1 - [12 - 9/2(1-4\omega^2) - (9/2\omega^2)(1-2c)]\Phi_2, \tag{65}$$

$$\omega_2/\omega = 1 - [12 - 9/2(1-4\omega^2) - (9/2\omega^2)(1-2c)]\Phi_1$$

$$- \tfrac{9}{2}[2 + 1/(1-4\omega^2) - (1-c)/\omega^2]\Phi_2. \tag{66}$$

Now, in order to have both ω_1 and ω_2/ω near 1, we *must* have $c = \tfrac{1}{2}$ *and* $\Phi_2 = O(h^2)$. Otherwise, at least one of the terms $(9/2\omega^2)(1-2c)\Phi_1$, $(9/2\omega^2)(1-2c)\Phi_2$ is of $O(1)$; therefore either ω_1 or ω_2/ω differs appreciably from 1.

Up to now we have not considered higher order terms in H^*. It is easy to see that H_6^* contains terms of the form Φ_1^3, $\Phi_1^2\Phi_2/\omega^2$, $\Phi_1\Phi_2^2/\omega^4$, Φ_2^3/ω^4, and $\Phi_1^{-1}\Phi_2^4$. If $\Phi_2 = O(h)$, we see that these terms give contributions of $O(1)$ in ω_1 and ω_2/ω. However, if $\Phi_2 = O(h^2)$ the contributions in ω_1 are of $O(h^2)$ and only the terms of the form $\Phi_1^2\Phi_2/\omega^2$ and $\Phi_1\Phi_2^2/\omega^4$ give contributions of $O(h)$ in ω_2/ω. These are found after several operations to be

$$\frac{693}{4}\left(\frac{\Phi_1^2\Phi_2}{\omega^2} - \frac{\Phi_1\Phi_2^2}{4\omega^4}\right), \tag{67}$$

and their contributions in ω_2/ω are

$$\frac{693}{4}\left(\frac{\Phi_1^2}{\omega^2} - \frac{\Phi_1\Phi_2}{2\omega^4}\right). \tag{68}$$

Those are the only terms of $O(h)$ beyond those of Eq. (63). In fact the contributions of H_8, etc., are of $O(h^2)$. We notice now that the terms (67) are of $O(h^3)$. Therefore, if we omit altogether terms of $O(h^3)$ from Eq. (61), we have

$$\Phi_1 + \Phi_2 - \tfrac{15}{2}\Phi_1^2 = h. \tag{69}$$

From this equation we find Φ_1 for every small Φ_2, of $O(h^2)$. Then keeping in ω_1 and ω_2/ω only terms of $O(h)$, we find

$$\omega_1 = 1 - 15\Phi_1, \tag{70}$$

$$\frac{\omega_2}{\omega} = 1 - \left(12 - \frac{9}{2(1-4\omega^2)}\right)\Phi_1 + \frac{9}{4\omega^2}\Phi_2 + \frac{693}{4}\left(\frac{\Phi_1^2}{\omega^2} - \frac{\Phi_1\Phi_2}{2\omega^4}\right), \tag{71}$$

and

$$\text{Rot} = \omega_1/\omega_2. \tag{72}$$

One can see that for $\Phi_2 = 0$ we have $\omega_1 < 1$, $\omega_2 > \omega$, and

Rot $< \omega^{-1}$; as Φ_2 increases, ω_1 increases, ω_2 decreases, and Rot increases considerably.

We find now the intersections of an orbit by the axis $z = 0$. It can be easily shown[3] that for $z = 0$, $\dot{z} > 0$, we have $b = 0$, $a + 1 = \rho$, $p_a = \dot{\rho}$ and $p_b = \dot{b}h_b^2 = \dot{z}\rho^3 > 0$. Therefore, $\theta_2 = 0$. Then from Eqs. (50), (58), and (60) we find

$$2\Phi_{12} = p_b^2 = 2\Phi_2 + (15 - 3\cos\theta_1 + \cdots)\Phi_1\Phi_2$$
$$+ (9/4\omega^2 + \cdots)\Phi_2^2 - 24\omega^2(2\Phi_1)^{1/2}\Phi_2\sin\theta_1 + \cdots$$

$$(73)$$

Thus $p_b^2|_{b=0}$ is constant to the lowest order,[3] but it has variations in the next order. In fact Φ_2 and $p_b^2|_{b=0}$ are of $O(h^2)$, while the variations of $p_b^2|_{b=0}$ are of $O(h^3)$. Of course, in resonant regions the value of p_b^2 is not even approximately constant.

The first equation (50) gives now

$$2\Phi_{12} = (\rho - 1)^2 + \dot{\rho}^2$$
$$= 2\Phi_1 + 4(2\Phi_1)^{3/2}\sin^3\theta_1 + \Phi_1^2(-25\cos2\theta_1$$
$$+ 13\cos4\theta_1 - 3\cos6\theta_1) \qquad (74)$$

up to terms of $O(h^2)$. If we solve this equation for Φ_1 and use Eqs. (46), we find, in this approximation,

$$2\Phi_1 = (\rho - 1)^2 + \dot{\rho}^2 - 4(\rho - 1)^3 + \tfrac{3}{4}\dot{\rho}^4 - \tfrac{29}{4}(\rho - 1)^4 \qquad (75)$$
$$+ \tfrac{39}{2}\dot{\rho}^2(\rho - 1)^2 + 3[\dot{\rho}^2 - (\rho - 1)^2]^3/[\dot{\rho}^2 + (\rho - 1)^2].$$

For a fixed value of Φ_1 this equation represents an invariant curve of the regular type. In the lowest order this is a circle around the point $\rho = 1$, $\dot{\rho} = 0$.

The points of intersection of the invariant curve by the axis ρ are found if we set $\dot{\rho} = 0$ in Eq. (75), or $\theta_1 = \pi/2, 3\pi/2$ in Eq. (74) ($\pi/2$ for $\rho > 1$, $3\pi/2$ for $\rho < 1$). In the same approximation as above we find

$$(\rho - 1)^2 = 2\Phi_1 + 41\Phi_1^2 \pm 4(2\Phi_1)^{3/2} \qquad (76)$$

(+ if $\rho > 1$, - if $\rho < 1$). We apply this method to one particular case studied by Dragt (private communication) where $h = 0.002\,907$; Dragt finds empirically two periodic orbits with rotation number Rot $= 6$, starting at $\rho = 0.9350$ and $\rho = 1.0897$. Using our formulas (35), (69), (70), (71), and (72), we find $\omega = 0.161\,75$, $\Phi_2 = 0.000\,17$, $\Phi_1 = 0.007\,80$, $\omega_1 = 0.9581$, and $\omega_2 = 0.1596$. Then from Eq. (76) we find $\rho = 0.935$ and $\rho = 1.087$. Thus the agreement with the numerical results is good.

At the boundary $\Phi_2 = 0$ we find $\Phi_1 = 0.002\,97$, and Rot $= 5.89$, while Dragt finds Rot $= 6.18$ by using his own theory. As we move inside the boundary, Rot becomes larger; therefore Dragt's theoretical value at the periodic orbit deviates further from Rot $= 6.00$. Thus the present theory constitutes an improvement.

However we must notice that in the present case z is small ($z_{max} \approx 0.1$). For larger z_{max} a larger number of terms is needed in order to find satisfactory results.[10]

We are presently applying a computer program to derive such higher order terms for the dipole field.

IV. SUMMARY

In order to construct integrals of motion for the dipole field, which are valid near resonances, we must first bring the Hamiltonian in a "regular" form

$$H = \tfrac{1}{2}(a^2 + p_a^2 + \omega^2 b^2 + p_b^2) + \text{higher order terms.} \qquad (77)$$

However, the Hamiltonian is given as

$$H = \tfrac{1}{2}(a^2 + p_a^2 + p_b^2) + \text{higher order terms.} \qquad (78)$$

In order to "construct" a term of the form $\tfrac{1}{2}\omega^2 b^2$ we notice that a higher order term of Eq. (78), namely $\tfrac{9}{2}b^2 p^2$ can be written in the form $c\tfrac{9}{2}b^2(2h - a^2 - p_a^2 - \cdots) + (1 - c)\tfrac{9}{2}b^2 p_b^2$, where h is a constant, equal to the numerical value of the Hamiltonian. Thus we can set $\omega^2 = 18ch$, and we prove that the new Hamiltonian gives the same orbits as the original one.

In order to make the best choice of the constant c, we bring the Hamiltonian to a normal form up to the terms of fourth degree in the variables. Then we compute the frequencies ω_1, ω_2 along the (new) a and b axes, and we require that these should be of the form: $\omega_1 = 1 +$ higher order terms, $\omega_2 = \omega(1 +$ higher order terms), as one should expect from the lowest order terms of the Hamiltonian (77). This requirement is satisfied only if $c = \tfrac{1}{2}$. Thus we take $\omega^2 = 9h$ and we can now proceed in constructing two integrals of motion, which can be applied both for nonresonance and resonance cases. Such integrals are particularly useful in explaining the forms of the orbits in near resonance cases, where the usual adiabatic invariants are not applicable.

It is clear that the above method can be applied to other Hamiltonians also of the form (78).

[1]C. Störmer, Arch. Sci. Phys. Nat. 24, 350 (1907).
[2]H. Alfvén, Cosmical Electrodynamics (Claredon Press, Oxford, 1950).
[3]A. Dragt, Geophys. Rev. 3, 255 (1965).
[4]See, e.g., A. Sommerfeld, Atombau und Spektralinien (F. Vieweg und Sohn, Braunschweig, 1919).
[5]G. Contopoulos, J. Math. Phys. 7, 788 (1966).
[6]B.V. Chirikov, Nucl. Phys. Inst. Siberian Sec. USSR Acad. Sci. Rep. No. 262; CERN Transl. 71-40 (1971).
[7]E.F. Jaeger and A.J. Lichtenberg, Ann. Phys. 71, 314 (1972).
[8]G. Contopoulos, Z. Astrophys. 49, 273 (1960).
[9]G. Contopoulos, Astron. J. 70, 562 (1965).
[10]G. Contopoulos, Astrophys. J. Suppl. 13, 122 (1966); Astrophys. J. 153, 83 (1968).
[11]See, e.g., Y. Hagihara, Celestial Mechanics, Vol. 2 (MIT Press, Cambridge, 1972). The von Zeipel method is not directly applicable when Rot is very close to a rational number n/m. However, unless this resonance affects the lowest order terms of the Hamiltonian H^* (this happens only if $\omega \simeq 1/2$), a truncated form of H^* can give the approximate positions of the resonances.

Second quantization of classical nonlinear relativistic field theory. I. Canonical formalism

T. Bałaban*

Institute of Mathematics, University of Warsaw, Warsaw, Poland

R. Rączka*

Institute for Nuclear Research, Warsaw, Poland
(Received 15 April 1974)

The canonical formalism for nonlinear relativistic classical field theory is presented. It is shown that the solutions $\Phi(x)$ of the nonlinear equation $(\Box + m^2)\Phi(x) = \lambda\Phi^3(x)$ as well as the asymptotic fields $\Phi_{in}(x)$ and $\Phi_{out}(x)$ are local relativistic fields with respect to Poisson brackets, with initial data as canonical variables. A convenient form for the generators of the Poincaré group is derived, and the properties of the scattering operator are discussed.

1. INTRODUCTION

We shall consider in this series of papers the explicit construction of interacting quantum scalar fields $\hat{\Phi}(x)$ which satisfy the nonlinear relativistic wave equation in four-dimensional space—time.

In the present work we elaborate the canonical formulation of the classical nonlinear relativistic field theory. We first show that if the initial Cauchy data for the classical field $\Phi(x)$ satisfying the equation

$$(\Box + m^2)\Phi(x) = \lambda\Phi^3(x), \quad \lambda < 0, \tag{1.1}$$

are properly chosen, then under the Poisson brackets the interacting field $\Phi(x)$ and the asymptotic fields $\Phi_{in}(x)$ and $\Phi_{out}(x)$ are canonical, i.e., $[\Pi(x) \equiv (\partial_t\Phi)(x)]$

$$\{\Phi(t, \mathbf{x}), \Pi(t, \mathbf{y})\} = \delta^{(3)}(\mathbf{x} - \mathbf{y}),$$

$$\{\Phi(t, \mathbf{x}), \Phi(t, \mathbf{y})\} = \{\Pi(t, \mathbf{x}), \Pi(t, \mathbf{y})\} = 0, \tag{1.2}$$

and

$$\{\Phi_{in \atop out}(x), \Phi_{in \atop out}(y)\} = \Delta(x - y : m). \tag{1.3}$$

This implies that the classical evolution operator $U(\tau, \tau_0)$ the Möller scattering operators $U(-\infty, \tau)$, $U(\tau, \infty)$, and the S operator $S = U(-\infty, \infty)$ are canonical transformations.

In addition we show that for the considered class of Cauchy data the generators of the Poincaré group, P_μ and $M_{\mu\nu}$, $\mu, \nu = 0, 1, 2, 3$, are equal to those associated with the free asymptotic fields Φ_{in} or Φ_{out}, i.e.,

$$P_\mu = P_\mu^{in} = P_\mu^{out}, \quad M_{\mu\nu} = M_{\mu\nu}^{in} = M_{\mu\nu}^{out}. \tag{1.4}$$

The fields $\Phi(x)$, $\Phi_{in}(x)$ and $\Phi_{out}(x)$ transform covariantly under the same representation $U_{(a,\Lambda)}$ of the Poincaré group. In particular, we have

$$\{\Phi, P_\mu\} = \partial_\mu\Phi, \quad \{\Phi, M_{\mu\nu}\} = (x_\mu\partial_\nu - x_\nu\partial_\mu)\Phi. \tag{1.5}$$

The classical S operator is invariant under the Poincaré group and differs from the identity.

We see therefore that this classical field theory with the canonical formulation satisfies most of the conditions which we usually impose in quantum field theory like, e.g., locality, relativistic covariance, and asymptotic conditions.

The next step is the passage from the canonical formulation of field theory to an operator quantization.

This consists in the construction of an operator representation of the Heisenberg Lie algebra given by Eq. (1.2). We carry out this program in Paper II. Finally in paper III we shall consider the quantum S operator formalism in the present framework.

Our works represent a continuation of Segal's program of the construction of an interacting quantum field as operators acting in the space of solutions of the corresponding classical nonlinear equations[1] (see also Streater[2]). The construction of an interacting quantum field is carried out by the quantization of solutions of Eq. (1.1). The alternative program of a direct quantization of dynamical equations (1.1) was considered by Rączka.[3]

2. CANONICAL FORMALISM

Consider the nonlinear relativistic wave equation

$$(\Box + m^2)\Phi(x) = \lambda\Phi^3(x), \quad \lambda < 0, \quad x = (t, \mathbf{x}) \in R^4, \tag{2.1}$$

with the initial conditions

$$\Phi(0, \mathbf{x}) = \varphi(\mathbf{x}), \quad \Pi(0, \mathbf{x}) = \pi(\mathbf{x}). \tag{2.2}$$

It was shown by Morawetz and Strauss[4] that for every given Cauchy data (2.2) defined and sufficiently regular on R^3 there exists the unique solution $\Phi(x)$ of Eq. (2.1) and the pair $\Phi_{in}(x)$ and $\Phi_{out}(x)$ of the solutions of the free Klein—Gordon equation such that

$$\Phi_{in}(t, \mathbf{x}) \underset{t \to -\infty}{\sim} \Phi(t, \mathbf{x}) \underset{t \to \infty}{\sim} \Phi_{out}(t, \mathbf{x}), \tag{2.3}$$

in the energy norm given by the formula

$$\|\Phi(t, \cdot)\|_E^2$$

$$= \int d^3\mathbf{x}[\Pi^2(t, \mathbf{x}) + |\nabla\Phi(t, \mathbf{x})|^2 + m^2\Phi^2(t, \mathbf{x})]. \tag{2.4}$$

It was shown in Ref. 4 that the functions $\Phi(\tau, \mathbf{x})$ and $\Pi(\tau, \mathbf{x})$ on a hyperplane $t = \tau$ belong to the Banach space of initial data and therefore may be used for the construction of a new free field $\Phi_\tau(t, \mathbf{x})$. This field is given by the formula

$$\Phi_\tau[t, \mathbf{x} \mid \varphi, \pi]$$

$$= \int d^3\mathbf{y} \, \Delta_R(t - t', \mathbf{x} - \mathbf{y})\overset{\leftrightarrow}{\partial}_{t'}\Phi[t', \mathbf{y} \mid \varphi, \pi]_{t'=\tau}, \quad t > \tau, \tag{2.5}$$

and we have

$$\Phi_\tau(t, \mathbf{x})$$

$$= -\int_\tau^\infty dt' \partial_{t'} \int d^3\mathbf{y} \, \Delta_R(t - t', \mathbf{x} - \mathbf{y})\overset{\leftrightarrow}{\partial}_{t'}\Phi(t', \mathbf{y})$$

$$= \Phi(t, \mathbf{x}) - \lambda \int_\tau^\infty dt' \int d^3\mathbf{y}\, \Delta_R(t - t', \mathbf{x} - \mathbf{y}) \Phi^3(t', \mathbf{y}). \quad (2.6)$$

Replacing in Eq. (2.5) Δ_R by Δ_A, we obtain the integral representation for $\Phi_\tau(t, x)$ field for $t < \tau$. We have

$$\Phi_{\substack{\text{in}\\\text{out}}} [t, \mathbf{x} \,|\, \varphi, \pi] = \lim_{\tau \to \mp \infty} \Phi_\tau[t, \mathbf{x} \,|\, \varphi, \pi], \quad (2.7)$$

in the energy norm.[4] The free solutions $\Phi_\tau(t, \mathbf{x})$ play an important role in the canonical formalism and in the scattering theory of classical fields.

The Cauchy data φ and π may be used as canonical variables in classical field theory. If $F(\varphi, \pi)$ is a smooth functional in the sense of Gateaux over the Banach space of Cauchy data then the functional derivative $\delta F/\delta\varphi(\mathbf{x})$ is defined by means of the Gateaux derivative by the formula

$$(\partial_s F)(\varphi, \pi) = \lim_{s \to 0} s^{-1}[F(\varphi + sX, \pi) - F(\varphi, \pi)]$$
$$= \left\langle \frac{\delta F}{\delta\varphi}, X \right\rangle = \int d^3\mathbf{x} \frac{\delta F}{\delta\varphi(\mathbf{x})} X(\mathbf{x}), \quad (2.8)$$

where $X(\mathbf{x}) \in C_0^\infty(R^3)$. If φ and π are functions which satisfy the smooth initial conditions of Morawetz and Strauss,[4] then the functions

$$\varphi + sX, \pi \quad \text{or} \quad \varphi, \pi + sX,$$
$$0 \leq s < \infty, \quad X \in C_0^\infty(R^3)$$

also represent smooth initial conditions and uniquely define solutions $\Phi[t, \mathbf{x} \,|\, \varphi + sX, \pi]$ or $\Phi[t\mathbf{x} \,|\, \varphi, \pi + sX]$ of Eq. (2.1) (cf. Appendix A).

In the canonical formalism presented below an important role is played by the Gateaux derivatives and related variational derivatives $\delta\Phi/\delta\varphi(\mathbf{x})$ and $\delta\Phi/\delta\pi(\mathbf{x})$ of solutions of Eq. (2.1). We have:

Proposition 1: The solutions $\Phi[\cdot \,|\, \varphi, \pi]$ of Eq. (2.1) have Gateaux derivatives with respect to φ and π in the topology defined by the energy norm (2.4).

Proof: Denote by Φ_s and Φ the solutions of Eq. (2.1) determined by the initial conditions $(\varphi + sX, \pi)$ and (φ, π) respectively and let $\vartheta_s = s^{-1}(\Phi_s - \Phi)$. The function ϑ_s satisfies the following equation:

$$(\Box + m^2)\vartheta_s = \lambda(\Phi_s^2 + \Phi_s\Phi + \Phi^2)\vartheta_s,$$

and the initial conditions $\vartheta_s(0, \mathbf{x}) = X(\mathbf{x})$, $(\partial_t \vartheta_s)(0, \mathbf{x}) = 0$. Because Φ has finite F-norm given by Eq. (A1) of Appendix A, the condition (B2) of Appendix B is satisfied. Hence by virtue of Lemma 1 of Appendix B one obtains

$$\|\vartheta_s(t, \cdot)\|_E^2 \leq C^2 \int (|\nabla X(\mathbf{x})|^2 + m^2 X^2) d^3\mathbf{x}.$$

Denote by ϑ_0 the solution of the equation

$$(\Box + m^2)\vartheta_0 = V\vartheta_0, \quad (2.9)$$

with $V(t, \mathbf{x}) = 3\lambda\Phi^2(t, \mathbf{x})$ and initial conditions $\vartheta_0(0, \mathbf{x}) = X(\mathbf{x})$ and $(\partial_t\vartheta_0)(0, \mathbf{x}) = 0$. We show that $\vartheta_s \to \vartheta_0$ at $s \to 0$, in the energy norm (2.4). For $\vartheta_s - \vartheta_0$ we have the following equation and initial conditions:

$$(\Box + m^2)(\vartheta_s - \vartheta_0)$$
$$= 3\lambda\Phi^2(\vartheta_s - \vartheta_0) + 3\lambda(\Phi_s + 2\Phi)(\Phi_s - \Phi)\vartheta_s, \quad (2.10)$$

$$(\vartheta_s - \vartheta_0)(0, \mathbf{x}) = 0, \quad [\partial_t(\vartheta_s - \vartheta_0)](0, \mathbf{x}) = 0.$$

Hence by virtue of Eq. (B3) of Appendix B we have

$$\|(\vartheta_s - \vartheta_0)(t, \cdot)\|_E$$
$$\leq c_1 |\lambda| \int_0^t \|[(\Phi_s + 2\Phi)(\Phi_s - \Phi)\vartheta_s](\tau, \cdot)\| d\tau$$
$$\leq c_2 |\lambda| (\sup|\Phi_s| + \sup|\Phi|) \sup|\Phi_s - \Phi|$$
$$\times \left| \int_0^t \|\vartheta_s(\tau, \cdot)\|_2 d\tau \right| \quad (2.11)$$
$$\leq c_3 |\lambda| |t| (\sup|\Phi_s| + \sup|\Phi|)$$
$$\times \left(\int [|\nabla X(\mathbf{x})|^2 + m^2 X^2(\mathbf{x})] d^3\mathbf{x} \right) \sup|\Phi_s - \Phi|.$$

It follows from the continuity of the solution $\Phi[t, \mathbf{x} \,|\, \varphi, \pi]$ with respect to the initial conditions that $\|\vartheta_s - \vartheta_0\|_E \to 0$ for $s \to 0$. The function ϑ_0 as the solution of a linear homogeneous equation with sufficiently regular coefficients depends linearly on X: hence $\Phi[\cdot \,|\, \varphi, \pi]$ is differentiable in the sense of Gateaux. Similarly one proves the existence of the Gateaux derivative for functionals $\Phi[\cdot \,|\, \varphi, \pi]$ with respect to π. \hfill QED

Remark: One may similarly prove the existence of arbitrary order Gateaux derivatives of the functional $\Phi[\cdot \,|\, \varphi, \pi]$ with respect to the variables φ and π.

The Poisson bracket $\{F, G\}$ of two smooth functionals of the canonical variables φ and π is defined by the formula

$$\{F, G\} = \int d^3\mathbf{x} \left(\frac{\delta F}{\delta\varphi(\mathbf{x})} \frac{\delta G}{\delta\pi(\mathbf{x})} - \frac{\delta F}{\delta\pi(\mathbf{x})} \frac{\delta G}{\delta\varphi(\mathbf{x})} \right). \quad (2.12)$$

In particular for δ functionals $F(\varphi, \pi) = \varphi(\mathbf{x})$ and $G(\varphi, \pi) = \pi(\mathbf{x})$ we obtain from formula (2.12)

$$\{\varphi(\mathbf{x}), \pi(\mathbf{y})\} = \delta^{(3)}(\mathbf{x} - \mathbf{y}), \quad \{\varphi(\mathbf{x})\varphi(\mathbf{y})\} = \{\pi(\mathbf{x}), \pi(\mathbf{y})\} = 0. \quad (2.13)$$

The following theorem shows that the evolution operator $U(\tau, \tau_0)$ for Eq. (2.1) is a canonical transformation. In fact we have

Theorem 2: The field $\Phi_\tau[t, \mathbf{x} \,|\, \varphi, \pi]$ satisfies the following Poisson bracket relations:

$$\{\Phi_\tau(x), \Phi_\tau(y)\} = -\Delta(x - y : m). \quad (2.14)$$

Proof: Let $\alpha(\mathbf{x}) \in C_0^\infty(R^3)$ and let $\alpha(t, \mathbf{x})$ be the solution of the Klein–Gordon equation satisfying at some $t = t_0$ the initial conditions $\alpha(t_0, \mathbf{x}) = 0$ and $(\partial_t\alpha)(t_0, \mathbf{x}) = \alpha(\mathbf{x})$. We have then

$$\int d^3\mathbf{x} \partial_s\Phi_\tau[t_0, \mathbf{x} \,|\, \varphi + sX, \pi]\alpha(\mathbf{x})$$
$$= \int d^3\mathbf{x} \{ \int d^3\mathbf{y}\, \Delta(t - t', \mathbf{x} - \mathbf{y})$$
$$\times \overleftrightarrow{\partial_{t'}} \partial_s\Phi[t', \mathbf{y} \,|\, \varphi + sX, \pi]_{t'=\tau} \} \overleftrightarrow{\partial_t} \alpha(t, \mathbf{x}) |_{t=t_0}$$
$$= \int d^3\mathbf{x} \partial_s\Phi(t, \mathbf{x}) \overleftrightarrow{\partial_t} \alpha(t, \mathbf{x}) |_{t=\tau}. \quad (2.15)$$

We now use the fact that the scalar product

$$(\alpha, \beta)_t = \int d^3\mathbf{x}\, \alpha(t, \mathbf{x}) \overleftrightarrow{\partial_t} \beta(t, \mathbf{x}) \quad (2.16)$$

is t-independent not only for solutions of the Klein–Gordon equation but also for the solution of a more general equation (2.9). Such an equation for $V = 3\lambda\Phi^2$ satisfies the function $\partial_s\Phi$. Take the solution u_τ of this equation satisfying the following initial conditions:

$$u_\tau(\tau, \mathbf{x}) = \alpha(\tau, \mathbf{x}),$$

$$(\partial_t u_\tau)(\tau, \mathbf{x}) = (\partial_t \alpha)(\tau, \mathbf{x}). \tag{2.17}$$

We have then

$$\int d^3\mathbf{x} \, \partial_s \Phi_\tau[t_0, \mathbf{x} \mid \varphi + sX, \pi]\alpha(\mathbf{x})$$
$$= (\partial_s \Phi(t, \cdot), u_\tau(t, \cdot))_{t=\tau}$$
$$= (\partial_s \Phi(t, \cdot), u_\tau(t, \cdot))_{t=0}$$
$$= \int d^3\mathbf{x}(\partial_t u_\tau)(0, \mathbf{x})X(\mathbf{x}). \tag{2.18}$$

Consequently,

$$\frac{\delta}{\delta\varphi(\mathbf{x})} \int d^3y \, \Phi_\tau[t_0, \mathbf{y} \mid \varphi, \pi]\alpha(y) = (\partial_t u_\tau)(0, \mathbf{x}). \tag{2.19}$$

One derives similarly the expression for the functional derivative $\delta\Phi/\delta\pi(\mathbf{x})$. Indeed let $\beta(t, \mathbf{x})$ be a solution of the free Klein—Gordon equation satisfying the initial condition

$$\beta(t_0, \mathbf{x}) = 0, \quad (\partial_t \beta)(t_0, \mathbf{x}) = \beta(\mathbf{x}),$$

$$\beta(\mathbf{x}) \in C_0^\infty(R^3).$$

Then, utilizing the same considerations as previously, one obtains

$$\int d^3\mathbf{x} \, \partial_s \Phi_\tau[t_0, \mathbf{x} \mid \varphi, \pi + sX]\beta(\mathbf{x})$$
$$= \int d^3\mathbf{x} \, \partial_s \Phi[t, \mathbf{x} \mid \varphi, \pi + sX]\overrightarrow{\partial_t}\beta(t, \mathbf{x}) \mid_{t=\tau}.$$

Defining now ϑ_τ as a solution of Eq. (2.9) with $V = 3\lambda\Phi^2$ satisfying at $t = \tau$ the same initial conditions as $\beta(t, \mathbf{x})$, we obtain

$$\int d^3\mathbf{x} \, \partial_s \Phi_\tau[t_0, \mathbf{x} \mid \varphi, \pi + sX]\beta(\mathbf{x})$$
$$= -\int d^3\mathbf{x} \, \vartheta_\tau(0, \mathbf{x})X(\mathbf{x}). \tag{2.20}$$

Hence

$$\frac{\delta}{\delta\pi(\mathbf{x})} \int d^3y \, \Phi_\tau[t_0, \mathbf{y} \mid \varphi, \pi]\beta(\mathbf{y}) = -\vartheta_\tau(0, \mathbf{x}). \tag{2.21}$$

It follows from Eqs. (2.19) and (2.21) that the Poisson bracket for the functionals

$$\int d^3\mathbf{x} \, \Phi_\tau[t, \mathbf{x} \mid \varphi, \pi]\alpha(\mathbf{x}),$$
$$\int d^3\mathbf{y} \, \Phi_\tau[r, \mathbf{y} \mid \varphi, \pi]\beta(\mathbf{y}) \tag{2.22}$$

is well defined by formula (2.12), i.e., the considered integral is convergent. We have, moreover

$$\{\int d^3\mathbf{x} \, \Phi_\tau[t, \mathbf{x} \mid \varphi, \pi]\alpha(\mathbf{x}), \int d^3\mathbf{y} \, \Phi_\tau[r, \mathbf{y} \mid \varphi, \pi]\beta(\mathbf{y})\}$$
$$= \int d^3\mathbf{z}[-(\partial_t u_\tau)(0, \mathbf{z})\vartheta_\tau(0, \mathbf{z}) + u_\tau(0, \mathbf{z})(\partial_t \vartheta_\tau)(0, \mathbf{z})]$$
$$= (u_\tau(t', \cdot), \vartheta_\tau(t', \cdot))_{t'=0}. \tag{2.23}$$

Using now the t-independence of the scalar product for solutions of Eq. (2.9) and the definition of u_τ and ϑ_τ, one obtains

$$(u_\tau(t', \cdot), \vartheta_\tau(t', \cdot))_{t'=0} = (u_\tau(t', \cdot), \vartheta_\tau(t', \cdot))_{t'=\tau}$$
$$= \int d^3\mathbf{z} \, \alpha(t', \mathbf{z})\overrightarrow{\partial_t}\beta(t', \mathbf{z}) \mid_{t'=\tau}$$
$$= (\alpha(t', \cdot), \beta(t', \cdot))_{t'=\tau} = (\alpha(t', \cdot), \beta(t', \cdot))_{t'=t}$$
$$= -\int d^3\mathbf{x} \, \alpha(\mathbf{x})\beta(t, \mathbf{x})$$
$$= \int d^3\mathbf{x} \alpha(\mathbf{x})(\int d^3\mathbf{y} \, \Delta(t - r, \mathbf{x} - \mathbf{y})\overrightarrow{\partial_r}\beta(r, \mathbf{y}))$$
$$= \int d^3\mathbf{x} \, d^3\mathbf{y} \, \alpha(\mathbf{x})\Delta(t - r, \mathbf{x} - \mathbf{y})\beta(\mathbf{y}). \tag{2.24}$$

This proves Eq. (2.14). QED

Corollary 1: The interacting field $\Phi(t, \mathbf{x})$ satisfies the following equal time commutation relations:

$$\{\Phi(t, \mathbf{x}), \Pi(t, \mathbf{y})\} = \delta^{(3)}(\mathbf{x} - \mathbf{y}), \tag{2.25}$$

$$\{\Phi(t, \mathbf{x}), \Phi(t, \mathbf{y})\} = \{\Pi(t, \mathbf{x}), \Pi(t, \mathbf{y})\} = 0.$$

Proof: This follows from Eq. (2.14) and the fact that for $t = \tau$ the values of the field $\Phi_\tau(\tau, \mathbf{x})$ and $\Pi_\tau(\tau, \mathbf{x})$ coincide with those of $\Phi(\tau, x)$ and $\Pi(\tau, \mathbf{x})$ respectively for the interacting field. QED

We now show that the asymptotic classical fields $\Phi_{in}(x)$ and $\Phi_{out}(x)$ are canonical. Indeed we have:

Theorem 3: The fields $\Phi_{in}(x)$ and $\Phi_{out}(x)$ satisfy the following commutation relations:

$$\{\Phi_{\substack{in\\out}}(x), \Phi_{\substack{in\\out}}(y)\} = -\Delta(x - y : m). \tag{2.26}$$

Proof: We first prove the uniform convergences of $\partial_s \Phi_\tau[t, \cdot \mid \varphi + sX, \pi]$ for $\tau \to -\infty$. By virtue of Eq. (2.6) we have

$$\partial_s \Phi_{\tau_1}(t, \mathbf{x}) - \partial_s \Phi_{\tau_2}(t, \mathbf{x})$$
$$= -3\lambda \int_{\tau_1}^{\tau_2} dt' \int d^3\mathbf{y} \, \Delta_R(t - t', \mathbf{x} - \mathbf{y})\partial_s \Phi(t', \mathbf{y})\Phi^2(t', \mathbf{y}). \tag{2.27}$$

Hence

$$\|\partial_s \Phi_{\tau_1}(t, \cdot) - \partial_s \Phi_{\tau_2}(t, \cdot)\|_E$$
$$\leq 3|\lambda| \int_{\tau_1}^{\tau_2} dt' \|\int d^3\mathbf{y} \, \Delta_R(t - t', \cdot - y)$$
$$\times \Phi^2(t', \mathbf{y})\partial_s \Phi(t', y)\|_E.$$

Using the definition of E norm and the properties of the Δ function, for $\tau_1 < \tau_2 < t$ we obtain

$$\|\int d^3\mathbf{y}\Delta(t - t', \cdot - y)\Phi^2(t', y)\partial_s \Phi(t', y)\|_E$$
$$= \|\Phi^2(t', \cdot)\partial_s \Phi(t', \cdot)\|_2. \tag{2.28}$$

Hence, using Lemma B1 of Appendix B, we obtain

$$\|\partial_s \Phi_{\tau_1}(t, \cdot) - \partial_s \Phi_{\tau_2}(t, \cdot)\|_E$$
$$\leq 3|\lambda| \int_{\tau_1}^{\tau_2} dt' \|\Phi^2(t', \cdot)\partial_s \Phi(t', \cdot)\|_2$$
$$\leq \frac{3|\lambda|}{m} \int_{\tau_1}^{\tau_2} \sup_x \Phi^2(t', \mathbf{x}) dt' \sup_{t' \in [\tau_1, \tau_2]} \|\partial_s \Phi(t', \cdot)\|_E$$
$$\leq C_0 \int_{\tau_1}^{\tau_2} \sup_x \Phi^2(t', \mathbf{x}) dt' \|\partial_s \Phi(0, \cdot)\|_E. \tag{2.29}$$

It follows from the convergence of the integral $\int_{-\infty}^\infty \sup_x \Phi^2(t, \mathbf{x}) dt$ which is a part of the F norm given by Eq. (A1) of Appendix A that

$$\|\partial_s \Phi_{\tau_1}(t, \cdot) - \partial_s \Phi_{\tau_2}(t, \cdot)\|_E \xrightarrow[\tau_1, \tau_2 \to -\infty]{} 0$$

uniformly with respect to s in a bounded interval. Hence

$$\partial_s \Phi_{in}[t, \mathbf{x} \mid \varphi + sX, \pi] = \lim_{\tau \to -\infty} \partial_s \Phi_\tau[t, \mathbf{x} \mid \varphi + sX, \pi], \tag{2.30}$$

in the energy norm. Using definition (2.5), we obtain for $s = 0$

$$\int d^3\mathbf{y} \frac{\delta\Phi_{in}[t, \mathbf{x} \mid \varphi, \pi]}{\delta\varphi(\mathbf{y})} X(\mathbf{y})$$

$$= \lim_{\tau \to -\infty} \int d^3 \mathbf{y} \frac{\delta \Phi_\tau[t, \mathbf{x} \mid \varphi, \pi]}{\delta \varphi(\mathbf{y})} X(\mathbf{y}). \qquad (2.31)$$

This expression converges in the enrgy norm: Consequently, it converges also after integration with a smooth function of variable $\mathbf{x}$.

Similarly one derives the formula

$$\int d^3 \mathbf{y} \frac{\delta \Phi_{in}[t, \mathbf{x} \mid \varphi, \pi]}{\delta \pi(\mathbf{y})} X(\mathbf{y})$$

$$= \lim_{\tau \to -\infty} \int d^3 \mathbf{y} \frac{\delta \Phi_\tau[t, \mathbf{x} \mid \varphi, \pi]}{\delta \pi(\mathbf{y})} X(\mathbf{y}). \qquad (2.32)$$

The above considerations show that we have convergence for

$$\frac{\delta}{\delta \varphi(\mathbf{y})} \int d^3 \mathbf{x} \, \Phi_\tau[t, \mathbf{x} \mid \varphi, \pi] \alpha(\mathbf{x})$$

$$- \frac{\delta}{\delta \varphi(\mathbf{y})} \int d^3 \mathbf{x} \, \Phi_{in}[t, \mathbf{x} \mid \varphi, \pi] \alpha(\mathbf{x}) \qquad (2.33)$$

and

$$\frac{\delta}{\delta \pi(\mathbf{y})} \int d^3 \mathbf{x} \, \Phi_\tau[t, \mathbf{x} \mid \varphi, \pi] \beta(\mathbf{x})$$

$$- \frac{\delta}{\delta \pi(\mathbf{y})} \int d^3 \mathbf{x} \, \Phi_{in}[t, \mathbf{x} \mid \varphi, \pi] \beta(\mathbf{x}), \qquad (2.34)$$

$\alpha, \beta \in C_0^\infty(R^3)$ in the sense of distributions from $D'(R^3)$.

We shall now analyze these convergences in detail; consider a point $t = t_0$ for which we shall consider the convergence of (2.33): Denote by $\alpha(t, \mathbf{x})$ and $\beta(t, \mathbf{x})$ the solutions of the free Klein—Gordon equation satisfying the initial conditions $\alpha(t_0, \mathbf{x}) = 0$, $(\partial_t \alpha)(t_0, \mathbf{x}) = \alpha(\mathbf{x})$, and analogously for $\beta(t, \mathbf{x})$. The functional derivative $[\delta/\partial \varphi(\mathbf{x})] \int d^3 \mathbf{y} \Phi_\tau \times [t_0 \mathbf{y} \mid \varphi, \pi] \alpha(\mathbf{y})$ satisfies Eq. (2.19). We now show that the functions u_τ are convergent for $\tau \to -\infty$ in the energy norm, uniformly with respect to t. The limit $u_{-\infty}$ is then the solution of Eq. (2.9) asymptotically convergent at $t \to -\infty$ to $\alpha(t, \mathbf{x})$.

By virtue of Lemma B1 of Appendix B we have

$$\| u_{\tau_1}(t, \cdot) - u_{\tau_2}(t, \cdot) \|_E$$

$$\leq C_1 \| u_{\tau_1}(\tau_2, \cdot) - u_{\tau_2}(\tau_2, \cdot) \|_E$$

$$= C_1 \| u_{\tau_1}(\tau_2, \cdot) - \alpha(\tau_2, \cdot) \|_E. \qquad (2.35)$$

If we take the solution α_2 of the free Klein—Gordon equation having the same initial conditions at $t = \tau_2$ as the function u_{τ_1}, then

$$\| u_{\tau_1}(\tau_2, \cdot) - \alpha(\tau_2, \cdot) \|_E = \| \alpha_2(\tau_2, \cdot) - \alpha(\tau_2, \cdot) \|_E$$

$$= \| \alpha_2(t, \cdot) - \alpha(t, \cdot) \|_E.$$

We shall now apply the same arguments as in the proof of Eq. (2.29). We have for the function α the following integral representation:

$$\alpha(t, \mathbf{x}) = u_{\tau_1}(t, \mathbf{x}) - 3\lambda \int_{\tau_1}^\infty dt' \int d^3 \mathbf{y}$$

$$\times \Delta_R(t - t', \mathbf{x} - \mathbf{y}) \Phi^2(t', \mathbf{y}) u_{\tau_1}(t', \mathbf{y}). \qquad (2.36)$$

For the function α_2 we have a similar representation with τ_2 instead of τ_1 in the lower limit of integration. Hence for $\tau_1, \tau_2 < t$ we have

$$\alpha_2(t, \mathbf{x}) - \alpha(t, \mathbf{x}) = -3\lambda \int_{\tau_1}^{\tau_2} dt' \int d^3 \mathbf{y}$$

$$\times \Delta(t - t', \mathbf{x} - \mathbf{y}) \Phi^2(t', \mathbf{y}) u_{\tau_1}(t', \mathbf{y}). \qquad (2.37)$$

Applying now the same evaluations as in Eq. (2.29), one obtains

$$\| \alpha_2(t, \cdot) - \alpha(t, \cdot) \|_E$$

$$\leq C_0 \left| \int_{\tau_1}^{\tau_2} \sup_\mathbf{x} \Phi^2(t, \mathbf{x}) \, dt \right| \| u_{\tau_1}(\tau_1, \cdot) \|_E. \qquad (2.38)$$

Consequently, from Eqs. (2.35), (2.36) and the last inequality one obtains

$$\| u_{\tau_1}(t, \cdot) - u_{\tau_2}(t, \cdot) \|_E$$

$$\leq C_2 \| \alpha \|_2 \left| \int_{\tau_1}^{\tau_2} \sup_\mathbf{x} \Phi^2(t, \mathbf{x}) \, dt \right|, \qquad (2.39)$$

from which the required convergence of u_τ follows.

In particular the derivative $(\partial_t u_\tau)(0, \mathbf{x})$ is convergent in $L^2(R^3)$ and by virtue of Eqs. (2.19) and (2.33) we have

$$\frac{\delta}{\delta \varphi(\mathbf{x})} \int d^3 \mathbf{y} \, \Phi_{in}[t_0, \mathbf{y} \mid \varphi, \pi] \alpha(\mathbf{y})$$

$$= \lim_{\tau \to -\infty} (\partial_t u_\tau)(0, \mathbf{x}). \qquad (2.40)$$

The derivation of the functional derivative with respect to π is similar and we point out the main steps only. We have

$$\int d^3 \mathbf{x} \partial_s \Phi_\tau[t_0, \mathbf{x} \mid \varphi, \pi + sX] \beta(\mathbf{x})$$

$$= \int d^3 \mathbf{x} \partial_s \Phi[t, \mathbf{x} \mid \varphi + sX] \overleftrightarrow{\partial_t} \beta(t, \mathbf{x}) \big|_{t=\tau}. \qquad (2.41)$$

Defining ϑ_τ as above and using Eqs. (2.34) and (2.21), one obtains

$$\frac{\delta}{\delta \pi(\mathbf{x})} \int d^3 \mathbf{y} \, \Phi_{in}[t_0, \mathbf{y} \mid \varphi, \pi] \beta(\mathbf{y}) = - \lim_{\tau \to -\infty} \vartheta_\tau(0, \mathbf{x}). \qquad (2.42)$$

Here we have convergence in $\| \cdot \|_2$ norm in $L^2(R^3)$.

Using the same steps as in the derivation of formula (2.24), one obtains

$$\left\{ \int d^3 \mathbf{x} \, \Phi_{in \atop out}[t, \mathbf{x} \mid \varphi, \pi] \alpha(\mathbf{x}), \right.$$

$$\left. \int d^3 \mathbf{y} \, \Phi_{in \atop out}[r, \mathbf{y} \mid \varphi, \pi] \beta(\mathbf{y}) \right\}$$

$$= - \int d^3 \mathbf{x} \, d^3 \mathbf{y} \, \alpha(\mathbf{x}) \Delta(t - r, \mathbf{x} - \mathbf{y}) \beta(\mathbf{y}),$$

which gives the assertion of Theorem 3. \hfill QED

3. LOCALITY

It is evident from Eq. (2.26) that the asymptotic fields Φ_{in} and Φ_{out} are local. The free fields $\Phi_\tau(x)$, $-\infty < \tau < \infty$, which are given by Eq. (2.5) are also local, by virtue of Eq. (2.14). We show now that the interacting field $\Phi(x)$ is also local. In fact we have:

Proposition 4: Let f, $g \in C_0^\infty(R^4)$ and let supports of f and g be spacelike separated. Then

$$\{ \Phi(f), \Phi(g) \} = 0, \qquad (3.1)$$

Proof: Because

$$\{ \Phi(f), \Phi(g) \}$$

$$= \int dr \, ds \{ \int d^3 \mathbf{x} \, \Phi(r, \mathbf{x}) f(r, \mathbf{x}), \int d^3 \mathbf{y} \, \Phi(s, \mathbf{y}) g(s, \mathbf{y}) \}, \qquad (3.2)$$

it is sufficient to show that

$$\{\int d^3\mathbf{x}\,\Phi(r,\mathbf{x})\alpha(\mathbf{x}),\ \int d^3\mathbf{y}\,\Phi(s,\mathbf{y})\beta(\mathbf{y})\}$$

$$= 0,\quad \alpha,\beta \in C_0^\infty(R^3), \tag{3.3}$$

when for $\mathbf{x} \in \mathrm{supp}\,\alpha,\ \mathbf{y} \in \mathrm{supp}\,\beta,\ |r-s| < |\mathbf{x}-\mathbf{y}|.$ (3.4)

Denoting by u_r, as in (2.17), a solution of Eq. (2.9) with initial conditions on $t = r$ equal to $(0,\alpha(\mathbf{x}))$ and by ϑ_s a solution of (2.9) with initial conditions on $t = s$ equal to $(0,\beta(\mathbf{y}))$, one obtains by virtue of Eq. (2.23)

$$\{\int d^3\mathbf{x}\,\Phi(r,\mathbf{x})\alpha(\mathbf{x}),\ \int d^3\mathbf{y}\,\Phi(s,\mathbf{y})\beta(\mathbf{y})\}$$

$$= (u_r(t,\cdot),\vartheta_s(t,\cdot))_{t=0} = (u_r(t,\cdot),\vartheta_s(t,\cdot))_{t=r}$$

$$= \int d^3\mathbf{x}\,\alpha(\mathbf{x})\partial_r\vartheta_s(r,\mathbf{x}). \tag{3.5}$$

Since ϑ_s satisfies hyperbolic equation of motion (2.9), the initial conditions for ϑ_s and the condition (3.4) imply that $\mathrm{supp}\,\partial_r\vartheta_s(r,\cdot)\cap\mathrm{supp}\,\alpha(\cdot)=0$. This by (3.5) implies that (3.3) and consequently (3.1) is satisfied.

QED

Thus the classical nonlinear field theory equipped with the Lie algebra structure provided by Poisson brackets is a local field theory, in which the interacting field possesses the local asymptotic fields $\Phi_{\rm in}$ and $\Phi_{\rm out}$.

4. RELATIVISTIC COVARIANCE

The nonlinear equation (2.1) may be derived from the following Lagrangian density:

$$\mathcal{L}(x) = \tfrac{1}{2}(\Phi_{,\mu}\Phi^{,\mu} + m^2\Phi^2) - \tfrac{1}{4}\lambda\Phi^4. \tag{4.1}$$

Using the standard technique one derives the following form for the energy-momentum tensor associated with the density (4.1):

$$T_{\mu\nu}(x) = \Phi_{,\mu}(x)\Phi_{,\nu}(x) - g_{\mu\nu}\mathcal{L}(x). \tag{4.2}$$

Let σ be a spacelike surface in the Minkowski space. Then the integrals

$$P_\mu(\sigma) = \int_\sigma d\sigma^\nu T_{\mu\nu},$$

$$M_{\mu\nu}(\sigma) = \int_\sigma d\sigma^\lambda(x_\mu T_{\nu\lambda} - x_\nu T_{\mu\lambda}) \tag{4.3}$$

are constants of motion. One verifies, using Eq. (2.25), that the quantities (4.3) satisfy the following commutation relations:

$$\{P_\mu,P_\nu\} = 0,\quad \{M_{\mu\nu},P_\lambda\} = g_{\nu\lambda}P_\mu - g_{\mu\lambda}P_\nu,$$

$$\{M_{\mu\nu},M_{\lambda\rho}\} = g_{\mu\rho}M_{\nu\lambda} + g_{\nu\lambda}M_{\mu\rho} - g_{\mu\lambda}M_{\nu\rho} - g_{\nu\rho}M_{\mu\lambda}, \tag{4.4}$$

which are the standard commutation relations for generators of the Poincaré Lie algebra.

Let $\Phi_{\rm in}(x)$ be the free asymptotic field associated with the interacting field $\Phi(x)$. The Lagrangian density for the $\Phi_{\rm in}$ field is given by the formula

$$\mathcal{L}_{\rm in}(x) = \tfrac{1}{2}(\Phi_{{\rm in},\mu}\Phi_{\rm in}{}^{,\mu} + m^2\Phi_{\rm in}^2). \tag{4.5}$$

Using the formula (4.2) and (4.3), one can calculate the corresponding generators $P_\mu^{\rm in}$ and $M_{\mu\nu}^{\rm in}$ of the Poincaré group for the free field $\Phi_{\rm in}(x)$. Similarly one can calculate the generators $P_\mu^{\rm out}$ and $M_{\mu\nu}^{\rm out}$ associated with the free field $\Phi_{\rm out}(x)$.

The commutation relations of P_μ and $M_{\mu\nu}$ with the field $\Phi(x)$ are directly obtained by using formula (2.25). One obtains

$$\{\Phi(x),P_\mu\} = \partial_\mu\Phi(x),$$

$$\{\Phi(x),M_{\mu\nu}\} = (x_\mu\partial_\nu - x_\nu\partial_\mu)\Phi(x). \tag{4.6}$$

The global transformations $(a,\Lambda) \to U_{(a,\Lambda)}$ of the Poincaré group $\mathcal{P}$ in the subset $\tilde{J}$ consisting of all solutions of Eq. (2.1) are given by the formula

$$(U_{(a,\Lambda)}\Phi)(x) = \Phi[\Lambda^{-1}(x-a)]. \tag{4.7}$$

The generators P_μ and $M_{\mu\nu}$ of $U_{(a,\Lambda)}$ by virtue of Eq. (4.6) are given by formulas (4.3).

The field $\Phi_{\rm in}$ may be expressed as a following scalar functional of

$$\Phi_{\rm in}(x) = \Phi(x) - \lambda\int\Delta_R(x-y)\Phi^3(y)\,d^4y. \tag{4.8}$$

The transformation (4.7) induces the transformation

$$\Phi_{\rm in}(x) \to (U_{(a,\Lambda)}\Phi_{\rm in})(x) = \Phi_{\rm in}[\Lambda^{-1}(x-a)]$$

of the field $\Phi_{\rm in}$. Hence we have

$$U_{(a,\Lambda)} = U_{(a,\Lambda)}^{\rm in}, \tag{4.9}$$

$$P_\mu = P_\mu^{\rm in},\quad M_{\mu\nu} = M_{\mu\nu}^{\rm in}. \tag{4.10}$$

It is instructive to derive the equality (4.10) directly. We show this in detail for the generator P_0. By virtue of Eq. (4.3) for a spacelike surface $\sigma(t)$ perpendicular to the time axis, we have

$$P_0(\sigma(t)) = \tfrac{1}{2}\int_{\sigma(t)}d^3\mathbf{x}[\Pi^2(t,\mathbf{x}) + |\nabla\Phi(t,\mathbf{x})|^2$$

$$+ m^2\Phi^2(t,\mathbf{x}) - \tfrac{1}{2}\lambda\Phi^4(t,\mathbf{x})]. \tag{4.11}$$

We shall evaluate the expression (4.11) for $t \to -\infty$. For the interaction term $\lambda\Phi^4$ utilizing the fact that $|\Phi(t,\mathbf{x})| < C|t|^{-3/2}$ for large t and that $\int d^3\mathbf{x}\Phi^2(t,\mathbf{x})$ is smaller than the total energy E, we have

$$\lim_{t\to-\infty}\int d^3\mathbf{x}\,\Phi^4(t,\mathbf{x})$$

$$\leq \lim\max_{\mathbf{x}}\Phi^2(t,\mathbf{x})\int d^3\mathbf{x}\,\Phi^2(t,\mathbf{x})$$

$$\leq \lim_{t\to-\infty}CE|t|^{-3} = 0. \tag{4.12}$$

Hence by virtue of Eq. (2.3) we obtain

$$\lim P_0(\sigma(t)) = \lim\tfrac{1}{2}\|\Phi(t,\cdot)\|_E^2$$

$$= \tfrac{1}{2}\|\Phi_{\rm in}(t,\cdot)\|_E^2 = P_0^{\rm in}.$$

Because P_0 is time independent, we have $P_0 = P_0^{\rm in}$. The derivation of Eq. (4.10) for remaining generators may be performed in a similar manner.

Using the Yang–Feldman, equation

$$\Phi_{\rm out}(x) = \Phi(x) - \lambda\int\Delta_A(x-y)\Phi^3(y)\,d^4y \tag{4.13}$$

and Eq. (3.7) one shows by analogous considerations that

$$U_{(a,\Lambda)} = U_{(a,\Lambda)}^{\rm out} \tag{4.14}$$

and

$$P_\mu = P_\mu^{\rm out},\quad M_{\mu\nu} = M_{\mu\nu}^{\rm out}. \tag{4.15}$$

Clearly using asymptotic properties of Φ one may derive formula (4.15) directly as in case of the $\Phi_{\rm in}$ field.

5. SCATTERING OPERATOR

It follows from Theorem 3 that the scattering operator defined in the subspace F of B_F by the formula $S: \Phi_{in} \to \Phi_{out}$ is canonical. In addition, since

$$\|\Phi_{out}\|_E^2 = \|S(\Phi_{in})\|_E^2 = 2P_0^{out} = 2P_0^{in} = \|\Phi_{in}\|^2, \qquad (5.1)$$

the S operator is isometric. It was proven in Ref. 4 that the Yang—Feldman equation (4.8) has the unique solution given by the limit of the iterative series. Hence the solution $\Phi(x)$ can be written in the form

$$\Phi(x) = \{\Phi_{in} + \lambda N_R \{\Phi_{in} + \lambda N_R \{\cdots\}\}\}(x), \qquad (5.2)$$

where $N_R(\cdot)$ is the nonlinear operator in the Banach space B_F, given by the formula

$$N_R(\psi)(x) = \int \Delta_R(x - y)\psi^3(y)\, d^4y. \qquad (5.3)$$

The convergence of the iterative series (5.2) in B_F holds in the F norm defined by the formula (A1) of Appendix A. We shall write for simplicity the iterative series (5.2) in the form

$$\Phi(x) = [(I - \lambda N_R)^{-1}(\Phi_{in})](x). \qquad (5.4)$$

Using the Yang—Feldman equation in the form $\Phi = \Phi_{out} + \lambda N_A(\Phi)$, where $N_A(\Phi) \equiv \Delta_A * \Phi^3$ and subtracting it from Eq. (4.8), one obtains

$$\Phi_{out} = \Phi_{in} + \lambda N(\Phi), \qquad (5.5)$$

where $N(\Phi) = \Delta * \Phi^3$. Utilizing Eq. (5.4), we find the following form of the classical scattering operator:

$$\Phi_{out} = S(\Phi_{in}) = [I + \lambda N(I - \lambda N_R)^{-1}](\Phi_{in}). \qquad (5.6)$$

We see that S operator is nonlinear in J and different from the identity.[5]

Using the fact that

$$(U_{(a,\Lambda)}\Delta_R)(x - y) = \Delta_R(x - y),$$

$$(U_{(a,\Lambda)}\Delta)(x - y) = \Delta(x - y),$$

one obtains

$$N_R(U_{(a,\Lambda)}\Phi)(x) = [U_{(a,\Lambda)}(N_R\Phi)](x),$$

$$N(U_{(a,\Lambda)}\Phi)(x) = [U_{(a,\Lambda)}(N(\Phi))](x). \qquad (5.7)$$

This by virtue of Eq. (5.6) implies the Poincaré invariance of scattering operator, i.e.,

$$U_{(a,\Lambda)}S = SU_{(a,\Lambda)} \quad \text{in } J. \qquad (5.8)$$

Finally expressing the interacting field Φ in terms of Φ_{out} field by means of the formulas $\Phi = [I - \lambda N_A]^{-1}(\Phi_{out})$ one obtains that the inverse to the scattering operator is given by the formulas

$$\Phi_{in} = S^{-1}(\Phi_{out}) = [I - \lambda N(I - \lambda N_A)^{-1}](\Phi_{out}). \qquad (5.9)$$

The above analysis shows that the classical S operator is a good candidate for the construction of a quantum $\hat{S}$ operator associated with solutions of a quantized version of the dynamical equation (1.1). We consider this problem in Paper III of our work.

6. DISCUSSION

A. We formulated the classical theory in the language

of Lie algebra, whose commutators are defined in terms of Poisson brackets. In Paper II we construct an operator representation of this Lie algebra. We obtain in this manner an interacting, local, relativistic quantum field $\hat{\Phi}(x)$ which satisfies the asymptotic conditions.

B. The canonical formalism discussed in the present paper may be extended to a class of nonpolynomial interactions where $F(\cdot)$ satisfies the conditions

(i) $F(\cdot)$ is an even analytic function,

(ii) $F'(u) = O(u^3)$ as $u \to 0$,

(iii) $|F(u)u^{-5}| \to 0$ as $|u| \to \infty$.

The extension of the present results may be proven by using, in the proofs of Theorem 2 and 3, the corresponding results for a classical nonlinear relativistic wave equation with an analytic nonlinear term.[4]

APPENDIX A

We summarize here the properties of the Banach space B_F which is a carrier space for solutions of Eq. (2.1). Let $\Phi_0(x)$ be a solution of the free Klein—Gordon equation, whose Cauchy data at $t = 0$ coincide with that of Φ. Define J_1 as the space of free solutions such that $\Phi_0(0, \mathbf{x}) = \varphi(\mathbf{x}) = \Phi(0, \mathbf{x})$ has third derivatives in $L_1(R^3)$ and second derivatives in $L_2(R^3)$, while $\Pi_0(0, \mathbf{x}) = \pi(\mathbf{x}) = \Pi(0, \mathbf{x})$ has second derivatives in $L_1(R^3)$ and first derivatives in $L_2(R^3)$. Then every element of J_1 is finite with respect to the following norm[4]:

$$\|\Phi\|_F^2 = \sup_t \|\Phi(t)\|_E^2 + \sup_{t,\mathbf{x}} |\Phi(t, \mathbf{x})|^2$$
$$+ \int_{-\infty}^{\infty} \sup_{\mathbf{x}} |\Phi(t, \mathbf{x})|^2 \, dt. \qquad (A1)$$

Denote by J the completion of J_1 in the norm (A1) and by B_F the Banach space of all functions with finite F norm. Then the main theorem of Ref. 4 asserts that if $\Phi_0(x) \in J$, then there exists a unique solution $\Phi(x)$ of Eq. (2.1) which is in B_F and free solutions $\Phi_{in}(x)$ and $\Phi_{out}(x)$ such that

$$\Phi_{in}(t, \mathbf{x}) \xleftarrow[t \to -\infty]{} \Phi(t, \mathbf{x}) \xrightarrow[t \to \infty]{} \Phi_{out}(t, \mathbf{x}) \qquad (A2)$$

in the energy norm (2.4). Moreover, if $\Phi_0(x) \in J_1$, then $\Phi(t, \mathbf{x})$ is uniformly $O(|t|^{-3/2})$.

APPENDIX B

We derive an useful norm inequality for solutions of Eq. (2.9).

Lemma B1: Let $\psi(t, \mathbf{x})$ be a regular and finite energy norm for $t = t_0$ solution of the equation

$$(\Box + m^2)\psi(t, \mathbf{x}) + V(t, \mathbf{x})\psi(t, \mathbf{x}) = f(t, \mathbf{x}), \qquad (B1)$$

1480   J. Math. Phys., Vol. 16, No. 7, July 1975

T. Bałaban and R. Rączka   1480

where $V(t, \mathbf{x})$ is a smooth function satisfying the condition

$$\int_{-\infty}^{\infty} \sup_{\mathbf{x}} |V(t, \mathbf{x})| \, dt < \infty \tag{B2}$$

and $f(t, \mathbf{x})$ has finite norm

$$\int_{-\infty}^{\infty} \|f(t, \cdot)\|_2 \, dt.$$

Then ψ has finite energy norm for all t and

$$\|\psi(t, \cdot)\|_E^2 \leq 4 \exp\left(\frac{2}{m} \int_{-\infty}^{\infty} \sup_{\mathbf{x}} |V(t, \mathbf{x})| \, dt\right)$$
$$\times \left[\|\psi(t_0, \cdot)\|_E^2 + \left(\int_{t_0}^{t} \|f(t', \cdot)\|_2 \, dt'\right)^2\right]. \tag{B3}$$

Proof: It is sufficient to prove the lemma for a solution ψ having a compact support with respect to the variable $\mathbf{x}$. Multiplying Eq. (B1) by $\partial_t \psi$ and integrating over $\mathbf{x}$, we obtain

$$\partial_t \int d\mathbf{x} [(\partial_t \psi)^2 + |\nabla \psi|^2 + m^2 \psi^2]$$
$$+ 2 \int d\mathbf{x} \, V(t, \mathbf{x}) \psi \partial_t \psi = 2 \int d^3 \mathbf{x} f(t, \mathbf{x}) \partial_t \psi.$$

Integrating this equality with respect to t in the interval $[t_0, t_1]$ we obtain

$$\|\psi(t_1, \cdot)\|_E^2$$
$$= \|\psi(t_0, \cdot)\|_E^2 - 2 \int_{t_0}^{t_1} dt \int d\mathbf{x} \, V(t, \mathbf{x}) \psi \partial_t \psi$$
$$+ 2 \int_{t_0}^{t_1} dt \int d\mathbf{x} f(t, \mathbf{x}) \partial_t \psi$$
$$\leq \|\psi(t_0, \cdot)\|_E^2 + m^{-1} \left| \int_{t_0}^{t_1} dt \sup_{\mathbf{x}} |V(t, \mathbf{x})| \right|$$
$$\times \|\psi(t, \cdot)\|_E^2 + 2 \left| \int_{t_0}^{t_1} dt \|f(t, \cdot)\|_2 \|\psi(t, \cdot)\|_E \right|. \tag{B4}$$

Denoting

$$\rho(t) = \sup_{t' \in [t_0, t]} \|\psi(t', \cdot)\|_E^2, \quad \sigma(t) = 2m^{-1} \sup_{\mathbf{x}} |V(t, \mathbf{x})|,$$

$$c = \left(2 \int_{t_0}^{t_1} dt \|f(t, \cdot)\|_2\right)^2,$$

we have the inequality

$$\rho(t_1) \leq 2\rho(t_0) + \left| \int_{t_0}^{t_1} \sigma(t)\rho(t) \, dt \right| + c.$$

The n-fold iteration of this inequality gives

$$\rho(t_1) \leq 2\rho(t_0) + c + \sum_{k=1}^{n} \left| \int_{t_0}^{t_1} ds_1 \int_{t_0}^{s_1} ds_2 \cdots \right.$$
$$\times \int_{t_0}^{s_{k-1}} ds_k \, \sigma(s_1)\sigma(s_2) \cdots \sigma(s_k)(2\rho(t_0) + c) \Big|$$
$$+ \left| \int_{t_0}^{t_1} ds_1 \cdots \int_{t_0}^{s_{n-1}} ds_n \sigma(s_1) \cdots \sigma(s_n)\rho(s_n) \right|,$$

but

$$\int_{t_0}^{t_1} ds_1 \cdots \int_{t_0}^{s_{k-1}} ds_k \sigma(s_1) \cdots \sigma(s_k) = (k!)^{-1} \left(\int_{t_0}^{t_1} \sigma(t) \, dt\right)^k,$$

$$\int_{t_0}^{t_1} ds_1 \cdots \int_{t_0}^{s_{n-1}} ds_n \sigma(s_1) \cdots \sigma(s_n)\rho(s_n) \underset{n \to \infty}{\to} 0;$$

hence

$$\rho(t_1) \leq \exp\left(\left| \int_{t_0}^{t_1} \sigma(t) \, dt \right|\right)(2\rho(t_0) + c),$$

what proves inequality (3). \hfill QED

[1](a) I.E. Segal, J. Math. Phys. 1, 468 (1960); (b) I.E. Segal, J. Math. Physics 5, 269 (1964); (c) I.E. Segal, J. de Math. 44 71, 107, (1965), Parts I, II; (d) I.E. Segal, "La variete des solutions d'une equation hyperbolique, nonlineare à ordre 2. III," Lecture notes, College de France, 1964-1965.
[2]R.F. Streater, Comm. Math. Phys. 2, 354 (1966).
[3]R. Rączka, J. Math. Phys. 16, 173 (1975).
[4]C.S. Morawetz and W.A. Strauss, Comm. Pure Appl. Math. 25. 1 (1972).
[5]C.S. Morawetz and W.A. Strauss, Comm. Pure Appl. Math. 26, 47 (1973).

Spectrum of the Kirkwood–Salsburg operator and phase transitions

William Klein*

Mathematics Department, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received 24 May 1974)

We prove that the spectrum of the Kirkwood–Salsburg equation in finite volumes is composed only of eigenvalues.

INTRODUCTION

In 1942 Mayer[1] suggested that phase transitions might be connected with bifurcation of the linear operators which relate the equilibrium distribution functions to each other. Other authors[2] have attempted to follow up this idea in various ways. Unfortunately, these attempts have required a multilation of whatever operator was being considered and, hence, weakened the argument.

In this paper we will put forth an approach which we hope will eventually lead to a rigorous justification of Mayer's conjecture.

In the main body of the paper we prove that the only spectral values of the Kirkwood–Salsburg (K–S) operator in compact volumes are eigenvalues.

In the discussion we also point out that those values of z, the activity, for which there exist solutions of the K–S equations which do not have full translational symmetry must belong to the spectrum of the K–S operator.

As was pointed out by Ruelle[3] the solutions of the K–S equation in finite volumes are elements of the linear vector space

$$\begin{pmatrix} \rho_1(x_1) \\ \rho_2(x_1, x_2) \\ \cdot \\ \cdot \\ \cdot \\ \rho_N(x_N) \end{pmatrix}$$

where the $\rho_N(\{x_N\})$ are bounded measurable functions. This is a Banach space under the norm

$$\sup_{\substack{\forall n \\ \forall \{x_n\}}} |\rho_n(\{x_n\})|.$$

The K–S operator is to be considered as acting on such a space.

I. REDEFINITION AND MODIFICATION OF THE K–S EQUATION

Before proving any theorems about the spectrum of the K–S operator, the operator must be defined more precisely. The ambiguity arises because of the symmetry under variable permutation of the expected solutions of the K–S equation.

The addition we will use will have the following form. Before inserting a vector

$$\zeta = \begin{pmatrix} \zeta_1(x_1) \\ \vdots \\ \zeta_N(\{x_N\}) \\ \vdots \end{pmatrix}$$

into the rhs of (II. 2), the variables will be cyclically permuted, i.e.,

$$\zeta_N(x_i, \ldots, x_N) \to \zeta_N(x_2, \ldots, x_N, x_i).$$

If $\zeta_N(\{x_N\})$ is to be multiplied by a kernel and integrated with respect to some of its variables, then x_1 will be replaced by a dummy variable.

If $\zeta_N(\{x_N\})$ is to be operated on as the first term in the sum in (II. 2), then

$$\zeta_{N-1}(x_1, \ldots, x_{N-1}) \to \zeta_{N-1}(x_2, \ldots, x_{N-1}, x_1)$$

and x_1 is replaced by x_N.

Clearly, the solutions of the original K–S equation of physical interest will also satisfy the equation as modified. [5]

The theorem we wish to prove is about the spectrum of the K–S operator on compact subspaces in $\mathbb{R}^N$ so that χ_Ω is to be thought of as a characteristic function of a region compact in $\mathbb{R}^N$.

II. SPECTRUM OF THE K–S OPERATOR

In everything that follows the K–S equation and operator will be understood to be modified as described in the previous section.

Theorem 1: The operator (K–S)

$$z\chi_\Omega K$$

has, for potentials which satisfy (II. 3) and the conditions

1. $\lim_+ \phi(r)$ and $\lim_- \phi(r)$,
   $\quad r \; r_0 \qquad\qquad r \; r_0$

exist and are not equal only at a finite number of r_0,

2. $\int |f_{ij}| \, dx_{ij} < \infty$,

3. $\phi(x_i - x_j) = \infty$, $\quad |x_i - x_j| < \sigma$, $\quad \sigma > 0$,

4. $\phi(x_i - x_j) > -B$, $\quad 0 < B < \infty$,

5. $\phi(x_i - x_j) = \phi(|x_i - x_j|)$,

6. $\phi(|x_i - x_j|) = 0$, $\quad |x_i - x_j| > \gamma < \infty$,

a spectrum composed entirely of eigenvalues.

Lemma 1: If we define an operator

$$z\chi_\Omega K^1 \begin{pmatrix} f_1(x_1) \\ \vdots \\ f_N(\{x_N\}) \\ \vdots \end{pmatrix} = \chi_\Omega \begin{pmatrix} 0 \\ (1+f_{i2})f_1(x_2) \\ \vdots \\ \prod_{j=2}^{N}(1+f_{ij})f_{N-1}(x_2\cdots x_N) \end{pmatrix} \quad (\text{II. 1})$$

(the variable displacement as was defined in the previous section) and an operator

$$z\chi_\Omega K_0 \begin{pmatrix} f_1(x_1) \\ \vdots \\ f_N(\{x_N\}) \end{pmatrix} = z\chi_\Omega \begin{pmatrix} f_1^1(x_1) \\ \vdots \\ f_N^1(\{x_N\}) \end{pmatrix}, \quad (\text{II. 2})$$

$f_1^1(x_1)$

$$= \chi_\Omega \sum_{n=1}^{\infty} \frac{1}{n!} \int f_n(x_2\cdots x_n, x_{n+1}) \prod_{j=2}^{n} f_{ij}\,dx_j,$$

$f_N^1(x_2\cdots x_N, x_1)$

$$= \chi_\Omega \prod_{j=2}^{N}(1+f_{ij}) \sum_{n=1}^{\infty} \frac{1}{n!} \int f_{N+n-1}(x_2\cdots x_N, \ldots, x_{N+n})$$

$$\otimes \prod_{j=N+1}^{N+n} f_{ij}\,dx_j,$$

where $f^1(x_2\cdots x_N, x_1)$ indicates the permutation that has taken place, then the K—S operator

$$z\chi_\Omega K$$

can be written as

$$z\chi_\Omega K = z\chi_\Omega K^1 + z\chi_\Omega K_0.$$

This is seen trivially by inspection. [5b]

Lemma 2: For all finite values of z the operator

$$I - z\chi_\Omega K^1$$

has a bounded inverse on the space of vectors

$$\zeta = \begin{pmatrix} \zeta_1(x_1) \\ \vdots \\ \zeta_N(\{x_N\}) \end{pmatrix},$$

N finite. This inverse maps this space onto itself. Note we are restricting the inverse to map into the space with the same number N of components.

Proof: Let

$$\phi = \begin{pmatrix} \phi_1(x_1) \\ \vdots \\ \phi_N(\{x_N\}) \end{pmatrix}$$

be an arbitrary vector in the specified space. Is

$$I - z\chi_\Omega K^1 \quad (\text{II. 3})$$

onto? We have operating with (II. 3) on ζ

$$\zeta_1(x_1) = \phi_i(x_1),$$

$$\zeta_2(x_1, x_2) - z\chi_\Omega(1+f_{12})\,\zeta_1(x_2) = \phi_2(x_1, x_2) \quad (\text{II. 4})$$

$$\vdots$$

$$\zeta_N(x_1\cdots x_N) - z\chi_\Omega \prod_{j=2}^{N}(1+f_{ij})\,\zeta_{N-1}(x_2\cdots x_N) = \phi_N(\{x_N\}).$$

Clearly, there exists a ζ such that the above set of equations is satisfied and it is bounded. If ϕ is picked to be the vector

$$\phi = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

then it is clear from (II. 4) that

$$\zeta = \begin{matrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{matrix}$$

is the only solution. Hence

$$I - z\chi_\Omega K^1$$

is one to one. The theorem is proved.

Lemma 3: If

$$I - z\chi_\Omega K^1$$

has an inverse, then the spectrum of

$$\lambda I - z\chi_\Omega K^1 - z\chi_\Omega K_0 = I - z\chi_\Omega K$$

is identical to the spectrum of

$$\lambda I - (I - z\chi_\Omega K')^{-1}z\chi_\Omega K_0$$

for $\lambda = 1$.

Proof: Since

$$I - (I - z\chi_\Omega K^1)^{-1} = A$$

maps the space of interest onto itself and is one to one, then we have the following:

1. If $(I - zK)$ is onto, i.e.,

$$(I - zK)\zeta = \alpha$$

has at least one solution for arbitrary α in the space, then

$$(I - zK)\,\zeta - \alpha = 0,$$

$$A[(I - zK)\,\zeta - 2] = 0,$$

$$(I - zA\chi_\Omega K_0)\,\zeta = A\alpha.$$

Since A is an onto mapping, clearly,

$$I - zA\chi_\Omega K_0$$

is onto.

2. If $I - z\chi_\Omega K$ is one to one, then

$$I - zA\chi_\Omega K_0$$

must be one to one because if it were not then

$$(I - zA\chi_\Omega K_0)\,\phi = 0, \quad \phi \neq 0;$$

then

$$A^{-1}(I - zA\chi_\Omega K_0)\,\phi = 0 \quad \text{or} \quad (I - z\chi_\Omega K)\,\phi = 0.$$

The converse of these are trivial to prove in exactly the same way as above. This clearly proves that the spectral values are the same. In the same manner as above, we can also trivially show that if z is such that

$$I - z\chi_\Omega K$$

is not one to one, then

$$I - zA\chi_\Omega K_0$$

is also not one to one, and the converse. It is also trivial to show the same relationship if

$$I - zA\chi_\Omega K_0$$

is not onto and the converse.

Before proving the next lemma it will be useful to introduce the following notation. An application of the operator $zA\chi_\Omega K_0$ to

$$\zeta = \begin{pmatrix} \zeta_1(x_1) \\ \vdots \\ \zeta_N(x_1 \cdots x_N) \end{pmatrix}$$

will produce a vector

$$\zeta = \begin{pmatrix} \zeta_1^1(\bar{x}_1) \\ \zeta_2^2(x_2, \bar{x}_1) \\ \vdots \\ \zeta_N^1(x_2 \cdots x_N, \bar{x}_1) \end{pmatrix}.$$

The bar over x_1 in the ζ_N^1 denotes that the variable dependence x_1 of the original ζ_N has been replaced by an integration variable and that the variable dependence x_1 of ζ_N^1 is expressed through a generalized kernel. This will be clear if we examine (II.1) and (II.2) for a space

$$\zeta = \begin{pmatrix} \zeta_1(x_1) \\ \zeta_2(x_1, x_2) \end{pmatrix}.$$

Upon application of (II.2) we have

$$\zeta_{1K_0}^1(\bar{x}_1)$$
$$= z\chi_\Omega [\int \zeta_1(x_2) f_{12} dx_2 + \int \zeta_2(x_2, x_3) f_{12} f_{13} dx_2 dx_3],$$
$$\zeta_{2K_0}^1(x_2, \bar{x}_1)$$
$$= z\chi_\Omega (1 + f_{12}) \int \zeta_2(x_2, x_3) f_{13} dx_3. \tag{II.5}$$

It is easily seen that the only dependence on x_1 of $\zeta_{K_0}^1$ is in a generalized kernel but not in ζ. The same of course cannot be said of the dependence on x_2. Applying

$$(I - z\chi_\Omega K^1)^{-1},$$

we have

$$\zeta_1^1(\bar{x}_1) = \zeta_{1K_0}^1(\bar{x}_1),$$
$$\zeta_2^1(x_2, \bar{x}_1) = \zeta_{2K_0}^1(x_2, \bar{x}_1) + z\chi_\Omega(1 + f_{12}) \zeta_{1K_0}^1(\bar{x}_2).$$

We retain the same properties of the variable dependence.

Lemma 4: If ζ^n is the result of applying $z(I - z\chi_\Omega K^1)^{-1}\chi_\Omega K_0$ n times to ζ, then the last n variables will have their dependence expressed through the kernels (of a generalized form) of the operator [i.e., $\zeta_4^3(x_4, \bar{x}_1, x_2, x_3)$]. If $n \geq N$, the number of variables, then ζ_N^n has all of its variable dependence of the specified form.

Proof: The proof proceeds by induction. For $n = 1$, an application of $z\chi_\Omega K_0$ [Eq. (II.2)] clearly produces a vector we will label $\zeta_{K_0}^1$ which has the required property.

The application of $(I - z\chi_\Omega K^1)^{-1}$ to $\zeta_{K_0}^1$ will map elements

$$\zeta_{N_{K_0}}^1(x_2 \cdots x_N, \bar{x}_1) \rightarrow \zeta_N^1(x_2 \cdots x_N, \bar{x}_1),$$

where

$$\zeta_N^1(x_2, \ldots, x_N, \bar{x}_1)$$
$$= \zeta_{N_{K_0}}^1(x_2 \cdots x_N, \bar{x}_1) + z\chi_\Omega \prod_{j=2}^N (1 + f_{ij}) \zeta_{N-1_{K_0}}^1$$
$$\times (x_3, \ldots, x_{N-1}, x_N, \bar{x}_2).$$

This clearly has the required property.

We assume that ζ^n has the required property, i.e.,

$$\zeta_N^n(\{x_N\}) = \zeta_N^n(x_{n+1}, \ldots, x_N, \bar{x}_1 \cdots \bar{x}_n)$$

and prove that ζ_N^{n+1} has this property. Applying the operator $z\chi_\Omega K_0$ requires that first

$$\zeta_N^n(\{x_N\}) \rightarrow \zeta_N^n(x_{n+2}, \ldots, x_N, x_1, \bar{x}_2, \ldots, \bar{x}_{n+1}).$$

Variable x_1 is now replaced by the dummy variable and the proper kernel is multiplied by $\zeta_N^n(\{x_n\})$ and integrated. This process produces a vector $\zeta_{K_0}^{n+1}$ which has the required property. An application of $(I - z\chi_\Omega K')^{-1}$ takes

$$\zeta_{N_{K_0}}^{n+1}(x_{n+2}, \ldots, x_N, \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{n+1})$$
$$\rightarrow \zeta_{N_{K_0}}^{n+1}(x_{n+2}, \ldots, x_N, \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{n+1})$$
$$- z\chi_\Omega \prod_{j=2}^N (1 + f_{ij}) \otimes \zeta_{N-1}^{n+1}(x_{n+3}, \ldots, x_N, \bar{x}_2, \ldots, \bar{x}_{n+2}).$$

This proves the theorem.

As this point is crucial it might help the reader if we work out the example begun in Eq. (II.5). Equation (II.5) shows that $\zeta_{1K_0}^1(\bar{x}_1)$ and $\zeta_{2K_0}^1(x_2, \bar{x}_1)$ have the form specified for one application of $z\chi_\Omega K_0$. Applying $(I - z\chi_\Omega K')^{-1}$ gives

$$\zeta_1^1(\bar{x}_1) = \zeta_{1K_0}^1(\bar{x}_1),$$
$$\zeta_2^1(x_2, \bar{x}_1) = \zeta_{2K_0}^1(x_2, \bar{x}_1) + z\chi_\Omega(1 + f_{12}) \zeta_{1K_0}^1(\bar{x}_2), \tag{II.6}$$

which also have the proper form since the variable x_1 dependence will no longer be contained in the original ζ vector. Applying $z\chi_0 K_0$ to ζ^1 gives

$$\zeta_{1K_0}^2(\bar{x}_1)$$
$$= z\chi_\Omega \int \zeta_1^1(\bar{x}_2) f_{12} dx_{12} dx_2 + \int \zeta_2^1(x_3, \bar{x}_2) f_{12} f_{13} dx_3 dx_3,$$
$$\zeta_{2K_0}^2(\bar{x}_1, \bar{x}_2)$$
$$= z\chi_\Omega (1 + f_{12}) \int \zeta_2^1(x_3, \bar{x}_2) f_{13} dx_3.$$

Since

$$\zeta_1^2(\bar{x}_1) = \zeta_{1K_0}^2(\bar{x}_1)$$

and

$$\zeta_2^2(\bar{x}_1, \bar{x}_2) = \zeta_{2K_0}^2(\bar{x}_1, \bar{x}_2) + z\chi_\Omega(1 + f_{12}) \zeta_{1K_0}^2(\bar{x}_2),$$

making the substitutions from (II.5) and (II.6) finally gives

$$\zeta_1^2(\bar{x}_1) = z\chi_\Omega \int z\chi_\Omega \int \zeta_1(x_3) f_{23} dx_3$$
$$+ \int \zeta_2(x_4, x_5) f_{24} f_{25} dx_4 dx_5 \, f_{12} dx_2$$
$$+ \int z\chi_\Omega(1 + f_{23}) \int \zeta_2(x_3, x_4) f_{24} dx_4 + z\chi_\Omega(1 + f_{23})$$
$$\times z\chi_\Omega \int \zeta_1(x_4) f_{34} dx_4 + \int \zeta_2(x_4, x_5) f_{34} f_{35} dx_4 dx_5 \, f_{12} f_{13},$$
$$\otimes dx_2 dx_3,$$

$$\zeta_2^2(\bar{x}_1, \bar{x}_2) = z\chi_\Omega(1 + f_{12}) \int z\chi_\Omega(1 + f_{23}) \int \zeta_2(x_3, x_4) f_{24} dx_4$$

$$+ z\chi_\Omega(1 + f_{23}) z\chi_\Omega \int \zeta_1(x_4) f_{34} dx_4$$

$$+ \int \zeta_2(x_4, x_5) f_{34} f_{35} dx_4 dx_5 f_{13} dx_3$$

$$+ z\chi_\Omega(1 + f_{12}) z\chi_\Omega \int dx_2 f_{23} z\chi_\Omega \int \zeta_1(x_4) f_{34} dx_4$$

$$+ \int \zeta_2(x_4, x_5) f_{34} f_{35} dx_4 dx_5 + z\chi \int z\chi_\Omega(1 + f_{34})$$

$$\times \int \zeta_2(x_4, x_5) f_{35} dx_5 f_{23} f_{24} dx_3 dx_4.$$

One can see from the above that the variable dependence of the original ζ vector appears in the above tow equations merely as an integration variable.

Before proving the next lemma we will introduce some simple terminology. We will call any function of the form

$$\int_V \zeta(\{x_j\}) \prod_{x_j \in \{x_N\}} \prod_{x_k \in \{x_m\}} f(x_j - x_k) dx_j$$

a ζ graph. The finite set $\{x_m\}$ will be called the root points and the finite set of integrated variables the field points. We also specify that $\zeta(\{x_N\})$ is a bounded function. V is an arbitrary but finite volume.

Lemma 5: Any ζ graph is a uniformly continuous function of its root points for the potentials specified in Theorem 1.

Proof: From the conditions imposed in Theorem 1 of this section

$$f(x_i - x_j) = \sum_{\alpha=0}^P h_\alpha O(\sigma_\alpha - x_{ij}).$$

Where

$$x_{ij} = |x_i - x_j|,$$

$$O(\sigma_0 - x_{ij})$$

is a continuous function of x_{ij},

$$h_0 = 1$$

and

$$\alpha \geq 1 \, O(\sigma_\alpha - x_{ij}) = \begin{cases} 1, & x_{ij} \leq \sigma_\alpha, \\ 0, & x_{ij} > \sigma_\alpha, \end{cases}$$

$$h_\alpha < c \quad \text{finite} \quad \forall \alpha.$$

An arbitrary product of f_{ij} can be written as a sum of products of $O(\sigma_\alpha - x_{ij})$. Since the product of f_{ij} is finite (i. e. , has a finite number of terms) the products of $O(\sigma_\alpha - x_{ij})$ have a finite number of terms as does the sum

$$\left| \int_V \zeta_N(y_1 \cdots y_N) \prod_{j,k} f(x_j - y_k) dy_k - \zeta_N(y_1 \cdots y_N) \prod_{j,k} (x_j' - y_k) \right.$$

$$\otimes \prod_k dy_k \Big|$$

$$\leq C \int_V \left| \prod_{j,k} f(x_j - y_k) - \prod_{j,k} f(x_j' - y_k) \right| dy_k$$

$$\leq C \sum_P \int_V \left| \prod_{j,k,\alpha} h_\alpha O(\sigma_\alpha - |x_j - y_k|) \right.$$

$$- \prod_{j,k,\alpha} O(\sigma_\alpha - |x_j' - y_k'|) \Big| \otimes \prod_k dy_k,$$

where $\sum_p$ denotes the sum over all possible products. Let us look at one such product:

$$\int_V \left| \prod_{j,k,\alpha} h_\alpha O(\sigma_\alpha - |x_j - y_k|) - \prod_{j,k,\alpha} h_\alpha O(\sigma_\alpha - |x_j' - y_k|) \right|$$

$$\otimes \prod_k dy_k. \tag{II. 7}$$

Clearly,

$$\left| \prod_{j,k} O(\sigma_0 - |x_j - y_k|) \right| < B < \infty,$$

where B is some positive constant. Therefore, the integrand of (II. 7) can be bounded by

$$\left| \prod_\alpha h_\alpha \right| \left(\sum_{\substack{\alpha, j, k \\ \alpha \neq 0}} |O(\sigma_\alpha - |x_j - y_k|) - O(\sigma_\alpha - |x_j' - y_k|)| B \right.$$

$$+ \left| \prod_{j,k} O(\sigma_0 - |x_j - y_k|) - O(\sigma_0 - |x_j' - y_k|) \right| \Bigg).$$

Therefore, (II. 7) is less than or equal to

$$\left| \prod_\alpha h_\alpha \right| \left| \int_V \left| \prod_{j,k} O(\sigma_0 - |x_j - y_k|) - O(\sigma_0 - |x_j' - y_k|) \right| \prod_k dy_k \right.$$

$$\times V^S \left| \prod_\alpha h_\alpha \right| \sum_{j,k,\alpha} \int_V |O(\sigma_\alpha - |x_j - y_k|) - O(\sigma_\alpha$$

$$- |x_j' - y_k|)| dy_k. \tag{II. 8}$$

The integrand of the first integral in (II. 8) is a continuous function (also bounded on compact support) and therefore the first integral is clearly zero uniformly in the $\lim x_j' \to x_j$. The integrand in the second integral the characteristic function of the volume of the union of two spheres minus their intersection. The two spheres have radius σ_α and are centered at x_j and x_j'. In the limit as $x_j' \to x_j$ the volume goes to zero and hence the integral. Since V^S (S equals the number of field points minus 1) and $\prod_\alpha |h_\alpha|$ are finite, and all sums and products are finite, the lemma is proved.

Lemma 6: Given a bounded sequence

$$||\zeta_{N_n}(\{x_N\})|| < c \quad \forall n,$$

where

$$||\zeta_N(\{x_N\})|| \equiv \sup_{\forall \{x_N\}} |\zeta_N(\{x_N\})|,$$

the sequence of ζ graphs, which we will call γ_n, is conditionally compact in the topology generated by the above norm.

Proof: Lemma 5 proves uniform continuity over a compact support. The fact that

$$||\zeta_{N_n}(\{x_N\})|| < c$$

then guarantees equicontinuity over compact support. By the Ascoli—Arzela[7] theorem the sequence γ_n is conditionally compact in the norm topology.

Lemma 7: Any bounded function multiplied by γ_n gives a sequence which is conditionally compact.

Proof: Since γ_n conditionally compact means that there exists a subsequence which converges in the norm to an element in the space of continuous functions, we must show that

$$T\gamma_{n_K} \to \gamma$$

in the space of bounded functions. We have

$$\lim_{n_K \to \infty} ||\gamma_{n_K} - \gamma^1|| = 0,$$

$$\lim_{n_K \to \infty} ||T\gamma_{nk} - T\gamma^1|| \leq B \lim_{n_K \to \infty} ||\gamma_{nk} - \gamma^1|| = 0,$$

where B is the bound for the $||T||$. The lemma is proved.

At this point we note that due to the presence of the hard core in the potential a volume Λ can have only a finite number of particles. Let us call this number N. Clearly the vector of distribution functions ζ will have only N components. From here on we consider all operators operating from the space of vectors

$$\zeta = \begin{pmatrix} \zeta_1(x_1) \\ \vdots \\ \zeta_N(x_N) \end{pmatrix}$$

with N components into the same space.

Lemma 7: The operator

$$(z(I - z\chi_\Omega K')^{-1} \chi_\Omega K_0)^N$$

maps a sequence of bounded vectors

$$\zeta_n = \begin{pmatrix} \zeta_{1n}(x_1) \\ \vdots \\ \zeta_{N_n}(\{x_n\}) \end{pmatrix},$$

$$||\zeta_n|| = \sup_{\substack{\forall \{x_N\} \\ \forall N}} |\zeta_{N_n}(\{x_N\})| < c \ \forall n$$

into a sequence of vectors

$$\phi_n \begin{pmatrix} \phi_{1n}(x_1) \\ \vdots \\ \phi_{N_n}(\{x_N\}) \end{pmatrix},$$

where each $\phi_{N_n}(\{x_N\})$ is a sum of a finite number of sequences of ζ graphs multiplied by bounded functions. (Note: the power N of the operator and the N of the last elements in the vectors ζ_n and ϕ_n are the same).

Proof: This is a trivial consequence of Lemma 4. This means that each $\phi_N(\{x_N\})$ is the sum of conditionally compact sequences in the space of bounded functions with the norm

$$||\phi_N(\{x_N\})|| = \sup_{\forall \{x_N\}} |\phi_N(\{x_N\})|.$$

Before proceeding to the next lemma we will want to define the concept of an ϵ net and state a well known theorem.

Definition: Let $\epsilon > 0$ be a given positive number and M a subset of the normed space X. The set M is called an ϵ net for the set E if there exists for every point X in E a point z in M such that

$$||x - z|| < \epsilon.$$

Theorem (Hausdorff): A necessary condition for a subset E of a normed space X to be conditionally compact is that for each $\epsilon > 0$ there exists in X a finite ϵ net

for E. The condition is also sufficient if X is a complete space.

Lemma 8: Finite sums of conditionally compact sequences in complete normed spaces are conditionally compact.

Proof: Assume there are P such sequences; then there exists an ϵ/P net for each sequence. We can then make an ϵ net for the sum of the P sequences by creating the set of all possible sums of the P ϵ/P nets one from each net. Since the space is complete by the Hausdorff theorem the sum is conditionally compact.

Lemma 9: A sequence of vectors (N finite)

$$\phi_n \begin{pmatrix} \phi_{1n}(x_1) \\ \vdots \\ \phi_{N_n}(\{x_N\}) \end{pmatrix},$$

where $\phi_{jn}(\{x_N\})$ are bounded and ϕ_n is an element of the Banach space with the norm

$$||\phi_n|| = \sup_{\substack{\forall \{x_N\} \\ \forall N}} |\phi_{N_n}\{x_N\}|,$$

is conditionally compact in the norm induced topology if each of the $\phi_{jn}\{x_j\}$ is conditionally compact in the respective norms:

$$||\phi_{jn}\{x_j\}|| = \sup_{\forall \{x_j\}} |\phi_{jn}\{x_j\}|.$$

Proof: Construct a finite ϵ net for the ϕ_n from all possible vectors of the form

$$\begin{pmatrix} \gamma_{1i}(x_1) \\ \vdots \\ \gamma_{NK}(\{x_N\}) \end{pmatrix},$$

where the γ_N's are members of the finite ϵ nets for the ϕ_{N_n}'s.

We have proven with all these lemmas that

$$(z(I - z\chi_\Omega)^{-1} \chi_\Omega K_0)^N \tag{II. 9}$$

maps bounded sequences

$$\zeta_n = \begin{pmatrix} \zeta_1(x_1) \\ \vdots \\ \zeta_N(\{x_N\}) \end{pmatrix}$$

into conditionally compact sequences. Hence the operator (II. 9) is compact. [9] But operators with compact products have the same spectrum as if they were compact. [10] Invoking Lemma 3 proves the theorem which we now restate.

The K—S operator on compact support has a spectrum composed only of eigenvalues.

III. RESULTS AND CONCLUSIONS

The theorem proven in the previous section is useful in two respects. (1) It indicates a possible path that might be taken to prove something rigorous about the

spectrum of the K—S operator in the thermodynamic limit. (2) The inclusion of the variable permutation in the definition of the K—S operator makes quite clear the importance of the symmetry of the distribution functions under variable exchange. It seems quite clear that without this variable permutation the K—S operator has quite different properties.

Although it is quite clear that the values of z^{-1} for which phase transitions occur must be elements of the spectrum of the K—S operator, the converse is not true. Consider a value of z for which there exists at least one solution of the K—S equations which does not have the full translational symmetry; then z^{-1} must be an element of the spectrum of the K—S operator since if θ is the translational operator

$$(I - zk)\rho = z,$$

$$\theta(I - zk)\theta^{-1}\theta\rho = \theta z.$$

Since in infinite volumes

$$\theta(I - zk)\theta^{-1} = (I - zk)$$

we have

$$(I - zk)\theta\rho = z.$$

If

$$\theta\rho \neq \rho,$$

then z^{-1} is an eigenvalue of the K—S operator. This is clearly independent of any boundary conditions as long as the volume is infinite.

This result though trivial is important for two reasons. First, it indicates that spectral values may not involve phase transitions in all cases, and, secondly, it indicates that there may be a difference in the z plane singularities of symmetry breaking and nonsymmetry breaking phase transitions.

*Present address: Institut für Theoretische Physik, Universtät zu Köln, 5 Köln 1, Universitätsstr. 14, Germany.
†Supported by NSF Grant GP 31674X.
[1]J. E. Mayer, J. Chem. Phys. 10, 629 (1942).
[2](a) J. D. Weeks, S. A. Rice, and J. J. Kozak, J. Chem. Phys. 52, 2416 (1970). (b) Ing Yih S. Cheng and J. J. Kozak, J. Math. Phys. 14, 682 (1973).
[3]D. Ruelle, Statistical Mechanics Rigorous Results (Benjamin, New York, 1969), Chap. 4.
[4]N. Dunford and J. T. Schwartz, Linear Operators, General Theory (Interscience, New York, 1967), Vol. I, p. 566.
[5]This symmetry property has been exploited by (a) Ruelle (Ref. 3) and (b) the author, W. Klein, J. Math. Phys. 14, 1049 (1973).
[6]See Ref. 4, p. 579.
[7]See Ref. 4, p. 266.
[8]L. V. Kantorovich and G. P. Akilov, Functional Analysis in Normed Spaces (MacMillan, New York, 1964), p. 19. (Note: Kantorovich uses compact where we have been using conditionally compact.)
[9]Ref. 4, p. 485.
[10]Ref. 4, p. 479, Theorems 5 and 6.

1487   J. Math. Phys., Vol. 16, No. 7, July 1975

William Klein   1487

Self-gravitating fluids with cylindrical symmetry

P. S. Letelier*

Department of Physics, Boston University, Boston, Massachusetts 02215
(Received 12 September 1974)

The general solution of the Einstein field equation is obtained under the assumptions that (1) the source of the gravitational field is a perfect fluid with pressure p, equal to energy density ρ, (2) the space-time is cylindrically symmetric, and (3) the metric is given by three functions of two variables. The coordinate transformation to comoving coordinates is discussed. The energy and the Hawkins-Penrose inequalities are studied. The singularities of a class of solutions is studied using the concept of velocity-dominated singularity. A relation between Einstein-Rosen waves and a class of solutions is shown.

1. INTRODUCTION

In a recent paper[1] it was found that Einstein's field equations for a self-gravitating perfect fluid with pressure p equal to rest energy ρ and four-velocity u_a is equivalent to the field equations[2]

$$R_{ab} = -2\sigma_{,a}\sigma_{,b} \tag{1a}$$

$$\Box\sigma = (\sqrt{-g}\,\sigma_{,a}g^{ab})_{,b}/\sqrt{-g} = 0, \tag{1b}$$

when irrotationality is imposed, i.e.,

$$u_a = \sigma_{,a}/\sigma_{,b}\sigma^{,b} \tag{2}$$

The units are chosen so that we have for the velocity of light $c = 1$ and Newton's constant of gravitation $G = 1/8\pi$. A comma means partial derivative with respect to the index.

The pressure p and the energy—momentum tensor T_{ab} are related to σ by

$$p = \rho = \sigma_{,a}\sigma^{,a}, \tag{3}$$

$$T_{ab} = 2\sigma_{,a}\sigma_{,b} - g_{ab}\sigma_{,c}\sigma^{,c} \tag{4}$$

It is the purpose of this paper to discuss the solution of Eqs. (1) when the space—time has cylindrical symmetry. The general metric with cylindrical symmetry is restricted to one with three unknown functions of the form

$$ds^2 = e^{2(\omega-\lambda)}(dt^2 - dr^2) - e^{2\mu}(r^2 e^{-2\lambda}d\theta^\lambda + e^{2\lambda}dz^2), \tag{5}$$

where ω, λ, and μ are functions of r and t. These assumptions enable us to reduce the problem to one of solving two linear equations and computing a line integral.

In Sec. 2 we find the solution of Eqs. (1) when the metric is (5). In Sec. 3 the coordinate transformation that enables us to write the solution in comoving coordinates is discussed. In Sec. 4 the energy and the Hawking—Penrose inequalities are studied.[3] In Sec. 5, using the concept of velocity-dominated singularity,[4,5] the singularities of a class of solutions are studied. In Sec. 6 a relation between Einstein—Rosen waves and a class of solutions is shown.

2. THE SOLUTION

The field equations (1) and the pressure (3) when the metric is (5) are

$$\omega_{00} - \omega_{11} - \omega_1/r - \lambda_{00} + \lambda_{11} + \lambda_1/r + 2(\mu_{00} + \mu_0^2 + \lambda_0^2)$$
$$+ 2\mu_1(\lambda_1 - \omega_1) + 2\mu_0(\lambda_0 - \omega_0) = -\sigma_0^2, \tag{6}$$

$$\mu_{10} + \mu_1\mu_0 + \lambda_0\lambda_1 + (\mu_0 - \omega_0)/2r + \mu_1(\lambda_0 - \omega_0) + \mu_0(\lambda_1 - \omega_0)$$
$$= -\sigma_0\sigma_1, \tag{7}$$

$$\omega_{11} - \omega_{00} - \omega_1/r + \lambda_{00} - \lambda_{11} - \lambda_1/r + 2(\mu_{11} + \mu_1/r + \mu_1^2 + \lambda_1^2)$$
$$+ 2\mu_1(\lambda_1 - \omega_1) + 2\mu_0(\lambda_0 - \omega_0) = -\sigma_1^2, \tag{8}$$

$$\mu_{11} - \mu_{00} + (3\mu_1 - \lambda_1)/r - \lambda_{11} + \lambda_{00} + 2(\mu_1^2 - \mu_0^2 + \mu_0\lambda_0 - \mu_1\lambda_1)$$
$$= 0, \tag{9}$$

$$\mu_{11} - \mu_{00} + (\mu_1 + \lambda_1)/r + \lambda_{11} - \lambda_{00} + 2(\mu_1^2 - \mu_0^2 - \mu_0\lambda_0 + \mu_1\lambda_1)$$
$$= 0, \tag{10}$$

$$\sigma_{00} - \sigma_{11} + 2\mu_0\sigma_0 - 2\mu_1\sigma_1 - \sigma_1/r = 0, \tag{11a}$$

$$p = \rho = \exp[-2(\omega - \lambda)](\sigma_0^2 - \sigma_1^2). \tag{12}$$

Where the indices 0 and 1 mean derivatives with respect to t and r, the comma is omitted for brevity.

Equations (9) and (10) are equivalent to

$$\lambda_{00} - \lambda_{11} + 2\mu_0\lambda_0 - 2\mu_1\lambda_1 - \lambda_1/r = 0, \tag{11b}$$

$$\mu_{00} - \mu_{11} - 2\mu_1/r + (\mu_0^2 - \mu_1^2) = 0. \tag{13}$$

The last equation can be easily integrated.[6] The integral is

$$e^{2\mu} = [F(t - r) + G(t + r)]/r, \tag{14}$$

where F and G are arbitrary functions of their arguments.

If μ is known in principle it is possible to find σ and λ, solving the linear Eq. (11). Then, to solve the field equation, it only remains to find ω. The system of Eqs. (6), (7), and (8) is equivalent to

$$2\mu_0\omega_0 + (1/r + 2\mu_1)\omega_1 = f + \sigma_0^2 + \sigma_1^2, \tag{15}$$

$$(1/r + 2\mu_1)\omega_0 + 2\mu_0\omega_1 = \varphi + 2\sigma_0\sigma_1, \tag{16}$$

$$\omega_{00} - \omega_{11} = h - \sigma_0^2 + \sigma_1^2, \tag{17}$$

where

$$f = \mu_{00} + \mu_{11} + \mu_1/r + \mu_0^2 + \mu_1^2 + \lambda_0^2 + \lambda_1^2 + 2\mu_0\lambda_0 + 2\mu_1\lambda_1, \tag{18}$$

$$\varphi = \mu_0/r + 2\mu_{10} + 2\mu_1\mu_0 + 2\lambda_0\lambda_0 + 2\lambda_0\mu_1 + 2\lambda_1\mu_0, \tag{19}$$

$$h = \lambda_{00} - \lambda_{11} - \lambda_1/r + \lambda_1^2 - \lambda_0^2 - \mu_{00} + \mu_{11} - \mu_0^2 + \mu_1^2. \tag{20}$$

Equation (17) follows from the other field equations. Equations (15) and (16) give ω as an integral,

$$\omega = \int \frac{1}{\Delta} \left\{ [2\mu_0(f + \sigma_0^2 + \sigma_1^2) - (1/r + 2\mu_1)(\varphi + 2\sigma_0\sigma_1)]\,dt \right.$$
$$\left. + [2\mu_0(\varphi + 2\sigma_0\sigma_1) - (1/r + 2\mu_1)(f + \sigma_0^2 + \sigma_1^2)]\,dr \right\}, \tag{21}$$

where

$$\Delta = (1/r + 2\mu_1)^2 - 4\mu_0^2. \tag{22}$$

The integrability conditions for ω are the Eqs. (11) and (13). Also the existence of ω is limited by the condition $\Delta \neq 0$. We notice that we can always add a constant to ω. This fact tells us that if g_{ab} is a solution determined by μ, λ, and σ, αg_{ab} is also a solution whenever α is a constant.

3. COMOVING COORDINATES

It can be easily verified that the coordinate R defined by

$$dR = e^{2\mu}r(\sigma_0\, dr + \sigma_1\, dt) \tag{23}$$

and $T = \sigma$ transform the 4-velocity u_a to $U_a = (U_0, 0, 0, 0)$; therefore, R is comoving. Equation (11a) guarantees that the differential that defines R is exact.

The Jacobian of the transformation to comoving coordinates is

$$J\,\frac{T, R, \theta, z}{t, r, \theta z} = e^{2\mu}r(\sigma_0^2 - \sigma_1^2),$$

which vanishes where $p = \rho = 0$ in the nonsingular region of space—time.

The line element in comoving coordinates is

$$ds^2 = [e^{2(\omega-\lambda)}/(\sigma_0^2 - \sigma_1^2)][dT^2 - (e^{-4\mu}/r^2)\,dR^2]$$
$$- e^{2\mu}(e^{-2\lambda}r^2\,d\theta^2 + e^{2\lambda}\,dz^2); \tag{24}$$

this line element has a singularity at $r = 0$.

4. THE REALITY CONDITIONS

In irrotational fluids with limit form of equation of state $p = \rho$, the energy condition $T_{ab}u^a u^b \geq 0$ and the Hawking—Penrose condition $(T_{ab} - \frac{1}{2}Tg_{ab})u^a u^b \geq 0$ Ref. 3 tell us the same, that

$$\rho = \exp[-2(\omega - \lambda)](\sigma_0^2 - \sigma_1^2) \geq 0.$$

It is clear that it is possible that ρ be negative in some regions of the space—time. The metric does not have necessarily a pathological behavior when this happens. The way of solving this problem is to fill the region where the energy is negative with a different kind of fluid.

We notice that $R_{,a}$ is orthogonal to $\sigma_{,a}$, $\theta_{,a}$, and $z_{,a}$. In the region where $\rho < 0$, $\sigma_{,a}$ is a spacelike vector and $R_{,a}$ a timelike one. Now let $\hat{R}_{,a}, \hat{\sigma}_{,a}$, $\hat{\theta}_{,a}$, and $\hat{z}_{,a}$ denote the corresponding unit vector fields. Using the fact that

$$g_{ab} = \hat{R}_{,a}\hat{R}_{,b} - \hat{\sigma}_{,a}\hat{\sigma}_{,b} - \hat{\theta}_{,a}\hat{\theta}_{,b} - \hat{z}_{,a}\hat{z}_{,b},$$

the stress—energy tensor can be written as

$$T_{ab} = -\sigma_{,c}\sigma^{,c}(\hat{R}_{,a}\hat{R}_{,b} + \hat{\sigma}_{,a}\hat{\sigma}_{,b} - \hat{\theta}_{,a}\hat{\theta}_{,b} - \hat{z}_{,a}\hat{z}_{,b}).$$

This stress-energy tensor is that of an anisotropic fluid with positive rest energy density $-\sigma_{,c}\sigma^{,c}$ and vanishing heat-flow vector. In this case both reality conditions are satisfied.

5. VELOCITY-DOMINATED SINGULARITIES

To study the solution singularities using the concept

of velocity-dominated singularity,[4,5] we must know the general solution of Eq. (11) when the coefficient μ is given by (14). We did not succeed in finding it, but we found the general solution when $e^{2\mu} = t$. This case is interesting because we still have two functions σ and λ as "parameters." The most general solution studied in this context is the Tabensky—Taub solution. This solution has only one "parameter" function.[5,7]

When $e^{2\mu} = t$, Eqs. (11) are

$$\lambda_{00} + \lambda_0/t = \lambda_{11} + \lambda_1/r, \tag{25a}$$

$$\sigma_{00} + \sigma_0/t = \sigma_{11} + \sigma_1/r. \tag{25b}$$

The solution of (25) has the integral representation

$$\int_0^\pi \int_0^\pi \Phi(t\cos u + r\cos v)\,du\,dv$$
$$+ \int_0^\infty \int_0^\infty \Psi(t\cosh u + r\cosh v)\,du\,dv, \tag{26}$$

as can be easily verified. Φ and Ψ are functions of the indicated variables. They are restricted only in that the integrals exist and the differentiation is allowed under the integral sign.

We shall study the behavior of the metric near the singularity $t = 0$. First we notice that[8]

$$\lim_{t\to 0} t\frac{\partial}{\partial t}\int_0^\infty\int_0^\infty \Psi(t\cosh u + r\cosh v)\,du\,dv$$
$$= -\int_0^\infty \Psi(r\cosh u)\,du. \tag{27}$$

Then near the singularity $t = 0$ we get

$$\lambda \doteq E(r)\ln t, \qquad \sigma \doteq \mathcal{E}(r)\ln t. \tag{28}$$

E and $\mathcal{E}$ are the right-hand side of (27) for each solution (26) of Eqs. (25a) and (25b), respectively.

The metric (24) when $e^{2\mu} = t$ is

$$ds^2 = [e^{2(\Omega-\lambda)}t^{-1/2}(t^2 - r^2)^{3/4}][dT^2 - dR^2/t^2r^2]$$
$$- t(e^{-2\lambda}r^2\,d\theta^2 + e^{2\lambda}\,dz^2), \tag{29a}$$

where

$$\Omega = \int \frac{tr}{t^2 - r^2}\{[t(\sigma_1^2 + \sigma_0^2 + \lambda_1^2 + \lambda_0^2) - 2(\lambda_0\lambda_1 + \sigma_0\sigma_1)$$
$$- r\lambda_1/t + \lambda_0]\,dr + [2t(\sigma_0\sigma_1 + \lambda_0\lambda_1) - r(\sigma_0^2 + \sigma_1^2 + \lambda_0^2 + \lambda_1^2)$$
$$- r\lambda_0/t + \lambda_1]\,dt\}. \tag{29b}$$

Near the singularity $t = 0$ from (28), we get

$$\Omega \doteq (E^2 + \mathcal{E}^2 + \mathcal{E})\ln t. \tag{30}$$

From (28) and (23) we realize that r is also a comoving coordinate, so the metric (29) in comoving coordinates, near the singularity $t = 0$, is

$$ds^2 \doteq \exp[2(E^2 + \mathcal{E}^2)\sigma/E]\,d\sigma^2/E^2 - \exp[2(E^2 - \mathcal{E}^2 - 1)\sigma/E]\,dr^2$$
$$- r^2\exp[(1 - 2\mathcal{E})\sigma/E]\,d\theta^2 - \exp[(1 + 2\mathcal{E})\sigma/E]\,dz^2. \tag{31}$$

The pressure near $t = 0$ is

$$p = \rho \doteq E^2\exp[-2(E^2 + \mathcal{E}^2)\sigma/E]. \tag{32}$$

The metric (31) tells us that the singularity $t = 0$ is velocity-dominated with symbol P,

$$P = \left(\frac{E^2 + \mathcal{E}^2 - 1}{E^2 + \mathcal{E}^2} \ ; \ \frac{1}{2} \frac{1 - 2\mathcal{E}}{E^2 + \mathcal{E}^2} \ ; \ \frac{1}{2} \frac{1 + 2\mathcal{E}}{E^2 + \mathcal{E}^2} \right). \tag{33}$$

We note that $\sum P_i = 1$, but $\sum P_i^2 \neq 1$. This fact defines a semi-Kasner-like velocity dominated singularity.[5]

The singularity $r = 0$ is also a spacelike one. Near $r = 0$ the role of r and t are interchanged, r is a time coordinate and t a space one. Also near $r = 0$ we have

$$\lambda = E(t)\ln r, \quad \sigma = \mathcal{E}(t)\ln r. \tag{34}$$

Via (23) we have $dR = tE(t)\,dt$ so t is a comoving spatial coordinate. In this case the P symbol is

$$P = \left(\frac{E^2 + \mathcal{E}^2 - \mathcal{E}}{E^2 + \mathcal{E}^2 - \mathcal{E} + 1} \ ; \ \frac{1 - \mathcal{E}}{E^2 + \mathcal{E}^2 - \mathcal{E} + 1} \ ; \ \frac{\mathcal{E}}{E^2 + \mathcal{E}^2 - \mathcal{E} + 1} \right). \tag{35}$$

This singularity is also a semi-Kasner-like one. It is interesting to remark that near the singularities the functions Φ are completely wiped out.

6. RELATION WITH CYLINDRICAL WAVES

When $\mu = 0$, the independent field Eqs. of (6) through (11) can be cast in the form

$$\omega_0 = 2r(\sigma_0\sigma_1 + \lambda_0\lambda_1), \tag{36a}$$

$$\omega_1 = r(\sigma_0^2 + \sigma_1^2 + \lambda_0^2 + \lambda_1^2), \tag{36b}$$

$$\lambda_{00} - \lambda_{11} - \lambda_1/r = 0, \tag{37a}$$

$$\sigma_{00} - \sigma_{11} - \sigma_1/r = 0. \tag{37b}$$

The field equations $R_{ab} = 0$ for cylindrical waves with metric

$$ds^2 = e^{2(\nu - \mu)}(dt^2 - dr^2) - r^2 e^{-2\mu}\,d\theta^2 - e^{2\mu}\,dz^2 \tag{38}$$

are[9,10]

$$\nu_0 = 2r\mu_0\mu_1, \quad \nu_1 = r(\mu_0^2 + \mu_1^2), \tag{39}$$

$$\mu_{00} - \mu_{11} - \mu_1/r = 0. \tag{40}$$

Putting $\lambda = a\sigma = \mu$, where a is a constant, Eqs. (37) tell us the same as (40). Putting $\omega = (a^2 + 1)\nu$ in Eqs. (36) they are reduced to (39). Then each solution of (36) and (37) generates a class of solutions of (39) and (40) whenever $\lambda = a\sigma$. Also, each solution of (39) and (40) generates a solution of (36) and (37) doing the same identification of functions. The pressure in this case is

$$p = \rho = \exp\{-2[(a^2 + 1)\nu - \mu]\}(\mu_0^2 - \mu_1^2)/a^2.$$

It interesting to remark that cylindrical gravitational waves are related to a special class of spherical and toroidal waves.[11,12] Also, this particular class of solutions can be easily related to spherical and toroidal waves using the previous identification of functions.

*On leave of absence from Instituto de Física, Universidad Católica de Chile, Santiago.

[1] R. Tabensky and A.H. Taub, Commun. Math. Phys. **29**, 61 (1973).

[2] In the present paper σ is the φ defined in the above paper.

[3] S.W. Hawking and R. Penrose, Proc. Roy. Soc. A **314**, 529 (1970).

[4] D. Eardley, E. Liang, and R. Sachs, J. Math. Phys. **23**, 99 (1972).

[5] E.P.T. Liang, J. Math. Phys. **13**, 386 (1972).

[6] The solution of (6) through (11) when $\lambda = 0$ was found by R. Tabensky (private communication).

[7] R. Tabensky, Ph.D. thesis, University of California, Berkeley, 1971.

[8] H. Lamb, *Hydrodynamics* (Dover, New York, 1932), p. 298.

[9] N. Rosen, Bull. Res. Council Israel **3**, 328 (1954).

[10] J.L. Synge, *Relativity: The General Theory* (North-Holland, Amsterdam, 1960), p. 352.

[11] L. Marder, Proc. Roy. Soc. A **313**, 83 (1969).

[12] L. Marder, Proc. Roy. Soc. A **327**, 123 (1972).

Solution of the three-body problem with inverse square potentials

Y. Avishai

Ben-Gurion University of the Negev, Beer-Sheva, Israel
(Received 13 September 1974)

The three-body problem with two-body inverse cube forces is solved by separation of the variables into an angular and a radial one. The angular equation is an integro–differential one, which can be solved by splines, while the radial equation is Bessel's.

In this note we show how to exactly solve the three-body Schrödinger equation with two-body inverse cube forces. It appears that in this special case there is a possibility for a separation of the variables. An inverse square potential is not realistic as far as elementary particles are concerned. However, there are other branches of physics in which such an interaction may play an important role. For example, the first correction to Newton's Law of Gravitation is a force of an inverse cube nature.

In the two-body problem with central force, an inverse square potential can be considered as an addition to the centrifugal force in each partial wave. This is not the case for a three-body system, since there is no partial wave analysis for each two-body channel. For simplicity we treat the problem of three identical Bosons in their s state but this should not be a limitation. The Faddeev equation in configuration space reads[1]

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + E - v(x)\right) U(x, y) = v(x)xy \int_{-1}^{1} \frac{U(x'y')}{x'y'} dt, \quad (1)$$

where

  $x =$ distance between any two particles,

  $y =$ distance between third particle and the center of mass of the other two,

  $v(x) =$ two body interaction,

  $E =$ total center of mass energy

and where

$$x'^2 = \tfrac{3}{4}x^2 + \tfrac{1}{4}y^2 - (\sqrt{3}/2)xyt,$$
$$y'^2 = \tfrac{1}{4}x^2 + \tfrac{3}{4}y^2 + (\sqrt{3}/2)xyt. \quad (2)$$

The boundary conditions at zero are

$$U(x, 0) = U(0, y) = 0. \quad (3)$$

While for large x and/or y the boundary conditions are dependent on the physical state described by the wavefunction. For a three-body bound state, U should fall exponentially with either x or y.

In order to solve Eq. (1) we first change into polar coordinates

$$x = R\cos\phi,$$
$$y = R\sin\phi,$$
$$0 \leqslant R < \infty,$$
$$0 \leqslant \phi \leqslant \pi/2. \quad (4)$$

Equation (1) then goes over into

$$\left(\frac{\partial^2}{\partial R^2} + \frac{1}{R}\frac{\partial}{\partial R} + \frac{1}{R^2}\frac{\partial^2}{\partial\phi^2} + E - v(R\cos\phi)\right) U(R, \phi)$$
$$= v(R\cos\phi) \int_{\Delta_-(\phi)}^{\Delta_+(\phi)} U(R, \theta) d\theta, \quad (5)$$

in which

$$\Delta_-(\phi) = |\pi/3 - \phi|,$$

$$\Delta_+(\phi) = \begin{matrix} \pi/3 + \phi, & 0 \leqslant \phi \leqslant \pi/6 \\ 2\pi/3 - \phi, & \pi/6 < \phi \leqslant \pi/2 \end{matrix} \quad (6)$$

are plotted versus ϕ in Fig. 1.

The boundary conditions Eq. (3) will now be

$$U(R, 0) = U(R, \pi/2) = 0. \quad (7)$$

Trying a solution of the form

$$U(R, \phi) = f(R)g(\phi) \quad (8)$$

and dividing by fg yields

$$\frac{1}{f(R)}\frac{\partial^2 f(R)}{\partial R^2} + \frac{1}{R}\frac{\partial f(R)}{\partial R} + \frac{1}{R^2 g(\phi)}\frac{\partial^2 g(\phi)}{\partial\phi^2} + E - v(R\cos\phi)$$
$$= \frac{v(R\cos\phi)}{g(\phi)} \int_{\Delta_-(\phi)}^{\Delta_+(\phi)} g(\theta) d\theta. \quad (9)$$

Up till now, the result is quite general. We now make the assumption that the interaction is of the inverse square form, namely,
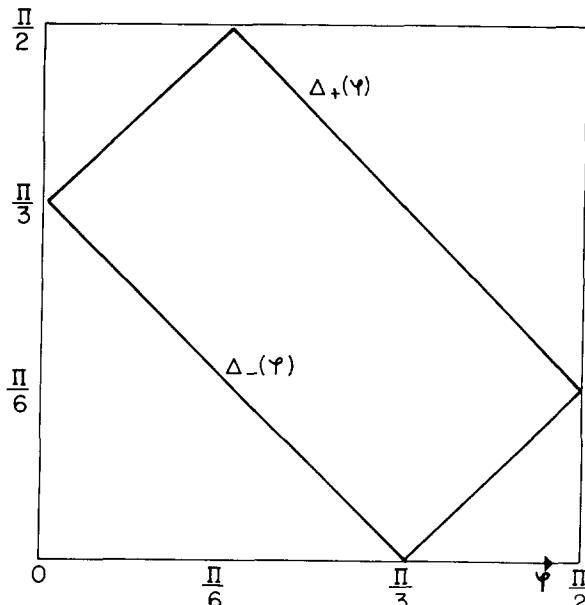
FIG. 1. The functions $\Delta_-(\phi)$ and $\Delta_+(\phi)$.

TABLE I. Eigenvalues of the angular Eq. (12) for different values of the potential strength.

λ \ c	-6.0	-5.0	-4.0	-3.0	-2.0	-1.0
λ_1	-56.01	-33.72	-18.66	-8.97	-1.87	3.19
λ_2	-11.04	-3.97	3.33	10.02	13.13	15.81
λ_3	13.91	18.65	24.71	29.37	32.13	35.72
λ_4	42.02	48.43	55.93	59.13	61.66	62.31

$$v(R\cos\phi) = \frac{c}{R^2\cos^2\phi} \tag{10}$$

with some constant c.

Substituting Eq. (10) into Eq. (9) and multiplying by R^2 leads into a complete separation

$$\frac{1}{f(R)}\left(R^2\frac{\partial^2 f(R)}{\partial R^2} + R\frac{\partial f(R)}{\partial R}\right) + R^2 E$$

$$= -\frac{1}{g(\phi)}\frac{\partial^2 g(\phi)}{\partial \phi^2} + \frac{c}{\cos^2\phi}\left(1 + \frac{1}{g(\phi)}\int_{\Delta_-(\phi)}^{\Delta_+(\phi)} g(\theta)\,d\theta\right) \tag{11}$$

$$= \lambda$$

say. First, one solves the angular eigenvalue problem

$$-\frac{\partial^2 g(\phi)}{\partial \phi^2} + c\frac{g(\phi)}{\cos^2\phi} + \frac{c}{\cos^2\phi}\int_{\Delta_-(\phi)}^{\Delta_+(\phi)} g(\theta)\,d\theta = \lambda g(\theta) \tag{12}$$

with

$$g(0) = g(\pi/2) = 0. \tag{13}$$

The eigenvalues λ_n are then put as constants into the radial equation

$$-\frac{\partial^2 f(R)}{\partial R^2} - \frac{1}{R}\frac{\partial f(R)}{\partial R} + \frac{\lambda_n}{R^2}f(R) = Ef(R), \tag{14}$$

or, with

$$f(R) = R^{-1/2}\psi(R), \tag{15}$$

we have the equation

$$-\psi'' + \frac{1}{R^2}(\lambda_n - \tfrac{1}{4})\psi = E\psi \tag{16}$$

which is related to Bessel's.

Because of Eq. (15) we require

$$\psi(0) = 0 \tag{17}$$

and for three-body bound state we must have rapid fall at large values of R.

While Eq. (16) is easily solved by standard techniques, this is not the case for Eq. (12), which seems unsolvable even numerically. However, the method of splines,[2] which for some reason is not much in use among theoretical physicists, allows an easy algorithm for solving such equations.

In Table I we have listed some of the eigenvalues of Eq. 12, for different potential strength c. In our choice of unites ($\hbar = m = 1$), the energy is given in fm^{-2}, and hence c as well as λn are dimensionless. We see that as $|c| \to 0$ $\lambda n \to 4n^2$ as it should be.

For $c > 0$ all the eigenvalues are greater than $\frac{1}{4}$ and the term $(\lambda n - \frac{1}{4})/R^2$ in Eq. (16) represents a repulsive force, and there are no bound states. On the other hand, when $c \lesssim -1\frac{1}{2}$, we have one eigenvalue less than $\frac{1}{4}$, and therefore one has an attractive term, with the possibility of having bound states.

As is well known,[3] an attractive inverse square potential is too singular at the origin, and one has to impose cutoff for small distances. If we assume an infinite three-body well at $R < a$, then the eigenvalues are found by the equation

$$H^{(1)}_{i\sqrt{\lambda n}}(i\sqrt{Bn}\,a) = 0. \tag{18}$$

Here $H^{(1)}$ is the cylindrical Hankel function of the first kind that solves Eq. 14 and falls off exponentially for large R.

[1]H. P. Noyes in *Three Body Problem in Nuclear and Particle Physics*, edited by J. S. C. McKee and P. M. Rolph (North-Holland, Amsterdam, 1970).
[2]J.H. Ahlberg, E. N. Nilson, and J. L. Walsh, *The Theory of Splines and Their Applications* (Academic, New York, 1967).
[3]L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Pergamon, New York), 2nd ed., p. 113.

1492   J. Math. Phys., Vol. 16, No. 7, July 1975

Y. Avishai   1492

Canonical transforms. III. Configuration and phase descriptions of quantum systems possessing an $sl(2,R)$ dynamical algebra

Charles P. Boyer and Kurt Bernardo Wolf

C.I.M.A.S., Universidad Nacional Autónoma de México, México D.F., Mexico
(Received 10 September 1974)

The purpose of this article is to present a detailed analysis on the quantum mechnical level of the canonical transformation between coordinate-momentum and number-phase descriptions for systems possessing an $sl(2,R)$ dynamical algebra, specifically, the radial harmonic oscillator and pseudo-Coulomb systems. The former one includes the attractive and repulsive oscillators and the free particle, each with an additional "centrifugal" force, while the latter includes the bound, free and threshold states with an added "centrifugal" force. This is implemented as a unitary mapping—canonical transform—between the usual Hilbert space L^2 of quantum mechanics and a new set of Hilbert spaces on the circle whose coordinate has the meaning of a phase variable. Moreover, the UIR's D_k^+ of the universal covering group of $SL(2,R)$ realized on the former space are mapped unitarily onto the latter.

1. INTRODUCTION

In this series of articles we have explored the question of canonical transformations in classical mechanics and their translation to quantum mechanics as unitary mappings between Hilbert spaces. These mappings have been given the general name of *canonical transforms*. In Ref. 1 we considered the set of (complex) linear transformations of phase space which preserved the Heisenberg algebra of coordinate and momentum variables (resp. operators) in classical (resp. quantum) mechanics, while in Ref. 2, upon examining the radial part of such an n-dimensional transformation, we found that the translation to quantum mechanics could be implemented asking for the preservation of a radial $sl(2,R) \approx su(1,1) \approx so(2,1)$ algebra built out of the n-dimensional underlying Heisenberg algebra. In this paper we will develop the unitary representation (canonical transform) of the transformation which can be formulated as follows.

Consider a classical system *possessing an $sl(2,R)$ dynamical algebra*. This means in our context that (i) there exist three quantities $\mathcal{G}_i(r,p_r)$, $i=1,2,3$ (where r and p_r are canonically conjugate variables: $\{r,p_r\}=1$) which under the Poisson bracket operation exhibit the $sl(2,R)$ Lie bracket relations

$$\{\mathcal{G}_1,\mathcal{G}_2\}=-\mathcal{G}_3, \quad \{\mathcal{G}_2,\mathcal{G}_3\}=\mathcal{G}_1, \quad \{\mathcal{G}_3,\mathcal{G}_1\}=\mathcal{G}_2, \qquad (1.1)$$

and such that (ii) the Hamiltonian H of the system belongs to the algebra, i.e., it can be written as a linear combination of the $\mathcal{G}_i(r,p_r)$. Now, through $SL(2,R)$ group transformations, we can always redefine the basis of the algebra so that H coincides with *one* of the three *orbit representatives* given by $\mathcal{G}_3, \mathcal{G}_1$, or $\mathcal{G}_1+\mathcal{G}_3$ corresponding, respectively, to elliptic, hyperbolic, or parabolic orbits. In each one of these cases we can define as action and phase variables.

$$p_\phi = \mathcal{G}_3, \quad \phi = \arctan(\mathcal{G}_2/\mathcal{G}_1), \qquad (1.2a)$$

$$p_\zeta = \mathcal{G}_1, \quad \zeta = \operatorname{arctanh}(\mathcal{G}_2/\mathcal{G}_3), \qquad (1.2b)$$

$$p_\xi = \mathcal{G}_1+\mathcal{G}_3, \quad \xi = \mathcal{G}_2/(\mathcal{G}_1+\mathcal{G}_3), \qquad (1.2c)$$

and in each of these cases one can verify that (1.1) implies that α and p_α ($\alpha=\phi,\zeta,\xi$) are canonically conjugate variables ($\{\alpha,p_\alpha\}=1$). The mapping (r,p_r) → (α,p_α) is a *canonical* transformation in the classical sense since the Heisenberg algebras are preserved, i.e., $\{r,p_r\}=1$ $\{\alpha,p_\alpha\}=1$, between the configuration and phase descriptions. The purpose of this article is to explore the quantum mechanical formulation of such canonical transformations. We shall see that the translation is possible when the Hamiltonian takes the standard from $\frac{1}{2}p_r^2+V(r)$ and the generators $\mathcal{G}_i(r,p_r)$ are *up-to-second order functions of p_r*. In this case (1.1) gives a set of coupled differential equations which severely restrict the types of potentials which can be considered, and in fact the possible realizations of the algebra (1.1) are essentially reduced to

$$\mathcal{G}_1 = \tfrac{1}{4}(p_r^2 - r^2 + gr^{-2}) = \tfrac{1}{4}(\mathbf{p}^2 - \mathbf{r}^2 + g|\mathbf{r}|^{-2}), \qquad (1.3a)$$

$$\mathcal{G}_2 = \tfrac{1}{2}rp_r = \tfrac{1}{2}\mathbf{r}\cdot\mathbf{p}, \qquad (1.3b)$$

$$\mathcal{G}_3 = \tfrac{1}{4}(p_r^2 + r^2 + gr^{-2}) = \tfrac{1}{4}(\mathbf{p}^2 + \mathbf{r}^2 + g|\mathbf{r}|^{-2}), \qquad (1.3c)$$

with arbitrary g, where $\mathbf{r}$ and $\mathbf{p}$ are n-dimensional vectors. The systems which can be described in this case are the attractive and repulsive harmonic oscillators and the free particle, all with an arbitrary additional "centrifugal" potential, corresponding to the elliptic, hyperbolic, and parabolic orbits mentioned above.

By *quantization* of (1.3) we mean the construction of self-adjoint operators on the usual Hilbert space of Lebesgue square-integrable functions $L^2(R^n)$. This procedure is unique[3,4] for (1.3) and yields an $sl(2,R)$ algebra of operators $I_i(r,\partial_r)$ under the commutator bracket, self-adjoint in the "radial" space $L^2(0,\infty)$. We will show in this article that we can perform a unitary mapping of $L^2(0,\infty)$ onto Hilbert spaces H_k^* (to be described below) where the operators $\hat{p}_\alpha$ defined in (1.2) are realized as $-i\partial/\partial\alpha$. The difficulties of giving a meaning in quantum mechanics to (1.2) can be seen clearly for the harmonic oscillator case (1.2a) to stem from the following problems: (i) The operator $-i\partial/\partial\phi$ is required to have a *discrete* spectrum which is incompatible with the existence of a phase operator "$\hat{\phi}$" such that $[\hat{\phi},\hat{p}_\phi]=\mathbb{1}$. (ii) When the operator $\hat{p}_\phi$ is realized as $-i\partial/\partial\phi$ on $L^2(-\pi,\pi)$, its spectrum turns out *not* to be positive-definite. The methods of treating these (and the related problem of angular momentum and angle observables) difficulties[5,6] have been through replacing the phase operator with some closely related

ones, e.g., Toeplitz operators[7] such as $\sin\phi$ and $\cos\phi$, and/or constructing a representation of the Heisenberg algebra which cannot be integrated to the group.[6]

In our construction, Hilbert spaces are constructed so that $\hat{p}_\alpha$ is a self-adjoint operator represented by $-i\partial/\partial\alpha$ *with the appropriate spectrum*. The phase variable α retains the meaning of an underlying space. Its operator realization (multiplication by α) is not Hermitean. The $sl(2,R)$ algebra and group representations are preserved and take the place of the Heisenberg algebra and Weyl group respectively in the definition and determination of the quantum canonical transformation corresponding to (1.2), as a unitary mapping between Hilbert spaces. The integral transform realization of such a mapping is the associated *canonical transform*. Furthermore, the unitary mapping is implemented for the pseudo-Coulomb system with the classical generators[8-10]

$$K_1 = \tfrac{1}{2}[r(\mathbf{p}^2 - 1) + g'r^{-1}], \qquad (1.4a)$$

$$K_2 = \mathbf{r}\cdot\mathbf{p}, \qquad (1.4b)$$

$$K_3 = \tfrac{1}{2}[r(\mathbf{p}^2 + 1) + g'r^{-1}], \qquad (1.4c)$$

by establishing the connection of this system with the harmonic oscillator. Although the complete dynamical groups for the two systems are different (the symplectic group $Sp(n,R)$ for the oscillator and $O(n,2)$ for the Coulomb system), the representations of the $\widetilde{SL}(2,R)$ subgroup are isomorphically related and appear to play a fundamental role in both systems.

The developments presented here have a group-theoretical significance of their own: On the algebra level, we connect the realization of the $sl(2,R)$ algebra generators on the line, as second-order differential operators, with their realization as first-order ones on the circle. On the group level, we relate the action of $\widetilde{SL}(2,R)$—the universal covering group of $SL(2,R)$— as conformal transformations of the circle with its nonlocal action on the line.

In Sec. 2 we construct the Hilbert spaces H_k^+ where $\hat{p}_\alpha$ has the required properties and its unitary mapping to $L^2(0,\infty)$. In Sec. 3 we relate the bound, free and threshold Coulomb systems with the three harmonic oscillator systems (1.2). In the Appendix we establish the connection between our spaces H_k^+ and the spaces of analytic functions on the disk[11-13] and half-plane[14] used for the description of the $sl(2,R)$ D_k^+ unitary irreducible representations (UIR's).

2. THE HARMONIC OSCILLATOR SYSTEMS AND THE CIRCLE

A. Elliptic case

We begin with the quantum Hamiltonian for the n-dimensional harmonic oscillator with an extra "centrifugal" potential of strength g

$$H = \tfrac{1}{2}(-\nabla^2 + r^2 + gr^{-2}), \qquad (2.1)$$

where ∇^2 is the n-dimensional Laplacian and $0 \le r^2 = |\mathbf{r}^2| < \infty$. Since we are interested in the radial part of H only, we separate (2.1) and its eigenfunctions into their radial and angular variables and write in place of

the angular part of (2.1) its well-known eigenvalues

$$\lambda = -L(L + n - 2), \qquad L = 0, 1, 2, \cdots, \qquad (2.2)$$

viz.

$$H = \tfrac{1}{2}\{-\partial_{rr} - [(n-1)/r]\,\partial_r + r^2 + (g - \lambda)/r^2\}. \qquad (2.3)$$

Now, the usual measure in n-dimensional radial configuration spaces is $r^{n-1}\,dr$; however, to facilitate our calculations, we can make the similarity transformation $H \to r^{(n-1)/2}Hr^{-(n-1)/2}$, which brings the measure to simply dr with the corresponding formal differential operator

$$I_3 \equiv \tfrac{1}{2}r^{(n-1)/2}Hr^{-(n-1)/2} = \tfrac{1}{4}\{-\partial_{rr} + r^2 + [(2k-1)^2 - \tfrac{1}{4}]/r^2\}, \qquad (2.4a)$$

where

$$2k = 1 \pm [(\tfrac{1}{2}n + L - 1)^2 + g]^{1/2}. \qquad (2.5)$$

Now, for $k \ge 1$, the spectral analysis of (2.4a) is well known[15] and there is a unique self-adjoint extension such that the normalized eigenvectors are

$$\psi_N^k(r) = [2N!/\Gamma(N + 2k)]^{1/2}e^{-r^2/2}r^{2k-1/2}L_N^{(2k-1)}(r^2), \qquad (2.6a)$$

where

$$I_3\psi_N^k(r) = (N + k)\psi_N^k(r), \qquad N = 0, 1, 2, \cdots, \qquad (2.6b)$$

and where $L_N^{(\alpha)}(z)$ are the associated Laguerre polynomials.[16] In the case that $(2k-1)^2 < 1$, both solutions to the eigenvalue problem for I_3 are square-integrable in the neighborhood of $r = 0$, and we must implement an additional boundary condition there. In this article we are interested in exploring the eigenvalue problems for I_3 whose spectra are bounded from below corresponding to the discrete series of representations D_k^+ of $SL(2,R)$. This corresponds, for the spectral analysis of I_3 with $\tfrac{1}{2} \le k < 1$, to implementing two different boundary conditions which yield $\{\psi_N^k\}$ and $\{\psi_N^{-k+1}\}$ separately as complete sets of orthonormal eigenvectors. The second set can be described equivalently by extending the range of k to $0 < k < 1$. Indeed the richer structure displayed in this interval has been noticed by Sally[13] and Montgomery and O'Raifertaigh.[17] Other boundary conditions corresponding to different self-adjoint extensions of I_3 give rise to the supplementary series of $\widetilde{SL}(2,R)$.

We complete the Lie algebra of $\widetilde{SL}(2,R)$ by adding the generators[10,18-20]

$$I_1 = \tfrac{1}{4}\{-\partial_{rr} - r^2 + [(2k-1)^2 - \tfrac{1}{4}]/r^2\}, \qquad (2.4b)$$

$$I_2 = -\tfrac{1}{2}i(r\partial_r + \tfrac{1}{2}). \qquad (2.4c)$$

It is straightforward to verify that (2.4) satisfy the well-known commutation relations

$$[I_1, I_2] = -iI_3, \quad [I_3, I_1] = iI_2, \quad [I_2, I_3] = iI_1 \qquad (2.7a)$$

and

$$I^2 = I_1^2 + I_2^2 - I_3^2 = k(1 - k). \qquad (2.7b)$$

The common invariant domain where the operators (2.3) as well as the Lie products (2.7) are densely defined is taken as $\{f \in L^2(0,\infty) : I_3^2 f \in L^2(0,\infty)\}$. Furthermore, as discussed previously, the generators (2.4) can be integrated[20] to a unique unitary representation of $\widetilde{SL}(2,R)$. For the general element of $SL(2,R)$

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, R), \quad ad - bc = 1. \tag{2.8a}$$

$\widetilde{SL}(2, R)$ is defined from the universal covering group of the compact subgroup $SO(2)$. Explicitly, for the matrix

$$\begin{pmatrix} \cos\tfrac{1}{2}\omega & \sin\tfrac{1}{2}\omega \\ -\sin\tfrac{1}{2}\omega & \cos\tfrac{1}{2}\omega \end{pmatrix} \longrightarrow \exp(-i\omega I_3), \tag{2.8b}$$

we now allow $-\infty < \omega < \infty$. The other one-parameter subgroups are, with their corresponding representations,

$$\begin{pmatrix} \cosh\tfrac{1}{2}\alpha & \sinh\tfrac{1}{2}\alpha \\ \sinh\tfrac{1}{2}\alpha & \cosh\tfrac{1}{2}\alpha \end{pmatrix} \longrightarrow \exp(-i\alpha I_1), \tag{2.8c}$$

where $0 \le \alpha < \infty$ and

$$\begin{pmatrix} e^{\beta/2} & 0 \\ 0 & e^{-\beta/2} \end{pmatrix} \longrightarrow \exp(-i\beta I_2). \tag{2.8d}$$

Associated with a general element of $SL(2, R)$ with $b \ne 0$ we have the group action[2,10,21]

$$(T_g f)(r)$$
$$= |b|^{-1} \exp(\mp i\pi k \, \mathrm{sgn}b) \int_0^\infty dr' (rr')^{1/2}$$
$$\times \exp\left(\frac{i}{2b}(ar'^2 + dr^2)\right) J_{2k-1}\left(\frac{rr'}{|b|}\right) f(r'), \tag{2.9}$$

where $f \in \underline{L}^2(0, \infty)$ and $g \in SL(2, R)$. The integral is understood to be in the sense of limit in the mean. Equation (2.9) can be extended to the entire range of the parameter ω in (2.8b) and thus to the whole universal covering group $\widetilde{SL}(2, R)$ through $\exp(-2i\pi I_3) = \exp(-2i\pi k)$. When $b = 0$, we have the local action

$$(T_g f)(r) = |a|^{-1/2} \exp[(ic/2|a|)r^2] f(|a|^{-1}). \tag{2.10}$$

We mention here that the ordinary $(g = 0)$ n-dimensional oscillator of angular momentum L belongs to the UIR of $\widetilde{SL}(2, R)$ with $k = \tfrac{1}{2}L + \tfrac{1}{4}n$, i.e., $D^+_{L/2+n/4}$. For $n = 3$, the oscillator states are spanned by the direct sum of UIR's $D^+_{3/4} \oplus D^+_{5/4} \oplus \cdots$. For the case $n = 1$ $(g = 0)$ the situation is somewhat different: The differential operator (2.4a) is no longer singular at the origin and the $O(n)$ rotational symmetry represented by the quantum number L is replaced by the two-element group C_2 of reflections, whose two representations are given by $L = 0$ and 1 in (2.2). The corresponding $\widetilde{SL}(2, R)$ UIR's are $D^+_{1/4}$ and $D^+_{3/4}$ corresponding to even and odd functions respectively.

We shall now construct a unitary isomorphism of the Lie algebra $sl(2, R)$ and covering group $\widetilde{SL}(2, R)$ representations on $\underline{L}^2(0, \infty)$ onto the corresponding algebra and group representations on the circle S^1 with a suitably defined inner product. Our realization for the Lie algebra $sl(2, R)$ on S^1 is the algebra of formal differential operators[11,12]

$$I_3 = -i\partial_\phi, \quad I_\pm = e^{\pm i\phi}(-i\partial_\phi \pm k), \tag{2.11}$$

where $I_\pm \equiv I_1 \pm iI_2$ and I_1, I_2, I_3 satisfy (2.7). For the discrete series of UIR's D^+_k $(k \ge \tfrac{1}{2})$ of $\widetilde{SL}(2, R)$, the infinitesimal generators satisfy the well-known relations

$$I_3 g^k_m = m g^k_m, \tag{2.12a}$$

$$I_+ g^k_m = \vartheta_m [m(m+1) + k(1-k)]^{1/2} g^k_{m+1}, \tag{2.12b}$$

$$I_- g^k_m = \frac{1}{\vartheta_{m-1}} [m(m-1) + k(1-k)]^{1/2} g^k_{m-1}, \tag{2.12c}$$

where $|\vartheta_m| = 1$, on a normalized set of basis vectors $\{g^k_m\}$ with the spectrum $m = k, k+1, \cdots$, thus $I_- g^k_k = 0$. Putting

$$g^k_m(\phi) = \gamma_m(k) e^{im\phi}, \tag{2.13a}$$

one can see after a straightforward calculation that

$$\gamma_m(k) = [\Gamma(m+k)/\Gamma(2k)\Gamma(m-k+1)]^{1/2} \tag{2.13b}$$

with $\vartheta_m = 1$.

We will now construct an inner product on S^1. This can be done by demanding that the $\{g^k_m\}$ form an orthonormal basis; however, we prefer to derive our inner product in the manner of Refs. 1 and 2, which elucidates the type of functions we are working with. We write down a general bilinear functional on a "nice" space of functions on S^1 and require the operators (2.11) to be Hermitean. It is easy to see that this inner product cannot be of the usual type for $\underline{L}^2(S^1)$, $(f,g) = \int_{-\pi}^\pi d\phi f(\phi)^* g(\phi)$ unless $k = \tfrac{1}{2} + i\rho$ with ρ real. This is the principal series of UIR's of $SL(2, R)$. Since we are treating the discrete series D^+_k this is not in general the case (except for $D^+_{1/2}$).

Now from the outset it is clear that we are dealing with multivalued representations, where the multivaluedness is determined by the real number $k \ge \tfrac{1}{2}$. We therefore consider the space $\mathcal{J}_k$ of infinitely differentiable functions on S^1 such that $f(\phi + 2\pi) = \exp(2\pi ik)f(\phi)$. Furthermore, consider the space $\Phi(\mathcal{J}_k)$ of continuous linear functionals[22] on $\mathcal{J}_k$

$$\Omega(f) \equiv (\Omega, f) = \int_{S^1} d\phi' \Omega(\phi, \phi') f(\phi'). \tag{2.14}$$

We can define the inner product

$$(f_1, f_2)_k = (f_1, \Omega(f_2))$$
$$= \int \int_{S^1} d\phi \, d\phi' \Omega(\phi, \phi') f_1(\phi)^* f_2(\phi'). \tag{2.15}$$

The $\Omega(\phi, \phi')$ can be determined from the hermiticity conditions for the generators (2.11). First, demanding the hermiticity of I_3, i.e., $(I_3 f_1, f_2)_k = (f_1, I_3 f_2)_k$, we find the conditions

$$\Omega(\phi, \phi') = \Omega(\phi - \phi'), \quad \Omega(\phi + 2\pi) = \exp(2\pi ik)\Omega(\phi). \tag{2.16}$$

Now any $f \in \mathcal{J}_k$ can be expanded uniformly in a Fourier series,

$$f(\phi) = \exp(ik\phi) \sum_{n=-\infty}^\infty a_n (2\pi)^{-1/2} \exp(in\phi), \tag{2.17}$$

and, by applying to it the lowering and raising operators $I_\pm$ it is clear that $\mathcal{J}_k$ is reducible since the subspace of functions $\mathcal{J}^+_k$ with Fourier coefficients $a_n = 0$ for n negative is invariant under the action of (2.11). The space $\mathcal{J}_k$ is not completely reducible, however, but the restriction of the Lie algebra representation (2.11) to $\mathcal{J}^+_k$ is irreducible.

For any $f_1, f_2 \in \mathcal{J}^+_k$ consider the hermiticity conditions

$$(f_1, I_\pm f_2)_k = (I_\mp f_1, f_2)_k. \tag{2.18}$$

A straightforward calculation involving integrations by parts yields the condition

$$\int \int_{S^1} d\phi\, d\phi' f_1(\phi)^* e^{i\phi'} f_2(\phi')[i(e^{i\theta} - 1)\Omega_\theta + (k-1)(e^{i\theta}+1)\Omega]$$
$$= 0. \qquad (2.19)$$

where $\theta \equiv \phi - \phi'$ and $\Omega_\theta = d\Omega/d\theta$. One is tempted to set the term in brackets in the integrand equal to zero and solve the resulting differential equation. Upon doing so, the solution is $\tilde{\Omega}(\theta) = c(1 - \cos\theta)^{k-1}$. We can verify that the Fourier expansion (2.17) of $\tilde{\Omega}$ contains only *negative-n* partial wave coefficients and thus is a member of $\Phi(\bar{\mathcal{J}}_k^+)$, where $\bar{\mathcal{J}}_k^+$ is the complement of $\mathcal{J}_k^+$ in $\mathcal{J}_k$. Hence $(f_1, \tilde{\Omega}(f_2)) = 0$ for *any* $f_1, f_2 \in \mathcal{J}_k^+$ and such a solution is worthless to us.

By inspecting (2.17) a bit closer it is seen that the vanishing of the terms in the square bracket of (2.19) is only a sufficient condition for the vanishing of the integral. Indeed, (2.17) is satisfied if the term in square brackets is orthogonal to $\exp(i\phi') f_2(\phi') \in \mathcal{J}_{k+1}^+$. So a *necessary and sufficient* condition for (2.17) to hold is

$$i[\exp(i\theta) - 1]\Omega_\theta + (k-1)[\exp(i\theta) + 1]\Omega = \tilde{\omega}(\theta) + c\,\exp(ik\theta),$$
$$(2.20)$$

where $\tilde{\omega}(\theta) \in \Phi(\bar{\mathcal{J}}_k^+)$ and c is a constant. Since any member of $\Phi(\bar{\mathcal{J}}_k^+)$ is useless to us as an inner product for $\mathcal{J}_k^+$ we discard $\tilde{\omega}(\theta)$ and look for a solution $\Omega \in \Phi(\mathcal{J}_k^+)$ of

$$i[\exp(i\theta) - 1]\Omega_\theta + (k-1)[\exp(i\theta) + 1]\Omega = c\,\exp(ik\theta). \qquad (2.21)$$

When we propose as a solution of (2.21) a series of the kind (2.17) with coefficients ω_n, we find this provides two independent solutions: One, for $n \geq 0$, yields the recursion $\omega_n = \omega_0 n!/(2k)_n$ in terms of the independent constant ω_0, while the second one, for $n < 0$, yields the recursion in terms of ω_{-1}. The latter series gives rise to $\tilde{\Omega}$ and we thus discard it. The former series is thus our solution $\Omega \in \Phi(\mathcal{J}_k^+)$ and, choosing $\omega_0 = 1/4\pi^2$,

$$\Omega(\theta) \equiv \sum_{m=k}^{\infty} \lambda_m(k) \exp(im\theta)$$
$$= \frac{1}{4\pi^2} \sum_{N=0}^{\infty} \frac{N!}{(2k)_N} \exp[i(k+N)\theta]$$
$$= \frac{1}{4\pi^2} \exp(ik\theta) F(1, 1; 2k; \exp(i\theta)). \qquad (2.22)$$

This series[16] converges absolutely for $k > 1$, conditionally for $\frac{1}{2} < k \leq 1$ (excluding $\theta = 0, 2\pi, \cdots$), and for $k = \frac{1}{2}$ it diverges on S^1. In the last case appropriate limiting arguments must be used in order to evaluate the double integral (2.15). For $0 < k < \frac{1}{2}$ the series (2.22) can still define a scalar product even though the series diverges.[13] Comparing the coefficients in (2.13) and (2.22), we find the important relation

$$\lambda_m(k) = [2\pi\gamma_m(k)]^{-2} \qquad (2.23)$$

which guarantees that $\{g_m^k\}$ is an orthonormal set under the scalar product (2.15). Equation (2.23) would have defined $\lambda_m(k)$ in the series (2.22) had we decided to find Ω from the requirement that $\{g_m^k\}$ form under (2.15) an orthonormal basis.

Now consider the inner product (2.15). We have for any $f_1, f_2 \in \mathcal{J}_k^+$, after some integrations,

$$(f_1, f_2)_k = \sum_{m=k}^{\infty} a_m^* b_m \lambda_m(k), \qquad (2.24)$$

where a_m and b_m are the ordinary Fourier coefficients for f_1 and f_2 respectively. We find from (2.22), for $k > \frac{1}{2}$ and $m = k + N$, N nonnegative integer, that $0 < 4\pi^2 \lambda_m(k) < 1$ and $\lambda_m(k) \to 0$ as $m \to 0$ while $4\pi^2 \lambda_m(\frac{1}{2}) = 1$. Thus the norm

$$0 < \|f\|_k^2 \equiv (f,f)_k = \sum_{m=k}^{\infty} |a_m|^2 \lambda_m(k) \leq \sum_{m=k}^{\infty} |a_m|^2$$
$$\leq \sum_{m=-\infty}^{\infty} |a_m|^2 < \infty \qquad (2.25a)$$

is dominated by the Hardy—Lebesgue norm[22] H^2 as well as $\mathcal{L}^2(-\pi, \pi)$. The members of H^2 are the boundary values almost everywhere on S^1 of functions analytic in the unit disc $|z| < 1$ completed with respect to the norm

$$\|f\|_k^2 = \sup_{0 < r < 1} \int_{-\pi}^{\pi} d\phi\, |f(r \exp(i\phi))|^2 = \sum_{m=k}^{\infty} |a_m|^2. \qquad (2.25b)$$

Thus, for $k = \frac{1}{2}$, closure gives the Hilbert space H^2. Notice also that when $0 < k < \frac{1}{2}$ the first inequality in (2.25a) is reversed; nevertheless the norm $\|f\|_k$ is defined by its series. Norms of this type were discussed by Sally[13] and are related to certain reproducing kernel spaces.

Using (2.23) and (2.24) we have

$$(f_1, f_2)_k = \sum_{m=k}^{\infty} (f_1, g_m^k)_k (g_m^k, f_2)_k \qquad (2.26)$$

for $f_1, f_2 \in \mathcal{J}_k^+$. Indeed, from (2.25) we can extend (2.26) to all functions $f_1, f_2 \in H^2$. Now H^2 is not closed with respect to the norm $\|f\|_k$, but by adjoining the limit points we obtain a Hilbert space which we denote by H_k^+. The connection between the Hilbert spaces H_k^+ and those of analytic functions on the disc will be elaborated upon in the Appendix.

Some further interesting properties of the linear functional $\Omega(f)$ defined by the kernel (2.22) can be seen by viewing Ω as a Hermitian operator on $\mathcal{L}^2(-\pi, \pi)$. It annihilates all $\tilde{f} \in \bar{\mathcal{J}}_k^+$ and hence all members of $\mathcal{L}^2(-\pi, \pi)$ which are limits of such $\tilde{f}$. For $k > \frac{1}{2}$ it is compact (completely continuous) and hence self-adjoint with eigenvalues $\lambda_m(k)$. Gel'fand and collaborators[14] have used such operators (for $k = 1, \frac{3}{2}, 2, \cdots$) to describe equivalences between representations labeled by k and $-k+1$.

Another linear functional in $\Phi(\mathcal{J}_k^+)$ which can be extended to all of H_k^+ is the reproducing functional given by the formal series

$$K(\phi, \phi') = \sum_{m=k}^{\infty} g_m^k(\phi) g_m^k(\phi')^*$$
$$= \exp[ik(\phi - \phi')](1 - \exp[i(\phi - \phi')])^{-2k}. \qquad (2.27)$$

Clearly this series diverges at $\phi = \phi'$, but nevertheless defines a continuous linear functional on $\mathcal{J}_k^+$, viz.,

$$f(\phi) = \int \int_{S^1} d\phi''\, d\phi'\, \Omega(\phi'' - \phi') K(\phi, \phi'') f(\phi'). \qquad (2.28)$$

We will now construct a unitary mapping which maps $\mathcal{L}^2(0, \infty)$ onto H_k^+ and the infinitesimal generators (2.4) onto (2.11) and conversely. The statement that the Hilbert space H_k^+ maps unitarily onto $\mathcal{L}^2(0, \infty)$ and conversely is almost trivial, since all separable Hilbert spaces

are unitarily equivalent. We see easily that $\underline{L}^2(0, \infty) \approx l^2$ $\approx \mathcal{H}_k^+$, where l^2 denotes the space of generalized Fourier coefficients $\{c_N\}$, $N = 0, 1, 2, \cdots$, such that $\sum_{N=0}^{\infty} |c_N|^2 < \infty$. We have for any $\psi \in \underline{L}^2(0, \infty)$

$$\psi(r) = \sum_{N=0}^{\infty} c_N \psi_N^k(r),\qquad (2.29)$$

where $\{\psi_N^k\}$ are given by (2.6) and convergence is in the mean. Thus $(\psi, \psi) = \sum_{N=0}^{\infty} |c_N|^2 < \infty$. But from (2.22) for any $\{c_N\} \in l^2$ we have an $f \in \mathcal{H}_k^+$ such that

$$f(\phi) = \sum_{N=0}^{\infty} c_N g_{k+N}^k(\phi)\qquad (2.30)$$

converges in the mean and hence $(f, f)_k = \sum_{N=0}^{\infty} |c_N|^2$ $= (\psi, \psi)$.

It is clear that the above statements are if and only if statements with the only proviso that both $\psi(r)$ and $f(\phi)$ are defined up to sets of measure zero. It is now a simple task to construct this mapping explicitly as

$$(A\psi)(\phi) = \text{l. i. m.} \int_0^{\infty} dr\, A(\phi, r) \psi(r)\qquad (2.31)$$

for $\psi \in \underline{L}^2(0, \infty)$, where

$$A(\phi, r) = \sum_{N=0}^{\infty} g_N^k(\phi) \psi_N^k(r)^*$$

$$= [2/\Gamma(2k)]^{1/2} r^{2k-1/2} \exp(ik\phi)[1 - \exp(i\phi)]^{-2k}$$

$$\times \exp[(r^2/2)(e^{i\phi} + 1)/(e^{i\phi} - 1)].\qquad (2.32)$$

This kernel is singular at $\phi = 0$, which in an intuitive sense is offset by the strong convergence in the $\mathcal{H}_k$ norm. The inverse mapping is given by

$$(A^{-1}f)(r) = \text{l. i. m.} \int \int_{S^1} d\phi\, d\phi'\, \Omega(\phi - \phi') A(\phi, r)^* f(\phi'),$$

$$(2.33)$$

for any $f \in \mathcal{H}_k^+$. We stress that the unitary transformation kernel $A(\phi, r)$ is a unitary representation in quantum mechanics of the classical canonical transformation (1.2a). This is what we call a unitary canonical transform.

Now the important consequence of the unitary mappings (2.31) and (2.33) is that the group representations, or equivalently the Lie algebra representations (2.4) and (2.11) are unitarily equivalent. A straightforward computation shows that the operators $I_\pm = I_1 \pm i I_2$ in the representation (2.4) satisfy the Lie algebra identities (2.12). Then using (2.31)—(2.33) and a simple integration by parts yields the desired results. The domain of the Lie algebra products is mapped onto each other and as a subspace of l^2 is given by all $\{c_N\} \in l^2$ such that $\sum_{m=k}^{\infty} m^4 |c_N|^2 < \infty$. Furthermore, the $\widetilde{SL}(2, R)$ group representation on $\mathcal{H}_k$ can be obtained from (2.9) and (2.31) by $U_g = A T_g A^{-1}$ yielding explicitly

$$(U_g f)(\phi) = |1 + \gamma^* \exp(-i\phi)|^{-2k} (1 - |\gamma|^2)^k$$

$$\times f(\exp(i\omega)[\gamma + \exp(i\phi)]/[1 + \gamma^* \exp(i\phi)]),$$

$$(2.34)$$

for $f \in \mathcal{H}_k^+$. Here we have used the $SU(1, 1)$ variables defined from (2.8a) as

$$\alpha = \tfrac{1}{2}[a + d + i(c - b)], \qquad \beta = \tfrac{1}{2}[a - d - i(b + c)],$$

$$(2.35)$$

$$\gamma = \beta/\alpha, \qquad \omega = 2\arg\alpha.$$

We mention that the representation (2.34) is equivalent to the representation $U^k(g, k)$ of Sally if we replace in (2.5.5) of Ref. 13 the complex variable z by its boundary $e^{i\phi}$ and perform the similarity transformation $\exp(ik\phi) U_g \exp(-ik\phi)$. For the connection between the representations described in this section and the usual treatment on the unit disc $\mathcal{M}$, the reader is referred to the Appendix.

We now pass to the description of a basis where a noncompact subgroup generator is diagonal.[23] As is well known, there are three orbits in the Lie algebra $sl(2, R)$ under the adjoint action of the group $SL(2, R)$. One of these orbits (the elliptic one) gives rise to the basis described previously (i. e. , I_3 is diagonal). We proceed to give a brief description of the remaining two cases.

B. Parabolic case

In this case an orbit representative of the generators (2.4) is given by the radial free Hamiltonian

$$I_1 + I_3 = \tfrac{1}{2}\{-\partial_{rr} + [(2k - 1)^2 - \tfrac{1}{4}]/r^2\}.\qquad (2.36)$$

The eigenvalue problem thus gives rise to the generalized orthonormal eigenfunctions

$$\psi_S^k(r) = (rs)^{1/2} J_{2k-1}(rs)\qquad (2.37)$$

with eigenvalues $\tfrac{1}{2}s^2$. We also mention that an orbit representative which is simpler but with no physical meaning is $I_3 - I_1 = \tfrac{1}{2}r^2$. The relation between the two is given by $\exp(i\pi I_3)(I_3 + I_1)\exp(-i\pi I_3) = I_3 - I_1$. We emphasize that harmonic analysis[2i,24] in terms of the latter is simpler than in terms of the former. Nevertheless, it is the former we are interested in, because of its physical meaning.

Our unitary mapping (2.31) can be extended in the usual way to operate on a suitable space of generalized functions[22] containing the eigenfunctions (2.37). This means that the generalized eigenfunctions have a meaning as the kernel of a particular transform (in this case the well-known Hankel transform) when applied to any $\psi \in \underline{L}^2(0, \infty)$. In this sense then the basis elements (2.37) are mapped unitarily onto generalized eigenfunctions $g_S^k(\phi)$ of the operator $I_1 + I_3$ realized on the circle. In terms of the realization (2.11) we find

$$I_1 + I_3 = -i[(1 + \cos\phi)\partial_\phi - k\sin\phi].\qquad (2.38)$$

This operator becomes more transparent under the stereographic projection of the circle onto the real line given by

$$\xi = \tan\tfrac{1}{2}\phi, \qquad -\pi \leq \phi < \pi, \quad -\infty < \xi < \infty.\qquad (2.39)$$

First we note that the space $\mathcal{J}_k^+$ on S^1 maps onto the space (called again $\mathcal{J}_k^+$) of infinitely differentiable functions which decrease at infinity as ξ^{-2k} (see the Appendix). The multivaluedness of functions on S^1 implies definite phase properties for the corresponding functions of ξ as $\xi \to \pm \infty$. This is specified by choosing the principal branch of $\ln z$ to correspond to the range $-\pi \leq \phi < \pi$, so that

$$\exp(ik\phi) = \exp\left(k \ln\frac{i - \xi}{i + \xi}\right) = \left(\frac{i - \xi}{i + \xi}\right)^k.\qquad (2.40)$$

Then (2.38) in the ξ-space realization becomes

C.P. Boyer and K.B. Wolf

$$I_1 + I_3 = -i\left(\partial_\xi - k\frac{2\xi}{\xi^2+1}\right). \qquad (2.41)$$

The generalized eigenfunctions of (2.41) then have the form of a multiplier times the Fourier transform kernel with the phase inherited from the unitary mapping (2.31). Actually it is a simple calculation to obtain the eigenfunctions directly by applying (2.31) to the orthonormal basis functions (2.37), viz.

$$g_s^k(\phi) \equiv (A\psi_s^k)(\phi) = \sum_{N=0}^{\infty} g_{N+k}^k(\phi)(\psi_N^k, \psi_s^k), \qquad (2.42a)$$

where (ψ_N^k, ψ_s^k) are the overlap functions between the canonical basis (2.6) and the parabolic basis (2.37). These overlap functions become trivial to calculate if we transform the ψ_s^k to a point on the orbit where $I_3 - I_1 = \frac{1}{2}r^2$ is diagonal with generalized eigenfunctions $\widetilde{\psi}_s^k(r)$ $= \exp(i\pi k)\delta(r - s)$. We find

$$(\psi_N^k, \psi_s^k) = (\exp(i\pi I_3)\psi_N^k, \exp(i\pi I_3)\psi_s^k)$$
$$= (\exp(i\pi N + k)\psi_N^k, \widetilde{\psi}_s^k) = \exp(-i\pi N)\psi_N^k(s)^*. \qquad (2.42b)$$

Hence, the properly normalized (including phase) generalized eigenfunctions on the circle are, using (2.32),

$$g_s^k(\phi) = \exp(i\pi k)A(\phi - \pi, s). \qquad (2.42c)$$

This calculation shows the close connection between the unitary mapping of $\mathcal{L}^2(0, \infty)$ onto $\mathcal{H}_k^+$ and the parabolic basis. In terms of the ξ-space realization we find the form

$$g_s^k(\phi(\xi)) = [\Gamma(2k)]^{-1/2}(\tfrac{1}{2}s)^{2k-1/2}(1 + \xi^2)^k \exp(\tfrac{1}{2}is^2\xi). \qquad (2.42d)$$

It is readily checked that these functions are eigenfunctions of (2.41) with eigenvalues $\frac{1}{2}s^2$. Actually, since $s^2 \geq 0$, this is the half-space Fourier transform which is in complete accord with the fact, as discussed in the Appendix, that the members $f(\xi) \in \mathcal{H}_k^+$ in the ξ-space realization are the boundary values of functions $f(w)$ analytic in the upper half-plane $\mathrm{Im}\,w > 0$ with $\mathrm{Re}\,w = \xi$.

C. Hyperbolic case

In this case an orbit representative is given by the generator I_1 which is one-half the Hamiltonian for the repulsive harmonic oscillator. The eigenvalue problem is

$$I_1\psi_\nu^k = \tfrac{1}{2}\nu\psi_\nu^k. \qquad (2.43a)$$

However, a much simpler orbit representative is given by the generator I_2 with the relation

$$\exp(\tfrac{1}{2}\pi i I_3)I_2 \exp(-\tfrac{1}{2}\pi i I_3) = I_1.$$

The eigenvalue problem for I_2 is

$$I_2\widetilde{\psi}_\nu^k(r) = \tfrac{1}{2}\nu\widetilde{\psi}_\nu^k(r), \qquad (2.43b)$$

with normalized generalized eigenfunctions given by the well-known Mellin transform kernel

$$\widetilde{\psi}_\nu^k(r) = (2\pi)^{-1/2}r^{i\nu-1/2}, \qquad (2.44)$$

with $-\infty < \nu < \infty$. Using (2.9) to transform these functions to the corresponding basis functions for I_1, we find

$$\psi_\nu^k(r) = (2\pi r)^{-1/2}\exp(i\pi k)\exp(\tfrac{1}{4}\pi\nu)2^{i\nu/2}$$
$$\times[\Gamma(k + \tfrac{1}{2}i\nu)/\Gamma(2k)]M_{i\nu/2,k-1/2}(-ir^2), \qquad (2.45)$$

where $M_{i\nu/2,k-1/2}(z)$ is a Whittaker function.[16] We wish to effect the mapping of the functions (2.45) to the generalized eigenfunctions on S^1. These will be eigenfunctions of the operator

$$I_1 = -i(\cos\phi\,\partial_\phi - k\sin\phi), \qquad (2.46)$$

which satisfy (2.43a). Again, using the stereographic projection given by

$$\zeta = \tan[\tfrac{1}{2}(\phi + \tfrac{1}{2}\pi)], \quad -\tfrac{3}{2}\pi \leq \phi < \tfrac{1}{2}\pi, \qquad (2.47)$$

we can write (2.46) as

$$I_1 = -i[\zeta\partial_\zeta - k(\zeta^2 - 1)/(\zeta^2 + 1)]. \qquad (2.48)$$

Now the unnormalized generalized eigenfunctions of (2.48) which satisfy (2.43a) are $(\zeta^2 + 1)^k\zeta^{-k}\zeta_\pm^{i\nu/2}$, where $\zeta_+ = \zeta$ for $\zeta > 0$ and 0 for $\zeta < 0$, while $\zeta_- = -\zeta$ for $\zeta < 0$ and 0 for $\zeta > 0$. The correct normalization and phase for these eigenfunctions can be determined from the mapping (2.31). Alternatively, following the same procedure as in the parabolic case, we can write the eigenfunctions on the circle in terms of the Mellin transform of A, viz.

$$g_\nu^k(\phi) = (2\pi)^{-1/2}\int_0^\infty dr\, A(\phi + \tfrac{1}{2}\pi, r)r^{i\nu-1/2}. \qquad (2.49)$$

Integrating this expression, we find explicitly

$$g_\nu^k(\phi) = \exp[\mp\tfrac{1}{4}i\pi\,(2k + i\nu)]\exp(i\pi k)2^{i\nu/2-k-1}[\pi\Gamma(2k)]^{-1/2}$$
$$\times \Gamma(k + \tfrac{1}{2}i\nu)\{\sin[\tfrac{1}{2}(\phi + \tfrac{1}{2}\pi)]\}^{-2k}|\tan[\tfrac{1}{2}(\phi + \tfrac{1}{2}\pi)]|^{k+i\nu/2},$$

$$(2.50)$$

where $\mp$ is taken for $-\tfrac{1}{2}\pi < \phi < \tfrac{1}{2}\pi$ and $-\tfrac{3}{2}\pi < \phi < -\tfrac{1}{2}\pi$ respectively. In terms of the variable ζ the eigenfunctions are

$$g_\nu^k(\phi(\zeta)) = \exp[\mp\tfrac{1}{4}i\pi\,(2k + i\nu)]\exp(i\pi k)2^{i\nu/2-k-1}[\pi\Gamma(2k)]^{-1/2}$$
$$\times \Gamma(k + \tfrac{1}{2}i\nu)(\zeta^2 + 1)^k|\zeta|^{i\nu/2-k}. \qquad (2.51)$$

We remark that in the process of evaluating the integral (2.49) we have evaluated the more difficult integral of $A(\phi, r)$ in (2.32) with the Whittaker basis functions (2.45). This demonstrates the power of the group theoretical approach in obtaining special functions relations and is in the spirit of Refs. 21 and 24, where more difficult integrals are obtained. One further point is that the multiplicity of the hyperbolic decomposition for the representations D_k^+ is *one* in contradistinction to multiplicity *two* for the principal series[11,23] of $\widetilde{SL}(2, R)$. This is apparent in the $\mathcal{L}^2(0, \infty)$ realization, but in S^1 it is deeply hidden in the nonlocal measure. For example, from (2.51) one is led to think that the multiplicity is two—one Mellin transform for each half-axis. However, as discussed in the Appendix, the Hilbert space $\mathcal{H}_k^+$ in the ζ-space realization consists of boundary values of functions $f(w)$ analytic in the upper half-plane $\mathrm{Im}\,w > 0$ with $\mathrm{Re}\,w = \zeta$; hence, one can relate the two apparently independent Mellin transforms by using Cauchy's integral formula.

3. THE PSEUDO-COULOMB SYSTEM

The Hamiltonian for the n'-dimensional Coulomb system with an extra centrifugal force of strength g' is given by

$$H = \tfrac{1}{2}(\mathbf{p}^2 + 2qr^{-1} + g'r^{-2}) \qquad (3.1)$$

where $r = |\mathbf{r}|$. It is to be noted that (3.1) is relevant in the relativistic Coulomb problem.[25] The standard trick[9] for introducing the $sl(2, R)$ Lie algebra is essentially to turn the standard eigenvalue problem for the energy (3.1) into an eigenvalue problem for the charge q by multiplying (3.1) by r, viz.,

$$(\tfrac{1}{2}r\mathbf{p}^2 - Er + \tfrac{1}{2}g'r^{-1} - q)\Phi(\mathbf{r}) = 0. \tag{3.2}$$

Then upon introducing the Lie algebra generators

$$K_1 = \tfrac{1}{2}[r(\mathbf{p}^2 - 1) + g'r^{-1}], \tag{3.3a}$$

$$K_2 = \mathbf{r} \cdot \mathbf{p} - i(n' - 2) = rp_r - i(n' - 2), \tag{3.3b}$$

$$K_3 = \tfrac{1}{2}[r(\mathbf{p}^2 + 1) + g'r^{-1}], \tag{3.3c}$$

Eq. (3.2) can be written as

$$[(\tfrac{1}{2} - E)K_3 + (\tfrac{1}{2} + E)K_1 - q]\Phi(\mathbf{r}) = 0. \tag{3.4}$$

A. Elliptic orbit (bound states)

There are three different solutions to (3.4) depending on which orbit the operator (3.4) lies. The case $E < 0$ gives rise to the bound state solutions of the H atom, while for $E > 0$ and $E = 0$ one finds the scattering and threshold solutions. For $E < 0$, the automorphism $\exp(i\theta I_2)$ called "tilting" by Barut and Kleinert,[9] where

$$\tanh\theta = \frac{E + \tfrac{1}{2}}{E - \tfrac{1}{2}} = \frac{|E| - \tfrac{1}{2}}{|E| + \tfrac{1}{2}}, \tag{3.5}$$

transforms (3.4) into

$$[(-2E)^{1/2}K_3 - q]\widetilde{\Phi}(\mathbf{r}) = 0, \tag{3.6}$$

where $\widetilde{\Phi}(\mathbf{r}) = \exp(i\theta I_2)\Phi(\mathbf{r})$.

Now we could insert (3.3c) into (3.6) and find the standard differential equation; however, we already know that for the UIR D_k^+ of $\widetilde{SL}(2, R)$ the spectrum of K_3 is simply $m = k + N$. Thus we have

$$q = (-2E)^{1/2}(k + N), \quad N = 0, 1, 2, \cdots, \tag{3.7a}$$

where

$$2k = 1 + [(n' + 2L' - 2)^2 + 4g']^{1/2}, \quad L' = 0, 1, 2, \cdots, \tag{3.7b}$$

[Note the *difference* between (3.7b) and (2.5)]. Turning Eq. (3.7a) around as an eigenvalue problem for E, we find the usual (at least for integer k, i.e., $g' = 0$) result

$$E = -\tfrac{1}{2}q^2/(k + N)^2. \tag{3.7c}$$

It is this interpretation of (3.6) as an eigenvalue problem for E which suggests the name *pseudo*-Coulomb[10] for Eq. (3.4) and the Lie algebra (3.3). Indeed, the transition between the two problems is canonical *only for fixed E*, as can be seen from the transformation of the coordinate r under the "tilting" operation

$$\exp(i\theta K_2)r\exp(-i\theta K_2) = (-2E)^{1/2}r \equiv \rho, \tag{3.8a}$$

and thus it is seen that $\widetilde{\Phi}(\mathbf{r}) = \Phi(\rho)$. Moreover, by using (3.3b) the canonical conjugate variable to r, p_r, transforms as

$$\exp(i\theta K_2)p_r \exp(-i\theta K_2) = (-2E)^{-1/2}p_r \equiv p_\rho. \tag{3.8b}$$

It is emphasized that what we have shown here is that the "tilting" operation of Barut and Kleinert is *equivalent* to the replacement of r and p_r in the generators

(3.3) by the pair ρ, p_ρ. Furthermore, if r, p_r is a canonical pair then ρ, p_ρ is a canonical pair *only when E is constant*. Again, turning the problem around, we can start with ρ, p_ρ as a canonical pair obtaining r, p_r as one only for constant E. This is the pseudo-Coulomb problem, and it is this problem which can be mapped canonically by a simple point transformation onto the multidimensional harmonic oscillator[10,26] and hence onto the circle S^1 through the analysis of the preceeding section. Nevertheless, this group-theoretical treatment[9] of the hydrogen atom has had remarkable success in calculating transition amplitudes, form factors, etc.

Rewriting the operators (3.3) in terms of the variables ρ, p_ρ defined in (3.8), we see that (3.4) becomes the differential equation for the radial part of $\Phi(\rho)$ which we denote by $\varphi(\rho)$,

$$\tfrac{1}{2}[-\rho\partial_{\rho\rho} - (n' - 1)\partial_\rho - (\lambda' - g')\rho^{-1} + \rho]\varphi(\rho) = m\varphi(\rho), \tag{3.9a}$$

where as before $m = k + N$ and is related to E through (3.7a), and, as in (2.2),

$$\lambda' = -L'(L' + n' - 2), \quad L' = 0, 1, 2, \cdots. \tag{3.9b}$$

Now again the spectral analysis of (3.9) with the proper boundary condition on $\varphi(\rho)$ yields the allowed values of m as

$$m = N + k = N + \tfrac{1}{2} + [(\tfrac{1}{2}n' + L' - 1)^2 + g']^{1/2}, \tag{3.10}$$

where we have introduced k in (3.7b). Equation (3.9a) can now be turned into the analog of Eq. (2.6) with an operator Hermitean with respect to the measure $d\rho$ ($\rho \in [0, \infty)$) through a similarity transformation mapping functions as $\varphi(\rho) \to \psi'(\rho) = \rho^{n/2-1}\varphi(\rho)$ and operators as $K_i \to K_i' = \rho^{n/2-1}K_i\rho^{-n/2-1}$, viz.,

$$K_3'\psi_N'(\rho) = m\psi_N'(\rho), \tag{3.11a}$$

$$K_3' = \tfrac{1}{2}[-\rho\partial_{\rho\rho} - \partial_\rho + \rho + (k - \tfrac{1}{2})^2\rho^{-1}], \tag{3.11b}$$

$$\psi_N'(\rho) = [2N!/\Gamma(2k + N)]^{1/2}\rho^{k-1/2}e^{-\rho}L_N^{(2k-1)}(2\rho), \tag{3.11c}$$

and similarly for the operators (3.3a, b):

$$K_1' = \tfrac{1}{2}[-\rho\partial_{\rho\rho} - \partial_\rho - \rho + (k - \tfrac{1}{2})^2\rho^{-1}], \tag{3.11d}$$

$$K_2' = -i(\rho\partial_\rho + \tfrac{1}{2}). \tag{3.11e}$$

It is to be noted that the ordinary ($g' = 0$) n'-dimensional pseudo-Coulomb problem with angular momentum L' has $k = L' + \tfrac{1}{2}(n' - 1)$ and thus belongs to the UIR $D_{L'+(n'-1)/2}^+$ of $\widetilde{SL}(2, R)$. For $n' = 3$, the bound states of the system belong to the direct sum $D_1^+ \oplus D_2^+ \oplus \cdots$.

We can now establish the link with the harmonic oscillator system. Indeed, if we take Eqs. (2.4a), (2.6a), and (2.6b) and effect the following:

(i) A change of variable $\rho = \tfrac{1}{2}r^2$ as suggested by the classical analogue (1.5); we obtain an operator (resp. eigenstates) Hermitean (resp. orthogonal) with respect to the measure $dr = (2\rho)^{-1/2}d\rho$ by simply following the chain rule for the derivatives.

(ii) A similarity transformation $\psi(\rho) \to \psi'(\rho) = (2\rho)^{-1/4}\psi(\rho)$ and $I_j \to K_j' = (2\rho)^{-1/4}I_j(2\rho)^{1/4}$ takes us to eigenstates (resp. operators) which are identical with (3.11c) [resp. (3.11b)] when

(iii) We identify

$$g = 4g', \quad L = 2L', \quad n = 2n' - 2. \tag{3.12}$$

Implementing this transformation, the spectrum-generating algebra of the pseudo-Coulomb system is obtained from the operators (2.4a, c) yielding precisely the operators (3.11d, e). We see that the ordinary ($g' = 0$) n'-dimensional pseudo-Coulomb system of angular momentum L' belongs to the UIR $D_{L'+(n'-1)/2}^+ = D_{L/2+n/4}^+$ of $\widetilde{SL}(2,R)$. Thus, for example, the states of the three-dimensional Hydrogen atom ($n' = 3$, $L' = 0, 1, 2, \cdots$) are mapped onto the even-angular momentum states of the *four*-dimensional harmonic oscillator[27] ($n = 4$, $L = 0, 2, 4, \cdots$) with the representation given by $D_1^+ \oplus D_2^+ \oplus \cdots$. We emphasize that the condition (3.12) and hence the mapping between the two systems is *not* a necessary one. Other possible mappings of the Hamiltonians were discussed in Ref. 26. Our choice (3.12) has the advantages of associating extra centrifugal potentials with each other as well as mapping states of zero angular momentum onto states of zero angular momentum. For $n' = 2$, $n = 2$, the mapping is the one described in Ref. 10.

A similar analysis can be effected for the two non-compact orbits.

B. Parabolic orbit (threshold states)

As the energy here is constant (zero), this is the only truly canonical mapping between the real Coulomb system and the system (2.4). In this case (3.4) becomes simply

$$[\tfrac{1}{2}(K_3 + K_1) - q]\Phi(\mathbf{r}) = 0, \tag{3.13}$$

where, from (3.3),

$$K_3 + K_1 = r\mathbf{p}^2 + g'r^{-1}. \tag{3.14}$$

Implementing the necessary similarity transformations which led to Eqs. (3.11) and replacing the variables r by ρ and p_r by p_ρ, the corresponding generator becomes

$$K_3' + K_1' = -\rho\partial_{\rho\rho} - \partial_\rho + (k - \tfrac{1}{2})^2\rho^{-1}. \tag{3.15}$$

Making again the simple change of variables as well as the similarity transformation (ii) and the identification (3.2), we find precisely the operator for the radial free particle (3.26) with the generalized eigenfunctions

$$\psi_s'^k(\rho) = s^{1/2}J_{2k-1}(s(2\rho)^{1/2}). \tag{3.16}$$

We mention here that in complete analogy with the parabolic orbit in Sec. 2 the harmonic analysis in terms of the operator $K_3' - K_1' = \rho$ is much simpler.

C. Hyperbolic orbit (scattering states)

The case $E > 0$ gives rise to the Coulomb scattering states.[9] Now Eq. (3.4) can be brought to the eigenvalue problem for K_1 by the "tilting" operator $\exp(i\theta K_2)$, where now

$$\tanh\theta = \frac{E + \tfrac{1}{2}}{E - \tfrac{1}{2}} = \frac{|E| + \tfrac{1}{2}}{|E| - \tfrac{1}{2}}, \tag{3.17}$$

and we arrive at

$$[(2E)^{1/2}K_1 - q]\widetilde{\Phi} = 0, \tag{3.18}$$

which again is equivalent to the replacement of r and p_r^2

by ρ and p_ρ^2 respectively. Under the transformation (3.16) we have

$$\rho = (2E)^{1/2}r, \quad p_\rho = (2E)^{-1/2}p_r, \tag{3.19}$$

in lieu of (3.8). Again it is emphasized that for calculation purposes it is much easier to deal with K_2 $= \exp(-\tfrac{1}{2}\pi i K_3)K_1\exp(\tfrac{1}{2}\pi i K_3)$ and the corresponding Mellin transform. Here we simply write down the eigenfunctions of the operator K_1' [(3.11d)] obtained from (2.45) by the point and similarity transformations (i) and (ii) described above with the identification (3.12), viz.,

$$\widetilde{\psi}_\nu'^k(\rho) = (2\pi)^{-1/2}\exp(i\pi k)\exp(\tfrac{1}{4}\pi\nu)2^{i\nu/2}\frac{\Gamma(k + \tfrac{1}{2}i\nu)}{\Gamma(2k)}(2\rho)^{-1/2}$$

$$\times M_{i\nu/2, k-1/2}(-2i\rho). \tag{3.20}$$

We have shown that the spectrum-generating algebra $so(2,1) \approx su(1,1) \approx sl(2,R)$ [as well as its universal covering group $\widetilde{SL}(2,R)$] for the pseudo-Coulomb problem maps unitarily onto the radial harmonic oscillator system and thus through the composition maps onto the circle S^1. It is emphasized that this rotor has a nonlocal scalar product in order to preserve the positive definiteness of the bound-state spectrum. A similar situation can be found in the original work of Barut and Kleinert,[9] and Fronsdal,[8] where a nonlocal scalar product appears on the Fock sphere to insure a unitary representation of the $SO(4,2)$ group, or equivalently the $SO(2,1)$ subgroup. Since we have singled out the latter by studying the radial problem, the symmetry group $SO(4)$ does not appear here. It should be mentioned that the stereographic projection of the circle S^1 can be related to the radial pseudo-Coulomb problem through a transform with a Fourier type kernel. The connection of this with the momentum space and the embedding in the Fock sphere will be studied elsewhere.

APPENDIX

We shall relate here the representation theory of $SL(2,R)$ on the circle S^1 as presented in Sec. 2 to the better known representation of the discrete series on the unit disc as described by Bargmann[11] for single-valued UIR's of $SL(2,R)$, by Sally[13] for the multivalued UIR's and Gel'fand[14] for single-valued representations on the complex upper half-plane.

Let $f \in J_k^+$; then we can expand f together with all its derivatives in a Fourier series with positive partial waves as

$$f^{(n)}(\phi) = \exp(ik\phi)\sum_{N=0}^\infty a_N^n\exp(iN\phi). \tag{A1a}$$

Moreover, for $z \equiv r\exp(i\phi)$ with $r \leq 1$,

$$f^{(n)}(\phi) \geq \exp(ik\phi)\sum_{N=0}^\infty a_N^n r^N\exp(iN\phi)$$

$$= \exp(ik\phi)\sum_{N=0}^\infty a_N^n z^N, \tag{A1b}$$

1500   J. Math. Phys., Vol. 16, No. 7, July 1975

C.P. Boyer and K.B. Wolf   1500

and hence the series

$$g(z) \equiv \sum_{N=0}^{\infty} a_n z^n \qquad (A2)$$

defines an analytic function whose radius of convergence is greater than 1. Thus for every $f \in \mathcal{J}_k^+$ we can associate a function g analytic in a region R_g containing the closed unit disc $\overline{\mathcal{M}} = \{z \in c : |z| \leqslant 1\}$ such that $\exp(-ik\phi)f(\phi)$ is the boundary value of $g(z)$ as $|z| \to 1$, and conversely, for every analytic function g in $R_g \supset \overline{\mathcal{M}}$ we can construct the uniformly converging series (A1). Following Bargmann,[11] we equip the space of analytic function on the open disc with the inner product

$$(g_1, g_2)_k = (2k-1)\pi^{-1} \int \int_{\mathcal{M}} r\, dr\, d\phi (1-r^2)^{2k-2} g_1(z)^* g_2(z) \qquad (A3)$$

and the norm $\|g\|_k = (g,g)_k^{1/2} < \infty$. The r integral is understood to be in the sense of limit in the mean. Now if g_1, g_2 are analytic on all of $\overline{\mathcal{M}}$, we can write a Cauchy integral representation

$$g(z) = \frac{1}{2\pi i} \oint_{|z'|=1} \frac{g(z')\, dz'}{z - z'} . \qquad (A4)$$

Substituting (A4) into (A3) and performing the r and ϕ integrals, we find

$$(g_1, g_2)_k = \int \int_{S^1} d\phi\, d\phi'\, \Omega(\phi - \phi') f_1(\phi)^* f_2(\phi') \qquad (A5)$$

with $f_i(\phi) = \exp(ik\phi) \lim_{|z| \to 1} g_i(z)$ and $\Omega(\phi - \phi')$ given precisely by (2.22). However, since the norms (A3) and (2.15) are equivalent on $\mathcal{J}_k^+$, mean convergence in one is the same as mean convergence in the other, and so the space of functions analytic in $\mathcal{M}$ with finite norm (A3) is a realization of the Hilbert space $\mathcal{H}_k^+$. The members of $\mathcal{H}_k^+$ on S^1 are the boundary values almost everywhere of analytic functions in $\mathcal{M}$ with finite norm (A3). Moreover, as demonstrated by Bargmann[11] and Sally,[13] mean convergence in $\mathcal{H}_k^+$ implies pointwise convergence of analytic functions in $\mathcal{M}$.

We can easily express the Lie algebra generators (2.8) and group representation (2.34) on $\mathcal{M}$ by replacing $\exp(i\phi)$ by z. Then the mapping (2.31) is a mapping from $L^2(0, \infty)$ to the $\mathcal{H}_k^+$ realization on the disc. This mapping was mentioned previously by Bargmann[28] and studied in detail by Sally.[13]

The well-known conformal mapping of the unit disc $\mathcal{M}$ onto the upper half-plane $C_+ = \{w \in C : \mathrm{Im}\, w > 0\}$ is the analog of the mapping (2.39). Explicitly, for $z \in C$ we write

$$z = \frac{i - w}{i + w}, \qquad w = i \frac{1 - z}{1 + z}. \qquad (A6)$$

Then it is easy to see that $|z| < 1$ implies $\mathrm{Im}\, w > 0$; moreover, the boundary $|z| = 1$ of $\mathcal{M}$ maps onto the real line $\mathrm{Im}\, w = 0$ including the point at infinity. Thus (A6) defines a homeomorphism of the closed unit disc $\overline{\mathcal{M}}$ onto the one-point compactification of the upper half-plane[14] $\overline{C}_+ = \{w \in C : \mathrm{Im}\, w \geqslant 0\} \cup \{\infty\}$. Under this mapping the scalar product (A3) becomes

$$(f_1, f_2)_k = (2k-1)\pi^{-1} \int_{C_+} dw (\mathrm{Im}\, w)^{2k-2} f_1(w)^* f_2(w), \qquad (A7a)$$

where

$$f_i(w) = 2^{2k-2}(i + w)^{-2k} g_i(z(w)). \qquad (A7b)$$

As a result f is analytic in $C_+(\overline{C}_+)$ when g is analytic in $\mathcal{M}(\overline{\mathcal{M}})$. Moreover, analyticity of $g \in \overline{\mathcal{M}}$ and therefore of $f \in \overline{C}_+$ implies the condition at infinity

$$f(w) \underset{|w| \to \infty}{\sim} |w|^{-2k}. \qquad (A8)$$

The realization of $\mathcal{J}_k^+$ on $\overline{C}_+$ is the space of all functions analytic in $\overline{C}_+$ satisfying the condition (A8). The realization of $\mathcal{H}_k^+$ on C_+ is the space of all functions analytic in C_+ with finite norm $\|f\|_k = (f,f)_k^{1/2}$ given from (A7a), and $\mathcal{H}_k^+$ is the completion of $\mathcal{J}_k^+$ with respect to this norm.

[1]K. B. Wolf, J. Math. Phys. 15, 1295 (1974).
[2]K. B. Wolf, J. Math. Phys. 15, 2102 (1974).
[3]R. Hermann, *Lie Groups for Physicists* (Benjamin, New York, 1966), Chap. 16.
[4]K. B. Wolf, "The Heisenberg-Weyl Ring in Quantum Mechanics" in *Group Theory and Applications. III*, edited by E. Loebl, (Academic, New York, to appear).
[5]D. Judge and J.T. Lewis, Phys. Lett. 5, 190 (1963); L. Susskind and J. Glogower, Physics 1, 49 (1964); P. Carruthers and M.M. Nieto, Rev. Mod. Phys. 40, 411 (1968); R. Jackiw, J. Math. Phys. 12, 1021 (1971); Ref. 4, Sec. VI.
[6]J.C. Garrison and J. Wong, J. Math. Phys. 11, 2242 (1970); here the number operator $-i\partial_\phi$ is not essentially self-adjoint on the commutator domain; moreover, the spectrum is empty there.
[7]M. Rosenblum, Proc. Am. Math. Soc. 13, 590 (1962).
[8]C. Fronsdal, Phys. Rev. 156, 1665 (1967); K.B. Wolf, Suppl. Nuovo Cimento 5, 1041 (1967).
[9]A.O. Barut and H. Kleinert, Phys. Rev. 156, 1541 (1967); 157, 1180 (1967); 160, 1149 (1967); H. Kleinert, "Group Dynamics of the Hydrogen Atom," Lectures in Theoretical Physics, Boulder, Col. (1967); A.O. Barut, SIAM J. Appl. Math. 25, 247 (1973).
[10]M. Moshinsky, T.H. Seligman, and K.B. Wolf, J. Math. Phys. 13, 901 (1972).
[11]V. Bargmann, Ann. Math. 48, 568 (1947).
[12]L. Pukánsky, Math. Ann. 156, 96 (1964); A.O. Barut and C. Fronsdal, Proc. Roy. Soc. London 287, 532 (1965).
[13]P.J. Sally Jr., Mem. Amer. Math. Soc. 69 (1967).
[14]I.M. Gel'fand, M.I. Graev, and N. Ya. Vilenkin, *Generalized Functions, Vol.* 5 (Academic, New York, 1966), Chap. VII.
[15]T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1966), Chap. VI, Sec. 4; E.C. Titchmarsh, *Eigenfunction Expansions, Vol.* 1 (Oxford U.P., Oxford, 1962).
[16]A. Erdelyi *et al.*, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vols. 1, 2.
[17]W. Montgomery and L. O'Raifertaigh, J. Math. Phys. 15. 380 (1974).
[18]H. Bacry and J.L. Richards, J. Math. Phys. 8, 2230 (1967); N. Mukunda and B. Radhakrishnan, J. Math. Phys. 14, 254 (1973).
[19]U. Niederer, Helv. Phys. Acta 46, 191 (1973); U. Niederer,

1501   J. Math. Phys., Vol. 16, No. 7, July 1975

C.P. Boyer and K.B. Wolf   1501

to appear in Helv. Phys. Acta.

[20]C. P. Boyer, Comunicaciones Técnicas CIMAS 5, No. 68 (1974), to appear in Helv. Phys. Acta; note that the group SL $(2, R) \otimes O(n)$ appears as the symmetry group of the time-dependent Schrödinger equation with Hamiltonian (2.3).

[21]E. G. Kalnins and W. Miller Jr., J. Math. Phys. 15, 1728 (1974).

[22]K. Yosida, *Functional Analysis* (Springer-Verlag, Berlin, 1965).

[23]N. Mukunda, J. Math. Phys. 8, 2210 (1967); A. O. Barut and E. C. Phillips, Commun. Math. Phys. 8, 52 (1968); G. Lindblad and B. Nagel, Ann. Inst. H. Poincaré A 13, 27

(1970); C. Itzykson, J. Math. Phys. 10, 1109 (1969).

[24]C. P. Boyer, E. G. Kalnins, and W. Miller Jr., J. Math. Phys. 16, 499 (1975).

[25]A. O. Barut and G. L. Bornzin, J. Math. Phys. 12, 841 (1971).

[26]D. Bergman and Y. Frishman, J. Math. Phys. 6, 1855 (1965).

[27]M. Boiteux, Physica 65, 381 (1973); for an alternative approach, see D. G. Ravenhall, R. T. Sharp, and W. J. Pardee, Phys. Rev. 164, 1950 (1967).

[28]V. Bargmann, Commun. Pure Appl. Math. 14, 187 (1961), p. 203.

1502   J. Math. Phys., Vol. 16, No. 7, July 1975

C.P. Boyer and K.B. Wolf   1502

Canonical realizations of the Poincare group. I. General theory

M. Pauri

Istituto di Fisica dell'Università, Parma, Italy

Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Milano, Italy

G. M. Prosperi

Istituto di Fisica dell'Università, Milano, Italy

Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Milano, Italy
(Received 24 June 1974)

The canonical realizations of the full Poincaré group in classical mechanics are studied by means of a general formalism introduced in preceding papers. The resulting classification displays significant analogies with the quantum one. The irreducible realizations, corresponding to positive, zero, and imaginary mass particles with or without spin are discussed. Also the irreducible realizations of the homogeneous Lorentz group are classified. Particular attention is given to the nonirreducible realization describing a system of two free particles and to a discussion of the physical meaning of the "center-of-mass" and "internal" variables. It is seen that this formalism provides a most natural framework for the introduction of a direct interaction between the particles according to the well-known prescription given by Bakamjian and Thomas. Finally, some simple models of relativistic "rigid" systems are discussed.

1. INTRODUCTION

A method developed in preceding papers[1-3] for the analysis, the classification and the construction of the realizations of local Lie groups by means of classical canonical transformations is applied in the present paper to the full Poincaré group.

In Sec. 2 we first discuss in general the problem of realizing the space and time reflections in a canonical or anticanonical way in analogy with the unitary and antiunitary way for the quantum representations; then we construct all the possible *schemes A* for the proper Poincaré group in the sense of Ref. 1.[4] Various aspects of the results are parallel to the quantum ones: In particular, the two *canonical invariants* can be chosen to be strictly similar to the corresponding quantum operators. The first invariant W is related to the little group of the energy—momentum 4-vector and it refers to the internal angular degrees of freedom of the dynamical system involved. The second invariant Π is a function of the space—time translations generators and it is related to the mass of the system or to its energy evaluated in the center-of-mass frame. Two main types of schemes A exist independently of the values of Π, and correspond to regular and *singular* realizations, respectively. The *singular* type is related in any case to trivial realizations of the little group. In the regular type, one of the little group generators can always be identified with the *helicity*. When Π is positive definite, an alternative form can be given in which the generators of the little group are independent of the values of the energy—momentum. For $\Pi \equiv 0$ two additional *singular* types of Schemes A exist; a first one in which the invariant W coincides with the helicity and a second one corresponding to the homogeneous Lorentz group.

Section 3 is devoted to the characterization of the irreducible realizations.

A first class corresponds to $\Pi = M^2 c^2 > 0$. In this case the invariant W is simply related to the magnitude S of the intrinsic angular momentum $\mathbf{S}$ of the system (free

particle); the realizations corresponding to $W = 0$ belong to the *singular* type and describe scalar particles; the realizations corresponding to $W > 0$ are regular and describe particles with spin. In the case $W = 0$, the components q_i of the canonical "center-of-mass" transform correctly as positions coordinates of the particle under any operation of the Poincaré group. In the case $W > 0$, q transforms as a position vector under space rotations and translations but not under special Lorentz transformations. In this case a new vector x can be constructed in terms of the canonical generators which has all the correct transformation properties and appears to be the only space vector entitled to be interpreted as the position vector of the free particle with spin; the components of this vector, however, cannot be assumed as canonical variables since their mutual Poisson brackets fail to vanish. Many of these last results are already known or strictly similar to corresponding quantum ones; yet they have been emphasized in our context in view of their relevance to more general questions discussed in the sequel of the investigation.

In a second class corresponding to $\Pi = -\Delta^2 c^2 < 0$ we find again a regular and a *singular* type of irreducible realizations, corresponding to four and three degrees of freedom, respectively. These canonical realizations, which are new to the authors knowledge, should describe free classical "tachyons." A canonical and covariant position vector can be defined within this class only in the *singular* realizations (free scalar tachyons).

Finally, in the third class $\Pi = 0$, we find four subclasses of irreducible realizations, one of the regular type and three of the *singular* type. The regular one is the classical analog of the so-called infinite-spin unitary representations of zero mass studied by Wigner. A first subclass of the *singular* type describes zero-mass particles with a given *helicity*; a second one, which may also be viewed as a particular case of the first, describes zero-mass scalar particles; finally, the third one corresponds to an identically vanishing energy—momentum 4-vector and provides the irreducible

canonical realizations of the homogeneous Lorentz group. Again a canonical and covariant position vector can be defined only in the scalar particle case.

Section 3 ends with a discussion of the realizations of the full Poincaré group within the phase spaces of the irreducible realizations already classified for the proper Poincaré group. The full group turns out to be realizable:

(a) in all of the irreducible realizations of the class $\Pi = M^2 c^2 > 0$, in the regular and in the singular scalar irreducible realizations of the class $\Pi = 0$—and in all these cases the space reflection and the time reflection can be realized in a canonical and anticanonical way, respectively;

(b) in the regular irreducible realizations of the class $\Pi = -\Delta^2 c^2 < 0$ with $W > 0$, in the scalar realizations of this same class and in the irreducible realizations of the homogeneous Lorentz group when the pseudoscalar invariant $\mathbf{J} \cdot \mathbf{K}$ vanishes; in all of these cases both space and time reflections can be realized canonically or anticanonically.

In all of the other cases there is always some reflection transformation which cannot be realized either in a canonical or an anticanonical way.

In Sec. 4 we consider the nonirreducible realization corresponding to a system of two free particles without spin, characterized by the canonical coordinates q_1, p_1, q_2, p_2. This realization is then reduced to the *typical form*,[1] after the introduction of "center-of-mass" and "internal" coordinates ($\mathbf{Q}, \mathbf{P}$ and ρ, π, respectively), linked to the original ones through a global canonical transformation. The physical meaning of these new variables is discussed with particular emphasis on the "internal" space vector ρ *which appears to be related in a significant way to the relative position of the particles in the center-of-mass system*. On the basis of these results a quite natural way for introducing an interaction between the particles consists in maintaining the formal structure of the canonical generators in terms of the basic variables $\mathbf{Q}, \mathbf{P}, \rho, \pi$ unaltered apart from the addition of an interaction "potential" $U(\rho, \pi)$ to the free particle expression of the center-of-mass energy $c\sqrt{\Pi}$. This procedure coincides with the prescriptions given by Bakamjian and Thomas[5] and Foldy.[6] A crucial consequence of the introduction of the interaction is that the original canonical variables $q_1(\mathbf{Q}, \mathbf{P}, \rho, \pi)$, $q_2(\mathbf{Q}, \mathbf{P}, \rho, \pi)$ no longer transform as covariant coordinates under special Lorentz transformations whatever the interaction potential is. This is a way of appearance of the well-known zero-interaction theorem.[7] Accordingly, the expressions q_1 and q_2 can be interpreted as position vectors of the particles only in the asymptotic region $|q_1 - q_2| \to \infty$ provided that the potential $U(\rho, \pi)$ vanishes for $|\rho| \to \infty$ in a suitable way. The whole physical interpretation of the theory rests on the construction of dynamical variables representing position vectors of the particles $\mathbf{x}_1(\mathbf{Q}, \mathbf{P}, \rho, \pi)$ and $\mathbf{x}_2(\mathbf{Q}, \mathbf{P}, \rho, \pi)$ which share all the desired transformation properties in the case $U(\rho, \pi) \neq 0$ thus allowing for a complete space—time description of two interacting particles within the Hamil-

tonian relativistic framework. This general problem is considered in detail in a separate paper.

In Sec. 5 a system of two points "rigidly" connected (relativistic linear rotator) is first discussed as a limiting case of a system of two particles interacting through a potential. In the case of equal rest-masses a moment of inertia can be defined and the typical form for this system closely resembles the corresponding form for the nonrelativistic linear rotator. The equations of motion of the system can be explicitly solved and it is interesting to note that the frequency of the internal rotation depends on the linear velocity of the center-of-mass according to the well-known relativistic formula for the retardation of moving clocks. Actually this system seems to be the most simple conceivable model of a relativistic clock. By means of similar formal modifications of the nonrelativistic formulas for a spherical, symmetrical and asymmetrical top, corresponding relativistic systems with six degrees of freedom are obtained. Presumably, the equations deduced by following this formal procedure describe actual relativistic "rigid" bodies whose internal mass distribution satisfies particular symmetry conditions.

2. GENERALITIES AND SCHEMES "A"

According to the usual parametrization (we adopt the *passive point of view* throughout), the operators of the infinitesimal transformations of the proper Poincaré group are

space rotations $\quad \mathcal{J} = -\mathbf{x} \wedge \dfrac{\partial}{\partial \mathbf{x}}$,

special Lorentz transformations $\quad K = -\left(\dfrac{1}{c^2} \mathbf{x} \dfrac{\partial}{\partial t} + t \dfrac{\partial}{\partial \mathbf{x}}\right)$,

space translations $\quad T = -\dfrac{\partial}{\partial \mathbf{x}}$,

time translations $\quad T_0 = -\dfrac{\partial}{\partial t}$.

Their commutation relations are

$$[\mathcal{J}_i, \mathcal{J}_j] = \epsilon_{ijk} \mathcal{J}_k, \quad [K_i, K_j] = -\frac{1}{c^2} \epsilon_{ijk} \mathcal{J}_k,$$

$$[\mathcal{J}_i, K_j] = \epsilon_{ijk} K_k, \quad [K_i, T_j] = \frac{1}{c^2} \delta_{ij} T_0,$$

$$[\mathcal{J}_i, T_j] = \epsilon_{ijk} T_k, \quad [K_i, T_0] = T_i, \tag{2.1}$$

$$[\mathcal{J}_i, T_0] = 0, \quad [T_i, T_j] = [T_i, T_0] = 0$$

$$(i, j, k = x, y, z).$$

Let us denote by $\mathbf{J}, \mathbf{K}, \mathbf{T}$, and $-H$, respectively, the generators of the corresponding transformations in a canonical realization. According to Ref. 1 these generators satisfy the Poisson bracket relations

$$\{J_i, J_j\} = \epsilon_{ijk} J_k + d_{J_i J_j}, \quad \{K_i, K_j\} = -\frac{1}{c^2} \epsilon_{ijk} J_k,$$

$$\{J_i, K_j\} = \epsilon_{ijk} K_k + d_{J_i K_j}, \quad \{K_i, T_j\} = -\frac{1}{c^2} \delta_{ij} H,$$

$$\{J_i, T_j\} = \epsilon_{ijk} T_k + d_{J_i T_j}, \quad \{K_i, H\} = -T_i - \frac{1}{2} \epsilon_{ihk} d_{J_h T_k} \tag{2.2}$$

$\{J_i, H\} = 0,$ $\qquad$ $\{T_i, T_j\} = \{T_i, H\} = 0$

$(i, j, k = x, y, z),$

where the Jacobi identity has already been enforced. By means of the substitutions[1-3,8]

$$J_i \rightarrow J_i + \tfrac{1}{2}\epsilon_{ihk}d_{J_hJ_k},$$
$$T_i \rightarrow T_i + \tfrac{1}{2}\epsilon_{ihk}d_{J_hT_k}, \qquad (2.3)$$
$$K_i \rightarrow K_i + \tfrac{1}{2}\epsilon_{ihk}d_{J_hK_k},$$

Eqs. (2.2) become

$$\{J_i, J_j\} = \epsilon_{ijk}J_k, \quad \{K_i, K_j\} = -\frac{1}{c^2}\epsilon_{ijk}J_k,$$

$$\{J_i, K_j\} = \epsilon_{ijk}K_k, \quad \{K_i, T_j\} = -\frac{1}{c^2}\delta_{ij}H,$$

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad (2.4)$$

$$\{J_i, T_j\} = \epsilon_{ijk}T_k, \quad \{K_i, H\} = -T_i,$$

$$\{J_i, H\} = 0, \qquad \{T_i, T_j\} = \{T_i, H\} = 0$$

$(i, j, k = x, y, z).$

Thus no constant $d_{\rho\sigma}$ remains, a result which is well-known for the case of unitary representations of the group.

The full Poincaré group contains the discrete transformations of space and time reflections besides the continuous transformations of the proper group. The corresponding operators are defined by

space reflection $\quad \mathcal{G}_s F(\mathbf{x}, t) = F(-\mathbf{x}, t),$

time reflection $\quad \mathcal{G}_t F(\mathbf{x}, t) = F(\mathbf{x}, -t),$

and the commutation relations (2.1) must be completed with

$$\mathcal{G}_s\mathcal{G}_t = \mathcal{G}_t\mathcal{G}_s \qquad \mathcal{G}_s T = -T\mathcal{G}_s,$$
$$\mathcal{G}_s K = -K\mathcal{G}_s, \qquad \mathcal{G}_s T_0 = T_0\mathcal{G}_s,$$
$$\mathcal{G}_t \mathcal{G} = \mathcal{G}\mathcal{G}_t, \qquad \mathcal{G}_t T = T\mathcal{G}_t, \qquad \mathcal{G}_s\mathcal{G}_t = \mathcal{G}_t\mathcal{G}_s, \quad (2.5)$$
$$\mathcal{G}_t K = -K\mathcal{G}_t, \qquad \mathcal{G}_t T_0 = -T_0\mathcal{G}_t.$$

Let us consider first the space reflection. We assume that this transformation is canonically realized,

$$q' = q'(q, p), \quad p' = p'(q, p),$$

and denote by I_s the operator which acts correspondingly on functions of the phase space

$$I_s f(q, p) = f(q', p'). \qquad (2.6)$$

Then, the first Eq. (2.5) implies

$$I_s\{\mathbf{J}, \cdots\} = \{\mathbf{J}, \cdots\}I_s. \qquad (2.7)$$

On the other hand, from the definition of a canonical transformation

$$\{f, g\}_{qp} = \{f, g\}_{q'p'}, \qquad (2.8)$$

we have

$$I_s\{f, g\} = \{I_s f, I_s g\} \qquad (2.9)$$

and Eq. (2.7) is equivalent to

$$\{I_s\mathbf{J}, I_s f\} = \{\mathbf{J}, I_s f\}. \qquad (2.10)$$

Being f an arbitrary function of the phase space, the last relation implies that $I_s\mathbf{J}$ and $\mathbf{J}$ differ by a constant

vector at most. Finally, since $I_s^2 = \mathbb{1}$ must hold, it follows

$$I_s\mathbf{J} = \mathbf{J}. \qquad (2.11)$$

In a similar way from the remaining Eqs. (2.5), we obtain

$$I_s\mathbf{K} = -\mathbf{K}, \quad I_s\mathbf{T} = -\mathbf{T}, \quad I_s H = H. \qquad (2.12)$$

It is apparent that the fundamental Poisson bracket relations (2.4) are not modified by the transformations (2.11) and (2.12).

Let us consider now the time inversion transformation. Denoting by I_t the operator which realizes it on the functions of phase space in a canonical way, the same procedure gives

$$I_t\mathbf{J} = \mathbf{J}, \quad I_t\mathbf{K} = -\mathbf{K}, \quad I_t\mathbf{T} = \mathbf{T}, \quad I_t H = -H. \qquad (2.13)$$

It will be noticed that this transformation cannot be defined within the canonical realizations of the Poincaré group which are most significant from a physical point of view: Actually, the generator H should be interpreted as the energy in these cases, and therefore it is forced to be a positive definite quantity—a fact which is not compatible with the last Eq. (2.13). The way out, however, is that we can also realize time reflection by means of an *anticanonical* transformation, i.e., a phase space mapping $q' = q'(q, p)$, $p' = p'(q, p)$ such that

$$\{f, g\}_{qp} = -\{f, g\}_{q'p'} \qquad (2.14)$$

for arbitrary functions f and g.[3] [The most general phase space mapping $q' = q'(q, p)$, $p' = p'(q, p)$ such that

$$q' = q'(q, p), \quad p' = p'(q, p), \qquad (2.14')$$

which leave the Hamiltonian character of the equations of motion

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \qquad (2.14'')$$

invariant independently of the particular form of H, coincide with the transformations for which

$$\{f, g\}_{qp} = \lambda\{f, g\}_{q'p'} \qquad (2.14''')$$

holds true, with λ a real constant. The canonical and the anticanonical transformations belong to this class for $\lambda = 1$ and $\lambda = -1$, respectively. On the other hand, any transformation satisfying Eq. (2.14''') can be reduced to a canonical or an anticanonical one by means of the trivial scale redefinition

$$q' \rightarrow q'' = \sqrt{|\lambda|}\,q', \quad p' \rightarrow p'' = \sqrt{|\lambda|}\,p'. \qquad (2.14'''')$$

Note the analogy with the unitary and antiunitary representations allowed for by the Wigner theorem in quantum mechanics.] Denoting by I_t^* the phase space operator which realizes the anticanonical time reflection, we have now

$$I_t^*\{f, g\} = -\{I_t^* f, I_t^* g\} \qquad (2.15)$$

instead of Eq. (2.9). Consequently, it follows

$$I_t^*\mathbf{J} = -\mathbf{J}, \quad I_t^*\mathbf{K} = \mathbf{K}, \quad I_t^*\mathbf{T} = -\mathbf{T}, \quad I_t^* H = H. \qquad (2.16)$$

It is evident that these relations differ from (2.13) by an overall change of sign and that the transformation I_t^* changes in sign the fundamental Poisson brackets

(2.4). In a similar way, we may consider the *anticanoni-cal* realization of space reflection. We summarize here for later use all the possible canonical or anticanonical realizations of the improper transformations:

Transformation	Canonical realization		Anticanonical realization	
space reflection	$J \to J$, $K \to -K$,	$T \to -T$ $H \to H$	$J \to -J$, $K \to K$,	$T \to T$ $H \to -H$
time reflection	$J \to J$, $K \to -K$,	$T \to T$ $H \to -H$	$J \to -J$, $K \to K$,	$T \to -T$ $H \to H$
space–time reflection	$J \to J$, $K \to K$,	$T \to -T$ $H \to -H$	$J \to -J$, $K \to -K$,	$T \to T$ $H \to H$

$$\text{(2.17)}$$

We turn now to the construction of the scheme A (Ref. 1) for the proper group. Recalling the Galilean case,[3] we put

$$\mathfrak{P}_1 = T_x, \quad \mathfrak{P}_2 = T_y, \quad \mathfrak{P}_3 = T_z. \tag{2.18}$$

Then in order to construct the canonically conjugate variables $\mathfrak{Q}_i$, we consider the expressions

$$R_i \equiv -\frac{c^2}{H} K_i, \quad (i = x, y, z). \tag{2.19}$$

One has

$$\{R_i, \mathfrak{P}_j\} = \delta_{ij}, \quad (i = x, y, z; j = 1, 2, 3); \tag{2.20}$$

however,

$$\{R_i, R_j\} = -\frac{c^2}{H} \epsilon_{ijk} [J_k - (R \bigwedge T)_k] \tag{2.21}$$

(ϵ_{ijk} Ricci pseudotensor normalized with $\epsilon_{123} = 1$).

Therefore, we write

$$\mathfrak{Q}_i \equiv R_i + \Lambda_i, \quad (i = 1, 2, 3) \tag{2.22}$$

and look for three independent functions Λ_i which have zero Poisson brackets with $\vec{\mathfrak{P}} \equiv (\mathfrak{P}_1, \mathfrak{P}_2, \mathfrak{P}_3)$ and satisfy the equations

$$\{\mathfrak{Q}_i, \mathfrak{Q}_j\} = 0, \quad (i, j = 1, 2, 3). \tag{2.23}$$

The system

$$\{\mathfrak{P}_i, \Phi\} = 0, \quad (i = 1, 2, 3) \tag{2.24}$$

has seven independent solutions. Besides T and H, a possible choice for the other ones is

$$\Omega \equiv J - R \bigwedge T. \tag{2.25}$$

Therefore, putting

$$\Lambda_i = \Lambda_i(T, H, \Omega) \tag{2.26}$$

in Eq. (2.22) and taking into account the Poisson bracket relations

$$\{R_i, T_j\} = \delta_{ij}, \quad \{R_i, H\} = \frac{c^2}{H} T_i,$$

$$\{R_i, R_j\} = -\frac{c^2}{H^2} \epsilon_{ijk} \Omega_k,$$

$$\{\Omega_i, R_j\} = -\frac{c^2}{H^2} (\Omega \cdot T \delta_{ij} - \Omega_i T_j), \tag{2.27}$$

$$\{\Omega_i, \Omega_j\} = \epsilon_{ijk} \left(\Omega_k - \frac{c^2}{H^2} \Omega \cdot T T_k \right),$$

$$\{\Omega_i, H\} = 0, \quad (i, j, k = x, y, z),$$

Eq. (2.23) becomes

$$\frac{\partial \Lambda_j}{\partial T_i} - \frac{\partial \Lambda_i}{\partial T_j} + \frac{c^2}{H} \left(\frac{\partial \Lambda_j}{\partial H} T_i - \frac{\partial \Lambda_i}{\partial H} T_j \right)$$

$$+ \frac{c^2}{H^2} (\Omega \cdot T) \left(\frac{\partial \Lambda_j}{\partial \Omega_i} - \frac{\partial \Lambda_i}{\partial \Omega_j} \right) - \frac{c^2}{H^2} \left(\Omega \cdot \frac{\partial \Lambda_j}{\partial \Omega} T_i \right.$$

$$\left. - \Omega \cdot \frac{\partial \Lambda_i}{\partial \Omega} T_j \right) + \frac{\partial \Lambda_i}{\partial \Omega_r} \frac{\partial \Lambda_j}{\partial \Omega_s} \epsilon_{rst} \left(\Omega_t - \frac{c^2}{H^2} (\Omega \cdot T) T_t \right)$$

$$= \frac{c^2}{H^2} \epsilon_{ijk} \Omega_k, \quad (i, j, k = x, y, z). \tag{2.28}$$

Finally, if we require that the expressions $\mathfrak{Q}_i$ ($i = 1, 2, 3$) transform as vector components under space rotations

$$\{J_i, \mathfrak{Q}_j\} = \epsilon_{ijk} \mathfrak{Q}_k, \quad (i = x, y, z; j, k = 1, 2, 3), \tag{2.29}$$

the functions Λ_i must be of the form

$$\Lambda_i = \alpha[T, H, \Omega, (\Omega \cdot T)] T_i + \beta[T, H, \Omega, (\Omega \cdot T)] \Omega_i$$

$$+ \gamma[T, H, \Omega, (\Omega \cdot T)] (T \bigwedge \Omega)_i,$$

$$T = |\mathbf{T}|, \quad \Omega = |\Omega|. \tag{2.30}$$

Then it can be easily checked that a solution of Eq. (2.28) is given by Eq. (2.30) with

$$\alpha \equiv \beta \equiv 0, \quad \gamma = -\frac{c^2}{\sqrt{H^2 - c^2 T^2} (H + \sqrt{H^2 - c^2 T^2})}. \tag{2.31}$$

Consequently, we obtain

$$\vec{\mathfrak{Q}} = -\frac{c^2}{H} K - \frac{c^2}{\sqrt{H^2 - c^2 T^2} (H + \sqrt{H^2 - c^2 T^2})} T \bigwedge \Omega$$

$$= -\frac{c^2}{\sqrt{H^2 - c^2 T^2}} K + \frac{c^2}{\sqrt{H^2 - c^2 T^2} (H + \sqrt{H^2 - c^2 T^2})} J \bigwedge T$$

$$+ \frac{c^4 (K \cdot T)}{H \sqrt{H^2 - c^2 T^2} (H + \sqrt{H^2 - c^2 T^2})} T. \tag{2.32}$$

Using Eqs. (2.11), (2.12), (2.16) one can also check that $\vec{\mathfrak{Q}}$ behaves as a position vector under the canonical space reflection and the anticanonical time reflection: precisely it remains unchanged under the latter and changes in sign under the former transformation. Conversely, assuming this behavior under the improper transformations, one is forced to choose $\alpha \equiv \beta \equiv 0$ in Eq. (2.30) because a pseudoscalar under anticanonical time reflection cannot be constructed out of $T, H, \Omega, \Omega \cdot T$; this makes the solution for $\vec{\mathfrak{Q}}$ unique. In order to proceed in constructing the scheme A we must look now for four functions having zero Poisson bracket with $\vec{\mathfrak{P}}$ and $\vec{\mathfrak{Q}}$. It is easy to see that, in analogy with the Galilean case, three independent functions satisfying this condition are

$$S_i \equiv J_i - (\vec{\mathfrak{Q}} \bigwedge \vec{\mathfrak{P}})_i = \frac{H}{\sqrt{H^2 - c^2 T^2}} J_i$$

$$+ \frac{c^2}{\sqrt{H^2 - c^2 T^2}} (K \bigwedge T)_i - \frac{c^2 (J \cdot T)}{\sqrt{H^2 - c^2 T^2} (H + \sqrt{H^2 - c^2 T^2})} T_i$$

$$(i = x, y, z) \tag{2.33}$$

Moreover, since Eqs. (2.27), (2.32) give

$$\{\mathfrak{Q}_i, H\} = \frac{c^2}{H} T_i, \tag{2.34}$$

a possible choice for the remaining one is

$$\Pi \equiv \frac{1}{c^2}(H^2 - c^2 T^2).\qquad(2.35)$$

Finally, one readily checks that

$$\{S_i, \Pi\} = 0, \quad \{S_i, S_j\} = \epsilon_{ijk} S_k \quad (i, j, k = x, y, z),\qquad(2.36)$$

which mean that the S_i's generate a canonical realization of the rotation group $SO(3)$. In conclusion, we can write (see Refs. 2, 3)

$$\mathfrak{P}_4 = S_z, \quad \mathfrak{I}_1 = S \quad \mathfrak{I}_2 = \sqrt{\Pi}$$

$$\mathfrak{Q}_4 = \arctan\frac{S_y}{S_x}$$

$$(S = |\mathbf{S}|),\qquad(2.37)$$

and the whole scheme A results

$\vec{\mathfrak{P}} = \mathbf{T}$	$\mathfrak{P}_4 = S_z$	$\mathfrak{I}_1 = S$	$\mathfrak{I}_2 = \sqrt{\Pi}$
$\vec{\mathfrak{Q}} = -\dfrac{c^2\mathbf{K}}{H} + \dfrac{c^2\mathbf{S}\wedge\mathbf{T}}{H(c\sqrt{\Pi}+H)}$	$\mathfrak{Q}_4 = \arctan\dfrac{S_y}{S_x}$		(A_1)

with $\mathbf{S}$ given by Eq. (2.33).

At this point we observe that the expressions $\vec{\mathfrak{Q}}$, S and $\sqrt{\Pi}$ in A_1 become complex if Π is a negative quantity; on the other hand, they are singular if $\Pi = 0$. This means that the scheme A_1 is significant only for canonical realizations in which Π is positive definite. If we want to obtain the scheme A in a form of general validity we are forced to follow a different approach and the new canonical variables $\mathfrak{Q}_1, \mathfrak{Q}_2, \mathfrak{Q}_3$ conjugate to $\mathfrak{P}_1, \mathfrak{P}_2, \mathfrak{P}_3$ cannot be chosen any more as components of a vector under space rotations and space and time reflections. We look now for this general form of the scheme A.

Leaving Eqs. (2.18) unchanged, we adopt a different choice for the independent solutions of the system (2.24). Precisely, we consider $\mathbf{T}$, H, and the expressions

$$W_T \equiv \frac{\mathbf{\Omega}\cdot\mathbf{T}}{T} = \frac{\mathbf{J}\cdot\mathbf{T}}{T},$$

$$W_u \equiv \frac{H}{c}\mathbf{\Omega}\cdot\mathbf{u},\qquad(2.38)$$

$$W_v \equiv \frac{H}{c}\mathbf{\Omega}\cdot\mathbf{v},$$

where $\mathbf{u}$ and $\mathbf{v}$ are unit vectors such that $(\mathbf{T}/T, \mathbf{u}, \mathbf{v})$ define an orthogonal right-handed triad; they can be written in turn by means of a fixed unit vector $\mathbf{n}$ as

$$\mathbf{u} = \frac{\mathbf{T}\wedge\mathbf{n}}{|\mathbf{T}\wedge\mathbf{n}|}, \quad \mathbf{v} = \frac{\mathbf{T}}{T}\wedge\mathbf{u}.\qquad(2.39)$$

The basic advantage of the new choice lies in that W_T, W_u, W_v satisfy simple Poisson bracket relations for any value of Π; actually,

$$\{W_T, W_u\} = W_v,$$

$$\{W_T, W_v\} = -W_u,\qquad(2.40)$$

$$\{W_u, W_v\} = \Pi W_T,$$

while the expression

$$W \equiv \Pi\Omega^2 + (\mathbf{\Omega}\wedge\mathbf{T})^2 = \Pi W_T^2 + W_u^2 + W_v^2\qquad(2.41)$$

has zero Poisson brackets with all of the W_i's (and with $\mathbf{T}$, of course),

$$\{W, W_i\} = 0, \quad (i = T, u, v).\qquad(2.42)$$

Hence we can put

$$\mathfrak{P}_4 = W_T, \quad \mathfrak{Q}_4 = \arctan(W_v/W_u).\qquad(2.43)$$

Then, writing

$$\vec{\mathfrak{Q}} = \mathbf{R} + \mathbf{\Lambda}(\mathbf{T}, H, W_T, W_u, W_v; \mathbf{n}),\qquad(2.44)$$

we look for a $\mathbf{\Lambda}$ such that the conditions

$$\{\vec{\mathfrak{Q}}, \mathfrak{Q}_4\} = 0, \quad \{\vec{\mathfrak{Q}}, \mathfrak{P}_4\} = 0,\qquad(2.45)$$

and

$$\{\mathfrak{Q}_i, \mathfrak{Q}_j\} = 0, \quad (i = 1, 2, 3)\qquad(2.46)$$

are satisfied.

Inserting Eq. (2.44) into Eqs. (2.45), we obtain the system

$$W_v\frac{\partial\mathbf{\Lambda}}{\partial W_u} - W_u\frac{\partial\mathbf{\Lambda}}{\partial W_v} = \frac{c}{HT}(W_u\mathbf{u} + W_v\mathbf{v}),$$

$$\frac{\partial\mathbf{\Lambda}}{\partial W_T} - \Pi\frac{W_T}{W_u^2 + W_v^2}\left(W_u\frac{\partial\mathbf{\Lambda}}{\partial W_u} + W_v\frac{\partial\mathbf{\Lambda}}{\partial W_v}\right)$$

$$= \frac{\mathbf{T}\cdot\mathbf{n}}{T\sqrt{T^2 - (\mathbf{T}\cdot\mathbf{n})^2}}\mathbf{u} - \Pi\frac{c}{HT}\frac{W_T}{W_u^2 + W_v^2}(W_u\mathbf{v} - W_v\mathbf{u}),$$

the general solution of which is

$$\mathbf{\Lambda} = \frac{\mathbf{T}\cdot\mathbf{n}}{T\sqrt{T^2 - (\mathbf{T}\cdot\mathbf{n})^2}}W_T\mathbf{u} + \frac{c}{HT}W_u\mathbf{v} - \frac{c}{HT}W_v\mathbf{u}$$

$$+ g_T(\mathbf{T}, W)\frac{\mathbf{T}}{T} + g_u(\mathbf{T}, W)\mathbf{u} + g_v(\mathbf{T}, W)\mathbf{v},\qquad(2.48)$$

where g_T, g_u, g_v are arbitrary functions of their arguments. It is easy to check that Eq. (2.46) is satisfied with the choice

$$g_T \equiv g_u \equiv g_v \equiv 0.\qquad(2.49)$$

Finally, via Eqs. (2.38), (2.39), (2.25), we have

$$\vec{\mathfrak{Q}} = -\frac{c^2}{H}\mathbf{K} - \frac{\mathbf{\Omega}\cdot\mathbf{T}}{T^2} + \frac{\mathbf{T}\cdot\mathbf{n}}{T(T^2 + (\mathbf{T}\cdot\mathbf{n})^2)}\frac{\mathbf{\Omega}\cdot\mathbf{T}}{T}\mathbf{T}\wedge\mathbf{n}$$

$$= -\frac{c^2}{H}\mathbf{K} + \frac{1}{T}(W_v\mathbf{u} - W_u\mathbf{v}) + \frac{\mathbf{T}\cdot\mathbf{n}}{T(T^2 - (\mathbf{T}\cdot\mathbf{n})^2)}W_T\mathbf{T}\wedge\mathbf{n}.\qquad(2.50)$$

We are left with the construction of the two invariants. Now, while the expression $\Pi = H^2/c^2 - T^2$ is still a good choice, it is clear that the second invariant must be essentially W. Actually, W is the only expression which has zero Poisson brackets with $\mathbf{T}$, H and W_T, W_u, W_v, separately. In order to verify that W commutes with $\vec{\mathfrak{Q}}$, it is sufficient to check that it commutes with $\mathbf{R}$, in force of Eq. (2.44). That this is true it can be seen directly from

$$\{R_i, W_T\} = \frac{c}{HT}(W_u u_i + W_v v_i),$$

$$\{R_i, W_u\} = -\Pi\frac{c}{HT}W_T u_i - \frac{\mathbf{T}\cdot\mathbf{n}}{T\sqrt{T^2 - (\mathbf{T}\cdot\mathbf{n})^2}}W_v u_i,\qquad(2.51)$$

$$\{R_i, W_v\} = -\Pi\frac{c}{HT}W_T v_i + \frac{\mathbf{T}\cdot\mathbf{n}}{T\sqrt{T^2 - (\mathbf{T}\cdot\mathbf{n})^2}}W_u u_i.$$

Note that from the fact that W is an invariant it follows

$$\{\vec{\Omega}, W_i\} = 0, \tag{2.52}$$

not only for $i = T$ but also $i = u, v$. In conclusion, the scheme A results

$$
\boxed{
\begin{array}{llll}
\vec{\mathfrak{P}} = \mathbf{T} & \mathfrak{P}_4 = W_T & \mathfrak{J}_1 = W & \mathfrak{J}_2 = \Pi \\[2mm]
\vec{\Omega} = -\dfrac{c^2}{H}\mathbf{K} + \dfrac{1}{T}(W_v \mathbf{u} - W_u \mathbf{v}) & \Omega_4 = \arctan\dfrac{W_v}{W_u} & & \\[4mm]
\quad + \dfrac{\mathbf{T}\cdot\mathbf{n}}{T^2 - (\mathbf{T}\cdot\mathbf{n})^2}\, W_T \mathbf{T} \wedge \mathbf{n} & & &
\end{array}
}
\tag{A_2}
$$

where $W = \Pi W_T^2 + W_u^2 + W_v^2$ and the W_i's are defined by Eqs. (2.38).

It is apparent that the scheme A_2 remains significant for any real value of the invariant Π. The price we had to pay for this result, however, is the occurrence of the fixed vector $\mathbf{n}$ with the consequence that the transformation properties of the expressions appearing in A_2, under space rotations, are more complicated than the corresponding ones for the expressions of A_1; in particular, $\vec{\Omega}$ is no longer a rotational vector, as it was expected.

It is also clear from the structure of the scheme A_2 that the expressions W_T, W_u, W_v, just as S_z, S_x, S_y in A_1, are simply related to the canonical generators of the little group of the 4-vector $(H/c, \mathbf{T})$. In order to display this relation explicitly we have to distinguish three cases:

(a) $\Pi = M^2 c^2$ definite positive,

(b) $\Pi = -\Delta^2 c^2$ definite negative,

(c) $\Pi = 0$.

(a) Putting

$$S_u = (1/Mc)W_u, \quad S_v = (1/Mc)W_v, \quad S_T = W_T, \tag{2.53}$$

from Eqs. (2.40), (2.41) it follows

$$\{S_i, S_j\} = \epsilon_{ijk}S_k, \quad (i, j, k = T, u, v) \tag{2.54}$$

and

$$S^2 \equiv S_u^2 + S_v^2 + S_T^2 = W/M^2 c^2, \tag{2.55}$$

i. e. , a canonical realization of the Lie algebra of the rotation group $SO(3)$ and the corresponding invariant.

(b) Putting

$$Z_0 = W_T, \quad Z_1 = (1/\Delta c)W_u, \quad Z_2 = (1/\Delta c)W_v, \tag{2.56}$$

from Eqs. (2.40), (2.41) it follows

$$\{Z_0, Z_1\} = Z_2, \quad \{Z_0, Z_2\} = -Z_1, \quad \{Z_1, Z_2\} = -Z_0, \tag{2.57}$$

and

$$Z \equiv Z_0^2 - Z_1^2 - Z_2^2 = -W/\Delta^2 c^2, \tag{2.58}$$

i. e. , a realization of the Lie algebra of the two-dimensional Lorentz group $SO(2,1)$ and its invariant.

(c) Equations (2, 40), (2, 41) directly give

$$\{W_T, W_u\} = W_v, \quad \{W_T, W_v\} = -W_u, \quad \{W_u, W_v\} = 0, \tag{2.59}$$

and

$$W = W_u^2 + W_v^2, \tag{2.60}$$

i. e. , a canonical realization of the Euclidean group of the plane $E(2)$ and the corresponding invariant.[9]

As regards case (a), we find that the expressions S_u, S_v, S_T satisfy the same Poisson bracket relations (2.54) as the expressions S_x, S_y, S_z in scheme A_1 (2.36). Actually, the two sets are related by

$$S_u = \mathbf{S}\cdot\mathbf{u}, \quad S_v = \mathbf{S}\cdot\mathbf{v}, \quad S_T = \mathbf{S}\cdot\mathbf{T}/T. \tag{2.61}$$

Therefore, for Π positive definite, the little group generators in the two different forms of the scheme A can be viewed as the components of the same vector $\mathbf{S}$ referred to two different bases. The fact that $\mathbf{S}$ can be referred to a $\mathbf{T}$-independent basis and the consequent vectorial character of $\vec{\Omega}$ are clearly related to the circumstance that, in the case $\Pi = M^2 c^2 > 0$, a Lorentz frame exists in which the 4-vector $(H/c, \mathbf{T})$ takes the form $(Mc, 0)$ in which any reference to the direction of $\mathbf{T}$ has disappeared.

We have so far exhausted the discussion of the schemes A associated with the *regular* realizations. The proper Poincaré group admits also nontrivial *singular* realizations corresponding to invariant manifolds of lower order of the *irreducible kernel* (see Refs. 1, 3) on which some of the expressions of the schemes A_1 or A_2 become singular.

In the cases $\Pi = M^2 c^2 > 0$ and $\Pi = -\Delta^2 c^2 < 0$, the only submanifolds of this kind correspond to

$$W_u = W_v = W_T = 0, \tag{2.62}$$

i. e. ,

$$\mathbf{\Omega} = 0, \tag{2.63}$$

or, more specifically, to

$$\mathbf{S} = 0 \tag{2.64}$$

for $\Pi > 0$ and to

$$Z_0 = Z_1 = Z_2 = 0 \tag{2.65}$$

for $\Pi < 0$. The reduced scheme A for these submanifolds can be written

$$
\boxed{
\begin{array}{ll}
\vec{\mathfrak{P}} = \mathbf{T} & \mathfrak{J} = \Pi \\[2mm]
\vec{\Omega} = -\dfrac{c^2}{H}\mathbf{K} &
\end{array}
}
\tag{A_3}
$$

This scheme can be considered as the reduction of the schemes A_1 or A_2, independently. Note that the expression $\vec{\Omega}$ in A_3 transforms as a vector under space rotations, independently of the sign of Π.

In the case $\Pi = 0$ we have[10]

$$H = cT, \quad W = W_u^2 + W_v^2 = |\mathbf{\Omega} \wedge \mathbf{T}|^2. \tag{2.66}$$

There are now three different *singular* invariant submanifolds and, correspondingly, three different classes of *singular* realizations. A first *singular* invariant submanifold is defined by

$$W = 0 \quad (\mathbf{\Omega} \wedge \mathbf{T} = 0), \tag{2.67}$$

i. e. ,

$$W_u = W_v = 0, \quad (\mathbf{\Omega} \cdot \mathbf{u} = \mathbf{\Omega} \cdot \mathbf{v} = 0). \tag{2.68}$$

Then the only component of $\mathbf{\Omega}$ which survives,

$$\Gamma \equiv W_T = \frac{\mathbf{\Omega} \cdot \mathbf{T}}{T}, \tag{2.69}$$

becomes an invariant of the corresponding class of realizations and the reduced scheme A can be written

$$\boxed{\begin{array}{ll} \vec{\mathfrak{P}} = \mathbf{T} & \mathfrak{I} = \Gamma \\[2mm] \vec{\mathfrak{Q}} = -\dfrac{c}{T}\mathbf{K} + \dfrac{\mathbf{T} \cdot \mathbf{n}}{T(T^2 - (\mathbf{T} \cdot \mathbf{n})^2)}\Gamma T \wedge \mathbf{n} \end{array}} \tag{A$_4$}$$

A special case of these singular realizations is obtained if

$$W_u = W_v = W_T = 0, \quad (\mathbf{\Omega} = 0), \tag{2.70}$$

i. e., if also $\Gamma = 0$. The reduced scheme A becomes

$$\boxed{\begin{array}{l} \vec{\mathfrak{P}} = \mathbf{T} \\[2mm] \vec{\mathfrak{Q}} = -\dfrac{c}{T}\mathbf{K} \end{array}} \tag{A$_5$}$$

Notice the strict similarity between this scheme and A$_3$. In particular, $\vec{\mathfrak{Q}}$ transforms again as a vector under space rotations.

Finally, a third singular invariant submanifold corresponds to

$$\mathbf{T} = 0, \quad H = 0. \tag{2.71}$$

In this case we are left with the canonical realizations of the homogeneous Lorentz group $SO(3,1)$. The scheme A for this group is derived in the Appendix A by following the standard method. It can be written in the form

$$\boxed{\begin{array}{llll} \mathfrak{P}_1 = J_z & \mathfrak{P}_2 = |\mathbf{J}| = \sqrt{J_x^2 + J_y^2 + J_z^2} & \mathfrak{I}_1 = \mathbf{J}^2 - \mathbf{K}^2 & \mathfrak{I}_2 = \mathbf{J} \cdot \mathbf{K} \\[2mm] \mathfrak{Q}_1 = \arctan\dfrac{J_y}{J_x} & \mathfrak{Q}_2 = \arctan\dfrac{\mathbf{J}(\mathbf{J} \wedge \mathbf{K})_z}{(\mathbf{J} \cdot \mathbf{K})J_z - J^2 K_z} & & \end{array}} \tag{A$_6$}$$

The calculations are quite parallel to those performed in the case $m = 0$, $\mathbf{T} = 0$ for the Galilei group [see Ref. 3, Eqs. (44'')]. No nontrivial *singular* realization exist in this case.

3. IRREDUCIBLE REALIZATIONS

A. Positive mass

The most interesting realizations from a physical point of view are those corresponding to $\Pi = M^2 c^2 > 0$. Among the irreducible realizations of this class, let us consider first the *singular* realizations for which Eq. (2.64) holds. Putting

$$\begin{array}{l} \vec{\mathfrak{P}} = \mathbf{p}, \quad \Pi = m^2 c^2, \\[2mm] \vec{\mathfrak{Q}} = \mathbf{q} \end{array} \tag{3.1}$$

in the scheme A$_3$, we obtain by inversion

$$\begin{array}{l} \mathbf{J} = \mathbf{q} \wedge \mathbf{p}, \quad \mathbf{K} = -\dfrac{H}{c^2}\mathbf{q}, \\[2mm] \mathbf{T} = \mathbf{p}, \quad H = \pm c\,(m^2 c^2 + \mathbf{p}^2)^{1/2} \equiv c p_0. \end{array} \tag{3.2}$$

If we choose the positive sign for H, the above expressions describe a *free spinless particle* characterized by a position $\mathbf{q}$ and a linear momentum $\mathbf{p}$. To realize this we should specify the transformation properties to be required for admissible relativistic position and momentum variables $\mathbf{x}$ and $\mathbf{p}$. First of all, we shall require the usual transformation laws under space translations and rotations

$$\{J_i, x_j\} = \epsilon_{ijk}x_k, \quad \{T_i, x_j\} = -\delta_{ij}, \tag{3.3}$$

$$\{J_i, p_j\} = \epsilon_{ijk}p_k, \quad \{T_i, p_j\} = 0, \quad (i,j,k = x,y,z). \tag{3.4}$$

Then we can argue in the following way for the transformation properties under special Lorentz transformations. An infinitesimal special Lorentz transformation for the space—time coordinates of a moving point can be written

$$\mathbf{x}'(t') = \mathbf{x}(t) - \delta\mathbf{v}t,$$

$$t' = t - \frac{1}{c^2}\delta\mathbf{v} \cdot \mathbf{x}(t), \tag{3.5}$$

i. e., neglecting infinitesimals of higher order,

$$\mathbf{x}'(t) = \mathbf{x}(t) + \dot{\mathbf{x}}(t)\frac{\delta\mathbf{v} \cdot \mathbf{x}(t)}{c^2} - \delta\mathbf{v}t. \tag{3.6}$$

Now, putting $t = 0$ in Eq. (3.6) and writing $\dot{\mathbf{x}} = \{\mathbf{x}, H\}$, we compare the resulting expression with the infinitesimal transformation as written in terms of the canonical generators

$$\mathbf{x}' = \mathbf{x} + \sum_k \delta v_k\{K_k, \mathbf{x}\}. \tag{3.7}$$

The result is

$$\{K_i, x_j\} = \frac{1}{c^2}x_i\{x_j, H\}. \tag{3.8}$$

In a similar way, the requirement that expressions of the form $[u_0(t), \mathbf{u}(t)]$ transform like a 4-vector gives

$$\begin{array}{l} \{K_i, u_j\} = \dfrac{1}{c^2}x_i\{u_j, H\} - \dfrac{1}{c}\delta_{ij}u_0, \\[3mm] \{K_i, u_0\} = \dfrac{1}{c^2}x_i\{u_0, H\} - \dfrac{1}{c}u_i. \end{array} \tag{3.9}$$

It will be noticed the significant difference between the usual transformation laws of the 4-vectors and Eqs. (3.9), or (3.8) for the space part of the space—time corodinate vector, which are typical of the "instant form" of dynamical description and involve the dynamics of the system explicitly.[11] Only in the case that $u_0(t)$, $\mathbf{u}(t)$ are constants of the motion do Eqs. (3.9) coincide with the usual ones. It is easy to check now that the variables $\mathbf{q}$ and $\mathbf{p}$ of the considered realization satisfy Eqs. (3.3), (3.8), and (3.4), (3.9), respectively, thus justifying the interpretation given above. Let us note finally that under an infinitesimal time translation we obtain

$$\begin{array}{l} \mathbf{q}' = \mathbf{q} + \delta\tau\{\mathbf{q}, H\} = \mathbf{q} + \delta\tau\dfrac{c\mathbf{p}}{p_0} \\[3mm] \mathbf{p}' = \mathbf{p} + \delta\tau\{\mathbf{p}, H\} = \mathbf{p}, \end{array} \tag{3.10}$$

as it should be expected.

We consider then the regular irreducible realizations;

putting

$$\vec{\mathfrak{P}} = \mathbf{p}, \quad \mathfrak{Q}_4 = p_\chi, \quad S = s, \quad \sqrt{\Pi} = mc,$$

$$\vec{\Omega} = \mathbf{q}, \quad \mathfrak{Q}_4 = \chi \tag{3.11}$$

in the scheme A_1, we obtain by inversion

$$\mathbf{J} = \mathbf{q} \wedge \mathbf{p} + \mathbf{S}, \quad \mathbf{K} = -\frac{H}{c^2}\mathbf{q} + \frac{\mathbf{S} \wedge \mathbf{p}}{mc^2 + H},$$

$$\mathbf{T} = \mathbf{p}, \quad H = cp_0 = +c\sqrt{m^2c^2 + \mathbf{p}^2}, \tag{3.12}$$

with

$$S_x = \sqrt{s^2 - p_\chi^2}\cos\chi,$$

$$S_y = \sqrt{s^2 - p_\chi^2}\sin\chi, \tag{3.13}$$

$$S_z = p_\chi, \quad (0 \le \chi < 2\pi, -s \le p_\chi \le +s).$$

(A single point of the phase space corresponds to the value $-s$ or $+s$ of p_χ for given $\mathbf{q}, \mathbf{p}$, independently of the value of χ.) What happens now is that while $\mathbf{q}$ continues to satisfy Eqs. (3.3) and (3.10), it no longer satisfies Eq. (3.8). Nevertheless, an expression satisfying also Eq. (3.8), can still be constructed in the form[11-13]

$$\mathbf{x} = \mathbf{p} + \frac{\mathbf{S} \wedge \mathbf{p}}{m(mc^2 + H)}. \tag{3.14}$$

Actually, this space vector turns out to be the most general solution of the conditions (3.3), (3.8), (3.10) within this class of irreducible realizations if the time-reflection condition ($I_t^* \mathbf{x} = \mathbf{x}$) is imposed (see Sec. 3f) and therefore it is the only entitled to be interpreted as the position variable of a particle. Note, however, that

$$\{x_i, x_j\} = \frac{1}{mH}\,\epsilon_{ijk}\left[S_k + \frac{\mathbf{S} \cdot \mathbf{p}}{m(mc^2 + H)}\,p_k\right], \tag{3.15}$$

so that the components of the vector $\mathbf{x}$ cannot be assumed as canonical variables unless the spin is zero. Furthermore, in the rest frame of the system ($\mathbf{p} = 0$) we have $\mathbf{x} = \mathbf{q}$. In the following we shall refer to $\mathbf{x}$ and $\mathbf{q}$ as the *covariant position vector* and the *canonical position vector*, respectively. Summarizing, the realization defined by Eqs. (3.12), (3.13) describes a *free particle with spin s*, position $\mathbf{x}$ and linear momentum $\mathbf{p}$.

Finally, for a better comparison with the other classes of regular irreducible realizations, it is profitable to give also the canonically equivalent form of the generators which is obtained by the inversion of the scheme A_2 rather than A_1. Putting

$$\vec{\mathfrak{P}} = \mathbf{p}, \quad \mathfrak{Q}_4 = p_\tau, \quad W = w, \quad \Pi = m^2c^2,$$

$$\vec{\Omega} = \vec{\mathfrak{q}}, \quad \mathfrak{Q}_4 = \tau \tag{3.16}$$

in the scheme A_2, we have

$$\mathbf{J} = \vec{\mathfrak{q}} \wedge \mathbf{p} + \frac{p W_T}{p^2 - (\mathbf{p} \cdot \mathbf{n})^2}\,(\mathbf{p} - (\mathbf{p} \cdot \mathbf{n})\mathbf{n}),$$

$$\mathbf{K} = -\frac{H}{c^2}\vec{\mathfrak{q}} + \frac{1}{cp}(W_u\mathbf{v} - W_v\mathbf{u}) + \frac{H}{c^2 p}\frac{(\mathbf{p} \cdot \mathbf{n})W_T}{\sqrt{p^2 - (\mathbf{p} \cdot \mathbf{n})^2}}\mathbf{u}, \tag{3.17}$$

$$\mathbf{T} = \mathbf{p}, \quad H = +c\sqrt{m^2c^2 + \mathbf{p}^2},$$

with

$$W_u = \sqrt{w - m^2c^2p_\tau^2}\cos\tau,$$

$$W_v = \sqrt{w - m^2c^2p_\tau^2}\sin\tau, \tag{3.18}$$

$$W_T = p_\tau, \quad \left(0 \le \tau < 2\pi, \quad -\frac{\sqrt{w}}{mc} \le p_\tau \le \frac{\sqrt{w}}{mc}\right),$$

$$\mathbf{u} = \frac{\mathbf{p} \wedge \mathbf{n}}{\sqrt{p^2 - (\mathbf{p} \cdot \mathbf{n})^2}}, \quad \mathbf{v} = \frac{\mathbf{p}}{p} \wedge \mathbf{u}, \tag{3.19}$$

and where

$$w = m^2c^2s^2. \tag{3.20}$$

Note that $\vec{\mathfrak{q}}$ does not transform as a vector under space rotations, in contrast with $\mathbf{q}$. Note, moreover, that, in order the realization is defined unambiguously, a single point of the phase space must correspond again to the value $\sqrt{w}/mc$ or $-\sqrt{w}/mc$ of p_τ, for given $\vec{\mathfrak{q}}$ and $\mathbf{p}$, independently of the value of τ.

B. Imaginary mass

Two types of irreducible realizations exist also for $\Pi = -\Delta^2c^2 < 0$, the *singular* realizations for which the little group variables Z_0, Z_1, Z_2 are identically zero [see Eq. (2.65)], and the regular realizations. The *singular* realizations correspond to the spin zero case for $\Pi > 0$ and should describe free scalar *tachyons*. These realizations can be obtained again by inversion of the scheme A_3 putting

$$\vec{\mathfrak{P}} = \mathbf{p}, \quad \mathfrak{J} = \Pi = -\delta^2c^2,$$

$$\vec{\Omega} = \mathbf{q}. \tag{3.21}$$

It follows

$$\mathbf{J} = \mathbf{q} \wedge \mathbf{p}, \quad \mathbf{K} = -\frac{H}{c^2}\mathbf{q},$$

$$\mathbf{T} = \mathbf{p}, \quad H = \pm c\sqrt{\mathbf{p}^2 - \delta^2c^2}. \tag{3.22}$$

The appearance of the double sign in front of the expression of H is a consequence of the fact that $H/|H|$ is no longer an invariant for $\Pi < 0$. The generators can be expressed unambiguously introducing a four-dimensional polar representation of the energy—momentum $(H/c, \mathbf{T})$ as follows:

$$\frac{H}{c} = \delta c \sinh p_\xi,$$

$$p_1 = \delta c \cosh p_\xi \sin p_\varphi \cos p_\psi,$$

$$p_2 = \delta c \cosh p_\xi \sin p_\varphi \sin p_\psi, \tag{3.23}$$

$$p_3 = \delta c \cosh p_\xi \cos p_\varphi,$$

with $-\infty < p_\xi < +\infty$, $0 \le p_\varphi \le \pi$, $0 \le p_\psi < 2\pi$. Then the corresponding "configurational" conjugate variables are[14]

$$\xi = -\frac{\partial\Phi}{\partial p_\xi} = -\delta c[q_1 \sinh p_\xi \sin p_\varphi \cos p_\psi$$

$$+ q_2 \sinh p_\xi \sin p_\varphi \sin p_\psi + q_3 \sinh p_\xi \cos p_\varphi],$$

$$\varphi = -\frac{\partial\Phi}{\partial p_\varphi} = -\delta c[q_1 \cosh p_\xi \cos p_\varphi \cos p_\psi \tag{3.24}$$

$$+ q_2 \cosh p_\xi \cos p_\varphi \sin p_\psi - q_3 \cosh p_\xi \sin p_\varphi],$$

$$\psi = -\frac{\partial\Phi}{\partial p_\psi} = -\delta c[-q_1 \cosh p_\xi \sin p_\varphi \sin p_\psi$$

$$+ q_2 \cosh p_\xi \, \sin p_\varphi \cos p_\psi]$$

$$[\Phi = \mathbf{p}(p_\xi, p_\varphi, p_\psi) \cdot \mathbf{q}]$$

The irreducible realizations of the regular type are defined by the inversion of the scheme A_2 with the same positions as in Eq. (3.16) except for $\Pi = m^2 c^2$ which is replaced by

$$\Pi = -\delta^2 c^2. \tag{3.25}$$

It follows[15]

$$\mathbf{J} = \vec{\mathbf{q}} \wedge \mathbf{p} + \frac{W_T p}{\mathbf{p}^2 - (\mathbf{p} \cdot \mathbf{n})^2} (\mathbf{p} - (\mathbf{p} \cdot \mathbf{n})\mathbf{n}),$$

$$\mathbf{K} = -\frac{H}{c^2} \vec{\mathbf{q}} + \frac{1}{cp}(W_u \mathbf{v} - W_v \mathbf{u}) + \frac{H}{c^2 p} \frac{(\mathbf{p} \cdot \mathbf{n}) W_T}{\sqrt{\mathbf{p}^2 - (\mathbf{p} \cdot \mathbf{n})^2}} \mathbf{u}, \tag{3.26}$$

$$\mathbf{T} = \mathbf{p}, \quad H = \pm c \sqrt{\mathbf{p}^2 - \delta^2 c^2},$$

with

$$W_T = p_\tau,$$

$$W_u = \sqrt{w + \delta^2 c^2 p_\tau^2} \cos \tau, \tag{3.27}$$

$$W_v = \sqrt{w + \delta^2 c^2 p_\tau^2} \sin \tau,$$

and $\mathbf{u}, \mathbf{v}$ are given by Eqs. (3.19). Note that using the natural invariant of the little group [Eq. (2.58)], we have

$$w = -\delta^2 c^2 \zeta \equiv -\delta^2 c^2 (Z_0^2 - Z_1^2 - Z_2^2). \tag{3.28}$$

The range of τ in Eq. (3.27) is obviously $0 \leqslant \tau < 2\pi$ as in Eq. (3.18). On the other hand, the range of p_τ and correspondingly the topological structure of the phase space are not unique: Three different subclasses correspond to the values $\zeta \gtrless 0$, i.e., to (Z_0, Z_1, Z_2) being a timelike, spacelike or a null vector in the three-dimensional Minkowsky space. In the case $\zeta > 0$ ($w < 0$), the open interval $(-\sqrt{\zeta}, +\sqrt{\zeta})$ must be excluded and the allowed range of p_τ is confined to $[\sqrt{\zeta}, +\infty)$, (or $(-\infty, -\sqrt{\zeta}]$) since the quantity $Z_0/|Z_0|$ is an additional invariant under the proper transformations. (Again, for given $\vec{\Omega}$ and $\mathbf{p}$, the value $p_\tau = \pm\sqrt{\zeta}$ has to be intended as specifying a single point of the phase space, independently of the value of τ). In the case $\zeta = 0$ ($w = 0$), three invariant manifolds exist under the proper transformations corresponding to $Z_0 > 0$, $Z_0 = 0$, $Z_0 < 0$. Obviously $Z_0 = 0$ implies now $Z_1 = Z_2 = 0$ with the consequence that $\vec{\Omega}$ and $\mathbf{p}$ transform according to the *singular* realizations. Apart from this singular manifold, already considered, the range of p_τ must be restricted to $(0, +\infty)$, [or $(-\infty, 0)$]. In the case $\zeta < 0$ ($w > 0$), no additional invariant occurs and the range of p_τ covers the entire real axis. Finally, we can dispose of the double valuedness of the generator H by introducing the variables (3.23), (3.24).

C. Zero mass

For $\Pi \equiv 0$ and $\mathbf{T} \neq 0$ we have three subclasses of irreducible realizations, a regular subclass corresponding to $w \neq 0$, and two singular subclasses corresponding to $w = 0$, $\Gamma \neq 0$ and $w = 0$, $\Gamma = 0$, respectively.

The *singular* realization corresponding to $\Gamma = 0$ is obtained by inverting the scheme A_5 in the usual way:

$$\mathbf{J} = \mathbf{q} \wedge \mathbf{p}, \quad \mathbf{K} = -(p/c)\mathbf{q},$$

$$\mathbf{T} = \mathbf{p}, \quad H = cp. \tag{3.29}$$

The *singular* realizations corresponding to $\Gamma \neq 0$ are obtained from the scheme A_4 with the position $\Gamma = \gamma$ (γ = arbitrary real constant)

$$\mathbf{J} = \vec{\Omega} \wedge \mathbf{p} + \frac{\gamma p}{\mathbf{p}^2 - (\mathbf{p}^2 \cdot \mathbf{n})^2}(\mathbf{p} - (\mathbf{p} \cdot \mathbf{n})\mathbf{n}),$$

$$\mathbf{K} = -\frac{p}{c}\vec{\Omega} + \frac{1}{c}\gamma \frac{\mathbf{p} \cdot \mathbf{n}}{\mathbf{p}^2 - (\mathbf{p}^2 \cdot \mathbf{n})^2} \mathbf{p} \wedge \mathbf{n}, \tag{3.30}$$

$$\mathbf{T} = \mathbf{p}, \quad H = cp.$$

Finally, the regular realizations are obtained putting

$$\vec{\mathfrak{P}} = \mathbf{p}, \quad \mathfrak{P}_4 = p_\tau, \quad \mathfrak{I}_1 \equiv W = w, \quad \mathfrak{I}_2 \equiv \Pi = 0,$$

$$\vec{\Omega} = \vec{\mathbf{q}}, \quad \Omega_4 = \tau, \tag{3.31}$$

in the scheme A_2. By inversion it follows

$$\mathbf{J} = \vec{\mathbf{q}} \wedge \mathbf{p} + \frac{W_T p}{\mathbf{p}^2 - (\mathbf{p} \cdot \mathbf{n})^2}(\mathbf{p} - (\mathbf{p} \cdot \mathbf{n})\mathbf{n}),$$

$$\mathbf{K} = -\frac{p}{c}\vec{\mathbf{q}} + \frac{1}{cp}(W_u \mathbf{v} - W_v \mathbf{u}) + \frac{1}{c}\frac{(\mathbf{p} \cdot \mathbf{n}) W_T}{\sqrt{\mathbf{p}^2 - (\mathbf{p} \cdot \mathbf{n})^2}} \mathbf{u}, \tag{3.32}$$

$$\mathbf{T} = \mathbf{p}, \quad H = cp,$$

where now

$$W_T = p_\tau,$$

$$W_u = \sqrt{w} \cos \tau, \tag{3.33}$$

$$W_v = \sqrt{w} \sin \tau, \quad (0 \leqslant \tau < 2\pi, \quad -\infty < p_\tau < +\infty),$$

and $\mathbf{u}, \mathbf{v}$ are given again by Eqs. (3.19). Note that W is directly the natural invariant of the little group and within these irreducible realizations it must be a positive number. The realizations (3.32) correspond to faithful realizations of the Euclidean group $E(2)$ [see Eqs. (3.33)] while the realizations (3.29) and (3.30) correspond to unfaithful realizations of the same group.

D. Zero energy-momentum: $H = 0$, $\mathbf{T} = 0$ (homogeneous Lorentz group)

Putting

$$\mathfrak{P}_1 = p_\alpha, \quad \mathfrak{P}_2 = p_\beta, \quad \mathfrak{I}_1 = j_1 \gtrless 0, \quad \mathfrak{I}_2 = j_2 \gtrless 0,$$

$$\Omega_1 = \alpha, \quad \Omega_2 = \beta, \tag{3.34}$$

in the scheme A_6 and inverting it, we obtain

$$J_x = \sqrt{p_\beta^2 - p_\alpha^2} \cos \alpha,$$

$$J_y = \sqrt{p_\beta^2 - p_\alpha^2} \sin \alpha,$$

$$J_z = p_\alpha,$$

$$K_x = \frac{1}{p_\beta^2}[j_2 \sqrt{p_\beta^2 - p_\alpha^2} \cos \alpha + \sqrt{p_\beta^2(p_\beta^2 - j_1) - j_2^2}$$

$$\times (p_\alpha \cos \alpha \cos \beta - p_\beta \sin \alpha \sin \beta)] \tag{3.35}$$

$$K_y = \frac{1}{p_\beta^2}[j_2 \sqrt{p_\beta^2 - p_\alpha^2} \sin \alpha + \sqrt{p_\beta^2(p_\beta^2 - j_1) - j_2^2}$$

$$\times (p_\alpha \sin \alpha \cos \beta + p_\beta \cos \alpha \sin \beta)]$$

$$K_z = \frac{1}{p_\beta^2}[j_2 p_\alpha - \sqrt{p_\beta^2 - p_\alpha^2}\sqrt{p_\beta^2(p_\beta^2 - j_1) - j_2^2} \cos \beta].$$

The global structure of these irreducible realizations is quite different according to whether $j_1 \gtrless 0$ and $j_2 = 0$ or $j_2 \neq 0$. However, we will not pursue further the discussion on this subject.

E. Limiting relations among the different classes of irreducible realizations

As we have seen, in the case $\Pi = m^2 c^2 > 0$ it is always possible to construct a covariant position vector, i. e., an expression $\mathbf{x}$ satisfying Eqs. (3.3), (3.8), (3.10). This vector coincides with $\mathbf{q}$ for $S = 0$ and it given by Eq. (3.14) for $S = s > 0$. On the contrary, if $\Pi = 0$ or $\Pi = -\delta^2 c^2 < 0$, a covariant position vector exists only in the "scalar" case where the little group is trivially realized ($W_T = W_u = W_v = 0$) and then coincides with the canonical vector $\mathbf{q}$. In the "nonscalar" case ($W \neq 0$) the canonical variables ($q_1 q_2 q_3$) themselves do not transform as vector components under space rotations. A quantity which closely resembles a position vector and does exist in any case is[11-13]

$$\mathbf{R} = -(c^2/H)\mathbf{K}. \qquad (3.36)$$

This vector satisfies Eqs. (3.3) and (3.10), but not, however, Eq. (3.8); i. e., it does not correspond to an invariantly defined world line. Nevertheless, it is convenient to reexpress the canonical generators for all the classes of irreducible realizations by replacing $\mathbf{R}$ for $\mathbf{q}$ ($\mathbf{q}$) as independent variables. Then we have in general

$$\mathbf{J} = \mathbf{R} \wedge \mathbf{p} + \mathbf{\Omega}, \quad \mathbf{K} = -\frac{H}{c^2}\mathbf{R},$$
$$\mathbf{T} = \mathbf{p}, \quad H = \pm c\sqrt{\Pi + \mathbf{p}^2}, \qquad (3.37)$$

where the double sign refers to the case $\Pi = -\delta^2 c^2 < 0$ only. Once this form is adopted, apart from the value of Π, the various realizations differ essentially for the expression of $\mathbf{\Omega}$. In the regular realizations $\mathbf{\Omega}$ assumes the common expression

$$\mathbf{\Omega} = p_\tau \frac{\mathbf{p}}{p} + \frac{c}{H}\sqrt{w - \Pi p_\tau}\ (\mathbf{u}\cos\tau + \mathbf{v}\sin\tau), \qquad (3.38)$$

valid for any value $\Pi \gtrless 0$. In the singular case $\Pi = 0$, $w = 0$, $\Gamma \neq 0$ [see Eqs. (3.30)] we have simply

$$\mathbf{\Omega} = \gamma(\mathbf{p}/p). \qquad (3.39)$$

Finally, in the three "scalar" singular cases ($\Pi \gtrless 0$) we have

$$\mathbf{\Omega} = 0, \quad \mathbf{R} = \mathbf{q}. \qquad (3.40)$$

The nontrivial Poisson bracket relations among the independent variables of the unified form (3.37) are

$$\{R_i, R_j\} = -\frac{c^2}{H^2}\epsilon_{ijk}\Omega_k, \quad \{R_i, p_j\} = \delta_{ij} \qquad (3.41)$$

[see Eqs. (2.27)] in all cases, and in addition

$$\{R_i, \tau\} = \frac{\mathbf{p} \circ \mathbf{n}}{p\sqrt{p^2 - (\mathbf{p} \cdot \mathbf{n})^2}}u_i$$
$$+ \frac{cH}{pH}\frac{p_\tau}{\sqrt{w - \Pi p_\tau^2}}\ (u_i \sin\tau - v_i \cos\tau) \qquad (3.42)$$

$$\{R_i, p_\tau\} = \frac{c}{pH}\sqrt{w - \Pi p_\tau^2}\ (u_i \cos\tau + v_i \sin\tau)$$

in the regular cases.

We want now to discuss the way the various irreducible realizations for zero mass can be recovered by means of a suitable limiting process from both the positive and the imaginary mass irreducible realizations. The problem is trivial for the "scalar" realization

(3.29) which is immediately obtained from Eqs. (3.2) for $m \to 0$ and from Eqs. (3.22) for $\delta \to 0$, when attention is paid to the fact that in the last case the limiting process brings to two distinct irreducible realizations corresponding to the double sign of H. On the other hand, two different limiting processes can be considered for the regular realizations corresponding to $\Pi = m^2 c^2 > 0$ and $\Pi = -\delta^2 c^2 < 0$. According to the first one, w is kept fixed when $\Pi \to 0$: Then Eqs. (3.32), (3.33) are again trivially obtained from Eqs. (3.17), (3.18) letting $m \to 0$ and from Eqs. (3.26), (3.27) letting $\delta \to 0$ *provided that the subclass* $w > 0$ ($\zeta < 0$) *is considered in the latter case.* In the second one, the natural little group invariants s and ζ are kept fixed when $\Pi \to 0$ with the consequences that $w \to 0$. In this case [*independently of the sign of* w (ζ) *if* $\Pi = -\delta^2 c^2 < 0$], ($W_u, W_v) \to (0, 0)$, $W_T = p_\tau$ becomes an invariant, the internal variables τ and p_τ go over into the *second set* of the scheme B (see Ref. 1) and the realization is no longer irreducible. If, however, we restrict ourselves to an invariant submanifold $W_T = \gamma = $ const, the internal variables are ruled out and the *singular* irreducible realizations (3.30), (3.29) are recovered. Of course these limiting processes can be directly discussed in terms of the unified form (3.27) and its particularizations to the various cases. Let us observe that keeping w fixed as $m \to 0$ in the case $\Pi = m^2 c^2 > 0$ amounts to let s going simultaneously to $+\infty$. For this reason the regular irreducible realizations for zero mass can be interpreted as describing a particle with infinite spin. On the other hand, the irreducible realizations corresponding to $\Pi = 0$, $w = 0$, $\Gamma = \gamma \neq 0$ can be naturally interpreted as describing zero mass particles with a fixed helicity. The analogy with the unitary quantum representations classified by Wigner[16] is apparent.

F. Reflection properties

We want to discuss now the improper transformations of the full Poincaré group. Our attitude here will be to seek for the irreducible realizations of them by means of canonical or anticanonical mappings *within the phase spaces of the irreducible realizations of the proper group* classified in the preceding subsections. Clearly, other irreducible realizations could always be constructed for all kinds of space and time reflections by allowing for suitable disconnected phase spaces. This point of view, however, seems to be quite unnatural in a classical framework.

As we have seen, a definite mapping of the space of the canonical generators onto itself exists for each canonical or anticanonical realization of the space and time reflections [see Table (2.17)]. Therefore, since every irreducible realization of the proper group is connected with an irreducible manifold in the generator space (invariant under the proper group), it is necessary that this manifold is transformed into itself also under the discrete mapping if the corresponding reflection has to be canonically (anticanonically) realized in the sense required above. {Note that a canonical or anticanonical transformation in the space of the generators of a Lie group necessarily maps an irreducible invariant manifold into another irreducible invariant manifold. Let $\{y_\rho, y_\sigma\} = c_{\rho\sigma}^\tau y_\tau + y_{\rho\sigma}$ be the Poisson brackets

among the generators and consider an invertible mapping $(y_1, \ldots, y_r) \to (y_1', \ldots, y_r')$ such that $\{y_\rho', y_\sigma'\}$ $= \pm (c_{\rho\sigma}^\tau y_\tau + d_{\rho\sigma})$. Let $\Im_1(y_1, \ldots, y_r), \ldots, \Im_k(y_1, \ldots, y_r)$ be a possible choice of the independent canonical invariants. Then also $\Im_1(y_1', \ldots, y_r'), \ldots, \Im_k(y_1', \ldots, y_r')$ are invariants because

$$\{\Im_t(y_1', \ldots, y_r'), y_\rho\} = \frac{\partial y_\rho}{\partial y_\sigma'}\{\Im_t(y_1', \ldots, y_r'), y_\sigma'\} = 0$$

$(t = 1, \ldots, k)$ holds true. Therefore, we must have

$$\Im_t(y_1', \ldots, y_r') = f_t[\Im_1(y_1, \ldots, y_r), \ldots, \Im_k(y_1, \ldots, y_r)]$$

and consequently the manifold $\Im_1(y_1, \ldots, y_r) = c_1, \ldots,$ $\Im_k(y_1, \ldots, y_k) = c_k$ is mapped into the manifold $\Im_1(y_1, \ldots, y_k) = f_1(c_1, \ldots, c_k), \ldots, \Im_k(y_1, \ldots, y_r)$ $= f_k(c_1, \ldots, c_k).\}$ Provided that the invariance of the irreducible manifolds in the generators space is assured also for the discrete operations, the transformation properties of the canonical variables q_i, p_j are uniquely determined by the equations

$$\mathfrak{Q}_i(\mathbf{J}, \mathbf{K}, \mathbf{T}, H) = q_i,$$
$$\mathfrak{P}_j(\mathbf{J}, \mathbf{K}, \mathbf{T}, H) = p_j \qquad (3.43)$$
$$(i, j = 1, \ldots, \tfrac{1}{2}\text{dim. kernel})$$

and the mappings (2.17).

Accordingly, we have the following situation:

1. Case $\Pi = m^2 c^2 > 0$

The manifold

$$\Pi = m^2 c^2, \qquad S = s \qquad (3.44)$$

is composed of two disconnected parts corresponding to different signs of H which are separately invariant under the proper group. All of the mappings classified in Table (2.17) leave the manifold (3.44) invariant. However, the anticanonical space reflection and the canonical time reflection change the sign of H. Therefore, only the *canonical space reflection* and the *anticanonical time reflection* can be realized in both the regular $(s \neq 0)$ and the *singular* $(s = 0)$ irreducible realizations.

2. Case $\Pi = \delta^2 c^2 < 0$

The reflection properties of the little group variables result [see Eqs. (2.17), (2.38), (2.39)]:

Transformation	Canonical real.	Anticanonical real.
Space reflection	$W_T \to -W_T$ $W_u \to -W_u$ $W_v \to W_v$	$W_T \to -W_T$ $W_u \to W_u$ $W_v \to W_v$
Time reflection	$W_T \to W_T$ $W_u \to -W_u$ $W_v \to -W_v$	$W_T \to W_T$ $W_u \to W_u$ $W_v \to -W_v$

$$(3.45)$$

We distinguish the following subcases:

(a) $\zeta = -w/\delta^2 c^2 < 0$: The manifold

$$\Pi = -\delta^2 c^2, \qquad Z = \zeta, \qquad (3.46)$$

is irreducible and invariant under all of the full group

transformations. Therefore, *all different kinds of reflections can be realized* within the corresponding irreducible realizations.

(b) $\zeta = -w/\delta^2 c^2 > 0$: The manifold (3.46) breaks into two disconnected parts corresponding to different signs of $Z_0 = W_T$, which are separately invariant under the proper group. Therefore, only the transformations which do not change this sign can be realized in our sense, namely *the canonical and the anticanonical time reflections*.

(c) $\zeta = -w/\delta^2 c^2 = 0$: The invariant manifold (3. 46) breaks into three disconnected parts corresponding to $Z_0 > 0$, $Z_0 < 0$, $Z_0 = 0$ and separately invariant under the proper group. In the first two cases only *the canonical and anticanonical time reflection* are admitted while in the third case, which corresponds to the singular realizations, *all kinds of reflections are realizable*.

3. Case $\Pi = 0$

The irreducible invariant manifolds (proper group) of the regular realizations are defined by

$$W = w > 0,$$
$$H/|H| = +1 \quad \text{or} \quad H/|H| = -1. \qquad (3.47)$$

Therefore, only the *canonical space reflection* and the *anticanonical time reflection* can be realized. The same situation occurs for the "scalar" singular case $w = 0$, $\gamma = 0$. On the other hand, in the singular case with a fixed helicity $(w = 0, \gamma \neq 0)$, the irreducible manifolds invariant under the proper group are

$$\Gamma \equiv W_T = \gamma,$$
$$H/|H| = +1 \quad \text{or} \quad H/|H| = -1, \qquad (3.48)$$

and only the *anticanonical time reflection* can be realized.

4. Case $\Pi = 0$; $H = 0$, $\mathbf{T} = 0$

Taking into account the reflection properties of the canonical invariant $\Im_2 = \mathbf{J} \cdot \mathbf{K}$, it is easily seen that *all kinds of reflection transformations can be realized* provided that $j_2 = 0$; on the other hand, if $j_2 \neq 0$, only *the canonical and anticanonical space—time reflections are realizable*.

In conclusion, we summarize the results in the tables at the end of this paper, where the transformation properties of the canonical variables are explicitly given. Note that the detailed transformations of the "internal" variables τ, p_τ, (χ, p_χ) in the various cases can be simply derived from Eqs. (3.45) on the basis of their geometrical interpretation: See in this connection Eqs. (2.43) and the discussions given in Ref. (2).

4. TWO-PARTICLE SYSTEM

A. Free particles

Let us introduce two sets of canonical variables q_1, p_1, q_2, p_2 and write

$$\mathbf{J} = q_1 \wedge p_1 + q_2 \wedge p_2,$$

$$\mathbf{K} = -\frac{p_{10}}{c}\mathbf{q}_1 - \frac{p_{20}}{c}\mathbf{q}_2, \tag{4.1}$$

$$\mathbf{T} = \mathbf{p}_1 + \mathbf{p}_2,$$

$$H = cp_{10} + cp_{20}, \quad p_{i0} = \sqrt{m_i^2 c^2 + \mathbf{p}^2} \quad (i = 1, 2).$$

This realization describes two free particles without spin and masses m_1, m_2. The variables $(\mathbf{q}_1, \mathbf{q}_2)$, $(p_{10}, \mathbf{p}_1; p_{20}, \mathbf{p}_2)$ satisfy Eqs. (3.3), (3.8) and (3.4), (3.9), respectively; then $\mathbf{q}_1$ and $\mathbf{q}_2$ can be assumed to represent the positions of the particles and $(p_{10}, \mathbf{p}_1)$, $(p_{20}, \mathbf{p}_2)$ their linear 4-momenta. In order to construct the scheme B (see Refs. 1—3), we first define the expressions

$$\mathbf{P} \equiv \vec{\mathfrak{P}} = \mathbf{p}_1 + \mathbf{p}_2, \tag{4.2}$$

$$\mathbf{Q} \equiv \vec{\mathfrak{Q}} = \frac{p_{10}(Mc^2 + H) - c\mathbf{p}_1 \cdot \mathbf{P}}{Mc(Mc^2 + H)}\mathbf{q}_1 + \frac{p_{20}(Mc^2 + H) - c\mathbf{p}_2 \cdot \mathbf{P}}{Mc(Mc^2 + H)}\mathbf{q}_2$$

$$+ \frac{c(\mathbf{q}_1 - \mathbf{q}_2) \cdot \mathbf{P}}{MH(Mc^2 + H)} (p_{20}\mathbf{p}_1 - p_{10}\mathbf{p}_2), \tag{4.3}$$

where $\Pi = M^2 c^2$ and $\vec{\mathfrak{P}}$ and $\vec{\mathfrak{Q}}$ are given in terms of the canonical generators according to the scheme A_1. These expressions represent "external" variables, being $\mathbf{P}$ the total momentum and $\mathbf{Q}$ the *canonical center of mass* (see below). Then, we have to construct six quantities which have zero Poisson brackets with $\mathbf{P}$ and $\mathbf{Q}$. A well-known possible choice is in the form of two vectors π and ρ obtained in the following way: First one defines π as the relative momentum in the center-of-mass system (more precisely in the center-of-momentum system $\mathbf{P} = 0$); then one performs the canonical transformation

$$\mathbf{Q} = \frac{\partial F}{\partial \mathbf{P}}, \quad \rho = \frac{\partial F}{\partial \pi}, \tag{4.4}$$

where $F = \mathbf{q}_1 \cdot \mathbf{p}_1(\mathbf{P}, \pi) + \mathbf{q}_2 \cdot \mathbf{p}_2(\mathbf{P}, \pi)$ (see also Refs. 5 and 17). The result is

$$\pi = \frac{p_{20}(Mc^2 + H) - c\mathbf{p}_2 \cdot \mathbf{P}}{Mc(Mc^2 + H)}\mathbf{p}_1 - \frac{p_{10}(Mc^2 + H) - c\mathbf{p}_1 \cdot \mathbf{P}}{Mc(Mc^2 + H)}\mathbf{p}_2, \tag{4.5}$$

$$\rho = \mathbf{q}_1 - \mathbf{q}_2 + (\mathbf{q}_1 - \mathbf{q}_2) \cdot \mathbf{P}\left[\frac{\mathbf{P}}{M(Mc^2 + H)} + \frac{1}{Mc}\left(\frac{1}{\pi_{10}} - \frac{1}{\pi_{20}}\right)\right.$$

$$\left. - \frac{c\pi \cdot \mathbf{P}}{H\pi_{10}\pi_{20}}\right)\pi\right] \quad (\pi_{i0} = \sqrt{m_i^2 c^2 + \pi^2}, \quad i = 1, 2). \tag{4.6}$$

For future reference we give also the inverted formulas:

$$\mathbf{q}_1 = \mathbf{Q} - \frac{\rho \cdot \mathbf{P}}{M(Mc^2 + H)}\frac{c\pi_{10}\pi_{20}}{\pi_{10}H + c\pi \cdot \mathbf{P}}\mathbf{P} + \frac{1}{Mc}\left(\pi_{20} - \frac{c\pi \cdot \mathbf{P}}{Mc^2 + H}\right)\rho$$

$$+ \frac{\rho \cdot \mathbf{P}}{M}\left(\frac{1}{Mc^2 + H} - \frac{\pi_{20}}{\pi_{10}H + c\pi \cdot \mathbf{P}}\right)\pi,$$

$$\mathbf{q}_2 = \mathbf{Q} + \frac{\rho \cdot \mathbf{P}}{M(Mc^2 + H)}\frac{c\pi_{10}\pi_{20}}{\pi_{20}H - c\pi \cdot \mathbf{P}}\mathbf{P} - \frac{1}{Mc}\left(\pi_{10} + \frac{c\pi \cdot \mathbf{P}}{Mc^2 + H}\right)\rho$$

$$+ \frac{\rho \cdot \mathbf{P}}{M}\left(\frac{1}{Mc^2 + H} - \frac{\pi_{10}}{\pi_{20}H - c\pi \cdot \mathbf{P}}\right)\pi, \tag{4.7}$$

$$\mathbf{p}_1 = \pi + \frac{\pi_{10}(Mc^2 + H) + c\pi \cdot \mathbf{P}}{Mc(Mc^2 + H)}\mathbf{P}, \quad p_{10} = \frac{\pi_{10}H + c\pi \cdot \mathbf{P}}{Mc^2},$$

$$\mathbf{p}_2 = -\pi + \frac{\pi_{20}(Mc^2 + H) - c\pi \cdot \mathbf{P}}{Mc(Mc^2 + H)}\mathbf{P}, \quad p_{20} = \frac{\pi_{20}H - c\pi \cdot \mathbf{P}}{Mc^2}.$$

The expressions (4.5), (4.6) together with $\mathbf{P}$ and $\mathbf{Q}$ provide a system of twelve canonical variables. As said above π is the relative momentum in the system $\mathbf{P} = 0$, while ρ is a kind of relative coordinate the meaning of which does not seem to have been exhaustively discussed in the literature. In terms of the variables $\mathbf{P}, \mathbf{Q}$ and π, ρ, the canonical generators (4.1) assume the form

$$\mathbf{J} = \mathbf{Q} \wedge \mathbf{P} + \mathbf{S},$$

$$\mathbf{K} = -\frac{H}{c^2}\mathbf{Q} + \frac{\mathbf{S} \wedge \mathbf{P}}{Mc^2 + H},$$

$$\mathbf{T} = \mathbf{P}, \tag{4.8}$$

$$H \equiv cP_0 = +c\sqrt{M^2 c^2 + \mathbf{P}^2},$$

with

$$\mathbf{S} = \rho \wedge \pi,$$

$$Mc \equiv \pi_{10} + \pi_{20} = +\sqrt{m_1^2 c^2 + \pi^2} + \sqrt{m_2^2 c^2 + \pi^2}. \tag{4.9}$$

Equations (4.8) are to be confronted with the corresponding expressions for the single free particle in the form (3.12). The construction of the remaining six variables of the scheme B in terms of the "internal" variables ρ, π *is essentially equivalent to the solution of a relativistic Hamilton—Jacobi problem for the internal motion.* This is done in the Appendix B. Finally, the whole scheme B results:

$P_1 = \mathfrak{P}_1 = P_x$	$P_2 = \mathfrak{P}_2 = P_y$	$P_3 = \mathfrak{P}_3 = P_z$	
$\mathfrak{Q}_1 = \mathfrak{Q}_1 = Q_x$	$Q_2 = \mathfrak{Q}_2 = Q_y$	$Q_3 = \mathfrak{Q}_3 = Q_z$	
$P_4 = \mathfrak{P}_4 = S_z$	$P_5 = \sqrt{\mathfrak{J}_1} = S$	$P_6 = \sqrt{\mathfrak{J}_2} = Mc$	(B_1)
$Q_4 = \mathfrak{Q}_4 = \arctan\dfrac{S_y}{S_x}$	$Q_5 = \arctan\dfrac{\pi_\theta\tan\theta - \rho\pi_\rho}{S + \rho\pi_\varphi\pi_\theta\tan\theta/S}$	$Q_6 = \dfrac{\rho\pi_\rho}{Mc}\dfrac{\pi_{10}\pi_{20}}{\pi_\rho^2 + S^2/\rho^2}$	

where S and Mc are given by Eqs. (4.9) and polar canonical coordinates have been introduced for the internal variables ρ, π. The expression of the scheme B in terms of the original variables $\mathbf{q}_1, \mathbf{q}_2, \mathbf{p}_1, \mathbf{p}_2$ is immediately obtained through the formulas (4.2), (4.3) and (4.5), (4.6).

We see from the scheme B_1 that none of the canonical invariants reduce to a constant and that the third and

fourth sets of the scheme B, in the sense of Ref. 1, are empty. The phase space contains invariants submanifolds of the form

$$P_5 \equiv S = \text{const}, \quad P_6 \equiv Mc = \text{const}, \tag{4.10}$$

and the realization corresponding to the two-particle system is *nonirreducible*. The variables Q_5 and Q_6 represent a simple relativistic generalization of the

corresponding variables for the Galilean two-particle system whose physical meaning has been discussed in Ref. 3, Sec. 7.

We want to discuss now with a greater detail the physical meaning of the variables $\mathbf{Q}$ and ρ. As it is well known, many definitions of center-of-mass have been proposed for a system of particles (mainly noninteracting) (see Refs. 11–13, 17, 18). A simple choice is provided by extending the nonrelativistic definition replacing the dynamical masses for the rest masses of the particles[11]

$$\mathbf{R} = \left(\frac{m_1}{\sqrt{1 - \dot{q}_1^2/c^2}} + \frac{m_2}{\sqrt{1 - \dot{q}_2^2/c^2}}\right)^{-1}\left(\frac{m_1}{\sqrt{1 - \dot{q}_1^2/c^2}}\mathbf{q}_1\right.$$
$$\left. + \frac{m_2}{\sqrt{1 - \dot{q}_2^2/c^2}}\mathbf{q}_2\right). \tag{4.11}$$

In our case of free particles this amounts to putting

$$\mathbf{R} = \frac{p_{10}\mathbf{q}_1 + p_{20}\mathbf{q}_2}{p_{10} + p_{20}} = -\frac{c^2}{H}\mathbf{K} \tag{4.12}$$

[see Eq. (3.36)]. This quantity changes in time according to the expected law

$$\mathbf{R}' = \mathbf{R} + \delta\tau\{\mathbf{R}, H\} = \mathbf{R} + \delta\tau\frac{c^2}{H}\mathbf{P}; \tag{4.13}$$

however, it is not invariantly defined: It does not satisfy Eq. (3.8). Another possibility, suggested by Fokker (see Refs. 11, 13), consists in defining first the above center-of-mass in the system $\mathbf{P} = 0$ and then in any other reference system by Lorentz transforming it. In our formalism this new quantity can be written

$$\mathbf{X} = \mathbf{Q} + \frac{\mathbf{S}\wedge\mathbf{P}}{M(Mc^2 + H)} = \mathbf{R} + \frac{\mathbf{S}\wedge\mathbf{P}}{MH}. \tag{4.14}$$

Actually, $\mathbf{X}$ reduces to $\mathbf{R}$ for $\mathbf{P} = 0$ and satisfies Eq. (3.8) [in this connection see Eq. (3.14)]. We have again

$$\mathbf{X}' = \mathbf{X} + \delta\tau\{\mathbf{X}, H\} = \mathbf{X} + \delta\tau\frac{c^2}{H}\mathbf{P}. \tag{4.15}$$

However, the Poisson brackets $\{R_i, R_j\}$ and, as already seen, the $\{X_i, X_j\}$'s are different from zero. Consequently, neither the components of $\mathbf{R}$ nor those of $\mathbf{X}$ can be used as canonical variables. A third possible definition of center-of-mass is provided by the canonical vector already introduced[11] [Eqs. (2.32), (2.43)]. This does not satisfy Eq. (3.8), however; it is simply related to $\mathbf{R}$ and $\mathbf{X}$, reduces to $\mathbf{R} \equiv \mathbf{X}$ for $\mathbf{P} = 0$, and its time evolution law is

$$\mathbf{Q}' = \mathbf{Q} + \delta\tau\{\mathbf{Q}, H\} = \mathbf{Q} + \delta\tau\frac{c^2}{H}\mathbf{P}. \tag{4.16}$$

We shall refer to $\mathbf{R}$ simply as to the center-to-mass, to $\mathbf{X}$ as to the covariant center-of-mass and to $\mathbf{Q}$ as to the canonical center-of-mass. Finally, notice the connection formulas

$$\mathbf{Q} = \mathbf{R} + \frac{c^2\mathbf{S}\wedge\mathbf{P}}{H(Mc^2 + H)} = \frac{H\mathbf{R} + Mc^2\mathbf{X}}{Mc^2 + H}, \tag{4.17}$$

which have been used in Eq. (4.14) implicitly. Before ending the discussion of the center-of-mass, let us observe that if the rest masses are used in its definition instead of the dynamical masses as in Eq. (4.11), the resulting center-of-mass is not in general at rest in the

system $\mathbf{P} = 0$. Pryce[11] stresses in addition that if the particles interact with one another this Newtonian center-of-mass does not move uniformly along a straight line in general. On the other hand, the definition (4.10) [and (4.13)] can be simply generalized to particles interacting through a field. It is worth stressing here that this generalization is not a minor problem in our present context. Actually, we will show elsewhere that *the identification of a uniformly moving center-of-mass in terms of the covariant position* coordinates of two particles interacting at a distance is an essential part in the choice of a dynamics.

Let us turn now to discuss the physical meaning of the variable ρ. We note first that in the system $\mathbf{P} = 0$ equation (4.6) becomes

$$\rho_{c.m.} = \mathbf{q}_{1_{c.m.}} - \mathbf{q}_{2_{c.m.}}. \tag{4.18}$$

On the other hand, let us consider the Lorentz transformation from the laboratory system to the system $\mathbf{P} = 0$:

$$\mathbf{x}_{c.m.}(t_{c.m.}) = \mathbf{x}(t) + \frac{\mathbf{P} \cdot \mathbf{x}(t) - t(Mc^2 + H)}{M(Mc^2 + H)}\mathbf{P},$$
$$t_{c.m.} = \frac{tH - \mathbf{P} \cdot \mathbf{x}(t)}{Mc^2}, \tag{4.19}$$

and the inverse transformation

$$\mathbf{x}(t) = \mathbf{x}_{c.m.}(t_{c.m.}) + \frac{\mathbf{P} \cdot \mathbf{x}_{c.m.}(t_{c.m.}) + t_{c.m.}(Mc^2 + H)}{M(Mc^2 + H)}\mathbf{P}$$
$$t = \frac{t_{c.m.}H + \mathbf{P} \cdot \mathbf{x}_{c.m.}(t_{c.m.})}{Mc^2}. \tag{4.20}$$

From Eq. (4.19), putting $t = 0$ and $\mathbf{x}_\tau(0) = \mathbf{q}_\tau$ ($\tau = 1, 2$), it follows

$$\mathbf{q}_{\tau, c.m.}(t^0_{\tau, c.m.}) = \mathbf{q}_\tau + \frac{\mathbf{P} \cdot \mathbf{q}_\tau}{M(Mc^2 + H)}\mathbf{P},$$
$$t^0_{\tau, c.m.} = -\mathbf{P} \cdot \mathbf{q}_\tau/Mc^2 \quad (\tau = 1, 2). \tag{4.21}$$

Then, using the relations

$$\mathbf{q}_{1, c.m.}(t) = \mathbf{q}_{1, c.m.} + \frac{c\pi}{\pi_{10}}t,$$
$$\mathbf{q}_{2, c.m.}(t) = \mathbf{q}_{2, c.m.} - \frac{c\pi}{\pi_{20}}t, \tag{4.22}$$

we obtain

$$\mathbf{q}_{1, c.m.} = \mathbf{q}_1 + \frac{\mathbf{P} \cdot \mathbf{q}_1}{M(Mc^2 + H)}\mathbf{P} + \frac{\mathbf{P} \cdot \mathbf{q}_1}{Mc}\frac{\pi}{\pi_{10}},$$
$$\mathbf{q}_{2, c.m.} = \mathbf{q}_2 + \frac{\mathbf{P} \cdot \mathbf{q}_2}{M(Mc^2 + H)}\mathbf{P} - \frac{\mathbf{P} \cdot \mathbf{q}_2}{Mc}\frac{\pi}{\pi_{20}}, \tag{4.23}$$

and

$$\mathbf{q}_{1, c.m.} - \mathbf{q}_{2, c.m.} = \mathbf{q}_1 - \mathbf{q}_2$$
$$+ \frac{(\mathbf{q}_1 - \mathbf{q}_2) \cdot \mathbf{P}}{M(Mc^2 + H)}\mathbf{P} + \left[\left(\frac{\mathbf{q}_1}{\pi_{10}} + \frac{\mathbf{q}_2}{\pi_{20}}\right) \cdot \mathbf{P}\right]\frac{\pi}{Mc}. \tag{4.24}$$

Finally, comparing with Eq. (4.6), we get

$$\rho = \mathbf{q}_{1, c.m.} - \mathbf{q}_{2, c.m.} - \left[\frac{\mathbf{q}_1 \cdot \mathbf{P}}{\pi_{10}\pi_{20}}\left(\frac{P_0\pi_{10} + \pi \cdot \mathbf{P}}{P_0}\right)\right.$$

$$+ \frac{q_2 \cdot P}{\pi_{10}\pi_{20}} \left(\frac{P_0\pi_{20} - \pi \cdot P}{P_0} \right) \Big] \frac{\pi}{Mc}$$

$$= q_{1,\,c.m.} - q_{2,\,c.m.} - \frac{1}{\pi_{10}\pi_{20}} \Big[\Big(\frac{p_{10}q_1 + p_{20}q_2}{P_0} \Big) \cdot P \Big] \pi$$

$$= q_{1,\,c.m.} - q_{2,\,c.m.} - \frac{1}{\pi_{10}\pi_{20}} (X \cdot P)\pi. \qquad (4.25)$$

Again, from Eq. (4.19), replacing $X(t)$ for $x(t)$ and putting $t = 0$, we can write

$$X_{c.m.}(t^0_{c.m.}) = X + \frac{X \cdot P}{M(Mc^2 + H)} P,$$

$$t^0_{c.m.} = - X \cdot P/Mc^2. \qquad (4.26)$$

Then, using Eqs. (4.22) and (4.26), [and the fact that $\{\pi, H\} = 0$ and $P \cdot \{K, \pi_j\} = 0$, see Eq. (4.38)], Eq. (4.25) can be written

$$\rho = q_{1,\,c.m.}(t^0_{c.m.}) - q_{2,\,c.m.}(t^0_{c.m.}) \qquad (4.27)$$

or, since the instant $t = 0$ has no privileged role,

$$\rho(t) = q_{1,\,c.m.}(t_{c.m.}) - q_{2,\,c.m.}(t_{c.m.}), \qquad (4.28)$$

with

$$X_{c.m.}(t_{c.m.}) = X(t) + \frac{P \cdot X(t) - t(Mc^2 + H)}{M(Mc^2 + H)} P,$$

$$t_{c.m.} = [tH - P \cdot X(t)]/Mc^2. \qquad (4.29)$$

Equations (4.27) or (4.28) provide the desired physical meaning of the internal variable ρ; that is, ρ *equals the difference between the position vectors of the particles evaluated in the system* $P = 0$ *at the time signed by the clock tied to the center-of-mass* X *when the laboratory clock points to the considered value* t.

A consequence of Eq. (4.28) is that the values of ρ in two reference frames connected by a Lorentz transformation in the direction of P are related by

$$\rho'(t') = \rho(t) \qquad (4.30)$$

with

$$X'(t') = X(t) + \left(\frac{1}{\sqrt{1 - \beta^2}} - 1 \right) \frac{X(t) \cdot P}{P} \frac{P}{P} - \frac{ct\beta}{\sqrt{1 - \beta^2}} \frac{P}{P},$$

$$t' = \frac{t - (\beta/c)X(t) \cdot P/P}{\sqrt{1 - \beta^2}}. \qquad (4.31)$$

In particular, for an infinitesimal transformation, putting $t = 0$, we can write

$$\rho'(\delta t) = \rho \quad \text{or} \quad \rho' = \rho(-\delta t), \qquad (4.32)$$

$$X'(\delta t) = X, \quad \delta t = - (\delta v/c^2)(X \cdot P/P). \qquad (4.33)$$

An equation corresponding to Eq. (4.32) in terms of canonical generators is

$$P \cdot \{K, \rho_j\} = (X \cdot P/c^2)\{\rho_j, H\}. \qquad (4.34)$$

This relation can be checked directly using Eqs. (4.8) and the first of Eqs. (4.9). Actually, from

$$\{K_i, \rho_j\} = \frac{\rho_i P_j - \delta_{ij}\rho \cdot P}{Mc^2 + H} + \frac{c^2}{H} \left(MQ_i + \frac{(S \wedge P)_i}{Mc^2 + H} \right) \frac{\partial M}{\partial \pi_j}$$

$$(4.35)$$

it follows that

$$P \cdot \{K, \rho_j\} = \frac{Mc^2}{H} Q \cdot P \frac{\partial M}{\partial \pi_j} = \frac{X \cdot P}{c^2} \{\rho_j, H\}. \qquad (4.36)$$

In the same way we have formally

$$\{K_i, \pi_j\} = \frac{\pi_i P_j - \delta_{ij}\pi \cdot P}{Mc^2 + H} - \frac{c^2}{H} \left(MQ_i + \frac{(S \wedge P)_i}{Mc^2 + H} \right) \frac{\partial M}{\partial \rho_j}$$

$$(4.37)$$

and

$$P \cdot \{K, \pi_j\} = (X \cdot P/c^2)\{\pi_j, H\}. \qquad (4.38)$$

It is important to stress the fact that the results expressed in the Eqs. (4.36), (4.38) hold true indepently of the particular form of the function M. Equations (4.34)—(4.38) are to be confronted with Eq. (3.8).

To conclude the present subsection, we note that, according to the results of Secs. 2 and 3f, the improper transformations are realized in the following way:

canonical space reflection
$$\begin{aligned} q_\tau &\to - q_\tau, & p_\tau &\to - p_\tau, & (\tau = 1, 2), \\ Q &\to - Q, & P &\to - P, \\ \rho &\to - \rho, & \pi &\to - \pi; \end{aligned} \qquad (4.39)$$

anticanonical time reflection
$$\begin{aligned} q_\tau &\to q_\tau, & p_\tau &\to - p_\tau, & (\tau = 1, 2), \\ Q &\to Q, & P &\to - P, \\ \rho &\to \rho, & \pi &\to - \pi; \end{aligned} \qquad (4.40)$$

anticanonical space—time reflection
$$\begin{aligned} q_\tau &\to - q_\tau, & p_\tau &\to p_\tau, & (\tau = 1, 2), \\ Q &\to - Q, & P &\to P, \\ \rho &\to - \rho, & \pi &\to \pi. \end{aligned} \qquad (4.41)$$

B. Interacting particles

As is well known, the introduction of an interaction in a two-particle classical relativistic system described within the Hamiltonian framework poses serious problems.[6,7,19-23] In particular, the zero-interaction theorem established by Currie, Jordan, and Sudarshan[7] prevents the possibility of modifying the expression of the canonical generators in any nontrivial way while leaving the conditions (3.8) simultaneously satisfied by the canonical vectors q_1 and q_2. This means that a relativistic theory of two interacting particles in which the coordinates x_1 and x_2 play the role of canonical variables cannot exist. Bakamjian and Thomas[5] (see also Foldy[6]), however, have proposed a formal procedure for introducing a direct interaction between the particles which appears quite natural in the context of our formalism. This procedure consists in preserving the structure of the canonical generators as given by Eqs. (4.8) and the first Eq. (4.9) and in modifying the second Eq. (4.9) by means of the *Ansatz*

$$Mc^2 = c(\pi_{10} + \pi_{20}) + U(\rho, \pi). \qquad (4.42)$$

Provided that $U(\rho, \pi)$ is a rotationally invariant function, i.e., a function of the variables $\rho = |\rho|$, $\pi = |\pi|$, $\sigma = \rho \cdot \pi$ only, the center-of-mass energy Mc^2 satisfies

$$\{S_i, Mc^2\} = 0 \qquad (4.43)$$

and the fundamental Poisson bracket relations (2.4) are preserved.

The function $U(\rho, \pi)$ plays here the role of the classical potential in the nonrelativistic theory and Eq. (4.42)

goes over into the corresponding expression for the Galilei group in the limit $c \to \infty$. If one assumes that the vector ρ shares a simple relation with the relative coordinate of the particles in the center-of-mass frame also for $U(\rho, \pi) \neq 0$, as it is strongly suggested by Eq. (4.36), one is lead to assume that $U(\rho, \pi)$ vanishes rapidly enough for $\rho \to \infty$. Under this condition, the quantities q_1, q_2 as defined by Eqs. (4.7) in terms of the basic variables Q, P, ρ, π will satisfy Eqs. (3.8) for large ρ values and will therefore be interpretable as relativistic coordinates of the particles at least in the asymptotic region. This should be sufficient, e.g., for a description of scattering processes within our canonical realization.[23,24]

We stress, however, that in order to achieve a satisfactory and complete physical interpretation of the formal prescription (4.42) it is necessary to construct dynamical variables $\mathbf{x}_1(Q, P, \rho, \pi)$ and $\mathbf{x}_2(Q, P, \rho, \pi)$ (obviously noncanonical!) which can be interpreted as physical position vectors of the interacting particles throughout the whole phase space because of their proper relativistic transformation properties. This problem, which is strictly connected with the formulation of an action-at-a-distance manifestly relativistic theory,[20-22,25] will be discussed and solved in a separate paper.

5. "RIGID" BODIES

The introduction of a direct interaction according to the line of the preceding section enables us to generalize the concept of rigid body to the relativistic case in a natural way by considering this system as a limiting case of a system of particles with suitable interactions. We shall confine the discussion to only two very simple examples, the second of which having a purely formal character.

A. Relativistic linear rotator

We consider a system of two interacting particles for which the potential is a function of ρ only:

$$Mc = \sqrt{m_1^2 c^2 + \pi^2} + \sqrt{m_2^2 c^2 + \pi^2} + \frac{1}{c} U(\rho) \quad \text{[see Eq. (4.42)].}$$
(5.1)

Introducing polar canonical coordinates for the internal variables $(\varphi, \theta, \rho, \pi_\varphi, \pi_\theta, \pi_\rho)$, we can write

$$\pi^2 = \pi_\rho^2 + S^2/\rho^2,$$
(5.2)

where

$$S_x = -\sin\varphi\,\pi_\theta - \cos\varphi \cot\theta\,\pi_\varphi,$$
$$S_y = \cos\varphi\,\pi_\theta - \sin\varphi \cot\theta\,\pi_\varphi,$$
(5.3)
$$S_z = \pi_\varphi.$$

A system of two heavy particles rigidly connected can be seen as a limit case of the above system with a potential $U(\rho)$ which constrains the two particles to a fixed distance ρ_0 in the center of mass system. Then we are led to set

$$\rho \to \rho_0, \quad \pi_\rho \to 0, \quad U(\rho_0) \to 0$$
(5.4)

in Eq. (5.1), with the consequence that

$$Mc = \sqrt{m_1^2 c^2 + S^2/\rho_0^2} + \sqrt{m_2^2 c^2 + S^2/\rho_0^2}.$$
(5.5)

In particular, *for the symmetric case* $m_1 = m_2 = m_0$, we have

$$Mc = \sqrt{m^2 c^2 + m S^2/I},$$
(5.6)

where we have put

$$m = 2m_0, \quad I = (m_0/2)\rho_0^2 = \mu\rho_0^2.$$
(5.7)

Then, one is lead to the Hamiltonian

$$H = c\sqrt{M^2 c^2 + \mathbf{P}^2} = c\sqrt{m^2 c^2 + m S^2/I + \mathbf{P}^2}.$$
(5.8)

Notice the close formal analogy between this system and the nonrelativistic rigid rod (nonrelativistic linear rotator) discussed in Refs. 2, 3. Actually, this last system can be obtained as a limit case for small P/mc and $S^2/mc^2 I$.

As Eq. (5.5) itself shows, however, the system defined by the Hamiltonian (5.8) cannot be considered as the most general relativistic rigid rod. As a matter of fact, it can at most correspond to a particular kind of rod in which the linear mass distribution satisfies special symmetry conditions.

The scheme B for the system (5.8) is identical to the scheme B for the nonrelativistic rigid rod apart from the structure of the fixed invariant. Besides the variables of the first set (see Ref. 1) which have a common form, there are variables in the second and third set according to

$P_5 \equiv \sqrt{\mathfrak{I}_1} = S$ $\qquad$ $\mathfrak{I}_2' \equiv M^2 c^2 - \dfrac{m S^2}{I} = m^2 c^2$
$Q_5 = \arctan \dfrac{\pi_\theta \tan\theta}{S}$

(B$_2$)

The time evolution of the variable Q_5 is given by

$$Q_5' = Q_5 + \delta\tau\{Q_5, H\} = Q_5 + \frac{1}{2I}\frac{mc^2}{H}\,\delta\tau,$$
(5.9)

so that, since H is a constant of the motion,

$$Q_5(t) = Q_5(0) + \frac{1}{2I}\frac{mc^2}{H}\,t.$$
(5.10)

Recalling the discussion given in the nonrelativistic case, and considering the physical meaning of the vector $\rho(t)$ in the present case, we realize that Q_5 is just the angle between the "rod" and the half-plane defined by the vector S and the z axis, in the center of mass system at the time t of the laboratory system. For $\mathbf{P} = 0$, Eq. (5.10) reduces to

$$Q_{5,\,\text{c.m.}}(t) = Q_{5,\,\text{c.m.}}(0) + \frac{1}{2I}\frac{m}{M}\,t.$$
(5.11)

The factor $Mc^2/H = \sqrt{1 - c^2\mathbf{P}^2/H^2}$ by which the time derivatives of Q_5 and $Q_{5,\,\text{c.m.}}$ differ expresses the relativistic time dilation. This model appears as the most simple conceivable example of a clock treated as a specifical dynamical system.

Notice finally that Eq. (5.10) reduces to the corresponding nonrelativistic one [see Ref. 3, Eq. (79)] for both P^2 and S^2 small. If P^2 is small but S^2 is not negligible, the angular speed of the "rod" in the relativistic case is smaller than in the nonrelativistic one.

B. Other rigid systems

The construction of more complex rigid systems by means of the same limiting procedure considered above would require complicated considerations involving a number of particles greater than two. We shall not do this explicitly. We observe, however, that we can give very simple formal generalizations of the Eq. (5.8) to systems with six degrees of freedom. Precisely, the expression

$$H = c\left[m^2c^2 + m\left(\frac{\Sigma_\xi^2}{2I_\xi} + \frac{\Sigma_\eta^2}{2I_\eta} + \frac{\Sigma_\zeta^2}{2I_\zeta}\right) + \mathbf{P}^2\right]^{1/2} \qquad (5.12)$$

could be interpreted as Hamiltonian for an asymmetrical top (rotator); and the two particular cases

$$H = c\left[m^2c^2 + \frac{m}{I}S^2 + m\left(\frac{1}{I_\zeta} - \frac{1}{I}\right)\Sigma_\zeta^2 + \mathbf{P}^2\right]^{1/2} \qquad (5.13)$$

and

$$H = c\left(m^2c^2 + \frac{m}{I}S^2 + \mathbf{P}^2\right)^{1/2} \qquad (5.14)$$

as Hamiltonians for the symmetrical and the spherical top, respectively (for the corresponding nonrelativistic cases see Ref. 3). [$\Sigma_\xi, \Sigma_\eta, \Sigma_\zeta$ denote the components of the intrinsic angular momentum referred to the body system.] However, as indicated by the case of the linear rotator, Eqs. (5.12)—(5.14) define relativistic objects which must be considered very particular examples of the systems mentioned above.

In the case of the spherical top, the relevant part of the scheme B is given by

$$P_5 = S \qquad \mathfrak{I}_2' = \mathfrak{I}_2 - \frac{m}{I}S^2 = m^2c^2 \quad P_6 = \Sigma_\zeta$$

$$Q_5 = \arctan\frac{\pi_\theta \tan\theta}{S - \pi_\varphi \pi_\psi/S\cos\theta} \qquad Q_6 = \arctan\frac{\Sigma_\eta}{\Sigma_\xi} \qquad \text{(B}_3\text{)}$$

In the case of the symmetrical top, instead, a possible choice is

$$P_5 = S \qquad P_6 = \frac{\mathfrak{I}_2}{m} - \frac{S^2}{I} = mc^2 + \frac{I - I_\zeta}{II_\zeta}\Sigma_\zeta^2$$

$$Q_5 = \arctan\frac{\pi_\theta \tan\theta}{S - \pi_\varphi \pi_\psi/S\cos\theta} \qquad Q_6 = \frac{II_\zeta}{I - I_\zeta}\frac{1}{2\Sigma_\zeta}\arctan\frac{\Sigma_\eta}{\Sigma_\xi} \qquad \text{(B}_4\text{)}$$

The expressions of the variables Q_5 and Q_6 in Eqs. (B$_3$) and (B$_4$) are identical to those of the corresponding nonrelativistic systems. Only the expressions of $\mathfrak{I}_2'$ in the scheme (B$_3$) and of P_6 in the scheme (B$_4$) are suitably modified. Also the physical meaning of the variables Q_5 and Q_6 is the same as in the nonrelativistic case. In particular, in the case of the symmetrical top, Q_5 and Q_6 are the precession angle and the proper rotation angle, respectively.

The time evolution of these variables is given by

$$Q_5(t) = Q_5(0) + \frac{1}{2I}\frac{mc^2}{H}t, \quad Q_6(t) = Q_6(0), \qquad (5.15)$$

in the case of the spherical top, and by

$$Q_5(t) = Q_5(0) + \frac{1}{2I}\frac{mc^2}{H}t, \quad Q_6(t) = Q_6(0) + \frac{1}{2}\frac{mc^2}{H}t, \qquad (5.16)$$

in the case of the symmetrical top, where we note again the relativistic time dilation factor.

APPENDIX A: CONSTRUCTION OF THE SCHEME A FOR THE HOMOGENEOUS LORENTZ GROUP

We have

$$\{J_i, J_j\} = \epsilon_{ijk}J_k,$$
$$\{J_i, K_j\} = \epsilon_{ijk}K_k, \qquad (A1)$$
$$\{K_i, K_j\} = -\frac{1}{c^2}\epsilon_{ijk}J_k.$$

By virtue of the results of Ref. 2, Sec. 2, for the rotation group, we can put

$$\mathfrak{P}_1 = J_z, \quad \mathfrak{Q}_1 = \arctan(J_y/J_x). \qquad (A2)$$

The system

$$\{\mathfrak{P}_1, \Phi(\mathbf{J}, \mathbf{K})\} = 0, \quad \{\mathfrak{Q}_1, \Phi(\mathbf{J}, \mathbf{K})\} = 0, \qquad (A3)$$

is complete and admits four independent solutions. It is evident that any scalar (pseudoscalar) built up from the generators is a solution of the system. For instance

$$J = |\mathbf{J}|, \quad K = |\mathbf{K}|, \quad \mathbf{J} \cdot \mathbf{K}. \qquad (A4)$$

A further independent solution remains to be determined which, obviously, cannot be a scalar. Since J_z, K_z and the scalars (A4) are five independent variables which have zero Poisson brackets with $\mathfrak{P}_1$, in order to find the remaining solution it is sufficient to look for a function Φ_1 of such variables having zero Poisson bracket with $\mathfrak{Q}_1$. We must have

$$\{\mathfrak{Q}_1, \Phi_1\} \equiv \frac{\mathbf{J} \cdot \mathbf{K} - J_z K_z}{J^2 - J_z^2}\frac{\partial \Phi_1}{\partial K_z} + \frac{\partial \Phi_1}{\partial J_z} = 0. \qquad (A5)$$

Thus Φ_1 can be assumed to be a function of only J_z, K_z, J and $\mathbf{J} \cdot \mathbf{K}$. Solving Eq. (A5) by the method of the characteristics,[26] we find

$$\Phi_1 = \frac{K_z - (\mathbf{J} \cdot \mathbf{K}/J^2)J_z}{(J^2 - J_z^2)^{1/2}} = \frac{[(\mathbf{J} \wedge \mathbf{K}) \wedge \mathbf{J}]_z}{J^2(J^2 - J_z^2)^{1/2}}. \qquad (A6)$$

Now we can choose $\mathfrak{P}_2 = J$ and look for a function $\mathfrak{Q}_2$ of J, Φ_1 and the remaining variables such that

$$\{\mathfrak{Q}_2, J\} = 1 \qquad (A7)$$

holds true. We have

$$\{\Phi_1, J\} = \frac{[J \wedge K]_z}{J(J^2 - J_z^2)^{1/2}},$$ (A8)

while, obviously, all of the scalars (A4) have zero Poisson brackets with J. The expression (A8) must be solution of the system (A3) by virtue of the Jacobi identity and therefore it could be reexpressed in terms of Φ_1 and the variables (A4). It is more convenient to denote the expression (A8) by Φ_2 and to take advantage of the fact that

$$\{\Phi_1, J\} = \Phi_2, \quad \{\Phi_2, J\} = -\Phi_1.$$ (A9)

This means that Ω_2 can be assumed to be a function of Φ_1, Φ_2 and J, and Eq. (A7) becomes

$$\Phi_2 \frac{\partial \Omega_2}{\partial \Phi_1} - \Phi_1 \frac{\partial \Omega_2}{\partial \Phi_2} = 1.$$ (A10)

Since a solution of the associated homogeneous equation is

$$\Psi = \Phi_1^2 + \Phi_2^2,$$ (A11)

using Φ_2 and Ψ as independent variables in place of Φ_1, Φ_2 we obtain

$$\frac{\partial \Omega_2}{\partial \Phi_2} = -\frac{1}{(\Psi - \Phi_2^2)^{1/2}},$$ (A12)

i.e.,

$$\Omega_2 = \arctan\left(-\frac{\Phi_2}{\Phi_1}\right) = \arctan\frac{J[J \wedge K]_z}{(J \cdot K)J_z - J^2 K_z}.$$ (A13)

We are left to find two independent functions of the variables (A4) and Φ_1 which have zero Poisson brackets with $\mathfrak{P}_2$ and Ω_2. Now $J \cdot K$ is easily seen to have this property, while

$$\{\Omega_2, K^2\} = 2J.$$ (A14)

Therefore, owing to Eq. (A7), it follows that the expression $J^2 - K^2$ provides a possible choice for the remaining independent function with the desired property. In conclusion, we have obtained the scheme A_6.

APPENDIX B: DETERMINATION OF THE VARIABLES Q_5 AND Q_6 OF THE TYPICAL FORM FOR THE TWO-FREE-PARTICLES SYSTEM

The variable Q_5

According to the constructive procedure given in Ref. 1, Theorem 2, we have to find a function $Q_5(\rho, \pi)$ such that

$$\{Q_4, Q_5\} = \{P_4, Q_5\} = 0,$$ (B1)

$$\{Q_5, P_5\} = 1,$$ (B2)

$$\{Q_5, P_6\} = 0,$$ (B3)

hold true. Introducing a polar representation for the "internal" vectors ρ, π, we have[2,3]

$$P_4 = S_z = \pi_\varphi,$$

$$Q_4 = \arctan\frac{S_y}{S_x} = \frac{\cos\varphi\pi_\theta - \cot\theta\sin\varphi\pi_\varphi}{-\sin\varphi\pi_\theta - \cot\theta\cos\varphi\pi_\varphi},$$ (B4)

$$P_5 \equiv S = \sqrt{\pi_\theta^2 + \pi_\varphi^2/\sin^2\theta},$$

$$P_6 \equiv Mc = \pi_{10} + \pi_{20}$$

$$= \sqrt{m_1^2 c^2 + \pi_\rho^2 + S^2/\rho^2} + \sqrt{m_2^2 c^2 + \pi_\rho^2 + S^2/\rho^2}.$$

From the results of Ref. 2, Sec. 3 and Appendix A, it follows that a solution Q_5^0 of Eqs. (B1) and (B2) is

$$Q_5^0 = \arctan(\pi_\theta \tan\theta/S).$$ (B5)

Then, putting

$$Q_5 = Q_5^0 + \Phi(\rho, \pi_\rho, S),$$ (B6)

we look for a function Φ such that Eq. (B3) is satisfied. We obtain the equation

$$\pi_\rho \frac{\partial \Phi}{\partial \rho} + \frac{S^2}{\rho^3} \frac{\partial \Phi}{\partial \pi_\rho} + \frac{S}{\rho^2} = 0.$$ (B7)

Since a solution of the associated homogeneous equation is

$$\Psi(\rho, \pi_\rho, S) = \pi_\rho^2 + S^2/\rho^2,$$ (B8)

using ρ and Ψ as independent variables in place of ρ and π_ρ, we obtain from (B7)

$$\frac{\partial \Phi}{\partial \rho} = -\frac{S}{\rho^2(\Psi - S^2/\rho^2)^{1/2}},$$ (B9)

a solution of which is

$$\Phi = -\arctan(\rho\pi_\rho/S).$$ (B10)

In conclusion, from Eqs. (B5), (B6), (B10), we have

$$Q_5 = \arctan\frac{\pi_\theta \tan\theta}{S} - \arctan\frac{\rho\pi_\rho}{S}$$

$$= \arctan\frac{\pi_\theta \tan\theta - \rho\pi_\rho}{S + \rho\pi_\rho\pi_\theta \tan\theta/S}.$$ (B11)

The variable Q_6

We must find a function Q_6 such that

$$\{Q_4, Q_6\} = \{P_4, Q_6\} = \{P_5, Q_6\} = 0,$$ (B12)

$$\{Q_5, Q_6\} = 0,$$ (B13)

$$\{Q_6, P_6\} = 1$$ (B14)

hold true. Equations (B12) are automatically satisfied if we consider a function $\Psi = \Psi(\rho, \pi_\rho, S)$. Equation (B13) then gives

$$\frac{\partial \Psi}{\partial S} + \frac{\rho S}{S^2 + \rho^2\pi_\rho^2} \frac{\partial \Psi}{\partial \rho} - \frac{\pi_\rho S}{S^2 + \rho^2\pi_\rho^2} \frac{\partial \Psi}{\partial \pi_\rho} = 0,$$ (B15)

which can be solved by the method of the characteristics.[26] Precisely, the general solution will be an arbitrary function of the independent integrals of the system

$$\frac{dS}{d\rho} = \frac{S^2 + \rho^2\pi_\rho^2}{\rho S},$$

$$\frac{d\pi_\rho}{d\rho} = -\frac{\pi_\rho}{\rho}.$$ (B16)

Now, two integrals of (B16) are

$$k_1(\rho, \pi_\rho, S) = \pi_\rho^2 + S^2/\rho^2,$$

$$k_2(\rho, \pi_\rho, S) = \rho\pi_\rho.$$ (B17)

Therefore, the general solution of Eqs. (B12), (B13) is

$$\Psi = \Psi(k_1, k_2).$$ (B18)

We are left with Eq. (B14). Since

$$P_6 = \sqrt{m_1^2 c^2 + k_1} + \sqrt{m_2^2 c^2 + k_2},$$

$$\{k_2, k_1\} = k_1,$$
(B19)

$$\{k_2, P_6\} = \frac{P_6}{\pi_{10}\pi_{20}} k_1,$$

it follows

$$\frac{\partial \Psi}{\partial k_2} = \frac{\pi_{10}\pi_{20}}{P_6} \frac{1}{k_1},$$
(B20)

and finally

$$Q_6 = \frac{\pi_{10}\pi_{20}}{P_6} \frac{\rho\pi_\rho}{\pi_\rho^2 + S^2/\rho^2}.$$
(B21)

The determination of the variables $P_4, P_5, P_6, Q_4, Q_5, Q_6$ in terms of ρ, π is essentially equivalent to the solution of the time-independent Hamilton—Jacobi problem for the "internal" motion with Hamiltonian $P_6 \equiv Mc = \pi_{10} + \pi_{20}$.

REFLECTION PROPERTIES

1. $\Pi = m^2 c^2 > 0$. Additional invariants of the proper group: $H/|H|$

(a) *Regular realizations*: $s > 0$

Transformation	Canonical realization	Anticanonical realization
Space reflection	$p \to -p$, $q \to -q$ $p_x \to p_x$, $\chi \to \chi$ $p \to -p$, $\vec{q} \to -\vec{q}$ $p_\tau \to -p_\tau$, $\tau \to \pi - \tau$	No
Time reflection	No	$p \to -p$, $q \to q$ $p_x \to -p_x$, $\chi \to \chi + \pi$ $p \to -p$, $\vec{q} \to \vec{q}$ $p_\tau \to p_\tau$, $\tau \to -\tau$
Space—time reflection	No	$p \to p$, $q \to -q$ $p_x \to -p_x$, $\chi \to \chi + \pi$ $p \to p$, $\vec{q} \to -\vec{q}$ $p_\tau \to p_\tau$, $\tau \to \tau + \pi$

(b) *Singular realizations*: $s = 0$

The internal variables χ, p_χ (τ, p_τ) are missing. The remaining transformations are the same as above.

2. $\Pi = \delta^2 c^2 < 0$. Additional invariants of the proper group: $W_T/|W_T|$ if $w \le 0$ with $W_T \ne 0$

(a) *Regular realizations*: $w \ne 0$ or $w = 0$ with $W_T \ne 0$

Transformation	Canonical realization	Anticanonical realization
	only if $w > 0$	only if $w > 0$
Space reflection	$p \to -p$, $\vec{q} \to -\vec{q}$ $p_\tau \to -p_\tau$, $\tau \to \pi - \tau$ $p_\xi \to p_\xi$, $\xi \to \xi$ $p_\varphi \to \pi - p_\varphi$, $\varphi \to -\varphi$ $p_\psi \to \pi + p_\psi$, $\psi \to \psi$	$p \to p$, $\vec{q} \to -\vec{q}$ $p_\tau \to -p_\tau$, $\tau \to \tau + \pi$ $p_\xi \to -p_\xi$, $\xi \to \xi$ $p_\varphi \to p_\varphi$, $\varphi \to -\varphi$ $p_\psi \to p_\psi$, $\psi \to -\psi$
Time reflection	$p \to p$, $\vec{q} \to \vec{q}$ $p_\tau \to p_\tau$, $\tau \to \tau$ $p_\xi \to -p_\xi$, $\xi \to -\xi$ $p_\varphi \to p_\varphi$, $\varphi \to \varphi$ $p_\psi \to p_\psi$, $\psi \to \psi$	$p \to -p$, $\vec{q} \to \vec{q}$ $p_\tau \to p_\tau$, $\tau \to -\tau$ $p_\xi \to p_\xi$, $\xi \to -\xi$ $p_\varphi \to \pi - p_\varphi$, $\varphi \to \varphi$ $p_\psi \to p_\psi + \pi$, $\psi \to -\psi$
	only if $w > 0$	only if $w > 0$
Space—time reflection	$p \to -p$, $\vec{q} \to -\vec{q}$ $p_\tau \to -p_\tau$, $\tau \to \pi - \tau$ $p_\xi \to -p_\xi$, $\xi \to -\xi$ $p_\varphi \to \pi - p_\varphi$, $\varphi \to -\varphi$ $p_\psi \to p_\psi + \pi$, $\psi \to \psi$	$p \to p$, $\vec{q} \to -\vec{q}$ $p_\tau \to -p_\tau$, $\tau \to \tau + \pi$ $p_\xi \to p_\xi$, $\xi \to -\xi$ $p_\varphi \to p_\varphi$, $\varphi \to -\varphi$ $p_\psi \to p_\psi$, $\psi \to -\psi$

(b) *Singular realizations*: $w = 0$, $W_T = 0$

The variables τ, p_τ are missing. The remaining transformations are the same as above and are all realizable.

3. $\Pi = 0$. Additional invariants of the proper group: $H/|H|$, Γ if $w = 0$

(a) *Regular realizations*: $w \ne 0$

Transformation	Canonical realization	Anticanonical realization
Space reflection	$p \to -p$, $\vec{q} \to -\vec{q}$ $p_\tau \to -p_\tau$, $\tau \to \pi - \tau$	No
Time reflection	No	$p \to -p$, $\vec{q} \to \vec{q}$ $p_\tau \to p_\tau$, $\tau \to -\tau$
Space—time reflection	No	$p \to p$, $\vec{q} \to -\vec{q}$ $p_\tau \to -p_\tau$, $\tau \to \tau + \pi$

(b) *Singular (helicity) realization*: $w = 0$, $\Gamma = \gamma \ne 0$

Transformation	Canonical realization	Anticanonical realization
Space reflection	No	No
Time reflection	No	$p \to -p$, $\vec{q} \to \vec{\Omega}$
Space—time reflection	No	No

(c) *Singular (scalar) realizations*: $w = 0$, $\Gamma = \gamma = 0$

The same possibilities exist as in the regular realization with the variables τ, p_τ missing.

4. $\Pi = 0$, $T = 0$. Additional invariants of the proper group: none

Transformation	Canonical realization	Anticanonical realization
	only if $j_2 = 0$	only if $j_2 = 0$
Space reflection	$p_\alpha \to p_\alpha$, $\alpha \to \alpha$ $p_\beta \to p_\beta$, $\beta \to \beta + \pi$	$p_\alpha \to -p_\alpha$, $\alpha \to \alpha + \pi$ $p_\beta \to p_\beta$, $\beta \to \pi - \beta$
	only if $j_2 = 0$	only if $j_2 = 0$
Time reflection	$p_\alpha \to p_\alpha$, $\alpha \to \alpha$ $p_\beta \to p_\beta$, $\beta \to \beta + \pi$	$p_\alpha \to -p_\alpha$, $\alpha \to \alpha + \pi$ $p_\beta \to p_\beta$, $\beta \to \pi - \beta$
Space—time reflection	$p_\alpha \to p_\alpha$, $\alpha \to \alpha$ $p_\beta \to p_\beta$, $\beta \to \beta$	$p_\alpha \to -p_\alpha$, $\alpha \to \alpha + \pi$ $p_\beta \to p_\beta$, $\beta \to -\beta$

[1] M. Pauri and G.M. Prosperi, J. Math. Phys. **7**, 366 (1966).
[2] M. Pauri and G.M. Prosperi, J. Math. Phys. **8**, 2256 (1967).
[3] M. Pauri and G.M. Prosperi, J. Math. Phys. **9**, 1146 (1968).
[4] In Refs. 1, 2, 3 the term "scheme A" means a certain set of independent functions of the canonical generators which satisfy canonical Poisson bracket relations. In a more precise mathematical language, our "scheme A" essentially represents a chart of canonical (symplectic) coordinates for the *dual space* of the Lie algebra of the considered group: This chart is adapted to the fibered structure defined in that space by the action of the co-adjoint representation [J.M. Souriau, *Structure des systèmes dynamiques* (Dunod, Paris, 1970); B. Kostant, *Lecture Notes in Mathematics*, edited by C.T. Taam (Springer, New York, 1970), Vol. 170]. It should be stressed that our characterization is only local in general.

1520 J. Math. Phys., Vol. 16, No. 7, July 1975

M. Pauri and G.M. Prosperi 1520

[5]B. Bakamjian and L.H. Thomas, Phys. Rev. **92**, 1300 (1953).

[6]L.L. Foldy, Phys. Rev. **122**, 275 (1961); see also P.A.M. Dirac, Rev. Mod. Phys. **21**, 392 (1949); L.H. Thomas, Rev. Mod. Phys. **17**, 182 (1945); Phys. Rev. **85**, 868 (1952); B. Bakamjian, Phys. Rev. **121**, 1849 (1961); T.F. Jordan, A.J. MacFarlane, and E.C.G. Sudarshan, Phys. Rev. **133**, B487 (1964).

[7]D.G. Currie, T.F. Jordan, and E.C.G. Sudarshan, Rev. Mod. Phys. **35**, 350 (1963); D.G. Currie, J. Math. Phys. **4**, 1470 (1963); J.T. Cannon and T.F. Jordan, J. Math. Phys. **5**, 299 (1964); H. Leutwyler, Nuovo Cimento **37**, 560 (1965); D.G. Currie and T.F. Jordan, "Interactions in relativistic classical particle mechanics," *Lectures at the Theoretical Physics Institute* (University of Colorado, Boulder, 1967).

[8]V. Bargmann, Ann. Math. **59**, 1 (1954); M. Hamermesh, *Group Theory* (Addison-Wesley, Reading, Mass., 1962).

[9]Note the relation between the vector Ω used in our derivation and the well-known Pauli—Lubansky 4-vector $\omega^\mu \equiv [\omega^0, \omega] : \omega^0 = \Omega \cdot \mathbf{T}$, $\omega = (H/c)\Omega$.

[10]It is worthwhile to note that, in spite of the fact that some variables become singular for $\Pi = 0$ in the scheme A, the equation $\Pi = 0$ does not define a *singular* invariant submanifold of the *irreducible kernel* (as it is apparent from A_2). Therefore, the realizations corresponding to $\Pi = 0$ have not to be considered in general as *singular* realizations in the sense of the given definition.

[11]A first derivation of Eq. (3.8) has been given by Pryce [M.H.L. Pryce, Proc. Roy. Soc. (Lond.) A**195**, 62 (1948); see also G. Møller, Ann. Inst. H. Poincaré **11**, 251 (1949); see also Refs. 7 for a detailed discussion.

[12]G. Lugarini and M. Pauri, Nuovo Cimento, **47**, 299 (1967).

[13]G. Fleming, Phys. Rev. **137**, B188 (1965); **139**, B963 (1965).

[14]Note that, in spite of the notation adopted, p_φ and p_ϑ are the *direction angles* of the vector $\mathbf{p}$.

[15]We use again the symbol $\vec{q}$ for emphasizing the nonvectorial character of the corresponding quantity under space rotations.

[16]E.P. Wigner, Ann. Math. **40**, 149 (1939); "Invariant quantum mechanical equations of motion" in *Theoretical Physics* (I.A.E.A. Vienna, 1963), p. 59; Iu. M. Shirokov, Zh. Eksp. Teor. Fiz. **33**, 1196 (1957) [Sov. Phys. JETP **6**, 919 (1958)]; D. Robinson, Helv. Phys. Acta **35**, 98 (1962); see also G.

Lugarini and M. Pauri, Ann. Phys. (N.Y.) **44**, 266 (1967).

[17]B. Barsella and E. Fabri, Phys. Rev. **126**, 1561 (1962); **128**, 451 (1962); A.J. MacFarlane, Rev. Mod. Phys. **34**, 41 (1962); J. Math. Phys. **4**, 490 (1963); H. Osborne, Phys. Rev. **176**, 1514 (1969).

[18]T.D. Newton and E.P. Wigner, Rev. Mod. Phys. **21**, 400 (1949); A.S. Wightman, Rev. Mod. Phys. **34**, 845 (1962); R.A. Berg, J. Math. Phys. **6**, 34 (1965); H. Bacry, Phys. Lett. **5**, 37 (1963); J. Math. Phys. **5**, 109 (1964).

[19]A. Peres, Phys. Rev. Lett. **27**, 1666 (1971).

[20]E.H. Kerner, J. Math. Phys. **6**, 1218 (1965).

[21]R.N. Hill and E.H. Kerner, Phys. Rev. Lett. **17**, 1156 (1966); R.N. Hill, J. Math. Phys. **8**, 1756 (1967).

[22]E.H. Kerner, Phys. Rev. Lett. **16**, 667 (1966); D.G. Currie, Phys. Rev. **142**, 817 (1966); D.G. Currie and T.F. Jordan, Phys. Rev. Lett. **16**, 1210 (1966); T.F. Jordan, Phys. Rev. **166**, 1308 (1968); R.N. Hill, J. Math. Phys. **11**, 1918 (1970); R.A. Rudd and R.N. Hill, J. Math. Phys. **11**, 2704 (1970); C.S. Shukre and T.F. Jordan, J. Math. Phys. **13**, 868 (1972).

[23]R. Fong and J. Sucher, J. Math. Phys. **5**, 456 (1964); N.D. Son and J. Sucher, Phys. Rev. **153**, 1496 (1967).

[24]S. Lardera, E. Montaldi, and G.M. Prosperi, Nuovo Cimento A **70**, 103 (1970).

[25]H. Tetrode, Z. Phys. **10**, 317 (1922); A.D. Fokker, Z. Phys. Phys., **58**, 386 (1929); Physica, **9**, 33 (1929); **12**, 145 (1932); J.A. Wheeler and R.P. Feynman, Rev. Mod. Phys. **17**, 157 (1945); **21**, 425 (1949); R.M.F. Houtappel, H. van Dam, and E.P. Wigner, Rev. Mod. Phys. **37**, 595 (1965); H. van Dam and E.P. Wigner, Phys. Rev. **138**, B1576 (1965); Phys. Rev. **142**, 838 (1966); J.V. Narlikar, Proc. Camb. Phil. Soc. **64**, 1071 (1968); E.H. Kerner, J. Math. Phys. **3**, 35 (1962); Phys. Rev. **125**, 2184 (1962); R.N. Hill, J. Math. Phys. **8**, 201 (1967); E.H. Kerner, J. Math. Phys. **9**, 222 (1968); A. Katz, J. Math. Phys. **10**, 1929, 2215 (1969); F. Rohrlich, *Classical Charged Particles* (Addison-Wesley, Reading, Mass., 1965), Chap. 7.; C.M. Andersen and H.C. von Baeyer, Ann. Phys. (N.Y.) **60**, 67 (1970); Phys. Rev. D **5**, 802 (1972); Phys. Rev. D **5**, 2470 (1972); A. Degasperis, Phys. Rev. D **3**, 273 (1971); P. Cordero and G.C. Ghirardi, ICTP/72/24-25.

[26]E. Goursat, *A Course in Mathematical Analysis* (Dover, New York, 1959), Vol. II, Part Two, Chap. II, Sec. 3.

Boundedness of linear integral operators in weighted L_p spaces

Yigal Ronen

Atomic Energy Commision, Nuclear Research Centre-Negev, Beer Sheva, Israel
(Received 19 July 1974; revised manuscript received 25 October 1974)

Boundedness properties play a role in the solutions of integral equations as well as in the question of error estimations. A test for the boundedness of linear integral operators in weighted L_p spaces is presented, and applied to the neutron energy dependent transport operator.

1. INTRODUCTION

The boundedness of integral operators in L_p spaces was discussed by Kato,[1] and for the transport operators of neutrons by Boffi and Spiga.[2] This question has also some importance for the theory of error estimations.[3-5] In their representation Boffi and Spiga dealt with the transformation

$$Kf = \int_X k(x, x') f(x') \, dx', \tag{1}$$

which is bounded in the $L_p(X)$ space, with

$$\|K\|_p \leqslant M_2^{1/p} M_1^{1-1/p}, \quad 1 \leqslant p \leqslant \infty. \tag{2}$$

The constants M_1 and M_2 satisfy the conditions

$$\int_X |k(x, x')| \, dx' \leqslant M_1 < \infty \tag{3}$$

for almost every $x' \in X$ and

$$\int_X |k(x, x')| \, dx \leqslant M_2 < \infty \tag{4}$$

for almost every $x \in X$.

However, when this test is applied to the energy dependent transport equation, the conditions required by Eqs. (3) and (4) are not satisfied.

However, boundedness can be achieved when a weighted space is applied. Such a test was given for the L_1 and L_2 spaces.[6,7] By applying a suitable weighting function the energy dependent transport transformation was found to be bounded in the L_1 and the L_2 spaces. By combining the results of these tests with those of Boffi and Spiga, a new test for the L_p weighted spaces was obtained.

2. THE TEST

Consider bounds of operators in the $L_p[X, w^p(x)]$, where the norm of a function in this space is defined with a positive weighting function $w(x)$.

$$\|f\|_p = [\int_X |f(x)|^p w^p(x) \, dx]^{1/p}. \tag{5}$$

Theorem 1: If X is a measurable space, k a measurable function on $X \times X$, w_1 and w_2 real positive measurable functions on X, and if

$$\sup_x \int_X |k(x, x')| w_1(x')/w_1(x) \, dx' \leqslant M_1 < \infty \tag{6}$$

and

$$\sup_x \int_X |k(x', x)| w_2(x')/w_2(x) \, dx' \leqslant M_2 < \infty, \tag{7}$$

then the transformation

$$Kf(x) = \int_X k(x, x') f(x') \, dx' \tag{8}$$

defines a bounded operator in the $L_p(X, [w_2^{1/p}(x)/w_1^{1/q}(x)]^p)$ space, where $1/p + 1/q = 1$, and

$$\|K\|_p \leqslant M_1^{1/q} M_2^{1/p} \tag{9}$$

for any $1 < p < \infty$.

Proof:

$$\|Kf\|_p = \{\int_X |\int_X k(x, x') f(x') \, dx'|^p [w_2^{1/p}(x)/w_1^{1/q}(x)]^p \, dx\}^{1/p}$$

$$\leqslant \{\int_X |\int_X |k(x, x')||f(x')| \, dx'|^p [w_2^{1/p}(x)/w_1^{1/q}(x)]^p dx\}^{1/p}$$

$$= \{\int_X |\int_X |k(x, x')|^{1/q} w_1(x')^{1/q} w_1(x')^{-1/q} |k(x, x')|^{1/p}$$

$$\times |f(x')| \, dx'|^p [w_2^{1/p}(x)/w_1^{1/q}(x)] dx\}^{1/p}. \tag{10}$$

By using Holder's inequality we obtain

$$\|Kf\|_p \leqslant \{\int_X [\int_X |k(x, x')| w_1(x') \, dx']^{p/q} [\int_X |k(x, x') w_1^{-p/q}(x')$$

$$\times |f(x')|^p \, dx'][w_2^{1/p}(x)/w_1^{1/q}(x)] dx\}^{1/p}$$

$$\leqslant M_1^{1/q} [\int_X \int_X |k(x, x')| w_1^{-p/q}(x') |f(x')|^p$$

$$\times w_2(x) w_2(x')/w_2(x') \, dx \, dx']^{1/p}$$

$$\leqslant M_1^{1/q} M_2^{1/p} [\int_X |f(x')|^p w_2(x')/w_1^{p/q}(x') \, dx']^{1/p}$$

$$= M_1^{1/q} M_2^{1/p} \|f\|_p. \tag{11}$$

For the case $p = 1$, the transformation K is bounded by[6]

$$\|K\|_1 \leqslant M_2. \tag{12}$$

In the special case that $M_1 = M_2 = M$ we obtain

$$\|K\|_p \leqslant M \tag{13}$$

for any $1 \leqslant p < \infty$.

Consider now the eigenvalue equations

$$\lambda_0 f_0(x) = \int_X k(x, x') f_0(x') \, dx' = K f_0(x) \tag{14}$$

and

$$\lambda_0 f_0^*(x) = \int_X k(x', x) f_0^*(x') \, dx' = K^* f_0^*(x), \tag{15}$$

where we suppose that the kernel $k(x, x')$ and its adjoint $k^*(x, x') = k(x', x)$ are real positive measurable functions on $X \times X$. There are cases in which the above equations represent physical situations. Thus, the eigenfunction f_0 and its adjoint f_0^* are physical quantities. In many

cases these physical quantities are known to be real positive measurable functions on X. These assumptions imply that the eigenvalue λ_0 is real. Now with these restrictions, in the space $L_p[X, (f^*_0{}^{1/p}(x)/f_0{}^{1/q}(x))]$ according to Theorem 1,

$$\|K\|_p \leqslant \lambda_0. \tag{16}$$

From Eq. (14) we obtain

$$|\lambda_0| \|f_0\|_p = \|Kf_0\|_p \leqslant \|K\|_p \|f_0\|_p. \tag{17}$$

Thus, in this space

$$\|K\|_p = \lambda_0. \tag{18}$$

The norm of the eigenfunction f_0 in this space is

$$\|f_0\|_p = [\int_X f_0{}^p(x)(f_0{}^{*1/p}(x)/f_0{}^{1/q}(x))^p \, dx]^{1/p}$$

$$= [\int_X f_0(x) f_0{}^*(x) \, dx]^{1/p}. \tag{19}$$

Inequalities like (17) hold for any eigenvalue. This fact, with Eq. (18) implies that λ_0 is the largest eigenvalue. Furthermore, there is no other eigenvalue for which the corresponding eigenfunction and its adjoint are real and positive.

3. AN EXAMPLE

The application of the presented test is illustrated by an example from neutron physics. Consider the integral transport operator for a single-region, homogeneous medium, without fission

$$K\phi = \int_V d^3r' \int_\epsilon dE' \int_\omega d^2\Omega' \frac{\exp[-\Sigma(E)|r - r'|]}{|r - r'|^2} \delta(\Omega_R - \Omega)$$

$$\times \Sigma_s(E' \to E, \Omega' \cdot \Omega)\phi(r', E', \Omega'). \tag{20}$$

The distribution function $\phi(r, E, \Omega)$ represents the neutron flux per unit energy, per unit volume and unit solid angle. This function is assumed to be well defined and measurable in the L_p space for $r \in V$, $E \in \epsilon$ and $\Omega \in \omega$, where V is the medium volume, $\epsilon = \{[0, \infty]\}$, and $\omega = \{(\mu, \phi) \in [-1, 1] \times [0, 2\pi]\}$. The total cross section $\Sigma(E)$ is expressed as the sum of the capture and of the scattering cross sections

$$\Sigma(E) = \Sigma_s(E) + \Sigma_c(E), \tag{21}$$

where

$$\Sigma_s(E) = \int_\epsilon dE' \int_\omega d^2\Omega' \Sigma_s(E \to E', \Omega \cdot \Omega')$$

$$= \int_\epsilon dE' \Sigma_s(E \to E'). \tag{22}$$

The scattering differential cross section obeys the detailed balance relation

$$EM(E)\Sigma_s(E \to E', \Omega \cdot \Omega') = E'M(E')\Sigma_s(E' \to E, \Omega' \cdot \Omega), \tag{23}$$

where the Maxwellian distribution is

$$M(E) = [\exp(-E/kT)]/(kT)^2. \tag{24}$$

The solid angle Ω_R is defined as

$$\Omega_R = (r - r')/|r - r'|. \tag{25}$$

The constants M_1 and M_2 of Eqs. (6) and (7) for the transport operator are

$$\sup_{r, E, \Omega} \int_V d^3r' \int_\epsilon dE' \int_\omega d^2\Omega' \frac{\exp[-\Sigma(E)|r - r'|]}{|r - r'|^2} \delta(\Omega_R - \Omega)$$

$$\times \Sigma_s(E' \to E, \Omega' \cdot \Omega)w_1(r', E', \Omega')/w_1(r, E, \Omega) \leqslant M_1 \tag{26}$$

and

$$\sup_{r, E, \Omega} \int_V d^3r' \int_\epsilon dE' \int_\omega d^2\Omega' \frac{\exp[-\Sigma(E')|r - r'|]}{|r - r'|^2} \delta(-\Omega_R - \Omega')$$

$$\times \Sigma_s(E - E', \Omega \cdot \Omega')w_2(r', E', \Omega')/w_2(r, E, \Omega) \leqslant M_2. \tag{27}$$

Substituting for $|r - r'| = R$ and for $d^3r' = R^2 \, dR d^2\Omega_R$, assuming that w_1 and w_2 are independent of r and Ω, and extending the volume integral over all space, Eqs. (26) and (27) are simplified to

$$\sup_{E, \Omega} \int_\epsilon dE' \int_\omega d^2\Omega' \frac{\Sigma_s(E' \to E, \Omega' \cdot \Omega)w_1(E')}{\Sigma(E)w_1(E)} \leqslant M_1 \tag{28}$$

and

$$\sup_{r, \Omega} \int_\epsilon dE' \int_\omega d^2\Omega' \frac{\Sigma_s(E \to E', \Omega \cdot \Omega')w_2(E')}{w_2(E)\Sigma(E')} \leqslant M_2. \tag{29}$$

There are many possibilities for choosing the weighting functions. These functions may be different for each physical problem. Chossing the weighting functions $w_1 = w_2 = 1$, unbounded constants arise. Choosing the weighting functions $w_1 = EM(E)$ and $w_2 = \Sigma(E)$, and, using the detailed balance relation and Eq. (22), we obtain

$$\|K\|_p \leqslant \sup_E [\Sigma_s(E)/\Sigma(E)] \leqslant 1, \quad 1 \leqslant p < \infty. \tag{30}$$

The norm of the flux in L_p space with these weighting functions is

$$\|\phi\|_p = \{\int_V d^3r \int_\epsilon dE \int_\omega d^2\Omega \phi^p(r, E, \Omega)$$

$$\times [\Sigma(E)^{1/p}/EM(E)^{1/q}]^p\}^{1/p}. \tag{31}$$

For the case that $p = 1$, $\|\phi\|_1$ is the total reaction rate in a given volume V.

[1]T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, Berlin, 1966).
[2]V.C. Boffi and G. Spiga, J. Math. Phys. 14, 1913 (1973).
[3]Y. Ronen, Nucl. Sci. Eng. 47, 195 (1972).
[4]Y. Ronen, Nucl. Sci. Eng. 52, 147 (1973).
[5]Y. Ronen, Nuovo Cimento 18, 354 (1973).
[6]Y. Ronen, J. Nucl. Energy 25, 27 (1971).
[7]Y. Ronen, Atomkernenergie 21, 183 (1973).

1523   J. Math. Phys., Vol. 16, No. 7, July 1975

Yigal Ronen   1523

Statistical methods in quantum field theory

Jerzy Hańćkowiak

Institute of Theoretical Physics, University of Wrocław, Wrocław, Poland
(Received 22 August 1974)

We make the construction of n-point functions describing quantum field theory with the aid of classical field equation solutions defined on five-dimensional space. We also introduce the generating functional for n-point functions, fulfilling the first-order functional equation, the same for all local interactions without derivatives.

I. INTRODUCTION

This work is devoted to adaptation of the methods used to describe classical random fields[1-3] to quantum field theory (QFT).

The main tool of these methods is the infinite set of n-point functions or functionals generating them.

Here we consider two functionals G and J.

It turns out that the functional G allows us to construct QFT by means of classical field equation solutions defined in five-dimensional space.

Unfortunately, up to now, there exist no effective methods of solving of equations fulfilled by functionals G in spite of particular results reached by some authors (e.g., see Ref. 4).

The second functional considered here, the functional J, is universal for all interactions without derivatives. QFT with given interaction can be described with the help of the functional J by the appropriate choice of its function argument. However, up to now, the functional J has been known in an unperturbative form in the case of trivial interaction only.

As for to an interpretation of the above functionals, one can say that in the case of random fields the functional G is connected with average products of fields, but the functional J is connected with so-called n-point probabilities.[3]

Now we sketch the content of present paper.

In Sec. 2 the methods of statistical classical fields are considered and the functional G is introduced.

Section 3 describes QFT by means of functional G. There, some identities are also derived.

The functional J is introduced in Sec. 4. It fulfills the first order functional equation, the same for any local interaction without derivatives.

It turns out that the functional J is invariant with respect of certain transformations of its function argument. This is result of the translational invariance of pseudomeasure into integral representation of J.

In Sec. 5 it is explained the role of the fifth dimension appearing in the classical equations (3.1)—(3.2).

2. STATISTICAL DESCRIPTION OF CLASSICAL FIELDS

Let φ be an r − component field

$$\varphi(s,x) \equiv (\varphi_1(s,x), \ldots, \varphi_r(s,x)) \qquad (2.1)$$

defined on $R^1 \times R^n$ manifold satisfying a partial differential equation which is of first order in s:

$$\frac{\partial}{\partial s} \varphi(s,x) = \Lambda[s,x;\varphi]. \qquad (2.2)$$

In general, Λ may be a nonlinear operator. The solutions of (2.2) have a physical meaning then only if they are experimentally verified.

That is not the case, for instance, for the Navier—Stokes equations, where some averaging of solutions and their products (n-point functions) have only a physical meaning.

Advantage tools to description of such theories are functionals used mainly to guarantee a compact description of equations fulfilled by n-point function and symmetricity of one.

First, let us consider the functional

$$G[s;j] \equiv \int \exp\left(i \sum_k^r \int j_k(x)\varphi_k[s,x;\alpha]\,dx\right) F[\alpha]\delta\alpha$$

$$\equiv \int e^{i(j,\varphi[s;\alpha])} F[\alpha]\delta\alpha, \qquad (2.3)$$

where φ are solutions of the starting equation (2.2) with marked explicitly initial and boundary conditions α, F is a smearing functional, $\delta\alpha$ denotes a pseudomeasure on the space of functions α.

G is related with n-point functions as follows:

$$(-i)^n \frac{\delta}{\delta j(x_1)} \cdots \frac{\delta}{\delta j(x_n)} G[s;j]\Big|_{j=0}$$

$$= \int \varphi[s,x_1;\alpha] \cdots \varphi[s,x_n;\alpha] F[\alpha]\delta\alpha. \qquad (2.4)$$

The implicit assumption included in (2.4) about the possibility of functional differentiation under integral sign in (2.3) is not necessary if from the beginning we consider the physical quantities only, namely n-point functions, or use so-called functional power series.[5]

Sometimes, however, it is useful to deal with the functional integrals like (2.3) since they lead more quickly to the final result.

The equation which fulfills functional G is obtained by differentiation of (2.3) with respect to parameter s and taking into account Eqs. (2.2):

$$\frac{\partial}{\partial s} G = i \sum_k^r \int dx\, j_k(x)\Lambda_k\left[s,x; -i\frac{\delta}{\delta j}\right] G. \qquad (2.5)$$

In Eq. (2.5) there do not appear the solutions of starting Eq. (2.2) and, *de facto*, they are not needed to find the functional G.

The representation of the functional G in the form

$$G[s;j] = \int \exp[i \int j(x)\alpha(x)\,dx]F[s;\alpha]\delta\alpha \qquad (2.6)$$

reveals a true amount of information taken from Eq. (2.2) to obtain n-point functions (2.4).

Indeed, substituting (2.6) into (2.5) and integrating by parts, we obtain for F the equation

$$\frac{\partial}{\partial s}F + \sum_k^r \int dx\, \Lambda_k[s,x;\alpha]\frac{\delta}{\delta\alpha_k(x)}F$$

$$= -F\sum_k^r \int dx\, \frac{\delta\Lambda_k[s,x;\alpha]}{\delta\alpha_k(x)}, \qquad (2.7)$$

which for fields with restriction

$$\sum_k^r \int dx\, \frac{\delta\Lambda_k[s,x;\alpha]}{\delta\alpha_k(x)} = 0 \qquad (2.8)$$

is an equation for integrals of motion of the differential equation (2.2).

The condition (2.8) is always fulfilled for the canonical systems [see (3.2)].

3. DESCRIPTION OF QFT WITH THE HELP OF THE FUNCTIONAL G

In this section we shall show that QFT of one scalar local field can be described by means of functional G connected with some classical equation; however, the smearing functional F appearing in the definition of G functional has no interpretation of the probability density.

For that purpose let us consider a self-interacting, scalar, classical field defined in $R^1 \times R^4$:

$$\left(\frac{\partial^2}{\partial s^2} + \Box - m^2\right)\chi(s,x) = -\mathcal{L}'(\chi(s,x)), \qquad (3.1)$$

where $(s,x) \in R^1 \times R^4$, $\mathcal{L}'$ denotes a derivative of Lagrangian density.

Equation (3.1) can be described in a canonical form by introducing new fields $\varphi_1 = \chi$ and $\varphi_2 = \dot{\chi}$:

$$\dot{\varphi}_1 = \frac{\delta H}{\delta\varphi_2} \equiv \Lambda_1 = \varphi_2,$$

$$\dot{\varphi}_2 = -\frac{\delta H}{\delta\varphi_1} \equiv \Lambda_2 = -(\Box - m^2)\varphi_1 - \mathcal{L}'(\varphi_1), \qquad (3.2)$$

where H is a Hamiltonian of the system (3.1):

$$H = \int dx\,[\tfrac{1}{2}(\varphi_2)^2 + \tfrac{1}{2}\varphi_1(\Box - m^2)\varphi_1 + \mathcal{L}(\varphi_1)]. \qquad (3.3)$$

Dots here mean derivatives with respect to s. Now, choosing the smearing functional

$$F = e^{iH}, \qquad (3.4)$$

we get for the functional G [see (2.3)] the following expression:

$$G \equiv \int e^{i(j,\varphi[s;\alpha])}e^{iH[\alpha]}\delta\alpha = \int e^{i(j,\alpha)}e^{iH[\alpha]}\delta\alpha. \qquad (3.5)$$

This equation can be understood by using the formula (2.6) and keeping in mind that (3.4) is an integral of motion of the canonical Eq. (3.2), which does not depend upon time. We remind that solutions of (3.2) fulfill the conditions $\varphi[s,x;\alpha]|_{s=0} = \alpha(x)$.

Performing now the integration with respect to α_2, we see that

$$G = e^{-i(j_2,j_2)/2}\tau[j_1], \qquad (3.6)$$

where

$$\tau[j_1] = \int e^{i(j_1,\alpha_1)}\exp\{i \int dx[\tfrac{1}{2}\alpha_1(\Box - m^2)\alpha_1 + \mathcal{L}(\alpha_1)]\}$$

is a generating functional for the time ordered Green's functions describing QFT.

This finishes our statement expressed at the beginning of this section.

The possibility of expressing G with the help of (3.5) gives some identity for the τ functional. Indeed, putting in (3.5) $j_2 = 0$, we get

$$\int e^{i(j_1,\varphi_1[s;\alpha])}e^{iH[\alpha]}\delta\alpha = \tau[j_1].$$

Hence

$$\exp\{i(j_1, \varphi_1[s; -i\delta/\delta\zeta])\}\, e^{-i(\zeta_2,\zeta_2)/2}\tau[\zeta_1]\big|_{\zeta=0} = \tau[j_1]. \qquad (3.7)$$

The formula (3.5) also shows that G is the stationary functional ($\dot{G} = 0$). However, it turns out that G may be expressed through different nonstationary functionals.

It is seen, for instance, by splitting the energy integral on free and interacting parts,

$$H = H_0 + H_{int}, \qquad (3.8)$$

and considering the functionals

$$G_n[s;j] \equiv \frac{i^n}{n!}\int e^{i(j,\varphi[s;\alpha])}H_{int}^n e^{iH_0[\alpha]}\delta\alpha. \qquad (3.9)$$

These functionals fulfill the nonstationary equation (2.5), because smearing functionals

$$F_n \equiv H_{int}^n\, e^{iH_0} \qquad (3.10)$$

are not (stationary) integrals of motion of Eq. (3.2). Of course,

$$\sum_n G_n[s;j] = G[j]. \qquad (3.11)$$

At the end of this section we would like to demonstrate the fulfilling of Eq. (2.7) by some integrals of motion in the case of the starting equation (3.2).

Since now (2.8) is fulfilled, (2.7) takes the form of equation for integrals of motion of Eq. (3.2):

$$\dot{F} + \int dx\left(\alpha_2\frac{\delta}{\delta\alpha_1} - [(\Box - m^2)\alpha_1 + \mathcal{L}'(\alpha_1)]\frac{\delta}{\delta\alpha_2}\right)F = 0. \qquad (3.12)$$

It turns out that in spite of the functional character of Eq. (3.12) some solutions can be easily found due to the Noether theorem or, directly, from Eq. (3.2).

The energy integral (3.3) or integral (3.4) fulfill (3.12), which is easily seen by substitution.

The momentum integrals

$$P_j[\alpha] \equiv \int dx\, \alpha_2\frac{\partial}{\partial x_j}\alpha_1 \qquad (3.13)$$

fulfill Eq. (3.12) because

$$\int dx\left(\alpha_2\frac{\partial}{\partial x_j}\alpha_2 - [(\Box - m^2)\alpha_1 + \mathcal{L}'(\alpha_1)]\frac{\partial}{\partial x_j}\alpha_1\right) = 0.$$

Indeed, the first and second terms disappear, for integration by parts gives the same expression but with

1525   J. Math. Phys., Vol. 16, No. 7, July 1975

Jerzy Hańćkowiak   1525

different signs. The third term

$$\int dx \, L'(\alpha_1) \frac{\partial}{\partial x_j} \alpha_1 = \int dx \frac{\partial}{\partial x_j} L(\alpha_1) = 0.$$

4. DESCRIPTION OF QFT WITH THE HELP OF FUNCTIONAL J

Let $U(\hat{x})$ be a function defined on a five-dimensional space: $\hat{x} = (\omega, x) \in R^1 \times R^4$.

The functional J, from definition, is equal to

$$J[U] \equiv \int \exp[i \int U(\alpha(x), x) dx]$$
$$\times \exp[\tfrac{1}{2}i \int \alpha(\square - m^2)\alpha \, dx] \delta\alpha \qquad (4.1)$$

Here the functional pseudomeasure $\delta\alpha$ is translational invariant, and α means a one-component function defined on R^4.

Of course,

$$J[u]\big|_{u=L (\omega) + \omega j_1(x)} = \tau[j_1], \qquad (4.2)$$

where τ is defined by (3.6).

The above equality shows that any local theory with Lagrangian which does not contain derivatives can be obobtained from the functional J.

Unfortunately, up to now, J is known only for functions

$$U(\omega, x) = \omega j(x) + b\omega^2,$$

which correspond to the trivial interaction (n-point functions obtained from the J functional are known. We do not know, however, the J functional given in a compact form.)

Now we shall derive some invariant properties of J, which are imposed by the translational invariance of pseudomeasure.

We have

$$J = \int e^{iU[\alpha]} \, e^{i\alpha(\square - m^2)\alpha/2} \delta\alpha$$
$$= \int e^{iU[\alpha + h]} \, e^{i(\alpha + h)(\square - m^2)(\alpha + h)/2} \delta\alpha, \qquad (4.3)$$

where we have used some obvious abbreviations of description.

Introducing the transformation Λ_h acting on a function U as follows,

$$\Lambda_h U(\omega, x)$$
$$= U(\omega + h(x), x) + \omega(\square - m^2)h(x), \qquad (4.4)$$

one can describe Eq. (4.3) in the compact way:

$$J[\Lambda_h u] = e^{-ih(\square - m^2)h/2} J[u], \qquad (4.5)$$

where h and u are any functions belonging to a certain class.

Equation (4.5) expresses the translational invariance of a functional pseudomeasure appearing in the definition of the J functional.

Differentiating (4.5) with respect to h and putting $h = 0$, one can get for J the first order functional differential equation

$$\int d\omega \left(\frac{\partial}{\partial \omega} U(\hat{x}) + \omega(\square - m^2) \right) \frac{\delta}{\delta U(\hat{x})} J = 0, \qquad (4.6)$$

where $\hat{x} = (\omega, x)$.

This equation is not equivalent to (4.5) since the functional

$$J = \exp[\int U(\hat{x})\varphi(x) d\hat{x}], \qquad (4.7)$$

where φ fulfills the Klein—Gordon equation:

$$(\square - m^2)\varphi(x) = 0$$

solves (4.6) but not (4.5).

For h fulfilling the Klein—Gordon equation: (4.4) becomes

$$\Lambda_h U(\omega, x) = U(\omega + h(x), x) \qquad (4.8)$$

and transformation Λ_h has the group property

$$\Lambda_{h'} \Lambda_h = \Lambda_{h' + h} \qquad (4.9)$$

Now Eq. (4.5) is the following:

$$J[\Lambda_h u] = J[u], \qquad (4.10)$$

where h is any solution of the Klein—Gordon equation.

5. FOUR-DIMENSIONAL FORMALISM

From Sec. 3 one can see that QFT of one scalar field defined on the *four-dimensional* space may be described with the help of the mathematical apparatus used in the case of statistical classical field considered on the *five-dimensional* space.

At first sight this enlarging of dimensions may be the result of using so-called space functional instead of space—time functionals (Ref. 2, Vol. 2, p. 614) leading to n-point functions with the same time (parameter s).

However, the considerations below show that the reasons are more intrinsical.

For that purpose let us consider a self-interacting classical field defined on the four-dimensional space:

$$(\square - m^2)\varphi(x) = -L'(\varphi(x)), \quad x = (t, \mathbf{x}) \in R^4. \qquad (5.1)$$

The space functional G (see Sec. 2) is defined as follows:

$$G[t; j] \equiv \int \exp\{i \int \varphi[t, \mathbf{x}; \alpha] j(\mathbf{x}) d\mathbf{x}\} F[\alpha] \delta\alpha. \qquad (5.2)$$

The definition of the space—time functional G leading to n-point functions with different times is

$$G[\mathcal{J}] \equiv \int \exp\{i \int \varphi[x; \alpha] \mathcal{J}(x) dx\} F[\alpha] \delta\alpha. \qquad (5.3)$$

The energy integral of Eq. (5.1) is an integral over R^3:

$$H = \int d\mathbf{x}[\tfrac{1}{2}(\varphi_2)^2 - \tfrac{1}{2}\varphi_1(\Delta - m^2)\varphi_1 + L(\alpha_1)]. \qquad (5.4)$$

Due to a three-dimensional integration, the dimension of H is

$$[H] = L^{-1}. \qquad (5.5)$$

Here we have assumed that the dimension of the classical field (5.1) is the same as quantum field.]

This fact makes it now impossible to postulate for the smearing functional F the simple expression (3.4), if we do not want to introduce some unphysical dimensional parameter.

The dimensionless of F in (5.4) may be realized with the help of more complicated integrals of motion, for example,

$$F = f(H/\sqrt{P_j P_j}),$$

(5.6)

but it is not known whether such functionals do lead to effective unitary S-matrix theory.

It is of interest to notice that assuming for a field φ such dimension that H is a dimensionless quantity leads to an incorrect dimension for S-matrix elements constructed from the generating functional

$$S[\alpha, \beta] \equiv \tau[(\Box - m^2)q_0[\alpha, \beta]]$$

$$= G[\mathcal{J}]\big|_{\mathcal{J}_2 = 0, \mathcal{J}_1 = (\Box - m^2)q_0[\alpha, \beta]}.$$

(5.7)

(see Ref. 6), or forces us again to introduce an un-

physical dimensional parameter.

In the case of the five-dimensional equation (3.1), the energy integral (3.3) as well as momentum integrals are dimensionless quantities, what permits us to construct from them smearing functionals without introducing any dimensional constants.

[1]E. Hopf, J. Ratl. Mech. Anal. 1, 87 (1952).
[2]A.S. Monin and A.M. Jaglom *Statistical Hydromechanics*, *Vols, I* and *II* (Moscow, 1965 and 1967).
[3]W.E. Brittin and W.R. Chappell, J. Math. Phys. 10, 661 (1969).
[4]B. Szafirski, Bull. Acad. Polon., Ser. Sci. Math. 19, 791, 799 (1971).
[5]D. Ruelle, Rev. Mod. Phys. 36, 580 (1964).
[6]J. Rzewuski, *Field Theory II* (Warsaw, 1969).

On the symmetries of the Racah coefficient

K. Srinivasa Rao, T. S. Santhanam, and K. Venkatesh

MATSCIENCE, The Institute of Mathematical Sciences, Madras-600020, India
(Received 23 July 1974)

Observations on the symmetries of the Racah coefficient are made in terms of its three allowed series representations. The Bailey's transform between the two terminating Saalschutzian $_4F_3(1)$ series, when applied to the $_4F_3(1)$ representation for the Racah coefficient, is shown to result (at best) in the allowed substitution $j \rightarrow -j -1$ for any one of the six angular momenta in the Racah coefficient.

The Racah coefficient[1] $W(abcd; ef)$ is defined as a single sum over an index P whose range is restricted to

$$\max \begin{bmatrix} a+b+e, a+c+f, \\ b+d+f, c+d+e \end{bmatrix}$$

$$\leqslant P \leqslant \min \begin{bmatrix} a+b+c+d, b+c+e+f, \\ a+d+e+f \end{bmatrix}. \tag{1}$$

By setting $n = a+b+c+d-P$, $n = b+c+e+f-P$, and $n = a+d+e+f-P$, in succession, we can obtain three series representations.

The well-known symmetries of the Racah coefficient are

(a) Permutations of the columns of $\begin{Bmatrix} abe \\ dcf \end{Bmatrix}$,

(b) interchange of any two elements in a row of $\begin{Bmatrix} abe \\ dcf \end{Bmatrix}$ with the corresponding elements in the other row, and

(c) the symmetries found by Regge.[2]

In literature,[2] the 24 symmetries of the types (a) and (b) are together called as tetrahedral symmetries. Since each of the six Regge symmetries can be superimposed on a given tetrahedral symmetry, we have, in all, 144 symmetries of the Racah coefficient.

The series representations of the Racah coefficient can be rearranged into the generalized hypergeometric series of unit argument, viz., $_4F_3(ABCD;EFG;1)$, defined in the usual manner,[3] as in Rose,[4] with

$$A = e-a-b, \quad B = e-c-d, \quad C = f-a-c, \quad D = f-b-d,$$

$$E = -a-b-c-d-1, \quad F = e+f-a-d+1,$$

$$G = e+f-b-c+1, \tag{2}$$

which corresponds to the series obtained by substituting $n = a+b+c+d-P = r-P$. Similarly, corresponding to the series obtained for substitutions $n = b+c+e+f-P = r-P$ and $n = a+d+e+f-P = r-P$, we have two $_4F_3$ (1)s whose numerator and denominator parameters are given, respectively, by

$$A = d-b-f, \quad B = a-b-e, \quad C = d-c-e,$$

$$D = a-c-f, \quad E = -b-c-e-f-1,$$

$$F = a+d-e-f+1, \quad G = a+d-b-c+1 \tag{3}$$

and

$$A = b-a-e, \quad B = b-d-f, \quad C = c-a-f,$$

$$D = c-d-e, \quad E = -a-d-e-f-1,$$

$$F = b+c-e-f+1, \quad G = b+c-a-d+1. \tag{4}$$

We notice that, for all the physical values of a, b, c, d, e, and f, all the numerator parameters, A, B, C, and D, are nonpositive and the denominator parameter E, being also negative, satisfies the condition $(A, B, C$ or $D) > E$, by virtue of the triangular conditions. For the $_4F_3$ (1) series to be convergent,[4] there must be a numerator parameter such that

$$(A, B, C, D) \geqslant (F, G). \tag{5}$$

However, comparison of the denominator parameters with the numerator parameters, along with the triangular conditons, yields the condition

$$(F, G) > (A, B, C, D) \tag{6}$$

in all the three cases. From (5) and (6) it follows that both F and G must be greater than zero for the $_4F_3$ (1) series to be convergent. The values of F and G depend upon the relative magnitudes of $a+d$, $b+c$, and $e+f$. It is straightforward to find that except for $a=b=c=d=e=f$ and $a+d=b+c=e+f$, when all the three $_4F_3$ (1)'s are convergent, for other physically allowed values of a, b, c, d, e and f, only one or two of the set of three $_4F_3$ (1)'s are convergent. Thus, the set of three $_4F_3$ (1) series representations are necessary.

Since a $_4F_3$ (1) = 1 when any one or more of the numerator parameters is zero, we can immediately obtain 24 formulas for special values of the arguments of the Racah coefficient, like the one given in Edmonds.[5]

Choosing any one of the three allowed $_4F_3$ (1) series representations for $\begin{Bmatrix} abe \\ dcf \end{Bmatrix}$, say, that with parameters given by (2), we find that symmetries of type (a) lead to either one of the other two $_4F_3$ (1) series [i.e., those with parameters given by (3) or (4)] or the same $_4F_3$ (1) series whose parameters are permuted as in $_4F_3$ $(ABDC; EGF;1)$. Symmetries of type (b), applied to any one of the three $_4F_3$ (1) series, lead to the same series but with numerator parameters alone permuted as in:

$$_4F_3 (CDAB; EFG; 1), \quad _4F_3 (DCBA; EFG; 1),$$

$$\text{and} \quad _4F_3 (BADC; EFG; 1).$$

All the other permutations of the numerator parameters lead to one of the five Regge symmetries on which a tetrahedral symmetry of type (a) and/or (b) is superimposed. It is straightforward to list the 24 possible numerator parameter permutations of a given $_4F_3$ (1) series and the corresponding symmetries of the Racah coefficient associated with them, with the denominator parameters kept fixed as EFG or permuted as EGF (E being identified as the obviously negative denominator parameter).

It is interesting to note that, of the $4! \times 3! = 144$ possible permutations of the numerator and denominator parameters of a given $_4F_3$ (1) series, only $4! \times 2 = 48$ permutations correspond to the known symmetries of the Racah coefficient. For, only two of the six possible denominator parameter permutations (EFG and EGF) correspond to meaningful symmetries of the Racah coefficient. For example, if we consider the case in which the denominator parameters alone are permuted to FEG, then the Racah coefficient $W(abcd; ef)$ is found to be related to

$$W(a, \tfrac{1}{2}[b - c - e - f - 2], \tfrac{1}{2}[-b + c - e - f - 2], d;$$

$$\tfrac{1}{2}[-b - c + e - f - 2], \tfrac{1}{2}[-b - c - e + f - 2]). \tag{7}$$

This is not a meaningful Racah coefficient, as long as we assume angular momenta to take positive integer and half-integer values only; since four of the six arguments in the coefficient (7) are strictly negative, they violate the defining triangular conditions on a, b, c, d, e, and f.

Therefore, of the 144 known symmetries of the Racah coefficients only 48—eight tetrahedral symmetries plus the five Regge symmetries on each of which eight tetrahedral symmetries are superimposed—can be accounted for by considering the numerator and denominator parameter permutations of any one of the three $_4F_3$ (1) series. This is due to the fact that there are three allowed series representations for a Racah coefficient and there are only eight tetrahedral symmetries which leave a given series invariant. To be explicit, the eight tetrahedral symmetries that leave the series given by (2) invariant are $\{^{abc}_{dcf}\}$ and $\{^{bae}_{cdf}\}$ on each of which symmetries of type (b) are superimposed. Again, four of the six Regge symmetries and four of the six column permutations take one series representation into another.

Minton[6] has tried to arrive at a new symmetry for the Racah coefficient by resorting to the Bailey transform[7] between the two terminating Saalschutzian $_4F_3$ (1) series:

$$_4F_3\,(ABCD; EFG; 1)$$

$$= \Gamma \begin{bmatrix} E + F - A - B - D, E + F - A - B - C, F - C - D, F \\ E + f - A - B, E + F - A - B - C - D, F - C, F - D \end{bmatrix}$$

$$\times {}_4F_3(E - B, E - A, C, D; E, E + F - A - B; E + - A - B; 1), \tag{8a}$$

where, we have used the notation,

$$\Gamma \begin{bmatrix} p, q, \cdots \\ r, s, \cdots \end{bmatrix} = \frac{\Gamma(p)\Gamma(q) \cdots}{\Gamma(r)\Gamma(s) \cdots}.$$

The above relation is invariant to interchange of A and B, C and D, and F and G. We observe that if we set $C = 0$, then the Γ factor in (8a) becomes 1 and the $_4F_3$ (1) series on the right and left sides of (8a) is 1. Choosing the set of parameters given by (8a) for the $_4F_3$ (1)

series, we find that

$$_4F_3 \begin{bmatrix} e - a - b, e - c - d, 0, a + c - b - d \\ -a - b - c - d - 1, c + e + 1 - d, a + e + 1 - b \end{bmatrix} = 1$$

$$= {}_4F_3 \begin{bmatrix} -e - 1 - a - b, -e - 1 - c - d, 0, a + c - b - d \\ -a - b - c - d - 1, c - e - d, a - e - b \end{bmatrix}. \tag{8b}$$

An identification of the corresponding numerator and denominator parameters on the right and left sides of (8b) clearly shows that the Bailey transform when applied to the $_4F_3$ (1) series given by (2) for the Racah coefficient leads to the substitution $e \rightarrow -e - 1$. This

$$j \rightarrow -j - 1 \tag{9}$$

substitution is mathematically allowed, since it leaves the eigenvalue of the square of the angular momentum operator as well as the matrix elements of the raising and lowering angular momentum operators unaltered. This property has been considered[8] in all the formulas of the Clebsch—Gordan and Racah coefficients, though physically negative angular momenta are not of significance.

When using the Bailey transform for the $_4F_3$ (1) series of the Racah coefficient, it should be noted that two of the numerator parameters (C and D) and a denominator parameter (E) are left unaltered. Since all the numerator parameters of the $_4F_3$ (1) series given by (2), (3), and (4) are negative, it follows that keeping two of them (C and D) fixed implies: min $(A, B) \le$ min (C, D). Since there is an arbitrariness in the naming of the numerator parameters as A, B, C, D and the denominator parameters as E, F, G, there arise ${}^4C_2 \times {}^3C_1 = 18$ possible ways of choosing (C and D) and E. Six of these 18 possibilities can be identified to be due to a $j \rightarrow -j - 1$ (or, equivalently, $-j \rightarrow j + 1$) substitution for one of the six angular momenta which occur in a Racah coefficient. Each of the twelve other cases leads to a $_4F_3$ (1) series whose parameters cannot be identified to be those of any of the three allowed $_4F_3$ (1) series (2), (3), or (4), under any substitution(s) of the type (9). Hence, it is clear that the $_4F_3$ (1) series that arises after the Bailey transformation in these 12 cases does not relate to a Racah coefficient. In fact, Minton made the choice

$$C = f - a - c, D = f - b - d \text{ and } E = e + f - a - d + 1$$

and was led to his "new" symmetry for the Racah coefficient which violates the defining triangular conditions.[9] Therefore, we conclude that it is not possible to arrive at meaningful new symmetries for the Racah coefficient by using the Bailey transform for a $_4F_3$ (1) series.

ACKNOWLEDGMENTS

[1]G. Racah, Phys. Rev. **62**, 438 (1942).

[2]T. Regge, Nuovo Cimento 11, 116 (1959).

[3]L. J. Slater (Cambridge U. P., Cambridge, 1966), Chap. 2. Also, L. C. Biedenharn, J. M. Blatt and M. E. Rose, Rev. Mod. Phys. **24**, 249 (1952).

[4]M. E. Rose, *Multipole Fields* (Wiley, New York, 1955), Appendix B, p. 92.

[5]A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton U. P., Princeton, N. J., 1957), p. 92.

[6]B. M. Minton, J. Math. Phys. **11**, 3061 (1970). If Minton had used the relation (8a), instead of the one he used with $E-B$ and $E-A$ permuted, then he would have arrived at the "new"

symmetry for the Racah coefficient:

$$W(abcd;ef) = W(a, \tfrac{1}{2}[b - c + e - f - 1], \tfrac{1}{2}[-b + c + e - f - 1], \ d;$$
$$\tfrac{1}{2}[b + c + e + f + 1], \quad \tfrac{1}{2}[-b - c + e + f - 1]).$$

[7]Reference 4, p. 64, Eq. (2.4.1.7).

[8]A. A. Bandzaitis, A. V. Karosene, A. Yu Savunikas, and A. P. Yutsis, Dokl. Akad. Nauk SSSR **154**, 812 (1964) [Sov. Phys. Doklady **9**, 139 (1964/65)]; also, Ya. A. Smorodinskii and L. A. Shelepin, Usp. Fiz. Nauk **106**, 3 (1972) [Sov. Phys. Usp. **15**, 1 (1972)]

[9]E. Yakimiw, J. Math. Phys. **12**, 1134 (1971). Also, Vinaya Joshi, J. Math. Phys. **12**, 1134 (1971).

1530   J. Math. Phys., Vol. 16, No. 7, July 1975

Rao, Santhanam, and Venkatesh   1530

Erratum: Lie theory and separation of variables. 3. The equation $f_{tt} - f_{ss} = \gamma^2 f$ [J. Math. Phys. 15, 1025 (1974)]

E. G. Kalnins and W. Miller, Jr.

Centre de Recherches Mathématiques, Université de Montréal, Montréal 101, P.Q., Canada
(Received 11 April 1975)

Dr. Charles Boyer has kindly pointed out an error in the computation of the spectrum of the operator $L_B = \partial_{xx} + \gamma^2 e^{2x}$ on $L_2(R)$, which corresponds to the Bessel function basis of solutions for the Klein—Gordon equation. This error, which is the responsibility of the second author, consists in the assertion that the self-adjoint extensions $L_{B,\alpha}$, have only discrete spectrum $\lambda = (2n + \alpha)^2$ and an orthonormal basis of eigenfunctions

$$f_n^{B,\,\alpha}(x) = \sqrt{2(\alpha + 2n)}\, J_{\alpha+2n}(\gamma e^x), \quad n = 0, 1, 2, \cdots,$$

where $0 < \alpha \le 2$. In fact, as is shown on pp. 93—95 of the book *Eigenfunction Expansions. Part One.* (Oxford U.P., Oxford, 1962), 2nd ed., by E.C. Titchmarsh, these operators also have continuous spectrum.

Taking the case $\alpha = 2$ for simplicity, we find that the operator L_B has discrete spectrum $\lambda = 4(n + 1)^2$, $n = 0, 1, \ldots$, and continuous spectrum $\lambda \le 0$ with generalized eigenfunctions

$$\tilde{f}_\lambda^B(x) = [J_{i\sqrt{-\lambda}}(\gamma e^x) + J_{-i\sqrt{-\lambda}}(\gamma e^x)]/2[\sinh(\pi\sqrt{-\lambda})]^{1/2},$$

$$\langle \tilde{f}_\lambda^B, \tilde{f}_{\lambda'}^B \rangle = \delta(\lambda - \lambda').$$

Here, $\langle \cdot\,, \cdot \rangle$ is the usual $L_2(R)$ inner product. The functions $\{f_n^B, \tilde{f}_\lambda^B\}$ together form a complete set for $L_2(R)$.

The separable solutions of the Klein—Gordon equation corresponding to the continuum basis are

$$\tilde{F}_\lambda^B(s, t) = [\sinh(\pi\sqrt{-\lambda})]^{-1/2}\,[J_{i\sqrt{-\lambda}}(\gamma u) + J_{-i\sqrt{-\lambda}}(\gamma u)]$$
$$\times K_{i\sqrt{-\lambda}}(-i\gamma v)$$

where
$$s = (u^2 + u^2 v^2 + v^2)/2uv, \quad t = (u^2 - u^2 v^2 + v^2)/2uv, \quad v > u > 0.$$

There are similar expressions for other regions of the (u, v) plane.

The error, while regrettable, in no way affects the principal methods and conclusions of this paper and of following papers in our series.

CUMULATIVE AUTHOR INDEX