Free particle systems and their dynamics in de Sitter space

Raymond C. Fabec

Department of Mathematics, Louisiana State University, College of Arts and Sciences, Baton Rouge, Louisiana 70803
(Received 29 November 1976)

A first step in determining all the global projective unitary representations describing free particle systems in imaginary Lobachevsky space is made. Essentially we determine explicitly the global form of any representation describing a free particle of spin j on the group generated by rotations and translations of space–time at time t and time translations. We also discuss whether or not positional observables should be preserved under physical equivalence and determine the effects this has on the representation theory of free particle systems.

In this paper free particle systems in imaginary Lobachevsky space with symmetry group $SO_o(1,4)$ are considered. We start by describing imaginary Lobachevsky space and its interpretation as space–time and its symmetry group. Space–time at time t is shown to be Euclidean space, and Euclidean motions, that is translations and rotations, form a subgroup of $SO_o(1,4)$. Their products with time translations form a subgroup $\overline{P}$ of $SO_o(1,4)$; $\overline{P}$ is a minimal parabolic subgroup of $SO_o(1,4)$.

Next we discuss free particles, physical equivalence, and equivalence. The difference between physical equivalence and equivalence is the latter insures preservation of the positional observables while the former does not. If one does not insist on the preservation of the positional observables, it will be shown that any representation describing any particle of spin j has the same explicit form on $\overline{P}$. Since $\overline{P}$ contains time translations, the Hamiltonians of all free particles of spin j are identical. This leads to some confusion in regards to their masses. There has already been some confusion in this regard as the Hamiltonians for systems using $SO_o(1,4)$ fail to have a minimum eigenvalue. If one insists the positional observables be preserved, then we show it is possible to find all particles of spin j once at least one is known from each physical equivalence class. In this case, however, there are many more free particles, and one can question their significance.

Our eventual goal will be to list the representations describing free particle systems having symmetries $SO_o(1,4)$. We have reduced it to finding the extensions of certain representations of $\overline{P}$ to the whole group $SO_o(1,4)$. We decided to use $SO_o(1,4)$ rather than $SO_o(3,2)$ for the former's global theory is better understood. Eventually we hope to look at similar problems for $SO_o(3,2)$.

There are several papers dealing with elementary particles in spaces of constant curvature. We mention those of Hannabus and Fronsdal. Hannabus has looked at the same problem we have and has claimed to have obtained all the real mass representations; see Ref. 1. The results here show there may be some question as to the meaning of real mass representations. Fronsdal has worked extensively in both positively and negatively curved Minkowski space, using $SO(3,2)$ as the invariance group. He however has studied elementary particles by analysing the differential operators of the group. For more details see Ref. 2.

We have included an appendix defining projective representations, induced representations, projection valued measures, and central projections for those who may be unfamiliar with these concepts.

1. IMAGINARY LOBACHEVSKY SPACE, ITS INTERPRETATION AS SPACE-TIME, AND THE SYMMETRY GROUP SO_o (1, 4)

Imaginary Lobachevsky space is the four-dimensional manifold M obtained by identifying antipodal points in the submanifold $x_0^2-x_1^2-x_2^2-x_3^2-x_4^2=-1$ of $\mathbf{R}^5$. The pseudometric $K^2(dx_0^2-dx_1^2-dx_2^2-dx_3^2-dx_4^2)$ turns M into a pseudo-Riemannian manifold with positive curvature K^2. This manifold is a space–time model for an infinite, expanding universe in general relativity; it is one of a collection of models known as De Sitter spaces. $SO_o(1,4)$, the component of the identity of the real 5 by 5 matrices preserving the quadratic form $x_0^2-x_1^2-x_2^2-x_3^2-x_4^2$, acts isometrically and transitively on the manifold M. It will be used as the symmetry group for quantum mechanical systems in M.

Absolute time in M is defined by $t(x)=\ln|x_0+x_1|$ where $x=\pm(x_0,x_1,x_2,x_3,x_4)\in M$. For $t>-\infty$, $M(t)=\{x\in M: t(x)=t\}=\{(\sinh t+\frac{1}{2}e^t|p|^2, \cosh t-\frac{1}{2}e^t|p|^2, e^t p): p\in\mathbf{R}^3\}$ is 3-space at time t. This coordinatization of $M(t)$ defines a diffeomorphism of $\mathbf{R}^3$ onto M_t. It suggests defining for each element $p+R$ of the Euclidean group E_3 the linear transformation $p+R$ of $\mathbf{R}^5$ by

$$(p+R)(x_0,x_1,q)=((1+\tfrac{1}{2}|p|^2)x_0+\tfrac{1}{2}|p|^2x_1 + p\cdot Rq,\ -\tfrac{1}{2}|p|^2x_0+(1-\tfrac{1}{2}|p|^2)x_1 - p\cdot Rq,\ (x_0+x_1)p+Rq).$$

That each element $p+R$ belongs to $SO_o(1,4)$ is easily shown. Hence $p+R$ is an isometry of M. It leaves each time-slice $M(t)$ invariant, and the diffeomorphism between $\mathbf{R}^3$ and $M(t)$ interwines rotation by R followed by translation by p with $p+R$. In particular $(p_1+R_1)(p_2+R_2)=p_1+R_1p_2+R_1R_2$. Hence $E_3=\{p+R: p\in\mathbf{R}^3,\ R\in SO(3)\}$; the Euclidean group, is contained in $SO_o(1,4)$.

Time translation which we denote by (I,t) or some-

 Copyright © 1977 American Institute of Physics

times just by t is the hyperbolic rotation $(x_0,x_1,p) \mapsto (x_0\cosh t + x_1 \sinh t, x_0 \sinh t + x_1 \cosh t, p)$ where $p \in \mathbb{R}^3$. Clearly $(I,t) \in SO_o(1,4)$ and $(I,t)\cdot M(s) = M(t+s)$.

The following identities are easily established:

(i) $(I,t)R = R(I,t)$ for $t \in \mathbb{R}$, $R \in SO(3)$,

(ii) $(I,t)(p+I) = (e^{-t}p+I)(I,t)$.

Hence the subgroup $\overline{P}$ of $SO_o(1,4)$ generated by E_3 and time translations is the group consisting of the products $(p+R)(I,t)$ $p \in \mathbb{R}^3$, $R \in SO(3)$, and $t \in \mathbb{R}$. We denote the element $(p+R)(I,t)$ of $SO_o(1,4)$ by $(p+R,t)$. By (i) and (ii) one sees $(p+R,t)(q+Q,s) = (p+e^{-t}Rq + RQ, t+s)$. This defines multiplication in $\overline{P}$. $\overline{P}$ is a closed seven-dimensional analytic subgroup of $SO_o(1,4)$.

2. FREE PARTICLE SYSTEMS IN M

The states of a quantum mechanical system are the one-dimensional subspaces of a complex Hilbert space H; the observables are the self-adjoint operators, and more specifically the questions are the observables given by orthogonal projections. Invariance of physical laws under space–time symmetries is reflected in the existence of a projective unitary representation U of $SO_o(1,4)$. Two quantum mechanical systems (U_1,H_1) and (U_2,H_2) are physically equivalent if there is a unitary map V from H_1 onto H_2 with $VU_1(x) = U_2(x)V$ for x in $SO_o(1,4)$. The equivalence class obtained is a quantum mechanical physical system.

If U is to describe a free particle, to each Borel subset E of $\mathbb{R}^3$, there is a projection P_E; it is the question: Is the particle in E? The map $E \to P_E$ is a projection-valued measure. Furthermore, U and P are related by the following:

(i) $U_{q+R}P_E U_{q+R}^{-1} = P_{(q+R)E}$,

(ii) If A is a bounded operator on H and $AU_{p+R} = U_{p+R}A$, $AP_E = P_E A$ for all p,R and E, then $A = cI$, where c is a scalar.

The probability the particle lies in set E when the system is in state φ is $(P_E\varphi,\varphi)/\|\varphi\|^2$. Condition (i) reflects invariance of this probability under changes of coordinates while condition (ii) states that the pair $(U|_{E_3}, P)$ is irreducible. Essentially this means that there are not two or more noninteracting particles appearing as a one-particle system. Using (i) and (ii) to describe free particles was developed by Wightman in Ref. 3.

Let (U^1,P^1) and (U^2,P^2) be two pairs satisfying (i) and (ii) above. They are said to be equivalent if there is a unitary map V from H_1 onto H_2 with $VU_x^1 = U_x^2 V$ for $x \in SO_o(1,4)$ and $VP_E^1 = P_E^2 V$ for Borel subsets E in $\mathbb{R}^3$. Equivalence classes under this equivalence are the distinct free particles in M. In this equivalence, the positional observables are preserved. Physical equivalence defines an equivalence relation on free particles. The equivalence classes under this equivalence are the free particle physical systems alluded to earlier.

3. RESTRICTIONS OF $SO_o(1,4)$ COCYCLES TO $\overline{P}$

Let U be a Borel projective unitary representation of $SO_o(1,4)$ with cocycle σ. Hence $Uxy = \sigma(x,y)UxUy$ where $|\sigma(x,y)| \equiv 1$ and σ is Borel. By Theorem 5.5, p.34 of Ref. 4, σ may be assumed to be locally trivial, that is, $\sigma(x,y) = 1$ for all x,y in a neighborhood of the identity of $SO_o(1,4)$.

To analyze σ further, a concrete realization of the universal covering group $\tilde{P}$ of $\overline{P}$ is given. It is well known that $SU(2) \simeq S^3$ is simply connected and is a 2-covering of $SO(3)$. Let $\varphi : SU(2) \to SO(3)$ be the covering map. $\mathrm{Ker}\varphi = \{\pm I\}$, the center of $SU(2)$. Set $\tilde{P} = \mathbb{R}^3 \times SU(2) \times \mathbb{R}$. Define $\tilde{\varphi} : \tilde{P} \to \overline{P}$ by $\tilde{\varphi}(p,u,t) = (p+\varphi(u),t)$. Multiplication in $\tilde{P}$ is given by $(p,u,t)(p',u',t') = (p + e^{-t}\varphi(u)p', uu', t+t')$. $(\tilde{P},\tilde{\varphi})$ is the universal covering group of $\overline{P}$. It is a 2-covering of $\overline{P}$ and $\mathrm{Ker}\tilde{\varphi} = \{(0, \pm I, 0)\}$, the center of $\tilde{P}$.

Let B be a Borel cross section of $SU(2)/\{\pm I\}$. B may be taken so that B contains a neighborhood of the identity of I. Let $\gamma = \varphi|_B^{-1}$. Define a cocycle τ on $\overline{P}$ by

$$\tau((p_1+R_1,t_1),(p_2+R_2,t_2)) = \begin{cases} 1 \text{ if } \gamma(R_1)\gamma(R_2) \in B, \\ -1 \text{ if } \gamma(R_1)\gamma(R_2) \notin B. \end{cases}$$

τ is locally trivial for B is a neighborhood of the identity I in $SU(2)$.

Proposition: Let σ be a locally trivial Borel cocycle on $\overline{P}$. Then there is a Borel coboundary β on $\overline{P}$ such that either $\beta(x,y)\sigma(x,y) = \tau(x,y)$ on $\overline{P}\times\overline{P}$ or $\beta(x,y)\sigma(x,y) = 1$ on $\overline{P}\times\overline{P}$.

Remark: β is Borel coboundary if there is a Borel function $b : \overline{P} \to C$, $|b(x)| = 1$ for all x, and $\beta(x,y) = b(x,y)\overline{b(x)}\overline{b(y)}$.

Proof: Let U be an irreducible σ representation of $\overline{P}$. Choose a neighborhood N of 1 in $\tilde{P}$ such that $\tilde{\varphi}|_N$ is a homomorphism and $\sigma|_{\tilde{\varphi}(N)\times\tilde{\varphi}(N)} \equiv 1$. Define $\tilde{U}_x = U_{\tilde{\varphi}(x)}$ for $x \in N$. $\tilde{U}$ is a local homomorphism of $\tilde{P}$ and hence has an extension to $\tilde{P}$. Since U is irreducible, $\tilde{U}$ is irreducible. If H_1 is the Hilbert space for U and π is the canonical map of the unitary group $\mathcal{U}(H)$ of H onto the projective unitary group $\mathcal{U}_\pi(H) = \mathcal{U}(H)/\{cI : |c| = 1\}$, one has $\pi(\tilde{U}_x) = U_{\tilde{\varphi}(x)}$ for all x. Since $\tilde{U}$ and U are both strongly Borel, there exists a Borel function b on $\tilde{P}$ with $|b| \equiv 1$ and $\tilde{U}_x = b(x)U_{\tilde{\varphi}(x)}$.

Let $\tilde{B} = \mathbb{R}^3 \times B \times \mathbb{R}$. $\tilde{B}$ is a Borel cross section for $\tilde{P}/_{\mathrm{Ker}\tilde{\varphi}}$. Let $\tilde{\gamma} = (\tilde{\varphi}|_{\tilde{B}})^{-1}$. Hence $U_y = \overline{b(\tilde{\gamma}(y))}\tilde{U}_{\tilde{\gamma}(y)}$ for $y \in \overline{P}$. Since $\mathrm{Ker}\tilde{\varphi} = \{(0,\pm I,0)\} = \{\pm I\}$ belongs to the center of $\tilde{P}$ and $\tilde{U}$ is irreducible, $\tilde{U}_{-I} = \pm I$. Now $U_{y_1y_2} = \sigma(y_1,y_2)U_{y_1}U_{y_2}$. Hence

$$\sigma(y_1,y_2)\overline{b(\tilde{\gamma}(y_1))}\overline{b(\tilde{\gamma}(y_2))}\tilde{U}_{\tilde{\gamma}(y_1)}\tilde{U}_{\tilde{\gamma}(y_2)} = \overline{b(\tilde{\gamma}(y_1y_2))}\tilde{U}_{\tilde{\gamma}(y_1y_2)}.$$

Hence

$$b(\tilde{\gamma}(y_1y_2))\overline{b(\tilde{\gamma}(y_1))}\overline{b(\tilde{\gamma}(y_2))}\sigma(y_1y_2)I = \tilde{U}_{\tilde{\gamma}(y_1y_2)\tilde{\gamma}(y_2)^{-1}\tilde{\gamma}(y_1)^{-1}} = \begin{cases} U_I \text{ if } \tilde{\gamma}(y_1)\tilde{\gamma}(y_2) \in \tilde{B} \\ U_{-I} \text{ if } \tilde{\gamma}(y_1)\tilde{\gamma}(y_2) \notin \tilde{B}. \end{cases}$$

Hence $b(\tilde{\gamma}(y_1y_2))\overline{b(\tilde{\gamma}(y_1))b(\tilde{\gamma}(y_2))}\sigma(y_1,y_2)=1$ if $\tilde{U}_{-I}=I$ while $b(\tilde{\gamma}(y_1y_2))b(\tilde{\gamma}(y_1))b(\tilde{\gamma}(y_2))=\tau(y_1,y_2)$ if $\tilde{U}_{-I}=-I$. Q.E.D.

4. LOCALIZABLE REPRESENTATIONS

In Ref. 3 Wightmann calls pairs (U,P) which satisfy (i) and (ii) of Sec. II localizable. By using Mackey's imprimitivity theorem he showed $U|_{E_3}$ is induced from an irreducible cocycle representation of SO(3) and P is the canonical projection-valued measure associated with the induction.

Let us assume (U,P) satisfies (i) and (ii) of Sec. II. By Sec. III we may assume the cocycle σ is either 1 or τ. By (i) and (ii) $(U|_{E_3},P)$ form an irreducible system of imprimitivity for the action $(p+R)\cdot x=p+Rx$. Since this action is transitive, one may apply the foremen-tioned imprimitivity theorem, p. 291 of Ref. 5. It yields an irreducible σ representation L of the stabi-lizer SO(3) of the point 0 and a unitary map V from H to $L^2(\mathbb{R}^3,H_L)$ such that $VU_{p+R}V^{-1}=(\mathrm{ind}_{\mathrm{SO}(3)}^{E_3}L)_{p+R}$ and $(VP_EV^{-1})f=\chi_E\cdot f$ where χ_E is the characteristic function of E. But the irreducible projective representations of SO(3) are known to be parametrized by nonnegative half-integers j. The representation D^j of SO(3) corre-sponding to j operates on the finite dimensional space C^{2j+1}; it has cocycle $\sigma=1$ when j is an integer and cocy-cle $\sigma=\tau$ when j is a half-integer. Hence in the case when when $L=D^j$ one has $(\mathrm{ind}L)_{p+R}f(x)=D_R^jf(R^{-1}(x-p))$. j is called the spin of the localizable system.

Proposition: Let U be a σ representation of $\mathrm{SO}_0(1,4)$. Let P be a projection-valued measure defined on the Borel subsets of $\mathbb{R}^3$. Assume the pair (U,P) satisfies (i) and (ii) of Sec. II. Then there is a nonnegative half-integer j and a unitary operator V from H to $L^2(\mathbb{R}^3,C^{2j+1})$ such that $VU_{p+R}V^{-1}f(x)=D_R^jf(R^{-1}(x-p))$ and $VP_EV^{-1}f=\chi_Ef$.

Furthermore V is unique up to a scalar of modulus 1.

Proof: All that remains to be shown is the uniqueness of V. Assume $W:H\to L^2(\mathbb{R}^3,C^{2j+1})$ is unitary and $WU_{p+R}W^{-1}=VU_{p+R}V^{-1}$, $WP_EW^{-1}=VP_EV^{-1}$. Then $V^{-1}W$ commutes both with $U|_{E_3}$ and P. By (ii) $V^{-1}W=cI$ where $|c|=1$. Q.E.D.

5. EXTENSIONS OF $\mathrm{IND}_{\mathrm{SO}(3)}^{E_3}D^j$ TO $\bar{P}$

Our goal will be to show that for each representation $\mathrm{ind}_{\mathrm{SO}(3)}^{E_3}D^j$ there is, up to unitary equivalence, one and only one extension to $\bar{P}$.

Theorem: Let U be a projective unitary representa-tion of $\bar{P}$ with cocycle $\sigma=1$ or $\sigma=\tau$. Assume $(U_{p+R}f)(x)=D_R^jf(R^{-1}(x-p))$ for $f\in L^2(\mathbb{R}^3,C^{2j+1})$. Then there is a unitary map V of $L^2(\mathbb{R}^3,C^{2j+1})$ such that $VU_{(p+R,t)}V^{-1}f(x)=\exp(3/2t)D_R^jf(e^tR^{-1}(x-p))$.

Proof: σ is locally trivial. Hence by the same argu-ment as in Sec. III, there is an ordinary representation $\tilde{U}$ of $\tilde{P}$ such that $\tilde{U}_x=U_{\tilde{\varphi}(x)}$ in a neighborhood of the iden-tity. Since $\sigma|_{\mathrm{SO}(3)\times\mathrm{SO}(3)}$ is locally trivial one also can obtain an ordinary representation $\tilde{D}^j$ of SU(2) with $\tilde{D}_u^j=D_{\varphi(u)}^j$ for u near the identity. $\therefore\tilde{U}_{(p,u,0)}f(x)=\tilde{D}_u^jf(\varphi(u)^{-1}(x-p))$ for (p,u) near $(0,I)$. $\therefore\tilde{U}_{(p,u,0)}f(x)=\tilde{D}_u^jf(\varphi(u)^{-1}(x-p))$ for all $p\in\mathbb{R}^3$, $u\in\mathrm{SU}(2)$.

Let F be the Fourier transform of $L^2(\mathbb{R}^3,C^{2j+1})$. $\therefore Ff(x)=\int_{\mathbb{R}^3}e^{-2\pi ix\cdot y}f(y)dy$ when $f\in L^1(\mathbb{R}^3,C^{2j+1})\cap L^2(\mathbb{R}^3,C^{2j+1})$. Then $F\tilde{U}_{(p,u,0)}F^{-1}f(x)=e^{-2\pi ix\cdot p}\tilde{D}_u^jf(\varphi(u)^{-1}x)$. Let $W_{(p,u,t)}=F\tilde{U}_{(p,u,t)}F^{-1}$. Then $W_{(0,I,t)}W_{(p,I,0)}W_{(0,I,t)}^{-1}=W_{(0,I,t)(p,I,0)(0,I,-t)}=W_{(e^{-t}p,I,0)}$. Define for each t the unitary operator H_t of $L^2(\mathbb{R}^3,C^{2j+1})$ by $H_tf(x)=e^{-3t/2}f(e^{-t}x)$. Again $H_tW_{(p,I,0)}H_{-t}=W_{(e^{-t}p,I,0)}$. $\therefore H_t^{-1}W_{(0,I,t)}$ commutes with $W_{(p,I,0)}$ for all $p\in\mathbb{R}^3$.

Consider the projection-valued measure P defined on $\mathbb{R}^3$ by $P_Ef=\chi_Ef$. One knows these projections are the central projections in the commuting ring of $W|_{\mathbb{R}^3}$; see p. 18 of Ref. 6. Hence $H_t^{-1}W_{(0,I,t)}P_E=P_EH_t^{-1}W_{(0,I,t)}$. Therefore, $W_{(0,I,t)}P_EW_{(0,I,t)}^{-1}=H_tP_EH_t^{-1}=P_{e^tE}$. Define a left action of $\tilde{P}$ on $\mathbb{R}^3$ by $(p,u,t)\cdot x=e^t\varphi(u)x$. Then $W_xP_EW_x^{-1}=P_{x\cdot E}$ for all x in $\tilde{P}$. Since the action is transitive on $\mathbb{R}^3/\{0\}$, the pair (W,P) forms a transitive system of im-primitivity. By Mackey's imprimitivity theorem, there is a unitary representation L of dimension $2j+1$ for the stabilizer subgroup H of $(1,0,0)$ and a unitary equivalence S on $L^2(\mathbb{R}^3,C^{2j+1})$ such that $SW_xS^{-1}=(\mathrm{ind}_H^{\tilde{P}}L)_x$ and $SP_ES^{-1}=P_E$. Here we have used the natural identification of $\tilde{P}/H$ and $\mathbb{R}^3/\{0\}$. Clearly $H=\{(p,u,0):\varphi(u)(1,0,0)=1\}$. Let $T=\{u\in\mathrm{SU}(2):\varphi(u)\cdot(1,0,0)=(1,0,0)\}$. Let B be a Borel cross section of the left coset space $\mathrm{SU}(2)/T$. Define $\gamma:\mathbb{R}^3/\{0\}\to\tilde{P}$ by $\gamma(x)=(0,\alpha(x),\ln|x|)$ where $\alpha(x)$ is the unique element in B such that $\varphi(\alpha(x))(1,0,0)=x/|x|$. Then $(\mathrm{ind}L)_{(p,u,t)}f(x)=e^{-3/2t}L_{\gamma(x)^{-1}(p,u,t)\gamma(e^{-t}\varphi(u)^{-1}x)}\times f(e^{-t}\varphi(u)^{-1}x)$. But $\gamma(x)=(0,\alpha(x),\ln|x|)$, $\gamma(e^{-t}\varphi(u)^{-1}x)=(0,\alpha(\varphi(u)^{-1}x),-t+\ln|x|)$. Hence $\gamma(x)^{-1}(p,u,t)\times\gamma(e^{-t}\varphi(u)^{-1}x)=(|x|\varphi(\alpha(x))^{-1}p,\alpha^{-1}(x)u\alpha(\varphi(u)^{-1}x),0)$.

Since $SP_ES^{-1}=P_E$ for all Borel sets E in $\mathbb{R}^3$, there is a Borel, unitary matrix-valued function $x\mapsto S(x)$ such that $(Sf)(x)=S(x)f(x)$; see Theorem P6, p. 92 of Ref. 6.

$$\therefore SW_{(p,I,0)}S^{-1}f(x)=e^{-2\pi ix\cdot p}f(x)=L_{(|x|\varphi(\alpha(x))^{-1}p,I)}f(x).$$

Hence $e^{-2\pi ix\cdot p}=L_{(|x|\varphi(\alpha(x))^{-1}p,I,0)}$ a.e. x. Since both are continuous, they are equal everywhere. But $\exp(-2\pi ix\cdot p)=\exp[-2\pi i(x/|x|)\cdot|x|p]=\exp[-2\pi i(1,0,0)\cdot|x|\varphi(\alpha(x))^{-1}p]$. Hence $L_{(p,I,0)}=e^{-2\pi i(1,0,0)\cdot p}$.

Since $SW_{(0,u,0)}S^{-1}f(x)=S(x)\tilde{D}_u^jS(\varphi(u)^{-1}x)^{-1}f(\varphi(u)^{-1}x)=L_{(0,\alpha^{-1}(x)u\alpha(\varphi(u)^{-1}x))}f(\varphi(u)^{-1}x)$, one has $S(x)\tilde{D}_u^jS(\varphi(u)^{-1}x)^{-1}=L_{(0,\alpha^{-1}(x)u\alpha(\varphi(u)^{-1}x))}$ a.e. x for each $u\in\mathrm{SU}(2)$. Choose $x_0\neq0$ such that

$$M=\{u\in\mathrm{SU}(2):S(x_0)\tilde{D}_u^jS(\varphi(u)^{-1}x)^{-1}=L_{(0,\alpha^{-1}(x_0)u\alpha(\varphi(u)^{-1}x_0))}\}$$

is conull in SU(2). Let $T_0=\{u\in\mathrm{SU}(2):\varphi(u)x_0=x_0\}$. Let $u_0\in T_0$. Since M is conull there exists a $u\in\mathrm{SU}(2)$ such that $u_0^{-1}u$ and u both belong to M.

$$\therefore S(x_0)\tilde{D}_u^jS(\varphi(u)^{-1}x_0)^{-1}=L_{(0,\alpha^{-1}(x_0)u\alpha(\varphi(u)^{-1}x_0))}$$

and

$$S(x_0)\tilde{D}_{u_0^{-1}u}^jS(\varphi(u)^{-1}x_0)^{-1}=L_{(0,\alpha^{-1}(x_0)u_0^{-1}u\alpha(\varphi(u)^{-1}x_0))}.$$

$$\therefore S(x_0)\tilde{D}_{u_0}^jS(x_0)^{-1}=L_{(0,\alpha^{-1}(x_0)u_0\alpha(x_0))}$$

$$\therefore S(x_0)\tilde{D}_{\alpha(x_0)}^j\tilde{D}_{\alpha(x_0)^{-1}u_0\alpha(x_0)}^j(S(x_0)\tilde{D}_{\alpha(x_0)}^j)^{-1}=L_{(0,\alpha(x_0)^{-1}u_0\alpha(x_0))}.$$

But $\alpha(x_0)^{-1}T_0\alpha(x_0)=T$. Hence $S(x_0)\tilde{D}^j_{\alpha(x_0)}\tilde{D}^j_u(S(x_0)\tilde{D}^j_{\alpha(x_0)})^{-1}$ $=L_{(0,u)}$ for $u\in T$.

Consider the unitary map R of $L^2(\mathrm{R}^3,C^{2j+1})$ defined by $(Rf)(x)=\tilde{D}^j_{\alpha(x)\alpha(x_0)^{-1}}S(x_0)^{-1}f(x)$. Since

$$\tilde{D}_{\alpha(x_0)^{-1}}S(x_0)^{-1}L_{(p,u)}S(x_0)\tilde{D}_{\alpha(x_0)}$$
$$=e^{-2\pi ip\cdot(1,0,0)}\tilde{D}^j_u \quad \text{for } p\in\mathrm{R}^3,\ u\in T,$$

one has

$$R(\mathrm{ind}L)_{(p,u,t)}R^{-1}f(x)$$
$$=\exp(-3/2t)\tilde{D}^j_{\alpha(x)}\tilde{D}^j_{\alpha(x_0)^{-1}}S(x_0)^{-1}$$
$$\times L_{(|x|\varphi(\alpha(x))^{-1}p,\,\alpha^{-1}(x)u\alpha(\varphi(u)^{-1}x))}$$
$$\times S(x_0)\tilde{D}^j_{\alpha(x_0)}\tilde{D}^j_{\alpha(\varphi(u)^{-1}x)^{-1}}f(e^{-t}\varphi(u)^{-1}x)$$
$$=\exp(-3/2t)\exp[-2\pi i(1,0,0)\cdot|x|\varphi(x)^{-1}p]$$
$$\tilde{D}^j_{\alpha(x)}\tilde{D}^j_{\alpha(x)^{-1}u\alpha(\varphi(u)^{-1}x)}$$
$$\times\tilde{D}^j_{\alpha(\varphi(u)^{-1}x)^{-1}}f(e^{-t}\varphi(u)^{-1}x)$$
$$=\exp(-3/2t)e^{-2\pi ix\cdot p}\tilde{D}^j_u f(e^{-t}\varphi(u)^{-1}x).$$

Hence

$$(F^{-1}RSF)\tilde{U}_{(p,u,t)}(F^{-1}RSF)^{-1}f(x)$$
$$=\int e^{2\pi ix\cdot y}e^{-3t/2}e^{-2\pi iy\cdot p}$$
$$\times\tilde{D}^j_u Ff(e^{-t}\varphi(u)^{-1}y)dy$$
$$=\exp(3/2t)\tilde{D}^j_u$$
$$\times\int\exp[2\pi ie^t\varphi(u)^{-1}(x-p)\cdot y]Ff(y)dy$$
$$=\exp(3/2t)\tilde{D}^j_u f(e^t\varphi(u)^{-1}(x-p)).$$

But $\tilde{U}_x=U_{\tilde{\varphi}(x)}$ for x near 1 and $\tilde{D}^j_u=D^j_{\varphi(u)}$ for u near I. Hence if $V=F^{-1}RSF$, $VU_{\tilde{\varphi}(p,u,t)}V^{-1}f(x)=\exp(3/2t)$ $\times D^j_{\varphi(u)}f(e^t\varphi(u)^{-1}(x-p))$.

$\therefore VU_{(p+R,t)}V^{-1}f(x)=\exp(3/2t)D^j_R f(e^tR^{-1}(x-p))$ for $(p+R,t)$ near $(I,0)$. Since both sides are σ representations, equality holds for all $(p+R,t)$. Q.E.D.

6. FREE PARTICLES AND FREE PARTICLE PHYSICAL SYSTEMS IN M.

From the theorem in Sec. V, one sees that every equivalence class defining a free particle physical system of spin j contains a σ representation U on $L^2(\mathrm{R}^3,C^{2j+1})$ such that $U(p+R,t)f(x)=e^{3t/2}D^j_R f(e^tR^{-1}(x-p))$. Hence U is determined on the group $\bar{P}$. As $\bar{P}$ contains time translations U_t, the free Hamiltonian may be calculated. It depends only on the spin j and the inherent geometry of M. Other information must then be determined from the behavior of U on the remaining part of $\mathrm{SO}_0(1,4)$. Some examples of free particle physical systems are given by Hannubus in Ref. 1. They are those described by the principal unitary series $N^{\mu,j}$, $\mu\geq 0$, $j=0,\frac{1}{2},\cdots$ of $\mathrm{SO}_0(1,4)$. They are defind on $\bar{P}$ by $N^{\mu,j}{}_{(p+R,t)}f(x)=e^{3t/2}e^{-i\mu t}D^j_R f(e^tR^{-1}(x-p))$. Hannubus, by comparing with the Poincaré group, concluded μ is the mass of the free particle.

Free particles in M are equivalence classes of pairs (U,P) which satisfy (i) and (ii) of Sec. II. Equivalent pairs define the same free particle physical system. Here we shall work with the converse problem. Given a free particle physical system F_s, determine all the free particles which define F_s.

Hence let F_s be a free particle physical system of spin j. Choose a representation U in F_s such that $U_{(p+R,t)}f(x)=e^{3t/2}D^j_R f(e^tR^{-1}(x-p))$ for f in $L^2(\mathrm{R}^3,C^{2j+1})$. Let P be the projection valued measure defined by $P_E=\chi_E f$. Let $\mathcal{U}(U)(\mathcal{U}^e(U))$ be the group of unitary interwining operators for U on $\mathrm{SO}_0(1,4)$ (on E_3). Clearly $\mathcal{U}(U)\subset\mathcal{U}^e(U)$. Let $Q(U)$ be the left quotient space $\mathcal{U}^e(U)/\mathcal{U}(U)$. For $x\in Q(U)$ consider the pair $({}^xU,P)$ where ${}^xU_g=WU_gW^{-1}$, $W\in x$.

The pair $({}^xU,P)$ certainly satisfies (i) and (ii) and hence determines a free particle.

Theorem: Any pair (V,Q) defining the free particle physical system F_s is equivalent to $({}^xU,P)$ for some unique x. In particular the $({}^xU,P)$ are pairwise nonequivalent.

Proof: Since (V,Q) defines F_s, U is unitarily equivalent to V. Hence we may assume $U=V$. By the proposition of Sec. IV, there is a unitary operator T of $L^2(\mathrm{R}^3,C^{2j+1})$ such that $T^{-1}U_{p+R}Tf(x)=D^j_R f(R^{-1}(x-p))$ $=U_{p+R}f(x)$ and $TQ_ET^{-1}=P_E$. Hence $T\in\mathcal{U}^e(U)$ and if $x=T\mathcal{U}(U)$, (U,Q) is equivalent to $({}^xU,P)$. To show x is unique assume $({}^xU,P)$ is equivalent to $({}^yU,P)$. Since ${}^xU|_E={}^yU|_E=U|_E$ and (U,P) satisfies (ii) of Sec. II, ${}^yU$ $={}^xU$ by the proposition of Sec. IV. $\therefore y=x$. Q.E.D.

If F_s is irreducible, one sees the defining free particles are in one-to-one correspondence with the unitary intertwining group for $\mathrm{ind}^{E_3}_{\mathrm{SO}(3)}D^j$.

Proposition: If V is unitary and V intertwines $\mathrm{ind}^{E_3}_{\mathrm{SO}(3)}D^j$, then $V=F^{-1}MF$ where F is the Fourier transform and $(Mf)(x)=M(x)f(x)$, M being a Borel, unitary matrix-valued function satisfying $M(Rx)$ $=D^j_RM(x)D^{j-1}_R$. Such functions M are in one-to-one correspondence with the Borel, diagonal, unitary matrix-valued functions on the nonnegative real line.

Proof: Let $U=F\,\mathrm{ind}D^jF^{-1}$. Then $(U_{p+R}f)(x)$ $=e^{-2\pi ix\cdot p}D^j_R f(R^{-1}x)$. Now FVF^{-1} commutes with $U|_{\mathrm{R}^3}$. Hence it commutes with the central projections P_E where $P_Ef=\chi_E\cdot f$. Hence $FVF^{-1}=M$ where $Mf=m(x)f(x)$ and $x\mapsto m(x)$ is a Borel, unitary matrix-valued function; see p. 92 of Ref. 6. Since $U_RM=MU_R$, $D^j_Rm(R^{-1}x)$ $=m(x)D^j_R$ a.e. x for each R. Therefore, $m(Rx)$ $=D^j_Rm(x)D^{j-1}_R$ a.e. R for almost all x. Let μ be the Haar measure on SO(3), $|\mu|=1$. Then $M(x)$ $=\int D^{j-1}_Rm(Rx)D^j_Rd\mu(R)$ is a Borel, unitary matrix-valued function; $M(Rx)=D^j_RM(x)D^{j-1}_R$, and $M=m$ a.e. Hence $V=F^{-1}MF$.

If $M(Rx)=D^j_RM(x)D^j_{R^{-1}}$, $M(x,0,0),x>0$ belongs to the commuting ring of $D^j|_T$; $T=\{R:R(1,0,0)=(1,0,0)\}$. T is commutative and $D^j|_T$ is a diagonal direct sum of distinct characters of T; see Ref. 7, Chapter 2. Hence $M(x,0,0)$ is diagonal. Q.E.D.

Consider a representation U of $\mathrm{SO}_0(1,4)$ with $U_{(p+R,t)}f(x)=e^{3t/2}D^j_Rf(e^tR^{-1}(x-p))$. Let $P_Ef=\chi_Ef$. Set $V=F^{-1}MF$ where $Mf(x)=|x|^{i\mu}f(x)$. A straightforward calculation shows $VU_{(p+R,t)}V^{-1}f(x)=e^{3t/2}e^{i\mu t}D^j_R$

$\times f(e^t R^{-1}(x-p))$. Hence V does not commute with U on $SO_o(1,4)$. Therefore, (U,P) and (VUV^{-1},P) define distinct free particles in M, however they define the same physical systems. What the mass is for such a system is not clear.

CONCLUSION

To find the free particles of spin j in imaginary Lobachevsky space, it suffices to find all the projective unitary representations of $SO_o(1,4)$ whose restriction to $\overline{P}$ have form $U_{(p\to R,t)}f(x)=e^{3t/2}D^j_R f(e^tR^{-1}(x-p))$. Having these will determine all free particle physical systems of spin j, and if one desires the positional observables be preserved, all free particle systems.

APPENDIX

Let G be locally compact, separable group. A projective unitary representation of G is a map $U:G\to \mathcal{U}(H)$ where H is a Hilbert space and $\mathcal{U}(H)$ is the group of unitary operators on H such that $U_e=I$, $U_{xy}=\sigma(x,y)U_xU_y$, and $x\to(U_x\varphi,\psi)$ is Borel measurable for all φ, $\psi\in H$. U is a unitary representation provided $\sigma\equiv 1$. σ is called the cocycle of U.

Let K be a closed subgroup of G. Then G/K, the left cosets of K in G, form a left Borel G-space. Indeed, let $\pi:G\to G/K$ be the map $\pi(g)=gK$. Then the smallest σ algebra on G/K making π Borel makes the map $x,gK\to xgK$ Borel. Let σ be a Borel cocycle on G. Let L be a projective representation on K with cocycle σ. There is a measure μ on G/K such that for each $g\in G$ there is a Borel function $\rho_g>0$ on G/K such that $\mu(g^{-1}E)=\int_E \rho_g(xK)d\mu(gK)$. There is also a Borel map $\gamma:G/K\to G$ such that $\pi(\gamma(gK))=gK$. A map with this property is called a Borel cross section. In specific cases both μ and γ may be determined from the structure of G and K. We now define $\mathrm{ind}_K^G L$. Let H be the Hilbert space for L. Then the Hilbert space H^L for $\mathrm{ind}L$ is $H^L=\{f\,|\,f:G/K\to H, xK\to f(xK),\varphi)$ is Borel for all φ, and $\int\|f(xK)\|^2 d\mu(xK)<\infty\}$. For $g\in G$, $(\mathrm{ind}_K^G L)(g)f(xK)=\rho_g(xK)^{1/2}\sigma(\gamma(xK),\gamma(xK))\overline{\sigma}(\gamma(xK)^{-1},g)\overline{\sigma}(\gamma(xK)^{-1}g,\gamma(g^{-1}xK))\times L_{\gamma(xK)^{-1}g\gamma(g^{-1}xK)}f(g^{-1}xK)$. The term involving ρ_g is the weight necessary to make the operator unitary, the terms involving σ and the fact that L is a σ representation force $\mathrm{ind}L$ to be a σ representation.

A projection valued measure on G/K is a map $E\to P(E)$ from the Borel subsets E of G/K to the orthogonal projections on a Hilbert space H such that $P(\emptyset)=0$, $P(G/K)=I$; $P(E\cap F)=P(E)P(F)$; and $P(\cup_{i=1}^{\infty}E_i)=\sum_{i=1}^{\infty}P(E_i)$ if $E_i\cap E_j=\emptyset$ for $i\neq j$. The canonical projection-valued measure associated with $\mathrm{ind}_K^G L$ is the projection-valued measure defined on G/K by $(P(E)f)(xK)=\chi_E(xK)f(xK)$ for $f\in H^L$ where $\chi_E(xK)$ is 1 or 0 according to whether or not xK is in E.

A central projection for a representation U is an orthogonal projection which commutes with all the bounded operators commuting with U.

For more information concerning group representations and particularly induced representations, we mention Refs. 5, 6, 8, 9, and 10.

[1] K. C. Hannabus, Proc. Camb. Philos. Soc. 70, 283–302 (1971).
[2] C. Fronsdal, Phys. Rev. D 10, 589–98 (1974).
[3] A. S. Wightman, Rev. Mod. Phys. 34, 845–72 (1962).
[4] K. R. Parthasarathy, *Multipliers on Locally Compact Groups* (Heidelberg, Berlin, Springer, New York, 1969).
[5] G. W. Mackey, Acta. Math. 99, 265–311 (1958).
[6] G. W. Mackey, "The Theory of Group Representations," notes by Dr. Fell and Dr. Lowdenslager, Department of Mathematics, University of Chicago, Summer 1955.
[7] N. J. Vilenkin, *Special Functions and the Theory of Group Representations* (Am. Math. Soc., Providence, Rhode Island, 1968).
[8] G. W. Mackey, Ann. Math. 2, 101–39 (1952).
[9] G. W. Mackey, *Induced Representations of Groups and Quantum Mechanics* (Benjamin, New York, 1968), 1st ed.
[10] A. Ramsay, Adv. Math. 6, 253–322 (1971).

A tachyon dust universe

Sushil K. Srivastava

Department of Mathematics, University of Gorakhpur, Gorakhpur-273001, India
(Received 26 January 1977)

In the present paper some investigations have been made on the model suggested by Ray for a tachyon dust universe and the results obtained have been compared with the results in the flat Friedmann universe filled with ordinary dust (here called bradyon dust) moving slower than light, by various scientists, on the ground that the role played by time for ordinary matter is played by spatial coordinates for tachyons. The effect of the cosmological constant (Λ) on the expanding tachyon universe also has been discussed here. The tetrad technique has been used as a mathematical tool for handling the problems of gravitational field equations and perturbation of momentum flux.

1. INTRODUCTION

A. Background

Many surprising astronomical discoveries have been made during the past several years. Many unusual physical phenomena happened in the universe that were guessed previously by theory and these phenomena have been explained by putting the general theory of relativity into service. For instance, the expansion of the universe neutron stars, 3 °K microwave background radiation, cosmic rays, pulsars, supernovae, quasars, and exploding galactic nuclei are some examples. It is with this background that we turn to a study of tachyon cosmologies. It is possible that the universe with all its diversity actually contains tachyonic matter as an important constituent.

In the present paper, we start with the spacelike line element proposed by Ray[1] which is based on Petrov classification[2] of gravitational fields by the symmetry of the space–time, regardless of the weak point that it allows dust solutions only. The energy–momentum tensor for dust is given by $T_{ij}=\rho u_i u_j$ where ρ is the momentum–flux of the tachyon dust and u_i the 4-velocity. For bradyons, ρ is the energy density.

In Sec. 1 B we give a brief note on comparison of the tachyons with ordinary matter, called bradyons here. In Sec. 1 C we describe the tetrad notation used in the paper which has been earlier used by Srivastava and Pathak.[3]

In Sec. 2, which is composed of four subsections, we deal with the theory. In Sec. 2 A we derive tetrad field equations and their solutions for tachyon dust. In Sec. 2 B we report the results obtained by Johri *et al.* in the flat Friedmann universe filled with dust like bradyons for the sake of comparison with our model. In Sec. 2 C we apply perturbation theory on those field equations. In the succeeding section also we report the results obtained for flat Friedmann universe after a slight perturbation for the same purpose.

In the concluding section we discuss the results obtained in the theory section. Here I have stressed the similarity between a tachyon dust universe and a flat Friedmann universe filled with bradyon dust.

B. A brief account on comparison of tachyons and bradyons

For our purpose let us choose a particular inertial frame s_0. The light speed c owing to its invariant character allows an exhaustive partition of frames $f\in\{I\}$ in two subclasses $\{s\}$, $\{S\}$ of frames having speeds $u<c$ and $U>c$ relative to s_0, respectively. For simplicity in the following we consider ourselves as "the observer s_0." Frames $s\in\{s\}$ will be called subluminal and frames $s\in\{S\}$ superluminal. The relative speed of two frames s_1, s_2 (or S_1, S_2) will always be smaller than c and the relative speed of two frames $\{s\},\{S\}$ will always be greater than c. The important point is that the above exhaustive partition is invariant when s_0 is made to vary inside $\{s\}$ (or inside $\{S\}$). On the contrary, when we pass from $s_0\in\{s\}$ to a frame $S_0\in\{S\}$, the subclasses $\{s\},\{S\}$ are interchanged one with other. One confusion may arise here on the physical ground, "how can an observer pass from subclass $\{s\}$ to subclass $\{S\}$." We do not actually mean that an observer of class $\{s\}$ goes to class $\{S\}$, but we mean that when we consider the observer of the class $\{s\}$ and afterwards the observer of class $\{S\}$, both are not the same observers.

Further it is well known that the linear transformations L, making the transition between two inertial frames f_1, f_2 must be such that

$$x_0'^2-\bar{x}'^2=\pm\,(x_0^2-\bar{x}^2)$$

for every 4-vector $x=(x_0,\bar{x})$ where x means either 4-position or 4-momentum or 4-velocity or 4-current density and so on. In particular

$$c^2t'^2-\bar{x}^2=\pm(c^2t^2-\bar{x}^2)$$

or

$$c^2t'^2+(i\bar{x}')^2=\pm(c^2t^2+(i\bar{x})^2).$$

For the physical validity it follows that objects must exist, which are at rest relative to S and tachyons relative to frames s. From the further fact that luxons l show the same velocity to any observer s or S it can be deduced that a bradyon relative to an S, $B(S)$ will be a tachyon relative to any s, $T(s)$ and vice versa,[4]

$$B(S)=T(s),\quad T(S)=B(s),\quad l(s)=l(S).$$

Now it is concluded that when frames s, S observe the same event "timelike" vectors transform into

 Copyright © 1977 American Institute of Physics

spacelike vectors and vice versa in going from s to S or from S to s. That is to say that the role of spacelike coordinates and timelike coordinates are interchanged while considering tachyons and bradyons. Whatever role the time coordinate plays for bradyons is interchanged with spatial coordinates for tachyons, and vice versa.

C. Notation

In this paper, space–time is represented as a four-dimensional Riemannian space with metric tensor g_{ij} of signature $(+,+,+,-)$. Covariant differentiation is indicated by a semicolon (;) and covariant differentiation along the lines (x^1, x^2, t), constant t, is indicated by a prime over the variable, i.e., prime denotes $\partial/\partial x^3$. Round brackets around the indices indicate symmetrization and square brackets antisymmetrization. Here we have taken $8\pi G = c^4 = 1$.

The Einstein's field equations for dust filled cosmological models are

$$R_{ab} - \tfrac{1}{2}Rg_{ab} + \Lambda g_{ab} = T_{ab} = \rho u_a u_b, \tag{1.1}$$

where u^a are the spacelike 4-velocities for the tachyon fluid so that $u^a u_a = 1$.

The acceleration of the fluid is

$$\dot{u}_i = u_{i;j}u^j, \quad \text{where the dot denotes } \partial/\partial t. \tag{1.2}$$

The velocity gradient may be further split up as

$$U_{i\ ;j} = W_{ij} + \sigma_{ij} + \tfrac{1}{3}\theta H_{ij} - \dot{U}_i U_j,$$

where $\theta = U^i_{;i}$ is the expansion scalar. $\sigma_{ij} = U_{(i;j)} + U'_{(i}U_{j)} - \frac{1}{3}\theta H_{ij}$ is the trace-free shear tensor. $W_{ij} = U_{i;j} + \dot{U}_i U_j$ is the vorticity tensor. Here H_{ij} is a tensor which projects a quantity from $x^3 = \text{constant}$ to $(x^1, x^2, t) = \text{const}$ defined by

$$H_{ij} = g_{ij} - U_i U_j,$$

or (1.3)

$$H_{ij}U^j = 0, \quad H^i_i = 3.$$

Here g_{ij} and U_i have their previous meaning.

The Ricci rotation coefficients are defined by

$$\Gamma_{abc} = e_a \nabla_b e_c = e^i_a e^c_{i;j} e^j_b$$

so that

$$\Gamma_{abc} + \Gamma_{cba} = 0.$$

Here e_a are four orthonormal vectors hereafter called tetrads of vectors which do not, in general, always remain the same.

The Lie derivative of e_b with respect to e_a is

$$\langle e_a, e_b \rangle = \gamma^c_{ab} e_c, \quad \gamma^c_{ab} = \gamma^c_{[ab]}.$$

It follows that γ^c_{ab} and Γ^c_{ab} are linearly dependent,

$$\gamma^c_{ab} = \Gamma^c_{ab} - \Gamma^c_{ba},$$

$$\Gamma_{abc} = \tfrac{1}{2}(\gamma_{abc} + \gamma_{cab} - \gamma_{bca}).$$

Now the Einstein field equation (1.1) can be written down in the tetrad form as

$$R_{bd} = \partial_d \Gamma^c_{ab} - \partial_c \Gamma^c_{ab} - \Gamma^c_{cg}\Gamma^g_{db} + \Gamma^g_{cb}\Gamma^c_{gd}$$
$$= -(\Lambda - \rho/2)H_{bd} - (\Lambda + \rho/2)U_b U_d. \tag{1.4}$$

The antisymmetry property of the curvature tensor is equivalent to the Jacobi identity

$$\partial_{[d}\gamma^f_{cb]} + \gamma^g_{[dc}\gamma^f_{b]g} = 0. \tag{1.5}$$

The tetrads are so chosen that the spacelike vector e_3 is the tachyon fluid flow vector U^3 so that

$$U^a = \delta^a_3, \quad U_a = \delta^3_a.$$

In a cosmological model filled with pressure-free tachyon fluid, the lines of flow are spacelike geodesics and the contracted Bianchi identities are

$$\rho' + \rho\theta = 0, \quad \text{where } \theta = \theta_1 + \theta_2 + \theta_4. \tag{1.6}$$

Suppose the perturbation of the model results in the formation of momentum flux $\rho + \delta\rho$ so that the ratio of increase in momentum flux to the model, the contrast momentum flux for tachyons is $K = \delta\rho/\rho$, the contrast energy density for bradyons is $\mu = \delta\rho/\rho$, and the relative expansion in this region is $-\delta\theta$.

Perturbation of (1.6) gives

$$\partial_3(\delta\rho) + \theta\delta\rho + \rho\delta\theta = 0. \tag{1.7}$$

Therefore,

$$\partial_3\left(\frac{\delta\rho}{\rho}\right) = \frac{\partial_3(\delta\rho)}{\rho} - \frac{\delta\rho}{\rho^2}\rho' = -\delta\theta. \tag{1.8}$$

This gives the ratio of growth of K with respect to x^3 in the condensation.

2. THEORY

A. Momentum flux of the tachyon dust universe

Let us consider the metric[1]

$$ds^2 = A^2(dx^1)^2 + C^2\exp[x^1](dx^2)^2 + (dx^3)^2 - C^2\exp[x^1](dx^4)^2 \tag{2.1}$$

as the metric for the background model.

The nonvanishing tetrad components corresponding to the components of the fundamental tensor in the line element (2.1) are given by

$$(e^1_1)_{t=0} = 1/A,$$
$$(e^2_2)_{t=0} = (e^4_4)_{t=0} = \frac{1}{C}\exp\left[-\frac{x'}{2}\right].$$

The components of γ^a_{bc} are given as

$$[\theta_1]_{t=0} = -\gamma^1_{\cdot 31} = A'/A, \tag{2.2}$$

$$[\theta_2]_{t=0} = -\gamma^2_{\cdot 32} = C'/C, \tag{2.3}$$

$$[\theta_4]_{t=0} = -\gamma^4_{\cdot 34} = C'/C, \tag{2.4}$$

and other components of $\gamma^a_{\cdot bc}$ vanish.

The tetrad field equations (1.4) for the tachyon dust model are

$$\begin{aligned} D\theta_1 + \theta_1(\theta_1 + \theta_2 + \theta_4) &= -\Lambda + \rho/2, \\ D\theta_2 + \theta_2(\theta_1 + \theta_2 + \theta_4) &= -\Lambda + \rho/2, \\ D(\theta_1 + \theta_2 + \theta_4) + (\theta_1^2 + \theta_2^2 + \theta_4^2) &= -\Lambda - \rho/2, \end{aligned} \tag{2.5}$$

$$D\theta_4 + \theta_4(\theta_1 + \theta_2 + \theta_4) = -\Lambda + \rho/2.$$

Here $D = \partial/\partial x^3$.

Computing the components T_{13} and T_{31} of the energy–momentum tensor for this model with the help of (1.1), we find that

$$T_{13} = T_{31} = \frac{1}{2}\left(\frac{C'}{C} - \frac{A'}{A}\right) = 0,$$

from which it immediately follows that

$$A = C, \text{ apart from a constant.} \tag{2.6}$$

Now (2.2), (2.3), (2.4), and (2.6) together yield

$$[\theta_1]_{t\neq 0} = [\theta_2]_{t\neq 0} = [\theta_4]_{t\neq 0} = A'/A = \theta_0 \text{ (say).} \tag{2.7}$$

Hence Eqs. (2.5) and (2.7) give

$$\begin{aligned} D\theta_0 + 3\theta_0^2 &= -\Lambda + \rho/2, \\ 3D\theta_0 + 3\theta_0^2 &= -\Lambda - \rho/2. \end{aligned} \tag{2.8}$$

These equations further yield

$$2D\theta_0 + 3\theta_0^2 = -\Lambda. \tag{2.9}$$

Case I: when $\Lambda = 0$

In this case (2.9) is reduced to

$$2D\theta_0 + 3\theta_0^2 = 0 \tag{2.10}$$

which is easily integrable, giving

$$A = G(x^3)^{2/3} + B,$$

where G and B are integration constants and

$$\theta_0 = \frac{2}{3(x^3)}. \tag{2.11}$$

This leads to the momentum flux variation relation for tachyons

$$\rho = \frac{4}{3(x^3)^2} \tag{2.12}$$

in the background model.

Case II: when $\Lambda \neq 0$

In this case the solution of Eq. (2.9) is given by

$$A^{3/2} = \sqrt{3/\Lambda}\sin\psi \quad \text{and} \quad \theta_0 = \sqrt{\Lambda/3}\cot\psi, \tag{2.13}$$

where

$$\psi = \frac{x^3\sqrt{3\Lambda}}{2} + \eta. \tag{2.14}$$

Here ψ is the measure of x^3, the proper distance for tachyons playing the same role as the proper time for bradyons. η is an integration constant.

In this case we find by solving Eq. (2.8) with the help of (2.13) that the momentum flux variation relation for the model is

$$\rho = \Lambda\,\mathrm{cosec}^2\psi. \tag{2.15}$$

B. Energy density variation relation for dustlike bradyons in the flat Friedmann universe

Case I: when $\Lambda = 0$

In this case the energy density ρ for the model is given by[5]

$$\rho = 4/3t^2. \tag{2.16}$$

Case II: when $\Lambda \neq 0$

In this case ρ is given by[6]

$$\rho = \Lambda\,\mathrm{cosech}^2(\sqrt{3\Lambda/4})t. \tag{2.17}$$

C. Perturbations of momentum flux in the proposed tachyon dust model

Dust is characterized as pressure-free fluid hence we can take $p = 0$. Moreover for the sake of simplicity we consider σ, w also vanishing. Hence Raychaudhuri's equation for tachyons,[3]

$$\theta' + \tfrac{1}{3}\theta^2 + 2(\sigma^2 - w^2) + \tfrac{1}{2}(\underset{\sim}{\mu} - 3p) + \Lambda = 0,$$

where $\underset{\sim}{\mu}$ is the total relativistic energy of matter measured by u^a given by the relation $\underset{\sim}{\mu} = \rho(1+\epsilon)$ where ϵ is the specific internal momentum flux of the tachyon fluid, is reduced to

$$\theta' + \tfrac{1}{3}\theta^2 + \tfrac{1}{2}\rho + \Lambda = 0. \tag{2.18}$$

After a slight perturbation, we have from this equation

$$D^2\theta + 2\theta_0 D\theta + \tfrac{1}{2}D\rho = 0. \tag{2.19}$$

Now with the help of (1.8), (2.19) reduces to

$$D^2K + \frac{2A'}{A}DK - \tfrac{1}{2}\rho K = 0. \tag{2.20}$$

Case I: when $\Lambda = 0$

Equations (2.11), (2.12), and (2.20) yield

$$D^2K + \frac{4}{3(x^3)}DK - \frac{2}{3(x^3)^2}K = 0$$

which gives an integration

$$K = E_1(x^3)^{2/3} + E_2(x^3)^{-1}, \tag{2.21}$$

where E_1 and E_2 are constants.

Case II: when $\Lambda \neq 0$

In this case Eqs. (2.13), (2.15), and (2.20) imply

$$D^2K + 2\sqrt{\Lambda/3}\cot\psi DK - (\Lambda/2)\,\mathrm{cosec}^2\psi K = 0. \tag{2.22}$$

Now changing the independent variable x^3 by ψ through the transformation $\psi = x^3\sqrt{3\Lambda}/2 + D$, we have Eq. (2.22) in the form

$$\frac{d^2K}{d\psi^2} + \tfrac{4}{3}\cot\psi\frac{dK}{d\psi} - \tfrac{2}{3}\mathrm{cosec}^2 K = 0. \tag{2.23}$$

On solving this equation we find two solutions

$$K_1 = F_1\cot\psi \tag{2.24}$$

and

$$K_2 = F_2\cot\psi\int_\beta^\psi \sin^{2/3}\psi\sec^2\psi\,d\psi, \tag{2.25}$$

where

$$\psi > \beta > 0.$$

D. Energy density perturbation for bradyons in the flat Friedmann universe

Case I: when $\Lambda = 0$

In this case after slight perturbation of the energy density, contrast density μ $(=\delta\rho/\rho)$ obtained by solving the differential equation[5]

$$\frac{d^2\mu}{dt^2}+\frac{4}{3t}\frac{d\mu}{dt}-\frac{2\mu}{3t^2}=0$$

as

$$\mu=B_1t^{2/3}+B_2t^{-1}. \tag{2.26}$$

Case II: when $\Lambda\neq 0$

In this case after perturbation of Raychaudhuri's field equation[7] the differential equation is obtained in the form[6]

$$\frac{d^2\mu}{dt^2}+\left(\frac{4\Lambda}{3}\right)^{1/2}\left[\coth\left(\frac{3\Lambda}{4}\right)^{1/2}t\right]\frac{d\mu}{dt}$$
$$-\frac{\Lambda}{2}\left[\operatorname{cosech}^2\left(\frac{3\Lambda}{4}\right)^{1/2}t\right]\mu=0,$$

having its solution as

$$\mu_1=F_1'\coth\tau, \tag{2.27}$$

$$\mu_2=F_2'\coth\tau\int_\gamma^\tau\sinh^{2/3}\tau\operatorname{sech}^2\tau\,d\tau, \tag{2.28}$$

where $\tau>\gamma>0$ and $\tau=t\sqrt{3\Lambda}/4$.

3. DISCUSSION

Equation (2.12) shows that, in the case of Λ vanishing, as x^3 increases the momentum flux of the model decreases and also space–time is singular at $x^3=0$. Applying a different method Ray has also derived a momentum flux variation relation for tachyons in this model when $\Lambda=0$, which is given as $\rho=(3/4B_0^2)\times(\sin\frac{1}{4}x^3)^2$ where B_0 is a constant. This relation can be interpreted to mean that space–time has a singularity at $x^3=0$ but for the maximum value of $\sin\frac{1}{4}x^3$, $\rho=3/4B_0^2$ which is a constant and minimum value of ρ, while our result yields that ρ would be zero when x^3 tends to infinity, i.e., the minimum value of ρ would be zero. Now from the comparison of the two results we find that the idea of infinite distance for tachyons carries no physical meaning and it is reasonable for a particle moving with such a high velocity. The analogous result for bradyons is given by Eq. (2.16) where there is an interchange of x^3 into t.

Equation (2.15) implies that in the case of nonvanishing Λ, momentum flux for tachyons decreases with the increase in ψ to its minimum value $\rho=\Lambda$ which is invariant and also space–time is singular at $\psi=0$. The corresponding relation for bradyons is given by (2.17) which is hyperbolic in nature showing a little difference between a tachyon dust universe and a bradyon dust universe.

One thing more is notable here; the nature of the momentum flux relations for tachyons in the case of Λ vanishing and in the case of nonvanishing Λ is different. If physical considerations are not seriously taken, according to the solution (2.12) the minimum value of ρ is zero in the first case while the minimum value of ρ is Λ in the second case. But near the singular point $x^3=0$, momentum flux becomes infinite in both cases. This implies that Λ has its effects at larger scales only.

To make the situation clearer, let us assume that the tachyons were created at or after a big bang along with many strange particles but they being superluminal soon went out of the boundary of bradyons due to repulsion between bradyons and tachyons,[4,8,9] where every physical phenomena is spacelike for a subluminal observer. This caused the continuous decrease of the momentum flux for tachyons and this decrease continued up to a fixed value of ρ and vanished afterwards with the result of expansion of the tachyon universe.

Another notable point here lies in the fact that Λ is more effective in the case of tachyons than in the case of bradyons. As it is evident from (2.16) and (2.17), up to the remote future the energy density of bradyons tends to zero in both cases (when $\Lambda=0$ and when $\Lambda\neq 0$) but in the case of tachyons it is not so.

Let us further assume that there exists a region at some distance from the point source of tachyons where its momentum flux is $\rho+\delta\rho$ against ρ. Now for further discussion we first take up the simpler case (when $\Lambda=0$). In this case we find that two relations for contrast momentum flux K $(=\delta\rho/\rho)$ for tachyons as $K_1\sim(x^3)^{2/3}$ and $K_2\sim(x^3)^{-1}$. $K_1\sim(x^3)^{2/3}$ shows that condensation of tachyons increases algebrically as $(x^3)^{2/3}$ with increase in x^3 which would create meta galaxies. But the solution $K_2\sim(x^3)^{-1}$ shows that the contrast momentum flux decreases with increasing x^3 meaning thereby that no condensation among tachyons would take place. But this goes against the fact that tachyons attracts tachyons. Now the question arises, "How is it proved that tachyons attract tachyons?" For answering this question we would have to go back to velocity addition formulas in the special theory of relativity and from those formulas we find that a tachyonic observer observes a tachyon moving with a subluminal velocity, not with a superluminal velocity. Therefore, we are able to say that tachyons are subluminal with respect to tachyons and superluminal with respect to bradyons. This fact has also been stressed in Sec. 1 B. On this ground there is no harm in taking the idea that tachyons attract tachyons because classification of the fundamental particles bradyons, tachyons, and luxons is primarily based on velocity consideration. It can also be proved by some analytical methods as considering spacelike geodesics, etc. It is why we discard the solution $K_2\sim(x^3)^{-1}$. The similar case happens for bradyons as it has been discussed by Johri[5] through Eq. (2.28).

In the case $\Lambda\neq 0$, from solution (2.24) we find that when $\psi\to 0$, $K\to\infty$ and when $\psi\to\pi/2$, $K\to 0$, and as ψ increases K decreases. It means that no condensation would occur. Hence this solution is also not of interest on the grounds mentioned above. But solution (2.25) follows in that, as ψ increases from a fixed value $\beta>0$, K increases which means that condensation takes place resulting in the formation of meta galaxies in this case also. The cosmical constant Λ is found effective here also if we compare the results in the two cases (when $\Lambda=0$ and when $\Lambda\neq 0$). In the first case contrast momentum flux of tachyon universe increases algebrically as $(x^3)^{2/3}$ which is faster than that in the second case where contrast momentum flux increases with the increase in

ψ (the measure of x^3). The parallel results also hold in the flat Friedmann universe filled with bradyon dust as it has been discussed by Pathak[6] through solutions (2.27) and (2.28).

Thus from the above discussions we find that a tachyon dust universe also expands as a bradyon dust universe and there is the possibility of the formation of galaxies as there is for the bradyon universe. Moreover, it is also found here that the cosmical constant Λ has similar effects in the tachyon universe at larger scales but it is somewhere still more effective.

ACKNOWLEDGMENTS

The author is thankful to Dr. V. B. Johri, Mr. Ramesh Chandra and Mr. Ravi P. Sinha for useful discussions and helpful suggestions. The financial assistance provided by UGC for the work is highly acknowledged.

[1]J.R. Ray, Lett. Nuovo Cimento 12, 249 (1975).
[2]A. Petrov, *Einstein Spaces* (Pergamon, New York, 1969).
[3]Sushil K. Srivastava and Mahendra P. Pathak, J. Math. Phys. 18, 483 (1977).
[4]E. Recami and R. Mignami, Nuovo Cimento 4, 209 (1974).
[5]V.B. Johri, Prog. Math. 3, 42 (1969).
[6]M.P. Pathak, thesis, Gorakhpur University, Gorakhpur, India (1975).
[7]A.K. Raychaudhuri, Phys. Rev. 98, 1123 (1955).
[8]A.K. Raychaudhuri, J. Math. Phys. 15, 856 (1974).
[9]P.C. Vaidya, Curr. Sci. 40, 651–3 (1974).

K matrix for the Woods–Saxon potential*

B. Talukdar and M. N. Sinha Ray

Department of Physics, Visva-Bharati University, Santiniketan 731235, West Bengal, India
(Received 13 December 1976)

The s-wave part of the off-shell K matrix elements for the Woods–Saxon potential has been obtained in terms of elementary transcendental functions by using the differential equation approach to off-shell scattering.

The purpose of the present note is to obtain the s-wave part of the off-shell two-body K matrix for the Woods–Saxon potential in close anlaogy to our recent work on the T matrix[1] (cited as paper I hereafter). Recent interest in the K matrix calculation has been stimulated because of its potential application to nuclear scattering reactions.[2]

The differential equation approach to off-shell scattering can be adapted to the K matrix calculation by imposing standing wave boundary conditions on the solution of the van Leeuwen and Reiner equation. For a central potential the relevant off-shell wavefunction regular at the origin is given by[3]

$$\phi(k,q,r)=-\frac{1}{4\pi}q\langle k|K(E)|q\rangle[f(k,r)+f(-k,r)] + \frac{1}{2i}[f(k,q,r)-f(k,-q,r)], \tag{1}$$

where $\langle k|K(E)|q\rangle$ represents the half off-shell K matrix written as

$$K(k,q,s)=\frac{2\,\mathrm{Im}f(k,v)}{\pi q|f(k)|\cos\delta(k)}\,. \tag{2}$$

In Eqs. (1) and (2) k is an on-shell momentum related to the energy by $s=k^2+i\epsilon=E$ with $\epsilon\ll 1$ and q is the off-shell momentum. The objects $f(k,r)$, $f(k,q,r)$, $f(k)$, and $f(k,q)$ are the appropriate Jost solutions and Jost functions. Here $\delta(k)$ stands for the s-wave phase shift.

In terms of the wavefunction $\phi(k,q,r)$ the off-shell K matrix can be written in the form

$$\langle p|K(s)|q\rangle\equiv K(p,q,s) =\frac{2}{\pi pq}\int_0^\infty dr\,\sin pr V(r)\phi(k,q,r). \tag{3}$$

Using the values of on- and off-shell Jost solutions $f(\pm k,r)$ and $f(k,\pm q,r)$ [Eqs. (14) and (16) of Ref. 1] for the Woods–Saxon potential,

$$V(r)=\frac{-V_0}{1+e^{(r-R)/a}}\,, \tag{4}$$

in Eq. (1) we get

$$\phi(k,q,r)=\sum_{m=0}^{\infty}\Bigg\{B(k,q)\Bigg[G_m^{(1)}e^{ik'r}\left(\frac{1}{1+e^{(r-R)/a}}\right)^{m+ia(k'-k)} + G_m^{(2)}e^{ik'r}\left(\frac{1}{1+e^{(r-R)/a}}\right)^{m+ia(k'+k)}\Bigg] + Q_2\sum_{n=0}^{\infty}\Bigg\{H_{m,n}(A,B,C,q,\sigma,\tau,R)e^{ik'r} \times\left(\frac{1}{1+e^{(r-R)/a}}\right)^{m+n+ia(k'-q)} - H_{m,n}(A,B,C,-q,\sigma',\tau',R)e^{ik'r} \times\left(\frac{1}{1+e^{(r-R)/a}}\right)^{m+n+ia(k'+q)}\Bigg\}\Bigg\}. \tag{5}$$

Each quantity in Eq. (5) has been defined in paper I except

$$G_m^{(1)}=\frac{e^{i(k-k')R}(A)_m(B)_m}{(C)_m m!}\,, \tag{6a}$$

$$G_m^{(2)}=\frac{e^{i(k+k')R}(A')_m(B')_m}{(C')_m m!}\,, \tag{6b}$$

and

$$B(k,q)=-\tfrac{1}{4}\pi q\langle k|K(s)|q\rangle \tag{6c}$$

with

$$A'=ia(k+k'),\quad B'=ia(k+k')+1,\quad C'=1+2ika. \tag{7}$$

For the sake of clarity we note that in writing Eq. (5) we have used the series representation for the hypergeometric function.[4] Substituting Eq. (5) into Eq. (3) we find

$$K(p,q,s)=\frac{2V_0}{\pi pq}\int_0^\infty dr\,\sin pr\sum_{m=0}^{\infty}\Bigg\{B(k,q) \times\Bigg[G_m^{(1)}e^{ik'r}\left(\frac{1}{1+e^{(r-R)/a}}\right)^{m+1+i(k'-k)a} + G_m^{(2)}e^{ik'r}\left(\frac{1}{1+e^{(r-R)/a}}\right)^{m+1+i(k+k')a}\Bigg] + Q_2\sum_{n=0}\Bigg\{H_{m,n}(A,B,C,q,\sigma,\tau,R)e^{ik'r} \times\left(\frac{1}{1+e^{(r-R)/a}}\right)^{m+n+1+i(k'-q)a} + H_{m,n}(A,B,C,-q,\sigma',\tau',R)e^{ik'r} \times\left(\frac{1}{1+e^{(r-R)/a}}\right)^{m+n+1+ia(k+q)}\Bigg\}\Bigg\}. \tag{8}$$

As usual we break up the integrals in Eq. (8) as

$$\int_0^\infty dr\,\cdots=\int_0^R dr\,\cdots+\int_R^\infty dr\,\cdots$$

and perform the latter integrals to obtain the K matrix in the final form,

$$K(p,q,s)=\frac{2V_0}{\pi pq}\sum_{m,s=0}^{\infty}\{B(k,q)[G_m^{(1)}(I^{(1)}_{m,0,s}(k)+I^{(2)}_{m,0,s}(k)) + G_m^{(2)}(I^{(1)}_{m,0,s}(-k)+I^{(2)}_{m,0,s}(-k))] + Q_2\sum_{n=0}^{\infty}[H_{m,n}(A,B,C,q,\sigma,\tau,R) \times(I^{(1)}_{m,n,s}(q)+I^{(2)}_{m,n,s}(q))$$

 Copyright © 1977 American Institute of Physics

$$- H_{m,n}(A, B, C, -q, \sigma', \tau', R)(I^{(1)}_{m,n,s}(-q) + I^{(2)}_{m,n,s}(-q))]\}, \quad (9)$$

where

$$I^{(1)}_{\mu,\nu,s}(\xi) \text{ and } I^{(2)}_{\mu,\nu,s}(\xi)$$

are exactly the same objects as defined in paper I. The triple sum in Eq. (9) is uniformly convergent.[1] Thus it can be used to check on programs which attempt to compute the K matrix by iteration techniques.[5]

*Work supported in part by the Department of Atomic Energy, Govt. of India.

[1]M.N. Sinha Ray, B. Talukdar, and D. Chattarji, J. Math. Phys. **17**, 1763 (1976).
[2]W. Tobocman and M.A. Nagarajan, Phys. Rev. **163**, 1011 (1967); D. Rodijak, W. Tobacman, and D.K. Tandon, Phys. Rev. C **6**, 1192 (1972); D.J. Ernest, C.M. Shakin, R.M. Thaler, and D.L. Weiss, Phys. Rev. C **8**, 2056 (1973).
[3]M.G. Fuda and J.S. Whiting, Phys. Rev. C **8**, 1255 (1973); B. Talukdar *et al.*, Phys. Rev. C **15**, 1252, 1257 (1977).
[4]Y.L. Luke, *The Special Functions and Their Approximations* (Academic, New York, 1969), Vol. 1.
[5]J. Hufner and R.H. Lemmer, Phys. Rev. **175**, 1394 (1958); J. Hufner and C.M. Shakin, *ibid.* **175**, 1350 (1968).

Stability, equilibrium and KMS for an infinite classical system

M. Pulvirenti

Istituto Matematico dell'Università di Camerino, Camerino, M.C., Italy
(Received 2 August 1976)

The stability condition as a property which characterizes the thermodynamic equilibrium is studied from an abstract point of view. Furthermore, an application of the main result to the case of an infinite classical harmonic system is given.

1. INTRODUCTION

The goal of equilibrium statistical mechanics is to give a reduced description of large and complex systems in terms of a few parameters which satisfy the laws of the thermodynamics. The macroscopic observables are usually obtained by averaging the microscopic ones (measurable functions on the phase space) with a measure which is required to describe the equilibrium state, and then by performing the (thermodynamic) limit for $N, V \to \infty$ (with some care), since the finite volume averages have the unpleasant property of depending on the particles number N and volume V. The use of such finite volume equilibrium measure is justified by the so-called ergodic hypothesis, which seems very hard to prove whenever it does not fail.[1]

Another point of view is to consider infinite systems directly. In this case the macroscopic equilibrium state is described by a measure on the phase space of the infinite system, that is the limit of finite volume equilibrium measures or the solution of the DLR equations; we will refer to such measures as Gibbs or equilibrium states for the system.[2] Obviously, besides the well-known results obtained by the equilibrium statistical mechanics, there remains the problem of justifying our making use of the Gibbs measures instead of the stationary (with respect to time evolution) ones. In this context it seems quite natural to look for some "physical condition" that forces any invariant measure satisfying it, to be a Gibbs state. Such kind of an approach was proposed by Haag, Kastler, and Trych-Polmeyer[3] (see also Ref. 4) for quantum systems. A classical analog of this result has been obtained by Aizemann, Gallavotti, Goldstein, and Lebowitz.[5] (Also see Ref. 6.)

The basic result in Ref. 5 may be summarized as follows. Let us consider an infinite classical particles system and a class $\mathcal{G}$ of stationary states on it, such that the following properties are satisfied:

(a) there exists an algebra $\mathcal{D}$ of functions on the phase space that is invariant for the time evolution of the system and on which the Poisson brackets make sense;

(b) each state of $\mathcal{G}$ is three-fold mixing with respect to the time evolution.

(c) $\mathcal{D}$ is supposed to verify the classical analog of the L_1 asymptotic Abelianness, the so-called dispersivity.

With these conditions it is shown in Ref. 5 that the only states in $\mathcal{G}$ which are stable under local perturbations are KMS states. The KMS condition is a good candidate to describe the thermodynamic equilibrium by itself and it proves in many cases to be equivalent to the Gibbs condition.[7]

Unfortunately, besides the free gas case, it is difficult to verify or disprove the assumptions (a), (b), and (c) in the case of interacting systems. In fact the ergodic properties as stated in (b) have not yet been proven,[8] and furthermore the choice of $\mathcal{D}$ seems problematic since it must be simultaneously large enough to satisfy (a) and small enough to satisfy (c).

In this paper the same result as in Ref. 5 is obtained in an abstract framework, without reference to any physical system by means of relaxed assumptions. More precisely it is required:

(a′) the existence of an algebra $\mathfrak{a}$ of functions on which the Poisson bracket makes sense, and which is a core for the dynamical infinitesimal generator;

(b′) the systems are assumed to be weakly mixing.

No dispersivity property is required.

Though the properties (a′) and (b′) are implied by (a) and (b), it is still hard to prove them in the most interesting physical cases. In particular, property (a′), that is the self-adjointness of the Liouville operator on the algebra of the cylindrical observables, seems to require a deeper knowledge of the infinite motion than we have thus far for the classical continuous system.[9]

We conclude by observing that we need to strengthen the stability assumption with respect to that used in Ref. 5. This means that, in selecting KMS states in the class of the invariant states satisfying (a′) and (b′), we need a stronger condition.

The main result (Theorem 2.1) of this paper is formulated in Sec. 2, and an application to a concrete system is discussed in Sec. 3. In Sec. 4 the proof of Theorem 2.1 is given, making use of the following theorem [it may be found in Ref. 10].

Theorem 1.1: Let (X, Σ, μ) be a probability space and H be a self-adjoint operator on $L_2(\Sigma, \mu)$.

Let us suppose the existence of a self-adjoint algebra of functions $\mathcal{A} \subset \mathcal{D}(H)$ such that

(i) $\mathcal{A}$ is a core for H,

(ii) $\forall\, f, g \in \mathcal{A}$ then $H(fg) = fHg + gHf$. Then $\exp(iHt)$ implements a group of automorphisms of $L_\infty(\Sigma, \mu)$.

 Copyright © 1977 American Institute of Physics

2. DEFINITIONS, NOTATIONS, RESULTS

Let (K, Σ, ω) be a Lebesgue probability space[11] and let α denote a representation of the real line $\mathbb{R}$ as measure preserving automorphisms of the measure algebra $\Sigma(\omega)$; the following measurability condition is required:

$$\text{if } A,\ B \in \Sigma(\omega),\ \text{then } \mathbb{R} \ni t \mapsto \omega(A \cap \alpha_t B) \text{ is a measurable function.} \tag{2.1}$$

On the basis of a well-known theorem of Halmos, Von Neumann, and Rokhlin,[11] we may think of α_t, $t \in \mathbb{R}$, as a family of almost everywhere defined point transformation on K. Furthermore the above condition (2.1) implies the existence of a weakly measurable (and hence strongly continuous) unitary group U_t on $L_2(\omega)$, such that

$$(U_t f)(\kappa) = f(\alpha_t \kappa)$$

for a.a. $\kappa \in K$, $\forall\ t \in \mathbb{R}$, $\forall\ f \in L_2(\omega)$.

Let H be the self-adjoint operator on $L_2(\omega)$ that generates U_t; we put $L = iH$ and denote by $D(L)$ its domain.

Let us suppose that $\mathfrak{a}$ is a self-adjoint algebra of functions with the following properties:

(i) $\mathfrak{a} \subset \bigcap_{p \geq 2} L_p(\omega) \cap D(L)$;

(ii) there exists a self-adjoint subalgebra $\beta \subset \mathfrak{a}$ of essentially bounded functions such that $I \in \beta$ and β is a core for L;

(iii) a bilinear form (the Poisson bracket) $\{\cdot, \cdot\}: \mathfrak{a} \times \mathfrak{a} \to \mathfrak{a}$ is defined such that for any f, y, $h \in \mathfrak{a}$ the following properties are verified:

(α) $\overline{\{f, g\}} = \{\bar{f}, \bar{g}\}$,

(β) $\{f, g\} = -\{g, f\}$,

(γ) $\{f, gh\} = \{f, g\}h + \{f, h\}g$;

(iv) $L\mathfrak{a} \subset \mathfrak{a}$;

(v) $\omega(\{Lg, h\}) = -\omega(\{g, Lh\})$, $g, h \in \mathfrak{a}$.

Definition 2.1: A system $Y \equiv (K, \omega, \alpha, \mathfrak{a}, \{\cdot, \cdot\})$ with the above properties is called a Poisson system.

Definition 2.2 (KMS): A Poisson system Y is said to verify the (static) KMS condition at some inverse temperature $\beta \in \mathbb{R}$ if for any f, $g \in \mathfrak{a}$

$$-\beta\omega(fLg) = \omega(\{f, g\}).$$

Definition 2.2 (KMS): A Poisson system Y is said to verify the (static) KMS condition at some inverse

(i) $\forall\ f \in \mathfrak{a}$, $\exists\ \delta : \forall\ \lambda \in (0, \delta)\ \exists\ \rho_{\lambda f} \in L_2(\omega)$

such that $\omega(\rho_{\lambda f}) = 1$ and the bounded measure defined as follows:

$$d\omega^{\lambda f} = \rho_{\lambda f} d\omega$$

is formally invariant for the perturbed dynamics, i.e.,

$$\omega^{\lambda f}(Lg + \lambda\{f, g\}) = 0, \quad g \in \mathfrak{a}.$$

(ii) $\lim_{\lambda \to 0}(\rho_{\lambda f} - I)/\lambda = \left.\frac{d\rho_{\lambda f}}{d\lambda}\right|_{\lambda=0}$ exists

in $L_2(\omega)$.

(iii) If $\{f_n\}_{n=1}^{\infty} \subseteq \mathfrak{a}$ is such that $f_n \to 0$ in $L_2(\omega)$,

then

$$\left.\frac{df_{\lambda f_n}}{d\lambda}\right|_{\lambda=0} \to 0 \text{ in } L_2(\omega).$$

The main result of this paper may be summarized in the following theorem:

Theorem 2.1: If a weakly mixing Poisson system Y is stable, then there exists an inverse temperature β for which Y is KMS.

We recall that the weak mixing condition means:

$$\lim_{T \to \infty} \frac{1}{T} \int_{-T}^{T} \left| (f, U_t g) - (f, E_0 g) \right| dt = 0$$

$\forall\ f, g \in L_2(\omega)$ and E_0 is the projection on the constants. In the proof, that we will give in Sec. 4, only the ergodicity and the equality $\mathrm{Sp}H = \mathbb{R}$ will be used as consequences of the weak mixing condition.[12]

3. HARMONIC SYSTEMS

In this section we will show by means of Theorem 2.1, that for some physical system, the only states of a certain class that are stable in the sense of Definition 2.3. are the equilibrium states.

The main difficulty we meet in using Theorem 2.1 is to describe the physical models in terms of Poisson systems and in particular to show property (ii) of Sec. 2, which, as we will see later (Remark 3.3), implies a unicity property of the time evolution with respect to the considered class of invariant states. We specify these considerations by studying the simple model of an infinite system of interacting harmonic oscillators. The equilibrium and the dynamics for this model have been studied in a more general context by Lanford and Lebowitz[13] and by Van Hemmen.[14]

Harmonic System (one dimension): The *phase space* K is $(\mathbb{R}^2)^{\mathbb{Z}}$, i.e., the space of all the sequences $x = \{x_\alpha\}_{\alpha \in \mathbb{Z}}$ where $x_\alpha \equiv (p_\alpha, q_\alpha)$ is the point of the one-particle phase space. The *interaction* is described by an infinite dimensional matrix $V_{\alpha,\beta}$ with the following properties:

(i) *translation invariance*:

$$V_{\alpha,\beta} = V_{\alpha+\gamma,\beta+\gamma}, \quad | \forall | \alpha, \beta, \gamma \in \mathbb{Z}.$$

(ii) *short range interaction*: For each $\alpha, \beta \in \mathbb{Z}$ there exists $r \in \mathbb{Z}^+$ such that $\beta > \alpha \pm r$ implies $V_{\alpha,\beta} = 0$.

(iii) *positivity*: For all finite nonzero sequences $\{q_\alpha\}$, one has $\sum_{\alpha,\beta} V_{\alpha\beta} q_\alpha q_\beta > 0$. Then the finite volume Hamiltonian is

$$H_n = \tfrac{1}{2} \sum_{\alpha=-n}^{n} P_\alpha^2 + \sum_{\alpha=-n}^{n} \sum_{\beta=-n}^{n} V_{\alpha\beta} q_\alpha q_\beta \geq 0.$$

The infinite equations of the motion are

$$\dot{p}_\alpha = \sum_\beta V_{\alpha\beta} q_\beta, \quad \dot{q}_\alpha = p_\alpha,$$

that may be written as

$$\dot{x} = Ax, \tag{3.1}$$

where $A: K \to K$ is the linear operator defined by

$$(Ax)_\alpha = (-\sum_\beta V_{\alpha,\beta} q_\beta, p_\alpha),$$

where

$$x_\alpha \equiv (p_\alpha, q_\alpha).$$

Let us denote by K_0 the subset of K such that for each $x \in K_0$ one has

$$\lim_{|\alpha| \to \infty} \exp(-|\alpha|)|x_\alpha| = 0, \tag{3.2}$$

where $|x_\alpha| = \max(|p_\alpha|, |q_\alpha|)$; then K_0 is a Banach space with the norm

$$\|\times\| = \sup_\alpha \exp(-|\alpha|)|x_\alpha|. \tag{3.3}$$

Let us set

$$\lambda = \sup_{\beta \in [-r, r]} |V_{0,\beta}| \quad \text{and} \quad \bar{\lambda} = \max(\lambda, 1).$$

Then we easily get for each $x \in K_0$,

$$\|A\times\| \leq (2r+1)\bar{\lambda}\,\|\times\|.$$

Hence the operator $e^{At}: K_0 \to K_0$ gives a solution of the problem (3.1) with initial data in K_0.

We now introduce the following *partial dynamics*. For all $n \in \mathbb{Z}^+$ let us define the projection operator on K_0,

$$(P_n x)_\beta = \begin{cases} x_\beta & \text{if} \quad |\beta| \leq n, \\ 0 & \text{if} \quad |\beta| > n. \end{cases}$$

Then by (3.2),

$$\|P_n x - x\| \to 0 \quad (n \to \infty) \quad \forall\ x \in K_0.$$

As a consequence, setting $A_n \equiv P_n A$ one has

$$\exp(A_n t)x \to \exp(At)x \quad (n \to \infty) \quad \forall\ x \in K_0.$$

$\exp(A_n t)x$ is obtained by evolving the oscillators in $[-n, n]$ under the action of themselves and freezing the oscillators in $[-n, n]^c$ in their initial positions.

Let $F^n: \mathbb{R}^{2(2n+1)} \to \mathbb{C}$ be an infinitely differentiable function such that for every $h \in \mathbb{Z}$ there exists a $k \in \mathbb{Z}$ for which $|\xi|^{-k}|(\partial^h F)(\xi)| \to 0$ as $|\xi| \to \infty$, where $\partial^h F$ denotes a derivative of order h and $\xi \in \mathbb{R}^{2(2n+1)}$. Starting from F^n we define the function $f \in K \mapsto \mathbb{C}$ as follows:

$$\forall\ x \in K, \quad f(x) = F^n(x_{-n} \cdots x_0 \cdots x_n). \tag{3.4}$$

We denote by $\mathfrak{a}$ the algebra of all functions f of this kind, and by $\mathcal{B} \subset \mathfrak{a}$ the subalgebra of all bounded functions with bounded derivatives.

The *Poisson bracket* is defined on $\mathfrak{a}$ by the usual formula:

$$\{f, g\} = \sum_\alpha \left(\frac{\partial F^n}{\partial p_\alpha} \frac{\partial G^m}{\partial q_\alpha} - \frac{\partial F^n}{\partial q_\alpha} \frac{\partial G^m}{\partial p_\alpha} \right), \tag{3.5}$$

where f and g are obtained by F^n and G^m via (3.4). The sums in (3.5) are finite by definition and the Poisson bracket is a bilinear form on $\mathfrak{a} \times \mathfrak{a}$ which verifies (iii) of Sec. 2.

The *Liouville operator* $\mathcal{L}$ is defined by the composition of the infinite formal Hamiltonian with the functions of $\mathfrak{a}$ via the Poisson bracket. By definition if F^n generates f as in (3.4), then $(\mathcal{L} f)(x) = \lim_k \{H_k, f\}(x)$ $= \langle \nabla_x F^n, P_n A x \rangle$ where $\langle\ ,\ \rangle$ denotes the following scalar product:

$$\langle x, y \rangle = \sum_\alpha x_\alpha \cdot y_\alpha$$

and $x_\alpha \cdot y_\alpha$ is the usual scalar product in $\mathbb{R}^2$.

We denote by $\mathcal{M}$ the set of all probability measures ω on K with the following properties:

(i) $\omega(K - K_0) = 0$;

(ii) ω is a regular Borel measure on K_0;

(iii) ω is invariant for the dynamics $\exp At$;

(iv) $\int \|\times\|^2 d\omega(x) < +\infty$;

(v) $\mathfrak{a} \subset L_2(\omega)$.

Proposition 3.1: If $\omega \in \mathcal{M}$ then $Y \equiv (K, \omega, \alpha, \mathfrak{a}, \{\cdot, \cdot\})$ is a Poisson system.

Proof: The measure space (K, ω) is Lebesgue, since K_0 is a separable complete metric space and ω satisfies the properties (i) and (ii) above. The algebra $\mathfrak{a}$ is an $L_2(\omega)$ dense self-adjoint algebra of unbounded functions. This may be shown by approximating any $\|\cdot\|$-continuous bounded function $f: K_0 \to \mathbb{C}$ with a sequence defined as $f_n(x) = f(P_n x)$ and then approximating any f_n with elements of $\mathcal{B}$. Denoting by $\mathfrak{a}_n$ the algebra of all $f \in \mathfrak{a}$ depending only on the first $2(2n+1)$ coordinates and moment, then $\mathcal{L}\mathfrak{a}_n \subset \mathfrak{a}_{n+r}$ and hence $\mathcal{L}\mathfrak{a} \subset \mathfrak{a}$ since $\mathfrak{a} = \cup_n \mathfrak{a}_n$. If $f \in \mathfrak{a}$, then for n large enough, $\mathcal{L} f = \{H_n, f\}$. The Jacobi identity, together with the invariance of the state, gives property (v) of Sec. 2.

Let us define

$$(U_t f)(x) = f(\alpha_t x), \quad (U_t^n f)(x) = f(\alpha_t^n x),$$

where $\alpha_t = \exp At$, $\alpha_t^n = \exp A_n t$. Then U_t is a strongly continuous unitary group by the invariance of ω and the continuity of the motion. To complete the proof we need $\mathcal{B}$ to be a core for $\mathcal{L}$.

If now $\tilde{\mathcal{L}}$ denotes the infinitesimal generator of U_t it remains to show that

$$\tilde{\mathcal{L}} = \mathcal{L}\,|\,^{**}_{\mathcal{B}}. \tag{3.6}$$

Since $\tilde{\mathcal{L}}\,|\,\mathcal{B} = \mathcal{L}\,|\,\mathcal{B}$, it is enough to prove that

$$U_t f \in \mathcal{D}(\mathcal{L}\,|\,^{**}_{\mathcal{B}}), \quad t \in \mathbb{R} \text{ and } f \in \mathcal{B}.^{15}$$

In other words, we must find a sequence $\{g_n\}_{n=1}^\infty \subset \mathcal{B}$, such that $g_n \to U_t f$ and $\mathcal{L} g_n \to \mathcal{L} U_t f$ in $L_2(\omega)$.

Let us put

$$g_n(x) = (U_t^n f)(x), \quad n \in \mathbb{Z}^+$$

since $f \in \mathcal{B}$ for n large enough g_n depends only on the variables in $[-n-r, n+r]$. Furthermore $g_n \in \mathfrak{a}_{n+r} \cap \mathcal{B}$ and $\|g_n\|_\infty = \|f\|_\infty$. Since $\alpha_t^n x \to \alpha_t x \ \forall\ x \in K_0$, then $g_n \to U_t f$, ω a.e.; the estimate $|U_t f - g_n| \leq 2\|f\|_\infty$ and the dominated convergence theorem imply that $g_n \to U_t f$ in $L_2(\omega)$. Furthermore we have

$$\begin{aligned}(\mathcal{L} g_n)(x) &= (\mathcal{L} U_t^n f)(x) = \langle \nabla_x U_t^n f, P_{n+r} A x \rangle \\ &= \langle \nabla_{\alpha_t^n x} f, \alpha_t^n P_{n+r} A x \rangle \\ &= \langle \nabla_{\alpha_t^n x} f, \alpha_t^n P_n A x \rangle \\ &\quad + \langle \nabla_{\alpha_t^n x} f, \alpha_t^n (P_{n+r} - P_n) A x \rangle.\end{aligned}$$

The first term in the rhs is $U_t^n \mathcal{L} f$ and the second one is bounded by

$$c\|(P_{n+r} - P_n) A x\|, \quad c \in \mathbb{R}$$

by the Schwarz inequality. Hence: $\mathcal{L}g_n \to \mathcal{L}U_t f\omega$ a.e. Finally the estimate:

$$|\mathcal{L}U_t f - \mathcal{L}g_n|(x) \leq |(U_t - U_t^n)\mathcal{L}f|(x) + 2c\|A\|\,\|\|\times\|$$
$$\leq 2\|\mathcal{L}f\|_\infty + 2c\|A\|\,\|\|\times\|$$

and the property (iv) for the measures in $\mathcal{M}$, allow the use of the dominated convergence theorem to achieve the proof. We can now use Theorem 2.1 to obtain:

Theorem 3.1: The KMS states are the only weakly mixing stable state of the class $\mathcal{M}$.

Remark 3.1: The above Proposition 3.1 and Theorem 3.1 hold for a γ-dimensional lattice since our dynamical arguments do not depend on the space dimension crucially.

Remark 3.2: In Refs. 13 and 14 the ergodic properties of the dynamical flow (K, ω, α_t) where ω is a KMS state, are also studied. In particular it has been shown that α_t is ergodic if and only if V has no point spectrum, and that α_t is Bernoulli if and only if V has absolutely continuous spectrum.

Remark 3.3: Let us consider the dynamical flow (K, ω, α_t) where ω is K.M.S. and α_t is a time evolution. The property $\mathcal{L}|_{\mathcal{B}}^{**} = \mathcal{L}$ implies that (K, ω, α_t) is the only flow constructed by solving Eq. (3.1). Otherwise $\mathcal{L}|_{\mathcal{B}}$ should have other self-adjoint extensions. Furthermore, the above self-adjointness property means that the dynamics is essentially local, since the knowledge of the infinitesimal generator on local quantities is enough to get the whole dynamics.

4. PROOF OF THEOREM 2.1

By Definition 2.3, (i) we have

$$\omega^{\lambda f}(\mathcal{L}g) = -\lambda\omega^{\lambda f}(\{f, g\}), \quad f, g \in \mathfrak{a}.$$

The invariance of ω then gives

$$\omega(([\rho_{\lambda f} - I]/\lambda)Lg) = -\omega^{\lambda f}(\{f, g\}).$$

Finally by Definition 2.3, (ii), one obtains

$$\omega\left(\frac{d\rho_{\lambda f}}{d\lambda}\bigg|_{\lambda=0}\mathcal{L}g\right) = -\omega(\{f, g\}). \tag{4.1}$$

Denoting by E_0 the projection on the constants, we put

$$\mathcal{H} = L_2(\omega) \ominus E_0 L_2(\omega).$$

Then by renormalization $\omega(d\rho_{\lambda f}/d\lambda|_{\lambda=0}) = 0$ and hence

$$\frac{d\rho_{\lambda f}}{d\lambda}\bigg|_{\lambda=0} \in \mathcal{H}. \tag{4.2}$$

By the ergodicity

$$[\mathrm{Ran}\mathcal{L}|^{\perp} = \ker(-\mathcal{L}) = E_0 L_2(\omega) \tag{4.3}$$

and so

$$\overline{\mathcal{L}\mathfrak{a}} = \mathcal{H} \tag{4.4}$$

because $\mathfrak{a}$ is a core for $\mathcal{L}$.

By (4.1), (4.2), and (4.4) we see that the mapping

$$f \mapsto \frac{d\rho_{\lambda f}}{d\lambda}\bigg|_{\lambda=0} \equiv \tilde{T}f, \qquad f \in \mathfrak{a}$$

is a well-defined linear operator. Furthermore Definition (2.3) (iii) forces $\tilde{T}$ to be continuous. Let T be the continuous extension of $\tilde{T}$ on all $L_2(\omega)$. We shall prove that T is self-adjoint and that it commutes with the spectral family of $H = -i\mathcal{L}$.

Putting $f, g \in \mathfrak{a}$, by (4.1),

$$\omega(\bar{T}f\mathcal{L}g) = -\omega(\{\bar{f}, g\}) = -\overline{\omega(\{f, \bar{g}\})}$$
$$= \overline{\omega(Tf\mathcal{L}g)} = \omega(\overline{Tf}\mathcal{L}g)$$

and hence

$$\overline{Tf} = T\bar{f}. \tag{4.5}$$

Furthermore if $f = \mathcal{L}h$, $h \in \mathfrak{a}$, by (4.1) and (4.5),

$$(f, T\mathcal{L}g) = (\mathcal{L}h, T\mathcal{L}g) = \omega(T\mathcal{L}g\mathcal{L}\bar{h})$$
$$= -\omega(\{\mathcal{L}g, \bar{h}\}) = \omega(\{g, \mathcal{L}\bar{h}\})$$
$$= -\omega(\{\bar{f}, g\}) = \omega(T\bar{f}\mathcal{L}g)$$
$$= (Tf, \mathcal{L}g). \tag{4.6}$$

Combining (4.2) and (4.6) we obtain $T = T^*$.

Finally $\forall\, f, g \in \mathfrak{a}$ we have

$$(Tf, \mathcal{L}g) = -\omega(\{\bar{f}, g\}) = \omega(\{g, \bar{f}\}) = -(T\bar{g}, \mathcal{L}\bar{f})$$
$$= -(\mathcal{L}f, Tg) = -(T\mathcal{L}f, g)$$

and so $Tf \in \mathcal{D}(\mathcal{L})$ and $\mathcal{L}Tf = T\mathcal{L}f$.

Let us now fix $u \in \mathcal{D}(\mathcal{L})$. We can choose a sequence $\{f_n\}_{n=1}^{\infty} \subset \mathfrak{a}$ such that

$$f_n \to u, \quad \mathcal{L}f_n \to \mathcal{L}u.$$

By the continuity of T

$$Tf_n \to Tu, \quad T\mathcal{L}f_n \to T\mathcal{L}u$$

and hence

$$Tu \in \mathcal{D}(L) \text{ and } \mathcal{L}Tu = T\mathcal{L}u.$$

Defining

$$R_\lambda = (\mathcal{L} + \lambda I)^{-1} \quad \forall\, \lambda \in \mathbb{R}$$

and putting

$$g = R_\lambda f \in \mathcal{D}(\mathcal{L}), \quad f \in L_2(\omega)$$

we have

$$TR_\lambda f = Tg = R_\lambda(\mathcal{L} + \lambda I)Tg = R_\lambda T(\mathcal{L} + \lambda I)g = R_\lambda Tf,$$

so we can conclude that T commutes with the spectral family of H.

The antisymmetric operator: $D_0 \equiv T\mathcal{L}|_{\mathcal{B}}$ acts as a derivation on $\mathcal{B}$ [that is $D_0(fg) = fD_0 g + gD_0 f$, $\forall\, f, g \in \mathcal{B}$] and this in virtue of (4.1) and property (γ) in the definition of the Poisson bracket (Sec. 2).

We want to show that $\bar{D}_0$ is essentially antiself-adjoint on $\mathcal{B}$ in order to show that $\exp\bar{D}_0 t$ implements an automorphism of $L_\infty(\omega)$, $t \in \mathbb{R}$. (See Theorem 1.1). Suppose the contrary. Then there exists $h \in \mathcal{D}(D_0^*)$ such that

$$D_0^* h = \pm h.$$

$\forall\, f \in \mathcal{B}$ we put $(\mathcal{L} + I)f = g \in \mathfrak{a}$; then

$$(h, D_0 f) = \mp(h, f),$$
$$(h, T\mathcal{L}R_1 g) = \mp(h, R_1 g),$$
$$-(R_{-1}h, T\mathcal{L}g) = \pm(R_{-1}h, g).$$

If f spans β the g's span a dense set by the core property of β and hence

$$\mathcal{L}TR_{-1}h = \mp R_{-1}h$$

This is a contradiction since $\mathcal{L}T$ defined on $\mathcal{D}(\mathcal{L})$ is antisymmetric.

Let us define the two parameters group of automorphisms of $L_\infty(\omega)$:

$$V(t,s) = \exp(Dt)\exp(\mathcal{L}s).$$

Then $\mathrm{Sp}V$, defined as

$$\mathrm{Sp}V \equiv \{\lambda \in \mathbb{R}^2 \mid \hat{f}(\lambda) = 0 \text{ for all } f \in L_1(\mathbb{R}^2) : V_f \equiv \int V(t) f(t)\, dt = 0\}$$

is additive.[16] (This may be seen by using Theorem 4.1 of Ref. 16, in Abelian case recognizing that in this context it is enough to have the commutativity of V_t with the ergodic flow U_t to obtain the conclusion.)

On the other hand,

$$\mathrm{Sp}V = \{\langle \lambda\mu, \lambda\rangle \mid \lambda \in \mathrm{Sp}H,\ \mu \in \mathrm{Sp}T;\ \lambda, \mu \neq 0\} \cup \{0\}. \quad (4.7)$$

Then combining the additivity of $\mathrm{Sp}V$ with (4.7) we can find, for fixed $\mu_1 \neq \mu_2 \in \mathrm{Sp}T$, a $\tilde{\mu} \in \mathrm{Sp}T$ with this form

$$\tilde{\mu} = \frac{\lambda_1\mu_1 + \lambda_2\mu_2}{\lambda_1 + \lambda_2} = \tfrac{1}{2}\left\{\mu_1 + \mu_2 + \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2}(\mu_1 - \mu_2)\right\}$$

where $\lambda_1, \lambda_2 \in \mathrm{Sp}H$. Since $\mathrm{Sp}H = \mathbb{R}$ by the weakly mixing assumption, the number $\tilde{\mu}$ may be arbitrarily large and this is in contradiction with the continuity of T. Hence it is possible only an eigenvalue β (different from 0) for T. So we conclude that

$$T = \beta(I - E_0)$$

and the KMS condition is proven.

ACKNOWLEDGMENTS

The author is indebted to G. Riela and L. Van Hemmen for useful discussions and to G. Gallavotti and E. Presutti for their patience in discussing step by step the content of this work.

[1] D. Ruelle, *Statistical Mechanics* (Benjamin, New York, 1969) 1969).
[2] R. L. Dobrushin, Funkts. Anal. Ego. 2, 31, 44 (1968); 3, 27 (1969); O. Lanford and D. Ruelle, Commun. Math. Phys. 13, 194 (1969).
[3] R. Haag, D. Kastler, E. Trych-Polmeyer, Commun. Math. Phys. 38, 173 (1974).
[4] O. Bratteli and D. Kastler, Commun. Math. Phys. 46, 37 (1976).
[5] M. Aizenman, G. Gallavotti, S. Goldstein, and J. L. Lebowitz, Commun. Math. Phys. 48, 1 (1976).
[6] M. Pulvirenti and G. Riela, "KMS condition for stable states of infinite classical systems," to appear in J. Math. Phys.
[7] G. Gallavotti and E. Verboren, Nuovo Cimento B 28, 274 (1975).
[8] S. Goldstein, J. L. Lebowitz, and A. Aizenman, Lect. Notes Phys. 38, 112 (1975).
[9] G. Gallavotti, Ann. Mat. Pura Appl. CVIII, 227 (1976).
[10] G. Gallavotti and M. Pulvirenti, Commun. Math. Phys. 46, 1 (1976).
[11] V. A. Rokhlin, Mat. Sb. 25, 1 (1949).
[12] Ya. Sinai, Aarhus Lect. Notes 23 (1970).
[13] O. Lanford and J. L. Lebowitz, Lect. Notes Phys. 38, 144 (1975).
[14] L. Van Hemmen, thesis, University of Groningen (1976).
[15] N. Dunford and J. T. Schwartz, *Spectral Theory* (Interscience Wiley, New York, 1963), Vol. 2.
[16] D. Kastler, "Equilibrium states of matter operator algebras," Proc. of the Rome Conf. on C^*-algebras, 1st Naz. di Alta Matematica (1975).

On notions of Markov property

G. O. S. Ekhaguere

Department of Mathematics, University of Ibadan, Ibadan, Nigeria
(Received 28 January 1977)

We exhibit the logical connection between two mathematically and physically interesting notions of Markov property due to Nelson [J. Funct. Anal. **12**, 97 (1973)] and to Wong [Ann. Math. Stat. **40**, 1625 (1969)], respectively, in the case of Gaussian generalized stochastic fields.

INTRODUCTION

Markov property, or probabilistic causality, is a statement of conditional stochastic independence of random variables and the notion of "conditioning" was first introduced by the Russian mathematician Markov.[1] His compatriot, Kolmogorov,[2] then gave this important concept a very rigorous mathematical basis by invoking measure theory. What are now known as Markov stochastic processes intervene in many important physical and mathematical considerations and a theory of these processes utilizing various mathematical methods is at an advanced stage of development. However, the definition of Markov property for stochastic processes indexed by R explicitly utilizes the ordering relation of points of R. In trying to extend the notion of Markov property so that it may apply to a stochastic field, generalized or ordinary, one is, therefore, initially handicapped by the absence of a corresponding ordering of the points of R^d, $d>1$. In spite of this apparent difficulty, one can still consider Markovicity of stochastic fields, but because different localizations are now possible, there are several notions of Markov property[3-5] for these fields. In Ref. 3, for example, Hegerfeldt discusses the connection between a notion of Markov property introduced by him and that due to Nelson. Here, we deal exclusively, with generalized stochastic fields, and we consider two notions of Markov property for these fields due to Nelson[4] and Wong,[5] respectively. Nelson's notion of Markov property has been employed in recent investigations in constructive quantum field theory[6,7] and it has led there to developments of no small significance. By imposing a regularity assumption on a Euclidean invariant (see below) scalar Gaussian generalized stochastic field, Wong was able to show,[5] much earlier than Nelson, that such a field would be Markov in his sense if and only if it is the same scalar generalized stochastic field which now leads, via Nelson's recent reconstruction theorem,[4] to the free massive scalar Wightman quantum field. It is therefore, mathematically and physically interesting to investigate the connection, if any, between these two notions of Markov property which we formulate below.

1. EUCLIDEAN COVARIANT GENERALIZED STOCHASTIC FIELDS

The d-dimensional Euclidean space E^d is a couple $(R^d, |\cdot|)$ where

$$|\cdot| : R^d \to R_+ = [0, \infty),$$
$$x=(x_1, \ldots, x_d) \mapsto |x| = \left(\sum_{i=1}^{d} x_i^2\right)^{1/2}.$$

E^d is a group space[8] whose group of transformations is the Euclidean group $M(d)$.[9] The latter is the semidirect product of the additive group E^d and the d-dimensional orthogonal group $O(d)$, i.e.,

$$M(d) = E^d \textcircled{s} O(d).$$

Thus, each $g \in M(d)$ is a pair $g=(\alpha, h)$ with $\alpha \in E^d$ and $h \in O(d)$. The orthogonal group $O(d)$ is compact and hence it admits finite-dimensional representations.[10]

Let $S(R^d)$ be Schwartz space of rapidly decreasing complex-valued C^∞ functions on R^d. Let (Ω, β, μ) be a probability space and let

$$H^0(\Omega) = \{\xi(f) = (\xi_i(f) : i=1, \ldots, N) : f \in S(R^d)\}$$

be an N-component generalized stochastic field[11] on (Ω, β, μ). We assume throughout that $H^0(\Omega)$ is of second order, i.e., $\xi(f)=(\xi_1(f), \ldots, \xi_N(f)) \in H^0(\Omega)$ implies $\xi_i(f) \in L^2(\Omega, \beta, \mu)$, $i=1, \ldots, N$. In that case, each $\xi(f) \in H^0(\Omega)$ necessarily has an expectation functional because $L^q(\Omega, \beta, \mu)$ contains $L^r(\Omega, \beta, \mu)$, for $q<r$.

Let $M(f)$ and $B(f^{(1)}, f^{(2)})$ denote the matrices whose entries are $M_i(f) = \langle 1, \xi_i(f)\rangle_{L^2(\Omega, \beta, \mu)}$ and $\langle \xi_i(f^{(1)}), \xi_j(f^{(2)})\rangle_{L^2(\Omega, \beta, \mu)}$, respectively. Thus the column matrix $M(f)$ contains the mean values of the components of $\xi(f) \in H^0(\Omega)$, while $B(f^{(1)}, f^{(2)})$ is the matrix of their correlation functionals. Let $h \mapsto V(h)$ be an N-dimensional unitary irreducible representation of $O(d)$. We are interested, in this paper, in N-dimensional generalized stochastic fields which transform covariantly with respect to $M(d)$ in the following way.

(1.1) Definition: Let $H^0(\Omega) = \{\xi(f) = (\xi_i(f);\ i=1, \ldots, N) : f \in S(R^d)\}$ and let $h \mapsto V(h)$, $h \in O(d)$, be as just described above. Then, we say that $H^0(\Omega)$ is a Euclidean covariant generalized stochastic field transforming according to the unitary irreducible representation $h \mapsto V(h)$, $h \in O(d)$, if it satisfies the following conditions:

(i) the induced action τ_g on $H^0(\Omega)$ of the transformation g of R^d onto itself is specified as follows

$$(\tau_g \xi(f))_i = \sum V_{ij}(h) \xi_j(V_g f), \quad i=1, \ldots, N.$$

(ii) the mean matrix $M(f)$ and the correlation matrix $B(f^{(1)}, f^{(2)})$ are such that

(a) $M(f) = V(h) M(V_g f)$

(b) $B(f^{(1)}, f^{(2)}) = V(h)\, B(V_g f^{(1)}, V_g f^{(2)})\, V(h)^{-1}$

where $(V_g f)(x) = f(g^{-1}x)$ and

$$g=(a, h) \in M(d) = E^d \textcircled{s} O(d).$$

 Copyright © 1977 American Institute of Physics

(1.2) Remark: It is clear that (ii) (a) is readily satisfied if $\mathcal{M}_i(f)=0$, $i=1,\ldots,N$. Hence in what follows, we assume this to be the case, i.e., that the components of each $\xi(f)\in H^0(\Omega)$ have zero mean. For a detailed analysis of the structure of the most general form of $B(f^{(1)}, f^{(2)})$ and the spectral representation for $\xi(f)\in H^0(\Omega)$ such that (ii) (b) holds, we refer the reader to Refs. 12 and 13.

2. THE SPACES $H(\mathcal{R}^d)$ AND $\mathcal{H}(\mathcal{R}^d)$

Let $H^0(\Omega)=\{\xi(f)=(\xi_j(f): j=1,\ldots,N): f\in\mathcal{S}(\mathcal{R}^d)\}$ be an N-component Euclidean covariant generalized stochastic field on the probability space $(\Omega, \mathcal{B}, \mu)$; let $\langle\xi_i(x), \xi_j(y)\rangle_{L^2(\Omega,\mathcal{B},\mu)}=B_{ij}(x-y)$, and let $\tilde{B}(p)$ denote the $N\times N$ matrix whose entries are given by

$$\tilde{B}_{ij}(p)=\int dx\, B_{ij}(x)\exp(ip\,.\,x).$$

Set $(\mathcal{S}(\mathcal{R}^d))^N=\mathcal{S}_N$.

The random-variable-valued generalized function ξ may be regarded as a bounded linear operator with domain $\mathcal{S}_N$ and range contained in $L^2(\Omega, \mathcal{B}, \mu)$ as follows[14]:

$$\xi:\mathcal{S}_N\to\xi(\mathcal{S}_N)\subset L^2(\Omega,\mathcal{B},\mu),$$
$$\mathbf{f}=(f_1,\ldots,f_N)\mapsto\xi(\mathbf{f})=\sum_{i=1}^{N}\xi_i(f_i).$$

Let

$$(\cdot,\cdot)_{\mathcal{S}_N}:\xi(\mathcal{S}_N)\times\xi(\mathcal{S}_N)\to\mathbb{C}$$

be the following sesquilinear functional:

$$\begin{aligned}(\xi(\mathbf{f}^{(1)}), \xi(\mathbf{f}^{(2)}))&\to(\xi(\mathbf{f}^{(1)}), \xi(\mathbf{f}^{(2)}))_{\mathcal{S}_N}\\&=\sum_{i,j=1}^{N}\langle\xi_i(f_i^{(1)}), \xi_j(f_j^{(2)})\rangle_{L^2(\Omega,\mathcal{B},\mu)}\\&=\sum_{i,j=1}^{N}\int dp\,\tilde{f}_i^{(1)}(p)\,\tilde{B}_{ij}(p)\,\tilde{f}_j^{(2)}(p).\end{aligned}$$

The positive-definiteness of $\tilde{B}(p)=(\tilde{B}_{ij}(p): i,j=1,\ldots,N)$ now allows us to assert that $(\cdot,\cdot)_{\mathcal{S}_N}$ is indeed an inner product, and $\|\xi(\mathbf{f})\|_{\mathcal{S}_N}=(\xi(\mathbf{f}), \xi(\mathbf{f}))^{1/2}_{\mathcal{S}_N}=0$, if and only if $\mathbf{f}=0$, and hence, if and only if $\xi(\mathbf{f})=0$.

Next, let $H(\mathcal{R}^d)$ denote the Hilbert space delivered by functional analytic completion of the pre-Hilbert space $(\xi(\mathcal{S}_N), (\cdot,\cdot)_{\mathcal{S}_N})$ in the norm topology derived from the following inner product:

$$\langle\cdot,\cdot\rangle_{H(\mathcal{R}^d)}:\xi(\mathcal{S}_N)\times\xi(\mathcal{S}_N)\to\mathbb{C}$$
$$\begin{aligned}(\xi(\mathbf{f}^{(1)}), \xi(\mathbf{f}^{(2)}))&\mapsto\langle\xi(\mathbf{f}^{(1)}), \xi(\mathbf{f}^{(2)})\rangle_{H(\mathcal{R}^d)}\\&=(\xi(f^{(1)}), \xi(f^{(2)}))_{\mathcal{S}_N}.\end{aligned}$$

$H(\mathcal{R}^d)$ is isometrically isomorphic to the Hilbert space $\mathcal{H}(\mathcal{R}^d)$ obtained by the completion of $\mathcal{S}_N=(\mathcal{S}(\mathcal{R}^d))^N$ in the norm-topology furnished by the inner product,

$$\langle\cdot,\cdot\rangle_{\mathcal{H}(\mathcal{R}^d)}:\mathcal{S}_N\times\mathcal{S}_N\to\mathbb{C},$$
$$\begin{aligned}(\mathbf{f}^{(1)}, \mathbf{f}^{(2)})&\mapsto\langle\mathbf{f}^{(1)}, \mathbf{f}^{(2)}\rangle_{\mathcal{H}(\mathcal{R}^d)}\\&=\langle\xi(\mathbf{f}^{(1)}), \xi(\mathbf{f}^{(2)})\rangle_{H(\mathcal{R}^d)}.\end{aligned}$$

$\mathcal{H}(\mathcal{R}^d)$ is called the index space for $H(\mathcal{R}^d)$.

Each of the Hilbert spaces $H(\mathcal{R}^d)$ and $\mathcal{H}(\mathcal{R}^d)$ has a quasilocal[15] structure. Take $H(\mathcal{R}^d)$, for example, and let D be any bounded open subset of $\mathcal{R}^d$ with complement D' and boundary ∂D. Define $H(D)$ as the completion in the topology of $L^2(\Omega, \mathcal{B}, \mu)$ of the linear space

$$\{\xi(\mathbf{f})\in\xi(\mathcal{S}_N): \mathbf{f}\in\mathcal{S}_N,\ \text{supp}\ \mathbf{f}\subset D\}.$$

The collection $\{H(D): D\subset\mathcal{R}^d\}$ of Hilbert spaces of random variables is a net whose ordering relation is isotonous inclusion,

$$\text{if } D_2\supset D_1,\ \text{then } H(D_1)\subset H(D_2).$$

It must now be clear that $H(\mathcal{R}^d)$ is the completion, in the norm topology of $L^2(\Omega, \mathcal{B}, \mu)$, of $\cup_{D\subset\mathcal{R}^d}H(D)$.

In case Λ is a bounded closed subset of $\mathcal{R}^d$, we define

$$H(\Lambda)=\bigcap_{D\supset\Lambda}H(D),$$

where the intersection is taken over all bounded open sets D in $\mathcal{R}^d$ which contain Λ.

Any linear space of random variables which possesses a quasilocal structure comes well equipped to sustain the important notion of Markov property. In this communication, we introduce two notions of Markov property and study the connection between them.

(2.1) Markovicity in the sense of Nelson[4]: Let $H^0(\Omega)=\{\xi(f)=(\xi_i(f): i=1,\ldots,N): f\in\mathcal{S}(\mathcal{R}^d)\}$ be a Euclidean covariant generalized stochastic field on $(\Omega, \mathcal{B}, \mu)$. Then, $H^0(\Omega)$ is said to be Markov in the sense of Nelson if, for any bounded open subset D of $\mathcal{R}^d$, and for all $u\in H(D)$,

$$(2.2)\quad E(u\,|\,H(D'))=E(u\,|\,H(\partial D)),$$

except possibly on a set of μ-measure zero.

For any bounded open or closed subset Λ of $\mathcal{R}^d$, let Q_Λ be the orthogonal projection of $H(\mathcal{R}^d)$ onto $H(\Lambda)$. Then,[1]

$$E(\cdot\,|\,H(\Lambda))=Q_\Lambda$$

and (2.2) admits the following abstract formulation:

$$(2.3)\quad Q_{D'}Q_D=Q_{\partial D}Q_D,$$

as an operator equation on $H(\mathcal{R}^d)$.

As already indicated in the Introduction, this notion of Markov property has led recently to the clarification of some mathematical and physical problems in the study of relativistic quantum fields. The other definition of Markov property of interest to us is that due to Wong.

(2.4) Markovicity in the sense of Wong[5]: Let $H^0(\Omega)=\{\xi(f)=(\xi_j(f): j=1,\ldots,N)\}$ be as in (2.1) above. Then, $H^0(\Omega)$ is said to be Markov in the sense of Wong if, for any nested increasing family ∂D_1, ∂D, ∂D_2 of boundaries in $\mathcal{R}^{d-1}$, we have that

$$H(\partial D_2)-P_{\partial D}H(\partial D_2)$$

is stochastically independent of $H(\partial D_1)$, where $P_{\partial D}$ is the projection of $H(\partial D_2)$ onto $H(\partial D)$.

(2.5) Remark: Wong's notion of Markov property was originally formulated[5] only for Gaussian generalized stochastic fields. The formulation given here is an extension and improvement on that in Ref. 5.

Most investigations[4,5,16,17] of the notion of Markov property have dealt with generalized or ordinary stochastic fields or processes obeying the Gaussian probability distribution law, because of the relatively

simple form to which the concept of stochastic independence reduces for this type of random variables. In what follows, we put the constraint on $H^0(\Omega)$ that finite collections of elements from $H^0(\Omega)$ have Gaussian probability distribution. With this assumption, stochastic independence of two members of $H^0(\Omega)$ is equivalent to their orthogonality in $L^2(\Omega, \beta, \mu)$.

It is useful to obtain an abstract formulation as in (2.3), for Wong's notion of Markov property. To this end, there is the following result:

(2.6) *Theorem*: Let ∂D_1, ∂D, and ∂D_2 be any increasing triplet of nested boundaries and let $H(\partial D_1)$, $H(\partial D)$, and $H(\partial D_2)$ be respectively, the associated boundary data Hilbert spaces. Let $P_{\partial D}$ denote the projection of $H(\partial D_2)$ onto $H(\partial D)$. Then Wong's notion of Markovicity is equivalent to the following condition:

$$Q_{\partial D_1} Q_{\partial D_2} = Q_{\partial D_1} P_{\partial D} Q_{\partial D_2}$$

as an operator equation on $H(R^d)$.

Proof: Wong's definition of Markovicity is the following statement: $H(\partial D_2) - P_{\partial D} H(\partial D_2)$ is always orthogonal to $H(\partial D_1)$. This is equivalent to the following:

$$\langle Q_{\partial D_2} u, Q_{\partial D_1} v\rangle_{H(R^d)} = \langle P_{\partial D} Q_{\partial D_2} u, Q_{\partial D_1} v\rangle_{H(R^d)}$$

for every u, v belonging to $H(R^d)$.

Thus, it follows that

$$\langle Q_{\partial D_1} Q_{\partial D_2} u, Q_{\partial D_1} v\rangle_{H(R^d)}$$
$$= \langle Q_{\partial D_1} P_{\partial D} Q_{\partial D_2} u, Q_{\partial D_1} v\rangle_{H(R^d)}.$$

Hence

$$\langle (Q_{\partial D_1} Q_{\partial D_2} - Q_{\partial D_1} P_{\partial D} Q_{\partial D_2}) u, v\rangle_{H(R^d)} = 0,$$

for every $u, v \in H(R^d)$. In particular, this is true for every $v \in H(R^d)$ and for arbitrary but fixed $u \in H(R^d)$.

Next, set

$$(Q_{\partial D_1} Q_{\partial D_2} - Q_{\partial D_1} P_{\partial D} Q_{\partial D_2}) u = \phi(u).$$

Then, since u is fixed in $H(R^d)$, so is $\phi(u)$. From the preceding, we have then that the bounded linear functional

$$F_u : H(R^d) \to \mathbb{C},$$
$$v \to F_u(v) = \langle \phi(u), v\rangle_{H(R^d)}$$

is identically zero on $H(R^d)$. Hence, by the uniqueness of any bounded linear functional on a Hilbert space, it follows that

$$\phi(u) = (Q_{\partial D_1} Q_{\partial D_2} - Q_{\partial D_1} P_{\partial D} Q_{\partial D_2}) u = 0$$

for all $u \in H(R^d)$.

Hence,

$$(*)\quad Q_{\partial D_1} Q_{\partial D_2} = Q_{\partial D_1} P_{\partial D} Q_{\partial D_2},$$

as claimed. ∎

(2.7*) Remark: If $H^0(\Omega)$ is not Gaussian as we assume it to be here, then (*) is a necessary but not sufficient condition for Markovicity in the sense of Wong.

(2.7) Definition: Say that $H^0(\Omega) = \{\xi(f) = (\xi_i(f): i = 1, \ldots, N): f \in \mathcal{S}(R^d)\}$ has the restricted Markov property of Wong if $E(uv \mid H(\partial D)) = E(u \mid H(\partial D)) E(v \mid H(\partial D))$ for every $u \in H(D)$ and $v \in H(D')$.

(2.8) Remark: There is the following connection between the restricted Markov property of Wong and Wong's notion of Markovicity.

(2.9) *Theorem*: Let $H^0(\Omega)$ have the restricted Markov property of Wong. Then, $H^0(\Omega)$ is Markov in the sense of Wong.

Proof: Let D_1, D, D_2 be bounded open subsets of R^d such that

$$D_2 \supset D \supset D_1.$$

Then, the boundaries ∂D_1, ∂D, ∂D_2 form an increasing family of nested surfaces in R^{d-1}.

Now, $H(\partial D)$, $H(\partial D)$, and $H(\partial D_2)$ are subspaces of $H(R^d)$. Furthermore, $H(\partial D_1)$ is a subspace of $H(D)$ and $H(\partial D_2)$ is a subspace of $H(D')$.

Hence, since by hypothesis $H^0(\Omega)$ has the restricted Markov property of Wong, we have

$$E(uv \mid H(\partial D)) = E(u \mid H(\partial D))\, E(v \mid H(\partial D))$$

for every $u \in H(\partial D_1) \subset H(D)$ and $v \in H(\partial D_2) \subset H(D')$. This last equation implies

$$\langle u, v\rangle_{L^2(\Omega, \beta, \mu)} = \langle u, E(v \mid H(\partial D))\rangle_{L(\Omega, \beta, \mu)}$$

or equivalently,

$$\langle u, v - E(v \mid H(\partial D))\rangle_{L^2(\Omega, \beta, \mu)} = 0.$$

Hence, we have that for every $u \in H(\partial D_1)$ and every $v \in H(\partial D_2)$, $v - E(v \mid H(\partial D))$ is stochastically independent of u. This is Wong's definition of Markovicity, and hence the claim is vindicated. ∎

(2.10) Remark: Let us now make contact with Nelson's notion of Markov property. To this end, we can assert as follows:

(2.11) *Theorem*: If $H^0(\Omega)$ is Markov in the sense of Nelson, then it has the restricted Markov property of Wong.

Proof: From Nelson's definition of Markov property, it follows that

$$E(uv \mid H(D')) = vE(u \mid H(D')) = vE(u \mid H(\partial D))$$

for all $u \in H(D)$ and all $v \in H(D')$. But $H(\partial D)$ is a subspace of $H(D')$. Hence,

$$E(E(uv \mid H(D')) \mid H(\partial D)) = E(u \mid H(D))\, E(v \mid H(D)),$$
by what precedes
$$= E(uv \mid H(\partial D)).$$

Thus, we have established that

$$E(uv \mid H(D)) = E(u \mid H(D))\, E(v \mid H(\partial D))$$

for all $u \in H(D)$ and all $v \in H(D')$. But this is the restricted Markov property of Wong as we have defined it in (2.7) above. Hence, our claim is justified. ∎

(2.12) Remark: We have obtained the following sequence of logical implications:

Nelson's notion of Markov property $\Longrightarrow$ restricted Markov property of Wong $\Longrightarrow$ Wong's notion of Markov property

That Wong's notion of Markov property is not only apparently, but indeed genuinely, weaker than Nelson's notion of Markov property emerges vividly in a paper, currently under preparation, by this author.

ACKNOWLEDGMENT

The work reported here is an infinitesimally tiny subset of the material presented by this author for the degree of Doctor of Philosophy at the University of London.[12] The author is grateful to his supervisor, Professor R. F. Streater, for the constructive suggestions and instructive criticisms he received from him.

[1]M. Loeve, *Probability Theory* (Van Nostrand, Princeton, N.J., 1963).
[2]A.N. Kolmogorov, "Uber die analytische methoden in der Wahrscheinlichkeitsrechung," Math. Ann. **104**, 415—458 (1931).
[3]G.C. Hegerfeldt, "From Euclidean to relativistic fields and on the notion of Markov fields," Commun. Math. Phys. **35**, 155—71 (1974).
[4]E. Nelson, "Construction of quantum fields from Markoff fields," J. Funct. Anal. **12**, 97—112 (1973).
[5]E. Wong, "Homogeneous Gauss—Markov fields," Ann. Math. Stat. **40**, 1625—35 (1969).
[6]G. Velo and A. Wightman, *Constructive Quantum Field Theory*, Lecture notes in Physics Vol. 25 (Springer, Berlin, 1973).
[7]B. Simon, *The $P(\phi)_2$ Euclidean (Quantum) Field Theory*, Princeton Series in Physics (Princeton U.P., Princeton, N.J., 1974).
[8]I.M. Gel'fand, M.I. Graev, and I.I. Pyatetskii-Shapiro, *Representation Theory and Automorphic Function* (Saunders, Philadelphia, 1969).
[9]N.Ya. Vilenkin, "Matrix elements of irreducible unitary representations of the group of real orthogonal matrices and the group of notions of $(d-1)$-dimensional Euclidean space" (Russian) Dokl. Akad. Nauk. SSSR **113**, 16—9 (1957).
[10]E. Hewitt and K.A. Ross, *Abstract Harmonic Analysis II* (Springer, Berlin, 1970).
[11]I.M. Gel'fand and N.Ya. Vilenkin, *Generalized Functions* 4: *Applications of Harmonic Analysis* (Academic, New York, 1964).
[12]G.O.S. Ekhaguere, "A Markov property for multicomponent Euclidean covariant, Gaussian generalized stochastic fields," thesis, University of London (1976).
[13]A.M. Yaglom, second order homogeneous random fields. Fourth Berkeley Symp. Math. Stat. Prob. 2, 593—622 (1961).
[14]A.S. Wightman, Partial differential equations and relativistic quantum field theory, in *Differential Equations Vol. II*, edited by A.K. Aziz (Van Nostrand, Princeton, N.J., 1969).
[15]K. Schmidt, Algebras with quasilocal structure and factorizable representations in *Mathematics of Contemporary Physics*, edited by R.F. Streater (Academic, London, 1972).
[16]H.P. McKean, Jr., "Brownian motion with a several dimensional time," Theory Prob. Appl. 8, 335—54 (1963).
[17]L.D. Pitt, "A Markov property for Gaussian Processes with a multidimensional parameter," Arch. Rat. Mech. Anal. **43**, 367—91 (1971).

Evolution of a stable profile for a class of nonlinear diffusion equations with fixed boundaries

James G. Berryman[a)]

Department of Physics, Mathematics Research Center, and Department of Electrical and Computer Engineering, University of Wisconsin, Madison, Wisconsin 53706
(Received 14 December 1976)

A class of quasilinear parabolic equations with fixed boundaries arising in studies of cross-field diffusion in toroidal multipole plasmas is presented. It is well known that these equations have separable solutions which decay in time. Surprisingly, both octupole and numerical experiments show, in particular cases, that the separable solution evolves from an arbitrary initial distribution of particles. The evolution and stability properties of these solutions are demonstrated in this paper. When the coefficients of the equations are independent of the spatial variable, infinitesimal perturbations decay as the fourth power (or higher) of the separable solution time dependence; the separable solution is therefore stable. When the initial particle distribution has no nulls except at the boundaries, an approximate analysis shows that large perturbations decay exponentially causing the rapid evolution of the separable solution. The analysis allows the asymptotic behavior of the system to be predicted approximately from knowledge of the initial particle distribution.

I. INTRODUCTION

A series of experiments on particle diffusion across magnetic fields in the Wisconsin toroidal octupole plasma containment device has been conducted by Drake, Greenwood, Navratil, and Post.[1] These experiments were performed with a purely poloidal field. They found that a density profile developed which decayed in amplitude but otherwise was essentially independent of time. This remarkable time independent density profile was dubbed the "normal mode" of the system and corresponds (as we will show) to the separable solution of the relevant nonlinear diffusion equation. Reduced to standard form and normalized units, the equation of interest is

$$F(x)\frac{\partial n}{\partial t}=\frac{\partial}{\partial x}\left[D(n)\frac{\partial n}{\partial x}\right] \text{ for } 0\le x\le 1, \tag{1}$$

where n is the particle density, x is the spatial variable in one-dimension, and t is the time. The geometrical factor $F(x)$ is a positive function determined by the octupole geometry. The diffusion coefficient $D(n)$ is a nonlinear function of the density. In the experiments of Ref. 1, D was found experimentally to scale like Okuda–Dawson diffusion[2]

$$D(n)\propto n^{-1/2}, \tag{2}$$

but in other density and field strength regimes the scaling is different. In general, the dependence of the diffusion coefficient on the density can be parametrized by taking

$$D(n)\propto n^{\delta}, \tag{3}$$

with $\delta\ge -1$. We may treat (1) with D specified by (3) as a mathematical model of the physical problem and determine what the predictions of that model are.

The present paper will treat the analytical properties of Eq. (1). The questions to be answered are: (a) What is the nature of the separable solution and how does the geometrical factor affect it? (b) Is the separable solution stable against infinitesimal perturbations? (c) Does an arbitrary initial distribution of particles evolve into the separable solution? These questions are given satisfactory answers in the sections which follow. The geometrical factor can cause the shape of the separable solution to be asymmetrical and affects the decay rate but has little effect on stability. Infinitesimal perturbations are shown to decay faster than the separable solution so stability is established. Furthermore, it is shown that the separable solution will rapidly evolve out of an arbitrary initial particle distribution as long as the density does not vanish anywhere except at the boundaries. Numerical experiments confirm this conclusion. In a second paper published elsewhere,[3] the model equation (1) is derived phenomenologically, and the theoretical predictions for the decay constants and shape of the separable solution are compared to the experimental results. A more detailed discussion of the effects of the geometrical factor $F(x)$ is also given there.

II. A MODEL OF NONLINEAR DIFFUSION

Equation (1) can be put into a form more convenient for both analytical and computational purposes. The proportionality constants for (3) can be chosen for convenience to give

$$D(n)=(1+\delta)n^{\delta} \text{ for } \delta>-1. \tag{4}$$

(We exclude the case $\delta=-1$. A similarity solution exists for this case,[4] but the boundary conditions and analysis must be treated by methods different from those required in the remainder of this paper.) A new dependent variable can be defined as

$$m(x,t)\equiv n^{1+\delta} \text{ for } \delta>-1, \tag{5}$$

which satisfies

$$F(x)(m^{q-1})_t=m_{xx}, \tag{6}$$

where

a)Present address: Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012.

 Copyright © 1977 American Institute of Physics

$$q \equiv (2+\delta)/(1+\delta) \tag{7}$$

and subscripts indicate partial derivatives. For $\delta > -1$, we see that $0 \le m(x,t) < \infty$ when $0 \le n(x,t) < \infty$ and furthermore $m = 0$ when $n = 0$. We will call $m(x,t)$ the pseudodensity associated with $n(x,t)$ since it is nonnegative and behaves very much like a density distribution.

The geometrical factor $F(x)$ is positive in the cases of physical interest and has a singularity at some point $x = x_s$. This singularity is integrable. (For the toroidal octupole, this singularity corresponds to the singularity in the function giving the flux tube volume per poloidal flux increment.) Typically, the singularity may be of either square root or logarithmic type. To include the case of physical interest in our studies of (6), we must consider F's which satisfy

$$F(x) > 0 \quad \text{for } 0 \le x \le 1, \tag{8}$$

$$\int_0^1 F(x)\,dx < \infty, \tag{9}$$

and allow F to have a singularity for some $x = x_s$. In addition, it is consistent and convenient to $F'(x) \ge 0$ for $x < x_s$ and $(x) \le 0$ for $x > x_s$, where $F' = d/dx$.

The general analysis that follows will apply for all $\delta > -1$. Since the case of most interest in Ref. 1 is $\delta = -\frac{1}{2}$, the detailed calculations will be restricted to treating this case as an example. If in addition we assume the geometrical factor $F(x)$ is not of crucial importance, we may study the especially simple case of $F(x) = 1$ and $\delta = -\frac{1}{2}$ or

$$2m\, m_t = m_{xx}. \tag{10}$$

The density $n(x,t) = m^2(x,t)$ [from (5) and (7) with $\delta = -\frac{1}{2}$ or $q = 3$] obtained by solving (10) for m will be shown to have many of the properties of the density profiles found experimentally.

Two types of boundary conditions are of interest: (a) For short times after the plasma has been injected into the octupole, the bulk of the particle density is localized and begins diffusing towards the boundaries. Thus, for short times the density is essentially unaffected by these boundaries. This situation is approximated by supposing there are no finite boundaries, i.e., consider $-\infty < x < \infty$. (b) For long times, a significant fraction of the particle density is in the vicinity of the boundaries. The experimental results on the toroidal octupole are well approximated by taking $n = 0$ at the boundaries. Using our transformed variables, this condition corresponds to the boundary condition for the pseudodensity given by

$$m(0,t) = m(1,t) = 0. \tag{11}$$

For $\delta > -1$, (11) is consistent with the physical requirement of finite flux.

Exact solutions to (6) can be found when $F(x) = 1$ assuming the boundary conditions discussed previously. The method of similarity transformation discussed by Boyer[5] and Ames[6] in a similar context (i.e., $\delta > 0$) can be employed. Since the method is well-documented elsewhere, we will simply quote the results here.

Assuming an $m(x,t)$ of the form

$$m(x,t) = S[x/R(t)]T(t) \tag{12}$$

satisfies (6), we find for $\alpha \neq 0$

$$S(z) = \begin{cases} S_0(1 - \sigma z^2/z_0^2)^{1/\sigma} \\ 0 \quad \text{for } \sigma z^2 > z_0^2 \text{ when } \sigma > 0, \end{cases} \tag{13}$$

and for all α

$$T(t) = \begin{cases} A_0(1 - r t/\tau)^{1/r} \\ 0 \quad \text{for } t > \tau/r \text{ when } r > 0, \end{cases} \tag{14}$$

and

$$R(t) = R_0 T^{-\alpha}(t), \tag{15}$$

where

$$z_0^2 = 2S_0^{\sigma}\lambda^{-1} \quad \text{and } \sigma = 2 - q, \tag{16}$$

$$\tau = (q-1)R_0^2 A_0^r \lambda^{-1} \text{ and } r = q - 2\alpha - 2. \tag{17}$$

In Eqs. (15)—(17), λ is the separation constant and α is determined by the boundary conditions. A_0, R_0, and S_0 are arbitrary constants.

Allowing x to range over all $-\infty < x < \infty$, the physical boundary condition[5] is that the total number of particles is conserved during the diffusion assuming no sources or sinks are present. This implies

$$N(t) = \int_{-\infty}^{\infty} n(x,t)\,dx = [T(t)]^{q-1} R(t) \int_{-\infty}^{\infty} S^{q-1}(z)\,dz \tag{18}$$

is a constant. Using (15), we find that (13) is a valid shape function if and only if

$$\alpha = q - 1. \tag{19}$$

For the second physically interesting boundary condition, $n = 0$ at the boundaries. It is easy to prove (and physically obvious) that R must be time independent so that

$$\alpha = 0. \tag{20}$$

This is the case of principal interest in this paper and is treated in Sec. III.

III. THE SEPARABLE SOLUTION

In Sec. II, we found that (6) could be separated for $0 \le x \le 1$. The general time dependence was calculated. A detailed analysis of the separable solution shape function is given in this section.

A. Exact results

For $\delta > -1$, $\alpha = 0$, and $F(x) = 1$, the shape function satisfies

$$S'' + \lambda S^{q-1} = 0, \tag{21}$$

where $' = d/dx$. Note that if $S_0(x)$ is a solution with separation constant λ_0, then $S(x) = aS_0(x)$ is also a solution with separation constant $\lambda = \lambda_0 a^{q-2}$, where a is any positive number. Therefore, we can scale S so that $0 \le S \le 1$ without loss of generality. Equation (21) can then be integrated to give the implicit relation

$$I(S) = \int_0^S \frac{dy}{(1 - y^q)^{1/2}} = \rho x \quad \text{for } 0 \le x \le \tfrac{1}{2}, \tag{22}$$

where

$$\rho^2 = 2\lambda/q. \tag{23}$$

Since S is necessarily symmetric around $x = \frac{1}{2}$, $S(x)$ for $\frac{1}{2} \le x \le 1$ can be found using the identity $S(1-x) = S(x)$.

The integral on the left can be expressed in terms of the incomplete beta function or in terms of Gauss' hypergeometric function as

$$I(S) = \frac{1}{q} B_S\left(\frac{1}{q}, \frac{1}{2}\right) = S^{1/q}\, {}_2F_1\left(\frac{1}{q};\ 1+\frac{1}{q};\ S\right). \tag{24}$$

No simple methods of inverting the hypergeometric function are known except for special values of its arguments. Therefore, Eqs. (22)—(24) are not very useful for determining $S(x)$ analytically. If desired, these equations can at least be used to determine the qualitative nature of $S(x)$.

However, certain special values can be obtained permitting the exact evaluation of some relevant constants. For example, from (22) we find

$$\rho = 2I(1) = \frac{2}{q}\frac{\Gamma(1/q)\Gamma(\frac{1}{2})}{\Gamma(\frac{1}{2}+1/q)} \tag{25}$$

where Γ is the gamma function. Furthermore, the eigenvalue λ is expressible in terms of ρ as

$$\lambda = \tfrac{1}{2} q \rho^2. \tag{26}$$

Thus the eigenvalue is known exactly for all q. In addition, one can show that

$$\gamma \equiv \int_0^1 S^{q-1}(x)\, dx = 4/q\rho \tag{27}$$

using (22). γ is just the integral of the physical density's shape function. The total number of particles in the separable solution obeys

$$N(t) = \int_0^1 n(x, t)\, dx = \gamma T^{q-1}(t) \tag{28}$$

for all time.

Table I gives values of q, ρ, γ, and λ for various values of the nonlinearity parameter δ. We note that as δ varies continuously between $+\infty$ and -1, q varies monotonically between 1 and $+\infty$, ρ varies between 4 and 2, γ between 1 and 0, and λ between 8 and $+\infty$. The values given for $\delta = -1$ should be thought of as limiting values as $\delta \to -1^+$ since the analysis in the previous paragraph is not valid for $\delta = -1$. The value of γ is important for reasons to be discussed in Sec. V. The values of λ are important for determining the time constant (17).

TABLE I. Values of q, ρ, λ, and γ for selected values of δ. The defining equations given in the text are (7), (25), (26), and (27), respectively. The slope of the separable solution shape function is $\pm\rho$ near the boundaries. The separation constant is λ. The total number of particles in the separable solution is proportional to γ.

δ	q	ρ	λ	γ
∞	1	4	8	1
1	$\frac{3}{2}$	3.4495	8.9242	0.77306
0	2	π	π^2	$2/\pi$
$-\frac{1}{2}$	3	2.8044	11.7967	0.47545
$-\frac{1}{3}$	4	2.6221	13.7504	0.38138
$-\frac{3}{7}$	5	2.5075	15.7184	0.31905
-1^+	∞	2	∞	0

B. Numerical methods

Since the shape function for the separable solution of (6) is not known in closed form even with $F(x)$ constant, it would be useful to construct numerical methods for calculating the shape functions. We construct a successive approximation method valid for arbitrary F satisfying Eqs. (8) and (9). The case $\delta = -\frac{1}{2}$ is treated as an example.

If the function $F(x)$ containing the geometrical effects in the diffusion problem satisfies conditions (8) and (9), it is no more difficult to construct an iterative method[7] for the more general problem

$$-S''(x) = \lambda f(x, S) \equiv \lambda F(x) S^{q-1}(x) \tag{29}$$

than it would be for (21). It is straightforward to show that (29) is the equation for the shape function which derives from (6).

It is natural to consider monotone iteration methods such as those proposed by Keller and Cohen[8] and reviewed in an abstract setting by Amann.[9] Straightforward application of these methods is not possible however. The convergence proofs given in Ref. 8 require two assumptions which are violated here: (i) continuity of $f(x, S)$ and (ii) $f(x, 0) > 0$. The singularity in the geometrical factor $F(x)$ violates (i). More importantly, the fact that $f(x, 0) = 0$ when $q \ge 2$ (thus admitting $S = 0$ as a solution) violates (ii). In fact, it is easy to construct examples where a straightforward iteration procedure with $\lambda = 1$ will rapidly generate the zero solution in a machine computation. A discussion of some relevant aspects of Amann's work is given later in this section.

The difficulties discussed in the previous paragraph can be avoided by using the following device. First, solve the linear boundary value problem

$$\Sigma_i''(x) = -F(x) S_{i-1}^{q-1}(x), \qquad \Sigma_i(0) = \Sigma_i(1) = 0 \tag{30}$$

taking $S_0(x) = \phi(x)$ where $0 \le \phi$ and ϕ is some continuous integrable function, not identically zero. Then, define

$$\lambda_i^{-1} = \max_{0 \le x \le 1} \Sigma_i(x) \tag{31}$$

and

$$S_i(x) = \lambda_i \Sigma_i(x). \tag{32}$$

Eqs. (31) and (32) merely normalize the amplitude of S_i so its maximum value is unity. Defining the Green's function

$$G(x, \xi) = \begin{cases} x(1-\xi) & \text{for } x \le \xi \\ \xi(1-x) & \text{for } x \ge \xi, \end{cases} \tag{33}$$

we can write the solution explicitly as

$$S_i(x) = \lambda_i \int_0^1 G(x, \xi) F(\xi) S_{i-1}^{q-1}(\xi)\, d\xi. \tag{34}$$

Thus, if $S_{i-1} = S_i$, (34) gives the solution of (29).

If the amplitude adjustment in (31) and (32) is not done, then the results obtained depend strongly on the amplitude of the initial guess $\phi(x)$. A small initial amplitude leads to a declining (diverging) sequence $\{S_i\}$ for $q > 2$ ($q < 2$) while a large amplitude produces the opposite result. On the other hand, if q is not too large, the normalization is not required at every step of the computation. With high precision, the *shape* produced at each stage is clearly the same with or without the normalization. However, searching for the maximum of Σ_i requires an insignificant expenditure of computer time compared to the integration in (34). Therefore, normalization at each step of the computation is recommended.

The method proposed in the preceding paragraphs has been tested numerically. It is both stable and convergent for $q = 3$. Whether this iteration procedure converges for all admissible q and $F(x)$ is not known. However, it is clear that the iteration process is stabilized by the normalization in (32): The possibility of generating either vanishing or diverging iterates is eliminated. It has been observed empirically that, starting with $\phi(x) = 4x(1-x)$, a monotonic decreasing sequence of iterates is generated when $q = 3$. It may be true that, for some choices of q and $F(x)$, this iteration procedure results in an oscillating sequence, i.e., $S_i \to S_{i+1}$ and $S_{i+1} \to S_i$. If this happens, convergence cannot occur. However, such oscillations have never been observed in the cases tested by Drake and the author.[3]

One additional fact about Eq. (34) is worth noting. Assume that $F(x)$ (a) is symmetric about $x = \frac{1}{2}$, (b) is monotonic nondecreasing on $(0, \frac{1}{2})$, and (c) satisfies Eqs. (8) and (9). Then a tedious graphical analysis of the class of functions S_i which can be generated from (34) shows that

$$4G(x, \tfrac{1}{2}) \le S_i(x) \le 4x(1-x) \tag{35}$$

for all $i \ge 1$ and all $1 \le q \le \infty$. The Green's function G is given by (33). The equalities in (35) are satisfied for all x if and only if (a) $F(x) = \delta(x - \frac{1}{2})$ for the left-hand equality [$\delta(x)$ is the Dirac delta function] or (b) $q = 1$ and $F(x) = \text{const}$ for the right-hand equality. Equation (35) implies that the slope of the shape function at $x = 0$ must be between 2 and 4. Table I shows that $2 \le \rho \le 4$ in agreement with this result.

Thus, the operator on the right-hand side of (34) maps its domain into a compact set. If a monotonic sequence S_i is generated by (34) [this is true for $q = 3$], that sequence must converge since it is bounded. A general proof of convergence could be obtained using Theorem 6.1 of Ref. 9 if we could show that the operator [including λ_i defined by (31)] on the right of (34) is "increasing." [An operator M is increasing if $u(x) \le v(x)$ implies $M(u) \le M(v)$.] The author has been unable to show that this is true for (34). Although such a theoretical result would be gratifying, it is not essential for practical computations. The computation itself will tell us very quickly whether or not convergence takes place.

C. The geometrical factor

The geometrical factor $F(x)$ can have three effects on the separable solution: (a) It can alter the shape of $S(x)$. If $F(x)$ is asymmetrical, then $S(x)$ is asymmetrical. In particular, if $F(x)$ has an integrable singularity for $x_s \ne \frac{1}{2}$, then one can give a variational argument which indicates that the peak of $S(x)$ lies somewhere in $[\frac{1}{2}, x_s]$, not necessarily at $x = x_s$. (b) As can be seen in (31), the eigenvalue λ depends inversely on the integral of $F(x)$. Hence, the decay rate of the separable solution is also dependent on the geometrical factor. (c) The stability of the separable solution can conceivably be affected by $F(x)$ if the decay rates of perturbations to the separable solution are altered substantially more than the separable solution's decay rate.

Of these three effects, the first two produce quantitative changes but only the third can cause a significant qualitative change in the temporal behavior of the density profile. The stability of the separable solution to infinitesimal perturbations is examined in the next section. As we shall see, the presence of nonconstant $F(x)$ does not appear to alter the qualitative behavior of the density profile. Even an integrably singular $F(x)$ does not change the decay rate of the perturbations enough to make the separable solution unstable for $\delta \ge -\frac{1}{2}$.

IV. STABILITY OF THE SEPARABLE SOLUTION

In this section the effects of infinitesimal perturbations on the stability of the separable solution are examined. For $\delta > -1$ and $F = 1$, Eq. (6) becomes

$$(m^{q-1})_t = m_{xx}. \tag{36}$$

Suppose that m is the separable solution plus a small perturbation

$$m(x, t) = S(x)T(t) + u(x)v(t). \tag{37}$$

The perturbation is assumed separable. This assumption does not restrict the generality of the arguments. Since the perturbation is small, linear equations for u and v are obtained by substituting (37) in (36) and linearizing. The result is

$$(q-1)T^{q-2}[\ln(vT^{q-2})]_t = u_{xx}S^{2-q}/u = -\kappa, \tag{38}$$

where T and S are given by (14), for $\alpha = 0$, and (22) respectively. κ is the separation constant for the perturbation.

The equation for v can be integrated and yields v within a multiplicative constant as

$$v(t) = T^p(t), \tag{39}$$

where

$$p = 2 - q + \kappa/\lambda. \tag{40}$$

Since T decreases as $t \to \infty$, v decreases as long as $p > 0$ and decreases faster than the separable solution for $p > 1$. Stability requires $p > 1$ for all perturbations which are not simply perturbations of the separable solution's amplitude.

κ (and therefore the time dependence of v) is determined by the eigenvalue problem

$$u_i''(x) + \kappa_i S^{q-2}(x) u_i(x) = 0 \tag{41}$$

with $u(0) = u(1) = 0$. Equation (41) is a Sturm—Liouville equation. A theorem of Coddington and Levinson[10] may be applied. The theorem states that there are an infinite number of eigenvalues $\kappa_0, \kappa_1, \cdots$ with $\kappa_0 \leq \kappa_1 \leq \kappa_2 \cdots$ and $\kappa_i \to \infty$ as $i \to \infty$. The eigenfunction u_i corresponding to κ_i has exactly i zeros in (0, 1). Notice that $u_0 = S$ satisfies (41) with $\kappa_0 = \lambda$. S has no zeros in (0, 1); therefore, $\kappa_0 = \lambda > 0$ is the smallest eigenvalue. Such a perturbation corresponds to a change in the initial amplitude of the separable solution. The exponent of the time dependence for this mode is $p = 3 - q$. For all perturbations,

$$p \geq \begin{cases} 1 & \text{for } \delta \geq 0 \\ 0 & \text{for } \delta \geq -\frac{1}{2}. \end{cases} \tag{42}$$

The separable solutions for $\delta \geq 0$ are obviously stable. For $0 \geq \delta \geq -1$, more argument is required.

To see that the lowest mode of (41) is just an amplitude perturbation and therefore ignorable, recall the form of the separable solution

$$m(x, t) = S(x)T(t) = S(x)A_0(1 - dA_0^{\gamma} t)^{1/r}, \tag{43}$$

where d is a constant. Then it is easy to show that

$$\frac{\partial m}{\partial A_0} = S(x) A_0^{q-3} T^{3-q}(t). \tag{44}$$

General arguments can be given to show that (44) must satisfy the linearized perturbation equation since it is the first (linear) term in a Taylor series expansion of (43) in the amplitude. The result $p = 3 - q$ is not surprising when seen in this light. This analysis also indicates that this lowest mode cannot affect the stability of the separable solution. Only perturbations with one or more zeros in (0, 1) are important to stability. Such perturbations must decay with exponents greater than $p_0 = 3 - q$. How much greater can only be determined by solving the eigenvalue problem (41).

If we can show that κ_1 is large enough so that $p_1 > 1$ for all δ, then stability will be established. We seek a solution of (41) which (i) vanishes only once in (0, 1) and (ii) also vanishes at the boundaries. It is easy to find a function which satisfies (i). The function $S'(x)$ obviously satisfies this condition but it does not satisfy (ii). Since S itself satisfies (ii), we try

$$u_1(x) = S(x)S'(x). \tag{45}$$

Using the identity $S'^2 = \rho^2(1 - S^q)$, we find (remarkably) that this ansatz does satisfy (41) for all q and has eigenvalue

$$\kappa_1 = q(\lambda + \rho^2) = (2 + q)\lambda. \tag{46}$$

Substituting (46) into (40), we find that *for all* q (!)

$$p_1 = 4. \tag{47}$$

Thus, the lowest nontrivial perturbation decays four times faster than the separable solution. All higher perturbations decay still faster. The stability of the separable solution has therefore been established in the geometry free $F(x) = 1$ case.

When $F(x) \neq \text{const}$, another argument is required. Note that substituting (37) into (6) gives an equation like (38) whose time dependent part is identical but whose spatially dependent part gives

$$u_i''(x) + \kappa_i F(x) S^{q-2}(x) u_i(x) = 0. \tag{48}$$

Clearly $u_0 = S$ and $\kappa_0 = \lambda$ is again a solution. The trick used above to find u_1 does not work in this case. Instead we will attempt to find a useful lower bound for p_1. Combining the equations (48) for $i = 0$ and $i = 1$, we find in general that

$$\lambda S(x) u_1''(x) = \kappa_1 u_1(x) S''(x). \tag{49}$$

We now assume $F(x)$ is symmetric about $x = \frac{1}{2}$. Integrating (49) on $[0, \frac{1}{2}]$ and using the facts that $S'(\frac{1}{2}) = 0$ and $S(\frac{1}{2}) = 1$, we find

$$\kappa_1/\lambda = 1 - u_1'(\tfrac{1}{2}) \left[\int_0^{1/2} u_1'(x) S'(x)\, dx \right]^{-1}. \tag{50}$$

For convenience, we make the convention that $u_1(x) \geq 0$ for $0 \leq x \leq \frac{1}{2}$. Then a simple graphical construction shows that

$$u_1'(\tfrac{1}{2}) < 0. \tag{51}$$

Furthermore, by integrating the denominator of (50) by parts, we find that

$$\int_0^{1/2} u_1'(x) S'(x)\, dx = \lambda \int_0^{1/2} F(x) S^{q-1}(x) u_1(x)\, dx > 0. \tag{52}$$

Hence, the ratio of these two factors is negative and the contribution to κ_1 is positive.

For $i = 1$, multiply (48) by $u_1'(x)$ and integrate. We find

$$u_1'^2(x) = u_1'^2(0) + \kappa_1 \int_0^x dx\, u_1^2 [F' S^{q-2} + (q-2) F S^{q-3} S'] - \kappa_1 F(x) S^{q-2}(x) u_1^2(x). \tag{53}$$

Assuming $F'(x) \geq 0$ for $0 \leq x \leq \frac{1}{2}$, we see that the integrand in (53) is nonnegative for $0 \leq x \leq \frac{1}{2}$. From the symmetry of the problem, $u_1(\frac{1}{2}) = 0$ and also $u_1''(\frac{1}{2}) = 0$. Thus, the last term in (53) vanishes at $x = \frac{1}{2}$ since $u_1''(\frac{1}{2}) = 0$ implies

$$\lim_{x \to 1/2} F(x) u_1(x) = 0. \tag{54}$$

We conclude that

$$u_1'^2(x) \leq u_1'^2(\tfrac{1}{2}) \tag{55}$$

for all $0 \leq x \leq 1$. Therefore, again using $S(\frac{1}{2}) = 1$, we find

$$\left| \int_0^{1/2} u_1'(x) S'(x)\, dx \right| < \left| u_1'(\tfrac{1}{2}) \right|. \tag{56}$$

Strict inequality in (56) is true because $u_1'(x)$ must change sign somewhere in $(0, \frac{1}{2})$. Using (56) in (50), we find

$$\kappa_1/\lambda > 2. \tag{57}$$

Equation (57) is a rather weak result since we have not used the oscillation of the integrand in (56) to improve the bound. However, it is sufficiently strong to prove stability for $\delta \geq -\frac{1}{2}$ because (57) implies

$$p_1 = 2 - q + \kappa_1/\lambda > 4 - q. \tag{58}$$

Thus, $p_1 > 1$ for all $q \leq 3$ regardless of the geometrical factor F. Stronger bounds can no doubt be found.

We have proven that the separable solution is stable against infinitesimal perturbations for all $\delta > -1$ when $F(x) = \text{const}$. We have also proven stability for $\delta \geq -\frac{1}{2}$ when $F(x)$ is symmetric and $F'(x) \geq 0$ for $0 \leq x \leq \frac{1}{2}$. We conjecture stability for all $\delta > -1$ and arbitrary $F(x)$ satisfying (8), (9), and the monotonicity condition $F'(x) \geq 0$ for $x \leq x_s$.

V. EVOLUTION OF THE SEPARABLE SOLUTION

Infinitesimal perturbations were shown to decay more rapidly than the separable solution in Sec. IV. Although this is an interesting result in itself, the physically relevant problem is to consider initial density distributions that are far from the separable solution, i.e., with finite perturbations. The experimental situation is clearly of this nature. Particles are initially injected into the containment device with some arbitrary spatial distribution. It is then experimentally observed[1] that after a finite time, the bulk of the particles appear to be in the "normal mode."

Numerical experiments have been performed for $\delta = -\frac{1}{2}$ with $F(x) = 1$ using Eq. (10). The initial values $m(x, 0)$ were chosen according to

$$m(x, 0) = \sum_{l=0}^{3} a_l \sin(l+1)x, \tag{59}$$

using four different sets of values for the $\{a_l\}$. The four cases tested were (a_1, a_2, a_3, a_4) equal to: (i) (1, 0.4, 0, 0), (ii) (1, 0, 0.3, 0), (iii) (1, 0, −0.3, 0), and (iv) (1, 0, 0, 0.225). Equation (10) was integrated using a linear three-level difference scheme developed by Lees[11] for quasilinear parabolic equations. In all four cases, the particle distribution decays into a distribution numerically indistinguishable from the separable solution by $t = 0.1$. In all four cases, all of the particles escape before $t = 0.2$.

Since both the plasma experiment and the numerical experiment indicate that the separable solution evolves from arbitrary initial data in a finite time, it would be satisfying to show analytically why this should happen. This analysis would be most convincing if we could predict the effective amplitude of the separable solution that evolves out of an initial particle distribution. Such an anlysis is given in this section.

A. Prediction of the asymptotic amplitude

An analysis predicting the final amplitude of the separable solution $\delta > -1$ is given here. This analysis is followed by comparison to the results of numerical experiments for $\delta = -\frac{1}{2}$.

Since (48) is a Sturm–Liouville equation,[12] the functions u_i form a complete, orthonormal set with

$$\int_0^1 dx\, F(x) u_0^{q-2}(x) u_i(x) u_j(x) = c\delta_{ij}, \tag{60}$$

where by assumption $0 \leq u_0 \leq 1$ and

$$\int_0^1 dx\, F(x) u_0^q(x) \equiv c. \tag{61}$$

The other u_i's are normalized to give (60).

An arbitrary function vanishing at $x = 0$ and $x = 1$ can be expanded in the complete set $\{u_i\}$. In particular,

$$m(x, t) = \sum_{i=0}^{\infty} a_i(t) u_i(x) \tag{62}$$

where the a_i's are time dependent amplitudes. The equation of motion for the a's is obtained by substituting (62) into (36) and using (41) to simplify the right-hand side. Noting that $S = u_0$, we obtain

$$(q-1)F(x)\sum_{j=0}^{\infty} \dot{a}_j(t) u_j(x) = -F(x) u_0^{q-2}(x) \sum_{i=0}^{\infty} a_i(t) \kappa_i u_i(x) \times \left[\sum_{l=0}^{\infty} a_l(t) u_l(x)\right]^{2-q}. \tag{63}$$

To simplify notation, a further definition is useful

$$\sum_{i=0}^{\infty} b_{ji}(t) u_i(x) \equiv u_j(x)[a_0(t) u_0(x)]^{q-2} \left[\sum_{i=0}^{\infty} a_i(t) u_i(x)\right]^{2-q}. \tag{64}$$

Equation (64) again uses the completeness of the set $\{u_i\}$. Equation (63) then becomes

$$F(x)\sum_j \dot{a}_j(t) u_j(x) = -(q-1)^{-1} F(x) a_0^{2-q} \sum_{ij} \kappa_j a_j b_{ji} u_i(x). \tag{65}$$

Multiplying by $u_0^{q-2} u_i$ and integrating with the help of (60), the equations of motion become

$$\dot{a}_i(t) = -(q-1)^{-1} a_0^{2-q} \sum_{j=0}^{\infty} \kappa_j b_{ji} a_j. \tag{66}$$

No approximations have been made in deriving (66), which gives an infinite set of nonlinear equations for the time dependent amplitudes. In this form, (66) is actually harder to solve exactly than the original nonlinear equation. However, it is not difficult to obtain approximations to the b_{ji}'s which permit progress to be made.

Assuming that $m(x, t)$ does not vanish in the interior of (0, 1), the product in (64) can be expanded as

$$[a_0 u_0(x)]^{q-2}\left[\sum_{i=0}^{\infty} a_i u_i(x)\right]^{2-q} \approx 1 + (2-q)\chi(x, t) + 1/2(2-q)(1-q)\chi^2(x, t) + \cdots, \tag{67}$$

where

$$\chi(x, t) = \sum_{i=1}^{\infty} a_i u_i / a_0 u_0. \tag{68}$$

We will assume that the expansion (67) converges and that the first two terms are dominant.

To obtain a first approximation to $a_0(t)$ for finite amplitude perturbations, we must retain terms to $O(a_j^2)$, $j \neq 0$, in (66) for $i = 0$. Since the right-hand side of (66) is of order $b_{ji} a_j$, we only need to retain $O(a_j)$ in computing b_{j0}. Furthermore, to obtain a_i $(i \neq 0)$ correct to $O(a_j)$, $j \neq 0$, in (66), we only need to retain terms of $O(1)$ in b_{ji} as is easily seen. Then retaining terms to the specified order, we find

$$b_{ji}(t) = \delta_{ji} + \frac{2-q}{c a_0(t)} \int_0^1 dx\, u_0^{q-3} u_j u_i F_{l=1} a_l u_l + O(a_k a_l) \cdots. \tag{69}$$

The most important special cases of (69) are, for $i=0$, when

$$b_{j0}(t)=\delta_{j0}+(2-q)a_j(t)/a_0(t)+O(a_k a_l) \tag{70}$$

and, for $i\neq 0$, $j\neq 0$, when

$$b_{ji}(t)\approx\delta_{ji}+O(a_k). \tag{71}$$

In (69)—(71), k and l satisfy $1\leq k,\ l\leq\infty$.

Substituting (70) and (71) into (66), we find

$$(q-1)a_0^{q-3}\dot{a}_0=-\lambda+(q-2)\sum_{i=1}^{\infty}\kappa_i\left(\frac{a_i}{a_0}\right)^2+\cdots \tag{72}$$

and

$$(q-1)\dot{a}_i=-p_i\lambda a_0^{2-q}a_i+\cdots, \tag{73}$$

where p_i is given by Eq. (40).

Equation (73) can be integrated to yield

$$a_i(t)\approx a_i(0)\exp\left[-\frac{p_i\lambda}{q-1}\int_0^t a_0^{2-q}(t)\,dt\right]. \tag{74}$$

If a_0 is nearly constant for small t, then

$$a_i(t)\approx a_i(0)\exp\{-[p_i\lambda/(q-1)]a^{2-q}(0)t\}. \tag{75}$$

Therefore, the perturbations of finite amplitude decay exponentially at times short enough for a_0 to be approximately constant.

Substituting (75) into (72) and integrating, a_0 is found to be

$$a_0(t)=a_0(0)\exp(-\lambda t)\quad\text{for } q=2 \tag{76}$$

and

$$a_0^{q-2}(t)\approx a_0^{q-2}(0)-\left(\frac{q-2}{q-1}\right)\lambda t+\frac{(q-2)^2}{2}a_0^{q-4}(0)\times\sum_{i=1}^{\infty}[a_i^2(0)-a_i^2(t)]\quad\text{for } q\neq 2. \tag{77}$$

The linear case has $q=2$, and in that case (76) is exact. Equation (77) is the principal result of this section. Since $a_i(t)$ vanishes exponentially from (75), the asymptotic time dependence of a_0 for $q\neq 2$ is

$$a_0(t)\approx A_0[1-rt/\tau]^{1/r}, \tag{78}$$

where $r=q-2$,

$$A_0\approx a_0(0)\left[1+\frac{r^2}{2}\sum_{i=1}\frac{a_i^2(0)}{a_0^2(0)}\right]^{1/r}, \tag{79}$$

and

$$\tau=(q-1)A_0^r\lambda^{-1};$$

cf. Eqs. (14) and (17).

In general, an arbitrary initial distribution of particles satisfying the boundary conditions will decay into the separable solution. Furthermore, Eq. (79) predicts the amplitude of the asymptotic state. A_0 can be calculated from the initial values of the pseudodensity m if $S=u_0$ is known. This conclusion follows from the fact that the integrals

$$\int_0^1 F(x)u^{q-1}(x)m(x,t)\,dx=ca_0(t), \tag{80}$$

$$\int_0^1 F(x)u^{q-2}(x)m^2(x,t)\,dx=c\sum_{i=0}^{\infty}a^2(t)\equiv cB(t), \tag{81}$$

and

$$\int_0^1 F(x)u_0^q(x)\,dx=c \tag{82}$$

can be found numerically if u_0 and m are known.

Certain features of the preceding analysis should be given some additional attention before proceeding to some numerical examples.

First, the principal approximation involves the expansion in (67). This expansion converges for all x when $2-q\geq 0$ and for all x except x's for which $m(x,t)=0$ when $2-q<0$. The convergence may be quite slow, however, when $2-q\ll -1$. Thus, we may argue that the approximations we have made are valid for $1\leq q\leq 3$ but will become less reliable for $q>3$.

Second, notice that it should be possible to obtain a linear stability argument from this analysis by taking $|a_i|\ll a_0$ for $i\neq 0$ and setting $a_0(t)=T(t)$ as given in Eq. (14). This is easily done by substituting T for a_0 in (74) and integrating. We find that, within a multiplicative constant, $a_i(t)$ is again given by (39), as we would expect. Since these results are independent of the detailed behavior of $F(x)$, we can conclude that the conjectures of stability for $1\leq q\leq 3$ in earlier parts of this paper are true for all physically reasonable choices of $F(x)$. However, because of the convergence questions raised in the previous paragraph, it is still questionable whether the separable solution is stable for all admissible $F(x)$ and $q>3$.

Finally, we remark that the existence of the approximation (79) for the asymptotic amplitude is rather surprising. Except for the approximation in (67), the arguments leading to (77) are short time approximations. It is not obvious that letting $t\to\infty$ in (77) should result in an expression for $a_0(t)$ which has exactly the right asymptotic character. We interpret this result to mean that, for this class of quasilinear equations, the solution of the initial boundary value problem is always close to the separable solution except for cases with extreme initial conditions.

B. Numerical comparisons

To demonstrate the accuracy of the amplitude prediction given by (79), the formulas were evaluated numerically for $\delta=-\frac{1}{2}$ with $F(x)=1$ and were compared to the results of numerical experiments.

In order to make this comparison, a test to determine how closely the particle distribution approximates the separable solution is required. First note that if the density *is* in the separable solution then

$$N(t)=\int_0^1 m^2(x,t)\,dx=\gamma T^2(t) \tag{83}$$

from Eq. (27). Recalling that for $\delta<0$, the separable solution vanishes in a finite time, what that time will be can be predicted using (17). The time at which the density is zero everywhere in a numerical experiment may be called the experimental termination time and

is given by

$$t_E = t + 2\lambda^{-1}[N(t)/\gamma]^{1/2}. \tag{84}$$

The constants λ and γ are known exactly (see Table I). The number t_E will not remain a constant during a numerical experiment unless the density is in the separable solution. However, as the perturbations decay, t_E rapidly approaches a constant. When t_E is stationary within some desired degree of accuracy, the perturbations are negligible to that degree of accuracy.

A gross estimate of the termination time can be found from the initial density distribution by finding t_E at $t = 0$,

$$t_0 = 2\lambda^{-1}[N(0)/\gamma]^{1/2}. \tag{85}$$

Equation (85) is only an order of magnitude estimate.

The theoretical estimate for the termination time obtained from (79) is easily shown to be

$$t_T = [B(0) + a_0^2(0)]/\lambda a_0(0). \tag{86}$$

The numbers $B(0)$ and $a_0(0)$ can be obtained from the initial distribution using (80)–(82) at $t = 0$.

The results of these calculations are compared in Table II. The values of t_E in the numerical experiment were stationary to six significant figures by $t = 0.10$. The order of magnitude estimate t_0 is seen to agree with t_E to two and sometimes three significant figures. The theoretical estimate t_T agrees very well with t_E being accurate to three and sometimes four significant figures. The approximations made in part A of this section are therefore valid to a satisfactory degree of accuracy in this particular case. Similar comparisons could be made for other values of δ to check the accuracy of the predictions in general. Such a systematic study will not be carried out here.

TABLE II. Values of the termination time from the order of magnitude estimate t_0 [Eq. (85)] and the theoretical estimate t_T [Eq. (86)] are compared to the value of t_E [Eq. (84)] found in numerical experiments for $\delta = -\frac{1}{2}$. The initial distribution of the pseudodensity is given by $m(x,0) = \sum_{l=0}^{3} a_l \sin(l+1)\pi x$ with the four cases of (a_1, a_2, a_3, a_4) equal to: (i) (1, 0.4, 0, 0), (ii) (1, 0, 0.3, 0), (iii) (1, 0, −0.3, 0), and (iv) (1, 0, 0, 0.225). The value of t_E in the numerical experiment was constant to six figures by $t = 0.1$.

Case	t_0	t_T	t_E
i	0.1873	0.1825	0.1847
ii	0.1815	0.1675	0.1677
iii	0.1815	0.1893	0.1895
iv	0.1782	0.1762	0.1762

VI. CONCLUSIONS

We have shown that for the geometry free case $[F(x) = 1]$ the qualitative behavior of plasma diffusing across a magnetic field can be successfully modelled using Eq. (6) and zero density boundary conditions when $\delta > -1$. The principal effect of the geometrical factor $F(x)$ is to make the separable solution shape function asymmetrical. A second effect of the geometry is to modify the eigenvalue λ [see (31)]. Since the rate at which plasma escapes from the containment device depends on λ, $F(x)$ also determines how long it takes for all the particles to escape.

When $F(x) = 1$, the separable solution has been shown to be stable against infinitesimal perturbations for all $\delta > -1$. The slowest decaying perturbation to the pseudodensity $m(x,t) = S(x)T(t)$ has time dependence $T^4(t)$. Thus, all perturbations decay at least four times faster than the separable solution. For finite perturbations an approximate analysis valid for all $F(x)$'s being considered [see Eqs. (8) and (9)] shows that initially the perturbations decay exponentially for all $\delta > -1$. Furthermore, an estimate (79) of the asymptotic separable solution amplitude was obtained. For $\delta = -\frac{1}{2}$, this estimate satisfactorily predicts the time it takes for all particles to escape in numerical diffusion experiments.

When $F(x) \neq$ constant but symmetric, stability has been rigorously established for $\delta > -\frac{1}{2}$ and is conjectured for all $\delta > -1$.

We conclude that the solutions of the model diffusion equation (6) behave qualitatively the same as plasma particle density diffusing across the magnetic field of a toroidal multipole. Stability of the separable solution is assured on theoretical grounds for $\delta \geq -\frac{1}{2}$. A detailed comparison with toroidal octupole experiments will determine whether the quantitative predictions concerning the shape function and decay rate are also in agreement. Such a comparison is made and favorable agreement is found in Ref. 3.

ACKNOWLEDGMENTS

The author thanks Professor A.C. Scott, Professor K.R. Symon, and Dr. J.R. Drake for helpful discussions during this work and for comments on the manuscript. The research was sponsored by ERDA under Contract No. E(11-1)-2387, by the United States Army under Contract No. DAAG29-75-C-0024, and by NSF under Grant No. ENG75-08492.

[1] J.R. Drake, Phys. Fluids **16**, 1554 (1973); J.R. Greenwood, Ph.D. thesis, University of Wisconsin (1975): J.R. Drake, J.R. Greenwood, G.A. Navratil, and R.S. Post, Phys. Fluids **20**, 148 (1977).
[2] H. Okuda and J.M. Dawson, Phys. Rev. Lett. **28**, 1625 (1972); and Phys. Fluids **16**, 408 (1973).
[3] J.R. Drake and J.G. Berryman, Phys. Fluids **20**, 851 (1977).
[4] K.E. Lonngren (private communication).
[5] R.H. Boyer, J. Math. & Phys. **40**, 41 (1961).
[6] W.F. Ames, *Nonlinear Partial Differential Equations in Engineering* (Academic, New York, 1965), Sec. 4.7.
[7] H. Levy and E.A. Baggot, *Numerical Studies in Differential Equations* (Watts, London, 1934), Sec. 5.17; W.E. Milne, *Numerical Solution of Differential Equations* (Dover, New York, 1970), Sec. 48.
[8] H.B. Keller and D.S. Cohen, J. Math. Mech. **16**, 1361 (1967); D.S. Cohen, *ibid.* **17**, 209 (1967).
[9] H. Amann, SIAM Rev. **18**, 620 (1976).
[10] E.A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations* (McGraw-Hill, New York, 1955), p. 212.
[11] M. Lees, Maths. Comput. **20**, 516 (1966).
[12] M.F. Weinberger, *Partial Differential Equations* (Blaisdell, Waltham, MA, 1965), Chap. VII.

Global structure of the "Kantowski–Sachs" cosmological models

C. B. Collins

Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
(Received 1 April 1977)

A discussion is given of the "Kantowski–Sachs" cosmological models; these are defined locally as admitting a four-parameter continuous isometry group which acts on spacelike hypersurfaces, and which possesses a three-parameter subgroup whose orbits are 2-surfaces of constant curvature (i.e., the models possess spherical symmetry, combined with a translational symmetry, and can thus be regarded as nonempty analogs of part of the extended Schwarzschild manifold). It is shown that all general relativistic models in which the matter content is a perfect fluid satisfying reasonable energy conditions are geodesically incomplete, both to the past and to the future, and that at each resulting singularity the fluid energy density is infinite. In the case where the fluid obeys a barotropic equation of state (which includes all known exact perfect fluid solutions) the field equations are shown to decouple to form a plane autonomous subsystem. This subsystem is examined using qualitative (Poincaré–Bendixson) theory, and phase-plane diagrams are drawn depicting the behavior of the fluid's energy density and shear anisotropy in the course of the models' evolution. Further diagrams depict the conformal structure of these universes, and a table summarizes the asymptotic properties of all physically relevant variables.

1. INTRODUCTION

In the study of spatially homogeneous cosmological models in general relativity, it is customary to formulate the subject in terms of the Bianchi classification, or of some modification of this (see, e.g., Ref. 1). This classification is based on the original works of Bianchi, who gave the first canonical reduction of three-parameter Lie groups. In general relativity, Lie groups most frequently appear when one postulates that the metric g_{ab} be invariant under a continuous transformation. The infinitesional generator of this transformation is known as a Killing vector, and the set of all continuous transformations that leave invariant a given metric forms a Lie group, known as an isometry group. Spatial homogeneity of a cosmological model is defined by requiring that (locally) the space–time be invariant under an isometry group which acts transitively on spacelike hypersurfaces.

For simplicity, we shall assume that the matter content of the models consists of a perfect fluid, satisfying the Einstein field equations with zero cosmological constant Λ. In this case the energy–momentum tensor T_{ab} can be written in the form

$$T_{ab} = (\mu + p)u_a u_b + p g_{ab},$$

where $\mathbf{u}$ is a unit timelike vector tangential to the fluid flow lines, $\mu > 0$ is the total energy density, and p is the isotropic pressure of the fluid. We shall also assume that the fluid obeys an equation of state of the form $p = p(\mu)$, satisfying the plausible inequalities $0 \leq p \leq \mu$ and $0 \leq dp/d\mu \leq 1$. Under these conditions, the vector field $\mathbf{u}$ can be characterized as the unique timelike eigenvector of the Ricci tensor R_{ab}.

Suppose that a cosmological model is spatially homogeneous. Then it is invariant under an r-parameter isometry group G_r whose orbits are spacelike hypersurfaces (so that $r \geq 3$). At any point q on any such hypersurface $\mathcal{S}$, there are three nonzero linearly independent Killing vector fields tangent to $\mathcal{S}$, and, if $r > 3$, there are additional linearly independent Killing vector fields tangent to $\mathcal{S}$, but which vanish at q. In this last case there is a continuous *isotropy* group H_q, which consists of those transformations that are isometries leaving the point q fixed. Each such transformation, $h \in H_q$, will generate a Lorentz transformation in the tangent space, T_q, of q, and this will leave invariant all intrinsically defined vectors in T_q. Thus h acts in the subspace of T_q orthogonal to $\mathbf{u}$, and so H_q is either one- or three-dimensional; consequently, if $r > 3$, then either $r = 4$ or $r = 6$. If $r = 6$, the spacetime is not only spatially homogeneous but also spatially isotropic, and belongs to the Friedmann–Robertson–Walker class.[1] If $r = 4$, the spacetime is locally rotationally symmetric.[2,3] If either $r = 4$ or $r = 6$, then, *in all but one case*, the full isometry group G_r admits a three-parameter subgroup which acts transitively on spacelike hypersurfaces, and the corresponding space–times can be regarded as being Bianchi models possessing additional symmetries. The exceptional case arises for the following reason. Any four-parameter Lie group G_4 admits a three-parameter subgroup G_3, whose orbits are either two- or three-dimensional.[4] In the latter case, the space–time belongs to the Bianchi class. In the former case, the orbits are necessarily of constant curvature. If this curvature is zero or negative, then it can be shown[4] that the group G_4 admits a second three-parameter subgroup, whose orbits are three-dimensional (so the space–time belongs to the Bianchi class), but if the curvature is positive there is no such subgroup, and G_3 is isomorphic to $SO(3, \mathbb{R})$, or Bianchi type IX. Although the detailed proofs of these statements have already been given by Kantowski,[4] this reference is not immediately available, and so similar, but slightly improved, versions are presented in Appendices A and B, where it is also shown that the four linearly independent Killing vectors ξ_1, ξ_2, ξ_3, and η can be chosen so that their Lie algebra is given by

$$\begin{aligned} &[\xi_1, \xi_2] = \xi_3, \quad [\xi_2, \xi_3] = \xi_1, \quad [\xi_3, \xi_1] = \xi_2, \\ &[\eta, \xi_1] = 0, \quad [\eta, \xi_2] = 0, \quad [\eta, \xi_3] = 0. \end{aligned} \tag{1.1}$$

 Copyright © 1977 American Institute of Physics

The three Killing vectors ξ_1, ξ_2, ξ_3 generate the subgroup G_3 with two-dimensional orbits. Appendices A and B are also given for didactic purposes, to exemplify the use and formalism of Lie group theory.

Perhaps the most familiar space–time admitting a four-parameter Lie group with Lie algebra (1.1) is the *Schwarzschild* solution, although here the vector field η is timelike outside the event horizon. The vectors ξ_1, ξ_2 and ξ_3 generate the spherical symmetry of the model. That part U of the analytically extended Schwarzschild space–time inside the event horizon has exactly the same symmetries as the models presently being discussed, since then all four group generators are spacelike. In effect, what we shall now do is to investigate what (nonempty) cosmological space–times locally possess the same symmetries as that of the region U of the extended Schwarzschild solution, and then analytically extend the resulting manifold to obtain its global structure. Space–times so determined are known most commonly as the *Kantowski–Sachs* models, following the studies by Kantowski and Sachs[5] in the case where the matter content is dust (i.e., incoherent matter, with equation of state $p=0$), and by Kantowski[4] in the case where the matter content is dust ($p=0$), radiation (ultrarelativstic Fermi gas, with equation of state $p=\frac{1}{3}\mu$) or Zeldovich's "stiff" matter (with equation of state $p=\mu$). Despite this nomenclature, such model universes appear to have been studied first by Kompaneets and Chernov,[6] in the case of dust and radiation. (I thank Dr. L. P. Grishchuk for informing me of this fact.) Further discussion of the Kantowski–Sachs models, including exact solutions, is given in the case of dust by Ellis[2] and Thorne,[7] of Zeldovich stiff matter by Thorne,[7] of a general perfect fluid by Stewart and Ellis,[3] of a perfect fluid and an electromagnetic field by Doroshkevich,[8] Thorne,[9] and Stewart and Ellis,[3] and of a pure electromagnetic field (the *Bertotti–Robinson* solutions) by Bertotti,[10] Robinson,[11] Stewart and Ellis,[3] and Thorne.[7,12] Ellis[13] has discussed the various topological structures that are possible in the Kantowski–Sachs models.

In the present article it is shown that certain features exhibited by the known exact solutions are in fact generic properties of the model. The most interesting such feature is that any perfect fluid model will be geodesically incomplete both to the future and to the past, and that the energy density of the fluid becomes infinite at both of these singularities.

In Sec. 2, we give the general form of the metric with the symmetries (1.1), together with the field equations. In Sec. 3, it is shown that all perfect fluid models possess both a past and a future singularity, at which the fluid's energy density becomes infinite. In Sec. 4, we examine the case where the perfect fluid obeys a barotropic equation of state of the form $p=(\gamma-1)\mu$, where $1\leq\gamma\leq 2$, and show that for each value of γ there are special models in which the fluid shear remains finite at the singularity, and that there are also models which are time symmetric about the moment of maximum expansion. Diagrams are drawn depicting the variety of possible types of evolution, and conformal representations of the models are given. A table summarizes the character of the different types of singularity, by exhibiting the asymptotic time dependence of all physically relevant variables. Section 5 draws the conclusions.

2. THE METRIC AND THE FORMALISM

We shall be concerned with the Kantowski–Sachs metric which admits four spacelike Killing vectors, ξ_1, ξ_2, ξ_3, and η, with Lie algebra (1.1). Since the space–time is spherically symmetric, it follows that there are coordinates $\{t, r, \theta, \phi\}$ such that the orbits of the Killing vectors ξ_1, ξ_2 and ξ_3 are the 2-surfaces $\{t, r=\text{const}\}$, in which the metric takes the form[14]

$$ds^2=-\frac{dt^2}{F^2(t,r)}+X^2(t,r)\,dr^2+Y^2(t,r)(d\theta^2+\sin^2\theta\,d\phi^2),$$

where the coordinates (t, r) are chosen in each two-surface $\{\theta, \phi=\text{const}\}$ so that they are orthogonal [i.e., $(\partial/\partial t)\cdot(\partial/\partial r)=0$], but are otherwise arbitrary. Since the vector field η is invariantly defined, it lies in the two-surfaces $\{\theta, \phi=\text{const}\}$, and we can choose the (t, r) coordinates so that $\eta=N(t, r, \theta, \phi)\partial/\partial r$. From Killing's equations we obtain $N=N(r)$, so, by redefining r, we have $\eta=\partial/\partial r$, i.e., without loss of generality, $N=1$. Killing's equations then necessitate $F=F(t)$, $X=X(t)$, and $Y=Y(t)$, and, by redefining t, we have, without loss of generality, $F=1$. Thus the final form of the Kantowski–Sachs metric is

$$ds^2=-dt^2+X^2(t)\,dr^2+Y^2(t)(d\theta^2+\sin^2\theta\,d\phi^2), \tag{2.1}$$

and the field equations are

$$\frac{2\dot{X}\dot{Y}}{XY}+\frac{1+\dot{Y}^2}{Y^2}=\mu, \tag{2.2}$$

$$\frac{2\ddot{Y}}{Y}+\frac{1+\dot{Y}^2}{Y^2}=-p, \tag{2.3}$$

and

$$\frac{\ddot{X}}{X}+\frac{\ddot{Y}}{Y}+\frac{\dot{X}\dot{Y}}{XY}=-p, \tag{2.4}$$

where a dot denotes differentiation with respect to t. Equations (2.2)–(2.4) are compatible if and only if the conservation equation

$$\dot{\mu}+(\mu+p)\left(\frac{\dot{X}}{X}+\frac{2\dot{Y}}{Y}\right)=0 \tag{2.5}$$

is satisfied, by virtue of the Bianchi identities.

These field equations can be expressed as constraint and propagation equations for the volume expansion $\theta=\dot{X}/X+2\dot{Y}/Y$ and shear tensor, σ_{ij}, which, with respect to any orthonormal frame whose "1"-direction is aligned along $\partial/\partial r$, is of form $\sigma_{ij}=\text{diag}(2\sigma/\sqrt{3}, -\sigma/\sqrt{3}, -\sigma/\sqrt{3})$, where $\sigma=(1/\sqrt{3})(\dot{X}/X-\dot{Y}/Y)$. Thus

$$\dot{\theta}+\tfrac{1}{3}\theta^2+2\sigma^2+\tfrac{1}{2}(\mu+3p)=0, \tag{2.6}$$

$$\dot{\sigma}+\sigma\theta-(1/\sqrt{3})/Y^2=0, \tag{2.7}$$

$$\tfrac{1}{3}\theta^2-\sigma^2+1/Y^2=\mu, \tag{2.8}$$

together with

$$\dot{\mu}+(\mu+p)\theta=0. \tag{2.9}$$

Equation (2.6) is the familiar Raychaudhuri's equa-

tion[15,16] of relativistic cosmology,[17] whereas Eq. (2.8) can be derived from the Gauss–Codacci equations relating the curvature tensor of the space–time to that of the hypersurfaces $\{t = \text{const}\}$, which therefore have Ricci scalar $R^* = 2/Y^2 > 0$.

Any metric (2.2) of our class (i.e., containing a perfect fluid with $\mu > 0$, $0 \le p \le \mu$, and $0 \le dp/d\mu \le 1$ and obeying Einstein's fluid equations with $\Lambda = 0$) is necessarily of Petrov type D (22) and not any specialization thereof. This is proved in Appendix C.

The covering manifold of the space sections $\{t = \text{const}\}$ of the Kantowski–Sachs metric has topology $S^2 \times \mathbf{R}^1$. One can obtain other topologies[13] by (a) a identification of points under a translation in the $\mathbf{R}^1$ (i.e., η) direction, or under a translation in the $\mathbf{R}^1$ direction together with a rotation or a reflection, (b) an identification of antipodal points in each S^2 $\{t, r = \text{const}\}$, and (c) a combination of (a) and (b).

3. SINGULARITIES–PAST AND FUTURE

If we define a characteristic length scale l by $l^3 = XY^2$ (so $\theta = 3\dot{l}/l$), it follows from Raychaudhuri's Eq. (2.6) that

$$3\ddot{l}/l + 2\sigma^2 + \tfrac{1}{2}(\mu + 3p) = 0. \tag{3.1}$$

Suppose that there is a time t_0 such that $\theta_0 \equiv \theta(t_0) > 0$. Then, if $l_0 \equiv l(t_0)$, $\mu_0 \equiv \mu(t_0)$, and $p_0 \equiv p(t_0)$, it follows from the energy conditions

$$p = p(\mu), \quad \mu > 0, \quad 0 \le p \le \mu, \quad \text{and } 0 \le dp/d\mu \le 1 \tag{3.2}$$

and from (2.9) that

$$0 < \mu_0 l_0^3/l^3 \le \mu \le \mu_0 l_0^6/l^6 \tag{3.3}$$

whenever $0 < l \le l_0$. Since (3.1) shows that $\ddot{l} < 0$, we deduce that $\exists\ T_1 < t_0$ such that $l \to 0$ and $\mu \to \infty$ as $t \to T_1+$, *provided that t can be extended that far*. It is conceivable that there is an intervening singularity which prevents us from extending the model as far back as time $t = T_1$. Such a situation is exemplified in the case of the Friedmann–Robertson–Walker models by simply cutting out and discarding a region $I^-(S)$ to the past of any Cauchy hypersurface S or, less artificially, in the case of certain anisotropic cosmological models by the onset of a milder type of singularity, where μ is finite.[18] In order to circumvent any "artificial" singularity, we shall simply assume that if it is possible to extend the universe to earlier times, then the universe is so extended: I.e., *we assume that the manifold is inextendible*. We now show that with this assumption there are no other "intervening" singularities; that is, if $l(t)$ is bounded away from zero at time T, it is possible to extend the space–time to values of $t < T$.

Theorem 1: If $\exists\ t_0$ such that $l_0 \equiv l(t_0) > 0$ and $\dot{l}(t_0) > 0$, then the Cauchy data for the field equations on the hypersurface $t = t_0$ is regular, and the model can be extended back to earlier times $t \le t_0$.

Proof: The proof of the theorem divides into two parts, depending on whether or not $Y_0 \equiv Y(t_0)$ is zero. We show first that the case $Y_0 = 0$ is untenable, and then proceed to examine the case $Y_0 > 0$.

Suppose that $Y_0 = 0$. Then $Y \to 0+$ as $t \to t_0+$. By inequalities (3.2) and (3.3) it follows that $\mu_0 \equiv \mu(t_0)$ and $p_0 = p(t_0)$ are bounded. By Eq. (2.3), if $\dot{Y}$ is ever zero, then $\ddot{Y} < 0$, and so as Y approaches zero it cannot do so in an oscillatory fashion; thus $\exists\ t_1 > t_0$ such that $\dot{Y} > 0$ for all $t \in (t_0, t_1)$. Choose $t_2 \in (t_0, t_1)$, and write $Y_2 \equiv Y(t_2) > 0$ and $\dot{Y}_2 \equiv \dot{Y}(t_2)$. By Eq. (2.3) it follows that

$$\frac{d}{dY}(Y\dot{Y}^2) = -(1 + pY^2) \tag{3.4}$$

and hence $\exists\ K \ge 1$ such that

$$Y_2(1 + \dot{Y}_2^2) - Y \le Y\dot{Y}^2 \le Y_2(K + \dot{Y}_2^2) - KY$$

whenever $Y < Y_2$. Consequently, $\dot{Y} \to +\infty$ as $t \to t_0+$, and, *a fortiori*, $\dot{Y}/Y \to +\infty$ as $t \to t_0+$. Recalling that μ is bounded, it follows that for all ε satifying $0 < \varepsilon < 3$, $\exists\ t_3 \in (t_0, t_2)$ such that

$$|\mu - 1/Y^2| < \varepsilon\dot{Y}^2/Y^2$$

for all $t \in (t_0, t_3)$. Using Eq. (2.2), we obtain

$$(3 - \varepsilon)\dot{Y}/Y < \theta = 3\dot{l}/l$$

$$= \frac{\mu - 1/Y^2 + 3\dot{Y}^2/Y^2}{\dot{Y}/Y} < (3 + \varepsilon)\dot{Y}/Y$$

for all $t \in (t_0, t_3)$. Since $Y \to 0+$ as $t \to t_0+$, it follows that $l \to 0$ as $t \to t_0+$, which contradicts our assumption that $l(t_0) > 0$. Hence, if $l(t_0) > 0$, it follows that $Y_0 \equiv Y(t_0) > 0$.

Now suppose that $l(t_0+) > 0$ and $Y_0 \equiv Y(t_0+) > 0$. Since $l^3 = XY^2$, it follows that $X(t_0+) > 0$, and $X(t_0+)$ is finite. By the inequalities (3.2) and (3.3), both $\mu_0 \equiv \mu(t_0+)$ and $p_0 \equiv p(t_0+)$ are finite. From Eq. (2.3) we again see that $\dot{Y} = 0$ implies $\ddot{Y} < 0$, and thus if $\dot{Y}(t_0+) = 0$, then $\exists\ t_0' > t_0$ such that $l(t_0'+) > 0$, $Y(t_0'+) > 0$ and $\dot{Y}(t_0'+) \ne 0$, i.e., without loss of generality, $\dot{Y}(t_0+) \ne 0$. From Eq. (2.2), it follows that $\dot{X}(t_0+)$ is finite. Using (2.3) and (2.4), we see that $\ddot{X}(t_0+)$ and $\ddot{Y}(t_0+)$ are bounded. Finally, from Eq. (2.5) we find that $\dot{\mu}(t_0+)$, and hence $\dot{p}(t_0+)$, is finite. Consequently, the metric components, X and Y, the components $\dot{X}/X$ and $\dot{Y}/Y$ of the second fundamental form of the hypersurfaces $\{t = \text{const}\}$, and the fluid's energy density and pressure are all regular for $t \ge t_0$. By our assumption of extendibility, the fact that the Cauchy data is well defined on $t = t_0$ means that the model can be extended to values of $t \le t_0$. This completes the proof of the theorem.

We shall now show that every model of our class is geodesically incomplete both to the future and the past. This is achieved by showing that, unless a singularity intervenes, $Y(t)$ tends to zero in the finite future and in the finite past and then by recalling from Theorem 1 that if $Y \to 0$, then $l \to 0$.

Theorem 2: The perfect fluid Kantowski–Sachs models are geodesically incomplete both to the past and future. At the associated singularities, $\mu \to +\infty$, $\theta \to \pm\infty$, and either $R^* \to R_0^* > 0$ and $\sigma \to \pm\infty$ or $R^* \to +\infty$ (R_0^* is a constant).

Proof: Equation (2.3) shows that if $\dot{Y}$ is ever zero and $l \ne 0$, then $\ddot{Y} < 0$ at that time (and so a maximum of

$Y = Y(t)$ is achieved). We show first that if $\exists\ t_0$ such that $Y_0 \equiv Y(t_0) > 0$ and $\dot{Y}_0 \equiv \dot{Y}(t_0) > 0$, then it is impossible to have $\dot{Y} > 0$ for all $t \geq t_0$, and conclude that the only alternative is that $\exists\ t_1 > t_0$ at which $\dot{Y}(t_1) = 0$, $l(t_1) \neq 0$, in which case a singularity is encountered at some later time, at which $l = 0$.

Suppose that $\dot{Y} > 0$ for all $t \geq t_0$. By Eq. (3.4) we have $(Y\dot{Y}^2)^{\cdot} = -\dot{Y}(1 + pY^2) \leq -\dot{Y}$, and hence $Y(1 + \dot{Y}^2) \leq Y_0(1 + \dot{Y}_0^2)$ for all $Y \geq Y_0$. Hence Y is bounded above by $Y_0(1 + \dot{Y}_0^2)$. Since $\dot{Y} > 0$, it follows that $\exists\ \bar{Y} > 0$ such that $Y \to \bar{Y}-$ as $t \to +\infty$. Define a new time variable τ by $\tau = \int_{t_0}^{t} (1/Y^{1/2})\, dt$. Since $Y \to \bar{Y}-$ as $t \to +\infty$, it follows that $\tau \to +\infty$, and, from Eq. (2.3), that $d^2Y/d\tau^2 \leq \frac{1}{2}$ for $\tau \geq 0$. Thus $Y \leq -\frac{1}{4}\tau^2 + \dot{Y}_0\tau + Y_0 \to -\infty$ as $\tau \to +\infty$, and this contradicts our assumption that $\dot{Y} > 0$ for all $t \geq t_0$. It is therefore false that $\dot{Y} > 0$ for all $t \geq t_0$; consequently, assuming that the manifold is inextendible, either $\exists$ a time $t_1 > t_0$ such that $Y(t_1-) \neq 0$ and $l(t_1-) = 0$, or $\exists$ a time $t_2 > t_0$ such that $\dot{Y}(t_2) = 0$ and $l(t_2) \neq 0$, in which case $\ddot{Y} < 0$ for $t \geq t_2$, and so $\exists$ a time $t_3 > t_2$ such that $Y(t_3-) = 0$ [and the time reverse of the argument in the proof of Theorem 1 shows that $l(t_3-) = 0$]. Hence, we have shown that if $\exists\ t_0$ such that $\dot{Y}_0 \equiv \dot{Y}(t_0) > 0$, then there is a singularity to the finite future of t_0, at which $l \to 0$ (and so $\theta = 3\dot{l}/l \to -\infty$), and, by the inequalities (3.3), at which $\mu \to +\infty$. It also follows that there is a singularity to the finite past of t_0, since Eq. (2.3) shows that $\ddot{Y} < 0$, and, by our extendibility assumption and the arguments in the proof of Theorem 1, the only possibilities are that $\exists\ t_4 < t_0$ such that $Y(t_4+) \neq 0$ and $l(t_4+) = 0$, or $\exists\ t_5 < t_0$ such that $Y(t_5+) = 0$ and $l(t_5+) = 0$. At this past singularity, $l \to 0$, $\theta = 3\dot{l}/l \to +\infty$, and $\mu \to +\infty$.

The time reverses of the above arguments show that if $\exists\ t_0'$ such that $\dot{Y}(t_0') < 0$, the same qualitative conclusions hold: There is a singularity to the finite future and finite past of t_0', at which $\mu \to +\infty$ and $\theta \to \pm\infty$.

In order to determine the possible behaviors of the fluid shear σ and the Ricci scalar R^* of the hypersurfaces $\{t = \text{const}\}$, we first recall that if $\dot{Y} = 0$ and $l \neq 0$, then, by Eq. (2.3), $\ddot{Y} < 0$, and so $R^* = 2/Y^2$ cannot oscillate. Hence either $R^* \to +\infty$ at the singularity, or R^* approaches a finite limit. In this last case, R^* cannot tend to zero, since this would require $Y \to +\infty$, which, as we have already seen, is impossible. For any singularity at which R^* tends to a finite (nonzero) limit, it follows from Eq. (2.7) that without loss of generality we can consider a time t_1 to the future of which σ does not change sign, and so $\exists\ K > 0$ such that $-K < [\ln(\sigma l^3)]^{\cdot} < K$ for all $t \geq t_1$. Writing $\sigma_1 \equiv \sigma(t_1)$, $l_1 \equiv l(t_1)$, we obtain $\exp[-K(t - t_1)] \leq \sigma l^3/\sigma_1 l_1^3 \leq \exp[K(t - t_1)]$ for $t \geq t_1$, and so, at the singularity to the future of t_1, σl^3 is bounded away from zero, i.e., $\sigma \to \pm\infty$. A similar argument shows that when R^* is finite, $\sigma \to \pm\infty$ at the singularity to the past of t_1. This completes the proof of Theorem 2.

In the next section we examine the subclass of models in which the equation of state is barotropic, and obtain examples of models with singularities at which $R^* \to R_0^* > 0$ and $\sigma \to \pm\infty$, and of models with singularities where $R^* \to +\infty$ and either $\sigma \to \pm\infty$ or $\sigma \to \sigma_0$, where σ_0 is a constant (possibly zero).

4. QUALITATIVE ANALYSIS OF THE MODELS

We shall consider in this section those Kantowski–Sachs models in which the perfect fluid content obeys a barotropic equation of state, of the form $p = (\gamma - 1)\mu$, where γ is a constant. For values of γ lying in the range $1 \leq \gamma \leq 2$, this equation of state is thought to be physically plausible. Of particular interest are the cases $\gamma = 1$ (dust), $\gamma = \frac{4}{3}$ (radiation) and $\gamma = 2$ (Zeldovich stiff matter). We shall describe the models in terms of variables β and Ω, defined by $X \equiv \exp(-\Omega + \beta)$ and $Y \equiv \exp(-\Omega - \beta/2)$ (and hence $l \equiv (XY^2)^{1/3} \equiv \exp(-\Omega)$). The quantity β is related to the fluid shear by $\sigma = \frac{1}{2}\sqrt{3}\dot{\beta}$. The variable Ω will be used as a time variable, and a prime will denote differentiation with respect to Ω; this description will then be valid except at the instant when $\dot{\Omega} = 0$, i.e., when the universe stops expanding and begins to contract. The fluid expansion rate is $\theta = -3\dot{\Omega}$, the dynamical importance of the shear is measured by $\sigma/\theta = (-1/2\sqrt{3})\beta'$, and the dynamical importance of the fluid by $x = 3\mu/\theta^2 = \mu/3\dot{\Omega}^2$. This notation agrees with that of previous works of a similar nature.[19–21]

The first step is to reexpress the field equations (2.3)–(2.5) in terms of coupled first-order differential equations with two dependent variables, β' and x, and with independent variable, Ω. Equation (2.2) becomes

$$\beta'^2 + 4x - 4 = (4/3\dot{\Omega}^2)\exp(2\Omega + \beta). \tag{4.1}$$

The $\ddot{\Omega}$ terms are eliminated from (2.3) and (2.4) to obtain

$$\beta'' = \tfrac{3}{8}\beta'(4 - \beta'^2) - \tfrac{3}{2}\beta' x(\gamma - 1) + (2 - \tfrac{1}{2}\beta') \cdot (1/3\dot{\Omega}^2)\exp(2\Omega + \beta). \tag{4.2}$$

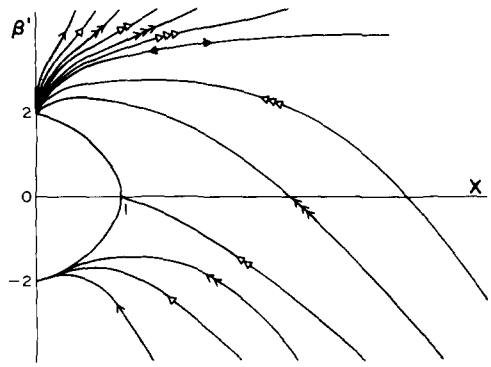
Substituting for the $\exp(2\Omega + \beta)$ term from (4.1) into (4.2), we have

$$\beta'' = \tfrac{1}{2}\beta'[4 - \beta'^2 - (3\gamma - 2)x] - \tfrac{1}{2}[4 - \beta'^2 - 4x]. \tag{4.3}$$

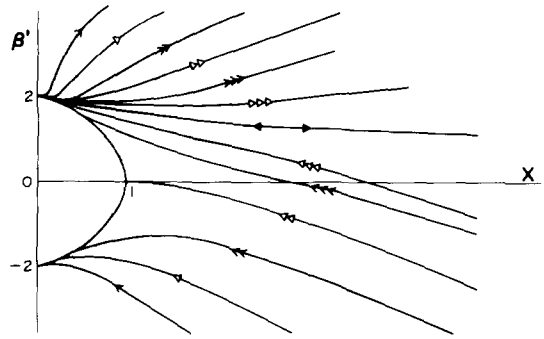
Equation (2.5), together with the expression for $\ddot{\Omega}$ obtained by eliminating the β'' terms from (2.3) and (2.4), yields

$$x' = x[(3\gamma - 2)(1 - x) - \beta'^2]. \tag{4.4}$$

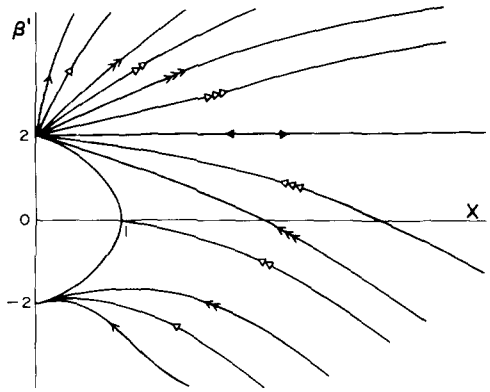
Equations (4.3) and (4.4) are exactly the same as the equations obtained in Ref. 19 for the locally rotationally symmetric Bianchi Type III models (where, in the notation of Ref. 19, $C = k = 1$). These equations form a plane autonomous system of ordinary differential equations, and the qualitative behavior of the solutions can be sketched in the x–β' phase plane (for a description of this procedure, and for its applications to general relativistic cosmology, see Refs. 18–24, and references cited to the standard literature). The resulting diagrams are, however, distinct from those of Bianchi type III, since in that case the region of interest is $\{(x, \beta') : x > 0$ and $\beta'^2 + 4x - 4 < 0\}$ whereas here we shall be interested in the region $\{(x, \beta') : x > 0$ and $\beta'^2 + 4x - 4 > 0\}$. This is evident from Eq. (4.1), and is directly related to the fact that in this case $R^* > 0$, whereas for Bianchi type III, $R^* < 0$; and more indirectly from the observation that the two classes of models can be related by a complex transformation[4] which involves $X \to X$, $Y \to iY$, and hence the transformations $\beta \to \beta - i\pi/3$, $\Omega \to \Omega - i\pi/3$.



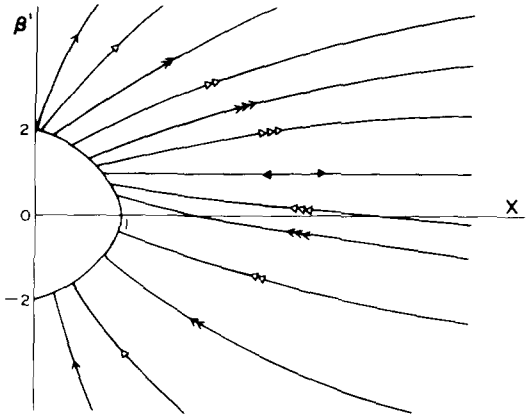
$(1 \leqslant \gamma < 4/3)$



$(4/3 < \gamma < 2)$



$(\gamma = 4/3)$



$(\gamma = 2)$

FIGS. 1—4. Qualitative description of the evolution of the Kantowski—Sachs models, in the case where the matter content is a perfect fluid obeying the equation of state $p = (\gamma - 1)\mu$ for the following (constant) values of γ: $1 \leq \gamma < \frac{4}{3}$ (Fig. 1), $\gamma = \frac{4}{3}$ (Fig. 2), $\frac{4}{3} < \gamma < 2$ (Fig. 3) and $\gamma = 2$ (Fig. 4). Each diagram refers to the evolution for a fixed value of γ, and each curve represents this evolution for a fixed set of initial conditions. The variable $x = 3\mu/\theta^2$ measures the relative dynamical importance of the matter, and the variable $\beta' = -2\sqrt{3}(\sigma/\theta)$ measures the relative dynamical importance of the fluid shear. The fact that the curves extend to infinite values of x signifies that the Ω variable ceases to be valid at the time of maximum radius $l = X^{1/3}Y^{2/3} = e^{-\Omega}$. The entire course of evolution is indicated by associated types of arrows. The time reverse of any of these models is also feasible. There is, for each value of γ, one model that is time symmetric; This is denoted by a solid arrow (→) in the figures.

The discussion now divides up into the two cases $1 \leq \gamma < 2$ and $\gamma = 2$, since in the special case $\gamma = 2$, Eqs. (4.3) and (4.4) become

$$\beta'' = \tfrac{1}{2}(\beta' - 1)(4 - 4x - \beta'^2)$$

and

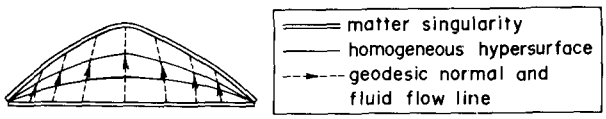
$$x' = x(4 - 4x - \beta'^2);$$

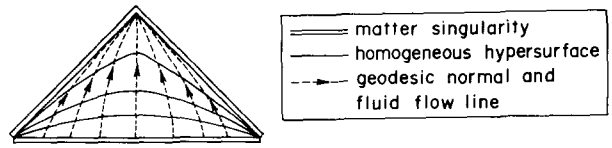
this means that when $\gamma = 2$, $(\beta' - 1)^2 = Kx$ (where $K \geq 0$ is constant) is a first integral, and that the integral curves of the system (4.3) and (4.4) are generically parabolas in the x—β' plane.

The integral curves of the system (4.3) and (4.4) are drawn in Figs. 1—4. Each diagram is drawn for a single value of γ, and represents the evolution of the class of models for that value of γ, under a variety of initial conditions. Slightly different qualitative pictures emerge for values of γ in the range $1 \leq \gamma < \frac{4}{3}$ (Fig. 1) $\gamma = \frac{4}{3}$ (Fig. 2) and $\frac{4}{3} < \gamma < 2$ (Fig. 3), whereas there is an entirely different pattern of evolution if $\gamma = 2$ (Fig. 4). In each diagram the integral curves extend out to infinitely large values of x (and, usually, of β'). The interpretation of this is that the Ω variable has ceased to be valid, since $x \to +\infty$ corresponds to $\dot{\Omega} \to 0-$, and that the universe model has halted its expansion, and is about to contract. Apart from certain special cases, the fluid shear, σ, will be nonzero at such a time, and so β' will be infinite. If, for instance, $\beta' \to +\infty$ at this time, then, since σ is continuous, this means that the continuation of our model into the contraction era is depicted in Figs. 1—4 by β' being large and negative. For this reason, the arrows in the diagrams depict the entire course of evolution (expansion and contraction), *despite* the fact that Ω is not a valid time variable throughout, although it should be recalled that the time reverses of these models are also feasible.

In each diagram it can be observed that, for one integral curve, β' remains finite as $x \to +\infty$. This requires that the fluid shear approach zero as $\dot{\Omega} \to 0-$, and has the interpretation that the model is time-symmetric or "momentarily static."[7]

The general features of the singularities are as follows. In the case $1 \leq \gamma < 2$, the models generically have either one "pancake" singularity ($X \to 0$, $Y \to Y_0 > 0$, where Y_0 is a constant) and one "cigar" singularity





FIGS. 5 and 6. Conformal ("Penrose") diagrams of the totally geodesic timelike 2-surfaces $\{\theta, \phi = \text{const}\}$ with metric induced by (2.1):

$$ds^2 = -dt^2 + X^2(t)dr^2.$$

Null lines are drawn at $\pm 45°$ and infinite distances are rescaled to finite coordinate values. Figure 5 displays the evolution of models whose singularities are points, barrels, or cigars. Figure 6 displays the evolution of those models which possess one pancake singularity, and indicates how particle horizons are thereby removed in the preferred $\partial/\partial r$ direction of the pancake.

$(X \to +\infty,\ Y \to 0)$ or they have two cigar singularities. The pancake singularities are represented by the point $x = 0,\ \beta' = -2$, and the cigar singularities by the point $x = 0,\ \beta' = +2$ in Figs. 1–3. At these general singularities the fluid shear σ and expansion rate θ are dominant in the Raychaudhuri Eq. (2.6), and the fluid energy density μ is insignificant. The models which possess two cigar singularities can be characterized as those in which there is an instant at which the fluid shear vanishes. There are also two special models. One of these has one "point" singularity $(X \to 0, Y \to 0)$, at which θ and μ dominate in (2.6) and σ is negligible, and the other singularity is a cigar singularity, where σ and θ dominate over μ. These singularities are represented by the points (1, 0) and (0, 2) respectively in the x–β' planes of Figs. 1–3. The other special model is the time-symmetric one, which begins and ends in a cigar singularity [with σ and θ again dominating μ, and represented by the point (0, 2) in Figs. 1–3].

In the case of a perfect fluid obeying the "stiff matter" equation of state $p = \mu$, the variable β' can attain any asymptotic value β_0', satisfying $-2 < \beta_0' < 2$, with a corresponding value of $x = 1 - \beta_0'^2/4$; the type of singularity depends on whether $\beta_0' \gtrless 1$. The general situation is that there is one cigar singularity $(\beta_0' > 1)$ and one point singularity $(\beta_0' < 1)$; at both singularities the quantities θ, σ, and μ are of equal importance in Raychaudhuri's Eq. (2.6). In some, but not all, of these models, there is an instant of time at which $\sigma = 0$. There is one special (time-symmetric) solution, given by $\beta' \equiv 1$, in which both singularities are "barrels" $(X \to X_0 > 0,\ Y \to 0$, where X_0 is constant), and μ, σ, and θ are equally significant. This special solution is not expanding in the X direction $(\partial/\partial r)$, and was first discovered by Thorne.[7]

It is of interest to note that all known exact solutions of our class (see references cited in Sec. 1) possess singularities as described above, and that the evolution of some of the dust, radiation, and stiff matter models,

TABLE I. This table summarizes the nature of the singularity, and gives the asymptotic behavior of all physically relevant variables: the fluid density μ, the fluid expansion θ, the fluid shear σ, the average length-scale l, in the rest space of the fluid, the length-scales $l_1 = X$ and $l_2 = Y$, in the directions parallel and orthogonal to the $\partial/\partial r$ direction, the integrated shear β, and the Ricci scaler R^*, of the spatially homogeneous hypersurfaces $\{t = \text{const}\}$. In each description the singularity is regarded as a past singularity labelled by $t = 0$ and f denotes "tends to finite nonzero limit." In the case when a point singularity arises for values of γ in the range $1 \leq \gamma < 2$, the value of β is finite. By relabelling β, Ω and r it is then possible to regard β as tending to zero, without any loss of generality.

Value of γ $[p = (\gamma - 1)\mu]$	(x, β')	Description of singularity	μ	θ	σ	l	$l_1 = X$	$l_2 = Y$	β	$R^* = 2/Y^2$	Dominant terms in Raychaudhuri Eq. (2.6)
	(1, 0)	Point matter singularity; particle horizons in all directions	$1/t^2$	$1/t$	$1/t^{(4-3\gamma)/\gamma}$	$t^{2/3\gamma}$	$t^{2/3\gamma}$	$t^{2/3\gamma}$	$t^{(2/3\gamma)(3\gamma-2)}$	$1/t^{4/3\gamma}$	$\mu, \theta^2 \propto 1/t^2$
$1 \leq \gamma < 2$	(0, −2)	Pancake matter singularity; particle horizon removed in $\partial/\partial r$ direction only	$1/t^\gamma$	$1/t$	$1/t$	$t^{1/3}$	t	f	$\ln t$	f	$\theta^2, \sigma^2 \propto 1/t^2$
	(0, 2)	Cigar matter singularity; particle horizons in all directions	$1/t^\gamma$	$1/t$	$1/t$	$t^{1/3}$	$1/t^{1/3}$	$t^{2/3}$	$-\ln t$	$1/t^{4/3}$	$\theta^2, \sigma^2 \propto 1/t^2$
$\gamma = 2$	(x_0, β_0') $-2 < \beta_0' < 2$, $x_0 = 1 - \frac{1}{4}\beta_0'^2$	Point $(-2 < \beta_0' < 1)$, barrel $(\beta_0' = 1)$, or cigar $(1 < \beta_0' < 2)$ matter singularity; particle horizons in all directions	$1/t^2$	$1/t$	$1/t$	$t^{1/3}$	$t^{(1-\beta_0')/3}$	$t^{(2+\beta_0')/6}$	$-\ln t (\beta_0' \neq 0)$; $t^{4/3}$ $(\beta_0' = 0)$	$1/t^{(2+\beta_0')/3}$	$\mu, \theta^2, \sigma^2 \propto 1/t^2$

graphed and examined numerically by Kantowski and Sachs,[4,5] is in agreement with that obtained herein.

We summarize our findings on the detailed nature of the singularities in Table I, and conclude with conformal diagrams for the totally geodesic timelike 2-surfaces $\{\theta, \phi = \text{const}\}$ in Figs. 5 and 6. In the case where one of the singularities is a pancake, a rather distinctive situation arises (see Fig. 6), and is related to the disappearance of particle horizons in the direction of $\partial/\partial r$.

5. CONCLUSION

We have shown that certain properties of known exact solutions of the Kantowski–Sachs cosmological models are general features of these models when there is no specialization made to a particular equation of state. In particular, we have shown that for any physically reasonable perfect fluid, there are matter singularities to the finite past and finite future. A detailed analysis was carried out in the case where the perfect fluid obeys a barotropic equation of state of the form $p = (\gamma - 1)\mu$, and an examination was made of the behavior of all physically relevant variables (summarized in Table I).

Although the Kantowski–Sachs models are of a very special kind, any may not be applicable as an adequate description of the real universe, the results herein are regarded as important, since they indicate the caution with which any examination of the singularity in general relativity should be carried out, and exemplify the type of careful argument that will be needed in more general situations (cf. Ref. 18).

ACKNOWLEDGMENTS

I am grateful to Dr. R. P. Geroch for some encouraging remarks, and for bring to my attention an error in the original version of the proof in Appendix A. This work was supported in part by an operating grant (No. A3978) from the National Research Council.

APPENDIX A: THE FOUR-PARAMETER LIE GROUPS ADMIT THREE-PARAMETER SUBGROUPS

Let L be the Lie algebra of the generators ξ_i $(i = 1, 2, 3, 4)$ of a four-parameter Lie group. Thus L is given by the commutator relations

$$[\xi_i, \xi_j] = C^k_{ij}\xi_k.$$

where the C^k_{ij} are the structure constants of the group, satisfying

$$C^k_{ij} = C^k_{[ij]}$$

and

$$C^i_{[jk}C^m_{l]i} = 0, \tag{A1}$$

which are known respectively as the first and second Jacobi identities. The derived algebra L' of L, i.e., the vector subspace spanned by the six vectors $C^k_{ij}\xi_k$ with the antisymmetric product $[\ ,\ ]$, has dimension $d \le 4$. If $d \le 3$, then any basis of L' can be extended to a basis of L, in such a way that a three-dimensional Lie subalgebra results. The only remaining case, where $d = 4$, does not exist. For in that case $L' = L$, and the six vectors $C^k_{ij}\xi_k$ span L, which has basis ξ_i and rank $(C^k_{ij}) = 4$, where (jk) is treated as a single index under the code $1 = (12)$, $2 = (13)$, $3 = (14)$, $4 = (34)$, $5 = (42)$, $6 = (23)$. Also if ε^{bcef} is the completely antisymmetric symbol ($\varepsilon^{bcef} = \varepsilon^{[bcef]}$, $\varepsilon^{1234} = 1$), $\varepsilon^{bcef}C^a_{ef}$ has rank 4. Contracting (A1), we have $C^e_{[ac}C^a_{d]e} = 0 \Leftrightarrow C^e_{cd}C^a_{ae} = 0 \Leftrightarrow C^a_{ae} = 0$ (since rank $(C^k_{ij}) = 4$). Thus

$$\varepsilon^{de[bc}C^{a]}_{de} = \tfrac{2}{3}\varepsilon^{bcad}C^e_{ed} = 0 \tag{A2}$$

and so

$$\varepsilon^{debc}C^f_{bc}C^a_{de}\xi_a = [\varepsilon^{debc}C^a_{de}C^f_{bc} - 2\varepsilon^{deca}C^b_{de}C^f_{cb}]\xi_a$$

$$= 3\varepsilon^{de[bc}C^{a]}_{de}C^f_{bc}\xi_a = 0 \text{ by (A2).}$$

Recalling that rank $(\varepsilon^{bcef}C^a_{ef}) = 4$, there are four linear combinations of the six vectors $C^a_{bc}\xi_a$ which vanish, and hence $4 = \dim L' \le 2$, a contradiction.

APPENDIX B: THE CASE OF THREE-PARAMETER SUBGROUPS WITH TWO-DIMENSIONAL ORBITS

It is well known that if an n-dimensional manifold admits a maximal continuous isometry group of $\frac{1}{2}n(n+1)$ parameters, then the manifold is of constant curvature.[25] Consequently, if a three-parameter isometry group G has two-dimensional orbits, those orbits are of constant curvature, which may be positive, zero, or negative. These orbits may be considered as the surfaces imbedded in flat three-dimensional space:

(a) $x^{1^2} + x^{2^2} + x^{3^2} = 1$ (positive curvature),

(b) $l_1x^1 + l_2x^2 + l_3x^3 = m$ (where $l_1^2 + l_2^2 + l_3^2 \neq 0$) (zero curvature),

and

(c) $x^{1^2} + x^{2^2} - x^{3^2} = -1$ (negative curvature).

In cases (a) and (b) the metric of the 3-space is $ds^2 = dx^{1^2} + dx^{2^2} + dx^{3^2}$ whereas in case (c) it is the indefinite form $ds^2 = dx^{1^2} + dx^{2^2} - dx^{3^2}$. The generators ξ_1, ξ_2, ξ_3 of the group G can be chosen to satisfy the Lie algebra

$$[\xi_1, \xi_2] = \xi_3, \quad [\xi_2, \xi_3] = k\xi_1, \quad [\xi_3, \xi_1] = \xi_2, \tag{B1}$$

where $k = +1$ [case (a)], $k = 0$ [case (b)], or $k = -1$ [case (c)]. The corresponding groups are of Bianchi types VII_0 $(k = 0)$, VIII $(k = -1)$ and IX $(k = +1)$ (see, e.g., Ref. 1). We now show that if this algebra is extended to a four-dimensional Lie algebra L, then, if $k = 0$ or -1, L will have a three-dimensional subalgebra which is not of type (B1), whereas if $k = +1$ the only three-dimensional subalgebra of L is of type (B1) with $k = +1$. This will suffice to show that if a three-dimensional manifold admits a four-parameter isometry group, then it is only in the case where there is a three-parameter subgroup with two-dimensional orbits that the manifold will not admit a simply transitive group.

Following Kantowski,[4] we introduce a fourth vector, η, which is an independent basis vector of the Lie algebra L. In addition to (B1), we must have relations of the form

$$[\eta, \xi_i] = A_i\eta + B^j_i\xi_j \quad (i = 1, 2, 3; \text{ summation over } j = 1, 2, 3)$$

to complete the Lie algebra. The quantities A_i and B^j_i are constants. The second Jacobi identities (B2) yield the following equations:

$$kA_1 = A_2 = A_3 = 0, \tag{B2}$$

$$B^1_3 - A_1 B^1_2 + kB^3_1 = 0, \tag{B3}$$

$$B^2_3 - A_1 B^2_2 + B^3_2 = 0, \tag{B4}$$

$$B^3_3 - B^2_2 - A_1 B^3_2 - B^1_1 = 0, \tag{B5}$$

$$kB^1_1 - kB^3_3 - kB^2_2 = 0, \tag{B6}$$

$$kB^2_1 + B^1_2 = 0, \tag{B7}$$

$$kB^3_1 + B^1_3 = 0, \tag{B8}$$

$$B^1_2 + A_1 B^1_3 + kB^2_1 = 0, \tag{B9}$$

$$B^2_2 + A_1 B^2_3 - B^1_1 - B^3_3 = 0, \tag{B10}$$

and

$$B^3_2 + A_1 B^3_3 + B^2_3 = 0. \tag{B11}$$

These equations are derived from the relation $[\eta, [\xi_1, \xi_2]] + [\xi_1, [\xi_2, \eta]] + [\xi_2, [\eta, \xi_1]] = \mathbf{0}$, together with its two counterparts obtained by cyclic permutation of ξ_1, ξ_2, and ξ_3.

Consider first the case where $k \neq 0$. Then from (B2), $A_1 = 0$, and Eqs. (B5), (B6) and (B10) imply $B^1_1 = B^2_2 = B^3_3 = 0$; the resulting equations give $B^1_3 = -kB^3_1$, $B^1_2 = -kB^2_1$ and $B^2_3 = -B^3_2$. Now instead of η choose the vector $\eta' = \eta - B^3_2\xi_1 + B^3_1\xi_2 - B^2_1\xi_3$. Then $[\eta', \xi_1] = [\eta', \xi_2] = [\eta', \xi_3] = 0$.

Next, we consider the case $k = 0$. From (B7) and (B8), we have $B^1_2 = B^1_3 = 0$. We can transform B^2_1 and B^3_1 to zero by the allowable transformation $\xi_1 \to \xi_1 - B^3_1\xi_2 + B^2_1\xi_3$, which preserves (B1). Thus

$$[\eta, \xi_1] = A_1\eta + B^1_1\xi_1,$$

$$[\eta, \xi_2] = B^2_2\xi_2 + B^3_2\xi_3,$$

and

$$[\eta, \xi_3] = B^2_3\xi_2 + B^3_3\xi_3.$$

From (B4) and (B11), $A_1(B^2_2 + B^3_3) = 0$, so either $A_1 \neq 0$ and $B^2_2 = -B^3_3$, or $A_1 = 0$. Suppose first that $A_1 \neq 0$. Replacing η by $\eta' = \eta + (B^1_1/A_1)\xi_1$ maintains $B^2_1 = B^3_1 = 0$, and transforms B^1_1 to zero. Then Eqs. (B5) and (B10) imply $B^3_2 + B^2_3 = 0$, which, from (B4) and (B11), requires $B^2_2 = B^3_3 = 0$. From Eqs. (B5) and (B10) it now follows that $B^3_2 = B^2_3 = 0$, and hence

$$[\eta', \xi_1] = A_1\eta', \quad [\eta', \xi_2] = [\eta', \xi_3] = \mathbf{0}.$$

Now suppose that $k = A_1 = 0$. By Eqs. (B5) and (B10) we obtain $B^1_1 = 0$ and $B^2_2 = B^3_3$; Eq. (B4) implies that $B^2_2 + B^3_3 = 0$. The transformations $\eta \to \eta' = (1/B^2_2)(\eta - B^3_2\xi_1)$ in the case $B^2_2 \neq 0$, and $\eta \to \eta' = \eta - B^3_2\xi_1$ in the case $B^2_2 = 0$, have the effect of transforming B^3_2 to zero and $B^2_2 = B^3_3$ to 1 (if $B^2_2 \neq 0$). Hence $[\eta, \xi_1] = 0$, $[\eta, \xi_2] = \epsilon\xi_2$, and $[\eta, \xi_3] = \epsilon\xi_3$, where $\epsilon = 0$ or 1.

We have thus determined that any four-dimensional Lie algebra containing Lie algebras of type (B1) has structure given by

(i) $k = +1$: $[\xi_1, \xi_2] = \xi_3$, $[\xi_2, \xi_3] = \xi_1$, $[\xi_3, \xi_1] = \xi_2$,
$[\eta, \xi_1] = \mathbf{0}$, $[\eta, \xi_2] = \mathbf{0}$, $[\eta, \xi_3] = \mathbf{0}$,

(ii) $k = -1$: $[\xi_1, \xi_2] = \xi_3$, $[\xi_2, \xi_3] = -\xi_1$, $[\xi_3, \xi_1] = \xi_2$,
$[\eta, \xi_1] = \mathbf{0}$, $[\eta, \xi_2] = \mathbf{0}$, $[\eta, \xi] = \mathbf{0}$,

(iii) $k = 0$: $[\xi_1, \xi_2] = \xi_3$, $[\xi_2, \xi_3] = \mathbf{0}$, $[\xi_3, \xi_1] = \xi_2$,
$[\eta, \xi_1] = A\eta$, $[\eta, \xi_2] = \epsilon\xi_2$, $[\eta, \xi_3] = \epsilon\xi_3$,

where either $\epsilon = 1$ and $A = 0$, or $\epsilon = 0$.

It is a straightforward but tedious computation to determine *all* possible three-dimensional subalgebras of these four-dimensional algebras. The results are

(i) $k = +1$: The only case is that spanned by ξ_1, ξ_2, ξ_3, with $[\xi_1, \xi_2] = \xi_3$, $[\xi_2, \xi_3] = \xi_1$, $[\xi_3, \xi_1] = \xi_2$, which is of Bianchi Type IX.

(ii) $k = -1$: In addition to the original subalgebra that is spanned by ξ_1, ξ_2, ξ_3, with structure $[\xi_1, \xi_2] = \xi_3$, $[\xi_2, \xi_3] = -\xi_1$, $[\xi_3, \xi_1] = \xi_2$ there is a family of subalgebras. This family of subalgebras is spanned by $\lambda_1 = \eta$, $\lambda_2 = \xi_1 + a\xi_2 + b\xi_3$, and $\lambda_3 = c\xi_2 + d\xi_3$, where the constants a, b, c, and d satisfy $c^2(1 - b^2) + 2abcd + d^2(1 - a^2) = 0$, $c^2 + d^2 > 0$. It is essentially a one-parameter family, since if $c \neq 0$ then, without loss of generality, $c = 1$ and $a = 0$ (which implies $b^2 = 1 + d^2$), and if $c = 0$ then without loss of generality $d = 1$ and $b = 0$ (and so $a = \pm 1$). The Lie algebra formed by λ_1, λ_2, λ_3 is given by $[\lambda_1, \lambda_2] = \mathbf{0}$, $[\lambda_2, \lambda_3] = -(ad - bc)\lambda_2 \pm (a^2 + b^2 - 1)^{1/2}\lambda_3$, $[\lambda_3, \lambda_1] = \mathbf{0}$. By appropriate linear transformations of λ_2 and λ_3 with either of the above specializations, this Lie algebra can be put into the canonical form $[\lambda_1, \lambda_2] = \mathbf{0}$, $[\lambda_2, \lambda_3] = \lambda_2$, $[\lambda_3, \lambda_1] = 0$, and is a special case of Type VI_h with $h = -1$, i.e., a special case of Bianchi Type III (cf. Ref. 1).

(iii) $k = 0$: There is in addition to the subalgebra spanned by ξ_1, ξ_2, ξ_3, with structure $[\xi_1, \xi_2] = \xi_3$, $[\xi_2, \xi_3] = \mathbf{0}$, $[\xi_3, \xi_1] = \xi_2$, a one-parameter family of subalgebras spanned by $\lambda_1 = \eta + a\xi_1$, $\lambda_2 = \xi_2$ and $\lambda_3 = \xi_3$, with structure $[\lambda_1, \lambda_2] = \epsilon\lambda_2 + a\lambda_3$, $[\lambda_2, \lambda_3] = 0$, $[\lambda_3, \lambda_1] = a\lambda_2 - \epsilon\lambda_3$. This is of Bianchi Type I if $a = \epsilon = 0$, Bianchi Type V if $a = 0$, $\epsilon = 1$, Type VII_0 if $a \neq 0$, $\epsilon = 0$, and Type VII_h if $a \neq 0$, $\epsilon = 1$. In the case $a \neq 0$, $\epsilon = 0$ the transformation $\lambda_1 \to a\lambda_1$ reduces a to the value $a = 1$. It follows that in the $k = 0$ and $k = -1$ cases, there are subalgebras which are not of the type (B1), whereas in the $k = +1$ case there are no such subalgebras. If a space–time admits a four-parameter isometry group acting multiply transitively on spacelike hypersurfaces, with a three-parameter subgroup whose orbits are two-dimensional, it follows that when $k = -1$ there is a three-dimensional isometry group of Type III acting simply transitively on the hypersurfaces, and that when $k = 0$ there are two three-dimensional isometry groups, of Types I and VII_0, or of Types V and VII_h, acting simply transitively on the hypersurfaces. These results are consistent with, but not necessitated by, those of Ellis and MacCallum.[1] In particular, there is only one case, that of $k = +1$, where there is no three-parameter group acting transitively on spacelike hypersurfaces.

APPENDIX C: THE PETROV TYPE OF THE KANTOWSKI-SACHS MODELS

We consider components of the Weyl tensor in the orthonormal tetrad $\mathbf{e}_0 = \partial/\partial t$, $\mathbf{e}_1 = (1/X)\partial/\partial r$, $\mathbf{e}_2 = (1/Y)\partial/\partial\theta$, $\mathbf{e}_3 = (1/Y\sin\theta)\partial/\partial\phi$. The commutation relations are

$$[\mathbf{e}_0, \mathbf{e}_1] = -(\dot{X}/X)\mathbf{e}_1, \quad [\mathbf{e}_0, \mathbf{e}_2] = -(\dot{Y}/Y)\mathbf{e}_2,$$

$$[\mathbf{e}_0, \mathbf{e}_3] = -(\dot{Y}/Y)\mathbf{e}_3, \quad [\mathbf{e}_1, \mathbf{e}_2] = 0,$$

$$[\mathbf{e}_2, \mathbf{e}_3] = -(1/Y)\cot\theta\,\mathbf{e}_3, \quad [\mathbf{e}_3, \mathbf{e}_1] = 0.$$

It follows, using a decomposition as in Refs. 1, 26, that the acceleration ($\dot{\mathbf{e}}_0$) and the vorticity of the fluid is zero, and that the orthonormal tetrad consists of Fermi-propagated shear eigenvectors that are Ricci eigenvectors of the homogeneous hypersurfaces. Decomposing the Weyl tensor $C_{abcd} = R_{abcd} + g_{a[d}R_{c]b} + g_{b[c}R_{d]a} + (R/3)g_{a[c}g_{d]b}$ into its "electric" and "magnetic" parts, E_{ab} and H_{ab}, defined by $E_{ab} = C_{apbq}u^p u^q$ and $H_{ab} = \frac{1}{2}\eta_{cdbs}C_a{}^{tcd}u_t u^s$, we find (by using the Ricci identity $u^a{}_{;d;c} - u^a{}_{;c;d} = R^a{}_{bcd}u^b$ for $\mathbf{u} = \mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2$, and $\mathbf{e}_3$ in turn) that the components in the orthonormal frame are $H_{ab} = 0$, $E_{ab} = \mathrm{diag}(0, 2E, -E, -E)$, where $E = -\frac{1}{2}[\dot{X}/X + \frac{1}{6}(\mu + 3p)]$. This means that the models are of Petrov type D, except in the case $E = 0$, which is in fact untenable. This is because the restriction $E = 0$, together with the field equations (2.2)–(2.4), necessitates $\mu = 3[(1+\dot{Y}^2)/Y^2]$. Differentiating this, and using (2.3), we obtain $\dot{\mu} = -(3\dot{Y}/Y)(\mu + p)$. Comparing this with the conservation equation (2.4), it follows that $\dot{X}/X = \dot{Y}/Y$, since $\mu + p > 0$. This is incompatible with Eq. (2.2). Consequently, $E \neq 0$, and the space–times must be of Petrov Type D.

In fact, by local rotational symmetry, the only *conceivable* Petrov Types are D and 0 (conformally flat). The above proof shows directly that the conformally flat case ($E = 0$) is impossible. This result can also be derived by recalling that if the space–time is conformally flat and contains a perfect fluid for which $\mu + p > 0$, then $\sigma_{ab} = 0$ (cf. Ref. 17), which contradicts (2.7).

This result is consistent with, but not necessitated by, a theorem of Wainwright.[27]

[1] G. F. R. Ellis and M. A. H. MacCallum, Commun. Math. Phys. **12**, 108 (1969).
[2] G. F. R. Ellis, J. Math. Phys. **8**, 1171 (1967).
[3] J. M. Stewart and G. F. R. Ellis, J. Math. Phys. **9**, 1072 (1968).
[4] R. Kantowski, Ph.D. thesis, University of Texas (1966); cf. Appendices A and B of the present paper.
[5] R. Kantowksi and R. K. Sachs, J. Math. Phys. **7**, 443 (1976).
[6] A. S. Kompaneets and A. S. Chernov, Zh. Eksp. Teor. Fiz. **47**, 1939 (1964) [Sov. Phys. JETP **20**, 1303 (1965)].
[7] K. S. Thorne, Ph.D. thesis, Princeton University (1965).
[8] A. G. Doroshkevich, Astrofiz. **1**, 255 (1965) [Astrophys. **1**, 138 (1965)].
[9] K. S. Thorne, Astrophys. J. **148**, 51 (1967).
[10] B. Bertotti, Phys. Rev. **116**, 1331 (1959).
[11] I. Robinson, Bull. Acad. Pol. Sci. **7**, 351 (1959).
[12] K. S. Thorne, "The Resistance of Magnetic Flux to Gravitational Collapse," in *Quasars and High Energy Astronomy; Including the Proceedings of the Second Texas Symposium on Relativistic Astrophysics, December* 15-19, 1964, edited by J. N. Douglas, I. Robinson, A. Schild, E. L. Schücking, J. A. Wheeler, and N. J. Woolf (Gordon and Breach, New York, London, and Paris, 1969).
[13] G. F. R. Ellis, Gen. Rel. Grav. **2**, 7 (1971).
[14] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time* (Cambridge U. P., Cambridge, 1973).
[15] A. Raychaudhuri, Phys. Rev. **98**, 1123 (1955).
[16] A. Raychaudhuri, Z. Astrophys. **43**, 161 (1957).
[17] G. F. R. Ellis, "Relativistic Cosmology," in *General Relativity and Cosmology*, XLVII Corso, Varenna, Italy, 1969, edited by R. K. Sachs (Academic, New York, 1971).
[18] C. B. Collins and G. F. R. Ellis, "Singularities in Bianchi Cosmologies," in *Bianchi Universes and Relativistic Cosmology*, edited by R. Ruffini (to be published) (1977).
[19] C. B. Collins, Commun. Math. Phys. **23**, 137 (1971).
[20] C. B. Collins, Commun. Math. Phys. **27**, 37 (1972).
[21] C. B. Collins, Commun. Math. Phys. **39**, 131 (1974).
[22] I. S. Shikin, Dokl. Akad. Nauk SSSR **176**, 1048 (1967) [Sov. Phys. Dokl. **12**, 950 (1968)].
[23] I. S. Shikin, Zh. Eksp. Teor. Fiz. **63**, 1629 (1972) [Sov. Phys. JETP **36**, 811 (1976)].
[24] I. S. Shikin, Zh. Eksp. Teor. Fiz. **68**, 1583 (1975) [Sov. Phys. JETP **41**, 794 (1976)].
[25] L. P. Eisenhart, *Riemannian Geometry* (Princeton U. P., Princeton, N.J., 1966).
[26] M. A. H. MacCallum, "Cosmological Models from a Geometric Point of View," in *Cargèse Lectures in Physics, Vol.* 6, edited by E. Schatzman (Gordon and Breach, New York, 1973).
[27] J. Wainwright, Commun. Math. Phys. **17**, 42 (1970).

A necessary condition for the validity of Huygens' principle on a curved space-time*

Riccardo Goldoni

University of Pisa, Istituto di Matematica "L. Tonelli," 56100 Pisa, Italy
(Received 29 December 1976)

It is proved that a necessary condition for the validity of Huygens' principle on a curved space-time V^4 is that V^4 be an Einstein space. In connection with this result some remarks about the strong and weak formulations of Mach's principle are also pointed out.

1. INTRODUCTION

The pioneering work about scalar Green's functions on a curved space-time was done by Hadamard,[1] Sobolev,[2] and Schwartz.[3] Tensorial Green's functions on a curved manifold were then introduced by Lichnerowicz[4-7] and DeWitt and Brehme[8] for different purposes. They have since been studied by several authors (cf. Refs. 9–12, among others).

One of the main peculiarities of such Green's functions is that, in general, they have a tail, i.e., in general, the support of the Green's function of a linear second order partial differential equation on an arbitrary space-time is not only the surface of the null cone, but also its interior,[8,13] so that Huygens' principle, in general, is not satisfied. This feature is present, for example, in the generally covariant formulations of the Einstein field equations,[14] so that, according to such formulations, the gravitational interaction is not forced to propagate along null geodesics.

A great deal of work has been done in order to find necessary conditions for the validity of Huygens' principle on a curved space-time V^4 (cf. Ref. 13 and references therein). In this paper we will prove that a necessary condition is that V^4 be an Einstein space.

In Sec. 2 we state the main theorems. In Sec. 3 we prove the main theorems in the case of a scalar kernel. In Sec. 4 we generalize the results of Sec. 3 to the case of tensorial kernels. Finally, Sec. 5 is devoted to some concluding remarks.

2. STATEMENT OF THE MAIN THEOREMS

Let (V^4, g) be a four-dimensional Riemannian manifold with signature -2 and D a causal domain of V^4. Moreover let

$$\sigma[\bar{x}, x] = \tfrac{1}{2} d^2[\bar{x}, x], \tag{2.1}$$

where $\bar{x}, x \in D$ and $d[\bar{x}, x]$ denotes the geodesic distance between $\bar{x}$ and x. Consider tailless kernels of the type

$$G_M^{(\pm)\bar{M}}(\bar{x}, x) = F(\bar{x}, x) g_M^{\bar{M}} \delta^{(\pm)}(\sigma[\bar{x}, x]), \tag{2.2}$$

where M and $\bar{M}$ denote the multi-indices

$$M = (\mu_1, \mu_2, \ldots, \mu_n), \tag{2.3}$$

$$\bar{M} = (\bar{\mu}_1, \bar{\mu}_2, \ldots, \bar{\mu}_n), \tag{2.4}$$

$F(\bar{x}, x)$ is a biscalar in $C^\infty(D \times D)$ not identically equal to zero on the surface $\sigma = 0$, $g_M^{\bar{M}}$ is defined according to the following equation:

$$g_{\bar{T}T} = g_{\bar{\alpha}\alpha} g_{\bar{\beta}\beta} \cdots g_{\bar{\eta}\eta}, \tag{2.5}$$

where

$$T = (\alpha, \beta, \ldots, \eta), \quad \bar{T} = (\bar{\alpha}, \bar{\beta}, \ldots, \bar{\eta}), \tag{2.6}$$

and $g_{\bar{\alpha}\alpha}$ denotes the parallel displacement bivector between $\bar{x}$ and x. $\delta^{(\pm)}$ denotes the advanced (+ sign) or retarded (− sign) Dirac distribution. Greek indices run from 0 to 3.

Finally let us consider the following linear second order differential equation

$$[P_M^{A\gamma\delta}(x)\nabla_\gamma\nabla_\delta + Q_M^{A\gamma}(x)\nabla_\gamma + U_M^A(x)] G_A^{(\pm)\bar{M}}(\bar{x}, x) = 8\pi g_M^{\bar{M}} \delta^4(\bar{x}, x), \tag{2.7}$$

where

$$A = (\alpha_1, \alpha_2, \ldots, \alpha_n) \tag{2.8}$$

and $\mathbf{P} = \{P_M^{A\gamma\delta}\}$, $\mathbf{Q} = \{Q_M^{A\gamma}\}$, and $\mathbf{U} = \{U_M^A\}$ are tensors in $C^\infty(D)$. ∇_μ denotes the covariant derivative with respect to x and $\delta^4(\bar{x}, x)$ is defined according to the following equation:

$$\langle \delta^4(\bar{x}, x), f(x) \rangle = f(\bar{x}). \tag{2.9}$$

The purpose of this paper is to prove the following theorem.

Main Theorem: Necessary conditions for the solvability of Eq. (2.7) with respect to the tensors $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{U}$ are

$$\hat{\mathbf{P}}(x) = K(x)\mathbf{g}(x), \tag{2.10}$$

$$\mathbf{R}(x) = c\mathbf{g}(x), \tag{2.11}$$

where

$$K(x) = [g_{\alpha\beta}(x)\hat{P}^{\alpha\beta}(x)]/4, \tag{2.12}$$

$$\mathbf{R} = \{R_{\alpha\beta}\}, \tag{2.13}$$

c is a constant and $\hat{\mathbf{P}}$ is defined as follows.

In the case of n even, say

$$n = 2m, \tag{2.14}$$

divide the set of $2m$ indices M into two sets of m indices, $M1$ and $M2$, each one ordered in an arbitrary way. Also define the multi-indices $A1$ and $A2$ obtained from A through a similar, but independent, procedure. $\hat{\mathbf{P}}$ turns out to be defined by the following equation:

$$\hat{P}^{\gamma\delta}(x) = g_{A1\,A2} g^{M1\,M2} [P_M^{A\gamma\delta}(x) + P_M^{A\delta\gamma}(x)]/2, \tag{2.15}$$

where $g_{A1\,A2}$ and $g_{M1\,M2}$ are defined according to Eq. (2.5). Of course, $\hat{\mathbf{P}}$ may depend on the choice of the multi-in-

 Copyright © 1977 American Institute of Physics

dices $M1$, $M2$, $A1$, and $A2$. Equation (2.10), however, is required to hold for any allowable choice.

In the case of n odd, say

$$n = 2m + 1 , \tag{2.16}$$

select an arbitrary index of A ($\alpha = \alpha_j$, say) and an arbitrary index of M ($\mu = \mu_k$, say) and then define the multi-indices N and B obtained from M and A by dropping the indices μ_k and α_j, respectively. Finally define the multi-indices $N1$, $N2$, $B1$, and $B2$ through the same procedure which allowed us to define the multi-indices $M1$, $M2$, $A1$, and $A2$ in the case of n even. In self-explanatory notation, $\hat{\mathbf{P}}$ turns out to be defined by the following equation:

$$\hat{P}^{\gamma\delta}(x) = g_{B1\,B2}\, g^{N1\,N2}[P^{B\alpha\gamma\delta}_{N\alpha}(x) + P^{B\alpha\delta\gamma}_{N\alpha}(x)]/2 . \tag{2.17}$$

Again $\hat{\mathbf{P}}$ may depend on the choice of α_j, μ_k, $N1$, $N2$, $B1$, and $B2$ and we require that Eq. (2.10) hold for any allowable choice.

From the main theorem the following theorem follows easily.

Let (V^4, g) denote a four-dimensional Riemannian manifold with signature -2 and let D be a causal domain of V^4. Consider the linear second order differential operator O defined by the equation

$$(\mathrm{O}f)_M = \nabla^\mu \nabla_\mu f_M + Q^{A\beta}_M \nabla_\beta f_A + U^A_M f_A , \tag{2.18}$$

where $\mathbf{f} = \{f_M\}$ is a tensor and Q and U are tensors in $C^\infty(D)$.

Theorem 1: A necessary condition for the differential operator O to satisfy Huygens' principle is that the Ricci tensor R is proportional to the metric tensor g by a constant factor.

Theorem 1 follows from the main theorem since adding a tail term to the kernels (2.2) gives the most general form of the retarded (– sign) or advanced (+ sign) Green's functions of the differential operator O.[12]

3. THE CASE OF A SCALAR KERNEL

We will now prove the main theorem in the case of a scalar kernel to better point out the technique of the proof.

Let (V^4, g) be a four-dimensional Riemannian manifold with signature -2 and let D denote a causal domain of V^4. Consider the scalar kernels

$$G^{(\pm)}(\bar{x}, x) = F(\bar{x}, x)\, \delta^{(\pm)}(\sigma[\bar{x}, x]) , \tag{3.1}$$

where $F(\bar{x}, x)$ denotes a biscalar in $C^\infty(D \times D)$ not identically equal to zero on the surface $\sigma = 0$. Finally consider the following linear second order differential equation:

$$[P^{\mu\nu}(x)\nabla_\mu \nabla_\nu + Q^\mu(x)\nabla_\mu + U(x)]G^{(\pm)}(\bar{x}, x) = 8\pi\delta^4(\bar{x}, x) , \tag{3.2}$$

where the tensors $\mathbf{P} = \{P^{\gamma\delta}\}$, $\mathbf{Q} = \{Q^\gamma\}$, and U are in $C^\infty(D)$. Moreover $P^{\gamma\delta}$ may be assumed to be a symmetric tensor without loss of generality.

Theorem 2: Necessary conditions for the solvability of Eq. (3.2) with respect to the tensor P, Q, and U are

$$\mathrm{P}(x) = K(x)\, \mathrm{g}(x) , \tag{3.3}$$

$$\mathrm{R}(x) = c\mathrm{g}(x) , \tag{3.4}$$

where

$$K(x) = [g_{\gamma\delta}(x)P^{\gamma\delta}(x)]/4 \tag{3.5}$$

and c is a constant.

To prove Theorem 2 let us first note that by a standard procedure Eq. (3.2) can be rewritten as

$$\{2P^{*\gamma\delta}\sigma_{;\gamma}\sigma_{;\delta} - \sigma P^{*\gamma\delta}\sigma_{;\gamma;\delta} - \sigma\sigma_{;\gamma}Q^{*\gamma} + U^*\sigma^2\}\delta^{(\pm)(2)}(\sigma[\bar{x}, x]) + 4\pi P^*\delta^4(\bar{x}, x) = 16\pi\delta^4(\bar{x}, x) , \tag{3.6}$$

where

$$P^{*\xi\xi} = P^{*\xi\xi}(\bar{x}, x) = F(\bar{x}, x)P^{\xi\xi}(x) , \tag{3.7}$$

$$Q^{*\xi} = Q^{*\xi}(\bar{x}, x) = Q^\xi(x) - 2P^{\xi\xi}(x)F_{;\xi}(\bar{x}, x) , \tag{3.8}$$

$$U^* = U^*(\bar{x}, x) = U(x) + P^{\xi\xi}(x)F_{;\xi;\xi}(\bar{x}, x) + Q^\xi(x)F_{;\xi}(\bar{x}, x) , \tag{3.9}$$

$$P^* = P^*(\bar{x}, x) = g_{\xi\xi}P^{*\xi\xi} , \tag{3.10}$$

$\delta^{(\pm)(2)}$ denotes the standard second order derivative of the $\delta^{(\pm)}$ distribution (cf., for example, Refs. 15 and 16). Moreover for simplicity the symbol ";" is used for the covariant differentiation.

Multiplying both sides of Eq. (3.6) by σ^t ($0 < t \ll 1$), one finds

$$\{2P^{*\gamma\delta}\sigma_{;\gamma}\sigma_{;\delta} - \sigma P^{*\gamma\delta}\sigma_{;\gamma;\delta} - \sigma\sigma_{;\gamma}Q^{*\gamma} + \sigma^2 U^*\} \times \sigma^t\delta^{(\pm)(2)}(\sigma[\bar{x}, x]) = 0 . \tag{3.11}$$

It is also easy to prove the following lemma.

Lemma 1: If $H = H(\bar{x}, x)$ is a biscalar, then

$$H\delta^{(\pm)(2)}(\sigma[\bar{x}, x]) = 0 \tag{3.12}$$

if and only if

$$H(\bar{x}, x) = L(\bar{x}, x)\sigma^2 , \tag{3.13}$$

with

$$L(\bar{x}, x) \xrightarrow{\sigma \to 0} 0 . \tag{3.14}$$

Lemma 1 essentially follows from the well-known identity

$$2\delta^{(\pm)}(\sigma) - \sigma^2\delta^{(\pm)(2)}(\sigma) = 0 . \tag{3.15}$$

From Lemma 1 and Eq. (3.11) it follows that

$$2P^{*\gamma\delta}\sigma_{;\gamma}\sigma_{;\delta} - \sigma P^{*\gamma\delta}\sigma_{;\gamma;\delta} - \sigma\sigma_{;\gamma}Q^{*\gamma} + U^*\sigma^2 = 2L\sigma^{2-t} , \tag{3.16}$$

where

$$L \xrightarrow{\sigma \to 0} 0 . \tag{3.14}$$

The left-hand side of Eq. (3.16) may be rewritten as

$$\sigma(4P^{*\gamma\delta}d_{;\gamma}d_{;\delta} - P^{*\gamma\delta}\sigma_{;\gamma;\delta} - Q^{*\gamma}\sigma_{;\gamma} + U^*\sigma) . \tag{3.17}$$

As a consequence the limit as $\bar{x} \to x$ (cf. Ref. 8) of the function

$$L\sigma^{1-t} \tag{3.18}$$

and of all its partial derivatives are finite.

Differentiating both sides of Eq. (3.16) and taking the limit as $\bar{x} \to x$ of both sides of the so-obtained equation, one finds an identity.

Differentiating once more and taking the limit as $\bar{x} \to x$ of both sides of the so-obtained equation, one finds

$$\mathrm{P}^*(x,x) = K'(x)\mathrm{g}(x)\,, \tag{3.19}$$

where

$$\mathrm{P}^* = \{P^{*\gamma\delta}\} \tag{3.20}$$

and

$$K'(x) = [\,g_{\gamma\delta}(x)P^{*\gamma\delta}(x,x)]/4\,. \tag{3.21}$$

Inserting Eq. (3.7) in Eq. (3.19), one then finds

$$\mathrm{P}(x) = K(x)\mathrm{g}(x)\,, \tag{3.3}$$

where

$$K(x) = [\,g_{\gamma\delta}(x)P^{\gamma\delta}(x)]/4\,. \tag{3.5}$$

Equation (3.3) is the first of the two necessary conditions stated in Theorem 2.

For the second necessary condition (3.4) one inserts Eqs. (3.3) and (3.7) in Eq. (3.16) and divides both sides by σ to obtain

$$FK(4 - \sigma^{;\alpha}{}_{;\alpha}) - Q^{*\alpha}\sigma_{;\alpha} + \sigma U^* = 2L\sigma^{1-t}\,. \tag{3.22}$$

Performing an expansion to the second order of the first term on the left-hand side of Eq. (3.22), one finds[8]

$$-\tfrac{1}{3}\,FKR^{\alpha\beta}\sigma_{;\alpha}\sigma_{;\beta}\,. \tag{3.23}$$

As a consequence, the limits as $\bar{x} \to x$ of the function S defined by

$$2S\sigma = 2L\sigma^{1-t} + Q^{*\alpha}\sigma_{;\alpha} \tag{3.24}$$

and of all its partial derivatives are finite.

Inserting Eq. (3.24) in Eq. (3.22), differentiating both sides of the so-obtained equation, and taking the limit as $\bar{x} \to x$, one finds an identity. Differentiating once more and again taking the limit as $\bar{x} \to x$ of both sides, one finds

$$\mathrm{R}(x) = q(x)\mathrm{g}(x)\,, \tag{3.25}$$

where

$$q(x) = 3[U^*(x,x) - 2S(x,x)]/2K'(x)\,. \tag{3.26}$$

From purely geometrical arguments it follows that $q(x)$ has to be a constant c,

$$q(x) = c\,, \tag{3.27}$$

so that the second necessary condition (3.4) is established.

As a further remark, it should be noted that other necessary conditions may follow from condition (3.27). This possibility, however, will not be investigated in this paper.

Finally we note that the procedure which allowed us to prove Theorem 2 shows that the hypotheses $F(\bar{x},x) \in C^\infty(D\times D)$ and $\mathbf{P},\mathbf{Q},U \in C^\infty(D)$ are stronger than necessary and have been made only for the sake of simplicity.

4. GENERALIZATION TO TENSORIAL KERNELS

To generalize the results obtained in Sec. 3 to tensorial kernels one must consider separately the cases when n is even, say

$$n = 2m\,, \tag{2.14}$$

or odd, say

$$n = 2m+1\,. \tag{2.16}$$

In the case $n = 2m$ the proof of the main theorem can be reduced to the proof of Theorem 2 by multiplying both sides of Eq. (2.7) by

$$g^{M1\,M2}g_{\bar{M}1\,\bar{M}2}\,, \tag{4.1}$$

where $g_{M1\,M2}$ is defined as in Sec. 2 and $g_{\bar{M}1\,\bar{M}2}$ is defined from the multi-index $\bar{M}$ through a similar, but independent, procedure.

The main theorem is proved by performing the whole procedure above for any possible choice of $M1$, $M2$, $\bar{M}1$, and $\bar{M}2$.

In the case $n = 2m+1$ let us rewrite Eq. (2.7) as

$$[P^{B\alpha\gamma\delta}_{N\mu}(x)\nabla_\gamma\nabla_\delta + Q^{B\alpha\gamma}_{N\mu}(x)\nabla_\gamma + U^{B\alpha}_{N\mu}]$$

$$\times F(\bar{x},x)\,g^{\bar{N}}_{B}\,g^{\bar{\mu}}_{\alpha}\,\delta^{(\pm)}(\sigma[\bar{x},x]) = 8\pi g^{\bar{M}}_{M}\,\delta^4(\bar{x},x)\,, \tag{4.2}$$

where the multi-indices N and B have been defined in Sec. 2, while the multi-index $\bar{N}$ is obtained from $\bar{M}$ by dropping the index $\bar{\mu} = \bar{\mu}_j$.

The next step is to multiply both sides of Eq. (4.2) by

$$g_{\bar{N}1\,\bar{N}2}\,g^{N1\,N2}\,, \tag{4.3}$$

where the multi-indices $N1$ and $N2$ have been defined in Sec. 2, while the multi-indices $\bar{N}1$ and $\bar{N}2$ are obtained from $\bar{N}$ through a similar, but independent, procedure.

The last step is to perform formally the covariant differentiations in the equation thus obtained and, finally, to multiply both sides of the so-obtained equation by $g^{\mu}_{\bar{\mu}}$.

This procedure allows us to reduce the proof of the main theorem to the proof of Theorem 2.

By performing the whole procedure above for any allowable choice of μ, $\bar{\mu}$, $N1$, $N2$, $\bar{N}1$, and $\bar{N}2$ the main theorem can be proved.

5. CONCLUDING REMARKS

The main purpose of this paper has been to prove that a necessary condition for the validity of Huygens' principle on a curved space–time V^4 is that V^4 be an Einstein space. The proof has been given for differential operators O of the type (2.18) on the basis of the main theorem and of the result that by adding a tail term to the kernels (2.2) one obtains the most general form of the retarded (– sign) or advanced (+ sign) Green's functions of the differential operator O.[12] We also noted that the procedure which allowed us to prove the main theo-

rem leaves open the possibility of obtaining further necessary conditions for the validity of Huygens' principle.

As already mentioned in Sec. 1, the kernels peculiar to the generally covariant formulations of the Einstein field equations have a support not only on the surface of the null cone but also in its interior,[14] so that, according to such formulations, the gravitational interaction is not forced to propagate along null geodesics.

Some authors[17-21] have made attempts to construct theories of gravitation based on the hypothesis that the gravitational interaction travels along null geodesics. This is the simplest assumption for the purpose of constructing a purely integral[17] or an integrodifferential[18-21] theory of gravitation in agreement with the strong Mach principle. A discussion about strong and weak Mach's principle is contained, for example, in Sec. 1 of Ref. 18. The problem which arises in this connection is whether a tailless kernel of the type (2.2) can be considered as the Green's function of a linear second order differential equation. In the case of a positive answer there would be no difference in principle between the weak formulations of Mach's principle and the strong ones based on tailless kernels of the type (2.2). Indeed one could reduce any integral or integrodifferential theory based on such kernels to a purely differential theory with suitable boundary conditions. The main theorem proves that this possibility is *not* allowed, at least if one limits oneself to linear second order differential equations.

Note added in proof: In the proof of Theorem 2 two assumptions, unfortunately *not* true in general, have been used, namely the hypotheses that the limits as $\bar{x} \to x$ of the functions $P^{*\alpha\beta} d_{;\alpha} d_{;\beta}$ and $R^{\alpha\beta} d_{;\alpha} d_{;\beta}$ and of all their partial derivatives are finite (hereafter referred to as assumptions A and B, respectively). For the necessary condition (3.3), however, assumption A is not needed since Eq. (3.16) may be easily proved also with $t = 0$. On the contrary, in the lack of assumption B, the general proof of Eq. (3.4) requires conditions on the conformal factor $K(x)$. Consequently, Theorem 1 becomes: A necessary condition for the differential operator O to satisfy Huygens' principle is that the Riemannian manifold (V^4, g) be conformal to an Einstein space.

ACKNOWLEDGMENTS

The author is very grateful to Professor A. Lichnerowicz for interesting discussions, suggestions, and critical remarks. He is also indebted to Professor Y. Choquet-Bruhat for critically reading the preliminary version of this work and to Dr. C. Moreno for friendly and interesting discussions. The kind hospitality of Professor A. Lichnerowicz in Paris is also acknowledged.

*This work was part of the author's activity within the "Gruppo Nazionale di Fisica Matematica" of the C. N. R., the Italian National Research Council.

[1] J. Hadamard, *Lectures on Cauchy's Problem in Linear Partial Differential Equations* (Yale U. P., New Haven, Connecticut, 1923).
[2] S. Sobolev, Mat. Sb. (N.S.) 1, 39 (1936).
[3] L. Schwartz, *Théorie des distributions* (Hermann, Paris, 1966).
[4] A. Lichnerowicz, C. R. Acad. Sci. Paris **249**, 1329 (1959).
[5] A. Lichnerowicz, C. R. Acad. Sci. Paris **249**, 2287 (1959).
[6] A. Lichnerowicz, C. R. Acad. Sci. Paris **250**, 3122 (1960).
[7] A. Lichnerowicz, Publ. Math. Inst. Hautes Etudes Paris **10**, 1 (1961).
[8] B. S. deWitt and R. W. Brehme, Ann. Phys. (N.Y.) **9**, 220 (1960).
[9] D. Robaschik, Acta Phys. Polon. **24**, 299 (1963).
[10] Y. Bruhat, Ann. Mat. Ser. 4, **64**, 191 (1964).
[11] H. P. Kunzle, Proc. Cambridge Philos. Soc. **64**, 779 (1968).
[12] F. G. Friedlander, "The wave equation on a curved space–time" (Cambridge Monographs on Mathematical Physics **2**, 1975).
[13] R. G. McLenaghan, Ann. Inst. Henri Poincaré, **XX**, 153 (1974).
[14] B. L. Altshuler, Zh. Eksp. Teor. Fiz. **51**, 1143 (1966); **55**, 1311 (1968) [Sov. Phys.—JETP **24**, 766 (1967); **28**, 687 (1969)]; D. Lynden-Bell, Mon. Not. R. Astron. Soc. **135**, 413 (1967); D. W. Sciama, P. C. Waylen, and R. C. Gilman, Phys. Rev. **187**, 1762 (1969).
[15] I. M. Guelfand and G. E. Chilov, *Les Distributions* (Dunod, Paris, 1962).
[16] J. M. Bird and W. G. Dixon, Ann. Phys. (N.Y.) **94**, 320 (1975), Appendix A.
[17] B. Bertotti, "On Mach's Principle," unpublished (1965).
[18] R. Goldoni, Gen. Rel. Grav. J. 7, 731 (1976).
[19] R. Goldoni, Gen. Rel. Grav. J. 7, 743 (1976).
[20] R. Goldoni, Istituto Matematico "L. Tonelli" preprint 77/3, to be published in Meccanica.
[21] R. Goldoni, Istituto Matematico "L. Tonelli" preprint 77/6, submitted to Rep. Math. Phys. C.

Perturbation solution of the Percus–Yevick equation for the square-mound potential

A. Fuliński and C. Jędrzejek

Department of Theoretical Chemistry, Jagellonian University, Cracow, Poland
(Received 20 February 1976; revised manuscript received 23 February 1977)

The perturbation solution of the Percus–Yevick equation, based on the known solution for hard-sphere interactions, is found for the square-mound potential. The correlation functions are expanded in powers of the parameter α, related to the height of the mound, i.e., describing the deviation of the interactions from the hard-sphere ones. An algorithm for the calculation of subsequent coefficients is given. Numerical calculations show, however, that the series converges slowly and thus a few terms approximate the whole series with sufficient accuracy for small (but still finite) values of α, i.e., for almost hard-sphere interactions, only. For mounds high enough the system behaves very similarly to the hard-sphere system, but it quickly loses its "hard" characteristics as the mound decreases. This result, in the light of the success of the recent perturbation theory of liquids, seems to suggest that, whereas fairly significant changes of the potential in the outside of the hard core may be treated as small perturbations, even a small change inside the hard core very strongly perturbs the properties of the system.

I. INTRODUCTION

The importance of integral equations in the theory of liquids results, at least partially, from the existence of a few analytic solutions to these equations. Moreover, one of these equations, the Percus–Yevick (PY) approximation,[1] has proved to be fairly accurate for hard-sphere systems. Hence the solutions of the PY equations obtained by Wertheim[2] and Thiele[3] for the one-component system of hard spheres, and by Lebowitz[4] for the mixture of hard spheres with additive diameters, are commonly used for the determination of the properties of the reference system in the perturbation theories of liquids.[5,6] These latter owe their success mainly to the fact of very quick convergence of the perturbation series when the repulsive part of the interparticle potential does not differ significantly from the hard-sphere potential, even if the attractive tail of the potential playing the role of the perturbation, cannot be treated as very small. In this paper we want to point out that the perturbation series behaves quite differently in the case of the short-ranged potential of the repulsive barrier of finite height. We have chosen this potential because it is possible to obtain for it the solution of the PY equation in the form of perturbation expansion of the relatively simple form.

Standard perturbation expansions[5] require for the determination of higher-order terms a knowledge of the n-particle ($n \geq 2$) distribution functions of the reference system. The idea of the use of integral equations for the calculation of the perturbation corrections was proposed by Lado,[7] who constructed an expansion in which it is not the bridge diagrams which are being neglected, but rather changes in them caused by the perturbing potential. Madden and Fitts[8] found the integral equations for the calculation of the first and second-order perturbation corrections to the radial distribution function $g(r)$. Recently, Kohler, Perram, and White[9] constructed a new method for the numerical solution of the PY equation, based on Baxter's formalism,[10] and used it for the determination of $g(r)$ for the repulsive part of the Lennard–Jones potential.

Baxter's formulation of the Ornstein–Zernike equation, used in Ref. 9, is also convenient for our purpose. It can be outlined as follows: Let $c(r)$ be the direct correlation function defined by the Ornstein–Zernike equation,

$$h(r) = c(r) + n \int d\mathbf{t}\, h(|\mathbf{r}-\mathbf{t}|)\, c(t)\,, \tag{1}$$

where n is the number density, and $h(r) = g(r) - 1$. The PY approximation,

$$\alpha(r)c(r) = [\alpha(r) - 1]g(r), \quad \alpha(r) = \exp[-u(r)/kT]\,, \tag{2}$$

implies that, when the two-particle potential $u(r)$ vanishes, $c(r)$ vanishes, too. Consider now the potentials $u(r)$ of finite range d, i.e., $u(r) \equiv 0$ for $r > d$. Baxter introduces a new function $q(r)$, by writing Eq. (1) in the form

$$rc(r) = -q'(r) + 2\pi n \int_r^d dt\, q'(t) q(t-r), \quad r \in [0, d]\,, \tag{3}$$

$$rh(r) = -q'(r) + 2\pi n \int_0^d dt\, (r-t) h(|r-t|) q(t), \quad r \geq 0\,, \tag{4}$$

where $q'(r)$ is the derivative of $g(r)$. The function $q(r)$ also vanishes identically for $r > d$, and is continuous. Combining Eqs. (1)–(4) we get one nonlinear integrodifferential equation for Baxter's function $q(r)$,

$$\begin{aligned} -r[\alpha(r) - 1] - q'(r) + 2\pi n \alpha(r) \int_r^d dt\, q'(t) q(t-r) \\ = -2\pi n [\alpha(r) - 1] \int_0^d dt\, (r-t) q(t) \\ + 2\pi n \int_0^d dt\, \mathrm{sign}(r-t) q(t) \alpha(|r-t|) \\ \times \{-q'(|r-t|) + 2\pi n \int_{|r-t|}^d dz\, q'(z) q(z - |r-t|)\}\,. \end{aligned} \tag{5}$$

II. SOLUTION OF THE PY EQUATION FOR SQUARE-MOUND POTENTIAL

The square-mound potential[11] can be defined by

$$\alpha(r) = \begin{cases} \alpha & \text{for } r < d\,, \\ 1 & \text{for } r > d\,. \end{cases} \tag{6}$$

The value $\alpha = 1$ (vanishing mound) describes the ideal gas, whereas the hard-sphere potential, for which $\alpha = 0$,

 Copyright © 1977 American Institute of Physics

corresponds to the limit of an infinitely high mound. We look for the solution of Eq. (5) in the form of the power series in α,

$$q(r;\alpha)=\sum_{i=0}^{\infty}\alpha^i q_i(r). \tag{7}$$

As $\alpha=0$ means the hard-sphere interaction, the series (7) forms the perturbation solution of the PY equation, the unperturbed system being that of hard spheres of the same diameter.[12]

The function $q(r;\alpha)$ is differentiable with respect to r. We look for such a solution for the functions $q_i(r)$ which would also be differentiable. Assuming that we can change the order of differentiation and summation in Eq. (7) we obtain the functions $q_i(r)$ as polynomials in accordance with the assumed property of $q(r;\alpha)$.

Substituting the series (7) into Eq. (5) and equating powers of α, we obtain the system of equations for the functions $q_i(r)$,

$$\begin{aligned}q_i'(r)=&-\delta_{1i}r+2\pi n\int_r^d dt\sum_{j=0}^{i-1}q_j'(t)q_{i-1-j}(t-r)\\&+2\pi n\int_0^d dt\,(r-t)[q_{i-1}(t)-q_i(t)]\\&-2\pi n\int_0^d dt\,\mathrm{sign}(r-t)\sum_{j=0}^{i-1}q_j(t)\\&\times\{-q_{i-1-j}'(|r-t|)+2\pi n\int_{|r-t|}^d dz\\&\times\sum_{l=0}^{i-1-j}q_l'(z)q_{i-1-j-l}(z-|r-t|)\}.\end{aligned} \tag{8}$$

The first of these equations immediately gives the hard-sphere function $q_0(r)$,

$$q_0(r)=\frac{1+2\eta}{2(1-\eta)^2}(r^2-d^2)-\frac{1.5\eta d}{(1-\eta)^2}(r-d), \tag{9}$$

where

$$\eta=\pi nd^3/6 \tag{10}$$

is the so-called packing density. Equation (8) together with Eq. (9) implies (by recurrence) that the function $q_i(r)$ is the polynomial of $(6i+2)$th degree in r,

$$q_i(r)=\sum_{l=0}^{6i+2}A_{il}r^l,\quad q_i'(r)=\sum_{l=0}^{6i+1}a_{il}r^l. \tag{11}$$

Equation (8) may be written in the form

$$q_i'(r)+2\pi n\int_0^d dt\,(r-t)q_i(t)=\sum_{l=0}^{6i-1}p_{il}r^l,\quad i=1,2,\ldots, \tag{12}$$

where the coefficients p_{il} are determined by integrating, according to Eq. (8), the combinations of functions $q_j(r)$ with $j<i$. Hence Eq. (8) may be solved for every i by recurrence, $q_0(r)$ is given by Eq. (9), coefficients p_{il} are given by coefficients A_{jl} $(j<i)$ determined in earlier steps, and the form of Eq. (12) implies $lA_{i,l+1}=a_{il}=p_{il}$ for $l=2,\ldots,6i+1$. The remaining coefficients, A_{i0}, $A_{i1}=a_{i0}$, $A_{i2}=a_{i1}/2$, are determined from Eq. (12) (for $l=0,1$) and from the condition $q(d)=0$, which must be fulfilled for every α. Hence, after some manipulations,

$$A_{i0}=-(A_{i1}+d^2A_{i2}/4+\sum_{l=3}^{6i+2}p_{i,l-1}d^l/l), \tag{13}$$

$$(2-4\pi nd^3/3)A_{i2}-\pi nd^2A_{i1}=p_{i1}+2\pi n\sum_{l=3}^{6i+1}p_{i,l-1}\frac{d^{l+1}}{l+1}, \tag{14}$$

$$\pi nd^4A_{i2}/2+(1+\pi nd^3/3)A_{i1}=p_{i0}-\pi n\sum_{l=2}^{6i+2}p_{i,l-1}\frac{d^{l+2}}{l+2}. \tag{15}$$

Let us note that for $\eta=1$ the principal determinant of the pair of equations (14), (15) is equal to zero. Hence the solution of the PY equation for square-mound potential behaves similarly to that for hard spheres, where the solution also diverges for $\eta=1$.

We may now calculate the remaining correlation functions. From Eq. (3) we have

$$c(r)=\sum_{i=0}^{\infty}\alpha^i c_i(r),\quad r\in[0,d], \tag{16}$$

with

$$rc_i(r)=-q_i(r)+2\pi n\sum_{j=0}^{i}\int_r^d dt\,q_j'(t)q_{i-j}(t-r). \tag{17}$$

The functions $c_i(r)$ are thus polynomials of the $(6i+3)$th degree, with coefficients determined—in terms of coefficients A_{jl}—directly from the relations (17) and (11).

The knowledge of $c(r)$ leads to the determination of $g(r)$. The Fourier transformation of Eqs. (1) and (16) gives

$$H(k)=\frac{C(k)}{1-nC(k)},\quad C(k)=\sum_{i=0}^{\infty}\alpha^i C_i(k), \tag{18}$$

and $H(k)$, $C(k)$, and $C_i(k)$ are the Fourier transforms of the functions $h(r)$, $c(r)$, and $c_i(r)$, respectively. Note that the functions $C_i(k)$ are well defined because $c_i(r)$ are (i) bounded and (ii) nonzero for $r\in[0,d]$ only. Expanding $H(k)$ as given by Eq. (18) in powers of α,

$$H(k)=\sum_{i=0}^{\infty}\alpha^i H_i(k), \tag{19}$$

we get the functions $H_i(k)$ as combinations of the products of functions $C_j(k)$ (with $j\leq i$) and finally

$$g(r)=1+\sum_{i=0}^{\infty}\alpha^i h_i(r), \tag{20}$$

where $h_i(r)$ are the inverse Fourier transforms of $H_i(k)$.

III. NUMERICAL RESULTS

To check the behavior of the perturbation solution obtained in the preceding section the numerical computations of the expansions (7), (16), and (20) of the distribution functions $c(r)$ and $g(r)$ have been performed up to the third order in the perturbation parameter α. The Mandel, Bearman, and Bearman[13] method (generalized for polynomials of arbitrary order) has been used for the inversion of the Fourier transforms $H_i(k)$, Eq. (19).

The numerical results show that the series of coefficients A_{il}, Eq. (11), is alternating for the initial terms of the series (7), the ratio $|A_{i+1,l}/A_{il}|$ being the greater, the higher the density of the system. The inequality $|A_{i+1,l}|<|A_{il}|$ is fulfilled for very low densities, $\eta<0.005$, only. The same behavior is found for the series $c_i(r)ih_i(r)$.

The PY approximation also enables us to check the accuracy of the solutions obtained by the present method. For this purpose, we introduce the function

$$y(r) = g(r)\exp[\beta u(r)], \tag{21}$$

continuous for all $r > 0$. Especially,

$$y(d_+) = y(d_-), \tag{22}$$

where

$$y(d_\pm) = \lim_{\epsilon \to 0} y(d \pm \epsilon).$$

From Eq. (2), $y(r)$ is also given by

$$y(r) = \frac{c(r)}{\exp[-\beta u(r)] - 1}. \tag{23}$$

The differences between the value of $y(r)$ calculated from Eq. (21) and Eq. (23), as well as between $y(d_+)$ and $y(d_-)$, may serve as the estimation of the errors resulting from the approximation of the whole series (7), (16), and (20) by the finite numbers of terms. It is found that these differences do not exceed 0.005 for $\alpha = 0.0075$, and 0.02 for $\alpha = 0.01$, for $nd^3 = 0.6$ (the estimated error of numerical computations is of the order of 0.0001). In Fig. 1 the square mound functions $c(r)$ and $g(r)$ for $nd^3 = 0.6$, $\alpha = 0.0075$ are compared with the hard-sphere ones for the same density. The value of $\alpha = 0.0075$, i.e., $\beta u(r) = 4.8$, may be recognized as the highest maximum value, α_m, for which the solution is still creditable, for the density $nd^3 = 0.6$, when the perturbation series is cut at α^3. For higher densities the value of α_m decreases: For example, for $nd^3 = 0.9$, $\alpha_m = 0.0015$, i.e., $\beta u(r) = 6.5$. For this value of α, $g(d_+) = 4.13$, which is to be compared with $g(d_+) = 4.42$ for hard spheres, $\alpha = 0$. For high densities the series (7), (16), and (20) converge very slowly, so that retaining a few more terms of higher order in α does not noticeably change the values of α_m.

Finally in Table I we illustrate the typical behavior of the α expansion by showing the several first coefficients b_{il} expressing the function $c_i(r)$, [Eq. (16)]

$$c_i(r) = \sum_{l=0}^{6i+3} b_{il} r^l, \tag{24}$$

for density $nd^3 = 0.764$ ($\eta = 0.4$). The character of the functions $q_i(r)$, $c_i(r)$, and $h_i(r)$ is similar within the whole physical range of density both for coefficients of the polynomials and values of the functions for particular r.

TABLE I. Coefficients b_{il} (listed only up to terms r^8) from Eq. (28) for density $nd^3 = 0.764$ ($\eta = 0.4$).

r^l	b_{0l}	b_{1l}	b_{2l}	b_{3l}	b_{4l}
r^0	−0.25E2	0.11E4	−0.73E5	0.56E7	−0.47E9
r^1	0.27E2	−0.12E4	0.77E5	−0.59E7	0.49E9
r^2	0.0	−0.24E3	0.21E5	−0.19E7	0.17E9
r^3	−0.50E1	0.31E3	−0.22E5	0.18E7	−0.16E9
r^4	0.0	0.15E3	−0.15E5	0.13E7	−0.12E9
r^5	0.0	−0.75E2	0.72E4	−0.64E6	0.59E8
r^6	0.0	−0.57E1	0.10E4	−0.11E6	0.12E8
r^7	0.0	0.61E1	−0.70E3	0.68E5	−0.65E7
r^8	0.0	0.0	−0.90E2	0.13E5	−0.15E7

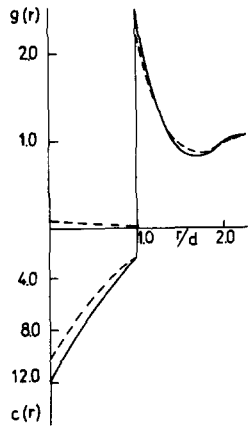


FIG. 1. Comparison of the hard-sphere ($\alpha = 0$, full line) and square mound ($\alpha = 0.0075$, dashed line) correlation functions $c(r)$ and $g(r)$ for $nd^3 = 0.6$.

IV. FINAL REMARKS

In this paper the solution of the PY equation for the square-mound potential has been given in the form of the perturbation series, the unperturbed system being that of hard spheres. The computation of the coefficients A_{il} for a given density, is sufficient for the determination of the solution for any value of α, the parameter α describing the deviation of the system from the hard-sphere one. This method is thus, in principle, both simpler and more general than the usual numerical iterative solutions of the PY equation,[14] which require carrying out the computation for a definite value of α.

In addition, the present method allows the detailed investigation of the behavior of the perturbation series, which is impossible in the numerical solutions of the PY equation by iteration methods. Such an investigation, reported in Sec. III, has shown that the series (20) converges very slowly, except for small values of α. Hence the practical use of our method is limited, for a given density, to the values of α smaller than some limiting value α_m, for which the perturbation series is still rapidly enough convergent. To this limiting value corresponds the respective limiting value u_m/kT of the potential $u(r)/kT$, for which the behavior of the square-mound system still reminds one of the hard sphere behavior. The very slow convergence of the perturbation series for $\alpha > \alpha_m$ seems to suggest that the system loses its "hard" characteristics for $u < u_m$.

The square-mound interactions are not very realistic ones. However, the method of solution of the PY equation for these interactions, presented in this paper, enabled us to show clearly that, whereas it is well known that quite serious changes in the longer-ranged attractive tail,[5] as well as in the outside region of the repulsive core,[9] may be treated as small perturbations, even relatively small changes of the interaction potential [or rather of the Mayer function $f(r) = \alpha(r) - 1$, associated with it—cf. Ref. 12] far inside the repulsive core strongly perturb the properties of the system.[15]

ACKNOWLEDGMENTS

The authors would like to thank Professor J. Siciak and Dr. J. Jurkiewicz and Dr. A. Korytowski for stimulating discussions.

[1]J. K. Percus and G. J. Yevick, Phys. Rev. **110**, 1 (1958).
[2]M. S. Wertheim, Phys. Rev. Lett. **10**, 321 (1963).
[3]E. Thiele, J. Chem. Phys. **39**, 474 (1963).
[4]J. L. Lebowitz, Phys. Rev. **133**, 895A (1964).
[5]D. Henderson and J. A. Barker, in *Physical Chemistry, An Advanced Treatise* (Academic, New York, 1971), Vol. VIIIA, Chapter 6.
[6]H. C. Andersen and D. Chandler, J. Chem. Phys. **57**, 1918 (1972).
[7]F. Lado, Phys. Rev. A **8**, 2548 (1973); J. Chem. Phys. **59**, 4830 (1973); **60**, 1686 (1974).
[8]W. G. Madden and D. D. Fitts, Mol. Phys. **28**, 1095 (1974); J. Chem. Phys. **61**, 4395 (1974).
[9]F. Kohler, J. W. Perram, and L. R. White, Chem. Phys. Lett. **32**, 42 (1975). Quite recently, the perturbation solution to the first order of the PY equation for the square-well potential has been published by D. Henderson and M. Chen, J. Math. Phys. **16**, 2042 (1975).
[10]R. J. Baxter, Aust. J. Phys. **21**, 563 (1968).
[11]This potential has been considered earlier in the literature. The first few virial coefficients were determined by others, S. Katsura and Y. Abe, J. Chem. Phys. **39**, 2068 (1963).
[12]Note that it is the function $\alpha(r)$, or rather the Mayer function (renormalized potential) $f(r)=\alpha(r)-1$, and not the potential $u(r)$ itself, which appears in all the formulas of statistical thermodynamics. The difference between the system of hard spheres and the system of spheres with finite square-mound potential with, e.g., $u(r)=5kT$, $r<d$, seems to be indefinitely great in terms of the corresponding potentials. However, the differences between the Mayer functions of both systems are small, of the order of 1%, and we expect that the corresponding differences of the thermodynamic properties will also be small. It is understandable because $u(r)\cong 5kT$ means that about 1% of interparticle collisions has enough energy to overcome the potential barrier, i.e., for about 99% of collisions, particles behave like hard spheres.
[13]F. Mandel, R. J. Bearman, and M. Y. Bearman, J. Chem. Phys. **52**, 3315 (1970).
[14]A. A. Broyles, J. Chem. Phys. **33**, 456 (1960); F. Lado, J. Chem. Phys. **47**, 4828 (1967); R. O. Watts, J. Chem. Phys. **48**, 50 (1968).
[15]J. S. Rowlinson, Mol. Phys. **7**, 349 (1964); **8**, 107 (1964).

Functional equations for extended hadrons

Z. Haba

Institute of Theoretical Physics, University of Wrocław, Wrocław, Poland
(Received 20 September 1976)

We consider the problem of solution of some functional equations occurring in the theory of extended hadrons. By means of stochastic methods solutions of these equations are obtained in the form of a contractive (Markov) semigroup in Hilbert space. Analytic continuation to a unitary group implementing time evolution is performed. The problem of unitary implementability of Lorentz and gauge symmetries, essential for physical interpretation, remains unsolved.

I. INTRODUCTION

The theory of extended hadrons, extensively studied recently,[1-3] is based on an extended position operator $X(\xi)$ depending on internal parameters ξ. Such an approach raises some difficult problems, which follow from the requirements of Lorentz and reparametrization (gauge) invariance of the theory.

This paper is devoted to mathematically rigorous investigation of functional equations occurring in the theory of extended hadrons.[1,3] In general such equations are highly complicated,[2] and it is hopeless to solve them. So, we consider only simplified equations, which in the above-mentioned models, hold in a special noncovariant gauge.[1] In such a case we are able to solve these equations by means of stochastic methods.[4-6] We apply the theory of Markov processes with values in infinite-dimensional spaces.[5,6] This allows us to solve the second order functional equation

$$-\frac{d}{dt}\psi_t(X)=H_0\left(X,\frac{\delta}{\delta X}\right)\psi_t(X) \tag{I.1}$$

where X belongs to the Schwartz space $\mathcal{S}'$ of tempered distributions. If the operator H_0 is positively definite in Hilbert space L^2_μ with μ as the Gaussian measure on $\mathcal{S}'$, then the solutions of Eq. (I.1) have analytic continuation in time to solutions of the Schrödinger equation. Then, we show how the functional equation with interaction $H=H_0+V$ can be solved by means of the Feynman-Kac method.[4,7]

Comparing with the Kaku and Kikkawa paper,[1] we can see that the transition function of the Markov process determined by Eq. (I.1) coincides with the Green's function in Ref. 1 describing a propagation of free string.

We do not consider here the problem of Lorentz and gauge invariance of the theory. These symmetries should be unitarily implementable, if this theory is to describe extended particles. There are serious difficulties on this way. We discuss briefly this problem in the last section.

II. WIENER PROCESS WITH VALUES IN $\mathcal{S}'$

It is known[4] that the transition function $p(t_0, \mathbf{x}, t, \Gamma)$, $\Gamma\subset R^n$, of the Wiener stochastic process is a solution of the equation

$$\frac{d}{dt}\varphi_t(\mathbf{x})=\sum_{m=1}^{n}\frac{\partial^2}{\partial x_m^2}\varphi_t(\mathbf{x}),\quad \mathbf{x}\in R^n$$

with initial condition

$$\varphi_{t_0}(\mathbf{x})=\chi_\Gamma(\mathbf{x})=\begin{cases}1, & \mathbf{x}\in\Gamma,\\ 0, & \mathbf{x}\notin\Gamma.\end{cases} \tag{II.1}$$

We will show that the transition function of Wiener process with values in $\mathcal{S}'$ is a solution of an analogous functional equation. The Wiener process W_t with values in $\mathcal{S}'$ is a stochastic process with Gaussian distribution of differences $W_t - W_s\in\mathcal{S}'$ $(t>s)$ with mean zero and covariance

$$E[(W_t(\xi)-W_s(\xi))(W_{t'}(\xi')-W_{s'}(\xi'))]$$
$$=\min(t-s,t'-s')\delta(\xi-\xi'),$$
$$t>s,\quad t'>s'. \tag{II.2}$$

So, the candidate for the transition function $\rho_W(t,Y,t',\Gamma)$ is the Gaussian measure with mean Y and covariance $t'-t$

$$\rho_W(t,Y,t',\Gamma)=\int_{\Gamma\subset\mathcal{S}'}\mu_{t'-t}(dY'-Y). \tag{II.3}$$

Let us introduce the *Banach space* $\mathcal{B}$ of functions F: $\mathcal{S}'\to C$ continuous under the norm $||F||=\sup_{Y\in\mathcal{S}'}|F(Y)|$ in the weak topology of $\mathcal{S}'$. Then we can prove

Theorem II.1: $\rho_W(t,Y,t',\Gamma)$ is a transition function of the Markovian contractive semigroup T_t^W in $\mathcal{B}$ defined by

$$T_t^WF(Y)=\int_{\mathcal{S}'}F(Y')\mu_t(dY'-Y),\quad t>0,\ F\in\mathcal{B}, \tag{II.4}$$

i.e.,

$$\text{(i) } T_s^WT_t^W=T_{s+t}^W,$$
$$\text{(ii) } T_t^W1=1, \tag{II.5}$$
$$\text{(iii) } \lim_{t\to+0}||T_t^WF-F||=0.$$

Proof: T_t^W is a contraction semigroup in $\mathcal{B}$ because

$$\left|\int_{\mathcal{S}'}F(Y')\mu_t(dY'-Y)\right|$$
$$\le\int\sup|F(Y')|\mu_t(dY'-Y)=\sup|F(Y')|.$$

Property (i) is equivalent to

$$\int_\Gamma\mu_{t+s}(dY'-Y)=\int_\Gamma\int_{\mathcal{S}'}\mu_t(dY'-Z)\mu_s(dZ-Y). \tag{II.6}$$

This equation can be checked by means of the Fourier transform

$$\int_{\mathcal{S}'}\exp[i(Y',h)]\mu_{t+s}(dY'-Y)$$
$$=\exp[-(t+s)\int h^2(\xi)d\xi]\exp[i(Y,h)].$$

 Copyright © 1977 American Institute of Physics

and

$$\int_{S'} \int_{S'} \exp[i(Y',h)]\,\mu_t(dY'-Z)\mu_s(dZ-Y)$$
$$= \int_{S'} \exp[-t\int h^2(\xi)d\xi]\exp[i(Z,h)]\,\mu_s(dZ-Y)$$
$$= \exp[-t\int h^2(\xi)d\xi]\exp[-s\int h^2(\xi)d\xi]\exp[i(Y,h)]\ .$$

So, Eq. (II.6) follows. The second equality in the formulas (II.5) is trivial, because the Gaussian measure is normalized $\int_{S'}\mu_t(dY)=1$. Let us consider now (iii):

$$|T_t^W F(Y)-F(Y)| = \left|\int_{S'}(F(Y')-F(Y))\mu_t(dY'-Y)\right|$$
$$\leq \int_{S_Y(\delta)}|F(Y')-F(Y)|\,\mu_t(dY'-Y)$$
$$+\int_{S'-S_Y(\delta)}|F(Y')-F(Y)|\,\mu_t(dY'-Y)\ , \tag{II.7}$$

where the ball $S_Y(\delta)$ [8] is defined by

$$S_Y(\delta)=\{Y'\in S';\ |(Y-Y',h)|\leq\delta \text{ for } h\in S \text{ with } \|h\|_n\leq 1\};$$

here $\|\ \|_n$ are seminorms defining the topology of S. From the continuity of F we get that for each ϵ there exists δ such that $Y'\in S_Y(\delta)$ implies $|F(Y')-F(Y)|<\epsilon$. The second term in Eq. (II.7) is smaller than

$$2\|F\|\int_{S'-S_Y(\delta)}\mu_t(dY'-Y)=2\|F\|\int_{S'-S_Y(\delta/t^{1/2})}\mu_1(dY'-Y)\ .$$

Now, if $t\to 0$ it goes to zero, because the measure of a set lying outside the ball $S_Y(R)$ goes to zero if the radius R goes to infinity.[8] This completes the proof.

Equations (II.5) are equivalent to the following properties of the transition function (II.3):

1. $\rho_W(t,Y,t',\Gamma)$ is a measure on a σ-algebra of sets $\Gamma\subset S'$.

2. $\int_{S'}\rho_W(\tau,Z,t',\Gamma)\rho_W(t,Y,\tau,dZ)=\rho_W(t,Y,t',\Gamma)$ for arbitrary $t\leq\tau\leq t'$. This is the Chapman–Kolmogorov equation.

3. $\rho_W(t,Y,t',S')=1$

4. $\rho_W(t,Y,t,\Gamma)=\chi_\Gamma(Y)$ (II.8)

The transition function ρ_W fulfills a functional equation, which is a generalization of Eq. (II.1). First, a precise definition of the infinite-dimensional Laplace operator is needed.[5,6] We define the functional (Frechet) derivative $F'(X)$ in the point $X\in S'$ as a linear functional belonging to S' and fulfilling the equation (if this limit exists in the norm topology of $\mathcal{B}$)

$$\lim_{s\to 0}(1/s)[F(X+sh)-F(X)]=(F'(X),h)\ . \tag{II.9}$$

Second order derivative is defined as an *operator* from S to S' obtained by differentiation of the right-hand side of Eq. (II.9)

$$\lim_{s\to 0}(1/s)[(F'(X+sg),h)-(F'(X),h)]=(F''(X)g,h); \tag{II.10}$$

here the limit is taken in the S' weak topology. The only way to obtain a scalar from an operator $F''(X)$: $S\to S'$ is to take its trace in the Hilbert space L^2 of square integrable functions [if the range of $F''(X)$ is contained in L^2]. We can prove now the following:

Theorem II.2: For a dense in $\mathcal{B}$ set of functions

$$F(Y)=\sum_{j=1}^{k} a_j\exp[i(Y,h_j)],\quad h_j\in S:$$

(i) The function $\varphi_t(Y)=T_t^W F(Y)$ is differentiable over $t>0$.

(ii) It is twice Frechet differentiable and the second Frechet derivative $\varphi_t''(Y)$ is a trace class operator in L^2.

(iii) $\varphi_t(Y)$ fulfills the functional equation

$$\frac{-d}{dt}\varphi_t(Y)=D^2\varphi_t(Y)$$
$$\text{where } D^2\varphi_t(Y)\overset{\text{def}}{=}-\mathrm{Tr}\varphi_t''(Y) \tag{II.11}$$

with initial condition $\varphi_0(Y)=F(Y)$.

Proof: It is sufficient to restrict ourselves to $F(Y)=\exp[i(Y,h)]$. Then

$$\varphi_t(Y)=T_t^W F(Y)=\int F(Y+Y')\mu_t(dY')$$
$$=\exp[i(Y,h)]\exp[-t(h,h)]\ . \tag{II.12}$$

So,

$$\frac{d}{dt}T_t^W F(Y)=-(h,h)T_t^W F(Y)\ ,$$
$$(F'(Y),f)=i(f,h)\exp[i(Y,h)]\exp[-t(h,h)]\ , \tag{II.13}$$
$$(F''(Y)g,f)=-(f,h)\,(g,h)\exp[i(Y,h)]\exp[-t(h,h)]\ . \tag{II.14}$$

The bilinear form (II.14) is continuous in g and f in the L^2 norm $|f|_L^2=\int f^2(\xi)d\xi$, because

$$|(F''(Y)g,f)|\leq|h|_L^2|f|_L|g|_L$$

and defines a bounded operator on L^2. Let us compute its trace:

$$\mathrm{Tr}F''(Y)=\sum_i(F''(Y)e_i,e_i)$$
$$=\sum_i -(h,e_i)(e_i,h)\cdot\exp[i(Y,h)]\exp[-t(h,h)]$$
$$=-(h,h)\exp[i(Y,h)]\exp[-t(h,h)]\ . \tag{II.15}$$

Comparing (II.15) and (II.13), we obtain that the equation $(d/dt)\varphi_t(Y)=\mathrm{Tr}\varphi_t''(Y)$ is fulfilled.

Remark: T_t^W is a contraction semigroup in $\mathcal{B}$, so it can be extended from a dense set on the whole space $\mathcal{B}$. However, $T_t^W F$ will not fulfill Eq. (II.11) if F does not belong to the domain of the operator D^2, i.e., if the second order Frechet derivative does not exist or if it is not a trace class operator. Necessary and sufficient conditions for existence of the trace were given by Gross.[5]

III. MARKOVIAN CONTRACTIVE SEMIGROUPS IN HILBERT SPACE

In physical applications we should find solutions of functional equations in *Hilbert* space. The Banach space in Sec. II (treated as a linear vector space) may be considered as a dense subspace of the Hilbert space L^2_μ of square integrable functions with regard to a Gaussian

measure μ on $\mathcal{S}'$. A Markovian contractive semigroup defined in $\mathcal{B}$ need not be a contraction on L^2_μ (in the L^2_μ norm $\|\ \|_\mu$), e.g., the semigroup defined by Eq. (II.4) is not a contraction in L^2_μ. A sufficient condition is given by the following:

Theorem III.1 (*Sin on*[9]): Markovian semigroup fulfilling [in addition to Eqs. (II.5)] the condition

$$T_t^\dagger 1 = 1 \tag{III.1}$$

is a contraction in L^2_μ, i.e.,

$$\|T_t F\|_\mu \leq \|F\|_\mu \ . \tag{III.2}$$

The semigroup (II.4) does not fulfill this condition. This can be shown as follows:

$$\frac{d}{dt}((T_t^W)^\dagger 1, F) = \frac{d}{dt}(1, T_t^W F) = (1, \mathcal{D}^{2\dagger} T_t^W F) = (\mathcal{D}^{2\dagger} 1, T_t^W F) \ , \tag{III.3}$$

where the operator $\mathcal{D}^2$ was defined by Eq. (II.11). The Hermitian conjugated operator $\mathcal{D}^{2\dagger}$ can be computed on the dense in L^2_μ set of functions depending on finite number n of variables $x_j = (X, h_j)$ (here the $h_j \in \mathcal{S}$ form a complete orthonormal set in L^2). Let us still fix the Gaussian measure μ by its covariance

$$E[X(f)X(g)] = \int X(f)X(g)\mu(dX) = (f, B^{-1}g) \ , \tag{III.4}$$

where B is a positive definite operator in L^2. Now, $\mathcal{D}^{2\dagger}$ is given by the formula

$$\begin{aligned}
&(\mathcal{D}^{2\dagger} F_1, F_2) \\
&= (F_1, \mathcal{D}^2 F_2) \\
&= -\int \mu(dX)\overline{F_1(X)}\,\mathrm{Tr}F_2''(X) \\
&= -\int \mu(dX)\overline{F_1(X)}\sum_{k=1}^{n}(\mathbf{h}_k, F_2''(X)h_k) \\
&= (4\pi)^{-n/2}(\det B)^{1/2}\int \overline{F}_1(x_1, \ldots, x_n)\Big(-\sum_{k=1}^{n}\frac{\partial^2}{\partial x_k^2} \\
&\quad \times F_2(x_1, \ldots, x_n)\Big)\exp\Big[-\frac{1}{4}\sum_{i,j=1}^{n} b_{ij}x_i x_j\Big]dx_1 \cdots dx_n \\
&= \int\Big\{\Big(\sum_{k=1}^{n} -\frac{\partial^2}{\partial x_k^2} + \sum_{k,i=1}^{n} b_{ik}x_i\frac{\partial}{\partial x_k} + \tfrac{1}{2}\sum_{i=1}^{n} b_{ii}\Big) \\
&\quad \times \overline{F}_1(x_1, \ldots, x_n)\Big\} F_2(x_1, \ldots, x_n)(4\pi)^{-n/2}(\det B)^{1/2} \\
&\quad \times \exp\Big[-\frac{1}{4}\sum_{i,j=1}^{n} b_{ij}x_i x_j\Big]dx_1 \cdots dx_n ;
\end{aligned} \tag{III.5}$$

here $b_{ij} = (h_i, Bh_j)$. From Eq. (III.5) it follows that $\mathcal{D}^{2\dagger} 1 \neq 0$, so from (III.3) $(T_t^W)^\dagger 1 \neq 1$. However, if we take the Hermitian part Δ of $\mathcal{D}^2$ (this is the Umemura Laplacian[10])

$$\Delta F(X) = -\mathrm{Tr}F''(X) + \tfrac{1}{2}(BX, F'(X)) \ , \tag{III.6}$$

then $\Delta 1 = \Delta^\dagger 1 = 0$. In such a case the condition (III.1) will be fulfilled for a semigroup $T_t F(X) = \varphi_t(X)$, solving the equation

$$-\frac{d}{dt}\varphi_t(X) = \Delta\varphi_t(X) \ . \tag{III.7}$$

The solution of Eq. (III.7) can be constructed by means of a stochastic process, which is a solution of a stochastic equation[4,6] for random variable X_t

$$dX_t = -\tfrac{1}{2}BX_t dt + dW_t \ ; \tag{III.8}$$

here W_t is the Wiener process. Equation (III.8) can be solved in elementary way. The result is

$$X_t = e^{-B(t-t_0)/2}X + \int_{t_0}^{t} e^{-B(t-s)/2}dW_s \ . \tag{III.9}$$

X_t given by Eq. (III.9) is a Gaussian process (it is an infinite dimensional counterpart of the Ornstein–Uhlenbeck velocity process[11]), because differences of Wiener process have Gaussian distribution. The mean and covariance can be computed from the formula (III.9)

$$m = E[X_t] = e^{-B(t-t_0)/2}X \ , \tag{III.10}$$

$$\begin{aligned}
\sigma^2 &= E[(X_t(\xi) - m)(X_t(\xi') - m)] \\
&= E\Big[\int_{t_0}^{t} e^{-B(t-s)/2}dW_s \int_{t_0}^{t} e^{-B(t-s')/2}dW_{s'}\Big] \\
&= B^{-1}(1 - e^{-B(t-t_0)})\delta(\xi - \xi') \ .
\end{aligned} \tag{III.11}$$

The transition function, being the transition amplitude to find X_t in Γ if it was in X at time t_0, is given for Gaussian process by the Gaussian measure $\mu_\sigma^{t-t_0, X}$ with mean m [(III.10)] and covariance σ^2 [(III.11)]

$$P(t_0, X, t, \Gamma) = \int_\Gamma \mu_\sigma^{t-t_0, X}(dX') \ . \tag{III.12}$$

It follows already from the proof of Theorem II.1 for Wiener process that P has the properties (II.8) of a transition function. We can further show the following:

Theorem III.2:

$$\varphi_t(X) = T_t F(X) = \int F(X')\mu_\sigma^{t-t_0, X}(dX'), \quad t - t_0 > 0 \ , \tag{III.13}$$

defines a Markovian semigroup in $\mathcal{B}$. The function $\varphi_t(X)$ fulfills Eq. (III.7) with initial condition $\varphi_{t_0}(X) = F(X)$ for $F(X)$ being linear combination of $\exp[i(X, h_j)]$.

Proof: A transition function always defines a Markovian semigroup.[4] It remains to show that Eq. (III.7) is fulfilled. The proof goes similarly as in Theorem II.2. We take $F(X) = \exp[i(X, h)]$; then

$$\begin{aligned}
\varphi_t(X) = T_t F(X) &= \exp[i(X, e^{-B(t-t_0)/2}h] \\
&\quad \times \exp[-(h, B^{-1}(1 - e^{-B(t-t_0)})h)]
\end{aligned}$$

and

$$\begin{aligned}
\frac{d}{dt}T_t F(X) &= \{-(i/2)(BX, e^{-B(t-t_0)/2}h) - (h, e^{-B(t-t_0)}h)\} \\
&\quad \times \exp[i(X, e^{-B(t-t_0)/2}h)] \\
&\quad \times \exp[-(h, B^{-1}(1 - e^{-B(t-t_0)})h] \ .
\end{aligned}$$

Computation of $\mathrm{Tr}\varphi_t''(X)$ gives

$$\begin{aligned}
\mathrm{Tr}\varphi_t''(X) &= -(h, e^{-B(t-t_0)}h)\exp[i(X, e^{-B(t-t_0)/2}h)] \\
&\quad \times \exp[-(h, B^{-1}(1 - e^{-B(t-t_0)})h] \ .
\end{aligned}$$

Finally

$$\begin{aligned}
(BX, \varphi_t'(X)) &= i(BX, e^{-B(t-t_0)/2}h)\exp[i(X, e^{-B(t-t_0)/2}h)] \\
&\quad \times \exp[-(h, B^{-1}(1 - e^{-B(t-t_0)})h)] \ .
\end{aligned}$$

So, the equation

$$\frac{d}{dt}\varphi_t(X) = \mathrm{Tr}\varphi_t''(X) - \tfrac{1}{2}(BX, \varphi_t'(X))$$

is fulfilled.

Theorem III.3: The Markovian semigroup T_t [Eq. (III.13)] is a contraction in L^2_μ.

Proof: follows from Theorem III.1, because

$$\frac{d}{dt}(T_t^\dagger 1, F) = \frac{d}{dt}(1, T_t F) = (1, \Delta T_t F) = (\Delta 1, T_t F) = 0 .$$

Therefore, $T_t^\dagger 1 = \text{const} = 1$.

Theorem III.4: The semigroup T_t has the form $T_t = e^{-t\Delta}$ with Δ as a self-adjoint positive definite operator. This semigroup can be analytically continued to a unitary group $U_t = e^{it\Delta}$, fulfilling the Schrödinger equation

$$-i\frac{\partial}{\partial t} U_t = \Delta U_t . \tag{III.14}$$

Proof: The semigroup (III.13) is self-adjoint, because its infinitesimal generator Δ [(III.6)] is symmetric. Now, the Hille–Yosida theorem[12] implies that $T_t = e^{-t\tilde{\Delta}}$, where $\tilde{\Delta}$ is a self-adjoint positive definite generator of T_t. So, $\tilde{\Delta}$ is the self-adjoint extension of Δ [(III.6)] to all F such that $\lim_{t\to 0}(1/t)(T_t - 1)F$ exists (we shall further identify Δ and $\tilde{\Delta}$). From positive definiteness of $\tilde{\Delta}$ follows the possibility of analytic continuation.

IV. QUANTUM MECHANICS OF EXTENDED PARTICLES

Due to Theorem III.4 we can continue analytically Eqs. (III.12) and (III.13). We can see that $\Theta(t - t_0) \times \rho(-it_0, X, -it, \Gamma)$ plays the role of the Green's function, but now it cannot be written as a density times Lebesgue measure as in the usual quantum mechanics of particles with finite degrees of freedom. A special choice of B in Eq. (III.4) will relate our Green's function (III.12) with that of the string model.[1] Let us assume that the internal parameter $0 \leq \xi \leq \pi$; then L^2 should be replaced by $L^2_{[0,\pi]}$. Let us choose $B = \frac{1}{2}(-d^2/d\xi_m^2)^{1/2} + 1$ and $h_n(\xi) = \pi^{-1/2}\exp[2in\xi]$ as an orthonormal basis in $L^2_{[0,\pi]}$. We can write then $\rho(-it_0, X, -it, \Gamma)$ in a formal way as an infinite product

$$\begin{aligned}
&\rho(-it_0, X, -it, \Gamma) \\
&= \prod_{m=1}^{\infty} \exp[\tfrac{1}{2} im(t - t_0)] \prod_{m=1}^{\infty} \exp[\tfrac{1}{4} m x_m^2] \\
&\times \int \prod_{m=1}^{\infty} dx'_m \exp[-\tfrac{1}{4} m x_m'^2] \left(\frac{4\pi}{m} i \sin[m(t - t_0)]\right)^{-1/2} \\
&\times \exp\left[-\frac{m}{4i\sin[m(t - t_0)]}\right. \\
&\left.\times (\cos[m(t - t_0)](x_m^2 + x_m'^2) - 2x_m x'_m)\right] .
\end{aligned} \tag{IV.1}$$

The density function in the formula (IV.1) coincides with the Green's function for strings in Ref. 1. However, it can be easily seen that we need not assume that we have a string. It is sufficient that B have the same spectrum as an oscillator consisting of positive integers. In particular, in the three-dimensional case we get a similar Green's function with $n\omega$ [we have put $\omega = 1$ in Eq. (IV.1)] replaced by $n_1\omega_1 + n_2\omega_2 + n_3\omega_3$.

We can still perturb the Hamiltonian Δ of a free extended particle by an external potential $V(X, t)$, which causes deformation of the particle. The wavefunction then fulfills the equation

$$-i\frac{\partial}{\partial t}\varphi_t(X) = (\Delta + V(X, t))\varphi_t(X) . \tag{IV.2}$$

Solutions of this equation can be obtained by means of the Feynman path integral (see Ref. 7 for the standard procedure). In particular, the fundamental solution of Eq. (IV.2) fulfilling the initial condition $\varphi_{t_0}(X) = \chi_\Gamma(X)$ has the form

$$\begin{aligned}
&\rho_V(-it_0, X, -it, \Gamma) \\
&= \int \exp[i\int_{t_0}^{t} V(X(\tau), \tau)d\tau]\chi_\Gamma(X(t))\mu_\sigma^{-i(t-t_0),X}(dX(\cdot)) ,
\end{aligned} \tag{IV.3}$$

where the measure $\mu_\sigma^{-i(t-t_0),X}$ is analytic continuation of the measure $\mu_\sigma^{t-t_0,X}$ appearing in the formula (III.12). The "paths" $X(\tau)$ in the integral (IV.3) are paths in the space S', because $X \in S'$. Again $\Theta(t - t_0)$ $\rho_V(-it_0, X, -it, \Gamma)$ plays the role of the retarded Green's function $[\rho_V^R(t - t_0, X, \Gamma)]$ for the theory with interaction and fulfills the equation

$$\left(-i\frac{\partial}{\partial t} - \Delta - V(X, t)\right)\rho_V^R(t - t_0, X, \Gamma) = -i\delta(t - t_0)\chi_\Gamma(X) . \tag{IV.4}$$

We are interested in computation of the transition amplitude from one configuration $\varphi_t^{\text{in}}(X)$ of the extended particle (φ^{in} is the state in $t = -\infty$ when interaction is turned off) to another $\varphi_t^{\text{out}}(X)$, when $t \to +\infty$. This amplitude is defined by (cf., e.g., Ref. 13)

$$\begin{aligned}
S &= \lim_{t\to\infty} (\varphi^{\text{out}}, \varphi_R) \\
&= \lim_{t\to\infty} \int \mu(dX)\varphi^{\text{out}}(t, X)\varphi_R(t, X) ,
\end{aligned} \tag{IV.5}$$

where $\varphi_R(t, x)$ is the retarded solution of Eq. (IV.2)

$$\begin{aligned}
\varphi_R(t, X) &= \varphi^{\text{in}}(t, X) + i\int dt' V(X', t')\varphi_R(t', X') \\
&\times \rho_V^R(t - t', X, dX') .
\end{aligned} \tag{IV.6}$$

It can be shown similarly as in the conventional quantum mechanics that the transition amplitude (IV.5) can be written by means of the Green's function in a simple form (this formula can be easily checked straightforwardly in perturbative calculation)

$$\begin{aligned}
S &= (\varphi^{\text{out}}, \varphi^{\text{in}}) + \int \mu(dX_1)\, dt_1 dt_2 \bar{\varphi}^{\text{out}}(X_1, t_1) \\
&\times \overrightarrow{K}_{X_1,t_1}\, \rho_V^R(t_2 - t_1, X_1, dX_2)\overleftarrow{K}_{X_2,t_2}\varphi^{\text{in}}(X_2, t_2) ,
\end{aligned} \tag{IV.7}$$

where $K_{X,t} = i\partial/\partial t + \Delta$, $\overline{K}_{X,t} = -i\partial/\partial t + \Delta$, and the functional derivative of a measure is defined in the same way as the derivative of a function:

$$\begin{aligned}
\left(\frac{\delta}{\delta X}\mu(dX), h\right) &= \lim_{s\to 0}\frac{1}{s}[\mu(d(X + sh)) - \mu(dX)] \\
&= \lim_{s\to 0}\frac{1}{s}\left[\frac{\mu(d(X + sh))}{\mu(dX)} - 1\right]\mu(dX) .
\end{aligned}$$

Here $\mu((d(X + sh))/\mu(dX)$ denotes the Radon-Nikodym derivative.

The formula (IV.7) provides us with the Green's function formulation of the scattering problem. By means of this formula multiple scattering (nth order term[13]) on

an external potential can be computed. Such calculations have been performed in the interaction picture[14,15] for four-dimensional position X_ν and external field $V(X) = \exp[\, ik_\nu X_t^\nu(0)]$ leading to the n-point Veneziano amplitude. However, this is beyond the scope of this paper.

V. FINAL REMARKS

After solution of the problem of time evolution of an extended object one should answer the question of Lorentz invariance of the theory. We could follow Refs. 1, 16 for the construction of generators of the Lorentz group. However, this construction works only in 26 space–time dimensions. This is the essential stumbling block of any, so-far proposed, model of field theory of strings.

It may be that this difficulty is related to another problem, which remained unsolved. Namely, any theory of extended particles should be invariant under a change of internal parameters. One expects that this (gauge) symmetry will be unitarily implemented in Hilbert space of solutions of the Schrödinger equation. It seems possible that Lorentz invariance lost by special choice of gauge[15,1] can be restored by the action of unitary operators of gauge transformations. This problem is now under investigation.[17] Our stochastic approach can be helpful in the problem of gauge symmetry, because the infinite-dimensional $SL(2, R)$ invariance[16] is an intrinsic property of infinite dimensional Wiener and Ornstein–Uhlenbeck processes (see Hida's lectures[18] for the finite-dimensional case; cf. also Ref. 19).

ACKNOWLEDGMENTS

I wish to acknowledge helpful discussions with J. Lukierski during the course of this work. I am also indebted to W. Woyczynski for an introduction to functional equations.

[1] M. Kaku and K. Kikkawa, Phys. Rev. D **10**, 1110 (1974).
[2] Ch. Marshall and P. Ramond, Nucl. Phys. B **85**, 375 (1975).
[3] G. Domokos, Phys. Rev. D **11**, 373 (1975).
[4] E. B. Dynkin, *Markov Processes* (Academic, New York, 1965).
[5] L. Gross, J. Funct. Anal. **1**, 123 (1967).
[6] Y. L. Daletski, Usp. Mat. Nauk **22**, 1 (1967) (in Russian).
[7] M. Kac, *Proc. 2nd Berkeley Sympos. Math. Stat. and Probab.* (Univ. of California Press, Berkeley, 1951), pp. 189–215.
[8] I. M. Gel'fand and N. Y. Vilenkin, *Generalized Functions, Vol.* 4 (Academic, New York, 1964).
[9] B. Simon, *The $P(\varphi)_2$ Euclidean (Quantum) Field Theory* (Princeton U. P., Princeton, N. J., 1974).
[10] Y. Umemura, J. Math. Kyoto Univ. **4**, 477 (1965).
[11] G. E. Uhlenbeck and L. S. Ornstein, Phys. Rev. **36**, 823 (1930).
[12] E. Hille and R. S. Phillips, *Functional Analysis and Semigroups*, Amer. Math. Soc. Colloq. **31** (Am. Math. Soc., Providence, R. I., 1957).
[13] I. Bialynicki-Birula and Z. Bialynicka-Birula, *Quantum Electrodynamics* (Pergamon, New York, 1975).
[14] Y. Miyamoto, Progr. Theor. Phys. **43**, 564 (1970).
[15] C. Rebbi, Phys. Rep. C **12**, 1 (1974).
[16] P. Goddard, J. Goldstone, C. Rebbi, and C. B. Thorn, Nucl. Phys. B **56**, 109 (1973).
[17] Z. Haba and J. Lukierski, in preparation.
[18] T. Hida, *Stationary Stochastic Processes* (Princeton U. P., Princeton, N. J., 1970), pp. 17 and 117.
[19] Y. Miyamoto, Progr. Theor. Phys. **43**, 1620 (1970).

Quantum counting processes

M. D. Srinivas

Department of Theoretical Physics, University of Madras, Guindy Campus, Madras-600 025, India
(Received 10 November 1976)

It is proposed that counting experiments in quantum physics should be analyzed in terms of point processes (QPP) defined in the framework of quantum probability theory. A coincidence approach is developed for a class QPP called the regular QPP. A counting formula is derived which determines completely the counting statistics of a regular QPP by means of a pair of "generators."

1. INTRODUCTION

The classical theory of counting or point processes has been successfully applied to a variety of phenomena[1] which involve a random sequence of events in time or a randomly located population. In particular, it forms the basis for many of the investigations[1-7] of counting experiments, where the random arrival times of a beam of elementary particles are observed by a system of detectors. The central objective of such investigations is to derive an expression (referred to as the counting formula) for the probability $p((t, t+T], m)$ that m counts occur in the interval $(t, t+T]$. One is thus led to the analysis of a situation where the detector (or a system of detectors) performs continuous observations on the system in the interval $(t, t+T]$.

It is well known[8] that in quantum theory the statistics of successive observations exhibits nonclassical features like the so called "interference of probabilities." Recent investigations[9,10] have led to a framework of quantum probability theory well suited to the analysis of statistics of successive observations in quantum theory. In this paper we undertake a study of point processes in quantum probability theory as a possible framework for analyzing counting experiments involving elementary particles.

We define a quantum point process (QPP) in such a way that it bears a close anology with classical point processes (CPP). We show that the so-called quantum stochastic processes (QSP) investigated by Davies,[11-13] are nothing but a certain restricted version of QPP in a Schrödinger picture formulation. We study a class of QPP called the regular QPP for which the counting statistics can be determined by a coincidence approach as for the regular CPP.[14] In particular, we show that a regular QPP can be characterized by a pair of "generators" and obtain a general counting formula in terms of these generators. We also make a few remarks on the physical interpretation of the formalism.

2. CLASSICAL POINT PROCESS

In this section we briefly outline the coincidence approach to classical point processes.[14] We shall only consider CPP defined on the real line R, which shall be taken to be the time axis; $\beta(R)$ denotes the σ-algebra of all Borel sets in R. If $(\Omega, \mathfrak{F}, \mu)$ is a probability space then a CPP may be defined[15] in terms of a function

$$\xi\colon \beta(R)\times\Omega\to Z^+,$$

where Z^+ is the set $\{0, 1, 2, \cdots\}$ of all nonnegative integers; the function $\xi(A, \omega)$ is required to satisfy the following properties:

(i) For each $A\in\beta(R)$, $\xi(A, \circ)$ is a Z^+-valued random variable on $(\Omega, \mathfrak{F}, \mu)$;

(ii) For each $\omega\in\Omega$ (except maybe a set of measure zero), $\xi(\cdot, \omega)$ is a Z^+-valued measure on $\beta(R)$.

In order to be able to extend this definition to quantum probability theory we have to characterize the random variables $\xi(A, \cdot)$ in terms of their inverse maps. We hence define a CPP as a collection of random variables $\{N_A \mid A\in\beta(R)\}$, which satisfy the following properties (CP1) and (CP2):

(CP1) for each $A\in\beta(R)$, N_A is a map

$$N_A\colon \beta(Z^+)\to\mathfrak{F},$$

where $\beta(Z^+)$ is the set of all subsets of Z^+. In order to be the inverse map of the random variable $\xi(A, \cdot)$, N_A has to satisfy[16] the following:

$$\text{(a) } N_A(\phi)=\phi; \tag{2.1}$$

$$\text{(b) } N_A(Z^+)=\Omega; \tag{2.2}$$

(c) If $\{X_i\}$ is a denumerable collection of mutually disjoint subsets of Z^+, then

$$N_A(X_i)\cap N_A(X_j)=\phi,$$

for all i, j and

$$N_A\Big(\bigcup_i X_i\Big)=\bigcup_i N_A(X_i). \tag{2.3}$$

(CP2) (a) For each $m\in Z^+$ and a denumerable collection $\{A_i\}$ of mutually disjoint sets $A_i\in\beta(R)$,

$$N_{\bigcup_i A_i}(\{m\})=\bigcup_{\sum m_i=m}\Big[\bigcap_i (N_{A_i}(\{m_i\}))\Big]. \tag{2.4}$$

(b) If $A_n\in\beta(R)$ and $A_n\downarrow\phi$, then

$$N_{A_n}(\{0\})\uparrow\Omega. \tag{2.5}$$

We thus see that a CPP on R is nothing but a family of Z^+-valued random variables which is indexed by the Borel sets $\beta(R)$ in such a way that a realization of the process is also a measure on $(R, \beta(R))$.

Just as in the case of a classical stochastic process, a CPP can also be characterized by its finite dimensional distributions, such as

$$p_r(A_1, k_1; A_2, k_2; \ldots; A_r, k_r)=\mu\Big\{\bigcap_{i=1}^{r} N_{A_i}(\{k_i\})\Big\}, \tag{2.6}$$

which gives the joint probability that k_i counts occur in

 Copyright © 1977 American Institute of Physics

the time period A_i. These joint probability distributions satisfy[15] the following relations:

(C1) For every permutation $(i_1, i_2, \ldots, i_r)$ of $(1, 2, \ldots, r)$,

$$p_r(A_1, k_1; A_2, k_2; \ldots; A_r, k_r) = p_r(A_{i_1}, k_{i_1}; A_{i_2}, k_{i_2}; \cdots; A_{i_r}, k_{i_r}); \quad (2.7)$$

(C2) $$\sum_{k_i=0}^{\infty} p_{r+1}(A_1, k_1; \ldots; A_{i-1}, k_{i-1}; A_i, k_i; A_{i+1}, k_{i+1}; \cdots) = p_r(A_1, k_1; \ldots; A_{i-1}, k_{i-1}; A_{i+1}, k_{i+1}; \cdots); \quad (2.8)$$

(C3) $$\sum_{k=0}^{\infty} p_1(A, k) = 1; \quad (2.9)$$

(C4) whenever $A_1, A_2, \ldots, A_r$ are mutually disjoint, then

(a) $$p_1\left(\bigcup_{i=1}^{r} A_i, k\right) = \sum_{k_1+\cdots+k_r=k} p_r(A_1, k_1; A_2, k_2; \ldots; A_r, k_r), \quad (2.10)$$

and

(b) $$p_{r+1}\left(\bigcup_{i=1}^{r} A_i, k; A_1, k_1; A_2, k_2; \ldots; A_r, k_r\right) = \delta_{\sum_{i=1}^{r} k_i, k}\, p_r(A_1, k_1; A_2, k_2; \ldots; A_r, k_r), \quad (2.11)$$

where

$$\delta_{\sum_{i=1}^{r} k_i, k} = \begin{cases} 0 & \text{when } \sum_{i=1}^{r} k_i \neq k, \\ 1 & \text{when } \sum_{i=1}^{r} k_i = k. \end{cases}$$

(C5) whenever $A_n \downarrow \phi$,

$$p_1(A_n, 0) \uparrow 1. \quad (2.12)$$

It is shown by Moyal[17] that a set of joint probability distributions $\{p_r\}$ that satisfy (C1)–(C5), characterize a unique CPP.

A majority of the theoretical investigations of counting phenomena in physics[2-5] are concerned with obtaining an expression (counting formula) for the probability $p((t, t+T], n)$ that n counts occur in the interval $(t, t+T]$. Such a counting formula is usually obtained[4,6,7] in terms of the so-called "exclusion" and "coincidence" probability densities, which in turn can be easily specified by the physical model under consideration. It has recently been shown by Macchi,[14] that for a class of CPP, which may be called the "regular CPP," both the exclusion probability densities (EPD), and coincidence probability densities (CPD), exist and characterize the process completely. In order to define these CPD and EPD, we first define the coincidence probability $H_r(A_1, A_2, \ldots, A_r)$ and the exclusion probability $\tilde{P}_r(A_1, A_2, \ldots, A_r)$, for a mutually disjoint collection of sets $\{A_1, A_2, \ldots, A_r\}$ by the relations

$$H_r(A_1, A_2, \ldots, A_r) = p_r(A_1, 1; A_2; 1; \ldots; A_r, 1), \quad (2.13)$$

and

$$\tilde{P}_r(A_1, A_2, \ldots, A_r) = p_{r+1}\left(A_1, 1; A_2, 1; \ldots; A_r, 1; R \setminus \bigcup_{i=1}^{r} A_i, 0\right). \quad (2.14)$$

In (2.14) we have assumed that the counting experiment is performed over the whole of the time axis R. The CPD $h_r(t_1, t_2, \ldots, t_r)$ and the EPD $\tilde{p}_r(t_1, t_2, \ldots, t_r)$ can now be defined by the formulas

$$h_r(t_1, t_2, \ldots, t_r) = \lim_{\tau_i \to 0} \frac{H_r((t_1, t_1+\tau_1], (t_2, t_2+\tau_2], \ldots, (t_r, t_r+\tau_r])}{\tau_1 \tau_2 \cdots \tau_r}, \quad (2.15)$$

and

$$\tilde{p}_r(t_1, t_2, \ldots, t_r) = \lim_{\tau_i \to 0} \frac{\tilde{P}_r((t_1, t_1+\tau_1], (t_2, t_2+\tau_2], \ldots, (t_r, t_r+\tau_r])}{\tau_1 \tau_2 \cdots \tau_r}. \quad (2.16)$$

For a regular CPP, both h_r and $\tilde{p}_r$ exist and satisfy the following relation,[14]

$$h_r(t_1, t_2, \ldots, t_r) = \sum_{j=0}^{\infty} \frac{1}{j!} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} d\theta_1 d\theta_2 \cdots d\theta_j \times \tilde{p}_{r+j}(t_1, \ldots, t_r, \theta_1, \ldots, \theta_j). \quad (2.17)$$

For the so-called "completely regular CPP," Macchi[14] has also shown that the relation (2.17) can be inverted to yield the formula

$$\tilde{p}_r(t_1, t_2, \ldots, t_r) = \sum_{j=0}^{\infty} \frac{(-1)^j}{j!} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} d\theta_1 d\theta_2 \cdots d\theta_j \times h_{r+j}(t_1, \ldots, t_r, \theta_1, \ldots, \theta_j). \quad (2.18)$$

We would like to emphasize that for a CPP both h_r and $\tilde{p}_r$ are symmetric nonnegative functions. This symmetry property [as well as the consistency relations (C1)–(C5)] is essential for the derivation of (2.17) and (2.18).

From the definitions (2.13) and (2.15), it is clear that $h_r(t_1, t_2, \ldots, t_r)$ is the joint probability density that one count occurs around each of the instants t_i $(i = 1, 2, \ldots, r)$, with nothing being specified about the rest of the duration of the experiment. Similarly, $\tilde{p}_r(t_1, t_2, \ldots, t_r)$ is the joint probability density that one count occurs around each of the instants t_i $(i = 1, 2, \ldots, r)$, and no count occurs in the rest of the duration of the experiment. Hence a counting formula can now be written in terms of the EPD as follows:

$$p((t, t+T], n) = \frac{1}{n!} \int_t^{t+T} \cdots \int_t^{t+T} \tilde{p}_n(t_1, \ldots, t_n) dt_1 \cdots dt_n. \quad (2.19)$$

In obtaining (2.19) the absence of any multiple occurance of counts as implied by the regularity of CPP is crucial. Also, for a completely regular CPP, the relation (2.18) can be used to rewrite (2.19) in terms of CPD alone, viz.

$$p((t, t+T], n) = \sum_{j=0}^{\infty} \frac{(-1)^j}{j!} \int_t^{t+T} \cdots \int_t^{t+T} dt_1 \cdots dt_n \times \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} d\theta_1 \cdots d\theta_j \times h_{n+j}(t_1, t_2, \ldots, t_n, \theta_1, \ldots, \theta_j). \quad (2.20)$$

One of the most commonly encountered CPP in physics

is the so-called doubly stochastic (or conditioned) Poisson process; the CPD of the process are specified as the correlation functions of a random intensity function $I(t)$, i.e.,

$$h_r(t_1, t_2, \ldots, t_r) = \langle I(t_1)I(t_2)\cdots I(t_r)\rangle . \tag{2.21}$$

From (2.20) and (2.21) we obtain the relation

$$p([t, t+T], n) = \langle (W^n/n!)e^{-W}\rangle , \tag{2.22}$$

where

$$W = \int_t^{t+T} I(t')dt' .$$

The counting formula (2.22) is nothing but the well-known Mandel formula[2] which describes the statistics of photon counting experiments as per classical theory.

3. DEFINITION OF A QUANTUM POINT PROCESS

In this section the formal definition of a QPP is obtained on the basis of an analogy with a CPP. For this we need to survey briefly the framework of quantum probability theory. The basic idea[10] is that to each event (or what is sometimes called an experimentally verifiable proposition), there is associated an experimental procedure, which in general alters the state of the system and is completely characterized by a "measurement transformation" or "operation." It can then be argued from very general empirical considerations that the space of events in quantum theory should have the structure of the *set of positive elements in the unit ball of an ordered Banach algebra.* This constitutes a major departure from the framework of classical probability theory where the space of events has the structure of a Boolean σ-algebra.

In the present investigation we restrict ourselves to the standard Hilbert space formulation of quantum theory, for which case, the structure of the space of events has been analyzed in detail in Ref. 10. If V is the Banach space (under the trace norm), of the set of all self-adjoint trace-class operators in a Hilbert space H, and V^+ is the closed cone of positive operators in V, then the space of events O is the set of positive, norm nonincreasing, linear operators on V (also called the set of operations), i.e., $\mathcal{E} \in O$ is a mapping

$$\mathcal{E}: V \to V$$

such that

$$\begin{aligned} &(O1)\ \mathcal{E} \text{ is linear;} \\ &(O2)\ \text{If } v \in V^+ \text{ then } \mathcal{E}(v) \in V^+ \text{ also;} \\ &(O3)\ \mathrm{Tr}[\mathcal{E}(v)] \leq \mathrm{Tr}v, \\ &\qquad \text{for all } v \in V^+. \end{aligned} \tag{3.1}$$

One of the special features of the quantum-event-space O is that it has a subset Σ (consisting in general of more than one element), of maximal elements which have the property[10]

$$\xi \in \Sigma \Rightarrow \mathrm{Tr}[\xi(v)] = \mathrm{Tr}v , \tag{3.2}$$

for all $v \in V^+$. On the other hand, there is a unique null element θ given by $\theta(v) = 0$ for all v.

In O the conjunction of two events $\mathcal{E}_1$, $\mathcal{E}_2$ is given by

$$(\mathcal{E}_1 \wedge \mathcal{E}_2)(v) = \mathcal{E}_2(\mathcal{E}_1(v)) , \tag{3.3}$$

and is noncommutative in general. It clearly corresponds to the experimental procedure where the system is subjected to the sequence of experimental procedures $\{\mathcal{E}_1, \mathcal{E}_2\}$ in that order.

A sequence of events $\{\mathcal{E}_i\}$ is said to be a disjoint sequence of events if

$$\sum_i \mathcal{E}_i \in O,$$

where the lhs is assumed to converge in strong topology. For a disjoint sequence $\{\mathcal{E}_i\}$ of events, the disjunction is defined by the relation

$$\bigvee_i \mathcal{E}_i = \sum_i \mathcal{E}_i . \tag{3.4}$$

By a quantum probability space we mean an ordered pair (O, μ) where μ is a "state" or a "measure" on O, i.e., μ is a mapping

$$\mu: O \to [0, 1],$$

which satisfies the following:

$$(QS1)\ \mu(\theta) = 0; \tag{3.5}$$

$$(QS2)\ \mu(\xi) = 1, \tag{3.6}$$

for all $\xi \in \Sigma$;

(QS3) If $\{\mathcal{E}_i\}$ is a disjoint sequence of events,

$$\mu\left(\bigvee_i \mathcal{E}_i\right) = \sum_i \mu(\mathcal{E}_i) . \tag{3.7}$$

A random variable X (or an instrument or observable) with value space R is a map

$$X: \mathcal{B}(R) \to O$$

which satisfies the following:

$$(QO1)\ X(\phi) = \theta; \tag{3.8}$$

$$(QO2)\ X(R) \in \Sigma; \tag{3.9}$$

(QO3) If $\{E_i\}$ is a sequence of mutually disjoint elements of $\mathcal{B}(R)$, then $\{X(E_i)\}$ is a disjoint sequence of events and

$$X\left(\bigcup_i E_i\right) = \bigvee_i X(E_i) . \tag{3.10}$$

Now, in order to define a QPP, we have to suitably generalize (CP1) and (CP2) to the quantum probability framework. (CP1) can be immediately generalized to the quantum case by considering $\{N_A\}$ to be quantum random variables with value space Z^+, which satisfy (QO1)–(QO3). As regards (CP2), we note that (CP2a) expresses the fact that, given a disjoint covering $\{A_i\}$ of $A \subset \mathcal{B}(R)$, the event that m counts occur in A can be expressed as a disjunction of all the events of the following type: m_1 counts occur in A_1, *and* m_2 counts occur in A_2 and $\cdots$, where $\Sigma m_i = m$. While generalizing this to quantum theory, we have to keep in mind the order in which the conjunction of events is considered, for this reflects the order in which the events occur. If $A_1, A_2 \in \mathcal{B}(R)$, we shall write

$$A_1 > A_2 ,$$

whenever

$$t_1 \in A_1 \text{ and } t_2 \in A_2 \Rightarrow t_1 > t_2 . \tag{3.11}$$

The preceding remarks suggest that for a QPP, it is natural to postulate a condition analogous to (CP2a) whenever $\{A_i\}$ is a collection of sets such that $A_1 > A_2 > A_3 > \cdots$. Finally we shall generalize (CP2b) in an obvious way and include a continuity requirement also in the definition of a QPP.

A QPP may thus be defined as a collection of quantum random variables $\{N_A \mid A \in \mathcal{B}(R)\}$, which satisfy the following conditions (QP1) and (QP2):

(QP1) Each N_A is a mapping

$$N_A : \mathcal{B}(Z^+) \to O$$

which satisfies

(a) $N_A(\phi) = \theta$; (3.12)

(b) $N_A(Z^+) \in \Sigma$; (3.13)

(c) If $\{E_i\}$ is a sequence of mutually disjoint subsets of Z^+, then $\{N_A(X_i)\}$ is a disjoint sequence of operations, and

$$N_A\left(\bigcup_i E_i\right) = \sum_i N_A(E_i), \tag{3.14}$$

where the rhs is assumed to converge in the strong operator topology.

(QP2) (a) If $\{A_i\}$ is a sequence of elements of $\mathcal{B}(R)$ such that $A_1 > A_2 > A_3 > \cdots$, then

$$N_{\bigcup_i A_i}(\{m\}) = \sum_{\sum m_i = m} \prod_i [N_{A_i}(\{m_i\})], \tag{3.15}$$

where on the rhs, the product is taken in such a way that $N_{A_i}(\{m_i\})$ precedes (is to the left of) $N_{A_{i+1}}(\{m_{i+1}\})$. The rhs is also assumed to converge in the strong operator topology.

(b) If $(t, s]$ is an interval in R, then both the maps

$$t \to N_{(t,s)}(Z^+)$$

and

$$s \to N_{(t,s]}(Z^+)$$

are continuous in the strong operator topology.

(c) If $A_n \downarrow A$ or $A_n \uparrow A$, then

$$s\text{-}\lim N_{A_n}(E) = N_A(E), \tag{3.16a}$$

for all $E \in \mathcal{B}(Z^+)$. Also

$$N_\phi(\{0\}) = I, \tag{3.16b}$$

where I is the identity operation.

It is clear that a QPP as defined by (QP1) and (QP2) is a natural generalization to quantum theory of a CPP. It may be noted that since $\mathcal{B}(R)$ is generated by intervals, it is sufficient to consider random variables N_A where A is an interval (like for example $(t, s]$); all other N_A can now be obtained by using (QP2). At this juncture, we would like to point out that the so-called "quantum stochastic processes" investigated by Davies,[11-13] are nothing but a certain restricted version of a QPP as defined above, but considered instead in the Schrödinger picture. We shall elaborate on this connection in the Appendix.

We now make the following identifications:

$$N_{(t,t']}(\{0\}) = S_{t,t'}; \tag{3.17}$$

$$N_{(t,t']}(Z^+) = T_{t,t'}. \tag{3.18}$$

From (QP2) we can conclude that both $\{S_{t,t'}\}$ and $\{T_{t,t'}\}$ are strongly continuous inhomogeneous semigroups (usually referred to as "propagators") of positive contraction operators in $L(V, V)$; in particular, they satisfy the relations

$$S_{t',t''} S_{t,t'} = S_{t,t''}; \tag{3.19a}$$

$$T_{t',t''} T_{t,t'} = T_{t,t''}, \tag{3.19b}$$

for $t'' \geq t' \geq t$;

$$\text{s-}\lim_{t' \to t} S_{t,t'} = \text{s-}\lim_{t' \to t} T_{t,t'} = I . \tag{3.20}$$

Relations analogous to (3.19) and (3.20), have been referred to as generalizations of the Chapman–Kolmogorov relation by Davies.[11] It should be emphasized that these relations are a natural generalization to quantum theory, of the consistency condition (CP2a) for the random variables of a CPP, and as such have nothing to do with Markovicity.[18]

4. QUANTUM COUNTING FORMULA

In this section we undertake a general analysis of QPP which leads us to a quantum counting formula valid for a large class of such processes. We first recall that if $A \in \mathcal{B}(R)$ then $N_A(\{m\})$ corresponds to the operation that m counts occur in the period A. We should remark at this juncture that in all our analysis up to now (and also in what follows) we have been employing the "Heisenberg picture" of evolution, as is the standard practice for a theory of stochastic processes. Here, of course, there is an added complication that our observables $\{N_A\}$ are indexed by Borel subsets $\{A\}$ of the time axis. Thus, under a time evolution (which includes any possible change in the nature of measurements performed, as well as the evolution of the system), the observable N_A will evolve into N_{A+t} where $A + t = \{x \mid (x - t) \in A\}$.

If $A_r > A_{r-1} > \cdots > A_1$, then we can write down the joint probability $p_r(A_1, m_1; A_2, m_2; \ldots; A_r, m_r)$ that m_1 counts occur in A_1 and m_2 counts occur in A_2 and $\cdots$ m_r counts occur in A_r *with no measurement being performed in the intervening periods* as follows:

$$p_r(A_1, m_1; A_2, m_1; \ldots; A_r, m_r) = \mu\left[\prod_{i=r}^{1} N_{A_i}(\{m_i\})\right], \tag{4.1}$$

where the order of factors in rhs is as indicated. The joint probability p_r can be written for any $\{A_i\}$ but it will not have the physical interpretation given above unless $A_r > A_{r-1} > \cdots > A_1$.

In what follows we shall consider only those states μ_ρ which are specified by a "density operator" ρ (i.e., $\rho \in V$ and $\mathrm{Tr}\rho = 1$), in the following sense:

$$\mu_\rho(\mathcal{E}) = \mathrm{Tr}[\mathcal{E}(\rho)]$$

for all $\mathcal{E} \in O$. For a state specified by the density op-

erator ρ, we can write Eq. (4.1) in the following form:

$$p_r(A_1, m_1; A_2, m_2; \ldots; A_r, m_r)$$

$$= \mathrm{Tr}\left[\left\{ \prod_{i=r}^{1} N_{A_i}(\{m_i\}) \right\} \rho \right]. \tag{4.2}$$

The joint probabilities (4.2) exhibit all the typical features of "interference of probabilities"[10] that is characteristic of quantum joint probability distributions. They are not symmetric in general; i.e., (C1) or (2.7) is not satisfied. Also, because the operations $N_{A_i}(\{m_i\})$ do not in general satisfy the "repeatability" property, (C4b) or (2.11) is not valid any more. Finally, the property (C2) or (2.8) is satisfied only when $i = r+1$, otherwise we have inequalities like

$$\sum_{k_1=0}^{\infty} p_2(A_1, k_1; A_2, k_2) \neq p_1(A_2, k_2), \tag{4.3}$$

in general. The relations which continue to be valid in quantum theory also, are the following:

$$\text{(Q1)} \sum_{k_{r+1}=0}^{\infty} p_{r+1}(A_1, k_1; \ldots; A_r, k_r; A_{r+1}, k_{r+1}) = p_r(A_1, k_1; \ldots; A_r, k_r); \tag{4.4}$$

$$\text{(Q2)} \sum_{k=0}^{\infty} p_1(A, k) = 1; \tag{4.5}$$

(Q3) whenever $A_r > A_{r-1} > \cdots > A_1$,

$$p_1\left(\bigcup_{i=1}^{r} A_i, m \right) = \sum_{\sum m_i = m} p_r(A_1, m_1; \ldots; A_r, m_r); \tag{4.6}$$

(Q4) If $A_n \downarrow \phi$, then

$$p_1(A_n, 0) \uparrow 1. \tag{4.7}$$

The fact that the joint probability distributions (4.2) in quantum theory are not symmetric and exhibit nonclassical features as in (4.3), lead us to suspect that CPD and EPD (which are after all derived from p_r) may exhibit similar nonclassical features. This would imply that the classical relations (2.17) and (2.18) [which were essential for the derivation of counting formulas like (2.22)] will not be valid in quantum theory. Before discussing these questions we should characterize a class of QPP for which CPD and EPD are well defined. Based on the pioneering work of Davies[11] on $Q \not\subseteq P$, it is possible to conclude that the CPD and EPD exist if certain conditions are imposed on the process. However, as our discussion in the Appendix shows, Davies' assumptions are extremely restrictive and do not certainly exhaust all the situations where the CPD and EPD exist. We shall instead consider the general case for which the coincidence approach is applicable.

We define a QPP to be *regular* if it satisfies the following conditions (R1) and (R2):

$$\text{(R1)} \; \text{s-}\lim_{\tau \to 0} \frac{N_{(t, t+\tau]}(Z^+ \setminus \{0, 1\})}{\tau} = 0; \tag{4.8}$$

$$\text{(R2)} \; \text{s-}\lim_{\tau \to 0} \frac{N_{(t, t+\tau]}(\{1\})}{\tau} = J_t, \tag{4.9}$$

exists, and $J_t \in L^+(V, V)$, for each t.

The coincidence probability $H_r(A_1, A_2, \ldots, A_r)$ can be defined when $A_r > A_{r-1} > \cdots > A_1$, by the formula

$$H_r(A_1, A_2, \ldots, A_r) = \mathrm{Tr}\left[\left(\prod_{i=r}^{1} N_{A_i}(\{1\}) \right) \rho \right]. \tag{4.10}$$

H_r is the joint probability that one count occurs in each of the A_i, with the *additional specification that no measurement is performed in between the periods* A_i. For a regular QPP the CPD exist and are given by the relation

$$h_r(t_1, t_2, \ldots, t_r) = \mathrm{Tr}[J_{t_r} \cdots J_{t_2} J_{t_1} \rho], \tag{4.11}$$

whenever $t_r > t_{r-1} > \cdots > t_1$. Again these CPD are the joint probabilities that one count occurs around each of t_i *with no measurement being performed in the rest of the duration.* They are also not symmetric in general. This is quite unlike the CPD in classical theory, which are actually obtained by summing over all the possibilities in the intervening periods. In quantum theory, because of the interference of probabilities we have relations like (4.3), which show that summing over all the possibilities for measurements performed in the intervening periods does not lead us to simple quantities like h_r as given by (4.11). In fact the joint probability density $\bar{h}_r$ that one count occurs around each t_i with the additional condition that continuous observations are performed in the intervening periods, and the number of counts may be anything (i.e., the measurements are nonselective), is given by

$$\bar{h}_r(t_1, t_2, \ldots, t_r) = \mathrm{Tr}[J_{t_r} N_{(t_{r-1}, t_r]}(Z^+) J_{t_{r-1}} \cdots J_{t_1} \rho]$$

$$= \mathrm{Tr}[J_{t_r} T_{t_{r-1}, t_r} J_{t_{r-1}} \cdots J_{t_1} \rho]. \tag{4.12}$$

In quantum theory $\bar{h}_r \neq h_r$, in general. The above discussion also shows that, quantities like $p((t, t+T], m)$, $\bar{h}_r$ and the EPD $\tilde{p}_r(t_1, t_2, \ldots, t_r)$ which refer to situations where continuous observations are performed over a finite interval, cannot be expressed in terms of $\{h_r(t_1, t_2, \ldots, t_r)\}$ alone; in other words, the CPD given by (4.11) *do not* characterize a QPP completely.

We shall now calculate the EPD and show that they do determine the counting statistics. For a regular QPP, it follows from (CP1), (CP2), and (R2) that the joint probability $\tilde{p}_r(t_1, t_2, \ldots, t_r)$ that one count occurs around each of the t_i and no count occurs in the rest of the interval $(t, t+T]$ is given by the formula

$$\tilde{p}_r(t_1, t_2, \ldots, t_r)$$

$$= \mathrm{Tr}[N_{(t_r, t+T]}(\{0\}) J_{t_r} N_{(t_{r-1}, t_r]}(\{0\}) \times J_{t_{r-1}} \cdots J_{t_1} N_{(t, t_1]}(\{0\}) \rho]$$

$$= \mathrm{Tr}[S_{t_r, t+T} J_{t_r} S_{t_{r-1}, t_r} J_{t_{r-1}} \cdots S_{t, t_1} \rho], \tag{4.13}$$

for $t + T > t_r > t_{r-1} > \cdots > t_1 > t$. The condition (R1) that we imposed on a regular QPP, rules out multiple occurrences of events. In fact, since

$$N_{(t, t+\tau]}(Z^+ \setminus \{0, 1\}) = \sum_{m=2}^{\infty} N_{(t, t+\tau]}(\{m\}),$$

the condition (R1) implies that

$$\text{s-}\lim_{\tau \to 0} \frac{N_{(t, t+\tau]}(\{m\})}{\tau} = 0, \tag{4.14}$$

for $m \geq 2$.

For a regular QPP we thus have the following relation:

$$p((t, t+T], r) = \int_t^{t+T} dt_r \cdots \int_t^{t_3} dt_2 \int_t^{t_2} dt_1 \times \tilde{p}_r(t_1, t_2, \ldots, t_r). \tag{4.15}$$

Equation (4.15) is the generalization of (2.19) to the case of a regular QPP, where the EPD $\tilde{p}_r$ as given by (4.13) are not symmetric and are meaningful physically only when $t_r > t_{r-1} > \cdots > t_1 > t$. From (4.15) and (4.13), we can write down a counting formula in terms of J_t and $S_{t,t'}$.

The conditions (R1) and (R2) also imply a general relation between J_t and the generator $\tilde{J}_t$ of the semigroup $\{S_{t,t'}\}$. In order to derive this, we start from the relation

$$T_{t,t+\tau} = N_{(t,t+\tau]}(Z^+) = \sum_{m=0}^{\infty} N_{(t,t+\tau]}(\{m\}). \tag{4.16}$$

Now, using (3.17), we get

$$\left(\frac{T_{t,t+\tau} - I}{\tau}\right)\rho = \left(\frac{S_{t,t+\tau} - I}{\tau}\right)\rho + \frac{N_{(t,t+\tau]}(\{1\})}{\tau}\rho + \frac{N_{(t,t+\tau]}(Z^+ \setminus \{0,1\})}{\tau}\rho. \tag{4.17}$$

Taking the limit $\tau \to 0$ and using (R1) and (R2), we first obtain that

$$\mathcal{D}(\Gamma_t) = \mathcal{D}(\tilde{J}_t), \tag{4.18}$$

where $\mathcal{D}(\ )$ denotes the domain of the operator and Γ_t is the generator of the semigroup $\{T_{t,t'}\}$. From (4.17) we also get

$$\Gamma_t \rho = J_t \rho + \tilde{J}_t \rho, \tag{4.19}$$

in $\mathcal{D}(\tilde{J}_t)$. Now since

$$\mathrm{Tr}[T_{t,t+\tau}\rho] = \mathrm{Tr}[N_{(t,t+\tau]}(Z^+)\rho] = \mathrm{Tr}\rho, \tag{4.20}$$

because of (QP1b), we obtain

$$\mathrm{Tr}(\Gamma_t \rho) = 0, \tag{4.21}$$

for all $\rho \in \mathcal{D}(\tilde{J}_t)$. We thus obtain the following relation:

$$\mathrm{Tr}[J_t \rho] = -\mathrm{Tr}[\tilde{J}_t \rho] \tag{4.22}$$

for all $\rho \in \mathcal{D}(\tilde{J}_t)$.

The semigroup $\{S_{t,t'}\}$ can be formally written in terms of its generator $\tilde{J}_t$ as

$$S_{t_1,t_2} = T\exp\left[\int_{t_1}^{t_2} \tilde{J}_t \, dt\right], \tag{4.23}$$

where the rhs actually stands for the strong limit of a sequence of terms[19,20] and T is the time-ordering operator which orders the operators in chronological order from right to left. For a regular QPP it is not true in general that the generators Γ_t and $\tilde{J}_t$ are densely defined for all t. However, in order to derive the counting formula we will now consider only those regular QPP for which these generators are densely defined. A set of regularity conditions on the propagators $\{T_{t,t'}\}$ and $\{S_{t,t'}\}$ can be obtained[19,20] which ensure that Γ_t and $\tilde{J}_t$ are densely defined for all t.

Then we can substitute (4.23) in (4.13) and (4.15) and obtain the relations

$$\tilde{p}_r(t_1, t_2, \ldots, t_r) = \mathrm{Tr}\left[T\left\{J_{t_1} J_{t_2} \cdots J_{t_r} \exp\int_t^{t+T} \tilde{J}_{t'} dt'\right\}\rho\right], \tag{4.24}$$

and

$$p((t, t+T], n) = \mathrm{Tr}\left[T\left\{\frac{(\int_t^{t+T} J_{t'} dt')^n}{n!} \exp\int_t^{t+T} \tilde{J}_{t'} dt'\right\}\rho\right]. \tag{4.25}$$

Equation (4.25) is the quantum counting formula for a regular QPP and it is remarkable that it looks quite a bit like the classical Mandel formula (2.22).

We have therefore shown that for a class of regular QPP there exist the "generators" $J_t \in L^+(V, V)$ and $\tilde{J}_t$ which generates a strongly continuous (inhomogeneous) semigroup of positive contraction operators in $L(V, V)$; the "generators" satisfy the relation (4.22). The counting formula is given by (4.25) in terms of the "generators." Conversely, given the "generators" $J_t \in L^+(V, V)$ and $\tilde{J}_t$ which generates a strongly continuous semigroup of operators in $\mathcal{O}$, then we can construct a regular QPP via the identification

$$N_{t,t+T]}(\{m\}) = T\left[\frac{(\int_t^{t+T} J_{t'} dt')^m}{m!} \exp\int_t^{t+T} \tilde{J}_{t'} dt'\right],$$

and verify that the counting statistics of this process is given by (4.25). The preceding analysis can be generalized to situations where the counts occur at different "locations" of a compact Hausdorff space X; then we have to consider J_t as a "bounded stochastic kernel" in the manner discussed by Davies.[11]

A wide variety of regular QPP can be constructed by suitably choosing $\{J_t\}$ and $\{\tilde{J}_t\}$ which satisfy the above conditions. For example the class of regular QPP discussed by Davies[11] (see also the discussion in the Appendix) correspond to the choice

$$\tilde{J}_t \rho = -\tfrac{1}{2}(R_t \rho + \rho R_t), \tag{4.26}$$

where R_t is the unique operator in $\mathcal{B}^+(\mathcal{H})$ defined by the relation

$$\mathrm{Tr}(R_t \rho) = \mathrm{Tr}(J_t \rho), \tag{4.27}$$

for all $\rho \in V$

In the general case $\{\tilde{J}_t\}$ are unbounded operators. We refer the reader to the extensive literature[19,20] on the study of evolution equations in a Banach space, for a study of sufficient conditions that a set of operators $\{\tilde{J}_t\}$ generates a contraction semigroup. We now consider the question as to whether a QPP can be shown to be regular if certain conditions are imposed on the semigroup $\{S_{t,t'}\}$ and $\{T_{t,t'}\}$. We first assume that the condition (R1) is satisfied by the QPP. It can then be shown, (following closely the line of argument outlined in Sec. 4 of Ref. 11), that the condition (R2) is also satisfied (i.e., the QPP is regular), if $\mathcal{D}(\Gamma_t) = \mathcal{D}(\tilde{J}_t)$ for each t.

5. DISCUSSION

In conclusion we would like to make a few remarks on the physical interpretation of the above formalism. We

have considered a situation where the detector is performing continuous observations on the system. From the general principles of quantum theory we know that each act of observation transforms the state of the system and hence the counting statistics is best discussed in terms of the measurement transformations corresponding to events where a certain number of counts have been recorded in an interval of time. We are thus led to the study of point processes in the framework of quantum probability theory. By imposing the requirement of regularity, we rule out the occurrence of multiple counts and also require that the operation J_t, which specifies the counting rate at t, is well defined; $\mathrm{Tr}[J_t\rho]$ is the probability density that a count is observed around t. For a regular QPP the counting statistics is determined by J_t and a semigroup of operations $\{S_{t,t'}\}$. $S_{t,t'}$ is the operation corresponding to the event that no counts are recorded when continuous observations are made in the interval $(t, t']$. Our discussion shows that a regular QPP is best specified in terms of the "generators" $\{J_t\}$ and $\{\tilde{J}_t\}$. Now, if the evolution of the system during the period when it is not subjected to observation, is given by a one-parameter group of unitary operators $\{U(t)\}$, then we can write

$$J_t\rho = U(t)[J\{U^{-1}(t)\rho U(t)\}]U^{-1}(t), \tag{5.1}$$

and

$$\tilde{J}_t\rho = U(t)[\tilde{J}\{U^{-1}(t)\rho U(t)\}]U^{-1}(t), \tag{5.2}$$

where J and $\tilde{J}$ are the "generators" in a Schrödinger picture description. Of course, such a relation is valid only when the nature of the measurement performed by the detector does not change with time; otherwise we shall have to include the effect of such a change also, in the specification of the time-dependence of J_t. If (5.1) and (5.2) are satisfied then we have the following:

(i) The nature of the measurement performed by the detector is completely characterized by J and $\tilde{J}$.

(ii) The statistics of counting, given by (4.25) together with (5.1) and (5.2), depends not only on the initial state ρ of the system but also on its dynamics as characterized by $U(t)$.

Finally there remains the question as to whether J_t and $\tilde{J}_t$ can be "derived" by assuming a particular form of interaction between the detector and the system. This is of course the well-known problem of measurement in quantum theory with the added complication that J_t and $\tilde{J}_t$ are related to measurement transformations that arise when continuous observations are made. However, one can consistently adopt the "operational" viewpoint that the measurement performed by the detector is to be characterized directly by means of the measurement transformations $N_{(t,t+\tau]}(\{m\})$ (or the "generators" $\{J_t\}$ and $\{\tilde{J}_t\}$). After all, what is always observed is the counting statistics and this is completely determined once such an association is made.

ACKNOWLEDGMENTS

The author is deeply indebted to Professor L. Mandel, Professor E. C. G. Sudarshan, and Professor E. Wolf, for patiently explaining to him several aspects of the theory of counting processes. He is grateful to the members of the Department of Theoretical Physics, University of Madras for their advice and encouragement.

APPENDIX: "QUANTUM STOCHASTIC PROCESSES" OF DAVIES

In this Appendix, we briefly describe the connection between the "quantum stochastic processes" investigated by Davies[11-13] and regular QPP. Davies considers essentially a family of instruments (random variables)

$$\mathcal{E}^t:\ \mathcal{B}(Z^+) \to \mathcal{O},$$

for each $t \geq 0$ which satisfy [apart from the usual relations (Q01)–(Q03)], the following properties:

(i) $\mathcal{E}^0(\{0\}) = I$; (A1)

(ii) For each $t > 0$

$t \to \mathcal{E}^t(Z^+)\rho$ is continuous;

(iii) $\mathcal{E}^{t+s}(\{m\}) = \sum_{m_1+m_2=m} \mathcal{E}^t(\{m_1\})\mathcal{E}^s(\{m_2\})$. (A2)

Let $\{U(t)\}$ be a one-parameter group of unitary operators characterizing the evolution of the system while it is left unobserved. We now define the random variables $\{N_{(t,t+\tau]}\}$ by the relation

$$N_{(t,t+\tau]}(\{m\})\rho = U(t)[\mathcal{E}^\tau(\{m\})\{U^{-1}(t)\rho U(t)\}]U^{-1}(t). \tag{A3}$$

From (i)–(iii) we can easily show that the random variables $\{N_{(t,t+\tau]}\}$ satisfy (QP1) and (QP2). Also a comparison of (A3) with (5.1) and (5.2) show that $\{\mathcal{E}^\tau\}$ are the random variables corresponding to $\{N_{(t,t+\tau]}\}$ in a Schrödinger picture description of the evolution. We again emphasize that a Heisenberg picture setting is more suitable for a study of stochastic processes in both classical and quantum probability theories as we will be able to accommodate any general time-evolution.

Let $\mathcal{B}_0(R)$ be the subset of $\mathcal{B}(R)$ consisting of ϕ, all the intervals $\{(t, t+\tau]\}$ and also every countable union of a disjoint sequence of such intervals. From (i)–(iii) it is possible to show that the set of random variables $\{N_{(t,t+\tau]}\}$ given by (A3), can be extended [by using (QP2)] into a set $\{N_A \mid A \in \mathcal{B}_0(R)\}$ of random variables which satisfy (QP1) and (QP2). We define a "restricted quantum point process" (RQPP) as a collection of random variables $\{N_A \mid A \in \mathcal{B}_0(R)\}$ which satisfy both (QP1) and (QP2). From our discussion it follows that a QSP as defined by (i)–(iii), corresponds to a RQPP. It is of course an open question as to whether a RQPP can be extended into a QPP (as defined in Sec. 3) and also whether such an extension would be unique.

In order to characterize the process $\{\mathcal{E}^\tau\}$ in terms of "generators" Davies imposes the following conditions:

(D1) (Assumption of bounded interaction rate)

$$\mathrm{Tr}[\mathcal{E}^\tau(Z^+ \setminus \{0, 1\})\rho] \leq K\tau\mathrm{Tr}(\rho), \tag{A4}$$

where K is a finite constant.

(D2) The semigroup of operators $\{S_t\}$, given by

$$S_t = \mathcal{E}^t(\{0\}), \tag{A5}$$

are such that they transform pure states into pure states. From (D1) and (D2) Davies shows, in particular, that

$$\text{s-lim}_{\tau\to 0} \frac{\mathcal{E}^\tau(Z^+\setminus\{0,1\})}{\tau} = \theta, \tag{A6}$$

and

$$\text{s-lim}_{\tau\to 0} \frac{\mathcal{E}^\tau(\{1\})}{\tau} = J, \tag{A7}$$

where $J \in L^+(V, V)$.

From (A6) and (A7) it follows that the corresponding RQPP determined by $\{N_{(t,\,t+\tau]}\}$ is regular, i.e., it satisfies the conditions (R1) and (R2). However, we would like to emphasize that the condition (D2) is extremely restrictive and is not motivated by any physical considerations. In fact, for the case of the photon counting problem, it can be shown[21] that the counting formula derived from (4.26) [which one obtains on the basis of (D2)], *does not* bear *any resemblence to the well-known* Glauber–Mandel formula[3,4] or its generalizations.[5] It would be of interest to consider more general semigroups $\{S_t\}$, which do not transform pure states into pure states only, and construct a general class of regular RQPP.

[1] P.A.W. Lewis, Ed. *Stochastic Point Processes Statistical Analysis, Theory and Applications* (Wiley, New York, 1972).
[2] L. Mandel, Proc. Phys. Soc. London 72, 1037 (1958).
[3] L. Mandel, E.C.G. Sudarshan, and E. Wolf, Proc. Phys. Soc. London 84, 435 (1964).
[4] P.L. Kelly and W.H. Kleiner, Phys. Rev. 136, 316A (1964).
[5] M. Lax and M. Zwanziger, Phys. Rev. A 7, 750 (1973).
[6] C. Benard and O. Macchi, J. Math. Phys. 14, 155 (1973).
[7] C. Benard, J. Math. Phys. 16, 710 (1975).
[8] L. de Broglie, *Physics and Microphysics* (Hutchinson, London, 1955).
[9] E.B. Davies and J.T. Lewis, Commun. Math. Phys. 17, 239 (1970).
[10] M.D. Srinivas, J. Math. Phys. 16, 1672 (1975).
[11] E.B. Davies, Commun. Math. Phys. 15, 277 (1969).
[12] E.B. Davies, Commun. Math. Phys. 19, 83 (1970).
[13] E.B. Davies, Commun. Math. Phys. 22, 51 (1971).
[14] O. Macchi, Adv. Appl. Prob. 7, 83 (1975).
[15] D.J. Daley and D. Vere-Jones, in Ref. 1, pp. 299–383.
[16] V.S. Varadarajan, *Geometry of Quantum Theory* (Von Nostrand, Princeton, N.J., 1968), Vol. I.
[17] J.E. Moyal, Acta Math. 108, 1 (1962).

[18] It may be noted that semigroup relations like (3.19) arise quite often in quantum probability theory in connection with situations where continuous observations are made. For each t, t' let $\mathcal{E}(t, t')$ be the operation corresponding to the event that a property P is satisfied in an interval $(t, t']$. If the property P is satisfied in an interval $(t, t'']$ if and only if it is satisfied in both the intervals $(t, t']$ and $(t', t'']$ for $t'' \geq t' \geq t$, then the operations $\mathcal{E}(t, t')$ satisfy the semigroup law

$$\mathcal{E}(t', t'')\,\mathcal{E}(t, t') = \mathcal{E}(t, t'').$$

[19] K. Yosida, *Functional Analysis* (Springer, New York, 1968), 2nd ed., pp. 424–38.
[20] M. Reed and B. Simon, *Methods of Mathematical Physics* (Academic, New York, 1975), Vol. II, p. 285.

[21] M. D. Srinivas, "Quantum Point Process Model for Photodetection" (to appear in the proceedings of the Fourth Rochester Conference on Coherence and Quantum Optics).

The gravitational influence of a beam of light of variable flux

Raymond W. Nackoney

Natural Science Department, Loyola University, Chicago, Illinois 60626
(Received 2 December 1976)

An exact solution is obtained for the Einstein field equations of a columnated, time varying beam of light. The beam is circular in cross section, infinite in path length, and is considered in the geometrical limit. The beam is described in a retarded time coordinate system. The flux density is dependent on the radial coordinate and on the retarded time. The solution is sufficiently general so as to describe a single pulse of light traveling through a vacuum. It also allows the description of acceleration fields which propagate in the direction of the beam at the speed of light. Geodesics are considered in order to test the interpretation of the solutions and the stability of the time varying beam.

INTRODUCTION

The present paper is a sequel to one published several years ago.[1] In that paper (here quoted as I) an exact solution is presented for a beam of light in general relativity. The beam was of constant circular cross section. It had an energy density which was constant in time but which could vary spatially within a cross section of the beam.

The current work extends the above ideas to a beam which varies with time. An exact solution is again obtained. The search for an exact solution was prompted by difficulties in trying to visualize the nature of a light beam and its field. It seemed apparent that a beam of variable flux density and cross section should be easily described in the retarded time metric of I. In addition, perturbations of the solution in I resulted in equations that implied trivial solutions if only variable energy density was incorporated. Finally, an attempt to find a Jean's instability in the first-order solutions resulted in failure. This oddity appeared logical in that light, although material in nature, is nonmassive and cannot be said to behave like a beam of dust or steel. Specifically, the beam of light does not ring when struck. The sum total implied that the field would share the beam's simplicity of description.

This paper presents the field equations and solution pertinent to this problem. The general solution is found to have a matter dependent part and a part which describes an acceleration field imposed along the axis of the beam. The basic form of the geodesics is not discussed in this paper since they are similar to those of Paper I. Geodesics are considered only to illucidate the nature of specific questions unique to this discussion.

1. THE BEAM METRIC

The analysis of this paper is similar to that of Paper I. A beam of light of circular cross section of radius R is propagating along the longitudinal axis, z, of a cylindrical coordinate system. The beam is directed toward positive z values. It is considered in the geometrical limit; as such, there is no diffraction at the edge of the beam and all of the beam's rays are strictly parallel to the z axis. The path length of the beam is infinite with both source and absorber infinitely far away. Hence the gravitational field of either region may be neglected in our calculations.

The energy density, ρ, of the beam is still chosen to be independent of the azimuthal coordinate, θ, and dependent on the radial coordinate, r. In addition, we wish to make the source of the light beam time dependent. This will result in an observer seeing a time dependent energy density. Since the beam is propagating along the z axis one would expect a z dependence in the observed energy density. In Paper I we found that the rays of the beam are unaffected by the gravitational field of the beam. They continue to travel parallel to the z axis at the unique speed of light. In a retarded time metric this implies that a cross section of the beam in the azimuthal plane can be parameterized solely by the retarded time l. We will assume that this result will be maintained in the variable flux problem. Therefore the energy density, ρ, in the current problem will be a function of r and l, but not of z. Essentially this says that each observer will see the same beam cross section at the same retarded time independent of his position along the z axis. This assumption will be later subjected to a check when we examine null geodesics to see if the beam maintains its unique speed and collinearity.

Within the above framework we include a further variation. We will allow the beam radius, R, to vary with time. This will allow a sausage shaped beam, and in the extreme case where R is equal to zero during different time intervals, a string of beads of light separated by vacuum results.

The gravitational influence of the beam shares with the beam density the property of being dependent only on r and l. In the same fashion as was previously discussed, no observer situated on the z axis is unique. He sees the same beam density and the same gravitational effect that any other observer sees at the same retarded time. Hence the gravitational metric is independent of z. The retarded time metric used is a form of Vaidya's Newtonian metric[2] discussed in I. In this case, the space–time interval

$$ds^2 = g_{ik}\,dx^i dx^k , \tag{1}$$

is given by the metric components

$$g_{00} = f, \quad g_{11} = -e^{-2\lambda}, \quad g_{22} = -S^2, \quad g_{33} = 0, \quad g_{03} = e^{\alpha} , \tag{2}$$

 Copyright © 1977 American Institute of Physics

where f, α, λ, and S are functions of r and t. All other components are zero. We could transform r and t in the form

$$t' = t'(r,t), \quad r' = r'(r,t)$$

so as to simplify the metric components. The first transformation would be unwise in that we would jeopardize the ability of the time coordinate to parameterize a cross section of the beam. As for the second, it would seem advisable to maintain a form of the metric which would result in a solution somewhat similar to that in I. By not specifying the radial coordinate too precisely at this time, we leave open its choice until the field equations are to be solved. We will assume that S goes to zero as r goes to zero or the nature of the azimuthal coordinate would change.

The contravariant metric tensor is the same as that given in I, Eq. (8), with S replacing r in the g_{22} component.

The Christoffel symbols are determined by the relation

$$\Gamma^i_{jk} = \tfrac{1}{2} g^{il}(g_{lj,k} + g_{lk,j} - g_{jk,l}).$$

For the metric of Eqs. (2) we find

$$\begin{aligned}
&\Gamma^0_{00} = \frac{\partial \alpha}{\partial t}, \quad \Gamma^0_{01} = \frac{1}{2}\frac{\partial \alpha}{\partial r}, \\
&\Gamma^1_{00} = \frac{1}{2}\frac{\partial f}{\partial r} e^{-2\lambda}, \quad \Gamma^1_{01} = \frac{\partial \lambda}{\partial t}, \\
&\Gamma^1_{03} = \frac{1}{2}\frac{\partial \alpha}{\partial r} e^{\alpha-2\lambda}, \quad \Gamma^1_{11} = \frac{\partial \lambda}{\partial r}, \\
&\Gamma^1_{22} = -S\frac{\partial S}{\partial r} e^{-2\lambda}, \quad \Gamma^2_{02} = \frac{1}{S}\frac{\partial S}{\partial t}, \\
&\Gamma^2_{12} = \frac{1}{S}\frac{\partial S}{\partial r}, \quad \Gamma^3_{00} = \frac{1}{2}\left(\frac{\partial f}{\partial t} - 2f\frac{\partial \alpha}{\partial t}\right)e^{-\alpha}, \\
&\Gamma^3_{01} = \frac{1}{2}\left(\frac{\partial f}{\partial r} - f\frac{\partial \alpha}{\partial r}\right)e^{-\alpha}, \quad \Gamma^3_{01} = \frac{\partial \lambda}{\partial t} e^{2\lambda-\alpha}, \\
&\Gamma^3_{13} = \frac{1}{2}\frac{\partial \alpha}{\partial r}, \quad \Gamma^3_{22} = S\frac{\partial S}{\partial t} e^{-\alpha}.
\end{aligned} \tag{3}$$

The Ricci tensor is obtained from the Christoffel symbols by the relation

$$R_{ik} = \Gamma^l_{kl,i} - \Gamma^l_{ik,l} + \Gamma^l_{mi}\Gamma^m_{kl} - \Gamma^m_{lm}\Gamma^l_{ik}.$$

The nonzero components in our case are

$$R_{22} = Se^{-2\lambda}\left[\frac{\partial^2 S}{\partial r^2} + \left(\frac{\partial \alpha}{\partial r} - \frac{\partial \lambda}{\partial r}\right)\frac{\partial S}{\partial r}\right],$$

$$R_{03} = -\tfrac{1}{2} e^{\alpha-2\lambda}\left[\frac{\partial^2 \alpha}{\partial r^2} + \left(\frac{\partial \alpha}{\partial r}\right)^2 - \frac{\partial \alpha}{\partial r}\frac{\partial \lambda}{\partial r} + \frac{1}{S}\frac{\partial S}{\partial r}\frac{\partial \alpha}{\partial r}\right],$$

$$R_{11} = \frac{\partial^2 \alpha}{\partial r^2} + \frac{1}{S}\frac{\partial^2 S}{\partial r^2} + \frac{1}{2}\left(\frac{\partial \alpha}{\partial r}\right)^2 - \frac{\partial \lambda}{\partial r}\frac{\partial \alpha}{\partial r} - \frac{1}{S}\frac{\partial S}{\partial r}\frac{\partial \lambda}{\partial r},$$

$$R_{01} = \frac{1}{2}\frac{\partial^2 \alpha}{\partial r \partial t} + \frac{1}{S}\frac{\partial^2 S}{\partial r \partial t} - \frac{1}{2S}\frac{\partial S}{\partial t}\frac{\partial \alpha}{\partial r} - \frac{1}{2}\frac{\partial \alpha}{\partial r}\frac{\partial \lambda}{\partial t} - \frac{1}{S}\frac{\partial S}{\partial t}\frac{\partial \lambda}{\partial t},$$

$$R_{00} = -\tfrac{1}{2} e^{-2\lambda}\left[\frac{\partial^2 f}{\partial r^2} + \frac{1}{S}\frac{\partial S}{\partial r}\frac{\partial f}{\partial r} - \frac{\partial \lambda}{\partial r}\frac{\partial f}{\partial r} - \frac{\partial \alpha}{\partial r}\frac{\partial f}{\partial r} + f\left(\frac{\partial \alpha}{\partial r}\right)^2\right] + \frac{\partial^2 \lambda}{\partial t^2} + \frac{1}{S}\frac{\partial^2 S}{\partial t^2} + \left(\frac{\partial \lambda}{\partial t}\right)^2 - \frac{\partial \alpha}{\partial t}\frac{\partial \lambda}{\partial t} - \frac{1}{S}\frac{\partial S}{\partial t}\frac{\partial \alpha}{\partial t}. \tag{4}$$

2. FIELD EQUATIONS

The energy–momentum tensor, T_{ik}, given in Paper I, Eq. (25), is for directed light radiation with a local energy density ρ. As in Paper I, the current beam has all of its flux directed in the positive z direction. Hence we obtain the same energy–momentum tensor. The only nonzero component of this tensor, Eq. (27) (of I) is

$$T_{00} = \rho f, \tag{5}$$

where ρ and f are now functions of r and t. The trace $T^i_{\ i} = 0$ as expected for null radiation.

Since the trace of the energy–momentum tensor is zero, the Einstein field equations with the cosmological constant set to zero are given by

$$R_{ik} = -\left(\frac{8\pi G}{c^4}\right)T_{ik}.$$

We scale the energy density as in I by the relation

$$m = \frac{4\pi G}{c^4}\rho. \tag{6}$$

Using this relation and Eqs. (4) and (5) in the field equations, we obtain

$$Se^{-2\lambda}\left[\frac{\partial^2 S}{\partial r^2} + \left(\frac{\partial \alpha}{\partial r} - \frac{\partial \lambda}{\partial r}\right)\frac{\partial S}{\partial r}\right] = 0, \tag{7}$$

$$e^{\alpha-2\lambda}\left[\frac{\partial^2 \alpha}{\partial r^2} + \left(\frac{\partial \alpha}{\partial r}\right)^2 - \frac{\partial \alpha}{\partial r}\frac{\partial \lambda}{\partial r} + \frac{1}{S}\frac{\partial S}{\partial r}\frac{\partial \alpha}{\partial r}\right] = 0, \tag{8}$$

$$\frac{\partial^2 \alpha}{\partial r^2} + \frac{1}{S}\frac{\partial^2 S}{\partial r^2} + \frac{1}{2}\left(\frac{\partial \alpha}{\partial r}\right)^2 - \frac{\partial \lambda}{\partial r}\frac{\partial \alpha}{\partial r} - \frac{1}{S}\frac{\partial S}{\partial r}\frac{\partial \lambda}{\partial r} = 0, \tag{9}$$

$$\frac{1}{2}\frac{\partial^2 \alpha}{\partial r \partial t} + \frac{1}{S}\frac{\partial^2 S}{\partial r \partial t} - \frac{1}{2S}\frac{\partial S}{\partial t}\frac{\partial \alpha}{\partial r} - \frac{1}{2}\frac{\partial \alpha}{\partial r}\frac{\partial \lambda}{\partial t} - \frac{1}{S}\frac{\partial S}{\partial r}\frac{\partial \lambda}{\partial t} = 0, \tag{10}$$

$$e^{-2\lambda}\left[\frac{\partial^2 f}{\partial r^2} + \frac{1}{S}\frac{\partial S}{\partial r}\frac{\partial f}{\partial r} - \left(\frac{\partial \alpha}{\partial r} + \frac{\partial \lambda}{\partial r}\right)\frac{\partial f}{\partial r} + \left(\frac{\partial \alpha}{\partial r}\right)^2 f\right] - 2\left[\frac{\partial^2 \lambda}{\partial t^2} + \frac{1}{S}\frac{\partial^2 S}{\partial t^2} + \left(\frac{\partial \lambda}{\partial t}\right)^2 - \frac{\partial \alpha}{\partial t}\frac{\partial \lambda}{\partial t} - \frac{1}{S}\frac{\partial S}{\partial t}\frac{\partial \alpha}{\partial t}\right] = 4mf. \tag{11}$$

Note that $\partial f/\partial t$ does not occur in these equations.

The nontrivial part of Eq. (7) may be rewritten

$$e^{\lambda-\alpha}\frac{\partial}{\partial r}\left(\frac{\partial S}{\partial r} e^{\alpha-\lambda}\right) = 0.$$

This may be integrated,

$$\frac{\partial S}{\partial r} e^{\alpha-\lambda} = F(t), \tag{12}$$

where F is an arbitrary function of t. Equation (8) may also be rewritten. Its nontrivial part gives

$$\frac{1}{S}\, e^{\lambda-\alpha}\, \frac{\partial}{\partial r}\left(S e^{-\lambda}\, \frac{\partial e^{\alpha}}{\partial r}\right)=0 .$$

We may integrate this and find

$$S e^{-\lambda}\, \partial e^{\alpha}/\partial r = J(t) , \tag{13}$$

where J is an arbitrary function of t. We now consider the combination of field equations

$$e^{2\lambda}/S^2\ \text{Eq. (7)} + e^{2\lambda-\alpha}\ \text{Eq. (8)} - \text{Eq. (9)} .$$

This results in the equation

$$\frac{1}{2}\, \frac{\partial \alpha}{\partial r}\left(\frac{\partial \alpha}{\partial r}+\frac{4}{r}\, \frac{\partial S}{\partial r}\right)=0 .$$

Two solutions are possible,

$$\frac{\partial \alpha}{\partial r} = -\frac{4}{S}\, \frac{\partial S}{\partial r} \tag{14}$$

or

$$\frac{\partial \alpha}{\partial r} = 0 . \tag{15}$$

Equation (14) integrates to

$$\alpha = \log[C(t)/S^4] ,$$

where C is an arbitrary function of t. The metric component determined by α then reads

$$g_{03} = C(t)/S^4 .$$

As noted in the paragraph following Eqs. (2), S goes to zero as r goes to zero. As such, this solution for α results in a fourth order pole in the metric. As discussed in I, such singularities are unacceptable. They result in unphysical solutions which have essential singularities. This leaves us with Eq. (15) which implies

$$\alpha = \alpha(t) . \tag{16}$$

This in Eq. (12) implies

$$\frac{\partial S}{\partial r}\, e^{-\lambda} = K(t) , \tag{17}$$

where K is an arbitrary function of t. Equation (16) may be used to simplify the field equation (10),

$$\frac{1}{S}\left(\frac{\partial^2 S}{\partial r\, \partial t}-\frac{\partial S}{\partial r}\, \frac{\partial \lambda}{\partial t}\right)=0 .$$

We may integrate this with respect to time,

$$\frac{\partial S}{\partial r}\, e^{-\lambda} = L(r) , \tag{18}$$

where L is an arbitrary function of r. Comparing Eqs. (17) and (18) we find

$$K(t) = L(r) .$$

That is,

$$\frac{\partial S}{\partial r}\, e^{-\lambda} = c , \tag{19}$$

where c is a constant.

We now consider the choice of coordinates more carefully in the hope of illuminating the remaining field equation. Consider a circle around the z axis formed by varying θ at constant r. The three-dimensional metric tensor gives the proper length, dl. In the plane of the circle dl is given by

$$dl^2 = e^{2\lambda} dr^2 + S^2 d\theta^2 .$$

The circumference of the circle is $2\pi S$. The radius of this circle is

$$\int e^{\lambda} dr .$$

Since space–time in a local region is flat, the ratio of these two should be equal to 2π for small r. Hence,

$$2\pi S = 2\pi \int e^{\lambda} dr, \quad r \text{ small} .$$

We may differentiate this and find

$$\frac{\partial S}{\partial r}\, e^{-\lambda} = 1 , \quad r \text{ small} .$$

This implies that the constant in Eq. (19) is equal to one. Equation (19) becomes

$$\frac{\partial S}{\partial r}\, e^{-\lambda} = 1 . \tag{20}$$

Using Eq. (16), the remaining field equation, Eq. (11) reads

$$e^{-2\lambda}\left(\frac{\partial^2 f}{\partial r^2}+\frac{1}{S}\, \frac{\partial S}{\partial r}\, \frac{\partial f}{\partial r}-\frac{\partial \lambda}{\partial r}\, \frac{\partial f}{\partial r}\right)$$
$$-2\left[\frac{\partial^2 \lambda}{\partial t^2}+\left(\frac{\partial \lambda}{\partial t}\right)^2+\frac{1}{S}\, \frac{\partial^2 S}{\partial t^2}\right]=4mf . \tag{21}$$

The above equation does not simplify using Eq. (20). We now impose coordinate conditions. For simplicity in the resulting equations we choose

$$S(r,t) = r . \tag{22}$$

This satisfies the condition on S for small r. Substituting this in Eqs. (20) and (21) gives

$$\lambda = 0 \tag{23}$$

and

$$\frac{\partial^2 f}{\partial r^2}+\frac{1}{r}\, \frac{\partial f}{\partial r}-4m(r,t) f = 0 . \tag{24}$$

Note that α no longer appears in this equation. The only condition on α is Eq. (16); hence, α is an arbitrary function of time. Since it determines the metric component g_{03} we may redefine t such that α equals zero without affecting the retarded time or modifying the quantity f. Our resulting metric comes from this and Eqs. (22), (23), and (24); namely,

$$ds^2 = f dt^2 + 2dt\, dz - dr^2 - r^2 d\theta^2 , \tag{25}$$

where

$$\frac{\partial^2 f}{\partial r^2}+\frac{1}{r}\, \frac{\partial f}{\partial r}-4m(r,t) f = 0 .$$

These are the same equations found in I with the sole addition that m is a function of t as well as r.

3. GEODESIC EQUATIONS

Using Eqs. (3) and the final metric form, Eq. (25), the nonzero Christoffel symbols may be written,

$$\Gamma^1_{00}=\frac{1}{2}\frac{\partial f}{\partial r},\quad \Gamma^1_{22}=-r,\quad \Gamma^2_{12}=\frac{1}{r},$$
$$\Gamma^3_{00}=\frac{1}{2}\frac{\partial f}{\partial t},\quad \Gamma^3_{01}=\frac{1}{2}\frac{\partial f}{\partial r}. \tag{26}$$

The geodesic equations are given by the expression

$$\frac{dv^k}{ds}+\Gamma^k_{il}v^iv^l=0, \tag{27}$$

where v^k is the 4-velocity of a test ray or particle. Combining Eqs. (26) and (27) we obtain

$$\frac{dv^0}{ds}=0, \tag{28}$$

$$\frac{dv^1}{ds}+\frac{1}{2}\frac{\partial f}{\partial r}v^0v^0-rv^2v^2=0, \tag{29}$$

$$\frac{dv^2}{ds}+\frac{2}{r}v^1v^2=0, \tag{30}$$

$$\frac{dv^3}{ds}+\frac{\partial f}{\partial r}v^0v^1+\frac{1}{2}\frac{\partial f}{\partial t}v^0v^0=0. \tag{31}$$

The first of this set has the solution

$$v^0=A, \tag{32}$$

where A is a constant. Equation (30) is in a standard form and has the integral

$$v^2=h/r^2, \tag{33}$$

where h is a constant and may be interpreted for nonnull geodesics as angular momentum per unit mass. To handle the remaining two equations we multiply Eq. (29) by v^1 and Eq. (31) by v^0. Then subtracting the first from the second we obtain

$$v^0\frac{dv^3}{ds}-v^1\frac{dv^1}{ds}+\frac{1}{2}\frac{\partial f}{\partial r}(v^0)^2v^1+\frac{1}{2}\frac{\partial f}{\partial t}(v^0)^3+r(v^2)^2v^1=0.$$

By using Eqs. (32) and (33) we get

$$A\frac{dv^3}{ds}-\frac{1}{2}\frac{d(v^1)^2}{ds}+\frac{A^2}{2}\left[\frac{\partial f}{\partial r}v^1+\frac{\partial f}{\partial t}v^0\right]+\frac{h^2}{r^3}v^1=0. \tag{34}$$

The quantity in the bracket may be rewritten as a total derivative with respect to s since f is a function of r and t

$$\frac{df}{ds}=\frac{\partial f}{\partial r}v^1+\frac{\partial f}{\partial t}v^0. \tag{35}$$

In Eq. (34) this makes each term a derivative with respect to s,

$$A\frac{dv^3}{ds}-\frac{1}{2}\frac{d(v^1)^2}{ds}+\frac{A^2}{2}\frac{df}{ds}-\frac{h^2}{2}\frac{dr^{-2}}{ds}=0.$$

In this form we may immediately integrate and obtain

$$2Av^3-(v^1)^2+A^2f-h^2/r^2=B, \tag{36}$$

where B is a constant. Using Eq. (25) one may check that the above relation is the contraction of the 4-velocity with itself, or the line element in 4-velocity form. As such, the constant B is determined,

$$B=\begin{cases}0, & \text{null trajectory},\\ 1, & \text{nonnull trajectory}.\end{cases} \tag{37}$$

One further integral of the geodesic equations remains. We may rearrange Eq. (31) so as to emphasize a total derivative with respect to s,

$$\frac{dv^3}{ds}+v^0\left(\frac{\partial f}{\partial r}v^1+\frac{\partial f}{\partial t}v^0\right)-\frac{1}{2}\frac{\partial f}{\partial t}(v^0)^2=0.$$

Using Eqs. (32) and (35)

$$\frac{dv^3}{ds}+A\frac{df}{ds}-\frac{1}{2}A^2\frac{\partial f}{\partial t}=0.$$

This does not have a simple integral with respect to s. The first two terms may be integrated, but the last term may only be written formally,

$$v^3=D-Af+\tfrac{1}{2}A^2\int\frac{\partial f}{\partial t}\,ds, \tag{38}$$

where D is a constant of integration. This may be substituted in Eq. (36) to obtain an expression for v^1,

$$v^1=\pm\left(-B+2AD-A^2f-h^2/r^2+A^3\int\frac{\partial f}{\partial t}\,ds\right)^{1/2}. \tag{39}$$

In summary, the geodesic equations are given by

$$v^0=A, \tag{40}$$

$$v^2=h/r^2, \tag{41}$$

$$v^3=D-Af+\tfrac{1}{2}A^2\int\frac{\partial f}{\partial t}\,ds, \tag{42}$$

$$v^1=\pm\left(-B+2AD-A^2f-h^2/r^2+A^3\int\frac{\partial f}{\partial t}\,ds\right)^{1/2}, \tag{43}$$

where A, B, D, and h are constants and B equals zero or one for null or nonnull trajectories, respectively. One may check that this reduces to the geodesic equations of I in which f is independent of time. As in Paper I, we will assume that a light ray or particle may pass through the beam without interaction. Hence all geodesics will be continued indefinitely.

4. ACCELERATION FIELDS

The simplest solution of the field equations results if the scaled energy density, m, is identically zero. Equation (24) becomes

$$\frac{\partial^2 f}{\partial r^2}+\frac{1}{r}\frac{\partial f}{\partial r}=0.$$

If f is to have no singularities of a logarithmic nature, then f is an arbitrary function of time, though positive definite, and independent of the radial coordinate.

To investigate such a field we may use the geodesic equations. With f independent of r, Eq. (35) reduces to

$$\frac{\partial f}{\partial t}=\frac{1}{A}\frac{df}{ds}.$$

This in Eqs. (42) and (43) results in

$$v^3=D-\tfrac{1}{2}Af \tag{44}$$

and

$$v^1=\pm(-B+2AD+h^2/r^2)^{1/2}. \tag{45}$$

In this situation v^1 is independent of f; therefore, it is independent of the retarded time. Only v^3 changes with the retarded time.

To investigate this more fully consider a particle which is at rest at t equal to zero. In Eq. (44) this gives

$$v^3(0) = D - \tfrac{1}{2} A f(0) = 0 .$$

This determines the constant D,

$$D = \tfrac{1}{2} A f(0) ,$$

which in turn determines v^3,

$$v^3 = \tfrac{1}{2} A (f(0) - f(t)) .$$

As f increases with time, v^3 becomes negative and the particle accelerates toward the negative z axis. If f then decreases to the previous value, $f(0)$, v^3 comes back to zero. At all times v^3 matches f in its time dependence. If we consider a set of particles all initially at rest, the relationship of a given particle to its neighbors is not changed with a change in f. The same conclusion results for particles not initially at rest. Light is a special case. Light rays moving parallel to the z axis are the only ones of interest since Eq. (45) (with $B = 0, h = 0$) states that the radial speeds are unaffected. Forward rays are given by $A = 0$; rearward rays are given by $D = 0$. The rearward rays are affected in the same manner as any particle trajectory. However the forward rays are unaffected by changes in f. The retarded nature of the coordinates is preserved. These results are consistent with f representing a true acceleration field. The field is homogeneous in the azimuthal plane, propagates at the speed of light along the z axis, and is directed toward increasing values of z.

The condition placed on f in I was that at $r = 0$, $f = 1$ [Paper I, Eq. (42)]. This would not hold in the above discussion. But an acceleration field is not considered to be physical because it cannot be detected by a set of neighboring particles or null rays. As such, we may transform the coordinates so as to absorb this nonphysical result. In so doing the field f becomes equal to one. The solution is then Minkowski space–time which is consistent with the absence of matter.

5. TIME DEPENDENT MATTER SOLUTION

A solution f of the field equation, Eq. (24), is generally a function of r and t. The only place in which time enters this homogeneous operator on f is in the scaled energy density, m. For each time t, f satisfies the equation for m at that time. This means that f follows m in its time dependence. Though the solution must be continuous in r, it may be discontinuous in t. This occurs because t denotes a wavefront of the beam; hence, m may be discontinuous in this variable. The discontinuity in f follows immediately from Eq. (24). In terms of physical considerations one should not expect the gravitational effect of a given cross section to outdistance the cross section itself. With the source of the beam being infinitely far away, the entire gravitational field wavefront has had time to become a plane wave. Both the beam and the gravitational field propagate at the speed of light, hence they stay together. There is, however, some surprise that the gravitational influence does not have a tail. The field f is dependent only on the current time and not on those times previous to or after this. The field then acts as a plane wave. This is somewhat similar to the plane wave appearance of the Schwarzschild field of a particle which is moving past an observer at nearly the speed of light.

Because of the above considerations, variations in the beam which are impossible to solve in other cylindrical problems can now be handled simply and analytically. For example, the radius, R, of the beam may be a function of time, $R(t)$. This includes beams in which R oscillates with time between two fixed values of r. This would describe a sausage shaped beam. It also includes cases where $R(t)$ is equal to zero for finite time intervals. These would include individual pulses of light separated by vacuum. The considerations of the previous section on acceleration fields would apply here.

As an example of a solution of Eq. (24) we consider a beam of light whose radial extent is nonzero but dependent on time. Its energy density within this radius is dependent only on time. The solution may be taken from the time independent case given in I,

$$\begin{aligned} f &= I_0(2m(t)^{1/2} r), \quad r < R(t) \\ &= I_0(2m(t)^{1/2} R(t)) \\ &\quad + 2m(t)^{1/2} R(t) I_1(2m(t)^{1/2} R(t)) \log\left(\frac{r}{R(t)}\right), \quad r > R(t) , \end{aligned} \tag{46}$$

where I_0 and I_1 are modified Bessel functions of order zero and one, respectively.

In short, the field differs from that caused by massive bodies. The field of a massive object envelopes the object like a peach surrounds its pit. The field in our case envelopes and travels with a cross section of the beam much like the caramel in a caramel cream.

To investigate this field further we look at the geodesic equations. In Eq. (42) we are left with an integral which is not readily accessible to integration. One way to investigate this equation is to compare v^3 at the current time t to v_0^3 at a previous time t_0. If f changes with time we may say that the value of f currently observed, by a test particle or ray, is accumulated from changes along a geodesic from a previous value $f(r_0, t_0)$

$$f(r,t) = f(r_0, t_0) + \int_{r_0, t_0}^{r, t} \frac{df}{ds}\, ds .$$

To simplify this further we assume that the trajectory is such that the radial coordinate, r, remains constant. Hence

$$f(r,t) = f(r, t_0) + \int_{t_0}^{t} \frac{df}{ds}\, ds .$$

In addition we may use the simplification of Eq. (35) as discussed in Sec. 4; namely,

$$\frac{df}{ds} = A \frac{\partial f}{\partial t} . \tag{47}$$

Using this in Eq. (42) we obtain

$$v^3(r,t) = D - \left[A f(t_0) + A \int_{t_0}^{t} \frac{df}{ds}\, ds \right] + \tfrac{1}{2} A \int_{t_0}^{t} \frac{df}{ds}\, ds ,$$

which reduces to

$$v^3(r,t) = v^3(r,t_0) - \tfrac{1}{2}A\int_{t_0}^{t}\frac{df}{ds}\,ds\,.$$

This may be integrated

$$v^3(r,t) = v^3(r,t_0) - \tfrac{1}{2}A(f(r,t) - f(r,t_0))\,. \tag{48}$$

The same analysis applied to Eq. (43) yields

$$v^1(r,t) = v^1(r,t_0)\,. \tag{49}$$

Since v^1 is initially zero, we obtain

$$v^1(r,t) = 0\,. \tag{50}$$

The first case to consider is that of a test ray of light traveling parallel to and in the same direction as the light beam. In this case both B and h are zero. When these are substituted in Eq. (43) we find that A is equal to zero. By Eq. (29) the radial acceleration is now zero, and by Eq. (48) the component v^3 is unchanged. Hence the ray of light is unaffected by the time dependence. Applying this to the beam itself we see that it is in equilibrium under self-gravitation. This result maintains the conclusion of Paper I. In addition, it means that unlike massive beams, no pinch or Jean's instability results from a sausage shaped beam.

The next simplest case to consider is a set of particles which are initially at rest. They are distributed with different distances, r, from the axis but at the same axial coordinate, z. We have found that the time dependent metric does not affect the radial component of the 4-velocity. Hence the particles may be held at their initial radial distances by rockets to offset the radial acceleration of the beam. This ploy will not modify the questions of interest. In this situation the initial values of v^3 are all zero. The change in v^3 is given by Eq. (48). If m has a net increase during this time interval interior to r, the field solution f at r grows in magnitude. The particles are attracted axially into the region of increased m with those farther out having the greater resulting speeds. If, with further changes, m returns to its previous distribution the particles will come to rest. However, since f is a function of r they do not come to rest at the same axial coordinate. Using Eqs. (47) and (48) the particles undergo a displacement

$$z - z_0 = -\tfrac{1}{2}\int_{t_0}^{t}[f(r,t) - f(r,t_0)]\,dt\,, \tag{51}$$

which is a monotonically increasing function of r for increasing m.

This effect can be further studied by observing an initially circular orbit around the beam in the azimuthal plane. Of the initial 4-velocity components only v^0 and v^2 are nonzero. However, as with the particles at rest, only the z component changes with time. With an increase in m the circular orbit changes to a helix which spirals toward the negative z axis. If, at a later time, f returns to its previous initial value the helix closes and returns to a circular orbit. The same may be stated if circular null orbits occur.

In all of the above cases particles and light rays are attracted toward the section of increased m as it comes upon the geodesic. As this section of increased m passes, the geodesics are again attracted toward the departing section. This is consistent with the Newtonian attraction of two bodies. It also is consistent with the result of Tolman, Ehrenfest, and Podolsky[3] in that the net acceleration is zero. However we have the simplification that the three-dimensional potentials are not necessary. The effect is a contact acceleration within a given two-dimensional azimuthal cross section.

One result seems to contradict the conclusions of Paper I. In that paper null helical trajectories can always occur with a backward moving spiral. Forward spiraling trajectories are restricted to sufficiently high energy densities. In the current paper we may start with a circular null orbit if we have a sufficiently high energy density. If the beam intensity decreases, a helix develops from the circular orbit. This is a forward spiraling case. If this were extended to its logical conclusion, we could have a beam of low density with a noticeably forward spiraling null geodesic. By moving in the direction of the beam's flux at the z component of the geodesic velocity, we transform to a new frame. In this new frame the helix is closed up to a circle. This new frame would share the same form of the metric as the preceding. But the energy density of the beam is Doppler shifted to a lower value. The net result is that any beam which changed its energy density with time could have light orbiting in closed circular trajectories. This would be true even if the beam were of very small energy density, say that of a flashlight. Obviously this does not occur in nature.

The paradox is resolved if we consider motion with respect to a material observer and not with respect to the coordinate system. We use an observer who is at rest at the same value of r as the circular orbit. The only geodesic equation of interest is that for the z component of the 4-velocity. We use this in the form of Eq. (51) which compares changes from one beam cross section to another. This result is independent of the value of A and hence independent of whether we talk of the observer or the null trajectory. The observer shares the axial motion of the helix. Hence they stay together and the observer will continue to see a null circular orbit. This is not to say that the change in beam intensity is not detectable. An observer who was initially at rest at a different radial coordinate will see a helical trajectory develop from the above changes. In fact, an observer on the z axis will see a rearward spiraling trajectory in keeping with the results in Paper I.

6. THE GENERAL SOLUTION

The solution of the preceding section is not the most general. The general solution of Eq. (24) is

$$f(r,t)G(t)\,, \tag{52}$$

where f is as given in the preceding section and G is an arbitrary function of time. This function G may be discontinuous but it must be positive definite. It results in an acceleration field which propagates with the beam along the beam axis. As such, this is similar to the section on acceleration fields.

To show the acceleration effect we consider the case

where m is independent of time; hence, $f = f(r)$. As before, we consider particles which are initially at rest and are kept at their radial positions by rockets. The argument leading to Eqs. (48) and (50) remain the same with $f(r)G(t)$ substituted for $f(r,t)$. We now obtain the relations

$$v^1(r,t) = 0\,,$$

and

$$v^3(r,t) = -\tfrac{1}{2} A f(r)[G(t) - G(t_0)]\,. \tag{53}$$

If only the acceleration field, $G(t)$, were present we would inquire about the z component of the 4-velocity as seen from an underlying Minkowski space. Here we apply the same method in a curvilinear space and hence inquire into proper velocities.

The proper distance along the z axis is given by

$$dl = (g_{00})^{-1/2}\, dz\,,$$

or in our case

$$dl = [\, f(r)G(t)]^{-1/2}\, dz\,.$$

The quantity A for a particle at rest is given by Paper I, Eq. (22) to be

$$A = [\, f(r)G(t_0)]^{-1/2}\,.$$

These last relations in Eq. (53) give

$$\frac{dl}{ds} = -\frac{1}{f(r)[G(t)G(t_0)]^{1/2}}\, f(r)[G(t) - G(t_0)] = -\frac{G(t) - G(t_0)}{[G(t)G(t_0)]^{1/2}}\,.$$

This is independent of the radial coordinate. All particles obtain the same proper velocity under the influence of G. Hence $G(t)$ is an acceleration field which may be superimposed on the actual metric. As before, we may neglect its contribution since its effect is the same on all particles.

In conclusion, we have obtained a general solution for a time dependent, columnated beam of light. As mentioned in I, the beam density needed to obtain physically measurable effects appears to be prohibitive. However, as in the last paper, the interpretations of the solution are particularly satisfying. Specifically, the solution itself has an amazing simplicity for a problem of this generality.

A system of linearized equations and some solutions are expected to be published soon.

ACKNOWLEDGMENTS

I wish to thank my wife, Patricia, and the members of the Department of Natural Science for their unfailing moral support.

[1]R.W. Nackoney, J. Math. Phys. 14, 1239 (1973).
[2]P.C. Vaidya, Nature 171, 260 (1953).
[3]R.C. Tolman, P. Ehrenfest, and B. Podolsky, Phys. Rev. 37, 602 (1931).

Spaces of positive and negative frequency solutions of field equations in curved space-times. I. The Klein-Gordon equation in stationary space-times

Carlos Moreno

Laboratoire de Physique Mathématique, Collège de France, Paris, France
(Received 12 November 1976)

In stationary space-times $V_n \times \mathbf{R}$ with compact space-section manifold without boundary V_n, the Klein–Gordon equation is solved by the one-parameter group of unitary operators generated by the energy operator $i^{-1}T^{-1}$ in the Sobolev spaces $H^1(V_n) \times H^{1-1}(V_n)$. The canonical symplectic and complex structures of the associated dynamical system are calculated. The existence and the uniqueness of the Lichnerowicz kernel are established. The Hilbert spaces of positive and negative frequency-part solutions defined by means of this kernel are constructed.

INTRODUCTION

Let V_{n+1} be a globally hyperbolic, $(n+1)$-dimensional Lorentzian manifold of class C^∞ with metric tensor g of the same class.

By definition, the *free* scalar neutral physical field of mass $m > 0$, in this space-time is described by the Klein–Gordon equation

$$(\Delta_{n+1} - m^2)u = 0 , \tag{1}$$

where Δ_{n+1} is the Laplace–Beltrami operator of V_{n+1} and u is a real distribution.

For Eq. (1), and in a neighborhood Ω of each point of V_{n+1}, the existence of two unique fundamental solutions is well known.[1-4] The propagator G of the equation is then the distribution kernel defined as the difference of the two fundamental solutions.

The Cauchy problem for Eq. (1) and for the spacelike hypersurface Σ is solved in Ω by the following expression[3,5]:

$$u(x) = \int_\Sigma \{u(y)\partial_\lambda G(x;y) - G(x;y)\partial_\lambda u(y)\} d\Sigma^\lambda(y) , \tag{2}$$

where $\partial_\lambda = \partial/\partial y^\lambda$, $\{y^\lambda\}$ are local coordinates of the point y of V_{n+1} and $\{d\Sigma^\lambda(y)\}$ the components of the surface vector of Σ.

Moreover, with Eq. (1), is canonically associated a skew-symmetric 2- form acting on its solutions; the local definition of this 2-form is

$$b(u,v) = \int_\Sigma \{u(y)\cdot\partial_\lambda v(y) - \partial_\lambda u(y)\cdot v(y)\} d\Sigma^\lambda(y) . \tag{3}$$

Expressions (2) and (3) are independent of the choice of the spacelike hypersurface Σ.

The primary elements required for the development of the standard quantization program for the field described by Eq. (1) are: (a) a suitable definition of the positive and negative frequency-part solutions, $u^\oplus$ and $u^\ominus$ of the real solution u of Eq. (1); (b) a Hilbert structure on the space of these solutions constructed consistently with (a). In other words,[6,7] Eq. (1) must provide a suitable real Hilbert structure on the space of its real solutions, and a suitable complex structure operator, orthogonal in this space.

This work deals only with points (a) and (b) and not with the quantized field.

Lichnerowicz, in his program[3,5,8] of quantization of free fields in curved space–times V_{n+1}, has set forth a method to introduce the elements (a) and (b), which may be summarized briefly as follows: In the Minkowski space–time M_{n+1} elements (a) and (b) are usually obtained by the Fourier transform.[9,10] In a general curved space–time V_{n+1}, this transform is not avilable. However a method lacking the Fourier transform may be sought to obtain these elements. To this end, consider the following two properties of the distribution kernel G^1 (or Δ^1, D_1, $\cdots$) in M_{n+1} [which is real, symmetric, and a solution of Eq. (1)]:

(i) the fundamental convolution relation

$$G(x;x') = \int_\Sigma \{G^1(x;y)\partial_\lambda G^1(x',y) - G^1(x',y)\cdot\partial_\lambda G^1(x;y)\} d\Sigma^\lambda(y) , \tag{4}$$

is satisfied.

(ii) the operator J between real solutions of Eq. (1), locally given by

$$(Ju)(x') = \int_\Sigma \{u(y)\partial_\lambda G^1(x^1;y) - G^1(x';y)\partial_\lambda u(y)\} d\Sigma^\lambda(y) , \tag{5}$$

is a complex-structure operator on a real space of real solutions of Eq. (1), which satisfies

$$b(JU;Jv) = b(u;v) ,$$

and the expression

$$\{u;v\} = b(u;Jv) , \tag{6}$$

is a scalar product in that real space.

The usual positive and negative frequency-part solutions of the real solution u of Eq. (1) in Minkowski space–time are given here by the expressions

$$u^\oplus = \tfrac{1}{2}(I + iJ)u , \quad u^\ominus = \tfrac{1}{2}(I - iJ)u ,$$

and the usual complex Hilbert structures of the spaces of frequency solutions are given here by the scalar product deduced by linearity from expression (6).

The Lichnerowicz method consists of seeking a distri-

 Copyright © 1977 American Institute of Physics

bution kernel G^1 [which is, symmetric, real, and a solution of Eq. (1)] in the space–time V_{n+1} with properties (i) and (ii) (obviously, these properties are formulable in V_{n+1}), and of obtaining elements (a) and (b) in the same way as in the Minkowski case, mentioned above.

The main goal of the present work is to prove the existence and the uniqueness of the distribution kernel G^1 in the case of stationary curved space–times, $V_{n+1} = V_n \times \mathrm{R}$, with close (i.e., compact and without boundary) space section V_n and to construct the Hilbert spaces of positive and negative frequency solutions following the foregoing Lichnerowicz method.

The mathematical treatment of the problem is greatly simplified, and the physical interpretation of the results is made possible, provided some metric conformal to the quotient metric (see Sec. 1) on space-section manifold V_n, is introduced from the start. The interest of such a metric in general relativity has already been remarked upon in literature[11-13] concerning problems lying far afield from the present one.

The solution provided for the problem has its source in physics: Stationary space–times admit time translations (in an adapted atlas) as a symmetry group, whereby the energy of the solutions of the Klein–Gordon equation may be considered. The Klein–Gordon equation is written in the Schrödinger form $(i^{-1}\partial_0 - H)f = 0$ with $f \equiv (u; \partial_0 u)$ as unknown, and solved in the Sobolev space $H^1(V_n) \times H^0(V_n)$ (which is endowed with the scalar product of the energy), by the group of unitary operators generated by the Hamiltonian H. With the help of the power operators T^{-K} (constructed by a method similar to that used by Seeley to construct complex-power operators of elliptic pseudodifferential operators[14,15]) scalar products for which H is self-adjoint, are introduced on the spaces $H^l \times H^{l-1}$, and the Klein–Gordon equation is solved by the group of unitary operators generated by the Hamiltonian H, in these spaces as well, wherein it defines an infinite-dimensional Hamiltonian system. Following Segal[6,7] a canonical complex structure can be determined on these spaces, from which the construction of the G^1 kernel follows. Under certain specified conditions, it will be unique.

An analogous theorem of existence and uniqueness of the G^1 kernel in stationary space–times has been established independently and simoultaneously by Chevalier.[16] In his work, the Klein–Gordon equation is taken as an ordinary first-order differential system $\partial_0 f = Af$ with unknown $f \equiv (u; \partial_n u)$ where (∂_n) is the normal derivative to the space section V_n. The semigroup theory of operators is used to solve the system on the Banach spaces $H^l \times H^{l-1}$. Operator A generates a semigroup of class C_0 (Hille–Yosida theorem). The construction of the G^1 kernel follows from the existence and the properties of the square root of the operator $-A^2$. However there are differences in the choice of certain technical elements between the Chevalier work and ours; namely, the metric induced on the manifold V_n is used by Chevalier in place of the above-mentioned conformal metric; in the Cauchy problem the normal derivative (∂_n) replaces the time derivative (∂_0) used in this paper and any particular scalar product on $H^l \times H^{l-1}$ is preferred to the energy scalar product employed here.

We believe that the choice of technical elements made in this paper makes it possible to obtain the results of the Chevalier work in a simple way. Additional results are also given (the infinite-dimensional dynamical system determined by the Klein-Gordon equation is studied, the power operators T^{-K} are constructed, the compactness of the operator T is exploited to give series expansions for the G and G^1 kernels).

In Minkowski space-time, Rideau[17] has given an uniqueness theorem for the G^1 kernel, using the Fourier transform. The conditions for the uniqueness of the G^1 kernel in the theorem of Chevalier and in the theorem presented here are the same, and are generalizations of the conditions given in the Rideau theorem.

From the construction of the fundamental solution of the Klein-Gordon equation in Minkowski space-time, given in Ref. 18, it is possible to obtain simultaneously a construction of the G and G^1 kernels in the space-time. E. Combet[19] has proved that this method can be generalized to construct these kernels in static space-times such that the Killing trajectories are geodesics.

Combet[20] has also proved the local existence of the G^1 kernel in static space-times. The Klein-Gordon equation is now written as $\partial_0^2 u + Bu = 0$, where B is a self-adjoint positive operator in $L^2(\Sigma)$, Σ being some open set of the space manifold. This equation may then be solved as a first-order differential system in the Hilbert space $H^1(\Sigma) \times L^2(\Sigma)$, by a group of unitary operators in this space. The propagator of the equation is given by the kernel of the operator $B^{-1/2}\sin(tB^{1/2})$, and the Lichnerowicz kernel is given by that of the operator $B^{-1/2}\cos(tB^{1/2})$.

In a work to be published in collaboration with Combet the existence and uniqueness of the G^1 kernel in any space–time will be considered. The Fourier integral operators technique will be employed. We will use the results contained in the work of Hörmander and Duistermaat[21] on the parametrix of the hyperbolic equations and some remarks of Segal on the same problem in Minkowski space–time.

A large number of references concerning the problem of field quantization in curved space-time is given in the Fulling thesis.[22] At present, a considerable effort is being devoted to this problem (Parker[23,24] Hawking,[25,26] Ruffini *et al.*,[27-29] B. deWitt,[30], Deser and Zumino,[31] Zeldo'vich,[32,33] etc), mainly in relation with matter creation in expanding universes, physical phenomena in the neighborhood of black holes and supersymetries of the gravitational field.

1. THE KLEIN-GORDON EQUATION IN STATIONARY SPACE-TIMES

(a) Let V_n, $n \neq 2$ be a compact manifold of class C^∞ without boundary, and $V_n \times \mathrm{R}$ be a stationary space–time. On the charts of any chosen adapted atlas, with coordinate functions x^α, $\alpha = 0, 1, \ldots, n$ ($x^i, i = 1, \ldots, n$ local coordinates in V_n and $x^0 = t$, canonical coordinate on R) the

metric tensor components $g_{\alpha\beta}(x^\nu)$ are independent of time, and the components of the Killing vector field satisfy

$$\xi_\alpha = g_{0\alpha}\,, \quad |\,\xi\,|^2 = g_{\alpha\beta}\,\xi^\alpha \cdot \xi^\beta = g_{00} > 0\,.$$

We suppose the $V_n \times \{t\}$ submanifolds to be spacelike (i.e., $g^{00} > 0$), and we endow them with the Riemannian manifold structure defined by the following metric, conformal to the quotient metric

$$\gamma_{ij} = |\,\xi\,|^{2/(n-2)}(g_{ij} - |\,\xi\,|^{-2} g_{0i} g_{0j})\,.$$

In the case $n = 3$, several authors[11-13] have pointed out the physical interest of this metric. In this work it appears as a valuable tool for detecting the essential features of the problem.

(b) The local expression of Eq. (1) is

$$-\frac{1}{\sqrt{g}}\;\partial_\alpha(g^{\alpha\beta}\partial_\beta u \cdot \sqrt{g}) - m^2 u = 0\,, \tag{7}$$

where

$$g = \det(g_{\alpha\beta}) \text{ and } \partial_\alpha = \frac{\partial}{\partial x^\alpha}\,.$$

In order to rewrite Eq. (1) in a convenient way, we introduce the following notations:

$$\sigma^2 = g_{00}^{1/(n-2)} \quad v^\alpha = \sigma^{-2} g^{0\alpha}\,, \quad \text{div } \mathbf{V} = -\frac{1}{\sqrt{\gamma}}\;\partial_i(v^i\sqrt{\gamma})$$

$$M = \frac{1}{v^0}\,[(-\Delta_n^*) + m^2\sigma^{-2}]\,, \tag{8}$$

$$N = \frac{1}{v^0}\,[2L(\mathbf{V}) - \text{div}\,\mathbf{V}]\,, \tag{9}$$

where Δ_n^* is the Laplace–Beltrami operator on V_n defined by the $\{\gamma_{ij}\}$ metric tensor; $L(\mathbf{V})$ is the Lie derivative with respect to the vector field $\mathbf{V}$ on V_n, $\gamma = \det(\gamma_{ij})$ and $\sqrt{\gamma} = \sigma^2\sqrt{g}$.

A straightfoward calculation from Eq. (7) leads to the following expression for the Klein–Gordon equation:

$$\partial_t^2 u(t) + N\partial_t u(t) + Mu(t) = 0\,.$$

It will be convenient to rewrite it in differential-system form

$$\frac{d}{dt}\begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} = T^{-1}\begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix}, \quad T^{-1} = \begin{pmatrix} 0 & I \\ -M & -N \end{pmatrix}. \tag{10}$$

2. THE T^{-1} OPERATOR

The functional framework to deal with Eq. (10) is that of Sobolev spaces $H^l(V_n)$ of the V_n manifold.[34,35]

M is an (elliptic) second-order differential operator in V_n; therefore it is a continuous operator from $H^l(V_n)$ to $H^{l-2}(V_n)$. N is a first-order differential operator and is therefore a continuous operator from $H^l(V_n)$ to $H^{l-1}(V_n)$.

Let us recall that the $C^\infty(V_n)$ space is the projective limit of the $H^l(V_n)$ spaces and $\mathfrak{D}'(V_n)$ is their inductive limit.

Relative to the scalar product

$$(u;v)_0 = \int_{V_n} u \cdot \bar{v} \cdot v^0 \eta(\gamma)\,, \quad u, v \in C^\infty(V_n)\,, \tag{11}$$

where $\eta(\gamma) = \sqrt{\gamma}\; dx^1 \wedge \cdots \wedge dx^n$; the completion of the $C^\infty(V_n)$ space is identical with $H^0(V_n)$. Consequently, the latter space becomes a Hilbert space with the scalar product (11).

A straightforward calculation from expressions (8), (9), and (11) allows us to prove

Proposition 1: (i) $\text{Ker}M = \{0\}$.

If $u, v \in H^2$, then

(ii) $(Mu;v)_0 = (u;Mv)_0$.

(iii) $(Mu;u)_0 \geq 0$, $(Mu;u) = 0 \Leftrightarrow u = 0$,

(iv) $\|(M - \lambda I)^{-1}\|_0 \leq |\lambda|^{-1}$ for $\lambda < 0$.

If $u, v \in H^1$, then

(v) $(Nu;v)_0 + (u;Nv)_0 = 0$. □

From (i) and (ii) of the preceding proposition and the general theory of elliptic operators on a manifold, the operator M is an isomorphism between H^{l+2} and H^l whose inverse operator M^{-1} is a pseudodifferential operator of order -2.

The complex poweres M^l, $l \in \mathbb{C}$ of the operator M exist,[15,36,37] since it is an invertible operator of positive order and its resolvent operator satisfies the inequality (iv) of Proposition 1. The M^l operator is a toplinear isomorphism between H^r and $H^{r-2\text{Re}(l)}$. Therefore, the expression

$$(u;v)_l = (M^{l/2}u; M^{l/2}v)_0\,, \quad u, v \in H^l(V_n) \tag{12}$$

is a scalar product in the H^l space, and defines the topology of this space. Thus the expression

$$((u_1;u_2);(v_1;v_2))_{l+1,\,l} = (u_1;v_1)_{l+1} + (u_2;v_2)_l\,, \tag{13}$$

where $u_1, v_1 \in H^{l+1}$ and u_2, $v_2 \in H^l$, is a scalar product in the space $H^{l+1} \times H^l$, and defines its topology.

Theorem 2: (i) The operator T^{-1} is a toplinear isomorphism between $H^{l+1} \times H^l$ and $H^l \times H^{l-1}$ with inverse operator

$$T = \begin{pmatrix} -M^{-1}N & -M^{-1} \\ I & 0 \end{pmatrix}.$$

(ii) In $H^1 \times H^0$, and relative to the scalar product (13), T^{-1} is a skew-adjoint operator with domain $H^2 \times H^1$.

(iii) T is a compact operator in $H^l \times H^{l-1}$.

(iv) $\mathbb{R} \subset \rho(T^{-1})$, $\mathbb{R} - \{0\} \subset \rho(T)$.

(v) If $\lambda \in \mathbb{R} - \{0\}$, then

$$\|(T^{-1} - \lambda I)^{-1}\|_{1,0} \leq |\lambda|^{-1}\,, \quad \|(T - \lambda I)^{-1}\|_{1,0} \leq |\lambda|^{-1}\,.$$

Proof: (i) Let $u \equiv (u_1;u_2) \in H^{l+1} \times H^l$. From the following obvious inequalities:

$$\|T^{-1}u\|_{l,\,l-1}^2 \leq \|u_2\|_l^2 + 2\,\|u_1\|_{l+1}^2 + \|Nu_2\|_{l-1}^2 + 2\,\|Mu_1\|_{l-1} \cdot \|Nu_2\|_{l-1}\,,$$

$$\|Nu_2\|_{l-1}^2 \leq K\,\|u_2\|_l^2\,, \quad K \in \mathbb{R}$$

$$\|Mu_1\|_{l-1}^2 + \|Nu_2\|_{l-1}^2 \geq 2\,\|Mu_1\|_{l-1} \cdot \|Nu_2\|_{l-1}\,,$$

we find a real constant K' such that

$$\| T^{-1}u \|^2_{l,\,l-1} \leq K' \| u \|^2_{l+1,\,l},$$

This inequality proves the continuity of T^{-1}. The continuity of T follows by a similar argument on T or by a well-known corollary of the open mapping theorem.[38]

Clearly $T \cdot T^{-1} = T^{-1} \cdot T = I$.

(ii) Let $v \equiv (v_1; v_2) \in H^2 \times H^1$; from expression (12) and (v) of Proposition 1 we obtain

$$(T^{-1}u, v)_{1,0} + (u\,; T^{-1}v)_{1,0} = [(u_2; v_1)_1 - (Mu_1; v_2)_0 - (Nu_2; v_2)_0] + [(u_1; v_2)_1 - (u_2; Mv_1)_0 - (u_2; Nv_2)_0].$$

(iii) From the Rellich theorem,[34,35] the inclusion mapping $H^{l+1} \times H^l \times H^{l-1}$ is compact; T is continuous from $H^l \times H^{l-1}$ to $H^{l+1} \times H^l$, therefore, T is compact in $H^l \times H^{l-1}$.

(iv) The spectrum of T^{-1} (and thus that of T) is the same in all the $H^l \times H^{l-1}$ spaces. The zero point belongs to the continuous spectrum of T. The assertion then follows from (ii).

(v) By (ii) we obtain

$$\| (T^{-1} - \lambda I)u \|^2_{1,0} = \| T^{-1}u \|^2_{1,0} + \lambda^2 \| u \|^2_{1,0} \geq \lambda^2 \| u \|^2_{1,0}.$$

The assertion then follows from (iv) and a similar argument on T instead of T^{-1}. □

3. T^{-1} OPERATORS

In this section we introduce the complex-power[15,37,39] operators of the T^{-1} operator [of the $(-T^{-1})$ operator]. By means of these powers, we provide the $H^r \times H^{r-1}$ space with a scalar product compatible with its topology. This space then becomes a Hilbert space, and T^{-1} is a skew-adjoint operator in it.

(a) For $\mathrm{Re}(l) < 0$, the $(T^{-1})^l$ operator is defined[14,15] by the Dunford formula

$$(T^{-1})^l = \frac{i}{2\pi} \int_\Gamma \lambda^l (T^{-1} - \lambda I)^{-1} d\lambda, \tag{14}$$

where $\Gamma =]-\infty\,; -\epsilon] \cup C(\epsilon) \cup [-\epsilon\,; -\infty[$, and $C(\epsilon)$ is a circle with center at the origin and radius ϵ, such that it does not contain points of the spectrum of T^{-1}.

From (v) of Theorem 2, integral (14) converges in norm in $H^l \times H^{l-1}$. A simple calculation gives

$$(T^{-1})^l = \begin{pmatrix} -\Lambda(\lambda)(N+\lambda) & -\Lambda(\lambda) \\ \Lambda(\lambda)M & -\lambda\Lambda(\lambda) \end{pmatrix}, \tag{15}$$

where

$$\Lambda(\lambda) = [M + \lambda N + \lambda^2]^{-1}.$$

Expression (14) then becomes

$$(T^{-1})^l = -R(l) \begin{pmatrix} N & I \\ -M & 0 \end{pmatrix} + R(l+1)I,$$

where

$$R(l) = \frac{i}{2\pi} \int_\Gamma \lambda^l [M + \lambda N + \lambda^2]^{-1} d\lambda.$$

Now, we can adapt to the $R(l)$ operator the process developed by Seeley[14,15] for proving the pseudodifferentiability of certain elliptic pseudodifferential operators. Most importantly we obtain[39]

Theorem 3: For $\mathrm{Re}(l) < 0$, $R(l+1)$ and $R(l)$ are two pseudodifferential elliptic operators of order l, $l-1$, respectively. □

Corollary 4: The $(T^{-1})^l$ operator given by expression (14) is continuous from $H^r \times H^{r-1}$ to $H^{r-\mathrm{Re}(l)} \times H^{r-\mathrm{Re}(l)-1}$.

The complex-power operators $(T^{-1})^l$ of any order, are now defined in the usual fashion.[15,36,37]

Definition 5: For $l \in \mathbb{C}$, we set

$$(T^{-1})^l = (T^{-1})^k \cdot (T^{-1})^{l-k},$$

with k an integer and $-1 \leq \mathrm{Re}(l) - k < 0$. □

By standard techniques we can prove[15,38]

Theorem 6: The $(T^{-1})^l$ operator is a toplinear isomorphism from $H^r \times H^{r-1}$ to $H^{r-\mathrm{Re}(l)} \times H^{r-\mathrm{Re}(l)-1}$ and $(T^{-1})^s \cdot (T^{-1})^l = (T^{-1})^{s+l}$, $l, s \in \mathbb{C}$. □

We shall write T^{-l} instead of $(T^{-1})^l$. Proposition 8 below will justify this new notation.

From Theorem 6 if follows that the expression

$$\{u; v\}_{l,\,l-1} = (T^{-l+1}u; T^{-l+1}v)_{1,0}, \tag{16}$$

$u, v \in H^l \times H^{l-1}$, is a scalar product defining the topology of $H^l \times H^{l-1}$, and from (ii) of Theorem 2, we obtain

Theorem 7: In $H^l \times H^{l-1}$ equipped with the scalar product (16), the T^{-1} operator is skew-adjoint with domain $H^{l+1} \times H^l$. In general,

$$\{T^r u; v\}_{l,\,l-1} = \{u; (-T)^{\bar{r}} v\}_{l,\,l-1}$$

with $u, v \in D(T^r)$, $r \in \mathbb{C}$. □

(b) Balakrishnan has defined[37] complex powers of certain operators on Banach spaces by means of an expression different from (14). We shall next prove both definitions coincide and a proposition that we shall need later.

Proposition 8: If $-1 < \mathrm{Re}(l) < 0$,

$$(T^{-1})^l = \frac{\sin\pi(-l)}{\pi} \int_0^{+\infty} \mu^{-l-1} (T + \mu I)^{-1} T d\mu. \tag{17}$$

Remark: This expression is the Balakrishnan definition of the T^{-l} operator.

Proof: The change $\lambda \to \mu = \lambda^{-1}$ in expression (14) leads to

$$(T^{-1})^l = \frac{i}{2\pi} \int_{\hat{\Gamma}(\mu)} \mu^{l-1} (T - \mu I)^{-1} T d\mu, \tag{18}$$

where $\hat{\Gamma}(\mu) =]0; -1/\epsilon] \cup C(1/\epsilon) \cup [-1/\epsilon, 0[$. The integral on $C(\epsilon)$ in expression (14) tends to zero when $\epsilon \to 0$, because $\| (T^{-1} - \lambda I)^{-1} \|_{r,\,r-1} < K$, for $|\lambda| \leq \epsilon$ and $-1 < \mathrm{Re}(l) < 0$. Then in expression (18) the integral on the circle $C(1/\epsilon)$ tends to zero when $\epsilon \to 0$. Thus

$$\frac{2\pi}{i} (T^{-1})^l = \int_0^{-\infty} (|\mu| e^{i\pi})^{-l-1} (T - \mu I)^{-1} T d\mu - \int_0^{-\infty} (|\mu| e^{i\pi})^{-l-1} (T - \mu I)^{-1} T d\mu$$

$$= -2i \sin\pi(-l) \int_0^{\infty} \mu^{-l-1} (T + \mu I)^{-1} T d\mu. \quad \square$$

Consequently

Proposition 9:

$$(-T^2)^{1/2} = T^{1/2} \cdot (-T)^{1/2}. \tag{19}$$

Proof (Also see Chevalier[16]): From expression (17) we obtain

$$\begin{aligned} T^{1/2} + (-T)^{1/2} &= \frac{1}{\pi}\int_0^\infty \mu^{-1/2}(T+\mu I)^{-1} T d\mu \\ &\quad + \frac{1}{\pi}\int_0^\infty \mu^{-1/2}(-T+\mu I)^{-1}(-T) d\mu \\ &= \frac{1}{\pi}\int_0^\infty \lambda^{-3/4}(-T^2+\lambda)^{-1}(-T^2) d\lambda \\ &= \sqrt{2}\,(-T^2)^{1/4}. \end{aligned}$$

The last equality is by definition. Operator $(-T^2)$ is bounded self-adjoint positive in the space $H^l \times H^{l-1}$ and it therefore has a unique positive $\frac{1}{4}$ power, here defined by the Balakrishnan formula. Relation (19) follows from the squares of this realtion.

4. SYMPLECTIC AND COMPLEX STRUCTURES ON SPACES OF REAL SOLUTIONS

(a) Let $\{U(t)\}$ (or $\{U_t\}$) be the one-parameter group of unitary operators in $H^l \times H^{l-1}$ generated by T^{-1}.

The expression

$$u(t') = U(t'-t)u(t), \quad u(t) \in H^l \times H^{l-1},$$

solves the Cauchy problem for the equation and Cauchy data in $H^l \times H^{l-1}$. We can then regard $H^l \times H^{l-1}$ as a Hilbert space of solutions of Eq. (10).

Since T^{-1} is a real operator, U_t is also a real operator.

For $r > l$, the restriction to $H^r \times H^{r-1}$ of the unitary group $\{U_t\}$ coincides with the unitary group of operators in $H^r \times H^{r-1}$ generated by T^{-1}. Moreover,

Proposition 10: (i) In the space $X \equiv C^\infty(V_n) \times C^\infty(V_n)$, the bounded operator T^{-1} generates a uniform one-parameter equicontinuous group which coincides with the restriction to X of the unitary group $\{U_t\}$. We designate it, too, by $\{U_t\}$, and we have $d^k/dt^k U_t = T^{-k} U_t$ in the sense of the uniform topology on the space of bounded linear operators on X, $L_b(X;X)$.

(ii) In the space $X' \equiv \mathcal{D}'(V_n) \times \mathcal{D}'(V_n)$, the bounded operator T^{-1} generates a uniform one-parameter equicontinuous group whose restriction to $H^l \times H^{l-1}$ coincides with the unitary group $\{U_t\}$. We designate it, too, by $\{U_t\}$, and we have $d^k/dt^k U_t = T^{-k} U_t$ in the sense of the uniform topology on the space of bounded linear operators on X', $L_b(X';X')$.

Proof: (i) The set of operators $\{(I - n^{-1}T^{-1})^{-m},\ n = \pm 1, \pm 2, \cdots,\ m = 0, 1, \cdots\}$ is equicontinuous in X because $\|(I - n^{-1}T^{-1})^{-m} u\|_{l,l-1} \leq \|u\|$ and the topology of X is defined by the set of norms $\{\|\ \|_{l,l-1}\}$. Now by a theorem from Yosida,[38] p. 246, T^{-1} is the infinitesimal generator of an equicontinuous group for class C_0, given by

$$U_t = \underset{n\to\infty}{\text{s-lim}} \left[\exp(-nt)\exp(nt(I - n^{-1}T^{-1})^{-1})\right]. \tag{20}$$

Here s-lim means strong limit; the convergence, for t in bounded sets of R is uniform. T^{-1} is bounded in X and $L_b(X;X)$ is complete; thus the convergence in expression (20) is uniform on bounded sets of X [i.e., it converges in the $L_b(X;X)$ space] and uniform for t in bounded sets of R. Expression (20) can thus be differentiated, with respect to t, an arbitrary number of times, precisely by differentiating each term of the sequence.

(ii) The space X' is the inductive limit of the space $H^l \times H^{l-1}$ which are barreled and bornological. Thus X' is also a barreled and bornological space. The set of operators $\{(I - n^{-1}T^{-1})^{-m},\ n = \pm 1, \pm 2, \cdots,\ m = 0, 1, \cdots\}$ is bounded in $L(X';X')$ for the simple convergence topology; by the Banach–Steinhauss theorem (X' barreled), this set is equicontinuous in X'. Thus, T^{-1} is the infinitesimal generator of an equicontinuous group of class C_0 of operators in X'. The restriction of this group to $H^l \times H^{l-1}$ is obviously the unitary group $\{U_t\}$. T^{-1} is bounded in X', and $L_b(X';X')$ is complete because X' is bornological and complete. The final part of the statement follows from this and from an argument similar to that used in (i). □

(b) From the Lagrangian density,

$$L(u) = \nabla^\lambda u \cdot \nabla_\lambda u - m^2 u^2, \quad m > 0,$$

we obtain (1), as Lagrange equations, and

$$T^{\alpha\beta}(u) = 2\nabla^\alpha u \cdot \nabla^\beta u - g^{\alpha\beta} L(u),$$

as the components, in natural frames, of the energy–momentum tensor. Here $\nabla^\alpha = g^{\alpha\beta}\nabla_\beta$ and ∇_β are covariant derivatives in V_{n+1}.

Let u, v be two real solutions of Eq. (10) in $C^\infty(V_n \times \mathrm{R})$ and let Σ be a spacelike hypersurface of $V_n \times \mathrm{R}$. The following two quantities:

$$K(u) = \int_\Sigma T_{\alpha\beta}(u)\xi^\beta d\Sigma^\alpha, \tag{21}$$

$$b(u;v) = \int_\Sigma (u\partial_\alpha v - \partial_\alpha u \cdot v) d\Sigma^\alpha, \tag{22}$$

are independent of the choice of the spacelike hypersurface Σ. Expression (21) is the definition of the energy[40] associated with a solution of Eq. (1) in a stationary space–time. Expression (22) is a bilinear skew-symmetric form defining the weakly-symplectic structure[41] of the dynamical system described by Eq. (10).

With $\Sigma \equiv V_n \times \{t\}$, a straightforward calculation leads to the following

Proposition 11: Let $u \equiv (u_1; u_2)$, $v \equiv (v_1; v_2)$ be two elements in $H^1 \times H^0$. Then:

(i) $K(u) = (u;u)_{1,0}$, $K(U_t u) = K(u)$,

(ii) $b(u,v) = (Tu, v)_{1,0} = (u_1; v_2)_0 - (u_2; v_1)_0 - (N u_1; v_1)_0$,

(iii) $b(U_t u, U_t v) = b(u;v)$. □

(c) Following Segal[6] (see also Chernoff and Marsden[41]), let us consider the unitary part J of the polar decomposition[42,43] of T in the real Hilbert space $H^l \times H^{l-1}$,

$$J = T(-T^2)^{-1/2} = T^{1/2} \circ (-T)^{-1/2} .$$

Proposition 9 justifies the second equality.

The proofs of the following two theorems are straightforward.

Theorem 12: On each real space $H^l \times H^{l-1}$ of real solutions of Eq. (10), the operator J defines a complex structure, i.e., (i) J is a toplinear automorphism of $H^l \times H^{l-1}$; (ii) $J^2 = -I$; (iii) $U_t J = J U_t$. □

Theorem 13: On the real space $H^l \times H^{l-1}$,

(i) The skew-symmetric form $\omega_l(u; v) = \{Ju; v\}_{l,l-1}$ is strongly symplectic.[41]

(ii) The skew-symmetric form $b_l(u; v) = \{Tu; v\}_{l,l-1}$ is weakly symplectic.[41]

(iii) The restriction of ω_l to the $H^{l+1/2} \times H^{l-1/2}$ subspace is $b_{l+1/2}$. □

A uniqueness theorem for the operator J now follows.

Theorem 14: On the real Hilbert space $H^l \times H^{l-1}$ of real solutions of Eq. (10) there is a unique complex structure operator J' satisfying: (i) $J'T = TJ'$; (ii) $\omega_l(u; J'u) > 0$ if $u \neq 0$. Hence $J' = J$.

Proof (also see Chevalier[16] and Combet[19]): From (i)

$$(I - JJ')(J - J') = 0 .$$

But $(I - JJ')$ is bound-symmetric, and from (ii) it has a strictly positive lower bound; then for $c < 2\|I - JJ'\|^{-1}_{l,l-1}$ the series

$$c \sum_{n=0}^{\infty} [I - c(I - JJ')]^n = (I - JJ')^{-1}$$

converges in norm in $H^l \times H^{l-1}$ to the inverse of $(I - JJ')$ which is a bounded positive symmetric operator. (See Ref. 43, pp. 263.) Hence $J = J'$. □

5. G AND G^1 KERNELS

(a) Let us write $U^1_t = U_t J = J U_t$,

$$U^1(t) = \begin{pmatrix} U^1_{11}(t) & U^1_{12}(t) \\ U^1_{21}(t) & U^1_{22}(t) \end{pmatrix},$$

and a similar matrix expression for U_t. $U^1_{ij}(t)$ and $U_{ij}(t)$ are continuous operators from $C^\infty(V_n)$ to $C^\infty(V_n)$ because they are continuous from $H^l(V_n)$ to $H^{l+j-i}(V_n)$ for any l.

From the kernels theorem, these operators determine the regular[44,45] distribution kernels $G^1_{ij}(t), G_{ij}(t)$, defined by linearity and contunity from the expression

$$\langle G^1_{ij}(t); \varphi \otimes \psi \rangle_{V_n \times V_n} = -(U^1_{ij}(t)\varphi; \bar{\psi})_0 ,$$

where $\varphi, \psi \in C^\infty(V_n)$, and similarly for $G_{ij}(t)$ and $U_{ij}(t)$.

Proposition 15: The mapping $G^1_{ij}(\cdot)$,

$$\mathbf{R} \to C^\infty(V_n) \hat{\otimes} \mathcal{D}'(V_n) ,$$

$$t \to G^1_{ij}(t) ,$$

is of class C^∞. The same is true of the $G_{ij}(\cdot)$ mapping.

Proof: From Proposition 10, the mapping $U^1_{ij}(\cdot): \mathbf{R} \to L_b(C^\infty(V_n); C^\infty(V_n))$ is of class C^∞. The proposition now follows by using the isomorphism $L_b(C^\infty(V_n); C^\infty(V_n)) \approx C^\infty(V_n) \hat{\otimes} \mathcal{D}'(V_n)$. □

Let φ be a mapping in $C^\infty(\mathbf{R}^2; C^\infty(V_n^2))$.

Corollary 16: The mapping on $\mathbf{R}^2$ with values in C defined by

$$(t'; t) \to \langle G^1_{ij}(t' - t), \varphi(t', t) \rangle_{V_n \times V_n}$$

is of class C^∞ and the Leibniz formula

$$\partial^k_{t'} \langle G^1_{ij}(t' - t); \varphi(t', t) \rangle_{V_n \times V_n}$$

$$= \sum_{l + l' = k} C_l \langle \partial^l_t G^1_{ij}(t' - t); \partial^{l'}_{t'} \varphi(t', t) \rangle_{V_n \times V_n} ,$$

follows. Similar properties hold for the $G_{ij}(\cdot)$ mapping.

Let ψ be a mapping of $C^\infty(\mathbf{R}; C^\infty(V_n))$.

Corollary 17: The mapping on $\mathbf{R}^2$ with values in $C^\infty(V_n)$ defined by

$$(t', t) \to \langle G^1_{ij}(t' - t); \psi(t', t) \rangle_{V_n \times V_n}$$

is of class C^∞, and the Leibniz formula for derivatives with respect to t or t' holds. □

(b) Let ϕ be a mapping in $C^\infty((V_n \times \mathbf{R}^2))$. Then, the mapping $\hat{\phi}$ on $\mathbf{R}^2$ with values in $C^\infty(V_n^2)$ canonically determined by ϕ is in $C^\infty(\mathbf{R}^2, C^\infty(V_n^2))$. Conversely, let $\hat{\phi}$ be a mapping in the latter space. Then the mapping $\hat{\phi}$ on $(V_n \times \mathbf{R})^2$, with values in C canonically defined by $\hat{\phi}$ is in $C^\infty((V_n \times \mathbf{R})^2)$. From this, and Corollaries 16 and 17, we obtain

Proposition 18: For $\phi \in \mathcal{D}((V_n \times \mathbf{R})^2)$ the expression

$$\langle \theta^{1k}_{ij}; \phi \rangle = \int_{\mathbf{R}^2} \left\langle \frac{\partial^k}{\partial t'^k} G^1_{ij}(t' - t); \frac{1}{g^{00}} \otimes \frac{1}{g^{00}} \hat{\phi}(t', t) \right\rangle_{V_n \times V_n} dt \cdot dt' , \tag{23}$$

defines a regular distribution kernel on $V_n \times \mathbf{R}$. Locally we have

$$\theta^{1k}_{ij}(t'\xi'; t\xi) = \partial^k_{t'} \theta^{10}_{ij}(t'\xi'; t\xi) .$$

The same statement holds for the $G_{ij}(t)$ and θ^k_{ij} kernels. □

The uniqueness of the solution of the Cauchy problem for Eq. (10) [see expression (2)] means that the θ^0_{12} kernel is the propagator of that equation.

(c) We shall see that the kernel $G^1 \equiv \theta^{10}_{12}$ satisfies the properties required in the Introduction. Also see Chevalier[16] and Combet.[19]

Proposition 19: The kernel G^1 is symmetric and the kernel G is skew-symmetric. Both are real.

Proof: The operators U_t, $t \in \mathbf{R}$, are real; the kernel G^1 and G are thus also real. The relation $U^{1*}_t = -U^1_{-t}$ and $T^{-1}U_t = U_t T^{-1}$ yield, in particular

$$(U^1_{12}(-t)\varphi; \psi)_1 = -(\varphi; U^1_{21}(t)\psi)_0$$

$$U^1_{21}(t) = -U^1_{12}(t)M ,$$

$\varphi, \psi \in C^\infty(V_n)$. Therefore,

$$(U^1_{12}(t)\varphi; M\psi)_0 = (\varphi, U^1_{21}(-t)M\psi)_0 ,$$

thus $(G^1_{12}(t))' = G^1_{12}(-t)$. (' stands for transposed kernel). Using this relation in definition (23) of the G^1 kernel,

we can find the symmetry of this kernel. Similarly we can find the skew-symmetry of the G kernel. □

From the relations

$$\left(T^{-1} - \frac{d}{dt}\right) U^1(t) = 0, \quad T^{-1}U^1(t) = U^1(t)T^{-1},$$

we obtain, in particular

$$(M \hat{\otimes} I)G^1_{12}(t'-t) + (N \hat{\otimes} I)\frac{d}{dt'} G^1_{12}(t'-t) + \frac{d^2}{dt'^2} G^1_{12}(t'-t) = 0, \tag{24}$$

$$(I \hat{\otimes} M)G^1_{12}(t'-t) + (I \hat{\otimes} N)\frac{d}{dt} G^1_{12}(t'-t) + \frac{d^2}{dt^2} G^1_{12}(t'-t) = 0, \tag{25}$$

where the time derivatives are taken in the sense of Proposition 15. Then,

Proposition 20: The kernels G and G^1 are solutions of the Klein–Gordon equation in a distributional sense, i.e.,

$$((\Delta_{n+1} - m^2) \hat{\otimes} I)G^1 = 0, \quad ((I \hat{\otimes} (\Delta_{n+1} - m^2))G^1 = 0$$

and a similar expression with G in place of G^1. □

(d) Let

$$(-T^2)^{1/2} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

be a matrix expression of the operator $(-T^2)^{1/2}$ on every $H^l \times H^{l-1}$ space. Then

$$J = \begin{pmatrix} -C & -D \\ MA+NC & MB+ND \end{pmatrix}.$$

Let $\hat{A}, \hat{B}, \ldots, \hat{M}, \cdots$ be the regular kernels on V_n determined by the $A, B, \ldots, M, \cdots$ operators. Given regularity, the convolution (*) of a finite number of kernels, in the sense of Volterra–Schwartz,[45] pp. 114–120, included among $\hat{A}, \hat{B}, \ldots, \hat{M}, G^1_{ij}(t)$, $G_{ij}(t)$, is well defined and associative.

Relations $U^1(t) = U(t)J = JU(t)$ and $U^1(t''-t')\cdot U^1(t'-t) = -U(t''-t)$ give, in particular,

$$G^1_{12}(t'-t) = G_{12}(t'-t) * \hat{M} * \hat{B} - \frac{d}{dt'} G_{12}(t'-t) * \hat{D},$$

$$G_{12}(t''-t) = G^1_{12}(t''-t') * \frac{d}{dt'} G^1_{12}(t'-t) + G^1_{12}(t''-t') * \hat{N} * G^1_{12}(t'-t) - \frac{d}{dt'} G^1_{12}(t''-t') * G^1_{12}(t'-t).$$

The former expression yields the G^1 kernel. The latter relation is the fundamental Lichnerowicz relation (4).

The first proof of this result in static space-times has been given by Combet.[2,19] We have thereby proved

Proposition 21: The calculated distribution kernel G^1 on a stationary space-time, satisfies the conditions required in the Introduction. Conditions for uniqueness are given in Theorem 14. The complex structure operator J [expression (5)] is given by the following matrix of kernels:

$$\begin{pmatrix} (I \hat{\otimes} N)G^1_{12}(t'-t) - \frac{d}{dt'} G^1_{12}(t'-t) & -G^1_{12}(t'-t) \\ (I \hat{\otimes} N)\frac{d}{dt'} G^1_{12}(t'-t) - \frac{d^2}{dt'^2} G^1_{12}(t'-t) & -\frac{d}{dt'} G^1_{12}(t'-t) \end{pmatrix}.$$

□

6. POSITIVE AND NEGATIVE FREQUENCY SPACES

On the complex space $H^l \times H^{l-1}$ of solutions, let us consider the operators[8]

$$\oplus = \tfrac{1}{2}(I + iJ), \quad \ominus = \tfrac{1}{2}(I - iJ).$$

Then

Proposition 22: (i) $\oplus$ and $\ominus$ are complementary projectors on $H^l \times H^{l-1}$ and

$$\oplus \cdot U_t = U_t \cdot \oplus, \quad \ominus \cdot U_t = U_t \cdot \ominus.$$

(ii) If $h \in H^l \times H^{l-1}$ is real,

$$2\{\oplus h; \oplus h\}_{l,l-1} = 2\{\ominus h; \ominus h\}_{l,l-1} = \{h; h\}_{l,l-1}.$$ □

By definition the closed subspaces

$$E^{\oplus}_l = \oplus(H^l \times H^{l-1}), \quad E^{\ominus}_l = \ominus(H^l \times H^{l-1})$$

are the Hilbert spaces of positive and negative frequency parts respectively.

7. THE SPECTRUM OF THE ENERGY OPERATOR

In this section we find the Fourier expansion for frequency solutions and kernels G, G^1, analogous to that given in the Minkowski space by means of the Fourier transform.

Basically we have exploited the compactness property of the inverse H^{-1} of the energy operator $H = i^{-1}T^{-1}$. See Theorems 2 and 7.

(a) Let us write Eq. (10) in the Schrödinger form

$$\frac{1}{i}\frac{d}{dt} u(t) = Hu(t), \quad H = i^{-1}T^{-1}. \tag{26}$$

H is the Hamiltonian operator[41] of the dynamical system defined by Eq. (10) and $H^l \times H^{l-1}$ is its "phase space." By Theorem 7, H is self-adjoint in $H^l \times H^{l-1}$ with respect to the scalar product (16), and by Theorem 2, H^{-1} is compact. Therefore, the spectrum of H is real, countable, and does not have any finite accumulation point; the proper subspaces are of finite dimension, and there is an orthogonal basis of $H^l \times H^{l-1}$ formed by eigenfunctions of H.

Proposition 23: Let l_p, $p \in \mathbb{Z}$ be any eigenvalue of H. The eigenspace E_{l_p} is a subspace of finite dimension of $C^\infty(V_n) \times C^\infty(V_n)$. $-l_p$ is also an eigenvalue of H. Complex conjugate elements of an orthonormal basis of E_{l_p} form (in $H^l \times H^{l-1}$) an orthonormal basis of E_{l_p}.

Proof: (i) Let $\{u_p\}$ be an orthonormal basis of $H^l \times H^{l-1}$ such that,

$$Hu_p = l_p u_p, \quad l_p \in \mathbb{R}, \quad p \in \mathbb{Z}$$

with $u_p \equiv ((u_p)_1; (u_p)_2)$. Thus

$$(M + il_pN - l_p^2)(u_p)_2 = 0, \quad (u_p)_2 = il_p(u_p)_1$$

Since $(u_p)_1$ and $(u_p)_2$ are in the kernel of an elliptic differential operator, both are in $C^\infty(V_n)$. □

(ii) Clearly we have: $H\bar{u}_p = -l_p\bar{u}_p$ and $(\bar{u}_p; \bar{u}_q)_{1,1-1} = \delta_{pq}$.

(b) Let $y_0 = \sum(\lambda_p u_p + \lambda_{\bar{p}} u_{\bar{p}})$, $\lambda_p, \lambda_{\bar{p}} \in \mathbb{C}$, $p, \bar{p} \in \mathbb{N}$ be, any element of $H^1 \times H^{1-1}$. The solution of Eq. (26) corresponding to the Cauchy data y_0 at time $t = 0$ is given by

$$y(t) = \sum_{p>0} (e^{il_pt}\lambda_p u_p + e^{-il_pt}\lambda_{\bar{p}} u_p), \quad (u_{\bar{p}} \equiv \bar{u}_p).$$

Thus the evolution operator U_t takes the form

$$U_t = \sum_{p>0} (e^{-il_pt} u_{\bar{p}} \otimes u_p + e^{il_pt} u_p \otimes u_{\bar{p}}), \tag{27}$$

where the series converges in the strong topology on bounded linear operators in $H^1 \times H^{1-1}$.

(c) Let us write

$$y^{\oplus}(t) = \sum_{p>0} e^{il_pt}\lambda_p u_p, \quad y^{\ominus}(t) = \sum_{p>0} e^{-il_pt}\lambda_{\bar{p}} u_{\bar{p}}, \tag{28}$$

$$y^1(t) = i(y^{\ominus}(t) - y^{\oplus}(t)). \tag{29}$$

Thus we obtain $y^1(t) = \tilde{U}^1(t)y_0$, where

$$\tilde{U}^1(t) = -i \sum_{p>0} (e^{-il_pt} u_{\bar{p}} \otimes u_p - e^{il_pt} u_p \otimes U_{\bar{p}}), \tag{30}$$

and the convergence is just as in expression (27).

Proposition 24: $Ju_p = -i\mathrm{sgn}l_p \cdot u_p$.

Proof:

$$(T^{-1})u_p = il_p u_p \Longleftrightarrow (-T^2)u_p = \frac{1}{l_p^2} u_p.$$

From the spectral theorem and the uniqueness of the positive square root of a positive self-adjoint operator, we obtain

$$(-T^2)^{1/2} u_p = \frac{1}{|l_p|} u_p.$$

The proposition now follows. □

The next proposition is now obvious.

Proposition 25: We have $\tilde{U}^1(t) = U^1(t)$ and expressions (28) and (29) give the frequency solutions defined in Proposition 22. □

(d) A simple calculation on expressions (27) and (30) shows that

$$G_{12}(t'-t) = i \sum_{p>0} l_p (e^{-il_pt}(\bar{u}_p)_1 \otimes e^{il_pt'}(u_p)_1 - e^{il_pt}(u_p)_1 \otimes e^{-il_pt'}(\bar{u}_p)_1),$$

$$G^1_{12}(t'-t) = \sum_{p>0} l_p (e^{-il_pt}(\bar{u}_p)_1 \otimes e^{il_pt'}(u_p)_1 + e^{il_pt}(u_p)_1 \otimes e^{-il_pt'}(\bar{u}_p)_1),$$

where the series converges in $\mathcal{D}'(V_n \times V_n)$ for fixed t and t'.

The function

$$f_p = e^{il_p(\cdot)}(u_p)_1$$

is in $C^\infty(V_n \times \mathbb{R})$. By Proposition 18, the series

$$G = i \sum_{p>0} l_p (\bar{f}_p \otimes f_p - f_p \otimes \bar{f}_p),$$

$$G^1 = \sum_{p>0} l_p (\bar{f}_p \otimes f_p + f_p \otimes \bar{f}_p),$$

are convergent in $\mathcal{D}'((V_n \times \mathbb{R})^2)$ and define the propagator and the Lichnerowicz kernel of the Klein–Gordon equation.

8. CONCLUSIONS

The space–times considered in this paper are stationary $V_n \times \mathbb{R}$ with compact space-section manifold without boundary.

In any adapted atlas, the Klein–Gordon equation becomes a time-independent evolution equation in spaces $H^1 \times H^{1-1}$. The complex powers of the T^{-1} operator have been defined. By means of these powers, a scalar product compatible with the topology has been introduced in the $H^1 \times H^{1-1}$ space. The energy operator $H = i^{-1}T^{-1}$ is self-adjoint in this space relative to the said scalar product. The Klein–Gordon equation has then been solved by the one-parameter group $\{U_t\}$ of unitary operators generated by T^{-1} in the Hilbert space $H^1 \times H^{1-1}$. In the spaces $C^\infty(V_n) \times C^\infty(V_n)$ and $\mathcal{D}'(V_n) \times \mathcal{D}'(V_n)$ this group is uniform equicontinuous. On each space $H^1 \times H^{1-1}$ a canonical complex structure has been defined. These structures have been simultaneously defined by a unique distribution kernel G^1, which satisfies the Lichnerowicz requirements. The Hilbert spaces of frequency solutions have thus been constructed by means of this kernel.

The spectrum of the energy operator $H = i^{-1}T^{-1}$ is real and symmetric with respect to the origin of $\mathbb{R}$. The only nonvoid part of the spectrum is the point spectrum, and it is countable, with no finite accumulation point. This property has been exploited to give the definition of frequency solutions in a form similar to that given in the Minkowski space by means of the Fourier transform.

[1] Y. Choquet-Bruat, in *Battelle Rencontres 1967. Lectures in Mathematics and Physics* (Benjamin, New York, 1968).
[2] E. Combet, *Solutions élémentaires des dalembertiens généralisés. Memorial des Sciences Mathématiques*, Fasc. CLX. (Gauthier-Villars, Paris, 1965), thesis.
[3] A. Lichnerowicz, "Propagateurs et commutateurs en Relativité Générale," I.H.E.S. publication no. 10 (1961).
[4] J. Leray, "Hyperbolic Differential Equations," Lectures, Institute for Advanced Study, Princeton (1954).
[5] A. Lichnerowicz, Bull. Soc. Math. France, **92**, 11 (1964).
[6] I. Segal, Seminaire Physique Mathématique, Collège de France (1964–65).
[7] I. Segal, J. Math. Phys. **1**, 468 (1960).
[8] A. Lichnerowicz, in *Relativity Groups and Topology, Les Houches 1963*, edited by DeWitt and DeWitt (Science, New York, 1964).
[9] D. Kastler, *Introduction à l'électrodynamique quantique* (Dunod, Paris, 1961).
[10] S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Harper and Row, New York, 1962).
[11] L. Bel, Gen. Rel. Grav. J. **1**, 334 (1971).
[12] V. Fock, *The Theory of Space–Time and Gravitation* (Pergamon, New York, 1962).
[13] F. Hennequin, thesis, University of Paris, Paris, 1950.
[14] R. Seeley, *Topics in Pseudo-Differential Operators* (C.I.M.E., Edizioni Cremonese, Rome, 1969).

[15]R. Seeley, Am. Math. Soc. Proc. Symp. Pure. Math., **10**, 288 (1968).
[16]M. Chevalier, J. Math. Pures Appl. **53**, 223 (1974).
[17]G. Rideau, C. R. Acad. Sci. Paris A **260**, 2719 (1965).
[18]I. Guelfand and G. Chilov, *Les distributions* (Dunod, Paris, 1962), Vol. 1.
[19]E. Combet, Seminaire Physique Mathématique, Collège de France, 1965.
[20]E. Combet, C. R. Acad. Sci. Paris A **262**, 1077 (1965).
[21]L. Hörmander and J. Duistermaat, Acta. Math. **128**, 183 (1972).
[22]S. Fulling, thesis, Princeton University, New Jersey, 1972.
[23]L. Parker, Phys. Rev. **183**, 5, 1057 (1969).
[24]L. Parker, Phys. Rev. D **3**, 346 (1971).
[25]S. Hawking, Commun. Math. Phys. **43**, 199 (1975).
[26]S. Hawking, Nature **248**, 30 (1974).
[27]R. Ruffini and N. Deruelle, Phys. Lett. B **52**, 434 (1974); **57**, 248 (1975).
[28]R. Ruffini and T. Damour, Phys. Rev. Lett. **35**, 463 (1975).
[29]T. Nakamura, R. Ruffini, and H. Sato, "Absortion of Massive Scalar Field by a Charged Black Hole," R.I.F.P. 231, Kyoto University.
[30]B. deWitt, "Quantum Field Theory in Curved Space–Time," preprint, University of Texas (1973).
[31]S. Deser and B. Zumino, Phys. Lett. B **62**, 335 (1976).
[32]B. Zel'dovich, Sov. Phys.—JETP Lett. **12**, 307 (1970).
[33]B. Zel'dovich and Starobinsky, Sov. Phys.—JETP Lett. **34**, 1159 (1972).
[34]L. Hörmander, *Linear Partial Differential Operators* (Springer, New York, 1963).
[35]R. Palais, Ed., *Seminar on the Atiyah—Singer Index Theorem*, Ann. of Math. Stud. *n°* **57** (Princeton U.P., Princeton, N. J., 1965).
[36]T. Burak, Ann. Scuola Normale Super. Pisa, **22**, 113 (1968).
[37]A. V. Balakrishnan, Pac. J. Math. **10**, 419 (1960).
[38]K. Yosida, *Functional Analysis* (Springer, New York, 1965).
[39]C. Moreno, thesis, University of Lyon, Lyon, France, 1975.
[40]A. Trautman, in *Gravitation: An Introduction to Current Research* (Wiley, New York, 1962).
[41]P. Chernoff and J. Marsden, *Properties of Infinite Dimensional Hamiltonian Systems*, Lectures Notes in Mathematics, no. 425 (Springer, New York, 1974).
[42]N. Dunford and J. Schwartz, *Linear Operators* (Interscience, New York, 1963), Vol. II.
[43]F. Riesz and B. SZ. Nagy, *Leçons d'analyse fonctionnelle* (Gauthier-Villars, Paris, 1968).
[44]L. Schwartz, *Théorie des distributions* (Hermann, Paris, 1966).
[45]L. Schwartz, Ann. Inst. Fourier, extraits des tomes VII et VIII (1959).

A note on the range of the applicability of the Ornstein–Zernike theory in the van der Waals model

Marek Napiórkowski

Institute of Theoretical Physics, Warsaw University, Hoza 69, 00681 Warsaw, Poland
(Received 9 December 1976)

We show that the condition determining the range of the applicability of the Ornstein–Zernike theory obtained previously by Hemmer for the van der Waals model is too strong. The weaker condition obtained by us is in agreement with the existing results for different physical systems, and it requires the temperature to satisfy the following inequality: $(T-T_c)/T_c \gg (\gamma d)^6$, where γd denotes the ratio of the short-range part of the potential to the range of the long-range part and T_c is the critical temperature predicted in the van der Waals model.

I. INTRODUCTION

In this paper we consider the van der Waals model of a fluid in which the interparticle potential $\Phi(r)$ is separated into a short-range repulsive part and a long-range attractive part,[1]

$$\Phi(r) = \Phi^{hc}(r) + \Phi^{attr}(r) \tag{1.1}$$

where $\Phi^{hc}(r)$ is the short-range part of the potential (of range d) while $\Phi^{attr}(r) = -\gamma^3\varphi(\gamma r)$ is the long-range part of the potential of range γ^{-1}. It is a well-known fact that the $\gamma \to 0$ limit of the van der Waals fluid pressure p is given by the following equation of state[1,2]:

$$p(\rho, T) = p^{hc}(\rho, T) + \tfrac{1}{2}\rho^2\tilde{\Phi}(0) \tag{1.2}$$

where p^{hc} is the pressure of the hard core fluid, ρ is the number density, and $\tilde{\Phi}(k)$ denotes the Fourier transform of the long-range part of the potential. The notation used in this paper is exactly the same as that introduced in Ref. 1.

The problem of determining the region near the critical point in which the two-particle correlation function can be approximated by the Ornstein–Zernike (O.Z.) function has been discussed for different physical systems.[3–7] The common conclusion obtained as a result of this analysis is that the O.Z. theory holds whenever the following condition is fulfilled:

$$\epsilon \gg l^3, \tag{1.3}$$

where l^{-1} denotes the range of the interparticle potential (measured in proper units) and $\epsilon^2 = (T - T_c)/T_c$.

The above problem has been also investigated by Hemmer[1] in the frame of the van der Waals model. The condition for the range of the applicability of the O.Z. theory was found to be

$$\epsilon \gg \gamma d \tag{1.4}$$

which disagrees with the previous results (1.3) in the sense that it represents a much stronger requirement (we always consider the $\gamma \to 0$ limit). It appears however that this exceptional status of the van der Waals model is apparent and results entirely from disregarding the behavior of the hard core system near the critical point of the fluid. We thus show that also in the van der Waals model the condition (1.3), i.e.,

$$\epsilon \gg (\gamma d)^3 \tag{1.5}$$

determines the range of the applicability of the O.Z. theory. This is done in Sec. II where the first two terms in the $\xi = (\gamma d)^3$ expansion of the two-particle correlation function $g(r)$ in the far range (i.e., for $r = O(\gamma^{-1})$) are considered and then the condition (1.5) is recovered. In Sec. III we analyze the contributions to the two-particle correlation function which are of an arbitrary order in ξ and we outline the proof of the validity of condition (1.5) for the general case.

II. THE ξ EXPANSION NEAR THE CRITICAL POINT AND THE O.Z. THEORY

We are interested in the behavior of the two-particle correlation function $g(r)$ in the far range $[r = O(\gamma^{-1})]$ in the vicinity of the critical point. We shall always approach the critical point along the critical isohore in the one-phase region and the parameter ϵ defined as[1]:

$$\epsilon^2 = 1 + \eta\tilde{\Phi}(0) = \left(\frac{\partial p}{\partial \rho}\right)_T \Big/ \left(\frac{\partial p^{hc}}{\partial \rho}\right)_T \sim T - T_c \tag{2.1}$$

will measure the distance from the critical point.

In this section we consider the first two terms in the ξ expansion of $g(r)$, i.e., $g^{(1)}$ and $g^{(2)}$. As is shown in Ref. 1 these two terms are given in the $\gamma \to 0$, $\epsilon \to 0$ limit by the following expressions (we always use the Fourier space representation):

$$\tilde{g}^{(1)}(k) \simeq \frac{6k_BT}{|\tilde{\Phi}(0)|}\,\frac{1}{k^2 + 6\epsilon^2}, \tag{2.2}$$

$$\tilde{g}^{(2)}(k) \simeq \left[9k_BT\left(\frac{\partial \eta}{\partial \rho}\right)_T^2\right]\frac{2\arctan(k/2\sqrt{6}\,\epsilon)}{\pi k(k^2 + 6\epsilon^2)^2}\,\gamma^3 + 18k_BTD\left(\frac{\partial^2\eta}{\partial\rho^2}\right)_T\left[\frac{1}{(k^2 + 6\epsilon^2)^2} - \frac{1}{36}\right], \tag{2.3}$$

where

$$D = -\left(\frac{\gamma}{2\pi}\right)^3 k_BT\eta^2\int d\mathbf{k}\,\frac{\tilde{\Phi}(k)}{1 + \eta\tilde{\Phi}(k)}.$$

If near the critical point we disregard the behavior of the hard core system represented in (2.3) by the functions η, $(\partial\eta/\partial\rho)_T$, and $(\partial^2\eta/\partial\rho^2)_T$, then for the wave vectors of the order of ϵ ($\mathbf{k} = \mathbf{q}\epsilon$) we obtain in the $\gamma \to 0$, $\epsilon \to 0$ limit the following result:

$$\tilde{g}^{(1)} \sim \epsilon^{-2}, \tag{2.4}$$

$$\tilde{g}^{(2)} \sim \gamma^3\epsilon^{-5}. \tag{2.5}$$

Thus the conclusion derived in Ref. 1 is that the second order terms are small in comparison with the first order term (O.Z. theory) if (1.4) is satisfied.

It appears however that the above line of reasoning is

 Copyright © 1977 American Institute of Physics

misleading because the function

$$\left(\frac{\partial\eta}{\partial\rho}\right)_T=\left(\frac{\partial p^{hc}}{\partial\rho}\right)_T^{-2}\left[\left(\frac{\partial p^{hc}}{\partial\rho}\right)_T-\rho\left(\frac{\partial^2 p^{hc}}{\partial\rho^2}\right)_T\right] \tag{2.6}$$

vanishes at the critical point. This follows from the equations determining the critical point

$$\left(\frac{\partial p}{\partial\rho}\right)_T=0,\quad \left(\frac{\partial^2 p}{\partial\rho^2}\right)_T=0$$

for the van der Waals equation of state (1.2).

The function $(\partial\eta/\partial\rho)_T$ enters into (2.3) through the fluctuation—dissipation theorem when considering the lowest order term in the γ expansion of the Fourier transform of μ_3^{hc},

$$\begin{aligned}&\int d\mathbf{r}\exp(i\mathbf{k}_2\cdot\mathbf{r}\gamma)\int d\mathbf{R}\exp(i\mathbf{k}_3\cdot\mathbf{R}\gamma)\mu_3^{hc}(0,\mathbf{r},\mathbf{R};\rho)\\&\quad=\int d\mathbf{r}\int d\mathbf{R}\mu_3^{hc}(0,\mathbf{r},\mathbf{R};\rho)[1+\gamma(i\mathbf{k}_2\cdot\mathbf{r}+i\mathbf{k}_3\cdot\mathbf{R})]\\&\quad\quad-\gamma^2\tfrac{1}{6}\int d\mathbf{r}\int d\mathbf{R}\,[k_2^2r^2+k_3^2R^2+6(\mathbf{k}_2\cdot\mathbf{r})(\mathbf{k}_3\cdot\mathbf{R})]\\&\quad\quad\times\mu_3^{hc}(0,\mathbf{r},\mathbf{R};\rho)+O(\gamma^3).\end{aligned} \tag{2.7}$$

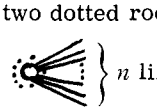
Since this lowest order term vanishes at the critical point and the linear terms also vanish for the symmetry reasons[8] then the first nonvanishing term is proportional to $\gamma^2\epsilon^2$ (after the change of variables $\mathbf{k}=\mathbf{q}\epsilon$). Thus near the critical point one has to use the expression proportional to $\gamma^2\epsilon^2$ instead of $(\partial\eta/\partial\rho)_T$. This means that, since the functions η and $(\partial^2\eta/\partial\rho^2)_T$ have finite limits at the critical point while $D\sim\gamma^3\epsilon$ in the $\gamma\to 0$, $\epsilon\to 0$ limit, the first term on the right-hand side of (2.3) is small in comparison with the second term.

Now the condition for $\tilde{g}^{(2)}$ to be small in comparison with $\tilde{g}^{(1)}$ reduces to the condition for the second term on the right-hand side of (2.3) to be small in comparison with (2.2) and this exactly takes the form (1.5).

III. THE GENERAL CASE

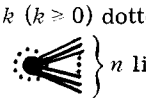
According to Ref. 1 the two-particle correlation function $g(r)$ is represented by the sum of all distinct irreducible generic 2-graphs. After the renormalization to all orders in the density is performed[1] each graph can be built from:

(a) two dotted root points



representing the functions $\tilde{\mu}_n^{hc}$, $n\geq 2$,

(b) k ($k\geq 0$) dotted field points



representing the functions $\tilde{\mu}_n^{hc}$, $n\geq 3$;

(c) p ($p\geq 1$) chain bonds—c—representing the functions

$$\tilde{C}(k)=-\frac{\tilde{\Phi}(k)}{k_BT+\tilde{\Phi}(k)\tilde{\mu}_2^{hc}(k)}$$

It is easy to check that the Fourier transforms $\tilde{\mu}_n^{hc}$ have finite limits at the critical point when $\gamma\to 0$, $\epsilon\to 0$, and $\mathbf{k}=\mathbf{q}\epsilon$ except $\tilde{\mu}_3^{hc}$ which is proportional to $\gamma^2\epsilon^2$ in the above limit. Moreover $\tilde{C}\sim\epsilon^{-2}$ and each wave vector integration appearing in the graph gives, after the change of variables, a factor $\gamma^3\epsilon^3$. It is also shown in Ref. 1 that the nth order terms in the ξ expansion of $g(r)$ are represented by the graphs in which there are n chain bonds more than the number of the dotted field points.

In this section we compare the γ and ϵ dependence of the expression represented by the graph of an arbitrary order in ξ with the O. Z. function. We shall show that if condition (1.5) is satisfied then the O. Z. function represents the dominant contribution to $g(r)$ in the limit $\gamma\to 0$, $\epsilon\to 0$. To this aim we apply the induction. It is shown in Sec. II that the second order terms are small in comparison with the O. Z. term if (1.5) is satisfied; now we assume that the nth order terms are small in comparison with the O. Z. term if (1.5) is satisfied and we prove that it is also true for the terms of order $n+1$. There are two methods of constructing the graph representing the term of order $(n+1)$ out of a given term of order n:

(I) We start from the graph representing the nth order term and we add to it one new chain bond in such a way that both ends of it are attached to the same dotted point;

(II) we start from the graph representing the nth order term and we add to it k ($k\geq 0$) new dotted field points and $k+1$ new chain bonds each of them joining two different dotted points.

An example of the method I is shown in Fig. 1 for the case $n=2$. In this method one gets the additional factor $\gamma^3\epsilon^3$ from the new wave vector integration and, in the most divergent case, the factor ϵ^{-4} from the two new chain bonds appearing; thus finally we obtain γ^3/ϵ. This means that since the nth order term was small in comparison with the O. Z. term, then the term of order $(n+1)$ obtained in the above way is also small in comparison with the O. Z. term when (1.5) holds.

In method II one can have *a priori* infinitely many ways of constructing the new graph. However in our analysis we shall consider only those graphs which represent the most divergent in ϵ expressions, i.e., the ones in which the root points are connected with the rest of the graph by the single chain bonds. This is schematically shown in Fig. 2. If it is not the case and a given graph contains a dotted root point not representing the function $\tilde{\mu}_2^{hc}$, then this graph is replaced by another one in which the root point is replaced by the field point connected with the root point by a single chain bond. An example is shown in Fig. 3. In this way we obtain the graph representing an expression which is more divergent in ϵ but of the same order in ξ. The observation that greatly simplifies the problem is the following: Since in each order we take into account only the graphs representing the most divergent terms then it is enough to consider only the cases for which $k\leq 2$ (k is the number of new dotted field points). All the other situations ($k\geq 3$) require the addition of at least $k+2$ new chain bonds in order not to violate the rules of construct-

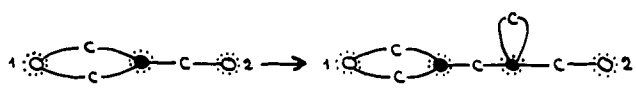


FIG. 1.

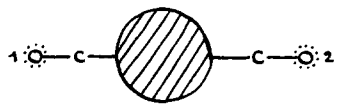


FIG. 2.

ing the new graphs and thus lead to the graphs representing the terms of at least order $n+2$. The proof of the above statement is based on the fact that the addition of each new dotted field point means the addition of $\tilde{\mu}_n^{hc}$, $n \geq 3$. Now it is easy to see that the addition of only one new dotted field point implies the addition of at least two chain bonds, and the addition of two new dotted field points implies the addition of at least three chain bonds; an example is shown in Fig. 4, and, consequently, the addition of three new dotted field points implies the addition of at least five chain bonds. This proves our statement for $k=3$ and simultaneously for $k>3$.

In order to prove the sufficiency of condition (1.5) we have to count the powers of γ and ϵ which are obtained in each of the possible ways of generating the $n+1$ order graph from the nth order graph according to the general method II. However, to simplify the calculation we shall use the parameters γ^3/ϵ and ϵ instead of γ and ϵ. In this case, to prove our hypothesis, it is enough to show that each of the above ways gives additional non-negative powers of ϵ.

Let us consider the graph of order n and let Γ_n be the ϵ power of an expression represented by this graph in the $\gamma \to 0$, $\epsilon \to 0$ limit enhanced by 2, i.e., we consider the ϵ power relative to the O.Z. function ($\sim \epsilon^{-2}$). In Sec. II we showed that $\Gamma_2 \geq 0$. According to our method we assume that $\Gamma_n \geq 0$ and we want to show that $\Gamma_{n+1} \geq 0$. To this aim we prove a stronger condition, namely that

$$\Gamma_n \geq \tfrac{5}{3}\Omega_n, \tag{3.1}$$

where Ω_n is the number of the dotted field points representing the functions $\tilde{\mu}_3^{hc}$ in a given graph of order n. It is easy to check (Ref. 1) that $\Gamma_2 \geq \frac{5}{3}\Omega_2$. Now we assume that $\Gamma_n \geq \frac{5}{3}\Omega_n$ and we prove that $\Gamma_{n+1} \geq \frac{5}{3}\Omega_{n+1}$.

This program will be realized by considering separately each way of constructing the new graph according to method II.

(1) We add one chain bond joining two dotted field points and

(a) none of the joined dotted field points represented the function $\tilde{\mu}_3^{hc}$; thus we get the factor $\gamma^3\epsilon^3\epsilon^{-2} = (\gamma^3/\epsilon)\epsilon^2$ which means that $\Gamma_{n+1} = \Gamma_n + 2$, $\Omega_{n+1} = \Omega_n$ and we see that $\Gamma_{n+1} \geq \frac{5}{3}\Omega_n + 2 = \frac{5}{3}\Omega_{n+1} + 2 \geq \frac{5}{3}\Omega_{n+1}$;

(b) only one of the joined dotted field points represented the function $\tilde{\mu}_3^{hc}$,

$$\gamma^3\epsilon^3\epsilon^{-2}(\gamma^2\epsilon^2)^{-1} = \left(\frac{\gamma^3}{\epsilon}\right)^{1/3}\epsilon^{-2/3},$$

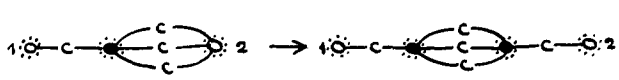


FIG. 3.

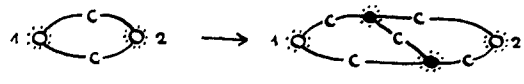


FIG. 4.

and

$$\left.\begin{aligned}\Gamma_{n+1} &= \Gamma_n - \tfrac{2}{3}\\ \Omega_n &= \Omega_{n+1} + 1\end{aligned}\right\} \to \Gamma_{n+1} \geq \tfrac{5}{3}\Omega_{n+1};$$

(c) each of the joined dotted field points represented the function $\tilde{\mu}_3^{hc}$,

$$\gamma^3\epsilon^3\epsilon^{-2}(\gamma^2\epsilon^2)^{-2} = \left(\frac{\gamma^3}{\epsilon}\right)^{-1/3}\epsilon^{-10/3},$$

and

$$\left.\begin{aligned}\Gamma_{n+1} &= \Gamma_n - \tfrac{10}{3}\\ \Omega_n &= \Omega_{n+1} + 2\end{aligned}\right\} \to \Gamma_{n+1} \geq \tfrac{5}{3}\Omega_{n+1}.$$

(2) We add one chain bond joining a dotted field point with a dotted root point and then we replace the dotted root point by the dotted field point connected with the new dotted root point by a single chain bond (to get the most divergent in ϵ expression).

(a) the joined dotted field point represented the function $\tilde{\mu}_3^{hc}$,

$$\gamma^3\epsilon^3\epsilon^{-2}(\gamma^2\epsilon^2)^{-1}(\gamma^2\epsilon^2)\epsilon^{-2} = \gamma^3/\epsilon,$$

and

$$\left.\begin{aligned}\Gamma_{n+1} &= \Gamma_n\\ \Omega_n &= \Omega_{n+1}\end{aligned}\right\} \to \Gamma_{n+1} \geq \tfrac{5}{3}\Omega_{n+1};$$

(b) the joined dotted field point did not represent the function $\tilde{\mu}_3^{hc}$,

$$\gamma^3\epsilon^3\epsilon^{-2}\gamma^2\epsilon^2\epsilon^{-2} = \left(\frac{\gamma^3}{\epsilon}\right)^{5/3}\epsilon^{8/3},$$

and

$$\left.\begin{aligned}\Gamma_{n+1} &= \Gamma_n + \tfrac{8}{3}\\ \Omega_n &= \Omega_{n+1} - 1\end{aligned}\right\} \to \Gamma_{n+1} \geq \tfrac{5}{3}\Omega_{n+1}.$$

(3) We add one dotted field point and two chain bonds in the following way: We "cut" one of the chain bonds existing in the graph by putting a dotted field point inside it and a new chain bond starts from this new dotted field point joining it with:

(a) a dotted field point representing $\tilde{\mu}_3^{hc}$,

$$\gamma^3\epsilon^3\epsilon^{-4}(\gamma^2\epsilon^2)(\gamma^2\epsilon^2)^{-1} = \gamma^3/\epsilon$$

and

$$\left.\begin{aligned}\Gamma_{n+1} &= \Gamma_n\\ \Omega_n &= \Omega_{n+1}\end{aligned}\right\} \to \Gamma_{n+1} \geq \tfrac{5}{3}\Omega_{n+1};$$

(b) a dotted field point not representing $\tilde{\mu}_3^{hc}$,

$$\gamma^3\epsilon^3\epsilon^{-4}\gamma^2\epsilon^2 = \left(\frac{\gamma^3}{\epsilon}\right)^{5/3}\epsilon^{8/3},$$

and

$$\left.\begin{aligned}\Gamma_{n+1} &= \Gamma_n + \tfrac{8}{3}\\ \Omega_n &= \Omega_{n+1} - 1\end{aligned}\right\} \to \Gamma_{n+1} \geq \tfrac{5}{3}\Omega_{n+1};$$

(c) a dotted root point and then we replace this dotted

FIG. 5.

root point by the dotted field point connected with the new dotted root point by a single chain bond (to get the most divergent in ϵ expression)

$$\gamma^3\epsilon^3\epsilon^{-4}(\gamma^2\epsilon^2)^2\epsilon^{-2}=\left(\frac{\gamma^3}{\epsilon}\right)^{7/3}\epsilon^{10/3},$$

$$\left.\begin{aligned}\Gamma_{n+1}&=\Gamma_n+\tfrac{10}{3}\\ \Omega_n&=\Omega_{n+1}-2\end{aligned}\right\}\rightarrow \Gamma_{n+1}\geq\tfrac{5}{3}\Omega_{n+1}.$$

(4) We add one chain bond joining two dotted root points and then we replace each of the dotted root points by the dotted field point connected with the new dotted root point by a single chain bond (to get the most divergent in ϵ expression),

$$\gamma^3\epsilon^3\epsilon^{-6}(\gamma^2\epsilon^2)^2=\left(\frac{\gamma^3}{\epsilon}\right)^{7/3}\epsilon^{10/3},$$

and

$$\left.\begin{aligned}\Gamma_{n+1}&=\Gamma_n+\tfrac{10}{3}\\ \Omega_n&=\Omega_{n+1}-2\end{aligned}\right\}\rightarrow \Gamma_{n+1}\geq\tfrac{5}{3}\Omega_{n+1}.$$

(5) We add the structure shown in Fig. 5 to the dotted root point,

$$\gamma_3\epsilon^3\epsilon^{-6}(\gamma^2\epsilon^2)^2=\left(\frac{\gamma^3}{\epsilon}\right)^{7/3}\epsilon^{10/3},$$

and

$$\left.\begin{aligned}\Gamma_{n+1}&=\Gamma_n+\tfrac{10}{3}\\ \Omega_n&=\Omega_{n+1}-2\end{aligned}\right\}\rightarrow \Gamma_{n+1}\geq\tfrac{5}{3}\Omega_{n+1}.$$

(6) We add two dotted field points and three chain bonds in the following way: We "cut" one or two chain bonds existing in the graph by putting two dotted field points inside them and then we add one chain bond joining those two new dotted field points

$$\gamma^3\epsilon^3\epsilon^{-6}(\gamma^2\epsilon^2)^2=\left(\frac{\gamma^3}{\epsilon}\right)^{7/3}\epsilon^{10/3}$$

and

$$\left.\begin{aligned}\Gamma_{n+1}&=\Gamma_n+\tfrac{10}{3}\\ \Omega_n&=\Omega_{n+1}-2\end{aligned}\right\}\rightarrow \Gamma_{n+1}\geq\tfrac{5}{3}\Omega_{n+1}.$$

Our analysis which is based on the counting of the powers of γ and ϵ shows that for all the graphs of the order $n+1$ obtained above, condition (3.1) is satisfied. Thus, near the critical point, the first term in the expansion of the two-particle correlation function is dominant in comparison with all the other terms if (1.5) is satisfied. This means that in the van der Waals model of a fluid close to the critical point we can use the Ornstein–Zernike function as long as the range of the potential γ^{-1} is so large that condition (1.5) is satisfied.

ACKNOWLEDGMENT

The author wishes to thank Dr. J. Piasecki for helpful discussions and for a critical reading of the manuscript.

[1]P.C. Hemmer, J. Math. Phys. 5, 75 (1964).
[2]J.L. Lebowitz and O. Penrose, J. Math. Phys. 7, 98 (1966).
[3]V.L. Ginzburg, Fiz. Tverd. Tela 2, 2031 (1960) [Sov. Phys. Solid State 2, 1824 (1961)].
[4]D.J. Thouless, Ann. Phys. (N.Y.) 10, 553 (1960); Phys. Rev. 181, 954 (1969).
[5]R. Abe, Progr. Theor. Phys. (Kyoto) 33, 600 (1965).
[6]R. Brout, *Phase Transitions* (Benjamin, New York, 1965).
[7]J.L. Lebowitz, G. Stell, and S. Baer, J. Math. Phys. 6, 1282 (1965); G. Stell, J.L. Lebowitz, S. Baer, and W. Theumann, J. Math. Phys. 7, 1532 (1966); G. Stell and W. Theumann, Phys. Rev. 186, 581 (1969).
[8]This can be easily checked with the help of the formula

$$\int d\mathbf{r}\,\mu_3(0,\mathbf{r},\mathbf{R};\rho)=k_BT\rho\left(\frac{\partial p}{\partial\rho}\right)_T^{-1}\frac{\partial x_2(0,\mathbf{R};\rho)}{\partial\rho}+k_BT\rho\left(\frac{\partial p}{\partial\rho}\right)_T^{-1}\delta(\mathbf{R}).$$

Generators of infinite direct products of unitary groups

K. Kraus, L. Polley, and G. Reents

Physikalisches Institut der Universität Würzburg, Würzburg, Germany
(Received 28 October 1976; revised manuscript received 17 March 1977)

Under suitable conditions, an infinite direct product $\otimes_n U_n(t)$ of continuous unitary one-parameter groups $U_n(t)$ is again a continuous unitary one-parameter group. This question is discussed here in terms of the generators A_n of $U_n(t)$. It is shown that the generator A of $\otimes_n U_n(t)$ has a total set of product vectors in its domain of definition. As examples, the particle number, energy–momentum, and angular momentum operators in non-Fock direct product representations of free fields are investigated. The spectra of these operators are determined.

1. INTRODUCTION

Infinite direct products of unitary groups emerge in a quite natural way in quantum field theory. Consider, for example, a free field represented in the Fock space $\mathcal{F}$, and let $\{I_n\}_{n=1}^{\infty}$ be a partition of momentum space into disjoint subsets I_n. If $\mathcal{F}_n$ denotes the "restricted" Fock space of particles with momenta in I_n, the total Fock space $\mathcal{F}$ may be written as an incomplete direct product[1] of the spaces $\mathcal{F}_n$. One can easily see that the free time evolution $U(t)$ in $\mathcal{F}$ becomes an infinite direct product

$$U(t) = \bigotimes_n U_n(t) \tag{1.1}$$

of the free time evolutions $U_n(t)$ in $\mathcal{F}_n$. The same decomposition applies to space translations and to gauge transformations of the first kind. It also holds true for spatial rotations if the subsets I_n are rotationally invariant.

Besides this, by taking suitable infinite direct products of the restricted Fock spaces $\mathcal{F}_n$ one also obtains new (non-Fock) representations which permit the solution of simple "infrared-divergent" models of quantum electrodynamics.[2–4] However, as soon as one leaves Fock space, the existence and unitarity of an operator like (1.1) can no longer be taken for granted. This problem has been investigated recently by one of us[5]; conditions have been formulated in terms of the unitary operators $U_n(t)$ which guarantee that the infinite direct product (1.1) again defines a continuous unitary one-parameter group.

A different approach has been adopted by Streit,[6] who has given equivalent conditions in terms of the infinitesimal generators A_n of the unitary groups $U_n(t)$, rather than of the groups themselves. Such conditions are particularly useful and transparent if, as usual, these infinitesimal generators have a direct physical interpretation. In Ref. 6 the discussion was restricted to the particular case where the operators $U_n(t)$ are exponentials of Schrödinger position or momentum operators, in which case (1.1) leads to Weyl operators in direct product representations of the canonical commutation relations. Actually, however, the results of Ref. 6 are completely general,[7] applying to arbitrary one-parameter groups after suitable reformulation. Such reformulation is presented here in Sec. 2 (Lemma 1), which also serves to collect—and to prove if necessary—further general statements about infinite direct products of unitary groups (Lemmas 2–6).

The results thus obtained are applied in Sec. 3 to energy–momentum, angular momentum, and particle number operators for free quantum fields represented on an infinite direct product of restricted Fock spaces. The spectra of these operators are also investigated in detail. Such applications, besides illustrating the general theory, have been important for the discussion of the infrared problem in Refs. 2–4.

2. DIRECT PRODUCTS OF UNITARY GROUPS AND THEIR GENERATORS

We consider a sequence $\{\mathcal{H}_n\}_{n=1}^{\infty}$ of separable Hilbert spaces and a family $U_n(t)$, $t \in \mathbb{R}$, $n = 1, 2, \cdots$, of (strongly) continuous unitary one-parameter groups on $\mathcal{H}_n$. Accordingly, $U_n(t)$ can be written as

$$U_n(t) = \exp(itA_n) = \int \exp(it\lambda)\, dE_n(\lambda)$$

with the self-adjoint generator A_n of $U_n(t)$ and the spectral resolution $E_n(\lambda)$ of A_n. With $\varphi_n \in \mathcal{H}_n$ and $\|\varphi_n\| = 1$ for all n, we denote by $\mathcal{H} = \otimes_n(\mathcal{H}_n, \varphi_n)$ von Neumann's incomplete direct product of the Hilbert spaces $\mathcal{H}_n$ with reference vector $\Phi = \otimes_n \varphi_n$.[1]

The following lemma is an immediate generalization of results of Ref. 6. It is obtained by replacing the Weyl operators $W_\nu(s_\nu t)$, the generators $s_\nu A_\nu$, and the spectral resolutions $s_\nu \int \xi\, dE_\nu(\xi)$ considered in Ref. 6 by $U_n(t)$, A_n, and $\int \lambda\, dE_n(\lambda)$, respectively.

Lemma 1: The infinite direct product $\otimes_n U_n(t)$ is a continuous unitary one-parameter group on $\otimes_n(\mathcal{H}_n, \varphi_n)$ if and only if there is a product vector $\Psi = \otimes_n \psi_n \in \otimes_n(\mathcal{H}_n, \varphi_n)$ with $\|\psi_n\| = 1$ for all n and

(i) $\sum_n |(\psi_n, A_n\psi_n)| < \infty$,

(ii) $\sum_n \|A_n\psi_n\|^2 < \infty$.

Moreover, Ψ may be chosen to satisfy

(iii) $(E_n(\epsilon) - E_n(-\epsilon))\psi_n = \psi_n$

for almost all n, with $\epsilon > 0$ arbitrary but fixed.

Proof: See Ref. 6, Corollary 1. The proof given there is not entirely complete since

$$\psi_n' = (E_n(\epsilon) - E_n(-\epsilon))\,\varphi_n \tag{2.1}$$

—corresponding to $E_\nu \chi_\nu$ of Ref. 6—might be zero for finitely many values of n. For such n we replace (2.1) by an arbitrary $\psi_n' \neq 0$ from D_{A_n}, the domain of definition of A_n, in order to obtain a nonzero $\Psi' = \otimes_n \Psi_n'$. The

 Copyright © 1977 American Institute of Physics

transition from ψ_n' to normalized $\psi_n = \psi_n'/\|\psi_n'\|$ leaves properties (i)–(iii) unaffected and with Ψ' also $\Psi = \otimes_n \psi_n$ belongs to $\otimes_n(H_n, \varphi_n)$.

With the mean-square deviations $\Delta_{\psi_n} A_n$ of A_n in the state ψ_n defined by

$$(\Delta_{\psi_n} A_n)^2 = \|A_n \psi_n\|^2 - (\psi_n, A_n \psi_n)^2, \tag{2.2}$$

condition (ii) may also be replaced by

$$\text{(ii)}' \quad \sum_n (\Delta_{\psi_n} A_n)^2 < \infty,$$

which is equivalent to (ii) if (i) holds true. Condition (ii) is replaced in Ref. 6 by

$$\sum_n (\psi_n, A_n^2 \psi_n) < \infty \tag{2.3}$$

which, requiring $\psi_n \in D_{A_n^2}$ rather than $\psi_n \in D_{A_n}$, is stronger than (ii). The proof of Lemma 1 is insensitive to this detail.[6]

Since D_{A_n} is dense in H_n, it is always possible to choose the reference vector $\Phi = \otimes_n \varphi_n$ so that $\varphi_n \in D_{A_n}$ for all n. Lemma 1 applied to Φ then yields a sufficient condition for the existence of $\otimes_n U_n(t)$. For a necessary condition involving the reference vector Φ see the theorem in Sec. 4 of Ref. 6.

Lemma 2: If $U(t) = \otimes_n U_n(t)$ exists as a continuous unitary one-parameter group on $\otimes_n(H_n, \varphi_n)$, then

$$U(t) = \text{s-}\lim_{k\to\infty} \overset{k}{\underset{n=1}{\otimes}} U_n(t) \otimes \mathbb{1}$$

for all t.[8]

Proof: See Ref. 5, or Ref. 6, Lemma 5.

Lemma 3: Any $\Psi = \otimes_n \psi_n$ fulfilling conditions (i) and (ii) of Lemma 1 is contained in the domain of definition D_A of the generator A of $U(t) = \otimes_n U_n(t) = \exp(itA)$. Expectation value and mean-square deviation of A in the state Ψ are given by

$$(\Psi, A\Psi) = \sum_n (\psi_n, A_n \psi_n), \tag{2.4}$$

$$(\Delta_\Psi A)^2 \equiv \|A\Psi\|^2 - (\Psi, A\Psi)^2 = \sum_n (\Delta_{\psi_n} A_n)^2. \tag{2.5}$$

Proof: Let $A^{(k)}$ be the generator of $U^{(k)}(t) \equiv (\otimes_{n=1}^k U_n(t)) \otimes \mathbb{1}$. Since $U^{(k)}(t)$ is the product of commuting unitary groups $U_n(t) \otimes \mathbb{1}$ $(n = 1, \ldots, k)$, $A^{(k)}$ is the closure of $\sum_{n=1}^k (A_n \otimes \mathbb{1})$.[8] Conditions (i) and (ii)′ of Lemma 1 imply that $A^{(k)}\Psi$ is a Cauchy sequence, since (with $j < k$)

$$\|A^{(k)}\Psi - A^{(j)}\Psi\|^2 = \left[\sum_{n=j+1}^{k} (\psi_n, A_n \psi_n)\right]^2 + \sum_{n=j+1}^{k} [\|A_n \psi_n\|^2 - (\psi_n, A_n \psi_n)^2] \to 0$$

if $j, k \to \infty$. By Lemma 2, $\exp(itA^{(k)})$ converges strongly to $\exp(itA)$ for all t. Hence, by Theorems VIII. 21 and VIII. 26 of Ref. 9, A is the strong graph limit of $A^{(k)}$. Since $\Psi \in D_A(k)$ for all k and $A^{(k)}\Psi$ converges strongly, we conclude that $\Psi \in D_A$ and $A\Psi = \text{s-}\lim_{k\to\infty} A^{(k)}\Psi$. Equations (2.4) and (2.5) are then proved by direct computation. ∎

With Lemma 3, it is easy to find a dense set of vectors χ in D_A with

$$A\chi = \text{s-}\lim_{k\to\infty} A^{(k)}\chi$$

which implies

$$(\chi, A\chi) = \lim_{k\to\infty} (\chi, A^{(k)}\chi) \quad \text{and} \quad \Delta_\chi A = \lim_{k\to\infty} \Delta_\chi A^{(k)}.$$

Namely, take $\Psi = \otimes_n \psi_n$ as in Lemma 3, let $\Psi' = \otimes_n \psi_n'$ with $\psi_n' \in D_{A_n}$ for all n and $\psi_n' = \psi_n$ for almost all n, and take for χ any finite linear combination of vectors Ψ' of this form. Since D_{A_n} is dense in H_n, the set of all such χ's is dense in $H = \otimes_n(H_n, \varphi_n)$.

The case of a $\otimes_n U_n(t)$ not unitary on $\otimes_n(H_n\ \varphi_n)$ but only on $\otimes_n^w(H_n, \varphi_n)$, constructed from the weak equivalence class of $\otimes_n \varphi_n$,[1] might be of physical interest since, roughly speaking, product vectors in $\otimes_n^w(H_n, \varphi_n)$ and in $\otimes_n(H_n, \varphi_n)$ differ from each other by "divergent phase factors" only. By means of a "renormalization"

$$U_n(t) \to \tilde{U}_n(t) = U_n(t) \exp(-it\alpha_n), \quad A_n \to \tilde{A}_n = A_n - \alpha_n, \tag{2.6}$$

with suitable real constants α_n, such phase factors may be incorporated into $\otimes_n U_n(t)$ so that the "renormalized" product $\tilde{U}(t) = \otimes_n \tilde{U}_n(t)$ is unitary on each incomplete direct product $\otimes_n(H_n, \varphi_n')$ contained in $\otimes_n^w(H_n, \varphi_n)$ and thus, in particular, on $\otimes_n(H_n, \varphi_n)$ itself. This follows immediately from:

Lemma 4: There exists a sequence $\{\alpha_n\}_{n=1}^\infty$ of real numbers so that $\tilde{U}(t) = \otimes_n U_n(t) \exp(-it\alpha_n)$ is a continuous unitary one-parameter group on $\otimes_n(H_n, \varphi_n)$ if and only if

$$\sum_n [1 - |(\varphi_n, U_n(t)\varphi_n)|] < \infty.$$

Proof: See Ref. 5, Theorem 2.5.

If suitably rewritten, the previous statements concerning $\otimes_n U_n(t)$ hold true for $\otimes_n \tilde{U}_n(t)$ also. Applying, in particular, Lemma 1 to the "renormalized" product $\otimes_n \tilde{U}_n(t)$, we obtain:

Lemma 5: There exists a real sequence $\{\alpha_n\}_{n=1}^\infty$ which renders $\tilde{U}(t) = \otimes_n U_n(t) \exp(-it\alpha_n)$ a unitary one-parameter group on $\otimes_n(H_n, \varphi_n)$ if and only if $\otimes_n(H_n, \varphi_n)$ contains a product vector $\Psi = \otimes_n \psi_n$ with $\|\psi_n\| = 1$ for all n and

$$\sum_n (\Delta_{\psi_n} A_n)^2 \equiv \sum_n [\|A_n \psi_n\|^2 - (\psi_n, A_n \psi_n)^2] < \infty.$$

Proof: Necessity follows from condition (ii)′ for the "renormalized" generators $\tilde{A}_n = A_n - \alpha_n$, since $(\Delta_{\psi_n} \tilde{A}_n)^2 = (\Delta_{\psi_n} A_n)^2$. Vice versa, take

$$\alpha_n = (\psi_n, A_n \psi_n); \tag{2.7}$$

then conditions (i) and (ii) of Lemma 1 are satisfied for the generators $\tilde{A}_n$, which proves sufficiency. ∎

Remark: The choice (2.7) for α_n is not the only possibility. Another sequence, $\{\alpha_n'\}_{n=1}^\infty$, serves the same

purpose if and only if[5]

$$\sum_n |\alpha_n - \alpha_n'| < \infty.$$

[In particular, $\alpha_n' = 0$ for all n is a possible choice if and only if $\sum_n |(\psi_n, A_n\psi_n)| < \infty$, so that Lemma 5 leads back to Lemma 1.]

A sufficient criterion for the existence of $\otimes_n \tilde{U}_n(t)$ [or, in case $\alpha_n = 0$, of $\otimes_n U_n(t)$] is given by:

Lemma 6: If there is a sequence $\{\alpha_n\}_{n=1}^{\infty}$ of real numbers and a product vector $\Psi = \otimes_n \psi_n \in \otimes_n (\mathcal{H}_n, \varphi_n)$, $\|\psi_n\| = 1$ for all n, such that

$$\sum_n (\psi_n, |A_n - \alpha_n| \psi_n) < \infty,$$

then $\tilde{U}(t) = \otimes_n U_n(t) \exp(-it\alpha_n)$ exists as a continuous unitary one-parameter group on $\otimes_n (\mathcal{H}_n, \varphi_n)$.

Proof: See Ref. 3, Lemma 4.

3. PHYSICAL APPLICATIONS

We now return to the example mentioned in the Introduction. There we have defined restricted Fock spaces $\mathcal{F}_n$ and unitary operators $U_n(t) = \exp(itH_n)$ implementing the time evolution of a free field in $\mathcal{F}_n$. We note that H_n is positive for all n so that Lemma 6 provides a sufficient condition under which

$$U_H(t) = \otimes_n \exp(itH_n)$$

exists as a continuous unitary one-parameter group on $\mathcal{H} = \otimes_n (\mathcal{F}_n, \varphi_n)$, namely

$$\text{there is a product vector } \Psi = \otimes_n \psi_n \in \mathcal{H} \text{ with } \psi_n \in D_{H_n},\ \|\psi_n\| = 1 \text{ for all } n, \text{ and } \sum_n (\psi_n, H_n\psi_n) < \infty. \tag{3.1}$$

This condition is also necessary due to Lemma 1, condition (i). If $U_H(t)$ exists, we denote its generator (the Hamiltonian) by H. It is positive because it is the limit of

$$H^{(k)} = \sum_{n=1}^{k} (H_n \otimes \mathbb{1})$$

in the strong resolvent sense, and $H^{(k)} \geq 0$ since $H_n \geq 0$ for all n (cf. Theorems VIII. 21, 24, and 33 of Ref. 9). If, as usual in relativistic theories, the Hamiltonian is required to be bounded from below, it is useless to look for phase factors $\exp(-it\alpha_n)$ in order that $\tilde{U}_H(t) = \otimes_n \exp[it(H_n - \alpha_n)]$ might become a unitary group with a "renormalized" Hamiltonian $\tilde{H}$ as generator. Namely, consider a product vector $\Psi = \otimes_n \psi_n$ as in Lemma 5 which, according to the remark following that lemma, satisfies

$$\sum_n |(\psi_n, H_n\psi_n) - \alpha_n| < \infty.$$

Then either $\sum_n (\psi_n, H_n\psi_n) < \infty$, in which case the "unrenormalized" Hamiltonian also exists, or $\sum_n (\psi_n, H_n\psi_n)$ diverges. In the latter case, let

$$\Psi_N = \left(\overset{N}{\underset{n=1}{\otimes}} \omega_n\right) \otimes \left(\underset{n>N}{\otimes} \psi_n\right),$$

with ω_n being the normalized Fock vacuum in $\mathcal{F}_n$. Then, by (the "renormalized" version of) Eq. (2.4),

$$(\Psi_N, \tilde{H}\Psi_N) = -\sum_{n=1}^{N} \alpha_n + \sum_{n>N} ((\psi_n, H_n\psi_n) - \alpha_n)$$

$$= \sum_n (\psi_n, H_n\psi_n) - \alpha_n) - \sum_{n=1}^{N} (\psi_n, H_n\psi_n).$$

The first term is finite and the second goes to $-\infty$ with $N \to \infty$, so that the "renormalized" Hamiltonian $\tilde{H}$ is not bounded from below.

The same reasoning also applies to the particle number operator N which we define, with the number operators N_n in $\mathcal{F}_n$, as the generator of $U_N(t) = \otimes_n \exp(itN_n)$. A "renormalized" $\tilde{N}$ (with $\sum_n |\alpha_n| = \infty$) thus cannot be interpreted as a particle number operator. However, if N itself exists on $\mathcal{H} = \otimes_n (\mathcal{F}_n, \varphi_n)$, then Lemma 1, condition (iii) with $\epsilon < 1$ and the known spectrum properties of N_n imply that $\mathcal{H}$ contains a product vector $\Psi = \otimes_n \psi_n$ with $\psi_n = \exp(i\beta_n)\omega_n$ for almost all n. This, in turn, implies that $\mathcal{H} = \otimes_n (\mathcal{F}_n, \varphi_n) = \otimes_n (\mathcal{F}_n, \psi_n)$ and the Fock space $\mathcal{F} = \otimes_n (\mathcal{F}_n, \omega_n)$ belongs to the same weak equivalence class,[1] which means that the representation of the CCRs on $\mathcal{H}$ is unitarily equivalent to the Fock representation.[10] With similar methods, a corresponding result for infinite direct products of Schrödinger representations was obtained by Chaiken.[11] If we are dealing with massive particles, the same arguments are applicable also to the Hamiltonian, since then there is a mass gap in the spectrum of each H_n and we can choose ϵ to be smaller than this mass gap. Thus the free Hamiltonian cannot exist in non-Fock direct product representations of the type studied here unless the particles considered have zero mass.

As far as the momentum operator $\mathbf{P}$ is concerned [again considered as the generator of $\otimes_n \exp(i\mathbf{a}\mathbf{P}_n)$ with the momentum operators $\mathbf{P}_n$ in $\mathcal{F}_n$], we may exploit the spectrum conditions in $\mathcal{F}_n$ which imply $|P_{ni}| \leq H_n$ for each component P_{ni} of $\mathbf{P}_n$. By (3.1) and Lemma 6, then, $\mathbf{P}$ exists when H exists. On the other hand, $\mathbf{P}$ may exist even if H does not, as the following example shows (cf. also Ref. 12). Consider particles of arbitrary mass $m > 0$, and take the subsets I_n of momentum space to be invariant under the reflection $\mathbf{p} \to -\mathbf{p}$. The (four-dimensional) spectrum of the Fock space energy–momentum operator $P_n = \{\mathbf{P}_n, H_n\}$ in $\mathcal{F}_n$ contains all finite sums of one-particle energy–momenta $p = \{\mathbf{p}, (\mathbf{p}^2 + m^2)^{1/2}\}$ with $\mathbf{p} \in I_n$. Therefore, in particular, it contains some point $q_n = \{\mathbf{O}, \epsilon_n\}$ with $\epsilon_n > \epsilon_0$, where $\epsilon_0 > 0$ is chosen independent of n. Denote by $E_n(\Delta)$ the spectral measure corresponding to the four-dimensional spectral representation of P_n, i.e.,

$$P_n = \int p E_n(d^4p),$$

and choose a reference vector $\Phi = \otimes_n \varphi_n$, $\|\varphi_n\| = 1$, which satisfies

$$E_n(\Delta_n)\varphi_n = \varphi_n$$

with four-dimensional spheres Δ_n of radii $1/n^2$ centered at q_n. Then

$$(\varphi_n, |P_{ni}| \varphi_n) \leq 1/n^2$$

for each component P_{ni} of $\mathbf{P}_n$, and thus $\mathbf{P}$ exists by Lemma 6. On the other hand, H cannot exist since any

vector $\Psi = \otimes_n \psi_n$ fulfilling condition (iii) of Lemma 1 satisfies $(\varphi_n, \psi_n) = 0$ for almost all n if $\epsilon < \epsilon_0$, and thus cannot be strongly equivalent to Φ. Although for this example "renormalized" Hamiltonians $\tilde{H}$ still exist (e.g., for $\alpha_n = \epsilon_n$), one may easily modify it so that not even $\tilde{H}$'s may be constructed.

In order to discuss the angular momentum operator, we assume that the subsets I_n are invariant under rotations. Then there exist continuous unitary representations $U_n(A)$ of SU(2), the covering group of the rotation group, in each restricted Fock space $\mathcal{J}_n$. The corresponding angular momentum components L_{ni} $(i = 1, 2, 3)$ generate the one-parameter subgroups $U_{ni}(t) = \exp(itL_{ni})$ of $U_n(A)$ on $\mathcal{J}_n$. In this case we can prove that $\otimes_n U_n(A)$ exists on $H = \otimes_n(\mathcal{J}_n, \varphi_n)$ as a continuous unitary representation $U(A)$ of SU(2) if and only if the infinite products $\otimes_n U_{ni}(t)$ define three continuous unitary one-parameter groups on H. It is even sufficient to assume this for $\otimes_n \tilde{U}_{ni}(t)$, where $\tilde{U}_{ni}(t) = U_{ni}(t) \times \exp(-it\alpha_{ni})$ with suitable real constants α_{ni}. (Thus, again, "renormalized" angular momenta need not be considered at all.) Morevoer, if one of these conditions is satisfied, then the reference vector $\otimes_n \varphi_n$ is strongly equivalent to a product vector $\otimes_n \sigma_n$ with $U_n(A)\sigma_n = \sigma_n$ for each n. The proof of these statements is given in the Appendix.

The previous results imply that particle number and angular momentum operators, if existing on $H = \otimes_n(\mathcal{J}_n, \varphi_n)$, have the same spectrum as the corresponding operators on Fock space. The same will now be shown for the energy–momentum operator $P = \{\mathbf{P}, H\}$. [Note that $U_{\mathbf{P}}(\mathbf{a}) = \exp(-i\mathbf{a}\mathbf{P}) = \otimes_n \exp(-i\mathbf{a}\mathbf{P}_n)$ and $U_H(t) = \exp(itH) = \otimes_n \exp(itH_n)$ commute.] Only massless particles need be considered, as remarked above. The Fock space spectrum of P then consists of the closed forward light cone $\bar{V}_+$. The generators $a_0 H - \mathbf{a}\mathbf{P}$ of space–time translations $U_H(a_0)U_{\mathbf{P}}(\mathbf{a})$ into timelike or lightlike directions (i.e., $a_0^2 - \mathbf{a}^2 \geq 0$) are easily seen to be positive (cf. the previous argument for H or Ref. 3). This implies that the spectrum of P is confined to $\bar{V}_+$. It remains to be shown that, vice versa, any $p \in \bar{V}_+$ belongs to the spectrum of P, and it suffices to consider the case $p \neq 0$. Any such p is a sum of two lightlike 4-vectors $p_i = \{\mathbf{p}_i, |\mathbf{p}_i|\}$, $i = 1, 2$, which may be interpreted as energy–momentum 4-vectors of single particles. If $\mathbf{p}_i \in I_{n_i}$, the 4-vector p_i belongs to the spectrum of the energy–momentum operator P_{n_i} in $\mathcal{J}_{n_i}$. First assume $n_1 \neq n_2$, decompose the space–time translation operators

$$U(a) = \exp(iaP) = U_H(a_0)U_{\mathbf{P}}(\mathbf{a}) = \otimes_n U_n(a)$$

(with $a = \{\mathbf{a}, a_0\}$, $aP = a_0 H - \mathbf{a}\mathbf{P}$) in the form

$$U(a) = U_{n_1}(a) \otimes U_{n_2}(a) \otimes \left(\otimes_{n \neq n_1, n_2} U_n(a) \right), \tag{3.2}$$

and denote the energy-momentum operator on $H' = \otimes'_n(\mathcal{J}_n, \varphi_n)$[13] corresponding to $U'(a) = \otimes'_n U_n(a)$ by P'. Since (3.2) implies that the spectrum of P is generated by adding the spectra of P_{n_1}, P_{n_2}, and P' (cf. Ref. 9, Theorem VIII. 33), the proof is complete if we show that $\{0, 0\}$ belongs to the spectrum of $P' = \{\mathbf{P}', H'\}$. It suffies to prove that 0 belongs to the spectrum of H', since the spectrum of P' is also confined to $\bar{V}_+$. Consider, therefore, $\exp(itH') = \otimes'_n \exp(itH_n)$. By Lemma 1, there is a vector $\Psi = \otimes'_n \psi_n \in H'$ with $\|\psi_n\| = 1$ and $\sum'_n (\psi_n, H_n \psi_n) < \infty$. For $N > n_1, n_2$, define $\Psi_N \in H'$ by $\Psi_N = (\otimes'_{n \leq N} \omega_n) \otimes (\otimes_{n > N} \psi_n)$ with the vacuum states ω_n in $\mathcal{J}_n$. With $H_n \omega_n = 0$ we get $(\Psi_N, H'\Psi_N) = \sum_{n>N} (\psi_n, H_n \psi_n)$, which converges to zero for $N \to \infty$. Together with the positivity of H' this yields the desired conclusion. If $n_1 = n_2$, (3.2) has to be replaced by $U(a) = U_{n_1}(a) \otimes (\otimes_{n \neq n_1} U_n(a))$. In this case, $p = p_1 + p_2$ belongs to the spectrum of P_{n_1}, and the rest of the proof goes through as above. We will finally show that, if P exists in $H = \otimes_n(\mathcal{J}_n, \varphi_n)$, its spectrum is purely continuous unless $\otimes_n \varphi_n$ is weakly equivalent to the Fock vacuum $\Omega = \otimes_n \omega_n$. Thus, in particular, translation invariant vectors do not exist in non-Fock representations of the type considered here. To prove this we need the following lemma:

Lemma 7: Let V_1 and V_2 be unitary operators on H_1 and H_2, respectively. Suppose that multiples of $\psi_1 \in H_1$ are the only eigenvectors of V_1. Then every eigenvector ψ of $V = V_1 \otimes V_2$ in $H = H_1 \otimes H_2$ is of the form $\psi = \psi_1 \otimes \psi_2$ where ψ_2 is an eigenvector of V_2.

Proof: We assume that $V\psi = \exp(i\alpha)\psi$, $\|\psi\| = 1$, and use the standard diagonal expansion[14]

$$\psi = \sum_i \lambda_i \varphi_i^1 \otimes \varphi_i^2$$

with orthonormal systems $\{\varphi_i^1\}_1^\infty$ and $\{\varphi_i^2\}_1^\infty$ in H_1 and H_2, respectively, and $\lambda_i \geq 0$. For every bounded operator T on H we have $(\psi, V^* T V \psi) = (\psi, T\psi)$. In particular, if $T = T_1 \otimes \mathbb{1}$ this implies

$$\sum_i \lambda_i^2 (\varphi_i^1, V_1^* T_1 V_1 \varphi_i^1) = \sum_i \lambda_i^2 (\varphi_i^1, T_1 \varphi_i^1),$$

or

$$\mathrm{Tr}(V_1^* T_1 V_1 W_1) = \mathrm{Tr}(T_1 W_1) \tag{3.3}$$

with

$$W_1 = \sum_i \lambda_i^2 |\varphi_i^1\rangle\langle\varphi_i^1|.$$

As T_1 is arbitrary, (3.3) leads to

$$V_1 W_1 V_1^* = \sum_i \lambda_i^2 |V_1 \varphi_i^1\rangle\langle V_1 \varphi_i^1| \equiv W_1.$$

Since $\sum_i \lambda_i^2 = 1$, and the only finite-dimensional subspace of H_1 invariant under V_1 is the one generated by ψ_1, it follows that only a single λ_i is $\neq 0$ and thus $= 1$, $\lambda_1 = 1$ say, and φ_1^1 is a multiple of ψ_1. Hence $\psi = \psi_1 \otimes \psi_2$, and then ψ_2 must be an eigenvector of V_2. ∎

We assume now that there exists an eigenvector Ψ of $U(a) = \otimes_n \exp(iaP_n)$ in $H = \otimes_n(\mathcal{J}_n, \varphi_n)$. We know that ω_n is the only eigenvector of $\exp(iaP_n)$. Writing $H = \mathcal{J}_1 \otimes (\otimes_{n>1}(\mathcal{J}_n, \varphi_n))$ and $U(a) = \exp(iaP_1) \otimes (\otimes_{n>1} \exp(iaP_n))$ we obtain, by Lemma 7, $\Psi = \omega_1 \otimes \Psi_2$. Applying the same argument to Ψ_2, etc., we conclude that Ψ is factorizable and, morevoer, $\Psi = \otimes_n \exp(i\beta_n)\omega_n$.[14] Since Ψ is strongly equivalent to $\Phi = \otimes_n \varphi_n$, the latter is weakly equivalent to $\Omega = \otimes_n \omega_n$. Vice versa, of course, weak equivalence of Ω and Φ implies that $H = \otimes_n(\mathcal{J}_n, \varphi_n)$ contains $\Psi = \otimes_n \exp(i\beta_n)\omega_n$ with suitable β_n, so that $U(a)$ exists in H and leaves Ψ invariant.

ACKNOWLEDGMENT

We want to thank Michael Everitt for his patient help concerning the English style of the manuscript.

APPENDIX

Consider Hilbert spaces H_n, $n=1, 2, \ldots$, and continuous unitary representations $U_n(A)$ of SU(2) on H_n. Denote by L_{ni}, $i=1, 2, 3$, the corresponding angular momentum components on H_n, which generate the one-parameter subgroups $U_{ni}(t)=\exp(itL_{ni})$ of $U_n(A)$. Let $H=\otimes_n(H_n, \varphi_n)$ with $\|\varphi_n\|=1$ for all n.

Theorem: The following statements are equivalent:

(a) H contains a product vector $\otimes_n\sigma_n$ with $\|\sigma_n\|=1$ for all n and $U_n(A)\sigma_n=\sigma_n$ for almost all n and all $A\in \mathrm{SU}(2)$.

(b) $\otimes_n U_n(A)$ exists on H as a continuous unitary representation $U(A)$ of SU(2).

(c) $\otimes_n U_{ni}(t)$ exist, for $i=1, 2, 3$, as continuous unitary one-parameter groups on H.

(d) There are three sequences $\{\alpha_{ni}\}_{n=1}^{\infty}$, $i=1, 2, 3$, of real numbers α_{ni}, so that the direct products $\otimes_n\tilde{U}_{ni}(t)$ of $\tilde{U}_{ni}(t)=U_{ni}(t)\exp(-it\alpha_{ni})$ define unitary one-parameter groups on H.

For the proof we need the following lemma. Consider a continuous unitary representation $D(A)$ of SU(2) on a Hilbert space D, and denote by L_i the corresponding angular momentum components.

Lemma 8: If D does not contain vectors $\varphi\neq 0$ with $D(A)\varphi=\varphi$ for all $A\in \mathrm{SU}(2)$, then

$$|(\psi, \chi)|\leq(1/\sqrt{2})\|\psi\|\,\|\chi\|$$

for arbitrary eigenstates ψ of L_i and χ of L_j, with $i\neq j$.

This lemma may be proved first for an irreducible representation with $\mathbf{L}^2\neq 0$. The general case then follows by decomposing D into irreducible subspaces. Both steps of the proof are elementary, and are thus omitted here.

Proof of the Theorem: Obviously it suffices to derive (b) from (a) and (a) from (d). Assume (a) to be true. Lemma 1 applied to $\otimes_n\sigma_n$ then implies that, for an arbitrary one-parameter subgroup A_t of SU(2), $\otimes_n U_n(A_t)$ is unitary and continuous on H. Since each $A\in \mathrm{SU}(2)$ belongs to some subgroup A_t, $U(A)=\otimes_n U_n(A)$ is unitary on H. The representation property of $U(A)$ is obvious, and continuity then follows from the continuity of $U(A_t)$ for one-parameter subgroups A_t.

Assume (d) to be true so that, in particular, $\otimes_n\tilde{U}_{n3}(t)$ exists. Then, with $\tilde{E}_{n3}(\lambda)$ denoting the spectral resolution of $\tilde{L}_{n3}=L_{n3}-\alpha_{n3}$ and $\epsilon>0$ fixed but arbitrary, there exists in H a product vector $\otimes_n\psi_n$, $\|\psi_n\|=1$, such that

$$(\tilde{E}_{n3}(\epsilon)-\tilde{E}_{n3}(-\epsilon))\psi_n=\psi_n$$

for almost all n [Lemma 1, condition (iii).] Since the spectrum of $\tilde{L}_{n3}$ is discrete with spacing $\geq\frac{1}{2}$, we may choose $\epsilon<\frac{1}{4}$ and conclude that, for all $n>N$ say,

$$L_{n3}\psi_n=m_n\psi_n.$$

Let $A_1\in \mathrm{SU}(2)$ correspond to a rotation by $\pi/2$ around the 1 axis. Then $\chi_n=U(A_1)\psi_n$ satisfies, for $n>N$,

$$L_{n2}\chi_n=m_n\chi_n.$$

Since $\otimes_n\tilde{U}_{n1}(\pi/2)$ is also unitary, the product vector $\otimes_n\tilde{\chi}_n=\otimes_n\exp[-i(\pi/2)\alpha_{n1}]\chi_n$ belongs to H, too, and therefore

$$\infty>\sum_n|1-(\psi_n, \tilde{\chi}_n)|\geq\sum_{n>N}[1-|(\psi_n, \chi_n)|].$$

Write, for $n>N$, $\psi_n=\psi_n^0+\psi_n^+$ and similarly for χ_n, with $\mathbf{L}_n^2\psi_n^0=0$ and $\psi_n^+\in D_n$, the subspace of H_n belonging to nonzero eigenvalues of $\mathbf{L}_n^2$. Then $\chi_n^0=U(A_1)\psi_n^0=\psi_n^0$ and $L_{n3}\psi_n^+=m_n\psi_n^+$, $L_{n2}\chi_n^+=m_n\chi_n^+$, and thus Lemma 8 implies, for $n>N$,

$$\begin{aligned}|(\psi_n, \chi_n)|&\leq\|\psi_n^0\|^2+|(\psi_n^+, \chi_n^+)|\\&\leq\|\psi_n^0\|^2+\frac{1}{\sqrt{2}}\|\psi_n^+\|^2\\&=\frac{1}{\sqrt{2}}+\left(1-\frac{1}{\sqrt{2}}\right)\|\psi_n^0\|^2\leq 1.\end{aligned}$$

Therefore,

$$\begin{aligned}\infty&>\sum_{n>N}[1-|(\psi_n, \chi_n)|]\\&\geq\left(1-\frac{1}{\sqrt{2}}\right)\sum_{n>N}(1-\|\psi_n^0\|^2).\end{aligned}$$

This implies $\psi_n^0\neq 0$ (and, consequently, $m_n=0$) for almost all $n>N$. Moreover, with

$$\sigma_n=\begin{cases}\psi_n^0/\|\psi_n^0\| & \text{if } n>N \text{ and } \psi_n^0\neq 0,\\ \text{any unit vector in } H_n & \text{if } n\leq N \text{ or } \psi_n^0=0,\end{cases}$$

the product vector $\otimes_n\sigma_n$ belongs to H. Choosing N' so that $\sigma_n=\psi_n^0/\|\psi_n^0\|$ for all $n>N'$, we have

$$\begin{aligned}\sum_{n>N'}(1-|\psi_n, \sigma_n)|)&=\sum_{n>N'}[1-(\psi_n, \psi_n^0)/\|\psi_n^0\|]\\&=\sum_{n>N'}(1-\|\psi_n^0\|)\\&\leq\sum_{n>N'}(1-\|\psi_n^0\|^2)<\infty,\end{aligned}$$

which proves the strong equivalence of $\otimes_n\psi_n$ and $\otimes_n\sigma_n$. ∎

Remarks: (1) If (as in the case $H_n=J_n$ considered here) each H_n contains vectors with $\mathbf{L}_n^2=0$, then we may achieve $\mathbf{L}_n^2\sigma_n=0$ for all n.

(2) It was sufficient for our proof to require, instead of (d), only the existence of $\otimes_n\tilde{U}_{ni}(t)$ for a single axis i and the unitarity of $\otimes_n\exp(i\beta_n)U_{nj}(\pi/2)$, with suitable real β_n, for a second axis $j\neq i$. However, this requirement is easily shown to be equivalent to (d).

[1]J. von Neumann, Compositio Math. 6, 1—77 (1938); *Collected Works* (Pergamon, New York, 1961), Vol. 3, pp. 323–99.
[2]G. Reents, J. Math. Phys. 15, 31—4 (1974).
[3]K. Kraus, G. Reents, and L. Polley, Ann. Inst. Henri Poincaré 26, 109—62 (1977).
[4]G. Reents, *Modelle zum Infrarotproblem in der Quantenelektrodynamik* (Habilitationsschrift, Universität Würzburg, 1977).
[5]G. Reents, Commun. Math. Phys. 39, 121—30 (1974).
[6]L. Streit, Commun. Math. Phys. 4, 22—31 (1967).
[7]This was not noticed by us in a previous version of this paper, which therefore contained independent proofs of the results mentioned. We want to thank the referee for drawing our attention to Ref. 6.
[8]Here $\mathbb{1}$ means $\otimes_{n>k}\mathbb{1}_n$. Similarly, we use $A_k\otimes\mathbb{1}$ as a shorthand notation for $(\otimes_{n<k}\mathbb{1}_n)\otimes A_k\otimes(\otimes_{n>k}\mathbb{1}_n)$.
[9]M. Reed and B. Simon, *Methods of Modern Mathematical Physics. Vol. I: Functional Analysis* (Academic, New York, 1972).
[10]J.R. Klauder, J. McKenna, and E.J. Woods, J. Math. Phys. 7, 822—8 (1966).
[11]J.M. Chaiken, Ann. Phys. (N.Y.) 42, 23—80 (1967); Commun. Math. Phys. 8, 164—84 (1968).
[12]L. Polley, *Quantisierte Maxwellfelder mit unendlich vielen Photonen endlicher Gesamtenergie* (Diplomarbeit, Würzburg, 1975).
[13]We abbreviate $\otimes_{n\neq n_1,n_2}$ and $\Sigma_{n\neq n_1,n_2}$ by $\otimes'_n$ and Σ'_n, respectively.
[14]H. Araki and E.J. Woods, Publ. RIMS Kyoto Univ. Ser. A 2, 157—242 (1966).

Sequences of transformations for particle-field Hamiltonians[a)]

Eugene P. Gross

Department of Physics, Brandeis University, Waltham, Massachusetts 02154
(Received 4 April 1977)

The nonrelativistic Hamiltonian for a particle interacting with a scalar field is studied by a method that involves a product of unitary transforms. We use intermediate coupling transforms for finite particle momenta to treat the interaction with oscillators with a range of wave vectors. Elimination of the oscillators in the range leads to a new Hamiltonian of the identical operator structure but with a modified interaction and mass. We show that with any finite number of divisions the self energy is always lower than the intermediate coupling result. In the limit of infinitely many divisions differential equations are derived for the interaction and effective mass as a function of the smallest wave vector of the oscillators that have been removed.

1. INTRODUCTION

In the present paper we study the nonrelativistic particle-field Hamiltonian from a point of view which is in the spirit of the ideas of Wilson[1] and of Anderson.[2] We note that the well known intermediate coupling theory[3] can be applied to study the interaction of a particle with a set of field oscillators with wave vectors that lie in a narrow band (in one dimension) or narrow shell (in three dimensions). This treatment is implemented by a simple, explicit canonical transformation. If one then takes the vacuum expectation value for the oscillators in question, one has "eliminated" these oscillators. One is left with a new Hamiltonian which has the following features. First there is a c number contribution to the self energy arising from the interaction of the particle with the oscillators in the band. Second there is an addition to the mass of the particle. Finally the interaction of the particle with the other oscillators is modified. The transformation that we use is sufficiently elementary that the new Hamiltonian has exactly the same operator structure as the old one.

The simplest illustration of the procedure is a one slice theory. Here the interaction of the particles with oscillators of wave vectors greater than $|K_1|$ is treated first by an intermediate coupling transformation. We then show that treating the residual Hamiltonian by a new transformation leads to a lower ground state energy than the one obtained with a single IC transformation. It is also possible to treat the residual Hamiltonian by standard strong coupling techniques and to show that the combination of steps improves strong coupling theory. We will however not go into this in the present paper.

A more interesting application results when one passes to the infinite limit of a many band theory. The theory then contains a function $m(k)$ which tends to the value of the bare mass as $k \to \infty$ and to the dressed mass as $k \to 0$. At a given value of $|k|$ it represents the effective mass arising from the elimination of oscillators from $|k|$ to ∞. There is a vertex correction $V(k) = V_0(k) \exp[-k^2 A(k)]$ where $V_0(k)$ is the original interaction and $A(k)$ is a function that tends to zero as $|k| \to \infty$. The functions $m(k)$ and $A(k)$ obey a set of coupled differential equations. The ground state energy has the IC form with $V(k)$ and $m(k)$ replacing the bare quantities in the integral over wave vectors.

Our results however yield only a modest improvement of standard theories. For example, for weak coupling the energy is lower than standard IC theory, but one does not obtain the full perturbation theoretic correction to that theory. The averaging procedure in the elimination of each band ignores certain correlations. In addition we do not obtain the strong coupling limit correctly with the simple procedure used in the present paper. To do that we need to diagonalize the Hamiltonian for bands of a finite width (that depends on the coupling constant) by methods more accurate than the one employed here. The use of repeated unitary transformations in the strong coupling limit requires further investigation.

The Hamiltonian that we will study has the form

$$H = p^2/2 + g \int [V_0(k)\, a(k) \exp(ikq) + \text{h. c.}]\, dk + \int \omega(k)\, a^*(k)\, a(k)\, dk \tag{1.1}$$

for the one-dimensional case. Here $\hbar = 1$ and the bare mass has been taken to be unity. $V_0(k)$ is the bare interaction and $\omega(k)$ the oscillator spectrum, and k ranges from $-\infty$ to $+\infty$. We operate in the continuum with

$$[a(k), a^*(l)] = \delta(k - l). \tag{1.2}$$

For the three-dimensional case we have

$$H = p^2/2 + g_3 \int [W_3(k)\, a(\mathbf{k}) \exp(i\mathbf{k}\cdot\mathbf{q}) + \text{h. c.}]\, d^3k + \int \omega(k)\, a(\mathbf{k})\, a(\mathbf{k})\, d^3k. \tag{1.3}$$

Because of the spherical averaging involved in our procedure the results for the three-dimensional case are simply obtained from the one-dimensional case by replacing $V_0(k)$ by $W_3(k)\, k^2$ and g by g_3 times a numerical factor. In some parts of the paper more explicit calculations are carried out for the special one-dimensional case $V_0(k) = \omega(k) = 1$.

2. TREATMENT OF A BAND OF OSCILLATORS

For the one-dimensional case, let

[a)]Work supported by the National Science Foundation under grant DMR-76-02048.

 Copyright © 1977 American Institute of Physics

$$H_\Delta = p^2/2m + g\int_\Delta [V_0(k)\, a(k)\exp(ikq) + \text{h. c.}]\, dk + \int_\Delta \omega(k)\, a^*(k)\, a(k)\, dk. \tag{2.1}$$

Here Δ includes the range K_1 to K_0, $K_1 < K_0$, and also the corresponding range $-K_1$ to $-K_0$.

Consider the unitary transform $U(\Delta)$,

$$U(\Delta) = \exp[iq\int_\Delta k a^* a\, dk], \tag{2.2}$$

$$U(\Delta)\, a(k)\, U^{-1}(\Delta) = a(k)\exp(-ikq), \tag{2.3}$$

for k inside the band Δ

$$U(\Delta)\, pU^{-1}(\Delta) = p - \int_\Delta k a^* a\, dk, \qquad U(\Delta)\, qU^{-1}(\Delta) = q. \tag{2.4}$$

Next we introduce a form of intermediate coupling theory appropriate for low particle momenta. It is an essential feature of the present approach that we need to perform a sequence of transforms for nonzero momenta even if our aim is to compute only the ground state energy. We use a function

$$f(k,p) = -gV_0(k)\, D[1 + kpD^2 1/m^*], \tag{2.5}$$

where

$$m^* = m + 2g^2\int_\Delta k^2 V_0^2(k)\, D^3\, dk \tag{2.6}$$

and

$$D(k,m) = [\omega + k^2/2m]^{-1}.$$

We perform the unitary transform

$$W(\Delta) = \exp\int_\Delta [a(k) - a^*(k)] f(k,p)\, dk, \qquad W(\Delta)\, a(k)\, W^{-1}(\Delta) = a(k) + f(k,p) \tag{2.7}$$

for k in the band Δ;

$$W(\Delta)\, p\, W^{-1}(\Delta) = p,$$

$$q' \equiv W(\Delta)\, q\, W^{-1}(\Delta) = q - i\int \frac{\partial f(k,p)}{\partial p}\,(a(k) - a^*(k))\, dk. \tag{2.8}$$

Note that $\partial f/\partial p$ is independent of p.

Next, we take a vacuum expectation value with respect to the oscillators in the band, viz. state vectors such that $a(k)\,\Phi_0 = 0$. We then have

$$E(\Delta) = \langle \Phi_0,\, W(\Delta)\, U(\Delta)\, H_\Delta U^{-1}(\Delta) W^{-1}(\Delta)\, \Phi_0\rangle, = p^2/2m^* - g^2\int_\Delta V_0^2\, dk\, D(k,m). \tag{2.9}$$

We now focus on the complete Hamiltonian, including the oscillators (labelled by λ) that lie outside the band considered. Since this Hamiltonian is additive in the oscillators the only coupling between bands occurs through the coordinate q. We have

$$\langle \Phi_0,\, W(\Delta)\, U(\Delta)\exp(i\lambda q)\, U^{-1}(\Delta)\, W^{-1}(\Delta)\, \Phi_0\rangle = \exp(i\lambda q)\exp(-\lambda^2\alpha/2), \tag{2.10}$$

where

$$\alpha = \frac{g^2}{m^{*2}}\int_\Delta V_0^2(k)\, D^4 k^2\, dk.$$

Thus the Hamiltonian after removal of the band Δ is

$$H = p^2/2m^* + g\int_\Lambda [V_0(\lambda)\exp(-\lambda^2\alpha/2)\, a(\lambda)\exp(i\lambda q) + \text{h. c.}]\, d\lambda + \int_\Lambda \omega a^* a\, d\lambda - g^2\int_\Delta V_0^2(k)\, D(k,m)\, dk, \tag{2.11}$$

where Λ refers to the remaining part of the range.

We will need to do this repeatedly, starting with a problem where there is an upper cutoff K_0. The starting interaction is $V_0(k)$ and bare mass is m_0. The define

$$\epsilon_j = -g^2\int_{\Delta_j} |V_{j-1}|^2 D_{j-1}(k)\, dk, \quad D_{j-1} = \left(\omega\frac{k^2}{2m_{j-1}}\right)^{-1}, \tag{2.12}$$

$$m_j - m_{j-1} = 2g^2\int_{\Delta_j} |V_{j-1}|^2 k^2 D_{j-1}^3\, dk, \tag{2.13}$$

$$m_j^2\alpha_j = g^2\int_{\Delta_j} V_{j-1}^2(k)\, k^2 D_{j-1}^4\, dk. \tag{2.14}$$

Here

$$V_j(k) = V_0(k)\exp[-(k^2/2)(\alpha_p + \alpha_2 + \cdots + \alpha_g)]. \tag{2.15}$$

After averaging over j bands we have the residual Hamiltonian

$$H_j = \frac{p^2}{2m_j} + g\int_{\Lambda_j} [V_j(k)\, a(k)\exp(ikq) + \text{h. c.}]\, dk + \int_{\Lambda_j} \omega a^* a\, dk + \sum_{n=0}^{j}\epsilon_n. \tag{2.16}$$

Here Λ_j refers to the remaining wave vector domain.

We now make N slices of the domain 0 to K_0 (with the corresponding negative part $-K_0$ to 0). In the limit of large N we have a function $m(k)$. We work with the positive wave vector range Δ_j^+ and double the integrals involved in m_j, α_j. Since m_j refers to lower k values than m_{j-1}, we have in the limit $N \to \infty$

$$-\frac{dm(k)}{dk} = +g^2 k^2 V^2(k)\, D^3(k)\cdot 4, \qquad D(k) \equiv \left(\omega(k) + \frac{k^2}{2m(k)}\right)^{-1}. \tag{2.17}$$

Each α_j is proportional to K_0/N. The sum

$$A_{j-1} \equiv \alpha_1 + \cdots + \alpha_{j-1} \tag{2.18}$$

becomes a function $A(k)$,

$$A(k) = 2g^2\int_k^{K_0} d\xi\, \frac{1}{m^2(\xi)}\, V^2(\xi)\, \xi^2 D^4(\xi)\, d\xi \tag{2.19}$$

and

$$V(k) = \exp(-k^2/2)A(k))V_0(k). \tag{2.20}$$

Thus we find the differential equations

$$-\frac{dA(k)}{dk} = \frac{2g^2}{m^2(k)}\, k^2 V_0^2(k)\exp[-k^2 A(k)]\, D^4(k), \tag{2.21}$$

$$-\frac{dm}{dk} = 4g^2 k^2 V_0^2(k)\exp(-k^2 A)\, D^3(k), \tag{2.22}$$

with

$$m(k=K_0)=m_0,\ A(k=K_0)=0.$$

Since

$$dm/dk<0,\ m(k=0)>m_0,\ \text{and } A(k=0)>0.$$

The expression for the total energy is

$$E=\frac{p^2}{2m(0)}-2g^2\int_0^{K_0}V^2(k)\,dk\,D(k). \tag{2.23}$$

Finally, we pass to the limit $K_0\to\infty$.

3. ENERGY FOR A SINGLE SLICE

We now examine the ground state energy for a single slice when both sides are treated by distinct IC transformations. The energy is

$$E=-2g^2\int_{K_1}^{\infty}dk\,V_0^2D_0(k)-2g^2\int_0^{K_1}V_0^2\times\exp(-k^2\alpha)\left(\omega+\frac{k^2}{2m_1}\right)^{-1}dk, \tag{3.1}$$

$$y^2\equiv\alpha=\frac{2g^2}{m_1^2}\int_{K_1}^{\infty}V_0^2k^2\,dk\,D_0^4,$$

$$m_1-1=4g^2\int_{K_1}^{\infty}V_0^2k^2\,dk\,D_0^3, \tag{3.2}$$

$$D_0(k)=(\omega+k^2/2)^{-1}.$$

We study the expression in detail for $\omega=V_0=1$.

Domain A: $K_1\gg1$

This is the most important domain. Define

$$\epsilon^2=\frac{g^2}{K_1^3}\,\frac{32}{3}. \tag{3.3}$$

Expanding $D(k)$ in the high k region

$$m_1\to1+\epsilon^2+\cdots,\quad\alpha\to\frac{1}{m_1^2}\,\frac{3}{5}\,\frac{\epsilon^2}{K_1^2},$$
$$(K_1y)^2\to\frac{3}{5}\,\frac{\epsilon^2}{(1+\epsilon^2)^2}\leq\sqrt{3/20}. \tag{3.4}$$

The first term in the energy is the standard IC contribution from the high k part. The second term represents a modified contribution from the low k part. It may be written as

$$\delta E=-2g^2K_1\int_0^1\frac{\exp[-\eta^2(K_1y)^2]\,d\eta}{1+\eta^2(K_1x)^2},\quad x^2=\frac{1}{2m_1}. \tag{3.5}$$

This is to be compared with the value $-g^2\,2\sqrt{2}\tan^{-1}(K_1/\sqrt{2})$ which is the contribution of this wave vector range to the self energy in standard IC theory.

Expanding the exponential we find

$$\delta E=-\frac{2g^2}{x}\left(\tan^{-1}K_1x-\frac{y^2}{x^2}[K_1x-\tan^{-1}K_1x]+\cdots\right). \tag{3.6}$$

There are a number of subcases.

1. $K_1x\gg1$

This condition is $(1+\epsilon^2)^{1/2}\ll K_1/\sqrt{2}$. It of course includes the situation $\epsilon^2\ll1$ when

$$\delta E\to-g^2\,2\sqrt{2}\left(\frac{\pi}{2}-\frac{\sqrt{2}}{K_1}\right)-g^2\sqrt{2}\,\frac{\epsilon^2}{2}. \tag{3.7}$$

The first term is the intermediate coupling contribution and the last term represents a lowering of the energy of order g^4. Note that $\epsilon^2\ll1$ is compatible with $g^2>1$ provided that $g^2<\frac{3}{32}K_1^3$. We can also have $\epsilon^2>1$, i.e., even stronger couplings. Then $x\to1/\sqrt{2}\epsilon$ and

$$\delta E\to-\sqrt{2}\,g^2\epsilon, \tag{3.8}$$

which is much lower than $-\sqrt{2}\,g^2\pi$.

2. $K_1x<1$

This is the domain of very large coupling strengths, viz. $g^2>K_1^5\,\frac{3}{32}$. We have

$$\delta E\to-2g^2K_1\left(1-\frac{(K_1x)^2}{3}+\cdots\right). \tag{3.9}$$

The leading term is the lower bound for the self energy noted by Lee and Pines.[4] It is the self energy of an infinitely massive source with a wave vector cutoff that is independent of coupling constant. It is interesting that this saturation of the interaction with a group of oscillators is obtained here as a result of the modification of the vertex and effective mass due to interaction.

Domain B: $K_1\ll1$

In this limit only a small part of the problem is treated with the corrected mass and interaction. We find that both K_1x and K_1y are less than unity. The energy is

$$\delta E=-2g^2K_1\left(1-\frac{K_1^2}{3}(x^2+y^2)+\cdots\right). \tag{3.10}$$

The IC result is

$$\delta E\to-2g^2K_1\left(1-\frac{K_1^2}{3}\,\frac{1}{2}+\cdots\right). \tag{3.11}$$

We write

$$m_1=1+g^2A_1,\quad y^2=\frac{2g^2}{m_1^2}\,B_1. \tag{3.12}$$

In the limit $K_1\ll1$,

$$A_1\to\int_0^{\infty}k^2\,dk\,D_0^3=\frac{\sqrt{2}\,\pi}{8},\quad B_1\to\int_0^{\infty}k^2\,dk\,D_0^4=\frac{2\pi}{\sqrt{2}}\,\frac{1}{16}. \tag{3.13}$$

In small g^2, $(x^2+y^2)\to\frac{1}{2}$ and we obtain the IC result. On the other hand, for $g^2\gg1$ we find

$$x^2+y^2\to\frac{A+B}{A^2}\,\frac{1}{2}\,\frac{1}{g^2}$$

which tends to zero. Thus the interaction saturates, i.e., reaches the lower bound. This is a general result for $g^2\gg1$. We have

$$y^2=\frac{g^2B_1}{(1+g^2A_1)^2}\to0$$

so that the energy approaches

$$\delta E\to-2g^2\sqrt{2m_1}\tan^{-1}\frac{K_1}{\sqrt{2m}}\to-2g^2K_1. \tag{3.14}$$

With a more general $\omega(k)$ and $V(k) \leq 1$ it is easy to treat the case $K_1 \gg 1$ when $K_1 y \ll 1$, $m_1 - 1 \ll 1$. The energy is lower than the IC result by the positive quantity

$$2g^2 \int_0^{K_1} V_0^2(k)\, k^2\, dk\, D(k) \int_{K_1}^{\infty} l^2 V_0^2(l)\, D_0^3(l)[D_0(k) - D_0(l)]\, dl \times D_0(k) = \left(\omega + \frac{k^2}{2}\right)^{-1} \tag{3.15}$$

The same considerations go through with trivial modifications in the three-dimensional case.

Finally, it is of interest to examine the ground state wavefunction corresponding to the two step IC theory. Standard IC theory works with states of total momentum p_0.

$$\begin{aligned}\Psi &= U^{-1} V^{-1} \Phi_0 \\ &= \exp(-iq \int ka^* a\, dk) \exp[\int (a^* - a) f(k,p)\, dk] \exp(ip_0 q)\, \Phi_0 \\ &= \exp[-\tfrac{1}{2} \int B^2(k,p_0)\, dk] \exp[a + (k) \\ &\quad \times \exp(-ikq) f(k,p_0)\, dk_1]\, \Phi_0 \exp(ip_0 q),\end{aligned} \tag{3.16}$$

where $f(k,p_0)$ contains the bare mass. In the two step theory

$$\Psi = U^{-1}(k)\, V^{-1}(k)\, U^{-1}(\lambda)\, V^{-1}(\lambda) \exp(ip_0 q)\, \Phi_0 \tag{3.17}$$

or

$$\begin{aligned}\Psi &= \exp[-iq \int ka^* a\, dk - iq'' \int \lambda a^* a\, d\lambda]\, V^{-1}(k)\, V^{-1}(\lambda) \\ &\quad \times \exp(ip_0 q)\, \Phi_0.\end{aligned} \tag{3.18}$$

Here again k refers to the high wave vector region and λ to the low wave vector region. We also have

$$\begin{aligned}q'' &= q + iG, \\ G &= \int \frac{\partial f(k,p)}{\partial p}\, [a(k) - a^*(k)]\, dk.\end{aligned} \tag{3.19}$$

We use $f_c(\lambda, p_0)$ to remind ourselves that in the λ region the corrected vertex and mass are used as parameters in the transformation. Defining

$$\begin{aligned}F^* &= \int a^*(k) \exp(-ikq) f(k,p_0) - \tfrac{1}{2} \int f^2\, dk, \\ F_c^* &= \int a^*(\lambda) \exp(-i\lambda q) f_c(\lambda, p_0) - \tfrac{1}{2} \int f_c^2\, d\lambda,\end{aligned} \tag{3.20}$$

we write

$$\Psi = \exp(\int \lambda a^* a\, d\lambda \cdot G) \exp(F^* + F_c^*) \exp(ip_0 q)\, \Phi_0. \tag{3.21}$$

It is clear that the type of correlations between bands in the present theory is different from conventional variational choices. This is even more pronounced in a multiband theory.

4. CONTINUOUS COMPOUNDING

We now return to Sec. 2 and examine the ground state energy for the continuum product of unitary transforms. We first study the differential equations in the limit of weak coupling and show that there is a g^4 lowering of the IC energy. The weak coupling limit is obtained by setting $A(k) = 0$, $m(k) = m_0 = 1$ on the right-band side of Eq. (2.21), (2.22). To order g^2, $m(k) = 1 + \delta m(k)$,

$$\delta m(k) = 4g^2 \int_k^{\infty} k^2 V_0^2 D_0^3(k)\, dk, \tag{4.1}$$

$$A(k) = 2g^2 \int_k^{\infty} l^2 V_0^2 D_0^4\, dl. \tag{4.2}$$

The energy to order g^4 is

$$E = \frac{p^2}{2m(k=0)} - 2g^2 \int_0^{\infty} D_0 V_0^2\, dk - 2g^2 \int_0^{\infty} V_0^2 k^2 D_0 [D_0 \delta m - A]\, dk. \tag{4.3}$$

The energy for $p = 0$ is the IC value plus the last term δE. The effective mass $m(k=0)$ is the IC value. We note first that δE is indeed negative since

$$D_0(k) \int_k^{\infty} V_0^2(l)\, l^2\, D_0^3(l)\, dl > \int_k^{\infty} V_0^2(l)\, l^2\, D_0^4(l)\, dl. \tag{4.4}$$

We have

$$\delta E = -4g^4 \left(\int_0^{\infty} k^2 V_0^2(k)\, D_0(k)\, dk \int_k^{\infty} V_0^2(l)\, l^2 D_0^3(l)\, dl - \int_0^{\infty} k^2 V_0^2(k)\, D_0(k) \int_k^{\infty} l^2 V_0^2\, D_0^4(l)\, dl \right). \tag{4.5}$$

For the case $\omega = V_0 = 1$ we do the integrals explicitly to find

$$\delta E = -4g^4 \left(\frac{3\pi^2}{32} - \frac{1}{4} \right). \tag{4.6}$$

This improvement of IC does not contain the entire g^4 correction. Perturbation theory yields

$$\delta E = -g^4 \left[\int V_0^2(k)\, D_0\, dk \right]^2 - g^4 \iint dk\, dl \frac{V_0^2(k)\, D_0(k)\, V_0^2(l)\, D_0(l)}{\omega(k) + \omega(l) + (k+l)^2/2}. \tag{4.7}$$

It is apparent that correlations are lost in the averaging process. This is already the case in one dimension. In three dimensions the loss is obvious because of angular averaging.

We now briefly analyze the differential equations for general coupling strengths. We scale variables by introducing $k^* = g^{2/3}$ and

$$\xi = k/k^*, \quad B = Ak^{*2}. \tag{4.8}$$

Then

$$-\frac{dm}{d\xi} = 4\xi^2 \left(\frac{1}{k^{*2}} + \frac{\xi^2}{2m} \right)^{-3} \exp(-\xi^2 B), \tag{4.9}$$

$$-\frac{dB}{d\xi} = \frac{2\xi^2}{m^2} \left(\frac{1}{k^{*2}} + \frac{\xi^2}{2m} \right)^{-4} \exp(-\xi^2 B), \qquad m(\xi \to \infty) = 1, \quad B(\xi \to \infty) = 0. \tag{4.10}$$

We note that both $m(\xi)$ and $B(\xi)$ are monotone functions.

There are two main regions. The outer region is defined by $\xi^2/2m \gg 1/k^{*2} = 1/g^{4/3}$. As the coupling increases $1/k^{*2}$ moves inward to a point $\ll 1$. We can show that $\xi^2/2m(\xi)$ is also a monotone function, increasing with increasing ξ. For small coupling constant g the point ξ_0 defined by $\xi_0^2/2m(\xi_0) = 1/g^{4/3}$ is at $\xi_0 \gg 1$. As g increases $\xi_0(g)$ moves closer to zero and ξ_0 becomes $\ll 1$.

In the large ξ region the differential equations

$$-\frac{dm}{d\xi} = 32\,\frac{m^3}{\xi^4}\exp(-\xi^2 B), \tag{4.11}$$

$$-\frac{dB}{d\xi} = \frac{32}{\xi^6}\,m^2\exp(-\xi^2 B), \tag{4.12}$$

are independent of the coupling constant. The equations may be integrated inward to some point ξ_0. In moderate and strong coupling ξ_0 is of order unity. This is easily done by noting that in the outer region

$$\frac{dm}{d\xi} = \xi^2 m\,\frac{dB}{d\xi} \tag{4.13}$$

so

$$m(\xi) = \exp\left(-\int_{\xi}^{\infty}\xi^2\,\frac{dB}{d\xi}\,d\xi\right). \tag{4.14}$$

We find the differential equation for $B(\xi)$,

$$\left(\frac{d}{d\xi}+\frac{6}{\xi}\right)\frac{dB}{d\xi} = \xi^2\,\frac{dB}{d\xi}\left(\frac{d}{d\xi}-\frac{2}{\xi}\right)B. \tag{4.15}$$

This yields the expansion for large ξ

$$B = \frac{32}{5}\,\frac{1}{\xi^5}\left(1+\frac{28}{3}\,\frac{1}{\xi^3}+\frac{32(28)(106)}{(66)(15)\,\xi^5}+\cdots\right), \tag{4.16}$$

$$m^2 = 1+\frac{64}{3}\,\frac{1}{\xi^3}+\frac{(64)}{3}(32)\left(\frac{2}{3}-\frac{1}{10}\right)\frac{1}{\xi^6}+\cdots. \tag{4.17}$$

In coupling strengths $g\sim 2-10$ these series may be used to find $B(\xi_0)$, $m(\xi_0)$ at $\xi_0\sim 4$ and the differential equations can be integrated numerically in the inner region.

In the extreme inner region we have

$$B - B(\xi_1) = -\frac{k^{*2}}{2}\,\frac{1}{m(\xi)}-\frac{1}{m(\xi_1)}, \tag{4.18}$$

where $\xi_1 \ll 1$. The region ξ_1 to ξ_0 can be treated numerically. The limiting values for small ξ are given by

$$\begin{aligned} m(\xi) &= m(0) - \tfrac{4}{3}\xi^3(k^*)^6+\cdots, \\ B(\xi) &= B(0) - \frac{2}{3}\,\frac{\xi^3(k^*)^8}{m(0)^2}+\cdots. \end{aligned} \tag{4.19}$$

It is not profitable to study the asymptotic forms in the limit of strong coupling since we do not expect the theory to be correct in that limit. The averaging procedure that has been used in the continuous compounding case is so drastic that it throws away important correlations. The present method also has no apparent relation to the well known physical picture of the strong coupling limit.[5]

5. CONCLUSIONS

We comment on extensions of the work reported here, some of which have been explored in detail. In the case of the single slice theory one can treat the low wave vector domain by more accurate methods than IC. For example, if K_1 is independent of a coupling constant a method such as that of Lee and Pines[4] yields the lower bound limit in strong coupling and is also accurate in the weak coupling limit. The Hamiltonian for this domain can also be treated by adiabatic strong coupling techniques of Pekar or Bogolyubov–Tyablikov.[5] If however we are concerned with a problem with an infinite cutoff we can take the slice K_1 to be coupling constant dependent. If K_1 increases at least as fast as g^2 the high momentum region can be accurately treated by IC. But then the Lee–Pines[4] method does not give the strong coupling limit for the low wave vector domain (which grows with increasing coupling). One must use the adiabatic method. The resulting theory is closely related but not identical to the author's transition theory[6] and also handles a certain class of ultraviolet divergent theories.

We have noted that in the continuous case, i.e., the use of infinitely many unitary transforms, one loses accuracy because of the averaging steps. It appears to be difficult to hold the operator parts and to organize them in a definite way as the unitary transforms are performed. This has to be done if the theory is to be systematic. An alternative approach is to take a finite number of bands whose width is coupling constant dependent. Each band is treated by canonical transformations and the modified Hamiltonian for the other bands can be obtained.

In summary, in the present paper we have studied an example in which explicit unitary transforms have been successful in treating the interaction of a particle with oscillators with a spectrum of wavelengths. We have shown that treating the interaction with different wavelength ranges sequentially leads to improved results.

[1] K. Wilson, Rev. Mod. Phys. 47, 773 (1975).
[2] P.W. Anderson, J. Phys. C 3, 2436 (1970).
[3] T.D. Lee, F.E. Low, and D. Pines, Phys. Rev. 90, 297 (1953); S.V. Tyablikov, Zh. Eksp. Teor. Fiz. 25, 688 (1953); M. Gurari, Phil. Mag. 44, 329 (1953).
[4] T.D. Lee and D. Pines, Phys. Rev. 92, 883 (1953).
[5] N.N. Bogolyubov and S.W. Tyablikov, Zh. Eksp. Teor. Fiz. 19, 256 (1949); G.R. Allcock, Adv. Physics 5, 412 (1956); S.I. Pekar, *Untersuchungen u.d. Electronetheorie der Kristalle* (Akademie-Verlag, Berlin, 1954).
[6] E.P. Gross, Ann. Phys. (N.Y.) 8, 78 (1959); 19, 219 (1962); Phys. Rev. D 12, 421 (1975).

Magnetic monopoles in SU(4) gauge theories

Y. Brihaye and J. Nuyts

Faculté des Sciences, Université de Mons, Belgium
(Received 28 December 1976)

All spherically symmetric magnetic poles are explictly constructed in SU(4). We show the relevance of the concept of the little group which transforms spherically symmetric solutions among themselves. Unlike the SU(2) and SU(3) situation this little group is not always Abelian in SU(4). On the other hand, it is shown that while the total strength of the pole is always quantized its projection along the Higgs field is sometimes arbitrary. This is also in contrast with the SU(2) and SU(3) cases.

I. INTRODUCTION

Many papers[1] generalizing 't Hooft[2] initial construction have solved the problem of finding magnetic poles in gauge theories with symmetry groups H larger than SU(2). We here present in great detail the SU(4) case.

We want to put the emphasis on the underlying invariance properties of the explicit solutions. In particular we will show the crucial role played by two groups and their algebra which we will call respectively the diagonal group D and the little group ℓ.

(a) The diagonal group D

Let E_i $(i=1,2,3)$ be a subalgebra of the algebra of H with the commutation relation of SU(2),

$$[E_i, E_j] = i\epsilon_{ijk}E_k. \tag{1.1}$$

This corresponds in the global group H to a subgroup which could be SU(2) or $SO(3)=SU(2)/Z_2$. Let J_i $(i=1,2,3)$ be the generators of the space rotation group (including spin if needed). Clearly

$$J_i = i\epsilon_{ijk}X_j\partial_k + (\text{spin part}), \tag{1.2}$$

$$[J_i, J_j] = i\epsilon_{ijk}J_k, \tag{1.3}$$

$$[J_i, E_j] = 0. \tag{1.4}$$

The diagonal group D has generators defined by

$$L_i = J_i + E_i \tag{1.5}$$

and has the commutation relations (1.1) of the algebra of SU(2). Use of (1.5) clearly mixes internal indices with space indices. The solutions we are looking for are static, i.e., they have no time dependence and are "spherically symmetric," i.e., they are annihilated by L_i. In other words, the solutions are time-independent singlets under the diagonal group.

(b) The little group ℓ

Let us consider the vector space of the algebra of H with position, $\mathbf{x}$, dependent scalars, and the subset ℓ of these operators which commute with the L_i,

$$[\ell, L_i] = 0. \tag{1.6}$$

These operators, once normalized, generate an algebra which will be called the little group algebra.

It is clear from their definition that little group transformations transform spherically symmetric solutions into spherically symmetric solutions. One particularly important operator belonging to ℓ is E, the radial component of E_i [see (2.9) for notations],

$$E = \hat{x}_i E_i. \tag{1.7}$$

We stress here that this little group ℓ is not the little group ℓ_Φ which is generally used. To make the connections more clearly we will introduce three other little groups:

(i) ℓ_Φ will denote the generators of the little group of the Higgs fields. ℓ_Φ correspond to the little group of the vacuum. Its generators annihilate Φ,

$$\ell_\Phi \Phi = 0. \tag{1.8}$$

Their direction in SU(4) correspond to the remaining mass zero vector bosons.

(ii) ℓ_G will denote the generators of the little group of $G_{\mu\nu}$,

$$\ell_G G_{\mu\nu} = 0. \tag{1.9}$$

(iii) ℓ_S will denote the intersection of ℓ_Φ and ℓ_G. It is the true little group of the full solution,

$$\ell_S = \ell_G \cap \ell_\Phi. \tag{1.10}$$

The paper is organized as follows. In Sec. II, we define our notation, the type of solution we try to find, the four inequivalent embeddings of the algebra E_i in SU(4), and show how the calculation can be simplified by subtraction from W_μ a specific term S_μ. We end with some general considerations. In Sec. III we study in turn the four embeddings. In every instance we present:

(i) the generators of SU(4) decomposed with respect to E_i,

(ii) the little group generators ℓ,

(iii) the natural basis vectors for $\overline{W}_\mu$ (2.16) and the change of basis suggested by ℓ. These basis vectors are identical for G_ρ (3.A.27) and $D_\mu\Phi$.

(iv) $G_{\mu\nu}$ and the Lagrangian,

(v) pointlike solutions for $G_{\mu\nu}$,

(vi) Higgs scalars-potential,

(vii) Higgs scalars, kinetic term,

(viii) discussion.

In Sec. IV all the solutions are discussed and the conclusions are obtained.

 Copyright © 1977 American Institute of Physics

II. NOTATION-SU(2) SUBALGEBRA–INHOMOGENEOUS SUBSTRACTION

In this section we recall some known facts. First we give the gauge invariant equation and discuss the static solutions we want to find. Then we identify in SU(4) all possible inequivalent SU(2) subalgebras [the E_i of (1.1)]. We find that there are four distinct such embeddings. Finally we explain why it is convenient to separate W_μ into two pieces, one piece S_μ the inhomogeneous substraction being written explicitly. This simplifies the computations considerably and makes the little group properties clearer.

A. Notation

In order to specify the notation let us recall the SU(4) gauge invariant Lagrangian L

$$L = -\tfrac{1}{4}(G^a_{\mu\nu})^2 - \tfrac{1}{2}(D_\mu\Phi)^2 - V(\Phi) \tag{2.1}$$

for the fifteen gauge fields W^a_μ and the Higgs fields Φ,

$$G^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + e f^{abc} W^b_\mu W^c_\nu, \tag{2.2}$$

where f^{abc} are the (completely antisymmetric) SU(4) structure constants. The covariant derivatives

$$D_\mu = \partial_\mu - ieW^a_\mu H^a, \tag{2.3}$$

with the H^a the appropriate generators corresponding to the representation chosen for the Higgs fields. To be specific we will later also assume that Φ^a belongs to the adjoint 15-dimensional representation. Many arguments in the paper however do not depend on this hypothesis.

Introducing λ^a ($a = 1, \ldots, 15$) as the (basic) four-dimensional representation of SU(4), it is convenient to associate in a one-to-one correspondence a four by four traceless matrix W_μ (or $G_{\mu\nu}$ or ϕ) to the 15 components W^a_μ (or $G^a_{\mu\nu}$ or Φ^a) by

$$W_\mu = \sum_{a=1}^{15} W^a_\mu \lambda_a, \tag{2.4}$$

$$W^a_\mu = \tfrac{1}{2}\,\mathrm{Tr}(W_\mu \lambda_a). \tag{2.5}$$

In this notation, for $G_{\mu\nu}$ and for a 15-dimensional Φ,

$$\begin{aligned} G_{\mu\nu} &= \partial_\mu W_\nu - \partial_\nu W_\mu + (e/i)[W_\mu, W_\nu], \\ D_\mu\Phi &= \partial_\mu\Phi + (e/i)[W_\mu, \Phi], \end{aligned} \tag{2.6}$$

while the space dependent SU(4) gauge transformations ($MM^* = 1$) are

$$\begin{aligned} W'_\mu &= MW_\mu M^* + (i/e)M\partial_\mu M^*, \\ G'_{\mu\nu} &= MG_{\mu\nu}M^*, \quad \Phi' = M\Phi M^*. \end{aligned} \tag{2.7}$$

We want to find asymptotic (pointlike) time independent solutions. By time independent we mean that

$$W_0 = 0 \tag{2.8}$$

and that no explicit time dependence appears (static solutions). The problem reduces to a three-dimensional x_i ($i = 1, 2, 3$) one. We choose the Euclidean metric. It is convenient to work with the following variables $\hat{x}_i, r$:

$$r^2 = \sum_i x_i^2, \quad \hat{x}_i = x_i/r. \tag{2.9}$$

Asymptotically the three terms in (2.1) will behave differently,

$$\begin{aligned} &V(\Phi) \;\; r \to \infty \;\; \text{const}, \\ &D_\mu\Phi \;\; r \to \infty \;\; \frac{\text{const}}{r}, \quad \text{(the constants may be } \hat{x}_i \text{ dependent)} \\ &G_{\mu\nu} \;\; r \to \infty \;\; \frac{\text{const}}{r^2}. \end{aligned} \tag{2.10}$$

Hence, in this approximation, all three terms in (2.1) have to be maximized separately. In all cases we will focus our attention first on $G_{\mu\nu}$, and only later study the two other terms. In particular, it will turn out in all instances, as is well known, that

$$D_\mu\Phi = 0, \tag{2.11}$$

but this equation will not be used as a starting point.

B. SU(2) subalgebra

In order to define the diagonal subgroup we need to know all possible ways to extract from the algebra of SU(4) the SU(2) subalgebra E_i. Up to an equivalence this is done quite easily by considering the four-dimensional (faithful) representation. Under SU(2) this representation splits into irreducible representations of SU(2). All possible embeddings thus correspond to partitions of 4, namely,

$$4 \to 4, \tag{2.12a}$$

$$4 \to 1 \oplus 3, \tag{2.12b}$$

$$4 \to 2 \oplus 2, \tag{2.12c}$$

$$4 \to 1 \oplus 1 \oplus 2, \tag{2.12d}$$

$$4 \to 1 \oplus 1 \oplus 1 \oplus 1. \tag{2.12e}$$

The possibility (2.12b), for example, means that 4 decomposes in a triplet plus a singlet under SU(2). The (2.12e) possibility is clearly excluded since it means that SU(2) does not act at all on the representation. It is easy to see that the other possibilities (2.12a)–(2.12d) can be realized. These decompositions are unique up to an equivalence (inner automorphism).

Since the representations of SU(2) are all self conjugate, the decomposition of $\bar{4}$ [the representation conjugate to 4 in SU(4)] is identical. From these remarks it is then easy to construct the decomposition of any other representation of SU(4). In particular since the adjoint (15-dimensional) representation is obtained in

$$4 \otimes \bar{4} = 1 \oplus 15. \tag{2.13}$$

The decomposition of 15 can be obtained by simple SU(2) decompositions as

$$15 \to 3 \oplus 5 \oplus 7, \tag{2.14a}$$

$$15 \to 1 \oplus 3 \oplus 3 \oplus 3 \oplus 5, \tag{2.14b}$$

$$15 \to 1 \oplus 1 \oplus 1 \oplus 3 \oplus 3 \oplus 3 \oplus 3, \tag{2.14c}$$

$$15 \to 1 \oplus 1 \oplus 1 \oplus 1 \oplus 2 \oplus 2 \oplus 2 \oplus 2 \oplus 3. \tag{2.14d}$$

These results will be used extensively later.

C. Inhomogeneous subtraction

We have seen that under SU(4) gauge transformations M the gauge fields transform (2.7) with an inhomogeneous part. For certain subsets of gauge transformations which will turn out to be those of the little group ℓ it is possible to find a solution S_μ of the following matrix equation:

$$S_\mu = MS_\mu M^\dagger - (i/e)M\partial_\mu M^\dagger. \tag{2.15}$$

If one then defines

$$\bar{W}_\mu = W_\mu + S_\mu \tag{2.16}$$

under the little group transformations one has

$$\bar{W}'_\mu = M\bar{W}_\mu M^\dagger \quad (M \in \ell), \tag{2.17}$$

i.e., there is no inhomogeneous part.

It turns out that for all known cases

$$S_\mu = (1/er)\epsilon_{\mu ij}\hat{x}_j E_i. \tag{2.18}$$

This is checked easily for the group transformations generated by E of (1.7).

It is precisely the property (2.17) which enables one to greatly simplify the computation if one works with $\bar{W}_\mu$ rather than W_μ.

D. General considerations

Let us now stress two important results which are valid when ϕ belongs to the adjoined representation.

First it will turn out in all instances that maximal $G_{\mu\nu}$ [G_ρ or better G; see (3.A.27) and (3.A.40) for notation] lie as ϕ in the little group ℓ directions and moreover that

$$[G, \phi] = 0. \tag{2.19}$$

This implies that they can be diagonalized together and the eigenvalues of both matrices [remember (2.4)] are of great importance.

Also the other quantity which plays a major role is

$$g = \frac{\mathrm{tr}(G\phi)}{[\mathrm{tr}(\phi^2)]^{1/2}}. \tag{2.20}$$

It is the strength of the magnetic pole in the direction of ϕ as opposed to the total strength as given by

$$g^2_{\mathrm{tot}} = \mathrm{tr}(G^2). \tag{2.21}$$

III. THE FOUR CASES

We now turn to the explicit solutions of the equations in the four cases A, B, C, and D. We discuss in turn the three parts of (2.1).

A. The 4 → 4 embedding

(i) Decomposition of the generators of SU(4)

There is only one way up to a SU(4) equivalence to embed the four-dimensional representation ($S = \frac{3}{2}$) of SU(2) in the four-dimensional representation of SU(4). A specific set of properties of E_i is given in Appendix A. The remaining 12 generators of SU(4) decompose according to (2.14a) in a 5-plet K_{ij} ($K_{ij} = K_{ji}$, $\sum_i K_{ii} = 0$),

$$K_{ij} = \tfrac{1}{2}(E_i E_j + E_j E_i) - \tfrac{1}{3}\delta_{ij}(E_m E_m), \tag{3.A.1}$$

and a completely symmetrical 7-plet

$$\begin{aligned} &N_{ijk}\Big(\sum_i N_{iik} = 0\Big), \\ &N_{ijk} = \tfrac{1}{6}(E_i E_j E_k + E_j E_k E_i + E_k E_i E_j \\ &\qquad + E_j E_i E_k + E_i E_k E_j + E_k E_j E_i) \\ &\qquad - \tfrac{41}{60}(\delta_{kj}E_i + \delta_{ki}E_j + \delta_{ij}E_k). \end{aligned} \tag{3.A.2}$$

The commutations relations of these operators and relevant traces are in Appendix A.

Let us also introduce the projections of these operators along the $\hat{x}_i$ direction (2.9) by

$$\begin{aligned} &E = \hat{x}_i E_i, \\ &K = \hat{x}_i K_i, \quad K_i = \hat{x}_j K_{ij}, \\ &N = \hat{x}_i N_i, \quad N_i = \hat{x}_j N_{ij}, \quad N_{ij} = \hat{x}_k N_{ijk}. \end{aligned} \tag{3.A.3}$$

(ii) The little group

The three operators $\{P\} = \{E, K, N\}$ generate the little group ℓ and commute. Hence

$$\ell = \{E, K, N\} = \mathrm{U}(1) \otimes \mathrm{U}(1) \otimes \mathrm{U}(1). \tag{3.A.4}$$

Yet it is useful to introduce a new basis for these generators by defining the three commuting generators $R \equiv \{A, B, C\}$ by

$$\begin{aligned} &A = -\tfrac{2}{5}E + N/3, \\ &B = \tfrac{3}{5}E + N/3, \\ &C = K/2. \end{aligned} \tag{3.A.5}$$

An infinitesimal transformation of ℓ is

$$M = 1 + i\pi P, \tag{3.A.6}$$

with the three *real r-independent* parameters $\pi \equiv \{\epsilon, \kappa, \eta\}$ or

$$M = 1 + i\rho R, \tag{3.A.7}$$

with the three *real* parameters $\rho \equiv \{\alpha, \beta, \gamma\}$

$$\begin{aligned} &\alpha = -\epsilon + \tfrac{9}{5}\eta, \\ &\beta = \epsilon + \tfrac{6}{5}\eta, \\ &\gamma = 2\kappa. \end{aligned} \tag{3.A.8}$$

It is then easy to check that formula (2.17) and (2.18) are valid for M of the form (3.A.6) and (3.A.7).

(iii) Basis vectors for W_μ

We now want to construct the most general basis for static spherically symmetric W^a_μ ($W^a_0 = 0$, $\mu = 1, 2, 3$, $a = 1, \ldots, 15$). There are nine independent orthogonal basic vectors. Their corresponding matrices (2.4) are naturally chosen as

$$T^P_{1\mu} = \hat{x}_\mu P, \quad \{P\} = \{E, K, N\}, \tag{3.A.9a}$$

$$T^P_{2\mu} = (\hat{x}_\mu \hat{x}_i - \delta_{\mu i})P_i, \quad \{P_i\} = \{E_i, K_i, N_i\}, \tag{3.A.9b}$$

$$T^P_{3\mu} = \epsilon_{\mu ij}\hat{x}_j P_i. \tag{3.A.9c}$$

Orthogonality is defined by the scalar product [see (A10) for the norms]

$$\sum_{\mu} \mathrm{Tr}(T^{P}_{i\mu} T^{P'}_{i'\mu}). \tag{3.A.10}$$

In this basis W_μ, $\bar{W}_\mu$ are written

$$W_\mu = W^P_i T^P_{i\mu}, \quad \bar{W}_\mu = \bar{W}^P_i T^P_{i\mu}, \tag{3.A.11}$$

where summation on P and i is understood.

The components W^P_i then depend upon r only. In writing (3.A.5) we have introduced the natural choice. In particular, in (3.A.9a) the tensor T_μ is in the direction of $\hat{x}_\mu$ while in (3.A.9b) and (3.A.9c) T_μ are orthogonal in directions orthogonal to $\hat{x}_\mu$. In other words, (3.A.9a) is the radial component and (3.A.9b) and (3.A.9c) are the tangential components. This distinction is well suited for spherically symmetric solutions.

Consideration of ℓ however suggests (see Appendix E) another choice for the basis vector of $\bar{W}_\mu$ [(2.16) and (2.17)],

$$\begin{aligned}\bar{W}_\mu &= \bar{W}_{(z)} \bar{T}^{(z)}_\mu \\ &= W_{(1P)} T^{(1P)}_\mu + \bar{W}_{(A)} \bar{T}^{(A)}_\mu + \bar{W}_{(AB)} \bar{T}^{(AB)}_\mu \\ &= W_{(1P)} T^{(1P)}_\mu + \tilde{W}^t \tilde{T}_\mu + \mathrm{tr}(\tilde{\tilde{W}}^t \tilde{\tilde{T}}_\mu),\end{aligned} \tag{3.A.12}$$

$\bar{T}^{(z)}_\mu$ is the new set of nine basis vectors ($\bar{T}^{(1P)}_\mu$, $\tilde{T} \equiv \bar{T}^{(A)}_\mu$, $\tilde{\tilde{T}} \equiv \bar{T}^{(AB)}_\mu$; $A, B = 1, 2$) which we now define.

The set $\bar{T}^z_\mu$ has *simple transformation properties* under ℓ in its new basis (A, B, C) (see Appendix E). It is composed of the *three singlets* $(T^{(1P)}_\mu)$

$$\bar{T}^{(1P)}_\mu \equiv T^P_{1\mu}, \tag{3.A.13}$$

with $P = \{E, K, N\}$ or $P = \{A, B, C\}$, and

$$\bar{W}_{(1P)} = \bar{W}^P_1, \tag{3.A.14}$$

of a *doublet* under the A transformation

$$\tilde{T}_\mu \equiv \bar{T}^{(A)}_\mu \begin{pmatrix} \bar{T}^{(1)}_\mu = \frac{2}{5} T^E_{2\mu} - T^W_{2\mu} \\ \bar{T}^{(2)}_\mu = -\frac{2}{5} T^E_{3\mu} + T^N_{3\mu} \end{pmatrix} \tag{3.A.15}$$

with

$$\tilde{W} \equiv \bar{W}_{(A)} \equiv \begin{pmatrix} \bar{W}_{(1)} = \bar{W}^E_2 - \frac{3}{5} \bar{W}^N_2 \\ \bar{W}_{(2)} = -\bar{W}^E_3 + \frac{3}{5} \bar{W}^N_3 \end{pmatrix}, \tag{3.A.16}$$

the remaining four tensors $\tilde{\tilde{T}}_\mu$ form a $(2 \otimes 2)$ *multiplet* under $B \otimes C$ transformations

$$\tilde{\tilde{T}}_\mu = \bar{T}^{(AB)}_\mu \equiv \begin{pmatrix} \bar{T}^{(11)}_\mu = \frac{3}{5} T^E_{3\mu} + T^N_{3\mu} & \bar{T}^{(12)}_\mu = T^K_{2\mu} \\ \bar{T}^{(21)}_\mu = -\frac{3}{5} T^E_{2\mu} - T^N_{2\mu} & \bar{T}^{(22)}_\mu = T^K_{3\mu} \end{pmatrix}, \tag{3.A.17}$$

with

$$\tilde{\tilde{W}} \equiv \bar{W}_{(AB)} \equiv \begin{pmatrix} \bar{W}_{(11)} = \bar{W}^E_3 + \frac{2}{5} \bar{W}^N_3 & \bar{W}_{(12)} = \bar{W}^K_2 \\ \bar{W}_{(21)} = -\bar{W}^E_2 - \frac{2}{5} \bar{W}^N_2 & \bar{W}_{(22)} = \bar{W}^K_3 \end{pmatrix}. \tag{3.A.18}$$

By considering global transformations

$$M_\rho = \begin{pmatrix} \cos\rho & \sin\rho \\ -\sin\rho & \cos\rho \end{pmatrix} \tag{3.A.19}$$

corresponding to two-dimensional representations for the three U(1) groups, then

$$\tilde{W}' = M_\alpha \tilde{W}, \tag{3.A.20}$$

$$\tilde{\tilde{W}}' = M_\beta \tilde{\tilde{W}} M^{-1}_\gamma. \tag{3.A.21}$$

This can be proved easily by using (2.17) and exponentiating the transformation (3.A.7) (3.A.8). These considerations will not be repeated for cases B, C, and D.

Finally before going to the $G_{\mu\nu}$ it is useful to factorize for all W a term $(er)^{-1}$,

$$W = (1/er) V. \tag{3.A.22}$$

Asymptotic conditions (2.10) will then correspond to constant V's.

Elementary group properties then tell us the results:

(a) The following quantities are ℓ invariant:

$$I_1 = \mathrm{Tr}(\tilde{\tilde{V}}^t \tilde{\tilde{V}}), \tag{3.A.23}$$

$$I_2 = \det(\tilde{\tilde{V}}), \tag{3.A.24}$$

$$I_3 = (\tilde{V}^t \tilde{\tilde{V}} \tilde{V}). \tag{3.A.25}$$

(b) Multiplication of $\tilde{V}$ (or $\tilde{W}$) on the left ($\tilde{\tilde{V}}$ or $\tilde{\tilde{W}}$ on either side) by $i\sigma_2$ does not change their transformation properties since

$$(i\sigma_2) M (i\sigma_2)^{-1} = M. \tag{3.A.26}$$

(iv) $G_{\mu\nu}$ and the Lagrangian

In this subsection we present the general form of the spherically symmetric static $G_{\mu\nu}$'s and find the asymptotic (maximal) solutions.

Since $G_{\mu\nu}$ is antisymmetric with nonvanishing space components only we define

$$G_\rho = \tfrac{1}{2} \epsilon_{\rho\mu\nu} G_{\mu\nu}, \quad G_{\mu\nu} = \epsilon_{\mu\nu\rho} G_\rho \tag{3.A.27}$$

whose transformation properties are

$$G'_\rho = M G_\rho M^+. \tag{3.A.28}$$

The basis for spherically symmetric G_ρ's will thus be identical to the basis for $\bar{W}_\mu$ [i.e. (3.A.9) or (3.A.13), (3.A.15), and (3.A.17)]. Hence under the little group the 15 components of G_ρ decompose as three singlets $G_{(1P)}$, an A doublet $\tilde{G}$, and a $B \otimes C$-$2 \otimes 2$ representation $\tilde{\tilde{G}}$. After some computations one finds [(2.6), (3.A.27)]

$$G_\rho = G_{(z)} \bar{T}^{(z)}_\rho \tag{3.A.29}$$

with the values [see (3.A.13)–(3.A.18) and (3.A.26)]. First

$$G_{(1E)} = \frac{1}{er^2} \left(\frac{3}{5} I_1 + \frac{2}{5} I_3 - 1 \right), \tag{3.A.30}$$

$$G_{(1K)} = \frac{3}{er^2} I_2, \tag{3.A.31}$$

$$G_{(1N)} = \frac{2}{er^2} (I_1 - I_3), \tag{3.A.32}$$

the three invariant components. The A doublet is

TABLE I. The six maximal solutions (a–f) [$\ell \equiv U(1) \otimes U(1) \otimes U(1)$] for the $(4 \to 4)$ spherically symmetric $G_{\mu\nu}$. The defining equations are: for I_i (3.A.23)–(3.A.25), for G (3.A.27), (3.A.40), (3.A.5), for g_i (3.A.41), for ℓ_G(1.9), and for g^2_{tot} (2.21), (3.A.42). $m \otimes n$ is a shorthand for the Lie algebra of $SU(m) \otimes SU(n)$ except when m or n equals one where it means the Lie algebra of U(1).

	a	b	$c^{\pm}$		$d^{\pm}$		e	f
I_1	1	$\frac{1}{3}$	$\frac{1}{3}$		$\frac{1}{6}$		0	0
I_2	0	0	$-\frac{1}{6}\epsilon$		$-\frac{1}{12}\epsilon$		0	0
I_3	1	0	$\frac{1}{2}$		0		$\frac{1}{4}$	0
G	0	$2A$	$-(B+\epsilon C)$		$\frac{1}{2}[3A-(B+\epsilon C)]$		$-\frac{3}{2}B$	$A-B$
			$\epsilon=+1$	$\epsilon=-1$	$\epsilon=+1$	$\epsilon=-1$		
g_1	0	-1	$-\frac{3}{2}$	$-\frac{1}{2}$	$-\frac{3}{2}$	-1	$-\frac{3}{2}$	$-\frac{3}{2}$
g_2	0	-1	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	-1	0	$-\frac{1}{2}$
g_3	0	1	$\frac{1}{2}$	$-\frac{1}{2}$	1	$\frac{1}{2}$	0	$\frac{1}{2}$
g_4	0	1	$\frac{1}{2}$	$\frac{3}{2}$	1	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$
ℓ_G	4	$2\otimes2\otimes1$	$3\otimes1$		$2\otimes1\otimes1$		$2\otimes1\otimes1$	$1\otimes1\otimes1$
g^2_{tot} unit e^{-2}	0	4	3		$\frac{9}{2}$		$\frac{9}{2}$	5

$$\tilde{G} = \frac{1}{er^2}[-(i\sigma_2)r\dot{\tilde{V}} - \overline{V}_{(1A)}\tilde{V}], \tag{3.A.33}$$

where the dot means differentiation with respect to r; while the $B\otimes C$, $2\otimes 2$ representation is [(E4)]

$$\tilde{\tilde{G}} = \frac{1}{er^2}[-(i\sigma_2)r\dot{\tilde{\tilde{V}}} + \overline{V}_{(1B)}\tilde{\tilde{V}} - \overline{V}_{(1C)}(i\sigma_2)\tilde{\tilde{V}}(i\sigma_2)]. \tag{3.A.34}$$

The covariance of the formulas under the ℓ transformation is obvious, (3.A.26).

We now turn to the Lagrangian and to maximization. First using (2.4), (3.A.27), and the normalization of the $\overline{T}$ tensors from Appendix A,

$$\begin{aligned} L &= -\tfrac{1}{4}G^{(a)}_{\mu\nu}G^{(a)}_{\mu} \\ &= -\tfrac{1}{4}\mathrm{Tr}(G_\rho G_\rho) \\ &= -\tfrac{1}{4}[5G^2_{(1E)} + 4G^2_{(1K)} + \tfrac{9}{5}G^2_{(1N)} \\ &\quad + 4(\tilde{G}^t\tilde{G}) + 6\mathrm{Tr}(\tilde{\tilde{G}}^t\tilde{\tilde{G}})]. \end{aligned} \tag{3.A.35}$$

As is clear from (3.A.30)–(3.A.34), the $\overset{\circ}{\overline{V}}_{(1P)}$ do not appear. This means that the $\overline{V}_{(1P)}$ are simply constrained. After some straightforward computation one can eliminate these components. At the same time three of the remaining six derivatives of $\tilde{V}$ and $\tilde{\tilde{V}}$ will be eliminated. Hence only three linearly independent combinations of these six derivatives will survive, namely $\dot{I}_1$, $\dot{I}_2$, and $\dot{I}_3$. Indeed if we define

$$\begin{aligned} C_1 &= -\mathrm{Tr}(\dot{\tilde{\tilde{V}}}\tilde{\tilde{V}}^t i\sigma_2), \\ C_2 &= -\mathrm{Tr}(\dot{\tilde{\tilde{V}}}(i\sigma_2)\tilde{\tilde{V}}), \\ C_3 &= -(\dot{\tilde{V}}^t i\sigma_2\tilde{V}), \end{aligned} \tag{3.A.36}$$

then

$$\begin{aligned} \overline{V}_{(1E)} &= \frac{r}{I_1^2 - 4I_2^2}(C_1 I_1 - 2C_2 I_2), \\ \overline{V}_{(1K)} &= \frac{r}{-2(I_1^2 - 4I_2^2)}(2C_1 I_2 - C_2 I_1), \\ \overline{V}_{(1N)} &= +\frac{r}{I_3}C_3. \end{aligned} \tag{3.A.37}$$

Replacing these values back in the Lagrangian one obtains

$$\begin{aligned} L = &-\frac{1}{4e^2r^2}\left[\frac{\dot{I}_3^2}{I_3} + \frac{6}{4}\left(\frac{\dot{I}_+^2}{I_+} + \frac{(\dot{I}_-^2)}{I_-}\right)\right] \\ &-\frac{1}{4e^2r^4}[\tfrac{1}{5}[3(I_+ + I_-) + 2I_3 - 5]^2 + 9(I_+ - I_-)^2 \\ &+ \tfrac{36}{5}[(I_+ + I_-) - I_3]^2] \end{aligned} \tag{3.A.38}$$

with

$$I_\pm = (I_1 \pm 2I_2)/2. \tag{3.A.39}$$

The form (3.A.38) together with the definition of the invariants (3.A.23)–(3.A.25), and (3.A.39) is particularly simple and elegant and shows the importance of the little group properties.

A final remark is useful here. Asymptotically [(2.10)] V will behave as a constant. Thus in this approximation G_ρ will correspond to a field which is of the magnetic type with its nonzero components given by (3.A.30)–(3.A.32). At this point G_ρ is still a matrix whose physical interpretation will be given later. Asymptotically

$$G_\rho = \hat{x}_\rho G/er^2 \tag{3.A.40}$$

and the eigenvalues of G are g_i:

$$\begin{aligned} g_1 &= \tfrac{3}{2}(I_1 + 2I_2 - 1), \\ g_2 &= \tfrac{1}{2}(-3I_1 - 6I_2 + 4I_3 - 1), \\ g_3 &= \tfrac{1}{2}(3I_1 - 6I_2 - 4I_3 + 1), \\ g_4 &= \tfrac{3}{2}(-I_1 + 2I_2 + 1). \end{aligned} \tag{3.A.41}$$

TABLE II. The six solutions (1–6) for the $(4\to 4)$ spherically symmetric Higgs fields. The defining equations are for Φ (3.A.5), for I_i (3.A.23–25) for Φ_i (3.A.44), for g^2 (2.20), for ℓ_Φ (1.8), and for $m\otimes n$ see Table I.

	1	2	3±		4±		5	6
Φ	0	$\Phi_A A$	$\Phi_B(B+\epsilon C)$		$\Phi_A A+\Phi_B(B+\epsilon C)$		$\Phi_B B+\Phi_C C$	
			$\epsilon=+1$	$\epsilon=-1$	$\epsilon=+1$	$\epsilon=-1$		
Φ_i	0	$\Phi_1=\Phi_2$	$\Phi_2=\Phi_3$	$\Phi_1=\Phi_2$	$\Phi_3=\Phi_4$	$\Phi_1=\Phi_2$	$\Phi_2=\Phi_3$	arbitrary
		$\Phi_3=\Phi_4$	$=\Phi_4$	$=\Phi_3$				
I_i	arbitrary	$I_3=0$	$I_1=-2\epsilon I_2$		$I_1=-2\epsilon I_2$		$I_1=I_2$	$I_1=I_2=I_3$
					$I_3=0$		$=0$	$=0$
ℓ_Φ	4	$2\otimes 2\otimes 1$	$3\otimes 1$		$2\otimes 1\otimes 1$		$2\otimes 1\otimes 1$	$1\otimes 1\otimes 1$

The norm of G is defined by

$$g_{\text{tot}}=[\text{tr}(GG)]^{1/2}=\left(\sum_i g_i^2\right)^{1/2}. \tag{3.A.42}$$

It will correspond to the strength of the monopole.

(v) Pointlike solution

Pointlike asymptotic solutions are obtained by maximalizing the nonderivative part of (3.A.38) with respect to the parameters $\vec{V}$ and $\vec{\tilde{V}}$. One obtains six different types of solutions which are given in Table I together with relevant properties. It is remarkable that the results are expressed most simply in the (A,B,C) basis (3.A.5).

(vi) Higgs scalars—potential term

The Higgs scalar ϕ could belong to any representation of SU(4). To be specific we will treat here the case where ϕ transforms with the adjoint representation of SU(4).

The most general form of a spherically symmetric solution is [see (3.A.5)]

$$\begin{aligned}
&\phi=\phi_E E+\phi_K K+\phi_N N\equiv\phi_P P\\
&\quad=\phi_A A+\phi_B B+\phi_C C\equiv\phi_R R,\\
&\phi_A=-\phi_E+\tfrac{9}{5}\phi_N,\\
&\phi_B=\phi_E+\tfrac{6}{5}\phi_N,\\
&\phi_C=2\phi_K,
\end{aligned} \tag{3.A.43}$$

with ϕ_P or ϕ_R depending on r only. General theorems about the maximalizations of the Higgs potential V, (2.1), tells us that ϕ is generally maximal when it has one or more equal eigenvalues. Nonequal eigenvalues are also possible but correspond to less interesting situations. The eigenvalues of ϕ are

$$\begin{aligned}
&\phi_1=\tfrac{3}{2}\phi_E+\phi_K+\tfrac{3}{10}\phi_N,\\
&\phi_2=\tfrac{1}{2}\phi_E-\phi_K-\tfrac{9}{10}\phi_N,\\
&\phi_3=-\tfrac{1}{2}\phi_E-\phi_K+\tfrac{9}{10}\phi_N,\\
&\phi_4=-\tfrac{3}{2}\phi_E+\phi_K-\tfrac{3}{10}\phi_N.
\end{aligned} \tag{3.A.44}$$

It is not difficult to discuss all the possibilities; there are four types:

$$\phi_3\neq\phi_1=\phi_2\neq\phi_4\neq\phi_3, \tag{3.A.45a}$$

$$\phi_1=\phi_2=\phi_3=-\tfrac{1}{3}\phi_4, \tag{3.A.45b}$$

$$\phi_1=\phi_2\neq\phi_3=\phi_4, \tag{3.A.45c}$$

$$\text{all } \phi_i\text{'s are different}, \tag{3.A.45d}$$

and all the distinct permutations of the indices (1,2,3,4). The little group ℓ_ϕ (or stability group) of ϕ always contains the ℓ group but is in general larger. Corresponding to the four cases just mentioned, the little group *algebras* are isomorphic to

$$\text{SU(2)}\otimes\text{U(1)}\otimes\text{U(1)} \text{ locally}, \tag{3.A.46a}$$

$$\text{SU(3)}\otimes\text{U(1)} \text{ locally}, \tag{3.A.46b}$$

$$\text{SU(2)}\otimes\text{SU(2)}\otimes\text{U(1)} \text{ locally}, \tag{3.A.46c}$$

$$\text{U(1)}\otimes\text{U(1)}\otimes\text{U(1)}\otimes\text{U(1)} \text{ locally}. \tag{3.A.46d}$$

We now turn immediately to the maximalization of the kinetic terms of the Higgs fields. We will show that its maximalization establishes a precise correspondence between the extremal values of $G_{\mu\nu}$ and of ϕ.

(vii) Higgs scalars—kinetic term

In terms of the definition (3.A.43), (2.16), (3.A.12), and (3.A.23)–(3.A.25), the kinetic term of the Higgs fields become [(2.6)]

$$\begin{aligned}
\tfrac{1}{2}D_\mu\phi^{(a)}D_\mu\phi^{(a)}&=\tfrac{1}{4}\text{Tr}(D_\mu\phi D_\mu\phi)\\
&=\tfrac{1}{4}\Big[5\dot{\phi}_E^2+4\dot{\phi}_K^2+\tfrac{9}{5}\dot{\phi}_N^2\\
&\quad+\frac{6}{r^2}((\phi_B^2+4\phi_K^2)I_1+4\phi_B\phi_C I_2)\\
&\quad+\frac{4}{r^2}\phi_A^2 I_3\Big].
\end{aligned} \tag{3.A.47}$$

Asymptotically again the derivatives do not contribute while the maximalization with respect to $\vec{V}$, $\vec{\tilde{V}}$, and ϕ considered as constant leads to Table II. In any instance, $D_\mu\phi$ turns out to be identically zero in agreement with the general theorems, and ϕ commutes with G [see (2.19)].

(viii) Discussion

Comparing Table I and Table II leads to a correspondence between the asymptotic solutions. This is given in Table III. Following 't Hooft, the component of $G_{\mu\nu}$ in the direction of ϕ in internal space is called the electromagnetic monopole, its strength (3.A.40) is

TABLE III. The correspondence between the solutions for $G_{\mu\nu}$ and Φ in the $(4\to 4)$ case. The defining equations are: for $\ell_\Phi, \ell_G, \ell_S$ (1.8)—(1.10), for g^2 (2.20), (3.A.48), $g_\perp^2$ is defined as the strength of the component of the G_ρ^a orthogonal to the direction of Φ^a, for $m \otimes n$ see Table I.

	1(a–f)	2b	2d±	2f	(3c)±	(3d)±	(4d)±	5e	5f	
ℓ_Φ	4	$2\otimes 2\otimes 1$	$2\otimes 2\otimes 1$	$2\otimes 2\otimes 1$	$3\otimes 1$	$3\otimes 1$	$2\otimes 1\otimes 1$	$2\otimes 1\otimes 1$	$2\otimes 1\otimes 1$	$1\otimes 1\otimes 1$
$\ell_S=\ell_G$	see Table I	ℓ_Φ	$2\otimes 2\otimes 1$	$1\otimes 1\otimes 1$	ℓ_Φ	$2\otimes 1\otimes 1$	ℓ_Φ	ℓ_Φ	$1\otimes 1\otimes 1$	ℓ_Φ
g^2 unit e^{-2}	/	4	4	4	3	3	$\frac{(2\Phi_A-3\Phi_B)^2}{\Phi_A^2+3\Phi_B^2-2\Phi_A\Phi_B}$	$\frac{9\Phi_B^2}{2\Phi_B^2+\Phi_C^2}$	$\frac{9\Phi_B^2}{2\Phi_B^2+\Phi_C^2}$	g^2
$g_\perp^2$ unit e^{-2}	see Table I	0	$\frac{1}{2}$	1	0	$\frac{3}{2}$	$\frac{9}{2}-g^2$	$\frac{9}{2}-g^2$	$5-g^2$	$5-g^2$
g_{tot}^2 unit e^{-2}	see Table I	4	$\frac{9}{2}$	5	3	$\frac{9}{2}$	$\frac{9}{2}$	$\frac{9}{2}$	5	5

$$g=\frac{\mathrm{Tr}(G\phi)}{[\mathrm{Tr}(\phi^2)]^{1/2}} \tag{3.A.48}$$

but as is seen in Table III and as was already known in SU(3), G_ρ does not always lie in the direction of ϕ. This is unlike the SU(2) case where G_ρ and ϕ are parallel. The components of G_ρ orthogonal to ϕ lie in the little group ℓ. At this point it is useful to introduce ℓ_S: the little group of the full solution , i. e., the stability group of G and ϕ. The part of G_ρ orthogonal to ϕ will always be called a monopole because the ℓ little group, being $\mathrm{U(1)}\otimes\mathrm{U(1)}\otimes\mathrm{U(1)}$, allows no freedom for an ℓ isopole.

In SU(3) the angles between ϕ and G_ρ in SU(3) could only take special value. This leads to a quantization of charge. Here, as seen in Table III, a new freedom appears because the angle between ϕ and G_ρ in ℓ is sometimes arbitrary. No direct quantization follows! In other words, once the direction of charge is fixed the direction of the magnetic poles is still arbitrary. This is due to the fact that since the little group ℓ_ϕ is larger than U(1) there may exist more than one mass zero vector boson compatible with electromagnetism.

B. The $4 \to 3+1$ embedding

We will now repeat more briefly the computation for the case $4\to 3+1$. The procedure will be very analogous though the details are quite different.

(i) Decomposition of the generators of SU(4)

The identification of a three-dimensional (vector) representation of SU(2) in the four-dimensional representation of SU(4) is again unique up to equivalence. The E_i are given in Appendix B. Taking into account (2.14b), the remaining twelve generators of SU(4) decompose into a singlet which will be called Z, two triplets K_i, N_i, and one 5-plet F_{ij} [$\sum_i F_{ii}=0$, $F_{ij}=F_{ji}$ (see Appendix B)]. As in the preceding case we also define the projections of these generators along $\hat{x}_i$ directions,

$$\begin{aligned} &E=\hat{x}_i E_i\,, \quad K=\hat{x}_i K_i\,, \\ &N=\hat{x}_i N_i\,, \quad F=\hat{x}_i F_i, \quad F_i=\hat{x}_j F_{ij}\,. \end{aligned} \tag{3.B.1}$$

(ii) The little group

The operators $\{Q\}=\{Z,\{P\}\}$, $\{P\}=\{E,K,N,F\}$ generate the little group ℓ. These generators close in the algebra of $\mathrm{SU(2)}\otimes\mathrm{U(1)}\otimes\mathrm{U(1)}$ as can be seen by a change of basis. Let

$$\begin{aligned} &A_1=K, \quad A_2=N, \\ &A_3=\tfrac{1}{3}Z-F, \quad \{R\}=(A_i/2,B,C)=\mathrm{SU(2)}\otimes\mathrm{U(1)}\otimes\mathrm{U(1)}, \\ &B=E, \quad C=\tfrac{1}{3}Z+2F\,. \end{aligned} \tag{3.B.2}$$

$A_i/2$ have the commutation relations of SU(2) and commute with B and C. Let $R=\{A_i,B,C\}$. An infinitesimal transformation of ℓ is written

$$M=1+i\xi Q=1+i\rho R \tag{3.B.3}$$

with five real independent parameters $\xi=(\zeta,\epsilon,\kappa,\eta,\phi)$, $\rho=\{\alpha_i,\beta,\gamma\}$. These two sets are related by obvious relations.

(iii) Basis vectors for W_μ

Quite naturally the 13 basis vectors for $\overline{W}_\mu$ are written

$$\begin{aligned} &T_{1\mu}^R=\hat{x}_\mu R, \quad \{(R)\}=\{A_i,B,C\}, \\ &T_{2\mu}^P=(\hat{x}_\mu\hat{x}_i-\delta_{\mu i})P_i, \quad \{P_i\}=\{E_i,K_i,N_i,F_i\}, \\ &T_{3\mu}^P=\epsilon_{\mu ij}\hat{x}_j P_i\,. \end{aligned} \tag{3.B.4}$$

With these definitions the basis vector for $\overline{W}_\mu$ are

$$\overline{T}_{1\mu}^R=\hat{x}_\mu R\,, \quad \{R\}=\{A_i,B,C\}. \tag{3.B.5}$$

Under the little group transformation $T_{1\mu}^B$ and $T_{1\mu}^C$ are singlets while $T_{1\mu}^{A_i}$ form a triplet (vector) under the SU(2) (A generators). The remaining eight components behave as a tensor

$$\overline{T}_\mu^{A,B,C}, \quad A=(1,2), \quad B=(1,2), \quad C=(1,2), \tag{3.B.6}$$

transforming as a $2\otimes 2\otimes 2$ representation of $\mathrm{SU(2)}\otimes\mathrm{U(1)}\otimes\mathrm{U(1)}(A_i,B,C)$. For example, for B and C fixed, $\overline{T}_\mu^{A,B,C}$ ($A=1,2$) behaves as a spinor under A_i transformations. The precise correspondence is given in (B18). At the same time [see (3.A.22), (2.16)]

$$\overline{V}_\mu=\overline{V}_{(R)}T_{1\mu}^R+\overline{V}_{A,B,C}\overline{T}_\mu^{A,B,C}\,. \tag{3.B.7}$$

It is also useful to write

TABLE IV. The four maximal solutions (a—d) for the $(4 \to 3 \otimes 1)$ spherically symmetric $G_{\mu\nu}$ [$\equiv SU(2) \otimes U(1) \otimes U(1)$]. The defining equations are: for I_i (3.B.9), for G (3.A.27), (3.A.40), for g_i (3.B.16), for ℓ_G (1.9), and for g_{tot} (2.21); for $m \otimes n$, see Table I.

	a	b	c	$d^{\pm}$	
I_1	2	1	0	$\frac{1}{2}$	
I_2	0	0	0	$\frac{\epsilon}{2}$	
I_3	0	1/2	0	$\frac{1}{8}$	
				$\epsilon=+1$	$\epsilon=-1$
g_1	0	$-1/2$	-1	$-1/2$	-1
g_2	0	1/2	1	1	$\frac{1}{2}$
g_3	0	1/2	0	0	$\frac{1}{2}$
g_4	0	$-1/2$	0	$-\frac{1}{2}$	0
ℓ_G	4	$2 \otimes 2 \otimes 1$	$2 \otimes 1 \otimes 1$	$2 \otimes 1 \otimes 1$	
g^2_{tot} unit e^{-2}	0	1	2	$\frac{3}{2}$	

$$\bar{V}_{A_1} = i\bar{V}_{21} = i\bar{V}_{12},$$
$$\bar{V}_{A_2} = \frac{1}{2i}(\bar{V}_{11} - \bar{V}_{22}), \tag{3.B.8}$$
$$\bar{V}_{A_3} = -\tfrac{1}{2}(\bar{V}_{11} + \bar{V}_{22}).$$

Hence, defining a symmetric matrix $\bar{V}_{AB} = \bar{V}_{BA}$. This corresponds mathematically to writing a vector as a symmetrized product of two spinors.

The following invariants are then constructed:

$$I_1 = \bar{V}_{A,B,C}\bar{V}_{A',B',C'}\epsilon_{AA'}\epsilon_{CC'}\delta_{BB'},$$
$$I_2 = \bar{V}_{A,B,C}\bar{V}_{A',B',C'}\epsilon_{AA'}\epsilon_{BB'}\delta_{CC'}, \tag{3.B.9}$$
$$I_3 = U_{AA'}U_{A''A'''}\epsilon_{AA''}\epsilon_{A'A'''},$$
$$U_{AA'} = \bar{V}_{A,B,C}\bar{V}_{A',B',C'}\epsilon_{BB'}\epsilon_{CC'} = U_{A'A}. \tag{3.B.10}$$

The metric has been used explicitly,

$$(i\sigma_2)_{AA'} = \epsilon_{AA'}. \tag{3.B.11}$$

(iv) $G_{\mu\nu}$ and the Lagrangian

Spherically symmetric G_ρ (3.A.27) have the same basis as $\bar{W}_\mu$. Hence er^2G_ρ is composed of two singlets G_B, G_C, and a A triplet G_{A_i}, i.e., the set $\{G_R\}$,

$$G_{A_1} = iU_{12}, \quad G_{A_2} = \frac{1}{2i}(U_{11} - U_{22}), \tag{3.B.12}$$
$$G_{A_3} = -\tfrac{1}{2}(U_{11} + U_{22}), \quad G_B = \tfrac{1}{2}I_1 - 1, \quad G_C = I_2.$$

We note that

$$G_{A_i}G_{A_i} = \tfrac{1}{2}I_3. \tag{3.B.13}$$

The remaining eight components contain derivatives and quadratic functions of $\bar{V}$'s,

$$G_{A,B,C} = -\epsilon_{BB'}\dot{\bar{V}}_{A,B',C} + \bar{V}_{(B)}\bar{V}_{A,B,C} - \bar{V}_{(C)}\epsilon_{BB'}\epsilon_{CC'}\bar{V}_{A,B',C'} - \bar{V}_{AA'}\epsilon_{A'A''}\epsilon_{BB'}\bar{V}_{A'',B',C}. \tag{3.B.14}$$

As in the preceding case, the Lagrangian does not depend on the derivatives of the $\{V_R\}$ components. These can be eliminated. The essential ingredient is that the Lagrange equations for $\{V_R\}$ are linear in $\{V_R\}$ and that the inhomogeneous terms are linear in $\dot{\bar{V}}_{A,B,C}$. We will not give these lengthy details here. It is sufficient for our purpose to note that the elimination of these five components eliminates five derivatives $\dot{\bar{V}}_{A,B,C}$. Since there were eight derivatives (considered as a vector space) to start with, there remain three derivatives, namely $\dot{I}_1$, $\dot{I}_2$, and $\dot{I}_3$. Hence the $G_{A,B,C}$ term only contributes to the kinetic term.

(v) Pointlike solutions

As in the preceding, the asymptotic pointlike solutions are obtained by maximalizing the nonderivative part of G_ρ, i.e., the set G_R. Using Appendix B, that part of the Lagrangian becomes [(3.B.12)],

$$L = -\frac{1}{4e^2r^4}\left(\frac{1}{2}(I_1 - 2)^2 + I_2^2 + I_3\right) + \text{derivatives}. \tag{3.B.15}$$

The eigenvalues of G [(3.A.40)] for asymptotic solutions are as follows:

$$g_1 = \tfrac{1}{2}(I_2 + I_1 - 2), \quad g_2 = \tfrac{1}{2}(I_2 - I_1 + 2),$$
$$g_3 = -\tfrac{1}{2}[I_2 - (2I_3)^{1/2}], \quad g_4 = -\tfrac{1}{2}[I_2 + (2I_3)^{1/2}] \tag{3.B.16}$$

Maximalization leads to four different solutions which are listed in Table IV. Again the strength of the monopole is defined by (3.A.42).

(vi) Higgs scalars—potential

A 15-plet of Higgs vectors reduces, for spherically symmetric solutions, to

$$\phi = \phi_R R, \quad \{R\} = \{A_i, B, C\}. \tag{3.B.17}$$

Let us define

$$\phi_A = (\phi^2_{A_1} + \phi^2_{A_2} + \phi^2_{A_3})^{1/2}, \tag{3.B.18}$$

the length of ϕ_{A_i} and also:

$$\phi_{A_1} = i\phi_{12} = i\phi_{21},$$
$$\phi_{A_2} = \frac{1}{2i}(\phi_{11} - \phi_{22}), \tag{3.B.19}$$
$$\phi_{A_3} = -\tfrac{1}{2}(\phi_{11} + \phi_{22}).$$

The eigenvalues of ϕ are then

$$\phi_1 = \phi_C + \phi_B, \quad \phi_2 = \phi_C - \phi_B,$$
$$\phi_3 = -\phi_C + \phi_A, \quad \phi_4 = -\phi_C - \phi_A. \tag{3.B.20}$$

They are clearly ℓ invariant.

According to the equalities between these eigenvalues the Higgs vector little group can belong to one of the sets (3.A.45) and (3.A.46). They correspond to maximalization of the potential.

(vii) Higgs scalar–kinetic terms

Owing again to the ℓ little group properties $D_\mu\phi$ the covariant derivative of ϕ can be written very simply

TABLE V. The six possible solutions (1—6) for the $(4 \to 3 \oplus 1)$ spherically symmetric Higgs fields. The defining equations are for Φ_i (3.B.20), for I_i (3.B.9), for g^2 (2.20), and for l_Φ (1.8); for $m \otimes n$ see Table I.

	1	2		3		4		5				6
Φ_i	$\Phi_i = 0$	$\Phi_1 = \Phi_3$ $\Phi_2 = \Phi_4$	$\Phi_2 = \Phi_3$ $\Phi_1 = \Phi_4$	$\Phi_1 = \Phi_2$ $= \Phi_3$	$\Phi_1 = \Phi_2$ $= \Phi_4$	$\Phi_2 = \Phi_3$ $= \Phi_4$	$\Phi_1 = \Phi_3$ $= \Phi_4$	$\Phi_1 = \Phi_3$	$\Phi_1 = \Phi_4$	$\Phi_2 = \Phi_3$	$\Phi_2 = \Phi_4$	arbitrary
I_i	arbitrary	$2I_3 = I_1^2$		$2I_3 = I_2^2$		arbitrary		$2I_3 = I_1^2 = I_2^2$				$I_i = 0$
g^2	/	1		0		$\frac{4}{3}$		arbitrary				arbitrary
l_Φ	4	$2 \otimes 2 \otimes 1$		$3 \otimes 1$		$3 \otimes 1$		$2 \otimes 1 \otimes 1$				$1 \otimes 1 \otimes 1$

with the same tensors as W_μ [(3.B.5), (3.B.6)]

$$D_\mu \phi = D_R \overline{T}^R_{1\mu} + D_{A,B,C} \overline{T}^{A,B,C}_\mu , \tag{3.B.21}$$

and

$$\begin{aligned} D_{A_i} &= \dot{\phi}_{A_i} + \frac{2}{r} \epsilon_{ijk} \overline{V}_{A_j} \phi_{A_k} , \\ D_B &= \dot{\phi}_B , \quad D_C = \dot{\phi}_C , \\ D_{A,B,C} &= \frac{1}{r} \left(-\phi_B \epsilon_{BB'} \overline{V}_{A,B',C} - 2\phi_C \epsilon_{CC'} \overline{V}_{A,B,C'} \right. \\ &\quad \left. - \phi_{AA'} \epsilon_{A'A''} \overline{V}_{A'',B,C} \right) . \end{aligned} \tag{3.B.22}$$

Again the covariance properties under ℓ ensure the correct transformation laws of the different terms. Maximalization leads to $D_\mu \phi = 0$. In view of the first equation of (3.B.22) this implies in particular that $\overline{V}_{A_i}$ and ϕ_{A_i} be parallel three vectors.

The results of the maximalization are given in Table V, and comparison between Tables IV and V are summarized in Table VI. The discussion is analogous to the discussion of the A case.

C. The 4 → 2 + 2 embedding

(i) Decomposition of the generators of SU(4)

The generators E_i forming a doublet of the spin-$\frac{1}{2}$ representation of SU(2) as a subset of the 4×4 SU(4) generators are unique up to an equivalence and are given in Appendix C. The remaining 12 generators decompose according to (2.14c) in three singlets,

$$K^{(k)} \quad (k = 1, 2, 3), \tag{3.C.1}$$

and three triplets

$$N_i^{(k)} \quad (k = 1, 2, 3). \tag{3.C.2}$$

The projections in the $\hat{x}_i$ direction of the triplets are

$$E = \hat{x}_i E_i , \quad N^{(k)} = \hat{x}_i N_i^{(k)} . \tag{3.C.3}$$

(ii) The little group

The seven operators $\{R\} = \{E, K^{(k)}, N^{(k)}\}$ generate the little group ℓ. They close under the commutation relations of

$$Q = \{E, K^{(k)} + N^{(k)}, K^{(k)} - N^{(k)}\} \equiv \mathrm{U(1)} \otimes \mathrm{SU(2)} \otimes \mathrm{SU(2)} . \tag{3.C.4}$$

We note that $\mathrm{SU(2)} \otimes \mathrm{SU(2)}$ is locally isomorphic to SO(4). And hence the antisymmetric tensor K^{AB} $(A, B = 1, 2, 3, 0)$,

$$K^{0k} = N^{(k)}, \quad K^{kl} = \epsilon^{lkp} K^{(p)} , \tag{3.C.5}$$

has the commutation relation of the Euclidean form $(g_{AB} = \delta_{AB})$ of the Lorentz group. The infinitesimal transformation of ℓ,

$$M = 1 + i\rho R , \tag{3.C.6}$$

depend on seven real parameters $\rho = \{\epsilon, \kappa^{(k)}, \eta^{(k)}\}$ or $\rho = \{\epsilon, \kappa^{AB}\}$ with the correspondence (3.C.5).

(iii) Basic vector for W_μ

The natural 15 basis vectors for $\overline{W}_\mu$ are

$$\begin{aligned} T^R_{1\mu} &= \hat{x}_\mu R, \quad \{R\} = \{E, K^{(k)}, N^{(k)}\} \\ T^P_{2\mu} &= (\hat{x}_\mu \hat{x}_i - \delta_{\mu i}) P_i, \quad \{P\} = \{E, N^{(k)}\} , \\ T^P_{3\mu} &= \epsilon_{\mu ij} \hat{x}_j P_i . \end{aligned} \tag{3.C.7}$$

TABLE VI. The correspondence between the solutions for $G_{\mu\nu}$ and Φ in the $(4 \to 3 \oplus 1)$ case. See Table III for notation.

	1abcd	2b	2c	$2^-\mathrm{d}^\pm$	3a	3c	$3^\pm\mathrm{d}^\mp$	4c	$4^\pm\mathrm{d}^\mp$	5c	$5\mathrm{d}^\pm$	6c
l_Φ	4	$2 \otimes 2 \otimes 1$	$2 \otimes 2 \otimes 1$	$2 \otimes 2 \otimes 1$	$3 \otimes 1$	$3 \otimes 1$	$3 \otimes 1$	$3 \otimes 1$	$3 \otimes 1$	$2 \otimes 1 \otimes 1$	$2 \otimes 1 \otimes 1$	$1 \otimes 1 \otimes 1$
l_G	see Table BI	l_Φ	$2 \otimes 1 \otimes 1$	$2 \otimes 1 \otimes 1$	4	$2 \otimes 1 \otimes 1$	$2 \otimes 1 \otimes 1$	$2 \otimes 1 \otimes 1$	$2 \otimes 1 \otimes 1$	$2 \otimes 1 \otimes 1$	$2 \otimes 1 \otimes 1$	$2 \otimes 1 \otimes 1$
l_S	l_G	l_Φ	$1 \otimes 1 \otimes 1$	$2 \otimes 1 \otimes 1$	l_Φ	$1 \otimes 1 \otimes 1$	$2 \otimes 1 \otimes 1$	$1 \otimes 1 \otimes 1$	$2 \otimes 1 \otimes 1$	$1 \otimes 1 \otimes 1$	$2 \otimes 1 \otimes 1$ or $1 \otimes 1 \otimes 1$	l_Φ
g^2 unit e^{-2}	/	1	1	1	0	0	0	$\frac{4}{3}$	$\frac{4}{3}$	g^2	g^2	g^2
$g_\perp^2$ unit e^{-2}	/	0	1	$\frac{1}{2}$	0	2	$\frac{3}{2}$	$\frac{2}{3}$	$\frac{1}{6}$	$2 - g^2$	$\frac{3}{2} - g^2$	$2 - g^2$
g_{tot}^2 unit e^{-2}	see Table BI	1	2	$\frac{3}{2}$	0	2	$\frac{3}{2}$	2	$\frac{3}{2}$	2	$\frac{3}{2}$	2

TABLE VII. The three maximal solutions (a–c) for the $(4 \to 2 \oplus 2)$ spherically symmetric $G_{\mu\nu}$ $[\ell = SU(2) \otimes SU(2) \otimes U(1)]$. The defining equations are: for I_i (3.C.13), for G (3.A.27), (3.A.40), for g_i (3.C.18), for ℓ_G (1.9), and for g^2_{tot} (2.21); for $m \otimes n$ see Table I.

	a	b	c
I_1	1	0	$\frac{1}{2}$
I_2	0	0	$\frac{1}{16}$
g_1	0	$\frac{1}{2}$	$\frac{1}{2}$
g_2	0	$\frac{1}{2}$	0
g_3	0	$-\frac{1}{2}$	0
g_4	0	$-\frac{1}{2}$	$-\frac{1}{2}$
ℓ_G	4	$2 \otimes 2 \otimes 1$	$2 \otimes 1 \otimes 1$
g^2_{tot} unit e^2	0	1	$\frac{1}{2}$

Under the little group transformations, the 15 basic tensors transform as a singlet,

$$\overline{T}^E_\mu = T^E_{1\mu} = \hat{x}_\mu E\,, \tag{3.C.8}$$

a six-dimensional $(1,0)+(0,1)$ representation of SU(2) $\otimes$ SU(2) represented by an antisymmetric tensor (3.C.5)

$$\overline{T}^{AB}_\mu = -\overline{T}^{BA}_\mu\,, \tag{3.C.9}$$

with

$$\overline{T}^{0k}_\mu = T^{N(k)}_{1\mu}\,, \quad \overline{T}^{kl}_\mu = \epsilon^{lkp} T^{K(p)}_{1\mu}\,. \tag{3.C.10}$$

The remaining eight tensors form a four-dimensional vector $(\frac{1}{2} \otimes \frac{1}{2})$ representation under SU(2)$\otimes$SU(2) which is also a doublet under the U(1),

$$\overline{T}^{A,\alpha}_\mu:\ T^{0,1}_\mu = T^E_{2\mu};\ T^{0,2}_\mu = -T^E_{3\mu}\ (A = 1,2,3,0),$$
$$T^{k,1}_\mu = T^{N(k)}_{3\mu};\ T^{k,2}_\mu = T^{N(k)}_{2\mu}\ (\alpha = 1,2). \tag{3.C.11}$$

With the same convention as before [(2.16), (3.A.22)],

$$\overline{V}_\mu = \overline{V}_E \overline{T}_\mu + \overline{V}_{AB}\overline{T}^{AB}_\mu + \overline{V}_{A,\alpha}\overline{T}^{A,\alpha}_\mu\,. \tag{3.C.12}$$

The relevant invariants are

$$I_1 = \overline{V}_{A,\alpha}\overline{V}_{A,\alpha}\,, \quad I_2 = \tfrac{1}{2} G_{AB} G_{AB}\,, \tag{3.C.13}$$

with

$$G_{AB} = \epsilon_{\alpha\beta}\overline{V}_{A\alpha}\overline{V}_{B\beta}\,. \tag{3.C.14}$$

(iv) $G_{\mu\nu}$ and the Lagrangian

Spherically symmetric G_ρ (3.A.27) have the same basis as $\overline{W}_\mu$. After some computation G_ρ is written,

$$G_\rho = \frac{1}{er^2}(I_1 - 1)\overline{T}_\rho + G_{AB}\overline{T}^{AB}_\rho + G_{A,\alpha}\overline{T}^{A,\alpha}_\rho\,, \tag{3.C.15}$$

with (3.C.14) and

$$G_{A,\alpha} = \epsilon_{\alpha\beta}\dot{\overline{V}}_{A,\beta} + \overline{V}_E \overline{V}_{A,\alpha} + \epsilon_{\alpha\beta}\overline{V}_{AB}\overline{V}_{B,\beta}\,. \tag{3.C.16}$$

As in the preceding cases the Lagrangian does not depend on the derivatives of $\overline{V}_E$ and $\overline{V}_{AB}$. These components can be eliminated together with six (and not seven) derivatives.

The nonderivative part of the Lagrangian which only matters for asymptotic solutions is then [(3.C.15), (3.C.13)]

$$L = -\tfrac{1}{4}[(I_1 - 1)^2 + 4 I_2] + \text{derivatives}\,. \tag{3.C.17}$$

(v) Pointlike solutions

Maximalization with respect to the relevant $\overline{V}_{A,\alpha}$ leads to the three solutions of Table VII. The asymptotic values of G (3.A.40) are

$$g_i = \epsilon \tfrac{1}{2}(I_1 - 1) + \epsilon'(I_2)^{1/2}\,, \tag{3.C.18}$$

where ϵ and ϵ' are arbitrary signs. Again the strength of the monopole is defined by (3.A.42).

(vi) Higgs scalars—potential

The 15-plet of Higgs scalars are written

$$\phi = \varphi_R R\,, \quad R = \{E, K^{(k)}, N^{(k)}\}\,, \tag{3.C.19}$$

for spherically symmetric solutions. The four eigenvalues of φ are

$$\phi_{\frac{1}{2}} = \tfrac{1}{2}\varphi_E \pm \left[\sum_k (\varphi_{K(k)} + \varphi_{N(k)})^2\right]^{1/2},$$
$$\phi_{\frac{3}{4}} = -\tfrac{1}{2}\varphi_E \pm \left[\sum_k (\varphi_{K(k)} - \varphi_{N(k)})^2\right]^{1/2}. \tag{3.C.20}$$

It is easy to verify that the values are ℓ invariant, by using the notation [see (3.C.5)] φ_{AB}. Equalities between these eigenvalues correspond to maximal potentials [see (3.A.45), (3.A.46)].

(vii) Higgs scalars—kinetic terms

The covariant derivatives of φ have the same basis vectors of $\overline{W}_\mu$ (3.C.12),

$$D_\mu \varphi = D_E \overline{T}_\mu + D_{AB}\overline{T}^{AB}_\mu + D_{A,\alpha}\overline{T}^{A,\alpha}_\mu \tag{3.C.21}$$

and

$$D_E = \dot{\varphi}_E\,,$$
$$D_{AB} = \dot{\varphi}_{AB} + \frac{1}{r}(\overline{V}_{AL}\varphi_{LB} - \varphi_{AL}\overline{V}_{LB})\,, \tag{3.C.22}$$
$$D_{A,\alpha} = \frac{1}{r}(\varphi_E \epsilon_{\alpha\beta}\overline{V}_{A,\beta} - \varphi_{AB}\overline{V}_{B,\alpha})\,.$$

Asymptotic maximization will again lead to $D_\mu \varphi = 0$. This implies relations between the directions of φ and G. The results are given in Tables VIII and IX.

TABLE VIII. The four solutions (1–4) for the $(4 \to 2 \oplus 2)$ spherically symmetric Higgs fields. The defining equations are: for Φ_i (3.C.20), for I_i (3.C.13), for g^2 (2.20), and for ℓ_Φ (1.8); for $m \otimes n$ see Table I.

	1	2	3	4
Φ_i	0	$\Phi_1 = \Phi_4$ $\Phi_2 = \Phi_3$	$\Phi_1 = \Phi_4$	arbitrary
I_i	arbitrary	arbitrary	$4\ I_2\ \ I_1^2$	$I_i = 0$
g^2	0	0	arbitrary $\otimes (2I_1 - 1)$	arbitrary
ℓ_Φ	4	$2 \otimes 2 \otimes 1$	$1 \otimes 1 \otimes 1$ could be larger	$1 \otimes 1 \otimes 1$

TABLE IX. The correspondence between the solutions for $G_{\mu\nu}$ and Φ in the $(4 \to 2 \oplus 2)$ case. See Table III for notations.

	1abc	2a	2b	2c	3b	3c	4b
ℓ_Φ	4	$2 \otimes 2 \otimes 1$	$2 \otimes 2 \otimes 1$	$2 \otimes 2 \otimes 1$	$2 \otimes 1 \otimes 1$	$2 \otimes 1 \otimes 1$	$1 \otimes 1 \otimes 1$
ℓ_G	see Table VII	4	$2 \otimes 2 \otimes 1$	$2 \otimes 1 \otimes 1$	$2 \otimes 2 \otimes 1$	$2 \otimes 1 \otimes 1$	$2 \otimes 2 \otimes 1$
ℓ_S	ℓ_G	ℓ_Φ	$1 \otimes 1 \otimes 1$	ℓ_G	$1 \otimes 1 \otimes 1$	$1 \otimes 1 \otimes 1$	ℓ_Φ
g^2 unit e^{-2}	/	0	0	0	g^2	0	g^2
g_I^2	/	0	1	$\frac{1}{2}$	$1-g^2$	$\frac{1}{2}$	$1-g^2$
g_{tot}^2 unit e^{-2}	see Table VII	0	1	$\frac{1}{2}$	1	$\frac{1}{2}$	1

D. The 4 → 2 + 1 + 1 embedding

(i) Decomposition of the generators of SU(4)

Up to an equivalence, the three E_i matrices are given in Appendix D. The remaining generators according to (2.14d) decompose in four doublets, $K_\alpha^{(k)}$ $(k=1,2,3,4,\ \alpha=1,2)$, and four singlets

$$N, F^{(k)} \quad (k=1,2,3)\,. \tag{3.D.1}$$

As usual the projection of E_i along the $\hat{x}_i$ direction is written by E,

$$E = \hat{x}_i E_i\,. \tag{3.D.2}$$

(ii) The little group

The five operators $\{R\} = \{E, N, F^{(k)}\}$ generate the little group ℓ. Their commutation relations are those of

$$\{R\} = \{E, N, F^{(k)}\} \equiv \mathrm{U}(1) \otimes \mathrm{U}(1) \otimes \mathrm{SU}(2)\,. \tag{3.D.3}$$

The infinitesimal transformation of ℓ,

$$M = 1 + i\rho R, \tag{3.D.4}$$

depend on the five real parameters $\rho = \{\epsilon, \eta, \phi^{(k)}\}$.

(iii) Basis vectors for W_μ

The seven basis vectors for $\overline{W}_\mu$ are

$$\begin{aligned} &T_{1\mu}^R = \hat{x}_\mu R, \quad \{R\} = \{E, N, F^{(k)}\}, \\ &T_{2\mu}^E = (\hat{x}_\mu \hat{x}_i - \delta_{\mu i}) E_i, \quad T_{3\mu}^E = \epsilon_{\mu ij} \hat{x}_j E_i\,. \end{aligned} \tag{3.D.5}$$

Under the little group transformations $T_{1\mu}^E$ and $T_{1\mu}^N$ transform as singlets. $T_{1\mu}^{F(k)}$ transforms as a triplet under the SU(2) and as a singlet under the two U(1). Finally T_μ^A,

$$T_\mu^A = \begin{pmatrix} T_\mu^1 = T_{2\mu}^E \\ T_\mu^2 = T_{3\mu}^E \end{pmatrix}, \tag{3.D.6}$$

form a doublet under the E transformation and are singlets under the second U(1) and SU(2). With the same convention as before

$$\overline{V}_\mu = \overline{V}_E T_{1\mu}^E + \overline{V}_N T_{1\mu}^N + \overline{V}_{F(k)} T_{1\mu}^{F(k)} + \overline{V}_A T_\mu^A\,. \tag{3.D.7}$$

The relevant invariant is

$$I = \overline{V}_A \overline{V}_A\,. \tag{3.D.8}$$

(iv) $G_{\mu\nu}$ and the Lagrangian

Spherically symmetric G_ρ have the same basis as $\overline{W}_\mu$. Computing (2.6) one finds

$$G_\rho = \frac{1}{er^2}\,(\overline{G}_E T_{1\rho}^E + G_A T_\rho^A)\,, \tag{3.D.9}$$

with

$$\overline{G}_E = I - 1, \quad G_A = -r\epsilon_{AB}\dot{\overline{V}}_B + \overline{V}_E \overline{V}_A\,. \tag{3.D.10}$$

Again the Lagrangian does not depend on the derivatives $\dot{\overline{V}}_E, \dot{\overline{V}}_N, \dot{\overline{V}}_{F(k)}$. In fact, as seen in (3.D.10), it does not depend on $\overline{V}_N$ and $\overline{V}_{F(k)}$ at all. The constraint equation on $\overline{V}_E$,

$$\overline{V}_E = r(V_A \epsilon_{AB} \dot{\overline{V}}_B)/I, \tag{3.D.11}$$

then enables one to write the Lagrangian

$$L = -\frac{1}{4e^2 r^4}\,[(\dot{I}^2)/4I + (I-1)^2]\,. \tag{3.D.12}$$

This is in fact the Lagrangian of 't Hooft's original paper.

(v) Pointlike solutions

The maximal asymptotic solutions of (3.D.12) correspond to $I=0$ and $I=1$, respectively. See Table X. The eigenvalues of G (3.A.40) are

$$g_{1,2} = \pm\tfrac{1}{2}(I-1), \quad g_{3,4} = 0\,. \tag{3.D.13}$$

Again the strength is defined by (3.A.42).

TABLE X. The two maximal solutions (a–b) for the $(4 \to 2 \oplus 1 \oplus 1)$ spherically symmetric $G_{\mu\nu}$ [$\ell = \mathrm{SU}(2) \otimes \mathrm{U}(1) \otimes \mathrm{U}(1)$]. The defining equations are: for I (3.D.8), for G (3.A.27), (3.A.40), for g_i (3.D.13), for ℓ_G (1.9), and for g_{tot}^2 (2.21); for $m \otimes n$ see Table I.

	a	b
I	1	0
g_1	0	$\frac{1}{2}$
g_2	0	$-\frac{1}{2}$
g_3	0	0
g_4	0	0
ℓ_G	4	$2 \otimes 1 \otimes 1$
g_{tot}^2 unit e^{-2}	0	$\frac{1}{2}$

TABLE XI. The two solutions (1–2) for the $(4 \to 2 \oplus 1 \oplus 1)$ spherically symmetric Higgs fields. The defining equations are: for Φ_i (3.D.16), for I (3.D.8), for g^2 (2.20), and for ℓ_Φ (1.8), for $m \otimes n$ see Table I.

	1	2
Φ_i	0	arbitrary
I_i	arbitrary	arbitrary
g^2	0	arbitrary $(I-1)^2$
ℓ_Φ	4	$1 \otimes 1 \otimes 1$ could be larger

(vi) Higgs scalars—potential

The 15-plet of spherically symmetric Higgs scalars is

$$\Phi = \Phi_R R, \quad R = \{E, N, F^{(k)}\}. \tag{3.D.14}$$

The four eigenvalues of Φ are written in terms of Φ_E, Φ_N, and

$$\Phi_F = \left(\sum_k \Phi^2_{F(k)}\right)^{1/2} \tag{3.D.15}$$

as

$$\underset{2}{\Phi_1} = \tfrac{1}{2}(\pm\Phi_E + \Phi_N), \quad \underset{4}{\Phi_3} = \tfrac{1}{2}(-\Phi_N \pm \Phi_F). \tag{3.D.16}$$

As usual a maximal potential corresponds to equalities between the eigenvalues [see (3.A.45), (3.A.46)].

(vii) Higgs scalars—kinetic terms

The covariant derivatives of Φ have the same basis vectors as $\overline{W}_\mu$ (3.D.7). Hence

$$D_\mu \Phi = D_E T^E_{1\mu} + D_N T^N_{1\mu} + D_{F(k)} T^{F(k)}_{1\mu} + D_A T^A_\mu, \tag{3.D.17}$$

$$D_E = \dot{\Phi}_E, \quad D_N = \dot{\Phi}_N,$$

$$D_{F(k)} = \dot{\Phi}_{F(k)} + \frac{1}{r} \epsilon_{klm} \overline{V}_{F(l)} \Phi_{F(m)}, \tag{3.D.18}$$

$$D_A = -\frac{1}{r} \epsilon_{AB} \overline{V}_B \Phi_E.$$

As is seen from the term $D_{F(k)}$, the minimization $D_\mu \Phi = 0$ implies in particular that $\overline{V}_{F(k)}$ and $\Phi_{F(k)}$ be parallel vectors. From D_A, one sees that Φ_E must be zero unless $I = 0$. Discussion of these cases and comparison with the extremal $G_{\mu\nu}$ are given in Tables XI and XII.

IV. CONCLUSION

By treating explicitly the problem of finding spherically symmetric monopoles in SU(4) we have obtained the following results:

(a) We have shown the relevance of the new group we have introduced, the little group ℓ which transforms spherically symmetric solutions into spherically symmetric solutions. Its use simplifies the problem greatly. In SU(2) and SU(3) this little group was always Abelian.[3] Here in some cases it is non-Abelian.

(b) We have found all the possible embeddings of SU(2) into SU(4) and in each case we have found a *discrete* set of asymptotic $G_{\mu\nu}$'s decreasing as $1/r^2$. This is without reference to the existence or transformation properties of the Higgs fields.

(c) In all instances we see that the eigenvalues of G [see (3.A.27), (3.A.40) and the tables] turn out always to be integers or half-integers. This is a particular case of results[4] which shows that the possible values of the magnetic poles are related to the Cartan subalgebras. The *direction* and *magnitude* of the poles correspond to weights of the *dual* group. Here the dual of SU(4) is $SU(4)/Z(4)$ and hence the results follow.

(d) Contrary to SU(2) and SU(3) the scalar product of G^a_ρ with a 15-plet (adjoint representation) of asymptotic Higgs vectors does not necessarily quantize the charge. The strengths g and $g_\perp$ (see the tables) of the component of G parallel and orthogonal to φ may vary continuously between prespecified limits.

The discussion of the smoothed out solution and of the related stability has not been given here. The discussion at the end of Ref. 1 applies *mutadis mutandis*.

ACKNOWLEDGMENTS

The authors thank their friends E. Corrigan, F. Englert, D. Fairlie, D. Olive, and P. Windey for numerous discussions on the subject.

APPENDIX A

We here collect some formulas related to case A. The commutation relations between E_i, F_{ij}, and N_{ijk} are

$$[E_i, E_j] = i\epsilon_{ijk} E_k, \tag{A1}$$

$$[E_i, K_{jk}] = i(\epsilon_{ijl} K_{lk} + \epsilon_{ikl} K_{jl}), \tag{A2}$$

$$[E_i, N_{jkl}] = i(\epsilon_{ijm} N_{mkl} + \epsilon_{ikm} N_{jml} + \epsilon_{ilm} N_{jkm}), \tag{A3}$$

$$[K_{kl}, K_{mn}] = \frac{3i}{5}\left(\sum^{5} \delta_{km}\epsilon_{lnr}\right) E_r + i\left(\sum^{4} \epsilon_{kmr} N_{lnr}\right), \tag{A4}$$

$$[K_{kl}, N_{mnp}] = \frac{i}{5}\left(\sum^{12} \epsilon_{kmr}\delta_{ln} K_{rp}\right) + \frac{i}{20}\left(\sum^{12} \epsilon_{kmr}\epsilon_{lnq}\epsilon_{qps} K_{rs}\right), \tag{A5}$$

TABLE XII. The correspondence between the solution for $G_{\mu\nu}$ and Φ in the $(4 \to 2 \oplus 1 \oplus 1)$ case. See Table III for notation.

	1 a–b	2a	2b
ℓ_Φ	4	$1 \otimes 1 \otimes 1$	$1 \otimes 1 \otimes 1$
ℓ_G	see Table X	4	$2 \otimes 1 \otimes 1$
ℓ_S	ℓ_G	ℓ_Φ	ℓ_Φ
g^2 unit e^{-2}	/	0	g^2
$g^2_\perp$ unit e^{-2}	/	0	$\frac{1}{2} - g^2$
g^2_{tot} unit e^{-2}	see Table X	0	$\frac{1}{2}$

$$[N_{klm}, N_{pqr}] = \frac{1}{20i}\left(\sum^{36} \epsilon_{kpa}\delta_{lq}N_{mra}\right) + \frac{3i}{25}\left(\sum^{12}\epsilon_{kpa}\delta_{lq}\delta_{mr}\right)E_a + \frac{1}{10i}\left(\sum^{6}\epsilon_{kpa}\epsilon_{lqb}\epsilon_{mrc}\right)N_{abc} + \frac{3}{100i}\left(\sum^{18}\epsilon_{kpa}\epsilon_{lqa}\epsilon_{mrc}\right)E_c . \quad \text{(A6)}$$

In these formulas $\sum^n$ means the sum of the *inequivalent* permutations on the indices needed to obtain the symmetry of the left-hand side, n is the number of terms in the sum. The traces of the matrices are

$$\mathrm{Tr}(E_i E_j) = 5\delta_{ij}, \quad \text{(A7)}$$

$$\mathrm{Tr}(K_{ij}K_{kl}) = 3(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) - \tfrac{2}{3}\delta_{ij}\delta_{kl}, \quad \text{(A8)}$$

$$\mathrm{Tr}(N_{ijk}N_{lmn}) = \tfrac{3}{4}\left(\sum^{6}\delta_{il}\delta_{jm}\delta_{kn}\right) - \tfrac{2}{5}\left(\sum^{9}\delta_{ij}\delta_{kl}\delta_{mn}\right). \quad \text{(A9)}$$

The cross traces between E, F, and N are zero.

With (3.A.9) and (A7)–(A9) one finds the norms of the basic vectors

$$W_\mu = W_i^P T_{i\mu}^P,$$

$$\mathrm{Tr}(W_\mu W^\mu) = 5(W_1^E)^2 + 4(W_1^K)^2 + \tfrac{9}{5}(W_1^N)^2 + 10[(W_2^E)^2 + (W_3^E)^2] + 6[(W_2^K)^2 + (W_3^K)^2] + \tfrac{12}{5}[(W_2^N)^2 + (W_3^N)^2]. \quad \text{(A10)}$$

For $T_\mu^{(A)}$ and $T_\mu^{(AB)}$ the norms are respectively 4 and 6.

APPENDIX B

We here collect some formulas related to the case B. The commutation relations between Z, E_i, K_i, N_i, and F_{ij} are the following:

$$[E_i, Z] = 0, \quad \text{(B1)}$$

$$[E_i, E_j] = i\epsilon_{ijk}E_k, \quad \text{(B2)}$$

$$[E_i, K_j] = i\epsilon_{ijk}K_k. \quad \text{(B3)}$$

$$[E_i, N_j] = i\epsilon_{ijk}N_k, \quad \text{(B4)}$$

$$[E_i, F_{jk}] = i(\epsilon_{ijl}F_{lk} + \epsilon_{ikl}F_{jl}), \quad \text{(B5)}$$

$$[Z, K_i] = 4iN_i, \quad \text{(B6)}$$

$$[Z, N_i] = -4iK_i, \quad \text{(B7)}$$

$$[K_i, K_j] = [N_i, N_j] = i\epsilon_{ijk}E_k, \quad \text{(B8)}$$

$$[K_i, N_j] = -2iF_{ij} + \tfrac{2}{3}i\delta_{ij}Z, \quad \text{(B9)}$$

$$[K_i, F_{kl}] = \tfrac{1}{2}i(\delta_{ki}N_l + \delta_{li}N_k - \tfrac{2}{3}\delta_{kl}N_i), \quad \text{(B10)}$$

$$[L_i, F_{kl}] = -\tfrac{1}{2}i(\delta_{ki}K_l + \delta_{li}K_k - \tfrac{2}{3}\delta_{lk}K_i), \quad \text{(B11)}$$

$$[F_{kl}, F_{mn}] = \tfrac{1}{4}i\left(\sum^{4}\delta_{nk}\epsilon_{lmq}E_q\right). \quad \text{(B12)}$$

The traces are

$$\mathrm{Tr}(Z^2) = 12, \quad \text{(B13)}$$

$$\mathrm{Tr}(E_i E_j) = 2\delta_{ij}, \quad \text{(B14)}$$

$$\mathrm{Tr}(K_i, K_j) = 2\delta_{ij}, \quad \text{(B15)}$$

$$\mathrm{Tr}(N_i, N_j) = 2\delta_{ij}, \quad \text{(B16)}$$

$$\mathrm{Tr}(F_{ij}F_{kl}) = \tfrac{1}{2}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk} - \tfrac{2}{3}\delta_{ij}\delta_{kl}). \quad \text{(B17)}$$

The correspondence between the natural tensors $T_{2\mu}^P$, $T_{3\mu}^P$ [(3.B.4)], and $T_\mu^{A,B,C}$ [(3.B.6)] is

$$T_{2\mu}^E = i(T_\mu^{1,2,2} + T_\mu^{2,2,1}), \quad T_{3\mu}^E = -i(T_\mu^{1,1,2} + T_\mu^{2,1,1}),$$
$$T_{2\mu}^K = T_\mu^{1,1,2} - T_\mu^{2,1,1}, \quad T_{3\mu}^K = T_\mu^{1,2,2} - T_\mu^{2,2,1},$$
$$T_{2\mu}^N = T_\mu^{1,1,1} + T_\mu^{2,1,2}, \quad T_{3\mu}^N = T_\mu^{1,2,1} + T_\mu^{2,2,2}, \quad \text{(B18)}$$
$$T_{2\mu}^F = -\frac{i}{2}(T_\mu^{1,1,1} - T_\mu^{2,1,2}), \quad T_{3\mu}^F = -\frac{i}{2}(T_\mu^{1,2,1} - T_\mu^{2,2,2}).$$

The norm of W_μ with the natural tensors [see (3.A.11) and (3.B.4)] is

$$\mathrm{Tr}(W_\mu W^\mu) = 12(W_1^s)^2 + 2[(W_1^E)^2 + (W_1^K)^2 + (W_1^N)^2] + \tfrac{2}{3}(W_1^F)^2 + 4[(W_2^E)^2 + (W_3^E)^2 + (W_2^K)^2 + (W_3^K)^2 + (W_2^N)^2 + (W_3^N)^2] + (W_2^F)^2 + (W_3^F)^2. \quad \text{(B19)}$$

For $T_\mu^{A,B,C}$ the norms are

$$\mathrm{Tr}(T_\mu^{A,B,C}T_\mu^{A',B',C'}) = 2\delta^{BB'}\epsilon^{AA'}\epsilon^{CC'}. \quad \text{(B20)}$$

APPENDIX C

We here present some formulas related to case C. The generators E_i, $K^{(k)}$, and $N_i^{(k)}$ take the simple form

$$E_i = \tfrac{1}{2}1 \otimes \sigma_i, \quad \text{(C1)}$$

$$K^{(k)} = \tfrac{1}{2}\sigma_k \otimes 1, \quad \text{(C2)}$$

$$N_i^{(k)} = \tfrac{1}{2}\sigma_k \otimes \sigma_i, \quad \text{(C3)}$$

in terms of the Pauli matrices σ_i and of the unit 2×2 matrix 1. The commutation relations between these generators are

$$[E_i, E_j] = i\epsilon_{ijk}E_k, \quad \text{(C4)}$$

$$[E_i, K^{(k)}] = 0, \quad \text{(C5)}$$

$$[E_i, N_j^{(k)}] = i\epsilon_{ijk}N_i^{(k)}, \quad \text{(C6)}$$

$$[K^{(k)}, K^{(l)}] = i\epsilon^{klm}K^{(m)}, \quad \text{(C7)}$$

$$[K^{(k)}, N_i^{(l)}] = i\epsilon^{klm}N_i^{(m)}, \quad \text{(C8)}$$

$$[N_i^{(k)}, N_j^{(l)}] = i(\delta^{kl}\epsilon_{ijm}E_m + \delta_{ij}\epsilon^{klm}K^{(m)}). \quad \text{(C9)}$$

The nonzero traces are

$$\mathrm{Tr}(E_i E_j) = \delta_{ij}, \quad \text{(C10)}$$

$$\mathrm{Tr}(K^{(l)}K^{(n)}) = \delta^{ln}, \quad \text{(C11)}$$

$$\mathrm{Tr}(N_i^{(k)}N_j^{(l)}) = \delta^{kl}\delta_{ij}. \quad \text{(C12)}$$

In the natural basis (3.C.7), the norm of W_μ [(3.A.11)] is

$$\mathrm{Tr}(W_\mu W^\mu) = (W_1^E)^2 + (W_1^{K(1)})^2 + (W_1^{N(1)})^2 + 2[(W_2^E)^2 + (W_3^E)^2 + (W_2^{N(k)})^2 + (W_3^{N(k)})^2]. \quad \text{(C13)}$$

APPENDIX D

Here are some formulas related to case D. The non-

zero commutation relations between E_i, N, and $F^{(k)}$ are

$$[E_i, E_j] = i\epsilon_{ijk}E_k , \tag{D1}$$

$$[F^{(k)}, F^{(l)}] = i\epsilon^{klm}F^{(m)} . \tag{D2}$$

The nonzero traces are chosen as

$$\mathrm{Tr}(E_i E_j) = \tfrac{1}{2}\delta_{ij} , \tag{D3}$$

$$\mathrm{Tr}(N^2) = 1 , \tag{D4}$$

$$\mathrm{Tr}(N^{(k)}N^{(l)}) = \tfrac{1}{2}\delta^{kl} . \tag{D5}$$

The norm of W_μ (3.A.11) in the natural basis (3.D.5) is

$$\mathrm{Tr}(W_\mu W^\mu) = \tfrac{1}{2}(W_1^E)^2 + (W_1^N)^2 + (W_1^{F(k)})^2 + (W_2^E)^2 + (W_3^E)^2 .$$

APPENDIX E

There are conflicting reasons for the two choices of basis vectors T_μ's and for the basis of the little group generators. Let us make some remarks.

(a) in the natural choice for T_μ, S_μ turns out to have always the same form,

$$S_\mu = (1/er)T^E_{3\mu} . \tag{E1}$$

The $\overline{T}_\mu$ becomes interesting once S_μ has been subtracted.

(b) For case A in the natural choice E, K, N are orthogonal in the sense

$$\mathrm{Tr}(P, P') = \alpha_P \delta_{PP'} . \tag{E2}$$

This is not the case for A, B, C since it is easy to verify that

$$\mathrm{tr}(AB) = -1 . \tag{E3}$$

(c) In view of Eqs. (3.A.33), (3.A.34), (3.A.8) it looks as if

$$\overline{V}_{(1B)} = \overline{V}_{(1E)} + \tfrac{6}{5}\overline{V}_{(1N)}, \quad \overline{V}_{(1A)} = -\overline{V}_{(1E)} + \tfrac{9}{5}\overline{V}_{(1N)} \tag{E4}$$

would be better tensor combinations to use. However the corresponding tensors $T_\mu^{(1A)}$ and $T_\mu^{(1B)}$ would then not be orthogonal.

[1] E. Corrigan, D. Fairlie, J. Nuyts, and D. Olive, Nucl. Phys. B **106**, 475 (1976).

[2] 't Hooft, Nucl. Phys. B **79**, 276 (1974).

[3] For SU(2), the little group contains E only [(1.7)]. For SU(3), the little groups in the $3 \to 3$ case has a $U(1)\otimes U(1)$ structure, in the $3 \to 2\oplus 1$ case a $U(1)\otimes U(1)\otimes U(1)$ structure. In all these case ℓ is Abelian. In SU(4) for the first time ℓ can be non-Abelian.

[4] F. Englert and P. Windey, Phys. Rev. D **14**, 2728 (1976). We thank these authors for showing us the general theorem they had found prior to publication and for discussion about weights and roots of the dual group. Related material is also found in D. Olive, Nucl. Phys. B **113**, 413 (1976) and in P. Goddard, J. Nuyts, and D. Olive, Nucl. Phys. B **125**, 1 (1977), where direct use is made of ℓ_0. See also A. Goldhaber and D. Wilkinson, "Spherically symmetric monopoles," Stony Brook (1977-30).

Stationary gravitational fields of a charged perfect fluid

S. Kloster and A. Das

Department of Mathematics, Simon Fraser University, Burnaby, British Columbia, V5A 1S6, Canada
(Received 30 December 1976)

Einstein's field equations which include electromagnetism are investigated when the metric admits a timelike Killing motion and the source is a charged perfect fluid under isometric motion. It is shown that the pressure must necessarily be a function of the electrostatic and gravitational potentials. A class of solutions is found under the following simplifying assumptions: (i) The pressure is a constant, (ii) the Lorentz force vanishes, and (iii) the magnetic and twist potentials are functionally related. In this class the ratio of $\sigma/(\rho+3p)$ is a constant and this resembles an equilibrium condition. Finally a four-parameter group (maximal) is supplied which can generate new solutions of this class.

1. INTRODUCTION

In this paper we investigate the stationary electrogravitational field equations in presence of a perfect fluid in isometric motion (that is, the fluid motion is along Killing direction). From the equation of motion in this case, one can show that the pressure must be a function of electrostatic and gravitational potentials, or in other words an "equation of state" must exist. If, moreover, the condition that $\sigma/(\rho+3p)$ is constant is stipulated (this resembles an equilibrium), then the pressure can be specified further.

Som and Ray Chaudhuri,[1] and Misra, Pandey and Srivastava,[2,3] have shown that the ratio σ/ρ is a constant for a stationary perfect fluid under the following three assumptions: (i) $p=0$, (ii) the Lorentz force vanishes, and (iii) the magnetic and twist potentials are functionally related. This result is generalized here (Sec. 3) to the constant pressure case: (i') $p=\text{const}$, where it is found that $\sigma/(\rho+3p)$ is constant.

Under the assumptions (i'), (ii), and (iii), it is shown that the field equations reduce to the equations for cosmological dust in isometric motion,[4] and hence any such solution can generate a charged perfect fluid solution. As a corollary, every static vacuum metric can generate a stationary charged dust solution. Thus we have generalized the results of Ehlers[4] for the uncharged case to the charged case.

Another class of solutions is found which depends on solutions of Laplace's equation in a three-dimensional space of constant negative curvature. A constructive method is given for finding these solutions. However, all of these solutions have some unphysical aspects.

Finally a table is furnished to show how new stationary charged perfect fluid solutions can be derived by the action of a four-parameter group on already known solutions.

2. FIELD EQUATIONS

In a previous paper,[5] it was shown that when a charged perfect fluid is in isometric motion one can define two complex potentials $\Gamma=e^{\omega}-\frac{1}{2}\kappa|\Phi|^2+i\chi$ and $\Phi=A+iB$, where $A_{,i}=-F_{4i}$ and $B_{,i}=F^*_{4i}$. The metric is given by

$$-e^{-\omega}(g_{\alpha\beta}dx^{\alpha}dx^{\beta})+e^{\omega}(a_{\alpha}dx^{\alpha}+dx^4)^2, \tag{2.1}$$

with $\tau_{\alpha}=-e^{2\omega}\,\mathrm{curl}(a_{\alpha})$ and $\chi_{,\alpha}=\tau_{\alpha}+\frac{1}{2}i\kappa(\Phi^*\Phi_{,\alpha}-\Phi\Phi^*_{,\alpha})$. Throughout this paper it is assumed that no monopoles are present; i.e., $\sigma^*=\sigma$. Then the field equations[5] become

$$\sigma_{\alpha\beta}\equiv R_{\alpha\beta}-\kappa e^{-\omega}\,\mathrm{Re}(\Phi_{,\alpha}\Phi^*_{,\beta}+2pg_{\alpha\beta})+\tfrac{1}{2}e^{-2\omega}\,\mathrm{Re}(\Gamma^*_{,\alpha}+\kappa\Phi\Phi^*_{,\alpha})(\Gamma_{,\beta}+\kappa\Phi^*\Phi_{,\beta})=0, \tag{2.2}$$

$$\nu\equiv\Delta_2\Phi-e^{-\omega}\Phi_{,\alpha}(\Gamma^{,\alpha}+\kappa\,\Phi^*\Phi^{,\alpha})-e^{-\omega/2}\sqrt{4\pi}\,\sigma=0, \tag{2.3}$$

$$\mu\equiv\Delta_2\Gamma-e^{-\omega}\Gamma_{,\alpha}(\Gamma^{,\alpha}+\kappa\Phi^*\Phi^{,\alpha})+\kappa(-3p-\rho+e^{-\omega/2}\sqrt{4\pi}\,\sigma\Phi^*)=0. \tag{2.4}$$

The number of unknowns is $13=6\ (g_{\alpha\beta})+2(\Gamma)+2(\Phi)+1(p)+1(\rho)+1(\sigma)$, and the number of equations in the system (2.2)—(2.4) is $13=6\ (\sigma_{\alpha\beta})+2(\nu)+2(\mu)+3$ (coordinate conditions). It would appear that the equation of state is already determined implicitly in this problem.

This fact is explicitly shown by the equation of motion[6]

$$p_{,\mu}=\tfrac{1}{2}(\rho+p)\,\omega_{,\mu}-\sqrt{4\pi}\,e^{-\omega/2}\sigma A_{,\mu}, \tag{2.5}$$

which shows than an "equation of state" must exist, namely,

$$p=p(A,\omega),\qquad p_{,A}=-\sqrt{4\pi}\,e^{-\omega/2}\sigma,\qquad p_{,\omega}=\tfrac{1}{2}(\rho+p). \tag{2.6}$$

Furthermore, if a condition analogous to equilibrium

$$\sigma/(\rho+3p)=b=\text{const} \tag{2.7}$$

is introduced, then (2.6) yields a partial differential equation

$$2b\sqrt{4\pi}\,e^{-\omega/2}p_{,\omega}+p_{,A}+2b\sqrt{4\pi}\,e^{-\omega/2}p=0. \tag{2.8}$$

The above equation can be integrated by using the characteristic curves[7] to obtain

$$p=e^{-\omega}f(\sqrt{4\pi}\,bA-e^{\omega/2}), \tag{2.9}$$

where f is an arbitrary differentiable function of $\sqrt{4\pi}\,bA-e^{\omega/2}$.

3. EXACT SOLUTIONS

Solutions are obtained by making the following two

 Copyright © 1977 American Institute of Physics

assumptions:

(i′) $p = \text{const}$ and (ii) $A = 0$.

The first assumption means that the mechanical force vanishes, while the second means that the Lorentz force vanishes. By the equation of motion (2.5), ω must be constant, say $\omega = 2\ln c$.

Now we have $\Phi = iB$ and $\Gamma = c^2 - \frac{1}{2}\kappa B^2 + i\chi$;

hence $\Gamma'^{\alpha} + \kappa\Phi^*\Phi'^{\alpha} = i\chi'^{\alpha}$. Equation (2.3) becomes

$$i\Delta_2 B + c^{-2}B_{,\alpha}\chi'^{\alpha} = c^{-1}\sqrt{4\pi}\,\sigma. \tag{3.1}$$

Thus we must have, from the imaginary and real parts,

$$\Delta_2 B = 0, \tag{3.2}$$

and

$$c^{-1}(4\pi)^{-1/2}B_{,\alpha}\chi'^{\alpha} = \sigma. \tag{3.3}$$

By using (3.2) and (3.3), Eq. (2.4) can be reduced to the following two equations:

$$\Delta_2\chi = 0, \tag{3.4}$$

$$-\kappa B_{,\alpha}B'^{\alpha} + c^{-2}\chi_{,\alpha}\chi'^{\alpha} = \kappa(3p + \rho). \tag{3.5}$$

Finally, Eq. (2.2) becomes

$$R_{\alpha\beta} = c^{-2}\kappa(B_{,\alpha}B_{,\beta} + 2pg_{\alpha\beta}) - \tfrac{1}{2}c^{-4}\chi_{,\alpha}\chi_{,\beta}. \tag{3.6}$$

This is a generalization of Eq. (2.29) of Ehlers.[4]

We now assume that χ and B are functionally related. Because of (3.2) and (3.4), this must be a linear relationship,[3] which we write as

$$\chi = ac\sqrt{2\kappa}\,B, \tag{3.7}$$

where a is a real constant (a possible additive constant has been absorbed by a gauge transformation). Then the system of Einstein–Maxwell equations (3.2)–(3.6) is reduced to the following:

$$\Delta_2 B = 0, \tag{3.2}$$

$$\sigma = 2aB_{,\alpha}B'^{\alpha}, \tag{3.3′}$$

$$3p + \rho = (2a^2 - 1)B_{,\alpha}B'^{\alpha}, \tag{3.5′}$$

$$R_{\alpha\beta} = c^{-2}\kappa[(1 - a^2)B_{,\alpha}B_{,\beta} + 2pg_{\alpha\beta}]. \tag{3.6′}$$

Furthermore, Eq. (3.2) follows from (3.6′) by using the contracted Bianchi identity in V_3. Note that

$$\frac{\sigma}{\rho + 3p} = \frac{2a}{2a^2 - 1}, \text{ a constant.} \tag{3.8}$$

To obtain solutions of this system, we need only solve Eq. (3.6′) since (3.3′) and (3.5′) may be regarded as definitions of σ and ρ.

Theorem 1: If $-g_{\alpha\beta}dx^{\alpha}\,dx^{\beta} + (\bar{a}_{\alpha}\,dx^{\alpha} + dt)^2$ is the local metric of a cosmological dust in isometric motion with twist potential $\bar{\chi}$ and cosmological constant Λ, then $-c^{-2}\,(g_{\alpha\beta}dx^{\alpha}\,dx^{\beta}) + c^2\,(a_{\alpha}\,dx^{\alpha} + dt)^2$ is the local metric of charged perfect fluid in isometric motion with twist potential $\chi = ac^2(a^2 - 1)^{-1/2}\bar{\chi}$, magnetic potential

$B = c[2\kappa(a^2 - 1)]^{-1/2}\bar{\chi}$, charge density $\sigma = 2aB_{,\alpha}B'^{\alpha}$,

and pressure $p = c^2\kappa^{-1}\Lambda$.

Proof: If we put the above values into Eq. (3.6′), we obtain

$$R_{\alpha\beta} = -\tfrac{1}{2}\bar{\chi}_{,\alpha}\,\bar{\chi}_{,\beta} + 2\Lambda g_{\alpha\beta}. \tag{3.9}$$

But this is just the equation for cosmological dust in isometric motion, as given by Eq. (2.29) of Ehlers. As examples of such solutions, we mention those of Wright.[8]

When $\Lambda = 0$, (3.9) is the static vacuum equation. Hence we obtain the following.

Corollary: If $-e^{-\bar{\chi}}(g_{\alpha\beta}dx^{\alpha}\,dx^{\beta}) + e^{\bar{\chi}}dt^2$ is locally static vacuum metric, then $-c^{-2}(g_{\alpha\beta}\,dx^{\alpha}\,dx^{\beta}) + c^2(a_{\alpha}\,dx^{\alpha} + dt)^2$ is the local metric of a charged dust in isometric motion, with χ and B as above and $p = 0$.

Another way to find solutions of (3.6′) is to put $a = 1$, giving

$$R_{\alpha\beta} = 2c^{-2}\kappa\,pg_{\alpha\beta}. \tag{3.6″}$$

This is just the equation[9] for a space of constant curvature $-c^{-2}\kappa p$. Thus solutions can be obtained as follows. Choose any nonnegative constant value for p. The associated V_3 is the space of constant curvature $-c^{-2}\kappa p$. Now Eq. (3.2) is not an automatic consequence of (3.6″) because it has become uncoupled. However, there is no difficulty in finding solutions of (3.2).

The metric of the V_3 can be written

$$ds^2 = k^{-1}[dr^2 + \sinh^2 r(d\theta^2 + \sin^2\theta\,d\phi^2)], \tag{3.10}$$

where $k = c^{-2}\kappa p$. Equation (3.2) becomes

$$(\sinh^2 rB_{,r})_{,r} + (\sin\theta)^{-1}(\sin\theta\,B_{,\theta})_{,\theta} + (\sin\theta)^{-2}B_{,\phi\phi} = 0. \tag{3.11}$$

Using separation of variables, the following general solution of (3.11) is found:

$$B(r, \theta, \phi) = \sum_{l,m}\alpha_{lm}f_l(R)\,Y_l^m(\theta, \phi) + \text{c.c.}, \tag{3.12}$$

where $R = \coth r$, $f_l(R) = (R^2 - 1)P_l'(R)$, P_l is the lth Legendre polynomial, $Y_l^m(\theta, \phi) \equiv P_l^m(\cos\theta)\,e^{im\phi}$, the spherical harmonics, and the α_{lm} are arbitrary complex constants.

Note that $B_{,R}$, $B_{,\theta}$, and $B_{,\phi}$ all involve power series in R, so there must be some radius $R > 1$ in which all three converge. Thus these three functions remain

TABLE I. a and d are arbitrary real parameters and δ is an arbitrary complex parameter.

	scale	rotational gauge	electromagnetic gauge
Γ'	$a^2\Gamma$	$\Gamma + id$	$\Gamma - 2\kappa\delta\Phi - 2\kappa\lvert\delta\rvert^2$
Φ'	$a\Phi$	Φ	$\Phi + 2\delta^*$
p'	a^2p	p	p
ρ'	$a^2\rho$	ρ	ρ
σ'	$a^2\sigma$	σ	σ

finite in the limit ($R \to 1^+$). But

$$B_{,\alpha}B^{,\alpha} = k(R^2-1)[(R^2-1)(B_{,R})^2 + (B_{,\theta})^2 + \sin^{-2}\theta\,(B_{,\phi})^2], \tag{3.13}$$

so $B_{,\alpha}B^{,\alpha} \to 0$ as $r \to \infty$. Then Eq. (3.5′) shows that if p is nonzero, ρ will become negative as $r \to \infty$.

In conclusion, we mention that all of these solutions may be subjected to the four-parameter group of transformations[5] which preserve charged fluid solutions having no monopole currents. These transformations are generated by the basic transformations in Table I.

ACKNOWLEDGMENTS

The authors are indebted to NRC Grant No. 3993.

[1]M.M. Som and A.K. Ray Chaudhuri, Proc. Roy. Soc. Lond. A 304, 81 (1968).
[2]R.M. Misra, D.B. Pandey, and D.C. Srivastava, Nuovo Cimento B 9, 64 (1972).
[3]R.M. Misra, D.B. Pandey, and D.C. Srivastava, Proc. Indian Natl. Acad. Sci. A 39, 132 (1973).
[4]J. Ehlers, Proc. Int. Conf. Rel Theories Grav. Lond. 2, 15 (1965).
[5]A. Das and S. Kloster, J. Math. Phys. 18, 586 (1977). These equations can be found in D. Kramer, G. Neugebauer, and H. Stephani, Forschr. Psysik 20, Heft 1, 1 (1972).
[6]Cf. M.L. Woolley, Proc. Roy. Soc. Lond. A 336, 273 (1974), Eq. (2.17).
[7]R. Dennemeyer, *Introduction to Partial Differential Equations and Boundary Value Problems* (McGraw-Hill, New York, 1968), p. 13.
[8]J.P. Wright, J. Math. Phys. 6, 103 (1965).
[9]T.J. Willmore, *Introduction to Differential Geometry* (Oxford U.P., London, 1959), p. 237.

Lee model with 3 V particles

Cvavb. Chandra Raju

C.T.A.E., University of Udaipur, Pratapnagar, Udaipur-313001, India
(Received 19 April 1976; revised manuscript received 1 March 1977)

The $3V$-particle Lee model is proposed. The $N\theta$ scattering amplitude, the $V\theta$ scattering amplitude and the $n\theta\theta$ production amplitude are evaluated. The integral equation in the $V\theta$ sector is solved. The usual deductive method cannot be applied here to solve the integral equation of the $V\theta$ sector. A nondeductive method is applied to solve the integral equation. The solution obtained correctly reduces to the solution of the $V\theta$ integral equation of the ordinary Lee model, whenever any two bare interaction constants in the Hamiltonian are switched off. The two resonances that appear in the $N\theta$ sector, again appear in the $V\theta$ sector at the same energies.

INTRODUCTION

A Lee model with more V particles is more interesting mathematically than the ordinary Lee model. Recently the generalized Källén–Pauli (GKP) equation in the $V\theta$ sector of the $2V$-particle Lee model was solved.[1] While generalizing the $2V$-particle Lee model to any nV- particle Lee model one comes across certain restrictions on the bare masses and the bare coupling constants. These restrictions do not arise in the $2V$-particle Lee model. Moreover the restrictions and the subtleties involved cannot be shown explicitly in the case of any nV-particle Lee model.

The necessity of these restrictions and other related matters can be better appreciated if we work out the $3V$-particle Lee model in detail. In this model we can explicitly derive these restrictions [see Eq. (39)] and justify their necessity. Once we follow the $3V$-particle model closely, it is a simple matter to generalize the results to any nV-particle Lee model.

In Sec. I the model is presented. In Secs. II and III the physical $|\mathbf{V}\rangle$ state and the properties of the related functions are discussed, whereas in Sec. IV the $N\theta$ scattering amplitude is evaluated. The GKP equation and its solution are presented in Secs. V, VI, and VII. In the last two sections we evaluate the $V\theta$ scattering amplitude and the $N\theta\theta$ production amplitude.

I. THE MODEL

In this model there are three V particles and all of them are subjected to the elementary interaction $V_1 \rightleftharpoons N\theta$, $V_2 \rightleftharpoons N\theta$, and $V_3 \rightleftharpoons N\theta$. The bare parameters are adjusted in such a way that only one of the V particles is stable, and the remaining two appear as resonances in the $N\theta$ sector. The model is described by the Hamiltonian, H, where.

$$H=\sum_{i=1}^{3} m_i V_i^{\dagger} V_i + m_N N^{\dagger} N + \sum_p \omega_p a_p^{\dagger} a_p + \sum_p \sum_{i=1} f(\omega_p) g_i V_i^{\dagger} N a_p + \sum_p \sum_{i=1} f(\omega_p) g_i V_i N^{\dagger} a_p^{\dagger}. \tag{1}$$

In Eq. (1), $\omega_p^2=\mu^2+p^2$, μ being the mass of the θ particle, and p is its linear momentum. Morevoer m_1, m_2, and m_3 are the real bare masses of V_1, V_2, and V_3. The bare interaction strengths g_1, g_2, and g_3 are real. In the Hamiltonian the mass of the N particle m_N will be set equal to zero. In Eq. (1), $f(\omega_p)$ is given by

$$f(\omega_p)=\frac{u(\omega_p)}{(2\omega_p\Omega)^{1/2}}. \tag{2}$$

The function $u(\omega_p)$ is so chosen as to make all the relevant integrals finite and so as not to allow any ghost states. Moreover Ω is the volume of quantization. The equal time commutators obeyed by the field operators are

$$[a_p, a_{p'}^{\dagger}]=\delta(\mathbf{p}-\mathbf{p}'),$$
$$[V_i, V_i^{\dagger}]=1, \quad \text{where } i=1,\ldots,3, \tag{3}$$

and

$$[N, N^{\dagger}]=1.$$

All other commutators between the field operators vanish. In Eq. (3) the first commutator should be taken as a Krönecker delta when the volume of quantization is finite.

II. THE PHYSICAL $|V\rangle$ STATE

The lowest sector is spanned by the physical $|\mathbf{V}\rangle$ particle state and the $N\theta$ scattering states. We choose the stable $|\mathbf{V}\rangle$ to have a physical mass equal to zero as in the $2V$-particle model. The physical $|\mathbf{V}\rangle$ state can be written as a linear combination of the bare states of the sector and we get

$$|\mathbf{V}\rangle=\sum_{i=1}^{3} \alpha_i V_i^{\dagger}|0\rangle + \sum_p \phi_1(\omega_p) N^{\dagger} a_p^{\dagger}|0\rangle. \tag{4}$$

The Schrödinger equation,

$$H|\mathbf{V}\rangle = E|\mathbf{V}\rangle, \tag{5}$$

yields,

$$(E-m_1)\alpha_1 - g_1 \sum_p f(\omega_p)\phi_1(\omega_p)=0, \tag{6}$$

$$(E-m_2)\alpha_2 - g_2 \sum_p f(\omega_p)\phi_1(\omega_p)=0, \tag{7}$$

$$(E-m_3)\alpha_3 - g_3 \sum_p f(\omega_p)\phi_1(\omega_p)=0, \tag{8}$$

and

$$\phi_1(\omega_p)=-(\alpha_1 g_1+\alpha_2 g_2+\alpha_3 g_3)\frac{f(\omega_p)}{\omega_p - E}. \tag{9}$$

In the above equations we have taken the physical mass of $|\mathbf{V}\rangle$ equal to E. Subsequently we will set E

 Copyright © 1977 American Institute of Physics

equal to zero. Inserting Eq. (9) into Eqs. (6)–(8) and taking the determinant of the coefficients of α_1, α_2, and α_3 equal to zero, we obtain,

$$H_3(E)=0, \tag{10}$$

where

$$H_3(E)=(E-m_1)(E-m_2)(E-m_3) + \sigma_3(E)I_1(E). \tag{11}$$

In Eq. (11) we have

$$\sigma_3(E)=g_1^2(E-m_2)(E-m_3) + g_2^2(E-m_3)(E-m_1)+g_3^2(E-m_1)(E-m_2), \tag{12}$$

and

$$I_1(E)=\frac{1}{4\pi^2}\int_\mu^\infty \frac{pu^2(\omega)}{\omega-E}d\omega. \tag{13}$$

By setting $E=0$ in Eq. (10) we obtain the eigenvalue condition corresponding to a $|\mathbf{V}\rangle$ state of physical mass zero. The eigenvalue condition is equivalent to the following requirement:

$$\left[\frac{g_1^2}{m_1}+\frac{g_2^2}{m_2}+\frac{g_3^2}{m_3}\right]^{-1}=I_1(0). \tag{14}$$

The normalization of the physical $|\mathbf{V}\rangle$ state is equivalent to the following requirement:

$$(\alpha_1 g_1+\alpha_2 g_2+\alpha_3 g_3)^{-2}=\left(\frac{g_1^2}{m_1^2}+\frac{g_2^2}{m_2^2}+\frac{g_3^2}{m_3^2}\right)\Big/\left(\frac{g_1^2}{m_1}+\frac{g_2^2}{m_2}+\frac{g_3^2}{m_3}\right)^2 + \frac{1}{4\pi^2}\int_\mu^\infty \left|\frac{pu^2(\omega)}{\omega^2}\right| d\omega. \tag{15}$$

It should be noted that the right-hand side of Eq. (15) is a positive real constant since g_1, g_2, g_3, m_1, m_2, and m_3 are all real parameters.

We define a function $H_3^+(\omega)$ where

$$H_3^+(\omega)=(\omega-m_1)(\omega-m_2)(\omega-m_3) + \sigma_3(\omega)\frac{1}{4\pi^2}\int_\mu^\infty \frac{p'u^2(\omega')}{\omega'-\omega-i\epsilon}d\omega'. \tag{16}$$

We follow the convention that $H_3^\pm(\omega)=H_3(\omega\pm i\epsilon)$. In Eq. (16), $\sigma_3(\omega)$ can be obtained from Eq. (12) by writing ω for E. The function $H_3^+(\omega)=0$ for $\omega=0$, as required by the Schrödinger equation. We introduce another function $G_3^+(\omega)$, where

$$G_3^+(\omega)=G_3(\omega+i\epsilon)=\frac{H_3(\omega+i\epsilon)}{\sigma_3(\omega)}. \tag{17}$$

Since $H_3(0)=0$, we require that

$$G_3(0)=0. \tag{18}$$

In order to insert the requirement of Eq. (18) into Eq. (17), we first find $G_3(0)$ from Eq. (17) and then subtract it from $G_3^+(\omega)$. This yields

$$G_3^+(\omega)=\omega h_3^+(\omega), \tag{19}$$

where

$$h_3^+(\omega)=a_1a_2\omega^2-\omega[(m_1+m_2+m_3)a_1a_2+m_1m_2m_3] + a_1a_2(m_1m_2+m_2m_3+m_3m_1)-(a_1+a_2)m_1m_2m_3 + \frac{1}{4\pi^2}\int_\mu^\infty \frac{p'u^2(\omega')}{\omega'(\omega'-\omega-i\epsilon)}d\omega'. \tag{20}$$

In Eq. (20) the two constants a_1 and a_2 are given by

$$a_1=\frac{g_1^2(m_2+m_3)+g_2^2(m_3+m_1)+g_3^2(m_1+m_2)+\sqrt{q}}{2(g_1^2+g_2^2+g_3^2)} \tag{21}$$

and

$$a_2=\frac{g_1^2(m_2+m_3)+g_2^2(m_3+m_1)+g_3^2(m_1+m_2)-\sqrt{q}}{2(g_1^2+g_2^2+g_3^2)}, \tag{22}$$

with

$$q=[g_1^2(m_2-m_3)]^2+[g_2^2(m_3-m_1)]^2+[g_3^2(m_1-m_2)]^2 - 2g_1^2g_2^2(m_2-m_3)(m_3-m_1)-2g_2^2g_3^2(m_3-m_1)(m_1-m_2) - 2g_3^2g_1^2(m_1-m_2)(m_2-m_3). \tag{23}$$

It will be quite useful to note the following relations which can be readily obtained from Eqs. (21), (22), and (23):

$$a_1+a_2=\frac{g_1^2(m_2+m_3)+g_2^2(m_3+m_1)+g_3^2(m_1+m_2)}{(g_1^2+g_2^2+g_3^2)}, \tag{24}$$

$$a_1a_2=\frac{g_1^2m_2m_3+g_2^2m_3m_1+g_3^2m_1m_2}{(g_1^2+g_2^2+g_3^2)}. \tag{25}$$

From Eq. (23) it is clear that q is not a perfect square. Hence the two constants a_1 and a_2 can have complex values as well. This must be clear from Eqs. (21) and (22) where the square root of q occurs. We can always choose the real bare masses and the real bare coupling constants in such a way that a_1 and a_2 are real. Even if a_1, and a_2 are complex, they are also complex conjugates of each other [see Eqs. (21) and (22)]. Hence the quantities a_1a_2 and (a_1+a_2) are always real. That this is so is also evident from Eqs. (24) and (25).

In Table I we give the values of a_1, a_2, (a_1+a_2), and a_1a_2 when one of the bare interaction constants is set equal to zero with the remaining two interaction constants being nonzero. The table is self explanatory.

TABLE I.

	a_1	a_2	a_1+a_2	a_1a_2
$g_1^2=0$, $g_2^2\neq 0$, $g_3^2\neq 0$	ω_{23}	m_1	$m_1+\omega_{23}$	$m_1\omega_{23}$
$g_2^2=0$, $g_3^2\neq 0$, $g_1^2\neq 0$	m_2	ω_{13}	$m_2+\omega_{13}$	$m_2\omega_{13}$
$g_3^2=0$, $g_2^2\neq 0$, $g_1^2\neq 0$	ω_{12}	m_3	$m_3+\omega_{12}$	$m_3\omega_{12}$

In Table I, we use the symbols ω_{ij}, where

$$\omega_{ij}=\frac{g_i^2 m_j+g_j^2 m_i}{g_i^2+g_j^2}. \tag{26}$$

It proves very useful to define a renormalized charge through Eq. (20). To this end, we put $\omega=0$ in Eq. (20). This yields

$$h_3(0)=a_1a_2(m_2m_3+m_1m_2+m_1m_3)-(a_1+a_2)m_1m_2m_3 + \frac{1}{4\pi^2}\int_\mu^\infty \frac{p'u^2(\omega')}{\omega'^2}d\omega'. \tag{27}$$

We take

$$h_3(0)=\frac{1}{g^2}, \tag{28}$$

where "g" is known as the renormalized charge. Comparing Eqs. (15), (27), and (28) one easily finds that

$$g^2=(\alpha_1 g_1+\alpha_2 g_2+\alpha_3 g_3)^2. \tag{29}$$

Whether a_1 and a_2 are real or complex, g^2 is always real and positive. From Eq. (15) we know that the right-hand side of Eq. (29) is a positive real constant. Moreover in Eq. (27), a_1 and a_2 occur in the combinations a_1a_2 and a_1+a_2. As mentioned earlier, when a_1 and a_2 are complex, they are also complex conjugates of each other. So the expression given by Eq. (27) is always a real quantity.

We now re-express $h_3^+(\omega)$ of Eq. (20) in such a way that the requirement of Eq. (28) is automatically taken care of. Thus we have

$$h_3^+(\omega)=h_3^+(\omega)-h(0)+\frac{1}{g^2}. \tag{30}$$

In Eq. (30) we use Eqs. (20) and (27) for the first two factors on the right-hand side and simplify it. We add $1/g^2$ to the resulting expression. This yields

$$h_3^+(\omega)=\frac{1}{g^2}\left[1+\frac{g^2(\omega^2B_1+\omega B_2)}{(\omega-a_1)(\omega-a_2)} + \frac{g^2\omega}{4\pi^2}\int_\mu^\infty \frac{p'u^2(\omega')}{\omega'^2(\omega'-\omega-i\epsilon)}d\omega'\right], \tag{31}$$

where

$$B_1=[(a_1a_2)^2-a_1a_2(m_1m_2+m_2m_3+m_3m_1) + (a_1+a_2)m_1m_2m_3]/(g_1^2+g_2^2+g_3^2)(a_1a_2)^2 \tag{32}$$

and

$$B_2=[a_1a_2(a_1+a_2)(m_1m_2+m_2m_3+m_3m_1)+a_1a_2m_1m_2m_3 - (a_1a_2)^2(m_1+m_2+m_3)-(a_1+a_2)^2m_1m_2m_3] \times[(g_1^2+g_2^2+g_3^2)(a_1a_2)^2]^{-1}. \tag{33}$$

It should again be noted that B_1 and B_2 are always real whether a_1 and a_2 are real or complex, since a_1 and a_2 occur in the combinations a_1a_2 and (a_1+a_2) in B_1 and B_2.

In Table II the values of B_1 and B_2 are given when one of the bare interaction constant out of the three is set equal to zero, with the remaining two interaction constants being nonzero. The table is self explanatory. The symbols ω_{ij} are given by Eq. (26).

III. THE PROPERTIES OF THE FUNCTION $h_3^+(\omega)$

Henceforth whenever we refer to the function $h_3^+(\omega)$ we mean Eq. (31) only. By $G_3^+(\omega)$ we mean $\omega h_3^+(\omega)$. The function $h_3^+(\omega)$ has simple poles at $\omega=a_1$ and $\omega=a_2$. The function $h_3^+(\omega)$ is known as a twice subtracted function. In fact, if we take $g_3^2=0$, the function $h_3^+(\omega)$ reduces to the function $h^+(\omega)$ of Ref. 1. The function $h^+(\omega)$ of Ref. 1 is given in Appendix B here. With the help of Tables I and II we can easily show that $h_3^+(\omega)$ reduces to the following function whenever $g_1^2=0$:

$$h_2^+(\omega)=\frac{1}{g'^2}\left[1-\frac{g'^2\omega B_1'}{(\omega_{23}-\omega)}+\frac{g'^2\omega}{4\pi^2}\int_\mu^\infty \frac{p'u^2(\omega')}{\omega'^2(\omega'-\omega-i\epsilon)}d\omega'\right], \tag{34}$$

where

$$g'^2=(\alpha_2g_2+\alpha_3g_3)^2 \tag{35}$$

and

$$B_1'=\frac{(\omega_{23}-m_2)(\omega_{23}-m_3)}{(g_2^2+g_3^2)\omega_{23}^2}. \tag{36}$$

The function $h_2^+(\omega)$ is the twice subtracted function that we would obtain in the $2V$-particle Lee model with bare interaction constants g_2 and g_3, and with bare masses m_2 and m_3. In an analogus manner, one can easily show that the function $h_3^+(\omega)$ reduces to a function similar in "content" to the function $h_2^+(\omega)$ of Eq. (34) whenever $g_2^2=0$. So the function $h_3^+(\omega)$ reduces to the twice subtracted function of the $2V$-particle Lee model whenever any one interaction constant, out of the three bare interaction constants, is switched off.

In Ref. 1 we have shown that the twice subtracted function of the $2V$-particle Lee model reduces to the twice subtracted function of the ordinary Lee model, whenever one of the interaction constants, out of the two is set equal to zero. From the foregoing discussion, it follows that the function $h_3^+(\omega)$ reduces to the following function of the ordinary Lee model, whenever any two of the bare interaction constants, out of the three, are switched off:

$$\mathcal{G}_1^+(\omega)=\frac{1}{g_0^2}\left[1+\frac{g_0^2\omega}{4\pi^2}\int_\mu^\infty \frac{p'u^2(\omega')}{\omega'^2(\omega'-\omega-i\epsilon)}d\omega'\right], \tag{37}$$

where

TABLE II.

	B_1	$-B_2$
$g_1^2=0$ $g_2^2\neq 0$ $g_3^2\neq 0$	$\frac{(\omega_{23}-m_2)(\omega_{23}-m_3)}{(g_2^2+g_3^2)\omega_{23}^2}$	$\frac{m_1(\omega_{23}-m_2)(\omega_{23}-m_3)}{(g_2^2+g_3^2)\omega_{23}^2}$
$g_2^2=0$ $g_3^2\neq 0$ $g_1^2\neq 0$	$\frac{(\omega_{13}-m_1)(\omega_{13}-m_3)}{(g_1^2+g_3^2)\omega_{13}^2}$	$\frac{m_2(\omega_{13}-m_1)(\omega_{13}-m_3)}{(g_1^2+g_3^2)\omega_{13}^2}$
$g_3^2=0$ $g_1^2\neq 0$ $g_2^2\neq 0$	$\frac{(\omega_{12}-m_1)(\omega_{12}-m_2)}{(g_1^2+g_2^2)\omega_{12}^2}$	$\frac{m_3(\omega_{12}-m_1)(\omega_{12}-m_2)}{(g_1^2+g_2^2)\omega_{12}^2}$

$$g_0^2=(\alpha_i g_i)^2. \tag{38}$$

In Eq. (38) there is no summation on i. If $g_2^2=0=g_3^2$ then $i=1$ in Eq. (38). On the other hand, if $g_3^2=0=g_1^2$, then $i=2$ in Eq. (38). Similarly, $i=3$, if $g_1^2=0=g_2^2$. The constant g_0 is known as the renormalized charge. The function $\mathcal{G}_1^+(\omega)$ is the twice subtracted function of the ordinary Lee model. In Refs. 2—5 the authors have used once subtracted function.

We adjust the bare parameters of the $3V$-particle Lee model in such a way that $a_1<\mu$ and $a_2<\mu$. This would enable us to find the solution of the integral equation in the $V\theta$ sector. Since μ is a real number we require that a_1 and a_2 must also be real. The constants a_1 and a_2 can be real if and only if

$$q\geq 0. \tag{39}$$

There are many free parameters in the theory, and we can always satisfy the condition imposed by Eq. (39). For example, Eq. (39) can be reduced to the following requirement:

$$\{g_1^2(m_2-m_3)-g_2^2(m_3-m_1)+g_3^2(m_1-m_2)\}^2 \geq 4g_1^2g_3^2[m_2m_1-m_2^2+m_3m_2-m_3m_1]. \tag{40}$$

In Eq. (40) if we take m_2 negative, and m_1 and m_3 positive, the right-hand side of Eq. (40) will be a negative number whereas the left-hand side is a positive number and hence Eq. (40) will always be true. In a similar manner we can obtain several other conditions analogous to Eq. (40) from Eq. (39), by a proper elimination of the parameters involved. The point is that the basic requirement of Eq. (39) can always be satisfied since we have six free parameters at our disposal, and the number of restrictions on them are less than six. Restrictions of the type of Eq. (39) are not at all necessary in the $2V$-particle Lee model. Such restrictions must also be present in any nV-particle Lee model with $n>3$. In general it will not be easy to derive conditions of the type of Eq. (39) for any nV-particle model, with $n>3$.

Suppose $q^{1/2}$ is not real. Then a_1 and a_2 will be complex numbers. What is the significance of these complex numbers? Do they have anything to do with the Lee—Wick[6,7] suggestions? These questions require a closer study of the model. For the time being we ignore these questions, and assume that $q^{1/2}$ is real.

The function $h_3^+(\omega)$ of Eq. (31) has a branch point at $\omega=\mu$. We attach a cut from $\omega=\mu$ to ∞.

The function $h_3^+(\omega)$ of Eq. (31) can be re-expressed as

$$h_3^+(\omega)=\mathcal{G}_3^+(\omega)+\frac{(\omega^2B_1+\omega B_2)}{(\omega-a_1)(\omega-a_2)}, \tag{41}$$

where

$$\mathcal{G}_3^+(\omega)=\frac{1}{g^2}\left(1+\frac{g^2\omega}{4\pi^2}\int_\mu^\infty \frac{p'u^2(\omega')}{\omega'^2(\omega'-\omega-i\epsilon)}d\omega'\right). \tag{42}$$

The function $\mathcal{G}_3^+(\omega)$ of Eq. (42) is similar, in so far as its analytical properties are concerned, to the function $\mathcal{G}_1^+(\omega)$ of Eq. (37). The difference lies only in the constant factors g^2 and g_0^2. It will be useful later to define a function $K_3^+(\omega)$, where

$$K_3^+(\omega)=\omega\mathcal{G}_3^+(\omega). \tag{43}$$

From Eqs. (31) and (42) we easily find that

$$\mathrm{Im}h_3^+(\omega)=\mathrm{Im}\mathcal{G}_3^+(\omega)=\frac{1}{4\pi\omega}pu^2(\omega)\theta(\omega-\mu). \tag{44}$$

From Eqs. (19) and (43) we observe that,

$$\mathrm{Im}G_3^+(\omega)=\mathrm{Im}K_3^+(\omega)=\frac{pu^2(\omega)}{4\pi}\theta(\omega-\mu). \tag{45}$$

Moreover from Eqs. (31), (42), (19), and (43) we easily find that

$$h_3(0)=\mathcal{G}_3(0)=\frac{1}{g^2} \tag{46}$$

and

$$G_3(0)=K_3(0)=0. \tag{47}$$

The integral representation for the inverse of the function $h_3^+(\omega)$ can be easily found with the help of the functions $h_d^+(\omega)$ and $G_d^+(\omega)$, where

$$h_d^+(\omega)=(\omega-a_1)(\omega-a_2)h_3^+(\omega) \tag{48}$$

and

$$G_d^+(\omega)=\omega h_d^+(\omega). \tag{49}$$

It should be noted that the function $h_d^+(\omega)$ does not have any poles. This will be clear if we insert $h_3^+(\omega)$ from Eq. (31) into Eq. (48). Moreover the function $[G_d^+(\omega)]^{-1}$ has a pole at $\omega=0$. Hence we find that

$$\frac{1}{G_d^+(\omega)}=\frac{g^2}{a_1a_2\omega}+\frac{1}{\pi}\int_\mu^\infty\left[\mathrm{Im}\frac{1}{G_d^+(\omega')}\right]\frac{1}{(\omega'-\omega-i\epsilon)}d\omega', \tag{50}$$

since

$$\frac{d}{d\omega}G_d^+(\omega)\Big|_{\omega\to 0}=\frac{a_1a_2}{g^2}. \tag{51}$$

The relation in Eq. (50) can be verified by doing the integral in Eq. (50) as a contour integral. The infinite circle gives no contribution because of the form factor $u(\omega)$ of Eq. (2). Multiplying Eq. (50) by ω, we obtain

$$\frac{1}{h_d^+(\omega)}=\frac{g^2}{a_1a_2}+\frac{\omega}{\pi}\int_\mu^\infty\left(\mathrm{Im}\frac{1}{G_d^+(\omega')}\right)\frac{1}{(\omega'-\omega-i\epsilon)}d\omega'. \tag{52}$$

With the help of Eqs. (52) and (48), one readily finds that

$$\frac{1}{h_3^+(\omega)}=(\omega-a_1)(\omega-a_2)\left[\frac{g^2}{a_1a_2}+\frac{\omega}{\pi}\int_\mu^\infty\left(\mathrm{Im}\frac{1}{G_d^+(\omega')}\right)\times\frac{d\omega'}{(\omega'-\omega-i\epsilon)}\right]. \tag{53}$$

Following Ref. 3, we can easily find an integral representation for the function $[\mathcal{G}_3^+(\omega)]^{-1}$. Thus,

$$\frac{1}{\mathcal{G}_3^+(\omega)}=g^2+\frac{\omega}{\pi}\int_\mu^\infty\left(\mathrm{Im}\frac{1}{K_3^+(\omega')}\right)\frac{1}{(\omega'-\omega-i\epsilon)}d\omega'. \tag{54}$$

Whenever any two bare interaction constants are switched off, $h_3^+(\omega)$ and $\mathcal{G}_3^+(\omega)$ both reduce to $\mathcal{G}_1^+(\omega)$. The inverse of $h_3^+(\omega)$, given by Eq. (53), also reduces to the inverse of $\mathcal{G}_1^+(\omega)$ whenever any two interaction constants are switched off (see Appendix C). However, to solve the integral equation in the $V\theta$ sector we find it more convenient to use an algebraic representation for the inverse of $h_3^+(\omega)$. By a mere rearrangement of Eq. (41) we find that

$$\frac{1}{h_3^+(\omega)} = \frac{1}{\mathcal{G}_3^+(\omega)}\left[1 - \frac{(\omega^2 B_1 + \omega B_2)}{h_d^+(\omega)}\right]. \tag{55}$$

Whenever any two interaction constants are switched off, B_1 and B_2 would go to zero and we would find from (55) that

$$\frac{1}{h_3^+(\omega)} \to \frac{1}{\mathcal{G}_3^+(\omega)} \to \frac{1}{\mathcal{G}_1^+(\omega)}. \tag{56}$$

The integral representation for $[\mathcal{G}_1^+(\omega)]^{-1}$ is identical in form to Eq. (54), only we have to interpret g^2 in Eq. (54) as g_0^2, and $K_3^+(\omega)$ as $K_1^+(\omega)$, where we have

$$K_1^+(\omega) = \omega \mathcal{G}_1^+(\omega). \tag{57}$$

When $\omega \to -\infty$ the functions $h_3^+(\omega)$ and $\mathcal{G}_3^+(\omega)$ tend to two related but different constants. Anyway, we do not use these representations.

IV. $N\theta$ SCATTERING

Consider a state consisting of an incident wave $N^+ a_{p_0}^+ |0\rangle$ plus additional terms which in the asymptotic region reduce to outgoing waves only. This state is designated by $|N\theta_{p_0}^{\text{in}}\rangle$. The energy of this state is ω_0, the incident energy of the θ particle since $m_N = 0$. The state vector in question can be expanded,

$$|N\theta_{p_0}^{\text{in}}\rangle = a_{p_0}^+ N^+ |0\rangle + \sum_{i=1}^{3} \beta_i(\mathbf{p}_0)\, V_i^+ |0\rangle + \int \alpha(\mathbf{p}_0, \mathbf{p}') a_{p'}^+ N^+ |0\rangle\, d^3p'. \tag{58}$$

From the Schrödinger equation,

$$H|N\theta_{p_0}^{\text{in}}\rangle = \omega_0 |N\theta_{p_0}^{\text{in}}\rangle, \tag{59}$$

one finds that

$$\beta_1(\mathbf{p}_0) = g_1(\omega_0 - m_2)(\omega_0 - m_3) f(\omega_0)[H_3^+(\omega_0)]^{-1}, \tag{60}$$

$$\beta_2(\mathbf{p}_0) = g_2(\omega_0 - m_3)(\omega_0 - m_1) f(\omega_0)[H_3^+(\omega_0)]^{-1}, \tag{61}$$

$$\beta_3(\mathbf{p}_0) = g_3(\omega_0 - m_1)(\omega_0 - m_2) f(\omega_0)[H_3^+(\omega_0)]^{-1}, \tag{62}$$

and

$$\alpha(\mathbf{p}, \mathbf{p}_0) = \frac{f(\omega_0) f(\omega_p)}{(\omega_0 - \omega_p + i\epsilon)}[G_3^+(\omega_0)]^{-1}. \tag{63}$$

The $N\theta$ scattering amplitude is defined by

$$S_{N\theta} = \langle N\theta_p^{\text{out}} | N\theta_{p_0}^{\text{in}}\rangle = \delta(\mathbf{p} - \mathbf{p}_0) + 2\pi i \delta(\omega - \omega_0) T_1(\omega_0). \tag{64}$$

The T matrix in the $N\theta$ sector is given by

$$T_1(\omega_0) = -f^2(\omega_0)[G_3^+(\omega_0)]^{-1}. \tag{65}$$

From Eq. (65) it is evident that $T_1(\omega_0)$ will be zero whenever the incident energy ω_0 has a value near or around a_1 or a_2. So in principle, from the $N\theta$ cross section and from the values of a_1, a_2, B_1, and B_2, we can determine the masses of the two resonances.

V. THE GKP EQUATION

For the scattering of a θ particle on a V particle we want an eigenstate of the total Hamiltonian of the form

$$|V\theta_{p_0}^{\text{in}}\rangle = a_{p_0}^+ |\mathbf{V}\rangle + |\chi^+\rangle, \tag{66}$$

with

$$H|V\theta_{p_0}^{\text{in}}\rangle = \omega_0 |V\theta_{p_0}^{\text{in}}\rangle. \tag{67}$$

The state $|\mathbf{V}\rangle$ denotes the stationary eigenstate of the total Hamiltonian describing a physical V particle of mass zero and

$$|\chi^+\rangle = \sum_{i=1}^{3} \int \psi_i(\mathbf{p}', \mathbf{p}_0) V_i^+ a_{p'}^+ |0\rangle\, d^3p' + \int \psi_4(\mathbf{p}', \mathbf{p}'', \mathbf{p}_0) |N\theta_{p'}, \theta_{p''}\rangle\, d^3p' d^3p'', \tag{68}$$

with only outgoing waves in ψ_1, ψ_2, ψ_3, and ψ_4.

The Schrödinger Eq. (67) yields

$$(\omega_0 - \omega - m_i)\psi_i(\mathbf{p}, \mathbf{p}_0) = -\frac{g g_i f(\omega_0) f(\omega)}{\omega} + 2g_i \int f(\omega')\psi_4(\mathbf{p}', \mathbf{p}, \mathbf{p}_0)\, d^3p' \tag{69}$$

and

$$2(\omega_0 - \omega' - \omega)\psi_4(\mathbf{p}', \mathbf{p}, \mathbf{p}_0) = \sum_{i=1}^{3} g_i[f(\omega)\psi_i(\mathbf{p}', \mathbf{p}_0) + f(\omega')\psi_i(\mathbf{p}, \mathbf{p}_0)]. \tag{70}$$

In Eq. (69) taking $i = 1-3$ we in fact obtain three equations for ψ_1, ψ_2, and ψ_3. We readily find from Eq. (69) that

$$\psi_2(\mathbf{p}, \mathbf{p}_0) = \frac{g_2}{g_1}\frac{(\omega_0 - \omega - m_1)}{(\omega_0 - \omega - m_2)}\psi_1(\mathbf{p}, \mathbf{p}_0) \tag{71}$$

and

$$\psi_3(\mathbf{p}, \mathbf{p}_0) = \frac{g_3}{g_1}\frac{(\omega_0 - \omega - m_1)}{(\omega_0 - \omega - m_3)}\psi_1(\mathbf{p}, \mathbf{p}_0). \tag{72}$$

Inserting Eqs. (71) and (72) into Eq. (70), we obtain

$$2(\omega_0 - \omega' - \omega + i\epsilon)\psi_4(\mathbf{p}', \mathbf{p}, \mathbf{p}_0) = f(\omega)\sigma_3(\omega_0 - \omega')\psi_1(\mathbf{p}', \mathbf{p}_0) + f(\omega')\sigma_3(\omega_0 - \omega)\psi_1(\mathbf{p}, \mathbf{p}_0). \tag{73}$$

In Eq. (72) the symbol σ_3 is given by the defining Eq. (12). Inserting Eq. (73) into Eq. (69) for ψ_1, we easily find that

$$\frac{H_3^+(\omega_0 - \omega)}{(\omega_0 - \omega - m_2)(\omega_0 - \omega - m_3)}\psi_1(\mathbf{p}, \mathbf{p}_0) = -\frac{g g_1 f(\omega_0) f(\omega)}{\omega} - f(\omega)\int \frac{f(\omega')\sigma_3(\omega_0 - \omega')\psi_1(\mathbf{p}', \mathbf{p}_0)\, d^3\mathbf{p}'}{(\omega' - \omega_0 + \omega)(\omega_0 - \omega' - m_2)(\omega_0 - \omega' - m_3)}. \tag{74}$$

Let

$$\psi_1(p, p_0) = \frac{g g_1 f(\omega_0) f(\omega)(\omega_0 - \omega - m_2)(\omega_0 - \omega - m_3)}{\sigma_3(\omega_0 - \omega)} \phi(\mathbf{p}, \mathbf{p}_0). \tag{75}$$

Inserting Eq. (75) into Eq. (73), we obtain

$$G_3^+(\omega_0 - \omega)\phi(\mathbf{p}, \mathbf{p}_0) = -\frac{1}{\omega} - \int \frac{f^2(\omega')}{(\omega' - \omega_0 + \omega - i\epsilon)} \phi(\mathbf{p}', \mathbf{p}_0)\, d^3\mathbf{p}'. \tag{76}$$

As in the usual Lee model we assume that $\phi(p, p_0)$ is a function of ω only. Doing the angular integral in Eq. (76), we find that

$$G_3^+(\omega_0 - \omega)\phi^-(\omega, \omega_0) = -\frac{1}{\omega} - \frac{1}{\pi}\int_\mu^\infty \frac{[\mathrm{Im} G_3^+(\omega')]}{(\omega' - \omega_0 + \omega - i\epsilon)} \phi^-(\omega', \omega_0). \tag{77}$$

Let

$$\phi(\omega, \omega_0) = -\frac{M(\omega, \omega_0)}{\omega(\omega_0 - \omega)}. \tag{78}$$

Inserting Eq. (78) into Eq. (77) yields the generalized Källén–Pauli (GKP) equation in the $3V$-particle Lee model,

$$h_3^+(\omega_0 - \omega)M^-(\omega, \omega_0) = 1 + \frac{\omega}{\pi}\int_\mu^\infty \frac{\mathrm{Im} h_3^+(\omega')}{(\omega' - \omega_0 - i\epsilon)} \frac{M^-(\omega', \omega_0)}{(\omega' - \omega_0 + \omega - i\epsilon)} d\omega'. \tag{79}$$

VI. SOLUTION OF THE INTEGRAL EQUATION

Like the celebrated Källén–Pauli equation, the GKP integral equation given by Eq. (79) is a singular integral equation. The function $h_3^+(\omega_0 - \omega)$ has simple poles at $\omega = \omega_0 - a_1$ and $\omega = \omega_0 - a_2$. By dividing Eq. (79) throughout by $h_3^+(\omega_0 - \omega)$, we find that

$$M^-(\omega, \omega_0) = \frac{1}{h_3^+(\omega_0 - \omega)}\left(1 + \frac{\omega}{\pi}\int_\mu^\infty \frac{\mathrm{Im} h_3^+(\omega')}{(\omega' - \omega_0 - i\epsilon)} \times \frac{M^-(\omega', \omega_0)}{\omega' - \omega_0 + \omega - i\epsilon} d\omega'\right). \tag{80}$$

Since the function $[h_3^+(\omega_0 - \omega)]^{-1}$ has zeroes at $\omega = \omega_0 - a_1$ and $\omega = \omega_0 - a_2$ [see Eq. (53)], we observe that

$$M^-(\omega, \omega_0) = 0 \quad \text{for} \quad \omega = \omega_0 - a_1 \quad \text{or} \quad \omega = \omega_0 - a_2. \tag{81}$$

Moreover, we know that, whenever any two bare interaction constants are switched off the function $h_3^+(\omega_0 - \omega)$ reduces to the function $\mathcal{G}_1^+(\omega_0 - \omega)$ of Eq. (37) with ω replaced by $(\omega_0 - \omega)$ there. That means the GKP Eq. (79) reduces to the ordinary Källén–Pauli (KP) equation whenever any two bare interaction constants, out of the three are switched off. In other words, the solution of Eq. (79) must reduce to the solution of the ordinary KP equation whenever any two bare interaction constants are switched off. Under these conditions condition (81) should simply disappear.

In a similar manner we note that the solution of Eq. (79) must also reduce to the solution of the $2V$-particle KP equation (Ref. 1) whenever any one interaction constant out of the three is set equal to zero.

To solve the integral Eq. (79) we exactly follow the procedure outlined in Ref. (1). To this end, we rewrite Eq. (79) with the help of Eq. (41) as

$$\mathcal{G}_3^+(\omega_0 - \omega)M^-(\omega, \omega_0) = -\frac{(\omega^2 B_1 + \omega B_2)}{(\omega - a_1)(\omega - a_2)} M^-(\omega_0, \omega_0) + 1 + \frac{\omega}{\pi}\int_\mu^\infty \frac{\mathrm{Im} h_3^+(\omega')}{(\omega' - \omega_0 - i\epsilon)} \frac{M^-(\omega', \omega_0)}{(\omega' - \omega_0 + \omega - i\epsilon)} d\omega'. \tag{82}$$

Dividing throughout by $\mathcal{G}_3^+(\omega_0 - \omega)$, we obtain

$$M^-(\omega, \omega_0) = -\frac{(\omega^2 B_1 + \omega B_2)}{(\omega - a_1)(\omega - a_2)} \frac{M^-(\omega, \omega_0)}{\mathcal{G}_3^+(\omega_0 - \omega)} + \frac{1}{\mathcal{G}_3^+(\omega_0 - \omega)}\left[1 + \frac{\omega}{\pi}\int_\mu^\infty \frac{\mathrm{Im} h_3^+(\omega')}{(\omega' - \omega_0 - i\epsilon)} \times \frac{M^-(\omega', \omega_0)}{(\omega' - \omega_0 + \omega - i\epsilon)} d\omega'\right]. \tag{83}$$

We suppose that the first term on the right-hand side is known to us. By following Refs. 1 and 2 closely, the quantity inside the brackets divided by $\mathcal{G}_3^+(\omega_0 - \omega)$ can at once be written down. Thus,

$$\frac{1}{\mathcal{G}_3^+(\omega_0 - \omega)}\left(1 + \frac{\omega}{\pi}\int_\mu^\infty \frac{\mathrm{Im} h_3^+(\omega')}{(\omega' - \omega_0 - i\epsilon)} \frac{M^-(\omega', \omega_0)}{(\omega' - \omega_0 + \omega - i\epsilon)} d\omega'\right) = C_1(\omega_0) + \frac{(\omega_0 - \omega)}{\pi}\int_\mu^\infty \left(\mathrm{Im} \frac{1}{K_3^+(\omega')}\right) \frac{X_3(\omega', \omega_0)}{\omega' - \omega_0 + \omega - i\epsilon} d\omega'. \tag{84}$$

In arriving at Eq. (84) we have used Eq. (44). In Eq. (84) $C_1(\omega_0)$ is a constant function of ω_0. The function $X_3(\omega', \omega_0)$ is still an unknown function. It has a left-hand cut.[2] Inserting Eq. (84) into Eq. (83) and after a little rearrangement, we find that

$$M^-(\omega, \omega_0) = \frac{\mathcal{G}_3^+(\omega_0 - \omega)}{h_3^+(\omega_0 - \omega)}\left(C_1(\omega_0) + \frac{(\omega_0 - \omega)}{\pi} \times \int_\mu^\infty \frac{X_3(\omega', \omega_0)\, d\omega'}{(\omega' - \omega_0 + \omega - i\epsilon)} \mathrm{Im} \frac{1}{K_3^+(\omega')}\right). \tag{85}$$

From Appendix B, we note that

$$X_3(\omega', \omega_0) = C_2(\omega_0) \frac{1}{h_3^+(\omega_0 - \omega')}, \tag{86}$$

where $C_2(\omega_0)$ is a constant function of ω_0. We hereafter suppress the arguments of the constants C_1 and C_2. Inserting Eq. (86) into Eq. (85) we readily find the solution of Eq. (79),

$$M^-(\omega, \omega_0) = \frac{\mathcal{G}_3^+(\omega_0 - \omega)}{h_3^+(\omega_0 - \omega)}\left(C_1 - (\omega_0 - \omega)C_2 I(\omega_0 - \omega)\right), \tag{87}$$

where

$$I(\omega_0 - \omega) = -\frac{1}{\pi}\int_\mu^\infty \frac{d\omega'}{(\omega' - \omega_0 + \omega - i\epsilon)} \frac{1}{h_3^+(\omega_0 - \omega')} \mathrm{Im} \frac{1}{K_3^+(\omega')}. \tag{88}$$

The solution will be complete if we can determine the constants C_1 and C_2. Once the solution is found the constants a_1 and a_2 may be allowed to take values higher than μ.

VII. EVALUATION OF THE CONSTANTS

To evaluate the constants C_1 and C_2 we need two equations connecting them. We evaluate $M(0)$ from Eqs. (79) and (87) and then equate. This yields

$$C_1 \mathcal{G}_0 - C_2 K_0 A_0 = 1, \tag{89}$$

where

$$h_0 = h_3^+(\omega_0), \quad \mathcal{G}_0 = \mathcal{G}_3^+(\omega_0), \quad K_0 = K_3^+(\omega_0), \tag{90}$$

and

$$A_0 = \frac{1}{\pi}\int_\mu^\infty \frac{1}{G_3^+(\omega_0-\omega')}\,\mathrm{Im}\frac{1}{K_3^+(\omega')}\,d\omega'. \tag{91}$$

Inserting the solution [Eq. (87)] into the original integral Eq. (78), we obtain

$$\begin{aligned}
&\mathcal{G}_3^+(\omega_0-\omega)[C_1-(\omega_0-\omega)C_2 I(\omega_0-\omega)] \\
&= 1 + \frac{\omega C_1}{\pi}\int_\mu^\infty \frac{\mathrm{Im}h_3^+(\omega')}{(\omega'-\omega_0-i\epsilon)}\,\frac{d\omega'}{(\omega'-\omega_0+\omega-i\epsilon)}\,\frac{\mathcal{G}_3^+(\omega_0-\omega')}{h_3^+(\omega_0-\omega')} \\
&\quad - \frac{\omega C_2}{\pi}\int_\mu^\infty \frac{\mathrm{Im}h_3^+(\omega')}{(\omega'-\omega_0-i\epsilon)}\,\frac{(\omega_0-\omega')\,d\omega'}{(\omega'-\omega_0+\omega-i\epsilon)} \\
&\quad \times \frac{\mathcal{G}_3^+(\omega_0-\omega')}{h_3^+(\omega_0-\omega')} I(\omega_0-\omega').
\end{aligned} \tag{92}$$

These integrals have been evaluated in Appendix A. Using these results we can write (92) as

$$\begin{aligned}
&\mathcal{G}_3^+(\omega_0-\omega)[C_1-(\omega_0-\omega)C_2 I(\omega_0-\omega)] \\
&= 1 + C_1\left[h_0 - \frac{\mathcal{G}_3^+(\omega)h_3^+(\omega_0-\omega)}{h_3^+(\omega)}\right] \\
&\quad + C_2\left[\frac{K_3^+(\omega)h_3^+(\omega_0-\omega)}{h_3^+(\omega)} I(\omega) + \frac{1}{h_3^+(\omega)} - g^2\right].
\end{aligned} \tag{93}$$

In Eq. (93) we now take $\omega = \omega_0$ and this yields, after some rearrangement,

$$1 + C_1\left[h_0 - \frac{\mathcal{G}_0}{h_0 g^2} - \frac{1}{g^2}\right] + C_2\left[\frac{K_0 A_0}{h_0 g^2} + \frac{1}{h_0} - g^2\right] = 0. \tag{94}$$

It is a simple matter to solve for C_1 and C_2 from Eqs. (89) and (94). We find that

$$\begin{aligned}
C_1 = &\left(K_0 A_0 \mathcal{G}_0 - \frac{K_0 A_0 \mathcal{G}_0}{h_0 g^2} + g^2 \mathcal{G}_0 - \frac{\mathcal{G}_0}{h_0}\right) \Big/ \\
&\mathcal{G}_0\left(\frac{K_0 A_0}{g^2} - K_0 A_0 h_0 + \mathcal{G}_0 g^2 - \frac{\mathcal{G}_0}{h_0}\right),
\end{aligned} \tag{95}$$

$$\begin{aligned}
C_2 = &\left(\mathcal{G}_0 + h_0 - \frac{\mathcal{G}_0}{h_0 g^2} - \frac{1}{g^2}\right) \Big/ \\
&\left(\frac{K_0 A_0}{g^2} - K_0 A_0 h_0 + \mathcal{G}_0 g^2 - \frac{\mathcal{G}_0}{h_0}\right).
\end{aligned} \tag{96}$$

The solution (87) of Eq. (79) satisfies all the conditions explained earlier. Whenever any two interaction constants are switched off, the factor outside the parenthesis in Eq. (87) simply cancels out and the solution reduces to the solution of the oridnary KP equation. The solution found here, and the solution of the $2V$-particle KP equation have the same "form." It is a simple matter to check that the solution presented here reduces to the corresponding solution of the $2V$-particle model. Finally Eq. (87) also satisfies the requirements of Eq. (81).

VIII. THE $V\theta$ S MATRIX

For computing the $V\theta$ scattering matrix we use the following definitions:

$$\begin{aligned}
S_{V\theta} &= \langle V\theta_p^{\mathrm{out}} \mid V\theta_{p_0}^{\mathrm{in}}\rangle \\
&= \delta(\mathbf{p}-\mathbf{p}_0) + 2\pi i\,\delta(\omega-\omega_0)T(\omega_0),
\end{aligned} \tag{97}$$

where

$$T(\omega_0) = \frac{f^2(\omega_0)M(\omega_0,\omega_0)}{\omega_0}. \tag{98}$$

From Eqs. (98) and (87) we find that

$$\begin{aligned}
T(\omega_0) &= \frac{f^2(\omega_0)C_1}{\omega_0} \\
&= \frac{f^2(\omega_0)[K_0 A_0 \mathcal{G}_0 - K_0 A_0\,\mathcal{G}_0/h_0 g^2 + g^2 \mathcal{G}_0 - \mathcal{G}_0/h_0]}{K_0[K_0 A_0/g^2 - K_0 A_0 h_0 + g^2 \mathcal{G}_0 - \mathcal{G}_0/h_0]}.
\end{aligned} \tag{99}$$

In Eq. (99) $f(\omega_0)$ is as defined in Eq. (2). From Eq. (99) we note that $T(\omega_0)=0$ for $\omega_0 = a_1$ or $\omega_0 = a_2$. This follows from the fact that

$$C_1 = 0 \quad \text{for} \quad \omega_0 = a_1 \quad \text{or} \quad \omega_0 = a_2. \tag{100}$$

This means that in the $V\theta$ cross section also we find the resonances, at the same energies at which they appear in the $N\theta$ cross section.

IX. THE PRODUCTION AMPLITUDE

The production amplitude $V\theta_0 - N\theta_1\theta_2$ can be easily computed from $\psi_4(\omega_1,\omega_2,\omega_0)$ of Eq. (73). Inserting the solution of ψ_1 into Eq. (73), we find that

$$\begin{aligned}
&\psi_4(\omega_1,\omega_2,\omega_0) \\
&= \frac{-gf(\omega_0)f(\omega_1)f(\omega_2)}{2(\omega_0-\omega_1-\omega_2+i\epsilon)}\left[\frac{M^-(\omega_1,\omega_0)}{\omega_1(\omega_0-\omega_1)} + \frac{M^-(\omega_2,\omega_0)}{\omega_2(\omega_0-\omega_2)}\right].
\end{aligned} \tag{101}$$

The production amplitude P defined by,

$$\begin{aligned}
S_{\mathrm{Prod}} &= \langle N\theta_1\theta_2^{\mathrm{out}} \mid V\theta_0^{\mathrm{in}}\rangle \\
&= 2\pi i\,\delta(\omega_1+\omega_2-\omega_0)P(\omega_1,\omega_0)
\end{aligned} \tag{102}$$

is given by the term containing $\delta(\omega_1+\omega_2-\omega_0)$ in $\psi_4(\omega_1,\omega_2,\omega_0)$ of Eq. (101). Substituting $\omega_2 = \omega_0 - \omega_1$ we find that

$$P = \frac{gf(\omega_1)f(\omega_0-\omega_1)f(\omega_0)}{\omega_1(\omega_0-\omega_1)}[M^-(\omega_1) + M^-(\omega_0-\omega_1)]. \tag{103}$$

The quantity inside the brackets in Eq. (103) has been evaluated in Appendix A. Inserting (A12) into Eq. (103), we find that

$$\begin{aligned}
&P(\omega_1,\omega_0) \\
&= gf(\omega_1)f(\omega_0-\omega_1)f(\omega_0)\left[\frac{C_2}{G_3^+(\omega_1)G_3^+(\omega_0-\omega_1)}\right. \\
&\quad \left. + \frac{(1+C_1 h_0 - C_2 g^2)}{\omega_1 G_3^+(\omega_0-\omega_1)}\right].
\end{aligned} \tag{104}$$

From Eq. (104) we note that the production amplitude goes to zero whenever

$$\omega_0 - a_1 = \omega_1 \quad \text{or} \quad \omega_0 - a_2 = \omega_1. \tag{105}$$

On the other hand, the first term in Eq. (104) goes to zero whenever

$$\omega_1 = a_1 \quad \text{or} \quad \omega_1 = a_2. \tag{106}$$

X. DISCUSSION

In the $3V$-particle model, as in the $2V$-particle model, the resonances that appear in the $N\theta$ sector, also appear in the $V\theta$ sector. In the light of the present solution, the Pierels[8] mechanism should be re-examined. The solution of the GKP equation here has the same form as the solution of the GKP equation of the $2V$-particle model. The ordinary deductive method usually employed to solve the KP equation is not applicable here. Hence a more general deductive method should be invented. It will be really a challenging task to prove the unitarity in this model. An algebraic method due to Bolsterli[9] may be employed to solve the integral equation here. This might help in checking our result, although we do not have any doubt about our solution.

APPENDIX A

Inserting the solution [Eq. (87)] into the integral Eq. (79), we obtain

$$\begin{aligned}
&\mathcal{G}_3(\omega_0-\omega)[C_1-(\omega_0-\omega)C_2 I(\omega_0-\omega)] \\
&\quad = 1+\frac{\omega C_1}{\pi}\int_\mu^\infty \frac{\mathrm{Im}h_3^+(\omega')}{(\omega'-\omega_0-i\epsilon)}\,\frac{d\omega'}{\omega'-\omega_0+\omega-i\epsilon}\,\frac{\mathcal{G}_3^+(\omega_0-\omega')}{h_3^+(\omega_0-\omega')} \\
&\qquad -\frac{\omega C_2}{\pi}\int_\mu^\infty \frac{\mathrm{Im}h_3^+(\omega')}{(\omega'-\omega_0-i\epsilon)}\,\frac{(\omega_0-\omega')\,d\omega'}{(\omega'-\omega_0+\omega-i\epsilon)} \\
&\qquad \times\frac{\mathcal{G}_3^+(\omega_0-\omega')}{h_3^+(\omega_0-\omega')} I(\omega_0-\omega').
\end{aligned} \tag{A1}$$

The integrals in Eq. (A1) can be performed easily by contour integration. The contour is an infinite circle with a cut from μ to ∞. Since the function $h_3(z)$ has poles at $Z=a_1$ and $Z=a_2$, these poles should be deleted from the contour of the integration by cross cuts from the infinite circle. If this is not done, the constants C_1 and C_2 will attain an indeterminate form when $\omega_0=a_1$ or $\omega_0=a_2$. We have found this by explicit calculation.

The first integral in (A1) can be written as

$$\frac{\omega C_1}{2\pi i}\int_C \frac{h_3(z)}{(z-\omega_0)}\,\frac{dz}{(z-\omega_0+\omega)}\,\frac{\mathcal{G}_3(\omega_0-z)}{h_3(\omega_0-z)}, \tag{A2}$$

which has simple poles at $z=\omega_0$ and $z=\omega_0-\omega$. So by the residue theorem, we have

$$\begin{aligned}
&\frac{\omega C_1}{2\pi i}\int_C \frac{h_3(z)}{(z-\omega_0)}\,\frac{dz}{(z-\omega_0+\omega)}\,\frac{\mathcal{G}_3(\omega_0-z)}{h_3(\omega_0-z)} \\
&\quad = \omega C_1\left[\frac{h_3^+(\omega_0)}{\omega}-\frac{\mathcal{G}_3^+(\omega)}{h_3^+(\omega)}\,\frac{h_3^+(\omega_0-\omega)}{\omega}\right].
\end{aligned} \tag{A3}$$

Inserting the definition of $I(\omega_0-\omega')$ into the second integral of (A1), interchanging the orders of integration, and after cancelling the factor $(\omega_0-\omega')$, we obtain the following integral:

$$\begin{aligned}
&-\frac{\omega C_2}{\pi}\int_\mu^\infty \mathrm{Im}\frac{1}{K_3^+(\omega'')}\,\frac{d\omega''}{h_3^+(\omega_0-\omega'')} \\
&\quad\times\frac{1}{\pi}\int_\mu^\infty \frac{\mathrm{Im}h_3^+(\omega')}{(\omega'-\omega_0+\omega-i\epsilon)}\,\frac{d\omega'}{\omega'-\omega_0+\omega''-i\epsilon}\,\frac{\mathcal{G}_3^+(\omega_0-\omega')}{h_3^+(\omega_0-\omega')}.
\end{aligned} \tag{A4}$$

The very last integral can be performed by residue method. So the residue method yields

$$\begin{aligned}
&\frac{1}{\pi}\int_\mu^\infty \frac{\mathrm{Im}h_3^+(\omega')}{(\omega'-\omega_0+\omega-i\epsilon)}\,\frac{d\omega'}{(\omega'-\omega_0+\omega-i\epsilon)}\,\frac{\mathcal{G}_3^+(\omega_0-\omega')}{h_3^+(\omega_0-\omega')} \\
&\quad = \frac{h_3^+(\omega_0-\omega)\mathcal{G}_3^+(\omega)}{h_3^+(\omega)(\omega''-\omega)}-\frac{h_3^+(\omega_0-\omega'')\mathcal{G}_3^+(\omega'')}{h_3^+(\omega'')(\omega''-\omega)}.
\end{aligned} \tag{A5}$$

Inserting (A5) into (A4), we obtain

$$\begin{aligned}
&C_2\frac{\omega\mathcal{G}_3^+(\omega)}{h_3^+(\omega)} I(\omega)h_3^+(\omega_0-\omega) \\
&\quad +\omega C_2\frac{1}{\pi}\int_\mu^\infty \mathrm{Im}\frac{1}{K_3^+(\omega')}\,\frac{d\omega'}{(\omega'-\omega-i\epsilon)}\,\frac{\mathcal{G}_3^+(\omega')}{h_3^+(\omega')},
\end{aligned} \tag{A6}$$

where

$$I(\omega) = -\frac{1}{\pi}\int_\mu^\infty \frac{1}{(\omega'-\omega-i\epsilon)}\,\frac{1}{h_3^+(\omega_0-\omega')}\,\mathrm{Im}\frac{1}{K_3^+(\omega')}\,d\omega'. \tag{A7}$$

The integral in (A6) can again be computed by the residue method. It has poles at $\omega'=0$ and at $\omega'=\omega$, so we have

$$\begin{aligned}
&\omega C_2\frac{1}{\pi}\int_\mu^\infty \mathrm{Im}\frac{1}{K_3^+(\omega')}\,\frac{d\omega'}{(\omega'-\omega-i\epsilon)}\,\frac{\mathcal{G}_3^+(\omega')}{h_3^+(\omega')} \\
&\quad = C_2\left[-\frac{1}{\mathcal{G}_3(0)}\,\frac{\mathcal{G}_3(0)}{h_3(0)}+\frac{1}{h_3^+(\omega)}\right].
\end{aligned} \tag{A8}$$

Inserting (A8) for the integral into (A6), and inserting the resulting expression of (A6) and (A3) into (A1), we find that

$$\begin{aligned}
&\mathcal{G}_3^+(\omega_0-\omega)[C_1-(\omega_0-\omega)C_2 I(\omega_0-\omega)] \\
&\quad = 1+C_1\left[h_0-\frac{h_3^+(\omega_0-\omega)\mathcal{G}_3^+(\omega)}{h_3^+(\omega)}\right] \\
&\qquad +C_2\left[\frac{K_3^+(\omega)h_3^+(\omega_0-\omega)}{h_3^+(\omega)}I(\omega)+\frac{1}{h_3^+(\omega)}-g^2\right]
\end{aligned} \tag{A9}$$

which is the desired result.

From the solution (87), we find that

$$\begin{aligned}
&M^-(\omega_1)+M^-(\omega_0-\omega_1) \\
&\quad = \frac{C_1}{h_3^+(\omega_0-\omega_1)h_3^+(\omega_1)}\left[\mathcal{G}_3^+(\omega_0-\omega_1)h_3^+(\omega_1)+\mathcal{G}_3^+(\omega_1)h_3^+(\omega_0-\omega_1)\right] \\
&\qquad -\frac{C_2}{h_3^+(\omega_0-\omega_1)h_3^+(\omega_1)}\left[K_3^+(\omega_0-\omega_1)I(\omega_0-\omega_1)h_3^+(\omega_1)\right. \\
&\qquad \left.+h_3^+(\omega_0-\omega_1)I(\omega_1)K_3^+(\omega_1)\right].
\end{aligned} \tag{A10}$$

On the other hand, from (A9) we find that (by putting $\omega=\omega_1$)

$$\begin{aligned}
&C_2+h_3^+(\omega_1)[1+C_1h_0-C_2g^2] \\
&\quad = C_1[\mathcal{G}_3^+(\omega_0-\omega)h_3^+(\omega_1)+h_3^+(\omega_0-\omega_1)\mathcal{G}_3^+(\omega_1)] \\
&\qquad -C_2[h_3^+(\omega_1)K_3^+(\omega_0-\omega_1)I(\omega_0-\omega_1) \\
&\qquad +K_3^+(\omega_1)h_3^+(\omega_0-\omega_1)I(\omega_1)].
\end{aligned} \tag{A11}$$

Inserting (A11) into (A10), we find that

$$M^-(\omega_1)+M^-(\omega_0-\omega_1)$$
$$=\frac{C_2}{h_3^+(\omega_0-\omega_1)h_3^+(\omega_1)}+\frac{(1+C_1h_0-C_2g^2)}{h_3^+(\omega_0-\omega_1)} \tag{A12}$$

which is the desired result.

APPENDIX B

In this appendix we propose an ansatz which will enable us to determine the function $X_3(z,\omega_0)$ of the text.

The solution of the ordinary Källén—Pauli equation is given by

$$M_1(z,\omega_0)=D\mathcal{G}_1(\omega_0-z)\times\left(D_1+(\omega_0-z)\frac{1}{\pi}\int_\mu^\infty\frac{X_1(\omega',\omega_0)}{(\omega'-\omega_0+z)}\,\mathrm{Im}\frac{1}{K_1^+(\omega')}\,d\omega'\right), \tag{B1}$$

where

$$D=\frac{1}{\mathcal{G}_1(\omega_0-z)}. \tag{B2}$$

An expression for the constant D_1 is given in Ref. 1. The function $\mathcal{G}_1(\omega_0-z)$ is the twice subtracted function of the ordinary Lee model. The function $X_1(z,\omega_0)$ can be found out by a deductive method.[2] We have

$$X_1(z,\omega_0)=\frac{D_2}{\mathcal{G}_1(\omega_0-z)}, \tag{B3}$$

where D_2 is a constant. The factor outside the parenthesis in Eq. (B1) is unity. But by comparing Eq. (B2) and Eq. (B3) we immediately find an interesting relation.

That is,

$$D\propto X_1(z,\omega_0). \tag{B4}$$

This ansatz appears to be true even in the nV-particle Lee model. So by a clear cut derivation, if we know a factor similar to D, we can always infer a factor like $X_1(z,\omega_0)$.

For example in the $2V$-particle model we quite generally find that,[1]

$$M_2(z)=\frac{\mathcal{G}_2(\omega_0-z)}{h(\omega_0-z)}\left(C_1'+(\omega_0-z)\frac{1}{\pi}\times\int_\mu^\infty\frac{X_2(\omega',\omega_0)}{(\omega'-\omega_0+z)}\,\mathrm{Im}\frac{1}{K_2^+(\omega')}\,d\omega'\right), \tag{B5}$$

where

$$h(z)=\frac{1}{\gamma^2}\left(1-\frac{\gamma^2zB}{\omega_c-z}+\frac{\gamma^2z}{4\pi^2}\int_\mu^\infty\frac{p'u^2(\omega')}{\omega'^2(\omega'-z)}\,d\omega'\right), \tag{B6}$$

$$\mathcal{G}_2(z)=\frac{1}{\gamma^2}\left(1+\frac{\gamma^2z}{4\pi^2}\int_\mu^\infty\frac{p'u^2(\omega')}{\omega'^2(\omega'-z)}\,d\omega'\right), \tag{B7}$$

and

$$K_2(z)=z\mathcal{G}_2(z). \tag{B8}$$

Here C_1' is a constant. Moreover the constants B and ω_c are given in Ref. 1. The renormalized charge γ^2 is given by

$$\gamma^2=(\alpha_1g_1+\alpha_2g_2)^2. \tag{B9}$$

Here the α's are the expansion coefficients of the physical $|\mathbf{V}\rangle$ state in the $2V$-particle model.

From the ansatz proposed above we easily infer that

$$X_2(z,\omega_0)=\frac{C_2'}{h(\omega_0-z)}, \tag{B10}$$

where C_2' is a constant. From the factor outside the large parentheses of Eq. (85) we readily conclude that

$$X_3(z,\omega_0)=\frac{C_2}{h_3(\omega_0-z)}. \tag{B11}$$

In general, for any nV-particle Lee model we can also infer the function $X_n(z,\omega_0)$ from the ansatz proposed here.

APPENDIX C

In this appendix we show that whenever any two interaction constants out of the three are switched off, relation (53) is an identity. As previously explained, the function $h_3(z)$ becomes identical to $\mathcal{G}_1(z)$ of Eq. (37) whenever any two interaction constants are switched off. Let us assume that $g_2^2=0=g_3^2$. Under these conditions, we have

$$a_1=m_2 \quad\text{and}\quad a_2=m_3. \tag{C1}$$

Moreover when any two interaction constants are zero, from Eqs. (48) and (49) we note that

$$G_d^+(\omega')=(\omega'-m_2)(\omega'-m_3)K_1^+(\omega'), \tag{C2}$$

where

$$K_1^+(\omega')=\omega'\mathcal{G}_1^+(\omega'). \tag{C3}$$

So Eq. (53) now reads

$$\frac{1}{\mathcal{G}_1(z)}=(z-m_2)(z-m_3)\left[\frac{g^2}{m_2m_3}+\frac{z}{\pi}\int_\mu^\infty\left(\mathrm{Im}\frac{1}{(\omega'-m_2)(\omega'-m_3)}\frac{1}{K_1^+(\omega')}\right)\frac{d\omega'}{(\omega'-z)}\right], \tag{C4}$$

where we have written z for $(\omega+i\epsilon)$ of Eq. (53). The integral in (C4) should be treated as a contour integral, the contour being an infinite circle with a cut from μ to ∞. We have

$$\frac{z}{\pi}\int_\mu^\infty\left(\mathrm{Im}\frac{1}{(\omega'-m_2)(\omega'-m_3)K_1^+(\omega')}\right)\frac{d\omega'}{(\omega'-z)}=\frac{z}{2\pi i}\int_C\frac{1}{(\omega'-m_2)(\omega'-m_3)}\frac{1}{K_1^+(\omega')}\frac{d\omega'}{(\omega'-z)}. \tag{C5}$$

The integrand in (C5) has poles at $\omega'=0$, since $K_1(0)=0$, and at $\omega'=z$. There are no poles at $\omega'=m_2$ and at $\omega'=m_3$, since these are the real masses of the now noninteracting V particles. Taking residues at these poles we obtain

$$\frac{z}{2\pi i}\int_c \frac{1}{(\omega'-m_2)(\omega'-m_3)}\frac{1}{K_1^+(\omega')}\frac{d\omega'}{(\omega'-z)}$$

$$= z\left(\frac{-g^2}{m_2 m_3 z}+\frac{1}{(z-m_2)(z-m_3)K_1(z)}\right). \tag{C6}$$

If we insert (C6) into (C4) we find that relation (C4) is an identity. In Ref. 1, below Eq. (47), we stated that although $h(z, B)$ reduces to $\mathcal{G}(z)$ when $B=0$, the inverse of $h(z, B)$ [given by Eq. (47), Ref. 1] does not reduce to the inverse of $\mathcal{G}(z)$. The statement is unfounded in view of the results of this appendix.

[1]Cvavb. Chandra Raju, J. Math. Phys. **17**, 6 (1976).
[2]A. Pagnamenta, J. Math. Phys. **6**, 955 (1965).
[3]A. Pagnamenta, J. Math. Phys. **7**, 356 (1966).
[4]E.M. Kazes, J. Math. Phys. **6**, 1772 (1965).
[5]C. Sommerfield, J. Math. Phys. **6**, 1170 (1965).
[6]T.D. Lee and G.C. Wick, Nucl. Phys. B **9**, 209 (1969).
[7]T.D. Lee and G.C. Wick, Nucl. Phys. B **10**, 1 (1969).
[8]R.F. Peierls, Phys. Rev. Lett. **6**, 641 (1961).
[9]M. Bolsterli, Phys. Rev. **166**, 1760 (1968).

Fermi–Bose and internal symmetries with universal Clifford algebras

Geoffrey Dixon

Department of Physics, Brandeis University, Waltham, Massachusetts 02154
(Received 21 November 1976)

Supersymmetry in four-dimensional space–time is approached from the theory of universal Clifford algebras. The representation produced results in a class of new algebras characterized by an index n which indicates the natural appearance of $su(2n)\times u(1)$ as a subalgebra interacting only with the spinor (or odd) parts of the surrounding Fermi–Bose algebra. Finally, a reformulation of the Dirac equation in this formalism is presented which is not plagued with the empirically untenable problem of a continuous range of eigenvalues.

A Clifford algebra for an n-dimensional real orthogonal space X (n finite), as defined by Porteous,[1] is a real associative algebra A with unity I containing isomorphic copies of R and X as linear subspaces such that for all x in X the algebra product $x^2 = -(x, x)I$, where $(\ ,\ )$ is the scalar product of the space X. Furthermore, A is generated as a real algebra by I and X. The universal Clifford algebra of X is that Clifford algebra which satisfies a certain universal mapping property, but it can also be characterized as the unique Clifford algebra of X with dimension 2^n.

Let $R^{p,q}$ be the $(p+q)$-dimensional real linear space with scalar product $(a, b) = -\sum_{i\in p} a_i b_i + \sum_{j\in q} a_{p+j} b_{p+j}$, and let $R_{p,q}$ be the universal Clifford algebra of $R^{p,q}$. If $n = p+q$ is even, then $R_{p,q}$ can be shown to be isomorphic to either $R(2^{n/2})$ or $H(2^{n/2-1})$, where $R(m)$ and $H(m)$ are the real algebras of $m\times m$ real and quaternion matrices, respectively. [Since the basic quaternions, denote them i, j, and k, can be represented as 2×2 complex matrices, i.e., as elements of $C(2)$, an $R_{p,q}$ isomorphic to $H(2^{n/2-1})$ could also be represented by elements of $C(2^{n/2})$, but since this algebra is 2^{n+1}-dimensional it is not isomorphic to $R_{p,q}$. The author has found the restriction to the canonical representations enlightening and these will be used throughout.]

$R_{1,3}$, the universal Clifford algebra of the Minkowski space $R^{1,3}$, is isomorphic to $H(2)$. One set of four basis matrices for $R^{1,3}$ in $R_{1,3}$ is $\gamma_0 = \left[\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right]$, $\gamma_1 = iC$, $\gamma_2 = jC$, $\gamma_3 = kC$, where $C = \left[\begin{smallmatrix}0 & 1\\ 1 & 0\end{smallmatrix}\right]$. (Basis matrices must anticommute.) The general element x in $R^{1,3}$ takes the form

$$x = \begin{bmatrix} x_0 & \mathbf{x} \\ \mathbf{x} & -x_0 \end{bmatrix}, \tag{1}$$

where $\mathbf{x} = x_1 i + x_2 j + x_3 k$. (Note that $\mathbf{xy} = \mathbf{x}\times\mathbf{x} - \mathbf{x}\cdot\mathbf{y}$.) Finally define $\gamma_5 = \gamma_1\gamma_2\gamma_3\gamma_0 = \left[\begin{smallmatrix}0 & 1\\ -1 & 0\end{smallmatrix}\right]$. In any representation $\gamma_5^2 = -1$.

The Clifford algebra of primary interest here is $R_{2,4} \cong H(4)$. A set of six basis matrices for $R^{2,4}$ in $R_{2,4}$ is $\Gamma_\mu = \left[\begin{smallmatrix}\gamma_\mu & 0\\ 0 & \gamma_\mu\end{smallmatrix}\right]$, $\mu = 0,1,2,3$, $\Delta_1 = \left[\begin{smallmatrix}0 & -\gamma_5\\ \gamma_5 & 0\end{smallmatrix}\right]$, $\Delta_2 = \left[\begin{smallmatrix}0 & \gamma_5\\ \gamma_5 & 0\end{smallmatrix}\right]$, where the γ's are the $H(2)$ matrices given above. Finally, let $\Delta = \Gamma_1\Gamma_2\Gamma_3\Gamma_0\Delta_1\Delta_2 = \left[\begin{smallmatrix}\gamma_5 & 0\\ 0 & -\gamma_5\end{smallmatrix}\right]$.

The fifteen matrices $\Gamma_\mu\Gamma_\nu$ ($\mu\neq\nu$), $\Gamma_\mu\Delta_\alpha$, and $\Delta_1\Delta_2$, together with Δ generate the Lie algebra $so(2,4)\times u(1) \cong su(2,2)\times u(1)$. This constitutes the even parts of a graded Lie algebra (GLA) first introduced by Wess and Zumino.[2] Denote the nontrivial components of this algebra by L_{-2}, L_{-1}, L_0, L_1, and L_2.

Let x be as in (1) and define a four-dimensional subset of $R^{2,4}$ consisting of elements of the form

$$X = \begin{bmatrix} x & x^2\gamma_5 \\ \gamma_5 & x \end{bmatrix}. \tag{2}$$

A basis for L_0 consists of the six $\Gamma_\mu\Gamma_\nu$ which generate Lorentz transformations on the Minkowski vector x in X, $\Delta_1\Delta_2$ which generates a dilation, and Δ which acts as a kind of phase transformation. A basis for L_2 is the four $\frac{1}{2}(\Gamma_\mu\Delta_2 - \Gamma_\mu\Delta_1)$ which generate translations on x in X, and for L_{-2} the four $\frac{1}{2}(\Gamma_\mu\Delta_2 + \Gamma_\mu\Delta_1)$ which generate space–time dependent dilations. It should be noted that elements of the form (2) constitute a subset of the null cone about the origin in $R^{2,4}$, and that whereas all the above conformal transformations map X into this null cone, only the Lorentz transformations and translations maintain the form of X given in (2).

To complete the GLA, $W \equiv L_{-2}\oplus L_{-1}\oplus L_0\oplus L_1\oplus L_2$, we must incorporate the odd elements. By maintaining the condition that $R_{2,4}$ be a real algebra a variation of the usual complex representation[3] results. In fact the algebra is radically altered and will be presented without the details of its development.

To incorporate the odd elements, the size of the matrices must be increased to $H(6)$ resulting in the even elements taking the forms

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \gamma_5 y & 0 & 0 \end{bmatrix} \text{ in } L_{-2}, \quad \begin{bmatrix} A & 0 & 0 \\ 0 & D & 0 \\ 0 & 0 & -A^* \end{bmatrix} \text{ in } L_0, \quad \begin{bmatrix} 0 & 0 & \gamma_5 x \\ 0 & 0 & \\ 0 & 0 & \end{bmatrix} \text{ in } L_2, \tag{3}$$

where x and y are of the form (1), A^* is the quaternion conjugate of the matrix A which has the form $uI + v\gamma_5$, u and v in H, and the form of D is yet to be determined.

The four basis matrices for L_1 are

$$S_a = 2^{-1/2}\begin{bmatrix} 0 & \gamma_a & 0 \\ 0 & 0 & C\gamma_a \\ 0 & 0 & 0 \end{bmatrix}, \quad a = 0,1,2,3. \tag{4}$$

The parameters of these super transformations are elements of a Grassmann algebra G, but it is necessary

 Copyright © 1977 American Institute of Physics

that these parameters be taken over the complex field with γ_5 representing the basis for the imaginary part. That is, if r_a and s_a are elements of G of odd degree, then the general element of L_1 takes the form

$$\tilde{R}=(r_a+\gamma_5 s_a)S_a=2^{1/2}\begin{bmatrix}0 & r+\gamma_5 s & 0\\ 0 & 0 & C(r-\gamma_5 s)\\ 0 & 0 & 0\end{bmatrix}$$
$$\equiv\begin{bmatrix}0 & R & 0\\ 0 & 0 & \overleftarrow{R}\\ 0 & 0 & 0\end{bmatrix}, \tag{5}$$

where r and s have the form (1).

Let $\tilde{R}'$ be the same as $\tilde{R}$ with primed components. Then

$$[\tilde{R},\tilde{R}']=\begin{bmatrix}0 & 0 & T\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix},$$

where

$$T=R\overleftarrow{R}'-R'\overleftarrow{R}=(r_0\mathbf{r}'+\mathbf{r}r_0')\gamma_0-(r_0r_0'+\mathbf{r}\cdot\mathbf{r}')C$$
$$-(\mathbf{r}\times\mathbf{s}')\gamma_0+(\text{similar terms}),$$

which has the form of an element of L_2 as required of such a commutator.

The four basis matrices of L_{-1} are

$$Q_a=2^{-1/2}\begin{bmatrix}0 & 0 & 0\\ \gamma_a & 0 & 0\\ 0 & -\gamma_a C & 0\end{bmatrix},\quad a=0,1,2,3.$$

The general element in this case must be defined through multiplication on the right by the parameters. That is,

$$\tilde{B}=Q_a(b_a-\gamma_5 c_a)=2^{-1/2}\begin{bmatrix}0 & 0 & 0\\ b+\gamma_5 c & 0 & 0\\ 0 & -(b-\gamma_5 c)C & 0\end{bmatrix}$$
$$\equiv\begin{bmatrix}0 & 0 & 0\\ B & 0 & 0\\ 0 & \vec{B} & 0\end{bmatrix}.$$

The commutator $[\tilde{R},\tilde{B}]$ must take the form of an element of L_0. From this it is possible to determine the form of the matrix D defined in (3),

$$[\tilde{R},\tilde{B}]=\begin{bmatrix}RB & 0 & 0\\ 0 & \overleftarrow{R}\vec{B}-BR & 0\\ 0 & 0 & -BR\end{bmatrix}.$$

It is not difficult to show $\vec{B}\overleftarrow{R}=(BR)^*$. The central term is

$$\overleftarrow{R}\vec{B}-BR=-(\mathbf{r}\times\mathbf{b})1+(\mathbf{r}c_0-r_0\mathbf{c})1+(r_0c_0-\mathbf{r}\cdot\mathbf{c})\gamma_5$$
$$+(\text{similar terms}).$$

This has the form of an element of su(2)×u(1) (complicated by Grassmann parameters). In fact, if we allow D in (3) to be an arbitrary matrix of the form $\begin{bmatrix}g & g\\ -g_0 & g\end{bmatrix}$, where the g_a are real or even elements of G, then the algebra remains closed and is now endowed with an extra, internal subalgebra su(2)×u(1) which interacts only with the odd components L_1 and L_{-1}.

This algebra is quite easily generalized. Instead of enlarging the representation of the even components from $H(4)$ to $H(6)$, as was done, enlarge it to $H(2n+4)$ in a similar fashion and define the general element of L_1 to have the form

$$\tilde{R}_n=\begin{bmatrix}0 & R_1 & R_2 & \cdots & R_n & 0\\ 0 & 0 & 0 & \cdots & 0 & \overleftarrow{R}_1\\ 0 & 0 & 0 & \cdots & 0 & \overleftarrow{R}_2\\ \vdots & \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & 0 & \overleftarrow{R}_n\\ 0 & 0 & 0 & \cdots & 0 & 0\end{bmatrix},$$

where R_i and $\overleftarrow{R}_i$ have the forms of R and $\overleftarrow{R}$ in (5). Define an element $\tilde{B}_n$ of the new L_{-1} by a similar generalization. Then

$$[\tilde{R}_n,\tilde{B}_n]=\begin{bmatrix}\sum_{i=1}^{n}R_iB_i & 0 & 0\cdots & 0\\ 0 & & & 0\\ 0 & & [\overleftarrow{R}_k\vec{B}_j-B_kR_j] & \vdots\\ \vdots & & & \\ 0 & 0 & 0\cdots & -\sum_{i=1}^{n}\vec{B}_i\overleftarrow{R}_i\end{bmatrix},$$

where $[\overleftarrow{R}_k\vec{B}_j-B_kR_j]$ is an element of $H(2n)$ whose $H(2)$ components are $\overleftarrow{R}_k\vec{B}_j-B_kR_j$. This central matrix has the form of an element of su(2n)×u(1) and again the algebra is closed if we allow this central matrix to vary independently in L_0. If this is the mathematical origin of internal symmetries it could have far reaching consequences.

To produce a Dirac equation we must adjust the formalism to accomodate the Clifford algebra formalism used here. We should expect phases to be generated by γ_5 and the lowest order spinors to be elements of $H(2)$.

Define $p=\gamma_\mu p^\mu=\begin{bmatrix}E & \mathbf{p}\\ \mathbf{p} & -E\end{bmatrix}$ and $M=p|_{p=0}=m\gamma_0=\begin{bmatrix}m & 0\\ 0 & -m\end{bmatrix}$, where E, $\mathbf{p}$, and m represent the energy, momentum, and rest mass of the particle to be described by the spinor. The variation on Dirac's equation is then

$$P\psi(x)\equiv\gamma_5\hbar\partial\psi(x)=M\psi(x), \tag{6}$$

where $\partial=\begin{bmatrix}\partial_t & -\nabla\\ \nabla & -\partial_t\end{bmatrix}$. A plane wave solution takes the form

$$\psi(x)=\exp(\gamma_5 p_\mu x^\mu)U, \tag{7}$$

where U is in $H(2)$. Substituting this into (6) gives

$$P\psi(x)=p\psi(x)=M\psi(x). \tag{8}$$

To obtain the momentum equation we must cancel the exponential in $\psi(x)$ [which could not be done were we to replace M with m in (8)] obtaining

$$pU=MU. \tag{9}$$

It is not difficult to show that $U=p$ is a solution of (9) unique up to multiplication on the right by an arbitrary element of $H(2)$ and that $m=0$. Furthermore, the helicity of the solution points along the momentum vector. Therefore, the only solution of (6) is a neutrino solution. Were it to have been demanded of the original Dirac equation that it play a more fundamental role in physics than that of a technical tool, then the continuous range of eigenvalues for m allowed by that equation should have been deemed unsatisfactory in light of the fact that nature has supplied us with no more than

countably many particle rest masses. Further, since massive spin-$\frac{1}{2}$ particles invariably attribute their rest masses to self fields, we should not expect to obtain massive solutions for a spin-$\frac{1}{2}$ equation if that equation incorporates none of these extra fields. The original Dirac equation differs from the above primarily in that it treated the Dirac matrices as generators of a complex rather than real Clifford algebra. It is seen here that if the theory is approached from the mathematical end, maintaining the mathematical consistency throughout, then complex analysis arises naturally and the theory takes on the semblance of fundamentality and provides a fresh starting point for the theory of fermions.

[1]I.A. Porteous, *Topological Geometry* (Van Nostrand, London, 1969).
[2]J. Wess and B. Zumino, Nucl. Phys. B 70, 39 (1974).
[3]L. Corwin, Y. Ne'eman, and S. Sternberg, Rev. Mod. Phys. 47, 573 (1975).

A Lie group framework for soliton equations. I. Path independent case

James Corones[a)]

Department of Mathematics, Iowa State University, Ames, Iowa 50011

B. L. Markovski and V. A. Rizov

Institute of Nuclear Research and Nuclear Energy, Sofia, Bulgaria

A general Lie group theoretic framework for the study of a class of nonlinear partial differential equations is presented. In two space-time dimensions this class includes soliton equations. The approach is applicable in $N \geq 2$ space-time dimensions. Eigenvalue problems and isospectral flows associated with equations have a natural group theoretic interpretation in this framework. A sequence of nonlocal exact 1-, 2-,...,$(N-1)$-forms are derived in N-dimensional space-time.

1. INTRODUCTION

This paper presents a general Lie group theoretical framework for a class of nonlinear partial differential equations. Members of this class are the integrability conditions for a certain system of first order differential equations in N-dimensional space–time. The solution of this system gives the parameters of a bilocal Lie group, that is, a Lie group described in terms of bilocal functions of space–time. A consequence of the existence of the bilocal Lie group is the existence of N-local conservation laws in N-dimensional space–time.

In two space–time dimensions the class of partial differential equations treated contains equations that have recently received considerable attention equations with soliton solutions.[1] These equations are known to have associated eigenvalue problems and isospectral flows and an infinite number of conservation laws. In the present approach the differential form of the linear action of the bilocal group on a linear representation space expresses the associated eigenvalue problem and isospectral flow.[1] Other equations of physical interest that have not yet been shown to have soliton properties are also contained in the general scheme. The main geometric feature of the present framework is that the bilocal group action on a linear representation space naturally defines a flat connection. The vanishing of the corresponding curvature tensor is a re-expression of the original nonlinear partial differential equation.

This work was motivated by the study of pseudopotentials. This idea was introduced in Ref. 2 and later studied for example in Refs. 3–8. The geometric picture associated with the vanishing of the curvature tensor extends previous work in Refs. 8 and 9 on particular soliton equations into a coherent group theoretic framework.

In the next section some standard results from Lie group theory are presented for future reference and to fix the notation. Section 3 introduces the central definitions, those of a bilocal Lie group and bilocal group parameters, and the connection with partial differential equations, particularly soliton equations, is made. This is followed in Sec. 4 by the introduction of local group parameters and the expression of bilocal parameters in terms of them. Group actions are then discussed in Sec. 5. It is shown how a representation space for the group is a set of pseudopotentials associated with the equation, and that the differential form of a linear group action is the associated eigenvalue problem and isospectral flow used to solve soliton equations by the inverse scattering method. A flat connection is defined by the linear group action. The final section gives the derivation of N-local conserved currents as well as a sequence of exact $(j+1)$-local j-forms, $j=1,\ldots,N-1$, in N-dimensional space–time. The forms are exact on the equations of motion, that is, on the integrability conditions for the bilocal group action.

2. LIE GROUPS

In this section some basic results from the theory of Lie groups are recounted. The reader is referred to any of the standard classical treatments of Lie theory for their derivation, see Refs. 10 and 11 for example. In the remainder of the work, group parameters of particular types are introduced and discussed, bilocal and local group parameters. The addition of space–time dependence to the group parameters produces a richer structure though all the basic results of Lie theory remain and indeed are often called upon to derive consequences of the space–time dependence. It is useful then to gather together the relevant results from the standard theory in order to delineate the boundary between them and those results that are due to the types of parametrization used and also to fix notation.

Let G be an f-parameter Lie group with parameters given by t^k, $k=1,\ldots,f$. The notation $g=\{t^k\}$ is used when $g \in G$ and the t^k are the corresponding parameters of g. If $g=\{t^k\}$ and g^{-1} is the element inverse to g, the notation $g^{-1}=\{\bar{t}^k\}$ is used. Let e denote the unit element of G. In the sequel, unless otherwise stated, sets of parameters are used that have the properties $e=\{t_e^k=0\}$, $k=1,\ldots,f$; if $g=\{t^k\}$, then

$$g^{-1}=\{\bar{t}^k=-t^k\}. \tag{2.1}$$

If $g_1, g_2 \in G$, $g_1=\{t_1^k\}$, $g_2=\{t_2^k\}$, and $g_1 g_2 = g_3$, then the parameters of g_3 are given by

$$t_3^k = R^k(t_1, t_2). \tag{2.2}$$

a)Supported in part by the National Science foundation.

 Copyright © 1977 American Institute of Physics

The functions R^k are called the *group composition law for G.*

The group composition law has the following properties:

$$R^k(t,\bar{t}) = R^k(\bar{t},t) = 0, \tag{2.3a}$$

$$R^k(t,0) = R^k(0,t) = t, \tag{2.3b}$$

$$R^k(t_1, R(t_2,t_3)) = R^k(R(t_1,t_2),t_3). \tag{2.3c}$$

The R^k are analytic functions of $2f$ arguments. The values of various derivatives of the R^k play a fundamental role in Lie theory; it is useful to introduce special notation for them. Thus define

$$A_l^k(t_1) \equiv \left.\frac{\partial R^k(t_1,t_2)}{\partial t_2^l}\right|_{t_2=0}, \tag{2.4a}$$

$$B_l^k(t_2) \equiv \left.\frac{\partial R^k(t_1,t_2)}{\partial t_1^l}\right|_{t_1=0}. \tag{2.4b}$$

It follows in the general theory of Lie groups that if

$$[A_k, A_l]^m \equiv \frac{\partial A_k^m}{\partial t^n} A_l^n - \frac{\partial A_l^m}{\partial t^n} A_k^n, \tag{2.5}$$

then

$$[A_k, A_l]^m = C_{kl}^n A_n^m, \tag{2.6a}$$

$$[B_k, B_l]^m = -C_{kl}^n B_n^m, \tag{2.6b}$$

where the constants C_{kl}^n are the structure constants of the group G. In (2.5) and throughout the text the summation convention is used. The A_l^k and B_l^k give two realizations of the Lie algebra of G. These realizations are clearly equivalent.

The sets of functions $A_l^k(t)$, $B_l^k(t)$ have the following properties:

$$A_l^k(0) = B_l^k(0) = \delta_l^k, \tag{2.7a}$$

$$A_l^k(t) = B_l^k(-t). \tag{2.7b}$$

It can also be shown that

$$[A_k, B_l]^m = 0. \tag{2.8}$$

Let $A(t)$ and $B(t)$ denote the matrices with matrix elements $A_l^k(t)$ and $B_l^k(t)$, respectively. The matrices A and B are invertible. Let V and W denote the matrices inverse to A and B, respectively. Thus

$$A(t)V(t) = V(t)A(t) = I, \tag{2.9a}$$

$$B(t)W(t) = W(t)B(t) = I, \tag{2.9b}$$

where I is the f-dimensional unit matrix. It follows from (2.7) and (2.9) that

$$V(t) = W(-t). \tag{2.10}$$

Using (2.6), (2.9), and (2.10) it is easily shown that

$$\frac{\partial V_l^k}{\partial t^m} - \frac{\partial V_m^k}{\partial t^l} = C_{np}^k V_m^n V_l^p, \tag{2.11a}$$

$$\frac{\partial W_l^k}{\partial t^m} - \frac{\partial W_m^k}{\partial t^l} = -C_{np}^k W_m^n W_l^p. \tag{2.11b}$$

If (2.3c) is differentiated with respect to t_3^k and the result evaluated at $t_3 = 0$, it is found that

$$\frac{\partial R^k(t_1,t_2)}{\partial t_2^l} A_n^l(t) = A_n^k(R(t_1,t_2)). \tag{2.12a}$$

The derivative of (2.3c) with respect to t_1 at $t_1 = 0$ is

$$\frac{\partial R^k(t_1,t_2)}{\partial t^l} B_n^l(t_1) = B_n^k(R(t_1,t_2)). \tag{2.12b}$$

Using (2.9) it readily follows that

$$V_l^k(R(t_1,t_2)) \frac{\partial R^l(t_1,t_2)}{\partial t_2^n} = V_n^k(t_2), \tag{2.13a}$$

$$W_l^k(R(t_1,t_2)) \frac{\partial R^l(t_1,t_2)}{\partial t_1^n} = W_n^k(t_1). \tag{2.13b}$$

If (2.13) are evaluated at $t_2 = \bar{t}_1 = -t_1$, one obtains

$$V_l^k(t_1) = \left.\frac{\partial R^k(t_1,t_2)}{\partial t_2^l}\right|_{t_1=\bar{t}_2}, \tag{2.14a}$$

$$W_l^k(t_1) = \left.\frac{\partial R^k(t_1,t_2)}{\partial t_1^l}\right|_{t_2=\bar{t}_1}, \tag{2.14b}$$

when (2.7) and (2.10) are used.

This ends the brief recital of standard results from Lie theory. In subsequent sections the structure is enriched by considering group parameters depending on two space–time points.

3. BILOCAL LIE GROUPS

Recent studies[3–5,7,8] have suggested that there is a close connection between Lie algebras, and therefore, between Lie groups and equations that have soliton solutions. These results consist of a series of computations and/or observations based on pseudopotentials.[2] Thus far no coherent basis for the existence of the Lie algebra structure has been developed. This section contains the principal definition needed for the construction of a Lie group framework for soliton equations; bilocal Lie groups. Following the definition the basic properties of bilocal group parameters are derived and the connection with soliton equations is established.

A bilocal Lie group G is given by a map

$$r : \mathbb{R}^N \times \mathbb{R}^N \to G, \tag{3.1}$$

i.e., if $(x,y) \in \mathbb{R}^N \times \mathbb{R}^N$, then the image of (x,y) under r is an element of $g \in G$ with parameters $r^k(x,y)$, $k = 1,\ldots,f$ (for some fixed parametrization of G). The expression "bilocal Lie group" is used to indicate that a Lie group G described in terms of bilocal group parameters is being considered.

We ask the following property for the map (3.1),

$$R^k(r(x,x'), r(x',y)) = r^k(x,y), \quad x,y,x' \in \mathbb{R}^N, \tag{3.2}$$

i.e., the product of two group elements having one common point x' does not depend on that point. It follows from (3.2) evaluated at $x = x'$ and from (2.3b) that $r^k(x,x) = 0$. If (3.2) is evaluated at $x = y$, (2.3a) and (2.1) can be used to show that $r^k(x,y) = -r^k(y,x)$.

It is necessary to derive numerous expressions involving derivatives of (3.2). The notation becomes cumbersome unless some conventions are used. For the purposes of differentiation let v^k, $k = 1,\ldots,f$, denote the first f arguments of R^k and u^k the second f arguments of R^k. The components of $x \in \mathbb{R}^N$ are denoted by x^λ, $\lambda = 0,\ldots,N-1$. In applications x^0 denotes the time variable.

Consider (3.2). Since the right-hand side is independent of x',

$$\frac{\partial R^k(r(x,x'),r(x',y))}{\partial u^l}\frac{\partial r^l(x,x')}{\partial x^\lambda}+\frac{\partial R^k(r(x,x'),r(x',y))}{\partial v^l}\times\frac{\partial r^l(x',y)}{\partial x^\lambda}=0. \tag{3.3}$$

Setting $x=x'$,

$$\frac{\partial R^k(u,r(x,y))}{\partial u^l}\bigg|_{u=0}\frac{\partial r^l(x,x')}{\partial x'^\lambda}\bigg|_{x=x'}+\frac{\partial R^k(u,r(x,y))}{\partial v^l}\bigg|_{u=0}\times\frac{\partial r^l(x',y)}{\partial x'^\lambda}\bigg|_{x'=x}=0, \tag{3.4a}$$

while setting $x'=y$ yields

$$\frac{\partial R^k(r(x,y),v)}{\partial u^l}\bigg|_{v=0}\frac{\partial r^l(x,x')}{\partial x'^\lambda}\bigg|_{x'=y}+\frac{\partial R^k(r(x,y),v)}{\partial v^l}\bigg|_{x=0}\times\frac{\partial r^l(x',y)}{\partial x'^\lambda}\bigg|_{x'=y}=0. \tag{3.4b}$$

These equations are a direct consequence of the bilocal composition law (3.2) and have a fundamental role in the present framework. They can be written in a more convenient form by introducing the notation

$$f^k_\lambda(y)\equiv\frac{\partial r^k(x,y)}{\partial y^\lambda}\bigg|_{x=y}, \tag{3.5a}$$

$$\hat f^k_\lambda(x)\equiv\frac{\partial r^k(x,y)}{\partial x^\lambda}\bigg|_{y=x}. \tag{3.5b}$$

Since $r^k(x,y)=-r^k(y,x)$ it follows from the definition that $\hat f^k_\lambda(x)=-f^k_\lambda(x)$. Using this fact together with the definitions of A and B given by (2.4), the pair (3.4) becomes

$$\frac{\partial r^k(x,y)}{\partial x^\lambda}+B^k_l(r(x,y))f^l_\lambda(x)=0, \tag{3.6}$$

$$\frac{\partial r^k(x,y)}{\partial y^\lambda}-A^k_l(r(x,y))f^l_\lambda(y)=0. \tag{3.7}$$

For a given set of functions $f^k_\lambda(x)$ and $B^k_l(r)$ $[A^k_l(t)]$, satisfying (2.6), the system (3.6) [(3.7)] can be treated as a set of equations for the $r^k(x,y)$. In this situation it is of interest to derive conditions for the existence of solutions of the equations; the integrability conditions. For (3.6), (3.7) these are

$$\frac{\partial^2 r^k(x,y)}{\partial x^\lambda\,\partial x^\mu}=\frac{\partial^2 r^k(x,y)}{\partial x^\mu\,\partial x^\lambda}, \tag{3.8a}$$

$$\frac{\partial^2 r^k(x,y)}{\partial y^\lambda\,\partial y^\mu}=\frac{\partial^2 r^k(x,y)}{\partial y^\mu\,\partial y^\lambda}, \tag{3.8b}$$

$$\frac{\partial^2 r^k(x,y)}{\partial x^\lambda\,\partial y^\mu}=\frac{\partial^2 r^k(x,y)}{\partial y^\mu\,\partial x^\lambda}. \tag{3.8c}$$

The first two sets of conditions give

$$A^k_l(r(x,y))\left\{\frac{\partial f^l_\mu}{\partial y^\lambda}-\frac{\partial f^l_\lambda}{\partial y^\mu}-C^l_{n\rho}f^n_\lambda f^\rho_\mu\right\}=0, \tag{3.9a}$$

$$B^k_l(r(x,y))\left\{\frac{\partial f^l_\mu}{\partial x^\lambda}-\frac{\partial f^l_\lambda}{\partial x^\mu}-C^l_{n\rho}f^n_\lambda f^\rho_\mu\right\}=0, \tag{3.9b}$$

where (2.6) has been used. Since A and B are invertible it follows that

$$\frac{\partial f^l_\mu}{\partial x^\lambda}-\frac{\partial f^l_\lambda}{\partial x^\mu}=C^l_{n\rho}f^n_\lambda f^\rho_\mu. \tag{3.10}$$

The final condition (3.8c) implies

$$f^k_\mu f^l_\lambda[A_l,B_k]^m=0. \tag{3.11}$$

The equality is satisfied due to (2.8). Thus all the restrictions put on f^k_λ so that (3.6), (3.7) is integrable are contained in (3.10). For a given set of f^k_λ, if (3.10) is satisfied the bilocal group parameters $r^k(x,y)$ exist. We note here that if f^k_λ satisfy (3.10) the functions ${}^1f^k_\lambda(y)\equiv E^k_j f^j_\lambda(y)$ have the same property if

$$E^k_l E^m_n C^j_{km}=C^m_{ln}E^j_m, \tag{3.12}$$

i.e., if the constants E^k_j form a matrix belonging to the adjoint representation of G. Therefore, linear transformations of f^k_l through the adjoint representation do not yield an essentially new solution of Eq. (3.10). We shall say that two solutions of (3.10) differing by transformation from the adjoint representation of G belong to the same class.

It is interesting to see how f^k_λ transform under change of parametrization of G. Suppose the new parameters $\tilde r^k(x,y)$, are expressed through the old by the formula

$$\tilde r^k(x,y)=\Psi^k(r^l(x,y))$$

with nonvanishing Jacobian $\|\partial\Psi^k/\partial r^l\|$. Following the definition (3.5) we obtain in the new parametrization

$$\tilde f^k_\lambda(y)=\frac{\partial\tilde r^k(x,y)}{\partial y^\lambda}\bigg|_{x=y}=\tilde E^k_j f^j_\lambda(y),$$

where

$$\tilde E^k_j=\frac{\partial\Psi^k(r)}{\partial r^j}\bigg|_{r=0}.$$

Let the new parameters be chosen so that the structure constants C^l_{nm} remain unchanged. Then rederivation of Eqs. (3.6) and (3.7) shows that $\tilde f^k_\lambda$ have to satisfy (3.10) again and that the constants $\tilde E^k_j$ obey (3.12). From here we conclude that transformations to new parameters which do not affect the structure constants change the functions f^k_λ in a trivial way—the latter are only linearly transformed according to the adjoint representation of G, and, therefore remain in the same class of solutions of (3.10).

This observation shows that in fact the functions are not tied to a particular choice of parameters (once C^k_{nm} are fixed). The independence of the f^k_λ is important in view of the applications to nonlinear partial differential equations, where they will be regarded as primary objects.

If we have a bilocal parameterization in two space–time dimensions, for some fixed group G, then with a suitable choice of f^k which has the form $f^k_\lambda=f^k(u,u_x,u_{xx},u,\cdots)$ (u is some function of x^0 and x^1), Eqs. (3.10) become known partial differential equations for u which possess the soliton solution. For example, in Ref. 5 (in a more restricted and rather *ad hoc* fashion) it was pointed out that all differential equations solvable by the inverse scattering method via the generalized Zakharov–Shabat eigenvalue problem can be cast into the form (3.12). As a particular case consider the identifications

$$f^1_1(x)=\tfrac{1}{2}i(u+ku^*);\quad f^2_1(x)=-\tfrac{1}{2}(u-ku^*),$$

$$f_1^3(x)=\lambda;\quad f_0^1(x)=\tfrac{1}{2}i\{2\lambda(u+ku^*)+i(u_x-ku_x^*)\},$$
$$f_0^2(x)=-\tfrac{1}{2}\{2\lambda(u-ku^*)+i(u_x+ku_x^*)\}; \tag{3.13}$$
$$f_0^3(x)=2\lambda^2+k|u|^2,$$
$$C_{lm}^n=2\epsilon_{nlm};\quad n,l,m=1,2,3.$$

Here u is a complex valued function of $x^0=t$, $x^1=x$; u^* is the complex conjugate of u, u_x denotes the derivative of u with respect to x^1, k and λ are constants. Furthermore the group G is taken to be SU(2) with structure constants $2\epsilon_{nlm}$ (ϵ_{klm} is the completely antisymmetric tensor $\epsilon_{123}=1$). With this choice of f_λ^k, (3.10) yields

$$iu_t+u_{xx}-2k|u|^2u=0, \tag{3.14a}$$

$$iu_t^*-u_{xx}^*+2k|u|^2u^*=0. \tag{3.14b}$$

Thus with the selection (3.14), (3.10) reduces in two space–time dimensions to the nonlinear Schrödinger equation.

Another example of physical interest is a two-dimensional σ model introduced in Ref. 13 and recently studied in Refs. 14 and 15. In Ref. 15 results are derived, in a very different context from the present one, that can easily be used to show that the model treated can be put in the form (3.10), with the C_{lm}^k the structure constants of O(4). This model has not yet been fully investigated.

An alternative interpretation of (3.10) is given by defining the "Yang–Mills" tensor

$$\mathfrak{F}_{\mu\lambda}^l\equiv\frac{\partial f_\mu^l}{\partial x^\lambda}-\frac{\partial f_\lambda^l}{\partial x^\mu}-C_{np}^l f_\lambda^n f_\mu^p, \tag{3.15}$$

through the "vector potentials" f_λ^l. The integrability conditions for the bilocal group action are then seen to be the conditions for the vanishing of the Yang–Mills tensor $\mathfrak{F}_{\mu\lambda}^l$ constructed for the gauge group G.[12,16] This interpretation will be commented on further below.

4. LOCAL LIE GROUPS AND FACTORIZATION

In this section a factorized form of the path independent map (3.1), (3.2) is introduced and its basic properties established.

Let s be the map

$$s:\mathbb{R}^N\to G, \tag{4.1}$$

such that if $x\in\mathbb{R}^N$ then the image of x under s is an element $g\in G$ with parameters $s^k(x)$. We are considering the *same* parametrization of G as in Sec. 3, i.e., with the same set of functions R^k as those in (2.2) for the group composition law.

With a given map s, we can explicitly construct a bilocal map r: $\mathbb{R}^N\times\mathbb{R}^N\to G$ satisfying (3.2). Indeed, defining the functions $r^R(x,y)$ by

$$r^k(x,y)=R^k(s(x),\bar{s}(y)), \tag{4.2}$$

we obtain the desired bilocal parametrization. In other words, we are considering pairs of maps (r,s) such that such that if $g_1=\{s^k(x)\}$, $g_2=\{s^k(y)\}$ the parameters of the product $g_1g_2^{-1}$ are

$$R^k(s(x),\bar{s}(y))=r^k(x,y).$$

If (4.2) is differentiated with respect to y and evaluated at $x=y$, it follows that

$$f_\lambda^k(y)=+V^k(\bar{s}(y))\frac{\partial\bar{s}^l(y)}{\partial y^\lambda}=-W_l^k(s(y))\frac{\partial s^l(y)}{\partial y_\lambda}, \tag{4.3}$$

where (3.5), (2.14), and (2.10) have been used.

As was shown in the previous section, in two-dimensional space–time it is possible to choose f_λ^k in the form $f_\lambda^k=f_\lambda^k(u,u_x,u_{xx},\cdots)$ so that the integrability conditions (3.10) become soliton equations for the function $u(x,t)$. From this point of view the equations (4.3) are a differential system for the $s^k(x)$ with the f_λ^k given and it is convenient to write (4.3) as

$$\frac{\partial s^k(x)}{\partial x^\lambda}=-B_l^k(s(x))f_\lambda^l(x). \tag{4.4}$$

This system for the $s^k(x)$ has a solution if (3.10) and (2.6) are satisfied.

There is a natural superposition principle for the integrability conditions (3.10) that is inherited from the group composition realized in local parameters. The logic of the superposition principle is as follows: Let $f_{1\lambda}^k$ and $f_{2\lambda}^k$ be two fixed solutions of (3.10) and $s_1^k(x)$, $s_2^k(x)$ are the corresponding solutions of (4.4). The functions $s_{1,2}^k(x)$ realize one and the same parametrization of G but with different space–time dependence, therefore the expression

$$s_3(x)=R^k(s_1(x),\bar{s}_2(x)) \tag{4.5}$$

is well defined and can be used to find the corresponding $f_{3\lambda}^k$.

The details of the calculation are the following. Let

$$\frac{\partial s_1^k(x)}{\partial x^\lambda}=-B_l^k(s_i(x))f_{i\lambda}^l(x),\quad i=1,2 \tag{4.6}$$

and suppose (4.5) is satisfied. Then one obtains

$$\begin{aligned}f_{3\lambda}^k(x)&=-W_l^k(s_3(x))\frac{s_3^l(x)}{\partial x^\lambda}\\&=-W_l^k(s_3(x))\left\{\frac{\partial R^l(s_1,\bar{s}_2)}{\partial u^m}\frac{\partial s_1^m}{\partial x^\lambda}+\frac{\partial R^l(s_1,\bar{s}_2)}{\partial v^m}\frac{\partial\bar{s}_2^m}{\partial x^\lambda}\right\}\\&=-W_l^k(s_3)\left\{B_n^l(s_3)W_m^n(s_1)\frac{\partial s_1^m}{\partial x^\lambda}+A_n^l(s_3)V_m^n(\bar{s}_2)\frac{\partial\bar{s}_2^m}{\partial x^\lambda}\right\}\\&=f_{1\lambda}^k(x)-W_l^k(s_3)A_n^l(s_3)f_{2\lambda}^n(x),\end{aligned} \tag{4.7}$$

when (2.10) is used. This equation indicates the possibility of finding new solutions of (3.10) in terms of two known solutions. This will be commented on further in the next section.

5. GROUP ACTION

In the present application of Lie groups to partial differential equations the properties of bilocal Lie groups viewed as continuous groups of transformations are of considerable importance. Considering the action of a group on a linear space defined over two-dimensional space–time establishes the connection of the new formulation presented here with the pseudopotential approach to nonlinear partial differential equations. Indeed it could be argued that the bilocal Lie group framework is the natural setting for the study of pseudopotentials.

More importantly study of the group action provides a group theoretic interpretation of the linear eigenvalue problem and isospectral flow used to solve soliton equations.[1] The basic properties of bilocal Lie groups of transformations are, of course, independent of the dimension of the underlying space–time. In this section the definitions of a global group action (written in bilocal parameters) is given and its consequences developed. In the bilocal case the general definition of group action, in arbitrary parameters, can be written in complete analogy with the usual case.[10,11]

Let Q be a d-dimensional space and let q be a map

$$q : \mathbb{R}^N \to Q. \tag{5.1}$$

The coordinates of $Q(x) \in Q$ are given by $q^a(x)$, $a = 1, \ldots, d$. A global action of G on Q is defined by

$$q^a(x) = F^a(r(x,y); q(y)). \tag{5.2}$$

This action is in general nonlinear in the sense that the F^a are not necessarily linear functions of q^a. For a set of functions F^a to realize a bilocal group action two conditions must be satisfied:

(1) If $g = \{r^k(x,y)\}$ and

$$q^a(x) = F^a(r(x,y); q(y)), \tag{5.3a}$$

then

$$q^a(y) = F^a(\bar{r}(x,y); q(x)). \tag{5.3b}$$

(2) If $g_1 g_2 = g_3$, $g_1 = \{r^k(y,x')\}$, $g_2 = \{r^k(x',x)\}$, and $g_3 = \{r^k(y,x)\}$, then

$$F^a(r(y,x'); F(r(x',x); q(x))) = F^a(r(y,x); q(x)). \tag{5.4}$$

The *generator functions* of the group action (4.1) are defined by

$$X_k^a(q) \equiv - \left. \frac{\partial F^a(r;q)}{\partial r^k} \right|_{r^k = 0}. \tag{5.5}$$

It follows from the general theory of Lie groups[10,11] that if

$$[X_k, X_l]^a \equiv \frac{\partial X_k^a}{\partial q^b} X_l^b - X_k^b \frac{\partial X_l^a}{\partial q^b}, \tag{5.6}$$

then

$$[X_k, X_l]^a = C_{kl}^m X_m^a, \tag{5.7}$$

where, again, the C_{kl}^m are the structure constants of G.

Consider the derivative of (5.2) with respect to x,

$$\frac{\partial q^a(x)}{\partial x^\lambda} = \frac{\partial F^a(r(x,y); q(y))}{\partial r^k} \frac{\partial r^k(x,y)}{\partial x^\lambda}. \tag{5.8}$$

If (5.8) is evaluated at $y = x$ and (5.2), (3.5), and (3.6) are used, it follows that

$$\frac{\partial q^a(x)}{\partial x^\lambda} = f_\lambda^k(x) X_k^a(q). \tag{5.9}$$

If (5.9) is treated as a system of equations for the $q^a(x)$ with given X_k^a and f_λ^k, then the integrability conditions for (5.9) are satisfied if X_k^a and f_λ^k obey (5.7) and (3.10), respectively, with the same structure constants C_{kl}^m.

Equations of the form (5.9) when $d = 1$ have been derived in connection with the study of simply pseudopotentials associated with the sine–Gordon, KdV, Hirota and other equations.[3,4] These results can now be properly interpreted. They are the differential form of the nonlinear action of a three-parameter bilocal Lie group on a one-dimensional representation space.

Let the F^a realize a linear action

$$F^a(r;q) = T_b^a(r) q^b, \tag{5.10}$$

where $T_b^a(r)$ are the elements of a $(d \times d)$-dimensional matrix function $T(r)$. In this case the generator functions of F^a are linear in q^a and are given by

$$X_k^a(q) = I_{kb}^a q^b, \tag{5.11}$$

where the I_{kb}^a are constants. Let I_k be the matrix with matrix elements I_{kb}^a. The set of I_k, $k = 1, \ldots, f$, satisfy the matrix commutation relations

$$I_k I_l - I_l I_k = C_{kl}^m I_m \tag{5.12}$$

and form a d-dimensional matrix representation of the Lie algebra of G.

The linear form of (5.9) is

$$\frac{\partial q^a(x)}{\partial x^\lambda} = f_\lambda^k(x) I_{kb}^a q^b(x). \tag{5.13}$$

Equations of this form have been derived in connection with the study of linear pseudopotentials.[3,5,7] It has been shown[5,7] that, e.g., for $G = \mathrm{SU}(2)$ with a suitable choice of the f_λ^k (for example that already made in Sec. 3 for the nonlinear Schrödinger equation) (5.13) reduces to the linear eigenvalue problem and associated isospectral flow used to solve the soliton equation given by the f_λ^k.

In fact for the case of the nonlinear Schrödinger equation using (3.13) and setting $I_{kb}^a = i(\sigma_k)_b^a$ (σ_k are the 2×2 Pauli matrices) one obtains

$$\frac{\partial}{\partial x^1} \begin{pmatrix} q^1 \\ q^2 \end{pmatrix} = \begin{pmatrix} -i\lambda & u \\ ku^* & i\lambda \end{pmatrix} \begin{pmatrix} q^1 \\ q^2 \end{pmatrix},$$

$$\frac{\partial}{\partial x^0} \begin{pmatrix} q^1 \\ q^2 \end{pmatrix} = \begin{pmatrix} -i(2\lambda^2 + k|u|^2) & 2\lambda u + iu_x \\ k(2\lambda u^* - u_x^*) & i(2\lambda^2 + k|u|^2) \end{pmatrix} \begin{pmatrix} q^1 \\ q^2 \end{pmatrix}.$$

It follows for a linear group action and from (4.3) that

$$T(r(x,y)) = T(R(s(x), \bar{s}(y))) = T(s(x)) T^{-1}(s(y)). \tag{5.14}$$

If $q(x)$ denotes the vector with components $q^1(x), \ldots, q^d(x)$, and (5.14) is used in (5.2), then

$$T^{-1}(s(x)) q(x) = T^{-1}(s(y)) q(y) = q_0, \tag{5.15}$$

where q_0 is some fixed constant vector in the space Q. Thus

$$q(x) = T(s(x)) q_0. \tag{5.16}$$

Differentiating with respect to x yields

$$\frac{\partial q(x)}{\partial x^\lambda} = \frac{\partial T(s(x))}{\partial x^\lambda} T^{-1}(s(x)) q(x). \tag{5.17}$$

This is another form of (5.13) so that we have the equality

$$\Gamma_\lambda(x) \equiv - \frac{\partial T}{\partial x^\lambda} T^{-1} = - f_\lambda^k I_k. \tag{5.18}$$

The quantities $\Gamma_\lambda(x)$ can be used to define a covariant derivative

$$\nabla_\lambda q(x) \equiv \frac{\partial q(x)}{\partial x^\lambda} + \Gamma_\lambda(x) q(x). \tag{5.19}$$

From (5.13) it follows that the covariant derivative of $q(x)$ is zero.

The $\Gamma_\lambda(x)$ give a natural linear connection of Q induced by the bilocal linear group action. This linear connection has the curvature tensor

$$R^a_{b\lambda\mu} \equiv \frac{\partial \Gamma^a_{\mu b}}{\partial x^\lambda} - \frac{\partial \Gamma^a_{\lambda b}}{\partial x^\mu} + \Gamma^a_{\lambda c}\Gamma^c_{\mu b} - \Gamma^a_{\mu c}\Gamma^c_{b}. \tag{5.20}$$

The definition (5.18) together with (5.12) and the integrability conditions for f^k_λ, (3.10) imply that

$$R^a_{b\lambda\mu} = 0. \tag{5.21}$$

When $N=2$, as has been discussed in Sec. 4, the f^k_λ can be chosen so that (3.10) reduces to the soliton equation. The vanishing of the curvature tensor (5.20) associated with the connection (5.18) is a restatement of this fact. The vanishing of a curvature tensor has previously been observed in connection with solitons in Refs. 8 and 9.

The fact that nonlinear equations possessing soliton solutions may be associated with a flat connection in the Q-space defined over two-dimensional space–time is, in the present general framework, of particular importance. The flatness of the connection is the principal geometric manifestation of the path independence of the bilocal group action, i.e., of the main property of the map (3.1) expressed by the composition law (3.2). A path dependent action leads to a nonzero curvature as is shown in Yang's integral formulation of gauge fields.[16] A discussion of nonvacuum gauge fields from the present point of view will be presented elsewhere.

Equation (5.18) shows that any set of functions that satisfies the integrability conditions (3.10) is a pure gauge term. Thus any two solutions can be connected by a gauge transformation. Alternatively put, consider

$$T_i(x) = T(s_i(x)), \quad i=1,2, \tag{5.22}$$

for two families of local group parameters $s_i(x)$, $i=1,2$ possessing the same composition law R^k. (See the end of Sec. 4.) If we construct

$$s^k_3(x) = R^k(s_1(x), s_2(x)), \tag{5.23}$$

then

$$\Gamma_{3\lambda}(x) = \Gamma_{1\lambda}(x) + T_1(x)\Gamma_{2\lambda}(x)T_1^{-1}(x). \tag{5.24}$$

This is the linear version of (4.7). Of course (5.23) [or (4.7)] may not produce a Γ_λ or f^k_λ that has the same *form* as the two original constituents, this would be necessary to obtain a third solution of a given nonlinear partial differential equation. However such superposition principles are contained in (5.23) and (4.7). The problem of the form invariance of the Γ_λ or f^k_λ is currently under investigation since its solution is closely related to finding the Backlund transformation for the original equation.

6. CONSERVATION LAWS

In this section, conservation laws for the equations of motion (3.10) are derived if the group under consideration is semisimple. All computations are performed in canonical parameters. In this case A and V have a simple closed form representation. To give this representation first define a matrix σ with matrix elements

$$\sigma^k_l = t^n C^k_{nl}. \tag{6.1}$$

With this notation[10,11]

$$V(t) = I + \tfrac{1}{2}\sigma + \tfrac{1}{3}\sigma^2 + \cdots. \tag{6.2}$$

As a formal series

$$V(t) = (e^\sigma - I)\sigma^{-1}. \tag{6.3}$$

Using (2.4) it follows from (6.3) that

$$A = \sigma(e^\sigma - 1)^{-1} = I - \tfrac{1}{2}\sigma + \sum_{j=1}^{\infty} \frac{B_{2j}\sigma^{2j}}{2j!}, \tag{6.4}$$

where B_{2j} are Bernoulli numbers.

It is useful to introduce the Killing tensor defined by

$$G_{kl} = C^m_{kp}C^p_{ml} \tag{6.5}$$

This tensor can be used to define objects with lowered indices, so that for example

$$t_k = G_{kl}t^l, \tag{6.6a}$$

$$C_{klm} = G_{kp}C^p_{lm}. \tag{6.6b}$$

It can be shown that in canonical parameters[10,11]

$$t_k V^k_l = t_k A^k_l = t_l. \tag{6.7}$$

In order to investigate the existence of conserved quantities consider the difference

$$\Delta_{\lambda\mu} = \frac{\partial}{\partial y^\lambda}(r^k(x,y)G_{kl}f^l_\mu(y)) - \frac{\partial}{\partial y^\mu}(r^k(x,y)G_{kl}f^l_\lambda(y)). \tag{6.8}$$

Clearly

$$\Delta_{\lambda\mu} = \left(\frac{\partial r^k(x,y)}{\partial y^\lambda}G_{kl}f^l_\mu(y) - \frac{\partial r^k(x,y)}{\partial y^\mu}G_{kl}f^l_\lambda(y)\right) + r^k(x,y)G_{kl}\left(\frac{\partial f^l_\mu(y)}{\partial y^\lambda} - \frac{\partial f^l_\lambda(y)}{\partial y^\mu}\right), \tag{6.9}$$

and using (3.7) and (3.10)

$$\begin{aligned}\Delta_{\lambda\mu} &= (A^k_n(r)f^n_\lambda G_{kl}f^l_\mu - A^k_n f^n_\mu G_{kl}f^l_\lambda) \\ &\quad + r^k G_{kl}C^l_{np}f^\mu_\lambda f^p_\mu \\ &= \{A^k_m G_{kp} - A^k_p G_{km} + r^k G_{kl}C^l_{mp}\}f^m_\lambda f^p_\mu.\end{aligned} \tag{6.10}$$

It is not difficult to show that $\Delta_{\lambda\mu}$ vanishes if the group G is semisimple. If G is semisimple, then G_{kl} is invertible, so that a matrix G^{kl} exists with the property that

$$G^{kl}G_{lm} = G_{ml}G^{lk} = \delta^k_m. \tag{6.11}$$

To show that $\Delta_{\lambda\mu}$ vanishes it is sufficient to show that

$$G_{kp}(\sigma^2)p_l = (\sigma^2)_{kl} = (\sigma^2)_{lk}. \tag{6.12}$$

This follows immediately from the complete antisymmetry of C_{klm}, and the symmetry of G_{kl} and G^{kl}. Thus

$$\Delta_{\lambda\mu} = 0. \tag{6.13}$$

In two-dimensional space–time (6.13) can be cast into the form

$$\nabla \cdot \mathbf{j}(x,y) = \frac{\partial j_0(x,y)}{\partial y^0} + \frac{\partial j_1(x,y)}{\partial y^1} = 0, \tag{6.14a}$$

where

$$j_0(x,y)=r^k(x,y)G_{kl}f_1^l(y), \tag{6.14b}$$

$$j_1(x,y)=-r^k(x,y)G_{kl}f_0^l(y). \tag{6.14c}$$

If

$$C(x;y_0)\equiv\int_{-\infty}^{\infty}r^k(x,y)G_{kl}f_1^l(y)\,dy^1 \tag{6.15}$$

exists and $j_1(x,y)$ vanishes at $|y^1|=\infty$, then

$$\frac{\partial C(x,y_0)}{\partial y^0}=0. \tag{6.16}$$

In fact the integral (6.15) can be explicitly calculated. Using (6.7) and (3.7),

$$\frac{1}{2}\frac{\partial r^k(x,y)G_{kl}r^l(x,y)}{\partial y^\lambda}=r^k(x,y)G_{kl}f_\lambda^l(y). \tag{6.17}$$

Thus

$$C(x)=\tfrac{1}{2}r^k(x,y)G_{kl}r^l(x,y)\Big|_{y^1=-\infty}^{y^1=\infty}. \tag{6.18}$$

To facilitate the derivation of conservation laws in more than two space–time dimensions certain differential 1-forms will be introduced. Define

$$\omega^k(y)=f_\lambda^k(y)\,dy^\lambda\quad(k=1,\ldots,f). \tag{6.19}$$

In this notation (3.10) becomes

$$d\omega^k(y)=\frac{1}{2}\left(\frac{\partial f_\lambda^k}{\partial y^\mu}-\frac{\partial f_\mu^k}{\partial y^\lambda}\right)dy^\lambda\wedge dy^\mu$$

$$=-\tfrac{1}{2}C_{lm}^k\omega^l\wedge\omega^m. \tag{6.20}$$

The fact that $\Delta_{\lambda\mu}$ vanishes is the coordinate statement that the 1-form

$$\Omega(x,y)=r^k(x,y)G_{kl}\omega^l(y), \tag{6.21}$$

is closed. Due to (6.17), $\Omega(x,y)$ is exact, with respect to y, in any space–time dimension.

In N space–time dimensions a conservation law of the form

$$\frac{\partial}{\partial y^\lambda}j^\lambda=0, \tag{6.22}$$

is expressed by giving an $(N-1)-$ form $\chi^{(N-1)}(y)$ such that

$$d\chi^{(N-1)}(y)=0. \tag{6.23}$$

It is easily seen that the N-local $(N-1)$-form

$$\tilde{\chi}^{(N-1)}(y)=\Omega(x_1,y)\wedge\Omega(x_2,y)\wedge\cdots\Omega(x_{N-1},y), \tag{6.24}$$

satisfies (6.23). Furthermore the $(j+1)$-local form

$$\tilde{\chi}^j(y)=\Omega(x_1,y)\wedge\Omega(x_2,y)\wedge\cdots\wedge\Omega(x_j,y),$$

$$j=1,\ldots,N-1 \tag{6.25}$$

has the property that

$$d\tilde{\chi}^j(y)=0,\quad j=1,\ldots,N-1. \tag{6.26}$$

If appropriate convergence conditions are satisfies, the $\tilde{\chi}^j(y)$ can be used to define quantities conserved on submanifolds of the N-dimensional space–time. A full discussion of the sequence of exact nonlocal forms associated with (3.10) will be presented elsewhere.

ACKNOWLEDGMENTS

We are grateful to Professor I. Todorov for his encouragement, his continued interest in this work, and for valuable discussions. We also thank Dr. L. Nikolova, Dr. P. Nikolov, and Dr. V. Gerdjikov for enlightening conversations.

One of us (J. C.) is deeply indebted to the National Academy of Sciences (U.S.A.) and to the Bulgarian Academy of Sciences for making his stay at the INRNE possible. He is also indebted to the Society of Physicists in Bulgaria for support. It is a great pleasure to thank Professor Ch. Christov and Professor I. Todorov and all members of the INRNE for their warm hospitality.

One of us (V. A. R.) thanks the organizing committee of the XVI Internationale Universitätswochen für Kernphysik in Schladming for making it possible for him to attend the school where part of this work was done.

[1]An introduction to solitons can be found in the review by A. C. Scott, F. Y. F. Chu and D. W. McLaughlin, Proc. IEEE 61, 1443 (1973).
[2]H. D. Wahlquist and F. B. Estabrook, J. Math. Phys. 16, 1 (1975).
[3]J. P. Corones and F. Testa in *Backlund Transformations*, Spring Lecture Notes in Mathematics No. 515, edited by R. Miura (Springer, New York, 1976).
[4]J. P. Corones, J. Math. Phys. 17, 756 (1976).
[5]J. P. Corones, J. Math. Phys. 18, 163 (1977).
[6]H. C. Morris, J. Math. 17, 1867, 1870 (1976); 18, 285, 530 (1977).
[7]H. C. Morris, J. Math. Phys. 18, 533 (1977).
[8]R. Hermann, Phys. Rev. Lett. 36, 835 (1976).
[9]M. Crampin, F. Pirani, and D. Robinson, "The Soliton Connection," Math. Phys. Lett. (to be published).
[10]L. P. Eisenhart, *Continuous Groups of Transformations* (Princeton U. P., Princeton, New Jersey, 1933).
[11]L. S. Pontryagin, *Topological Groups* (Gordon and Breach, New York, 1966).
[12]C. N. Yang and R. L. Mills, Phys. Rev. 96, 191 (1954).
[13]K. Pohlmeyer, Commun. Math. Phys. 46, 207 (1976).
[14]L. Lund and T. Regge, Phys. Rev. D 14, 1524 (1976).
[15]L. Lund, "Note on the Geometry of the Nonlinear σ-model," to appear in Phys. Rev. D.
[16]C. N. Yang, Phys. Rev. Lett. 33, 445 (1974).

Discrete path approach to linear recursion relations[a)]

Adel F. Antippa

Département de Physique, Université du Québec à Trois-Rivières, Trois-Rivières, Québec G9A 5H7, Canada
(Received 9 March 1977)

It is shown that the solutions of linear homogeneous recursion relations, with arbitrarily specified boundary conditions, are related, by a mapping, to the totality of discrete paths joining the two ends of an interval and made up of a predetermined set of directed segments. We study the dependence of these solutions on the way the boundary conditions are specified. When the boundary conditions are given as initial conditions, the present approach reduces to the formalism already developed for that specific case, and which is based on the partitions of an interval into classes.

I. INTRODUCTION

The original motivation for the study of multiterm linear recursion relations[1] with variable coefficients, was to obtain an explicit solution of the Schrödinger equation with a linear central potential.[2] This quark confining potential has recently acquired the same importance in particle physics[3] as the Coulomb potential in atomic physics.

As is well known, explicit solutions of the Schrödinger equation with central potentials of the form r^n only existed for $n=-1$, 0, and 2, the Coulomb, constant, and harmonic potentials, respectively. This is related to the fact that, only in these three cases are the expansion coefficients, of the series solution, given by a two-term recursion relation. This type of recursion relation has two important characteristics: (i) it admits an explicit solution (thus giving explicit wave functions), and (ii) it can be cut off by a proper choice of parameters (thus providing an energy eigenvalue equations).[4]

For other values of n, one obtains a three-term recursion relation for the expansion coefficients. These cannot be cut off and their explicit solutions were, until recently, not known. However solutions of this type of recursion relation were recently developed,[1] in terms of so called "combinatorics functions," and the corresponding energy eigenvalue problem resolved without the need of a cutoff.[2] This provided an exact analytic solution of the linear potential problem.[3]

The scope of the method developed in Refs. 1 and 2, goes beyond the problem of the linear potential, and actually has a bearing on the whole subject of ordinary differential equations. As a second application of the method, an exact analytic solution of the combined Coulomb and linear potentials has been obtained.[5] This latter type of quark confining potential is indicated by the requirements of asymptotic freedom and infrared slavery.[3] Previously, this problem could only be handled in the framework of perturbation theory.

The solutions of two-term recursion relations are given essentially in terms of Pochhammer symbols, or equivalently in terms of gamma functions. More precisely, they are given in terms of factorial expressions.[6] It is thus not surprising that the solutions of multiterm linear recursion relations should be given by a sort of highly generalized factorial expressions, which is what the combinatorics functions are. As for the energy eigenvalue equation it is given in terms of so called conjugate combinatoric functions. These are solutions of the inverted recursion relation with the boundary conditions imposed at infinity.[2,5]

Thus at present the solution of multiterm linear homogeneous recursion relations with variable coefficients is known for the two most commonly occurring types of boundary conditions; initial conditions,[1] and final conditions.[2] The solution in both cases is basically a finite sum of terms, each of which is a certain product of the coefficients of the recursion relation. These are the only two types of boundary conditions for which the solution is given in terms of a finite number of terms. As will be seen, in all other cases the number of terms is infinite.

In this paper we would like to modify the formalism used in Refs. 1 and 2 in order to make it suitable for handling the case of arbitrarily specified boundary conditions. The modified formalism, in addition to permitting more flexibility in handling recursion relations, provides a unified treatment of the two previously studied cases of initial and final conditions, and gives a solution for the important case of mixed boundary conditions; partly initial and partly final.

The solutions given previously[1,2] are related by a specific mapping to the totally of partitions of an interval (m,j) into parts $a_1, a_2, \ldots, a_N$ all of which are positive. In the modified formalism, these are replaced by the totality of discrete paths, starting at m and terminating at j, and made up of the directed segments $a_0, a_1, a_2, \ldots, a_N$. When all the segments have the same sign, the discrete paths from m to j become partitions of the interval (m,j).

The discrete paths that have been introduced here are essentially a generalization of one-dimensional random walks.[7] In a conventional random walk, the steps could be forward or backward, but are all equal in magnitude. For discrete paths, on the other hand, the steps are variable. Moreover the general approach presented here for the solution of recursion relations is essentially a flow graph topological approach.[7] In a traditional flow graph approach linear recursion relations would be

a)Work supported in part by the National Research Council of Canada.

 Copyright © 1977 American Institute of Physics

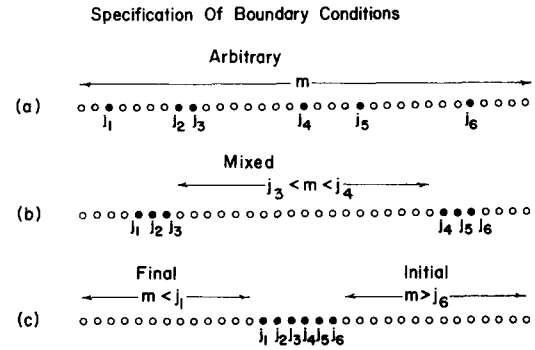


FIG. 1. The three basic ways of specifying the elements of the set $\mathcal{J} = \{j_1, j_2, \ldots, j_h\}$. The boundary conditions are the values of b_{j_α}, $\alpha = 1, 2, \ldots, h$.

treated as an infinite set of simultaneous linear relations, and the problem is too difficult to handle. Our contribution consists of having utilized the functional dependence of the coefficients, to relate the flow graph topology of the problem to a much simpler underlying topology, that of discrete paths. The latter is relatively easy to handle.

To set up the problem we write the general multiterm homogeneous linear recursion relation with variable coefficients as

$$\sum_{k=0}^{N} g_{a_k}(m) b_{m+a_k} = 0, \tag{1.1}$$

where the coefficients $g_{a_k}(m)$ are arbitrary functions of the level m, and the numbers a_k are integers assumed to be ordered according to

$$a_0 < a_1 < \cdots < a_{r-1} < a_r < a_{r+1} < \cdots < a_N, \tag{1.2a}$$

with

$$a_r = 0. \tag{1.2b}$$

Depending on how Eq. (1.1) is written r can take on any value between O and N. The order h of the equation is given by

$$h = a_N - a_0. \tag{1.3}$$

The boundary conditions we will be concerned with specify h different values of the solution b_m for certain values of the index m, given by the elements of the set

$$\mathcal{J} = \{j_1, j_2, \ldots, j_h\}, \tag{1.4}$$

according to

$$b_{j_\alpha} = \lambda_{j_\alpha}, \quad \alpha = 1, 2, \ldots, h. \tag{1.5}$$

The points on the real axis whose coordinates are given by j_α will be referred to as "terminal points." It is convenient to separate the study of boundary conditions into the three cases shown in Fig. 1. These are:

(i) The terminal points are distributed in an arbitrary fashion.

(ii) The terminal points form two separate sets of consecutive integers. If the sets are j_1 to j_s and j_{s+1} to j_h, then for $j_s < m < j_{s+1}$, the boundary conditions are mixed, partly initial and partly final.

(iii) The terminal points are h consecutive integers. In this case, for $m > j_h$ the boundary conditions are initial conditions, while for $m < j_1$ they are final conditions.

In Sec. II we will introduce the terminology of discrete paths, how to operate on them, and their mapping into combinatorics functions. In Sec. III we develop the solution of linear recursion relations, and in Sec. IV we will analyze a special class of paths, those which are closed loops. A concrete example is worked our in the Appendix.

II. DISCRETE PATH APPROACH

A. Discrete paths and terminal points

Consider a one-dimensional real space, which we take to be the x axis, a set of directed segments

$$\bar{\mathcal{A}} = \{a_0, a_1, \ldots, a_{r-1}, a_r, a_{r+1}, \ldots, a_N\}, \tag{2.1}$$

and a set of terminal points

$$\mathcal{J} = \{j_1, j_2, \ldots, j_h\}. \tag{2.2}$$

The elements of $\bar{\mathcal{A}}$ are ordered by Eq. (1.2) and h is given by Eq. (1.3). From the elements of $\bar{\mathcal{A}}$, we construct a new set $\mathcal{A}$, which contains all the elements of $\bar{\mathcal{A}}$ except the element $a_r = 0$. That is,

$$\mathcal{A} = \{a_0, a_1, \ldots, a_{r-1}, a_{r+1}, \ldots, a_N\}. \tag{2.3}$$

A point m on the real axis is said to be joined to a terminal point j, provided there exists a set of segments $(\delta_1, \delta_2, \ldots, \delta_l)$ satisfying the following conditions,

$$\sum_{i=1}^{l} \delta_i = j - m, \tag{2.4a}$$

and

$$\delta_i \in \bar{\mathcal{A}}, \quad i = 1, 2, \ldots, l. \tag{2.4b}$$

The different points on the paths are given by

$$s_1 = m, \quad s_{i+1} = \sum_{k=1}^{i} \delta_k, \quad s_{l+1} = j. \tag{2.4c}$$

1. Zero length segments

It is evident that, from a given path, we can derive an infinite number of essentially identical paths, all passing through exactly the same points on the real axis, by adding segments a_r of zero length to the original path. The need to keep the element $a_r = 0$ in the set of segments is in order to be able to join a terminal point to itself. As will be seen later on, this plays an important role in determining the boundary conditions associated with a given solution. When paths are eventually mapped into functions, all paths that only differ by the number of zero segments they contain must map into the same function. This restriction normalizes the solution and uniquely determines the boundary conditions associated with a given solution. If we artificially eliminate a_r from the set of segments, rather than keeping it and rendering it ineffective, and use the set $\mathcal{A}$ instead of $\bar{\mathcal{A}}$ from the beginning, we lose this vital information, on the normalization of the solutions. Having said this we will from here on only deal with paths which do not contain any zero segments, and justify this *a posteriori* by

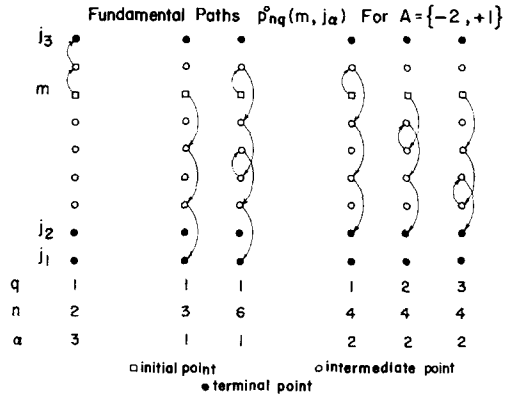


FIG. 2. An example of fundamental paths $p^0_{nq}(m,j_\alpha)$ for $A=\{-2,+1\}$ and mixed boundary conditions. m is the initial point, j_α the terminal point, n the number of parts, and q labels the paths with the same value of m, j_α, and n.

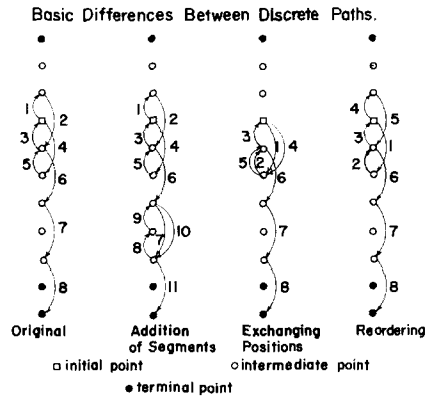


FIG. 3. From a given path we can generate other paths by three methods: adding segments, exchanging the positions (in space) of unequal segments, or reordering segments without exchanging their positions.

choosing our mapping to be independent of the number of zero segments in a path. Thus except for the bath (j_α,j_α), the segments of a path will belong to A rather than $\mathcal{A}$. That is, Eq. (2.4b) will be replaced by

$$\delta_i \in A, \quad i=1,2,\ldots,l. \tag{2.4d}$$

2. Initial, intermediate, and terminal points

A path starting at m terminates whenever it reaches one of the terminal points j_α. Thus none of the intermediate points of a path can be a terminal point,

$$s_i \notin \mathcal{J}, \quad i=2,3,\ldots,l, \tag{2.4e}$$

where l is the number of segments in the path. As for the initial point the situation is a little more subtle. Consider a path from m to j_α with the intermediate points obeying Eq. (2.4e). If m is a terminal point j_β, then the path terminates at its starting point j_β and cannot join to j_α. The one exception to this, is when $\beta=\alpha$, for by terminating at j_β the path would have reached j_α. Thus for a path from m to j_α we have the condition

$$m \neq j_\beta \text{ unless } \beta=\alpha. \tag{2.4f}$$

It is important to note that the path (j_α,j_α) cannot have any closed loops, for then j_α becomes an intermediate point as well, in violation of condition (2.4e).

3. Fundamental paths and closed loops

The discrete paths joining m to j fall into two main categories according to whether or not they have closed loops. The paths without closed loops will be referred to as "fundamental paths" and labeled by two indexes n and q. n is the number of segments in the fundamental path, and q labels the different fundamental paths with the same number of segments. Paths with loops will have an additional index α, which will be used to classify the loops that are superimposed on the fundamental path. In Fig. 2 we give an example of a set of fundamental paths built up from the two segments of the set $A=\{-2,+1\}$.

Corresponding to a given path $p^\alpha_{nq}(m,j)$, joining m to j, we introduce the vectors $\mathbf{\Delta}^\alpha_{nq}$ and $\mathbf{S}^\alpha_{nq}$ by

$$\mathbf{\Delta}^\alpha_{nq}(j-m)=(\delta_1,\delta_2,\ldots,\delta_l), \tag{2.5a}$$

and

$$\mathbf{S}^\alpha_{nq}(m,j)=(s_1,s_2,\ldots,s_l,s_{l+1}). \tag{2.6a}$$

The components of $\mathbf{\Delta}^\alpha_{nq}$ are the segments of the path, and those of $\mathbf{S}^\alpha_{nq}$ are the coordinates of the successive points on the path. For a fundamental path $\alpha=0$, and $l=n$, thus

$$\mathbf{\Delta}^0_{nq}(j-m)=(\delta_1,\delta_2,\ldots,\delta_n), \tag{2.5b}$$

$$\mathbf{S}^0_{nq}(m,j)=(s_1,s_2,\ldots,s_n,s_{n+1}). \tag{2.6b}$$

4. Topology of paths

The totality of discrete paths joining m to j, can be separated into subsets according to their topology. The paths in each subset have the same topology but differ in the ordering of their segments.

Distinct fundamental paths necessarily differ in their topology. On the other hand, paths with loops may have the same topology and still differ by the ordering of their segments. From a given path, there are three basic ways of generating other paths, distinct from it, as shown in Fig. 3. These are:

(i) adding segments,

(ii) exchanging the positions of segments of different length,

(iii) changing the ordering of segments having the same length but positioned differently.

On the other hand, exchanging the ordering of two segments having the same length and position (two segments joining i to $i+3$ for example) does not generate a distinct path. The addition of segments of zero length does not generate distinct paths either, as has already been mentioned. Methods (i) and (ii) above alter the topology of the path, while (iii) leaves in intact.

B. The fusion and union of partial paths

For the purpose of constructing and analyzing paths, it is convenient to introduce the symbol $\otimes$ to indicate

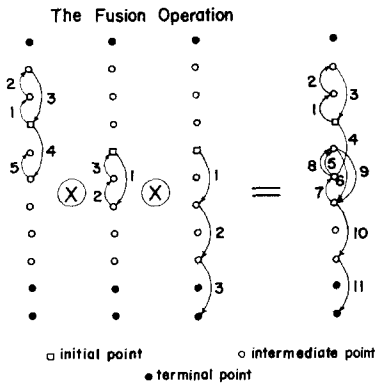


FIG. 4. The noncommutative fusion operation $\otimes$ on partial paths.

the operation of "fusion" of two partial paths as shown in Fig. 4. A partial path does not need to end at a terminal point. The fusion operation is noncommutative but will be assumed to be distributive with respect to the union operation. If the partial path p_{ab} is obtained by taking the partial path p_a and prolongating it by the partial path p_b, then

$$p_{ab} = p_a \otimes p_b \neq p_b \otimes p_a. \tag{2.7a}$$

Furthermore if p_i, $i = 1, 2, \ldots, n$ is a set of partial paths, then

$$p_a \otimes \bigcup_{i=1}^{n} p_i = \bigcup_{i=1}^{n} p_a \otimes p_i. \tag{2.7b}$$

Using the fusion operation we can write a given path as the product of its segments. Thus for a fundamental path, for example, we have

$$p^0_{nq}(m,j) = (m, s_2) \otimes (s_2, s_3) \otimes \cdots \otimes (s_{n-1}, s_n) \otimes (s_n, j), \tag{2.8a}$$

where (s_i, s_{i+1}) is the segment of the path between the points s_i and s_{i+1}. For a nonfundamental path with l segments we have

$$p^\alpha_{nq}(m,j) = (m, s_2) \otimes (s_2, s_3) \otimes \cdots \otimes \cdots \otimes (s_l, j). \tag{2.8b}$$

We will designate by $P(m,j)$ the set of all distinct discrete paths joining m to j subject to conditions (2.4a), (2.4d), (2.4e), and (2.4f),

$$P(m,j) = \{p^\alpha_{nq}(m,j);\ n \in \mathcal{N}(j-m),\ q = 1, \ldots, q_{\max}(n),\ \alpha = 0, 1, 2, 3, \cdots\} \tag{2.9}$$

where

$$\mathcal{N}(j-m) = \{n_1, n_2, \cdots\}, \tag{2.10}$$

and n_i is a possible number of segments for a fundamental path joining m to j. When no paths joining m to j exist, $P(m,j)$ is empty. In particular, if j_α and j_β are two different terminal points, then

$$P(j_\alpha, j_\beta) = \emptyset, \quad \alpha \neq \beta. \tag{2.11}$$

The subset of $P(m,j)$ consisting of all paths having a_k as their first segment, will be denoted by $P_{a_k}(m,j)$. That is,

$$p^\alpha_{nq}(m,j) \in P_{a_k}(m,j) \Leftrightarrow \{p^\alpha_{nq}(m,j) \in P(m,j) \text{ and } \delta_1 = a_k\}. \tag{2.12}$$

Since the segments a_k are all different, then

$$P_{a_i}(m,j) \cap P_{a_k}(m,j) = \emptyset. \tag{2.13}$$

Furthermore since the first segment of every path joining m to j must be, like all the other segments, an element of $\mathcal{A}$, then the set of all paths from m to j subdivides into N disjoint subsets, each containing the paths that start with the same segment. That is,

$$P(m,j) = \bigcup_{a_k \in \mathcal{A}} P_{a_k}(m,j). \tag{2.14}$$

Since the subsets $P_{a_k}(m,j)$ are exhaustive and mutually exclusive, there is a one to one correspondence between the elements of the right- and left-hand sides of Eq. (2.14).

Using the fusion operation, the definition of $P_{a_k}(m,j)$, and the representation of a segment, we have

$$P_{a_k}(m,j) = (m, m+a_k) \otimes P(m+a_k, j). \tag{2.15}$$

Since $P(m+a_k, j)$ is the union of all its elements, then the segment $(m, m+a_k)$ fuses at the beginning of each element of the set $P(m+a_k, j)$ to produce the set $P_{a_k}(m,j)$. Combining the above equation with Eq. (2.14), we finally obtain the basic result

$$P(m,j) = \bigcup_{a_k \in \mathcal{A}} (m, m+a_k) \otimes P(m+a_k, j). \tag{2.16}$$

C. The mapping of paths

We can transform Eq. (2.16) into the recursion relation (1.1) by a mapping of paths into functions in the following way (see Fig. 5),

(i) A segment (s_i, s_{i+a_k}) is mapped into the function $f_{a_k}(x)$ evaluated at s_i:

$$(s_i, s_{i+a_k}) \to f_{a_k}(s_i). \tag{2.17a}$$

(ii) The fusion operation $\otimes$ is mapped into the function multiplication operation

$$\otimes \to X. \tag{2.17b}$$

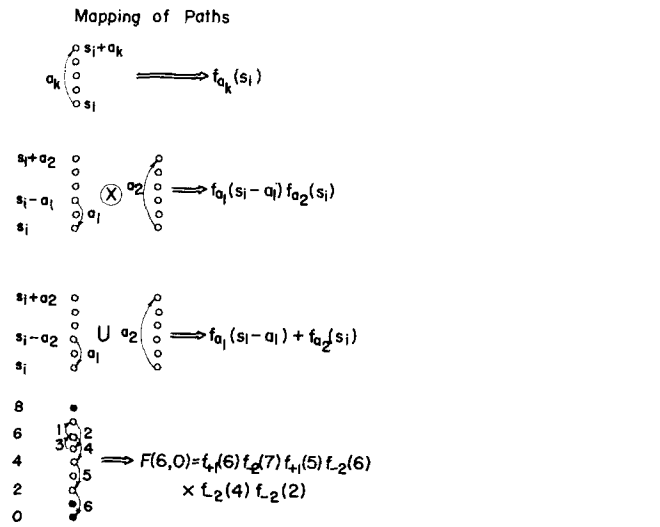


FIG. 5. The mapping that relates paths to combinatorics functions and the solutions of recursion relations.

(iii) The union operation is mapped into the operation of summation

$$\cup \to \sum. \tag{2.17c}$$

(iv) A segment of zero length is mapped into unity,

$$f_{a_r}(s_i) \equiv f_0(s_i) = 1. \tag{2.17d}$$

Condition (iv) is necessary to guarantee that the mapping is independent of the number of segments of zero length that the path contains, and consequently justifies our neglecting these segments when discussing paths.

D. Combinatorics functions

We define the functions $F^{\alpha}_{nq}(m,j)$ and $C(m,j)$ by

$$F^{\alpha}_{nq}(m,j) = \prod_{i=1}^{l} f_{\delta_i}(s_i) \tag{2.18}$$

and

$$C(m,j) = \sum_{n \in \mathcal{N}} \sum_{q=1}^{q_{\max}} \sum_{\alpha} F^{\alpha}_{nq}(m,j), \tag{2.19}$$

where δ_i and s_i, $i = 1, 2, 3, \ldots, l$, are the components of $\Delta^{\alpha}_{nq}(j-m)$ and $\mathbf{S}^{\alpha}_{nq}(m,j)$, respectively. Due to the combinatorial way in which the s_i are evaluated, we refer to $C(m,j)$ as a combinatorics function. Using Eqs. (2.8), (2.9), and the fact that a set is the union of all its elements, we obtain the mappings of the path $p^{\alpha}_{nq}(m,j)$ and the set $\mathcal{P}(m,j)$ as

$$p^{\alpha}_{nq}(m,j) \to F^{\alpha}_{nq}(m,j), \tag{2.20a}$$

and

$$\mathcal{P}(m,j) \to C(m,j). \tag{2.20b}$$

Due to the commutativity of the multiplication of the functions $f_{a_k}(s_i)$, all paths having the same topology map into the same function $F_{nq}(m,j)$.

If there are no paths joining m to j, then, on the one hand $\mathcal{P}(m,j)$ is empty, and on the other hand, Eq. (2.19) gives $C(m,j) = 0$. Hence

$$\phi \to 0. \tag{2.20c}$$

Combining Eqs. (2.11) and (2.20c) we have

$$C(j_\alpha, j_\beta) = 0 \quad \text{for } \alpha \neq \beta. \tag{2.21a}$$

On the other hand, for $\alpha = \beta$, the path joining j_α to j_β exists, has one segment of length zero, and no loops. Thus the set $\mathcal{P}(j_\alpha, j_\alpha)$ contains one element $p^0_{01}(j_\alpha, j_\alpha)$ and according to Eqs. (2.17a) and (2.17d), this maps into unity. Hence

$$C(j_\alpha, j_\alpha) = 1. \tag{2.21b}$$

This will actually determine the normalization of the solution. Combining the above two equations, we have

$$C(j_\alpha, j_\beta) = \delta_{\alpha\beta}. \tag{2.22}$$

Finally by making use of Eqs. (2.17) and (2.20) we find that Eq. (2.16) maps into

$$C(m,j) = \sum_{a_k \in \mathcal{A}} f_{a_k}(m) C(m + a_k, j). \tag{2.23}$$

III. LINEAR RECURSION RELATIONS

The preceding analysis provides a general solution for linear homogeneous recursion relations with nonconstant coefficients and arbitrarily specified boundary conditions.

A. General solution

Theorem: The general solution of the recursion relation

$$b_m = \sum_{a_k \in \mathcal{A}} f_{a_k}(m) b_{m+a_k}, \tag{3.1}$$

subject to the initial conditions

$$b_{j_\alpha} = \lambda_{j_\alpha}, \quad \alpha = 1, 2, \ldots, h, \tag{3.2}$$

is given by

$$b_m = \sum_{\alpha=1}^{h} \lambda_{j_\alpha} C(m, j_\alpha). \tag{3.3}$$

Proof: Substituting Eq. (2.23) in Eq. (3.3) we obtain

$$b_m = \sum_{\alpha=1}^{h} \lambda_{j_\alpha} \sum_{a_k \in \mathcal{A}} f_{a_k}(m) C(m + a_k, j_\alpha). \tag{3.4a}$$

Since the summation over a_k and α are independent of each other, we can exchange them to obtain

$$b_m = \sum_{a_k \in \mathcal{A}} f_{a_k}(m) \sum_{\alpha=1}^{h} \lambda_{j_\alpha} C(m + a_k, j_\alpha), \tag{3.4b}$$

and making use of Eq. (3.3) one more time we obtain Eq. (3.1). As for the boundary conditions, Eq. (3.3) gives

$$b_{j_\beta} = \sum_{\alpha=1}^{h} \lambda_{j\alpha} C(j_\beta, j_\alpha). \tag{3.5}$$

Substituting Eq. (2.22) in Eq. (3.5) gives back Eq. (3.2). This comples the proof of the theorem.

B. Boundary conditions

We will discuss the three types of boundary conditions separately.

1. *Arbitrary specification*: In this case the boundary conditions b_{j_α} are specified at h arbitrarily distributed points j_α as shown in Fig. 1(a). The main characteristic of the resulting solutions is that for every value of m there are in general, an infinite number of loops, and the loops can be infinite in length. The usefulness of the method in this case depends on the speed of convergence of the infinite series giving the combinatorics functions. Each term in the series corresponds to one path.

2. *Mixed specification*: In this case the values of j_α separate into two sets of consecutive integers, $(j_1, \ldots, j_s)$ and $(j_{s+1}, \ldots, j_h)$ as shown in Fig. 1(b). For $m < j_1$ and $m > j_h$ the situation is not very different than the case of arbitrary specification discussed above. But the region $j_s < m < j_{s+1}$, is interesting, and of practical importance. The boundary conditions, when m is in this region, become partly initial and partly final, and we refer to them as mixed.

Since the recursion relation is of order h one needs all the h boundary values b_{j_α}, $\alpha = 1, 2, \ldots, h$, to obtain a unique solution. Thus there must be paths terminating on every one of the j_α's. But since no path can have a terminal point as one of its intermediate points, then

to reach the terminal points j_1 and j_h we need segments having the values $a_-=j_1-(j_s+1)$ and $a_+=j_h-(j_{s+1}-1)$, respectively. Since $j_s-j_1=(s-1)$ and $(j_h-j_{s+1})=h-(s+1)$, then

$$a_-=-s, \quad a_+=h-s, \tag{3.6a}$$

and

$$a_+-a_-=h. \tag{3.6b}$$

The above equations guarantee that, in general, every terminal point can be reached by a path starting at m, and consequently that there are h arbitrary constants in Eq. (3.3). To find out whether the specification of the boundary conditions is compatible with the recursion relation for a specific problem, we need to determine whether Eq. (1.1) can be rewritten in the form (3.1) with the set $\mathcal{A}$ containing the elements a_- and a_+. Equivalengtly we need to determine whether Eq. (1.1) can be rewritten in such a way as to include elements a_+, a_-, and $a_r=0$ in the set $\bar{\mathcal{A}}$. As will be seen shortly this is not always possible.

Comparing Eqs. (1.2) and (1.3) with Eqs. (3.6) we find that the only way to include both a_+ and a_- in $\mathcal{A}$ is to set

$$a_0=a_- \text{ and } a_N=a_+. \tag{3.7}$$

Once the value of a_0 is determined, then the values of all the other segments are also determined through Eq. (1.1). Due to Eqs. (1.3) and (3.6b) this automatically sets $a_N=a_+$. On the other hand, there is no guarantee that among the remaining segments there is one of length zero. Thus this is the crucial test of compatibility of the boundary conditions with the recursion relation.

When there is compatibility, the main characteristic of the solutions is that all paths are bounded, from below by j_1 and from above by j_h. No path can bypass the lower set of terminal points since the most negative segment is $a_0=-s$, and thus cannot make the $(s+1)$ downward jump which would have been necessary to bypass. Similarly the largest positive segment is $a_N=h-s$ and cannot make the $(h-s+1)$ jump necessary to bypass the upper set of terminal points. A direct consequence of this is that the number of fundamental paths in this case is finite. On the other hand, the closed loops, even though they are limited to the region $[j_s+1, j_{s+1}-1]$ can have an infinite number of segments. An illustrative example is solved in the Appendix.

3. *Initial and final specification*: In this case the coordinates j_α of the terminal points are h consecutive integers as shown in Fig. 1(c). Thus the boundary conditions are initial for $m>j_h$ and final for $m<j_1$.

(a) Initial conditions: In order to have paths starting at $m>j_h$ reach all the terminal points; we need a segment $a_-=j_1-(j_h+1)=-h$. In addition we need a segment of length zero. From Eq. (1.3) it is seen that both these requirements can be satisfied by choosing $a_0=a_-=-h$, leading to $a_N=0$. Hence the boundary conditions are compatible with the recursion relation. Since all the elements of $\mathcal{A}$ are negative in this case, then the coordinates s_i of the successive points on the path are decreasing functions of i. Consequencely no closed loops are possible, and each path from m to j_α determines a partition of the interval (j_α, m) with the first part greater than j_h-j_α. Hence this formalism reduces to that developed, in Ref. 1, for the case of initially specified boundary conditions.

(b) Final conditions: To reach all terminal points from $m<j_1$, we now need a segment $a_+=j_h-(j_1-1)=+h$. This and the requirement of having a segment $a_r=0$ can both be satisfied by choosing $a_0=a_r=0$ which leads via Eq. (1.3) to $a_N=a_+=+h$. The compatibility of the boundary conditions with the recursion relation is thus guaranteed. Since the elements of $\mathcal{A}$ are now all positive, s_i is an increasing function of i. Thus again no closed loops are possible, and each path from m to j_α determines a partition of the internal (j_α, m) with the first part greater than $j_\alpha-j_1$ as in the formalism of Ref. 2.

IV. ANALYSIS OF CLOSED LOOPS

As an illustration of the technique of evaluating discrete paths with loops, we consider the very simple, but typical, problem of a three-term recursion relation

$$b_m=f_{+1}(m)b_{m+1}+f_{-1}(m)b_{m-1}, \tag{4.1}$$

subject to the boundary conditions

$$b_{j_1}=\lambda_{j_1} \text{ and } b_{j_2}=\lambda_{j_2}. \tag{4.2}$$

The set $\mathcal{A}$ contains two elements, $\mathcal{A}=\{-1, +1\}$. Thus corresponding to every point m, included between j_1 and j_2, there are only two fundamental paths; one joining it to j_1 and the other to j_2.

To these fundamental paths we can adjoin closed loops. The basic building block of these closed loops, in the special case considered here, is a circle passing through two successive points. We will designate by $\lambda^n(i, i+1)$ a closed loop having n circles passing through the points i and $i+1$,

$$\lambda(i, i+1)=(i, i+1)\otimes(i+1, i), \tag{4.3a}$$

$$\lambda(i+1, i)=(i+1, i)\otimes(i, i+1), \tag{4.3b}$$

$$\lambda^n(i, i+1)=\lambda(i, i+1)\otimes\lambda^{n-1}(i, i+1). \tag{4.3c}$$

Using Eqs. (2.17) we find that λ^n maps as follows:

$$\left.\begin{matrix}\lambda^n(i, i+1)\\ \lambda^n(i+1, i)\end{matrix}\right\rangle \xi^n(i, i+1), \tag{4.4a}$$

where

$$\xi(i, i+1)=f_{+1}(i)f_{-1}(i+1). \tag{4.4b}$$

The set of all closed loops which are made up of circles passing through the points i and $i+1$ and having i as initial point will be denoted by $\Lambda(i, i+1)$,

$$\Lambda(i, i+1)=\{\lambda^n(i, i+1);\ n=1, 2, \ldots, \infty\}, \tag{4.5a}$$

or equivalently

$$\Lambda(i, i+1)=\bigcup_{n=1}^{\infty}\lambda^n(i, i+1). \tag{4.5b}$$

The set Λ is graphically represented by a dark circle, as shown in Fig. 6, and maps as follows:

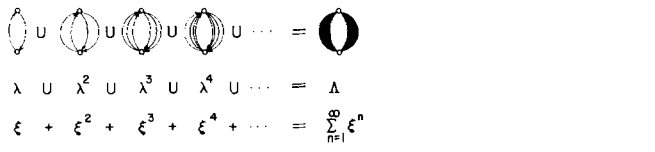


FIG. 6. The graphic representation of the set Λ, its elements λ^n, and its mapping.

$$\Lambda(i,i+1)\Rightarrow\eta(i,i+1)=\sum_{n=1}^{\infty}\xi^n(i,i+1)=\frac{\xi(i,i+1)}{1-\xi(i,i+1)}, \tag{4.6}$$

where in performing the sum we have assumed that $|\xi(i,i+1)|<1$.

The set of closed loops which is next in complexity, is the one formed by adjoining circles passing through $i-1$ and i to circles passing through i and $i+1$. We denote the set of all loops thus formed by $\Gamma^{\alpha}(i-1,i,i+1)$, where $\alpha=-$, 0, or +, depending on whether the loop is entered from the fundamental path at the point $i-1$, i, or $i+1$, respectively. Γ^{α} can be divided into subsets each of which contains the loops which have the same topology. The loops in each subset differ by the ordering of their segments. The subset $\Gamma^{\alpha}_{nm}(i-1,i,i+1)$ contains all loops which are made up of n circles $\lambda(i-1,i)$ and m circles $\lambda(i,i+1)$. The graphic representation of the set Γ^{α} and its subsets Γ^{α}_{nm} is shown in Fig. 7.

The elements of the subset Γ^{α}_{nm} will be denoted by γ^{α}_{nml}, and they all map into the same function,

$$\gamma^{\alpha}_{nml}(i-1,i,i+1)\rightarrow\xi^n(i-1,i)\xi^m(i,i+1). \tag{4.7a}$$

Thus the subset Γ^{α}_{nm} maps into

$$\Gamma^{\alpha}_{nm}(i-1,i,i+1)\rightarrow d^{\alpha}_{nm}\xi^n(i-1,i)\xi^m(i,i+1), \tag{4.7b}$$

where d^{α}_{nm} is the number of elements in the subset Γ^{α}_{nm}. Since the subsets Γ^{α}_{nm} are exhaustive and mutually exclusive, then Γ^{α} maps according to

$$\Gamma^{\alpha}(i-1,i,i+1)\rightarrow\sigma_{\alpha}(i)=\sum_{n}\sum_{m}d^{\alpha}_{nm}\,\xi^n(i-1,i)\xi^m(i,i+1). \tag{4.8}$$

For $\alpha=0$, the loop is entered through its central point i. In this case d^{α}_{nm} is the number of ways in which we

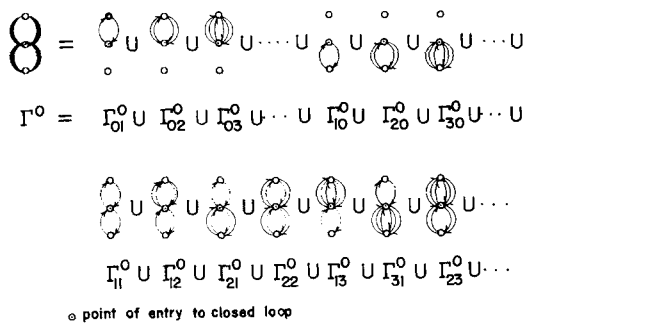


FIG. 7. The graphic representation of the set Γ^0 and its topological subset Γ^0_{nm}.

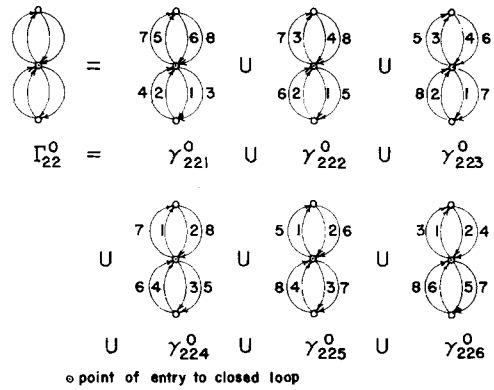


FIG. 8. The graphic representation of the topological subset Γ^0_{22} and its six elements γ^0_{22l}, $l=1,2,\ldots,6$.

can order $n+m$ circles, n of which are of one kind and m of another kind (see Fig. 8). Hence[8]

$$d^0_{nm}=\frac{(n+m)!}{n!\,m!}, \tag{4.9a}$$

and the summation over n and m in Eq. (4.8) is from 0 to ∞ subject to the condition $n+m\geq 1$. For $\alpha=-1$, the point of entry to the loop is $i-1$, and hence the m upper circles $\lambda(i,i+1)$ can only be reached from the n lower circles $\lambda(i-1,i)$. Hence d^-_{nm} is the number of ways we can adjoin m circles to n ordered circles (see Fig. 9). This is the number of m-combinations with repetitions of elements of n types, and is given by[8]

$$d^-_{nm}=\frac{(m+n-1)!}{m!\,(n-1)!}. \tag{4.9b}$$

Furthermore in the summation of Eq. (4.8) m varies from 1 to infinity and m from 0 to infinity. Similarly for $\alpha=+1$, the point of entry is $i+1$, m varies from 1 to infinity, n varies from 0 to infinity, and

$$d^+_{nm}=\frac{(n+m-1)!}{n!\,(m-1)!}. \tag{4.9c}$$

Using Eqs. (4.9), and the above results on the limits of the summation of n and m for the different values of α, we have,

$$\sigma_0(i)=\sum_{\substack{n=0\\ n+m\geq 1}}^{\infty}\sum_{m=0}^{\infty}\frac{(n+m)!}{n!\,m!}\,\xi^n(i-1,i)\xi^m(i,i+1), \tag{4.10a}$$

$$\sigma_-(i)=\sum_{n=1}^{\infty}\sum_{m=0}^{\infty}\frac{(m+n-1)!}{m!\,(n-1)!}\,\xi^n(i-1,i)\xi^m(i,i+1), \tag{4.10b}$$

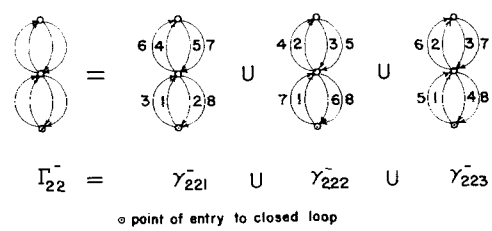


FIG. 9. The graphic representation of the topological subset Γ^-_{22} and its three elements γ^-_{22l}, $l=1,2,3$.

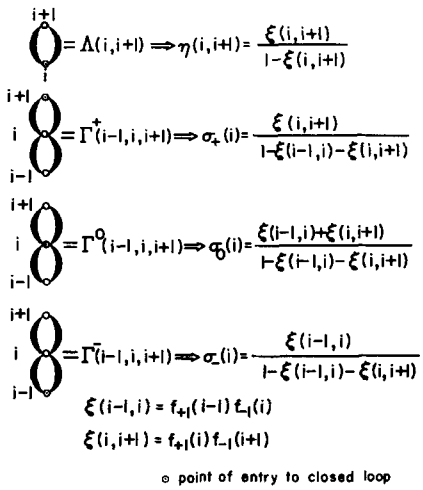


FIG. 10. Summary of the graphic representations, symbols, and mappings of, up to second order, loop sets arising in the study of the recursion relation $b_m = f_{+1}(m) b_{m+1} + f_{-1}(m) b_{m-1}$.

$$\sigma_{+}(i) = \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} \frac{(n+m-1)!}{n!\,(m-1)!}\, \xi^{n}(i-1,i)\, \xi^{m}(i,i+1). \quad (4.10c)$$

Making the change of variable $m \to q = m+n$ in Eq. (4.10a) we obtain

$$\sigma_{0}(i) = \sum_{q=1}^{\infty} \sum_{n=0}^{q} \frac{q!}{n!\,(q-n)!}\, \xi^{n}(i-1,i)\, \xi^{q-n}(i,i+1)$$

$$= \sum_{q=1}^{\infty} [\xi(i-1,i) + \xi(i,i+1)]^{q}.$$

Assuming that $|\xi(i-1,i) + \xi(i,i+1)| < 1$, the above sum adds up to

$$\sigma_{0}(i) = \frac{\xi(i-1,i) + \xi(i,i+1)}{1 - \xi(i-1,i) - \xi(i,i+1)}. \quad (4.11a)$$

Similarly we find that

$$\sigma_{-}(i) = \frac{\xi(i-1,i)}{1 - \xi(i-1,i) - \xi(i,i+1)}, \quad (4.11b)$$

and

$$\sigma_{+}(i) = \frac{\xi(i,i+1)}{1 - \xi(i-1,i) - \xi(i,i+1)}. \quad (4.11c)$$

It is interesting to note that

$$\sigma_{0}(i) = \sigma_{-}(i) + \sigma_{+}(i). \quad (4.12)$$

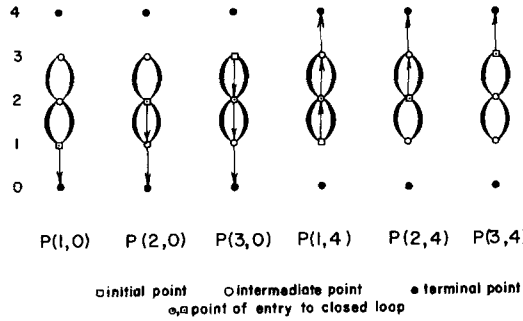


FIG. 11. The six sets of paths arising in the solution of the recursion relation $b_m = f_{+1}(m) b_{m+1} + f_{-1}(m) b_{m-1}$ subject to the mixed boundary conditions specification $j_1 = 0$, $j_2 = 4$.

A summary of the notation used, and results obtained, in this section is given in Fig. 10.

For loops of higher order, the analysis becomes correspondingly more complicated, and we will not go into it here. Instead, in the Appendix, we will apply the results already developed, for up to second order loops, to obtain an exact solution for the case $j_1 = 0$, $j_2 = 4$.

V. CONCLUSION

The solution of linear recursion relations has been shown to be related by a mapping to the problem of joining two points by a discrete path made up of a predetermined set of directed segments.

We have found that the way in which the recursion relation should be written depends on the way in which the boundary conditions are specified. For initially specified conditions, each term should be given in terms of lower level terms, while, for finally specified conditions, each term is written in terms of higher level terms. In both these cases we found that the specification of the boundary conditions is compatible with the recursion relation, that the paths have no closed loops, and that the present approach using paths reduces to the formalism using partitions.[1,2]

For mixed boundary conditions, part of the conditions are specified initially and part finally. We have found that not all ways of effecting the separation into initial and final conditions is compatible with the recursion relation. When there is compatibility, the paths are bounded by the initial conditions from below and the

TABLE I. The sets of discrete paths $P(m, j_\alpha)$ and related combinatorics functions $C(m, j_\alpha)$ involved in the evaluation of b_m for $1 \le m \le 3$. "I" is the identity element with respect to the operation of adjoining ⊗.

m	j_α	$P(m, j_\alpha)$	$C(m, j_\alpha)$
1	0	$[I \cup \Gamma^{-}(1,2,3)] \otimes (1,0)$	$[1+\sigma_{-}(2)] f_{-1}(1)$
2	0	$[I \cup \Gamma^{0}(1,2,3)] \otimes (2,1) \otimes (1,0)$	$[1+\sigma_{0}(2)] f_{-1}(2) f_{-1}(1)$
3	0	$(3,2) \otimes [I \cup \Gamma^{0}(1,2,3)] \otimes (2,1) \otimes (1,0)$	$f_{1}(3)[1+\sigma_{0}(2)] f_{-1}(2) f_{-1}(1)$
1	4	$(1,2) \otimes [I \cup \Gamma^{0}(1,2,3)] \otimes (2,3) \otimes (3,4)$	$f_{+1}(1)[1+\sigma_{0}(2)] f_{+1}(2) f_{+1}(3)$
2	4	$[I \cup \Gamma^{0}(1,2,3)] \otimes (2,3) \otimes (3,4)$	$[1+\sigma_{0}(2)] f_{+1}(2) f_{+1}(3)$
3	4	$[I \cup \Gamma^{+}(1,2,3)] \otimes (3,4)$	$[1+\sigma_{+}(2)] f_{+1}(3)$

final conditions from above. None the less they may have an infinite number of loops, and these loops may be infinite in length.

Finally we have worked out a specific example of mixed boundary conditions to show how, in practice, discrete paths can be handled, classified, and summed. In more complicated cases one may decide to sum paths up to a given order, that, is develop a perturbation theory approach. From the point of view of insight into a given problem, this remains far superior to a purely numerical approach.

APPENDIX

Given the recursion relation

$$b_m = f_{+1}(m) b_{m+1} + f_{-1}(m) b_{m-1}, \tag{A1}$$

subject to the boundary conditions

$$b_0 = \lambda_0 \quad \text{and} \quad b_4 = \lambda_4, \tag{A2}$$

we want to find an explicit expression for b_m, in the range $1 \leq m \leq 3$. The relevant paths are shown in Fig. 11, and are given with their corresponding combinatorics functions in Table I.

To obtain numerical values for the combinatorics functions we have to make a specific choice for the coefficients $f_{+1}(m)$ and $f_{-1}(m)$. We choose

$$f_{+1}(i) = \frac{1}{i+1} \quad \text{and} \quad f_{-1}(i) = 1. \tag{A3}$$

Then, according, to Eq. (4.4b),

$$\xi(i-1, i) = \frac{1}{i} \quad \text{and} \quad \xi(i, i+1) = \frac{1}{i+1}. \tag{A4}$$

When substituted in Eqs. (4.11) these lead to:

$$\sigma_+(i) = \frac{i}{i^2 - i - 1}, \quad \sigma_+(2) = 2, \tag{A5a}$$

$$\sigma_0(i) = \frac{2i+1}{i^2 - i - 1}, \quad \sigma_0(2) = 5, \tag{A5b}$$

$$\sigma_-(i) = \frac{i+1}{i^2 - i - 1}, \quad \sigma_-(2) = 3. \tag{A5c}$$

The resulting values of $C(m, j_\alpha)$ are:

$$\begin{aligned} &C(1,0) = 4, \quad C(2,0) = 6, \quad C(3,0) = 6, \\ &C(1,4) = \tfrac{1}{4}, \quad C(2,4) = \tfrac{2}{4}, \quad C(3,4) = \tfrac{3}{4}. \end{aligned} \tag{A6}$$

Combining Eqs. (3.3) and (A6) we obtain the required solution in the form

$$b_1 = 4\lambda_0 + \tfrac{1}{4}\lambda_4, \tag{A7a}$$

$$b_2 = 6\lambda_0 + \tfrac{2}{4}\lambda_4, \tag{A7b}$$

$$b_3 = 6\lambda_0 + \tfrac{3}{4}\lambda_4. \tag{A7c}$$

As a check on this result we note that if the specific form of the coefficients f_{+1} and f_{-1}, as given by Eq. (A3), is used, Eq. (A1) can be rewritten in the form

$$b_m = m(b_{m-1} - b_{m-2}). \tag{A8}$$

If in addition we use as boundary conditions, $b_0 = \lambda_0$ and $b_1 = 4\lambda_0 + \frac{1}{4}\lambda_4$, Eq. (A8) reproduces the values of b_2, b_3, and b_4 as given in Eqs. (A7b), (A7c), and (A2), respectively.

Finally, by way of a numerical example, if $\lambda_0 = 1$ and $\lambda_4 = -12$, then Eqs. (A7) give $b_1 = 1$, $b_2 = 0$, and $b_3 = -3$.

[1] A. F. Antippa and A. J. Phares, J. Math. Phys. 18, 173 (1977).

[2] A. F. Antippa and A. J. Phares, "The linear potential: A solution in terms of combinatorics functions," J. Math. Phys. (to be published).

[3] T. Applequist and H. D. Politzer, Phys. Rev. Lett. **34**, 43 (1975); A. De Rujula and S. L. Glashow, *ibid.* **34**, 46 (1975); T. Applequist, A. De Rujula, H. D. Politzer, and S. L. Glashow, *ibid.* **34**, 365 (1975); B. J. Harrington, S. Y. Park, and A. Yildiz, *ibid.* **34**, 168, 706 (1975); E. Eichten, K. Gottfried, T. Kinoshita, K. D. Lane, and T. M. Yan, *ibid.* **36**, 500 (1976); E. P. Tryon, *ibid.* **36**, 455 (1976); J. F. Gunion and L. F. Li, Phys. Rev. D **12**, 3583 (1975); J. F. Gunion and R. S. Willey, *ibid.* **12**, 174 (1975); C. L. Critchfield, J. Math. Phys. **17**, 261 (1976); E. P. Tryon, Phys. Rev. Lett. **28**, 1605 (1972); A. Casher, J. Kogut, and L. Susskind, *ibid.* **31**, 792 (1973); J. Kogut and L. Susskind, Phys. Rev. D **9**, 3501 (1974); H. D. Politzer, Phys. Rev. Lett. **30**, 1346 (1973); D. J. Gross and F. Wilczek, *ibid.* **30**, 1343 (1973).

[4] See, for example, E. Merzbacher, *Quantum Mechanics* (Wiley, New York, 1961), pp. 186–190.

[5] A. F. Antippa and A. J. Phares (to be published).

[6] L. M. Milne-Thomson, *The Calculus of Finite Differences* (Macmillan, London, 1951), p. 25.

[7] See, for example, Frank Harary, *Graph Theory and Theoretical Physics* (Academic, New York, 1967).

[8] See, for example, N. Ya. Vilenkin, *Combinatorics* (Academic, New York, 1971), Chaps. 1 and 2.

The translation kernel in the n-dimensional scattering problem

Marcel Coz

Department of Physics and Astronomy, University of Kentucky, Lexington, Kentucky 40506

Pierre Rochus[a)]

Physique Nucléaire Théorique, Université de Liège,
Institut de Physique au Sart-Tilman, B—4000 Liège 1, Belgium
(Received 4 February 1977)

Radial wavefunctions are defined for the n-dimensional scattering problem ($n > 1$) with spherical symmetry by conditions of regularity at the origin or by conditions of behavior at infinity. The existence of translation kernels can therefore be discussed in both instances. The problem of representing regular solutions appears to be essentially different from that of representing irregular solutions. The essential difference originates from the type of domain used in the representation: It is bounded in the first case and unbounded in the second. If one can still compare the ranges of validity of the two types of representation when one is dealing with a scalar situation, upon proceeding to a matrix situation, a comparison is no longer possible.

1. INTRODUCTION

Our first intention was to bring new elements into the discussion of the role of the translation kernel in the potential scattering discussion begun in Refs. 1—3. The first motivation for seeking new elements was noted in Ref. 3: it was the existence of a connection between the partial wave translation kernel and the partial wave Green function. This connection appears when the equations

$$\psi_l(x) = \phi_l(x) + \int K(l,x,y)\phi_l(y)\,dy, \tag{1}$$

$$\psi_l(x) = \phi_l(x) + \int G(l,k;x,y)V_l(y)\phi_l(y)\,dy \tag{2}$$

containing ϕ_l, the radial solution for the "reference" Schrödinger equation, and ψ_l, the radial solution for the complete equation, are considered simultaneously. A second motivation was the technique due to Blazěch.[4] From the partial operator K_l one can construct a global transformation operator κ using any of the two equations

$$\kappa(x,y;\mathbf{n}_1\cdot\mathbf{n}_2) = \sum_l (2l+1)K_l(x,y)P_l[\cos(\mathbf{n}_1\cdot\mathbf{n}_2)] \tag{3a}$$

or

$$\kappa(\mathbf{x},\mathbf{y}) = \sum_l (2l+1)K_l(x,y)P_l[\cos(\mathbf{x}\cdot\mathbf{y})]. \tag{3b}$$

Along these lines in Ref. 2(a) we investigated the conditions for the existence of K_l for the three-dimensional potential scattering problem when the potential possesses spherical symmetry. A first extension of the study was obtained in Ref. 2(b) where the many (finite)-channel scattering problem was solved. A second possible extension may be investigated. It is realized when one wants to consider the n-dimensional problem ($n > 0$). In the study the separation of the $n = 1$ case from the $n > 1$ cases becomes necessary. While for $n > 1$, a radiation condition emerges as a prerequisite,[5] in the $n = 1$ case, such a condition cannot be verified. The $n = 1$ case was therefore separated and studied in Ref. 3. However, in Refs. 2 and 3, we kept the restrictive condition that the number of channels remain finite. The original purpose of the present paper was therefore to subject the $n > 1$ dimensional scattering problem to discussion and to present indications on how the restriction upon the finite number of channels could be removed. However, during the work it became more and more evident that the representation of regular solutions (Gel'fand—Levitan representation) and of irregular solutions defined by an asymptotic condition (Marchenko representation) were not two parallel problems. Consequently while considering possible extensions to the cases considered in Refs. 2(a) and 2(b) the interest shifted to the differences between the two types of representations. Reasons for the differences were brought to light. They were two in number. The first: the Marchenko contour is not closed whereas the Gel'fand—Levitan is closed. This latter is built up of four segments and it is one of these segments which is troublesome. The second: the essential element for the discussion, the Riemann solution itself, has different analytical properties in the two cases. As a by-product of our study, appears the necessity of choosing a complete system of functions which may lead either to a finite set of differential equations or to an infinite set of equations which can be truncated. In Ref. 6 this problem of truncating an infinite number of channels has received some consideration. In the results presented here the condition for the existence of a translation kernel are expressed in terms of requirements imposed on the matrix elements of the potential between two channels.

To avoid any misunderstanding we want to emphasize that we have made no attempt in this paper to apply hyperspherical systems to the many-body problem. (The poor convergence of hyperspherical systems[7] and the lack of "compactness" of the Lippman—Schwinger equation forbid the use of translation operators in the many-body problem.

The present paper is divided into four sections: The Introduction is in Sec. 1, the study of the radial equation for the n-dimensional Schrödinger equation is found in Sec. 2 where the condition for the existence of the

[a)]Chercheur IISN.

 Copyright © 1977 American Institute of Physics

solutions is stated. In Sec. 3, translation operators are applied to the n-dimensional Schrödinger equation, while in Sec. 4 the existence of translation operators for the matrix n-dimensional radial equations are discussed.

To fix the notations throughout this work the number n denotes the dimension of the space.

As usual with any function of one or many variables, one associates its absolute value by $f \to |f|$. To each matrix A, one will associate the matrix of its absolute value $|A|$ defined by

$$|A|_{ij} = |A_{ij}|,$$

and its Marchenko's norm

$$\|A\| \equiv \sup_i \sum_{j=1} |A_{ij}|.$$

2. THE RADIAL EQUATION

One considers the time independent n-dimensional Schrödinger equation

$$\Delta_n \psi(\mathbf{x}) + [k^2 - V(\mathbf{x})]\psi(\mathbf{x}) = 0. \tag{4}$$

In Eq. (4) one has

$$\Delta_n = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$$

and $\psi(\mathbf{x})$ denotes the n-dimensional scattering wave. Suppose the interaction potential $V(\mathbf{x})$ possesses spherical symmetry and satisfies the following conditions[7,8]: At the origin and at infinity

$$\int_0^r s\,|V(s)|\,ds < \infty \quad \text{and} \quad \int_r^\infty |V(s)|\,ds < \infty.$$

Then with a sufficiently large $r = |x|$, one can find a solution of Eq. (4) of the form,[10]

$$\psi(\mathbf{x}) = \exp[i(\mathbf{k}\cdot\mathbf{x})] + f(k,\theta)\frac{\exp(ikr)}{r^{(n-1)/2}} + O(r)^{-(n+1)/2}. \tag{5}$$

An alternate way of posing the problem is to seek a solution $u(\mathbf{x}) = \psi(\mathbf{x}) - \exp(i\mathbf{k}\cdot\mathbf{x})$ which satisfies a finiteness condition:

$$|r^{(n-1)/2} u(\mathbf{x})| \leq \text{const}, \tag{6a}$$

and verifies uniformly in all directions the Sommerfeld radiation condition[11,5]

$$\lim_{r\to\infty} r^{(n-1)/2}\left(\left|\frac{d}{dr}u(\mathbf{x}) - iku(\mathbf{x})\right|\right) = 0. \tag{6b}$$

This condition is not satisfied by the one-dimensional Schrödinger equation solution.[3] The Sommerfeld condition means, in physical terms, that no energy can be radiated in from infinity.[11] A more satisfying version of conditions 3(a), 3(b) was given by Rellich,[12]

$$\lim_{R\to\infty} \int_{|x|=R} dS \left|\frac{\partial}{\partial r}\psi(\mathbf{x}) - ik\psi(\mathbf{x})\right|^2 = 0. \tag{6c}$$

Condition $\{[(6a) \text{ and } 6(b)] \equiv 6(c)\}$ expresses simply the characterization of the wave function $\psi(\mathbf{x})$ as a vehicle for the energy. It does not tell whether $\psi(\mathbf{x})$ is a solution of a Schrödinger equation for an energy operator H, or whether it is not. The conditions for the existence of H are not hard to satisfy (see Ref. 13); but even if H does not exist, condition (6) may be valid.

In what follows, we use Eq. (5) without assuming the existence of H. Let us set

$$\mathbf{k}\cdot\mathbf{x} = kr\cos\theta.$$

The n-dimensional plane wave expansion follows:

$$\exp(ikr\cos\theta) = \left(\frac{2}{kr}\right)^{(n-2)/2}\Gamma[\tfrac{1}{2}(n-2)] \times \sum_{p=0}^{\infty} i^p\left(p + \frac{n-2}{2}\right)J_{p+(n-2)/2}(kr) \times C_p^{(n-2)/2}(\cos\theta), \tag{7}$$

with the Gegenbauer polynomial $C_p^{(n-2)/2}(\cos\theta)$, see Ref. 14, defined by the generating function

$$\sum_{p=0}^{\infty} r^p C_p^{(n-2)/2}(\cos\theta) = (1 - 2r\cos\theta + r^2)^{-(n-2)/2}.$$

Gegenbauer polynomials are related to Jacobi polynomials by $(\lambda + \frac{1}{2})_p C_p^\lambda(x) = (2\lambda)_p P_p^{\alpha\alpha}(x)$.[15] Therefore, expansions may be found in the literature which employ the Jacobi polynomials.[16,17]

Obviously $C_p^{1/2} = P_p$, see Ref. 14.

Since

$$J_{r\to\infty\nu} \sim \frac{1}{\sqrt{\pi}}\left(\frac{2}{kr}\right)^{1/2}\left(\sin\, kr - (\nu - \tfrac{1}{2})\frac{\pi}{2}\right)(18),$$

one has

$$J_{\substack{p+(n-2)/2\\ r\to\infty}} \sim \frac{1}{\pi^{1/2}}\left(\frac{2}{kr}\right)^{1/2}\sin\left[kr - \left(p + \frac{n-3}{2}\right)\frac{\pi}{2}\right].$$

One introduces hyperspherical Bessel functions by

$$j_p(kr) = \tfrac{1}{2}\Gamma[\tfrac{1}{2}(n-2)]\left(\frac{2}{kr}\right)^{(n-2)/2} J_{p+(n-2)/2}(kr). \tag{8}$$

They behave asymptotically as

$$j_p(kr) \sim \tfrac{1}{2}\Gamma[\tfrac{1}{2}(n-2)]\left(\frac{2}{kr}\right)^{(n-2)/2}\left(\frac{1}{\pi^{1/2}}\right)\left(\frac{2}{kr}\right)^{1/2} \times \sin\left[kr - \left(p + \frac{n-3}{2}\right)\frac{\pi}{2}\right].$$

Using Eq. (8) the plane wave expansion of Eq. (7) reads as

$$\exp(ikr\cos\theta) = \sum_p (2p + n - 2)i^p j_p(kr) C_p^{(n-2)/2}(\cos\theta). \tag{9}$$

To obtain the scattering amplitude, the asymptotic form for Eq. (9) is used,

$$(\phi_{\substack{\text{PW}\\ r\to\infty}}) \sim \frac{1}{\pi^{1/2}}\sum_p (2p + n - 2)i^p\left(\frac{2}{kr}\right)^{1/2}\left(\frac{2}{kr}\right)^{(n-2)/2} \times \Gamma[\tfrac{1}{2}(n-2)]\frac{i}{2}\left(\exp\left\{-i\left[kr - \left(p + \frac{n-3}{2}\right)\frac{\pi}{2}\right]\right\} - \exp\left\{+i\left[kr - \left(p + \frac{n-3}{2}\right)\frac{\pi}{2}\right]\right\}\right)C_p^{(n-2)/2}(\cos\theta). \tag{10}$$

In Eq. (10) we have used the subscripts PW to denote the words plane wave. From Eq. (10) we get $\psi_{\text{SCATTERED}}$ defined by the equation

$$\psi_{\text{SC}} = \frac{\exp(ikr)}{r^{(n-1)/2}} f(k,\theta);$$

by identification the transition amplitude $f(k,\theta)$ is

obtained,

$$f(k,\theta)=i\frac{2^{(n-5)/2}}{\pi^{1/2}}\frac{\Gamma[(n-2)/2]}{k^{(n-1)/2}}\sum_p (2p+n-2)\times\exp[i(n-3)\pi/4](S_p-1)C_p^{(n-2)/2}(\cos\theta). \quad (11)$$

Equation (11) can be assumed therefore from the scattering radiation condition without the adjunction of the assumption that some Hamiltonian exists.

In our work, a Hamiltonian is assumed and Eq. (11) comes from the partial wave decomposition of the scattering solution in a way similar to that followed for the $n=3$ dimensional case.

A time dependent Schrödinger equation is considered,

$$i\hbar\frac{\partial}{\partial t}\phi(\mathbf{x},t)=-\frac{\hbar^2}{2M}\Delta_n\Phi(\mathbf{x},t)+U(\mathbf{x})\Phi(\mathbf{x},t).$$

We write $\phi(\mathbf{x},t)=\psi(\mathbf{x})g(t)$ to obtain the stationary equation,

$$-\frac{\hbar^2}{2M}\Delta_n\psi(\mathbf{x})+U(\mathbf{x})=E\psi(\mathbf{x}),$$

from which the reduced equation results,

$$\Delta_n\psi(\mathbf{x})+k^2\psi(\mathbf{x})-V(\mathbf{x})\psi(\mathbf{x})=0.$$

Let us assume a special form for the interaction $V(\mathbf{x})$,

$$V(x)=V(\|\mathbf{x}\|). \quad (12)$$

We indicate now the behaviors of the radial functions used in the scattering descriptions when r goes to zero.

For the Bessel and Newman functions, as r goes to zero

$$J_{p+1/2}\sim r^{p+1/2} \quad \text{or} \quad J_\nu\sim r^\nu,$$
$$N_{p+1/2}\sim r^{-p-1/2} \quad \text{or} \quad N_\nu\sim r^{-\nu}.$$

When n is even, $\nu=p+(n-2)/2$ is an integer and the Newman function N_ν is defined by $\lim_{\epsilon\to0}(1/\epsilon)\times[J_{\nu+\epsilon}-(-1)^\nu J_{-\nu-\epsilon}]$.

For hyperspherical Bessel functions the behavior at the origin is

$$j_p\sim\left(\frac{1}{r}\right)^{(n-2)/2}J_{p+(n-2)/2}=r^{-(n-2)/2}\gamma^p\gamma^{n/2}\gamma^{-1}\sim r^p,$$
$$\eta_p\sim\left(\frac{1}{r}\right)^{(n-2)/2}N_{p+(n-2)/2}\sim r^{-(n-2)/2}r^{-p}r^{-n/2}r^{+1}\sim r^{-(p+n-2)},$$

while for the Riccati—Bessel's it is

$$u_p\sim j_p r^{(n-1)/2}\sim r^{p+(n-1)/2},$$
$$v_p=\eta_p r^{(n-1)/2}\sim r^{-p-(n-3)/2}. \quad (13)$$

We had introduced previously the hyperspherical Bessel functions. This is their equation[18,19]:

$$\left[\frac{d^2}{dr^2}+\frac{(n-1)}{r}\frac{d}{dr}+k^2-\frac{p(p+n-2)}{r^2}\right]\frac{j_p}{\eta_p}=0. \quad (14)$$

To arrive at the analogous one for the Riccati—Bessel's we use the reducing factor $f(r)=r^{-(n-1)/2}$ and we get the reduced equation

$$\left[\frac{d^2}{dr^2}+k^2-\frac{p(p+n-2)}{r^2}-\frac{1}{4r^2}(n-1)(n-3)\right]\frac{u_p}{v_p}=0. \quad (15)$$

It is obvious that the product of the behaviors of u_p, v_p at the origin, namely the product $[p+(n-1)/2]\times[-p-(n-3)/2]$ should equal the coefficient of r^{-2} in Eq. (15).

Together with the Riccati—Bessel functions we should introduce the Riccati—Hankel functions which behave like imaginary exponentials when r goes to infinity.

After these preliminaries we can concern ourselves with the solutions for the radial equation

$$\left[\frac{d^2}{dr^2}+k^2-\frac{(\nu^2-\frac{1}{4})}{r^2}-V(r)\right]\psi(p,k,r)=0 \quad (16)$$

with $\nu=p+(n-2)/2$. In the discussion of Eq. (16) the authoritative treatment of Newton in Ref. 20 is followed. Regular and irregular solutions $\psi(p,k,r)$, $f(p,\pm k,r)$ are respectively defined by conditions at the origin and conditions as r goes to infinity. We write the integral equations which define these solutions,

$$\phi(p,k,r)=u(p,k,r)+\int_0^r g(p,k;r,s)V(s)\phi(p,k,s)\,ds, \quad (17)$$

$$f(p,\pm k,r)=w(p,\pm k,r)-\int_r^\infty g(p,k;r,s)V(s)f(p,\pm k,s)\,ds, \quad (18)$$

where g is the Green function for Eq. (15).

Equation (16) including the factor $(\nu^2-\frac{1}{4})$ allows the use of Levinson—Newton upper bounds[21] which read, with the notation

$$k=x+iy \quad \text{and} \quad \nu=a+ib, \quad a>0,$$
$$|u(p,k,r)|\le C\left[\frac{r}{1+|k|r}\right]^{a+1/2}\exp(|y|r),$$
$$|v(p,k,r)|\le C\left[\frac{1+|k|r}{r}\right]^{a-1/2}\exp(|y|r),$$
$$g(p,k;r,s)=\{u(p,k,r)v(p,k,s)-u(p,k,s)v(p,k,r)\}/W_r(p,k),$$

where $W_r(p,k)$ is the Wronskian of u and v,

$$|g(p,k;r,s)|\le C\frac{1}{|W_r(p,k)|}\left[\frac{r}{1+|k|r}\right]^{a+1/2}\times\left[\frac{1+|k|s}{s}\right]^{a-1/2}\exp(|y|r).$$

With these bounds we can obtain the existence for the regular and irregular solutions provided that

$$\int_0^\infty|V(s)|\,ds<\infty.$$

3. THE TRANSLATION KERNEL (UNCOUPLED EQUATIONS)

Having established the existence of regular and irregular solutions for the reduced radial equation, we turn to the existence of translation kernels.

While the existence of the solutions depends on properties of the Bessel solutions which do not discriminate between values of p, the problem of the existence of the translation kernel is dependent upon the properties of the Legendre functions, with an order p which can be either an integer or half an integer.

The translation kernels are used as in Refs. 1–3 either for the representation of the irregular solutions $f(f,\pm k,r)$,

$$f(p,\pm k,r)\sim \exp\{\pm (ikr)\}\exp i[(n-3)\pi/4]\exp ip\pi/2,$$

or for the representation of the regular solution (solution which behaves at the origin as a Riccati–Bessel).

We introduce first the operator

$$L(x)=\frac{d^2}{dx^2}-\frac{(p(p+n-2))}{x^2}-\frac{(n-1)(n-3)}{4x^2}$$
$$=\frac{d^2}{dx^2}-\frac{\lambda(\lambda+1)}{x^2}, \tag{19}$$

where we have defined $\lambda=\nu-\frac{1}{2}=p+(n-3)/2$.

The translation kernel $K(x,y)$ is the solution of the Darboux equation[1]

$$L(x)K(x,y)=L(y)K(x+y)+V(x)K(x,y)$$

with appropriate boundary conditions which depend upon whether the regular solution or an irregular solution is represented. These conditions are:

Irregular case	*Regular case*	
(Marchenko)	(Gel'fand–Levitan)	
$\lim_{y\to\infty} K_M(x,y)$	$\lim_{y\to 0} K_G(x,y)=0$	
$=\lim_{y\to\infty}\frac{\partial}{\partial y}K_M(x,y)=0$		
$K_M(x,x)=\frac{1}{2}\int_x^\infty V(s)\,ds$	$K_G(x,x)=\frac{1}{2}\int_0^x V(s)ds$	(20)
$y\geq x$	$x\geq y$	

The kernels $K(x,y)$ are the solutions of integral equations which incorporate their boundary conditions.[3]

To find these integral equations one uses Riemann's method. Let $R(x,y;s,u)$ be the solution of the equation

$$L(s)R(x,y;s,u)=L(u)R(x,y;s,u)$$

with the conditions

$$R(x,y;x,y)=1,$$

$$\frac{\partial R}{\partial s}=\frac{\partial R}{\partial u} \quad \text{when} \quad x+y=u+s,$$

$$\frac{\partial R}{\partial s}=-\frac{\partial R}{\partial u}, \quad \text{when} \quad x-y=u-s.$$

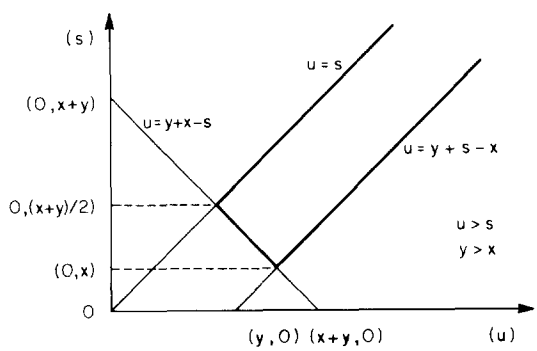


FIG. 1. Marchenko domain.

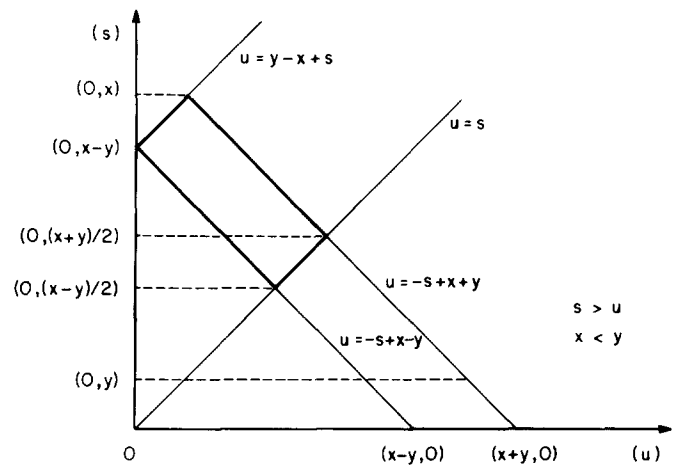


FIG. 2. Gel'fand–Levitan domain.

Let $\mathcal{D}$ be any of the two domains $\mathcal{D}_M$, $\mathcal{D}_G$ specified in Figs. 1 and 2 and let C_+ be the boundary of the domain. By Green's theorem one obtains

$$\iint_{\mathcal{D}} R[L(s)-L(u)]K\,du\,ds$$
$$=\int_{C_+}\left(\frac{\partial}{\partial u}R\,ds+\frac{\partial}{\partial s}R\,du\right)K$$
$$-\int_{C^+} R\,\frac{\partial}{\partial u}K\,ds+\frac{\partial}{\partial s}K\,du\ . \tag{21}$$

The left-hand side is replaced by $\int\int_{\mathcal{D}} RVK\,du\,ds$ and the integrations are performed on the right-hand side. By so doing, one obtains the integral equations.

Irregular case:

$$K_M(x,y)=\tfrac{1}{2}\int_{(x+y)/2}^{\infty} ds\,R(x,y;s,s)V(s)\,ds$$
$$+\tfrac{1}{2}\int\int_{\mathcal{D}_M} R(x,y;s,u)V(s)K(s,u)\,du\,ds; \tag{22a}$$

$\mathcal{D}_M$ is the Marchenko domain shown in Fig. 1 where we have

$$s\geq x; \quad u-s\leq y-x; \quad u+s\geq y+x; \quad y>x; \quad u\geq s.$$

Regular case:

$$K_G(x,y)=\frac{1}{2}\int_{(x-y)/2}^{(x+y)/2} ds\,R(x,y;s,s)V(s)\,ds$$
$$+\int_{(x-y)/2}^{(x-y)}\frac{d}{ds}[R(x,y;s,-s+x-y)]$$
$$\times K(s,-s+x-y)\,ds$$
$$+\frac{1}{2}\iint_{\mathcal{D}_G} R(x,y;s,u)V(s)K(s,u)\,du\,ds. \tag{22b}$$

The domain $\mathcal{D}_G$ in Eq. (22b) is the Gel'fand–Levitan domain of Fig. 2 where we have

$$s\leq x; \quad u-s\geq y-x; \quad u+s\leq x+y; \quad s\geq u;$$

$$s+u\geq x-y; \quad x\geq y.$$

At this point a word on Eqs. (22a) and (22b) is necessary.

Equation (22a) was obtained under the assumption that the contribution of the contour C_+ is negligible when y goes to infinity. Remembering that the kernel $K_M(x,y)$

is such that

$$\lim_{u\to\infty} K_M(s,u) = \lim_{u\to\infty} \frac{\partial}{\partial u} K_M(s,u) = 0,$$

the conditions upon $R(x,y;s,u)$ so that

$$\int_{C_+(u+s\to\infty)} \left(R(x,y;s,u)\frac{d}{du}K(s,u) - \frac{d}{du}R(x,y;s,u)K(s,u)\,du \right) = 0$$

are not particularly stringent.

The contour C_+ of Eq. (22b) is closed at finite distance. It is built up of four segments. Two of these segments are characteristic segments. A third one carries the boundary condition

$$\frac{d}{dr}K(r,r) = \tfrac{1}{2}V(r).$$

We are therefore left with a fourth segment whose contribution will in general be different from zero. However in the case of uncoupled equations considered in this section in the presence of the centripetal potential only, the Riemann function is

$$\begin{aligned} R &= \Pi\begin{pmatrix} \lambda, \lambda \\ x_1, x_2 \end{pmatrix} \\ &= P_\lambda(1-2x_1) - 2x_2 \int_0^1 P_\lambda(1-2x_1+2x_1t)P'_\lambda(1-2x_2t)\,dt \\ &= P_\lambda(1-2x_2) - 2x_1 \int_0^1 P_\lambda(1-2x_2t)P'_\lambda(1-2x_1+2x_1t)\,dt \\ &= P_\lambda(1-2x_1-2x_2+2x_1x_2). \end{aligned} \tag{23}$$

In Eq. (23), P_λ is the Legendre function of order λ [$\lambda = p + (n-3)/2$], which is an integer or half an integer number according to whether the dimension of the space is odd or even. The Chaundy variables x_1, x_2 are defined[22] as follows:

$$\begin{aligned} x_1 &= \frac{(u+s-x-y)(x-y-s+u)}{4xs}, \\ x_2 &= \frac{(x+y-u-s)(x-y+u-s)}{4uy}, \\ Z &= 1-2x_1-2x_2+2x_1x_2. \end{aligned} \tag{24}$$

Along the curve $(x-y=u+s)$ the value of Z is -1. Consequently when λ is an integer (space with odd dimension)

$$R_\lambda(x,y;s,-s+x-y) = P_\lambda(-1) = (-1)^\lambda.$$

The derivative of R_λ with respect to s vanishes and the integral equation has the known reduction

$$\begin{aligned} K_G(x,y) &= \tfrac{1}{2}\int\int_{\partial_G} ds\, R_\lambda(x,y;s,s)V(s) \\ &\quad + \tfrac{1}{2}\int\int_{\partial_G} du\,ds\, R_\lambda(x,y;s,u)V(s)K_G(s,u). \end{aligned} \tag{22c}$$

If λ is half an integer, P_λ is not defined for the argument $Z=-1$. The representation of regular solutions for spaces with even dimension will not be discussed here. More trivially, one can verify that applying

$$L(x) - L(y) - V(x)$$

to both sides of Eq. (22c) gives zero only if

$$\left(\frac{\partial}{\partial s} - \frac{\partial}{\partial u}\right)R(x,y;s,u) = 0 \quad \text{for} \quad u+s = x-y,$$

$$\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right)R(x,y;s,u) = 0 \quad \text{for} \quad x-y = s+u.$$

Such a circumstance happens when R is the Legendre function $P_\lambda(Z)$ and λ is an integer. One can see it by using characteristic variables

$$\eta = x+y, \quad \eta_0 = s+u, \quad \xi = x-y, \quad \xi_0 = s-u,$$

and

$$Z = 1 + \frac{(\xi^2-\xi_0^2)(\eta^2-\eta_0^2)}{(\eta^2-\xi^2)(\eta_0^2-\xi_0^2)},$$

one gets

$$\frac{\partial R_\lambda}{\partial x} + \frac{\partial R_\lambda}{\partial y} = \frac{\partial R_\lambda}{\partial \eta} = \frac{\partial P_\lambda}{\partial \eta} = \frac{\partial P}{\partial Z}\frac{\partial Z}{\partial \eta}.$$

Since

$$\frac{\partial Z}{\partial \eta} = \frac{(\xi^2-\xi_0^2)}{(\xi^2-\xi_0^2)}\,\frac{2\eta[\eta_0^2-\xi^2]}{(\eta^2-\xi^2)^2},$$

$$\frac{\partial R_\lambda}{\partial \eta} = 0 \quad \text{for} \quad \begin{cases} \xi = \xi_0 \text{ or } x-y = s-u, \\ \eta_0 = \xi \text{ or } s+u = x-y. \end{cases}$$

The result of Ref. 2(a) concerning the extension of the representation from the s-wave to the higher l-waves was dependent upon this circumstance. This important fact was not pointed out at the time. We can consider now the two types of representation.

A. Marchenko representation

In the representation of the irregular solutions the argument Z of the Legendre function is greater than one. Since

$$\begin{aligned} Z &= 1-2x_1-2x_2+2x_1x_2 \\ &= 1 + \frac{(u+s-x-y)(s-u+y-x)(u+y+s+x)(u+y-s-x)}{8usxv}, \end{aligned} \tag{25}$$

and since the four factors in the numerator of Eq. (25) are all positive, we have $Z \geqslant 1$.

On the other hand, in the Marchenko domain we have

$$\begin{aligned} 0 \leqslant -x_1 &= \frac{(u+s-x-y)(y-x+s-u)}{4xs} \\ &\leqslant \frac{2(s-x)2(s-x)}{4xs} \leqslant \frac{s}{x} \end{aligned}$$

and

$$0 \leqslant x_2 = \frac{(u+s-x-y)(y-x+s-u)}{4uy} \leqslant \frac{2uy}{4uy} \leqslant \tfrac{1}{2}.$$

We can write

$$Z = 1-2x_1-2x_2(1-x_1) \leqslant 1-2x_1 \leqslant 1+2s/x \leqslant 3s/x.$$

Since $Z > 1$, one can use the Laplace integral representation[23]

$$P_\lambda(Z) = \frac{1}{\pi}\int_0^\pi [Z + (Z^2-1)^{1/2}\cos\phi]^\lambda\, d\phi \tag{26}$$

for all values of λ.

According for all values of n, the upper bound

$$|P_\lambda(Z)| \leqslant \left(\frac{6s}{x}\right)^\lambda$$

holds and can be used in the study of the representation of the irregular solutions.

From the study of Ref. 2(a), the existence of translation kernels results.

Theorem 1: The translation kernel $K(x,y)$ used in the representation of irregular solutions is uniformly bounded, if the two integrals

$$\sigma_1(x)=\int_x^\infty s\,|\bar V(s)|\,ds \quad \text{and} \quad \sigma_\lambda(x)=\int_x^\infty s^\lambda\,|\bar V(s)|\,ds \tag{27}$$

coverge.

In the theorem we have defined $\bar V=(6)^\lambda|V(s)|$. To be precise one obtains the bound

$$|K(x,y)|<\tfrac{1}{2}(1/x)^\lambda\sigma_\lambda\left(\frac{x+y}{2}\right)\exp\sigma_1(x). \tag{28}$$

The existence of K is proved for any λ. Therefore, K exists for any n-dimensional space with $1<n<\infty$.

B. Gel'fand-Levitan representation

In the representation of the regular solution, considering Z defined by

$$Z=1-[(x+y-u-s)(u-s+x-y)(u+y+s+x)\times(s+x-u-y)(8usxy)^{-1}],$$

with four positive factors in the numerator, we find $Z\leq 1$.

On the other hand, since, in the Gel'fand–Levitan domain we have

$$0\leq -x_1=\frac{(x+y-u-s)(x-y+u-s)}{4xs}\leq\frac{2x(x-y)}{4xs}\leq\frac{2x\cdot 2u}{4xs}\leq 1 \tag{29}$$

and

$$0\leq x_2=\frac{(x+y-u-s)(x-y+u-s)}{4uy}\leq\frac{2y\cdot 2u}{4uy}\leq 1, \tag{30}$$

we also can write

$$z=1-2x_2-2x_1(1-x_2)\geq 1-2x_2\geq -1. \tag{31}$$

The argument Z of the Legendre function being between $+1$ and -1, one must set apart the cases where λ is an integer from the cases where λ is half an integer; the latters will not be considered.

Using the method of Ref. 2(b), when λ is an integer we state the theorem.

Theorem 2: If λ is an integer (n odd) and if the moments σ_0 and σ_1 exist, the kernel $K(x,y)$ used for the representation of the regular solution is bounded. One has

$$|K(x,y)|\leq\tfrac{1}{2}\sigma_0\left(\frac{x+y}{2}\right)\exp\sigma_1(x) \tag{32}$$

with

$$\lim_{y\to 0}K(x,y)=0 \quad \text{if } V \text{ is continuous.}$$

The existence of K is proved when the dimension of the space is odd. In addition the integral

$$\int_0^x|K(x,y)|\,dy \tag{33}$$

exists when x goes to infinity.

The present Sec. 3 on uncoupled equations has separated even dimensional from odd dimensional spaces when the potential possesses spherical symmetry. The study already made can be extended without difficulty to the generalized axially symmetric Hamiltonian of Gilbert.[28] For the extension one has simply to consider the Hamiltonian equation

$$\left[\Delta_n+\frac{s}{x_n}\frac{\partial}{\partial x_n}+k^2-V(r)\right]\psi(\mathbf{x})=0, \tag{34}$$

where s is a fixed parameter. The asymptotic behavior of the solution $\psi(\mathbf{x})$ of Eq. (34) is this time

$$\lim_{|r|\to\infty}\psi(\mathbf{x})=\exp(i\mathbf{k}\cdot\mathbf{x})+\frac{f(k,\theta)\exp(ikr)}{r^{1/2(n+s-1)}}+O[r^{-1/2(n+s+1)}]. \tag{35}$$

Equations (34) and (35) together with the method of Sec. 2 lead to the construction of a reduced radial equation,

$$\left(\frac{d^2}{dr^2}+k^2-\frac{p(p+n-2+s)}{r^2}-\frac{1}{4r^2}(n-1+s)(n-3+s)-V(r)\right)\psi(p,k,r)=0. \tag{36}$$

Translation operators can be applied for representing irregular solutions. The Legendre functions to be used in the representation are never polynomials (their order is no longer an integer). The result expressed in Theorem 1 continues to hold provided an appropriate value is given to λ. Needless to say the representation of the regular solutions for this generalized axially symmetric equation is excluded.

4. TRANSLATION KERNEL (MATRIX EQUATIONS)

In Sec. 4, the results of Sec. 3 are extended to matrix differential equations. Solutions, as are customary in the case, are matrices which are built up using vector solutions. After a discussion of an important Riemann solution in Sec. 4 A, the extension to the n-dimensional problem is attacked in Sec. 4 B. Existence theorems are formulated when a many-channel approach is menaingful.

Before proceeding further, we feel it is useful to discuss the introduction of hyperspherical systems in m-body problems.[29,30] Although the n-dimensional space shares with the m-body ($3m=n+3$) the same number degrees of freedom, the hyperspherical scattering solutions do not possess the exact asymptotic form. Furthermore, the kernel of the Lippman–Schwinger equation associated with the m-body problem is not *compact*.[31] Faddeev equations (or any equivalent) are needed to reach compactness.[32] Whether the method of Zachariev[33] and Raynal[34] removes the "dangerous contributions" is not the point here. The fact is that Refs. 33 and 34 construct nonhomogeneous differential systems, unfitted by nature to linear translation operators. Consequently the m-body problem is not discussed here except when it reduces trivially to a many-channel problem.[35] We will allow, however the number of channels to become infinite.

A. The Riemann solution $C(l, m)$

The solution $C(l,m)$ is constructed because of its use in the following subsection.

We consider the two operators

$$A \equiv \left[\frac{d^2}{dx^2} + k^2 - l^2\right] - V(l,m,x),$$
$$A_0 \equiv \left[\frac{d^2}{dx^2} + k^2 - l^2\right], \tag{37}$$

where l^2 may be a continuous or simply a discrete threshold energy. The partial differential equations for the continuous "matrix" $K(l,m;x,y)$ are

$$\frac{\partial^2}{\partial x^2}K(l,m;x,y) - l^2K(l,m;x,y) + m^2K(l,m;x,y)$$
$$= \frac{\partial^2}{\partial y^2}K(l,m;x,y) + \int V(l,n,x)K(n,m;x,y)\,dn. \tag{38}$$

In the case of a discrete index the integration is replaced by an ordinary summation.

To Eq. (38) boundary conditions are to be added. We are led to consider the Riemann solutions for the

$$\left[\frac{\partial^2}{\partial x^2} - (l^2 - m^2)\right]C(l,m;x,y) = \frac{\partial^2}{\partial y^2}C(l,m;x,y). \tag{39}$$

These Riemann solutions which we briefly denote $C(l,m)$ are given in Ref. 2(b) where we followed Ref. 36. In characteristic variables they are

$$C(l,m) \equiv J_0[-(4v)^{1/2}] \tag{40}$$

with

$$v = (l^2 - m^2)(\xi - \xi_0)(\eta - \eta_0)$$
$$= (l^2 - m^2)[(x-s)^2 - (y-u)^2].$$

Since the inequality $(x-s)^2 > (y-u)^2$ holds for both the Marchenko and the Gel'fand—Levitan representation, one has

$$C(l,m) = I_0(2\sqrt{v})$$

and therefore

$$|C(l,m)| \leq 1 \quad \text{for} \quad |m| \geq |l|, \tag{41}$$
$$|C(l,m)| \leq \exp 2\sqrt{v} \quad \text{for} \quad |m| \leq |l|. \tag{42}$$

From Eqs. (41) we can obtain the bounds to be used later:

Marchenko representation: Since $s > x$, we have

$$|C(l,m)| < \exp 2|l|s, \quad \forall\, m. \tag{43}$$

Gel'fand—Levitan representation: One can use $u + y \geq x - s > 0$,

$$v < (u+y)^2(l^2 - m^2),$$
$$|C(l,m)| \leq \exp 2|l|u \exp 2|l|y, \quad \forall\, m. \tag{44}$$

B. The n-dimensional space

One expands the wavefunction in hyperspherical variables,

$$\psi(\mathbf{x}) = \sum_{\nu\alpha} P_{\nu\alpha}(\Omega)R_\nu(r)\frac{1}{r^{(n-1)/2}}. \tag{45}$$

In Eq. (45) we included in the hyperspherical definition the "order" ν and the set α of all the "quantum numbers" necessary for their definition. See Ref. 30(d).

Noting $(\nu,\alpha) \equiv i$ and $(\mu,\beta) \equiv j$, we define the elements of the matrix potential between the i and j "channels" by the integrals

$$V_{ij}(r) = \int P_{\nu\alpha}(\Omega)V(x)P_{\mu\alpha}(\Omega)\,d\Omega. \tag{46}$$

Recalling the definition

$$\lambda_1 = p_i + (n-3)/2, \tag{47}$$

we introduce *a priori* the system of coupled equations,

$$L_i(x)u_i(x) + V_{ii}(x)u_i(x)$$
$$= \left[\frac{d^2}{dx^2} + k_i^2 - \frac{\lambda_i(\lambda_i+1)}{x^2} + V_{ii}(x)\right]u_i(x)$$
$$= \sum_{i\neq j} V_{ij}(x)u_j(x). \tag{48}$$

In Eq. (48) one has

$$k_i^2 = (E - E_i)\frac{\hbar^2}{2M}. \tag{49}$$

In Eq. (49), E is the incident energy and E_i is the threshold energy of the ith "channel." The case where all the E_i are set equal to zero (all the k_i^2 are equal to k^2) is discussed first. The extension to the general case which makes use of Sec. 4 A follows.

In the physical applications the system of equations (48) may be infinite. Concerning infinite systems, the question of the existence of translation kernels may be raised. In the following the conditions for the existence of translation kernels are stated first when the order of the system is finite. We indicate hereafter how the conditions should be supplemented when infinite systems are considered.

To obtain our results we use the Riemann solutions $R_{ij}(x,y;s,u)$. They satisfy the equations:

$$L_i(x)R_{ij}(x,y;s,u) = L_j(y)R_{ij}(x,y;s,u), \tag{50}$$

$$R_{ij}(x,y;s,u)$$
$$= \Pi\begin{pmatrix}\lambda_i, \lambda_j \\ x_1, x_2\end{pmatrix}$$
$$= P_{\lambda_i}(1-2x_1) - 2x_2\int_0^1 P_{\lambda_i}(1-2x_1+2x_1t)$$
$$\times P'_{\lambda_j}(1-2x_2t)\,dt$$
$$= P_{\lambda_j}(1-2x_2) - 2x_1\int_0^1 P_{\lambda_j}(1-2x_2t)$$
$$\times P'_{\lambda_i}(1-2x_1+2x_1t)\,dt, \tag{51}$$

where x_1, x_2 have already been defined in Eqs. (24).

Again we separate the representation of solutions defined by a condition at infinity (Marchenko representation) from the representation of solutions defined by a condition at the origin (Gel'fand—Levitan representation). When the latter representation is considered, along the segment $u + s = x - y$, x_2 equals 1. In the last equation, (51), P_{λ_j} is defined only if λ_j is an integer, that is if n is odd; the representation of regular solutions for even n is again therefore excluded.

(a) *Marchenko representation*: The representation of irregular matrix solutions is discussed for even and odd dimensional spaces. Since we have $0 \leq -x_1 \leq s/x$

and $0 \leq x_2 \leq \frac{1}{2}$, all the arguments of the Legendre functions which appear in the last of the Eqs. (51), are positive. Consequently, we can use the bounds derived from the Laplace integral. This can be done whether the dimension of the space is odd or even. Following the methods used in Ref. 2(b), we obtain

$$|R_{ij}| \leq \left(\frac{7s}{x}\right)^{\lambda_i}. \tag{52}$$

We introduce now the notation

$$\begin{aligned}
&\tilde{V}_{ij}(s) = (7s)^{\lambda_i} V_{ij}(s),\\
&W_{ij}(s) = (7s)^{\lambda_i} V_{ij}(s) \frac{1}{(s)^{\lambda_j}},\\
&\eta\left(\frac{x+y}{2}\right) = \int_{(x+y)/2}^{\infty} \|\tilde{V}(s)\|\, ds,\\
&\xi(x) = \int_x^{\infty} s\, \|W(s)\|\, dx,\\
&D_{ij}(x) = x^{\lambda_i}\delta_{ij}
\end{aligned} \tag{53}$$

and we can state the existence theorem when the system (48) is finite and all the k_i^2 are set equal.

Theorem 3: If the system considered is finite and if all k_i^2 are equal, and if in addition the two integrals $\sigma_y^{(0)}$ and $\sigma_y^{(1)}$ converge as follows:

$$\begin{aligned}
&\sigma_{ij}^{(0)}(x) = \int_x^{\infty} t^{\lambda_i} |V_{ij}(t)|\, dt < \infty,\\
&\sigma_{ij}^{(1)}(x) = \int_x^{\infty} t\, |V_{ij}(t)|\, t^{|\lambda_i - \lambda_j|}\, dt < \infty,
\end{aligned} \tag{54}$$

a translation kernel for the n-dimensional irregular matrix solutions exists. To be precise one has

$$D^{-1}(x)\|K(x,y)\| \leq \tfrac{1}{2}\eta\left(\frac{x+y}{2}\right) \exp\xi(x). \tag{55}$$

By Theorem 3, a restriction has been set on the non-diagonal elements of the matrix potential $V_{ij}(x)$. In addition to this restriction, the theorem requires the existence of absolute moments for the matrix $V_{ij}(x)$. When the order of the matrices introduced in the proof is finite, these matrices have a finite norm but this is not necessarily so when their order is infinite. Then one needs additional requirements. First one must assume

$$\|V\| = \sup_i \sum_{j=1} |V_{ij}| < \infty, \tag{56}$$

so that V possesses a finite norm. Afterwards one should require the existence of the following limits:

$$\lim_{i,j\to\infty} \int_x^{\infty} t^{|\lambda_i - \lambda_j| + 1} |V_{ij}(t)|\, dt < \infty, \tag{57}$$

$$\lim_{i,j\to\infty} \int_x^{\infty} t^{\lambda_i} |V_{ij}(t)|\, dt < \infty. \tag{58}$$

The exponents (λ_i, λ_j) of Eqs. (57) and (58) represent the angular momenta of the i and j channels. They may remain finite even if the number of channels become infinite.

In the same way as the norm for V was defined in Eq. (55), norms for the infinite matrices V and W, Eq. (53), have to be introduced and should be assumed to exist. Writing

$$\xi(x) = \int_x^{\infty} s\, \|W(s)\|\, ds < \infty, \tag{59}$$

$$\eta(x) = \int_x^{\infty} \|V(s)\|\, ds < \infty, \tag{60}$$

and assuming inequalities (57), (58), (59), and (60) are satisfied one takes into account the possibility for i and j to assume infinite values. In this way the upper bound expressed by Eq. (55) remains finite. Now according to Eqs. (48) and (49) threshold energies are included. Assuming the existence of an upper bound K on the set of positive numbers K_i^2 defined by

$$K_i^2 = -k_i^2, \tag{61}$$

and using the Riemann solutions $C(K_i, K_j)$ of Eq. (43), the existence of a translation kernel can be proved in the extended case where the k_i^2 are different. The extension is obtained at the price of requiring an additional exponential decrease of the matrix potential; the measure of which is expressed by

$$4 \sup_{i,j} (|k_i^2 - k_j^2|)^{1/2}. \tag{62}$$

As one realizes, the system of operators introduced in Eq. (37) is a discrete system.

(b) *Gel'fand–Levitan representation*: We recall a first restriction. No consideration is given to even dimensional spaces. The Riemann solutions introduced in the possible integral representation are defined in Eqs. (44) and (51). When $x_2 = 1$ these $R_{ij}(x,y;s,u)$ are no longer constant. The contribution of the segment $u+s = x-y$ to the integral representation has to be included. We are therefore obliged to consider the full integral equations, namely

$$\begin{aligned}
K_{ij}(x,y) = {}& \tfrac{1}{2}\int_{(x-y)/2}^{(x+y)/2} R_{ij}(x,y;s,s) V_{ij}(s)\\
&+ \tfrac{1}{2}\iint_{\partial_G} R_{ij}(x,y;s,u) \sum_k V_{ik}(s)\\
&\times K_{kj}(s,u)\, du\, ds\\
&+ \int_{(x-y)/2}^{(x-y)} \frac{d}{ds}[R_{ij}(x,y;s,-s+x-y)]\\
&\times K_{ij}(s,-s+x-y)\, ds.
\end{aligned} \tag{63}$$

In this form Eq. (63) does not seem suitable for the method of successive approximations. A variant is given which is obtained by integrating the last term of Eq. (63) by parts. Denoting this last term by I, one gets

$$\begin{aligned}
I = {}& R_{ij}(x,y;x-y,0) K_{ij}(x-y,0) - R_{ij}\left(x,y;\frac{x-y}{2},\frac{x-y}{2}\right)\\
&\times K_{ij}\left(\frac{x-y}{2},\frac{x-y}{2}\right) - \int_{(x-y)/2}^{(x-y)} R_{ij}(x_1,x_2)\\
&\times \frac{d}{ds} K_{ij}(s,-s+x-y)\, ds.
\end{aligned} \tag{64}$$

Using the boundary conditions $K_{ij}(x,y)$ has to satisfy, Eq. (64) reduces to

$$\begin{aligned}
I = {}& -\tfrac{1}{2} R_{ij}\left(x,y;\frac{x-y}{2},\frac{x-y}{2}\right)\int_0^{(x-y)/2} V_{ij}(s)\, ds\\
&- \int_{(x-y)/2}^{(x-y)} R_{ij}(x_1,x_2)\frac{d}{ds} K_{ij}(s,-s+x-y)\, ds.
\end{aligned} \tag{65}$$

Equations (64) and (65) suggest that under proper conditions of convergence the kernel $K_{ij}(x,y)$ may exist in very general circumstances. When all the "channels" are coupled at the same angular momentum λ_i and all have a zero threshold energy the value of the term I is exactly zero and the contribution of the segment $u+s+x-y$, to Eq. (63) disappears. Then under the simple conditions that the matrix potential possesses finite moments of order zero and order one, a translation kernel exists. Due to the difficulty of providing the existence of $K(x,y)$ for the most general form of Eq. (63), no more consideration is given to the Gel'fand–Levitan representation. From now on we will consider only the Marchenko method.[37]

(c) *Physical application*: A physical application concerning the many-channel case for the n-dimensional scattering problem is now discussed. The basic idea is to solve the A-body Hamiltonian

$$H_A\phi_n(\zeta)=\epsilon_n\phi_n(\zeta),$$

before considering the $(A+1)$ Hamiltonian H_{A+1}. Its solutions $\phi_n(\zeta)$ are separated in radial and angular coordinates using hyperspherical variables

$$\phi_n(\zeta)=\sum_{m'} P_{m'}(\Omega)u_{m'n}(\zeta).$$

Channels can be defined by coupling target-generalized angular momenta m' to the incident projectile ones m'' to a total angular momentum m,

$$[P_{m''}(\Omega)\times P_{m'}(\Omega)u_{m'n}(\zeta)]_m.$$

It is assumed, as is the case in three-dimensional nuclear problems that the target momenta m' are finite. The projectile angular momenta m'' which can be coupled to some m' to construct m (Clebsch–Gordon generalizations) are finite. The set of λ_i's present in Eq. (48) possesses therefore an upper bound. The left-hand side of expressions Eqs. (57), (58), (59), and (60) can be finite even under the assumption of an infinite number of channels.

If we assume, in addition, that closed channels can be safely neglected if their respective threshold energies are large[6,38] and the constants l and m which appear in the bounds of Sec. 4A for the Riemann solution $C(l,m)$ are themselves bounded.

Theorem 3 can be used to assert the existence of a Marchenko kernel provided the matrix potential has, in addition to the conditions it expresses, adequate exponential decrease. See Eq. (62).

[1]M. Coz and C. Coudray, J. Math. Phys. **14**, 1574–1978 (1973).
[2](a) M. Coz and C. Coudray, J. Math. Phys. **17**, 888–93 (1976); (b) **17**, 894–9 (1976).
[3]M. Coz, J. Math. Anal. Appl. **60** (1977) in press.
[4]M. Blazech, Commun. Math. Phys. **3**, 282–91 (1966).
[5]G. Helwig, *Partial Differential Equation* [Glinn (Blaisdell, Boston, 1960)], Chap. 2.
[6]R.B. Gerber and N.C. Rosenbach, Phys. Rev. A **9**, 305–15 (1974).
[7]L.M. Delves, (II), Nucl. Phys. **20**, 275–308 (1960).
[8]T.A. Green and I.E. Lanford, III, J. Math. Phys. **1**, 139 (1960).
[9]S.T. Kuroda, J. Math. Phys. **3**, 933 (1962).
[10]R. P. Gilbert, *Function Theoretic Methods in Partial Differential Equations* (Academic, New York, 1963), p. 262.
[11]A. Sommerfeld, Jahresber. Deut. Math. Ver. **21**, 309–53 (1912).
[12]F. Rellich, Jahresber. Deut. Math. Ver. **53**, 57–65 (1943).
[13]J.M. Jauch, *Dispersion Relation*, edited by G.R. Screaton (Oliver and Boyd, Edinburgh, 1961), p. 283.
[14]A. Erdelyi, "Higher Transcendental Functions," Bateman manuscript II, p. 177.
[15]A. Erdelyi, Bateman manuscript II, p. 174.
[16]M. Fabre de la Ripelle, IPNO/TH 57, Orsay, pp. 6 and 16.
[17]Ref. 10, p. 289.
[18]Ref. 14, p. 85.
[19]Ref. 10 contains typographical errors which are corrected here.
[20](a)R.G. Newton, J. Math. Phys. **1**, 322–3 (1960); (b) *The Scattering of Waves and Particles* (McGraw-Hill, New York, 1966), p. 372.
[21](a) N. Levinson, Det. Kgl. Danske Videnskab. Selskab., Mat.-fys. Medd. **25**, n°9 (1949); R.G. Newton, Phys. Rev. **100**, 412 (1955).
[22]T.W. Chaundy, Q.J. Math. Oxford Ser. **10**, 219–40 (1939).
[23]P.S. Laplace, *Mecanique celeste*, Tome V, Livre XL, Chap. II, p. 41 (1825).
[24]E.W. Hobson, *Spherical and Ellipsoidal Harmonics* (Chelsea, New York, 1955), pp. 252 and 254.
[25]Ref. 2, p. 895–6.
[26]W. Magnus, F. Oberhettinger, and R.P. Soni, *Formulae and Special Functions of Mathematical Physics* (Springer-Verlag, New York, 1966).
[27]E.W. Hobson, p. 310.
[28]Ref. 10, p. 264
[29]M. Fabre de la Ripelle, IPNO/TH 71-13, Orsay, 1971.
[30](a) N.A. Vilenkin, G.I. Kuznetsov, and Ya.A. Somorodinskii, Sov. J. Nucl. Phys. **2**, 645–52 (1966); (b) Yu.A. Simonov, Sov. J. Nucl. Phys. **3**, 461–6 (1966); (c) A.M. Badalyan and Yu.A. Simonov, Sov. J. Nucl. Phys. **3**, 755–64 (1966); (d) V.V. Pustovalov and Yu.A. Simonov, Sov. Phys. JETP **24**, 230–39 (1967); (e) H. Letz, Nuovo Cimento B **26**, 522–36 (1975); (f) H. Letz, Nuovo Cimento B **30**, 348–63 (1975).
[31]Ref. 20(b), p. 555.
[32]L.D. Faddeev, "Mathematical aspects of the three-body problem in the quantum scattering theory," IPST Jerusalem (1965).
[33]B.N. Zachariev, V.V. Pustovalov, and V.D. Efros, Sov. J. Nucl. Phys. **8**, 234–8 (1969).
[34]J. Revai and J. Raynal, D.Ph. 71/40, Centre d'Etudes Nucleaires de Saclay, France.
[35]M. Fabre de la Ripelle and J.S. Levinger, Nuovo Cimento A25, 555–70 (1975).
[36]G. Valiron, *Equations fonctionnelles* (Masson, Paris, 1950), p. 562.
[37]L.S. Agranovitch and V.A. Marchenko, *The Inverse Problem in Scattering Theory* (Gordon and Breach, New York, 1963).
[38]R.G. Newton, Ann. Phys. (N.Y.) 4, 29–56(1958), see p. 32.

Translation kernels for velocity dependent interactions

M. Coz

Department of Physics and Astronomy, University of Kentucky, Lexington, Kentucky 40506

P. Rochus[a)]

Physique Nucléaire Théorique, Université de Liège, Institut de Physique au Sart Tilman, B—4000 Liège 1, Belgium

(Received 9 March 1977)

We reconsider first the representation of solutions defined by a condition at the origin and recognize the difficulty of extending the representation outside simple cases. By eliminating the study of solutions defined at the origin in further studies the translation kernels for velocity dependent interactions are constructed only for solutions defined by their behavior at infinity. Two methods are proposed. Their domains of extension are compared and shown to be different.

1. INTRODUCTION

The translation kernels were brought to consideration in the early 1950's within the framework of the inverse scattering problem.[1] Their use by Gel'fand—Levitan[2] and Marchenko[3] was decisive in the mathematical solution of the inverse problem for systems without singularities; Ref. 3 contains, in addition to this mathematical solution, an indirect attempt to solve a particular case of a system with singularity, the particular case which appears in the deuteron problem with a tensor for force. Almost at the same time, within his theory of perturbations, Friedrichs[4] introduces the idea of similitude between operators.

Let A_1 and A_2 be two operators with the same domain. If there exists an operator U with a bounded inverse U^{-1} such that the equation

$$UA_1 = A_2U \tag{1}$$

holds, the operators A_1 and A_2 will be called similar and U a similitude transformation. The existence of U^{-1} leads to the equation

$$UA_1U^{-1} = A_2. \tag{2}$$

If A_1 and A_2 operate on a Hilbert space and are both self-adjoint, Eq. (2) leads to the equation

$$UU^{-1}A_1U^*U = A_1 \tag{3}$$

from which it follows that U is a unitary transformation. The point is important when the spectra of A_1 and A_2 are compared.

The translation operators satisfy Eq. (1); in addition we require they be integral transformations and belong to the category of Volterra operators.

Recently the Clarkson school[5] while developing an idea of Lax[6] constructed a class of nonlinear equations whose solution is connected to that of an inverse scattering problem. By so doing they renewed interest in the inverse problem and it becomes more compelling to specify the class of equations for which a translation kernel exists.

In a first work on the inverse problem,[7] the theory of translation operators was re-examined so as to use them in the inverse problem at fixed energy. Recently, we moved to problems at fixed angular momentum with singularities but we restricted our investigation to differential operators and to systems of coupled differential operators of the form

$$\frac{d^2}{dx^2}\, I + V(x) = \lambda, \tag{4}$$

where I is the unit matrix and λ is a constant diagonal matrix.[8-10] The matrix $V(x)$ called potential was separated into

$$V(x) = V_0(x) + V_N(x) \tag{5}$$

with a reference potential $V_0(x)$ and a nuclear potential $V_N(x)$. In most of the cases we took the reference potential to be

$$V_0(x) = \frac{\alpha(\alpha+1)}{x^2} + \frac{\eta}{x}. \tag{6}$$

When V_0 is defined by Eq. (6), it contains a centripetal and a Coulomb part, the usual singularities of the nuclear problem.

In each case we examined, the task was to specify the conditions the nuclear potential $V_N(x)$ has to satisfy for a bounded translation kernel to exist. The conditions were, of course, sufficient conditions. However, the method of proof used for their specification was a *constructive* method. The translation kernel was, in principle, constructed. It is a general remark that sufficient conditions obtained by a constructive method are hard to improve. The construction was obtained by transforming partial differential equations of hyperbolic type into Volterra integral equations which incorporate their boundary equations. This latter transformation has, in our opinion, its own importance; therefore another distinct example which uses different boundary conditions is included in this present paper. Extensions of the transformations may be obtained by using the appropriate elementary solutions which are here the Riemann solutions.

To specify the notations and the object of this present paper, we give the following definitions. Let A_0 and A_1 be two (systems of) differential operators together with

[a)]Chercheur de l'Institut Interuniversitaire des Sciences Nucléaires.

 Copyright © 1977 American Institute of Physics

boundary conditions necessary to specify the solutions their respective equations may possess.

The two types of boundary conditions we studied were conditions of regularity at the origin or conditions of behavior at infinity. The new case discussed in the present paper incorporates conditions at the origin which do not involve regularity. The two first types of conditions are the usual conditions considered in scattering problems. A translation operator X is defined as a bounded operator with an inverse X^{-1},

$$X = 1 + K, \quad X^{-1} = 1 + \tilde{K}.$$

The operator K of the definition is always an integral operator. As a consequence of the definition, X transforms the solutions of the differential operator A_0 into the solutions of the operator A_1. The operator A_1 was written as $A_0 + V_N$; A_0 contains the reference potential which may be zero. If the solutions thus transformed are the regular solutions, we call the kernel K a Gel'fand—Levitan kernel: The notation of the Gel'fand—Levitan representation is retained here, but it is extended here to any representation of solutions specified by a condition at the origin. In the case of solutions defined by their behavior at infinity, we will have a Marchenko kernel and a Marchenko representation.

In all the cases we investigated, the conditions were expressed in terms of conditions the nuclear potential V_N should satisfy; whether it was attractive or repulsive was unimportant. Its strength also was irrelevant. The conditions were dependent upon the *absolute* moments of the potential V_N as are the usual conditions for the existence of solutions of the Schrödinger equation. It is important to note that they were, in addition, dependent upon the choice of the reference potential.

Systems of coupled differential operators defined by Eq. (4) belong to a restricted class of operators. However when, for the first time, Levitan[11] defined the concept of translation operation, he stated its use for differential operators of the more general form:

$$a(x)\frac{d^2}{dx^2} + b(x)\frac{d}{dx} - V(x) + \lambda. \tag{7}$$

The extension was obtained, he said through a straightforward extension of the Liouville transformation.[12] As we will see later, he was a little optimistic; the Liouville transformation cannot be used directly except for the $b = 0$ cases, but the idea of Levitan can be pursued and a transformation shown to exist. Already a desire for such an extension appeared in Ref. 7. At that time it was simply recognized that the conditions for existence depend upon the solution of a Cauchy problem but this quest was not pursued. Now many years have passed, progress has been made; not only do we possess, in the Riemann method, a tool for specifying the conditions a nuclear potential has to satisfy, but also we possess the motivation for pursuing the physical application of the extension which once seemed far away and has now become real. While the forms of operators discussed in Ref. 5 are related to the problem of the string amplitude,[13,14] a subject dear to dual theorists,[15] the Sturm—Liouville operators for Eq. (7) contain the differential operators with velocity dependent interactions.[16] The latter have been shown to be equivalent to static interactions with a hard core.[17] After additional work on the inverse problem by Marchenko[18] and Faddeev,[19] Zachariev[20] had just formulated the inverse problem for potentials which depend on the velocity. He has, in addition, proposed a solution for the approximation which involves the replacement of the differential operator by a difference operator.

Attention cannot be restricted however to differential operators; systems of coupled differential operators described by Eq. (8) have to be treated. The extension to such systems is actually realized in the present paper.

The reader is warned that the advice of Levitan of using a Liouville transformation is not followed all the time; we found it more convenient to use a simpler transformation, precisely the one advocated for the construction of equivalent local potentials in Ref. 21, In Sec. 2 the Riemann solutions method is reviewed and a new representation for a solution which is not regular but is defined by a condition at the origin is given for the $l = 0$ case.

In Sec. 3, systems of coupled equations with velocity dependent interactions are discussed. The systems are transformed into simpler ones prior to being subjected to the translation operators techniques. The use of a more direct method is discussed in the conclusion.

2. INTRODUCTION OF THE RIEMANN'S SOLUTION

Before proceeding to this introduction, we consider the construction of an irregular solution for a local Schrödinger equation defined by a condition at the origin. For this construction we consider the Volterra equation

$$\xi_l(k,x) = v_l(k,x) + \int_0^x g_l(k;x,y)\, V(y)\, \xi_l(k,y)\, dy \tag{8}$$

with

$$g_l(k;x,y) = [u_l(k,x)\, v_l(k,y) - u_l(k,y) \times v_l(k,x)]/W_r(u_l, v_l). \tag{9}$$

In the Green function g_l of Eq. (9) the Riccati—Bessel u_l and v_l are inserted together with their Wronskian. One has

$$\lim_{x\to 0} v_l(k,x) = \frac{1}{(2l-1)!!\,(kx)^l},$$

$$\lim_{x\to 0} u_l(k,x) = \frac{(kx)^{l+1}}{(2l+1)!!}.$$

Obviously, Eq. (8) is meaningless for $l \neq 0$ except for special classes of interactions.

We will limit ourselves to the $l = 0$ (s-wave) case. Then Eq. (8) reads

$$\xi_0(k,x) = \cos kx + \int_0^x \frac{\sin k(x-y)}{k}\, V(y)\, \xi_0(k,y)\, dy. \tag{10}$$

Together with Eq. (10), we consider the possibility of representing $\xi_0(k,x)$ by an integral

$$\xi_0(k,x) = \cos kx + \int_0^x K_+(x,y)\cos ky\, dy. \tag{11}$$

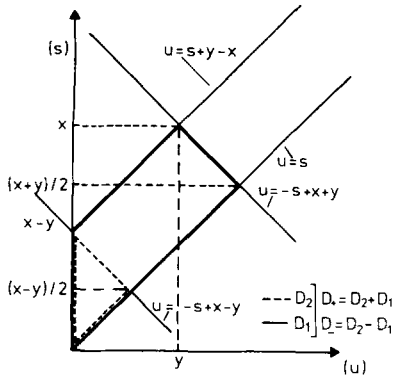


FIG. 1.

In order to exist, the kernel $K_+(x,y)$ has to satisfy the partial differential equation

$$\left(\frac{\partial^2}{\partial x^2}-\frac{\partial^2}{\partial y^2}\right) K_+(x,y)-V(x)K_+(x,y)=0, \tag{12}$$

$$K_+(x,x)=\tfrac{1}{2}\int_0^x V(s)\,ds, \tag{13}$$

$$\frac{\partial}{\partial y}K_+(x,y)\big|_{y=0}=0. \tag{14a}$$

In order to prove the existence of $K_+(x,y)$ one has to transform Eqs. (12), (13), and (14a) into a Volterra equation.

This can be done by using the identity

$$\frac{\sin k(x-s)}{k}\cos ku=\frac{1}{2}\int_{u+s-x}^{u+x-s}\cos kt\,dt. \tag{15}$$

After trivial interchanges of variables valid if $y\geq x/3$ and which are justified by the Tonelli—Fubini's theorems, one obtains

$$\begin{aligned}K_+(x,y)=&\tfrac{1}{2}\Big[\int_0^{(x+y)/2}V(s)\,ds+\int_0^{(x-y)/2}V(s)\,ds\Big]\\&+\tfrac{1}{2}\Big[\int_{(x+y)/2}^{x}V(s)\,ds\int_0^{x-s+y}K_+(s,u)\,du\Big]\\&+\tfrac{1}{2}\Big[\int_{(x-y)/2}^{(x-y)}V(s)\,ds\int_0^{x-s-y}K_+(s,u)\,du\\&+\int_0^{x-y}V(s)\,ds\int_0^{s}K_+(s,u)\,du\\&+\int_0^{(x-y)/2}V(s)\,ds\int_0^{s}K_+(s,u)\,du\\&-\int_{(x+y)/2}^{x}V(s)\,ds\int_0^{s}K_+(s,u)\,du\\&+\int_{x-y}^{(x+y)/2}V(s)\,ds\int_{y+s-x}^{s}K_+(s,u)\,du\Big].\end{aligned} \tag{16}$$

As we said in the introduction we decide to call this representation a Gel'fand—Levitan representation, denoting $K_\pm(x,y)$ the kernels of the two Gel'fand—Levitan representations. Equation (16) of this paper and Eq. (16) of Ref. 8 can be summarized as follows:

$$\begin{aligned}K_\pm(x,y)=&\tfrac{1}{2}\Big[\int_0^{(x+y)/2}V(s)\,ds\pm\int_0^{(x-y)/2}V(s)\,ds\Big]\\&+\tfrac{1}{2}\Big[\int_{(x+y)/2}^{x}V(s)\,ds\int_0^{x-s+y}K_\pm(s,u)\,du\\&\pm\int_{(x-y)/2}^{x-y}V(s)\,ds\int_0^{x-s-y}K_\pm(s,u)\,du\\&+\int_0^{x-y}V(s)\,ds\int_0^{s}K_\pm(s,u)\,du\pm\int_0^{(x-y)/2}V(s)\,ds\\&\times\int_0^{s}K_\pm(s,u)\,du-\int_{(x+y)/2}^{x}V(s)\,ds\int_0^{s}K_\pm(s,u)\,du\\&+\int_{(x-y)}^{(x+y)/2}V(s)\,ds\int_{y+s-x}^{s}K_\pm(s,u)\,du\Big].\end{aligned} \tag{17}$$

K_- must satisfy the same equations (12) and (13) as K_+ but the condition (14a) is replaced by

$$K_-(x,y)\big|_{y=0}=0. \tag{14b}$$

At this point it is interesting to visualize the two domains. Let D_1 and D_2 be the two domains described in Fig. 1,

$$D_\pm=D_1\pm D_2,$$

so Eq. (17) reads

$$\begin{aligned}K_\pm(x,y)=&\tfrac{1}{2}\Big[\int_0^{(x+y)/2}V(s)\,ds\pm\int_0^{(x-y)/2}V(s)\,ds\Big]\\&+\tfrac{1}{2}\Big[\int\int_{D_1}V(s)K(s,u)\,du\,ds\pm\int\int_{D_2}V(s)K(s,u)\,du\,ds\Big].\end{aligned} \tag{18}$$

Retaining the same segments $(u=0, u=s, u+s=x+y, u-s=y-x)$, $(u=0, u=s, s+y=x-y)$ as boundary, D_1 and D_2 can be defined for $0\leq y\leq x$. With this extension, one can prove that Eq. (18) is quite general and valid for $0\leq y\leq x$.

Using the method of successive approximations, the reader can verify that the two kernels K_+ and K_- exist under the conditions that the local potential V possesses absolute moments of order zero and of order one. Interest in the two kernels K_+ and K_- has been displayed by Mehta[22] and more recently by Dyson[23] but until now no integral representation has been given for both these kernels. Before proceeding further a discussion of the Gel'fand—Levitan domains is in order. Both are bounded at finite distance and are built up with more than three segments. When one desires to extend the representation to a reference potential $p(x)\neq 0$, the Riemann solution R for the operator

$$\frac{\partial^2}{\partial x^2}-\frac{\partial^2}{\partial y^2}+p(x)-p(y)$$

enters into the picture. Then the Riemann—Green formula for the domain D and its contour C_+ is used,

$$\int_{C_+}\left[R\left(\frac{\partial}{\partial u}K\,ds+\frac{\partial}{\partial s}K\,du\right)\right]-\left[\frac{\partial}{\partial u}R\,ds+\frac{\partial}{\partial s}du\right]K=\iint_D RVK(s,u)\,du\,ds.$$

Two segments C_+ are characteristics, a third $u=s$ carries the boundary condition $K(x,x)$. Consequently, except in special circumstances, the reduction

$$\begin{aligned}K_\pm(x,y)=&\tfrac{1}{2}\Big[\int_0^{(x+y)/2}RV\,ds\pm\int_0^{(x-y)/2}RV\,ds\Big]\\&+\tfrac{1}{2}\Big[\int\int_{D_1}RVK\,du\,ds\pm\int\int_{D_2}RVK\,du\,ds\Big]\end{aligned}$$

will not result. Such a special circumstance occurs when the reference potential is the scalar centripetal barrier

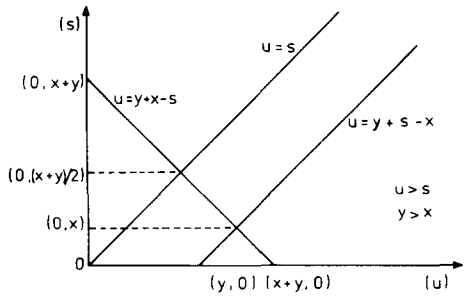


FIG. 2. Marchenko domain.

$$p(x) \equiv -l(l+1)/x^2$$

and when the K_- kernel is considered.

Because of this fundamental deficiency of the two Gel'fand–Levitan representations only the Marchenko representation is considered in the rest of the present paper.

Although Riemann's method can be introduced, as it has been done by Riemann himself, a simpler heuristic presentation is given here. In the following, D denotes Marchenko's domain (Fig. 2).

Let L denote the *scalar* partial differential operator,

$$L = \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \{p(x) - p(y)\} \tag{19}$$

and finally let $R(x, y; s, u)$ be Riemann's solution for Eq. (20a), which follows:

$$\begin{aligned} &LR = 0, \\ &\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right) R = 0 \quad \text{if } y - x = u - s, \\ &\left(\frac{\partial}{\partial x} - \frac{\partial}{\partial y}\right) R = 0 \quad \text{if } y + x = u + s, \\ &R(s, u; s, u) = 1. \end{aligned} \tag{20}$$

From the last three Eqs. (20) it follows that

$$R(x, y; s, u) = 1 \quad \text{if } y - x = u - s.$$

One can now prove the two following equalities using Eqs. (19) and (20):

$$L \int\int_D R(x, y; s, u)\, W_1(s, u)\, du\, ds = 2 W_1(x, y), \tag{21}$$

$$L \int_{(x+y)/2}^{\infty} R(x, y; s, s)\, W_2(s)\, ds = 0. \tag{22}$$

The conditions for Eqs. (21) or (22) to be valid are simply the usual conditions for the differentiation under the integral sign.

In addition to Eqs. (21) and (22), provided that the convergence of the integrals is uniform, the following limits exist and have a common value which is zero:

$$\lim_{y \to \infty} \int\int_D R(x, y; s, u)\, W_1(s, u)\, du\, ds, \tag{23}$$

$$\lim_{y \to \infty} \int_{(x+y)/2}^{\infty} R(x, y; s, s)\, W_2(s)\, ds, \tag{24}$$

$$\lim_{y \to x} \int\int_D R(x, y; s, u)\, W_1(s, u)\, du\, ds. \tag{25}$$

In addition to these zero-value limits one has

$$\lim_{y \to x} \int_{(x+y)/2}^{\infty} R(x, y; s, s)\, W_2(s)\, ds = \int_x^{\infty} W_2(s)\, ds. \tag{26}$$

The Marchenko integral representation for the translation kernel is obtained as a consequence of these equations. Let us consider the partial differential equation the kernel has to satisfy:

$$\begin{aligned} &LK(x, y) = V(x)\, K(x, y), \\ &\lim_{y \to \infty} K(x, y) = 0 = \lim_{y \to \infty} \frac{\partial}{\partial y} K(x, y), \\ &K(x, x) = \tfrac{1}{2} \int_x^{\infty} V(s)\, ds. \end{aligned} \tag{27}$$

In view of Eqs. (21)–(26), Eq. (27) is equivalent to the Volterra integral equation

$$\begin{aligned} K(x, y) = {} & \tfrac{1}{2} \int_{(x+y)/2}^{\infty} R(x, y; s, s)\, V(s)\, ds \\ & + \tfrac{1}{2} \int\int_D R(x, y; u, s)\, V(s)\, K(s, u)\, du\, ds. \end{aligned} \tag{28}$$

Since we emphasized the dependence of the conditions on the nuclear potential on the choice of the reference potential, we report the following statement.

If the two potentials V_1 and V_2 belong to the class of acceptable potentials for the Marchenko representation, the reference potential V_0 being chosen, then the potentials $V_1 \pm V_2$ are members of the same class. We write $V_1, V_2 \in C(V_0)$ to denote this property. Furthermore the potential $V_2 \in C(V_1)$, resp. $V_1 \in C(V_2)$, belongs to the class of acceptable potentials, the reference potential being V_1, resp. V_2. The result can be obtained by estimates on the Riemann's solutions involved; it also results from a simpler argument.[24]

If V_1 and $V_2 \in C(V_0)$, then

$$\phi_1(x) = \phi_0(x) + \int_x^{\infty} K_{10}(x, y)\, \phi_0(y)\, dy, \tag{29}$$

$$\phi_0(x) = \phi_2(x) + \int_x^{\infty} K_{02}(x, y)\, \phi_2(y)\, dy. \tag{30}$$

Equation (30) follows the fact that the existence of K_{20} implies that of the inverse kernel K_{02}.

We consider now the integral operator

$$\begin{aligned} A(x, y) = {} & K_{10}(x, y) + K_{02}(x, y) \\ & + \int_x^y K_{10}(x, z)\, K_{02}(z, y)\, dz. \end{aligned} \tag{31}$$

From Eq. (31) one gets

$$\begin{aligned} &\int_x^{\infty} A(x, y)\, \phi_2(y)\, dy \\ &\quad = \int_x^{\infty} K_{10}(x, y)\, \phi_2(y)\, dy + \int_x^{\infty} K_{02}(x, y)\, \phi_2(y)\, dy \\ &\qquad + \int_x^{\infty} \int_x^{y} K_{10}(x, z)\, K_{02}(z, y)\, \phi_2(y)\, dz\, dy. \end{aligned} \tag{32}$$

Using Eq. (30) and a permutation of variables, one gets

$$\begin{aligned} \text{rhs} \equiv {} & \int_x^{\infty} K_{10}(x, y)\, \phi_2(y)\, dy + \phi_0(x) - \phi_2(x) \\ & + \int_x^{\infty} dz \int_z^{\infty} K_{10}(x, z)\, K_{02}(z, y)\, \phi_2(y)\, dy \\ \equiv {} & \int_x^{\infty} K_{10}(x, y)\, \phi_2(y)\, dy + \phi_0(x) - \phi_2(x) \\ & + \int_x^{\infty} K_{10}(x, z)[\phi_0(z) - \phi_2(z)]\, dz \equiv \phi_1(x) - \phi_2(x). \end{aligned}$$

The law of composition of Marchenko's kernel is thus established and the statement proved.

The kernel $A(x,y)$ is identical to the kernel $K_{12}(x,y)$ we wanted to construct.

An interesting feature of Eq. (31) follows. Let us assume $V_1 = V_2$, then

$$0 = K_{10}(x,y) + K_{01}(x,y) + \int_x^y K_{10}(x,z)\,K_{01}(z,y)\,dz. \tag{33}$$

Equation (33) is an integral equation for the inverse kernel $K_{01}(x,y)$ of $K_{10}(x,y)$.

3. THE TRANSLATION KERNEL FOR VELOCITY DEPENDENT INTERACTIONS

In this section the use of the transformation for Schrödinger equations with velocity dependent interaction (effective mass dependence) is discussed. Two types of methods are used. The first one transforms the velocity dependent potential into an energy dependent operator. The second one transforms the same potential into an angular momentum dependent operator. Advantages of both methods are compared and an extension is proposed for systems of coupled differential equations.

In the n-dimensional space, the Schrödinger equation reads

$$-\frac{\hbar^2}{2\mu}\{\nabla[1+p(r)]\nabla - U(r) + E\}\psi(x_1, x_2, \ldots, x_n) = 0. \tag{34}$$

In Eq. (34) the following notation was used,

$$\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \cdots, \frac{\partial}{\partial x_n}\right), \quad r^2 = \sum_{i=1}^{n} x_i^2. \tag{35}$$

As in the three-dimensional case, a partial wave decomposition can be used followed by the construction of the reduced radial equation $(u_m(r)/r)$ in the three-dimensional case. For the reduced radial function $u_m(r)$ one has the equation valid for all values of n,

$$\left(\frac{d}{dr}\,[1+p(r)]\right)\frac{d}{dr} - [1+p(r)]\frac{\nu(\nu+1)}{r^2} - W + k^2]\,u_m(r) = 0. \tag{36}$$

In Eq. (36) we have denoted by $W(r)$ the potential

$$W(r) = U(r) + \frac{(n-1)}{2}\,\frac{p'}{r},$$

and by the index ν, the number obtained from the integer m by

$$\nu = m + (n-3)/2.$$

In the physical situations the form factor $p(r)$ proportional to the density $\rho(r)$ of the medium, is

$$p(r) \sim \rho(r).$$

This proportionality leads to these two consequences:

$$p'(0) = 0 \tag{37}$$

and

$$\lim_{r\to\infty} r^2 p(r) = 0. \tag{38}$$

In addition to this assumption, we impose a positivity condition

$$1 + p(r) > 0 \quad \forall\, r. \tag{39}$$

If this positivity condition is assumed, no singularity is brought into the operator through the effective mass. Also if Eq. (36) is a matrix equation and $p(r)$ denotes the diagonal matrix

$$p_i(r)\,\delta_{ij} = (p(r))_{ij}, \tag{40}$$

one can define two diagonal matrices,

$$[1+p(r)]^{+1/2} \quad \text{and} \quad [1+p(r)]^{-1/2}. \tag{41}$$

In what follows, solutions may be scalars or matrices. When they will be scalars, the index m is used.

A. Method number one

Let us define (see Refs. 21 and 25)

$$u_m(r) = [1+p(r)]^{-1/2}\,v_m(r). \tag{42}$$

The factor $[1+p(r)]$ is the Wronskian of two linearly independent solutions of Eq. (36). The transformation defined by Eq. (42) is exactly the transformation from a nonlocal potential to its local equivalent used in Ref. 21.

The new radial equation for $v_m(r)$ follows,

$$\left(\frac{d^2}{dr^2} - \frac{\nu(\nu+1)}{r^2} - W(r,k^2) + k^2\right) v_m(r) = 0.$$

We have therefore defined:

$$W(r,k^2) \equiv W_0(r) + W_1(r) + k^2 W_2(r),$$

$$W_0(r) = -[1+p(r)]^{-1}\left(-\tfrac{1}{2}p'' + \tfrac{1}{4}p'^2[1+p(r)]^{-1} - \frac{(n-1)}{2}\,\frac{p'}{r}\right),$$

$$W_1(r) = [1+p(r)]^{-1/2}\,U(r)[1+p(r)]^{-1/2},$$

$$W_2(r) = p(r)[1+p(r)]^{-1}.$$

When the equation for $v_m(r)$ has been obtained one is in a situation to inquire about the existence of translation operators. For this inquiry one introduces the two operators

$$A_0 \equiv \frac{d^2}{dx^2} - \frac{\nu(\nu+1)}{x^2} + k^2,$$

$$A \equiv \frac{d^2}{dx^2} - \frac{\nu(\nu+1)}{x^2} - W(x,k^2) + k^2,$$

and defines the solutions $v_{0m}(x)$, $v_m(x)$ specified by their behavior at infinity, by

$$A_0 v_{0m}(x) = 0, \quad A v_m(x) = 0.$$

The existence of the integral representation

$$v_m(x) = v_{0m}(x) + \int_x^\infty K(k^2;x,y)\,v_{0m}(y)\,dy,$$

can be discussed. The kernel $K(k^2;x,y)$ satisfies the partial differential equation

$$\left[\frac{\partial^2}{\partial x^2} - \frac{\nu(\nu+1)}{x^2} - W(x,k^2)\right] K(k^2;x,y) = \left[\frac{\partial^2}{\partial y^2} - \frac{\nu(\nu+1)}{y^2}\right] K(k^2,x,y),$$

$$\lim_{y\to\infty} K(k^2,x,y) = \lim_{y\to\infty} \frac{\partial}{\partial y}\,K(k^2,x,y) = 0, \tag{43}$$

$$K(k^2,x,x) = \tfrac{1}{2}\int_x^\infty W(s,k^2)\,ds.$$

With the help of the Riemann solution R_ν,

$$R_\nu \equiv P_\nu(1-2x_1-2x_2+2x_1x_2),$$

for the equation

$$\left[\frac{\partial^2}{\partial x^2}-\frac{\nu(\nu+1)}{x^2}\right] R_\nu=\left[\frac{\partial^2}{\partial y^2}-\frac{\nu(\nu+1)}{y^2}\right] R_\nu,$$

Eq. (43) with its boundary condition can be transformed into a Volterra equation. Using Eq. (28), one has

$$K(k^2;x,y)=\tfrac{1}{2}\int_{(x+y)/2}^{\infty} R_\nu(x,y;s,x)\,W(s,k^2)\,ds + \tfrac{1}{2}\int\int_{D} R_\nu(x,y;s,u)\,W(s,k^2)\,K(k^2;s,u)\,du\,ds. \tag{44}$$

According to previous studies the existence of K follows the possession of moments of order 1 and order ν by the energy-dependent potential $W(x,k^2)$.

Since

$$W(x,k^2)=W_0(x)+W_1(x)+k^2W_2(x),$$

and W_0 and W_2 contain the density related form factor $p(r)$ and its derivatives, the condition reduces simply to the possession by the original nuclear potential $U(s)$ of moments of order 1 and order ν.

A natural question arises, that of the analytical dependence of K with respect to the energy $\lambda=k^2$.

Let us define a new function

$$L(\lambda,x,y)\equiv\frac{\partial}{\partial\lambda}K(x,y;k^2).$$

The integral equation for L is

$$L(\lambda,x;y)=\tfrac{1}{2}\int_{(x+y)/2}^{\infty} R_\nu(x,y;s,s)\,p(s)[1+p(s)]^{-1}\,ds + \tfrac{1}{2}\int\int_{D} R_\nu(x,y;s,u)\,p(s)[1+p(s)]^{-1} \times K(\lambda;s,u)\,ds\,du + \tfrac{1}{2}\int\int_{D} R_\nu(x,y;s,u) \times W(s,\lambda)\,L(\lambda;s,u)\,ds\,du. \tag{45}$$

Equation (45) can be solved by the method of successive approximations. The upper bound

$$(x)^\nu\,|K(\lambda;x,y)|\le\tfrac{1}{2}\sigma_\nu\left(\frac{x+y}{2}\right)\exp\sigma_1(x)$$

can be used to obtain an estimate for the zero-order term

$$L_0(\lambda,x,y)=\tfrac{1}{2}\int_{(x+y)/2}^{\infty} ds\,R_\nu(x,y;s,s)\,p(s)[1+p(s)]^{-1} + \tfrac{1}{2}\int\int_{D} R_\nu(x,y;s,u)\,p(s)[1+p(s)]^{-1} \times K(\lambda,s,u)\,du\,ds. \tag{46}$$

By looking to Eqs. (45) and (46) it is obvious that the analyticity of K follows its very existence.

The method just outlined extends itself to systems of coupled differential equations of the form:

$$\frac{d}{dr}[1+p_i(r)]\frac{d}{dr}u_i(r)-[1+p_i(r)]\frac{\nu_i(\nu_i+1)}{r^2}u_i(r) - \frac{(n-1)}{2}\frac{p_i'(r)}{r}u_i(r)-\sum_j U_{ij}(r)\,u_j(r)+k_i^2u_i(r)=0, \tag{47}$$

where we have separated the diagonal operators from the nondiagonal ones. Defining the matrix solution

$$v_i(x)=[1+p_i(x)]^{-1/2}u_i(x),$$

and operating as we did, we obtain

$$\frac{d^2}{dx^2}v_i(x)-\left[\frac{\nu_i(\nu_i+1)}{x^2}+W_i^{(0)}(x)+k_i^2W_i^{(2)}(x)\right]v_i(x) - \sum_j W_{ij}^{(1)}(x)\,v_j(x)+k_i^2v_i(x)=0. \tag{48}$$

Definitions used in Eq. (48) are

$$W_i^{(0)}(x)=[1+p_i(x)]^{-1}\left(-\tfrac{1}{2}p_i''+\tfrac{1}{4}p_i'^2[1+p_i(x)]^{-1} - \frac{(n-1)}{2}\frac{p_i'(x)}{x}\right),$$

$$W_{ij}^{(1)}(x)=[1+p_i(x)]^{-1/2}U_{ij}(x)[1+p_j(x)]^{-1/2},$$

$$W_i^{(2)}(x)=p_i(x)[1+p_i(x)]^{-1}.$$

The matrices

$$W^{(0)}(x)=W_i^{(0)}(x)\,\delta_{ij},$$

$$W^{(1)}(x)=W_{ij}^{(1)}(x),$$

$$W^{(2)}(x)=W_i^{(2)}(x)\,\delta_{ij},$$

$$\Lambda=\nu_i(\nu_i+1)\,\delta_{ij},$$

$$k^2=k_i^2\delta_{ij}$$

are defined and the two operators

$$A_0=\frac{d^2}{dx^2}+k^2-\frac{\Lambda}{x^2},$$

$$A=\frac{d^2}{dx^2}+k^2-\frac{\Lambda}{x^2}+W^{(0)}(x)+k^2W^{(2)}(x)+W^{(1)}(x) \tag{49}$$

are introduced together with the matrix solutions $u_{0\Lambda}, u_\Lambda$.

The translation kernel K which is now a matrix satisfies:

$$\left(\frac{\partial^2}{\partial x^2}+k^2-\frac{\Lambda}{x^2}-W^{(0)}(x)-k^2W^{(2)}(x)-W^{(1)}(x)\right)K(x,y) = \left(\frac{\partial^2}{\partial y^2}\right)K(x,y)+K(x,y)\left(k^2-\frac{\Lambda}{y^2}\right),$$

$$\lim_{y\to\infty}K(x,y)=0=\lim_{y\to\infty}\frac{\partial}{\partial y}K(x,y),$$

$$K(x,x)=\tfrac{1}{2}\int_x^\infty[W^{(0)}(s)+W^{(1)}(s)+k^2W^{(2)}(s)]\,ds. \tag{50}$$

For the discussion of the existence of K, the Riemann solutions

$$\Pi\begin{pmatrix}\nu_i,\nu_j\\ x_1,x_2\end{pmatrix}$$

are used.[26] Conclusions are identical to the ones we reported in Ref. 9; they don't need to be repeated in the present paper.

When the matrix kernel $K(k^2;x,y)$ for the operators defined in Eq. (50) has been obtained, one writes the final representation for the matrix solution $u_\Lambda(x)$,

$$u_\Lambda(x)=[1+p(x)]^{-1/2}\{v_{0\Lambda}(x)+\int_x^\infty K(k^2;x,y)\,v_{0\Lambda}(y)\,dy\}. \tag{51}$$

B. Method number two (the Liouville transformation)

Although less extensive than the first method, this second method has the advantage of not introducing an energy dependence into the transformed interaction and consequently into the transformation kernel. Its lack of extension comes first from the requirement of the same effective mass in all the channels. It comes also from the appearance of potential decreasing like x^{-3} infinity. The second method has nonetheless enough interest to be developed for its own merits. The Liouville transformation takes place in two steps. In the first step, one defines a new radial variable x by

$$x = \int_0^r \frac{dt}{[1+p(t)]^{1/2}} . \tag{52}$$

The definition implies again $1+p(t) > 0$. Then one has a one to one mapping between r and x. With this definition and the assumption that $p(t)$ goes to zero when t goes to infinity,

$$\lim_{r\to\infty} \frac{dx}{dr} = \frac{1}{[1+p(r)]^{1/2}} = 1.$$

So instead of Eq. (52) we write definition (53) for x which is its equivalent

$$x = r + \int_0^r \{(1+p(t))^{-1/2} - 1\} dt. \tag{53}$$

From Eq. (53) one sees that r goes to zero with x and x with r.

Using Eq. (52) one gets

$$\frac{d}{dr}(1+p)\frac{d}{dr} = \frac{d^2}{dx^2} + \frac{1}{2}\frac{1}{1+p}\dot{p}\frac{d}{dx} . \tag{54}$$

In Eq. (54) $\dot{p}$ denotes $(d/dx)\,p(r(x))$ (not to be confused with dp/dr which we denoted earlier p').

With the help of Eq. (54), Eq. (36) becomes

$$\left(\frac{d^2}{dx^2} + \frac{1}{2}\frac{\dot{p}}{1+p}\frac{d}{dx} - (1+p)\frac{\nu(\nu+1)}{r^2(x)} - W(x) + k^2\right) u_m(x) = 0 \tag{55}$$

with

$$W(x) = U(x) + \frac{(n-1)}{2}\frac{p'}{r} = U(x) + \frac{(n-1)}{2}\frac{\dot{p}}{r}(1+p)^{-1/2} . \tag{56}$$

Again the index m denotes scalar solutions.

The second step of the Liouville transformation is to renormalize the radial wavefunction so as to eliminate the velocity dependence from the equations.

For this purpose one defines

$$u_m(x) = (1+p)^{-1/4} v_m(x) \tag{57}$$

and uses

$$\frac{d}{dx}u = (1+p)^{-1/4}\frac{d}{dx}v - \tfrac{1}{4}\dot{p}(1+p)^{-5/4}v, \tag{58}$$

$$\frac{d^2}{dx^2}u = (1+p)^{-1/4}\frac{d^2}{dx^2}v - \tfrac{1}{2}\dot{p}(1+p)^{-5/4}\frac{d}{dx}v + \tfrac{5}{16}\dot{p}^2(1+p)^{-9/4}v - \tfrac{1}{4}(1+p)^{-5/4}\ddot{p}v. \tag{59}$$

Inserting Eqs. (57)–(59) into Eq. (55) gives

$$\left(\frac{d^2}{dx^2} - [1+p(x)]\frac{\nu(\nu+1)}{r^2(x)} - V + k^2\right) v_m = 0 \tag{60}$$

with

$$V = (1+p)^{1/4}U(x)(1+p)^{-1/4} + \frac{(n-1)}{2}\frac{\dot{p}}{r}(1+p)^{-1/2} - \tfrac{1}{16}[3\dot{p}^2(1+p)^{-1} - 4\ddot{p}](1+p)^{-1}. \tag{61}$$

As x goes to infinity the centripetal barrier goes to

$$\frac{\nu(\nu+1)}{(x+c)^2} ,$$

where c is the constant defined by

$$c = \int_0^\infty \{1 - [1+p(t)]^{-1/2}\} dt.$$

In the same way, by a Taylor expansion, as x goes to zero

$$r(x) \to x\left.\frac{dr}{dx}\right|_{x=0} = x[1+p(0)]^{1/2} .$$

The centripetal barrier at the origin is therefore

$$[1+p(0)]\frac{\nu(\nu+1)}{x^2[1+p(0)]} \equiv \frac{\nu(\nu+1)}{x^2} .$$

Equation (60) is first rewritten as

$$\left[\frac{d^2}{dx^2} - \frac{\nu(\nu+1)}{x^2} - \left(V + [1+p(x)]\frac{\nu(\nu+1)}{r^2(x)} - \frac{\nu(\nu+1)}{x^2}\right) + k^2\right] v_m(x) = 0. \tag{62}$$

To shorten the notations, we define

$$V^{(1)} = [1+p(x)]\frac{\nu(\nu+1)}{r^2(x)} - \frac{\nu(\nu+1)}{x^2} .$$

Notice that the potential $V^{(1)}$ does not possess any singular point at the origin and decreases like x^{-3} at infinity.

Back to Eq. (61) where V is defined, and still assuming the relationship between $p(r)$ and the density, we have

$$\dot{p}(0) = 0.$$

The only singularities at the origin or at infinity of $V + V^{(1)}$ are those possessed by the original potential $U(r)$.

According to the normal procedure the two differential operators

$$A \equiv \frac{d^2}{dx^2} - \frac{\Lambda}{x^2} - (V^{(1)} + V) + k^2,$$

$$A_0 \equiv \frac{d^2}{dx^2} - \frac{\Lambda}{x^2} + k^2,$$

where V may be a $j\times j$ matrix. If present, nonzero threshold energies would require an exponential decrease from $V^{(1)}$ and V. The x^{-3} decrease of $V^{(1)}$ obliges k^2 to be a scalar. Furthermore, σ_ν found in the bound

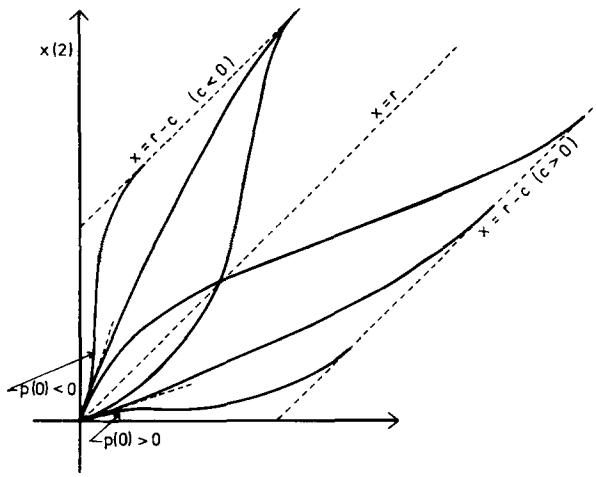


FIG. 3.

for K, together with the x^{-3} decrease, restrict the extension to systems with maximum $\nu_j = 3$. Matrix solutions of the differential matrix equations

$$A v_\Lambda(x) = 0, \quad A_0 v_{0\Lambda}(x) = 0,$$

are now related through an integral transformation. We write this transformation as

$$v_\Lambda(x) = v_{0\Lambda}(x) + \int_x^\infty K(x,y)\, v_{0\Lambda}(y)\, dy. \tag{63}$$

The equations for $K(x,y)$ are similar to Eq. (27), within the restricted class of operators specified earlier. $K(x,y)$ exists if the original potential U possesses the appropriate absolute moments.[9] From Eq. (63) one returns to the original variable x by

$$u_\Lambda(r) = [1+p(r)]^{-1/4}[v_{0\Lambda}(x) + \int_x^\infty K(x,y)\, v_{0\Lambda}(y)\, dy] \tag{64}$$

and

$$u_\Lambda(r) = [1+p(r)^{-1/4}\{v_{0\Lambda}[x(r)] + \int_x^\infty K[x(r), y(s)] \times v_{0\Lambda}[y(s)] \times [1+p(s)]^{-1/2}\, ds\}. \tag{65}$$

From Eq. (53) one obtains the asymptotic relation

$$x = r - c + \int_r^\infty \{1 - [1+p(s)]^{-1/2}\}\, ds = r - c + O[\int_r^\infty ds f(s)]. \tag{66}$$

For illustration in Fig. 3, typical curves $x(r)$ are constructed. Returning to the scalar case we have the following asymptotic behaviors:

$$u_m(r) \sim a_m \sin\left(kr - \frac{m\pi}{2} + \delta_m\right), \qquad \phi_m(x) \sim b_m \sin\left(kx - \frac{m\pi}{2} + \eta_m\right). \tag{67}$$

Using Eq. (66), one sees that the limits (67) as x and r go together to infinity imply the simple equation

$$k(r-c) + \eta_m = kr + \delta_m$$

or

$$\eta_m = \delta_m + kc. \tag{68}$$

Equations (66) and (68) are the basis for Calogero choosing a new variable $x = r - c$ and identifying velocity dependent with static hard core interactions.

4. CONCLUSION

In the paper we assumed Schrödinger equations with an effective mass and showed the existence of translation operators, in the sense of Marchenko, between the free Schrödinger equation and the Schrödinger equation with an effective mass in addition to the nuclear potential. Two methods were used to achieve this; they had a common characteristic, the full Schrödinger equation was subjected to some transformation prior to being considered for translation kernel purposes. One may wonder whether this preliminary transformation is necessary: The answer is no, but it is convenient as the following will show. The two operators

$$A_0 = a_0(x)\frac{d^2}{dx^2} + b_0(x)\frac{d}{dx} + c_0(x), \tag{69}$$

$$A = a(x)\frac{d^2}{dx^2} + b(x)\frac{d}{dx} + c(x), \tag{70}$$

may be directly considered. When Eq. (1), $XA_0 = AX$, is developed one obtains the partial differential equation for the kernel

$$K(x,y).$$

They are

$$\frac{\partial^2}{\partial y^2}[K(x,y)\, a_0(y)] - \frac{\partial}{\partial y}[K(x,y)\, b_0(y)] + K(x,y)\, c_0(y) = a_1(x)\frac{\partial^2}{\partial x^2} K(x,y) + b_1(x)\frac{\partial}{\partial x} K(x,y) + c_1(x) K(x,y), \tag{71}$$

with complicated boundary conditions. The restriction for the translation operator to be of the Volterra type leads to the following two constraints on the coefficients of Eqs. (69) and (70):

$$a_1(x) = a_0(x) \tag{72}$$

and

$$b_1(x) - b_0(x) = -K(x,x)\, a_0(x) + a_1(x) K(x,x). \tag{73}$$

Inspection of the system thus obtained, convinced us it was not worth pursuing except when the two conditions

$$a = a_0 \equiv \text{scalar}, \quad b = b_0 = 0$$

are realized.

At the termination of this paper we can assert the validity of the Marchenko representation in a wide variety of physical situations: many-channel scattering in a n-dimensional space with centrifugal force or Coulomb force or even both, with or without an effective mass. Along our study we were led to abandon the Gel'fand—Levitan representation. It is not such a damageable result, the connection of the regular physical solution with the physical solution is not maintained when the passage from a one-channel to a many-channel problem is operated. In addition one should say the Marchenko approach is the one which is appealing to the creators of the solution of nonlinear problems via inverse scattering methods.[5]

We must emphasize that the determination of the conditions the nuclear potential should satisfy for the existence of the translation kernel, is only a first step

into the inverse problem. When translation kernels for a class of potentials have been proved to exist, one obtains upper bounds for these translation kernels. Within this class of potentials, it may be possible to construct a fundamental equation between $K(x,y)$ and some spectral matrix $F(x,y)$ as Marchenko or Gel'fand–Levitan did. The upper bounds satisfied by $K(x,y)$ induce upper bounds that the spectral matrix $F(x,y)$ itself verifies. Considering now the fundamental equation as an equation for $K(x,y)$, the necessary bound, which the spectral matrix verifies, becomes the important element in deciding whether or not the fundamental equation possesses a unique solution. This discussion is the essence of the inverse problem.

[1]B.M. Levitan, *Generalized Translation Operators* (IPST, Jerusalem, 1964).
[2]I.M. Gel'fand and B.M. Levitan, "On the Determination of Differential Equation from its Spectral Function," Am. Math. Soc. Transl. Sec. II, 1 (1955).
[3]Z.S. Agranovitch and V.A. Marchenko, *The Inverse Problem of Scattering Theory* (Gordon and Breach, New York, 1963).
[4]K.O. Friedrichs, Commun. Pure Appl. Math. **1**, 361 (1948).
[5]M.J. Ablowitz, D.J. Kaup, A.C. Newell, and H. Segur, Stud. Appl. Math. **53**, 249 (1976).
[6]P.D. Lax, Commun. Pure Appl. Math. **21**, 467 (1968).
[7]C. Coudray and M. Coz, Ann. Phys. (N.Y.) **6**, 488 (1970).
[8]M. Coz and C. Coudray, J. Math. Phys. **17**, 888 (1976).
[9]M. Coz and P. Rochus, J. Math. Phys. **17**, 899 (1976).
[10]M. Coz, J. Math. Anal. Appl. **60** (1977) (to be published).
[11]Ref. 1, pp. 37–8.
[12]K. Yosida, *Lectures on Differential and Integral Equations* (Wiley-Interscience, New York, 1960), p. 110; F.G. Tricomi, *Differential Equations* (Blackie, London, 1961), p. 163.
[13]W. Feller, J. Math. Mech. **8**, 339 (1959).
[14]I.S. Kac, Am. Math. Soc. Transl. **62**, 204 (1967).
[15]M. Jacob, Ed., *Dual Theories* (North-Holland, Amsterdam, 1974).
[16]E.H. Canfield, thesis UCLA 6652; M. Ravazy, G. Field, and J.S. Levinger, Phys. Rev. **125**, 269 (1962); O. Rojo and L.M. Simmons, Phys. Rev. **125**, 273 (1962); O. Rojo and J.S. Levinger, Phys. Rev. **123**, 2177 (1961).
[17]F. Calogero, Lett. Nuovo Cimento **2**, 553 (1967).
[18]V.A. Marchenko, Naukova Dumka Kiev, 1972.
[19]L.D. Faddeev, VINITI **3**, 93 (1974).
[20]B.N. Zachariev, Sov. J. Nucl. Phys. **22**, 149 (1976).
[21]M. Coz and A.D. MacKellar, Ann. Phys. **58**, 504 (1970); **59**, 219 (1970).
[22]M.L. Mehta, *Random Matrices* (Academic, New York, 1967).
[23]F. Dyson, Commun. Math. Phys. **47**, 171 (1976).
[24]V. De Alfaro and T. Regge, *Potential Scattering* (North-Holland, Amsterdam, 1965), p. 149.
[25]A.M. Green, Nucl. Phys. **33**, 218 (1962).
[26]T.W. Chaundy (see also Ref. 7), Q.J. Math. Oxford Ser. **10**, 219 (1939).

ERRATA

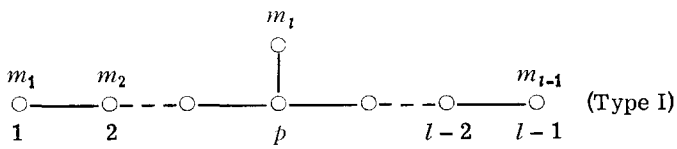
Erratum: Higher weights of semisimple Lie algebras [J. Math. Phys. 18, 1162 (1977)]

W. Laskar

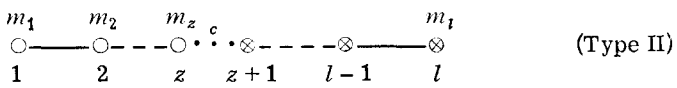
24, rue de la Distillerie, 44000 Nantes, France

(Received 24 June 1977)

Page 1163, just before Eq. (11): the following are the corrected versions of Types I and II:



(Type I)



(Type II)

Page 1163, at the end of line (– 2) A_l should belong to line (–3).

Page 1165, line 1, left column: The word "will" was blotted out in the printing process.

Page 1167, first line immediately after Eq. (35): the lhs of the equations should read

$$\delta(S_{r-1}^{(i)}) \quad \text{and} \quad \delta(\lambda_{r-1}^{(i)}).$$

Page 1168, Table V: the second line should read

$$+\sum_{k=p}^{l-1}\left(\sum_{i=1}^{p-1}(l-k)\,2im_i+\sum_{i=p}^{i=k}[p^2+(2-p)\,i](l-k)\,m_i\right.$$
$$\left.+\sum_{i=k+1}^{l-1}[p^2+(2-p)\,k](l-i)\,m_i\right.$$

Page 1169, Table VII, rightmost part of lines for $l=7$: The numbers blotted out in the printing process are 16 (first line) and 15 (second line).

Page 1169, Table VII, last column for $l=8$ (last entry under 3): the condition $1 \leq i \leq 16$ should apply to both equations in that last $l=8$ entry.

Page 1170, Table VIII: the large brace should only extend to include the equation $\alpha_4 = \frac{1}{2}(e_1 - e_2 - e_3 - e_4)$.

Page 1171, Table IX, left-hand entry 2: the first factor on the rhs of the $\Pi_2(D_l)$ equation should read

$$(m_l + 1).$$

Page 1172, Eq. (46): The lhs partially blotted out in the printing process, should read: $(H_{\alpha_k})^{r,i}_{r,i}$.

Page 1173, Table XI, bottom line of the E_6 entry: C_{21}^2 and C_{27}^2 should read C_{27}^2 and C_{27}^3, respectively.

Page 1174, Table XII, under 1st labeling:

$$\left(2\,\frac{l-1}{2l}\right) \quad \text{should read} \quad 2\left(\frac{l-1}{2l}\right).$$

Page 1175, the third to last line should not include ⊞.

Page 1176: the rhs of H_{α_2} equation should read

$$\text{diag}(m_2/c,\, m_2/c+1,\, m_2/c+1-2/c,\ \text{etc.}).$$

Copyright © 1977 American Institute of Physics

into the inverse problem. When translation kernels for a class of potentials have been proved to exist, one obtains upper bounds for these translation kernels. Within this class of potentials, it may be possible to construct a fundamental equation between $K(x,y)$ and some spectral matrix $F(x,y)$ as Marchenko or Gel'fand–Levitan did. The upper bounds satisfied by $K(x,y)$ induce upper bounds that the spectral matrix $F(x,y)$ itself verifies. Considering now the fundamental equation as an equation for $K(x,y)$, the necessary bound, which the spectral matrix verifies, becomes the important element in deciding whether or not the fundamental equation possesses a unique solution. This discussion is the essence of the inverse problem.

[1]B.M. Levitan, *Generalized Translation Operators* (IPST, Jerusalem, 1964).
[2]I.M. Gel'fand and B.M. Levitan, "On the Determination of Differential Equation from its Spectral Function," Am. Math. Soc. Transl. Sec. II, 1 (1955).
[3]Z.S. Agranovitch and V.A. Marchenko, *The Inverse Problem of Scattering Theory* (Gordon and Breach, New York, 1963).
[4]K.O. Friedrichs, Commun. Pure Appl. Math. **1**, 361 (1948).
[5]M.J. Ablowitz, D.J. Kaup, A.C. Newell, and H. Segur, Stud. Appl. Math. **53**, 249 (1976).
[6]P.D. Lax, Commun. Pure Appl. Math. **21**, 467 (1968).
[7]C. Coudray and M. Coz, Ann. Phys. (N.Y.) **6**, 488 (1970).
[8]M. Coz and C. Coudray, J. Math. Phys. **17**, 888 (1976).
[9]M. Coz and P. Rochus, J. Math. Phys. **17**, 899 (1976).
[10]M. Coz, J. Math. Anal. Appl. **60** (1977) (to be published).
[11]Ref. 1, pp. 37–8.
[12]K. Yosida, *Lectures on Differential and Integral Equations* (Wiley-Interscience, New York, 1960), p. 110; F.G. Tricomi, *Differential Equations* (Blackie, London, 1961), p. 163.
[13]W. Feller, J. Math. Mech. **8**, 339 (1959).
[14]I.S. Kac, Am. Math. Soc. Transl. **62**, 204 (1967).
[15]M. Jacob, Ed., *Dual Theories* (North-Holland, Amsterdam, 1974).
[16]E.H. Canfield, thesis UCLA 6652; M. Ravazy, G. Field, and J.S. Levinger, Phys. Rev. **125**, 269 (1962); O. Rojo and L.M. Simmons, Phys. Rev. **125**, 273 (1962); O. Rojo and J.S. Levinger, Phys. Rev. **123**, 2177 (1961).
[17]F. Calogero, Lett. Nuovo Cimento **2**, 553 (1967).
[18]V.A. Marchenko, Naukova Dumka Kiev, 1972.
[19]L.D. Faddeev, VINITI **3**, 93 (1974).
[20]B.N. Zachariev, Sov. J. Nucl. Phys. **22**, 149 (1976).
[21]M. Coz and A.D. MacKellar, Ann. Phys. **58**, 504 (1970); **59**, 219 (1970).
[22]M.L. Mehta, *Random Matrices* (Academic, New York, 1967).
[23]F. Dyson, Commun. Math. Phys. **47**, 171 (1976).
[24]V. De Alfaro and T. Regge, *Potential Scattering* (North-Holland, Amsterdam, 1965), p. 149.
[25]A.M. Green, Nucl. Phys. **33**, 218 (1962).
[26]T.W. Chaundy (see also Ref. 7), Q.J. Math. Oxford Ser. **10**, 219 (1939).

ERRATA

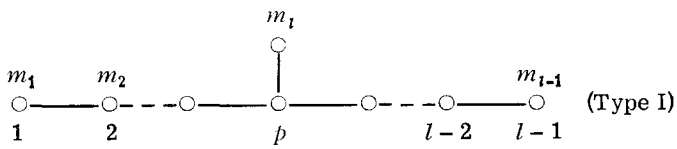
Erratum: Higher weights of semisimple Lie algebras [J. Math. Phys. 18, 1162 (1977)]

W. Laskar

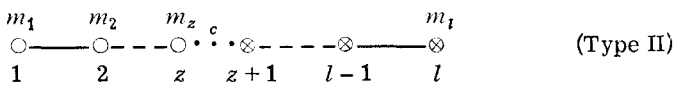
24, rue de la Distillerie, 44000 Nantes, France

(Received 24 June 1977)

Page 1163, just before Eq. (11): the following are the corrected versions of Types I and II:



(Type I)



(Type II)

Page 1163, at the end of line (– 2) A_l should belong to line (–3).

Page 1165, line 1, left column: The word "will" was blotted out in the printing process.

Page 1167, first line immediately after Eq. (35): the lhs of the equations should read

$$\delta(S_{r-1}^{(i)}) \quad \text{and} \quad \delta(\lambda_{r-1}^{(i)}).$$

Page 1168, Table V: the second line should read

$$+\sum_{k=p}^{l-1}\left(\sum_{i=1}^{p-1}(l-k)\,2im_i+\sum_{i=p}^{i=k}[p^2+(2-p)\,i](l-k)\,m_i\right.$$
$$\left.+\sum_{i=k+1}^{l-1}[p^2+(2-p)\,k](l-i)\,m_i\right.$$

Page 1169, Table VII, rightmost part of lines for $l=7$: The numbers blotted out in the printing process are 16 (first line) and 15 (second line).

Page 1169, Table VII, last column for $l=8$ (last entry under 3): the condition $1 \leq i \leq 16$ should apply to both equations in that last $l=8$ entry.

Page 1170, Table VIII: the large brace should only extend to include the equation $\alpha_4 = \frac{1}{2}(e_1 - e_2 - e_3 - e_4)$.

Page 1171, Table IX, left-hand entry 2: the first factor on the rhs of the $\Pi_2(D_l)$ equation should read

$$(m_l + 1).$$

Page 1172, Eq. (46): The lhs partially blotted out in the printing process, should read: $(H_{\alpha_k})^{r,i}_{r,i}$.

Page 1173, Table XI, bottom line of the E_6 entry: C_{21}^2 and C_{27}^2 should read C_{27}^2 and C_{27}^3, respectively.

Page 1174, Table XII, under 1st labeling:

$$\left(2\,\frac{l-1}{2l}\right) \quad \text{should read} \quad 2\left(\frac{l-1}{2l}\right).$$

Page 1175, the third to last line should not include ⊞.

Page 1176: the rhs of H_{α_2} equation should read

$$\text{diag}(m_2/c,\, m_2/c+1,\, m_2/c+1-2/c,\ \text{etc.}).$$

Copyright © 1977 American Institute of Physics