A criterion for reducibility of a relativistic wave equation*

E. C. G. Sudarshan, M.A.K. Khalil, and W. J. Hurley

Center for Particle Theory, Department of Physics, University of Texas at Austin, Austin, Texas 78712

(Received 18 June 1976; revised manuscript received 30 June 1976)

In general when one writes a relativistic wave equation of the form $(-i\Gamma\cdot\partial+m)\psi(x)=0$, that transforms covariantly under some representation $\Lambda\rightarrow T(\Lambda)$ of SL(2,$\mathbb{C}$), it is nontrivial to determine whether or not the equation is irreducible or to avoid ending up with a reducible equation; especially if $T(\Lambda)$ contains repeating irreducible representations. In this paper a simple(st) criterion is given by which one can determine whether or not an equation is irreducible. It is shown that if Γ_μ have any invariant subspace at all, then that subspace must be a representation space of some combination of SL(2,$\mathbb{C}$) representations in $T(\Lambda)$. Knowing this, it is shown that a wave equation is reducible if and only if there exists some idempotent projector $\tilde{P}$ such that $(1-\tilde{P})\Gamma_0\tilde{P}=0$ other than $\tilde{P}=0$ or I. A method for constructing all possible admissable $\tilde{P}$'s is given. A simple example of the technique is also given.

I. INTRODUCTION

Relativistic wave equations of the form:

$$(-i\Gamma\cdot\partial+m)\psi(x)=0$$

can be reducible or irreducible. The meaning of "reducible"[1] in the context of relativistic wave equations is precisely formulated in the next section.

It turns out that reducible equations have particular properties that make theories based on such equations equivalent to simpler theories, both in the free field and interacting cases.[2,3] It is therefore important to know when a given equation is reducible, and hence, possibly equivalent to a simpler equation. The structure of reducible equations has been studied in Ref. 2.

When one constructs a wave equation, it is in general nontrivial to insure that the equation is irreducible. The main concern of this paper is to formulate the simplest possible criterion by which one can determine whether a given wave equation is reducible or not. Such a criterion is formulated in the next section.

Finally, in Sec. III, a simple example is considered that illustrates the use of the criterion.

II. CRITERION FOR REDUCIBILITY

$$(-i\Gamma_\mu\partial^\mu+m)\psi(x)=0 \tag{1}$$

is a relativistic wave equation that transforms covariantly under a representation of SL(2,$\mathbb{C}$), $\Lambda\rightarrow T(\Lambda)$:

$$T(\Lambda)=\bigoplus_{j=1}^{n}\alpha_j T_j(\Lambda). \tag{2}$$

The set of matrices $\{\Gamma_\mu\}$ may be regarded as a set of linear transformations over a linear space $R(N)$, where N is the number of rows (or columns) of Γ_μ,

$$\{\Gamma_\mu\}=\{\Gamma_0,\Gamma_1,\Gamma_2,\Gamma_3\}.$$

Definition 1: $\{\Gamma_\mu\}$ is called a *reducible set* $\Longleftrightarrow\exists$ a proper subspace $R_1\subset R(N)\ni$

$$\Gamma R_1\subseteq R_1\subset R(N)\quad\forall\,\Gamma\in\{\Gamma_\mu\}. \tag{3}$$

The subspace R_1 is called an *invariant subspace of* $\{\Gamma_\mu\}$ (IS of $\{\Gamma_\mu\}$).

Definition 2: Suppose R_j with $j=1,\ldots,L$ is a collection of all the invariant subspaces of $\{\Gamma_\mu\}$; then

$$R_0=\bigcup_{j=1}^{L}R_j \tag{4}$$

is an invariant subspace of $\{\Gamma_\mu\}$ called the *maximal invariant subspace of* $\{\Gamma_\mu\}$.

The space $R(N)$ is a representation space of $\Lambda\rightarrow T(\Lambda)$, i.e., $T(\Lambda)$ act as linear transformations on $R(N)$.

Definition 3: If R_s is an IS of $\{\Gamma_\mu\}$ and $TR_s=R_s$, then R_s is called an *invariant* SL(2,$\mathbb{C}$) *subspace of* $\{\Gamma_\mu\}$.

Lemma 1: $R_0\subset R(N)$ in Eq. (4) is an invariant SL(2,$\mathbb{C}$) subspace of $\{\Gamma_\mu\}$.

Proof: $\Gamma_\mu R_0\subseteq R_0\quad\forall\,\Gamma_\mu$.

Recall that

$$T^{-1}\Gamma_\mu T=\Lambda_\mu{}^\nu\Gamma_\nu, \tag{5}$$

$$\Gamma_\mu TR_0\subseteq T\Lambda_\mu{}^\nu\Gamma_\nu R_0. \tag{6}$$

Suppose ϕ_0 is any vector in R_0;

$$\Gamma_\mu T\phi_0=T\Lambda_\mu{}^\nu\Gamma_\nu\phi_0, \tag{7}$$

now

$$\Lambda_\mu{}^\nu\Gamma_\nu\phi_0\in R_0 \tag{8}$$

since for any value of $\mu=0,1,2,3$ the right-hand side of (7) is a linear combination of Γ_ν acting on ϕ_0, and each $\Gamma_\nu\phi_0\in R_0$ hence (8) follows. Now according to (7)

$$\Gamma_\mu T\phi_0=T\phi_0',\quad\phi_0,\phi_0'\in R_0$$

or

$$\Gamma_\mu TR_0\subseteq TR_0. \tag{9}$$

Therefore, TR_0 is also an invariant subspace of Γ_μ, but R_0 is a maximal invariant subspace of Γ_μ, hence

$$TR_0\subseteq R_0. \tag{10}$$

Recall that T are nonsingular transformations so

$$TR_0=R_0.\ \blacksquare \tag{11}$$

In the following discussion a criterion for determining whether or not $\{\Gamma_\mu\}$ is a reducible set, will be formulated. For later convenience the bases for $R(N)$ will be chosen to be completely reducible bases (CRB's). A CRB is any basis in which $T(\Lambda)$ is block diagonal, and each block corresponds to an irreducible representation, $T_j(\Lambda)$ of SL(2, $\mathbb{C}$), in Eq. (2). Recall that altogether there are M irreducible representations where $M=\sum_{j=1}^{n}\alpha_j$ [Eq. (2)], there are α_j copies of the irreducible representation $T_j(\Lambda)$ for each j; so $T(\Lambda)$ is $M\times M$ in block form. Similarly $R(N)$ is a direct sum of M subspaces, each being a representation space of some irreducible representation of SL(2, $\mathbb{C}$) in (2),

$$R(N)=\bigoplus_{j=1}^{n}\alpha_j R_{(j)}. \tag{12}$$

One can now define SL(2, $\mathbb{C}$) *projectors* represented by Hermitian matrices in some CRB, that are idempotent and act on $R(N)$ such that

$$P_\alpha R(N)=R_{[\alpha]}, \tag{13}$$

where $R_{[\alpha]}$ is an SL(2, $\mathbb{C}$) subspace of $R(N)$, i.e., a direct sum of some $R_{(j)}$ in (2),

$$R_{[\alpha]}=\bigoplus_{j=1}^{n}a_j R_{(j)}, \tag{14a}$$

$$0\le a_j\le\alpha_j. \tag{14b}$$

The subscript $[\alpha]$ denotes all the different combinations of a_j that satisfy (14b). P_α can be written in $n\times n$ block form, where each block j corresponds to the connection of the α_j representations T_j and is thus an $\alpha_j\times\alpha_j$ block matrix that is idempotent. The P_α do not mix vectors from spaces corresponding to different representations of SL(2, $\mathbb{C}$) in (2), but can mix vectors corresponding to the same representation of which there are α_j copies for a given T_j in (2). In this form it is obvious that

$$[P_\alpha, T(\Lambda)]=0 \quad \forall\Lambda\in \mathrm{SL}(2,\mathbb{C}). \tag{15}$$

Since R_0 is an SL(2, $\mathbb{C}$) subspace of $R(N)$, there exists projectors of the type described above, P_0, such that

$$P_0R(N)=R_0 \tag{16}$$

and every vector $\phi_0\in R_0$ can be written as $P_0\phi$ for some vector $\phi\in R(N)$.

Lemma 2: $\{\Gamma_\mu\}$ reducible $\Longleftrightarrow\exists$ some $\tilde{P}_0$ satisfying (16) that is idempotent such that

$$(I-\tilde{P}_0)\Gamma_\mu\tilde{P}_0=0. \tag{17}$$

Proof: $\tilde{P}_0$ is not required to be Hermitian. Suppose $\{\Gamma_\mu\}$ is reducible, then by Lemma 1 $\exists$ an SL(2, $\mathbb{C}$) subspace R_0 of $R(N)\ni\Gamma_\mu R_0\subseteq R_0$. Choose a particular basis for $R(N)$ where $\phi\in R(N)$ is in the form

$$\phi=\left[\frac{\chi_0}{\psi}\right], \tag{18}$$

where every vector of R_0 can be written

$$\phi_0\in R_0,\quad \phi_0=\left[\frac{\chi_0}{0}\right]. \tag{19}$$

Now in this basis Γ'_μ still has the property

$$\Gamma'_\mu R_0\subseteq R_0,$$

since this property is basis independent. Now clearly one may pick a $\tilde{P}_0^J$ in this basis such that

$$\tilde{P}_0^J=\begin{array}{|c|c|}\hline I & 0\\ \hline 0 & 0\\ \hline\end{array}\ . \tag{20}$$

I is a $q\times q$ identity matrix where $q=\dim R_0$. Note that (20) is an idempotent operator (actually in its Jordan canonical form). Apply the matrix $[\Gamma'_\mu\tilde{P}_0^J-\tilde{P}_0^J\Gamma'_\mu\tilde{P}_0^J]$ to an arbitrary vector $\phi\in R(N)$,

$$\Gamma'_\mu\tilde{P}_0^J\phi-\tilde{P}_0^J\Gamma'_\mu\tilde{P}_0^J\phi=\Gamma'_\mu\phi_0-\tilde{P}_0^J(\Gamma'_\mu\phi_0)=0.$$

The last step follows because $\Gamma'_\mu\phi_0=\phi'_0\in R_0$ and in this basis $\tilde{P}_0^J\phi'_0=\phi'_0$ for any vector $\phi'_0\in R_0$. The only matrix that maps every vector $\phi\in R(N)$ into 0 is the zero matrix

$$\Gamma'_\mu\tilde{P}_0^J-\tilde{P}_0^J\Gamma'_\mu P_0^J=0. \tag{21}$$

Since an idempotent projector exists in one basis satisfying (21) (by construction) such an operator exists in all bases; i.e., for any $\Gamma_\mu=V\Gamma'_\mu V^{-1}$, the operator $V\tilde{P}_0^J V^{-1}$ is such an operator. In particular a $\tilde{P}_0$ exists in all CRB's satisfying (17). The quality of $\tilde{P}_0$ being idempotent is preserved by all nonsingular transformations but the hermiticity is not. So in general for any CRB, only $\tilde{P}_0^2=\tilde{P}_0$ will be required.

Now assume that some SL(2, $\mathbb{C}$) invariant operator $\tilde{P}_0$ exists such that (17) holds. Suppose ϕ is any vector in $R(N)$, then $\tilde{P}_0\phi=\phi_0\in R_0$ and by (17),

$$\Gamma_\mu P_0\phi=\tilde{P}_0\Gamma_\mu\tilde{P}_0\phi \iff \Gamma_\mu\phi_0=\tilde{P}_0\Gamma_\mu\phi_0.$$

Now $\Gamma_\mu\phi_0\in R(N)$ since $\Gamma_\mu:R(N)\to R(N)$. The right-hand side of the above equation,

$$\tilde{P}_0\Gamma_\mu\phi_0=\tilde{P}_0\phi'=\phi'_0\in R_0,$$

so one can see that Γ_μ: $\phi_0\to\phi'_0$, ϕ_0, $\phi'_0\in R_0$. Since every vector $\phi_0\in R_0$ can be written as $\tilde{P}_0\phi$ for some $\phi\in R(N)$ one concludes that Γ_μ maps every vector $\phi_0\in R_0$ into some other vector of R_0, therefore $\Gamma_\mu R_0\subseteq R_0$ and $\{\Gamma_\mu\}$ is a reducible set. ■

Lemma 3: $(1-\tilde{P}_0)\Gamma_\mu\tilde{P}_0=0\iff(1-\tilde{P}_0)\Gamma_0\tilde{P}_0=0$.

Proof: It is obvious that $(1-\tilde{P}_0)\Gamma_\mu\tilde{P}_0=0 \Longrightarrow(1-\tilde{P}_0)\Gamma_0\tilde{P}_0=0$.

On the other hand, suppose $(1-\tilde{P}_0)\Gamma_0\tilde{P}_0=0$, then since $\Gamma_i=i\Gamma_0N_i-iN_i\Gamma_0$ where N_i are the generators of the boosts in the i direction for the representation $\Lambda\to T(\Lambda)$, one notices that

$$(1-\tilde{P}_0)\Gamma_i\tilde{P}_0=i[(1-\tilde{P}_0)\Gamma_0N_i\tilde{P}_0-(1-\tilde{P}_0)N_i\Gamma_0\tilde{P}_0].$$

Since

$$[\tilde{P}_0,N_i]=0,$$
$$(1-\tilde{P}_0)\Gamma_i\tilde{P}_0=i[(1-\tilde{P}_0)\Gamma_0\tilde{P}_0N_i-N_i(1-\tilde{P}_0)\Gamma_0\tilde{P}_0]=0.$$

The conclusion one can draw is that $\{\Gamma_\mu\}$ is a reducible set $\iff$ there exists an idempotent SL(2, $\mathbb{C}$) projector $\tilde{P}_0$ such that

$$(1-\tilde{P}_0)\Gamma_0\tilde{P}_0=0. \tag{22}$$

$\tilde{P}_0$, being an SL(2, $\mathbb{C}$) projector, is of the following form in the CRB indicated:

$$\tilde{P}_0 = \begin{array}{cccccc} \alpha_1 T_1 & \alpha_2 T_2 & \cdots & \alpha_n T_n & \\ \left[\begin{array}{cccc} \tilde{P}_0^{11} & & & \\ & \tilde{P}_0^{22} & & \\ & & \ddots & \\ & & & \tilde{P}_0^{nn} \end{array}\right] & & & & \begin{array}{c} \alpha_1 T_1 \\ \alpha_2 T_2 \\ \vdots \\ \alpha_n T_n \end{array} \end{array},$$

$[\tilde{P}_0^{jj}]^2 = \tilde{P}_0^{jj}$ for each $j = 1, \ldots, n$.

If one finds that no such projector exists other than I or 0, then one may conclude that $\{\Gamma_\mu\}$ is an irreducible set.

Relativistic wave equations where $\{\Gamma_\mu\}$ is a reducible set are called *reducible wave equations*.

III. AN EXAMPLE

Consider any two irreducible, interlocking representations of $\mathrm{SL}(2, \mathbb{C})$ denoted A and B. For illustrating the technique consider any wave equation that can be constructed so as to transform under $T(\Lambda) = A \oplus B \oplus B$, then Γ_0 is the following:

$$\Gamma_0 = \begin{array}{cc} \begin{array}{ccc} A & B & B \end{array} & \\ \left[\begin{array}{ccc} & aD_1 & bD_1 \\ cD_2 & & \\ dD_2 & & \end{array}\right] & \begin{array}{c} A \\ B \\ B \end{array} \end{array}, \tag{23}$$

where a, b, c, and d are complex numbers, assumed to be nonzero (no requirements of unique mass or spin are imposed). The most general allowed $\tilde{P}$ is

$$\tilde{P} = \begin{array}{cc} \begin{array}{ccc} A & B & B \end{array} & \\ \left[\begin{array}{ccc} 1 & & \\ & \alpha & \beta \\ & \gamma & \rho \end{array}\right] & \begin{array}{c} A \\ B \\ B \end{array} \end{array} \qquad \begin{bmatrix} \alpha & \beta \\ \gamma & \rho \end{bmatrix}^2 = \begin{bmatrix} \alpha & \beta \\ \gamma & \rho \end{bmatrix}, \tag{24}$$

where α, β, γ, and ρ are complex multiples of the appropriate dimensional identity matrices.

The criterion $(1 - \tilde{P})\Gamma_0 \tilde{P} = 0$ yields

$$(1 - \alpha)c - \beta d = 0, \tag{25a}$$

$$-\gamma c + (1 - \rho)d = 0. \tag{25b}$$

If $\begin{pmatrix} \alpha & \beta \\ \gamma & \rho \end{pmatrix}$ is a nontrivial idempotent operator ($\neq 0, I$), then

$$\alpha + \rho = 1, \tag{26a}$$

$$\alpha\rho - \beta\gamma = 0. \tag{26b}$$

The condition $(1 - \tilde{P})\Gamma_0 \tilde{P} = 0$ and Eqs. (25) can be rewritten, using (26a), as

$$\rho c - \beta d = 0, \quad \alpha d - \gamma c = 0,$$

but (26b) assures us that there is always a nontrivial solution

$$\beta = \frac{(1 - \alpha)c}{d}, \quad \gamma = \alpha \frac{d}{c}.$$

So, whatever be the specific nonzero values of c, d (and any values of a, b) there exists a family of projectors (one for each choice of α),

$$\tilde{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \alpha & (1 - \alpha)c/d \\ 0 & \alpha(d/c) & (1 - \alpha) \end{bmatrix}, \tag{3.5}$$

such that $\tilde{P}^2 = \tilde{P}$ and $\tilde{P} \neq 0$ or I, and $(1 - \tilde{P})\Gamma_0 \tilde{P} = 0$. In case $c = 0$, choose $\alpha = 0$; similarly if $d = 0$, choose $\alpha = 1$ (c and d cannot both be zero). Therefore, any equation transforming under $A \oplus B \oplus B$ must be a reducible equation. The structure of such equations, and general theorems regarding the condition on $T(\Lambda)$ when one is forced into reducible equations are discussed elsewhere.[3,4]

An example of the use of these results to prove a given equation to be irreducible can be found in the references.[5]

ACKNOWLEDGMENT

We are grateful to Professor N. Mukunda for critical reading of the manuscript.

*Research supported in part by the Energy Research and Development Administration E(40-1)3992.

[1] M.A. Naimark, *Linear Representations of the Lorentz Group* (Pergamon, New York, 1964).
[2] M.A.K. Khalil, "Reducible Relativistic Wave Equations," CPT, University of Texas preprint ORO 260, 1976 (to be published).
[3] M.A.K. Khalil, "Relativistic Wave Equations," unpublished Doctoral Dissertation (University of Texas, Austin, 1976).
[4] M.A.K. Khalil, "The Structure of Barnacled Relativistic Wave Equations," CPT, University of Texas preprint CPT 261, 1975 (to be published).
[5] M.A.K. Khalil, "Properties of a 20-component Spin 1/2 Relativistic Wave Equation," CPT, University of Texas preprint ORO 253, 1976 (to be published in Phys. Rev. D).

Asymmetric gas thermodynamics

M. Shahinpoor*

College of Engineering, Pahlavi University, Shiraz, Iran
(Received 27 April 1976)

We consider a gas in which the density and temperature fields are spatially and timewise nonuniform. We then show that on the basis of the entropy principle the stress tensor is nonsymmetric. This nonsymmetry of the stress tensor is shown to be the driving force behind the creation of microrotational fields and dynamic spatial polarizations in the gas. It is also shown that the nonsymmetric part of the stress tensor arises out of the interaction of gradient fields of temperature and density.

1. INTRODUCTION

The modern theories of fluids with microstructure have not allocated any space for gases in which the state of the stress is antisymmetric. The antisymmetric part of the stress tensor in such theories[1,2] strongly depends on the higher gradients of the deformation rate tensor D_{ij} and spin rate tensor ω_{ij}, where

$$D_{ij}=\tfrac{1}{2}(v_{i,j}+v_{j,i}), \tag{1.1}$$

$$\omega_{ij}=\tfrac{1}{2}(v_{i,j}-v_{j,i})+\Omega_{ji} \tag{1.2}$$

where v_i is the velocity vector and a comma denotes partial differentiation with respect to a rectangular Cartesian coordinate system x_i, i.e., $v_{i,j}=\partial v_i/\partial x_j$, and Ω_{ij} is some intrinsic microrotation tensor.

The purpose of the present work is to show that for gases with spatially and timewise nonuniform density and temperature fields the stress tensor is asymmetric while the couple stress tensor is zero. The asymmetry in the stress tensor is due to asymmetric interaction of gradient fields of density ρ and temperature θ in the form of $(\rho_{,j}\theta_{,k}-\rho_{,k}\theta_{,j})$.

First we write down the governing equations of motion and energy for a gas in the presence of asymmetric fields and then we employ a generalized entropy principle to arrive at the pertinent constitutive equations. We also make use of Truesdell's equipresence principle and let all constitutive variables appear in all constitutive functional representations.

By employing a method of Lagrange multipliers we arrive at the expressions for the stress and the couple stress. Finally we relate the asymmetric part of the stress tensor to the field of microrotations.

2. GOVERNING EQUATIONS OF ASYMMETRIC GAS DYNAMICS

Generally from a mathematical point of view, a microcontinuum is the one possessing internal degrees of microfreedom or internal fields[3] $B(x_i, t)$, such that B could be of any tensorial character. Among all microfluids the simplest emerges as the one we are going to consider. Namely, we consider gases with only one internal polar field $\omega_i(x_j, t)$. Further, ω_i characterizes a field of rigid rotations throughout the body $V(t)$ enclosed by a regular surface $S(t)$. The existence of $\omega_i(x_j, t)$ could imply the rigid microinclusions in the body which creates poles by rotation while moving with the gas. Therefore, in the above sense the gas is polar. We note that the presence of the ω_i field also corresponds to the presence of a rigidly rotating and moving frame comparable to the rigid triad of E. and F. Cosserat.[4] Shahinpoor and Ahmadi[5] have presented a simple approach in deriving the governing equations of motion and energy for such a polar medium and from them we find that, in the absence of body force, body couple, and any heat source, they reduce to

$$\dot{\rho}+\rho v_{k,k}=0, \tag{2.1}$$

$$\rho\dot{v}_i-\tau_{ki,k}=0, \tag{2.2}$$

$$\rho J\dot{\omega}_i-\mu_{ki,k}-\epsilon_{ikl}\tau_{kl}=0, \tag{2.3}$$

$$\rho\dot{\epsilon}+q_{k,k}-\tau_{ij}v_{j,i}-\mu_{ij}\omega_{j,i}=0, \tag{2.4}$$

where $\rho(x_i, t)$ is the mass density, a dot denotes total differentiation in time, v_i is the material velocity vector, ω_i is the local angular spin vector, J is the local constant micro-inertia, τ_{ki}, μ_{ki} are the stress and the couple stress tensors, respectively, defined as forces and moments per unit surface area, ϵ is the internal energy per unit mass, and q_k is the heat flux vector.

We, furthermore, employ a generalized entropy principle proposed by Müller.[6,7] It states that in every body there exists an additive scalar quantity, which is called entropy, which has a positive-definite production, so that if there is no supply of energy[8] in the body then

$$\rho\dot{\eta}+\Phi_{k,k}\geqslant 0, \tag{2.5}$$

where η is the specific entropy and Φ_k is its flux.

Our objective is to find thermodynamic fields in the gas by determining eight unknowns $\{\rho, v_i, \omega_i, \theta\}$ where the last unknown is the absolute temperature. Clearly the set (2.1)–(2.5) is not deterministic for the fields of unknowns and further relations are needed. These relations are the constitutive equations. Since we are particularly concerned with the effects of the spatial and timewise variations in density and temperature we consider the following set of independent variables:

$$\{\rho, \dot{\rho}, \rho_{,k}, \theta, \dot{\theta}, \theta_{,k}\}. \tag{2.6}$$

Employing Truesdell's equipresence principle[9] we consider the following constitutive relations:

$$\tau_{ij}=\hat{\tau}_{ij}(\rho, \dot{\rho}, \rho_{,k}, \theta, \dot{\theta}, \theta_{,k}), \tag{2.7}$$

$$\mu_{ij}=\hat{\mu}_{ij}(\rho, \dot{\rho}, \rho_{,k}, \theta, \dot{\theta}, \theta_{,k}), \tag{2.8}$$

$$q_i=\hat{q}_i(\rho, \dot{\rho}, \rho_{,k}, \theta, \dot{\theta}, \theta_{,k}), \tag{2.9}$$

$$\epsilon=\hat{\epsilon}(\rho, \dot{\rho}, \rho_{,k}, \theta, \dot{\theta}, \theta_{,k}), \tag{2.10}$$

$$\eta = \hat{\eta}(\rho, \dot{\rho}, \rho_{,k}, \theta, \dot{\theta}, \theta_{,k}), \tag{2.11}$$

$$\Phi_i = \hat{\Phi}_i(\rho, \dot{\rho}, \rho_{,k}, \theta, \dot{\theta}, \theta_{,k}). \tag{2.12}$$

Here τ_{ij} and μ_{ij} are generally nonsymmetric tensors. We further note that through the constitutive dependence on both ρ and $\dot{\rho}$ the constitutive functions (2.7)–(2.12) also depend on the gradients of dilatation $v_{k,k}$.

The constitutive functions have been assumed to be objective under Galilean transformations, which implies that they are isotropic with respect to the orthogonal group of transformations as is discussed by Noll.[10] This also implies that for any material there exists an isotropy group in the form of a group of orthogonal transformations which allows all admissible processes to remain admissible after a change of frame. Representation theorems are well known for (2.7)–(2.12) and, for example,

$$q_i \equiv -\kappa(\rho, \dot{\rho}, \theta, \dot{\theta}, I_1, I_2, I_3)\theta_{,i} + D(\rho, \dot{\rho}, \theta, \dot{\theta}, I_1, I_2, I_3)\rho_{,i}, \tag{2.13}$$

$$\Phi_i \equiv \phi(\rho, \dot{\rho}, \theta, \dot{\theta}, I_1, I_2, I_3)\theta_{,i} + \psi(\rho, \dot{\rho}, \theta, \dot{\theta}, I_1, I_2, I_3)\rho_{,i}, \tag{2.14}$$

where

$$I_1 = \theta_{,k}\theta_{,k}, \quad I_2 = \rho_{,k}\rho_{,k}, \quad I_3 = \theta_{,k}\rho_{,k}. \tag{2.15}$$

We call every set of solutions of (2.1)–(2.12) existing in the close neighborhood $N(x_i, t)$, a thermodynamic set and every solution in $N(x_i, t)$ a thermodynamic process for the nonsimple gaseous medium under consideration. Not allowing any shock wave dissipation of energy assures us that the functions in (2.7)–(2.12) are analytic.

Once (2.7)–(2.12) are defined over the space of all admissible functions ρ, v_i, ω_i, and θ, then the system of equations (2.1)–(2.5) gives rise to acceptable solutions for which the inequality (2.5) must hold for all ρ, v_i, ω_i, and θ. In effect, (2.5) puts additional restrictions on all constitutive relations (2.7)–(2.12). To evaluate such restrictions we choose to employ the method of Lagrange multipliers proposed by Liu,[11] according to which inequality (2.5) is equivalent to

$$\begin{aligned}&\rho\dot{\eta} + \Phi_{k,k} - \Lambda^{\rho}(\dot{\rho} + \rho v_{k,k}) - \Lambda^{v_i}(\rho\dot{v}_i - \tau_{ki,k}) \\ &\quad - \Lambda^{\omega_i}(\rho J\dot{\omega}_i - \mu_{ki,k} - \epsilon_{ikl}\tau_{kl}) \\ &\quad - \Lambda^{\epsilon}(\rho\dot{\epsilon} + q_{k,k} - \tau_{ij}v_{j,i} - \mu_{ij}\omega_{j,i}) \geq 0,\end{aligned} \tag{2.16}$$

holding for all admissible arbitrary values of ρ, $\dot{\rho}$, $\rho_{,i}$, θ, $\dot{\theta}$, $\theta_{,i}$, v_i, ω_i, $\dot{v}_i$, $\dot{\omega}_i$, $\dot{v}_i$, $\dot{\omega}_i$, $\ddot{\rho}$, $\dot{\rho}_{,k}$, $\ddot{\theta}$, $\dot{\theta}_{,k}$, $\rho_{,kl}$, $\dot{\theta}_{,k}$, $\theta_{,kl}$, $v_{k,l}$, and $\omega_{k,l}$. The set $\{\Lambda^{\rho}, \Lambda^{v_i}, \Lambda^{\omega_i}, \Lambda^{\epsilon}\}$ comprises a set of Lagrange multipliers over the eight-dimensional space of $(\rho, v_i, \omega_i, \theta)$ and are in general functions of the above mentioned variables.

Since the laws governing the thermodynamical behavior of the gas, namely Eqs. (2.7)–(2.12), do not depend on

$$\left\{\frac{\partial^2\rho}{\partial t^2}, \frac{\partial\rho_{,k}}{\partial t}, \frac{\partial^2\theta}{\partial t^2}, \frac{\partial\theta_{,k}}{\partial t}, \rho_{,kl}, \theta_{,kl}, v_{k,l}, \frac{\partial v_k}{\partial t}, \omega_{k,l}, \frac{\partial\omega_k}{\partial t}\right\}, \tag{2.17}$$

therefore arbitrary variations of such quantities should not alter the nature of inequality (2.5). Thus the coefficients of all such fields in (2.16) must be set equal to zero. This procedure yields:

$$\frac{\partial\hat{\eta}}{\partial\dot{\rho}} - \Lambda^{\epsilon}\frac{\partial\hat{\epsilon}}{\partial\dot{\rho}} = 0, \tag{2.18}$$

$$\rho\left(\frac{\partial\hat{\eta}}{\partial\rho_{,k}} - \Lambda^{\epsilon}\frac{\partial\hat{\epsilon}}{\partial\rho_{,k}}\right) + \left(\frac{\partial\hat{\Phi}_k}{\partial\dot{\rho}} - \Lambda^{\epsilon}\frac{\partial\hat{q}_k}{\partial\dot{\rho}}\right) + \Lambda^{v_i}\frac{\partial\hat{\tau}_{ki}}{\partial\dot{\rho}} + \Lambda^{\omega_i}\frac{\partial\hat{\mu}_{ki}}{\partial\dot{\rho}} = 0, \tag{2.19}$$

$$\frac{\partial\hat{\eta}}{\partial\dot{\theta}} - \Lambda^{\epsilon}\frac{\partial\hat{\epsilon}}{\partial\dot{\theta}} = 0, \tag{2.20}$$

$$\rho\left(\frac{\partial\hat{\eta}}{\partial\theta_{,k}} - \Lambda^{\epsilon}\frac{\partial\hat{\epsilon}}{\partial\theta_{,k}}\right) + \left(\frac{\partial\hat{\Phi}_k}{\partial\dot{\theta}} - \Lambda^{\epsilon}\frac{\partial\hat{q}_k}{\partial\dot{\theta}}\right) + \Lambda^{v_i}\frac{\partial\hat{\tau}_{ki}}{\partial\dot{\theta}} + \Lambda^{\omega_i}\frac{\partial\hat{\mu}_{ki}}{\partial\dot{\theta}} = 0, \tag{2.21}$$

$$\frac{\partial\hat{\Phi}_{(k}}{\partial\rho_{,j)}} + \Lambda^{v_i}\frac{\partial\hat{\tau}_{(ki}}{\partial\rho_{,j)}} + \Lambda^{\omega_i}\frac{\partial\hat{\mu}_{(ki}}{\partial\rho_{,j)}} - \Lambda^{\epsilon}\frac{\partial\hat{q}_{(k}}{\partial\rho_{,j)}} = 0, \tag{2.22}$$

$$\frac{\partial\hat{\Phi}_{(k}}{\partial\theta_{,j)}} + \Lambda^{v_i}\frac{\partial\hat{\tau}_{(ki}}{\partial\theta_{,j)}} + \Lambda^{\omega_i}\frac{\partial\hat{\mu}_{(ki}}{\partial\theta_{,j)}} - \Lambda^{\epsilon}\frac{\partial\hat{q}_{(k}}{\partial\theta_{,j)}} = 0, \tag{2.23}$$

$$\left(\frac{\partial\hat{\Phi}_k}{\partial\dot{\rho}} - \Lambda^{\epsilon}\frac{\partial\hat{q}_k}{\partial\dot{\rho}}\right)\rho_{,j} + \left(\frac{\partial\hat{\Phi}_k}{\partial\dot{\theta}} - \Lambda^{\epsilon}\frac{\partial\hat{q}_k}{\partial\dot{\theta}}\right)\theta_{,j} - \rho\Lambda^{v_i}v_k - \rho\delta_{kj}\Lambda^{\rho} + \Lambda^{\epsilon}\hat{\tau}_{kj} = 0, \tag{2.24}$$

$$\Lambda^{v_i} = 0, \tag{2.25}$$

$$\Lambda^{\epsilon}\hat{\mu}_{ki} = 0, \tag{2.26}$$

$$J\Lambda^{\omega_i} = 0. \tag{2.27}$$

The inequality (2.16) now shrinks to

$$\begin{aligned}&\rho\left(\frac{\partial\hat{\eta}}{\partial\theta} - \Lambda^{\epsilon}\frac{\partial\hat{\epsilon}}{\partial\theta}\right)\dot{\theta} + \left(\frac{\partial\hat{\Phi}_k}{\partial\rho} + \Lambda^{v_i}\frac{\partial\hat{\tau}_{ki}}{\partial\rho} + \Lambda^{\omega_i}\frac{\partial\hat{\mu}_{ki}}{\partial\rho} - \Lambda^{\epsilon}\frac{\partial\hat{q}_k}{\partial\rho}\right)\rho_{,k} + \left(\frac{\partial\hat{\Phi}_k}{\partial\theta} + \Lambda^{v_i}\frac{\partial\hat{\tau}_{ki}}{\partial\theta} + \Lambda^{\omega_i}\frac{\partial\hat{\mu}_{ki}}{\partial\theta} - \Lambda^{\epsilon}\frac{\partial\hat{q}_k}{\partial\theta}\right)\theta_{,k} \\ &\quad + \left[\rho\left(\frac{\partial\hat{\eta}}{\partial\rho} - \Lambda^{\epsilon}\frac{\partial\hat{\epsilon}}{\partial\rho}\right) - \Lambda^{\rho}\right]\dot{\rho} + \Lambda^{\omega_i}\epsilon_{ikl}\hat{\tau}_{kl} \geq 0.\end{aligned} \tag{2.28}$$

We note from (2.25) that $\Lambda^{v_i} = 0$. However, from (2.27) we conclude that $\Lambda^{\omega_i} = 0$ only if the microinertia J is different from zero. From (2.24) it is clear that the stress can be nonsymmetric. We assume $J > 0$ and thus $\Lambda^{\omega_i} = 0$. Thus, inequality (2.28) reduces to

$$\rho\left(\frac{\partial\hat{\eta}}{\partial\theta} - \Lambda^{\epsilon}\frac{\partial\hat{\epsilon}}{\partial\theta}\right)\dot{\theta} + \left(\frac{\partial\hat{\Phi}_k}{\partial\theta} - \Lambda^{\epsilon}\frac{\partial q_i}{\partial\theta}\right)\theta_{,i} + \left(\frac{\partial\Phi_k}{\partial\rho} - \Lambda^{\epsilon}\frac{\partial\hat{q}_k}{\partial\rho}\right)\rho_{,k} + \left[\rho\left(\frac{\partial\hat{\eta}}{\partial\rho} - \Lambda^{\epsilon}\frac{\partial\hat{\epsilon}}{\partial\rho}\right) - \Lambda^{\rho}\right]\dot{\rho} \geq 0. \tag{2.29}$$

Furthermore, the conditions (2.18)–(2.29) must hold for arbitrary values of ρ, θ, $\dot{\rho}$, $\rho_{,i}$, $\dot{\theta}$, $\theta_{,i}$, v_i, and ω_i. We then conclude that Λ^{ρ} and Λ^{ϵ} are scalar isotropic functions of the variables $\rho, \dot{\rho}, \rho_{,i}, \theta, \dot{\theta}, \theta_{,i}$ and in particular

$$\Lambda^{\epsilon} = \Lambda^{\epsilon}(\rho, \dot{\rho}, \theta, \dot{\theta}, I_1, I_2, I_3). \tag{2.30}$$

We explore the case when $\hat{\mu}_{ik} = 0$ and thus $\Lambda^{\epsilon} \neq 0$.

With representations (2.13) and (2.14) Eqs. (2.22) and (2.23) lead to

$$\Phi_i = \Lambda^\epsilon q_i , \tag{2.31}$$

and the conclusion that Λ^ϵ is independent of I_1, I_2, and I_3 provided that $\kappa \neq 0$. This being true, we can write

$$\Lambda^\epsilon = \Lambda^\epsilon(\rho, \dot{\rho}, \theta, \dot{\theta}). \tag{2.32}$$

3. GENERAL STRUCTURE OF THE STRESS AND THE COUPLE STRESS TENSORS

From (2.26) if $\Lambda^\epsilon \neq 0$, then $\hat{\mu}_{ki} = 0$. Now from (2.31) and (2.24) it is found that

$$\hat{\tau}_{kj} = \frac{\Lambda^\rho}{\Lambda^\epsilon}\rho\delta_{kj} + Q_1\theta_{,k}\rho_{,j} - Q_1\left(\frac{D}{\kappa}\right)\rho_{,k}\rho_{,j} + Q_2\theta_{,k}\theta_{,j} - Q_2\left(\frac{D}{\kappa}\right)\rho_{,k}\theta_{,j}, \tag{3.1}$$

where

$$Q_1 = \kappa\frac{\partial}{\partial\dot{\rho}}(\ln\Lambda^\epsilon), \quad Q_2 = \kappa\frac{\partial}{\partial\dot{\theta}}(\ln\Lambda^\epsilon). \tag{3.2}$$

By means of the trace of (2.1) we can eliminate Λ^ρ and rewrite (3.1) in the form

$$(\tau_{kj} - \tfrac{1}{3}\tau_{ii}\delta_{kj}) = Q_1\theta_{,k}\rho_{,j} - Q_1\left(\frac{D}{\kappa}\right)(\rho_{,k}\rho_{,j} - \tfrac{1}{3}\rho_{,i}\rho_{,i}\delta_{kj}) + Q_2(\theta_{,k}\theta_{,j} - \tfrac{1}{3}\theta_{,i}\theta_{,i}\delta_{kj}) - Q_2\left(\frac{D}{\kappa}\right)\rho_{,k}\theta_{,j} - \tfrac{1}{3}\left[Q_1 - Q_2\left(\frac{D}{\kappa}\right)\right]\rho_{,i}\theta_{,i}\delta_{kj}. \tag{3.3}$$

This is the deviatoric part of the stress tensor. As can be seen the stress tensor is generally in this case nonsymmetric. The antisymmetric part of the stress tensor can be found from either (3.1) or (3.3) to be

$$\tau_{[kj]} = \tfrac{1}{2}\left[Q_1 + Q_2\left(\frac{D}{\kappa}\right)\right](\rho_{,j}\theta_{,k} - \rho_{,k}\theta_{,j}). \tag{3.4}$$

Note from (2.3) that

$$\tau_{[kj]} = \tfrac{1}{2}\rho J\epsilon_{ikj}\dot{\omega}_i . \tag{3.5}$$

It must be mentioned here that if $\tau_{[kj]}$ is assumed to be equal to zero, then it can be easily shown that on the foundation of the validity of Fourier's law of conduction with θ interpreted as a temperature field the following representations are valid:

$$\epsilon = \epsilon(\rho, \theta, \dot{\theta}, I_1), \tag{3.6}$$

$$q_i = -\kappa(\rho, \theta, \dot{\theta}, I_1)\theta_{,i}, \tag{3.7}$$

$$\tau_{kj} = -p(\rho, \dot{\rho}, \theta, \dot{\theta}, I_1, I_2, I_3)\delta_{kj} + Q_2(\rho, \theta, \dot{\theta}, I_1)\theta_{,k}\theta_{,j}, \tag{3.8}$$

$$\eta = \eta(\rho, \theta, \dot{\theta}, I_1). \tag{3.9}$$

$$\Phi_i = -\Lambda^\epsilon(\theta, \dot{\theta})\kappa(\rho, \theta, \dot{\theta}, I_1)\theta_{,i}, \tag{3.10}$$

$$\frac{\partial\eta}{\partial\dot{\theta}} = \Lambda^\epsilon\frac{\partial\epsilon}{\partial\dot{\theta}}, \tag{3.11}$$

$$\frac{\partial\eta}{\partial I_1} = \Lambda^\epsilon\frac{\partial\epsilon}{\partial I_1} + \frac{\kappa}{2\rho}\frac{\partial\Lambda^\epsilon}{\partial\dot{\theta}} = \Lambda^\epsilon\left(\frac{\partial\epsilon}{\partial I_1} + \frac{Q_2}{2\rho}\right), \tag{3.12}$$

$$\frac{Q_2}{\kappa} = -\frac{\partial Q_2/\partial\dot{\theta}}{2\rho\,\partial\epsilon/\partial I_1 + Q_2}, \tag{3.13}$$

$$\frac{\partial}{\partial\theta}\left(\ln\frac{\partial\Lambda^\epsilon}{\partial\dot{\theta}}\right) = -\left(2\rho\frac{\partial\epsilon}{\partial I_1} - \frac{\partial\kappa}{\partial\dot{\theta}}\right)\Big/\kappa, \tag{3.14}$$

and the residual inequality

$$\rho\left[\frac{\partial\eta}{\partial\rho} - \Lambda^\epsilon\left(\frac{\partial\epsilon}{\partial\rho} - \frac{p}{\rho^2}\right)\right]\dot{\rho} + \rho\left(\frac{\partial\eta}{\partial\theta} - \Lambda^\epsilon\frac{\partial\epsilon}{\partial\theta}\right)\dot{\theta} - \kappa\frac{\partial\Lambda^\epsilon}{\partial\theta}\theta_{,i}\theta_{,i} \geq 0, \tag{3.15}$$

holds.

In this particular case $\Lambda^\epsilon(\theta, \dot{\theta})$ is the "coldness" as suggested by Müller.[6] This universal function will equal the inverse absolute temperature at equilibrium, i.e., when equality holds for (3.15).

4. GOVERNING THERMODYNAMIC EQUATIONS FOR AN IDEAL ASYMMETRIC GAS

Let us define an ideal asymmetric gas by assuming the following relations to be true:

$$\epsilon = \alpha_1\theta + \alpha_0, \quad \alpha_1, \alpha_0 \equiv \text{constants} > 0, \tag{4.1}$$

$$q_i = -\kappa\theta_{,i} + D\rho_{,i}, \quad \kappa, D \equiv \text{constants}, \quad \kappa > 0, \tag{4.2}$$

$$-\tfrac{1}{3}\tau_{ii} = \alpha_2\rho\theta, \tag{4.3}$$

$$\Lambda^\epsilon = \alpha_3\theta^{-1}e^{\alpha_4\dot{\theta}/\kappa}, \quad Q_1 = 0, \quad Q_2 = \alpha_4 = \text{const}, \quad \alpha_3 = \text{const}. \tag{4.4}$$

The stress tensor now reduces to

$$\tau_{kj} = -\alpha_2\rho\theta\delta_{kj} + \alpha_4(\theta_{,k}\theta_{,j} - \tfrac{1}{3}\theta_{,i}\theta_{,i}\delta_{kj}) - \alpha_4\left(\frac{D}{\kappa}\right)\rho_{,k}\theta_{,j} + \tfrac{1}{3}\alpha_4\left(\frac{D}{\kappa}\right)\rho_{,i}\theta_{,i}\delta_{kj}, \tag{4.5}$$

and the complete set of governing thermodynamic equations for an ideal asymmetric gas reduce to

$$\dot{\rho} + \rho v_{k,k} = 0, \tag{4.6}$$

$$\rho\dot{v}_j = \left[-\alpha_2\rho_{,k}\theta - \alpha_2\rho\theta_{,k} - \frac{2}{3}\alpha_4\theta_{,ik}\theta_i + \frac{1}{3}\alpha_4\left(\frac{D}{\kappa}\right)\rho_{,ik}\theta_{,i} + \frac{1}{3}\alpha_4\left(\frac{D}{\kappa}\right)\rho_{,i}\theta_{,ik}\right]\delta_{kj} + \alpha_4(\theta_{,kk}\theta_{,j} + \theta_{,k}\theta_{,jk}) - \alpha\left(\frac{D}{\kappa}\right)(\rho_{,kk}\theta_{,j} + \rho_{,k}\theta_{,jk}), \tag{4.7}$$

$$\rho J\dot{\omega}_i = \tfrac{1}{2}\epsilon_{ikj}\alpha_4\left(\frac{D}{\kappa}\right)[\rho_{,j}\theta_{,k} - \rho_{,k}\theta_{,j}], \tag{4.8}$$

$$\rho\alpha_1\dot{\theta} = +\kappa\theta_{,kk} - D\rho_{,kk} + v_{j,k}\Big\{\Big[-\alpha_2\rho\theta - \tfrac{1}{3}\alpha_4\theta_{,i}\theta_{,i} - \tfrac{1}{3}\alpha_4\left(\frac{D}{\kappa}\right)\rho_{,i}\theta_{,i}\Big]\delta_{kj} + \alpha_4\theta_{,k}\theta_{,j} - \alpha_4\left(\frac{D}{\kappa}\right)\rho_{,k}\theta_{,j}\Big\}.$$

Thus we have a set of eight nonlinear coupled partial differential equations in the eight unknowns ρ, v_i, ω_i, and θ for the governing thermodynamic equations of an ideal asymmetric gas.

5. CONCLUDING REMARKS

As can be clearly seen from the general expressions (3.1), (3.3), (3.4) or the special expression (4.5) for the stress tensor in asymmetric gases the asymmetry in the stress tensor arises because of the presence of an asymmetric differential operator $(\rho_{,j}\theta_{,k}-\rho_{,k}\theta_{,j})$ in the expression for the stress tensor. This is what we meant in the Introduction by "the asymmetric interaction of gradient fields of density and temperature."

For the case of an ideal asymmetric gas the assumptions (4.1) and (4.3) are clearly motivated by the classical assumptions for the thermodynamic behavior of an ideal gas with $(-\frac{1}{3}\tau_{ii})$ interpreted as the hydrostatic pressure p. The assumption (4.2) is the simplest extension to the classical Fourier's law of heat conduction in order to have nonsymmetric stress tensor fields. Clearly D being zero would reduce the stress tensor to a symmetric tensor unless $Q_1 \neq 0$. The assumption (4.4) is motivated by the fact that if both Q_1 and Q_2 are zero, then Λ^{ϵ} would reduce to an inverse temperature scale which is typical for ideal gases.

*Professor of Mechanical and Aerospace Engineering. Presently: Principal Research Scientist, Department of Mechanics and Materials Science, The Johns Hopkins University, Baltimore, Md.

[1]T. Ariman, M. A. Turk, and N. D. Sylvester, Int. J. Eng. Sci. **11**, 905–30 (1973).
[2]M. Shahinpoor, Iran. J. Sci. Technol. (4), 133–42 (1975).
[3]We are using rectangular Cartesian coordinates $x_i \equiv (x_1, x_2, x_3)$ and time t. Commas shall denote partial differentiations with respect to x_i.
[4]E. and F. Cosserat, *Théorie des corps deformables* (Hermann, Paris, 1909).
[5]M. Shahinpoor and G. Ahmadi, Arch. Ration. Mech. Anal. **47**, 188–94 (1972).
[6]I. Müller, Arch. Ration. Mech. Anal. **41**, 391–22 (1971).
[7]I. Müller, Habilitationsschrift an der RWTH Aachen (1970); Arch. Ration. Mech. Anal. **40**, 1–36 (1971).
[8]Chemical reactions, radiations, etc.
[9]C. Truesdell and W. Noll, *Handbuch der Physik*, *III*/3 (Springer-Verlag, Berlin, 1965).
[10]W. Noll, Arch. Ration. Mech. Anal. **52**, 62–92 (1973).
[11]I. Shih Liu, Arch. Ration. Mech. Anal. **46**, 131–48 (1972).

Classical systems of infinitely many noninteracting particles

J. V. Pulè* and A. Verbeure

Universiteit Leuven, Belgium †
(Received 27 April 1976)

It is proved that the C^*-algebra of observables of an infinite classical system is isomorphic to the group algebra on the test function space $\mathcal{D}$. The physical dynamical system consisting of infinitely many noninteracting particles is studied. A particular class of states, called the quasifree states, is exhibited and their properties are studied. Some results on the spectral properties of monoparticle evolutions are obtained. Finally we give explicitly a solution of the classical KMS condition for these evolutions.

1. INTRODUCTION

Since the work of Haag and Kastler[1] much progress has been made in the mathematically rigorous description of quantum systems with infinitely many degrees of freedom. In particular, statistical mechanics benefitted from their approach. In this context of special importance, for physics as well as mathematics, was the work of Haag, Hugenholtz, and Winnink[2] where they studied the properties of infinite systems satisfying the KMS condition.

On the other hand, for classical infinite systems much less has been done in this direction. We mention the results of the algebraic approach on thermodynamic functions to be found in Ruelle's book.[3] A new impetus has been given to the study of classical systems by the recent work on the classical KMS condition.[4–6] Motivated by these works, and by the fact that to our knowledge the only Hamiltonian evolution known for an infinite system is that of infinitely many free particles,[7] we describe explicitly the class of states which naturally should describe noninteracting infinitely many particles. We call these states, in analogy with quantum mechanics,[8] quasifree states. To that end we construct in Sec. 2, in a "canonical way" the algebra, generated by the exponential of unbounded observables on the configuration space, in analogy with the Weyl algebra. We prove that it is isomorphic to the group algebra on the set of test functions $\mathcal{D}$. This enables us to work with the latter algebra making unnecessary the use of the classical phase space, which is rather difficult to handle.

On this algebra the set of quasifree states is defined and its properties are studied; in particular we give its GNS representation on Fock space.

In Sec. 4 the quasifree dynamics is introduced through monoparticle evolutions, and some results on the spectrum are obtained.

We prove that the spectrum of the Liouvillian for the infinite free system is absolutely continuous except for the point zero (Theorem 4.2).

Finally in Sec. 5 we study the quasifree states satisfying the KMS conditions and reduce the problem of the infinite system to the one-particle case. We prove that for quasifree evolutions given by a Hamiltonian there always exists a solution of the KMS equation which is quasifree. We do not go into the uniqueness of these solutions.

2. THE ALGEBRA OF OBSERVABLES AS A GROUP ALGEBRA

By analogy with the Weyl algebra in quantum mechanics, we construct a new algebra of observables for the classical infinite systems, which has the advantage of being generated by the exponential function of test functions. Our algebra is contained in the one of Ruelle.[3]

Let $\mathcal{K}$ be the set of infinite countable subsets x of $R^d \times R^d$ $d \in N$ such that for every bounded subset V of R^d, $x \cap V \times R^d$ contains only a finite number of elements. $\mathcal{K}$ is called the set of configurations. Let $\mathcal{D}$ be the set of real C^∞ functions of bounded support on $R^d \times R^d$. For each element $f \in \mathcal{D}$ define the function Sf from $\mathcal{K}$ to R by

$$(Sf)(x) = \sum_i f(x_i), \quad x \in \mathcal{K}.$$

Denote by $W(f)$ the bounded function on $\mathcal{K}$ defined by

$$W(f) = \exp iSf$$

and let $\mathcal{A}$ be the complex Abelian C^*-algebra generated by the set $\{W(f) \mid f \in \mathcal{D}\}$ with respect to the usual pointwise multiplication of functions, involution the complex conjugation, and norm the supremum norm over $\mathcal{K}$. We call $\mathcal{A}$ the set of observables of the infinite classical system.

Now we define a particular class of states on the algebra $\mathcal{A}$, through its family of density distributions $(\mu_\Lambda^n)_{n=0}^\infty$ for each bounded open set Λ of R^d. Let σ be any nonnegative Lebesgue measurable function on $R^d \times R^d$ such that

$$w(\Lambda) \equiv \int_{\Lambda \times R^d} \sigma(x)\, dx$$

is finite, then let the family of density distributions be given by

$$d\mu_\Lambda^n(x_1, \ldots, x_n) = \frac{\exp[-w(\Lambda)]}{n!} \prod_{i=1}^n \sigma(x_i)\, dx_i$$

for all $n \geq 1$ and $\mu_\Lambda^0 = \exp[-w(\Lambda)]$.

It is clear that (μ_Λ^n) satisfies:

(a) the normalization condition.

$$1 = \sum_{n=0}^\infty \int_{(\Lambda \times R^d)^n} d\mu_\Lambda^n(x_1, \ldots, x_n),$$

(b) the compatibility condition, i.e., for each $\Lambda' \supset \Lambda$

$$d\mu_\Lambda^n(x_1, \ldots, x_n) = \sum_{p=0}^\infty \frac{(n+p)!}{n!\,p!} \int \cdots \int_{x_{n+1}, \ldots, x_{n+p} \in \Lambda' \setminus \Lambda \times R^d} d\mu^{n+p}(x_1, \ldots, x_{n+p}).$$

Given any family of density distributions (μ^n_Λ), they determine a state ω on $\mathcal{A}$ by

$$\omega(X)=\sum_{n=0}^{\infty}\int d\mu^n_\Lambda(x_2,\ldots,x_n)X(x_1,\ldots,x_n)$$

for $X\in\mathcal{A}$ with "support" in Λ. Any state determined by a family of density distributions of the above type will be called a product state.

It is easily checked that if ω is a product state determined by the function σ then for all $f\in\mathcal{D}$

$$\omega(\exp iSf)=\exp\left(\int_{R^d\times R^d}\sigma(x)[\exp(if(x))-1]\,dx\right). \qquad (1)$$

Now we define the group algebra of $\mathcal{D}$. Let $\Delta(\mathcal{D})$ be the set of complex valued functions on $\mathcal{D}$ which vanish except on a finite number of elements of $\mathcal{D}$; $\Delta(\mathcal{D})$ equipped with the following addition, multiplication, involution, and norm:

$$(a+b)(f)=a(f)+b(f),$$
$$(\alpha a)(f)=\alpha a(f),$$
$$(a\cdot b)(f)=\sum_{g\in\mathcal{D}}a(g)b(f-g),$$
$$a^*(f)=\overline{a(-f)},$$
$$\|a\|_1=\sum_{f\in\mathcal{D}}|a(f)|,$$

for all $a,b\in\Delta(\mathcal{D})$, $\alpha\in C$, $f,g\in\mathcal{D}$, is a normed involutive algebra. Denote by δ_f the element of $\Delta(\mathcal{D})$ given by

$$\delta_f(g)=0 \quad \text{if } f\neq g,$$
$$\delta_f(g)=1 \quad \text{if } f=g.$$

Note that $\delta_f\delta_g=\delta_{f+g}$ for all $f,g\in\mathcal{D}$, δ_0 is the unit element of the algebra $\Delta(\mathcal{D})$, $\delta_f^*=\delta_{-f}$ and that the set $\{\delta_f\,|\,f\in\mathcal{D}\}$ is a basis for $\Delta(\mathcal{D})$.

Denote by $\Delta(\mathcal{D})_1$ the closure of $\Delta(\mathcal{D})$ with respect to the norm $\|\circ\|_1$. For each element $a\in\Delta(\mathcal{D})$ the map from $\Delta(\mathcal{D})$ into R,

$$\|\cdot\|:a\to\|a\|=\sup_\pi\|\pi\|(a)\|,$$

where the sup means the supremum over all nondegenerate representations π of $\Delta(\mathcal{D})$; $\|\circ\|$ is a C^* norm.

Consider now $\overline{\Delta(\mathcal{D})}$, the completion of $\Delta(\mathcal{D})$ with respect to this norm; it is a C^*-algebra, called the C^*-group algebra of $\mathcal{D}$.

In the following theorem we prove that C^*-algebras $\mathcal{A}$ and $\overline{\Delta(\mathcal{D})}$ are isomorphic. This result simplifies the study of classical dynamical systems when the configuration space is not involved, in particular the study of states on the algebra.

Theorem 2.1: The C^*-algebra of classical observables $\mathcal{A}$ is isomorphic to the group algebra $\overline{\Delta(\mathcal{D})}$ of $\mathcal{D}$.

Proof: Let ϕ be the linear map of $\Delta(\mathcal{D})$ into $\mathcal{A}$ given by

$$\phi(\delta_f)=\exp iSf.$$

It is clear that ϕ is a $*$-homomorphism of $\Delta(\mathcal{D})$ into $\mathcal{A}$. Now we prove that ϕ is injective. Let $f_1,\ldots,f_n$ be different elements of $\mathcal{D}$, $a_1,\cdots,a_n$ be complex numbers and suppose that

$$\phi\left(\sum_{j=1}^{n}a_j\delta_{f_j}\right)=0.$$

Then

$$\sum_{j=1}^{n}a_j\exp iSf_j=0$$

and for any k

$$a_k+\sum_{j\neq k}a_j\exp[iS(f_j-f_k)]=0. \qquad (*)$$

Consider the product states ω_λ determined by the function $\sigma_\lambda(q,p)=\lambda\exp(-p^2)$ with λ a positive number. Then applying the state ω_λ on (*), and using (1), one gets

$$a_k+\sum_{j\neq k}a_j\exp\{-\lambda[\tfrac{1}{2}\int e^{-p^2}\,dq\,dp\,|\exp[i(f_j-f_k)]-1|^2$$
$$-i\,\mathrm{Im}\int e^{-p^2}\,dq\,dp\,(\exp[i(f_j-f_k)]-1)]\}.$$

Since the functions f_j are continuous $\int\exp(-p^2)\,dq\,dp\,|\exp[i(f_j-f_k)]-1|^2\neq0$ for $j\neq k$, hence by letting λ tend to infinity one gets $a_k=0$. As this is true for all k, we proved that ϕ is injective.

Since ϕ is a $*$-homomorphism of a dense subalgebra $\Delta(\mathcal{D})$ onto a dense subalgebra $\phi(\Delta(\mathcal{D}))$, we have that for any $a\in\Delta(\mathcal{D})$,

$$a^*a\leq\|a\|^2 \quad\text{and}\quad \phi(a^*a)\leq\|a\|^2.$$

Therefore,

$$\|\phi(a)\|^2=\|\phi(a)^*\phi(a)\|=\|\phi(a^*a)\|\leq\|a\|^2.$$

Since ϕ is a $*$-isomorphism, the inverse ϕ^{-1} also satisfies

$$\|\phi^{-1}(a)\|\leq\|a\|,$$

hence

$$\|a\|=\|\phi(a)\|$$

and the theorem follows by continuity.

Q.E.D.

3. QUASIFREE STATES

Denote by $\tilde{\mathcal{D}}$ the set of functions given by $\tilde{\mathcal{D}}=\{f+ig\,|\,f,g\in\mathcal{D}\}$, i.e., $\tilde{\mathcal{D}}$ is the complex algebra generated by the algebra $\mathcal{D}$.

Theorem 3.1: Let ω_0 be any positive linear functional on $\tilde{\mathcal{D}}$, and let ω be the linear functional on $\Delta(\mathcal{D})$ defined by

$$\omega(\delta_f)=\exp[\omega_0(\exp(if)-1)], \quad f\in\mathcal{D}.$$

Then

(i) ω extends to a state (i.e., positive normalized linear form) on $\overline{\Delta(\mathcal{D})}$

(ii) for all f and g in $\mathcal{D}$ with disjoint supports

$$\omega(\delta_f\delta_g)=\omega(\delta_f)\,\omega(\delta_g).$$

Proof: (i) Let $c_1,\ldots,c_N$ be arbiitrary complex numbers, and $g_1,\ldots,g_N$ be arbitrary in $\mathcal{D}$,

$$\sum_{n,m=1}^{N}c_n c_m\omega(\delta_{g_n-g_m})$$
$$=\sum_{n,m=1}^{N}c_n c_m\exp[\omega_0(\exp[i(g_n-g_m)]-1)]. \qquad (\alpha)$$

But, using

$$[\exp(ig_n)-1][\exp(-ig_m)-1]-(1-\exp ig_n)-[1-\exp(-ig_m)] \\ =\exp[i(g_n-g_m)]-1.$$

(α) becomes

$$\sum_{n,m=1}^{N} \mu_n\mu_m \exp(\psi_m, \psi_n), \qquad (\beta)$$

where $\mu_n = c_n \exp[\omega_0(\exp(ig_n)-1)]$ and $(\cdot,\cdot)$ in the sesquilinear form on $\tilde{D}$ given by

$$(\psi,\phi)=\omega_0(\bar{\psi}\phi) \quad \psi,\phi\in\tilde{D}$$

and $\psi_n=\exp(ig_n)-1$.

It is well known from the theory of positive definite type functions that the expression (β) is positive. This proves the positivity of the linear functional ω. Furthermore, as $\omega(\delta_0)=1$ and the algebra is Abelian, ω is a continuous positive normalized linear form on $\Delta(D)$. Therefore, ω extends to a state on $\overline{\Delta(D)}$.

(ii) If f and g have disjoint supports then $[\exp(ig)-1][\exp(if)-1]=0$ and therefore $\exp[i(f+g)]-1=[\exp(if)-1]+[\exp(ig)-1]$ and (ii) follows.

Q.E.D.

This theorem shows that any positve linear functional ω_0 on the one-particle space $\tilde{D}$ gives rise to a state ω of the infinite system. All states of this form will be called quasifree states

Furthermore, any state ω on $\overline{\Delta(D)}$ such that the map $\lambda\in R\to\omega(\delta_{\lambda f+g})$ is continuous for all f and g in D, will be called a regular state.

Let ω be a regular state and (π,H,Ω) be its GNS representation, then it is easily checked that for all $f\in D$, $\{\pi(\delta_{\lambda f})\mid\lambda\in R\}$ is a strongly continuous one-parameter group of unitaries on H. Hence by Stone's theorem there exists a self-adjoint operator $B(f)$ on H such that

$$\pi(\delta_f)=\exp[iB(f)]$$

and $B(\circ)$ is linear on D; $B(\cdot)$ is called the classical field operator

Moreover, any state ω on $\Delta(D)$ such that the map $\lambda\in R\to\omega(\delta_{\lambda f+g})$ is infinitely differentiable for all f and g in D will be called a C^∞ state.

A C^∞ state is always a regular state, moreover it shares the property that for all $f_1,\dots,f_n$ and g in D:

$B(f_1)\cdots B(f_n)\Omega$ belongs to the domain of $B(g)$. For each C^∞ state we can define the truncated functions ω_T by the following recurrence relation: For each set $g_1,\dots,g_n$ in D,

$$(\Omega, B(g_1)\cdots B(g_n)\Omega)=\sum\omega_T(g_1\cdots)\cdots\omega_T(\cdots g_{i_n}), \qquad (2)$$

where the summation is over all possible partitions $(i_1,\cdots),\dots,(\cdots,i_n)$ of the set $\{1,\dots,n\}$, within each cluster the original order being preserved.

Theorem 3.2: Let ω be a quasifree C^∞ state, determined by the functional ω_0 on $\tilde{D}$, then for each set $g_1,\dots,g_n$ in D one has $\omega_T(g_1\cdots g_n)=\omega_0(g_1\cdots g_n)$.

Proof: For quasifree states the differentiability of $\lambda\to\omega(\delta_{\lambda f+g})$ ie equivalent with the differentiability of $\lambda\to\omega_0(\exp[i(\lambda f+g)]-1)$, and for $n=1$, the relation between ω_T and ω_0 follows by differentiating $\lambda\to\omega(\delta_{\lambda g})$ at $\lambda=0$. From the definition of ω_T [see Eq. (2)],

$$\omega_T(g_1\cdots g_n)=(\Omega,B(g_1)\cdots B(g_n)\Omega) \\ -\sum{}'\omega_T(g_{i_1}\cdots)\cdots\omega_T(\cdots g_{i_n}),$$

where $\sum'$ stands for the summation $\sum$ without the trivial partition $(1,\dots,n)$. As now

$$(\Omega,B(g_1)\cdots B(g_n)\Omega)=\frac{d^n}{d\lambda_1\cdots d\lambda_n}\,\omega(\delta_{\sum_{j=1}^{n}\lambda_j g_j})\Big|_{\lambda_1=\cdots=\lambda_n=0} \\ =\sum\omega_0(g_{i_1}\cdots)\cdots\omega_0(\cdots g_{i_n})$$

by induction on n, the theorem follows.

Q.E.D.

Now we turn to the study of representations of the algebra of observables induced by quasifree states. Let ω be the quasifree state determined by the positive linear form ω_0 on $\tilde{D}$, let N be the kernel of the quadratic form $(\psi,\phi)=\omega_0(\bar{\psi}\phi)$ on $\tilde{D}$ and let H be the completion of $\tilde{D}/N$ with respect to the inner product induced by ω_0. Denote by $\mathcal{F}(H)$ the symmetrized Fock space on H, i.e.,

$$\mathcal{F}(H)=S\bigoplus_{n=0}^{\infty}H^{(n)}$$

where

$$H^{(n)}=\bigotimes_{n\text{ times}}H.$$

Consider the total set of coherent vectors $\{\Omega(h)\mid h\in H\}$ where $\Omega(h)$ is the vector with nth component,

$$\Omega(h)^{(n)}=\frac{i^n}{\sqrt{n!}}\exp\left[-\left(\frac{\|h\|^2}{2}\right)\right]h\otimes h\otimes\cdots\otimes h.$$

We note that $\Omega(h)$ can be written in the form

$$\Omega(h)=\exp[iR(h)]\,\Omega(0),$$

where $R(\circ)$ is the canonical field operator for bosons on the Fock space $\mathcal{F}(H)$.[9]

Define the map π of $\overline{\Delta(D)}$ into the linear operators on $\mathcal{F}(H)$ by:

$$\pi(\delta_g)\Omega(h)=\exp\{i\mathrm{Im}[(\psi_{-g},\ h)+\omega_0(\psi_g)]\}\circ\Omega(\psi_g(h+1)+h), \qquad (3)$$

where $\psi_g=\exp(ig)-1$.

It is easily checked that

(i) $(\Omega(0),\ \pi(\delta_g)\Omega(0))=\omega(\delta_g)$.

(ii) $(\Omega(h),\ \pi(\delta_g)\Omega(f))=(\pi(\delta_{-g})\Omega(h),\ \Omega(f))$ for all $f,\ h\in H$ and $g\in D$.

Hence π is a $*$ representation of $\overline{\Delta(D)}$ into the bounded operators on $\mathcal{F}(H)$

(iii) Now we prove that $\Omega(0)$ is cyclic for the representation π.

Let F_0 be the dense set of $\mathcal{F}(H)$ consisting of the vectors $\psi=\bigoplus_n\psi^{(n)}$ such that $\psi^{(n)}=0$ except for a finite number of components.

Then for all $\psi\in F_0$ and $g\in D$

$$\lim_{\lambda\to 0}\frac{\exp[iR(\psi_{\lambda g})]-1}{i\lambda}\psi=R(ig)\psi, \quad (4)$$

$$\lim_{\lambda\to 0}\frac{\exp[iR(\psi_{\lambda g})]-1-i\lambda R(ig)}{i\lambda^2}\psi=\tfrac{1}{2}R(g^2)\psi. \quad (5)$$

Therefore, the vectors $R(ig)\psi$ and $R(g^2)\psi$ are in the closure of the linear hull, X, generated by vectors of the type

$$\prod_{j=1}^{n}\exp[iR(\psi_{g_j})]\Omega(0),$$

where g_j are elements of $\mathcal{D}$. Now, an arbitrary element f of $\widetilde{\mathcal{D}}$, can be written as

$$f=(g_1)^2-(g_2)^2+ig_3,$$

where g_1, g_2, and g_3 are in $\mathcal{D}$. Hence, since

$$R(f)\psi=R(g_1^2)\psi+R(ig_3)\psi,$$

the vector $R(f)\psi$ is in X. As $\Omega(0)$ is cyclic for the field $\{R(f)\mid f\in\mathcal{D}\}$ the cyclicity of $\Omega(0)$ for the representation π follows.

Results (i), (ii), and (iii) are now formulated in the following theorem.

Theorem 3.3: Let ω be a quasifree state on $\overline{\Delta(\mathcal{D})}$ determined by the positive linear functional ω_0 on $\widetilde{\mathcal{D}}$. Let $\mathcal{H}$ be the completion of $\widetilde{\mathcal{D}}/\mathcal{N}$ with respect to the scalar product $(f,g)=\omega_0(\bar{f}g)$, $\mathcal{N}$ being the kernel of ω_0, and $\mathcal{F}(\mathcal{H})$ the symmetric Fock space on $\mathcal{H}$.

Then π, given by (3) is the GNS representation of the state ω on $\mathcal{F}(\mathcal{H})$, with the Fock vacuum as cyclic vector.

We note that if ω_0 is continuous with respect to the supremum norm on $\widetilde{\mathcal{D}}$, we can give a simpler form of the the representation.

Let $\widetilde{\widetilde{\mathcal{D}}}$ be the involutive algebra $\widetilde{\mathcal{D}}+C$ and let $\widetilde{\omega}_0$ be the extension of ω_0 given by

$$\widetilde{\omega}_0(x+\lambda)=\omega_0(x)+\lambda\,\|\omega_0\|$$

for all $x\in\widetilde{\mathcal{D}}$ and $\lambda\in C$. Denote by $\widetilde{\mathcal{H}}$ the completion of $\widetilde{\mathcal{D}}/\widetilde{\mathcal{N}}$ with respect to the scalar product $(f,\ g)=\widetilde{\omega}_0(\bar{f}g)$, $\widetilde{\mathcal{N}}$ being the kernel of $\widetilde{\omega}_0$. It can be checked that the Hilbert space $\widetilde{\mathcal{H}}$ contains the constant functions. This allows us to diagonalize the representation π.

Let

$$V:\mathcal{F}(\mathcal{H})\to\mathcal{F}(\mathcal{H})$$

be the unitary operator given by

$$V\Omega(h)=\Omega(h-1)\exp[i\mathrm{Im}(1,h)]$$

for all $h\in\mathcal{D}$ define $\widetilde{\pi}(\delta_f)$ by

$$\widetilde{\pi}(\delta_f)=V^*\pi(\delta_f)V.$$

Then

$$\widetilde{\pi}(\delta_f)\,\Omega(h)=\Omega((\exp if)h).$$

The cyclic vector for the representation $\widetilde{\pi}$ is now $V^*\Omega(0)=\Omega(1)$.

4. MONOPARTICLE EVOLUTIONS

Now we introduce the dynamics of a particular kind, namely the monoparticle evolutions. Let $(T_t)_t$ be a one-parameter group of automorphisms on $\mathcal{D}$; this induces a one-parameter group of automorphisms $(\alpha_t)_t$ on $\overline{\Delta(\mathcal{D})}$ by

$$\alpha_t(\delta_f)=\delta_{T_t f},\ \ f\in\mathcal{D}.$$

Any time evolution α_t of this kind is called a quasifree evolution; T_t is called the monoparticle evolution.

In the following we will only be interested in time invariant quasifree states, i.e., states ω determined by functionals ω_0, invariant under T_t extended by linearity to $\widetilde{\mathcal{D}}$. In this case T_t extends to a one-parameter group of unitaries on $\mathcal{H}$ which is also denoted by T_t.

As the state ω is invariant there exists a group of unitaries $(U_t)_t$ on the representation space $\mathcal{F}(\mathcal{H})$ such that

$$\pi(\alpha_t X)=U_t\pi(X)U_{-t},\ \ X\in\overline{\Delta(\mathcal{D})},$$

$$U_t\Omega(0)=\Omega(0),$$

for all $t\in R$; it can be checked that U_t is given on the coherent vectors by

$$U_t\Omega(h)=\Omega(T_t h),\ \ h\in\mathcal{H}.$$

Furthermore, it can also be checked that, if $t\to T_t$ is strongly continous, than $t\to U_t$ is strongly continuous, and therefore by Stone's theorem there exists a self-adjoint operator $\mathcal{L}$, called the Liouville operator, on $\mathcal{F}(\mathcal{H})$ such that $U_t=\exp it\mathcal{L}$. In the following we assume that $t\to T_t$ is strongly continuous.

Let us introduce the time reversal operator θ on $\mathcal{H}$ by,

$$(\theta f)(q,\ p)=f(q,-p).$$

The evolution T_t will be called "time reversal invariant" if $\theta T_t=T_{-t}\theta$. This is the case when T_t is given by a Hamiltonian of the form $H=p^2/2+V(q)$.

Theorem 4.1: (i) If the spectrum of T_t is continuous, then 1 is the only discrete point in the spectrum of U_t, moreover it is a simple eigenvalue.

(ii) If T_t is time reversal invariant and $\Omega(0)$ is the only invariant vector of U_t, then the spectrum of U_t is continuous, except for the point one.

Proof: (i) Suppose first that the spectrum of T_t is continuous. Then

$$\mathcal{M}\ |(T_{\hat{t}}g,h)|=0\ \text{ for all } g,h\in\mathcal{H},$$

where

$$\mathcal{M}f(\hat{t})=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}f(t)\,dt,$$

but

$$\begin{aligned}
&|\mathcal{M}\,(\Omega(h),U_{\hat{t}}\Omega(g))-(\Omega(h),\Omega(0))(\Omega(0),\Omega(h))|\\
&\quad=\exp(-\tfrac{1}{2}(\|h\|^2+\|g\|^2)\,|\mathcal{M}\{\exp[-(h,T_{\hat{t}}g)]-1\}|)\\
&\quad\leqslant\exp(-\tfrac{1}{2}(\|h\|^2+\|g\|^2)\mathcal{M}\,|\exp[-h,T_{\hat{t}}g)]-1|)\\
&\quad\leqslant\exp(-\tfrac{1}{2}(\|h\|^2+\|g\|^2)\mathcal{M}\{|(h,T_{\hat{t}}g)\exp|(h,T_{\hat{t}}g)|\})\\
&\quad\leqslant\exp(-\tfrac{1}{2}(\|h\|-\|g\|)^2\mathcal{M}\,|(h,T_{\hat{t}}g)|)=0.
\end{aligned}$$

Hence

$$\mathcal{M}(\Omega(h),U_{\hat{t}}\Omega(g))=(\Omega(h),\Omega(0))(\Omega(0),\Omega(g)).$$

Since

$\{\Omega(h) | h \in \mathcal{H}\}$

is a total set in $\mathcal{J}(\mathcal{H})$, this proves that $\Omega(0)$ is the only invariant vector of U_t.

Analogously, for all vectors $\psi \in \mathcal{J}(\mathcal{H})$ such that $(\psi, \Omega(0)) = 0$, $|\mathcal{M}(\psi, U_t\phi)| = 0$ for arbitrary $\phi \in \mathcal{J}(\mathcal{H})$. Hence 1 is the only eigenvector of U_t.

(ii) Suppose now $\Omega(0)$ is the only invariant vector and suppose that the spectrum of T_t is not continuous so that there exists an eigenvalue λ and eigenvector f. For $h \in \mathcal{H}$, let

$$R^{\pm}(h) = \tfrac{1}{2}[R(h) \mp iR(ih)]$$

be the usual creating and annihilation operators on $\mathcal{J}(\mathcal{H})$ and let

$$\phi = R^{+}(f) R^{+}(\theta f)\, \Omega(0).$$

Then

$$U_t \phi = R^{+}(T_t f) R^{+}(T_t \theta f)\, \Omega(0) = |\lambda|^2 \phi = \phi.$$

Since $\phi \neq \Omega(0)$, $\Omega(0)$ is not the only invariant vector of U_t.

Q.E.D.

In the rest of this section we specialize the free evolution and the state ω induced by ω_0, where ω_0 is given by

$$\omega_0(f) = \rho \int f(q,p)\, d\mu(q,p), \tag{6}$$

where $d\mu(q,p) = \exp(-\beta p^2/2)\, dq\, dp$; ρ and β are positive real numbers corresponding to the density and inverse temperature respectively; the state ω is a KMS state for the free evolution (see Sec. 5). As the measure μ is absolutely continuous with respect to Lebesque measure $dqdp$ we can just as well work with Lebesque measure, so that $\mathcal{H} = \mathcal{L}^2(R^d \times R^d)$. The infinitesimal generator L of T_t on $\mathcal{H}$ is well known to be self-adjoint[9] and given formally by

$$L = -ip \frac{\partial}{\partial q}. \tag{7}$$

As the representation U_t is given by

$$U_t \Omega(h) = \Omega(T_t h), \quad h \in \mathcal{H}$$

it is well known that

$$U_t = \exp[it\, d\Gamma(L)],$$

where $d\Gamma(L)$ on $D_L \cap \mathcal{H}^{(n)}$, with

$$D_L = \{\psi \in F_0 \,|\, \psi^{(n)} \in \textstyle\bigotimes_{k=1}^{n} C_0^\infty(R^{2d}) \text{ for all } n\}$$

is given by

$$L \otimes 1 \otimes \cdots \otimes 1 + \cdots 1 \otimes \cdots \otimes L.$$

Consider the unitary map U of $\mathcal{H}$ onto $\mathcal{H}$,

$$(Uf)(k,p) = \tilde{f}(k,p) = \left(\frac{1}{2\pi}\right)^{d/2} \int e^{iqk} f(q,p)\, dq.$$

Then ULU^{-1} is the multiplication operator by $-pk$. One checks that the time evolution is given by

$$(U_t\psi)^{(n)}(k_1p_1, \ldots, k_np_n) = \exp[i(k_1 \cdot p_1 + \cdots + k_n \cdot p_n)t]\, \psi^{(n)}(k_1p_1, \ldots, k_np_n),$$

i.e., on the orthogonal complement of $\Omega(0)$, U_t is a multiplication operator, its infinitesimal generator is also a multiplication operator by a real-valued continuously differentiable function with nonzero gradient almost everywhere. Hence $\mathcal{L}$ is spectrally absolutely continuous (see Ref. 10, p. 518) except for the point zero.

Theorem 4.2: For the infinite classical system of free particles in the quasifree state determined by ω_0 as in (6), the spectrum of the Liouvillian $\mathcal{L}$ is absolutely continuous except for zero.

5. KMS STATES

In this section we have to introduce the Poisson brackets, therefore we turn for the moment to the original algebra $\mathcal{A}$. The Poisson bracket between two generators of the algebra $\exp iSf$ and $\exp iSg$, $f, g \in \mathcal{D}$ is given by

$$\{\exp iSf,\ \exp iSg\} = -S(\{f,g\}) \exp[iS(f+g)], \tag{8}$$

where

$$\{f,g\} = \frac{\partial f}{\partial q} \circ \frac{\partial q}{\partial p} - \frac{\partial f}{\partial p} \circ \frac{\partial g}{\partial q}$$

is again an element of $\mathcal{D}$. Let ω be a regular state on $\overline{\Delta(\mathcal{D})}$, using the isomorphism between $\mathcal{A}$ and $\overline{\Delta(\mathcal{D})}$ (Theorem 2.1), formula (8) suggest that we define the Poisson bracket between two generators in the GNS representation π as follows:

$$\{\pi(\delta_f),\ \pi(\delta_g)\} = -B(\{f,g\})\, \pi(\delta_{f+g}), \tag{9}$$

where B is the associated field.

Definition 5.1: The state ω on the algebra $\Delta(\mathcal{D})$ is KMS with respect to the monoparticle evolution $(\alpha_t)_t$ at inverse temperature β if

(i) ω is a C^∞ state,

(ii) ω is time invariant,

(iii) for all f and g in $\mathcal{D}$, $t \to \omega(\delta_f \alpha_t \delta_g)$ is differentiable and

$$\beta \frac{d}{dt} \omega(\delta_f \alpha_t \delta_g) = -(\Omega, \{\pi(\delta_f),\ \pi(\alpha_t \delta_g)\}\Omega).$$

Definition 5.2: The positive linear functional ω_0 in $\mathcal{D}$ is said to be KMS for the evolution $(T_t)_t$ at inverse temperature β if

(i) ω_0 is T_t invariant,

(ii) for all f and g in $\tilde{\mathcal{D}}$: $t \to \omega_0(fT_t g)$ is differentiable, and

$$\beta \frac{d}{dt}\, \omega_0(fT_t g) = -\omega_0(\{f, T_t g\}).$$

Now we have the following theorem.

Theorem 5.3: Let ω be a quasifree state on $\overline{\Delta(\mathcal{D})}$ determined by the positive linear functional ω_0 on $\tilde{\mathcal{D}}$, then ω is KMS for the evolution α_t if and only if ω_0 is KMS for the corresponding evolution $(T_t)_t$.

Proof: Suppose first ω_0 is KMS, then

$$\beta \frac{d}{dt}\, \omega(\delta_f \alpha_t \delta_g)$$

$$= \beta \frac{d}{dt} \omega(\delta_{f+T_t g})$$

$$= \beta \frac{d}{dt} \exp[\omega_0(\exp[i(f+T_t g)] - 1)]$$

$$= \omega(\delta_{f+T_t g}) \beta \frac{d}{dt} \omega_0([\exp(if) - 1]\, T_t(\exp(ig) - 1)),$$

where the invariance of ω_0 was used.

Since ω_0 is KMS,

$$\beta \frac{d}{dt} \omega(\delta_f \alpha_t \delta_g)$$

$$= -\omega(\delta_f \alpha_t \delta_g) \omega_0(\{\exp if,\ \exp i T_t g\})$$

$$= \omega(\delta_f \alpha_t \delta_g) \omega_0(\{f, T_t g\} \exp i(f+T_t g)).$$

Now as for all $h, k \in \mathcal{D}$

$$(\Omega, B(h)\pi(\delta_k)\Omega)$$

$$= \frac{d}{id\lambda} \exp \omega_0(\exp[i(k+\lambda h)] - 1)\Big|_{\lambda=0}$$

$$= \omega(\delta_k)\omega_0(h \exp ik),$$

we have

$$\beta \frac{d}{dt} \omega(\delta_f \alpha_t \delta_g) = -(\Omega, \{\pi(\delta_f),\ \pi(\alpha_t \delta_g)\}\Omega).$$

Conversely, suppose that ω is KMS, then as above,

$$\omega(\delta_{f+T_t g})\beta \frac{d}{dt} \omega_0([\exp(if) - 1] T_t(\exp(ig) - 1))$$

$$= \beta \frac{d}{dt} \omega(\delta_f \alpha_t \delta_g)$$

$$= -(\Omega,\ \{\pi(\delta_f),\ \pi(\alpha_t \delta_g)\}\Omega)$$

$$= -\omega(\delta_{f+T_t g}) \cdot \omega_0(\{\exp if,\ \exp i T_t g\}).$$

Since $\omega(\delta_k) \neq 0$ for all $k \in \mathcal{D}$, this relation is equivalent with ω_0 satisfying KMS, because all functions of the type $\exp(if) - 1$ with $f \in \mathcal{D}$ generate $\tilde{\mathcal{D}}$, and the invariance of ω_0 follows immediately from the invariance of ω.

Q.E.D.

By Theorem 5.3, the problem of solving the KMS condition is reduced to solving the KMS condition for the functional ω_0 on $\mathcal{D}$.

If T_t is given by a Hamiltonian $H(q,p)$ which is continuously differentiable, and if we suppose that ω_0 is absolutely continuous with respect to Lebesque measure, then by partial integration the ω_0 KMS condition has a unique solution up to the parameter ρ, given by

$$\omega_0(h) = \rho \int \exp(-\beta H)\, h dq dp,$$

where ρ is a positive constant. Hence we may conclude that for quasifree evolution induced by monoparticle evolutions given by a Hamiltonian there exists at least one solution of the KMS condition. We treat the uniqueness problem of these solutions elsewhere.

Note added in proof: After the completion of this work, we received the thesis of J. L. van Hemmen, "Dynamics and ergodicity of the infinite harmonic crystal," University of Groningen, where he considered Gaussian states for classical harmonic crystals.

ACKNOWLEDGMENTS

The authors wish to thank A. Van Daele and M. Fannes for very useful remarks and stimulating discussions.

*On leave of absence from the Royal University of Malta.
†Postal address: Instituut voor Theoretische Fysica, Celestijnenlaan 200 D, B-3030 Heverlee, Belgium.

[1]R. Haag and D. Kastler, J. Math. Phys. 5, 848 (1964).
[2]R. Haag, N. M. Hugenholtz, and M. Winnink, Commun. Math. Phys. 5, 215 (1967).
[3]D. Ruelle, *Statistical Mechanics* (Benjamin, New York, 1969).
[4]G. Gallavotti and E. Verboven, Nuovo Cimento 28, 274 (1975).
[5]G. Gallavotti and M. Pulvirenti, Commun. Math. Phys. 46, 1 (1976).
[6]M. Aizenman, G. Gallavotti, S. Goldstein, and J. L. Lebowitz, "Stability and Equilibrium States of Infinite Classical Systems," preprint.
[7]O. E. Lanford, III, Commun. Math. Phys. 8, 176 (1968); 11, 257 (1969).
[8]D. W. Robinson, Commun. Math. Phys. 1, 159 (1965).
[9]M. Reed and B. Simon, *Methods of Modern Mathematical Physics* (Academic, New York, 1975), Vol. II.
[10]T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, Berlin, 1966).

On a family of interior solutions for relativistic fluid spheres with possible applications to highly collapsed stellar objects*

Patrick G. Whitman

North Texas State University, Denton, Texas 76203
(Received 7 September 1976)

A one-parameter family of interior solutions to Einstein's field equations for a static spherical fluid is given. It is shown that for various values of the parameter and choices of the constants of integration, several previously known solutions for static fluids are contained therein. This family of solutions can be joined continuously to the Schwarzchild exterior solution, and as such may be applicable to the investigation of stellar interiors where high central densities and pressures are of interest.

I. INTRODUCTION

The field equations of general relativity in the presence of matter form a highly nonlinear system of equations. For this reason, few exact solutions have been obtained, even for the simplest cases. The most notable of these is the constant density solution for a spherical distribution of matter.[1]

For the case of a static spherically symmetric fluid of density ρ and pressure P, the field equations reduce to a set of three coupled ordinary differential equations involving these fluid variables and two metric functions. In order to solve this system, it is necessary to specify in some manner one of the unknowns, or to introduce a subsidiary relation between two of them, i.e., specify an equation of state. Such an assumption removes the indeterminancy of the system. We now consider these equations in more detail.

The general relativistic field equations for a static spherically symmetric line element in the presence of a perfect fluid

$$ds^2 = \gamma(r)^2 dt^2 - \tau(r)^{-1} dr^2 - r^2(d\theta^2 + \sin^2(\theta)\, d\varphi^2) \quad (1.1)$$

can be written as one equation relating the metric functions γ and τ, and two equations which can be taken as the definitions of the fluid variables in terms of the metric functions. The relationship between γ and τ is given by the expression

$$\tau(r) = \exp[-F(r)]\left[\int^r \exp[F(u)]\, g(u)\, du + c\right] \quad (1.2)$$

with the functions g and F being defined as

$$g(r) = -2\gamma / r(\gamma + r\gamma'),$$

$$F(r) = \int^r g(u)\gamma^{-1}(\gamma + u\gamma' - u^2\gamma'')\, du.$$

C is a constant of integration to be fixed by the boundary conditions. Here the prime refers to differentiation of the function in question. This equation, in slightly altered form, is due to Adler.[2]

The two remaining field equations define for us the pressure and density in terms of the metric functions. These are

$$8\pi r^2 P(r) = (\gamma + 2r\gamma')(\tau/\gamma) - 1 \quad (1.3)$$

and

$$m(r) = 4\pi \int_0^r \rho(u)\, u^2\, du, \quad (1.4)$$

where $m = \frac{1}{2} r(1 - \tau^{-1})$ expresses how the mass is distributed throughout the fluid.

Analytic specification of the equation of state does not always lead to a tractable solution, and numerical or graphic techniques must be applied. Exact solutions in terms of known functions are most easily obtained by requiring one of the field variables to satisfy some subsidiary condition which simplifies the full set of equations. Once the field equations are solved in this manner, an equation of state can then be extracted. Such solutions may be useful in understanding a system in the extreme relativistic limit where we cannot specify a priori what the equation of state might be.

II. A SPECIFIC CHOICE OF $\gamma(r)$

As stated above, the set of equations (1.2)—(1.4) cannot be solved without either choosing an equation of state or making a specific assumption on one of the functions P, ρ, γ, or τ. After Adler, we note the Eq. (1.2) is linear in τ if γ is a known function. This being the case, we choose γ in such a manner that Eq. (1.2) can be immediately integrated. Such a choice is that γ satisfies the Cauchy equation:

$$r^2\gamma'' - r\gamma' + (1 - \alpha^2)\gamma = 0, \quad 0 \leqslant \alpha \leqslant 1. \quad (2.1)$$

This expression can be immediately integrated. It yields

$$\gamma(r) = ar^i + br^j, \quad (2.2)$$

where $i = 1 + \alpha$, $j = 1 - \alpha$. In this solution, a and b are constants of integration.

When Eq. (2.1) is used in conjunction with Eq. (1.2) τ can be readily obtained

$$\tau(r) = s^{-1} + cr^{2s/l}\,[akr^{2\alpha} + bl]^{-2s/lk}, \quad (2.3)$$

where $s = 2 - \alpha^2$, $k = 2 + \alpha$, $l = 2 - \alpha$.

Since γ and τ are now known, they may be used to solve the remaining two equations (1.3) and (1.4). We find the pressure, Eq. (1.3), is given by

$$8\pi r^2 P(r) = \tau(r)\,[anr^{2\alpha} + bq][ar^{2\alpha} + b]^{-1} - 1 \quad (2.4)$$

and the density, Eq. (1.4), is

$$8\pi r^2\rho(r) = 1 - \tau(r) - 2(s\tau(r) - 1) \times (ar^{2\alpha} + b)(akr^{2\alpha} + bl)^{-1}, \qquad (2.5)$$

where $n = 3 + 2\alpha$.

The three constants of integration can now be determined by matching the solutions at the boundary to the Schwarzchild exterior solution. The result is

$$a = (1 - q\gamma_s^2)(4\alpha R^l \gamma_s)^{-1}, \qquad (2.6a)$$

$$b = -(1 - n\gamma_s^2)(4\alpha R^j \gamma_s)^{-1}, \qquad (2.6b)$$

$$c = (\gamma_s^2 - 1/s)[(1 + \gamma_s^2)(2\gamma_s R^3)^{-1}]^{2s/lk}, \qquad (2.6c)$$

with $\gamma_s^2 = 1 - 2M/R$ and $q = 3 - 2\alpha$. R refers to the radius of the fluid and M to the total mass:

$$M = 4\pi \int_0^R \rho(u) u^2 \, du. \qquad (2.7)$$

Note that now $0 < \alpha \leq 1$. The solution with $\alpha = 0$ cannot be matched to the Schwarzchild with a finite boundary.

III. CONCLUSION

The solutions discussed above have the following properties:

(1) The pressure and density diverge at the origin for $0 < \alpha < 1$, but the ratio of central values

$$P_c/\rho_c = (1 - \alpha)(1 + \alpha)^{-1} \qquad (2.8)$$

remains finite and is independent of the total mass and radius of the fluid.

(2) Though the density diverges at the origin, its integral remains finite:

$$4\pi \int_0^\epsilon \rho(r) r^2 \, dr \leq M, \quad \epsilon \leq R. \qquad (2.9)$$

(3) If the fluid is considered adiabatic, the velocity of sound is given by the relation

$$\frac{d\rho}{dP} = -\Gamma^2, \qquad (2.10)$$

where Γ is the speed of sound in the fluid. Requiring that this quantity be less than one places restrictions on the values of M/R.

For particular values of the parameter α and integration constants a, b, and C, several previously known solutions for static fluids are contained herein. The limiting value of $\alpha = 1$ is the solution given by Adler.[2] This is the only solution of the family which does not diverge at the origin. Three solutions published by Tolman are also in this family.[3] These are the solutions listed as Tolman numbers 1, 5, and 6. Tolman number 1 is also known as the Einstein Universe. Tolman numbers 5 and 6 were compared to the numerical solution for a degenerate quantum gas.[4]

The family of solutions described in this paper, though singular at the origin, may be useful in the investigation of massive stars. They allow the investigator to vary the equation of state in a continuous manner by changing the value of the parameter α. It is interesting to note that this is the only family of interior solutions for relativistic fluid spheres in which none of the field variables are considered constant.

ACKNOWLEDGMENT

The author is greatly indebted to I. R. Jackson Jr. II for running a computer check on the solutions and their associated boundary conditions. He further wishes to acknowledge helpful and informative discussions with Donald Rapp, and Istvan Ozsvath, and Don Kobe.

*Supported in part by the Robert A. Welch Foundation, Grant No. B-516.

[1]R. C. Tolman, *Relativity, Thermodynamics and Cosmology* (Oxford University, Oxford, 1934), pp. 246–247.

[2]R. Adler, J. Math. Phys. 15, 727 (1974). Also, there is an erratum to this paper published in J. Math. Phys. 17, 158 (1976).

[3]R. C. Tolman, Phys. Rev. **55**, 364 (1939).

[4]J. R. Oppenhimer and G. M. Volkoff, Phys. Rev. **55**, 374 (1939).

Group theory of the collective model of the nucleus

E. Chacón* and M. Moshinsky*†

Instituto de Física, Universidad de México (UNAM), México 20, D.F., Mexico

(Received 12 July 1976)

In the present paper we extend the group theoretical analysis of a previous publication to obtain explicitly, as a polynomial in $\sin\gamma$, $\cos\gamma$, the function $\phi_K^{\lambda\mu L}(\gamma)$ required in the discussion of the quadrupole vibrations of the nucleus. The states appearing in the collective model $|\nu\lambda\mu LM\rangle = F_l^\lambda(\beta)\Sigma_K \phi_K^{\lambda\mu L}(\gamma)\, D_{MK}^{L*}(\vartheta_i)$, $l=(\nu-\lambda)/2$, are then defined, as $F_l^\lambda(\beta)$, $D_{MK}^{L*}(\vartheta_i)$ are well known. All matrix elements required in the collective model of the nucleus are related then with the expression $(\lambda\mu L;\lambda'\mu'L';\lambda''\mu''L'')$ = $\int_0^\pi \Sigma_{KK'K''}\left(\begin{smallmatrix}L & L' & L''\\ K & K' & K''\end{smallmatrix}\right)\phi_K^{\lambda\mu L}(\gamma)\phi_{K'}^{\lambda'\mu'L'}(\gamma)\phi_{K''}^{\lambda''\mu''L''}(\gamma)\sin 3\gamma\, d\gamma$, which is a reduced $3j$-symbol in the $O(5)\supset O(3)$ chain of groups.

1. INTRODUCTION

In a recent publication[1] the authors, in collaboration with Sharp, developed a procedure for the exact solution of the quantum mechanical problem associated with the quadrupole vibrations of the nucleus. This problem, originally discussed by Bohr,[2] played a very important role in the development of the collective model of the nucleus. It is related with the liquid drop model in which the surface of the nucleus is given by the equations

$$R=R_0\left(1+\sum_m \alpha^m Y_{2m}(\theta,\varphi)\right), \tag{1.1}$$

where R_0 is the radius in the absence of deformation, Y_{2m} is a spherical harmonic, and α^m, $m=2,1,0,-1,-2$ the contravariant form of the generalized coordinates describing the collective motion. The Hamiltonian of the problem, in appropriate units,[1] can be written as[1-3]

$$H_0=\tfrac{1}{2}\sum_m(\pi_m\pi^m+\alpha_m\alpha^m),\quad \alpha_m=(-)^m\alpha^{-m},$$

$$\pi_m=\frac{1}{i}\frac{\partial}{\partial\alpha^m},\quad \pi_m=(-)^m\pi^{-m}. \tag{1.2}$$

We can then pass to the coordinates fixed in the body through the transformation[1-3]

$$\alpha_m=\sum_{m'}D^{2*}_{mm'}(\vartheta_i)a_{m'}, \tag{1.3a}$$

$$a_2=a_{-2}=(1/\sqrt{2})\beta\sin\gamma,\quad a_1=a_{-1}=0,\quad a_0=\beta\cos\gamma, \tag{1.3b}$$

where ϑ_i, $i=1,2,3$, are the Euler angles, β,γ the remaining coordinates, and $D^l_{mm'}(\vartheta_i)$ the Wigner functions that are the irreducible representations of the O(3) group. The eigenstates of the Hamiltonian (1.2) associated with the number ν of quanta and of definite seniority λ, angular momentum L and projection M can be denoted by the ket

$$|\nu\lambda\mu LM\rangle=F_l^\lambda(\beta)\sum_K\phi_K^{\lambda\mu L}(\gamma)D^{L*}_{MK}(\vartheta_i) \tag{1.4}$$

where

$$l\equiv\tfrac{1}{2}(\nu-\lambda) \tag{1.5}$$

and μ indicates the remaining quantum number required to fully characterize the states of the Hamiltonian (1.2). In (1.4) $F_l^\lambda(\beta)$ are well known[1-3] functions of β associated with the radial part of a five-dimensional oscillator. Thus the essential remaining point is to determine the γ-dependent functions $\phi_K^{\lambda\mu L}(\gamma)$ that have the symmetry properties

$$\phi_K^{\lambda\mu L}(\gamma)=0 \text{ if } K \text{ is odd},$$

$$\phi_K^{\lambda\mu L}(\gamma)=(-)^L\phi_{-K}^{\lambda\mu L}(\gamma)\ \text{ if } K \text{ is even}, \tag{1.6a}$$

$$K=L,L-1,L-2,\ldots,-L,\quad 0\le\gamma\le\frac{\pi}{3}. \tag{1.6b}$$

Already in 1959, Bes gave a recursive technique for deriving $\phi_K^{\lambda\mu L}(\gamma)$ for $L=0,2,3,4,5,6$. In the present paper we shall use the procedure of Ref. 1 to determine $\phi_K^{\lambda\mu L}(\gamma)$ explicitly for arbitrary L and λ.

At this point the reader may well ask whether it is worthwhile to make the considerable algebraic effort needed to achieve this purpose. We shall answer the question by discussing some of the problems that appear in the collective model of the nucleus and showing that they require the reduced Wigner coefficients for the chain of groups $O(5)\supset O(3)$. These coefficients in turn can be determined by an integral involving the functions $\phi_K^{\lambda\mu L}(\gamma)$. Thus many aspects of the collective model of the nucleus reduct to a problem of group theory.

In the original work of Bohr and Mottelson,[2] two points of view were stressed. In one of them the Hamiltonian (1.2) was used to describe the low energy vibrations of the nucleus giving rise to equidistant energy levels. In the other β,γ were not considered as dynamical variables, but were substituted by the numbers β_0, γ_0 giving the minimum of the potential energy surface. The operator (1.2) becomes then a function of the Euler angles only, giving rise to the Hamiltonian of the symmetric ($\gamma_0=0$ or $\pi/3$) or asymmetric ($0<\gamma_0<\pi/3$) top. The rotational levels obtained are afterwards modified by the rotation vibration interaction and also by coupling with the single particle degrees of freedom.[2]

The two approaches indicated in the previous paragraph were combined in a single one in the work of Greiner and his collaborators.[4,5] They consider a Hamiltonian

$$H=T+V(\beta,\gamma) \tag{1.7}$$

in which T is the kinetic energy $\sum_m\pi^m\pi_m$ in (1.2) but may also include higher order corrections and the potential energy $V(\beta,\gamma)$ (which is taken as a polynomial function of the α_m invariant under rotations) can then be written as[1,4,5]

$$V(\beta,\gamma)=\sum_{\rho\mu} U_{\rho\mu}\{2,0\}^\rho\{3,0\}^\mu$$
$$=\sum_{\rho\mu}(-\sqrt{2})^\rho U_{\rho\mu}\beta^{2\rho+3\mu}x^\mu$$
$$=\sum_{\rho\mu} V_{\rho\mu}\beta^{2\rho+3\mu}P_\mu(x), \tag{1.8}$$

where in the work of Greiner and collaborators $2\rho+3\mu \leq 6$. In (1.8) the $U_{\rho\mu}$ are some constant coefficients and the elementary permissible diagrams $\{2,0\}$, $\{3,0\}$ are polynomials in the α_m discussed in (5.1), (5.2) of Ref. 1, of the form

$$\{2,0\}=\sqrt{5}[\alpha\times\alpha]_0^0=\beta^2 \tag{1.9a}$$

$$\{3,0\}=\sqrt{7}[[\alpha\times\alpha]^2\times\alpha]_0^0=-\sqrt{2}\beta^3 x, \tag{1.9b}$$

where

$$x\equiv\cos 3\gamma \tag{1.9c}$$

and the square brackets with multiplications signs indicate the coupling of the α's to the definite angular momentum given at the upper right hand side corner. We can write the potential energy $V(\beta,\gamma)$ in terms of powers of $\beta^{2\rho+3\mu}$ and Legendre polynomials $P_\mu(x)$ with constant coefficients which we now call $V_{\rho\mu}$. The advantage of this is that from the form of the Casimir operator Λ^2 of O(5) given in Eq. (2.14) of Ref. 1, the $P_\mu(x)$ is an eigenfunction of it with eigenvalue $\lambda(\lambda+3)$, where $\lambda=3\mu$. Thus the potential energy (1.8) can be expressed as a linear combination of irreducible tensors of O(5) whose row is characterized by $L=M=0$.

The problem of determining the eigenvalues of the Hamiltonian (1.7) reduces then to the evaluation of the matrix elements of $\beta^{2\rho+3\mu}P_\mu(x)$ with respect to the states (1.4). Once they are available we have the matrix of the H of (1.7) as function of the parameters $V_{\rho\mu}$. Greiner and his collaborators[4,5] were able to determine these parameters by diagonalizing the matrix and comparing the resulting eigenvalues of H with the low lying energy levels of even-even nuclei. Thus they could obtain the potential energy surfaces (PES) associated with $V(\beta,\gamma)$ and find the minima β_0, γ_0 for this potential which indicates the type of deformation of the nuclei in question.

Before proceeding to determine the matrix elements we are interested in we briefly indicate how they were obtained by Greiner and collaborators.[4,5] The states (1.4) are characterized by irreducible representations (IR) or the chain of groups[1,4,5]

$$\underset{\nu}{U(5)}\supset\underset{\lambda}{O(5)}\supset\underset{L}{O(3)}\supset\underset{M}{O(2)}, \tag{1.10}$$

where underneath each one of them we have put the corresponding quantum number. In the analysis of Greiner and collaborators[4,5] the states were originally characterized by IR of the chain of groups[4,5]

$$U(5)\supset O(5)\supset SU(2)\times SU(2) \tag{1.11}$$

used by Hecht.[6] Then numerically they obtained the states characterized by IR of the chain of groups (1.10) as linear combinations of those of (1.11) by essentially diagonalizing L^2 in the latter basis.

The procedure followed by Greiner and collaborators allowed them to obtain results of considerable physical significance. These include not only the potential energy surfaces mentioned above but also the problem of giant dipole resonances that require the matrix elements of α_m with respect to the states (1.4)[4,5]; the collective nuclear excitation by electron scattering; the fission problem when described as two nuclei with quadrupole excitations joining to form a compound nucleus with the same type of excitation[5,7]; the problem of back bending related with the crossing of two bands with different moments of inertia,[5,8] etc. There are even applications outside nuclear physics as for example in the Jahn–Teller effect in crystals discussed by Judd[9] where again matrix elements of α_m with respect to the states (1.4) are required.

The importance of the problems indicated in the previous paragraph, has prompted the present authors to reduce them essentially to the determination of the reduced Wigner coefficients in the chain $O(5)\supset O(3)$, then in turn to express these coefficients in terms of an integral involving the functions $\phi_K^{\lambda\mu L}(\gamma)$ and finally to evaluate explicitly these functions and the integral in which they appear.

The matrix elements required in all the problems mentioned above are related with those of operators that are homogeneous polynomials of degree λ in the α_m's corresponding to definite angular momentum L and projection M and that satisfy the five dimensional Laplace equation. As the generators of the O(5) group can be expressed as

$$\Lambda_m^{m'}=\alpha_m\frac{\partial}{\partial\alpha_{m'}}-\alpha^{m'}\frac{\partial}{\partial\alpha^m}, \tag{1.12}$$

the Casimir operator becomes

$$\Lambda^2=\tfrac{1}{2}\sum_{mm'}\Lambda_m^{m'}\Lambda_{m'}^m=\mathcal{N}(\mathcal{N}+3)-\beta^2\nabla^2, \tag{1.13a}$$

where

$$\mathcal{N}=\sum_m\alpha_m\frac{\partial}{\partial\alpha_m},\quad \nabla^2=\sum_m(-)^m\frac{\partial^2}{\partial\alpha_m\partial\alpha_{-m}} \tag{1.13b}$$

Thus the homogeneous polynomials in the α_m's mentioned above are eigenfunctions of Λ^2 with eigenvalue $\lambda(\lambda+3)$ and can be designated by

$$T_M^{\lambda\mu L}(\alpha_m)=\beta^\lambda T_M^{\lambda\mu L}\left(\frac{\alpha_m}{\beta}\right), \tag{1.14}$$

where the polynomial function $T_M^{\lambda\mu L}(\alpha_m/\beta)$ depends only on γ and the Euler angles ϑ_i, and is associated with irreducible representations of the chain of groups

$$\underset{\lambda}{O(5)}\supset\underset{L}{O(3)}\supset\underset{M}{O(2)}. \tag{1.15}$$

We have added an extra index μ in $T_M^{\lambda\mu L}(\alpha_m/\beta)$ as we shall show in Sec. 3 that the most general polynomials in the α_m requires it in the same way as the state (1.4).

From (1.3) we see that

$$T_M^{\lambda\mu L}\left(\frac{\alpha_m}{\beta}\right)=\sum_K T_K^{\lambda\mu L}\left(\frac{a_m}{\beta}\right)D_{MK}^{L*}(\vartheta_i), \tag{1.16}$$

where $T_K^{\lambda\mu L}(a_m/\beta)$ is the same polynomial as the one appearing in (1.14) but now of the ratio a_m/β. We note then from (1.3b) that $T_K^{\lambda\mu L}(a_m/\beta)$ is a function of γ only and thus we can write

$$T_K^{\lambda\mu L}(a_m/\beta) \equiv \phi_K^{\lambda\mu L}(\gamma). \tag{1.17}$$

We use the same notation as for the function ϕ appearing in (1.4) because we shall show in Sec. 3 that the identification is justified.

All of the operators required in Refs. 4–9 are of the form (1.14) multiplied by some power of β^2. Therefore the most general matrix element we need is given by

$$\langle \nu''\lambda''\mu''L''M''\left|\beta^{2\rho+\lambda}T_M^{\lambda\mu L}\left(\frac{\alpha_m}{\beta}\right)\right|\nu'\lambda'\mu'L'M'\rangle = \int_0^\infty F_{l''}^{\lambda''}(\beta)\beta^{2\rho+\lambda}F_{l'}^{\lambda'}(\beta)\beta^4\,d\beta\cdot 8\pi^2(-)^{L''+M''}\begin{pmatrix} L & L' & L'' \\ M & M' & -M''\end{pmatrix}$$

$$\times\int_0^\pi \sum_{KK'K''}\begin{pmatrix} L & L' & L'' \\ K & K' & K''\end{pmatrix}\phi_K^{\lambda\mu L}(\gamma)\phi_{K'}^{\lambda'\mu'L'}(\gamma)\phi_{K''}^{\lambda''\mu''L''}(\gamma)\sin 3\gamma\,d\gamma$$

$$=\Big\{2\pi[l'!\,\Gamma(l'+\lambda'+\tfrac{5}{2})]^{-1/2}[l''!\,\Gamma(l''+\lambda''+\tfrac{5}{2})]^{-1/2}(-)^{l'+l''}$$

$$\times\frac{\Gamma[\frac{1}{2}(\lambda'+\lambda''+\lambda+2\rho+5)]\Gamma[\frac{1}{2}(2\rho+\lambda-\lambda'+\lambda'')+1]\Gamma[\frac{1}{2}(2\rho+\lambda-\lambda''+\lambda')+1]}{\Gamma[\frac{1}{2}(2\rho+\lambda-\lambda'+\lambda'')-l'+1]\Gamma[\frac{1}{2}(2\rho+\lambda-\lambda''+\lambda')-l''+1]}$$

$$\times\,{}_3F_2\left[\begin{matrix}\frac{1}{2}(2\rho+\lambda+\lambda'+\lambda''+5), & -l', & -l'' \\ \frac{1}{2}(2\rho+\lambda-\lambda'+\lambda'')-l'+1, & \frac{1}{2}(2\rho+\lambda-\lambda''+\lambda')-l''+1 & \end{matrix};1\right]\Big\}$$

$$\times(\lambda\mu L;\lambda'\mu'L';\lambda''\mu''L'')8\pi^2(-)^{M''}\begin{pmatrix} L' & L & L'' \\ M' & M & -M''\end{pmatrix}. \tag{1.18}$$

In (1.18) we made use of (1.4) and (1.16), (1.17) to have an integral over Euler angles of the product of three D_{MK}^L functions which give the ordinary $3j$-coefficients in the integral over γ. The radial integral over β appears separately and is evaluated explictly in the curly bracket at the right hand side of (1.18) in terms of gamma and hypergeometric functions of the arguments indicated in which, as in (1.5), $l'=\frac{1}{2}(\nu'-\lambda')$ $l''=\frac{1}{2}(\nu''-\lambda'')$. The radial part is also a result of an analysis associated with a O(2,1) group whose generators are

$$I_1=\tfrac{1}{4}\sum_m(\pi_m\pi^m-\alpha_m\alpha^m),\quad I_2=\tfrac{1}{4}\sum_m(\alpha_m\pi^m+\pi^m\alpha_m),$$

$$I_3=\tfrac{1}{4}\sum_m(\pi_m\pi^m+\alpha^m\alpha_m)=\tfrac{1}{2}H_0 \tag{1.19}$$

as shown in several publications.[10–12]

Our remaining concern is then the determination of the reduced Wigner coefficient in the chain $O(5)\supset O(3)$ given by the expression

$$(\lambda\mu L;\lambda'\mu'L';\lambda''\mu''L'')$$

$$\equiv\int_0^\pi\sum_{KK'K''}\left[\begin{pmatrix} L & L' & L'' \\ K & K' & K''\end{pmatrix}\phi_K^{\lambda\mu L}(\gamma)\right.$$

$$\left.\times\phi_{K'}^{\lambda'\mu'L'}(\gamma)\phi_{K''}^{\lambda''\mu''L''}(\gamma)\right]\sin 3\gamma\,d\gamma, \tag{1.20}$$

where we have made use of the symmetry properties (1.6) and the fact that the $\phi_K^{\lambda\mu L}(\gamma)$ we shall determine later are real.

What are the types of irreducible tensors $T_K^{\lambda\mu L}(a_m/\beta)=\phi_K^{\lambda\mu L}(\gamma)$ we have encountered so far? For the potential energy $V(\beta,\gamma)$ of (1.8) we see that $L=M=0$ and then from (2.23) of Ref. 1 we have

$$\phi_0^{3\mu\mu 0}(\gamma)=P_\mu(x),\quad \lambda=3\mu,\quad \mu \text{ integer}, \tag{1.21}$$

where P is a Legendre polynomial and $x=\cos 3\gamma$ as indicated in (1.9c). For the quadrupole matrix elements between collective states, required in many applications discussed by Greiner and his collaborators,[4–8] Judd[9] and in the recent papers of Iachello and Arima,[13] one considers the operators

$$\alpha_m, \tag{1.22a}$$

$$\sqrt{7}[\alpha\times\alpha]_m^2. \tag{1.22b}$$

The irreducible tensors corresponds then to $L=2$, $M=m$ and $\lambda=1$ or 2 respectively with the value of μ, from the discussion given in the following sections, being 0. In fact from (1.3) the $\phi_K^{102}(\gamma)$ corresponding to α_m has the components

$$\phi_2^{102}(\gamma)=\phi_{-2}^{102}(\gamma)=(1/\sqrt{2})\sin\gamma,$$
$$\phi_1^{102}(\gamma)=\phi_{-1}^{102}(\gamma)=0,\quad \phi_0^{102}(\gamma)=\cos\gamma, \tag{1.23a}$$

while $\phi_K^{202}(\gamma)$ associated with $\sqrt{7}[\alpha\times\alpha]_m^2$ takes the form

$$\phi_2^{202}(\gamma)=\phi_{-2}^{202}(\gamma)=\sin 2\gamma,$$
$$\phi_1^{202}(\gamma)=\phi_{-1}^{202}(\gamma)=0,\quad \phi_0^{202}=-\sqrt{2}\cos 2\gamma. \tag{1.23b}$$

The purpose of the present paper is to determine the $\phi_K^{\lambda\mu L}(\gamma)$ explicitly for arbitrary values of λ, μ, L and then to discuss the reduced Wigner coefficient in the chain $O(5)\supset O(3)$ given by (1.20). With other collaborators we intend to publish later a book of tables and programs for the evaluation of all matrix elements of the form (1.18). We hope that this book may play for the collective model the role that the *Tables of Transformation Brackets* of Brody and Moshinsky,[14] first edited in 1960, played for the harmonic oscillator shell model.

We shall start our analysis by indicating that the states discussed in Ref. 1 in terms of traceless boson creation operators can also be expressed in terms of annihilation operators or in a mixed traceless creation and annihilation form. It is the latter that will prove particularly convenient for the definition of the $\phi_K^{\lambda\mu L}(\gamma)$ which we want to determine.

2. THE STATES IN A TRACELESS "PARTICLE-HOLE" PICTURE

In the discussion carried out in Ref. 1 (to be designated below by I, with specific equations indicated by the number followed by I) we noted that we only need to analyze states (1.4) in which $\nu=\lambda$, $M=L$. The states with arbitrary ν can then be obtained[1] from those mentioned by introducing an associated Laguerre polynomial in β^2, while for the matrix elements we require only the reduced ones which can be calculated from states with $M=L$, as the others can be obtained from them using the Wigner—Eckart theorem of O(3).

We showed in (4.12I) that a complete, though not orthonormalized set of states with $\nu=\lambda$, $M=L$, and L even, can be written in the operator form

$$|\nu=\lambda,\lambda,\mu,L,M=L\rangle \equiv |\lambda\mu L) = [1,2]^\sigma[2,2]^\tau[3,0]^\mu|0\rangle, \tag{2.1}$$

where $[\nu, L]$ are certain elementary polynomial functions (epd)[1] of the traceless boson creation operators

$$a_m^+ = \eta_m - \Big[\sum_{m'}(-)^{m'}\eta_{m'}\eta_{-m'}\Big](2N+5)^{-1}\xi_m\Big| \qquad m, m' = 2, 1, 0, -1, -2, \tag{2.2}$$

and, in turn, η_m and ξ_m are the standard creation and annihilation operators

$$\eta_m = (1/\sqrt{2})(\alpha_m - i\pi_m), \tag{2.3a}$$

$$\xi_m = (1/\sqrt{2})(\alpha_m + i\pi_m), \tag{2.3b}$$

and $N=\sum_m \eta_m\xi_m$ is the number operator.

The epd are given in (3.10I), (3.20I) when we replace η_m by a_m^+, but as we shall make extensive use of them we give explicitly those appearing in (2.1), i.e.,

$$[1,2] = a_2^+, \tag{2.4a}$$

$$[2,2] = 2\sqrt{2}a_2^+a_0^+ - \sqrt{3}(a_1^+)^2, \tag{2.4b}$$

$$[3,0] = -\sqrt{2}(a_0^+)^3 - 3\sqrt{3}a_2^+(a_{-1}^+)^2 - 3\sqrt{3}a_{-2}^+(a_1^+)^2 + 3\sqrt{2}a_1^+a_0^+a_{-1}^+ + 6\sqrt{2}a_2^+a_0^+a_{-2}^+ \tag{2.4c}$$

We note finally from (4.12aI), that σ, τ, are nonnegative integers satisfying

$$\sigma+\tau = L/2, \tag{2.5a}$$

$$\sigma+2\tau+3\mu=\lambda, \tag{2.5b}$$

$$\sigma,\ \tau,\ \mu \geq 0, \tag{2.5c}$$

where in this and the following sections we shall assume L even, reserving to Sec. 8 the extension of the results to L odd.

We showed in Appendix A of I that from (2.5) the number of states (2.1) consistent with a given value of λ is equal to the dimension d_λ of the IR of O(5) associated with λ.

Are there other ways of constructing all the states of given seniority λ and angular momentum L for which $\nu=\lambda$, $M=L$? We note first that if $\sigma=\tau=0$ the state (2.1) becomes

$$|3\mu,\mu,0\rangle = [3,0]^\mu|0\rangle \tag{2.6}$$

and it corresponds to angular momentum 0 (remember L is even here) and seniority $\lambda=3\mu$. It is explicitly given as a function of β, γ in (2.26I). States of arbitrary angular momentum were obtained then by applying powers of the epd of traceless boson operators $[1,2]$ and $[2,2]$. The essential point for these states to be of definite seniority λ, is that when applying the operator[1]

$$(\overline{2,0}) \equiv \sum_m(-)^m\xi_m\xi_{-m} \tag{2.7}$$

we get zero. As $(\overline{2,0})$ commutes with ξ_m, if we apply any homogeneous polynomial function of the ξ_m to (2,6) we still get a state of definite seniority. This suggests immediately that we can define the states

$$|\lambda\mu L] \equiv (\overline{1,2})^\tau(\overline{2,2})^\sigma[3,0]^{\mu+\sigma+\tau}|0\rangle, \tag{2.8}$$

where $(\overline{1,2})$, $(\overline{2,2})$ are the epd (2.4a), (2.4b) in which ξ_m replaces a_m^+. The total number of quanta of the state (2.8) continues to be given by the λ of (2.5b) while the angular momentum L is related to σ, τ through (2.5a). We have then the same number d_λ of states for a given λ as in the case (2.1).

The states (2.1) are given in what could be called a traceless creation or particle picture, while those of (2.8), which are distinguished by a square instead of round bracket, are in the traceless annihilation or hole picture. If for a given L, λ there is only one possible[1] value of μ the states (2.1) and (2.8) are proportional to each other. If there are several values of μ the states $|\lambda\mu L]$ are linear combinations of $|\lambda\mu' L)$ and vice versa.

Can we construct states in a mixed traceless creation and annihilation operator form or what could be called a traceless "particle—hole" picture? We note that if we apply powers of ξ_2 to the state (2.6) we get a state of definite seniority as when we apply $(\overline{2,0})$ of (2.7) to it we continue to get 0. If we then apply powers of the traceless creation operator a_2^+ to the state just mentioned, the resulting final state continues to be of given seniority and angular momentum because the arguments of Sec. 4 of I. This suggests that we can define a traceless "particle—hole" state

$$|\lambda\mu L\} = (a_2^+)^\sigma\xi_2^\tau[3,0]^{\mu+\tau}|0\rangle. \tag{2.9}$$

As in (2.1) and (2.8) we have that for (2.9) the total number of quanta is still $\lambda=\sigma+2\tau+3\mu$ and the angular momentum is $L=2(\sigma+\tau)$. We have then for the kets (2.9) the same number of states d_λ for a given λ as in the case of (2.1) or (2.8). There is still the question whether all the states (2.9) are independent. The proofs for this proposition based solely on the form (2.9) of the states seems very cumbersome and so we will postpone the discussion of this point to a later publication, where we plan to give compact expressions for the scalar products of the states (2.9) with different μ's and the same L, λ.

We shall use the definition (2.9) of the states, which we characterize by a curly rather than round or square bracket, in the procedure of determining explicitly the corresponding function $\phi_K^{\lambda\mu L}(\gamma)$ in an expansion similar to (1.4). The reason is that in the traceless "particle–hole" picture of the states both ξ_2 and a_2^* involve only first order derivatives[1] with respect to the variable γ. On the other hand in the "particle" or "hole" pictures (2.1) or (2.8) the operators $[2,2]$, $(\overline{2,2})$ involve second order derivatives in γ. Thus it is considerably simpler to proceed in a recursive fashion to obtain $\phi_K^{\lambda\mu L}(\gamma)$ if we are dealing with the traceless "particle–hole" states (2.9).

3. THE POLYNOMIAL IN THE α_m APPEARING IN THE STATE $|\lambda\mu L\}$

When we study the states of the three dimensional oscillator $|\nu LM\rangle$, where ν is the number of quanta and L, M are the angular momentum and its projection, we learn that

$$\nu = L, L, M\rangle = A_L r^L Y_{LM}(\theta,\varphi)\exp(-r^2/2)$$

$$\equiv A_L \mathcal{Y}_{LM}(\mathbf{r})\exp(-r^2/2), \tag{3.1}$$

where A_L is a normalization constant and $\mathcal{Y}_{LM}(\mathbf{r})$ is a solid spherical harmonic which is a homogeneous polynomial of degree L in the components x_q, $q=1,0,-1$ of the position vector. Furthermore, $\mathcal{Y}_{LM}(\mathbf{r})$ satisfies the Laplace equation $\nabla^2\mathcal{Y}_{LM}(\mathbf{r})=0$.

The chain of groups here is $U(3)\supset O(3)$, but it is well known that in the general case $U(n)\supset O(n)$, the states in which the IR of $U(n)$ is the same as that of $O(n)$ (corresponding to the $\nu=L$ of this case) can also be expressed as homogeneous polynomials in the components of the position vector multiplied by a Gaussian. Thus we are certain that we can write

$$|\lambda\mu L\} = P_{\lambda\mu L}(\alpha_m)\exp(-\beta^2/2), \tag{3.2}$$

where $P_{\lambda\mu L}(\alpha_m)$ is an homogeneous polynomial of degree λ in the α_m associated with the angular momentum L and projection $M=L$. Therefore $P_{\lambda\mu L}(\alpha_m)$ satisfies the equations

$$\sum_m \alpha_m \frac{\partial P_{\lambda\mu L}}{\partial\alpha_m} = \lambda P_{\lambda\mu L}, \tag{3.3a}$$

$$L_1 P_{\lambda\mu L} = 0, \tag{3.3b}$$

$$L_0 P_{\lambda\mu L} = L P_{\lambda\mu L}, \tag{3.3c}$$

where[1,15]

$$L_q = \sum_{mm'} \sqrt{6}\langle 21mq|2m'\rangle \alpha_{m'}\frac{\partial}{\partial\alpha_m}, \quad q=1,0,-1. \tag{3.4}$$

The problem we have to solve in (3.3) is thus identical to the one we discussed in Sec. 3 of I only it is now expressed in terms of the α_m variables instead of the η_m. We note though that Eqs. (3.3) do not completely define the $P_{\lambda\mu L}(\alpha_m)$ as there remains the condition[1]

$$\sum_m (-)^m \xi_m \xi_{-m}|\lambda\mu L\} = 0 \tag{3.5}$$

which from (2.3b), (3.2) leads to the equation

$$\sum_m (-)^m \frac{\partial^2 P_{\lambda\mu L}(\alpha_m)}{\partial\alpha_m\partial\alpha_{-m}} = 0. \tag{3.6}$$

If we were able to find the polynomials $P_{\lambda\mu L}(\alpha_m)$ that satisfy the equations (3.3), (3.6), we could then use the relations (1.3) to express them in the form

$$P_{\lambda\mu L}(\alpha_m) = \beta^\lambda \sum_K \phi_K^{\lambda\mu L}(\gamma) D_{LK}^{L*}(\vartheta_i) \tag{3.7}$$

and thus determine the $\phi_K^{\lambda\mu L}(\gamma)$ we are searching for. Unfortunately, while the most general solution of Eqs. (3.3) was given in Sec. 3 of I, the further restriction imposed by (3.6) is very difficult to satisfy. Thus while we shall make use of the solutions of Eqs. (3.3) to introduce some convenient functions of the variable

$$x \equiv \cos 3\gamma \tag{3.8}$$

we shall determine the $\phi_K^{\lambda\mu L}(\gamma)$ directly from the form (2.9) of the states $|\lambda\mu L\}$ in the traceless particle–hole picture.

Returning now to the equations (3.3) we note from (3.28aI) that

$$P_{\lambda\mu L}(\alpha_m) = \sum_{r,s} A_{rs}^{\lambda\mu L}\{1,2\}^{L-\lambda+2r+3s}$$

$$\times\{2,2\}^{\lambda-L/2-3s-2r}\{2,0\}^r\{3,0\}^s, \tag{3.9}$$

where $A_{rs}^{\lambda\mu L}$ are so far arbitrary constants and $\{1,2\}$, $\{2,2\}$, $\{3,0\}=$ are the epd (2.4) in which α_m replaces a_m^*. Finally

$$\{2,0\} = \sum_m (-)^m \alpha_m\alpha_{-m} = \beta^2. \tag{3.10}$$

We note furthermore that as in the $\{3,0\}$ epd we have the α_m coupled to zero angular momentum, we can replace them by the a_m of (1.3b) and thus get, as in (5.2I),

$$\{3,0\} = -\sqrt{2}\beta^3\cos 3\gamma = -\sqrt{2}\beta^3 x. \tag{3.11}$$

Introducing now the index n by the definition

$$n \equiv \lambda - L/2 - 3s - 2r \tag{3.12}$$

we can write

$$P_{\lambda\mu L}(\alpha_m) = \sum_n \{1,2\}^{\sigma+\tau-n}\{2,2\}^n \beta^{3\mu+\tau-n} f_n^{\sigma\tau\mu}(x), \tag{3.13}$$

where for later notational convenience we introduce the indices σ, τ related with λ, L, μ through (2.5). The $f_n^{\sigma\tau\mu}(x)$ is an as yet undetermined polynomial in the variable x.

In the next sections we shall discuss procedures for determining $f_n^{\sigma\tau\mu}(x)$, but here we will continue the analysis of the $P_{\lambda\mu L}(\alpha_m)$ to relate through (3.7), (3.13) the $\phi_K^{\lambda\mu L}(\gamma)$ with the $f_n^{\sigma\tau\mu}(x)$.

We note from the definition of $\{1,2\}$, $\{2,2\}$ and the relations (1.3), (3.10cI) that

$$\beta^{-1}\{1,2\} = \alpha_2/\beta$$

$$= (1/\sqrt{2})[D_{22}^{2*}(\vartheta_i) + D_{2-2}^{2*}(\vartheta_i)]\sin\gamma$$

$$+ D_{20}^{2*}(\vartheta_i)\cos\gamma, \tag{3.14}$$

$$\beta^{-2}\{2,2\}=\sqrt{7}\sum_m D^{2*}_{2m}(\vartheta_i)\beta^{-2}[a\times a]^2_m$$

$$=(-\sqrt{2})\left\{\frac{1}{\sqrt{2}}\left[D^{2*}_{22}(\vartheta_i)+D^{2*}_{2-2}(\vartheta_i)\right]\sin(-2\gamma)+D^{2*}_{20}(\vartheta_i)\cos(-2\gamma)\right\}, \tag{3.15}$$

where $[a\times a]^2_m$ indicates the coupling of two a_m of (1.3) to total angular momentum 2 and projection m. As we have

$$D^{L*}_{MK}=\exp(iM\vartheta_1)\,d^L_{MK}(\vartheta_2)\exp(iK\vartheta_3), \tag{3.16}$$

with a well-known[11] expression for $d^L_{MK}(\vartheta_2)$, we immediately obtain from (3.14) that

$$[\beta^{-1}\{1,2\}]^r=\sum_K D^{2r*}_{2r,K}(\vartheta_i)S^{2r}_K(\gamma), \tag{3.17}$$

where the function $S^{2r}_K(\gamma)$ takes the form

$$S^{2r}_K(\gamma)=\left[\frac{(2r+K)!\,(2r-K)!}{(4r)!}\right]^{1/2}\sum_q\binom{r}{2q-K/2}\binom{2q-K/2}{q}\cdot(\sqrt{6})^r\left(\frac{1}{2\sqrt{3}}\right)^{2q-K/2}(\cos\gamma)^{r+K/2-2q}(\sin\gamma)^{2q-K/2}$$

$$=\left[\frac{(2r+K)!\,(2r-K)!}{(4r)!}\right]^{1/2}\frac{r!\,(\sqrt{6})^r(1/2\sqrt{3})^{K/2}}{(K/2)!\,(r-K/2)!}\cdot(\cos\gamma)^{r-K/2}(\sin\gamma)^{K/2}\,{}_2F_1\left(-\frac{r}{2}+\frac{K}{4},\,-\frac{r}{2}+\frac{K}{4}+\frac{1}{2};\,\frac{K}{2}+1;\,\tfrac{1}{3}\tan^2\gamma\right), \tag{3.18}$$

where ${}_2F_1$ is an hypergeometric function and K is restricted to even values. We note also that if we replace K by $-K$ and $2q$ by $2q-K$ in (3.18) we get an identical expression and thus we have the property

$$S^{2r}_K(\gamma)=S^{2r}_{-K}(\gamma). \tag{3.19}$$

From the above results and (3.15) we then immediately see that

$$[\beta^{-2}\{2,2\}]^n=(-\sqrt{2})^n\sum_K D^{2n*}_{2n,K}(\vartheta_i)S^{2n}_K(-2\gamma). \tag{3.20}$$

Thus we have that the product

$$[\beta^{-1}\{1,2\}]^{L/2-n}[\beta^{-2}\{2,2\}]^n=\sum_K G^{nL}_K(\gamma)D^{L*}_{LK}(\vartheta_i), \tag{3.21}$$

where

$$G^{nL}_K(\gamma)=(-\sqrt{2})^n\sum_{K',K''}\langle L-2n,2n,K',K''\,|\,LK\rangle\times S^{L-2n}_{K'}(\gamma)S^{2n}_{K''}(-2\gamma). \tag{3.22}$$

In (3.22) $\langle\;|\;\rangle$ is an ordinary Clebsch–Gordan coefficient of O(3) and in deriving the result we made only use of the well known decomposition of products[16] of the $D^L_{MK}(\vartheta_i)$.

Remembering that we are discussing the case when L is even we see, from (3.19) and the symmetry properties of Clebsch–Gordan coefficients,[16] that

$$G^{nL}_K(\gamma)=G^{nL}_{-K}(\gamma). \tag{3.23}$$

Introducing now (3.21) into (3.13) and comparing with (3.7), we obtain

$$\phi^{\lambda\mu L}_K(\gamma)=\sum_n G^{nL}_K(\gamma)f^{\sigma\tau\mu}_n(x), \tag{3.24}$$

where $G^{nL}_K(\gamma)$ is the completely defined function of γ of (3.22) while $f^{\sigma\tau\mu}_n(x)$ has as yet to be determined. The σ, τ, are related to λ, L, through (2.5).

In the following sections we indicate procedures for determining the $f^{\sigma\tau\mu}_n(x)$.

4. THE SYSTEM OF COUPLED ORDINARY DIFFERENTIAL EQUATIONS FOR THE $f^{\sigma\tau\mu}_n(x)$

We indicated in the previous section that the polynomial $P_{\lambda\mu L}(\alpha_m)$ given by (3.13) must also satisfy Eq. (3.6), i.e., the five dimensional Laplacian applied to it gives zero. This immediately allows us to obtain a set of coupled ordinary differential equations for $f^{\sigma\tau\mu}_n(x)$. We only require the knowledge that

$$\frac{\partial\{1,2\}}{\partial\alpha_m}=\delta_{m2},\quad \frac{\partial\{2,2\}}{\partial\alpha_m}=2\sqrt{2}\alpha_0\delta_{m2}+2\sqrt{2}\alpha_2\delta_{m0}-2\sqrt{3}\alpha_1\delta_{m1},$$
$$\frac{\partial\beta}{\partial\alpha_m}=\frac{\alpha^m}{\beta},\quad \frac{\partial x}{\partial\alpha_m}=-\frac{1}{\sqrt{2}}\frac{\partial\beta^{-3}\{3,0\}}{\partial\alpha_m}=-\frac{1}{\sqrt{2}\beta^3}\frac{\partial\{3,0\}}{\partial\alpha_m}-\frac{3\alpha^m x}{\beta^2}, \tag{4.1}$$

where $\partial\{3,0\}/\partial\alpha_m$ can be obtained immediately from (2.4c) if we replace there a^+_m by α_m. A straightforward calculation gives then

$$\sum_m(-)^m\frac{\partial^2P_{\lambda\mu L}}{\partial\alpha_m\partial\alpha_{-m}}=\sum_n\{1,2\}^{L/2-n}\{2,2\}^n\beta^{\lambda-L/2-n-2}\left\{9\frac{d}{dx}(1-x^2)\frac{df^{\sigma\tau\mu}_n(x)}{dx}-12\sqrt{2}(n+1)\frac{df^{\sigma\tau\mu}_{n+1}(x)}{dx}-3\sqrt{2}(\sigma+\tau-n+1)\frac{df^{\sigma\tau\mu}_{n-1}(x)}{dx}-3(2\sigma+2\tau+2n)x\frac{df^{\sigma\tau\mu}_n(x)}{dx}+8(n+1)(n+2)f^{\sigma\tau\mu}_{n+2}(x)+\left[(\tau+3\mu-n)(\tau+3\mu-n+3)+(2\sigma+2\tau+2n)(\tau+3\mu-n)\right]f^{\sigma\tau\mu}_n(x)\right\}=0, \tag{4.2}$$

where n is restricted to the interval

$$0\leq n\leq L/2. \tag{4.3}$$

Thus the set of coupled ordinary differential equations for the $f^{\sigma\tau\mu}_n(x)$ is obtained when we set the expression inside the large curly bracket equal to zero for all integer n in the interval (4.3). This set of equations gives, for any given even L and arbitrary λ, results equivalent to those that Bes[3] was able to write explicitly only for L up to 6. He later used them[17] to evaluate particular matrix elements of interest in the collective model.

The set of coupled ordinary differential equations obtained from (4.2) is very difficult to solve in general and in fact Bes[3] for $L\leq 6$ solves them only for small λ's. Thus in our search for the general $f^{\sigma\tau\mu}_n(x)$ we shall follow a different procedure.

We shall first in the next section obtain the states $|\lambda\mu L\}$ of (2.9) as polynomials in the epd of the normal creation operators η_m of (2.3a). Then by arguments similar to those indicated at the beginning of Sec. 3 we show that these polynomials are proportional to the $P_{\lambda\mu L}(\alpha_m)$ of (3.9) if we replace α_m by η_m. Thus we have the coefficients $A_{rs}^{\lambda\mu L}$ appearing in (3.9) and we can determine in Sec. 7 $f_n^{\sigma\tau\mu}$ explicitly as a polynomial in x.

5. THE EXPRESSION OF THE STATES $|\lambda\mu L\}$ IN TERMS OF POLYNOMIALS IN THE EPD'S OF CREATION OPERATORS

The state $|\lambda\mu L\}$ of (2.9) is given as a product of elementary permissible diagrams (epd) in traceless boson creation and annihilation operators. In this section we wish to express them as a polynomial in the epd of the ordinary creation operator η_m acting on the ground state. To achieve this purpose we start from the state with $L=0$ which from (3.28aI) must have the form

$$|3\mu,\mu,0\}=[3,0]^\mu|0\rangle=\sum_r B_r^\mu(2,0)^{3r}(3,0)^{\mu-2r}|0\rangle. \tag{5.1}$$

The coefficients B_r^μ were obtained by Cowan and Sharp[18] by applying $\sum_m(-)^m\xi_m\xi_{-m}$ to (5.1), which must then vanish, and getting a recursion relation for the B_r^μ which they solved. Another procedure consists in remembering that the $P_{\lambda\mu L}(\alpha_m)$ of Sec. 3 for $L=0$ has, from (2.26I), the form

$$\begin{aligned}P_{3\mu,\mu,0}(\alpha_m)&=\beta^{3\mu}P_\mu(x)\\&=\beta^{3\mu}\sum_r(-)^r\frac{(2\mu-2r-1)!!}{2^r r!(\mu-2r)!}x^{\mu-2r}\\&=\frac{(-)^\mu}{2^{\mu/2}}\sum_r\frac{(-)^r(2\mu-2r-1)!!}{r!(\mu-2r)!}\{2,0\}^{3r}\{3,0\}^{\mu-2r},\end{aligned} \tag{5.2}$$

where we made use of (3.10), (3.11).

Comparing now (5.1) and (5.2), where the polynomials satisfy the same equations, only that in the first case they are functions of the η_m and in the second of the α_m, we conclude that we can take

$$B_r^\mu=\frac{(-)^r(2\mu-2r-1)!!}{r!(\mu-2r)!} \tag{5.3}$$

as the states $|\lambda\mu L\}$ are not normalized and thus have an arbitrary multiplicative constant.

We want now to extend the development (5.1) to states with arbitrary even L. We shall indicate in Sec. 8 how to extend all the results obtained so far to odd angular momentum.

As the states $|\lambda\mu L\}$ can be written in the form

$$|\lambda\mu L\}=(a_2^+)^\sigma\xi_2^\tau|3\mu,\mu,0\} \tag{5.4}$$

we see that the application of operator ξ_2 to (5.1) is simple as $\xi_2=\partial/\partial\eta_{-2}$. On the other hand, from the form (2.2) of a_m^+ we see that application of powers a_m^+ to polynomials in the η's is more complicated. As a first step we must put $(a_2^+)^\sigma$ in a convenient form. The expression (2.2) suggests that we can write

$$(a_2^+)^\sigma=\sum_{n=0}^{\sigma}\eta_2^{\sigma-n}(2,0)^n R_n^\sigma(N)\xi_2^n, \tag{5.5}$$

where $R_n^\sigma(N)$ is some function of the number operator alone. In fact for $\sigma=1$ we get from (2.2) that

$$R_0^1(N)=1,\quad R_1^1(N)=-(2N+5)^{-1}. \tag{5.6}$$

We shall prove (5.5) by induction, getting a recursion relation for the $R_n^\sigma(N)$ which we can solve and thus determining it explicitly. We note that

$$\begin{aligned}(a_2^+)^{\sigma+1}&=[\eta_2-(2,0)(2N+5)^{-1}\xi_2]\\&\quad\times\sum_n\eta_2^{\sigma-n}(2,0)^n R_n^\sigma(N)\xi_2^n.\end{aligned} \tag{5.7}$$

Developing this result and making use repeatedly of the relations

$$\xi_2R_n^\sigma(N)=R_n^\sigma(N+1)\xi_2,\quad \eta_2R_n^\sigma(N)=R_n^\sigma(N-1)\eta_2 \tag{5.8a}$$

for arbitrary functions of the number operator N we obtain that

$$\begin{aligned}(a_2^+)^{\sigma+1}&=\sum_{n=0}^{\sigma}\eta_2^{\sigma+1-n}(2,0)^nR_n^\sigma(N)\\&\quad\times[1-2n(2N+2\sigma+2n+3)^{-1}]\xi_2^n-\sum_{n=1}^{\sigma+1}\eta_2^{\sigma+1-n}(2,0)^n\\&\quad\times(2N+2\sigma+2n+3)^{-1}R_{n-1}^\sigma(N+1)\xi_2^n,\end{aligned} \tag{5.8b}$$

which leads to the recursion relation

$$\begin{aligned}&(2N+2\sigma+2n+3)R_n^{\sigma+1}(N)\\&\quad=(2N+2\sigma+3)R_n^\sigma(N)-R_{n-1}^\sigma(N+1),\end{aligned} \tag{5.8c}$$

satisfied by

$$R_n^\sigma(N)=(-)^n\binom{\sigma}{n}\frac{(2N+2\sigma+1)!!}{(2N+2\sigma+2n+1)!!} \tag{5.9}$$

As $R_n^1(N)$, $n=0,1$ gives precisely (5.6) we have the $R_n^\sigma(N)$ we require.

We can now write the state (5.4) in the form

$$|\lambda\mu L\}=\sum_n\eta_2^{\sigma-n}(2,0)^nR_n^\sigma(N)\xi_2^{\tau+n}|3\mu,\mu,0\}. \tag{5.10}$$

Applying ξ_2 as the derivative $\partial/\partial\eta_{-2}$ and noting that

$$\begin{aligned}&\frac{\partial}{\partial\eta_{-2}}(2,0)=2(1,2),\quad\frac{\partial}{\partial\eta_{-2}}(3,0)=3(2,2),\\&\frac{\partial(1,2)}{\partial\eta_{-2}}=\frac{\partial(2,2)}{\partial\eta_{-2}}=0,\end{aligned} \tag{5.11}$$

we obtain finally that

$$\begin{aligned}|\lambda\mu L\}&=\sum_{r,n}C_{rn}^{\sigma\tau\mu}(1,2)^{\sigma+\tau-n}(2,2)^n\\&\quad\times(2,0)^{3r-\tau+n}(3,0)^{\mu+\tau-2r-n}|0\rangle,\end{aligned} \tag{5.12}$$

where the constants $C_{rn}^{\sigma\tau\mu}$ have the form

$$C_{rn}^{\sigma\tau\mu}=B_r^{\mu+\tau}2^{\tau-n}3^n\frac{(3r)!\,(\mu+\tau-2r)!}{(\mu+\tau-2r-n)!}\sum_s 2^s\binom{\tau+s}{n}\frac{R_s^\sigma(3\mu+2\tau-s)}{(3r-\tau+n-s)!}$$

$$=\frac{3^n\sigma!\,\lambda!\,(-)^r2^r(2\mu+2\tau-2r)!\,(3r)!}{2^{\mu+n}n!\,(2\lambda+1)!\,r!\,(\mu+\tau-r)!\,(\mu+\tau-n-2r)!}\sum_s\frac{(-)^s4^s(\tau+s)!\,(2\lambda+1-2s)!}{s!\,(\sigma-s)!\,(\tau-n+s)!\,(3r-\tau+n-s)!\,(\lambda-s)!}\qquad(5.13)$$

as $B_r^{\mu+\tau}$, $R_s^\sigma(3\mu+2\tau-s)$ are in turn given by (5.3), (5.9).

We have thus obtained $|\lambda\mu L\}$ for L even as the polynomial (5.12) in the elementary permissible polynomials (epd) of the η's. We note that by construction this state satisfies the equations

$$N|\lambda\mu L\}=\lambda|\lambda\mu L\},\qquad(5.14a)$$

$$L_1|\lambda\mu L\}=0,\qquad(5.14b)$$

$$L_0|\lambda\mu L\}=L|\lambda\mu L\},\qquad(5.14c)$$

$$\sum_m(-)^m\xi_m\xi_{-m}|\lambda\mu L\}=0,\qquad(5.15)$$

which are similar to Eqs. (3.3), (3.6) satisfied by $P_{\lambda\mu L}(\alpha_m)$. In the following section we shall prove that $P_{\lambda\mu L}(\alpha_m)$ is in fact given by (5.12) if we replace the round epd (ν,L) of η_m by the curly ones $\{\nu,L\}$ which are functions of α_m.

6. RELATIONS BETWEEN THE POLYNOMIALS IN η_m AND α_m FOR STATES OF DEFINITE SENIORITY

For the eventual determination of the $f_n^{\sigma\tau\mu}(x)$ we require to prove that for an r-dimensional harmonic oscillator (and thus in particular for $r=5$), the states of given seniority, i.e., those satisfying the equation

$$\sum_{i=1}^r\xi_i\xi_iP(\eta_j)|0\rangle=0,\quad NP(\eta_j)|0\rangle=\lambda P(\eta_j)|0\rangle\qquad(6.1a,b)$$

can be written as

$$P(\eta_j)|0\rangle=\pi^{-r/4}2^{\lambda/2}P(\alpha_j)\exp(-\beta^2/2),\qquad(6.2)$$

where as before

$$\eta_j=(1/\sqrt{2})(\alpha_j-i\pi_j),\quad\xi_j=(1/\sqrt{2})(\alpha_j+i\pi_j),\quad j=1,2,\ldots,r$$

though now the generalized coordinates, momenta, creation, and annihilation operators are given in Cartesian and not spherical components so that the number operator and β^2 become

$$N=\sum_{i=1}^r\eta_i\xi_i,\quad\beta^2=\sum_{i=1}^r\alpha_i^2.\qquad(6.3)$$

We note that in (6.2) we are assuming the same polynomial P on the left and right hand sides.

The proof of (6.2) was given by Dragt[19] and because of its shortness and importance for our analysis we reproduce it here.

We remember that for the one-dimensional oscillator

$$\eta^n|0\rangle=\pi^{-1/4}2^{-n/2}H_n(\alpha)\exp(-\alpha^2/2)\qquad(6.4)$$

where H_n is an Hermite polynomial whose leading terms are[20]

$$H_n(\alpha)=2^n\alpha^n-2^{n-1}\binom{n}{2}\alpha^{n-2}+\cdots.\qquad(6.5)$$

Thus for the r-dimensional oscillator the state

$$\sum_{n_i}A_{n_1\ldots n_r}\eta_1^{n_1}\cdots\eta_r^{n_r}|0\rangle$$

$$=\pi^{-r/4}2^{-\lambda/2}\sum_{n_i}A_{n_1\ldots n_r}H_{n_1}(\alpha_1)\cdots H_{n_r}(\alpha_r)\exp(-\beta^2/2)$$

$$=\pi^{-r/4}2^{\lambda/2}\sum_{n_i}A_{n_1\ldots n_r}(\alpha_1^{n_1}\alpha_2^{n_2}\cdots\alpha_r^{n_r}+\cdots)\exp(-\beta^2/2),\qquad(6.6)$$

where

$$n_1+n_2+\cdots+n_r=\lambda\qquad(6.7)$$

and the last dots in the bracket in (6.6) stand for monomials in α_i of degree smaller than λ.

In the state (6.6) is of given seniority, i.e., if it also satisfies Eq. (6.1a), then the polynomial in the α's is homogeneous of degree λ and it satisfies the r-dimensional Laplace equation in the α_i. Thus the $A_{n_1n_2\ldots n_r}$ in (6.6) must then be such that all the dotted terms disappear and Eq. (6.2) is proved.

The relation (6.2) holds also when the polynomials are given in terms of η_m,α_m in spherical components as these variables are in turn linear combinations of the Cartesian η_i,α_i, $i=1,\ldots,5$. Thus we immediately have that

$$|\lambda\mu L\}=P_{\lambda\mu L}(\alpha_m)\exp(-\beta^2/2)$$

$$=\pi^{-5/4}2^{\lambda/2}\sum_{r,n}[C_{rn}^{\sigma\tau\mu}\{1,2\}^{\sigma+\tau-n}\{2,2\}^n\{2,0\}^{3r-\tau+n}\times\{3,0\}^{\mu+\tau-2r-n}]\exp(-\beta^2/2)\qquad(6.8)$$

satisfies Eqs. (3.3), (3.6) when the $C_{rn}^{\sigma\tau\mu}$ are given by (5.13) and λ, L are related to σ, τ, μ through (2.5).

Making use of (6.8) we give in the next section the explicit expression of $f_n^{\sigma\tau\mu}(x)$.

7. THE EXPLICIT EXPRESSION OF THE FUNCTIONS $f_n^{\sigma\tau\mu}(x)$

Turning now our attention to the polynomial $P_{\lambda\mu L}(\alpha_m)$ of (6.8) and making use of the expression (3.10), (3.11) of the epd $\{2,0\}$, $\{3,0\}$ in terms of β and x we obtain that

$$P_{\lambda\mu L}(\alpha_m)=\pi^{-5/4}2^{\lambda/2}\sum_n\{1,2\}^{\sigma+\tau-n}\{2,2\}^n\beta^{3\mu+\tau-n}\times\left[\sum_r C_{rn}^{\sigma\tau\mu}x^{\mu+\tau-n-2r}\right](-\sqrt{2})^{\mu+\tau-2r-n}.\qquad(7.1)$$

Thus comparing this expression with $P_{\lambda\mu L}(\alpha_m)$ of (3.13) we obtain

$$f_n^{\sigma\tau\mu}(x)=\pi^{-5/4}(-)^{\mu+\tau-n}2^{(\lambda+\mu+\tau-n)/2}\times\sum_r C_{rn}^{\sigma\tau\mu}2^{-r}x^{\mu+\tau-n-2r},\qquad(7.2)$$

where the $C_{rn}^{\sigma\tau\mu}$ are given by (5.13).

The $f_n^{\sigma\tau\mu}(x)$ is a polynomial as r cannot exceed $\frac{1}{2}(\mu+\tau-n)$ and from (4.3) in turn n is limited to the interval $0 \leq n \leq L/2$. The polynomial $f_n^{\sigma\tau\mu}(x)$ is even (odd) when $\mu+\tau-n$ is even (odd).

Having obtained the $f_n^{\sigma\tau\mu}(x)$ and knowing the $G_K^{nL}(\gamma)$ of (3.22) we then get an explicit expression for the $\phi_K^{\lambda\mu L}(\gamma)$ given by (3.24).

All of the previous results were obtained for L even. In the next section we proceed to extend them to odd angular momentum.

8. EXTENSION OF THE ANALYSIS TO ODD ANGULAR MOMENTA

In the previous sections we restricted ourselves to even angular momentum and we shall now generalize our results to the odd case. It is convenient to designate systematically in this section with a bar above the magnitude related with the problem when the angular momentum is odd. Thus when we write L we mean even and $\bar{L}$ odd values of the angular momenta. We designate also by λ the seniority for even L and by $\bar{\lambda}$ that for odd $\bar{L}$ states. The analysis in Sec. 4 of I indicates then that we can write our states for odd $\bar{L}$ as

$$|\bar{\lambda}\mu\bar{L}) = [1,2]^\sigma[2,2]^\tau[3,3][3,0]^\mu|0\rangle, \qquad (8.1)$$

where from (3.11bI)

$$[3,3] = 2a_{-1}^\dagger(a_2^\dagger)^2 - \sqrt{6}\,a_2^\dagger a_1^\dagger a_0^\dagger + (a_1^\dagger)^3 \qquad (8.2)$$

and the other epd are given in (2.4). We note furthermore that

$$\bar{L} = 2(\sigma+\tau) + 3 \equiv L + 3, \qquad (8.3a)$$

$$\bar{\lambda} = \sigma + 2\tau + 3\mu + 3 \equiv \lambda + 3, \qquad (8.3b)$$

where we introduce an auxiliary even L and a corresponding λ by the definitions in (8.3).

The discussion in Sec. 2 of the present paper indicates immediately that we can have for $\bar{L}$ odd a traceless "particle–hole" state

$$|\bar{\lambda}\mu\bar{L}\} = (a_2^\dagger)^\sigma \xi_2^\tau [3,3][3,0]^{\mu+\tau}|0\rangle. \qquad (8.4)$$

Turning now to Sec. 3 and using (3.28bI) we see that we can also write the state (8.4) in terms of a polynomial α_m to obtain

$$|\bar{\lambda}\mu\bar{L}\} = \{3,3\}\bar{P}_{\lambda\mu L}(\alpha_m)\exp(-\beta^2/2), \qquad (8.5)$$

where

$$\bar{P}_{\lambda\mu L}(\alpha_m) = \sum_n \{1,2\}^{\sigma+\tau-n}\{2,2\}^n\beta^{3\mu+\tau-n}\bar{f}_n^{\sigma\tau\mu}(x) \qquad (8.6)$$

and σ, τ, μ continue to be related with the *unbarred* L, λ defined in (8.3) by (2.5). In (8.5), $\{3,3\}$ is given by (8.2) when we replace $a_m^\dagger$ by α_m.

The *barred* $\bar{f}_n^{\sigma\tau\mu}(x)$ satisfy now a set of ordinary coupled differential equations that come from the equation

$$\sum_m (-)^m \frac{\partial^2}{\partial\alpha_m\partial\alpha_{-m}}[\{3,3\}\bar{P}_{\lambda\mu L}(\alpha_m)]$$
$$= 6\{3,3\}\left[\frac{1}{\beta}\frac{\partial\bar{P}_{\lambda\mu L}}{\partial\beta} - \frac{1}{\beta^2}x\frac{\partial\bar{P}_{\lambda\mu L}}{\partial x}\right]$$
$$+ \{3,3\}\sum_m(-)^m\frac{\partial^2\bar{P}}{\partial\alpha_m\partial\alpha_{-m}} = 0, \qquad (8.7a)$$

where we made use of the explicit form of the epd and of the fact that $\nabla^2\{3,3\}=0$ and from (4.1) we have

$$\sum_m(-)^m\frac{\partial\{3,3\}}{\partial\alpha_m}\frac{\partial Q}{\partial\alpha_{-m}} = \begin{cases} 0 & \text{for } Q = \{1,2\},\{2,2\},\{3,0\}, \\ 6\{3,3\} & \text{for } Q=\{2,0\} \end{cases} \qquad (8.7b)$$

From (4.2), (8.7) we see then immediately that the set of coupled ordinary linear differential equations that the $\bar{f}_n^{\sigma\tau\mu}(x)$ satisfy becomes

$$9\frac{d}{dx}(1-x^2)\frac{d\bar{f}_n^{\sigma\tau\mu}(x)}{dx} - 12\sqrt{2}(n+1)\frac{d\bar{f}_{n+1}^{\sigma\tau\mu}(x)}{dx} - 3\sqrt{2}(\sigma+\tau-n+1)\frac{d\bar{f}_{n-1}^{\sigma\tau\mu}(x)}{dx} - 3(2\sigma+2\tau+2n)x\frac{d\bar{f}_n^{\sigma\tau\mu}(x)}{dx}$$
$$+8(n+1)(n+2)\bar{f}_{n+2}^{\sigma\tau\mu}(x) + [(\tau+3\mu-n)(\tau+3\mu-n+3)+(2\sigma+2\tau+2n)(\tau+3\mu-n)]\bar{f}_n^{\sigma\tau\mu}(x)$$
$$+6(3\mu+\tau-n)\bar{f}_n^{\sigma\tau\mu}(x) - 6x\frac{d\bar{f}_n^{\sigma\tau\mu}(x)}{dx} = 0. \qquad (8.8)$$

As in (3.7) we can now write

$$\{3,3\}\bar{P}_{\lambda\mu L}(\alpha_m) = \beta^{\bar{\lambda}}\sum_{\bar{K}}\phi_{\bar{K}}^{\bar{\lambda}\mu\bar{L}}(\gamma)D_{\bar{L}\bar{K}}^{\bar{L}*}(\vartheta_i) \qquad (8.9)$$

and thus from the discussion in Sec. 3 we immediately obtain

$$\phi_{\bar{K}}^{\bar{\lambda}\mu\bar{L}}(\gamma) = \sum_n G_{\bar{K}}^{n\bar{L}}(\gamma)\bar{f}_n^{\sigma\tau\mu}(x), \qquad (8.10)$$

where

$$G_{\bar{K}}^{n\bar{L}}(\gamma) = \sum_{Kk}\langle L3Kk|\bar{L}\bar{K}\rangle G_K^{nL}(\gamma)g_k^{(3)}(\gamma), \qquad (8.11)$$

in which $\langle\ |\ \rangle$ is a Clebsch–Gordan coefficient of O(3), $G_K^{nL}(\gamma)$ is given by (3.22), and from (3.10bI), (1.3) we have

$$g_k^{(3)}(\gamma) = \sqrt{14/3}\,\beta^{-3}[[a\times a]^2\times a]_k^3$$
$$= (1/\sqrt{3})\sin 3\gamma(\delta_{k2}-\delta_{k,-2}). \qquad (8.12)$$

From (8.12) and the properties of Clebsch–Gordan coefficients we conclude that

$$G_{\bar{K}}^{n\bar{L}}(\gamma) = -G_{-\bar{K}}^{n\bar{L}}(\gamma). \qquad (8.13)$$

We now turn our attention to the development for odd angular momentum of the states $|\bar{\lambda}\mu\bar{L}\}$ in terms of polynomials in the epd of the creation operators alone. As a first step we require the development of the state of lowest odd angular momentum, i.e., $\bar{L}=3$ in terms of epd in the η_m and from (3.28bI) it takes the form

$$|3\mu+3,\mu,3\}=(3,3)\sum_r \bar{B}_r^\mu(2,0)^{3r}(3,0)^{\mu-2r}|0\rangle. \qquad (8.14)$$

To determine the coefficient $\bar{B}_r^\mu$ we recall that the polynomial $P_{3\mu+3,\mu,3}(\alpha_m)$ is, from (2.31), given by

$$P_{3\mu+3,\mu,3}(\alpha_m)$$

$$=\beta^{3\mu+3}P^1_{\mu+1}(x)[D^{3*}_{32}(\vartheta_i)-D^{3*}_{3-2}(\vartheta_i)]$$

$$=-\beta^{3\mu+3}(1-x^2)^{1/2}[D^{3*}_{32}(\vartheta_i)-D^{3*}_{3-2}(\vartheta_i)]$$

$$\times\sum_r(-)^r\frac{(2\mu+1-2r)!!}{(\mu-2r)!(2r)!!}x^{\mu-2r}$$

$$=\beta^{3\mu+3}\sin 3\gamma[D^{3*}_{32}(\vartheta_i)-D^{3*}_{3-2}(\vartheta_i)](-)^{\mu+1}2^{\mu/2}$$

$$\times\sum_r\frac{(-)^r(2\mu+1-2r)!!}{(\mu-2r)!r!}\{2,0\}^{3r}\{3,0\}^{\mu-2r}, \qquad (8.15)$$

where we made use of the expansion of the associated Legendre polynomial $P^1_{\mu+1}(x)$ and Eqs. (3.10), (3.11). From the discussion in section 6 that relates polynomials in the α_m of the type (8.15) with those in the η_m of the form (8.14) and using (8.12) we conclude that we can take for $\bar{B}_r^\mu$ the value

$$\bar{B}_r^\mu=(-)^r\frac{(2\mu+1-2r)!!}{r!(\mu-2r)!}\,. \qquad (8.16)$$

Now, to calculate $|\bar{\lambda}\mu\bar{L}\}$, the procedure is exactly the same as in Sec. 5, and noting in particular that

$$[\xi_2,(3,3)]=0 \qquad (8.17)$$

we obtain immediately that

$$|\bar{\lambda}\mu\bar{L}\}$$

$$=(3,3)\sum_{r,n}\bar{C}^{\sigma\tau\mu}_{rn}(1,2)^{\sigma+\tau-n}(2,2)^n$$

$$\times(3,0)^{\mu+\tau-2r-n}(2,0)^{3r-\tau+n}|0\rangle, \qquad (8.18)$$

where

$$\bar{C}^{\sigma\tau\mu}_{rn}=\bar{B}_r^{\mu+\tau}2^{\tau-n}3^n\frac{(3r)!(\mu+\tau-2r)!}{(\mu+\tau-2r-n)!}$$

$$\times\sum_s\binom{\tau+s}{n}2^s\frac{R_s^\sigma(3\mu+2\tau+3-s)}{(3r-\tau+n-s)!}$$

$$=\frac{3^n\sigma!(\lambda+3)!(-)^r2^r(2\mu+2\tau+1-2r)!(3r)!}{2^{\mu+n}n!(2\lambda+7)!r!(\mu+\tau-r)!(\mu+\tau-n-2r)!}$$

$$\times\sum_{s=0}\frac{(-4)^s(\tau+s)!(2\lambda+7-2s)!}{s!(\sigma-s)!(\tau-n+s)!(3r-\tau+n-s)!(\lambda+3-s)!}\,. \qquad (8.19)$$

Obviously the polynomial $\{3,3\}\bar{P}_{\lambda\mu L}(\alpha_m)$ of (8.5) has the same form as (8.18) but with the curly epd's function of α_m replacing the round ones depending on η_m that appear in the latter. Considering the relation (8.6) that defines $\bar{f}_n^{\sigma\tau\mu}(x)$ we conclude that

$$\bar{f}_n^{\sigma\tau\mu}(x)=\pi^{-5/4}(-)^{\mu+\tau-n}2^{(\mu+\tau-n+\lambda)/2}\sum_r\bar{C}^{\sigma\tau\mu}_{rn}2^{-r}x^{\mu+\tau-n-2r}, \qquad (8.20)$$

where $\bar{C}^{\sigma\tau\mu}_{rn}$ is given by (8.19).

Thus all the results we have derived for even angular momentum can also be obtained in the odd case.

9. THE REDUCED WIGNER COEFFICIENTS IN THE CHAIN O(5) ⊃ O(3)

In Sec. 1 we indicated that the matrix elements we are interested in for the collective model of the nucleus are of the type (1.18) where the operator $T_L^{\lambda\mu L}(\alpha_m)$ is a polynomial in the α_m that satisfies Eqs. (3.3), (3.6) and thus is identical to the polynomial $P_{\lambda\mu L}(\alpha_m)$. Thus we are justified in writing the relation

$$T_M^{\lambda\mu L}(\alpha_m)=\beta^\lambda\sum_K\phi_K^{\lambda\mu L}(\gamma)D^{L*}_{MK}(\vartheta_i) \qquad (9.1)$$

and the matrix element (1.18) requires then for its full determination the reduced Wigner coefficient (RWC) in the $O(5)\supset O(3)$ chain given by (1.20), i.e., a single integral in γ of the product of three $\phi_K^{\lambda\mu L}(\gamma)$ functions with an ordinary $3j$ symbol. Making use of the expression (3.24) we can also write it as

$$(\lambda\mu L;\lambda'\mu'L';\lambda''\mu''L'')$$

$$=\sum_{nn'n''}\int_0^\pi\left\{\left[\sum_{KK'K''}\begin{pmatrix}L&L'&L''\\K&K'&K''\end{pmatrix}G_K^{nL}(\gamma)G_{K'}^{n'L'}(\gamma)G_{K''}^{n''L''}(\gamma)\right]\right.$$

$$\left.\times f_n^{\sigma\tau\mu}(x)f_{n'}^{\sigma'\tau'\mu'}(x)f_{n''}^{\sigma''\tau''\mu''}(x)\right\}\sin 3\gamma\,d\gamma. \qquad (9.2)$$

In (9.2) $\left(\begin{smallmatrix}L&L'&L''\\K&K'&K''\end{smallmatrix}\right)$ stands for a $3j$-symbol while the $G_K^{nL}(\gamma)$, $f_n^{\sigma\tau\mu}(x)$ are given by (3.22), (7.2), respectively. Note that the RWC are symmetric, except for a phase, under exchange of the triplets $\lambda\mu L$ and thus they correspond to a kind of $3j$ symbol for the chain $O(5)\supset O(3)$ rather than the RWC proper for which certain orthonormalization conditions are required which (9.2) does not satisfy. We shall continue though to call them RWC.

To evaluate (9.2) we note that the square bracket appearing there must be a function of $x=\cos 3\gamma$ only because of the symmetry properties of the functions $G_K^{nL}(\gamma)$ which follow from those of $\phi_K^{\lambda\mu L}(\gamma)$ discussed in Refs. 2 and 15. We can then write

$$\sum_{KK'K''}\begin{pmatrix}L&L'&L''\\K&K'&K''\end{pmatrix}G_K^{nL}(\gamma)G_{K'}^{n'L'}(\gamma)G_{K''}^{n''L''}(\gamma)$$

$$=\sum_r M_r(nL,n'L',n''L'')P_r(x), \qquad (9.3)$$

where the coefficients $M_r(nL,n'L',n''L'')$ will be discussed in another publication and $P_r(x)$ are Legendre polynomials of order r. Thus we can finally write for the RWC in the $O(5)\supset O(3)$ chain the expression

$$(\lambda\mu L;\lambda'\mu'L';\lambda''\mu''L'')$$

$$=\sum_{nn'n''}\sum_r M_r(nL,n'L',n''L'')$$

$$\times\int_{-1}^1 P_r(x)f_n^{\sigma\tau\mu}(x)f_{n'}^{\sigma'\tau'\mu'}(x)f_{n''}^{\sigma''\tau''\mu''}(x)\,dx. \qquad (9.4)$$

If we put in the explicit form of the polynomial functions appearing in (9.4) all integrals reduce to the trivial one

$$\int_{-1}^1 x^p\,dx=[1+(-1)^p](p+1)^{-1}, \qquad (9.5)$$

and thus we obtain the RWC in the $O(5)\supset O(3)$ chain as an explicit summation of products of factorials.

Programs are being developed for the evaluation of (9.3), (9.4) and they will be published together with pertinent tables in the book mentioned in the introduction. We wish to stress that for the problem of potential energy surfaces discussed in Sec. 1 we make use only of the RWC

$$(3\mu, \mu, 0; \lambda'\mu' L; \lambda''\mu'' L), \tag{9.6}$$

while for the quadrupole transitions we require only

$$(1, 0, 2; \lambda'\mu' L'; \lambda''\mu'' L''), \tag{9.7a}$$

$$(2, 0, 2; \lambda'\mu' L'; \lambda''\mu'' L''). \tag{9.7b}$$

We are also interested in the particular case of RWC (9.6) corresponding to $\mu = 0$, i.e.,

$$(000; \lambda\mu' L; \lambda\mu'' L). \tag{9.8}$$

This is not as simple as the $3j$ symbol

$$\begin{pmatrix} 0 & L & L \\ 0 & M & -M \end{pmatrix} = (2L+1)^{-1/2}(-1)^M \tag{9.9}$$

of the O(3) group. We require its evaluation in compact form for the discussion of the linear independence of the states $|\lambda\mu L\}$ with fixed λ, L, but different μ as mentioned in Sec. 2. Furthermore, programs for it allow us to pass from the complete though not orthonormal set of states $|\lambda\mu L\}$ to an orthonormal one.

While some of the problems mentioned in this paper require further mathematical analysis, we hope to have established to the satisfaction of the reader that many problems in the collective model of the nucleus are essentially of a group theoretical nature.

*Member of the Instituto Nacional de Energía Nuclear (México).
†Member of El Colegio Nacional.

[1] E. Chacón, M. Moshinsky, and R.T. Sharp, J. Math. Phys. **17**, 668 (1976).
[2] A. Bohr, Kgl. Dan. Videnskab. Selk. Mat. Fys. Medd. **26**, 14 (1952); "Rotational States in Atomic Nuclei," thesis, Copenhagen (1954); A. Bohr and B. Mottelson , Kgl. Dan. Videnskab. Selsk. Mat. Fys. Medd. **27**, 16 (1953).
[3] D.R. Bes, Nucl. Phys. **10**, 373 (1959).
[4] G. Gneuss and W. Greiner, Nucl. Phys. A **171**, 449 (1971).
[5] L. von Bernus *et al.*, "A Collective Model for Transitional Nuclei," in *Heavy-Ion, High-Spin States and Nuclear Structure* (International Atomic Energy Agency, Vienna, 1975).
[6] K.T. Hecht, Nucl. Phys. **63**, 177 (1965).
[7] A. Kappatsch, thesis, University Frankfurt/Main (1974).
[8] L. von Bernus, thesis, University Frankfurt/Main (1975).
[9] B.R. Judd, "Lie groups and the Jahn—Teller effect for a color center," Fourth International Colloquium on Group Theoretical Methods in Physics, University of Nijmegen. June 1975; B.R. Judd and E.E. Vogel, Phys. Rev. B **11**, 2427 (1975); B.R. Judd, Can. J. Phys. **52**, 999 (1974).
[10] L. Armstrong, J. Math. Phys. **12**, 953 (1971).
[11] C. Quesne and M. Moshinsky, J. Math. Phys. **12**, 1780 (1971).
[12] E. Chacôn, D. Levi, and M. Moshinsky, J. Math. Phys. **17**, 1919 (1976).
[13] F. Iachello and A. Arima, Phys. Lett. B **53**, 309 (1974); A. Arima and F. Iachello, Phys. Lett. B **57**, 39 (1975).
[14] T.A. Brody and M. Moshinsky, *Tables of Transformation Brackets* (Gordon & Breach, New York, 1967).
[15] J.M. Eisenberg and W. Greiner, *Nuclear Models* (North-Holland, Amsterdam, 1970), p. 60.
[16] M.E. Rose, *Elementary Theory of Angular Momentum* (Wiley, New York, 1957), pp. 58, 52, 41.
[17] G.G. Dussel and D.R. Bes, Nucl. Phys. A **143**, 623 (1970).
[18] W. Cowan and R.T. Sharp (private communication).
[19] A.J. Dragt, J. Math. Phys. **6**, 533 (1965).
[20] I.S. Gradshteyn and I.M. Ryzhik, *Tables of Integrals, Series and Products* (Academic, New York, 1965).

A systematic investigation of the Petrov G_4 types

J. R. Ray and John Charles Zimmerman

Department of Physics and Astronomy, Clemson University, Clemson, South Carolina 29631
(Received 24 May 1976)

In order to investigate exact solutions in general relativistic cosmology, one usually assumes the spacetime possesses symmetry. Here, we study exact solutions for the Petrov four-parameter Lie groups G_4, acting on nonnull hypersurfaces where the energy–momentum tensor is that of a pressureless perfect fluid (a so-called "dust"). We find that the preponderance of solutions are for a spacelike dust and, in several cases, are able to give their explicit forms. Among these are spacelike versions of previously known timelike matter cosmological models.

1. INTRODUCTION

In recent years, there has been an increasing interest in anisotropic and partially anisotropic cosmological models on the part of general relativity theorists. This new attention paid to an area once considered purely as a study of the mathematical properties of the Einstein field equations has been stimulated for several reasons: the search for a method of particle creation in the hadron era, the possible presence of a primordial magnetic field, the affects of shear in the early evolution of the universe, and as a means of explaining the presently observed isotropy of the 2.7 K background radiation, galaxy counts, and the distribution of cosmic ray particles.

To find an exact solution of the field equations of general relativity,

$$R_{ij} - \tfrac{1}{2}g_{ij}R + \Lambda g_{ij} = T_{ij}, \tag{1}$$

where R_{ij} is the Ricci tensor, g_{ij} the metric tensor with signature $(+ + + -)$, Λ the cosmological constant (possibly nonzero),[1] and T_{ij} the energy–momentum tensor, one may assume that the spacetime possesses some type of symmetry. Such an assumption gives us, for example, the well-known Schwarzschild solution. In cosmology, the same approach may also be employed where the symmetries are assoicated with the homogeneity and isotropy of the model. The assumption that a Bianchi three-parameter group acts on a three-dimensional spacelike hypersurface (homogeneity) has been made by numerous authors, beginning with Taub[2] and Heckmann and Schucking[3] (who investigated anistropic cosmological models).

In this paper, we will search for exact solutions to cosmologies where the spacetime admits a four-parameter group of isometries acting on three-dimensional nonnull hypersurfaces in which the matter is described by a pressureless perfect fluid energy–momentum tensor ("dust") and where the cosmological constant may be nonzero.

In Sec. 2, we will give the general form of the metric based on the above assumption, as well as the symmetries of $U^i(x^j)$, the four-velocity of the dust source. Section 3 will contain the itemized results of our study, while in Sec. 4 we discuss the consequences of these findings: In particular, since in all but one case a spacelike dust is allowed if not, in fact, required, some thought must be given to the place of such solutions in general relativity.

2. THE FORM OF THE METRIC AND THE CONDITIONS ON $U^i(x^j)$

Our cosmological solutions will be based on only three assumptions, the first of which is the validity of the description of the gravitational field by Eq. (1). We next require that the spacetime admit a four-parameter Lie group of continuous transformations where the infinitesimal generators of these mappings are written

$$X_A = \underset{(A)}{\xi^i}\frac{\partial}{\partial x^i} \quad (i=1,2,3,4 \quad A=1,2,3,4), \tag{2}$$

where

$$[X_A, X_B] = C^D_{AB}X_D \tag{3}$$

and where C^D_{AB} are the structure constants of the group. The generators satisfy Killing's equation

$$\underset{(A)}{\xi}{}_{i;j} + \underset{(A)}{\xi}{}_{j;i} = 0, \tag{4}$$

which implies that they form a four-parameter group of isometries (the semicolon denotes a covariant derivative). The group (in Petrov's[4] notation a G_4) will act on a three-dimensional nonnull invariant variety V_3 (or "orbit"), and thereby possess a one-dimensional isotropy group, I_1.[5–8] The groups G_4I–G_4VII have three-parameter subgroups which act on a V_3, while the subgroup of G_4VIII acts on a V_2 and is therefore the maximal group on the V_2.[9] In group theoretic terms, G_4I–G_4V have a subalgebra composed of the Killing vectors X_1 and X_2, G_4VI$_i$ have an Abelian subalgebra of (X_1, X_2, X_3), while G_4VII and G_4VIII have three-parameter semisimple subalgebras formed by (X_1, X_2, X_3).

Now, let us find the general form for the metric from our second assumption. Since the order of G_r is four, the dimension of spacetime is four, and the rank q of the matrix M, where

$$M = \left|\underset{(A)}{\xi^i}\right|, \tag{5}$$

is three, the group acts transitively on a V_3 (see Refs. 10 and 11 for a G_4 on a null V_3). Choosing coordinates so that these hypersurfaces are described by $x^4 = \text{const}$, the definition of an absolute invariant implies that $\xi^4 = 0$, and consideration of Killing's equation demonstrates that $g_{4i} = 0$.[12] Then the metric is written as

$$ds^2 = h_{\alpha\beta}(x^i)dx^\alpha dx^\beta + e_4(dx^4)^2 \quad (\alpha, \beta = 1, 2, 3), \tag{6}$$

where $e_4 = \pm 1$ is determined from the condition that the determinant of the metric must be negative, i.e., principle of equivalance is satisfied.

Our third assumption is that the source of the gravitational field is a pressureless perfect fluid, so that

$$T_{ij} = \rho U_i U_j, \tag{7}$$

where ρ is the invariant density of the dust. The motion of the dust in spacetime is restricted by four conditions:

$$\text{(1)}\quad \underset{(A)}{\mathcal{L}_{\xi}} U^i = 0, \tag{8}$$

where $\underset{(A)}{\mathcal{L}_{\xi}}$ is the Lie derivative with respect to the Killing vectors. For a proof of Eq. (8) see Ref. 13

$$\text{(2)}\quad U_i U^i = \begin{cases} +1\text{—spacelike dust,} \\ -1\text{—timelike dust.} \end{cases} \tag{9}$$

The choice of this sign is not arbitrary, as it will be dictated by the signature $(+++-)$, condition (1), and the fact that the determinant of the metric be negative. This is a significant comment in view of our results.

$$\text{(3)}\quad U^i{}_{;j} U^j = 0, \tag{10}$$

which is the geodesic equation.

$$\text{(4)}\quad g_{ij} U^i \underset{(A)}{\xi^j} = \text{const.} \tag{11}$$

These four conditions will give us the form for $U^i(x^j)$ even if we are unable to solve the field equations in a closed form. See Ref. 14 for a brief summary of our results.

3. THE PETROV G_4 TYPES

(i) G_4I

The metric and Killing vectors are correctly listed in the English edition of Ref. 4 where $c = 0$ in the commutation relations when the group is assumed to act on a V_3. The line element may be written

$$ds^2 = 2A(x^4)dx^1dx^3 + B(x^4)[dx^2 + x^1dx^3]^2 + (dx^4)^2, \tag{12}$$

where x^1 is a null coordinate and (x^2, x^3, x^4) are spacelike. Here e_4 must be chosen to be $+1$ to make the determinant negative. We find, by utilizing the conditions on the four-velocity, that

$$U^i = \delta^i_4. \tag{13}$$

Thus, even though neither exact vacuum nor dust solutions are known, any such solutions with a pressureless perfect fluid would represent a spacelike dust.

(ii) G_4II

This algebra, when applied to a V_3, does not allow a gravitational field since all of the metric coefficients vanish.

(iii) G_4III

The Russian edition of Petrov's book[15] contains an error in X_4, whereas Ref. 4 has the correct form for both the metric and the Killing vectors. The 4-velocity is purely timelike with $U^i = \delta^i_4$. There are no known exact vacuum or dust solutions.

(iv) G_4IV

This metric has been extensively investigated by Ray and Foster.[16] They found that

$$U^i = \alpha(x^3)\delta^i_1 + \beta(x^3)\delta^i_3\ , \tag{14}$$

where $\alpha(x^3)$, $\beta(x^3)$ are unknown functions and (x^1, x^2, x^3) are spacelike coordinates. One may perform a transformation to align $U^i(x^j)$ along either the x^1 or x^3 axis, which facilitates solving Eq. (1).

(v) G_4V

This G_4 contains a Bianchi G_3V subgroup acting simply transitively on the V_3 which is a spacelike hypersurface of negative curvature. This metric may also be obtained by a noncentral extension of $G_3\text{VIII}_0$ in the Bianchi–Behr classification.[17] In general, $U^i(x^j)$ has both spacelike and timelike components

$$U^i = \alpha(x^4)\delta^i_1 + \beta(x^4)\delta^i_4 \tag{15}$$

(where α, β are arbitrary functions of x^4, the timelike coordinate) and thus is often referred to as a "tilted" cosmological model,[18] which is of considerable interest in connection with a whimper singularity.[19,20] It is the only such titled dust model among the G_4 on V_3 metrics. The timelike dust solution was found by Farnsworth.[21] It should be noted that a transformation may be performed resulting in $U^i = \delta^i_1$, and a spacelike dust solution analogous to Farnsworth's may be obtained.

(vi) $G_4\text{VI}_1$

By assuming a four-parameter group G_4 which contains a three-parameter Abelian subgroup acting simply transitively on a V_3, one can alter the form of the fourth Killing vector X_4 and change the entire cosmological model. This has been done by Petrov but, unfortunately, a considerable amount of confusion has ensued as reflected in a comparison of the group structures as listed in the English,[4] German,[22] and Russian[15] editions of his book. Because of this difficulty, in this and each of the following three subsections, we will not only present new solutions, but endeavor to establish a uniform notation of $G_4\text{VI}_i$ by listing both the metric and X_4. The first three Killing vectors remain unchanged: They are

$$X_1 = \frac{\partial}{\partial x^2}, \quad X_2 = \frac{\partial}{\partial x^3}, \quad X_3 = -\frac{\partial}{\partial x^1}. \tag{16}$$

For $G_4\text{VI}_1$,

$$ds^2 = A^2(x^4)(dx^1)^2 + 2B^2(x^4)dx^2dx^3 + (dx^4)^2, \tag{17}$$

$$X_4 = x^2\frac{\partial}{\partial x^2} - x^3\frac{\partial}{\partial x^3}, \tag{18}$$

which is denoted by $G_4\text{VI}_1$ (with $k = -1$, $l = 0$, $\epsilon = 1$) in Ref. 15, $G_4\text{VI}_1$ [with case (v) plus $e = 0$, $c = -1$] in Ref. 22, and does not appear in the English edition.[4] The metric may be easily diagonalized to demonstrate that it represents a static spacetime where

$$\begin{gathered} ds^2 = A^2(x^3)(dx^1)^2 + B^2(x^3)(dx^2)^2 \\ + (dx^3)^2 - B^2(x^3)(dx^4)^2, \\ X_1 = \frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial x^2} + \frac{\partial}{\partial x^4}\right), \quad X_2 = \frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial x^2} - \frac{\partial}{\partial x^4}\right), \\ X_3 = -\frac{\partial}{\partial x^1}, \quad X_4 = x^4\frac{\partial}{\partial x^2} + x^2\frac{\partial}{\partial x^4}, \end{gathered} \tag{19}$$

and

$$U^i = \alpha(x^3)\delta_1^i + \beta(x^3)\delta_3^i, \quad (20)$$

admitting a spacelike dust only (after the transformation, x^4 is timelike, x^1, x^2, x^3 are spacelike coordinates, and α, β are arbitrary functions). From Eq. (1), $G_{13} = \rho\alpha\beta = 0$, so we have two possible cases: $\alpha = 0$ or $\beta = 0$.

$\alpha = 0$: For $\Lambda = 0$ and $U^i = \delta_3^i$,

$$\begin{aligned} &A(z) = C_1 z^{2/3} + C_2 z^{-1/3}, \quad B(z) = (3z)^{2/3}, \\ &\rho(z) = \tfrac{4}{3}C_1(C_1 z^2 + C_2 z)^{-1} = 4C_1 3^{1/3}[AB]^{-1}, \end{aligned} \quad (21)$$

where C_1 are constants and $x^3 = z$. If $C_1 = 0$ and $C_2 \neq 0$, then we have a Kasner vacuum solution with a spacelike variable.[23-25] For $C_1 \neq 0$ and $C_2 = 0$, then

$$A \sim z^{2/3}, \quad B \sim z^{2/3}, \quad \rho \sim z^{-2}, \quad (22)$$

which is an Einstein–deSitter model for spacelike dust.

$\beta = 0$: If we allow for the possibility that Λ may be nonzero and negative then we find

$$A = \text{const} \neq 0, \quad B = \exp[-(-\Lambda)^{1/2} z], \quad \rho = -2\Lambda \quad (23)$$

and $R = 6\Lambda$.

There also exist solutions for $\alpha = 0$ $(\Lambda \neq 0)$ and $\beta = 0$ $(\Lambda = 0)$, but they are either not easily integrable or represent flat vacuum solutions. The existence of these latter solutions will be found in most of the cases which we consider, but they will not be explicitly indicated.

(vii) $G_4\mathrm{VI}_2$

Here

$$X_4 = \epsilon x^1 \frac{\partial}{\partial x^1} + k x^2 \frac{\partial}{\partial x^2} + (x^2 + k x^3)\frac{\partial}{\partial x^3} \quad (k, \epsilon = \text{const}) \quad (24)$$

which, when substituted into Killing's equation, shows that this group does not admit a gravitational field because this forces the determinant of the metric to be zero. This conclusion is correctly attributed to $G_4\mathrm{VI}_2$ in Ref. 22, but is listed as $G_4\mathrm{VI}_3$ in Ref. 15, and in the English edition,[4] the correct commutators are given for $G_4\mathrm{VI}_2$, but the group is later mislabeled as $G_4\mathrm{VI}_3$.

(viii) $G_4\mathrm{VI}_3$

As in the previous subsections, we first need to establish a uniform notation. What we will denote by $G_4\mathrm{VI}_3$ is similarly marked in Ref. 22, labeled as $G_4\mathrm{VI}_4$ in Ref. 15, whereas Ref. 4 has the $G_4\mathrm{VI}_3$ commutators mixed with the $G_4\mathrm{VI}_4$ metric. In all cases, $k = \epsilon = 0$. Following the German edition, $G_4\mathrm{VI}_3$ is the group with

$$ds^2 = A(x^4)[2\,dx^1 dx^2 + (dx^3)^2] + B(x^4)(dx^2)^2 + dx^4)^2, \quad (25)$$

$$X_4 = x^2 \frac{\partial}{\partial x^3} - x^3 \frac{\partial}{\partial x^1}, \quad (26)$$

where the conditions on $U^i(x^j)$ yield

$$U^i = \alpha(x^4)\delta_1^i + \delta_4^i. \quad (27)$$

The field equations (1) require $\alpha(x^4) = 0$, so $U^i = \delta_4^i$ which is a spacelike 4-velocity.

For $\Lambda = 0$, $x^4 = \tau$,

$$\begin{aligned} &A(\tau) = (k_1\tau + k_2)^{4/3}, \\ &B(\tau) = C_1(k_1\tau + k_2)^{4/3} + C_2(k_1\tau + k_2)^{1/3}, \\ &\rho = \tfrac{4}{3}k_1^2(k_1\tau + k_2)^{-2}, \end{aligned} \quad (28)$$

where C_i and k_i are constants.

For $\Lambda \neq 0$, we find two sets of solutions,

$$\begin{aligned} &A(\tau) = C_3 \exp[-\tfrac{2}{3}(3\Lambda)^{1/2}\tau], \\ &B(\tau) = C_4 \exp[\tfrac{1}{3}(3\Lambda)^{1/2}\tau] + C_5 \exp[-\tfrac{2}{3}(3\Lambda)^{1/2}\tau], \\ &\rho = 0, \end{aligned} \quad (29)$$

and

$$\begin{aligned} &A(\tau) = C_6 \exp[\tfrac{2}{3}(3\Lambda)^{1/2}\tau], \\ &B(\tau) = C_7 \exp[\tfrac{2}{3}(3\Lambda)^{1/2}\tau] + C_8 \exp[-\tfrac{1}{3}(3\Lambda)^{1/2}\tau], \\ &\rho = 0, \end{aligned} \quad (30)$$

both of which represent curved vacuum solutions.

(ix) $G_4\mathrm{VI}_4$

Here,

$$ds^2 = A^2(x^4)(dx^1)^2 + B^2(x^4)[(dx^2)^2 + (dx^3)^2] - (dx^4)^2, \quad (31)$$

$$X_4 = -x^3 \frac{\partial}{\partial x^2} + x^2 \frac{\partial}{\partial x^3}, \quad (32)$$

which is listed at $G_4\mathrm{VI}_4$ in Ref. 22, $G_4\mathrm{VI}_2$ in Ref. 15, and the $G_4\mathrm{VI}_4$ commutators are with the $G_4\mathrm{VI}_{1,2}$ metric in Ref. 4. Besides having a Bianchi type I subgroup, the Killing vectors (X_1, X_2, X_4) form the group $G_3\mathrm{VII}$ which, when it acts on a V_2, represents the case of planes symmetry.[17,26] We find that $U^i(x^j)$ has both a spacelike and timelike component,

$$U^i = \alpha(x^4)\delta_1^i + \beta(x^4)\delta_4^i, \quad (33)$$

where $\alpha(x^4)$ and $\beta(x^4)$ are arbitrary functions of the timelike coordinate x^4. However, Eq. (1) leads to

$$G_{14} = \rho A^2(x^4)\alpha\beta = 0, \quad (34)$$

so that we have two possible cases: $\alpha = 0$ or $\beta = 0$.

$\alpha = 0$: This case represents a timelike dust. For $\Lambda = 0$, $x^4 = t$,

$$\begin{aligned} &A(t) = K_1(C_1 t + C_2)^{-1/3} + K_2(C_1 t + C_2)^{2/3}, \\ &B(t) = (C_1 t + C_2)^{2/3}, \\ &\rho = \tfrac{4}{3}C_1^2 K_2 [AB^2]^{-1}, \end{aligned} \quad (35)$$

where C_i and K_i are constants and which, for $K_2 = 0$, leads to the Kasner vacuum solution with a timelike variable. This solution has been studied extensively by Thorne,[27] Zel'dovich,[28] Jacobs,[29] and Vajk and Eltgroth.[30]

For $\Lambda \neq 0$,

$$\begin{aligned} &A(t) = K_3 \exp[\tfrac{1}{3}(3\Lambda)^{1/2} t] - K_4 \exp[-\tfrac{2}{3}(3\Lambda)^{1/2} t], \\ &B(t) = K_5 \exp[\tfrac{1}{3}(3\Lambda)^{1/2} t], \quad \rho = 2\Lambda K_4 [AB^2]^{-1}. \end{aligned} \quad (36)$$

As far as the authors are aware, the existence of this solution has been discussed,[3,31,32] but never explicity stated. For $K_4 = 0$, we have a conformally flat deSitter vacuum solution.

$\beta = 0$: For $\Lambda \neq 0$, we find the following spacelike dust solution where $U^i = \delta_1^i$:

$$\begin{aligned} &A(t) = \text{const} \neq 0, \quad B(t) = k \exp[-(\Lambda)^{1/2} t], \\ &\rho = -2\Lambda, \end{aligned} \quad (37)$$

which is similar to a static solution found for G_4VI_1.

(x) G_4VII

This case, and the following one, are the well-known Kantowski–Sachs models,[8,9] where

$$U^i = \alpha(x^4)\delta_1^i + \beta(x^4)\delta_4^i, \tag{38}$$

where (x^1, x^2, x^3) are spacelike and $x^4 = t$ is timelike and α, β are arbitrary functions. The field equations lead to the possibility of two cases: $\alpha = 0$ or $\beta = 0$. The $\alpha = 0$ ($U^i = \delta_4^i$) case has been investigated. For $\beta = 0$, $\Lambda \neq 0$, we find

$$\begin{aligned} &A = \text{const} \neq 0, \\ &B(t) = \Lambda^{-1/2}\sinh(\Lambda)^{1/2}(t + t_0), \\ &\rho = -2\Lambda\coth(\Lambda)^{1/2}(t + t_0), \end{aligned} \tag{39}$$

as an expanding spacelike dust solution ($t_0 = \text{const}$).

(xi) G_4VIII

Proceeding as in the previous section, we again obtain (38) for the four-velocity. The $\alpha = 0$ solutions have been studied by Kantowski and Sachs. For $\beta = 0$, $\Lambda = 0$ there do not exist any real solutions, while for $\beta = 0$, $\Lambda \neq 0$ we find

$$\begin{aligned} &A = \text{const} \neq 0, \\ &B(t) = \tfrac{1}{2}\{\exp[(\Lambda)^{1/2}(t - t_0)] + \Lambda^{-1}\exp[-(\Lambda)^{1/2}(t - t_0)]\}, \\ &\rho = -2\Lambda \end{aligned} \tag{40}$$

which represents an expanding spacelike dust with $U^i = \delta_1^i$ and $x^4 = t$ being a timelike coordinate. For $\Lambda = 1$, $B(t) = \cosh(t - t_0)$.

4. CONCLUSIONS

What we have done is to analyze, in a systematic way, all of the Petrov four-parameter Lie isometry groups which act on a nonnull V_3 and have a dust source, i.e., to solve

$$G_{ij} + \Lambda g_{ij} = \rho U_i U_j, \tag{41}$$

by having made the physically reasonable assumptions that (1) spacetime has an intrinsic symmetry and, (2) the matter in the model universe may be represented by a pressureless perfect fluid. Our approach has encountered solutions which have been previously investigated, as well as demonstrated the existence of spacelike dust versions of the Einstein–deSitter and Farnsworth models.

For a nonzero Λ, and assuming that machine calculations would yield solutions in G_4I and G_4III, we find that spacelike dust models occur for all groups with the exception of G_4III. Whether the affects of Λ are physically observable or it has arisen purely from a mathematical motivation, it should be noted that for a vanishing Λ, we only have spacelike dust solutions in G_4IV, G_4V, G_4VI_1, and G_4VI_3.

The results of our research raise an important question: What is the place of the spacelike dust solutions (sometimes referred to as "tachyons") in general relativistic cosmology? The existence of such tachyon cosmologies is beyond doubt when one considers the work of this paper and those of Gott[33] and Davies.[34] In much of the previous research, one initially proposes a tachyon model and then proceeds to analyze the consequences. Here, we never intended to hunt for such solutions—they just naturally sprang from our assumptions. Also, the tachyons in these models generate the metric and are not solely a test field. It appears that tachyon solutions permeate the set of exact solutions with dust sources especially in the Petrov G_4 on V_3 type gravitational fields. It is interesting that the only homogeneous spacelike dust solutions known are (37), (39), and (40). From one point of view the existence of solely spacelike dust solutions, e.g., G_4IV, for a given metric might imply that the metric is unphysical and should be disregarded. It has been the purpose of this paper to discover in a systematic manner what types of dust solutions arise in gravitational fields having certain symmetries. Whether the suggestion of so many tachyon solutions has physical significance must await future results of experimental physics.

[1] J.E. Gunn and B.M. Tinsley, Nature 257, 454 (1975).
[2] A.H. Taub, Ann. Math. 53, 472 (1951)
[3] O. Heckmann and E. Schucking, in *Gravitation: An Introduction to Current Research*, edited by L. Witten (Wiley, New York, 1962).
[4] A.Z. Petrov, *Einstein Spaces* (Pergamon, London, 1969).
[5] L.P. Eisenhart, *Continuous Groups of Transformations* (Princeton U.P., Princeton, N.J., 1933), Chap. V.
[6] H. Goenner and J. Stachel, J. Math. Phys. 11, 3358 (1970).
[7] M.A.H. MacCallum, in *Cargese Lectures in Physics*, edited by E. Schatzman (Gordon and Breach, New York, 1973).
[8] R. Kantowski, unpublished dissertation, University of Texas, 1966 (available from Univ. Microfilms).
[9] R. Kantowski and R.K. Sachs, J. Math. Phys. 7, 443 (1966).
[10] G.J. Kruchkovich, Mat. Sb. 41, (83), 195 (1957).
[11] W.T. Lauten, III and J.R. Ray, Lett. Nuovo Cimento 14, 63 (1975); J. Math. Phys. 18, 885 (1977).
[12] L.P. Eisenhart, *Riemannian Geometry* (Princeton U.P., Princeton, N.J., 1949), Chap. VI.
[13] J.C. Foster, Unpublished dissertation, Clemson University, 1973 (available from Univ. Microfilms).
[14] J.R. Ray and J.C. Zimmerman, Lett. Nuovo Cimento 15, 457 (1976).
[15] A.Z. Petrov, *Novye metody v obshehyi teorii otnositel'nosti* (Nauka, Moscow, 1966).
[16] J.R. Ray and J.C. Foster, Gen. Rel. Grav. 4, 371 (1973).
[17] I.S. Shikin, Commun. Math. Phys. 26, 24 (1972).
[18] A.R. King and G.F.R. Ellis, Commun. Math. Phys. 31, 209 (1973).
[19] G.F.R.Ellis and A.R. King, Commun. Math. Phys. 38, 119 119 (1974).
[20] C.B. Collins, Commun. Math. Phys. 39 131 (1974)
[21] D.L. Farnsworth, J. Math. Phys. 8, 2315 (1967).
[22] A.Z. Petrov, *Einstein-Raume* (Akademie-Verlag, Berlin, 1964).
[23] A.Z. Petrov, in *Recent Developments in General Relativity*, (Pergamon, New York, 1962).
[24] R.A. Harris and J.D. Zund, Tensor, N.S. 29, 103 (1975).
[25] L.D. Landau and E.M. Lifshitz, *The Classical Theory of Fields* (Pergamon, New York, 1971), Chap. 11.
[26] M. Demianski and L.P. Grishchuk, Commun. Math. Phys. 25, 233 (1972).
[27] K.S. Thorne, Appl. J. 148, 51 (1967).
[28] Ya.B. Zel'dovich, Astron. Zh. 41, 873 [Sov. Astron. AJ 8, 700 (1965)].
[29] K.C. Jacobs, unpublished dissertation, California Institute of Technology, 1969 (available from Univ. Microfilms).
[30] J.P. Vajk and P.G. Eltgroth, J. Math. Phys. 7, 2212 (1970).
[31] G.F.R. Ellis, J. Math. Phys. 8, 1171 (1967).
[32] G.F.R. Ellis and M.A.H. MacCallum, Commun. Math. Phys. 12, 108 (1969).
[33] J.R. Gott, III, Appl. J. 187, 1 (1974).
[34] P.C.W. Davies, Nuovo Cimento B 25, 571 (1975).

Investigations of space–times with four-parameter groups of motions acting on null hypersurfaces

W. T. Lauten, III* and J. R. Ray

Department of Physics and Astronomy, Clemson University, Clemson, South Carolina 29631
(Received 12 July 1976)

An investigation of all metrics having a four-parameter group of symmetries with null three-dimensional orbits is made. An attempt to solve the Einstein field equations using various simple energy–momentum tensors, with one exception, gives incompatible sets of equations. The exception is a solution for a null fluid possessing the $G_4\mathrm{I}_1$ group of symmetries.

I. INTRODUCTION

In a previous paper[1] we investigated space—times having a particular four-parameter group of motions ($G_4\mathrm{VII}_1$) acting on null three-dimensional hypersurfaces. In that work we attempted to solve the Einstein field equations using a number of simple energy—momentum tensors. We obtained the result that no gravitational field was compatible with the symmetry under consideration. We have since investigated all metrics obtained by allowing a four-parameter group to act on null three-dimensional hypersurfaces in an attempt to determine whether this unusual behavior, which we found in space—times with one group of symmetries, was characteristic of spacetimes with other symmetries of the same class.

Briefly, the investigation involved obtaining the form of the Killing vectors from the equations of structure of the group and the assumption that the orbits of the group were null hypersurfaces. Killing's equations,

$$\xi_{i;j}+\xi_{j;i}=0,$$

were then used to obtain the general form of the metric. We then calculated the components of the Einstein tensor and attempted to solve the Einstein field equations,

$$G_{ij}=T_{ij}.$$

Actually Kruchkovich[2] and Petrov[3] have obtained the Killing vectors and general metrics for all four-parameter groups whose orbits are three-dimensional null hypersurfaces. Thus it was only necessary to calculate the Einstein tensor and attempt to solve the fluid equations.

The results of this work are discussed in Sec. II and tabulated in Table I. In Sec. III we present an example which illustrates how the results in Table I were obtained. We have chosen the $G_4\mathrm{I}_1$ symmetry as our example since of all the cases studied (54 in total) this symmetry contains the only case for which a solution was found. Other examples for the $G_4\mathrm{VII}_1$ symmetry have previously been discussed.[1]

II. REVIEW OF RESULTS

Table I presents a tabulation of the results of this work. Listed in the left hand column are all symmetry groups of the type under consideration. The notation is that of Kruchkovich[2] and Petrov.[3] $G_4\mathrm{I}$, for example, means the group has four parameters and is the first

TABLE I. The results of attempting to solve the Einstein equations for various four-parameter groups which generate three-dimensional null surfaces are tabulated below. The crosses imply that a consistent set of equations does not exist. A solution was found for the G_4I_1 symmetry with a perfect fluid as the source.

	Traceless Fields	Dust	Perfect Fluid	Massless Scalar Field	Massive Scalar Field	Massive Vector Field	Dust and E & M	E & M and Massless S.F.	Dust and Massless S.F.
$G_4\mathrm{I}_1$	X	X	soln.	X	X	X	X	X	X
$G_4\mathrm{I}_2$				unphysical $g>0$					
$G_4\,\mathrm{II}$			does not generate null hypresurfaces						
$G_4\,\mathrm{III}$				incomplete group					
$G_4\,\mathrm{IV}$				unphysical $g>0$					
$G_4\,\mathrm{V}$	X	X	X	X	X	X	X	X	X
G_4 VI_2 & VI_4				incomplete group					
$G_4\,\mathrm{VII}_1$	X	X	X	X	X	X	X	X	X
$G_4\,\mathrm{VII}_2$	X	X	X	X	X	X	X	X	X
$G_4\,\mathrm{VIII}_2$	X	X	X	X	X	X	X	X	X
$G_4\,\mathrm{VIII}_2$	X	X	X	X	X	X	X	X	X

one listed in the classification of such groups. The top row indicates the various energy—momentum tensors that were used as sources in the Einstein field equations (e.g., dust, perfect fluid). We also tried massive fermions in a few cases but found no solutions.

One notes that in a number of cases there are two metrics and two sets of Killing vectors corresponding to one set of structure equations, for example $G_4\mathrm{VII}_1$ and $G_4\mathrm{VII}_2$. These different cases arise depending on whether or not a subgroup of the group contains a null Killing vector.[2]

In two cases, $G_4\mathrm{I}_2$ and $G_4\mathrm{IV}$, the metrices obtained are unphysical since their determinants are positive and hence they are not Minkowskian.

There is no group with the $G_4\mathrm{II}$ equations of structure whose operators have null three-dimensional hypersurfaces as their orbits. The structure equations, when solved with the restriction that the Killing vectors act in the null hypersurface, in one case lead to a contradiction and in a second case give Killing vectors which produce a metric with determinant zero.

In three cases, those with the $G_4\mathrm{III}$, $G_4\mathrm{VI}_1$, and $G_4\mathrm{VI}_2$ symmetry, the four-parameter group corresponding to the metric obtained by solving Killing's equations is not the complete group. Upon solving Killing's equations for the Killing vectors, using the obtained metric, we find that the metric actually admits more than the original four Killing vectors. Thus we say the group is not complete.

The crosses (X) in the blocks indicate that a compatible set of Einstein field equations does not exist for those cases. We do not mean that the field equations could not be solved, rather that an attempt to solve the equations resulted in a contradiction. As we mentioned earlier, in only one case, that of a perfect fluid in a space—time with the $G_4\mathrm{I}_1$ symmetry, was a solution found.

III. EXAMPLE

In this section we give an example which illustrates the results tabulated in Table I. We have chosen the $G_4\mathrm{I}_1$ symmetry for this purpose.

The null hypersurface generated by the Killing vectors, in the coordinate system used by Kruchkovich and Petrov, is given by $x^4=\text{const}$. The metric for the $G_4\mathrm{I}_1$ case is

$$ds^2=y^2\exp(-2x^3)[2dx^1\,dx^4+(dx^2)^2]+z^2(dx^3)^2, \tag{3.1}$$

where the functions y and z depend only on x^4.

The Killing vectors are

$$X_1=p_1, \tag{3.2a}$$

$$X_2=p_2, \tag{3.2b}$$

$$X_3=x^2p_1-x^4p_2, \tag{3.2c}$$

$$X_4=2x^1p_1+x^2p_2+p_3, \tag{3.2d}$$

where $p_i=\partial/\partial x^i$. Note that X_1 is a null Killing vector and that none of the Killing vectors have a component in the x^4 direction since they all must lie in the null hypersurface $x^4=\text{const}$.

We diagonalize the metric and obtain

$$ds^2=y^2\exp(-2x^3)[(dx^1)^2+(dx^2)^2-(dx^4)^2]+z^2(dx^3)^2 \tag{3.3}$$

where now the functions y and z depend on x^1-x^4.

The transformed Killing vectors are

$$X_1=\frac{1}{\sqrt{2}}\,(p_1+p_4), \tag{3.4a}$$

$$X_2=p_2, \tag{3.4b}$$

$$X_3=\frac{1}{\sqrt{2}}\,x^2(p_1+p_4)-\frac{1}{\sqrt{2}}\,(x^1-x^4)p_2, \tag{3.4c}$$

$$X_4=(x^1+x^4)(p_1+p_4)+x^2p_2+p_3. \tag{3.4d}$$

The Einstein tensor in the coordinate basis of the metric (3.3) has components

$$G_{11}=\frac{2y'^2}{y^2}-\frac{y'}{y}-\frac{z'}{z}+\frac{2y'z'}{yz}+\frac{3y^2}{z^2}\exp(-2x^3), \tag{3.5a}$$

$$G_{13}=-2z'/z, \tag{3.5b}$$

$$G_{14}=\frac{-2y'^2}{y^2}+\frac{y''}{y}+\frac{z''}{z}-\frac{2y'z'}{yz}, \tag{3.5c}$$

$$G_{22}=\frac{3y^2}{z^2}\exp(-2x^3), \tag{3.5d}$$

$$G_{33}=3, \tag{3.5e}$$

$$G_{34}=2z'/z, \tag{3.5f}$$

$$G_{44}=\frac{2y'^2}{y^2}-\frac{y''}{y}-\frac{z''}{z}+\frac{2y'z'}{yz}-\frac{3y^2}{z^2}\exp(-2x^3), \tag{3.5g}$$

where the prime indicates differentiation with respect to x^1-x^4. All other components are zero. All calculations of the Einstein tensor have been checked by computer calculations.

The scalar curvature is

$$R=-12/z^2. \tag{3.6}$$

The Einstein field equations are

$$G_{ij}=T_{ij}, \tag{3.7}$$

where T_{ij} is the energy—momentum tensor.

It is immediately clear that there are no solutions for traceless fields (e.g., vacuum, electrovac, neutrinos, etc.) since by (3.6) and (3.7) the trace of the energy—momentum tensor can never be zero.

The next case we try is dust with energy—momentum tensor

$$T_{ij}=\rho u_iu_j, \tag{3.8}$$

where ρ is the density and u_i is the 4-velocity of the dust.

The Einstein field equations are

$$G_{11}=\rho(u_1)^2, \tag{3.9a}$$

$$G_{13}=\rho u_1u_3, \tag{3.9b}$$

$$G_{14}=\rho u_1u_4, \tag{3.9c}$$

$$G_{22}=\rho(u_2)^2, \tag{3.9d}$$

$$G_{33}=\rho(u_3)^2, \tag{3.9e}$$

$$G_{34}=\rho u_3u_4, \tag{3.9f}$$

$$G_{44}=\rho(u_4)^2. \tag{3.9g}$$

From (3.9d) we see that u_2 cannot be zero. But $G_{12}=0$ implies $\rho u_1 u_2=0$ or $u_1=0$. If $u_1=0$, then (3.9a) and (3.9c) imply

$$3(y^2/z^2)\exp(-2x^3)=0$$

which is a contradiction.

We now turn to the perfect fluid with energy–momentum tensor given by

$$T_{ij}=(W+P)u_i u_j+g_{ij}P,$$

where W is the energy density and P is the pressure.

The field equations are

$$G_{11}=(W+P)(u_1)^2+y^2\exp(-2x^3)P, \tag{3.10a}$$

$$G_{13}=(W+P)u_1u_3, \tag{3.10b}$$

$$G_{14}=(W+P)u_1u_4, \tag{3.10c}$$

$$G_{22}=(W+P)(u_2)^2+y^2\exp(-2x^3)P, \tag{3.10d}$$

$$G_{33}=(W+P)(u_3)^2+z^2P, \tag{3.10e}$$

$$G_{34}=(W+P)u_3u_4, \tag{3.10f}$$

$$G_{44}=(W+P)(u_4)^2-y^2\exp(-2x^3)P. \tag{3.10g}$$

From Eqs. (3.10a), (3.10c), and (3.10g) we can easily show

$$u_1=-u_4. \tag{3.11}$$

Then subtracting (3.10g) from (3.10a) results in

$$P=3/z^2. \tag{3.12}$$

Now Eqs. (3.10d) and (3.10e) imply

$$u_2=u_3=0, \tag{3.13}$$

which with (3.10b) give

$$z=\text{const} \quad \text{and} \quad P=\text{const}.$$

The Einstein equations now are reduced to one equation

$$2y'^2/y^2-y''/y=f(x^1-x^4), \tag{3.14}$$

where $f(x^1-x^4)=(W+P)(u_1)^2$.

Making the substitution

$$y=1/A,$$

we obtain the equation

$$A''-fA=0 \tag{3.15}$$

which may be solved once the function f is specified.

Since the pressure in this solution is constant we may interpret it as a cosmical constant in the Einstein equations. Furthermore, the 4-velocity of the fluid is null as implied by Eqs. (3.11) and (3.13), hence our solution describes a null fluid. Null fluid solutions have previously been studied by Bonnor.[4] The g_{33} component of the metric is constant, however, no additional symmetry is introduced by this. The solution is not flat as is clear from (3.6).

We next attempt to find a solution for the massless scalar field with energy–momentum tensor

$$T_{ij}=\phi_{,i}\phi_{,j}-\tfrac{1}{2}g_{ij}\phi_{,k}\phi^{,k}$$

where $\phi(x^i)$ is a scalar function of the coordinates, and the comma denotes partial differentiation with respect to the coordinates x^i.

The Einstein equations are

$$G_{11}=(\phi_{,1})^2-\frac{y^2}{2}\exp(-2x^3)\,\phi_{,k}\phi^{,k}, \tag{3.16a}$$

$$G_{13}=\phi_{,1}\phi_{,3}, \tag{3.16b}$$

$$G_{14}=\phi_{,1}\phi_{,4}, \tag{3.16c}$$

$$G_{22}=(\phi_{,2})^2-\tfrac{1}{2}y^2\exp(-2x^3)\,\phi_{,k}\phi^{,k}, \tag{3.16d}$$

$$G_{33}=(\phi_{,3})^2-\frac{z^2}{2}\,\phi_{,k}\phi^{,k}, \tag{3.16e}$$

$$G_{34}=\phi_{,3}\phi_{,4},$$

$$G_{44}=(\phi_{,4})^2+\frac{y^2}{2}\exp(-2x^3)\,\phi_{,k}\phi^{,k}. \tag{3.16g}$$

We can show using Eqs. (3.16a), (3.16c), and (3.16g) that

$$\phi_{,1}=-\phi_{,4}. \tag{3.17}$$

Then subtracting (3.16g) from (3.16a) gives

$$\phi_{,k}\phi^{,k}=-6/z^2, \tag{3.18}$$

but (3.17), (3.18), (3.16d), and (3.16e) imply $\phi_{,k}\phi^{,k}=0$ which is a contradiction.

The massive scalar field gives a result similar to the massless scalar field and can easily be worked out by the interested reader.

We next turn to the massive vector field with energy–momentum tensor

$$T_{ij}=g_{ik}F_{jr}F^{kr}-\tfrac{1}{4}g_{ij}F_{kr}F^{kr}+m^2A_iA_j-\tfrac{1}{2}g_{ij}m^2A_kA^k,$$

where F_{ij} and A_i are related by the equation

$$F_{ij}=A_{j,i}-A_{i,j}. \tag{3.19}$$

In this case we make the reasonable assumption that the vector potential A_i has the G_4I_1 symmetry, that is, the Lie derivative of A_i with respect to the Killing vectors is zero. This is the first time we have had to make any extra assumption of this type. A similar assumption is also made in some other cases, for example, the case of the electromagnetic field coupled with dust. Without such an assumption we cannot make any progress towards a solution. The assumption that the Lie derivative vanishes is written

$$L_\xi A_i=0. \tag{3.20}$$

Equation (3.20) implies that

$$A_i=\alpha(\delta_i^1-\delta_i^4)+\beta\delta_i^3, \tag{3.21}$$

where $\alpha=\alpha(x^1-x^4)$ and $\beta=\beta(x^1-x^4)$. From Eq. (3.19) we find that the only nonvanishing components of F_{ij} are

$$F_{13}=F_{34}=\beta', \tag{3.22}$$

where, again, the prime indicates differentiation with respect to $x^1 - x^4$. One can show that

$$F_{ij}F^{ij} = {}^*F_{ij}F^{ij} = 0,$$

where the star denotes the duality operation. Hence, the field is null.

The field equations are

$$G_{11} = \frac{\beta'^2}{z^2} + m^2\alpha^2 - \frac{y^2}{2z^2}\exp(-2x^3)\, m^2\beta^2, \tag{3.23a}$$

$$G_{13} = m^2\alpha\beta, \tag{3.23b}$$

$$G_{14} = -\beta'^2/z^2 - m^2\alpha^2, \tag{3.23c}$$

$$G_{22} = -\frac{m^2\beta^2}{2}\ \frac{y^2}{z^2}\exp(-2x^3), \tag{3.23d}$$

$$G_{33} = m^2\beta^2/2, \tag{3.23e}$$

$$G_{44} = +\frac{\beta'^2}{z^2} + m^2\alpha^2 + \frac{y^2}{2z^2}\exp(-2x^3)\, m^2\beta^2. \tag{3.23f}$$

Equations (3.23d) and (3.23e) lead to a contradiction.

Finally we attempt to solve the Einstein equations with a coupled electromagnetic field and dust as the source. The energy–momentum tensor is

$$T_{ij} = \rho u_i u_j + g_{ik}F_{jm}F^{km} - \tfrac{1}{4}g_{ij}F_{km}F^{km}.$$

The assumption we make in this case is that the electromagnetic field tensor, F_{ij}, has the G_4I_1 symmetry, that is

$$L_\xi F_{ij} = 0. \tag{3.24}$$

It has been shown by Wainwright and Yaremovicz[5] that for nonnull fields $L_\xi F_{ij} = \tilde{K}\,{}^*F_{ij}$ where $\tilde{K}$ is a scalar. We assume $\tilde{K} = 0$.

Solving for F_{ij} in Eq. (3.24) we obtain

$$F_{ij} = \begin{pmatrix} 0 & \gamma\exp(-x^3) & \nu & 0 \\ -\gamma\exp(-x^3) & 0 & 0 & \gamma\exp(-x^3) \\ -\nu & 0 & 0 & \nu \\ 0 & -\gamma\exp(-x^3) & -\nu & 0 \end{pmatrix}, \tag{3.25}$$

when ν and γ are functions of $x^1 - x^4$. We find the field is null, that is,

$$F^{ij}F_{ij}\ {}^*F_{ij} = 0.$$

The Einstein equations are

$$G_{11} = \rho(u_1)^2 + \frac{\gamma^2}{y^2} + \frac{\nu^2}{z^2}, \tag{3.26a}$$

$$G_{13} = \rho u_1 u_3, \tag{3.26b}$$

$$G_{14} = \rho u_1 u_4 - \frac{\gamma^2}{y^2} - \frac{\nu^2}{z^2}, \tag{3.26c}$$

$$G_{22} = \rho(u_2)^2, \tag{3.26d}$$

$$G_{33} = \rho(u_3)^2, \tag{3.26e}$$

$$G_{34} = -m^2\alpha\beta, \tag{3.26f}$$

$$G_{44} = \rho(u_4)^2 + \frac{\gamma^2}{y^2} + \frac{\nu^2}{z^2}. \tag{3.26g}$$

Equations (3.26a), (3.26c), and (3.26g) give

$$u_1 = -u_4. \tag{3.27}$$

Then subtracting (3.26g) from (3.26a) gives

$$\frac{y^2}{z^2}\exp(-2x^3) = 0$$

which is not allowed.

The remaining cases of the coupled electromagnetic and scalar fields and the coupled dust and scalar field work in a similar way, although there need be no assumption of symmetry in the case of the coupled dust and scalar field.

IV. CONCLUSION

We have investigated all space–time metrics having a four-parameter group of symmetries whose orbits are null three-dimensional hypersurfaces. We have attempted to solve the Einstein equations for each metric using various simple energy–momentum tensors as sources. We have found, with one exception, that each set of Einstein field equations is incompatible.

The one exception is the set of field equations having the G_4I_1 group of symmetries and a perfect fluid as the source of the gravitational field. The perfect fluid turns out to be a null fluid and the constant pressure is interpreted as a cosmical constant in the Einstein tensor.

When we began our investigation of gravitational fields having the four-parameter symmetry on null hypersurfaces, we expected to obtain many solutions similar to the perfect fluid solution, or solutions for plane wave gravitational radiation in matter. It is still not clear why so many cases in this class of symmetries give incompatible field equations. The Bondi–Robinson plane gravitational waves are of type G_4VI_2 which allows a fifth Killing vector.

The Einstein theory of gravitation occasionally presents one with other curious results such as the existence of closed timelike lines, multiple universes, tachyons, and "ghost nuetrinos.[6]" It has been shown that in the Einstein–Cartan theory "ghost neutrino" solutions are not allowed.[7,8] It might be that the Einstein–Cartan theory would admit solutions for some of the symmetries discussed in the paper. This possibility is being investigated by Kuchowicz.[8]

*Present address: Department of Physics, Sweet Briar College, Sweet Briar, Virginia 24595.

[1] W.T. Lauten and J.R. Ray, Lett. Nuovo Cimento 14, 63 (1975).
[2] G.I. Kruchkovich, Math. Sb. 41, 195 (1957).
[3] A.Z. Petrov, *Einstein Spaces* (Pergamon, New York, 1969).
[4] W.B. Bonnor, Commun. Math. Phys. 13, 163 (1969).
[5] J. Wainwright and P.E.A. Yaremovicz, Gen. Rel. Grav. 7, 595 (1976).
[6] T.M. Davis and J.R. Ray, Phys. Rev. D 9, 331 (1974).
[7] P.S. Letelier, Phys. Lett. A 54, 351 (1975).
[8] B. Kuchowicz, private communication.

Lorentz transformations as space–time reflections

Jorge Krause

Departmento de Física Aplicada, Facultad de Ingeniería, Universidad Central de Venezuela, Caracas, Venezuela

(Received 4 August 1976)

A rank-two tensor is built out of the 4-velocities of two inertial observers, which corresponds precisely to the most general Lorentz matrix connecting the two Cartesian frames of the observers. The Lorentz tensor is then factorized as the product of two "complementary" space–time reflections. It is shown that the first tensorial factor performs the very essential task (i.e., FitzGerald contraction and time dilation) of the corresponding Lorentz transformation, while the second factor is just an internal reflection performed in one and the same inertial frame. Thus, in its essential features, a Lorentz transformation between two *different* inertial frames obtains upon performing just *one* space–time reflection. It is also shown that the (same) Lorentz tensor of two inertial observers can be factorized into "complementary" reflections either by two hyperplanes with spacelike normals, or else by two hyperplanes with timelike normals, which geometric meaning is rather simple. An application of the presented formalism to Dirac's 4-spinor transformation law is also briefly discussed.

1. INTRODUCTION

In the present communication we discuss the kinematics of Lorentz transformations from the point of view of the Minkowski geometry. Therefore, we shall adopt from the beginning the absolute 4-geometry standpoint based on the Lorentz transformation themselves, while characterizing a proper orthochronous Lorentz transformation by means of a space–time *tensor*, i.e.,

$$L^{\mu}{}_{\nu}=\delta^{\mu}_{\nu}-(u^{\mu}-v^{\mu})u_{\nu}$$
$$-(u^{\rho}v_{\rho}+1)^{-1}(u^{\mu}+v^{\mu})(\delta^{\lambda}_{\nu}-u^{\lambda}u_{\nu})v_{\lambda}. \qquad (1.1)$$

In Appendix A we present some "vierbein" projection manipulations for the construction of this tensor. L^{μ}_{ν} is a rank-two tensor built exclusively out of the 4-velocities u^{μ} and v^{μ} of two inertial observers,[1] and it represents the most general Lorentz transformation (keeping aside improper and antichronous transformations) connecting the old v-frame with the new u-frame. Once found, one may easily check that Eq. (1.1) indeed corresponds to a Lorentz matrix. Moreover, with the aim of properly interpreting the Lorentz tensor L^{μ}_{ν}, let us briefly consider the active transformation of events

$$x'^{\mu}=L^{\mu}{}_{\nu}x^{\nu} \qquad (1.2)$$

from the v-frame standpoint; namely, we define $x^{\mu}=(t,\mathbf{x})$, $x'^{\mu}=(t',\mathbf{x}')$, $v^{\mu}=(1,\mathbf{0})$, and $u^{\mu}=\gamma(\mathbf{V})(1,\mathbf{V})$, where $\mathbf{V}$ is the 3-velocity of the u-observer relative to the v-frame, and $\gamma(\mathbf{V})=(1-\mathbf{V}^2)^{-1/2}$, as usual. Thus, from Eqs. (1.1) and (1.2), we get

$$t'=\gamma(\mathbf{V})(t-\mathbf{V}\cdot\mathbf{x}), \qquad (1.3)$$
$$x'^{i}=[\delta^{i}_{j}-\mathbf{V}^{-2}(1-\gamma)V^{i}V_{j}](x^{j}-V^{j}t),$$

that is, a proper orthochronous Lorentz transformation. The transformed event x'^{μ} has precisely the same space–time coordinates in the old v-frame as the object event x^{μ} would have once transformed to the new u-frame. This is the well-known feature relating active and passive transformations.[2]

The algebraic form of Eq. (1.1) is perhaps unnecessarily cumbersome; we keep it as it stands, however, for the explicit appearance of the vectors $u^{\mu}-v^{\mu}$ and $u^{\mu}+v^{\mu}$. The orthogonal projector $\delta^{\mu}_{\nu}-u^{\mu}u_{\nu}$, helps clarify the space–time geometry involved in the L^{μ}_{ν} tensor. Indeed, this note aims to show that the above Lorentz tensor can be factorized into two very special "complementary" space–time reflections (along two spacelike directions, or else, for that matter, along two timelike directions), and the rather simple geometric meaning of the involved reflections are exhibited. The group theoretic features related to this work are somewhat well known today, after algebraic investigations on orthogonal[3] and pseudo-orthogonal[4] groups in n-dimensional spaces.[5] In this sense, it should be mentioned here that, in their formal content, our results are special cases of much more general results which hold good for noncompact groups. Our emphasis in this communication, however, is not the group theoretic aspects, instead, it lies in the explicit absolute space–time characterization, of the issues involved.

Going back to Eq. (1.1) we observe that

$$L^{\mu}{}_{\nu}u^{\nu}=v^{\mu}, \qquad (1.4)$$

as expected, while

$$L^{\mu}{}_{\nu}v^{\nu}=[v^{\mu}v_{\nu}-(\delta^{\mu}_{\nu}-v^{\mu}v_{\nu})]u^{\nu}. \qquad (1.5)$$

We see that L^{μ}_{ν} produces the same hyperrotation on the 4-velocities u^{μ} and v^{μ}. This hyperrotation lies in the 2-flat defined by u^{μ} and v^{μ}. In other words, only the hyperplane spanned by u^{μ} and v^{μ} is to be "turned" by the right "angle", while the two-dimensional subspace orthogonal to that hyperplane remains fixed. This picture characterizes the uniqueness of the result obtained. Furthermore, we see that $L^{\mu}_{\nu}L^{\nu}_{\lambda}u^{\lambda}$ corresponds to a reflection of u^{μ} in the hyperplane orthogonal to the (u,v)-flat. Incidentally, as for infinitesimal Lorentz transformations, we write

$$u^{\mu}=v^{\mu}+\epsilon w^{\mu} \qquad (1.6)$$

(with $\epsilon>0$ a parameter of smallness, and $v^{\mu}w_{\mu}=0$), and we readily obtain, from Eq. (1.1),

$$L^{\mu}{}_{\nu}=\delta^{\mu}_{\nu}+\epsilon(v^{\mu}w_{\nu}-w^{\mu}v_{\nu}), \qquad (1.7)$$

to the first order of approximation. Clearly, in the limit $u^\mu \to v^\mu$, we have $L^\mu_\nu = \delta^\mu_\nu$, as it should be. Equation (1.7) neatly shows the well known fact that the skews-symmetric infinitesimal generators of the proper orthochronous Lorentz transformations of a v-frame are necessarily the Fermi–Walker "propagators" along the v^μ world line.[6]

Finally, we wish to mention here that an equivalent representation of a proper orthochronous Lorentz transformation by means of a space–time tensor has been discussed some years ago by Basanski,[7] in the context of the null tetrad formalism.[8] Moreover, a decomposition of the Lorentz transformation matrix into skew-symmetric tensors was attained by Basanski, while showing that any matrix describing a finite proper orthochronous Lorentz transformation of the null tetrad in Minkowski-space–time may be written as a polynomial of the second order in skew-symmetric tensors. Of course, both tensor representations (Basanski's and ours) of the Lorentz matrix are equivalent for they just correspond to a change of the space–time basis used therefor,[9] i.e., instead of null tetrads, we use orthonormal tetrads in this note. However, when one comes to the decomposition of the Lorentz tensor L^μ_ν, substantial differences appear between both geometric approaches. First, Basanski's decomposition obtains in terms of skew-symmetric tensors (i.e., space–time rotations), while ours, as we shall see presently, is attained in terms of symmetric tensors (i.e., space–time reflections). Hence, the geometric meaning of these decompositions is quite different. Next, while Basanski's approach is able to produce a decomposition of a null rotation transformation[10] as the product of two second-rank tensor factors,[11] it fails to produce such a factorization for the most general proper orthochronous Lorentz transformation; namely, Basanski's second order polynomial expression for the L^μ_ν tensor[12] may not in general be factorized into the product of two transformations, each represented by a second rank tensor. The orthonormal tetrad approach (see Appendix A), on the other hand, gives us a representation of the L^μ_ν Lorentz tensor [Eq. (1.1)] which can be factorized quite generally into the product of two "complementary" space–time symmetric reflection tensors with a rather simple geometric meaning.

We end up this note with a brief Appendix B whose only purpose is to show the handiness of the covariant tool presented here, while reviewing the transformation law of Dirac spinors.[13]

2. SPACE-TIME REFLECTIONS

In order to clarify the geometric content of the tensor defined in Eq. (1.1), we are going to show that L^μ_ν can be factorized as the product of two very special and simple space–time reflections.[14] Let N^μ_ν be a space–time reflection tensor by a hyperplane with spacelike normal; say

$$N^\mu_\nu = \delta^\mu_\nu + 2n^\mu n_\nu, \tag{2.1}$$

with $n^\mu n_\mu = -1$. It is immediate that N^μ_ν defines a Lorentz matrix. Let us next define the spacelike unit vector

$$n^\mu = [2(u^\lambda v_\lambda - 1)]^{-1/2}(u^\mu - v^\mu), \tag{2.2}$$

and consider the corresponding reflection tensor,

$$N^\mu_\nu = \delta^\mu_\nu + (u^\lambda v_\lambda - 1)^{-1}(u^\mu - v^\mu)(u_\nu - v_\nu). \tag{2.3}$$

If we explicitly analyze the reflection

$$x''^\mu = N^\mu_\nu x^\nu, \tag{2.4}$$

from the v-frame standpoint, we readily obtain $t'' = t'$, as in Eq. (1.3), and we also get

$$x''^i = [\delta^i_j - \mathbf{V}^{-2}(1+\gamma)V^i V_j](x^j - V^j t). \tag{2.5}$$

The meaning of Eq. (2.5) is easy to grasp. Indeed, if we decompose the second equation in (1.3) into longitudinal and transverse components relative to the velocity $\mathbf{V}$, we obtain the well known result

$$\mathbf{x}'_L = \gamma(\mathbf{V})(\mathbf{x}_L - \mathbf{V}t), \quad \mathbf{x}'_T = \mathbf{x}_T, \tag{2.6}$$

while if we decompose (2.5) in the same manner, we get, instead,

$$\mathbf{x}''_L = -\gamma(\mathbf{V})(\mathbf{x}_L - \mathbf{V}t), \quad \mathbf{x}''_T = \mathbf{x}_T. \tag{2.7}$$

Hence we conclude that, besides a space reflection in the plane orthogonal to $\mathbf{V}$, the essential features of the Lorentz transformation (between the v-frame and the u-frame) are already contained in the space–time reflection (2.4), whose manifestly covariant tensor N^μ_ν has the simple geometric structure presented in Eq. (2.3). Figure 1 is a sketchy space–time diagram representing this reflection.

Finally, then, let us write the longitudinal reflection

$$t' = t'', \quad \mathbf{x}'_L = -\mathbf{x}''_L, \quad \mathbf{x}'_T = \mathbf{x}''_T \tag{2.8}$$

in a manifestly covariant manner, in order to recover the Lorentz transformation tensor L^μ_ν presented in Eq. (1.1). For that matter we define in the v-frame the spacelike vector $V^\mu = (0, \mathbf{V})$. This means that we introduce, in space–time, the projection

$$V^\mu = (u^\lambda v_\lambda)^{-1}(\delta^\mu_\nu - v^\mu v_\nu)u^\nu, \tag{2.9}$$

such that $V^\mu V_\rho = -\mathbf{V}^2$ is given by the 4-scalar

$$V_\mu V^\mu = (u_\lambda v^\lambda)^{-2}[1 - (u_\mu v^\mu)^2]. \tag{2.10}$$

Let M^μ_ν be the reflection tensor

$$M^\mu_\nu = \delta^\mu_\nu + 2m^\mu m_\nu, \tag{2.11}$$

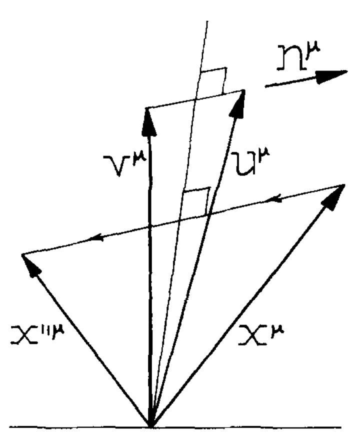

FIG. 1. Space–time diagram of the first reflection: $x''^\mu = N^\mu{}_\nu x^\nu$.

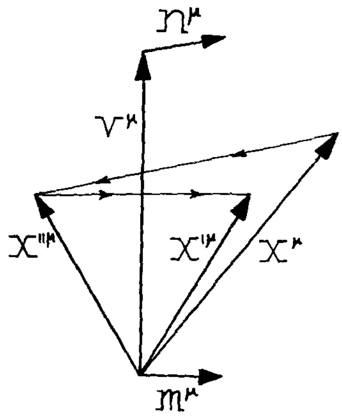

FIG. 2. Space—time diagram of the complementary reflection: $x'^\mu = M^\mu{}_\nu x''^\nu = L^\mu{}_\nu x^\nu$.

with $m^\mu = (-V_\nu V^\nu)^{-1/2} V^\mu$, that is

$$m^\mu = [(u_\lambda v^\lambda)^2 - 1]^{-1/2}(\delta^\mu_\nu - v^\mu v_\nu)u^\nu. \tag{2.12}$$

In this manner, if we now work out explicitly the space—time reflection

$$x'^\mu = M^\mu_\nu x''^\nu, \tag{2.13}$$

in the v-frame, we clearly obtain the longitudinal space-reflection stated in Eq. (2.8). Figure 2 is a space—time diagram representing the reflection (2.13).

3. CONCLUDING REMARKS

Thus we have shown in terms of absolute flat space—time geometry that the Lorentz transformation tensor L^μ_ν can be factorized as the product of two very special space—time reflections; namely

$$L^\mu{}_\nu = (\delta^\mu_\lambda + 2m^\mu m_\lambda)(\delta^\lambda_\nu + 2n^\lambda n_\nu), \tag{3.1}$$

by two hyperplanes whose spacelike normals, n^μ and m^μ, are given in Eqs. (2.2) and (2.12), respectively. In effect, the first factor N^μ_ν is a reflection by the hyperplane orthogonal to the "relative 4-velocity" $u^\mu - v^\mu$, while the second factor M^μ_ν corresponds to a reflection by the hyperplane orthogonal to the V^μ 4-vector. The noncommutative product of these two special reflections simply characterizes the proper orthochronous transformation between the v-frame and the u-frame, i.e., Eq. (1.1). Furthermore, we observe that the second reflection M^μ_ν plays a very secondary role, for it corresponds to a longitudinal reflection of the space coordinates performed in one and the same inertial frame. Hence, essentially, up to this second space reflection, a Lorentz transformation between two *different* frames is already performed by the very simple (first) space—time reflection, i.e.,

$$x''^\mu = [\delta^\mu_\nu + (u_\lambda v^\lambda - 1)^{-1}(u^\mu - v^\mu)(u_\nu - v_\nu)]x^\nu, \tag{3.2}$$

in the hyperplane orthogonal to $u^\mu - v^\mu$ (as shown in Fig. 1). The expected FitzGerald contraction and time dilation are correctly performed by reflection (3.2). The second reflection factor in L^μ_ν merely corresponds to an internal "correction" (say) of the space coordinates taking place within the new inertial frame; i.e., M^μ_ν performs an internal transformation of coordinates (in Møller's sense[15]) within the u-frame.

It must be borne in mind that the two reflections into which we factorize a proper orthochronous Lorentz transformation are not independent. Indeed, it is clear that, given the 4-velocity v^μ and a spacelike unit vector n^μ, which is future-pointing in the v-frame (that is, $v_\mu n^\mu > 0$), we have one and only one 4-velocity vector u^μ satisfying Eq. (2.2); namely

$$u^\mu = 2(v_\nu n^\nu)n^\mu + v^\mu. \tag{3.3}$$

Therefore, it can be shown that the m^μ unit vector *complementary* to the set $\{v^\mu, n^\mu\}$, i.e., able to produce a proper orthochronous Lorentz transformation as in Eq. (3.1), is given by

$$m^\mu(v,n) = [(v_\lambda n^\lambda)^2 + 1]^{-1/2}(\delta^\mu_\nu - v^\mu v_\nu)n^\nu. \tag{3.4}$$

Thus we project u^μ orthogonal to v^μ and normalize for a spacelike unit vector. Hence, every set $\{v^\mu, n^\mu\}$, with $v_\mu n^\mu \geqslant 0$, defines one and only one proper orthochronous Lorentz transformation tensor, as expected. This fact obviously corresponds to the six degrees of freedom of the Lorentz matrix. Clearly, the limit $v_\mu n^\mu = 0$ affords the identity.

Finally, let us briefly comment on the fact that the Lorentz tensor presented in Eq. (1.1) can also be factorized into two complementary space—time reflections by two hyperplanes with timelike normals. It is a matter of straightforward calculation (we leave the details to the reader) to show that

$$L^\mu{}_\nu = [\delta^\mu_\nu - 2v^\mu v_\lambda)] \times [\delta^\lambda_\nu - (1 + u_\rho v^\rho)^{-1}(u^\lambda + v^\lambda)(u_\nu + v_\nu)], \tag{3.5}$$

where, clearly, the first performed reflection is along the direction of the "mean 4-velocity" $u^\mu + v^\mu$ of the two inertial observers. Figure 3 sketchily represents these reflections. Since the second timelike reflection just corresponds to an inversion of the time coordinate taking place within the v-frame, we again conclude that the essential features of the Lorentz transformation (between the v-frame and the u-frame) are already presented as a consequence of performing one simple reflection by the hyperplane orthogonal to the "mean 4-velocity" of the observers.

APPENDIX A: COVARIANT FORMULATION OF LORENTZ TRANSFORMATIONS

In absolute flat space—time let us consider two inertial observers, with given 4-velocities u^μ and v^μ, and their attached orthonormal tetrads, say $\{\alpha^\mu_{(\lambda)}\}$ and $\{\beta^\mu_{(\lambda)}\}$,

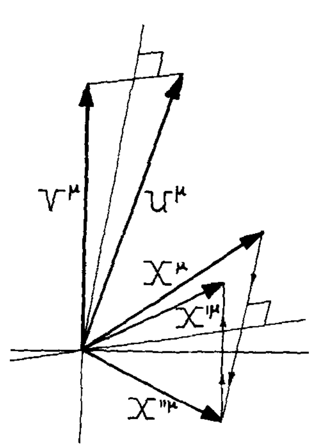

FIG. 3. Space—time diagram of the two timelike reflection factorization stated in Eq. (3.5).

respectively.[16] These tetrads satisfy the orthogonality conditions, as well as the relations of completeness. Furthermore, we assume these tetrads to be such that, by construction,

$$\begin{aligned}\alpha^{\mu}_{(0)} &= v^{\mu}, \quad \beta^{\mu}_{(0)} = u^{\mu}, \\ \alpha^{\mu}_{(1)} &= a^{\mu}, \quad \beta^{\mu}_{(1)} = b^{\mu}, \\ \alpha^{\mu}_{(2)} &= \rho^{\mu}, \quad \beta^{\mu}_{(2)} = \rho^{\mu}, \\ \alpha^{\mu}_{(3)} &= \sigma^{\mu}, \quad \beta^{\mu}_{(3)} = \sigma^{\mu},\end{aligned} \tag{A1}$$

where a^μ and b^μ are spacelike unit vectors belonging in the 2-flat defined by u^μ and v^μ (and, clearly, orthogonal to v^μ and u^μ, respectively), while the spacelike unit vectors ρ^μ and σ^μ are introduced to simultaneously complete both orthonormal sets. This clarifies the task we set ourselves; namely, to construct the one space–time mapping that turns a given timelike future-pointing unit vector u^μ into another given timelike future-pointing unit vector v^μ, i.e., $L^{\mu}_{\nu}u^{\nu} = v^{\mu}$, while the tensor L^{μ}_{ν} must be a covariant function of the vectors u^μ and v^μ, exclusively, which preserves the Minkowskian norm of these vectors. This problem has but one solution [cf. Eq. (1.1)].

It is a well known feature of the "vierbein" tool that these orthonormal bases are related by means of a Lorentz transformation. Indeed, we have that

$$L^{\mu}{}_{\nu} = \alpha^{\mu}_{(\lambda)}\beta^{(\lambda)}_{\nu} \tag{A2}$$

is a Lorentz matrix. Therefore, we immediately recognize in these absolute scaffoldings a Lorentz covariant specification of a special Lorentz transformation, while interpreting the triads $\{\alpha^{\mu}_{(i)}\}$ and $\{\beta^{\mu}_{(i)}\}$ as the rectangular Cartesian basis used by the inertial observers in their respective 3-spaces. Hence, according to Eq. (A2), the corresponding special Lorentz transformation is performed by a Lorentz covariant matrix (i.e., a rank-two 4-tensor) of the form

$$L^{\mu}{}_{\nu} = v^{\mu}u_{\nu} - a^{\mu}b_{\nu} - (\rho^{\mu}\rho_{\nu} + \sigma^{\mu}\sigma_{\nu}). \tag{A3}$$

Since, by construction, a^μ and b^μ are linear combinations of u^μ and v^μ (which shall be worked out presently), in order to get rid of $\rho^\mu\rho_\nu + \sigma^\mu\sigma_\nu$ terms in (A3) we use the fact that L^{μ}_{ν} leaves the Minkowski metric invariant. At once we get

$$L^{\mu}{}_{\nu} = \delta^{\mu}_{\nu} - (u^{\mu} - v^{\mu})u_{\nu} - (a^{\mu} - b^{\mu})b_{\nu}. \tag{A4}$$

Let us then write, *ex hypothesis*,

$$a^{\mu} = Av^{\mu} + Bu^{\mu}, \quad b^{\mu} = Cv^{\mu} + Du^{\mu}. \tag{A5}$$

After some calculations, using the facts $v_{\mu}a^{\mu} = 0$ and $a_{\mu}a^{\mu} = -1$, the following projection obtains

$$a^{\mu} = k^{-1}(\delta^{\mu}_{\nu} - v^{\mu}v_{\nu})u^{\nu}, \tag{A6}$$

where we define

$$k = [(u_{\lambda}v^{\lambda})^2 - 1]^{1/2} > 0. \tag{A7}$$

(Clearly, $k^2 > 0$ because of the Schwarz inequality for timelike vectors.) By the same token, *mutatis mutandis*, we get the projection

$$b^{\mu} = -k^{-1}(\delta^{\mu}_{\nu} - u^{\mu}u_{\nu})v^{\nu}. \tag{A8}$$

Therefore,

$$(a^{\mu} - b^{\mu})b_{\nu} = (u_{\rho}v^{\rho} + 1)^{-1}(u^{\mu} + v^{\mu})(\delta^{\lambda}_{\nu} - u^{\lambda}v_{\nu})v_{\lambda}. \tag{A9}$$

So we have the final answer to our problem in the manifestly Lorentz covariant formula (1.1) for the L^{μ}_{ν} tensor connecting both tetrads. Hence, this tensor represents the most general Lorentz transformation (keeping aside improper and antichronous transformations) between two inertial frames with given 4-velocities u^μ and v^μ. In effect, the spatial orientations of the triads $\{\alpha^{\mu}_{(i)}\}$ and $\{\beta^{\mu}_{(i)}\}$ have been completely eliminated from the formalism.

APPENDIX B: DIRAC FOUR-SPINOR TRANSFORMATION LAW REVISITED

As a simple application of the covariant reflection factorization of proper orthochronous Lorentz transformations presented in this paper, let us consider the well know instance of the Lorentz covariance of the Dirac equation. The usefulness of the proposed complementary factorization of the Lorentz tensor will become apparent through the compactness of the method afforded for handling Dirac spinors. In particular, it seems interesting to remark that we shall arrive at the finite transformation law for 4-spinors (under proper orthochronous Lorentz transformations) without recourse to the infinitesimal transformations and the ensuing rather lengthy iterative integration process,[17] as is usually done in Lie group theory and its applications. Indeed, we first tackle the task for a space–time reflection, as in Eq. (2.1), and then we are ready to solve the problem by just two finite steps, one for each complementary reflection, as in Eq. (3.1).

As is well known, the Dirac equation will be form invariant under Lorentz transformations provided

$$\gamma^{\mu} = L^{\mu}{}_{\nu}S(L)\gamma^{\nu}S^{-1}(L), \tag{B1}$$

i.e., the Dirac matrices γ^μ, $\mu = 0,1,2,3$, remain unaltered under Lorentz transformation L^{μ}_{ν}. Let us examine invariance under space–time reflection by an hyperplane with spacelike unit normal n^μ, Eq. (2.1), say. We have

$$\gamma^{\mu} = (\delta^{\mu}_{\nu} + 2n^{\mu}n_{\nu})S(n)\gamma^{\nu}S^{-1}(n), \tag{B2}$$

hence $S(n)$ must be such that

$$[S(n),\ \gamma^{\mu}] = -2n^{\mu}S(n)\not{n}, \tag{B3}$$

where, as usual, $\not{n} = n_{\mu}\gamma^{\mu}$. In order to solve condition (B3) for $S(n)$, we consider the Dirac algebra

$$\gamma^{\mu}\gamma^{\nu} + \gamma^{\nu}\gamma^{\mu} = 2\eta^{\mu\nu}. \tag{B4}$$

Thus, we immediately get the anticommutation relation

$$[\not{n},\ \gamma^{\mu}]_{+} = -2n^{\mu}\not{n}\not{n} \tag{B5}$$

(since $\not{n}\not{n} = -1$) which resembles (B3). We next transform the anticommutator into a commutator using a well known trick. We define

$$\gamma^{5} = \gamma^{0}\gamma^{1}\gamma^{2}\gamma^{3}, \tag{B6}$$

so we have

$$\gamma^{5}[\not{n},\ \gamma^{\mu}]_{+} = [\gamma^{5}\not{n},\ \gamma^{\mu}]. \tag{B7}$$

Therefore, upon left multiplication by γ^5, Eq. (B5) becomes

$$[\gamma^5 \not{n},\ \gamma^\mu] = -2n^\mu(\gamma^5 \not{n})\not{n}, \tag{B8}$$

which is precisely of the form (B3). This determines $S(n)$ up to a scalar factor. But since for a reflection we obviously require $S^{-1}(n) = S(-n)$, the scalar factor must be ± 1 (recall $\gamma^5\gamma^5 = -1$). Hence we take

$$S(n) = \gamma^5 \not{n}. \tag{B9}$$

This elemental result answers our first problem: For a spacelike reflection we project the Dirac 4-vector-spinor γ^μ on the reflection vector n^μ, and then we multiply (to the left) this projection scalar by γ^5 in order to have a pseudoscalar spinor matrix.

Incidentally, for a reflection by an hyperplane with timelike normal,

$$u_\mu u^\mu = 1,$$

i.e.,

$$U^\mu_\nu = \delta^\mu_\nu - 2u^\mu u_\nu. \tag{B10}$$

Say, we get, instead of (B3), the commutator

$$[S(u),\ \gamma^\mu] = 2u^\mu S(u)\not{u}. \tag{B11}$$

If we define $n^\mu = iu^\mu$, we formally arrive at (B3) again. Therefore, we have, for a timelike reflection,

$$S(u) = i\gamma^5 \not{u}. \tag{B12}$$

One may easily check that the transformation spinor matrices (B9) and (B12) correspond to the usual ones for special reflections.[18]

As for the Lorentz transformation (B1), we simply observe that, by the same token, for the complementary reflection M^μ_ν, cf. Eq. (2.11), we must have

$$S(m) = \gamma^5 \not{m}, \tag{B13}$$

and thus, to the Lorentz tensor L^μ_ν in Eq. (3.1), we associate the spinor matrix

$$S(L) = S(m)S(n) = \not{m}\not{n} = m_\mu n_\nu \gamma^\mu \gamma^\nu. \tag{B14}$$

This formula answers our problem in a compact fashion: We project the Dirac 4-vector-spinor γ^μ upon the complementary reflection vectors n^μ and m^μ, and take the ordered product of these projections.

Finally, using Eqs. (2.2) and (2.12) for n^μ and m^μ, respectively, after a straightforward calculations, we may write Eq. (B14) more explicitly in the form

$$S(L) = -[2(1 + u_\lambda v^\lambda)]^{-1/2}(1 + \not{v}\not{u}). \tag{B15}$$

The reader may easily convince himself that this 4-scalar matrix corresponds to the usual result by considering some special cases.[18] The manifest covariance of the whole procedure leading to Eq. (B14) should be observed, for it is essentially on this geometric formality that the presented tool's power is resting.

[1]In this article we let Greek indices run over the range 0, 1, 2, 3, and Latin indices over 1, 2, 3. We adopt signature (-2) for the Minkowski metric, i.e., $n_{\mu\nu} = n^{\mu\nu} = (+---)$. We set $c = 1$, throughout.

[2]Cf. J.L. Synge, *Relativity: The Special Theory* (North-Holland, Amsterdam, 1965), 2nd ed., p. 75.

[3]H. Weyl, *The Classical Groups* (Princeton U.P., Princeton, New Jersey, 1946).

[4]V. Bargmann, Ann. Math. **59**, 1 (1954).

[5]As a good general reference see, for instance, M. Hamermesh, *Group Theory* (Addison-Wesley, Reading, Mass., 1962).

[6]See J.L. Synge, *Relativity: The General Theory* (North-Holland, Amsterdam, 1966). See also J. Krause, Int. J. Theor. Phys. **12**, 35 (1975), for some remarks concerning the Fermi–Walker "propagator."

[7]S.L. Basanski, J. Math. Phys. **6**, 1201 (1965).

[8]R.K. Sachs, Proc. R. Soc. London A **264**, 309 (1961).

[9]Cf. S.L. Basanski, Ref. 7, e.g., Eq. (2.2) of that paper.

[10]Null tetrad rotations are well discussed by H. Bondi, F.A.E. Pirani, and I. Robinson, Proc. R. Soc. London A **251**, 519 (1959), and also by R.K. Sachs, Ref. 2.

[11]See S.L. Basanski, Ref. 7, Eqs. (2.7), (3.1), and (3.2).

[12]Cf. S.L. Basanski, Ref. 7, Eq. (3.7) of that article.

[13]Other physical applications of the presented formalism will be published elsewhere.

[14]Perhaps we need to remark that the problem we tackle in this section is *not* to obtain this (proper orthochronous) Lorentz tensor $L^\mu{}_\nu$ as merely the product of two tensors, each of which corresponds to an improper but orthochronous Lorentz transformation. That this can be done, and that the choice of factors is far from unique, is obvious. Hence, we stress the very *special geometric character* of our decomposition.

[15]C. Møller, *The Theory of Relativity* (Oxford U.P., New York, 1969).

[16]Although the orthonormal tetrad approach is clearly not the only way of arriving at the desired result, i.e., Eq. (1.1), we choose this method for it is geometrically neat and appealing. We denote an orthonormal tetrad by $\{\alpha^\mu_{(\nu)}\}$, say, where μ is a tensor index, while (ν) stands for a label of the components of the tetrad. Thus, we denote with (0) the timelike component, and with (i) the spacelike components, of the tetrad. Let also recall that we can raise and lower the labels by means of the Minkowski matrix $n^{(\mu)(\nu)} = n_{(\mu)(\nu)} = \mathrm{diag}(+---)$, as if they were true tensorial indices.

[17]S.S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Harper and Row, New York, 1962), p. 74.

[18]See, for instance, N.N. Bogolioubov and D.V. Chirkov, *Introduction a la théorie quantique des champs* (Dunod, Paris, 1960), p. 55.

Construction of the Yukawa$_2$ field theory with a large external field

Lon Rosen*

Department of Mathematics, University of British Columbia, Vancouver, B.C., Canada V6T 1W5
(Received 9 December 1976)

We consider the Yukawa$_2$ model with (relativistic) interaction density $\lambda\bar{\psi}\Gamma\psi\phi+\mu\phi$, where $\Gamma=1$ or γ_5. For sufficiently large μ, we apply the Glimm–Jaffe–Spencer cluster expansion to construct the infinite volume theory satisfying the Wightman and Osterwalder–Schrader axioms including a positive mass gap.

Much of the recent progress in the study of the Yukawa model (Y_2) has been based on the Matthews–Salam–Seiler (MSS) formula[1] for the Schwinger functions. The virtue of this formula is that the Fermi fields have been "integrated out," and thus one can apply to Y_2 The Euclidean Q-space techniques which have proved so successful for the $P(\phi)_2$ model. In particular, starting with the MSS formula, Magnen and Sénéor[2] and Cooper and Rosen[3] have adapted the Glimm–Jaffe–Spencer cluster expansion[4] to the weakly coupled Y_2 model. In this note we show that Spencer's extension[5] of the cluster expansion to the $P(\phi)_2$ model with a strong external field has a particularly simple analog in the case of Y_2.

As in Ref. 5, we consider the theory in a finite $l_0\times l_1$ rectangle $\Lambda\subset\mathbb{R}^2$ with periodic B.C. on $\partial\Lambda$. This choice of B.C. facilitates shifts in the field and in the mass. With apologies for the notational complexity, we now write down the MSS formula for the Schwinger function for n bosons and m fermion–antifermion pairs:

$$S_\Lambda^P(\mu)=Z_\Lambda^{-1}\int\phi(h_1)\cdots\phi(h_n)\times\det S'(f_i,g_j;\phi)\,\rho_\Lambda(\phi)\exp[\mu\phi(\chi_\Lambda)]\,d\mu_\Lambda^P, \tag{1}$$

where $\mu\in\mathbb{R}$ is the external field, χ_Λ is the characteristic function of Λ, and where:

(a) $f=(f_1,\dots,f_m)$, $g=(g_1,\dots,g_m)$, and $h=(h_1,\dots,h_n)$ are suitable test functions, e.g., h_j in $\mathcal{S}(\Lambda)=\{h\in\mathcal{S}(\mathbb{R}^2)\mid \operatorname{supp}h\subset\Lambda\}$ and f_i,g_j in $\mathcal{S}(\Lambda)\oplus\mathcal{S}(\Lambda)$.

(b) ϕ is the free boson field and $d\mu_\Lambda^P=d\mu_{m_b,\Lambda}^P$ the free boson measure on $\mathcal{S}'(\Lambda)$ with mean 0 and covariance

$$\int\phi(f)\phi(g)\,d\mu_\Lambda^P=(f,(-\Delta_\Lambda^P+m_b^2)^{-1}g)_{L^2(\Lambda)},$$

where $m_b>0$ is the boson mass and $-\Delta_\Lambda^P$ is the Laplacian with periodic B.C. on $\partial\Lambda$.

(c) $S'(f_i,g_j;\phi)=(f_i,(1-\lambda K)^{-1}Sg_j)_0$, where $(\cdot,\cdot)_0$ denotes the inner product on $\mathcal{H}_0=L^2(\Lambda)\oplus L^2(\Lambda)$, $\lambda\in\mathbb{R}$ is the coupling constant, S denotes the integral operator on $\mathcal{H}_0$ with kernel given by the two-point Schwinger function for the fermions with mass $m_f>0$,

$$S(x,y)=\frac{1}{|\Lambda|}\sum_p\exp[ip\cdot(x-y)]\frac{\not{p}+m_f}{p^2+m_f^2}\Gamma. \tag{2}$$

Here $p=(p_0,p_1)$ runs over the appropriate lattice $(2\pi/l_0)\mathbb{Z}\times(2\pi/l_1)\mathbb{Z}$, $\not{p}=p_0\gamma_0+p_1\gamma_1$, and $\Gamma=a+b\gamma_2$ where[2] $a,b\in\mathbb{R}$, $\gamma_2\equiv i\gamma_5\equiv-\gamma_0\gamma_1$, and γ_0,γ_1 are the 2×2 Euclidean γ matrices. Choosing a concrete representation, we may take

$$\gamma_0=\begin{pmatrix}0&i\\i&0\end{pmatrix},\quad\gamma_1=\begin{pmatrix}i&0\\0&-i\end{pmatrix},\quad\gamma_2=\begin{pmatrix}0&-1\\1&0\end{pmatrix}. \tag{3}$$

The operator K is defined by the kernel

$$K(x,y)=S(x,y)\phi(y). \tag{4}$$

(d)

$$\rho_\Lambda(\phi)=\det\nolimits_3(1-\lambda K)\exp[-\lambda^2B(K)], \tag{5}$$

where

$$\det\nolimits_3(1-A)=\exp[\operatorname{Tr}(\ln(1-A)+A+A^2/2)] \tag{6}$$

and

$$B(K)=\tfrac{1}{2}:\operatorname{Tr}(K^2+K^\dagger K):, \tag{7}$$

where $K^\dagger$ is the adjoint of K as an operator on $\mathcal{H}_0$ and $:\;:$ denotes Wick ordering with respect to $d\mu_\Lambda^P$.

(e)

$$Z_\Lambda=\int\rho_\Lambda\exp[\mu\phi(\chi_\Lambda)]\,d\mu_\Lambda^P.$$

We refer the reader to Seiler's paper[1] for a rigorous justification of the above formalism (our use of periodic instead of free B.C. represents only a minor difference). Seiler's main result is that the product $\det S'\cdot\rho_\Lambda$ occurring in (1) is in $L^p(d\mu_\Lambda^P)$ for any $p<\infty$. Of course, there are infinite counterterms to be cancelled in the Y_2 model [e.g., there is a cancellation between the infinite quantities $\operatorname{Tr}K^2$ and $\operatorname{Tr}K^\dagger K$ in (5)]. These cancellations are accomplished in the standard way by introducing momentum cutoffs, performing the cancellations, and then removing the cutoffs. In this note, in order not to obscure the simplicity of the argument, we shall perform the cancellations without explicitly going through this procedure. We can now state our result:

Theorem 1: Let λ, $m_b>0$, $m_f>0$ be given, and consider the Schwinger functions $S_\Lambda^P(\mu)$ defined in (1). There is a $\mu_0=\mu_0(\lambda,m_b,m_f)$ such that for $|\mu|\geq\mu_0$:

(a) The infinite volume Schwinger functions $S(\mu)=\lim_{\Lambda\to\mathbb{R}^2}S_\Lambda^P(\mu)$ exist.

(b) The $S(\mu)$ satisfy the Osterwalder–Schrader axioms[7] including exponential decoupling and hence the corresponding relativistic theory satisfies the Wightman axioms including a positive mass gap.

Is it possible to use this large μ result to establish the existence of the infinite volume limit for arbitrary μ? This question is especially interesting in view of

Fröhlich's result[8] that pseudoscalar Y_2 exhibits symmetry breaking at $\mu=0$ (provided it exists!). In the case of $P(\phi)_2$, Fröhlich and Simon[9] have shown how to go from large μ to small μ by means of the FKG inequalities. Unfortunately, according to some elementary calculations we have made, FKG inequalitites do not seem to hold for Y_2.

Given the machinery of Ref. 2 and 3, the idea of the proof of Theorem 1 is very simple: Starting with the theory with parameters (λ, m_b, m_f, μ), we shift the field ϕ and mass m_b in order to express the Schwinger functions in terms of a theory with parameters $(\lambda, \tilde{m}_b, \tilde{m}_f, 0)$, where $\tilde{m}_b(\mu)$ and $\tilde{m}_f(\mu)$ become arbitrarily large as $|\mu| \to \infty$. But we know[2,3] that the cluster expansion applies to this latter theory in the "weak coupling" region $\{(\lambda, \tilde{m}_b, \tilde{m}_f) \mid |\lambda/\tilde{m}_b| < \epsilon_b,\ |\lambda/\tilde{m}_f| < \epsilon_f\}$, where $\epsilon_b > 0$ and $\epsilon_f > 0$, and this yields the theorem. Actually the above description is an oversimplification for two reasons: (i) Because of the use of periodic B.C. we can only shift to a theory whose parameters *approach* $(\lambda, \tilde{m}_b, \tilde{m}_f, 0)$ as $\Lambda \to \mathbb{R}^2$; (ii) as is clear from the Lagrangian, the new "mass" $\tilde{m}_f$ is in fact a 2×2 matrix except in the special case of scalar Y_2. Fortunately, neither of these complications seriously affects the proof.

We begin by noting the following elementary formulas for the change in measure under the shifts $\phi \to \phi - c$ and $m_b^2 \to \tilde{m}_b^2$ (see, e.g., Ref. 10):

$$d\mu_\Lambda^P(\phi - c) = \text{const} \cdot \exp[c m_b^2 \phi(\chi_\Lambda)]\, d\mu_\Lambda^P(\phi), \tag{8}$$

$$d\mu_{\tilde{m}_b,\Lambda}^P = \text{const} \cdot \exp[-\tfrac{1}{2}(\tilde{m}_b^2 - m_b^2) : \phi^2 : (\chi_\Lambda)]\, d\mu_{m_b,\Lambda}^P. \tag{9}$$

In (9), : : denotes Wick ordering with respect to $d\mu_{m_b,\Lambda}^P$ (or, if we wish, with respect to $d\mu_{\tilde{m}_b,\Lambda}^P$ since changing the Wick ordering only introduces a constant). We apply (8) to the MSS formula (1), where the constant c is to be determined below:

$$\begin{aligned} S_\Lambda(\mu) = \int \prod_i [\phi(h_i) + c \textstyle\int h_i] \det S'(f_i, g_j; \phi + c)\, \rho_\Lambda(\phi + c) \\ \times \exp[(\mu - c m_b^2)\, \phi(\chi_\Lambda)]\, d\mu_\Lambda^P \Big/ \\ \int \rho_\Lambda(\phi + c) \exp[(\mu - c m_b^2)\phi(\chi_\Lambda)]\, d\mu_\Lambda^P. \end{aligned} \tag{10}$$

Now by the definition of S' [see (c) above; we set $\lambda = 1$]

$$\begin{aligned} S'(f_i, g_j; \phi + c) &= (f_i, [1 - S(\phi + c)]^{-1} S g_j) \\ &= (f_i, [1 - (1 - cS)^{-1} S\phi]^{-1} (1 - cS)^{-1} S g_j) \\ &= (f_i, [1 - \tilde{K}]^{-1} \tilde{S} g_j) \\ &\equiv \tilde{S}'(f_i, g_j; \phi) \end{aligned} \tag{11}$$

where $\tilde{S} = (1 - cS)^{-1} S$ and $\tilde{K} = \tilde{S}\phi$. Using the formula

$$(-\not{p} + m_f)(\not{p} + m_f) = p^2 + m_f^2, \tag{12}$$

we calculate $\tilde{S}$ in momentum space

$$\begin{aligned} &\left(1 - c\frac{\not{p} + m_f}{p^2 + m_f^2}\Gamma\right)^{-1} \frac{\not{p} + m_f}{p^2 + m_f^2}\Gamma \\ &= [1 - c(-\not{p} + m_f)^{-1}\Gamma]^{-1} (-\not{p} + m_f)^{-1}\Gamma \\ &= (-\not{p} + \tilde{m}_f)^{-1}\Gamma, \end{aligned}$$

where $\tilde{m}_f \equiv m_f - c\Gamma$. In analogy to (12) we easily check that, for $\Gamma = a + b\gamma_2$,

$$(-\not{p} + \tilde{m}_f)(\not{p} + \tilde{m}_f^*) = p^2 + \tilde{m}_f \tilde{m}_f^* \tag{13}$$

with

$$\begin{aligned} \tilde{m}_f \tilde{m}_f^* &= (m_f - ca - cb\gamma_2)(m_f - ca + cb\gamma_2) \\ &= (m_f - ca)^2 + (cb)^2 \equiv |\tilde{m}_f|^2, \end{aligned} \tag{14}$$

a multiple of the identity. Thus

$$\tilde{S}(x,y) = \frac{1}{|\Lambda|} \sum_p \exp[ip \cdot (x - y)] \frac{\not{p} + m_f^*}{p^2 + |m_f|^2}\Gamma, \tag{15}$$

where for large $|c|$, $|\tilde{m}_f| \approx |c| \cdot |\Gamma|$, where $|\Gamma| = (a^2 + b^2)^{1/2}$. This will provide the large fermion mass $\tilde{m}_f$.

More interesting, perhaps, is the manner in which the renormalization prescription provides the large boson mass $\tilde{m}_b$: We (formally) rewrite (5) as

$$\rho_\Lambda(\phi) = \text{const} \cdot \det(1 - K) \exp(\operatorname{Tr} K - \tfrac{1}{2} : \operatorname{Tr} K^\dagger K :), \tag{16}$$

where, as we warned above, each factor in (16) is actually infinite, so that the worried reader may wish to supply and then remove appropriate cutoffs. Under $\phi \to \phi + c$, K is replaced by $K + cS$ so that

$$\begin{aligned} \rho_\Lambda(\phi + c) &= \text{const} \cdot \det[(1 - cS)(1 - \tilde{K})] \\ &\quad \times \exp[\operatorname{Tr} K - (c/2) \operatorname{Tr}(K^\dagger S + S^\dagger K) - \tfrac{1}{2} : \operatorname{Tr} K^\dagger K :] \\ &= \text{const} \cdot \det(1 - \tilde{K}) \exp[\alpha^P \phi(\chi_\Lambda) - \beta^P : \phi^2 : (\chi_\Lambda)], \end{aligned} \tag{17}$$

where the (infinite) constants α^P and β^P are given by

$$\alpha^P = 2(a m_f - c|\Gamma|^2) |\Lambda|^{-1} \sum_p (p^2 + m_f^2)^{-1},$$

$$\beta^P = |\Gamma|^2 |\Lambda|^{-1} \sum_p (p^2 + m_f^2)^{-1}.$$

By similar calculations using (13) we find that

$$\operatorname{Tr} \tilde{K} - \tfrac{1}{2} \operatorname{Tr} \tilde{K}^\dagger \tilde{K} = \tilde{\alpha}^P \phi(\chi_\Lambda) - \tilde{\beta}^P : \phi^2 : (\chi_\Lambda), \tag{18}$$

where

$$\tilde{\alpha}^P = 2(a m_f - c|\Gamma|^2) |\Lambda|^{-1} \sum_p (p^2 + |\tilde{m}_f|^2)^{-1},$$

$$\tilde{\beta}^P = |\Gamma|^2 |\Lambda|^{-1} \sum_p (p^2 + |\tilde{m}_f|^2)^{-1}.$$

Combining (17) and (18), we obtain, for the interaction factor in (10),

$$\begin{aligned} &\rho_\Lambda(\phi + c) \exp[(\mu - c m_b^2)\phi(\chi_\Lambda)] \\ &= \text{const} \cdot \tilde{\rho}_\Lambda(\phi) \exp[\gamma^P \phi(\chi_\Lambda) - \delta^P : \phi^2 : (\chi_\Lambda)], \end{aligned} \tag{19}$$

where as in (5)

$$\begin{aligned} \tilde{\rho}_\Lambda(\phi) &= \text{const} \cdot \det(1 - \tilde{K}) \exp(\operatorname{Tr} \tilde{K} - \tfrac{1}{2} : \operatorname{Tr} \tilde{K}^\dagger \tilde{K} :) \\ &= \det{}_3(1 - \tilde{K}) \exp[-B(\tilde{K})], \end{aligned}$$

and where the constants

$$\gamma^P = \mu - c m_b^2 + \alpha^P - \tilde{\alpha}^P, \quad \delta^P = \beta^P - \tilde{\beta}^P, \tag{20}$$

are *finite*. (19) is the desired identity involving finite (a.e.) quantities.

We now explain our choice of the constant c. Since we have used periodic B.C. the constants γ^P and δ^P are Λ dependent. In order that the new boson mass $\tilde{m}_b$ not depend on Λ, we consider instead the corresponding constants γ^F and δ^F for free B.C. These are defined as in (20) but with $\alpha^P, \tilde{\alpha}^P, \beta^P, \tilde{\beta}^P$ replaced by the corresponding $\alpha^F, \tilde{\alpha}^F, \beta^F, \tilde{\beta}^F$ defined in the obvious way, e.g., β^F

$= |\Gamma|^2(2\pi)^{-2}\int dp(p^2+m_f^2)^{-1}$. By Lemma III. 3 of Ref. 10,

$$(2\pi)^{-2}\int dp(p^2+m_f^2)^{-1} - |\Lambda|^{-1}\sum_p (p^2+m_f^2)^{-1}$$
$$=\lim_{y\to x}[(-\Delta+m_f^2)^{-1}(x,y)-(-\Delta_\Lambda^P+m_f^2)^{-1}(x,y)]$$
$$=O(\exp[-m_f\min(l_0,l_1)]).$$

Hence $c_1(\Lambda)\equiv\gamma^P-\gamma^F$ and $c_2(\Lambda)\equiv\delta^F-\delta^P$ satisy

$$c_j(\Lambda)=O(\exp[-\min(m_f,|\tilde{m}_f|)\cdot\min(l_0,l_1)]). \tag{21}$$

We choose c (independently of Λ) so that $\gamma^F=\mu-cm_b^2+\alpha^F-\tilde{\alpha}^F=0$. By (21) this leaves a small linear term in (19). Since, for large c, $\alpha^F-\tilde{\alpha}^F\sim|c|\log|\tilde{m}_f|\sim|c|\log|c|$, we have $|c|\sim|\mu|/\log|\mu|$ for large μ. Therefore, $\delta^F=\beta^F-\tilde{\beta}^F\sim\log|c|\sim\log|\mu|$ for large μ, and this gives the large boson mass $\tilde{m}_b$. That is, we combine the identity (19) with the mass shift formula (9), setting $\tilde{m}_b^2-m_b^2=2\delta^F$, to rewrite (10) as

$$S_\Lambda^P(\mu)=\frac{\int\Pi[\phi(h_i)+c\int h_i]\det\tilde{S}'(f_i,g_j;\phi)\tilde{\rho}_\Lambda(\phi)\exp[Q_\Lambda(\phi)]d\mu^P_{\tilde{m}_b,\Lambda}}{\int\tilde{\rho}_\Lambda(\phi)\exp[Q_\Lambda(\phi)]d\mu^P_{\tilde{m}_b,\Lambda}}, \tag{22}$$

where

$$Q_\Lambda(\phi)=c_1(\Lambda)\phi(\chi_\Lambda)+c_2(\Lambda):\phi^2:(\chi_\Lambda). \tag{23}$$

We summarize the discussion to this point in:

Lemma 2: The Schwinger functions (1) for the Y_2 theory with parameters (λ,m_b,m_f,μ) can be expressed as in (22) as a linear combination of the Schwinger functions for a theory with parameters $(\lambda,\tilde{m}_b(\mu),\tilde{m}_f(\mu),0)$ and with an additional interaction term (23), where now $\tilde{m}_f(\mu)$ is a 2×2 matrix of the form $m_f-c(\mu)\Gamma$ [see (15)] and where

(i) as $|\mu|\to\infty$, $\tilde{m}_b(\mu)$ and $|\tilde{m}_f(\mu)|\to\infty$,

(ii) for fixed μ, $c_j(\Lambda)\to 0$ exponentially as $\Lambda\to\mathbb{R}^2$ [see (21)].

We are now in a position to prove our main result:

Sketch of the proof of Theorem 1: As in Spencer's case,[5] the proof follows from Lemma 2 and from the convergence of the cluster expansion for the weakly coupled theory (Refs. 2 and 3). Owing to our greater familiarity with the second of these references, we shall base our discussion on the version of the cluster expansion to be found there. We work with the "transformed theory" [i.e., in the form (22)] and we consider only rectangles Λ sufficiently large so that $|c_j(\Lambda)|\leq 1$. For purposes of establishing Euclidean covariance we also introduce an additional spatial cutoff function $g(x)$ into (1) or (22) by replacing each $\phi(x)$ by $g(x)\phi(x)$ except for the $\phi(h_j)$'s. Here $g(x)$ is a measurable function of compact support satisfying $0\leq g(x)\leq 1$.

The steps in the argument are:

1. Given any $\alpha>0$ we choose $\mu_0=\mu_0(\alpha,\lambda,m_b,m_f)$ so that if $|\mu|\geq\mu_0$, then the transformed theory satisfies the cluster property with rate $\exp(-\alpha d)$. More precisely,[11] if $A,B\subset\mathbb{R}^2$ and if F_A is a trace-class operator on $\wedge^m\mathcal{H}$ which is a function of ϕ localized in A, and similarly for F_B, then

$$\langle\tau_{m_A+m_B}(F_A\wedge F_B)\rangle^P_{\Lambda,g,Q}-\langle\tau_{m_A}(F_A)\rangle^P_{\Lambda,g,Q}\langle\tau_{m_B}(F_B)\rangle^P_{\Lambda,g,Q}\leq\beta\exp[-\alpha d(A,B)], \tag{24}$$

where β is a constant independent of Λ,g and of translations of A and B, $d(A,B)$ is the distance between A and B, τ_m is the functional of trace-class operators on $\wedge^m\mathcal{H}$ defined by

$$\tau_m(F)=m!\,\mathrm{Tr}_{\wedge^m\mathcal{H}}(\wedge^m[1-K(g\phi)]^{-1}\cdot F)$$

and the expectation

$$\langle\cdot\rangle^P_{\Lambda,g,Q}=\frac{\int\cdot\tilde{\rho}_\Lambda(g\phi)\exp[Q_\Lambda(g\phi)]d\mu^P_{\tilde{m}_b,\Lambda}}{\int\tilde{\rho}_\Lambda(g\phi)\exp[Q_\Lambda(g\phi)]d\mu^P_{\tilde{m}_b,\Lambda}}.$$

The proof of (24) is based on the cluster expansion as in Sec. V of Ref. 3. However, several new features of the present situation deserve to be mentioned: The additional cutoff function g causes no change in the proof nor does the additional boson self-interaction (23) since the coefficients $|c_j(\Lambda)|\leq 1$. The fact that the fermion mass $\tilde{m}_f$ has become a 2×2 matrix is also not a serious complication since the denominator $p^2+|\tilde{m}_f|^2$ in (15) has the same form as before. As for the use of periodic B.C., we can still obtain the needed L^p_{loc} estimates on derivatives of the covariance as in Sec. VI of Ref. 3 by using the method of images. In particular we use the following formula: If $C^{P,\partial\Lambda}_{D,\gamma}(x,y)$ denotes the Green's function for $-\Delta+m^2$ with periodic B.C. on $\partial\Lambda$ and Dirichlet B.C. on a set $\gamma\subset\Lambda$, then

$$C^{P,\partial\Lambda}_{D,\gamma}(x,y)=\sum_n C_{D,\Gamma}(x,y_n),$$

where the sum takes place over all translations y_n of y in the sides of Λ and $C_{D,\Gamma}$ is the Green's function with Dirichlet B.C. on $\Gamma=\cup\gamma_n$, the union of all translations of γ. We omit further details.

2. We next take $\Lambda\to\mathbb{R}^2$ with g fixed. From the clustering (24) and the decay (21) it follows by a standard proof[3] that (for suitable arguments)

$$\langle\cdot\rangle^P_{\Lambda,g,Q}\to\langle\cdot\rangle_g=\frac{\int\cdot\tilde{\rho}(g\phi)d\mu_{\tilde{m}_b}}{\int\tilde{\rho}(g\phi)d\mu_{\tilde{m}_b}}, \tag{25}$$

the expectation for the theory with free B.C., spatial cutoff g, and no boson self-interaction terms. In particular we have convergence of the Schwinger functions for the (λ,m_b,m_f,μ) theory: $S^P_\Lambda\to S_g$, where S_g is the Schwinger function with free B.C. and spatial cutoff g.

3. Obviously the expectation $\langle\cdot\rangle_g$ in (25) also satisfies the cluster property (24) and so we may take the limit $g\to 1$ with the convergence of S_g and $\langle\cdot\rangle_g$ to infinite volume quantities S and $\langle\cdot\rangle$. This establishes part (a) of Theorem 1.

4. Since we are at liberty to choose g spherically symmetric, it is clear that the infinite volume theory is Euclidean covariant; and the cluster property of the form (24) goes over to the infinite volume. In this way we verify the Osterwalder—Schrader and Wightman

axioms including a positive mass gap α (where the mass gap can be made arbitrarily large by taking μ large).

Note, incidentally, that we have shown that the limits of the finite volume theories with both free and periodic B.C. exist and are the same.

ACKNOWLEDGMENT

I wish to thank Alan Cooper for useful discussions concerning the region of validity of the cluster expansion.

*A. Sloan Foundation Fellow. Research partially supported by the National Research Council of Canada.

[1]E. Seiler, Commun. Math. Phys. **42**, 163 (1975).

[2]J. Magnen and R. Sénéor, "The Wightman Axioms for the Weakly Coupled Yukawa Model in Two Dimensions," Commun. Math. Phys., to appear.

[3]A. Cooper and L. Rosen, "The Weakly Coupled Yukawa$_2$ Field Theory: Cluster Expansion and Wightman Axioms," Trans. Amer. Math. Soc., to appear.

[4]J. Glimm, A. Jaffe, and T. Spencer, in *Constructive Quantum Field Theory*, edited by G. Velo and A. Wightman (Springer-Verlag, Berlin, 1973), pp. 199–242.

[5]T. Spencer, Commun. Math. Phys. **39**, 63 (1974).

[6]Γ is usually taken to be the identity (scalar Y_2) or the matrix γ_2 (pseudoscalar Y_2), but the cluster expansion and the results of this note are valid for general $\Gamma = a + b\gamma_2$.

[7]K. Osterwalder and R. Schrader, Commun. Math. Phys. **31**, 83 (1973); **42**, 281 (1975).

[8]J. Fröhlich, "Phase Transitions in Two Dimensional Quantum Field Models," ZiF, Univ. of Bielefeld preprint.

[9]J. Fröhlich and B. Simon, "Pure States for General $P(\phi)_2$ Theories: Construction, Regularity and Variational Equality," Princeton Univ. preprint.

[10]F. Guerra, L. Rosen, and B. Simon, "Boundary Conditions for the $P(\phi)_2$ Euclidean Field Theory," Ann. Inst. H. Poincaré, to appear.

[11]We are essentially using the notation of Ref. 3: $\mathcal{H} = \mathcal{H}_{1/2} \oplus \mathcal{H}_{1/2}$, where $\mathcal{H}_{1/2}$ is the Sobolev space $\mathcal{H}_{1/2} = L^2((p^2 + 1)^{1/2} d^2p)$; $\wedge$ denotes antisymmetric tensor product; $\mathrm{Tr}_{\wedge_m}$ is the trace on $\wedge^m \mathcal{H}$. For technical reasons we also must restrict the operators F_A and F_B in (24) to have the special product form dictated by Theorem VIII. 14′ of Ref. 3; i.e., $F = F_1 \wedge \cdots \wedge F_m$, where each $F_j \in \mathcal{C}_{1,q}$ for any $q < \infty$, and each F_j is localized in the sense that $F_j \chi_j$, where χ_j is the characteristic function of some finite region. It is easy to see[3] that the Schwinger functions can be expressed in the form $\langle \tau_m(F) \rangle$ in terms of such F's and that the proof of the convergence of the infinite volume limit involves only such F's.

The bundle boundary in some special cases*

Russell A. Johnson

University of Wisconsin-Parkside, Kenosha, Wisconsin 53140
and University of Southern California, University Park, Los Angeles, California 90007
(Received 2 December 1976)

We examine a class of two-dimensional Lorentz manifolds which are "singular" in a certain sense. It is shown that, for such a manifold (M,g), the bundle boundary is a single point whose only neighborhood is all of $\bar{M}$ [the bundle completion of M; see B. G. Schmidt, Gen. Rel. Grav. 1, 269–80 (1971)]. The four-dimensional Schwarzschild and Friedmann–Robertson–Walker solutions are then investigated. We show that the bundle completions of these spaces are not Hausdorff.

1. PRELIMINARIES

We first review some standard definitions and results concerning the bundle boundary.

Let M be an n-dimensional C^∞ manifold with C^∞ Lorentz metric g. Let $L(M)$ be the bundle of Lorentz-orthonormal frames on M, $\pi: L(M)\to M$ the projection. The Lorentz group $O(1,n-1)$ acts on $L(M)$ on the right; we use $u\to u\cdot g$ or R_g to denote the right translation by $g\in O(1,n-1)$. The metric g induces a torsion-free connection on $L(M)$. Define vector fields $(B_i)_{i=1}^n$ on $L(M)$ by (i) $\pi_*(B_i(u))=e_i$, and (ii) $B_i(u)$ is horizontal, if $u=(e_1,\ldots,e_i,\ldots,e_n)$; the B_i are the *standard horizontal vector fields*. Choose a basis $(\tilde{E}_j)_{j=1}^l$ $[l=(n-1)(n-2)\cdot\frac{1}{2}]$ of the Lie algebra $o(1,n-1)$, and let $(E_j)_{j=1}^l$ be the corresponding vertical vector fields on $L(M)$ [thus $E_j(u)=(d/dt)(R_{a(t)}\cdot u)_{t=0}$, where $a(t)$ is the one-parameter subgroup generated by $\tilde{E}_j$].

1.1. *Definition* [1]: Let $u\in L(M)$, and let $X=\sum_{i=1}^n X_iB_i(u)+\sum_{j=1}^l X_{j+n}E_j(u)$, $Y=\sum_{i=1}^n Y_iB_i(u)+\sum_{j=1}^l Y_{j+n}E_j(u)$. Define $\gamma(X,Y)=\sum_{i=1}^{n+l}X_iY_i$.

It is easily seen that γ is a C^∞ Riemannian metric on $L(M)$. It induces a topological metric d on $L(M)$ such that the metric topology and the original topology coincide.

1.2. *Lemma*: Let $K\subset O(1,n-1)$ be compact. There are constants $\alpha>0$, $\beta>0$ such that $\alpha d(u,v)\le d(u\cdot g, v\cdot g)\le\beta d(u,v)$ $[g\in K,\ u,v\in L(M)]$.

Proof: Fix $u\in L(M)$. Let S be the γ-unit sphere in $T_u(L(M))$. The map $\sigma:(g,X)\to\gamma(R_{g*}X,R_{g*}X):K\times S\to IR$ is jointly continuous; hence, if

$$\alpha=\min_{g\in K,X\in S}\sigma(g,X),\quad \beta=\max_{g\in K,X\in S}\sigma(g,X),$$

then $(*)$ $0<\alpha\le\gamma(R_{g*}X,R_{g*}X)\le\beta<\infty (g\in K)$. The definition of γ and the transformation properties of the B_i and E_j may be used to show that $(*)$ holds for *every* X in the tangent bundle to $L(M)$ which satisfies $\gamma(X,X)=1$ (see Ref. 2). The lemma follows.

Let $\overline{L(M)}$ be the completion of $L(M)$ with respect to the metric d, and let $\dot{L}(M)=\overline{L(M)}\setminus L(M)$. Using 1.2, it may be shown that each R_g extends to a homeomorphism of $\overline{L(M)}$, and that $u\cdot(g\cdot h)=(u\cdot g)\cdot h$ $[u\in L(M),\ g,h\in O(1,n-1)]$ (see Ref. 2). Observe that Lemma 1.2 applies if u and/or v is in $\dot{L}(M)$.

1.3. *Proposition*: The object $(\overline{L(M)},\ O(1,n-1))$ is a transformation group; that is, the map $\overline{L(M)}\times O(1,n-1)\to\overline{L(M)}:(u,g)\to u\cdot g$ is jointly continuous.

Proof: Let $u_n\to u$ in $\overline{L(M)}$, $g_n\to g$ in $O(1,n-1)$, and let $\epsilon>0$ be given. We assume $u\in\dot{L}(M)$, since otherwise the conclusion follows from the definition of the topology on $L(M)$. By the triangle inequality, $d(u_ng_n,ug)\le d(u_ng_n,ug_n)+d(ug_n,ug)$. Let K be a compact neighborhood of g. Choose N_1 such that $n\ge N_1\Rightarrow g_n\in K$. By 1.2, there exists $N_2\ge N_1$ such that $n\ge N_2\Rightarrow d(u_ng_n,ug_n)<\epsilon/2$.

We complete the proof by showing that, for n large, $d(ug_n,ug)<\epsilon/2$. Choose a sequence (u_r) in $L(M)$ such that $u_r\to u$. Then $d(ug_n,ug)\le d(ug_n,u_rg_n)+d(u_rg_n,u_rg)+d(u_rg,ug)$. If $n\ge N_2$, then 1.2 implies that the first and third terms are $\le\text{const}\cdot d(u,u_r)$, where the constant does not depend on n. Choose r_0 so that $\text{const}\cdot d(u,u_{r_0})<\epsilon/6$, then choose $N_3\ge N_2$ so that $n\ge N_3\Rightarrow d(u_{r_0}g_n,u_{r_0}g)<\epsilon/6$. If $n\ge N_3$, then $d(ug_n,ug)<\epsilon/2$.

1.4. *Definition*: Let $\bar{M}=\overline{L(M)}/O(1,n-1)$ with the quotient topology; $\bar{M}$ is the Schmidt or bundle or b-completion of M. The b-boundary of M is $\dot{M}=\bar{M}\setminus M$. Let $\pi:\overline{L(M)}\to\bar{M}$ be the projection.

The next result states that any point of $\dot{M}$ is determined by the horizontal lift of some curve in M. It is easily seen that any Cauchy sequence (u_n) in $L(M)$ such that $\lim_{n\to\infty}u_n\in\dot{L}(M)$ may be assumed to lie on some C^∞ curve in $L(M)$ of finite bundle length (i.e., of finite length as defined by γ). Call a C^∞ curve $\eta:[s_1,s_2)\to L(M)$ *b-incomplete* if its bundle length is finite and $\lim_{s\to s_2}\eta(s)\in\dot{L}(M)$.

1.5. *Proposition*: Let $\eta:[s_1,s_2)\to L(M)$ be a b-incomplete curve with $\lim_{s\to s_2}\eta(s)=u\in\dot{L}(M)$. Let $\bar{\eta}$ be the horizontal lift of $\pi\circ\eta$. Then there exists $g_0\in O(1,n-1)$ such that $\lim_{s\to s_2}\bar{\eta}(s)=u\cdot g_0$.

Proof: The argument given on pp. 278–80 of Ref. 1 shows that $\bar{\eta}$ has finite bundle length and that $\eta(s)=\bar{\eta}(s)\cdot g(s)$, where $\{g(s):s_1\le s<s_2\}$ has compact closure in $O(1,n-1)$ (the argument is given for dimension 4, but applies also in dimension n). Choose a sequence $s_n\to s_2$ such that $g(s_n)$ converges to some g_0^{-1}. Then (1.3) $u=[\lim_{s\to s_2}\bar{\eta})]\cdot g_0^{-1}$, proving 1.5.

2. MORE PRELIMINARIES

We restrict attention to the following class of Lorentz

2-manifolds $(M, g): M$ is a cylinder with coordinates (r, ϕ), where ϕ is taken mod 2π and $0 < r < R < \infty$; and g is given by $ds^2 = -b^2(r)\,dr^2 + a^2(r)\,d\phi^2$, where a, b are positive C^∞ functions of r on $0 < r < R$. Particular cases are the "Schwarzschild" metric $(a(r) = r, b(r) = (2/r - 1)^{-1/2}, R = 1)$ and the FRW metrics $(b(r) \equiv 1;\ a(r) \to 0,\ a'(r) \to \infty$ as $r \to 0)$. Note that, in all of these examples, the metric is "singular" at $r = 0$. Our interest is in precisely this situation. We will show that, under assumptions which are satisfied by Schwarzschild and many of the FRW's, the b-boundary $\dot{M}$ of M is (essentially) a single point, the only neighborhood of which is M itself.

Fix a Lorentz manifold (M, g) as above. Such an M is orientable and time-orientable. Hence, $L(M)$ has four connected components, corresponding to the four components of $O(1,1)$. Any component L of $L(M)$ is acted on by the subgroup G of $O(1,1)$ consisting of isochronous matrices of determinant $+1$ [thus

$$G = \left\{ \begin{pmatrix} \cosh\lambda & \sinh\lambda \\ \sinh\lambda & \cosh\lambda \end{pmatrix} : \lambda \in \mathbf{R} \right\} \Bigg].$$

Restrict γ, d to L, and let $\bar{L}$ be the b-completion of L. It is easily seen that 1.2, 1.3, and 1.5 hold if $L(M)$, $\overline{L(M)}$, and $O(1, n-1)$ are replaced by L, $\bar{L}$, G, and that $\overline{L(M)}/O(1,1) = \bar{L}/G$. [Regarding 1.5, observe that any curve in $L(M)$ which intersects L is contained in L.]

2.1: From what has just been said, we may restrict attention to one component L of $L(M)$. We let L be that component containing all the frames $(e_1(m), e_2(m))$ $(m \in M)$, where

$$e_1(m) = \frac{1}{b(r)} \frac{\partial}{\partial r}(m), \quad e_2(m) = \frac{1}{a(r)} \frac{\partial}{\partial \phi}(m).$$

Clearly e_1 and e_2 are C^∞ vector fields on M. Letting $\pi : \bar{L} \to \bar{L}/G \approx \bar{M}$ be the projection, we observe that each element of $\pi^{-1}(m)$ may be written $\{\cosh\lambda\, e_1(m) + \sinh\lambda\, e_2(m),\ \sinh\lambda\, e_1(m) + \cosh\lambda\, e_2(m)\}$ for some $\lambda \in \mathbf{R}$ $(m \in M)$. Give L the coordinates (r, ϕ, λ). Note that $(r, \phi, \lambda) \cdot g = (r, \phi, \lambda + \lambda_1)$, where

$$g = \begin{pmatrix} \cosh\lambda_1 & \sinh\lambda_1 \\ \sinh\lambda_1 & \cosh\lambda_1 \end{pmatrix}.$$

Now consider the curves $\eta_1(s) = (s, \phi_0)$ $(0 < s < R)$ and $\eta_2(s) = (r_0, \phi_0 + s)$ $(-\infty < s < \infty)$ in M, where r_0, ϕ_0 are fixed. Their horizontal lifts through $(r_0, \phi_0, \lambda_0) \in L$ are

$$\bar{\eta}_1(s) = (s, \phi_0, \lambda_0), \tag{1}$$

$$\bar{\eta}_2(s) = (r_0, \phi_0 + s, \lambda_0 - [a'(r_0)/b(r_0)]s). \tag{2}$$

From (1) and (2), we can compute the distribution $u \to H_u$ of horizontal subspaces correponding to the torsion-free connection defined by g. We obtain

$$H_u \text{ is spanned by } \frac{\partial}{\partial r} \text{ and } \frac{\partial}{\partial \phi} - \frac{a'(r_0)}{b(r_0)} \frac{\partial}{\partial \lambda} \tag{3}$$

if $u = (r_0, \phi_0, \lambda_0)$. If we think of e_1 and e_2 as vector fields on L, the spanning vector are

$$e_1 \text{ and } e_2 - \frac{a'}{ab} e_3, \quad \text{where } e_3 = \frac{\partial}{\partial \lambda}. \tag{4}$$

We may define the bundle metric γ using e_3 as the vertical vector field; then

$$\gamma(v, v) = (\cosh\lambda_0 v_1 - \sinh\lambda_0 v_2)^2 + (\sinh\lambda_0 v_1 + \cosh\lambda_0 v_2)^2 + (v_3 + a' v_2/ab)^2, \tag{5}$$

where $v = \sum_{i=1}^{3} v_i e_i$ is a tangent vector at (r_0, ϕ_0, λ_0).

2.2: We now state the (five) assumptions referred to at the beginning of the section:

(A1) $\dfrac{db}{dr} \geq 0$ on $(0, R)$;

(A2) $\lim_{r \to 0^+} \dfrac{b(r)}{a'(r)} = 0$, and $\dfrac{d}{dr}\left(\dfrac{b(r)}{a'(r)}\right) \geq 0$ on $(0, R)$;

(A3) $\dfrac{a(r)}{a'(r)} \geq C \cdot r$ on $(0, R)$, where $C > 0$ is a constant;

(A4) For each $r_0 \in (0, R)$, there are positive C^∞ functions $\tilde{a}(r)$, $\tilde{b}(r)$ defined on $-\infty < r < \infty$ such that $\tilde{a}|_{[r_0, R)} = a$, $\tilde{b}|_{[r_0, R)} = b$, and $\tilde{a}(r) \equiv 1 \equiv \tilde{b}(r)$ except on a compact set;

(A5) For some sequence $r_n \to 0$, $\sum_{n=1}^{\infty} \int_0^{r_n} b(r)\,dr < \infty$ and $\sum_{n=1}^{\infty} \int_0^{r_n} \dfrac{b(r)}{a(r)}\,dr = \infty$.

2.3. *Remarks*: (a) Condition (A2) says that there is a singularity at $r = 0$.

(b) Using condition (A2) (and perhaps decreasing R), we may assume $a' > 0$ on $(0, R)$.

(c) We will only use condition (A5) in 3.9.

(d) Condition (A4) is meant to ensure that the only "interesting" elements of $\bar{L} \setminus L$ occur at $r = 0$. Consider the metric $ds^2 = \tilde{b}^2\,dr^2 + \tilde{a}^2\,d\phi^2$ on the cylinder $\tilde{M}: -\infty < r < \infty$, ϕ taken mod 2π. Define $\tilde{L}$ over $\tilde{M}$ just as L was defined over M. Let $\tilde{\gamma}$ be the corresponding bundle metric on $\tilde{L}$, and let γ_f be the metric induced on $\tilde{L}$ by the flat connection. Using (3), one can show that $\tilde{\gamma}$ and γ_f are equivalent, i.e., there are constants $c_1, c_2 > 0$ such that $c_1 \tilde{\gamma} \leq \gamma_f \leq c_2 \tilde{\gamma}$. This implies that $(\tilde{L}, \tilde{\gamma})$ is complete. Now if $\eta(s) = (r(s), \phi(s), \lambda(s))$ is a b-incomplete curve in L such that $r(s) \geq r_0 > 0$ on η, then clearly $r(s) \to R$, $\phi(s) \to \phi_0$, $\lambda(s) \to \lambda_0$. Thus the points of $\bar{L}$ defined by such curves are uninteresting. We therefore make the following conventions.

2.4. *Definition*: Let $\dot{L} = \{p \in \bar{L} : p = \lim_{s \to s_2} \eta(s)$, where $\eta : [s_1, s_2) \to L$ is a b-incomplete curve such that r is not bounded away from zero along $\eta\}$.

With this notation, $\dot{L}$ is a proper subset of $\bar{L} \setminus L$. It is the portion of $\bar{L} \setminus L$ which is "at $r = 0$."

2.5. *Definition*: The *essential boundary* of M is $\pi(\dot{L})$.

3. RESULTS

We maintain the notation of Sec. 2.

3.1: There is a natural class of incomplete curves in L, namely the radial curves $\eta_\phi(s) = (s, \phi, 0)$, where ϕ is fixed. These curves are the horizontal lifts of their projections to M (see (1)). Let $p(\phi) = \lim_{s \to 0} \eta_\phi(s)$, and let $P = \{p(\phi) : 0 \leq \phi < 2\pi\}$.

We outline what follows. We first show that $P = \{\bar{p}\}$, a point (3.6). It is then shown that $\dot{L} = \{\bar{p}\}$ (3.7); see

2.3d and 2.4. Finally, it is demonstrated that $\bar{M}$ is the only open set containing $\pi(\bar{p})$ (3.8).

3.2. *Lemma*: The map $S^1 \to P: \phi \to p(\phi)$ is continuous (hence P is compact).

Proof: Let $\phi_n \to \phi$. Fix $\epsilon > 0$, and choose r_0 so that $r_0 b(r_0) < \epsilon/3$ if $r \leq r_0$ (A1). Draw a curve between $p(\phi_n)$ and $p(\phi)$ in three pieces: (i) $\eta_1(s) = (s, \phi_n, 0)$, $0 < s \leq r_0$; (ii) $\eta_2(s) = (r_0, \phi_n + s(\phi - \phi_n), 0)$, $0 \leq s \leq 1$; (iii) $\eta_3(s) = (r_0 - s, \phi_0, 0)$, $0 \leq s < r_0$. Using (5), one finds that the lengths of η_1 and η_3 are given by $\int_0^{r_0} b(r)\,dr$, which by (A1) is $\leq r_0 b(r_0) < \epsilon/3$. Choose N so that $n \geq N = |\phi_n - \phi| < \epsilon/3 \cdot [1 + a'(r_0)/a(r_0)b(r_0)]^{-1/2}$. By (5), η_2 has length $< \epsilon/3$. We conclude that $d(p(\phi_n), p(\phi)) < \epsilon$; hence $p(\phi_n) \to p(\phi)$.

The method of proof used in 3.2 yields

3.3. *Lemma*: If $r_n \to 0$ and $\phi_n \to \phi$, then $(r_n, \phi_n, 0) \to p(\phi)$.

3.4. *Lemma*: Each fiber $\pi^{-1}\pi(\bar{p})$ ($\bar{p} \in P$) reduces to a point. That is, $\bar{p} \cdot g = \bar{p}$ for all $\bar{p} \in P$, $g \in G$.

Proof: Let

$$g = \begin{pmatrix} \cosh\lambda_0 & \sinh\lambda_0 \\ \sinh\lambda_0 & \cosh\lambda_0 \end{pmatrix},$$

and let $\bar{p} \equiv p(\bar{\phi})$. We will find a sequence $p_n \to \bar{p}$ such that $p_n \cdot g \to \bar{p}$; if this is done, 1.3 implies that $\bar{p} \cdot g = \bar{p}$.

For each $r > 0$, let $(\Delta\phi)_r = \lambda_0 b(r)/a'(r)$. By (A2), $(\Delta\phi)_r \to 0$ as $r \to 0$. Assume $\lambda_0 \geq 0$ (the proof is similar if $\lambda_0 < 0$). Let $r_n \to 0$, let $\Delta_n \equiv (\Delta\phi)_{r_n}$, and let $p_n = (r_n, \bar{\phi} - \Delta_n, 0)$. By 3.3, $p_n \to \bar{p}$. Note $p_n \cdot g = (r_n, \bar{\phi} - \Delta_n, \lambda_0)$. Consider the curve *in M* $\eta(s) = (r_n, \bar{\phi} - \Delta_n + s)$, $0 \leq s \leq \Delta_n$. The horizontal lift $\bar{\eta}$ of this curve through $p_n \cdot g$ satisfies $\bar{\eta}(\Delta_n) = (r_n, \bar{\phi}, 0)$ [see (2)]. Also, the length of $\bar{\eta}$ is $\leq (\sqrt{2}\cosh\lambda_0)a(r_n)\Delta_n$, which $\to 0$ as $n \to \infty$. Since $(r_n, \bar{\phi}, 0) \to \bar{p}$ (3.3), we conclude that $d(p_n \cdot g, \bar{p}) \to 0$.

We are going to prove that $|P| = 1$. Consider the two systems of equations

$$\dot{r} = 1, \quad \frac{d\varphi}{dr} = \frac{b}{a}, \quad \frac{d\lambda}{dr} = \frac{-a'}{a}; \quad \varphi(r_0) = \varphi_0, \quad \lambda(r_0) = \lambda_0, \tag{6}$$

and

$$\dot{r} = 1, \quad \frac{d\varphi}{dr} = \frac{-b}{a}, \quad \frac{d\lambda}{dr} = \frac{a'}{a}; \quad \varphi(r_0) = \varphi_0, \quad \lambda(r_0) = \lambda_0. \tag{7}$$

Here $a' \equiv da/dr$; solutions are parametrized by r. If $\eta(r) = (r, \phi(r), \lambda(r))$ is a solution to (6) [(7)], then $\pi \circ \eta$ is lightlike with tangent vector field parallel to $e_1 + e_2(e_1 - e_2)$. The length of the segment of η traced out as r ranges from r_0 to 0 may be computed from (5):

$$\begin{aligned} l(r_0) &= [\sqrt{2}/a(r_0)]\exp(-\lambda_0)\int_0^{r_0} a(r)b(r)\,dr \\ &\leq \sqrt{2}\exp(-\lambda_0)\int_0^{r_0} b(r)\,dr \end{aligned} \tag{8}$$

if η is a solution to (6) (see 2.3b for the $\leq$);

$$l(r_0) \leq \sqrt{2}\exp(\lambda_0)\int_0^{r_0} b(r)\,dr \tag{9}$$

if η is a solution to (7). The change in angle along η is

$$\Delta\varphi \equiv \left|\varphi(r_0) - \lim_{r\to 0}\varphi(r)\right| = \int_0^{r_0}[b(r)/a(r)]\,dr, \tag{10}$$

which may be $+\infty$.

3.5. *Lemma*: If η is a solution to (6) or (7), then $\lim_{r\to 0}\eta(r) \in P$.

Proof: We consider only the case when η solves (6). For fixed r, use (5) to compute the length of the path going radially inward from $\eta(r)$; then apply 3.4 to see that

$$d(\eta(r), p(\varphi(r))) \leq \left[\sqrt{2}\int_0^r b(\bar{r})\,d\bar{r}\right] \cdot \cosh\lambda(r).$$

But

$$\int_0^r b(\bar{r})\,d\bar{r} \leq [\text{by (A1)}]\; rb(r) \leq [\text{by (A3)}]\; \frac{1}{C}\frac{b(r)a(r)}{a'(r)},$$

and $\lambda(r) = \lambda_0 + \ln[a(r_0)/a(r)]$. Using these facts and 2.3b, it is easily seen that

$$d(\eta(r), p(\varphi(r))) \leq (\sqrt{2}/C)a(r_0)[b(r)/a'(r)]\cosh\lambda_0$$

if $r \leq r_0$. Let $r \to 0$; the lemma follows from compactness of P.

3.6. *Theorem*: P consists of a single point.

Proof: We show that if $\epsilon > 0$ and $\bar{\varphi} \in (0, 2\pi)$, then $d(p(0), p(\bar{\varphi})) < \epsilon$ (which, of course, implies 3.6).

Suppose first that $\int_0^r [b(\bar{r})/a(r)]\,d\bar{r} = \infty$ for every $r > 0$. Choose r_0 so that (i) the solution η to (6) with initial conditions $\phi(r_0) = 0 = \lambda(r_0)$ satisfies $l(r_0) < \epsilon/3$ [see (8)]; (ii) if $r \leq r_0$, then $d(\eta(r), p(\phi(r))) < \epsilon/3$ (see 3.8). By (6) and (10), there exists $\bar{r} < r_0$ such that $\phi(\bar{r}) = \bar{\phi}$; clearly, one has $d(p(0), p(\bar{\phi})) < 3 \cdot \epsilon/3 = \epsilon$.

If $\int_0^r [b(\bar{r})/a(\bar{r})]\,d\bar{r} < \infty$ for some (hence all) $r > 0$, we use (A5). Choose a sequence $r_n \to 0$ such that

$$\sum_{n=1}^{\infty}\sqrt{2}\int_0^{r_n} b(r)\,dr < \frac{\epsilon}{2}, \quad \sum_{n=1}^{\infty}\int_0^{r_n}\frac{b(r)}{a(r)}\,dr = \infty.$$

Let $(\Delta\phi)_n = \int_0^{r_n}[b(r)/a(r)]\,dr$. Perhaps decreasing some of the r_n, we may find $N \geq 1$ such that $\sum_{n=1}^N (\Delta\phi)_n = \phi/2$. Now let $p_1 = p(0)$, $p_i = p(\sum_{j=1}^{i-1} 2(\Delta\phi)_j)$ $(2 \leq i \leq N+1)$, $q_1 = (r_1, (\Delta\phi)_1, 0)$, $q_i = (r_1, (\Delta\phi)_1 + \sum_{j=1}^{i-1} 2(\Delta\phi)_j, 0)$; p_i and q_i may be joined by the solution η_i of (9) satisfying $\eta_i(r_i) = q_i$, while q_i and p_{i+1} may be joined by the solution σ_i of (10) satisfying $\sigma_i(r_i) = q_i$. By Eq. (A.1) [or (A.2)], the length of each $\eta_i(\sigma_i)$ is bounded by $\sqrt{2}\int_0^{r_i} b(\bar{r})\,d\bar{r}$. Since $p(\bar{\phi}) = p_N$, we have

$$\begin{aligned} d(p(0), p(\bar{\phi})) &\leq \sum_{i=1}^N d(p_i, q_i) + \sum_{i=1}^N d(q_i, p_{i+1}) \\ &\leq 2\sum_{i=1}^N \sqrt{2}\int_0^{r_i} b(\bar{r})\,d\bar{r} < \epsilon. \end{aligned}$$

Let $P = \{\bar{p}\}$.

3.7. *Theorem*: $\dot{L} = \{\bar{p}\}$ (see 2.4 for the definition of $\dot{L}$). Thus the essential boundary of M (2.5) is a point.

Proof: Let $\tilde{\eta}(s) = (\tilde{r}(s), \tilde{\phi}(s), \tilde{\lambda}(s))$ $(0 \leq s < \bar{s}$, $s = $ arc length) be any b-incomplete curve in L for which there is a sequence $s_n \to \bar{s}$ with $r(s_n) \to 0$. We can assume that $\lim_{s\to\bar{s}}|\lambda(s)| = \infty$ (see 1.3). Assume $\lambda \to +\infty$. Let $r_n = r(s_n)$; for fixed n, consider the solution $\eta(r)$ $(0 < r \leq r_n)$ of (6) satisfying $\phi(r_n) = \tilde{\phi}(r_n)$, $\lambda(\Gamma_n) = \tilde{\lambda}(\Gamma_n)$. We may assume $\lambda(r_n) \geq 0$. From (8), then, the length of η is $\leq 2\int_0^{r_n} b(r)\,dr \to 0$. But $\lim_{r\to 0}\eta(r) = \bar{p}$(3.5, 3.6), which means $d(\tilde{\eta}(s_n), \bar{p}) \to 0$.

3.8. *Theorem*: The only open set in $\bar{M}$ containing $\pi(\bar{p})$ is $\bar{M}$ itself.

Thus $\pi(\bar{p})$ can be Hausdorff separated from no other point in $\bar{M}$.

Proof: Let $(r_0, \phi_0) \in M$. It suffices to show that the constant sequence (m_n), all of whose terms are (r_0, ϕ_0), converges to $\pi(\bar{p})$. Consider the sequence in L whose nth term is (r_0, ϕ_0, n) $(n \geq 1)$. Let $\eta_n(r)$ be the solution of (6) satisfying $\phi(r_0) = \phi_0$, $\lambda(r_0) = n$. By (8), the length of η_n is $\leq 2e^{-n}\int_0^{r_0} b(r)\,dr \to 0$. Hence $d(u_n, \bar{p}) \to 0$, i.e., $u_n \to \bar{p}$. So $(r_0, \phi_0) = \pi(u_n) \to \pi(\bar{p})$.

4. FOUR-DIMENSIONAL SOLUTIONS

In this section, the Schwarzschild and closed FRW solutions are considered. It is shown that, in both cases the bundle completion is not Hausdorff.

4.1: We take the Schwarzschild metric in the form $ds^2 = -b^2(r)\,dr^2 + b^{-2}(r)\,dt^2 + r^2(d\theta^2 + \sin^2\theta\,d\phi^2)$, where $b(r) = (2/r - 1)^{-1/2}$. Put the restriction $0 < r < 1$ on the manifold M. Define vector fields

$$\tilde{e}_1 = \frac{1}{b}\frac{\partial}{\partial r}, \quad \tilde{e}_2 = b\frac{\partial}{\partial t}, \quad \tilde{e}_3 = \frac{1}{r}\frac{\partial}{\partial \theta}, \quad \tilde{e}_4 = \frac{1}{r\sin\theta}\frac{\partial}{\partial \phi}.$$

These vector fields form an orthonormal frame $u(m)$ at each point $m \in \tilde{M} = \{m \in M \mid \theta \neq 0, \ \theta \neq \pi \text{ at } m\}$. Let δ be the bundle metric [Riemannian, not topological] on $L(M)$. We may assume that the element

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

of the Lie algebra $o(1,3)$ is used to construct one of the six vertical vector fields on $L(M)$ which enter into the definition of δ (see 1.1).

4.2: Consider the submanifold M_0 of M defined by $\theta = \pi/2$, $t = t_0 = \text{const}$. This submanifold is a cylinder of the type considered in Secs. 2 and 3. The metric is given by $ds^2 = -b^2\,dr^2 + r^2\,d\phi^2$, which, it may be checked, satisfies (A1)–(A5) of 2.2. Let $L \subset L(M_0)$ be the space defined in Sec. 2, with γ the b-metric. Define a map $i: L \to L(\tilde{M}): (r, \phi, \lambda) \to u(m)\cdot A_\lambda$, where $m = (r, t_0, \pi/2, \phi)$, $u(m)$ is the frame $(\tilde{e}_1, \tilde{e}_2, \tilde{e}_3, \tilde{e}_4)(m)$, and $A_\lambda \in O(1,3)$ is the matrix

$$\begin{pmatrix} \cosh\lambda & 0 & 0 & \sinh\lambda \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh\lambda & 0 & 0 & \cosh\lambda \end{pmatrix}.$$

4.3. *Proposition*: The map i is an isometry of (L, γ) into $(L(\tilde{M}), \delta)$.

Proof: Let $e_1 = (1/b)\partial/\partial r$, $e_2 = (1/r)\partial/\partial\phi$, and $\partial/\partial\lambda$ be vector fields on L, and let E be the vertical vector field on $L(\tilde{M})$ defined by

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \in o(1,3).$$

Clearly $i_*(e_1) = \tilde{e}_1$, $i_*(e_2) = \tilde{e}_4$. A straightforward computation shows that $i_*(\partial/\partial\lambda) = E$. Now e_1, e_2, and $\partial/\partial\lambda$ are γ-orthonormal, while $\tilde{e}_1$, $\tilde{e}_4$, and E are δ-orthonormal. Hence i is an isometry.

4.4. *Proposition*: Let d be the topological metric on L defined by γ, $\tilde{d}$ the topological metric on $L(M)$ defined by δ. Then $\tilde{d}(i(x), i(y)) \leq d(x, y)$.

Proof: Both $\tilde{d}$ and d are defined by an infimum of a length functional over a class of curves. By 4.3, the class for $\tilde{d}$ contains (essentially) the class for d. The proposition follows.

In analogy with the two-dimensional situation, there is a natural subset P of $\overline{L(M)}\setminus L(M)$ defined by horizontally lifting curves $t = \text{const}$, $\theta = \text{const}$, $\phi = \text{const}$. Consider the subset P_0 of P defined by those curves lying in $i(L)$. By 3.7, 4.3, and 4.4, $P_0 = \{\bar{p}\}$, a single point. Write π for the projection $\overline{L(M)} \to \bar{M}$.

4.5. *Proposition*: Every neighborhood of $\pi(\bar{p})$ contains M_0.

Proof: Let (m_n) be the constant sequence all of whose terms are $(r_0, t_0, \pi/2, \phi_0) \in M_0$. By 4.3, 4.4, and the proof of 3.8, the sequence $\{i(r_0, \phi_0, n)\}$ $(n \geq 1)$ converges to $\bar{p}$. This proves 4.5.

4.6. *Theorem*: Let $m \in M$. Then there exists $q \in \dot{M}$ such that m is in every neighborhood of q.

Proof: Since SO(3) acts on M by isometries, we may assume that $\theta = \pi/2$ at q. Apply 4.5.

4.7: Consider the FRW metric in the form (see Refs. 1 or 3) $ds^2 = dt^2 + S^2(t)\,d\sigma^2$, where (i) $d\sigma^2$ is the metric on a standard 3-sphere, and (ii) $S = a(1 - \cos\tau)$, $t = a(\tau - \sin\tau)$. We may write $d\sigma^2 = d\psi^2 + \sin^2\psi(d\theta^2 + \sin^2\theta\,d\phi^2)$. Consider the submanifold $\psi = \theta = \pi/2$. The induced metric is $ds^2 = -dt^2 + S^2(t)\,d\phi^2$; conditions (A1)–(A5) of 2.2 are satisfied. The various steps of 4.1–4.6 may now be carried out to yield results analogous to 4.3, 4.4, 4.5, and 4.6 [in 4.6, SO(3) should be replaced by SO(4)].

4.8. *Remarks*: (a) The techniques used in Sec. 4 cannot be applied to the Reissner–Nordstrom solution, nor to the Kerr solution.

(b) Presumably non-Hausdorffness, particularly of the type considered in 4.6, is not a desirable property of the b-completion. One has a situation in which points of M, which intuitively are not particularly "close" to the singularity, nevertheless are in every neighborhood of some singular point. Perhaps a way around this problem would be to put something other than the quotient topology on $\bar{M}$. However, see the next remark.

(c) Refer to Sec. 3 for a moment: One can show that if

$$\int_0^R \left| \ln\left(\frac{a'}{ab}\right) \frac{d}{dr}\left(\frac{b}{a'}\right) \right| < \infty,$$

then ϕ has a limit along any curve $\eta(s)$ such that $\dot{r}(s) \leq 0$. It might be argued that, in this case, $\dot{M}$ should be a circle, not a point. Thus, even if the topological difficulties with $\dot{M}$ could be resolved, questions as to whether the b-boundary is a "good" boundary would remain.

(d) Two conjectures concerning the b-boundary: for Schwarzchild, $\dot{M}$ is a line; for FRW, $\dot{M}$ is a point.

ACKNOWLEDGMENT

The author wishes to thank Professor Robert Geroch for generous aid and valuable correspondence.

APPENDIX

As in Sec. 2, we consider a cylinder $0 < r < R < \infty$, φ taken mod 2π, with metric $ds^2 = -b^2(r)\,dr^2 + a^2(r)\,d\varphi^2$. Let η be a horizontal, b-incomplete curve in L, and let s = arclength along η. Suppose that η "goes directly to $r=0$" in the sense that $\dot{r}(s) \leq 0$ for all s (observe that a C^1, timelike curve satisfying $\inf_s r(s) = 0$ must also satisfy $\dot{r}(s) \leq 0$ for some choice of arclength parameter s). We will prove a theorem which gives a bound on the growth of $\lambda(s)$ along η. The theorem is included here because it is felt that bundle techniques will have further application in the study of singularities. For, the bundle length may be thought of as generalized affine parameter.[1] Hence, if a "singularity" is thought of as an equivalence class of curves in M which are "incomplete" in some sense, then the class of b-incomplete curves is a natural object to consider.

Let $\eta(s) = (r(s), \varphi(s), \lambda(s))$ be horizontal and b-incomplete, s = arclength. Write $\dot{r}$, $\dot{\varphi}$, $\dot{\lambda}$ for dr/ds, $d\varphi/ds$, $d\lambda/ds$ (as before, let $a' = da/dr$, $b' = db/dr$).

By (2),

$$\dot{\lambda} = [-a'(r(s))/b(r(s))]\dot{\varphi}. \tag{A1}$$

The square of the length of $\dot{\eta}(s)$ is [see (5)]

$$\|\dot{\eta}(s)\|^2 = [\cosh\lambda \cdot b(r)\dot{r} - \sinh\lambda \cdot a(r)\dot{\varphi}]^2. \tag{A2}$$

The proof of the following lemma is omitted.

A.1. *Lemma*: (a) $(a\dot{\varphi} - b\dot{r})^2 \sinh^2\lambda(s) \leq \|\dot{\eta}(s)\|^2$ if $\lambda \geq 0$,

(b) $b^2\dot{r}^2(\cosh\lambda(s) - \sinh\lambda(s))^2 \leq \|\dot{\eta}(s)\|^2$ if $\lambda \geq 0$,

(c) $(a\dot{\varphi} + b\dot{r})^2 \sinh^2\lambda(s) \leq \|\dot{\eta}(s)\|^2$ if $\lambda \leq 0$,

(d) $b^2\dot{r}^2(\cosh\lambda(s) + \sinh\lambda(s))^2 \leq \|\dot{\eta}(s)\|^2$ if $\lambda \leq 0$.

A.2. *Remark*: Note that 5.1 really states a fact about lengths of horizontal vectors; hence it holds for any horizontal curve η with any parametrization s.

A.3. *Theorem*: Let $\eta: [0, \bar{s}) \to L$ be horizontal and b-incomplete with $\dot{r}(s) \leq 0$ (s = arclength). Let $A(s) = b(r(s))a(r(s))/a'(r(s)) \cosh\lambda(s)$. Then $\lim_{s\to\bar{s}} A(s) = 0$.

Proof: Let $I(s) = \int_0^s f(u)\,du$, where

$$f(s) = \begin{matrix} (a\dot{\varphi} - b\dot{r}) \sinh\lambda(s) & \text{if } \lambda(s) \geq 0, \\ (a\dot{\varphi} + b\dot{r}) \sinh\lambda(s) & \text{if } \lambda(s) < 0. \end{matrix}$$

Then f is continuous; we will write (abusing notation) $f(s) = (a\dot{\varphi} \mp b\dot{r}) \sinh\lambda(s)$. By 5.1a and 5.1c, $\lim_{s\to\bar{s}} I(s)$ exists. From Eq. (A.1), $I(s) = \int_0^s \mp b\dot{r} \sinh\lambda(u)\,du - \int_0^s (ab/a')[\dot{\lambda} \sinh\lambda(u)]\,du$, $0 \leq s < \bar{s}$. Integrate by parts in the second integral to obtain $I(s) = \int_0^s b\dot{r}(\cosh\lambda \mp \sinh\lambda)\,du - ab/a' \cosh\lambda]_0^s + \int_0^s [ab'/a' - aba''/a'^2]\dot{r} \cosh\lambda\,du$. Let $J(s) = \int_0^s b\dot{r}(\cosh\lambda \mp \sinh\lambda)\,du$. By 5.1b and 5.1d, $\lim_{s\to\bar{s}} J(s)$ exists. We have

$$(*)\ I(s) - J(s) = -\{A(s) - A(0) + \int_0^s [a''/a' - b'/b]\dot{r}A(u)\,du\}.$$

Now $a''/a' - b'/b = (d/dr)\ln(a'/b) \leq 0$ (A2), and $\dot{r} \leq 0$. Hence the integrand in $(*)$ is (for $0 \leq s < \bar{s}$) positive. Suppose that $A(s) \geq \epsilon > 0$ for some ϵ. Then $I(s) - J(s) \leq -\{\epsilon - A(0) + \epsilon \ln[a'(r(s))/b(r(s))] + \epsilon \ln[b(r(0))/a'(r(0))]\}$.

By 2.3d and the assumption $\dot{r} \leq 0$, $r(s) \to 0$ as $s \to \bar{s}$. By (A2), the quantity in brackets tends to ∞. This contradicts the existence of limits for I and J as $s \to \bar{s}$. Returning to $(*)$ and using $A(s_n) \to 0$ together with the fact that the integrand is positive shows that $A(s) \to 0$.

*This research was supported by NSF Grant No. MCS76-07195.

[1] S.W. Hawking and G.F.R. Ellis, *The Large-Scale Structure of Space Time* (Cambridge, London, 1973).
[2] B.G. Schmidt, J. Gen. Rel. Gravitation 1, 269–80 (1971).
[3] R.P. Geroch, J. Math. Phys. 9, 450–65 (1968).

Phase-integral calculation of quantal matrix elements without the use of wavefunctions

Nanny Fröman and Per Olof Fröman

Institute of Theoretical Physics, University of Uppsala, Uppsala, Sweden
(Received 3 December 1976)

Simple phase-integral formulas for the calculation, without the use of wavefunctions, of quantal matrix elements of multiplicative and differential operators are given for the case of bound states in a single-well potential. The matrix elements are obtained to within the accuracy corresponding to any conveniently chosen order of the kind of phase-integral approximations used.

1. MATRIX ELEMENTS OF AN UNSPECIFIED OPERATOR A

Consider the Schrödinger equation

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

where, with obvious notations,

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

The function $V(z)$ may be the actual physical potential or an effective potential. Thus, if we are concerned with the radial Schrödinger equation, $V(z)$ is assumed to include also the centrifugal term.

When the differential equation (1) is solved approximately by means of the $(2N+1)$th order of the kind of phase-integral approximations considered in Refs. 1–3 and on pp. 126–31 in Ref. 4, there appears a function $q(z,E)$ defined by

$$q(z,E)=Q_{\rm mod}(z,E)\sum_{n=0}^{N} Y_{2n},\quad N=0,1,2,\cdots, \tag{3}$$

the explicit expressions for Y_{2n} up to Y_8 being given in Ref. 5 and up to Y_{20} in Ref. 6. The function $Q_{\rm mod}(z,E)$ is either identical to $Q(z,E)$ (unmodified case) or is another function (modified case) chosen such that the first-order phase-integral approximation becomes good at certain points where it would fail, if the function $Q(z,E)$ were used (cf. Ref. 3 and pp. 126–31 in Ref. 4).

Let us now assume that we have a single-well potential and are concerned with a bound-state problem, for the solution of which we can use the above-mentioned "symmetric" (cf. the terminology introduced by McHugh on p. 280 in his review article[7]) phase-integral approximations of the order $2N+1$. The complex z plane is assumed to be cut along the real axis between the two generalized classical turning points, i.e., the two real zeros of $Q^2_{\rm mod}(z,E)$. It has recently been shown by N. Fröman[8] that the expectation value of an analytic function $f(z)$ with respect to a bound state with an energy E_n of a given single-well potential $V(z)$ can be obtained by means of the approximate formula

$$\langle n|f(z)|n\rangle=\int_{\Gamma_n} f(z)\frac{dz}{q(z,E_n)}\bigg/\int_{\Gamma_n}\frac{dz}{q(z,E_n)}, \tag{4}$$

where Γ_n is a closed contour of integration encircling the two generalized classical turning points, i.e., the two real zeros of $Q^2_{\rm mod}(z,E_n)$, as well as the zeros of $q(z,E_n)$ associated with these two generalized classical turning points, but encircling no other zeros of $Q^2_{\rm mod}(z,E_n)$ and $q(z,E_n)$. The formula (4), in which the analytic function $f(z)$ is assumed to be regular on and within the contour of integration Γ_n, is very accurate for important classes of physical problems.

According to (2) the function $Q^2(z,E)$ is real on the real axis, and therefore it is no restriction to assume the solutions of the differential equation (1) to be real on the real axis. If $\psi(z,E_n)$ and $\psi(z,E_{n'})$ are two normalized eigenfunctions (real on the real axis) corresponding to the eigenvalues E_n and $E_{n'}$, respectively, the matrix element, with respect to those states, of an operator A is defined by

$$\langle n|A|n'\rangle=\int\psi(z,E_n)A\,\psi(z,E_{n'})\,dz, \tag{5}$$

where the integration is to be performed along the part of the real z axis which is appropriate to the range of the physical variable z.

For the sake of simplicity, and for didactic reasons, we shall now for a moment assume that the bound state with the energy E_n is the ground state. In this particular case the wavefunction $\psi(z,E_n)$ is different from zero everywhere on the appropriate part of the real axis, and therefore we can write (5) as follows:

$$\begin{aligned}\langle n|A|n'\rangle&=\int\psi(z,E_n)\,\frac{A\psi(z,E_{n'})}{\psi(z,E_n)}\,\psi(z,E_n)\,dz\\&=\left\langle n\left|\frac{A\psi(z,E_{n'})}{\psi(z,E_n)}\right|n\right\rangle .\end{aligned} \tag{6}$$

Using (4), we obtain from (6) the approximate formula

$$\langle n|A|n'\rangle=\int_{\Gamma_n}\frac{A\psi(z,E_{n'})}{\Psi(z,E_n)}\,\frac{dz}{q(z,E_n)}\bigg/\int_{\Gamma_n}\frac{dz}{q(z,E_n)}, \tag{7}$$

where $A\psi(z,E_{n'})$ is assumed to be a regular analytic function on and within the contour of integration Γ_n. This formula can be proved quite generally, i.e., without the use of the simplifying assumption that the state n is the ground state. We shall, however, not dwell on this question in the present paper, since a general proof will be given in a forthcoming publication by the present authors.

When $Q_{\rm mod}(z,E)$ is chosen to be positive on the upper lip of the cut between the two real zeros of $Q^2_{\rm mod}(z,E)$, and z lies sufficiently far away from the classically allowed region, the normalized eigenfunctions $\psi(z,E_n)$ and $\psi(z,E_{n'})$, which are real on the real axis, are given by the approximate formulas (cf. Eqs. (30a, c) in Ref. 9; cf. also Ref. 10)

$$\psi(z, E_n) = \frac{\exp(\frac{1}{4}\pi i)\, q^{-1/2}(z, E_n) \exp[-iw(z, E_n)]}{(\int_{\Gamma_n} dz/q(z, E_n))^{1/2}} \tag{8a}$$

and

$$\psi(z, E_{n'}) = \frac{\exp(\frac{1}{4}\pi i)\, q^{-1/2}(z, E_{n'}) \exp[-iw(z, E_{n'})]}{(\int_{\Gamma_{n'}} dz/q(z, E_{n'}))^{1/2}}, \tag{8b}$$

where the contours of integration Γ_n and $\Gamma_{n'}$ encircle, in the negative sense, the generalized classical turning points and the associated zeros of $q(z, E)$ and where

$$w(z, E) = \tfrac{1}{2} \int q(z, E)\, dz \tag{9}$$

with the path of integration in (9) defined either as a nonclosed loop around the left-hand generalized classical turning point (cf. in Ref. 1 Eq. (11a) and Fig. 1b), which corresponds to the situation that all wavefunctions are chosen to have the same sign in the classically forbidden region to the left of that turning point, or as a nonclosed loop around the right-hand generalized classical turning point, which corresponds to the situation that all wavefunctions are chosen to have the same sign in the classically forbidden region to the right of that turning point. The contours Γ_n and $\Gamma_{n'}$ in (7) and (8a), (8b) can be replaced by a single contour Γ which encircles (in the negative sense) the two real zeros of $Q^2_{\text{mod}}(z, E_n)$ and the associated zeros of $q(z, E_n)$, as well as the two real zeros of $Q^2_{\text{mod}}(z, E_{n'})$ and the associated zeros of $q(z, E_{n'})$, but does not encircle any other zero of the functions $Q^2_{\text{mod}}(z, E_n)$, $q(z, E_n)$, $Q^2_{\text{mod}}(z, E_{n'})$, and $q(z, E_{n'})$. The contour Γ shall thus enclose the generalized classical turning points, as well as the associated zeros of $q(z, E)$, corresponding to both of the bound states under consideration. We can also impose the condition that all points z on Γ shall lie far away from the classically allowed regions corresponding to the energies E_n and $E_{n'}$. We can then use (8a), (8b) for obtaining an approximate explicit expression for the quotient $A\psi(z, E_{n'})/\psi(z, E_n)$ in (7). The use of this approximate expression in the integrand of (7) is allowed only when $A\psi(z, E_{n'})/\psi(z, E_n)$ is not a too large and too rapidly varying function of z on the contour of integration Γ.

2. THE CASE WHEN $A = g(z)$

Letting A be a multiplicative operator $g(z)$, which is a regular analytic function on and within the contour of integration Γ, we obtain, in the way just described, the approximate formula

$$\langle n|g|n'\rangle = \int_\Gamma g(z) \frac{\exp\{i[w(z, E_n) - w(z, E_{n'})]\}}{q^{1/2}(z, E_n)\, q^{1/2}(z, E_{n'})}\, dz \Big/ \left(\int_\Gamma \frac{dz}{q(z, E_n)}\right)^{1/2} \left(\int_\Gamma \frac{dz}{q(z, E_{n'})}\right)^{1/2}, \tag{10a}$$

which is valid provided that the function $g(z)\exp\{i[w(z, E_n) - w(z, E_{n'})]\}$ is not too large and too rapidly varying on the contour Γ. The larger and more rapidly varying this function is on Γ, the less accurate formula (10a) is expected to be for a fixed order $2N+1$ of the phase-integral approximations.

Interchanging the indices n and n', and noting that the matrix element $\langle n|g|n'\rangle$ is symmetric in these indices, we obtain from (10a) the alternative approximate formula

$$\langle n|g|n'\rangle = \int_\Gamma g(z) \frac{\exp\{-i[w(z, E_n) - w(z, E_{n'})]\}}{q^{1/2}(z, E_n)\, q^{1/2}(z, E_{n'})}\, dz \Big/ \left(\int_\Gamma \frac{dz}{q(z, E_n)}\right)^{1/2} \left(\int_\Gamma \frac{dz}{q(z, E_{n'})}\right)^{1/2}, \tag{10b}$$

which is valid if the function $g(z)\exp\{-i[w(z, E_n) - w(z, E_{n'})]\}$ is not too large and too rapidly varying on the contour Γ. The larger and more rapidly varying this function is on Γ, the less accurate formula (10b) is expected to be for a fixed order $2N+1$ of the phase-integral approximations.

Formula (10a) is expected to be more accurate than (10b) when, on the contour of integration Γ, the function $g(z)\exp\{+i[w(z, E_n) - w(z, E_{n'})]\}$ is smaller than and less rapidly varying than the function $g(z)\exp\{-i[w(z, E_n) - w(z, E_{n'})]\}$, which is the case when $E_n > E_{n'}$ if $g(z)$ is slowly varying on Γ. The accuracy of both (10a) and (10b) is expected to increase when the quantum numbers n and n' increase, while $|n-n'|$ is kept fixed. On the other hand, the accuracy can be expected to decrease when $|n-n'|$ increases, which is, however, not a serious disadvantage, since the matrix elements then, in general, decrease and become small.

The importance of choosing the appropriate sign in the function $\exp\{\pm i[w(z, E_n) - w(z, E_{n'})]\}$, occurring in an integrand in the matrix element formula [(10a) or (10b)], has been clearly illustrated by calculations on the harmonic oscillator.[11] The integrand in question is not large and has few oscillations when the appropriate sign is chosen in the exponential, but becomes very large and has many oscillations when the other sign is chosen. For the same relative error of the integrand, i.e., for a given order $2N+1$ of the phase-integral approximations, one therefore obtains a much more accurate value of the integral when the appropriate sign is chosen in the exponential.

For operators such as $g(z) = z^p$ and $g(z) = \exp(ikz)$, which may become large and rapidly varying on Γ when p or $|k|$, respectively, becomes large, the accuracy is, in accordance with what has just been said, expected to decrease when p or $|k|$ increases. Test calculations on the harmonic oscillator[11] confirm this assertion.

When (10a) and (10b) are of comparable accuracy, it is reasonable to combine these two formulas to give the following alternative approximate formula:

$$\langle n|g|n'\rangle = \int_\Gamma g(z)\,\frac{\cos[w(z,E_n)-w(z,E_{n'})]}{q^{1/2}(z,E_n)\,q^{1/2}(z,E_{n'})}\,dz \Bigg/ \left(\int_\Gamma \frac{dz}{q(z,E_n)}\right)^{1/2} \left(\int_\Gamma \frac{dz}{q(z,E_{n'})}\right)^{1/2}. \tag{10c}$$

Test calculations on the harmonic oscillator and on a hydrogenlike ion[11] show that when n and n' are large, while $|n-n'|$ is small, (10c) may be considerably more accurate than (10a) or (10b). In more general cases (10c) is, however, not as accurate as the appropriate one of the formulas (10a) or (10b).

If we specialize formula (10c) to the case of the *first-order* JWKB approximation and assume n and n' to be large but only slightly different, formula (10c) can be transformed into the matrix element formulas obtained in the early days of quantum mechanics with the aid of arguments closely connected with the classical limit and the correspondence principle (cf., e.g., pp. 172–73 in Ref. 12, pp. 173–74 in the Geiger–Scheel edition of Ref. 13, p. 95 in the Flügge edition of Ref. 13, Refs. 14, 15, and pp. 178–83 in Ref. 16). However, such arguments are very unsatisfactory when the quantum numbers are small, and furthermore there occur ambiguities in the resulting matrix element formulas due to the fact that one is concerned with four classical turning points (cf. the discussion of Debye's matrix element formula below). The arguments in question would, by the way, fail completely when one uses the higher-order phase-integral approximations with their strong singularities at the classical turning points and their extremely great accuracy when these points are not approached.

To illuminate what has just been said, we refer, for instance, to Debye's formula for the matrix element of a variable x (corresponding to z in our notations). In Debye's matrix element formula (12) in Ref. 12 the interval (x_1, x_2) is not well defined, since the classically allowed region is different for each one of the two quantal states involved. If one chooses x_1 and x_2 to be the interior turning points, Debye's formula should be correct to the order of magnitude. If, however, one choses x_1 and x_2 to be the outer turning points, Debye's formula may be unreliable, and it turns out that the matrix element of a power of x may be wrong by several orders of magnitude, if one of the states is excited and the other is low-lying. This has been checked by calculations on the harmonic oscillator.[11]

One can draw quite wrong conclusions when one tries to calculate quantal matrix elements by means of phase-integral approximations without sufficient mathematical rigor and understanding of the delicate properties of those approximations. To demonstrate this, we shall consider the diagonal matrix element, i.e., the expectation value, of an analytic function $g(z)$ with respect to a quantal state $|n\rangle$ of a single-well potential. By definition this expectation value is

$$\langle n|g|n\rangle = \int g(z)[\psi(z,E_n)]^2\,dz, \tag{11}$$

where $\psi(z,E_n)$ is the exact, normalized wavefunction (real on the real axis), and the path of integration in (11) is the interval of the real axis appropriate to the physical problem in question. Since the exact wavefunction is analytic, and the same is assumed to be true for the function $g(z)$, one can without any approximation replace the original path of integration in (11) by a contour which lies everywhere far away from the classically allowed region, i.e., which lies in the region where the normalized wavefunction $\psi(z,E_n)$ is approximately given by (8a). On this new contour of integration the relative error of the approximate wavefunction (8a) is very small (especially for large quantum numbers n), and therefore one might be tempted to substitute the approximation (8a) for the exact, normalized wavefunction directly in the integrand of (11), after the path of integration has been deformed as just described. In this way one would conclude that the expectation value $\langle n|g|n\rangle$ were approximately equal to the expression

$$i\left(\int_{\Gamma_n} \frac{dz}{q(z,E_n)}\right)^{-1} \int g(z)\exp[-2iw(z,E_n)]\,\frac{dz}{q(z,E_n)}. \tag{12}$$

One can, however, realize that this conclusion is erroneous by arguing as follows. Let us, in particular, choose $g(z)$ to be equal to $Q^2(z,E_n)$ and use the first-order JWKB approximation, which means that $dw(z,E_n)/dz = Q(z,E_n)$. The expression (12) can then be written as follows:

$$i\left(\int_{\Gamma_n} \frac{dz}{Q(z,E_n)}\right)^{-1} \int Q^2(z,E_n)\exp[-2iw(z,E_n)]\,\frac{dz}{Q(z,E_n)} = -\frac{1}{2}\left(\int_{\Gamma_n} \frac{dz}{Q(z,E_n)}\right)^{-1} \int dz\,\frac{d}{dz}\exp[-2iw(z,E_n)] = 0, \tag{13}$$

the last member of (13) being obtained as a consequence of the assumption that $\exp[-iw(z,E_n)]$ tends to zero as z approaches the end points of the contour of integration and the fact that $\int_{\Gamma_n} dz/Q(z,E_n)$ is different from zero. One would thus, in this way, obtain the value zero for $\langle n|Q^2(z,E_n)|n\rangle$ when the first-order JWKB approximation is used. By considering the harmonic oscillator we can in the following way see that this result is not correct. With a conveniently chosen dimensionless linear coordinate z for the harmonic oscillator we have $Q^2(z,E_n) = (2n+1) - z^2$, and the exact expression for the expectation value of $Q^2(z,E_n)$ is found to be equal to $n+\frac{1}{2}$. This value of $\langle n|Q^2(z,E_n)|n\rangle$ is also obtained for all orders of the phase-integral approximations from the correct formula (4). The reason why an incorrect result was obtained in the above way of arguing is that, in a large part of the region, in which (8a) is valid with a very small relative error, the wavefunction is oscillating with a very large amplitude, and when one calculates the integral in (11) by integrating in such a region, it is very dangerous to introduce any approximation in the integrand.

3. THE CASE WHEN $A = g(z)\, d/dz$

Letting now A be the operator $g(z)d/dz$, where $g(z)$ is an analytic function which is regular on and within the contour of integration Γ, we obtain, in the way described at the end of Section 1, the approximate formula

$$\left\langle n \left| g(z)\frac{d}{dz} \right| n' \right\rangle = \int_\Gamma \left(-iq(z,E_{n'}) - \frac{1}{2q(z,E_{n'})}\frac{dq(z,E_{n'})}{dz}\right) g(z) \frac{\exp\{i[w(z,E_n)-w(z,E_{n'})]\}}{q^{1/2}(z,E_n)\, q^{1/2}(z,E_{n'})}\, dz \Bigg/ \left(\int_\Gamma \frac{dz}{q(z,E_n)}\right)^{1/2} \left(\int_\Gamma \frac{dz}{q(z,E_{n'})}\right)^{1/2} \tag{14}$$

which is valid provided that the function $q(z,E_{n'})g(z)\exp\{i[w(z,E_n)-w(z,E_{n'})]\}$ is not too large and too rapidly varying on the contour of integration. For the validity of (14) it is therefore in general required that $E_n \gtrsim E_{n'}$, unless one uses phase-integral approximations of sufficiently high orders. This prescription has been confirmed by calculations on the harmonic oscillator. [11]

Starting from the definition (5) with $A = g(z)d/dz$, and using partial integration, one easily obtains the exact relation

$$\left\langle n \left| g\frac{d}{dz} \right| n' \right\rangle = -\left\langle n' \left| \frac{dg}{dz} \right| n \right\rangle - \left\langle n' \left| g\frac{d}{dz} \right| n \right\rangle, \tag{15}$$

which, in combination with (14) with the indices n and n' interchanged, is often useful for the calculation of $\langle n | g(z)d/dz | n' \rangle$ when $E_n < E_{n'}$.

To illuminate the consistency of (14), we shall now consider the particular case of this approximate formula obtained when $n' = n$:

$$\left\langle n \left| g(z)\frac{d}{dz} \right| n \right\rangle = \int_\Gamma \left(-iq - \frac{1}{2q}\frac{dq}{dz}\right) g \frac{dz}{q} \Bigg/ \int_\Gamma \frac{dz}{q} = \left[-i\int_\Gamma g\,dz + \int_\Gamma \tfrac{1}{2} g \frac{d}{dz}\left(\frac{1}{q}\right) dz \right] \Bigg/ \int_\Gamma \frac{dz}{q}. \tag{16}$$

Since $g(z)$ is a regular analytic function within and on the closed contour Γ, and since $q(z,E)$ is single-valued on this contour, we can, after a partial integration, write the approximate formula (16) as follows:

$$\left\langle n \left| g(z)\frac{d}{dz} \right| n \right\rangle = \int_\Gamma -\frac{1}{2}\frac{dg}{dz}\frac{dz}{q} \Bigg/ \int_\Gamma \frac{dz}{q}. \tag{17}$$

This approximate formula can also be obtained directly by using the exact relation (15) with $n' = n$ and the approximate expectation value formula (4).

4. GENERALIZATION

For the sake of simplicity we have in the present paper assumed E_n and $E_{n'}$ to be energy levels in the same single-well potential. Our results are, however, applicable also when E_n and $E_{n'}$ are energy levels in different, neighboring single-well potentials. Thus, to mention a particular application, when we consider bound states in different, but neighboring, single-well potentials and put $g(z) \equiv 1$, the formulas (10a), (10b), or (10c) can be used for calculating Franck–Condon factors.

[1] N. Fröman, Ark. Fys. 32, 541 (1966).
[2] N. Fröman, Ann. Phys. (N.Y.) 61, 451 (1970).
[3] N. Fröman and P.O. Fröman, Ann. Phys. (N.Y.) 83, 103 (1974).
[4] N. Fröman and P.O. Fröman, Nuovo Cimento B 20, 121 (1974).
[5] N. Fröman and P.O. Fröman, Nucl. Phys. A 147, 606 (1970).
[6] J.A. Campbell, J. Comp. Phys. 10, 308 (1972).
[7] J.A.M. McHugh, Arch. Hist. Exact Sci. 7, 277 (1971).
[8] N. Fröman, Phys. Lett. A 48, 137 (1974).
[9] P.O. Fröman, Ann. Phys. (N.Y.) 88, 621 (1974).
[10] N. Fröman and P.O. Fröman, J. Math. Phys. 18, 96 (1977).
[11] B. Arvidsson, P.O. Fröman, F. Karlsson, and W. Mrazek (unpublished results).
[12] P. Debye, Physik. Zeitschr. XXVIII, 170 (1927).
[13] W. Pauli, *Die allgemeinen Prinzipien der Wellenmechanik, Handbuch der Physik* (*Geiger–Scheel*), Band XXIV (Springer-Verlag, Berlin, 1933), 1st ed.; *Encyclopedia of Physics*, edited by Flügge (Springer-Verlag, Berlin, Göttingen, Heidelberg, 1958), Vol. V, Part 1.
[14] L.D. Landau, Physik Z. Sowjetunion 1, 88 (1932).
[15] L.D. Landau, Physik Z. Sowjetunion 2, 46 (1932).
[16] L.D. Landau and E.M. Lifshitz, *Quantum Mechanics* (Non-Relativistic Theory). Course of Theoretical Physics, Vol. 3 (Pergamon, London and Paris, 1958).

Energy forms, Hamiltonians, and distorted Brownian paths*

Sergio Albeverio* and Raphael Høegh-Krohn

ZiF, University of Bielefeld, Germany
and Insititute of Mathematics, University of Olso, Blindern, Oslo 3, Norway

Ludwig Streit

Fakultät für Physik, University of Bielefeld, Germany
(Received 25 October 1976)

We study the Hamiltonians for nonrelativistic quantum mechanics—and for the heat equation—in terms of energy forms $\int \nabla f \nabla f d\mu$, where $d\mu$ is a positive, not necessarily finite measure on R^n. We cover the cases of very singular interactions (e.g., N particles in R^3 interacting by two-body "δ potentials"). We also exhibit, on the other hand, regularity conditions for μ in order that H be realized as a perturbation of the Laplacian by a measurable or generalized function. The Hamiltonians defined by energy forms always generate Markov semigroups, and the associated processes are symmetric homogeneous strong Markov diffusion Hunt processes with continuous paths realizations. Ergodicity, transiency, and recurrency are also discussed. The associated stochastic differential equation is discussed in the situation where μ is finite but the drift coefficient is only restricted to be in $L_2(R^n, d\mu)$. These results provide a large class of examples where solutions of the heat equation can be expressed by averages with respect to the constructed Hunt processes, rather than with respect to Brownian motion. This is discussed in relation to recent work of Ezawa, Klauder, and Shepp, as well as of Hida and Streit.

1. INTRODUCTION

There is a well-known relation between the theory of diffusion processes, Markov semigroups, second order elliptic and parabolic partial differential equations, and potential theory on the Euclidean space R^n, see, e.g., Refs. 1–7. New momentum to this study, especially for the case of symmetric Markov processes, was given more recently by Fukushima, see Refs. 8–10 and references therein. This approach, continued particularly by Silverstein,[11,12] is of the analytic potential theoretic nature, based on the study of general Dirichlet forms, by methods extended from the previous classical work of Beurling and Deny.[13,14] In Refs. 15–18 two of us have adapted this approach to perform a detailed study of the energy forms on infinite dimensional spaces, with a view to applications in quantum field theory.

The results and methods also have new implications in the finite dimensional case, as remarked in Refs. 15, 16, and 19, however a systematic and detailed study of this case is only undertaken in the present paper. It turns out that the approach permits us, in particular, to define positive self-adjoint energy operators, Hamiltonians from the point of view of quantum mechanics, for very singular perturbations of the Laplacian (e.g., sums of two-body "δ-interactions" in three dimensions), and, on the other hand, to give a natural incorporation of boundary conditions. Moreover, all Hamiltonians constructed generate Hunt diffusion processes with continuous path realizations, having the strong properties studied in the quoted work of Fukushima and Silverstein. The finite dimensionality of the space, as opposed to the infinite dimensionality of the cases discussed in the preceding work (Refs. 15–18) allows for stronger results and in particular we can deal with forms which are given by nonfinite measures (i.e., nonnormalizable ground states).

Let us mention that a very detailed study of the case where $n=1$, the measures are finite and smooth and the potentials are lower bounded polynomials, is contained in Refs. 20 and 21. We also refer the reader to these references for many questions which can be profitably discussed in that case and are not treated in our more general situation. From another point of view, stimulus for this work also has its roots in the discussions on the relations between stochastic processes and quantum mechanics, see, e.g., Refs. 22–24, and more particularly in the recent work of Ezawa, Klauder, and Shepp.[25] The latter authors consider the stochastic equation corresponding to the heat equation, under the usual smoothness and growth conditions on the drift coefficient. They emphasize and discuss particularly the replacement of averages with respect to the Brownian motion by averages with respect to the process given by the solution of the stochastic equation ("distorted Brownian path" picture). One of us, in collaboration with Hida,[19] has already given a detailed mathematical analysis of this interpretation in the case of Gaussian processes, in relation both to quantum mechanics and quantum field theory. In this work it was shown in particular that the transformation which realizes the distorted Brownian path picture is the canonical one in the usual technical sense of Gaussian processes. Moreover, the advantages of the distorted Brownian paths picture as a formulation of the dynamics were emphasized and illustrated by examples. One of the results of the present paper in relation of the "distorted Brownian path" picture of Ref. 25 is a precise formulation of this picture for cases where the drift coefficient is not restricted by the usual Hölder and growth conditions of the theory of stochastic equations (in particular it may have singularities and at infinity can grow faster than linearily). In terms of the associated potential V, entering in the Hamiltonian as a perturbation of the Laplacian, we allow singularities (in fact, V need not even be a measurable function, it can be a distribution) and more than quadratic growth at infinity. This illustrates the usefulness of considering the ground state—rather than the potential—as dynamical input. This point of view has its historical roots in Ref. 26. Let us now go briefly through the content of the different sections in the present paper.

In Sec. 2 we study the energy forms in R^n given by $\int \nabla f \nabla g \, d\mu$, where ∇ is the gradient, f and g are $C^1(R^n)$ functions of compact support, and μ is a positive regular σ finite Borel measure on R^n. When the form is closable in $L_2(R^n, d\mu)$, then it defines a self-adjoint positive operator H, the energy operator. Theorems 2.1–2.4 contain sufficient conditions for the measure μ in order that the energy form be closable, hence H well defined, and in fact equal, on a dense domain in $L_2(R^n, d\mu)$, to $-\nabla - \beta(x)\cdot\nabla$, where Δ is the Laplacian and $\beta(x)$ is related to the measure μ by $\beta(x) = 2\nabla\phi(x)/\phi(x)$, with $d\mu(x) = \phi(x)^2 dx$. Sufficient conditions on the measure μ in order that H be realizable in the form $-\nabla + V$ with V a generalized respectively measurable function are in Theorem 2.3, resp. Theorems 2.2 and 2.4. Theorem 2.5 provides a relation between the growth at infinity of $\phi(x)$ and the location of the lower end of the spectrum of H. Essential self-adjointness of H on a natural domain is examined in Theorem 2.6. Finally Theorem 2.7 extends the preceding results to the case of operators of the form $-\sigma(x)\Delta - \beta(x)\cdot\nabla$, where $\sigma(x)$ is a positive function, and to arbitrary open domains instead of R^n. A considerable part of Sec. 2 is devoted to the discussion of several examples which illustrate the results and the range of application, such as boundary conditions (Examples 1, 2, 8), "δ potentials" (Examples 3 and 4 and Lemma 2.6) and other singular perturbations (Examples 5, 7, 9) (Example 9 is related to Ref. 27).

In Sec. 3 we show that the energy forms defined in Sec. 2 fall, from the potential theoretic point of view, into the class covered by the Beurling–Deny and Fukushima methods,[8–10,13,14] and therefore generate, in particular, Markov semigroups $\exp(-tH)$, $t \geq 0$. Moreover we show that the energy forms are regular local Dirichlet forms in the sense of Fukushima and Silverstein,[28,11] hence there exist Hunt[3] processes ξ_t which are properly associated with them and have continuous trajectories outside a set of capacity zero. The results of Silverstein[11] are then applied to discuss ergodicity, transiency, and recurrency of the energy forms and the associated processes.

In Sec. 4 we derive the stochastic differential equation satisfied by the process ξ_t associated, by the construction of Sec. 3, with the energy form. In the case where $d\mu(x) = \phi(x)^2 dx$, with ϕ such that $\phi, \nabla\phi \in L_2 \in L_2(R^n, dx)$, ξ_t satisfies the stochastic differential equation

$$d\xi_t = \beta(\xi_t)\,dt + dw_t,$$

where w_t is the Brownian motion in R^n with unit coefficient of diffusion, and $\beta = 2\phi^{-1}\nabla\phi$. We also show uniqueness of the solutions within the class of Markov processes with invariant distribution μ. We end with a remark concerning the "distorted Brownian path" picture discussed in Refs. 25 and 19.

2. ENERGY FORMS IN R^n

Let μ be a positive regular σ finite Borel measure on R^n. Let $L_2(d\mu)$ be the Hilbert space of real square integrable functions and consider R^n as a real Hilbert space with its natural Hilbert structure. For f in the space $C_0^1(R^n)$ of continuously differentiable functions of compact support we define the linear operator ∇ from $L_2(d\mu)$ to the Hilbert space $L_2(d\mu)\otimes R^n$ by $\nabla = \{\partial f/\partial x_1, \cdots, \partial f/\partial x_n\}$. We define the energy form E given by μ by

$$E(f,g) = (\nabla f, \nabla g) = \sum_{i=1}^{n} \int \frac{\partial f}{\partial x_i}\,\frac{\partial g}{\partial x_i}\,d\mu, \tag{2.1}$$

for f and g in $D(\nabla)$.

We shall say that μ is admissible if E is a closable form in $L_2(d\mu)$, which is equivalent to ∇ being a closable operator from $L_2(d\mu)$ to $L_2(d\mu)\otimes R^n$. In this case we shall denote the respective closures of E and ∇ again by E and ∇. From now on we shall assume that μ is admissible and E and ∇ are both closed.

Let $E_1(f,g) = E(f,g) + (f,g)$, where (f,g) is the inner product in $L_2(d\mu)$, and set $|f|_1^2 = E_1(f,f)$. Then $|f|_1$ is the graph norm of ∇ and it organizes $D(\nabla)$ to a complete Hilbert space. Since E is a symmetric nonnegative closed form it defines uniquely a nonnegative self-adjoint operator H on $L_2(d\mu)$, and we shall call H the energy operator given by the admissible measure μ. Since ∇ is closed it has an adjoint ∇^* which is densely defined and one has $H = \nabla^*\nabla$.

Theorem 2.1: If μ is absolutely continuous with respect to the Lebesgue measure such that the density $\rho(x) = d\mu/dx$ has a logarithm $\ln\rho(x)$ in $D(\nabla)$, then μ is admissible. On the other hand, if μ is admissible and the coordinate functions are in $D(H)$, then $\ln\rho(x) \in D(\nabla)$.

Proof: Let $\beta(x) = \nabla\ln\rho(x)$. Then one observes that $-\nabla - \beta(x)$ is a densely defined adjoint of ∇, from which it follows that ∇ is closable. If, on the other hand, μ is admissible, then $H = \nabla^*\nabla$ and $Hx_i = \nabla^*\{0, 0, \ldots, 1, 0, \ldots, 0\}$. Hence for $f \in C_0^\infty(R^n)$ we have $(f, Hx_i) = \int(\partial f/\partial x_i)\,d\mu$, since $C_0^\infty(R^n) \subset D(\nabla)$. Since μ is σ finite, it defines a distribution μ and in fact we have $Hx_i = -(\partial/\partial x_i)\mu \in L_2(d\mu)$, in the sense of distributions. It is well known from the theory of distributions that this implies that μ is absolutely continuous with respect to the Lebesgue measure so that $d\mu = \rho dx$ and moreover $\partial\rho/\partial x_i = (\partial/\partial x_i)\mu$ in the sense of distributions. Hence $(f, Hx_i) = -\langle f, \partial\rho/\partial x_i\rangle$, where the left-hand side is the $L_2(d\mu)$ inner product. Thus we have $Hx_i\,d\mu = -\partial\rho/\partial x_i$ as distributions, i.e., $Hx_i\,\rho(x) = -\partial\rho/\partial x_i(x)$, and therefore $\nabla\rho/\rho \in L_2(d\mu)$. This proves the theorem. ∎

Let us now assume that μ is absolutely continuous with respect to the Lebesgue measure, i.e., $d\mu(x) = \rho(x)\,dx$ and set $\phi(x) = \rho(x)^{1/2}$. Then $\phi \in L_2^{\text{loc}}(dx)$. $\ln\rho(x) \in D(\nabla)$ is equivalent to $\ln\phi(x) \in D(\nabla)$ and in fact

$\beta = 2\nabla \ln\phi$. Now

$$\|\beta\|^2 = 4\|\nabla \ln\phi\|^2 = 4\int \frac{|\nabla\phi|^2}{\phi^2}\, d\mu = 4\int |\nabla\phi|^2\, dx. \tag{2.2}$$

So we see that $\beta \in L_2(d\mu)$ iff $\nabla\phi \in L_2(dx)$.

We observe now that if we only have $\nabla\phi \in L_2^{\mathrm{loc}}(R^n)$ then, with $\beta(x) = 2\nabla\ln\phi(x)$, we still have that $-\nabla - \beta(x)$ is defined on $C_0^1(R^n)$ and is a formal adjoint of ∇. Since $C_0^1(R^n)$ is dense in $L_2(R^n)$, ∇ is closable and thus μ is admissible. Moreover, in this case we also have that ∇ maps $C_0^2(R^n)$ into $C_0^1(R^n) \subset D(\nabla^*)$, since $\nabla^* = -\nabla - \beta(x)$. Therefore, we have that $C_0^2(R^n) \subset D(H)$ and, for $f \in C_0^2(R^n)$,

$$Hf = -\Delta f - \beta(x)\cdot\nabla f(x), \tag{2.3}$$

where $\Delta f = \sum_{i=1}^n \partial^2 f/\partial x_i^2$. If, in addition, we have $\phi \in L_2(dx)$, then μ is bounded and $1 \in L_2(d\mu)$ and obviously $H1 = 0$.

On the other hand, if ϕ, $\nabla\phi$, $\nabla\phi/\phi$, and $\Delta\phi/\phi$ are all in $L_2^{\mathrm{loc}}(R^n)$, where the derivatives have to be understood in the sense of distributions, then for $f \in C_0^2(R^n)$ we have that $\phi^{-1}f \in D(H)$, because

$$\nabla(\phi^{-1}f) = \phi^{-1}\nabla f - \phi^{-1}\frac{\nabla\phi}{\phi}f \tag{2.4}$$

is in $L_2(d\mu)$, so that $\phi^{-1}f \in D(\nabla)$ and

$$E(\phi^{-1}f,\ \phi^{-1}f) = \int_{R^n} f(x)(-\Delta + V(x))f(x)\,dx, \tag{2.5}$$

with

$$V(x) = \frac{\Delta\phi}{\phi}(x) = \tfrac14\beta(x)\cdot\beta(x) + \tfrac12\nabla\cdot\beta(x). \tag{2.6}$$

Since $V = \Delta\phi/\phi \in L_2^{\mathrm{loc}}(R^n)$, it follows that $C_0^2(R^n) \subset D(-\Delta + V)$, so that by (2.5) and the definition H we have $\phi^{-1}f \in D(H)$. Hence $f \longrightarrow \phi^{-1}f$ is an isometry between $L_2(dx)$ and $L_2(d\mu)$ which takes H into a self-adjoint extension of $-\Delta + V(x)$ as defined on $C_0^2(R^n)$. In fact the form

$$\int_{R^n} f(x)(-\Delta + V(x))f(x)\,dx \tag{2.7}$$

restricted to $f = \phi g$ with $g \in C_0^1(R^n)$ is a closable nonnegative form and the operator defined by its closure is the image of H under the isometry $f \longrightarrow \phi^{-1}f$ between $L_2(dx)$ and $L_2(d\mu)$. If, in addition, $V(x) \geq -c|x|^2$ for some constant c, then $-\Delta + V$ is essentially self-adjoint on $C_0^2(R^n)$ [29] so that in this case the image of H in $L_2(dx)$ is the closure of $-\Delta + V$. Other conditions for $-\Delta + V$ to be essentially self-adjoint on $C_0^2(R^n)$ are well known, e.g., $V \in L_2(R^n) + L_\infty(R^n)$ for $n \leq 3$ or $V \in L_p(R^n) + L_\infty(R^n)$ for some $p > n/2$ for $n \geq 4$.[30] We formulate these results in the following theorem.

Theorem 2.2: 1. Let μ be absolutely continuous with respect to the Lebesgue measure and set $d\mu(x) = \phi^2(x)\,dx$.

If $\nabla\phi$ is in $L_2^{\mathrm{loc}}(R^n)$, where the derivative is taken in the sense of distributions, then μ is admissible. Moreover, $C_0^2(R^n) \subset D(H)$, and for $f \in C_0^2(R^n)$ we have

$$Hf = -\Delta f - \beta\cdot\nabla f,$$

where $\beta(x) = 2\phi^{-1}(x)\nabla\phi(x)$. If, in addition, $\phi \in L_2(dx)$, then $1 \in L_2(d\mu)$ and $H\cdot 1 = 0$.

2. Let μ be equivalent with the Lebesgue measure and let $\nabla\phi$, $\phi^{-1}\nabla\phi$, and $\phi^{-1}\Delta\phi$ be all in $L_2^{\mathrm{loc}}(R^n)$, where the derivatives are taken in the sense of distributions. Then $\phi^{-1}C_0^2(R^n) \subset D(H)$ and for $f \in C_0^2(R^n)$

$$H\phi^{-1}f = \phi^{-1}(-\Delta + V)f,$$

where $V(x) = (\Delta\phi/\phi)(x) = \frac14\beta(x)\cdot\beta(x) + \frac12\nabla\cdot\beta(x)$.

Moreover the isometry $f \longrightarrow \phi^{-1}f$ between $L_2(dx)$ and $L_2(d\mu)$ takes H into a self-adjoint extension of $-\Delta + V(x)$ on $C_0^2(R^n)$. If, in addition, $V(x) \geq -c|x|^2$ for some constant c then $-\Delta + V$ is essentially self-adjoint and hence $\phi H\phi^{-1}$ is the closure of $-\Delta + V$. ∎

We shall now consider in more detail the case where μ is equivalent with the Lebesgue measure, and we assume in addition that $\nabla\phi$ and $\phi^{-1}\nabla\phi$ are in $L_2^{\mathrm{loc}}(R^n)$. In this case it follows from (2.4) that $\phi^{-1}C_0^1(R^n) \subset D(\nabla)$, and for f in $C_0^\infty(R^n)$ we have

$$E(\phi^{-1}f,\ \phi^{-1}f) = \int_{R^n}\left[\nabla f\cdot\nabla f - 2\nabla f\frac{\nabla\phi}{\phi}f + \left(\frac{\nabla\phi}{\phi}\right)^2 f^2\ dx\right.$$
$$= \int_{R^n}\nabla f\cdot\nabla f\,dx$$
$$-\tfrac12\int_{R^n}\nabla(f^2)\circ\beta(x)\,dx + \tfrac14\int\beta^2 f^2\,dx, \tag{2.5'}$$

which is to say that for $f \in C_0^\infty(R^n)$

$$E(\phi^{-1}f,\ \phi^{-1}f) = \int_{R^n}\nabla f\cdot\nabla f\,dx + \int_{R^n}(\tfrac12\nabla\cdot\beta + \tfrac14\beta^2)f^2\,dx, \tag{2.8}$$

where the derivative $\nabla\beta(x)$ is to be taken in the sense of distributions. So introducing the distribution $V(x) = \frac12\nabla\beta + \frac14\beta^2$ we have that, for $f \in C_0^\infty(R^n)$,

$$E(\phi^{-1}f,\ \phi^{-1}f) = \int_{R^n}\nabla f\cdot\nabla f\,dx + \int f\,Vf\,dx. \tag{2.9}$$

We state these results in the following theorem.

Theorem 2.3: Let μ be equivalent with the Lebesgue measure and set $d\mu(x) = \phi^2(x)\,dx$, and assume that $\nabla\phi$ and $\beta(x) = 2\phi^{-1}(x)\nabla\phi(x)$ are in $L_2^{\mathrm{loc}}(R^n)$, where the derivatives are taken in the sense of distributions. Then $\phi^{-1}C_0^1(R^n) \subset D(\nabla) = D(H^{1/2})$. The distribution $V(x) = \frac12\nabla\beta + \frac14\beta^2$ is a continuous linear functional on $C_0^1(R^n)$ and for $f \in C_0^1(R^n)$ we have

$$E(\phi^{-1}f,\ \phi^{-1}f) = \int_{R^n}|\nabla f|^2\,dx + \int_{R^n} Vf^2\,dx,$$

where V is here considered as an element in the dual of $C_0^1(R^n)$. Moreover, the right-hand side in the equation above is a form defined on $C_0^1(R^n)\times C_0^1(R^n)$ which is nonnegative and closable in $L_2(dx)$, and this form is the restriction of the form of the self-adjoint operator $\phi H\phi^{-1}$ in $L_2(dx)$ to $C_0^1(R^n)\times C_0^1(R^n)$. ∎

There is another version of the first part of Theorem 2.2 which we shall give, but first we need some notations. If T is a distribution we shall say that $T \in L_p^{\mathrm{loc}}(R^n - A)$, where A is a closed set, iff for any $f \in C_0^\infty(R^n - A)$ we have that the distribution $f(x)T(x)$ is actually given by a function in $L_p(R^n)$. We then have

Theorem 2.4: Let $d\mu(x) = \phi^2(x)\,dx$ where $\phi \in L_2^{\mathrm{loc}}(R^n)$. If there is a closed set N of Lebesgue measure zero so that the distributions $\nabla\phi(x)$ are in $L_2^{\mathrm{loc}}(R^n - N)$, then μ is admissible and for any $f \in C_0^2(R^n - N)$ we have

$$Hf = -\Delta f - \beta\cdot\nabla f,$$

where $\beta(x) = 2\nabla\phi(x)/\phi(x)$.

Proof: Let $f\in C_0^1(R^n)$ and $g\in C_0^1(R^n-N)$. Then

$$\int \nabla f g\phi^2 dx = -\langle f, \nabla(g\phi^2)\rangle.$$

Now since $\phi\in L_2^{\mathrm{loc}}(R^n)$ we get that

$$\nabla(g\phi\,\phi)=\nabla(g\phi)\,\phi+g\phi\,\nabla\phi$$
$$=\nabla g\,\phi^2+2g\phi\,\nabla\phi$$

in the sense of distributions by utilizing the fact that the derivatives of distributions are the weak derivatives with respect to translations, and that the product of two L_2-convergent sequences converges in L_1, which implies distributional convergence. From this it follows that

$$\int \nabla f\cdot g\phi^2 dx = -\int f(\nabla g+\beta g)\phi^2 dx$$

and obviously $\beta(x)\cdot g(x)\in L_2(d\mu)$, so that $-\nabla-\beta(x)$ defined on $C_0^1(R^n-N)$ is a formal adjoint of ∇. Since $C_0^1(R^n-N)$ is dense in $L_2(d\mu)$ we have that ∇ is closable, and therefore μ is admissible. That for $f\in C_0^2(R^n-N)$ we have that $Hf=-\Delta f-\beta\nabla f$ follows from the fact that ∇ maps $C_0^2(R^n-N)$ into $C_0^1(R^n-N)$ which is in the $D(\nabla^*)$ and where ∇^* takes the form $-\nabla-\beta$, together with the fact that $H=\nabla^*\nabla$. ∎

We shall now give some examples to illustrate the strength of the theorems above. We shall not give all calculations in connection with the different examples, but in each case the statements can be easily verified.

Example 1: Let $\Omega\subset R^n$ be compact with $\partial\Omega$ piecewise C^1. Then Δ_0 in $L_2(\Omega)$ with Dirichlet boundary conditions on $\partial\Omega$ is well defined as a self-adjoint operator and it has discrete spectrum. Let λ_0 be its lowest eigenvalue and ϕ_0 be its lowest eigenvalue and ϕ_0 be its corresponding normalized eigenfunction. It is well known that ϕ_0 may be chosen nonnegative in which case it is unique. We define $\phi(x)=0$ for $x\notin\Omega$ and $\phi(x)=\phi_0(x)$ for $x\in\Omega$, then $\phi(x)$ is continuous and piecewise C^1. We define $d\mu=\phi^2dx$. It follows now from Theorem 2.2 that μ is admissible. And by utilizing that $\phi_0(x)$ is in fact strictly positive in the interior of Ω we actually find that

$$\phi H\phi^{-1}=-\Delta_0+\lambda_0$$

as a self-adjoint operator in $L_2(\Omega, dx)$.

Example 2: Let Ω be as in Example 1 and let λ_k be any eigenvalue of Δ_0 and ϕ_k be its normalized eigenfunction. Again we extend ϕ_k to R^n by setting it equal to zero outside Ω. Let $d\mu=\phi^2dx$ where $\phi=\phi_k$. Again we have that μ is admissible and in fact we find that in this case

$$\phi H\phi^{-1}=-\Delta_k+\lambda_k,$$

where Δ_k is the self-adjoint Laplacian with Dirichlet boundary conditions on the set $\phi_k(x)=0$. This means that $-\Delta_k$ is the operator given by the closure of the form $\int_\Omega|\nabla g|^2dx$, where $g\in\phi_kC_0^1(R^n)$.

Example 3: Let $n=3$ and $\phi(x)=(4\pi|x|)^{-1}\exp(-m|x|)$ with $m\geq 0$. Then $\nabla\phi(x)=-(4\pi|x|^3)^{-1}(x-mx|x|)\times\exp(-m|x|)$ which obviously is in $L_2^{\mathrm{loc}}(R^3-\{0\})$. Hence we have by Theorem 2.4 that μ is admissible for all $m\geq 0$.

Let $f\in C_0^1(R^3)$ such that $f(0)=0$, we then easily find that $\phi^{-1}f\in C_0^1(R^3)\subset D(\nabla)$. By computation we have

$$E(\phi^{-1}f,\ \phi^{-1}f)=\int_{R^n}(|\nabla f|^2+m^2|f|^2)dx.$$

Moreover, it follows from this equation that if $f\in C_0^2(R^3)$ with $f(0)=0$ we have that $\phi^{-1}f\in D(H)$ and

$$(\phi H\phi^{-1})f=(-\Delta+m^2)f,$$

so that in this case $\phi(H-m^2)\phi^{-1}$ is a self-adjoint extension of $-\Delta$ defined on $f\in C_0^2(R^3)$ with $f(0)=0$. This self-adjoint extension $-\Delta_m$ is actually different for different values of m since it has $(4\pi|x|)^{-1}\exp(-m|x|)$ as an eigenfunction with eigenvalue m^2 for $m>0$ and no eigenfunction for $m=0$. Hence Δ_m is a one parameter family of self-adjoint extensions of Δ on $f\in C_0^2(R^3)$ with $f(0)=0$. For a complete characterization of all the self adjoint extensions of this operator see Ref. 31.

Example 4: Let us set $n=3N$ and let us write $x\in R^n$ as $\{x_1,\ldots,x_N\}$ with $x_i\in R^3$. Let Δ_i be Δ_{x_i} so that $\Delta=\sum_{i=1}^n\Delta_i$. Set

$$\phi(x)=\frac{1}{4\pi}\sum_{i<j}^N|x_i-x_j|^{-1}\exp(-m|x_i-x_j|$$

with $m\geq 0$. We then obviously have that $\phi\in L_2^{\mathrm{loc}}(R^n)$. Let $D=\{x\,|\,x_i=x_j \text{ for some } i\neq j\}$, then obviously D is a closed set of measure zero and it follows from the formula for $\nabla\phi$ in the previous example that $\nabla\phi(x)\in L_2^{\mathrm{loc}}(R^n-D)$. Hence we have by Theorem 2.4 that μ is admissible for all $m\geq 0$.

We observe now that, for $x\in R^n-D$, $\Delta\phi(x)=(1/4\pi)\times\sum_{i<j}^N(\Delta_i+\Delta_j)\times|x_i-x_j|^{-1}\exp(-m|x_i-x_j|$, so that for $x\in R^n-D$ we have $\Delta\phi=2m^2\phi$. Therefore, let $f\in C_0^1(R^n-D)$, we then easily find that

$$E(\phi^{-1}f,\ \phi^{-1}f)=\int_{R^n}(|\nabla f|^2+2m^2|f|^2)dx,$$

or if $f\in C_0^2(R^n-D)$ we have that $\phi^{-1}f\in D(H)$ and

$$(\phi H\phi^{-1})f=(-\Delta+2m^2)f.$$

So in this case $\phi(H-2m^2)\phi^{-1}$ is a self-adjoint extension of $-\Delta$ defined on $C_0^2(R^n-D)$. These self-adjoint extensions are actually all different from the usual Laplacian since any $f\in C_0^2(R^n)\subset D(H)$ implies that ϕf for $f\in C_0^2(R^n)$ is in the domain of $\phi H\phi^{-1}$ and ϕf is evidently not in the domain of Δ. In fact from Theorem 2.5, to be proven subsequently, it follows that H has 0 as a greatest lower bound for its spectrum, so the extension $\phi(H-2m^2)\phi^{-1}$ has $-2m^2$ as a greatest lower bound for its spectrum. Hence the extensions for different values of m are all unitarily inequivalent and therefore different for different values of m. We therefore have:

Let

$$\Delta_m=\phi(H-2m^2)\phi^{-1}.$$

Then Δ_m is a self-adjoint extension of the restriction of Δ to $C_0^2(R^n-D)$ where $x\in D \iff x_i=x_j$ for some $i\neq j$, such that $-2m^2$ is the greatest lower bound for its spectrum. So obviously $\Delta_{m_1}=\Delta_{m_2}\Longrightarrow m_1=m_2$ and moreover Δ_m for $m=0$ is still different from Δ. Hence Δ_0, i.e., Δ_m for $m=0$, is an extension different from Δ which is nonnegative. Moreover, Δ_m is obviously invariant under arbitrary translations or permutations and under $O(3)$, since ϕ is invariant under these groups.

For previous discussions on the self-adjoint extensions of the restrictions of Δ to $C_0^2(R^n - D)$ in the case of $N=3$ see Refs. 32 and 33.

Example 5: Let $\phi(x)=1$ for $|x|\leq 1$ and $\phi(x)=|x|^{-(n-2)}$ for $|x|\geq 1$, $x\in R^n$. Then obviously $\phi\in L_2^{\mathrm{loc}}(R^n)$. Moreover $\nabla\phi=0$ for $|x|\leq 1$ and $\nabla\phi=-(n-2)\,|x|^{-(n-1)}(x/|x|)$ for $|x|\geq 1$ is also in $L_2^{\mathrm{loc}}(R^n)$. Also $\beta(x)=\phi^{-1}\nabla\phi=0$ for $|x|\leq 1$ and $\beta(x)=-2(n-2)\,|x|^{-2}x$ for $x\geq 1$ is also in $L_2^{\mathrm{loc}}(R^n)$. We continue now according to Theorem 2.3 and calculate the distribution $V(x)=\frac{1}{2}\nabla\cdot\beta+\frac{1}{4}\beta^2$. We find that $\frac{1}{2}\nabla\cdot\beta=(2n-4)\delta(x^2-1)-(n-2)^2|x|^{-2}\,\theta(x^2-1)$, $\frac{1}{4}\beta^2=(n-2)^2\,|x|^{-2}\,\theta(|x|^2-1)$, where $\delta(x^2-1)$ is the measure concentrated on the sphere of radius 1 and $\theta(s)=1$ for $s>0$ and 0 for $s\leq 0$. From this we see that $V(x)==2(n-2)\delta(x^2-1)$. Hence we get from Theorem 2.3 that μ is admissible, that $\phi^{-1}C_0^1(R^n)\subset D(H^{1/2})$ and

$$E(\phi^{-1}f,\ \phi^{-1}f)=\int_{R^n}|\nabla f|^2\,dx+2(n-2)\int_{|x|=1}f(x)^2\,ds.$$

If we replace $\phi(x)$ by $\phi(x)+\alpha$, then $2(n-2)$ would go to $\lambda=2(n-2)/(1+\alpha)$ in the formula above. We easily see that $\phi^{-1}C_0^1(R^n)$ is dense in the graph norm in $D(H^{1/2})$, from which it follows that the closure of the form above gives the operator $\phi H\phi^{-1}$.

Example 6: Let $\phi(x)=\exp(x^2/2)$, then $\nabla\phi=x\exp(x^2/2)$ and $\beta(x)=2x$ are all in $L_2^{\mathrm{loc}}(R^n)$ and also $V(x)=x^2+n$ is in $L_2^{\mathrm{loc}}(R^n)$. Hence according to Theorem 2.2, part 2 we know that $\phi^{-1}C_0^2(R^n)\subset D(H)$ and

$$\phi H\phi^{-1}=-\Delta+x^2+n$$

on $C_0^2(R^n)$. On the other hand, the operator $-\Delta+x^2-n$ is the harmonic oscillator which is essentially self-adjoint on $C_0^2(R^n)$ and has zero as a lowest eigenvalue with corresponding eigenfunction $\exp(-x^2/2)$. Hence we wee that this furnishes us with an example where the spectrum of H does not go down to zero but in fact $H\geq 2n$. It is also an example of a case where $\phi\in L_2(R^n)$ but where H still has discrete spectrum.

That the spectrum of H in this case does not start at zero has to do with the growth of ϕ at infinity. We shall see this in the next theorem. First we need a definition.

Let μ be absolutely continuous with respect to the Lebesgue measure with $d\mu=\phi^2dx$ and $\phi\in L_2^{\mathrm{loc}}(R^n)$. Let $\alpha_r=\mu(\{x;\ |x|\leq r\})$. Then for any $r>0$ we have that $\alpha_r^{-1}\alpha_{r+1}-1$ is an upper bound for the infimum of the spectrum of H. We shall say that μ has growth of exact order k iff $\alpha_r=Cr^k+O(r^k)$ for some $k\geq 0$. We see that if α_r has growth of exact order k then $\alpha_r^{-1}\alpha_{r+1}-1$ converges to zero as $r\to\infty$. This gives us the following theorem.

Theorem 2.5: Let $d\mu=\phi^2dx$ with $\phi\in L_2^{\mathrm{loc}}(R^n)$. If μ has growth of exact order k and μ is admissible then the infimum of the spectrum of H is zero. ∎

Let now

$$\phi(x)=\frac{1}{4\pi}\sum_{i<j}^{N}|x_i-x_j|^{-1}\exp(-m|x_i-x_j| \tag{2.10}$$

with $x\in R^n$, $n=3N$ and $x_i\in R^3$. Then

$$\phi(x)^2=\frac{1}{16\pi^2}\sum_{\substack{i<j\\k<l}}|x_i-x_j|^{-1}|x_k-x_l|^{-1}$$
$$\times\exp(-m|x_i-x_j|\exp(-m|x_k-x_l|.$$

Then obviously

$$\int_{|x|\leq r}\phi(x)^2\,dx=\frac{1}{16\pi^2}\sum_{i<j}\int_{|x|\leq r}|x_i-x_j|^{-2}$$
$$\times\exp(-2m|x_i-x_j|)\,dx+O(r^{n-6}) \tag{2.11}$$

or after a change of variables $y=(1/\sqrt{2})(x_i-x_j)$ and $z=x\ominus y\in R^{n-3}$,

$$\int_{|x|\leq r}\phi^2(x)\,dx=\frac{N(N-1)}{64\pi^2}\int_{|y|^2+|z|^2\leq r^2}|y|^{-2}$$
$$\times\exp(-2\sqrt{2}\,m|y|)\,dy\,dz+O(r^{n-6}). \tag{2.12}$$

Now set $s=|y|$ and $t=|z|$,

$$\int_{|y|^2+|z|^2\leq r^2}|y|^{-2}\exp(-2\sqrt{2}\,m|y|)\,dy\,dz \tag{2.13}$$
$$=|S_2|\ |S_{n-4}|\iint_{s^2+t^2\leq r^2}\exp(-2\sqrt{2}\,ms)\,t^{n-4}\,ds\,dt,$$

where $|S_k|$ is the area of the unit k-sphere. Equation (2.13) is equal to

$$|S_2|\,|S_{n-4}|\frac{r^{n-3}}{n-4}\int_0^r\exp(-2\sqrt{2}\,ms)\left[1-\left(\frac{s}{r}\right)^2\right]^{(n-3)/2}ds$$
$$=[(n-4)2\sqrt{2}\,m]^{-1}\,|S_2|\,|S_{n-4}|\,r^{n-3}+O(r^{n-4}). \tag{2.14}$$

Hence we have proved the following lemma.

Lemma 2.6: If $d\mu=\phi^2dx$ where

$$\phi(x)=\frac{1}{4\pi}\sum_{i<j}^{N}|x_i-x_j|^{-1}\exp(-m|x_i-x_j|,$$

then

$$\int_{|x|\leq r}d\mu=N(N-1)[64\pi^2\cdot(3N-4)2\sqrt{2}\,m]^{-1}$$
$$\times|S_2|\,|S_{n-4}|\,r^{3(N-1)}+O(r^{3N-4})$$

so that μ has growth of exact order $3(N-1)$, hence by Theorem 2.5 the infimum of the spectrum of H in Example 4 is zero. ∎

Example 7: Set $\phi(x)=|x|^{-\alpha}$ with $x\in R^n$. Then for $\alpha<n/2$ we have that $\phi\in L_2^{\mathrm{loc}}(R^n)$ and $\nabla\phi=-\alpha x|x|^{-\alpha-2}$ is also in $L_2^{\mathrm{loc}}(R^n)$ for $\alpha<n/2-1$. $\beta(x)=-2\alpha x|x|^{-2}$ is in $L_2^{\mathrm{loc}}(R^n)$ for $n\geq 3$. In this case we find $V=\frac{1}{2}\nabla\beta+\frac{1}{4}\beta^2$ which gives $V(x)=-\alpha(n-2-\alpha)|x|^{-2}$. Hence by Theorem 2.3 we have that, for $\alpha<n/2-1$ and $n\geq 3$, $d\mu=\phi^2dx$ is admissible. $\alpha<n/2-1$ is equivalent to $\lambda=\alpha(n-2-\alpha)<\frac{1}{4}(n-2)^2$, so by Theorem 2.3 we have for $f\in C_0^1(R^n)$ that

$$E(\phi^{-1}f,\ \phi^{-1}f)=\int_{R^n}|\nabla f|^2\,dx-\lambda\int_{R^n}|x|^{-2}\,|f|^2\,dx$$

for $n\geq 3$ and $\lambda<\frac{1}{4}(n-2)^2$. Since the form above is the restriction of E to $C_0^1\times C_0^1$ it is bounded below, hence it is obviously also bounded below for $\lambda=\frac{1}{4}(n-2)^2$, which is a classic inequality of Rellich.

If, however, $n/2 - 1 \leq \alpha < n/2$, then we still have that $\phi \in L_2^{\mathrm{loc}}(R^n)$ and $\nabla\phi \in L_2^{\mathrm{loc}}(R^n - \{0\})$. Hence by Theorem 2.4 $d\mu = \phi^2 dx$ is still admissible, and in this case $\frac{1}{4}(n-2)^2 \leq \lambda < \frac{1}{4}(n-2)^2 + n/2 - 1$. By Theorem 2.4 we get that the formula above for the energy form still holds but now only for $f \in C_0^2(R^n - \{0\})$. Hence we find the rather interesting result that for any n the form

$$\int_{R^n} |\nabla f|^2 - \lambda \int_{R^n} |x|^{-2} |f|^2 dx$$

defined for $\lambda < \frac{1}{4}(n-2)^2 + n/2 - 1$ and for $f \in C_0^2(R^n - \{0\})$ is bounded below and in fact it is the restriction of the self-adjoint operator $\phi H \phi^{-1}$ to $C_0^2(R^n - \{0\})$. For other discussions of the above form see Refs. 30, 34, 35 (Chap. X), 36, 37, and 38.

Example 8: Let $D \subset R^n$ be open and set $\phi(x) = 1$ for $x \in D$ and $\phi(x) = 0$ for $x \notin D$. Then $\phi \in L_2^{\mathrm{loc}}(R^n)$ and $\nabla\phi \in L_2^{\mathrm{loc}}(R^n - \partial D)$ and since ∂D is a closed set of measure zero we have that $d\mu = \phi^2 dx$ is admissible and obviously for any $f \in C_0^1(R^n)$

$$E(f,f) = \int_D |\nabla f|^2 dx \quad \text{while} \quad (f,f) = \int_D |f|^2 dx,$$

from which it follows that H is the Laplacian with Neumann boundary conditions in $L_2(D)$.

Example 9: Consider the measure $d\mu(x) = \phi(x)^2 dx$ on the real line, where $\phi(x)$ is the solution of the equation

$$(-\Delta + gV)\phi = 0,$$

with $V(x) = |x|^{-\alpha}$, $\alpha > 0$ for $x \neq 0$, and g a positive constant. As α grows we have interactions of increasing singularity. We have, for $\alpha \neq 2$,

$$\phi(x) = a|x|^{1/2} I_{|2-\alpha|^{-1}}\left(\frac{2\sqrt{g}}{|2-\alpha|}\,|x|^{(2-\alpha)/2}\right) + b|x|^{1/2} K_{|2-\alpha|^{-1}}\left(\frac{2\sqrt{g}}{|2-\alpha|}|x|^{(2-\alpha)/2}\right)$$

where I and K are Bessel functions, and for $\alpha = 2$,

$$\phi(x) = a|x|^{1/2+\sqrt{g+1/4}} + b|x|^{1/2-\sqrt{g+1/4}}.$$

We list the results of Theorems 2.1–2.4 (which all imply that μ is admissible) and their applicability to the present example in Table I.

The potentials of Example 9 are standard examples of singular perturbations, of e.g., Ref. 27. We shall see that the corresponding dynamics may be expressed directly in terms of stochastic processes.

Theorem 2.6: Let $d\mu(x) = \phi^2(x)dx$ with $\phi(x)$ bounded below by a positive constant on each compact, and assume that $\nabla\phi$ and $\Delta\phi$ are in $L_2^{\mathrm{loc}}(R^n)$, where $\nabla\phi$ and $\Delta\phi$ are the distributional derivatives of ϕ. Then μ is admissible and $\phi^{-1} C_0^\infty(R^n) \subset D(H)$ where H is the corresponding energy operator. For any $f \in \phi^{-1} C_0^\infty(R^n)$, H has the form

$$Hf = -\Delta f - \beta \cdot \nabla f.$$

Assume now, in addition, that $\beta = 2\phi^{-1}\nabla\phi$ is in $L_4^{\mathrm{loc}}(R^n)$ and satisfies the condition $\nabla\beta = V_1(x) + V_2(x)$ with $V_2 \in L_p(R^n)$ with $p \geq 2$ for $n \leq 3$, $p > 2$ for $n = 4$ and $p \geq n/2$ for $n \geq 5$, and $V_1 \in L_2^{\mathrm{loc}}(R^n)$ with $V_1(x) \geq -cx^2 - d$, for some constants c and d. Then H is essentially self-adjoint on $\phi^{-1} C_0^\infty(R^n)$.

Proof: By the fact that ϕ is bounded below on compacts and $\nabla\phi$ as well as $\Delta\phi$ are in $L_2^{\mathrm{loc}}(R^n)$ we have that $\nabla\phi^{-1} \in L_2^{\mathrm{loc}}(R^n)$ and $\Delta\phi^{-1} = \phi^{-1}V$ where $V = \frac{1}{2}\nabla\beta + \frac{1}{4}\beta^2$ is in $L_2^{\mathrm{loc}}(R^n)$. From this we also get that $\beta = 2\phi^{-1}\nabla\phi \in L_2^{\mathrm{loc}}(R^n)$. So by Theorem 2.3 μ is admissible and $\phi^{-1}C_0^\infty(R^n)$ is in $D(H^{1/2})$, moreover we have

$$E(\phi^{-}f, \phi^{-1}f) = \int_{R^n} |\nabla f|^2 dx + \int_{R^n} Vf^2 dx = \int_{R^n} f(-\Delta + V) f\, dx.$$

Now under the assumptions of the Theorem $-\Delta + V$ is essentially self-adjoint on $C_0^\infty(R^n)$. This follows from the Faris–Lavine theorem (Theorem X.38 [35]) and its corollary. Hence the form above has a unique closure so that

$$\phi H \phi^{-1} = -\Delta + V$$

TABLE I.

Conditions	Result on H	Example 9
A. $d\mu = \phi^2 dx$ (absolute equivalence)		
Theorem 2.4: $\phi \in L_2^{\mathrm{loc}}$, $\phi \in L_2^{\mathrm{loc}}(R^n - N)$	$H\Big\vert_{C_0^2(R^n - N)} = -\Delta - \beta \cdot \nabla$	$\alpha > 2$ with $a = 0$ $\alpha = 2$ with $b = 0$ if $g \geq \frac{3}{4}$ $\alpha < 2$
Theorem 2.2.1: $\phi \in L_2^{\mathrm{loc}}$, $\nabla\phi \in L_2^{\mathrm{loc}}$ Theorem 2.1: $\phi \in L_2^{\mathrm{loc}}$, $\nabla\phi \in L_2$	$H\Big\vert_{C_0^2(R^n)} = -\Delta - \beta \cdot \nabla$	$\alpha > 2$ with $a = 0$ $\alpha = 2$ " $b = 0$ $\alpha < 2$ " $b \neq 0$
B. $d\mu \sim dx$ (equivalence)		
Theorem 2.3: $\phi, \nabla\phi, \phi^{-1}\nabla\phi \in L_2^{\mathrm{loc}}$	$E(\phi^{-1}f, \phi^{-1}f) = \int [(\nabla f)^2 + Vf^2]\, dx$, $f \in C_0^1$	$\alpha < 1$ with $b \neq 0$
Theorem 2.2.2: $\phi, \nabla\phi, \phi^{-1}\nabla\phi, \phi^{-1}\nabla\phi \in L_2^{\mathrm{loc}}$	$H\phi^{-1}\Big\vert_{C_0^2} = \phi^{-1}(-\Delta + V)$	
	with $V = \phi^{-1}\Delta\phi = \frac{1}{4}\beta \cdot \beta + \frac{1}{2}\nabla \cdot \beta$	$\alpha < \frac{1}{2}$ with $b \neq 0$

as a self-adjoint operator. However $-\Delta+V$ is essentially self-adjoint on C_0^∞ so that H is essentially self-adjoint on $\phi^{-1}C_0^\infty$. Moreover, the identity $Hf=-\Delta f - \beta\nabla f$, $f=\phi^{-1}g$ follows from the fact that $H\phi^{-1}g = \phi^{-1}(-\Delta+V)g$, together with the definition of V. This proves the theorem. ∎

Let $\Omega\subset R^n$ be an open subset of R^n. Let μ and ν be positive regular σ finite Borel measures on Ω and μ absolutely continuous with respect to ν. For f,g in $C^1(\Omega)$ we define the energy form given by μ by

$$E(f,g)=\int_\Omega \nabla f\cdot\nabla g\,d\mu. \tag{2.15}$$

We shall say that μ is ν admissible if the form (2.15) is closable in $L_2(\Omega,d\nu)$. We then have the following generalization of Theorem 2.4.

Theorem 2.7: Let μ be absolutely continuous with respect to ν, such that $d\mu=\phi^2dx$ and $d\nu=\psi^2dx$ in Ω, with ϕ,ψ, and $\phi\psi^{-1}$ in $L_2^{\text{loc}}(\Omega,dx)\equiv L_2^{\text{loc}}(\Omega)$, where dx is the Lebesgue measure in Ω. If there is a closed set $N\subset\Omega$ of zero Lebesgue measure and $\nabla\phi$ as well as $\phi\cdot(\nabla\psi/\psi)$ are in $L_2^{\text{loc}}(\Omega-N)$, then μ is ν admissible, i.e., the form $E(f,g)=\int_\Omega\nabla f\cdot\nabla g\,d\mu$ as defined on $C_0^1(\Omega)\times C_0^1(\Omega)$ is closable in $L_2(d\nu)$. Let H be the operator associated with the closure of this form. If $\beta=2\phi\nabla\phi/\psi^2$ is in $L_2^{\text{loc}}(\Omega-N)$, then for $f\in C_0^2(\Omega-N)$ we have that

$$Hf=-\sigma\,\Delta f-\beta\cdot\nabla f,$$

where $\sigma=(d\mu/d\nu)\cdot(\phi^2/\psi^2)$.

Proof: Let $f\in C_0^1(\Omega)$ and $g\in C_0^1(\Omega-N)$, then

$$E(f,f)=\int\left(\frac{\phi}{\psi}\nabla f\right)\cdot\left(\frac{\phi}{\psi}\nabla f\right)\psi^2\,dx$$

so that $E(f,f)$ is the square of the graph norm of the operator $f\to\phi\psi^{-1}\nabla f$ in $L_2(d\nu)$. So E being closable in $L_2(d\nu)$ is equivalent with $\phi\psi^{-1}\nabla$ defined on $C_0^1(\Omega)$ being a closable operator. We find as in Theorem 2.4 that

$$-\nabla-\left(\frac{\nabla\phi}{\psi}+\frac{\phi}{\psi}\,\frac{\nabla\psi}{\psi}\right)$$

on $C_0^1(\Omega-N)$ is a densely defined adjoint and therefore $\phi\psi^{-1}\nabla$ must be closable. The formula for H follows by direct calculation. This proves the theorem. ∎

3. DIFFUSION PROCESSES IN R^n

As in the final part of Sec. 2, let Ω be an open subset of R^n and let $d\mu<d\nu<dx$, i.e., $d\mu=\phi^2dx$, $d\nu=\psi^2dx$ with $\phi\psi^{-1}\in L_2^{\text{loc}}(\Omega)$, such that μ is ν admissible in Ω. Let $\exp(-tH)$ be the strongly continuous semigroup generated by H, where H is the self-adjoint operator associated with the closure of the energy form $E(f,g)=\int\nabla f\cdot\nabla g\,d\mu$ in $L_2(\Omega,d\nu)$. We know that $\exp(-tH)$ is a strongly continuous contraction semigroup in $L_2(d\nu)$. $\operatorname{Exp}(-tH)$ is called Markov if $0\le f\le 1$ implies $0\le\exp(-tH)f\le 1$. A nonnegative symmetric bilinear form ϵ on $L_2(d\nu)$ is called Markov if for any $\delta>0$ there exists a nondecreasing function $\phi_\delta(t)$, $t\in R$ with $\phi_\delta(t)=t$ for $0\le t\le 1$, $|\phi_\delta(t)|\le|t|$, and $-\delta\le\phi_\delta(t)\le 1+\delta$ for all $t\in R$, such that for any function $f\in D(\epsilon)$, where $D(\epsilon)$ is the domain of ϵ, we also have $\phi_\delta(f)\in D(\epsilon)$ and

$$\epsilon(\phi_\delta(f),\phi_\delta(f))\le\epsilon(f,f). \tag{3.1}$$

Theorem 3.1 (*Fukushima*): If ϵ is Markov and closable then the self-adjoint operator associated with the closed form is the infinitesimal generator for a semigroup $\exp(-tH)$ which is Markov. Moreover, if a strongly continuous semigroup $\exp(-tH)$ is Markov, then the closed form defined by its generator is Markov.

For the proof of this theorem see Ref. 9.

Theorem 3.2: Let Ω be an open subset of R^n and let $d\mu<d\nu<dx$, such that μ is ν admissible. Then the energy form $E(f,f)=\int\nabla f\cdot\nabla f\,d\mu$ defined on $C_0^1(\Omega)$ is a closable Markov form. So the energy operator H given by the closure of E in $L_2(d\nu)$ is the infinitesimal generator of a strongly continuous Markov semigroup $\exp(-tH)$.

Proof: Let $\Psi_\delta\in C_0^1(R)$ be such that $0\le\Psi_\delta\le 1$ and $\Psi_\delta(t)=1$ for $0\le t\le 1$, $\Psi_\delta(t)=0$ for $t\in(-\delta,1+\delta)$. Note that for any $\delta>0$ there exists such a Ψ_δ. Let $\phi_\delta(t)=\int_0^t\Psi_\delta(\tau)\,d\tau$. We then have for any $f\in C_0^1(R^n)$ that

$$\begin{aligned}E(\phi_\delta(f),\phi_\delta(f))&=\int|\Psi_\delta(f)|^2\,|\nabla f|^2\,d\mu\\&\le\int|\nabla f|^2\,d\mu,\end{aligned}$$

which proves that E is Markov. The rest of the theorem follows from Theorem 3.1. ∎

We shall now assume that the measure ν is everywhere dense in Ω, i.e., that $\psi(x)>0$ a.e. with respect to the Lebesgue measure, where $d\nu=\psi^2dx$. A regular Dirichlet form in the sense of Fukushima[9] is a closed Markov form such that $D(\epsilon)\cap C_0(\Omega)$ is dense in $D(\epsilon)$ in the $\epsilon(f,f)+(f,f)$ norm and also closed in $C_0(\Omega)$ in the supremum norm, where $C_0(\Omega)$ are the continuous functions in Ω which vanish at $\partial\Omega$ or at infinity. Since obviously $C_0^1(\Omega)$ is dense in $C_0(\Omega)$ and also by definition in $D(\bar E)$, the domain of the closed energy form, we have that the closure $\bar E$ of E is a regular Dirichlet form.

Lemma 3.3: The closure $\bar E$ of E is a regular Dirichlet form in the sense of Fukushima. ∎

Let A be open in Ω. We define the capacity of A by

$$\operatorname{Cap}(A)=\begin{cases}\inf_{u\in L_A}E(u,u)+(u,u), & \text{if } L_A\ne\emptyset,\\ \infty & \text{if } L_A=\emptyset,\end{cases}$$

where $L_A=\{u\in D(E);\ u\ge 1\ \nu\text{-a.e. on } A\}$.

The capacity of any subset B of Ω is then defined as

$$\operatorname{Cap}(B)=\inf\{\operatorname{Cap}(A);\ B\subset A \text{ and } A \text{ is open}\}. \tag{3.2}$$

$\operatorname{Cap}(B)$ is a strongly subadditive Choquet capacity.[39]

We say that a set B in Ω is polar if it is a Borel set and $\operatorname{Cap}(B)=0$.

For any Borel set $A\subset\Omega$, we denote by $\mathcal{B}(A)$ the σ-algebra of Borel subsets of A, while $B(A)$ will stand for the space of all $\mathcal{B}(A)$ measurable bounded functions in A.

Suppose now that we have a Markov process $\xi_t(\omega)$ with state space $(Y,\mathcal{B}(Y))$ where Y is some Borel subset of Ω. We adjoin a "death" point ∂ to Y and regard $Y\cup\partial$ as a topological subspace of the one-point compactification $\Omega\cup\partial$ of Ω. The transition semigroup given by the process ξ_t will be denoted by P_t: $P_tf(x)=E_x(f(\xi_t))$ for

$x \in Y$ and $f \in B(Y)$. Following Fukushima,[9] the Markov process ξ_t is then said to be properly associated with the energy form E if

$$\mathrm{Cap}(\Omega - Y) = 0 \tag{3.3}$$

and $P_t f$ is a quasicontinuous version of $\exp(-tH)f$ for each $f \in L_2(d\nu) \cap B(Y)$ and $t > 0$. A function g is said to be quasicontinuous if it is defined in the complement of a set of capacity zero and if for any $\delta > 0$ there is an open set G of $\mathrm{Cap}(G) < \delta$ such that the restriction of g to $\Omega - G$ is continuous and continuously extendable to $\Omega \cup \partial - G$ by setting $g(\partial) = 0$.[9]

We now have the following theorem, as an immediate consequence of a corresponding theorem of Fukushima.

Theorem 3.4: If μ is ν admissible then there exists a Hunt process ξ_t properly associated with E ∎

For the definition of a Hunt process see Ref. 3, Chapter XIV.

A regular Dirichlet form ϵ is said to be local if for any f_1 and f_2 in $D(\epsilon) \cap C_0(\Omega)$ we have that $\epsilon(f_1, f_2) = 0$ whenever f_1 and f_2 have disjoint supports. It is easy to see that the closure of E is local. It follows then from a theorem of Fukushima and Silverstein (see Refs. 28 and 11, Theorem 11.10, p. 124) that the process ξ_t of the previous theorem can be so chosen that it has continuous trajectories on $Y \cup \partial$. We therefore have

Theorem 3.5: If μ is ν admissible then the process ξ_t of the previous theorem has continuous trajectories on $Y \cup \partial$, or equivalently $\xi_t(\omega)$ is continuous in Y for almost all ω and for $0 \leq t < \zeta$, where ζ is the lifetime of the trajectory, i.e.,

$$\zeta(\omega) = \inf\{t > 0 \text{ such that } \xi_t(\omega) = \partial\}. \qquad ∎$$

It follows from the positivity and symmetry of $\exp(-tH)$ that it extends to contraction semigroups in $L^p(d\nu)$ for $1 \leq p \leq \infty$ which are strongly continuous contraction semigroups in the cases $1 \leq p < \infty$. We shall say that the energy form E is ergodic if the corresponding semigroup $\exp(-tH)$ is ergodic. We now have the following lemma which is a direct consequence of Ref. 11, corollary 1.5, p. 12.

Lemma 3.6: If E is ergodic and $f, g > 0$ ν-almost everywhere and in $L_1(d\nu)$, then

$$\{x : (H^{-1}f)(x) < \infty\} = \{x : (H^{-1}g)(x) < \infty\},$$

where H is the infinitesimal generator of $\exp(-tH)$ in L_1. ∎

If E is ergodic we say that the energy form E is *transient* if $H^{-1}f$ is finite ν-almost everywhere for all f in $L_1(d\nu)$ and *recurrent* if $H^{-1}f = \infty$ ν-a.e. for all f in $L_1(d\nu)$ and *recurrent* if $H^{-1}f = \infty$ ν-a.e. for all $f \geq 0$ and nontrivial in $L_1(d\nu)$.

We have the following theorem which follows from Silverstein (Ref. 11, Theorem 1.6, p. 13).

Theorem 3.7: If E is ergodic and recurrent, then $\exp(-tH)1 = 1$ for all $t \geq 0$. ∎

We introduce now the extended space $D^e(E)$ by saying that $f \in D^e(E)$ if there exists a sequence $f_n \in D(E) = D(H^{1/2})$ such that f_n is Cauchy in $E(f,f)^{1/2}$ and $f_n \to f$ ν-a.e. on Ω, and we have the following lemma, a consequence of results of Silverstein (Ref. 11, Lemma 1.7, p. 15).

Lemma 3.8: Let $f \in D^e(E)$ and let f_n be as above. Then $\lim_{n\to\infty} E(f_n, f_n)$ exists and is independent of the choice of f_n. We may therefore extend E to $D^e(E)$ by setting $E(f,f) = \lim_{n\to\infty} E(f_n, f_n)$. We then have that $D(E) = D^e(E) \cap L_2(d\nu)$. ∎

We now have the following theorem (Ref. 11, Lemma 1.8, p. 16, Theorem 8.6, p. 93).

Theorem 3.9: If E is ergodic and recurrent, then $1 \in D^e(E)$ and $E(1,1) = 0$, and if E is transient, then $\lim_{t\to\infty} \exp(-tH) = 0$ in the strong operator topology in $L_2(d\nu)$. Moreover if E is transient then $D^e(E)$ is a Hilbert space, and if E is recurrent then $D^e(E)$ is not a Hilbert space. ∎

From Propositions 4.16 and 4.5 of Ref. 11, p. 58 and p. 43, we have the following two theorems.

Theorem 3.10: If E is recurrent, then if M is a Borel set of positive capacity we have that

$$\Pr\{\xi_t(\omega) \in M,\ t \geq n \mid \xi_0(\omega) = x\} = 1$$

for all $x \notin \Omega$ outside a set of capacity zero, and for all n. ∎

Theorem 3.11: If M is polar, then

$$\Pr\{\xi_t(\omega) \in M \text{ for some finite } t \mid \xi_0(\omega) = x\} = 0$$

for all x outside some polar set. ∎

4. THE STOCHASTIC DIFFERENTIAL EQUATION

Let Ω be an open subset of R^n. Let μ be a probability measure on Ω. In relation to Theorem 2.7 we assume, for simplicity of notation, $\mu = \nu$. Otherwise we assume the same as in Theorem 2.7, i.e., $d\mu = \phi^2 dx$ with ϕ in $L_2^{\mathrm{loc}}(\Omega, dx)$ and such that there exists a closed set $N \subset \Omega$ of zero Lebesgue measure such that $\nabla\phi$ is in $L_2^{\mathrm{loc}}(\Omega - N)$. Then we know by Theorem 2.7 that μ is admissible. Let H be the energy operator, i.e., the operator associated with the closure of the form $E(f,g) = \int_\Omega \nabla f \cdot \nabla g \, d\mu$ as defined on $C_0^1(\Omega) \times C_0^1(\Omega)$ in $L_2(\Omega, d\mu)$. We also assume that $\beta = 2\nabla\phi/\phi$ is $L_2(\Omega, d\mu)$.

Since μ is a probability measure we have $1 \in L_2(\Omega, d\mu)$ and $\exp(-|t|H_1) = 1$, so that the corresponding process may be taken as a homogeneous Markov process ξ_t, $-\infty < t < \infty$, with invariant measure μ.

Let $L_2(d\omega)$ be the L_2 space of all the L_2 functions which are measurable with respect to the process, i.e., with respect to the functions $\xi_t(\omega)$. The time translation $\xi_t(\omega) \to \xi_{t+s}(\omega)$ induces a measurable transformation on the ω space which leaves invariant the probability measure $d\omega$. Hence it generates a strongly continuous unitary group U_t on $L_2(d\omega)$. Let $f \in L_2(d\mu)$, then $f(\xi_t) \in L_2(d\omega)$ with $\int |f(x)|^2 d\mu = E(f^2(\xi_t))$ and $U_s f(\xi_t) = f(\xi_{t+s})$. Moreover for $f \in D(H)$ we have that $f(\xi_t)$ is in the domain of the infinitesimal generator of U_t and we have

$$\frac{d}{dt} f(\xi_t) = -(Hf)(\xi_t), \tag{4.1}$$

where the derivative is in the strong $L_2(d\omega)$ sense.

Now let

$$\chi_s \equiv -\int_0^s \beta(\xi_\tau)\,d\tau, \tag{4.2}$$

then χ_s is strongly $L_2(d\omega)$ differentiable, $\beta(\xi_\tau)$ being strongly $L_2(d\omega)$ continuous. Let $\chi_s^h \equiv \chi_{s+h} - \chi_s$, then by Taylor's theorem we have, for $f\in C_0^2(\Omega)$,

$$|f(\xi_t+\chi_{s+h}) - f(\xi_t+\chi_s) - \nabla f(\xi_t+\chi_s)\chi_s^h| \\ \le \tfrac{1}{2}\,\|f\|_{2,\infty}|\chi_s^h|^2, \tag{4.3}$$

where $\|f\|_{2,\infty}$ is the norm in $C_0^2(\Omega)$. We now have that $h^{-1}\chi_s^h$ converges to $-\beta(\xi_s)$, strongly in $L_2(d\omega)$ as $h\to 0$, thus $\nabla f(\xi_t+\chi_s)(1/h)\chi_s^h$ converges strongly in $L_2(d\omega)$ to $-\nabla f(\xi_t+\chi_s)\cdot\beta(\xi_s)$, as $h\to 0$. On the other hand, $h^{-1}|\chi_s^h|^2$ converges strongly to zero in $L_1(d\omega)$, because $h^{-1}\chi_s^h$ is uniformly $L_2(d\omega)$ bounded and χ_s^h converges strongly to zero in $L_2(d\omega)$ as $h\to 0$. By (4.3) we then obtain that $f(\xi_t+\chi_s)$ as a function of s from $[0,\infty)$ into $L_1(d\omega)$ is strongly differentiable, with derivative given by $-\nabla f(\xi_t+\chi_s)\cdot\beta(\xi_s)$.

Now let F_t be the conditional expectation with respect to the past of t. Conditioned with respect to the past of t, $f(\xi_{t+h}+\chi_t)$ is a function of ξ_{t+h} alone for $h>0$, because χ_t is sure if conditioned with respect to the past of t. From this it easily follows that, for $h>0$,

$$F_t[f(\xi_{t+h}+\chi_t)] \tag{4.4}$$

is strongly $L_2(d\omega)$ differentiable with respect to h, with forward derivative at zero equal to

$$F_t[-(Hf)(\xi_t+\chi_t)]. \tag{4.5}$$

Since strong $L_2(d\omega)$ differentiability implies strong $L_1(d\omega)$ differentiability, $d\omega$ being a probability measure, we have that (4.4) also has a strong forward $L_1(d\omega)$ derivative at $h=0$ given by (4.5). Hence $F_t[f(\xi_{t+h}+\chi_{t+h})]$ has a strong $L_1(d\omega)$ forward derivative at $h=0$, by the results above and Leibnitz' rule of strong differentiation, with derivative given by

$$F_t[-(Hf)(\eta_t) - \beta(\xi_t)\cdot(\nabla f)(\eta_t)], \tag{4.6}$$

with $\eta_t \equiv \xi_t + \chi_t$.

Recalling the argument preceding formula (4.5) we have that Hf is to be computed by considering f as the function $f(x+\chi_t)$, with χ_t sure, so that Hf is actually given, by the results of the previous section, as $-\Delta f - \beta(x)\nabla f$, where β is to be evaluated at the point x. Hence we have that, in (4.6), $(Hf)(\eta_t) = -(\Delta f)(\eta_t) - \beta(\xi_t)\cdot\nabla f(\eta_t)$. Thus (4.6) takes the form

$$F_t[(\Delta f)(\eta_t)], \tag{4.7}$$

where

$$\eta_t = \xi_t - \int_0^t \beta(\xi_s)\,ds. \tag{4.8}$$

Hence we have proven that, if $\eta_t(\omega)$ is defined by (4.8), the integral being understood in the strong $L_2(d\omega)$ sense, then for any $f\in C_0^2(\Omega)$ we have that $f(\eta_t)$ has a strong forward derivative in the sense that $F_t[f(\eta_{t+h})]$ has a strong one-sided $L_1(d\omega)$ derivative at $h=0$, which is equal to $F_t[(\Delta f)(\eta_t)]$. By (4.8) we have that F_t is also the conditional expectation given by the past of the process $\eta_t(\omega)$.

Let us now assume that $\Omega = R^n$ and consequently that $\mu=\nu$ is a probability measure which is admissible in the sense of Sec. 2. Since by assumption $\beta\in L_2(R^n)$ we easily see that $S(R^n)\subset D(H)$, where $S(R^n)$ is the Schwartz test function space. It follows then from the argument above that (4.7) holds also for all $f\in S(R^n)$, if $\beta\in L_2(d\mu)$. An obvious extension of Lemma 3.1 of Ref. 16 then gives us that η_t is the Brownian motion w_t with initial distribution given by μ. Hence in this case we have that ξ_t satisfies the stochastic differential equation

$$d\xi_t = \beta(\xi_t)\,dt + dw_t, \tag{4.9}$$

where w_t is the standard Brownian motion in R^n. We can now state the following theorem.

Theorem 4.1: Let $d\mu(x)=\phi(x)^2\,dx$ be a probability measure on R^n such that μ is admissible and such that $\beta(x)\equiv 2\nabla\ln\phi(x)$ is in $L_2(d\mu)$ or equivalently, $\nabla\phi$ is in $L_2(dx)$. Then there exists a solution η_t of the stochastic differential equation

$$d\xi_t = \beta(\xi_t)\,dt + dw_t,$$

where w_t is the standard Brownian motion in R^n, with continuous paths, such that ξ_t is a Hunt process. Moreover if $\phi(x)$ satisfies the conditions of Theorem 2.6, then there is only one nonanticipating solution of the above stochastic equation, in the class of Markov processes with invariant distribution μ.

Proof: We have already proven the first part of the theorem; we shall now also prove the remaining "moreover part" of the theorem. Let ξ_t be a Markov process which solves the stochastic equation with the prescribed initial distribution $d\mu$. Then by the fact that ξ_t solves the stochastic equation we get

$$\xi_t = \int_0^t \beta(\xi_s)\,ds + w_t, \tag{4.10}$$

where w_t is the Brownian motion with initial distribution μ. Since μ is an invariant measure for ξ_t, the above integral can be taken in the strong $L_2(d\omega)$ sense, as in the proof of the first part of the theorem. For $f\in S(R^n)$, by an argument similar to the one preceding the theorem, we have that $f(\int_0^t\beta(\xi_s)\,ds + w_t)$ has a forward derivative in the strong $L_1(d\omega)$ sense, because by an adaptation of Ito's results, ξ_s being by assumption a nonanticipating solution, we have that

$$F_t[f(\int_0^t \beta(\xi_s)\,ds + w_{t+h})] \tag{4.11}$$

is strongly $L_2(d\omega)$ differentiable with respect to h at $h=0$, with derivative

$$F_t[(\Delta f)(\int_0^t \beta(\xi_s)\,ds + dw_t)],$$

where F_t is the conditional expectation with respect to the past of t. Then, we get, as in the argument preceding the theorem, that $F_t[f(\xi_{t+h})]$ has a right-derivative given by

$$F_t[\Delta f(\xi_t) + \beta(\xi_t)\cdot\nabla f(\xi_t)] = \Delta f(\xi_t) + \beta(\xi_t)\nabla f(\xi_t). \tag{4.12}$$

Hence the Markov semigroup $P_t f = E_0[f(\xi_t)]$, where E_0 is the conditional expectation with respect to ξ_0, has $-\Delta - \beta\cdot\nabla$ as an infinitesimal generator on $S(R^n)$. However by Theorem 2.6 we have that the only semigroup with such a property is the one given by the energy form and it is constructed in Sec. 2. The process associated with this semigroup is the one described in Sec. 3. Thus we have $P_t = \exp(-tH)$, where H is the energy operator given by μ. This completes the proof of the theorem. ∎

Remark: Let us now make some brief comments in relation with Ref. 25 (also see Ref. 19). Theorem 4.1 implies, in particular, that under the stated general assumptions the weak $L_2(R^n, d\mu)$ solution $\psi(t,x)$ of the heat equation

$$\frac{\partial}{\partial t}\psi(t,x) = \Delta\psi(t,x) - V\psi(t,x) \tag{4.13}$$

with initial condition

$$\psi(0,x) = f(x) \tag{4.14}$$

in $L_2(R^n, d\mu)$ is given in terms of the process ξ_t by $f(\xi_t)(\cdot)$, in the sense that for any $g \in L_2(R^n, d\mu)$ we have

$$\int g(x)\psi(t,x)\, d\mu(x) = E(g(\xi_0)f(\xi_t)) \tag{4.15}$$

and

$$\frac{d}{dt}E(g(\xi_0)f(\xi_t)) = -E(g(\xi_0)(Hf)(\xi_t))$$
$$= -\int g(x)(H\psi)(t,x)\, d\mu(x), \tag{4.16}$$

where E is the expectation in $L_2(\Omega, d\omega)$. The solution of the heat equation is thus given by integrals over the sample paths of the process ξ_t, related to the Brownian motion w_t by the stochastic equation

$$d\xi_t = \beta(\xi_t)\, dt + dw_t.$$

In the pictorial language of Ref. 25 we have an expression of the solutions of the Schrödinger equation as averages with respect to "distorted Brownian paths." Let us recall that our result holds under the only assumptions that the measure $d\mu(x) = \phi(x)^2 dx$ is admissible (see Theorems 2.1–2.4 for conditions for this to happen) and that the drift coefficient β is in $L_2(R^n, d\mu)$, i.e., equivalently that $\nabla\phi \in L_2(R^n, dx)$. Thus we need neither local Hölder continuity nor restriction to linear growth at infinity for β (take, e.g., $\phi = \exp(-|x|^\alpha)$, $\alpha > 1$). In our case the potential $V = \frac{1}{4}\beta^2 + \nabla\beta = \Delta\phi/\phi$ need not necessarily exist as a measurable function. On the other hand, also strongly growing potentials like, e.g., lower bounded polynomials are allowed. In the cases where the potential V exists as a measurable function and is such that the Feynman–Kac formula holds, we have of course that (4.12) is also given by

$$E_w\left(\exp\left[-\int_0^t V(w_\tau)\, d\tau\right]\right)^{-1} E_w(g(w_0)f(w_t)$$
$$\times \exp\left[-\int_0^t V(w_\tau)\, d\tau\right]),$$

where E_w is expectation with respect to the Wiener process w_t. We recall that sufficient conditions for the Feynman–Kac formula to hold are, e.g., $V \in (L_p + L_\infty) \times (R^n, dx)$ for some $p > n/2$ if $n \neq 3$ and $p = 2$ for $n = 3$, or $V \geq 0$ and $V \in L_1^{\mathrm{loc}}(R^n, dx)$, see, e.g., Refs. 35 (Chapter X, p. 279) and 40.

ACKNOWLEDGMENT

It is a real pleasure for the two first-named authors to thank the Zentrum für interdisziplinäre Forschung at the University of Bielefeld for the kind hospitality which has made this work possible.

*Work supported in part by the Norwegian Research Council for Science and the Humanities.

[1]E.B. Dynkin, *Markov Processes I, II* (Springer, Berlin, 1965) (transl.).
[2]P.A. Meyer, *Probability and Potentials* (Blaisdell, Waltham, 1966); C. Dellacherie and P.A. Meyer, *Probabilités et potentiel* (Hermann, Paris, 1975).
[3]P.A. Meyer, *Processus de Markov* (Springer, Berlin, 1967).
[4]R.M. Blumenthal and R.K. Getoor, *Markov Processes and Potential Theory* (Academic, New York, 1968).
[5]K. Ito and H.P. McKean, Jr., *Diffusion Processes and Their Sample Paths* (Springer, Berlin, 1965).
[6]I.I. Gihman and A.V. Skorohod, *Stochastic Differential Equations* (Springer, Berlin, 1972) (transl.).
[7]H.P. McKean, Jr., *Stochastic Integrals* (Academic, New York, 1969).
[8]M. Fukushima, J. Math. Soc. Jpn. **21**, 58–93 (1969).
[9]M. Fukushima, *Proceedings of the Second Japan—USSR Symposium on Probability Theory* (Lecture Notes in Mathematics 330), edited by G. Maruyama and Yu.V. Prokhorov (Springer, Berlin, 1973).
[10]M. Fukushima, Trans. Am. Math. Soc. **162**, 185–224 (1971).
[11]M.L. Silverstein, *Symmetric Markov Processes* (Lecture Notes in Mathematics 426) (Springer, Berlin, 1974).
[12]M.L. Silverstein, *Boundary Theory for Symmetric Markov Processes* (Lecture Notes in Mathematics 516) (Springer, Berlin, 1976).
[13]A. Beurling and J. Deny, Proc. Nat. Acad. Sci. USA **45**, 208–15 (1959).
[14]J. Deny, "Méthodes Hilbertiennes en théorie du potentiel," in *Potential Theory* (C.I.M.E., Ed. Cremonese, Rome, 1970), pp. 121–201.
[15]S. Albeverio and R. Høegh-Krohn, Colloq. Int. CNRS **248**, 11–59 (1976).
[16]S. Albeverio and R. Høegh-Krohn, "Dirichlet Forms and Diffusion Processes on Rigged Hilbert Spaces," Oslo University Preprint, December 1975 (to be published).
[17]S. Albeverio and R. Høegh-Krohn, "Hunt Processes and Analytic Potential Theory on Rigged Hilbert Spaces," ZiF preprint, to appear. See also the authors' contribution to the "IV. International Symposium on Information Theory," Leningrad, June 1976, Akad. Nauk SSSR, Moscow-Leningrad (1976).
[18]S. Albeverio and R. Høegh-Krohn, "Canonical Quantum Fields in Two Space—Time Dimensions," Oslo University Preprint, March 1976 (to be published).
[19]T. Hida and L. Streit, "On Quantum Theory in Terms of White Noise," ZiF preprint, University of Bielefeld, to appear.
[20]P. Courrège and P. Renouard, Astérisque **22–23**, 3–245 (1975).
[21]P. Priouret and M. Yor, Astérisque **22–23**, 247–90 (1975).
[22]E. Nelson, *Dynamical Theories of Brownian Motion* (Princeton U.P., Princeton, N.J., 1967).
[23]F. Guerra, "On stochastic field theory, Proc. II Aix-en-Provence Intern. Conf. Elem. Part., 1973," J. Phys. Suppl. T 34, Fasc. 11-12, Colloque, C-1-95-98.
[24]S. Albeverio and R. Høegh-Krohn, J. Math. Phys. **15**, 1745–7 (1974).
[25]H. Ezawa, J.R. Klauder, and L.A. Shepp, Ann. Phys. **88**, 588 (1974).
[26]F. Coester and R. Haag, Phys. Rev. **117**, 1137–1145 (1960); H. Araki, J. Math. Phys. **1**, 492–504 (1960).
[27]H. Ezawa, J.R. Klauder, and L.A. Shepp, J. Math. Phys. **16**, 783–99 (1975); H. Narnhofer and J.R. Klauder, J. Math. Phys. **17**, 1201–9 (1976).

[28]M. Fukushima, Z. Wahrsch. u.v. Geb., **29**, 1–6 (1974).
[29]W. Faris and R. Lavine, Commun. Math. Phys. **35**, 39–48 (1974).
[30]E. Nelson, J. Math. Phys. **5**, 332–43 (1964). (See also, e.g., Refs. 34 and 35.)
[31]F.A. Berezin and L.D. Faddeev, Sov. Math. Dokl. **2**, 372–5 (1961) (transl.)
[32]R.A. Minlos and L.D. Faddeev, Dokl. Akad. Nauk **141**, 1335–8 (1961) [Sov. Phys. Dokl. **6**, 1072–4 (1962)].
[33]J.G. Flamand, in *Cargèse Lectures in Theoretical Physics*, edited by F. Lurçat (Gordon and Breach, New York, 1967).
[34]W. Faris, *Self-Adjoint Operators* (Springer, Berlin, 1975).
[35]M. Reed and B. Simon, *Methods of Modern Mathematical Physics, II. Fourier Analysis, Self-Adjointness* (Academic, New York, 1975).
[36]T. Kato, in *Physical Reality and Mathematical Description*, edited by C.P. Enz and J. Mehra (Reidel, Dordrecht, 1974).
[37]C. Radin, J. Math. Phys. **16**, 544–7 (1975).
[38]H. Narnhofer, "Quantum Theory for $1/r^2$ – Potentials," Universität Wien Preprint.
[39]M. Fukushima, Trans. Am. Math. Soc. **162**, 185–224 (1971).
[40]S. Albeverio and R. Høegh-Krohn, J. Funct. Anal. **16**, 39–82 (1974).

Instability of the continuous spectrum: The N-band Stark ladder*

J. E. Avron[†‡] and J. Zak

Department of Physics, Technion-Israel Institute of Technology, Haifa, Israel

A. Grossmann

Centre de Physique Theorique, C.N.R.S., Marseille, France

L. Gunther[§]

Department of Physics, Tufts University, Medford, Massachusetts
(Received 11 December 1975; revised manuscript received 12 April 1976)

It is shown that the energy spectrum of the Bloch electron in an external field is *continuous*. Furthermore, it is shown that all approximations which take into account interband coupling within groups of finite number of bands (the N-band approximation) lead to a *pure-point* spectrum of intertwining Wannier–Stark ladders. This instability of the continuous spectrum under the N-band approximation is related to a theorem due to Weyl and von Neumann. Approximation methods for dealing with interband coupling within a group of finite number of bands are given.

1. INTRODUCTION

Weyl and von Neumann pointed out that the continuous spectrum is very unstable in that arbitrarily small perturbations can turn it into a pure point spectrum. On the contrary, the common belief is that physically meaningful perturbations do not cause such pathologies: They leave the continuous spectrum continuous and the discrete spectrum discrete. Instabilities of the continuous spectrum have been considered to be of mathematical interest having little or nothing to do with physics.

We shall show that a natural physical approximation widely used in solid state physics is of this nature: The approximation discards a presumably small part of the Hamiltonian, thereby changing the original continuous spectrum into a pure point spectrum. The example is taken from the theory of the one-dimensional Bloch electron in an external field. In Sec. 2 it is shown that the spectrum is absolutely continuous from $-\infty$ to ∞ and that it has no gaps. The absolute continuity of the spectrum follows from a general theorem in Sec. 2. This theorem generalizes a known result of Dunfold and Schwartz to potentials that are not necessarily monotonic at infinity.

It is an experimental fact that, even though the band index is no longer a constant of motion, the electron leaks out of the band very slowly for external fields that are not too strong. A natural approximation is to assume that the band index is a true constant of motion (and then correct perturbatively). This is the "single band approximation, " and it leads to the pathological character described in the Weyl—von Neumann theorem: The exact continuous spectrum is made a pure point spectrum by it.

The point spectrum (eigenvalues) of the single band approximation has been of considerable interest because it is related to the Wannier—Stark ladder,[1] which consists of an infinite set of eigenvalues with spacing Ea. E is the force field and a is the lattice spacing. In some sense, the Stark ladder is the analog of Landau levels in external magnetic fields. However, whereas the latter has at least a sound experimental basis, this is not the case for the Stark ladder which is more of a problem than a well-established effect.[2] The point spectrum is an artifact of the single band approximation since the exact spectrum is absolutely continuous. However, this *does not* prove that there is no Stark ladder effect in the sense of a periodic structure in the physical (optical, say) spectrum. In particular, this does not mean that the single band approximation is "large" or unphysical. On the contrary, the lesson of the Weyl—von Neumann theorem is that the spectrum, in its set theoretic sense, is a very sensitive object.

In Sec. 4 we discuss the N-band Hamiltonian. We prove that the spectrum consists of N intertwined Wannier—Stark ladders. Thus a finite interband interaction preserves the discrete spectrum. We also consider methods of approximations for the interband coupling. An analogy between the time-dependent Schrödinger equation and the N-band Hamiltonian is used to apply the adiabatic approximation to obtain the eigenvalues of the N-band Hamiltonian. This analogy also leads to a conservation law of probability in k.

2. THE MODEL

Consider the one-dimensional single particle Hamiltonian:

$$H=p^2/2m+V(x)+Ex, \quad x\in \mathbf{R}^1. \tag{1}$$

$V(x)$ is periodic and twice differentiable.

This Hamiltonian describes the motion of a charged particle in a periodic crystaline field and in an external constant electric (or gravitational) field. The crystal is assumed to be infinitely big. H is self-adjoint by standard arguments.[3]

The spectrum of H is the real energy axis; i.e., it is continuous stretching from $-\infty$ to ∞. Before proving this, let us consider the simpler Hamiltonian:

$$H_0=p^2/2m+Ex. \tag{2}$$

The unitary transformation

$$U=\exp[ip^3/6m] \tag{3}$$

transforms H_0 into the multiplication operator Ex. H_0 and Ex are thus unitarily equivalent so that in particular they have the same spectrum. The spectrum of an operator which is a multiplication by a function is the range of values this function assumes. In our case the spectrum of Ex, and hence of H_0, stretches continuously from $-\infty$ to ∞. Note that H_0 has no eigenvalues embedded in the continuous spectrum.[4]

It is perhaps physically obvious that the addition of a bounded periodic function to H_0 does not change the nature of the spectrum and in particular that the Hamiltonian (1) has no bound states.[5] Mathematically, a problem arises because the periodic potential remains finite at infinity. It may then happen that interference due to the wiggling of the potential produces bound states. Examples of such bound states were given by von Neumann and Wigner.[6]

The following theorem guarantees the absolute continuity of the spectrum for potentials that go to $-\infty$ in one direction (at least) with no assumption of monotonicity.[7]

Theorem: Let there be given the second order differential operator

$$-\frac{d^2}{dx^2}-q(x) \tag{4}$$

on the interval $[\alpha,\infty)$. Assume that:

(a) $q(x)$ is positive for x large enough,

(b) $\int^\infty[(q'/q^{3/2})'+\frac{1}{4}(q')^2q^{-5/2}]\,dx<\infty$,

(c) $\int^\infty q^{-1/2}\,dx=\infty$,

(d) $|q(x+a)|>|q(x)|$ for $a>a_0$.

Then the spectrum of any self-adjoint extension of the operator is entirely continuous and covers the whole real axis. In particular, this is the case for the Hamiltonian in Eq. (1).

The above theorem is a standard result in analysis, except for condition (d), which is customarily replaced by a much stronger condition of monotonicity of $q(x)$ (monotonicity is equivalent to $a_0=0$). The proof of this theorem is identical to the proof of Corrolary XIII. 6.21 in Dunford and Schwartz.[8] In other words, Dunford and Schwartz prove a stronger result than the one they state. We shall only point out that (d), or equivalently the assumption of monotonicity, is used once in the proof, to show the absence of L^2 solutions.

To summarize:

(a) H has absolutely continuous spectrum from $-\infty$ to $+\infty$ for $E\neq 0$. In particular it has no gaps of forbidden energies characteristic to the free ($E=0$) Bloch Hamiltonian.

(b) H has no eigenvalues (bound states), not even eigenvalues embedded in the continuous spectrum. In particular there is no ladder structure for the eigenvalues.

3. THE SINGLE BAND APPROXIMATION

Here we shall briefly review the "single band approximation," which reflects the idea that bands are meaningful objects even in the presence of external fields of force.[9] The approximation involves adding a term to the Hamiltonian which makes the band index a true constant of motion. The matrix elements of the position operator x in the nk representation are[5,10]

$$(nk|x|mk')=i\delta_{m,n}\frac{d}{dk}\delta(k-k')+x_{mn}(k)\delta(k-k'); \tag{5}$$

$x_{mn}(k)$ are continuous functions in the absence of bands crossing.[11] (This is the generic situation in one dimension.[14] Moreover,

$$\begin{aligned} x_{mn}&=\frac{i\hbar}{m}\frac{p_{mn}(k)}{\epsilon_m(k)-\epsilon_n(k)}, \\ \frac{1}{2m}\sum_m|p_{mn}(k)|^2&\leq\epsilon_n(k)+\sup_x|V(x)|. \end{aligned} \tag{6}$$

Consider the Hermitian operator A with matrix elements

$$\begin{aligned} A_{mn}(k)&=x_{mn}(k), \quad m\neq n, \\ A_{mn}(k)&=0. \end{aligned} \tag{7}$$

In the single band approximation, the Hamiltonian H in (1) is replaced by

$$H_{\mathrm{SB}}=H-EA. \tag{8}$$

One expects that H_{SB} is an approximation to H if A is in some sense small. Phenomenologically, A is associated with tunneling which is a very slow process on atomic scale for large band gaps[12] [see Eq. (6)]. H_{SB} assumes the simple form of an infinite number of decoupled, first-order differential operators, with the operator corresponding to the nth band being

$$iE\frac{d}{dk}+\epsilon_n(k)+Ex_{nn}(k). \tag{9}$$

H_{SB} is diagonal in the band index n and has pure point spectrum. The eigenvalues have two quantum numbers—the band index n and a ladder index ν. The eigenfunctions and eigenvalues are, respectively,

$$\begin{aligned} \psi_{n\nu}(mk)&=\frac{\delta_{n,m}}{\sqrt{2\pi/a}}\exp\left(-\frac{i}{E}\int_{-\pi/a}^{k}dk'[\lambda_{n\nu}-\epsilon_n(k') \right. \\ &\qquad \left. -Ex_{nn}(k')]\right), \\ \lambda_{n\nu}&=\nu Ea+\frac{a}{2\pi}\int_{-\pi/a}^{+\pi/a}dk[\epsilon_n(k)+Ex_{nn}(k)], \end{aligned} \tag{10}$$

where $\nu=\{0,\pm1,\pm2,\cdots\}$, $n=\{1,2,\cdots\}$. $\psi_{na}(k,m)$ is just a phase in k. This is a consequence of a conservation law which will be discussed in Sec. 4.

For fixed n, the eigenvalues are equally spaced, hence the name "Stark ladder." The infinite number of ladders, corresponding to the infinite number of bands, are intertwined. There is thus an *infinite* number of eigenvalues within each *energy interval* Ea.

4. N-BAND HAMILTONIAN

A. A spectral theorem

Consider the N-band Hamiltonian H_{NB} in $E_N\otimes L^2(B)$,

E_N an N-dimensional vector space and B the Brillouin zone. H_{NB} is given by

$$H_{NB}\psi(n,k)=iE\frac{d}{dk}\psi(n,k)+\epsilon_n(k)\psi(n,k)$$
$$+E\sum_{m=1}^{N}x_{nm}(k)\psi(n,k),\quad n=1,\ldots,N. \tag{11}$$

H_{NB} describes N bands coupled by the interband interaction $x_{mn}(k)$.[13] Our purpose is to show that, for any finite N (and in the absence of band crossing), H_{NB} has a purely discrete spectrum in the form of N Stark ladders. Only for $N=\infty$ is the continuous spectrum recovered.

H_{NB} has discrete spectrum by the following argument: The N-band Hamiltonian with $x_{mn}(k)$ set equal to zero has a compact resolvent [this follows from Eq. (10)]. If there is no band crossing, it follows from (6) that the interband interaction $x_{mn}(k)$ is a bounded operator. By a basic theorem of Rellich[14] H_{NB} has a compact resolvent, and so H_{NB} has a purely discrete spectrum with isolated eigenvalues accumulating only at infinity.

The pathological behavior of the spectrum under the perturbation of interband coupling is peculiar to the full infinite bands Hamiltonian. That is, only in this case does the interband coupling make a continuous spectrum discrete or vice versa. In particular, no N-band model recovers the absolute continuity of the true spectrum.

Let $\mathrm{h}(k)$ be the operator

$$\mathrm{h}\psi(n,k)=\epsilon_n(k)\psi(n,k)+E\sum_{m=1}^{N}x_{nm}(k)\psi(m,k) \tag{12}$$

and $h_{mn}(k)$ its matrix elements.

Hermiticity and periodicity give two global characteristics of the solutions of N-band Hamiltonians.

(a) The spectrum of eigenvalues has the form of N intertwined Stark ladders.

This follows from periodicity in k-space: If $\psi_\lambda(n,k)$, $\{k\in B, n=1,\ldots,N\}$ is an eigenvector with eigenenergy λ, then $\exp(i\nu ka)\psi_\lambda(k,n)$ is an eigenvector with eigenenergy $\lambda+\nu Ea$. A simple continuity argument shows that there are N such ladders: Let the interbands coupling shrink to zero. This shifts the eigenvalues up or down but it does not annihilate or create eigenvalues. Since there are N ladders for zero coupling, there are N ladders also for any non-zero interband coupling.

(b) Probability conservation in k-space: If $\psi_\lambda(k,n)$ is an eigenvector of the N-band Hamiltonian then $\sum_{n=1}^{N}|\psi_\lambda(k,n)|^2$ is constant, independent of k.

To show this let $(\psi(k),\phi(k))$ denote scalar product in the N-dimensional vector space, i.e.,

$$(\psi(k),\phi(k))\equiv\sum_{n=1}^{N}\psi^*(n,k)\phi(n,k). \tag{13}$$

The eigenvector ψ_λ of the N-band Hamiltonian satisfies

$$iE\frac{d}{dk}\psi_\lambda(k,n)=[\lambda-\mathrm{h}(k)]\psi_\lambda(k,n). \tag{14}$$

Since λ is real and $\mathrm{h}(k)$ self-adjoint, the result follows by Stone's theorem.[15] This conservation relation is the analog of the conservation of probability for the time-dependent Schrödinger Hamiltonian.

B. Methods of approximation

The N-band Hamiltonian cannot be solved exactly in general and one must resort to approximations.[13] Although perturbation expansion in the interband interaction is in principle possible, this is not the most convenient method.[16] In particular, the perturbation expansion does not preserve property (b)—conservation of probability in k. [It does preserve property (a).] A more natural approximation is the adiabatic method.[17] This approximation is exact either when $\mathrm{h}(k)$ is independent of k or when $\mathrm{h}(k)$ is diagonal. The approximation proceeds from the aforementioned k-t analogy.

Let $\lambda_n(k)$ and $\psi_n(m,k)$, $\{n,m=1,\ldots,N\}$ be the eigenvalues and eigenvectors of the matrix $\mathrm{h}(k)$ in E_N. Furthermore, let

$$\bar{\lambda}_n=\frac{a}{2\pi}\int_{-\pi/a}^{+\pi/a}dk\,\lambda_n(k).$$

Then, in the adiabatic approximation, the eigenvalues and eigenfunctions of the N-band Hamiltonian are respectively

$$\lambda_{n\nu}=\bar{\lambda}_n+\nu Ea,$$
$$\psi_{n\nu}^{NB}(m,k)=\psi_n(m,k)\exp(i\nu ka)\exp\left(\frac{i}{E}\int_{-\pi/a}^{k}\times dk[\lambda_n(k)-\bar{\lambda}_n]\right),$$

where $n=1,\ldots,N$ and $\nu=0,\pm1,\ldots,\pm\infty$. The adabatic approximation satisfies both properties (a) and (b), i.e., it has the spectrum of N ladders and it conserves probability in k.

SUMMARY

We have shown that the spectrum of the Bloch electron in an external field is continuous. Under a perturbation corresponding to the accounting for interband coupling within groups consisting of a finite number of bands, the spectrum has been shown to become discrete, consisting of intertwined Stark ladders. This phenomenon is related to a theorem of Weyl and von Neumann.[18] Finally we have shown that probability in k is conserved by eigenfunctions of the N-band Hamiltonian, and introduced a method of approximation for the interband interaction which preserves this property.

ACKNOWLEDGMENTS

J.E.A. thanks Professor B. Simon for his kind advice and the CPT, CNRS, Marseille for financial support.

*Partially supported by the European Research Office of the U.S. Army under contract No. DAJA-71-C-1977.

†This paper is partly based upon a thesis submitted by J.E.A. in partial fulfillment of the requirements for the D.Sc. degree in Physics at the Technion, Israel.

‡Present address: Department of Physics, Princeton University, Princeton, New Jersey.

§Work carried out also at the Technion, Israel and partially supported by National Science Foundation Grant No. 16025.

[1] G. Wannier, Phys. Rev. **117**, 432 (1960).
[2] J. Zak, Phys. Rev. Lett. **20**, 1477 (1968); W. Shockley, Phys. Rev. Lett. **28**, 349 (1972); C.A. Moyer, Phys. Rev. B **7**, 5025 (1973); R.W. Koss and L.M. Lambert, Phys. Rev. B **5**, 1479 (1972).
[3] T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1966).
[4] ϵ is an eigenvalue only if its associated eigenfunction is normalizable. Physicists often use a different terminology, with "eigenvalue" denoting any point of the spectrum.
[5] L.D. Landau and E.M. Lifshitz, *Statistical Physics* (Pergamon, Oxford, 1958).
[6] J. von Neumann and E. Wigner, Z. Physik **30**, 465 (1929).
[7] After this work was completed we learned of a similar result by J. Walter, Math. Z. **129**, 83–94 (1972).
[8] N. Dunford and J.T. Schwartz, *Linear Operators* (Interscience, New York, 1958).
[9] E.N. Adams, J. Chem. Phys. **21**, 2013 (1953).
[10] E.I. Blount, *Solid State Physics*, *Vol.* 13 (Academic Press, New York, 1962).
[11] J.E. Avron, thesis (Technion, Israel, 1975), unpublished.
[12] C. Zener, in *Mathematical Physics in One Dimension*, edited by E. Lieb and D. Mattis (Academic, New York, 1966).
[13] Approximate solutions to the two-band Hamiltonians were considered by Yu. A. Bychkov and A.M. Dykhne, Zh. Eksp. Theor. Fiz. **48**, 1174 (1965) [Sov. Phys. JETP **21**, 783 (1965)].
[14] F. Rellich, Math. Ann. **118**, 462 (1942).
[15] M. Reed and B. Simon, *Methods of Modern Mathematical Physics* (Academic, New York, 1975), Vols. I, II.
[16] Some of the difficultities which arise in perturbation treatment of the interband interaction are discussed in G.H. Wannier and J.P. van Dyke, J. Math. Phys. **9**, 899 (1968).
[17] L. Schiff, *Quantum Mechanics* (Wiley, New York, 1958).
[18] In this connection one should mention the work of N. Aroszian, Am. J. Math. **79**, 597 (1957), where the Sturm–Liouville equation for the semi-infinite line is considered subject to mixed boundary conditions at zero. Two examples are constructed having a continuous differentiable potential. In the first, a change in the boundary conditions transforms the singular continuous spectrum into a discrete spectrum of isolated eigenvalues. In the second example, a discrete spectrum which is dense on the real line becomes continuous.

Synchronized solitons

Alan C. Newell

Department of Mathematics, Clarkson College of Technology, Potsdam, New York 13676
(Received 1 June 1976)

We find that, under certain conditions, two solitons of an integrable system can form a synchronized bound state when the system is perturbed.

1. INTRODUCTION

In this paper we investigate the behavior of certain solutions to the "double" sine–Gordon equation,

$$u_{XX}-u_{TT}=u_{xt}=\sin u+\epsilon\lambda_0\sin(u/2),\quad 0<\epsilon\ll 1,\qquad (1.1)$$

where $x=(X+T)/2$, $t=(X-T)/2$, by mapping the solutions $u(x,t)$ of (1.1) into the scattering functions associated with the integrable sine–Gordon equation,[1]

$$u_{XX}-u_{TT}=u_{xt}=\sin u.\qquad (1.2)$$

This particular investigation is part of a general effort to develop a singular perturbation theory for the inverse scattering transform in which one constructs uniform asymptotic expansions for the scattering functions over time scales inversely proportioned to the small coupling coefficient ϵ multiplying the "nonintegrable" terms in the equation. Of particular importance is the understanding of the analog of "resonance" in strongly nonlinear systems.

A weakly nonlinear system is usually analyzed by calculating the slow changes, due to the weak nonlinear coupling, which occur in the fixed parameters (amplitudes A_i, wave vectors $\mathbf{k}_i$ and frequencies ω_i, $\omega_i=\omega(\mathbf{k}_i)$, $i=1,\ldots,N$) associated with the normal modes of the linear system. A solution is sought in the form of an asymptotic expansion in powers of the coupling parameter, the leading order term of which is given by a linear combination of the linear normal modes. In general, this expansion is nonuniform in time due to resonances between waves satisfying conservation of momenta $\sum_{j=1}^{n}\mathbf{k}_j=0$ and energy $\sum_{j=1}^{n}\omega_j=0$, $n=3,4,\ldots,N$, criteria.[†] The expansions can be made uniform (renormalized) by allowing the amplitudes (and in some cases the wave vectors and frequencies as well) to be slowly varying functions of time.[2] In this way, one obtains a description of how the energy in a given normal mode is affected by the nonlinear coupling with the other normal modes. It should be noted that the strength with which the secular terms arise depends on the class of solutions $u(x,t)$ which are sought; for example, in the ϕ^4 model of quantum field theory, if $u(x,t)$ is periodic over a finite interval, quartet resonances occur on a time scale ϵ^{-2}, ϵ being a perturbation amplitude; on the other hand, if $u(x,t)$ is a stationary random function of position, the interaction time for quartet resonances is ϵ^{-4}. In both cases, however, the mechanism for energy transfer is the same.

In this paper, we will discuss what happens if the leading approximation is no longer linear but belongs, instead, to the class of *integrable* nonlinear partial differential equations. The general solution of equations of this class consists of two components; one is associated with a continuum of wavenumbers and has (in most regions) a behavior analogous to that of dispersive waves in a linear system. The second component has no linear analog and consists of special isolated pulses known as solitons. A soliton is a local, permanent, traveling wave solution of one of the integrable equations and is distinguished by the remarkable property that when it interacts with other solution components, it remerges from the interaction with its identity (speed, shape, amplitude) intact. It does, however, undergo a phase shift in position relative to where it would have been had it travelled unimpeded. In (1.2), the soliton (kink, 2π pulse) solution is,

$$u(x,t)=\pm 4\tan^{-1}\exp(\pm 2\theta),\quad \theta=\theta_0-\eta x-t/4\eta,\qquad (1.3)$$

with the various plus–minus combinations referring to $\pm 2\pi\to 0$ or $0\to\pm 2\pi$ transitions as the real space variable $X=x+t$ sweeps from $-\infty$ to $+\infty$. The invariance of the identity of the soliton follows from the fact that each soliton is associated with a complex wavenumber $\zeta=i\eta$ (that is, a complex eigenvalue of an appropriate operator; the real spectrum is associated with the solution component analogous to dispersive waves), which for the integrable system (1.2) is a constant of the motion.

What we intend to explore is what happens when we add to the integrable system (1.2) a perturbation term. It is to be expected that the noninteracting normal modes of the integrable system will now be coupled. It is the aim of this paper to study the nature of the interaction between these nonlinear normal modes and in particular between two solitons which in the nonperturbed system would simply pass through each other. The mechanism for a strong interaction (i.e., an order one change in the parameters describing each soliton, a change which can be produced even by a weak coupling given a sufficiently long time) cannot be one of simple resonance as that concept is primarily a linear one. The analog of resonance in nonlinear systems is synchronization or phase-locking and, under certain circumstances, this is exactly what the two solitons do. The two, previously noninteracting, solitons form a common synchronized state in which the notion of stability or binding energy can be precisely defined. The solution we find closely resembles the "wobbler" solutions of (1.1) observed numerically by Bullough and Caudrey[3] and discussed by them at the University of Arizona Conference held at Tucson in January 1976.

2. ANALYSIS

Consider the eigenvalue problem

$$
\begin{aligned}
v_{1x}+i\zeta v_1 &= -\tfrac{1}{2}u_x v_2,\\
v_{2x}-i\zeta v_2 &= \tfrac{1}{2}u_x v_1,
\end{aligned}
\tag{2.1}
$$

where $u(x,t)$ is a solution of (1.1). The eigenvalue problem (2.1) provides a means of mapping the potential $u_x(x,t)$ into the scattering data associated with (2.1). The scattering data consists of (a) the spectrum of (2.1), i.e., the real ζ axis and a set of discrete imaginary eigenvalues $\zeta_k=i\eta_k$ and a set of discrete paired eigenvalues $(\zeta_k, -\zeta_k^*)$ in the complex plane and (b) the asymptotic behavior (as $x\to+\infty$) of the oscillatory and bound state eigenfunctions. We shall not go into details here. Instead we shall refer the reader to Refs. 4–6 and here simply write down the time rate of change of the scattering data:

$$
\zeta_{kt}=\frac{1}{2\beta_k a_k'^2}\int_{-\infty}^{\infty} u_{xt}(\psi_1^2+\psi_2^2)_k\,dx,
\tag{2.2}
$$

$$
\beta_{kt}=\frac{1}{2a_k'^2}\int_{-\infty}^{\infty} u_{xt}\times\left\{\frac{\partial}{\partial\zeta}(\psi_1^2+\psi_2^2)_k-\frac{a_k''}{a_k'}(\psi_1^2+\psi_2^2)_k\right\}dx,
\tag{2.3}
$$

$$
(b^*/a)_t=\frac{1}{2a^2}\int_{-\infty}^{\infty} u_{xt}(\psi_1^2+\psi_2^2)\,dx,
\tag{2.4}
$$

where $\{\zeta_k\}_{k=1}^N$ are the discrete eigenvalues of (2.1) and $(\psi_1(x,\zeta),\psi_2(x,\zeta))$ is the solution to (2.1) which is analytic for $\mathrm{Im}\,\zeta>0$, and in particular at each ζ_k, and which has the asymptotic behavior $(0,1)\exp(i\zeta x)$ as $x\to+\infty$. The scattering function $a(\zeta,t)$ plays a central role in the theory (see Ref. 7) and in particular its zeros in the upper half ζ-plane are the bound state eigenvalues. It can be written in terms of its zeros and its behavior for real ζ,

$$
a(\zeta)=\prod_{j=1}^{N}\left(\frac{\zeta-\zeta_j}{\zeta-\zeta_j^*}\right)\exp\frac{1}{2\pi i}\int_{-\infty}^{\infty}\frac{1}{\xi-\zeta}\ln[aa^*(\xi)]\,d\xi.
\tag{2.5}
$$

The quantity b^*/a is the reflection coefficient and defined on the real spectrum. The parameter β_k is given by $(b_k a_k')^{-1}$, $a_k'=(\partial a/\partial\zeta)_k$, and is related to the asymptotic behavior of $\psi_1(x,\zeta_k)$ as $x\to-\infty$; in fact, $\lim_{x\to-\infty}\times\psi_1(x,\zeta_k)\exp(i\zeta_k x)=b_k^*$. It is also the residue of the analytic extension of b^*/a at ζ_k, when this quantity admits such an extension.

We take as our leading approximation the multisoliton solution which corresponds to the scattering data,

$$
\begin{aligned}
&\zeta_1=i\eta_1,\quad \zeta_2=i\eta_2,\quad \beta_1,\quad \beta_2,\\
&b\equiv 0,\qquad a=\frac{\zeta-i\eta_1}{\zeta+i\eta_1}\,\frac{\zeta-i\eta_2}{\zeta+i\eta_2}.
\end{aligned}
\tag{2.6}
$$

The corresponding potential is given by

$$
u(x)=4\tan^{-1}\frac{1-[(\eta_1-\eta_2)/(\eta_1+\eta_2)]^2\exp(2\theta_1+2\theta_2)}{\exp(2\theta_1)+\exp(2\theta_2)},
\tag{2.7}
$$

where

$$
\theta_j=-\eta_j x+\theta_{j0}
$$

and

$$
\beta_j=(1/2i\eta_j a_j'^2)\exp(-2\theta_{j0}),\quad j=1,2.
\tag{2.8}
$$

Now, if we were simply solving Eq. (1.2), we would find

$$
\begin{aligned}
&\zeta_{kt}=0,\\
&\beta_{kt}=(i/2\zeta_k)\,\beta_k,\\
&b_t(\zeta,t)=0.
\end{aligned}
$$

Then the phases in (2.8) have the structure

$$
\theta_j=\hat\theta_{j0}-\eta_j x-(1/4\eta_j)t
\tag{2.9}
$$

in light cone coordinates and the structure

$$
\begin{aligned}
\theta_j&=-\tfrac{1}{2}(\eta_j+1/4\eta_j)(X-X_j-V_jT),\\
\hat\theta_{j0}&=\tfrac{1}{2}(\eta_j+1/4\eta_j)X_j
\end{aligned}
\tag{2.10}
$$

in real space–time coordinates. Here V_j is the velocity of the pulse corresponding to η_j and is given by

$$
V_j=-1+2/(4\eta_j^2+1).
\tag{2.11}
$$

In order to understand the motion, let us imagine that $\eta_1>\eta_2$ and that the η_1 pulse starts far to the right (in X space) of the η_2 pulse. Then, when $\theta_1=0(1)$, $\theta_2=-\infty$ and

$$
u(x,t)=4\tan^{-1}\exp(-2\theta_1),
\tag{2.12}
$$

which is a kink with a transition from 0 to 2π as X sweeps from left to right through X_1, the center of the pulse. Similarily, when $\theta_2=0(1)$, $\theta_1=+\infty$ and in the neighborhood of

$$
X_2'=X_2+\frac{8\eta_2}{4\eta_2^2+1}\ln\left|\frac{\eta_1-\eta_2}{\eta_1+\eta_2}\right|,
$$

the center of the η_2 pulse,

$$
\begin{aligned}
u(x,t)&=-4\tan^{-1}\left(\frac{\eta_1-\eta_2}{\eta_1+\eta_2}\right)^2\exp(2\theta_2),\\
&=-4\tan^{-1}\exp\left(2\theta_2+2\ln\left|\frac{\eta_1-\eta_2}{\eta_1+\eta_2}\right|\right).
\end{aligned}
\tag{2.13}
$$

This pulse makes the transition from -2π as $X-X_2'\to-\infty$ to 0 as $X-X_2'\to+\infty$.

Now, as $T\to+\infty$, the pulses change places since $\eta_1>\eta_2$ implies that $V_1<V_2$ and so the η_1 pulse will eventually be to the left of the η_2 pulse and the solution locally in its neighborhood is given by

$$
u(x,t)=-4\tan^{-1}\exp\left(2\theta_1+2\ln\left|\frac{\eta_1-\eta_2}{\eta_1+\eta_2}\right|\right),
$$

which shows that the $0\to2\pi$ kink has switched to a $-2\pi\to0$ kink and its center is now at

$$
X=X_1+V_1T+\frac{8\eta_1}{4\eta_1^2+1}\ln\left|\frac{\eta_1-\eta_2}{\eta_1+\eta_2}\right|.
$$

Similarly the η_2 pulse has become a $0\to2\pi$ transition with a center at

$$
X=X_2+V_2T,
$$

and each pulse has undergone a phase shift. After infinite time, the pulses will be infinitely far apart.

Now, we will look at the effect of adding the perturbation term $\epsilon\lambda_0 \sin(u/2)$ to (1.2). The right-hand sides of (2.2) and (2.3) no longer give simple expressions but contain the extra terms

$$\frac{\epsilon\lambda_0}{2\beta_k a_k'^2}\int_{-\infty}^{\infty}\sin\frac{u}{2}\cdot(\psi_1^2+\psi_2^2)_k\,dx$$

and

$$\frac{\epsilon\lambda_0}{2a_k'^2}\int_{-\infty}^{\infty}\sin\frac{u}{2}\left\{\frac{\partial}{\partial\zeta}(\psi_1^2+\psi_2^2)_k-\frac{a_k''}{a_k'}(\psi_1^2+\psi_2^2)_k\right\}dx \tag{2.14}$$

respectively. We solve the system of equations (2.2), (2.3) iteratively by calculating the squared eigenfunctions which correspond to the exact multisoliton solution (2.7) of (1.2). Following the ideas outlined by Zakharov and Shabat,[8] we define the quantities

$$\sqrt{\gamma_j}\ \psi_j(\zeta_k)=\sqrt{\gamma_j}\ \psi_{jk}=u_{jk},$$

$$\sqrt{\gamma_j}\ \exp(i\zeta_j x)=\lambda_j,\quad \gamma_j=1/\beta_j a_j'^2, \tag{2.15}$$

and find the equations

$$\begin{bmatrix} I & B\\ & \\ -B^* & I\end{bmatrix}\begin{bmatrix}u_{11}\\ u_{12}\\ u_{21}^*\\ u_{22}^*\end{bmatrix}=\begin{bmatrix}0\\ 0\\ \lambda_1^*\\ \lambda_2^*\end{bmatrix}$$

with

$$B=[b_{jk}]=\left[\frac{\lambda_j\lambda_k^*}{\zeta_j-\zeta_k^*}\right],\quad j,k=1,2. \tag{2.16}$$

Using these expressions and the leading approximation (2.7) for $u(x,t)$, we may calculate the expressions (2.14). The computations are lengthy but straightforward and involve integrals of the form

$$\int_0^{\infty}\frac{C+Dy+Ey^2+Fy^3+Gy^4}{(y+A)^3(y+B)^3}dy$$

and

$$\int_0^{\infty}\ln y\frac{C+Dy+Ey^2+Fy^3+Gy^4}{(y+A)^3(y+B)^3}dy. \tag{2.17}$$

It turns out that the terms (2.14) lead to secular contributions (namely, the eigenvalues ζ_j and the coefficients of $\exp(t/2\eta_j)$ in β_j grow proportional to ϵt) only when η_1 is close to η_2. We, therefore, assume the parameter

$$\beta=(\eta_1-\eta_2)/(\eta_1+\eta_2) \tag{2.18}$$

to be small and keep only the leading terms in the two expressions (2.14). These will turn out to be $O(1)$ and $O(\beta^{-1})$. We will show how β and ϵ are related when we analyze these perturbation equations in the next section. It also turns out that the scattering function connected with the continuous spectrum remains small.

3. ANALYSIS OF THE PERTURBATION EQUATIONS

With the addition of the perturbation terms $\epsilon\lambda_0\sin u/2$, the equations for the evolution of the scattering data $\eta_1,\eta_2,\beta_1,\beta_2$ associated with the double soliton solution are given by

$$\eta_{1t}=-\eta_{2t}=-\frac{\epsilon\lambda_0}{2}\frac{\beta_1-\beta_2}{\beta_1+\beta_2}+O(\epsilon\beta), \tag{3.1}$$

$$\beta_{1t}=\frac{1}{2\eta_1}\beta_1-\frac{\epsilon\lambda_0}{2\eta\beta}\beta_1-\frac{1}{\eta\beta}\eta_{1t}\beta_1+O(\epsilon\ln\beta), \tag{3.2}$$

$$\beta_{2t}=\frac{1}{2\eta_2}\beta_2+\frac{\epsilon\lambda_0}{2\eta\beta}\beta_2+\frac{1}{\eta\beta}\eta_{2t}\beta_2+O(\epsilon\ln\beta). \tag{3.3}$$

In these expressions, we have only included the two leading order terms and have also set $\eta=(\eta_1+\eta_2)/2$. The first observation of interest is that the total energy of the two pulses,

$$\int_{-\infty}^{\infty}u_x^2\,dx=16(\eta_1+\eta_2),$$

is a constant of the motion. This may also be verified directly from (1.1). The second observation is that if the η_1 pulse starts far to the right of the η_2 pulse, then

$$(\beta_1-\beta_2)/(\beta_1+\beta_2)\approx-1\quad\text{and}\quad \eta_{1t}=-\eta_{2t}=\epsilon\lambda_0/2. \tag{3.4}$$

Note that this agrees precisely with (1.1) by integrating from $x=-\infty$ to between the two pulses and again from between the two pulses to $x=+\infty$. We find

$$\frac{\partial}{\partial t}\int u_x^2\,dx=16\frac{\partial}{\partial t}\eta=-4\epsilon\lambda_0\left[\cos\frac{u}{2}\right]_{u=a}^{u=b}. \tag{3.5}$$

Because the rightmost pulse η_R has a $0-2\pi$ transition and the leftmost pulse has a $-2\pi\to 0$ transition, we find (in terms of real time T)

$$\eta_{Rt}=-2\eta_{RT}=\epsilon\lambda_0/2$$

and

$$\eta_{Lt}=-2\eta_{LT}=-\epsilon\lambda_0/2. \tag{3.6}$$

Thus, the rightmost (leftmost) pulse decreases (increases) its velocity when $\lambda_0<0$, and in this situation we expect that the pulses can become phase locked. The role of the sign of λ_0 is important, but we note its role in Eq. (1.1) is small since it can be changed by letting $u\to u+2\pi$.

Returning to a perturbation analysis of (3.1)–(3.3), we see that if η_1 and η_2 are significantly different, the velocities are different and the interaction time is so short that no significant energy exchange occurs. On the other hand, for small β, a significant interaction does take place. The relation between β and ϵ is determined by balancing the difference in the frequencies $1/2\eta_1$ and $1/2\eta_2$ with the ϵ/β term arising from the perturbation. The correct choice is clearly $\beta=O(\epsilon^{1/2})$ and thus it is clear that the relevant interaction time scale is also $\tau=\epsilon^{1/2}t$. We introduce the following changes of variables,

$$\beta=\epsilon^{1/2}b(\tau),\quad \beta_j=b_j(\tau)\exp(t/2\eta) \tag{3.7}$$

and obtain the equations describing the slow changes in the parameters η_1,η_2 (and β), b_j, $j=1,2$. These equations are (set $\lambda_0=-\mu_0$)

$$b_\tau=\frac{\mu_0}{2\eta}\frac{b_1-b_2}{b_2+b_1}, \tag{3.8}$$

$$b_{1\tau}=-\frac{1}{2\eta}\left(b-\frac{\mu_0}{b}\right)b_1-\frac{b_\tau}{b}b_1, \tag{3.9}$$

$$b_{2\tau}=\frac{1}{2\eta}\left(b-\frac{\mu_0}{b}\right)b_2-\frac{b_\tau}{b}b_2. \tag{3.10}$$

We may integrate this third order system twice. Define

$$y=b_1-b_2, \quad x=b_1+b_2$$

and find after a little calculation that

$$x=A\exp[(-1/2\mu_0)b^2], \quad y=(2\eta/\mu_0)xb_\tau, \tag{3.11}$$

and that $b^2=z$ satisfies the equation

$$z_{\tau\tau}-\frac{1}{2\mu_0}z_\tau^2+\frac{\mu_0}{2\eta^2}(z-\mu_0)=0. \tag{3.12}$$

Equation (3.12) may be integrated once,

$$\left(\frac{dz}{d\tau}\right)^2=\frac{\mu_0^2}{\eta^2}z+\left(C-\frac{\mu_0^3}{\eta^2}\right)\exp[+(1/\mu_0)(z-\mu_0)], \tag{3.13}$$

where C may be calculated from the initial data,

$$z(0)=b^2(0)$$

and

$$\left(\frac{dz(0)}{d\tau}\right)^2=4b^2(0)b_\tau^2(0)$$
$$=\frac{\mu_0^2}{\eta^2}b^2(0)\left(\frac{b_1(0)-b_2(0)}{b_1(0)+b_2(0)}\right)^2=\frac{\mu_0^2}{\eta^2}\nu^2b^2(0), \tag{3.14}$$

where

$$\nu=[b_1(0)-b_2(0)]/[b_1(0)+b_2(0)], \quad -1<\nu<1.$$

Hence,

$$C=\frac{\mu_0^3}{\eta^2}-\frac{\mu_0^2}{\eta^2}b^2(0)(1-\nu^2)\exp\left(-\frac{1}{\mu_0}(z(0)-\mu_0)\right),$$

and

$$\frac{\eta^2}{\mu_0^2}\left(\frac{dz}{d\tau}\right)^2=z-b^2(0)(1-\nu^2)\exp\left(\frac{1}{\mu_0}[z-b^2(0)]\right). \tag{3.15}$$

The analysis of (3.15) is straightforward. There are always two real positive roots z_1 and z_2 of the right-hand side of (3.15) since at $z=b^2(0)$ the right-hand side is positive whereas at $z=0$ and ∞ it is negative. The solution oscillates between the two roots z_1 and z_2 ($z_1<z_2$) and the phase plane of $v=(\eta/\mu_0)\,dz/d\tau$ and z is shown in Fig. 1.

If $P=b^2(0)(1-\nu^2)\exp[-b^2(0)/\mu_0]$ is almost equal to μ_0e^{-1} (namely, $b^2(0)\approx\mu_0$, $\nu^2\approx 0$), then the motion is an elliptical one in the neighborhood of the center $z=\mu_0$, $dz/d\tau=0$. On the other hand, if ν^2 is just less than unity (i.e., the pulses start very far apart), then the two roots of the right-hand side are very small (of order P) and very large respectively. However, we note that z is always greater than zero. For intermediate values of the parameter P, we obtain a typical orbit given by the curve $ABCD$.

Beginning at C where v/z is maximal [it should be noted that maximal v does not mean the two pulses are apart by the maximum distance since $(b_1-b_2)/(b_1+b_2)=y/x=v/z$; see (3.11)], the two pulses are far apart with pulse 2 lying to the left of pulse 1. The motion proceeds in real time (recall real time T is related to light cone time by $\partial/\partial t=-2\partial/\partial T$) along CB. The square of the difference in the two pulse velocites increases and reaches a maximum at B when the two pulses are close together. Thereafter, pulse 1 ($\eta_1>\eta_2>\frac{1}{2}$, $|V_1|>|V_2|$) is the leftmost one and is slowed down as it travels along BA till the maximum separation is again reached. Along AD, pulse 2, moving leftward overtakes pulse 1 at D and continues ahead to C. The reason for the asymmetry in the difference in pulse velocities is that if no perturbation were present, pulse 1 would be the fastest (moving leftward) pulse.

The predictions of this theory agree closely with the observations of Bullough and Caudrey who have examined Eq. (1.1) numerically. Ablowitz, Kruskal, and Ladik[9] have also seen a similar behavior in their numerical experiments. We stress that the solutions discussed here are strictly only valid for time scales $O(\epsilon^{-1/2})$. It is possible that other weaker secular terms may cause these structures to collapse by emitting radiation over longer times.

4. SUMMARY

Whereas in the noninteracting system (1.2), the two pulses were separate and noninteracting, in the perturbed system the two separate pulses form one synchronized state. We can define the stability of this state, or the binding energy of the two pulses, to be the minimal distance between the orbit $ABCD$ and the limiting orbit $OMNP$. If we had included more soliton pairs in our basic description, then the initial conditions change and it may be possible to knock the two pulses out of their synchronized state by collision with other synchronized states.

In this way, we can build up classes of particles (synchronized states) which can interact (scatter) in a nontrivial way. The calculations are lengthy but presently we are making some progress on the interaction

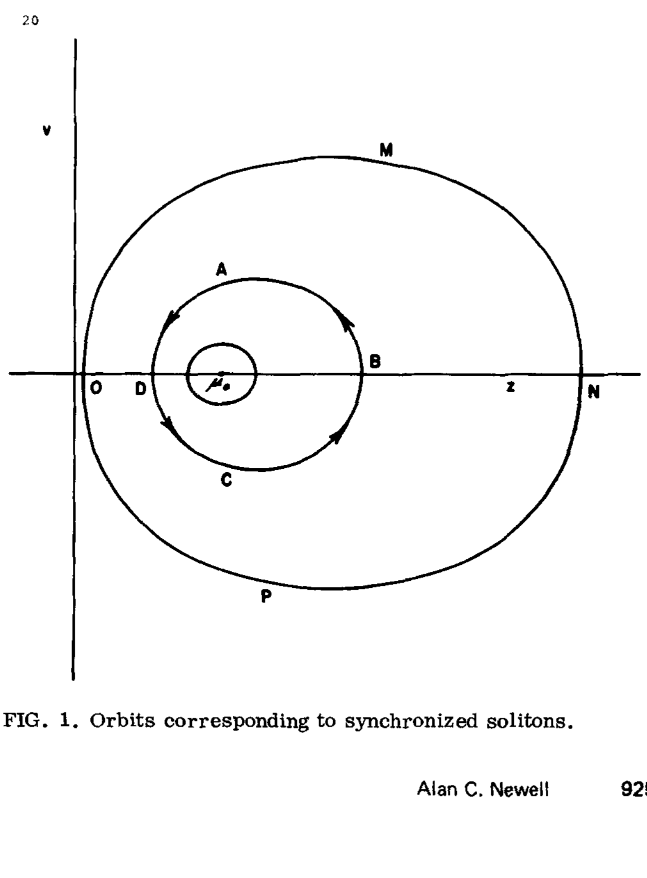

FIG. 1. Orbits corresponding to synchronized solitons.

of many particles. If such a system were to provide a useful model of a classical field theory, it would suggest that the building blocks of the "real" universe consist of synchronized states of more elementary and noninteracting building blocks generated by the field equation of some "perfect" universe.

ACKNOWLEDGMENTS

I wish to thank Philip Caudrey and Robin Bullough whose numerical experiments suggested to me that the double sine—Gordon equation might provide a useful yet simple model in which to study strong interactions between solitons. The work is partially supported by the National Science Foundation, Grant Nos. MPS75—07568 and DES75—06537.

†*Note added in proof*: Recently, there has been the remarkable discovery that this criterion also governs the strong interaction of solitons in higher spatial dimensions. Miles (to be published shortly) has shown that the Hiroto two-soliton solution for the Kadomtsev—Petviashvili equation breaks down when this criterion with $n=3$ is satisfied. Furthermore, Newell and Redekopp (also to be published shortly) have shown that when in fact the criterion obtains, the multisoliton solutions of all systems integrable by the general Zakharov—Shabat scheme break down and this in turn leads to the possibility of soliton creation.

[1]M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur, Phys. Rev. Lett. 30, 1462 (1973).
[2]B. B. Varga and S. O. Aks, J. Math. Phys. 15, 149 (1974).
[3]R. Bullough and P. Caudrey, private communication.
[4]M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur, Stud. Appl. Math. 53, 249 (1974).
[5]D. J. Kaup and A. C. Newell, Advan. Math. (1977) (to be published).
[6]D. J. Kaup, SIAM J. Appl. Math. 31, 121 (1976).
[7]H. Flaschka and A. C. Newell, *Lecture Notes in Physics*, *Vol.* 38, edited by J. Moser (Springer, New York, 1975), pp. 355ff.
[8]V. E. Zakharov and A. B. Shabat, Zh. Eksp. Teor. Fiz. 61, 118 (1971) [Sov. Phys. JETP 34, 62 (1972)].
[9]M. J. Ablowitz, M. D. Kruskal, and J. Ladik "Numerical Studies of Interacting Solitary Waves in Nonintegrable Equations" (to be published).

T matrix and effective range function for Coulomb plus rational separable potentials especially for $l=1$

H. van Haeringen

Natuurkundig Laboratorium der Vrije Universiteit, Amsterdam, The Netherlands
(Received 20 October 1976)

The off-shell $l=1$ T matrix in the momentum representation for the pure Coulomb potential and for the Coulomb plus a rational separable potential of the Yamaguchi type is obtained in closed form. The amplitude, the effective range function, and the effective range parameters are derived from the T matrix and are given in closed form. For a large number of rational separable potentials we prove that the effective range function is real analytic at zero energy. We give, however, an example of a potential for which this effective range function has a *pole* at the origin. From these effective range functions a certain function W is extracted which does not depend either on l or on the particular potential. This function W is studied in detail. We indicate how the results of this paper can be generalized to arbitrary values of l and to all Coulomb plus rational separable potentials.

1. INTRODUCTION

In Sec. 2 we present a number of results for scattering by a potential which is the sum of the Coulomb potential and a rational separable potential[1] in the $l=1$ partial wave space. Analogous results for $l=0$ have been published in Refs. 1 and 2.

We give a closed formula for the pure Coulomb T matrix for $l=1$ in Eq. (2.1). Further we consider a rank-one separable potential with form factor of the type $p^l(p^2+\beta_l^2)^{-l-1}$. For $l=1$ we obtain the T matrix for the Coulomb potential plus a potential of the above type. By applying the asymptotic states defined and studied by the author to this T matrix we obtain the $l=1$ partial wave projected physical scattering amplitude. The effective range parameters a_1 and r_1 are derived from the amplitude in the well-known way and given in closed form, see Eqs. (2.9) and (2.10).

The larger part of this paper, Secs. 3–6, is mainly concerned with the so-called effective range function K and related functions. The function K, being analytic at zero energy ($k^2=0$), can be expanded in a Taylor series,

$$K_l(k^2) = -a_l^{-1} + \tfrac{1}{2} r_l k^2 - \cdots,$$

where a_l, r_l, $\cdots$ are real. An analytic function with real expansion coefficients is called real analytic.

In Ref. 2 we *assumed* that K_0 is real analytic, and we derived closed expressions for a_0 and r_0. In the present paper we *prove* the real analyticity of K_l for a large number of rational separable potentials [see Eq. (3.11)]. To our knowledge, no such proof has been given before when the additional potential is nonlocal. Only for Coulomb plus *local* short-range potentials the analytic properties of the effective range function K_l have been studied and the real analyticity of K_l has been proved, see Hamilton *et al.*[3] and Cornille and Martin.[4]

In the following we shall use several variables. In the first place we have the strength s of the Coulomb potential. It is real and in this paper it is kept fixed; however, both $s>0$ and $s<0$ will be considered. Secondly we have the strengths λ_l of the separable potential which play only a role of minor importance. Further we have α and β which are parameters related to the range of the rational separable potential. Finally we have the wavenumber k which is equal to the square root of the energy. It is often convenient to use instead of k, α, β the variables γ, μ, ν defined by

$$\mu\alpha \equiv \nu\beta \equiv k\gamma \equiv -s.$$

We shall consider functions of k and also functions of k^2. It is important to note that the so-called physical complex k^2-plane corresponds to the upper half of the complex k-plane, i.e., $\mathrm{Im}k>0$. Consequently, the physical complex γ-half-plane depends on the sign of the Coulomb strength s. It is determined by $s\,\mathrm{Im}\gamma>0$. The Coulomb bound states occur only if $s>0$ and are given by $\gamma=in$, $n=1,2,\cdots$.

In Sec. 3 we extract from the effective range functions K_l (corresponding to different potentials) a certain function $W(\gamma;\mu,\nu)$ which depends neither on l nor on the particular potential. If W is real analytic so is K_l, with the exception of possible poles, for a large number of potentials defined in Sec. 3. We claim that this even holds for all rational separable potentials. We shall prove that $W(\gamma;\mu,\nu)$ is indeed real analytic at $\gamma^{-2}=0$ for real μ and ν.

Related to W is the hypergeometric function $F_{l\gamma}(AB) \equiv {}_2F_1(1, i\gamma; 1+i\gamma; AB)$ [with $A\equiv(\alpha+ik)/(\alpha-ik)$, $B\equiv(\beta+ik)/(\beta-ik)$] that we encountered before.[1] In Sec. 4 we shall study $F_{l\gamma}(AB)$ for real positive k, α and β, i.e., for real γ, μ and ν. In Sec. 5 we investigate $W(\gamma;\mu,\nu)$ for real μ,ν and complex γ. We introduce there the function $W(\gamma;\xi)$ which is related to but simpler than $W(\gamma;\mu,\nu)$. This function $W(\gamma;\xi)$ is useful for the exact numerical computation of $W(\gamma;\mu,\nu)$ and therefore of the effective range functions, see Eq. (5.11).

Eventually in Sec. 6 all variables are taken complex and the proof of the real analyticity of $W(\gamma;\mu,\nu)$ at $k=0$ (for real μ and ν) is given. To achieve this, related functions $V(\gamma;\mu,\nu)$ and $V(\gamma;\xi)$ will be introduced which are analytic in all their variables. Section 7 summarizes the results.

2. *T* MATRIX, AMPLITUDE AND EFFECTIVE RANGE PARAMETERS

In this section we present the Coulomb T matrix, the T matrix for the Coulomb plus Yamaguchi-type potential and the corresponding amplitude and effective range parameters in explicit form, all for $l=1$. The analogous $l=0$ quantities have been presented in Refs. 1 and 2. We use here the same notation.

The $l=1$ pure Coulomb T matrix is obtained from Eq. (24) of Ref. 1, in the same way as $T_{c,l=0}$ has been derived. The result is as follows:

$$\langle p'|T_{c,l=1}(k^2)|p\rangle$$
$$=\frac{ik}{\pi pp'}\frac{1}{1+\gamma^2}\left[2i\gamma(1-i\gamma)+i\gamma\frac{p^2-k^2}{2pk}\frac{p'^2-k^2}{2p'k}\ln\left(\frac{p'+p}{p'-p}\right)^2\right.$$
$$+\left(i\gamma-\frac{p^2+k^2}{2pk}\right)\left(i\gamma-\frac{p'^2+k^2}{2p'k}\right)F_{i\gamma}(aa')$$
$$+\left(i\gamma+\frac{p^2+k^2}{2pk}\right)\left(i\gamma+\frac{p'^2+k^2}{2p'k}\right)F_{i\gamma}\left(\frac{1}{aa'}\right)$$
$$+\left(i\gamma-\frac{p^2+k^2}{2pk}\right)\left(i\gamma+\frac{p'^2+k^2}{2p'k}\right)F_{i\gamma}\left(\frac{a}{a'}\right)$$
$$\left.+\left(i\gamma+\frac{p^2+k^2}{2pk}\right)\left(i\gamma-\frac{p'^2+k^2}{2p'k}\right)F_{i\gamma}\left(\frac{a'}{a}\right)\right]. \qquad (2.1)$$

Here $F_{i\gamma}(\cdot)$ is the hypergeometric function ${}_2F_1(1,i\gamma;1+i\gamma;\cdot)$ and $a=(p-k)/(p+k)$, $a'=(p'-k)/(p'+k)$.

In Ref. 1 we have introduced what we call the rational separable potentials in the $l=0$ space. The definition of a rational separable potential can be extended to all l in an obvious way. Here we consider only the form factor

$$\langle p|g_{\beta,l}\rangle=(2/\pi)^{1/2}p^l(\beta^2+p^2)^{-l-1}. \qquad (2.2)$$

This form factor is often proposed to describe nucleon—nucleus and nucleon—nucleon scattering (e.g., Cattapan[5] and Črepinšek[6]). We would like to have closed formulas for the T matrices corresponding to potentials $V=V_c+V_s$, where V_c is the Coulomb potential and V_s a separable potential of finite rank with form factors of the type of Eq. (2.2). According to Sec. 5 of Ref. 1 it is for this purpose sufficient to derive closed formulas for the following two objects:

$$\langle p|g^c_{\beta,l}(k^2)\rangle\equiv\langle p|[1+T_{c,l}(k^2)G_{0,l}(k^2)]|g_{\beta,l}\rangle$$
and
$$\langle g_{\alpha,l}|G_{c,l}(k^2)|g_{\beta,l}\rangle\equiv\langle g_{\alpha,l}|G_{0,l}(k^2)|g^c_{\beta,l}(k^2)\rangle.$$

With the help of Eq. (2.1) this plan of action has been carried out for $l=1$. By applying the straightforward method of Secs. 6 and 7 of Ref. 1 we find, with $A=(\alpha+ik)/(\alpha-ik)$ and $B=(\beta+ik)/(\beta-ik)$,

$$\langle p|g^c_{\beta,l=1}(k^2)\rangle$$
$$=\frac{(2/\pi)^{1/2}p}{(\beta^2+p^2)^2}-\frac{(2/\pi)^{1/2}i\gamma k^2}{p(\beta^2+k^2)^2}+\frac{(2/\pi)^{1/2}\beta k\gamma(p^2-k^2)}{p(\beta^2+p^2)(\beta^2+k^2)^2}$$
$$+\frac{(2/\pi)^{1/2}k^2}{p(\beta^2+k^2)^2}\left[\left(i\gamma-\frac{p^2+k^2}{2pk}\right)F_{i\gamma}(Ba)\right.$$
$$\left.+\left(i\gamma+\frac{p^2+k^2}{2pk}\right)F_{i\gamma}\left(\frac{B}{a}\right)\right], \qquad (2.3)$$

and

$$\langle g_{\alpha,l=1}|G_{c,l=1}|g_{\beta,l=1}\rangle$$
$$=\frac{2ik^3}{(\alpha^2+k^2)^2(\beta^2+k^2)^2}\left[1-(1+\gamma^2)F_{i\gamma}(AB)\right]$$
$$+\frac{ik}{2(\alpha+\beta)^2(\alpha-ik)^2(\beta-ik)^2}-\frac{1}{2(\alpha+\beta)^3(\alpha-ik)(\beta-ik)}$$
$$+\frac{k\gamma}{2(\alpha+\beta)^2(\alpha^2+k^2)(\beta^2+k^2)}$$
$$-\frac{(k\gamma)^2}{(\alpha+\beta)(\alpha+ik)(\beta+ik)(\alpha^2+k^2)(\beta^2+k^2)}. \qquad (2.4)$$

In order to obtain the scattering amplitude from the off-shell T matrix, one should take the physical on-shell T matrix, i.e., sandwich the T matrix with the appropriate asymptotic states as has been discussed by the author.[7]

Let V_s from now on be restricted to a rank-one potential defined in the partial wave space characterized by l,

$$V_{s,l}=-\lambda_l|g_{\beta,l}\rangle\langle g_{\beta,l}|, \qquad (2.5)$$

where the form factor is given by Eq. (2.2). Suppressing now the subscript β, we get for the physical on-shell Coulomb-modified T matrix,

$$t_{cs,l}(k)\equiv\langle k\infty-|T_{cs,l}(k^2)|k\infty\rangle$$
$$=-\tau^c_l\langle k\infty-|g^c_l\rangle\langle g^c_l|k\infty\rangle, \qquad (2.6)$$

with

$$\tau^{c-1}_l=\lambda^{-1}_l+\langle g_l|G_c|g_l\rangle,$$

and the amplitude is proportional to $t_{c,l}+t_{cs,l}$.

At this point we are able to give a closed formula for the amplitude for $l=1$. Indeed, from Ref. 7 we derive

$$\langle g^c_l|k\infty\rangle=\langle g_l|kl+\rangle_c, \qquad (2.7)$$

where the right-hand side is known,

$$\langle g_l|kl+\rangle_c=(2/\pi)^{1/2}[i^{i\gamma}(l+i\gamma)!/l!]k^l(\beta^2+k^2)^{-l-1}B^{-i\gamma}. \qquad (2.8)$$

We checked this expression explicitly for $l=1$, using Eq. (2.3) (cf. Ref. 8). Furthermore, a closed form for $\tau^c_{l=1}$ is obtained from Eq. (2.4).

In particular we are now able to express the Coulomb-modified low energy scattering parameters for $l=1$ in terms of known functions. For a repulsive Coulomb potential ($\nu=k\gamma/\beta>0$) the results are

$$-a^{-1}_{cs,l=1}=2\beta^3\nu^3\Gamma(0,4\nu)$$
$$+\tfrac{1}{16}\beta^3\exp(-4\nu)(16\beta^5/\lambda-1+2\nu-8\nu^2), \qquad (2.9)$$
$$\tfrac{1}{2}r_{cs,l=1}=2\beta\nu\Gamma(0,4\nu)$$
$$+\tfrac{1}{6}\beta\nu+\exp(-4\nu)[-\tfrac{9}{16}\beta+(4+\tfrac{4}{3}\nu)\beta^6/\lambda]. \qquad (2.10)$$

For an attractive Coulomb potential ($\nu<0$) the incomplete gamma function must be replaced by its real part in both equations. The explicit derivation of these formulas will be given in Sec. 4.

For vanishing Coulomb strength: $k\gamma\to0$, i.e., $\nu\to0$, Eqs. (2.9) and (2.10) become

$$-a^{-1}_{s,l=1}=\beta^8/\lambda-\beta^3/16,$$
$$\tfrac{1}{2}r_{s,l=1}=4\beta^6/\lambda-9\beta/16.$$

These expressions just give the effective range parameters for the rank-one potential of Eq. (2.5).

3. THE EFFECTIVE RANGE FUNCTION

It is well known that in the theory of scattering by a short-range potential the so-called effective range function

$$K_l(k^2) = k^{2l+1} \cot\delta_l(k) \tag{3.1}$$

plays an important role. It has been proved that (under certain conditions on the potential) this function is real analytic at the origin. Its first two expansion coefficients are related to the scattering length a_l and the effective range r_l according to[9]

$$K_l(k^2) = -a_l^{-1} + \tfrac{1}{2} r_l k^2 - \cdots. \tag{3.2}$$

If the potential is equal to $V = V_c + V_s$, where V_c is the repulsive Coulomb potential, the effective range function is modified, and may be taken as

$$K_{cs,l}(k^2) = k^{2l+1} \binom{l+i\gamma}{l} \binom{l-i\gamma}{l} \times \left[2\gamma h(\gamma) + \frac{2\pi\gamma}{\exp(2\pi\gamma)-1} \left(\cot\delta_l^c - i\right)\right], \tag{3.3}$$

where the function $h(\gamma)$ is defined by[10]

$$h(\gamma) \equiv -2 \int_0^\infty \frac{t\,dt}{(t^2-\gamma^2)[\exp(2\pi t)-1]}, \quad \mathrm{Re}\, i\gamma > 0. \tag{3.4}$$

If the Coulomb potential is attractive the function $h(\gamma)$ should be replaced by

$$h(\gamma) + i\pi \coth\pi\gamma.$$

Now we know from the work of Cornille and Martin[4] and that of Hamilton *et al.*[3] that for certain classes of *local* potentials the function $K_{cs,l}(k^2)$ is again a real analytic function of k^2 at $k^2 = 0$ with a branch cut on (part of) the negative real axis and possibly with isolated poles in the cut complex k^2 plane. The first two expansion coefficients are related to the Coulomb-modified scattering length $a_{cs,l}$ and effective range $r_{cs,l}$,

$$K_{cs,l}(k^2) = -a_{cs,l}^{-1} + \tfrac{1}{2} r_{cs,l} k^2 - \cdots. \tag{3.5}$$

In this section we shall discuss the effective range function for the case that V_s is a *separable* interaction of finite rank with *rational* form factors. For simplicity we shall restrict ourselves mainly to the simple potential of Eqs. (2.2) and (2.5), and we shall discuss only a few examples of more general rational separable potentials.

The main purpose of the rest of this paper will be to prove for those rational separable potentials the real analyticity of $K_{cs,l}(k^2)$ at $k=0$ for real α and β, to derive a method for exact practical calculations of $K_{cs,l}$, and to investigate it in general.

First we note that the function [cf. also Eq. (27) of Ref. 4]

$$H(\gamma) \equiv \psi(i\gamma) + (2i\gamma)^{-1} - \ln[-i\gamma\, \mathrm{sgn}(s)], \tag{3.6}$$

where ψ is the digamma function, is more useful than $h(\gamma)$. Indeed, by using[11]

$$\psi(i\gamma) = \psi(-i\gamma) - (i\gamma)^{-1} + i\pi \coth\pi\gamma, \tag{3.7}$$

we find that substitution of $h(\gamma)$ by $H(\gamma)$ in Eq. (3.3) yields the general formula for $K_{cs,l}$ which is valid for a repulsive ($s<0$) as well as for an attractive ($s>0$) Coulomb potential. (Recall $k\gamma \equiv -s$, so that on the physical domain $\mathrm{Im}k>0$ we have $\mathrm{Re}\,i\gamma > 0$ if $s<0$ and $\mathrm{Re}\,i\gamma<0$ if $s>0$). Furthermore, by taking the limit $s \to 0$ in Eq. (3.3) [with $H(\gamma)$ instead of $h(\gamma)$] we obtain Eq. (3.1), independently of the sign of s. This follows from the fact that

$$\lim_{s\to 0} 2\gamma H(\gamma) = i.$$

Some useful equalities in connection with Eq. (3.3) are

$$2\pi\gamma[\exp(2\pi\gamma)-1]^{-1} = \exp(-\pi\gamma)\,\Gamma(1+i\gamma)\,\Gamma(1-i\gamma),$$

and

$$\binom{l+i\gamma}{l}\binom{l-i\gamma}{l} = \prod_{m=1}^{l} (1+\gamma^2/m^2).$$

Now we have to express $\cot\delta_l^c$ occurring in Eq. (3.3) in terms of known functions. The relation of the phase shifts σ_l and δ_l^c with the physical on-shell t matrices is as follows:

$$\begin{aligned} t_{c,l}(k) &= [i/(\pi k)] \exp(2i\sigma_l), \\ t_{cs,l}(k) &= t_{c,l}(k)[\exp(2i\delta_l^c)-1]. \end{aligned} \tag{3.8}$$

It is now easy to derive

$$\cot\delta_l^c - i = -2\exp(2i\sigma_l)/(\pi k t_{cs,l}),$$

and, with the help of Eq. (2.6) for $t_{cs,l}$, $K_{cs,l}$ is obtained in closed form. For the simple Yamaguchi-type potential of Eqs. (2.2) and (2.5) this yields

$$K_l(k^2) = 2\gamma H(\gamma)\, k^{2l+1} \binom{l+i\gamma}{l}\binom{l-i\gamma}{l} + (\beta^2+k^2)^{2l+2} B^{2i\gamma} \{\lambda_l^{-1} + \langle g_l | G_{c,l}(k^2) | g_l\rangle\}, \tag{3.9}$$

where the subscripts cs and β have been suppressed. An explicit expression for $\langle g_l | G_{c,l} | g_l\rangle$ is known[1] in the case $l=0$, and for $l=1$ we found the expression (2.4).

For a general rational separable potential the function K_l is much more complicated than the one of Eq. (3.9). We claim, however, that, for *any* rational separable potential, K_l can be expressed[12] in terms of simple real analytic functions and a certain function W which we define by

$$W(\gamma; \mu, \nu) \equiv (i\gamma)^{-1} A^{i\gamma} B^{i\gamma} [F_{i\gamma}(AB) - \tfrac{1}{2}] + H(\gamma). \tag{3.10}$$

This function depends on $\mu = k\gamma/\alpha$ and $\nu = k\gamma/\beta$ through A and B respectively. A warning is appropriate here, that $A^{i\gamma}B^{i\gamma}$ is not everywhere equal to $(AB)^{i\gamma}$. This will be discussed in Secs. 4 and 5. Note that W is independent of l and of the particular potential employed. We have worked out three examples for different types of potentials, in order to make the above conjecture plausible. These three cases describe the principal generalizations of the Yamaguchi potential that one can imagine.

(i) The rank-one potential of Eqs. (2.2) and (2.5) for all l, that is, Eq. (3.9). In this case we find

$$K_l = R_l^{(1)} + R_l^{(2)} W(\gamma; \nu, \nu), \tag{3.11}$$

with

$$R_l^{(2)} = 2\gamma k^{2l+1} \binom{l+i\gamma}{l} \binom{l-i\gamma}{l},$$

and $R_l^{(1)}$ is regular at zero energy.

(ii) A rank-one potential with form factor

$$g(p) = (p^2+\alpha^2)^{-1} + b(p^2+\beta^2)^{-1},$$

for $l=0$ only (b is a real parameter). This yields the expression

$$K_0 = R^{(0)} + R^{(\mu)} W(\gamma;\mu,\mu) + R^{(\nu)} W(\gamma;\nu,\nu) + R^{(\mu\nu)} W(\gamma;\mu,\nu). \tag{3.12}$$

(iii) A rank-two potential with form factors of the type of Eq. (2.2), for $l=0$ only. Then we get

$$K_0 = 2k\gamma \times \frac{[R_\alpha + W(\gamma;\mu,\mu)][R_\beta + W(\gamma;\nu,\nu)] - [R_{\alpha\beta} + W(\gamma;\mu,\nu)]^2}{R_\alpha + R_\beta - 2R_{\alpha\beta} + W(\gamma;\mu,\mu) + W(\gamma;\nu,\nu) - 2W(\gamma;\mu,\nu)}. \tag{3.13}$$

The R's denote simple real analytic functions. Their dependence on k (or γ) is contained in the quantities $A^{i\gamma}$, $B^{i\gamma}$, and rational functions of k^2 with real coefficients. We shall derive in Sec. 4 the equality [recall $A=(1+i\mu/\gamma)/(1-i\mu/\gamma)$]

$$A^{i\gamma} = \exp[-2\gamma \arctan(\mu/\gamma)] = \exp[(2s/k)\arctan(k/\alpha)],$$

from which it easily follows that $A^{i\gamma}$ and similarly $B^{i\gamma}$ is real analytic at $k=0$ for real α and β respectively.

Our main task will be to prove that $W(\gamma;\mu,\nu)$ is a real analytic function of γ^{-2} at $\gamma^{-2}=0$. Once this proof has been given, it is relatively easy to investigate the effective range function itself. We shall find that the only singularities of W are the branch cuts $-\infty<k^2<-\alpha^2$ and $-\infty<k^2<-\beta^2$. The only additional singularities of K_l can be (isolated) poles of finite order, i.e., K_l is a meromorphic function in the cut k^2 plane. The position of these poles depends on the particular potential and cannot be predicted in general. We have been able to show that K_l of Eq. (3.11) is real analytic at $k=0$. However, in general even a pole at $k=0$ can occur although this is exceptional. It may be interesting to consider an example in some detail.

We take the rank-one potential of Eq. (2.5) for $l=0$, with form factor

$$g(p^2) = \left(\frac{2}{\pi}\right)^{1/2} \left(p^2 - \frac{\beta^2 s}{\beta - s}\right) (p^2+\beta^2)^{-2}. \tag{3.14}$$

With this form factor g we obtain from Eq. (2.8)

$$\langle g | k+\rangle_c = (2/\pi)^{1/2} i^{i\gamma} \Gamma(1+i\gamma) k^2 (\beta^2+k^2)^{-2} B^{-i\gamma}. \tag{3.15}$$

Utilizing Eqs. (86)—(88) of Ref. 1 we derive $\langle g|G_c|g\rangle$, and for the effective range function we get then

$$K_0 = 2k\gamma\, W(\gamma;\nu,\nu) + k^{-4}R, \tag{3.16}$$

where R is a certain real analytic function of k^2 which is regular and different from zero at $k=0$. Consequently, K_0 has a pole of fourth order at the origin.

We like to discuss a few properties of the Coulomb-modified phase shift δ_l^c. At a bound state of V_c+V_s we have in general

$$\cot\delta_l^c = i, \quad \delta_l^c = -i\infty. \tag{3.17}$$

This corresponds with the situation for a pure short-range potential, where we have

$$t_{s,l}(k) = [1 - \exp(2i\delta_{s,l})]/(i\pi k). \tag{3.18}$$

In general $t_{s,l}(k)$ has a pole at the bound state, so that

$$\cot\delta_{s,l} = i, \quad \delta_{s,l} = -i\infty. \tag{3.19}$$

Since $t_{s,l}(k)$ is always real for negative energy, it also follows from Eq. (3.18) that $\delta_{s,l}(k)$ is (purely) imaginary when k is imaginary. In contrast, this does *not* hold for $\delta_l^c(k)$; from Eq. (3.3) we see that the expression

$$(\cot\delta_l^c - i)/[\exp(2\pi\gamma) - 1] \tag{3.20}$$

must be real for negative energy because $h(\gamma)$ is then real.

Finally we note that the central function on the right-hand side of Eq. (3.10) is $F_{i\gamma}(AB)$. We have met this function before.[1,2] Its behavior at $k=0$ is particularly interesting but complicated. Now it seems that just this function appears in the off-shell T matrix formula for the Coulomb plus any rational separable potential. For the case $l=0$ this has been proved by van Haeringen and van Wageningen.[1] See also the recent paper by Bajzer.[13] Equations (2.3) and (2.4) of the present paper suggest that the same holds for $l=1$ and we have reasons to believe that it is true for all l. Moreover, it is very likely that also the *pure* Coulomb transition matrices $T_{c,l}$ contain functions $F_{i\gamma}$ with a similar structure, see, e.g., Eq. (2.1). For these reasons we devote the following section to an investigation of $F_{i\gamma}(AB)$.

4. THE FUNCTION $F_{i\gamma}(AB)$

As a first step in our proof we shall consider in this section the function $F_{i\gamma}(AB)$ for real positive k, α, and β, so γ is real. The series representation

$$F_{i\gamma}(z) = i\gamma \sum_{n=0}^{\infty} z^n/(n+i\gamma) \tag{4.1}$$

reminds us of the logarithmic function series

$$-\ln(1-z) = \sum_{n=1}^{\infty} z^n/n.$$

Indeed we have (Ref. 11, pp. 13 and 49)

$$\lim_{\gamma\to 0} (1 - F_{i\gamma}(z))/(i\gamma) = \ln(1-z)$$

and

$$\lim_{z\to 1}[\ln(1-z) + (1/i\gamma) F_{i\gamma}(z)] = -C - \psi(i\gamma) = \sum_{n=0}^{\infty} \left(\frac{1}{n+i\gamma} - \frac{1}{n+1}\right). \tag{4.2}$$

Here ψ is the digamma function as before, $\psi(z) = \Gamma'(z)/\Gamma(z)$ and $C=0.5772\ldots$ is the constant of Euler or Mascheroni. We have substituted $-C$ for $\psi(1)$.

The infinite series in Eq. (4.1) is convergent if $|z|\le 1$, $z\ne 1$. When $|z|>1$, but z not real positive, one can find an expression for $F_{i\gamma}(z)$ by applying the very useful formula [see Eq. (32) of Ref. 1]

$$F_{i\gamma}(z) + F_{-i\gamma}(1/z) = 1 + \Gamma(1+i\gamma)\Gamma(1-i\gamma)(-z)^{-i\gamma}, \tag{4.3}$$

since the series (4.1) for $F_{-i\gamma}(1/z)$ converges in that case.

Now we have $A=(\alpha+ik)/(\alpha-ik)$, $B=(\beta+ik)/(\beta-ik)$ with $\alpha,\beta,k>0$, so $A^*=A^{-1}$, $B^*=B^{-1}$ and Eq. (4.3) gives

$$2\,\mathrm{Re}F_{i\gamma}(AB)=F_{i\gamma}(AB)+F_{-i\gamma}(A^{-1}B^{-1})$$
$$=1+|\Gamma(1+i\gamma)|^2(-AB)^{-i\gamma}. \tag{4.4}$$

However, $\mathrm{Im}F_{i\gamma}(AB)$ is somewhat more complicated. In Ref. 2 we have obtained the coefficients c_0 and c_2 (in the restricted case $A=B$) of the asymptotic expansion

$$\gamma^{-1}\,\mathrm{Im}F_{i\gamma}(AB)=\mathrm{Re}(i\gamma)^{-1}F_{i\gamma}(AB)$$
$$=c_0+c_2(i\gamma)^{-2}+O(\gamma^{-4}),\quad \gamma^2\to\infty.$$

Since then we have found that the coefficients d_{2n} of the asymptotic expansion

$$A^{i\gamma}B^{i\gamma}\,\mathrm{Re}(i\gamma)^{-1}F_{i\gamma}(AB)$$
$$=d_0+d_2(i\gamma)^{-2}+d_4(i\gamma)^{-4}+O(\gamma^{-6}),\quad \gamma^2\to\infty, \tag{4.5}$$

have simpler closed expressions than c_{2n}. We shall derive d_0, d_2, and d_4 in explicit form. Let us first investigate the factor $A^{i\gamma}B^{i\gamma}$. We use for convenience the parameters μ and ν, defined before by

$$\mu\alpha\equiv\nu\beta\equiv k\gamma\equiv -s. \tag{4.6}$$

If arctan denotes the principal value determined by

$$-\tfrac{1}{2}\pi<\arctan x<\tfrac{1}{2}\pi,\quad -\infty<x<\infty,$$

then the following equalities hold for $k>0$:

$$\begin{aligned}(1/2i)\ln(-AB)&=\arctan[(k^2-\alpha\beta)/k(\alpha+\beta)]\\&=\arctan[(\mu\nu-\gamma^2)/\gamma(\mu+\nu)]\\&=-\tfrac{1}{2}\pi+\arctan(k/\alpha)+\arctan(k/\beta)\\&=\tfrac{1}{2}\pi-\arctan(\gamma/\mu)-\arctan(\gamma/\nu),\end{aligned} \tag{4.7}$$

$$\begin{aligned}(1/2i)\ln(AB)&=\arctan[k(\alpha+\beta)/(\alpha\beta-k^2)]\\&=\arctan[\gamma(\mu+\nu)/(\gamma^2-\mu\nu)]\\&=-\pi+\arctan(k/\alpha)+\arctan(k/\beta)\quad\text{if } k^2>\alpha\beta\\&=\arctan(k/\alpha)+\arctan(k/\beta)\quad\text{if } k^2<\alpha\beta.\end{aligned} \tag{4.8}$$

Therefore

$$(-AB)^{i\gamma}=\exp(\pi\gamma)A^{i\gamma}B^{i\gamma}, \tag{4.9}$$

$$\begin{aligned}(AB)^{i\gamma}&=\exp(2\pi\gamma)A^{i\gamma}B^{i\gamma}\quad\text{if } k^2>\alpha\beta,\ \text{i.e., if } \gamma^2<\mu\nu,\\&=A^{i\gamma}B^{i\gamma}\quad\text{if } k^2<\alpha\beta,\ \text{i.e., if } \gamma^2>\mu\nu,\end{aligned} \tag{4.10}$$

which shows that indeed $A^{i\gamma}B^{i\gamma}$ is not everywhere equal to $(AB)^{i\gamma}$, cf. Sec. 3. We shall work in the region defined by $0<k<\min(\alpha,\beta)$, which means that we can write $(AB)^{i\gamma}$ for $A^{i\gamma}B^{i\gamma}$, and we can expand $\arctan(k/\alpha)$ and $\arctan(k/\beta)$ at $k=0$ in the well-known way. From

$$A^{i\gamma}=\exp[-2\gamma\arctan(\mu/\gamma)]=\exp[(2s/k)\arctan(k/\alpha)] \tag{4.11}$$

we see that $A^{i\gamma}$ is real analytic at $k=0$, and so is $(AB)^{i\gamma}$. This means in particular that $A^{i\gamma}$ and $(AB)^{i\gamma}$ are real for real k, α, and β.

After these introductory formulas we are now in the position to derive the coefficients d_{2n} of Eq. (4.5). The integral representation

$$F_{i\gamma}(AB)=i\gamma\int_0^1 t^{i\gamma-1}(1-ABt)^{-1}\,dt,$$

which holds for $\mathrm{Re}\,i\gamma>0$, can be recast into the form

$$(AB)^{i\gamma}[F_{i\gamma}(AB)-1]=i\gamma\int_0^{AB}t^{i\gamma}(1-t)^{-1}\,dt. \tag{4.12}$$

This equation is valid for $0<k<\alpha$, $0<k<\beta$. We differentiate both sides of this equation with respect to μ. By splitting the derivatives into real and imaginary parts we find, using Eq. (4.5),

$$\frac{(AB)^{i\gamma}}{\mu+\nu}\,\frac{\mu\nu-\gamma^2}{\mu^2+\gamma^2}=d_0'+d_2'(i\gamma)^{-2}+d_4'(i\gamma)^{-4}+O(\gamma^{-6}), \tag{4.13}$$

with $d_{2n}'\equiv d_{2n}'(\mu,\nu)\equiv(d/d\mu)\,d_{2n}(\mu,\nu)$. By taking the limit for $\gamma^{-2}\to 0$ of both sides of Eq. (4.13) we obtain

$$d_0'=-\exp(-2\mu-2\nu)/(\mu+\nu). \tag{4.14}$$

By inserting this expression into Eq. (4.13) and multiplying with $\exp(2\mu+2\nu)$ we get

$$\begin{aligned}\exp(2\mu+2\nu)\sum_{n=1}^{\infty}d_{2n}'(i\gamma)^{-2n}&=\frac{1}{\mu+\nu}\,[1-(AB)^{i\gamma}\exp(2\mu+2\nu)]\\&\quad+\frac{\mu}{\mu^2+\gamma^2}(AB)^{i\gamma}\exp(2\mu+2\nu).\end{aligned} \tag{4.15}$$

(The series $\sum d_{2n}'(i\gamma)^{-2n}$ converges, although $\sum d_{2n}(i\gamma)^{-2n}$ diverges.)

Now Eq. (4.8) yields

$$\begin{aligned}(AB)^{i\gamma}&\exp(2\mu+2\nu)\\&=\exp\left[2\mu+2\nu-2\gamma\arctan\frac{(\mu+\nu)/\gamma}{1-\mu\nu/\gamma^2}\right]\\&=\exp\left[(2\mu+2\nu)\left(\frac{-\mu\nu/\gamma^2}{1-\mu\nu/\gamma^2}+\frac{1}{3}\,\frac{(\mu+\nu)^2/\gamma^2}{(1-\mu\nu/\gamma^2)^3}\right.\right.\\&\qquad\left.\left.-\frac{1}{5}\,\frac{(\mu+\nu)^4/\gamma^4}{(1-\mu\nu/\gamma^2)^5}+\cdots\right)\right].\end{aligned}$$

Consequently, we have the interesting relation

$$(AB)^{i\gamma}\exp(2\mu+2\nu)=1-(2\mu+2\nu)\sum_{n=1}^{\infty}b_{2n}(i\gamma)^{-2n}, \tag{4.16}$$

where $b_{2n}\equiv b_{2n}(\mu,\nu)$ are certain symmetric *polynomials* in μ and ν, of degree $3n-1$. Therefore, the factor $1/(\mu+\nu)$ in Eq. (4.15) cancels. Upon substitution of Eq. (4.16) into Eq. (4.15) we find

$$\exp(2\mu+2\nu)\,d_{2n}'=2b_{2n}-\mu^{2n-1}+(2\mu+2\nu)\sum_{m=2}^{n}b_{2m-2}\,\mu^{2n-2m+1}.$$

So d_{2n}' is equal to $\exp(-2\mu-2\nu)$ times some polynomial in μ and ν of degree $3n-1$. For $n=1$ and $n=2$ we have obtained

$$\begin{aligned}d_2'&=\tfrac{1}{3}\exp(-2\mu-2\nu)[2\mu^2+2\nu^2-\mu(2\nu+3)],\\d_4'&=-\tfrac{1}{45}\exp(-2\mu-2\nu)[10\mu^5-2\mu^4(5\nu+24)\\&\quad+\mu^3(10\nu^2+18\nu+45)+2\mu^2\nu^2(5\nu-9)\\&\quad-2\mu\nu^3(5\nu+6)+2\nu^4(5\nu-9)].\end{aligned} \tag{4.17}$$

Integration of d_{2n}' with respect to μ yields d_{2n} up to some function of ν. Because $d_{2n}(\mu, \nu)$ must be symmetric in μ and ν, it follows from the special form of d'_{2n} that this function can only be a constant. This integration constant will be determined below.

It is easy to see that the integration of an expression of the form $\exp(z)\,\mathrm{Pol}(z)$ yields $\exp(z)$ times a polynomial of the same degree. Therefore, d_{2n} is (just as d'_{2n}) equal to $\exp(-2\mu-2\nu)$ times a polynomial of degree $3n-1$ (for $n>0$). This polynomial is now, fo course, symmetric in μ and ν. An economical way to obtain d_{2n} from d'_{2n} is to make use of the equalities

$$\frac{d}{dz}\,\Gamma(n+1, \lambda z) = -\lambda(\lambda z)^n \exp(-\lambda z)$$

and

$$\Gamma(n+1,\ z) = n!\, e_n(z)\exp(-z). \tag{4.18}$$

Here the polynomial e_n is defined by (Ref. 11, p. 338)

$$e_n(z) \equiv \sum_{m=0}^{n} z^m/m! \tag{4.19}$$

and $\Gamma(n+1, z)$ is the incomplete gamma function.

We have obtained the following explicit expressions:

$$\begin{aligned} d_0 &= \mathrm{Re}\Gamma(0, 2\mu+2\nu),\\ d_2 &= \tfrac{1}{12}\exp(-2\mu-2\nu)[1+(2\mu+2\nu)\\ &\quad + (1/2!)(2\mu+2\nu)^2 - 6(\mu^2+\nu^2)]\\ &= \tfrac{1}{12}\exp(-2\mu-2\nu)[e_2(2\mu+2\nu) - 6(\mu^2+\nu^2)],\\ d_4 &= -\tfrac{1}{120}\exp(-2\mu-2\nu)[1+(2\mu+2\nu)+(1/2!)(2\mu+2\nu)^2\\ &\quad + (1/3!)(2\mu+2\nu)^3 + (1/4!)(2\mu+2\nu)^4 + 30(\mu^4+\nu^4)\\ &\quad - \tfrac{40}{3}(\mu^5 - \mu^4\nu + \mu^3\nu^2 + \mu^2\nu^3 - \mu\nu^4 + \nu^5)]\\ &= \tfrac{1}{4}\exp(-2\mu-2\nu)[B_4 e_4(2\mu+2\nu) - \mu^4 - \nu^4\\ &\quad + \tfrac{4}{9}(\mu^5 - \mu^4\nu + \mu^3\nu^2 + \mu^2\nu^3 - \mu\nu^4 + \nu^5)]. \end{aligned} \tag{4.20}$$

The expression for d_0 is in agreement with the expression for c_0 given in Eq. (31) of Ref. 2. [Notice that $\Gamma(0, z)$ is real for real positive z. Along the negative real axis it has a branch cut, but the real part of $\Gamma(0, z)$ is continuous across this cut so that $\mathrm{Re}\Gamma(0, z)$ is well defined for $z<0$, cf. Eq. (6.3)]. It is interesting to compare Eq. (4.20) with the general formula for d_{2n} in Eq. (5.20) below. In Eq. (4.20) the correct integration constants have already been inserted in the expressions for d_2 and d_4. Now we are going to determine these constants. This will be done by considering Eq. (4.5) for $\mu, \nu \to 0$ and $\alpha, \beta \to \infty$, such that $\mu\alpha = \nu\beta = k\gamma = -s$ remains constant. Utilizing

$$\lim_{\alpha\to\infty} A^{i\gamma} = \lim_{\beta\to\infty} B^{i\gamma} = 1$$

we get from Eq. (4.5)

$$\begin{aligned} &\lim_{\substack{\mu,\nu\to 0\\ \alpha,\beta\to\infty}} \mathrm{Re}(1/(i\gamma)\,F_{i\gamma}(AB) - d_0)\\ &\quad = \lim_{\mu,\nu\to 0}(d_2(i\gamma)^{-2} + d_4(i\gamma)^{-4} + \cdots). \end{aligned} \tag{4.21}$$

From Eq. (25) of Ref. 2 [see Eq. (4.28) below] it follows that

$$\begin{aligned} 1/(i\gamma)\,F_{i\gamma}(AB) &= \psi(1) - \psi(i\gamma) - \ln(1-AB)\\ &\quad + O(1-AB),\quad AB \to 1. \end{aligned}$$

Using Eqs. (29) and (30) of Ref. 2, we get [recall that $\psi(1) = -C$]

$$\begin{aligned} &\lim_{\mu,\nu\to 0} \mathrm{Re}(1/(i\gamma)\,F_{i\gamma}(AB) - d_0) = \mathrm{Re}[\ln\gamma - \psi(i\gamma)]\\ &\quad \sim \sum_{n=1}^{\infty}(i\gamma)^{-2n} B_{2n}/(2n), \end{aligned} \tag{4.22}$$

where B_{2n} are the Bernoulli numbers. The symbol ~ denotes an asymptotic expansion. Note that the infinite series in Eq. (4.22) is divergent for all finite γ. Comparison of Eq. (4.22) with Eq. (4.21) yields

$$\lim_{\mu,\nu\to 0} d_{2n}(\mu,\nu) = B_{2n}/(2n),\quad n = 1, 2, \cdots,$$

and this determines the above mentioned constant of the μ integration, if the symmetry with respect to μ and ν is taken into account.

We summarize the results obtained so far. The coefficients $d_{2n} = d_{2n}(\mu, \nu)$ of the asymptotic expansion

$$(AB)^{i\gamma}\,\mathrm{Re}\left(\frac{1}{i\gamma}\,F_{i\gamma}(AB)\right) \sim \sum_{n=0}^{\infty} d_{2n}(i\gamma)^{-2n} \tag{4.23}$$

are symmetric functions of μ and ν. For $n>0$ we have

$$d_{2n}(\mu,\nu) = \exp(-2\mu-2\nu)\,P_{3n-1}(\mu,\nu), \tag{4.24}$$

where P is a certain polynomial of degree $3n-1$ and symmetric in μ and ν. Its value for $\mu=\nu=0$ can be expressed in terms of the Bernoulli numbers B_{2n},

$$d_{2n}(0,0) = P_{3n-1}(0,0) = B_{2n}/(2n),\quad n = 1, 2, \cdots. \tag{4.25}$$

Equation (4.20) gives d_0, d_2, and d_4 in closed form. The expressions for c_0 and c_2, which follow from $d_0(\nu,\nu)$ and $d_2(\nu,\nu)$, agree with Eqs. (31) and (32) of Ref. 2. The effective range parameters for $l=1$ given in Eqs. (2.9) and (2.10) can easily be obtained with the help of the expressions for d_0 and d_2.

We conclude this section with a few formulas that are useful for the high-energy limit: $k\to\infty$. It turns out that $\mathrm{Re}F_{i\gamma}(AB)$ can be expanded in a power series at $\gamma=0$. Let $|\gamma|<1$, $|\gamma|<|\mu|$, and $|\gamma|<|\nu|$. Apply Eqs. (4.4) and (4.7) and recall that γ, μ, and ν have equal signs since k, α, and β are real positive by definition. With the help of the well-known equality

$$\begin{aligned} &\exp(\pi\gamma)\,\Gamma(1+i\gamma)\,\Gamma(1-i\gamma)\\ &\quad = 1 + \pi\gamma + \sum_{n=1}^{\infty}(2\pi\gamma)^{2n}B_{2n}/(2n)!,\quad |\gamma|<1\\ &\quad = \sum_{n=0}^{\infty}(-2\pi\gamma)^n B_n/n!,\quad |\gamma|<1, \end{aligned} \tag{4.26}$$

the following interesting equality is readily established:

$$\begin{aligned} &2\,\mathrm{Re}F_{i\gamma}(AB)\\ &\quad = 1 + \exp\left[-2\gamma\arctan\frac{\gamma(\mu+\nu)}{\mu\nu-\gamma^2}\right]\sum_{n=0}^{\infty}(-2\pi\gamma)^n\frac{B_n}{n!},\quad |\gamma|<1. \end{aligned} \tag{4.27}$$

The exponential function here is recast into [cf. Eq. (4.7)]

$$\exp\{-2\gamma[\arctan(\gamma/\mu) + \arctan(\gamma/\nu)]\},$$

and this function can easily be expanded in powers of γ^2. Alternatively one can start from Eq. (25) of Ref. 2 which can be rewritten as

$$F_{i\gamma}(z)=i\gamma\sum_{n=0}^{\infty}\binom{n+i\gamma-1}{n}(1-z)^n$$
$$\times[\psi(n+1)-\psi(n+i\gamma)-\ln(1-z)],$$
$$|1-z|<1,\quad |\arg(1-z)|<\pi. \tag{4.28}$$

Further one has to use the Laurent expansion of the digamma function $\psi(z)$ at $z=0$ (Ref. 11, p. 13, corrected for the misprint),

$$\psi(z)=-\frac{1}{z}-C-\sum_{n=1}^{\infty}\zeta(n+1)(-z)^n,\quad 0<|z|<1, \tag{4.29}$$

where ζ is Riemann's zeta function. After a few manipulations one arrives at the expansion

$$F_{i\gamma}(AB)=1+\tfrac{1}{2}\pi\gamma+\gamma^2\left(\frac{\pi^2}{6}-\frac{1}{\mu}-\frac{1}{\nu}\right)-i\gamma\ln\left(\frac{2\gamma}{\mu}+\frac{2\gamma}{\nu}\right)$$
$$+O(\gamma^3\ln|\gamma|),\quad \gamma\to 0. \tag{4.30}$$

The real part of this expression agrees with the second-order approximation obtained from Eq. (4.27).

5. THE FUNCTIONS $W(\gamma;\mu,\nu)$ AND $W(\gamma;\xi)$

In this section we are going to investigate the function $W(\gamma;\mu,\nu)$ introduced in Eq. (3.10). We shall *assume* in this section that $W(\gamma;\mu,\nu)$ is real analytic at γ^{-2} for real μ and ν. The proof will be given in Sec. 6A. The three independent variables γ, μ, and ν are related to k, α, and β respectively through Eq. (4.6), where the strength s is supposed to be fixed.

It turns out that it is useful to investigate in addition a closely related but somewhat simpler function, which we denote by $W(\gamma;\xi)$. This function is also real analytic at $\gamma^{-2}=0$ (which will be proved in Sec. 6B). We shall obtain all its expansion coefficients in closed form. In this whole section we still take μ, ν, and ξ real, and only (k and) γ complex.

A. The function $W(\gamma;\mu,\nu)$

The defining expression for $W(\gamma;\mu,\nu)$ is obtained from Eqs. (3.6) and (3.10),

$$W(\gamma;\mu,\nu)=(i\gamma)^{-1}A^{i\gamma}B^{i\gamma}[F_{i\gamma}(AB)-\tfrac{1}{2}]$$
$$+\psi(i\gamma)+(2i\gamma)^{-1}-\ln[-i\gamma\,\mathrm{sgn}(s)], \tag{5.1}$$

where the last three terms are equal to $H(\gamma)$. Since we have assumed that W is real analytic at $\gamma^{-2}=0$, we may write

$$W(\gamma;\mu,\nu)=\sum_{n=0}^{\infty}w_{2n}(\mu,\nu)(i\gamma)^{-2n}, \tag{5.2}$$

where the coefficients $w_{2n}(\mu,\nu)$ are real symmetric functions of μ and ν.

We are interested in the radius of convergence of the expansion (5.2) and in closed expressions for $w_{2n}(\mu,\nu)$. When we take γ real for the moment, we can write

$$W(\gamma;\mu,\nu)=\mathrm{Re}[H(\gamma)]+\mathrm{Re}\{(i\gamma)^{-1}A^{i\gamma}B^{i\gamma}[F_{i\gamma}(AB)-\tfrac{1}{2}]\},$$

because $W(\gamma;\mu,\nu)$ is a real-valued function for real γ, μ, and ν. From Eq. (30) of Ref. 2 and Eq. (3.6) we have the asymptotic expansion

$$\mathrm{Re}H(\gamma)\sim-\sum_{n=1}^{\infty}(i\gamma)^{-2n}B_{2n}/(2n). \tag{5.3a}$$

Notice that the infinite series diverges for all finite γ. Further Eq. (4.23) yields

$$\mathrm{Re}\{(i\gamma)^{-1}A^{i\gamma}B^{i\gamma}[F_{i\gamma}(AB)-\tfrac{1}{2}]\}\sim\sum_{n=0}^{\infty}d_{2n}(i\gamma)^{-2n}. \tag{5.3b}$$

Addition of Eqs. (5.3a) and (5.3b) gives $W(\gamma;\mu,\nu)$. Because the asymptotic series expansions are unique, it follows that

$$w_0=d_0=\mathrm{Re}\Gamma(0,2\mu+2\nu),$$
$$w_{2n}=d_{2n}-B_{2n}/(2n),\quad n=1,2,\cdots. \tag{5.4}$$

It is interesting that the summation of the two divergent asymptotic series in Eqs. (5.3a) and (5.3b) yields a convergent series, that is, the power series (5.2).

In order to investigate the radius of convergence of the latter, we shall study the singularities of the functions occurring on the right-hand side of Eq. (5.1). We discern four sources of singularities namely those originating from

(i) $A^{i\gamma}B^{i\gamma}$,

(ii) the hypergeometric function $F_{i\gamma}(AB)$,

(iii) the logarithmic function,

(iv) the digamma function $\psi(i\gamma)$.

(i) In the first place, $A^{i\gamma}\equiv\exp(i\gamma\ln A)$ has a branch cut for real negative A and similarly $B^{i\gamma}$ for real negative B. The location of these branch cuts in the complex k plane is easily found. For real positive α and complex k we have

$$A=\frac{\alpha+ik}{\alpha-ik}=\frac{\alpha^2-|k|^2+2i\alpha\,\mathrm{Re}k}{(\alpha+\mathrm{Im}k)^2+(\mathrm{Re}k)^2}.$$

Since the denominator is clearly always positive (or zero), it follows that A is real and negative if and only if $\mathrm{Re}k=0$ and $|k|>\alpha$. The branch cut, therefore, consists of the two intervals $(-i\infty,-i\alpha)$ and $(i\alpha,i\infty)$ along the imaginary k axis. The product $A^{i\gamma}B^{i\gamma}$ has in the k plane the following four branch cuts:

$$(i\alpha,i\infty),\quad(-i\infty),-i\alpha),\quad(i\beta,i\infty),\quad(-i\infty,-i\beta). \tag{5.5}$$

We have used $(-AB)^{i\gamma}$ in Eqs. (4.4) and (4.9). It is useful to know the branch cuts of this quantity. They are determined by $\mathrm{Im}AB=0$, $\mathrm{Re}AB>0$. Now

$$\mathrm{Im}AB=D^{-1}(\alpha+\beta)(\alpha\beta-|k|^2)\,\mathrm{Re}k,$$

where D is real nonnegative. Therefore, AB is real if and only if either $|k|^2=\alpha\beta$ or $\mathrm{Re}k=0$. Assuming $\alpha<\beta$ for definiteness, one can easily verify:

$$AB>1\Longleftrightarrow k\in(-i\infty,-i\beta)\cup(-i\alpha,0), \tag{5.6a}$$
$$AB>0\Longleftrightarrow k\in(-i\infty,-i\beta)\cup(-i\alpha,i\alpha)\cup(i\beta,i\infty), \tag{5.6b}$$
$$AB<0\Longleftrightarrow k\in(-i\beta,-i\alpha)\cup(i\alpha,i\beta)\text{ or }|k|^2=\alpha\beta. \tag{5.6c}$$

Therefore $(-AB)^{i\gamma}$ has the three branch cuts of Eq. (5.6b). Furthermore, we have in the k plane, cut along the imaginary axis,

$$(-AB)^{i\gamma}=\exp(\pm\pi\gamma)A^{i\gamma}B^{i\gamma},\quad \mathrm{Re}k\gtrless 0. \tag{5.7a}$$

Note that the origin is an exceptional point here since $k=0$ is an isolated essential singularity of $\exp(\pm\pi\gamma)$.

Although $(AB)^{i\gamma}$ plays no role in the physical quantities, it is interesting to compare also this function with $A^{i\gamma}B^{i\gamma}$. The branch cut of $(AB)^{i\gamma}$ is given by Eq. (5.6c). It has a rather peculiar shape. There are four branch points namely $\pm i\alpha$, $\pm i\beta$ and the cut connecting them is only one curve, composed of a circle and two finite intervals. Inside the circle $|k|^2=\alpha\beta$ we have

$$(AB)^{i\gamma}=A^{i\gamma}B^{i\gamma}, \quad |k|^2<\alpha\beta, \tag{5.7b}$$

and, outside the circle,

$$(AB)^{i\gamma}=\exp(\pm 2\pi\gamma)A^{i\gamma}B^{i\gamma}, \quad \mathrm{Re}k \gtreqless 0, \quad |k|^2>\alpha\beta. \tag{5.7c}$$

(ii) Secondly, $F_{i\gamma}(AB)$ has a branch cut for real AB with $1<AB<\infty$. According to Eq. (5.6a) AB is real and larger than one if and only if k lies in either the interval $(-i\infty, -i\beta)$ or the interval $(-i\alpha, 0)$. The discontinuity across the cut $(-i\alpha, 0)$ of the expression

$$(i\gamma)^{-1}A^{i\gamma}B^{i\gamma}F_{i\gamma}(AB)$$

in Eq. (5.1) is equal to $2\pi i$; see Eq. (6.19) below. It is remarkable that this discontinuity is independent of α and β, despite the fact that the expression itself does depend on α and β.

Further $F_{i\gamma}(z)/\Gamma(1+i\gamma)$ is an entire function of $i\gamma$ for fixed z. This implies that $F_{i\gamma}(z)$ has simple poles at $i\gamma=-n$ for $n=1,2,\cdots$.

(iii) In the third place $\ln[-i\gamma\,\mathrm{sgn}(s)]$ yields a branch cut for $0<ik<\infty$, that is, for k on the negative imaginary axis. The discontinuity across the cut is $-2\pi i$. If we combine this with the branch cut $0<ik<\alpha$ from (ii), we see that the discontinuities cancel. Therefore, we call $(-i\alpha, 0)$ a "removable branch cut" of $W(\gamma;\mu,\nu)$. Below we shall find other removable singularities.

The above considerations lead us to the conjecture that the infinite series in Eq. (5.2) converges provided that

$$|k|<\min(\alpha,\beta), \quad \text{i.e.,} \quad |\gamma|^2>\max(\mu^2,\nu^2). \tag{5.8}$$

(iv) In the fourth place the digamma function $\psi(i\gamma)$ is a meromorphic function having simple poles at $i\gamma=-n$, $n=0,1,2,\cdots$. The pole at $i\gamma=0$ is always located outside the domain defined by Eq. (5.8). Now $F_{i\gamma}(AB)$ has simple poles at $i\gamma=-n$ for $n=1,2,\cdots$ [see (ii)]. It will be shown below that the residues of $\psi(i\gamma)$ at these poles cancel the residues of $1/(i\gamma)A^{i\gamma}B^{i\gamma}F_{i\gamma}(AB)$, despite the fact that this latter expression depends also on α and β. Consequently, $W(\gamma;\mu,\nu)$ can be made regular at these "Coulomb bound-state poles." We may call them "removable poles." The limit for $n\to\infty$ is particularly interesting. This point $i\gamma=-\infty$ is the origin $k=0$, and we will find that this singularity is removable as well.

We conclude this section now with a short derivation of the value of $W(\gamma;\mu,\nu)$ at $i\gamma=-n$, and for comparison at $i\gamma=n$, for $n=1,2,\cdots$. In the limit $n\to\infty$ we shall obtain in both cases the value w_0 which is just what we expect from Eq. (5.2).

Utilizing

$$\lim_{x\to -n}[\psi(x)+1/(x+n)]=\psi(n+1)$$

and l'Hôspital's theorem, we find

$$\lim_{i\gamma\to -n}\left[\psi(i\gamma)+\frac{z^{i\gamma}}{i\gamma}F_{i\gamma}(z)\right]$$

$$=\psi(n+1)+\ln z+\sum_{\substack{m=0\\ m\neq n}}^{\infty}\frac{z^{m-n}}{m-n}$$

$$=-C-\ln\frac{1-z}{z}+\sum_{m=1}^{n}\frac{1-z^{-m}}{m}.$$

Let $0<s<\alpha,\beta$ and so $-1<\mu,\nu<0$. (Note that the poles occur only in case of attraction, i.e., if s is positive). Then we obtain from Eq. (5.1) with the help of the above formulas,

$$\lim_{i\gamma\to -n}W(\gamma;\mu,\nu)=-C-\ln(-2\mu-2\nu)$$

$$+\ln\left[\left(\frac{n+\mu}{n}\right)\left(\frac{n+\nu}{n}\right)\right]-\sum_{m=1}^{n}{}'\frac{1}{m}\left[\left(\frac{n-\mu}{n+\mu}\,\frac{n-\nu}{n+\nu}\right)^m-1\right]. \tag{5.9a}$$

The prime in $\sum'$ means that the last term $(m=n)$ should be divided by 2. The limit for $n\to\infty$ of the right-hand side of Eq. (5.9a) must be equal to $W(\gamma;\mu,\nu)$ at $k=0$, that is, $w_0(\mu,\nu)$ for which we have the closed form (5.4). That this limit has indeed the correct value can be checked as follows. The formula

$$\lim_{n\to\infty}\sum_{m=1}^{n}\frac{1}{m}\left[\left(1+\frac{x}{n}\right)^m-1\right]=\sum_{l=1}^{\infty}\frac{x^l}{l\cdot l!}$$

can be proved by using the binomial theorem for $(1+x/n)^m$, interchanging the two finite summations and applying the equality

$$\sum_{m=j}^{n}\frac{(m-1)!}{(m-j)!}=\frac{n!}{j(n-j)!}.$$

By applying the above formula to Eq. (5.9a) we obtain

$$\lim_{n\to\infty}\lim_{i\gamma\to -n}W(\gamma;\mu,\nu)=-C-\ln(-2\mu-2\nu)$$

$$-\sum_{l=1}^{\infty}\frac{(-2\mu-2\nu)^l}{l\cdot l!}.$$

This expression is indeed equal to $w_0=d_0=\mathrm{Re}\Gamma(0,2\mu+2\nu)$ according to Eq. (29) of Ref. 2.

One should keep in mind that for $s<0$ there are no poles in the physical region, neither in $\psi(i\gamma)$ nor in $F_{i\gamma}(AB)$. So $W(\gamma;\mu,\nu)$ is regular at $i\gamma=n$, for $n=1,2,\cdots$. Nevertheless, it is interesting to calculate the value of W at $i\gamma=n$. We find

$$W(-in;\mu,\nu)=(1-z^n)/(2n)-\ln n+\psi(n)+\sum_{m=n}^{\infty}z^m/m, \tag{5.9b}$$

with

$$z=\frac{n-\mu}{n+\mu}\,\frac{n-\nu}{n+\nu}.$$

In order to derive the limit of $W(-in;\mu,\nu)$ for $n\to\infty$ we proceed as follows. First, observe that

$$n^{-1}\sum_{m=n}^{\infty}(1-x/n)^{m-1}=x^{-1}(1-x/n)^{n-1},$$

where the convergence is uniform for $0<\epsilon\le x\le n$. Integration with respect to x yields

$$\sum_{m=n}^{\infty}m^{-1}(1-x/n)^m=\int_x^n t^{-1}(1-t/n)^{n-1}\,dt, \quad 0<x\le n.$$

We now apply the inequalities

$$0 < \exp(-t) - (1-t/y)^y \leq 1/(ey), \quad 0 < t \leq y,$$

to the above integral and obtain (which follows also from Tannery's theorem)

$$\lim_{n\to\infty} \int_x^n t^{-1}(1-t/n)^{n-1}\,dt = \int_x^\infty t^{-1}\exp(-t)\,dt, \quad x>0.$$

This latter integral equals $\Gamma(0,x)$ [cf. Eq. (5.15a)], so that

$$\lim_{n\to\infty} \sum_{m=n}^{\infty} m^{-1}(1-x/n)^m = \Gamma(0,x), \quad x>0.$$

Finally we utilize the well-known fact

$$\lim_{n\to\infty}[\psi(n) - \ln n] = \lim_{n\to\infty}\left(-C - \ln n + \sum_{m=1}^{n-1} m^{-1}\right) = 0.$$

In this way we obtain

$$\lim_{n\to\infty} W(-in;\mu,\nu) = \Gamma(0, 2\mu+2\nu) = w_0(\mu,\nu),$$

where now $\mu>0$ and $\nu>0$ since $s<0$.

One should be careful in applying here the equality $W(\gamma;\mu,\nu) = W(-\gamma;\mu,\nu)$, since W depends also on the sign of s, and therefore on the sign of μ and ν. In Eq. (5.9a) μ and ν are negative, in Eq. (5.9b) μ and ν are positive. It can be shown that the expression of Eq. (5.9a) becomes equal to the expression of Eq. (5.9b) if one replaces $\ln(-2\mu-2\nu)$ by $\ln(2\mu+2\nu)$.

B. The function $W(\gamma;\xi)$

We define $W(\gamma;\xi)$ for real ξ by

$$W(\gamma;\xi) \equiv (i\gamma)^{-1}\exp(-\xi)\{F_{i\gamma}[\exp(i\xi/\gamma)] - \tfrac{1}{2}\} + H(\gamma). \tag{5.10}$$

The function $W(\gamma;\mu,\nu)$ is obtained from $W(\gamma;\xi)$ if one takes

$$\xi = 2\gamma[\arctan(\mu/\gamma) + \arctan(\nu/\gamma)].$$

One easily verifies that with this expression for ξ one has $\exp(i\xi/\gamma) = AB$ and $\exp(-\xi) = A^{i\gamma}B^{i\gamma}$, cf. Eq. (4.11). The following important equality holds, therefore:

$$W(\gamma;\mu,\nu) = W(\gamma; 2\gamma[\arctan(\mu/\gamma) + \arctan(\nu/\gamma)]). \tag{5.11}$$

Now in order to calculate $W(\gamma;\mu,\nu)$ for some value of γ, μ, and ν, one first has to calculate ξ from the above expression, and this value of ξ has to be used then in $W(\gamma;\xi)$. In Eq. (5.14) we give closed expressions for the coefficients of the power series of $W(\gamma;\xi)$ at $\gamma^{-2}=0$. Therefore, $W(\gamma;\xi)$ is useful for the exact numerical computation of $W(\gamma;\mu,\nu)$. For real μ, ν, and γ one has clearly $|\xi| < 2\pi|\gamma|$. We shall find that $W(\gamma;\xi)$ is analytic in γ on this same domain, see Eqs. (6.32) and (6.34). For values of ξ, μ, and ν which are small compared with γ the functions $W(\gamma;\xi)$ and $W(\gamma;\mu,\nu)$ are comparable according to

$$\xi \to 2\mu+2\nu:\quad W(\gamma;\xi) \approx W(\gamma;\mu,\nu). \tag{5.12}$$

Notice in particular that ξ, μ, and ν have equal signs here.

By repeating the procedure that we described for $W(\gamma;\mu,\nu)$ we obtain an asymptotic expansion and a Taylor series for the ξ functions which are the analogs of Eqs. (4.23) and (5.2) respectively. The coefficients are denoted by $d_{2n}(\xi)$ and $w_{2n}(\xi)$. At the "Coulomb-bound-state poles" $i\gamma=-n$ we have the analog of Eq. (5.9a),

$$\lim_{i\gamma\to -n} W(\gamma;\xi) = -C - \ln n - \ln[1-\exp(\xi/n)] + \xi/n - \sum_{m=1}^{n}{}' m^{-1}[\exp(-\xi m/n) - 1],$$

where $\xi<0$ since $s>0$ (compare $\mu,\nu<0$). At zero energy we get

$$\lim_{n\to\infty}\lim_{i\gamma\to -n} W(\gamma;\xi) = -C - \ln(-\xi) - \sum_{l=1}^{\infty}\frac{(-\xi)^l}{l\cdot l!} = \mathrm{Re}\,\Gamma(0,\xi) = w_0(\xi).$$

We have obtained the following simple closed formulas for $d_{2n}(\xi)$:

$$d_{2n}(\xi) = [B_{2n}/(2n)!]\,\Gamma(2n,\xi), \tag{5.13a}$$

$$= [B_{2n}/(2n)!]\exp(-\xi)\,\xi^{2n-1}\,{}_2F_0(1, 1-2n; -1/\xi), \tag{5.13b}$$

$$= [B_{2n}/(2n)]\exp(-\xi)\,e_{2n-1}(\xi), \quad n>0, \tag{5.13c}$$

$$= -[B_{2n}/(2n)]\exp(-\xi)\,L_{2n-1}^{(-2n)}(\xi), \quad n>0. \tag{5.13d}$$

Here ${}_2F_0$ is a generalized hypergeometric function, e_{2n-1} is the exponential polynomial of Eq. (4.19), and L is the generalized Laguerre polynomial. Equations (5.13) are valid for all real ξ if $n>0$, and for positive ξ if $n=0$. When ξ is negative, d_0 is equal to the real part of the right-hand side of Eqs. (5.13a) or (5.13b). For the coefficients $w_{2n}(\xi)$ of the expansion

$$W(\gamma;\xi) = \sum_{n=0}^{\infty} w_{2n}(\xi)(i\gamma)^{-2n}, \quad |\xi| < 2\pi|\gamma|,$$

we have obtained:

$$w_0(\xi) = d_0(\xi) = \mathrm{Re}\,\Gamma(0,\xi), \tag{5.14a}$$

$$w_{2n}(\xi) = -[B_{2n}/(2n)!]\,\gamma(2n,\xi), \quad n>0, \tag{5.14b}$$

$$= -[B_{2n}/(2n)!][\xi^{2n}/(2n)]\exp(-\xi)\,{}_1F_1(1;2n+1;\xi), \quad n>0, \tag{5.14c}$$

$$= -[B_{2n}/(2n)!][\xi^{2n}/(2n)]\,{}_1F_1(2n;2n+1;-\xi), \quad n>0. \tag{5.14d}$$

These Eqs. (5.13) and (5.14) have been derived with the help of Eqs. (6.31) and (6.36). The incomplete gamma functions are defined by

$$\Gamma(2n,\xi) = \int_\xi^\infty \exp(-t)\,t^{2n-1}\,dt, \tag{5.15a}$$

$$\gamma(2n,\xi) = \int_0^\xi \exp(-t)\,t^{2n-1}\,dt. \tag{5.15b}$$

One should not confuse $\gamma(2n,\xi)$ and the variable γ. From Eqs. (5.13a) and (5.14b) it follows that

$$w_{2n}(\xi) = d_{2n}(\xi) - B_{2n}/(2n), \quad n>0, \tag{5.16}$$

which should be compared with Eq. (5.4). The polynomial in Eq. (5.13c) is just a cutoff Taylor series expansion of $\exp(\xi)$. This polynomial is multiplied by $\exp(-\xi)$ and therefore we have

$$d_{2n}(\xi) = B_{2n}/(2n) + O(\xi^{2n}), \quad \xi\to 0, \tag{5.17}$$

and so

$$w_{2n}(\xi)=O(\xi^{2n}),\quad \xi\to 0, \tag{5.18}$$

which follows also from Eqs. (5.14c) and (5.14d). Now if we replace ξ by $2\mu+2\nu$ according to Eq. (5.12), we get analogous formulas for $d_{2n}(\mu,\nu)$ and $w_{2n}(\mu,\nu)$,

$$d_{2n}(\mu,\nu)=B_{2n}/(2n)+O_{2n}(\mu,\nu),\quad \mu,\nu\to 0,$$
$$w_{2n}(\mu,\nu)=O_{2n}(\mu,\nu),\qquad \mu,\nu\to 0,$$

where $O_{2n}(\mu,\nu)$ contains terms of degree $\geq 2n$ in μ and ν together. The proof of this remarkable fact will be given in Sec. 6. There we will obtain the more precise expression [cf. Eq. (6.29)]

$$w_{2n}(\mu,\nu)=-\frac{1}{2n}\left(\mu^{2n}+\nu^{2n}\right)+O_{2n+1}(\mu,\nu),\quad \mu,\nu\to 0, \tag{5.19}$$

valid for $n>0$. By combining this result with the already known properties of d_{2n} we arrive at the following interesting expression:

$$d_{2n}(\mu,\nu)=(2n)^{-1}\exp(-2\mu-2\nu)\times[B_{2n}e_{2n}(2\mu+2\nu)-(\mu^{2n}+\nu^{2n})+f_{2n}(\mu,\nu)] \tag{5.20}$$

for $n>0$. Here $f_{2n}(\mu,\nu)$ is a certain symmetric polynomial in μ and ν with the property that the degree of its terms is at least $2n+1$ and at most $3n-1$. Obviously $f_2=0$ and f_4 has only terms of degree 5. In the particular cases $n=1$ and $n=2$ we have checked Eqs. (5.19) and (5.20) explicitly. The expressions for d_2 and d_4 have been given in Eq. (4.20).

We conclude this section with two remarks.

(i) Equation (5.14d) can also be derived directly from Eq. (5.10) by using the last formula on p. 33 of Ref. 11. We note that this formula must be corrected as follows: Replace $\Gamma(m+2)$ by $(m+1)\Gamma(m+2)$. The function $F_{i\gamma}(z)$ is simply connected to Lerch's function Φ,

$$F_{i\gamma}(z)=i\gamma\,\Phi(z,1,i\gamma). \tag{5.21}$$

We have then

$$(i\gamma)^{-1}z^{i\gamma}F_{i\gamma}(z)=-C-\psi(i\gamma)-\ln(\ln z^{-1})-\sum_{n=1}^{\infty}\frac{(\ln z)^n}{n\cdot n!}B_n(i\gamma),\quad |\ln z|<2\pi, \tag{5.22}$$

where $B_n(i\gamma)$ are the Bernoulli polynomials. We insert Eq. (5.22) with $z=\exp(i\xi/\gamma)$ into Eq. (5.10). Comparison with

$$W(\gamma;\xi)=\sum_{n=0}^{\infty}(i\gamma)^{-2n}w_{2n}(\xi) \tag{5.23}$$

yields after some manipulations agreement with the expression (5.14d) for $w_{2n}(\xi)$.

(ii) The integral in Eq. (3.4) for $h(\gamma)$ is reminiscent of the Mellin–Barnes integral representation[11] of $F_{i\gamma}$,

$$F_{i\gamma}(z)=i\gamma\int_{-\infty}^{\infty}\frac{dt}{t+\gamma}\,\frac{(-z)^{it}}{2\sinh\pi t},\quad |\arg(-z)|<\pi. \tag{5.24}$$

The path of integration is chosen such that the points $t=0,\ -i,\ -2i,\cdots$, are under the contour and the point $t=-\gamma$ is above the contour. Therefore γ cannot be equal to 0, i, $2i,\cdots$. It might be that this integral is a good starting point to prove the real analyticity of $W(\gamma;\mu,\nu)$. However, we have not been able to take advantage of the similarity.

6. PROOF OF THE ANALYTICITY OF V AND W

In Sec. 6A we shall prove that $W(\gamma;\mu,\nu)$ is a real analytic function of γ^{-2} at $\gamma^{-2}=0$ when μ and ν are real. For this purpose it is convenient to introduce a closely related function $V(\gamma;\mu,\nu)$ which is analytic in the three complex variables γ, μ and ν, on the domain defined by $|\mu/\gamma|<1$, $|\nu/\gamma|<1$.

In Sec. 6B we shall introduce the function $V(\gamma;\xi)$, which is similarly related to $W(\gamma;\xi)$, and prove that it is analytic in γ and in ξ on the domain defined by $|\xi/\gamma|<2\pi$. We will obtain simple closed expressions for the expansion coefficients $v_{2n}(\xi)$ (expansion in powers of γ^{-2}), and for $\bar{v}_n(\gamma)$ (expansion in powers of ξ).

A. The functions $V(\gamma;\mu,\nu)$ and $W(\gamma;\mu,\nu)$

The function $W(\gamma;\mu,\nu)$ has been defined in Eq. (5.1) for real μ and ν. Let us first take $\nu=0$. Then we have

$$W(\gamma;\mu,0)=(i\gamma)^{-1}A^{i\gamma}[F_{i\gamma}(A)-\tfrac{1}{2}]+H(\gamma), \tag{6.1}$$

with $A=(1+i\mu/\gamma)/(1-i\mu/\gamma)$ and [Eq. (3.6)]

$$H(\gamma)=\psi(i\gamma)+1/(2i\gamma)-\ln[-i\gamma\,\mathrm{sgn}(s)].$$

Assuming first μ real, we define

$$V(\gamma;\mu,0)\equiv W(\gamma;\mu,0)+C+\ln[-2\mu\,\mathrm{sgn}(s)]. \tag{6.2}$$

The Euler constant C has been added for convenience only, but the term $\ln[-2\mu\,\mathrm{sgn}(s)]$ has the effect to cancel the singularity of $w_0(\mu,0)=\mathrm{Re}\Gamma(0,2\mu)$, see Eq. (5.4). In fact we have

$$\Gamma(0,z)+\ln z+C=-\sum_{n=1}^{\infty}\frac{(-z)^n}{n\cdot n!}. \tag{6.3}$$

This is an entire function for which we have the following useful integral representations:

$$\Gamma(0,z)+\ln z+C=\int_0^z\frac{dt}{t}\,[1-\exp(-t)]=\int_0^1\frac{dt}{t}\,[1-\exp(-zt)]. \tag{6.4}$$

The combination of Eqs. (6.1), (6.2), and (3.6) yields

$$V(\gamma;\mu,0)=(1/i\gamma)A^{i\gamma}[F_{i\gamma}(A)-\tfrac{1}{2}]+\psi(i\gamma)+1/(2i\gamma)+C+\ln(2\mu/i\gamma). \tag{6.5}$$

Notice that sgn(s) has disappeared. Below we shall find that the right-hand side of Eq. (6.5) can be analytically continued into the complex μ plane on the domain defined by

$$|\mu/\gamma|<1. \tag{6.6}$$

We shall derive now a simple integral representation for $V(\gamma;\mu,0)$ from which the analytical properties can easily be obtained.

We differentiate Eq. (6.5) with respect to μ. Utilizing the following integral representation [cf. Eq. (4.12)],

$$\frac{1}{(i\gamma)}z^{i\gamma}F_{i\gamma}(z)=\int_0^z\frac{t^{i\gamma-1}}{1-t}\,dt,\quad \mathrm{Re}\,i\gamma>0, \tag{6.7}$$

we obtain

$$\frac{d}{d\mu}\, V(\gamma;\mu,0)=\frac{1}{\mu}\left(1-\frac{A^{i\gamma}}{1+\mu^2/\gamma^2}\right). \tag{6.8}$$

Further we know that $w_0(\mu,0)=\mathrm{Re}\Gamma(0,2\mu)$ and so

$$\lim_{\mu\to 0} V(\gamma;\mu,0)=0, \tag{6.9}$$

where we utilized Eq. (6.3). From Eqs. (6.8) and (6.9) we obtain the desired integral representation

$$V(\gamma;\mu,0)=\int_0^\mu \frac{dt}{t}\left[1-\frac{1}{1+t^2/\gamma^2}\left(\frac{1+it/\gamma}{1-it/\gamma}\right)^{i\gamma}\right], \tag{6.10}$$

which can be recast into the form

$$V(\gamma;\mu,0)=\int_0^\mu \frac{dt}{t}\left[1-\frac{1}{1+t^2/\gamma^2}\exp[-2\gamma\arctan(t/\gamma)]\right] \tag{6.11}$$

or

$$V(\gamma;\mu,0)=\int_0^1 \frac{dt}{t}\left[1-\frac{1}{1+\mu^2t^2/\gamma^2}\left(\frac{1+i\mu t/\gamma}{1-i\mu t/\gamma}\right)^{i\gamma}\right]. \tag{6.12}$$

The integrand of the integral in Eq. (6.12) is analytic provided that

$$0<|t|<|\gamma/\mu|,$$

and it can be analytically continued to $t=0$. It follows that $V(\gamma;\mu,0)$ is an analytic function of γ and μ on the domain defined by Eq. (6.6), that is, $|\mu|<|\gamma|$. By making the substitution $\gamma\to-\gamma$ in either of the Eqs. (6.10)–(6.12) we see that $V(\gamma;\mu,0)$ is actually a function of γ^2 rather than of γ.

We have obtained in Eqs. (6.5) and (6.12) two important expressions for the function $V(\gamma;\mu,0)$. Since the expressions (6.5) on the one hand and (6.10)–(6.12) on the other hand look quite different, they deserve a detailed investigation. In particular we shall compare their singularities. In Eq. (6.5) we see simple poles at $i\gamma=-n$, for $n=1,2,\cdots$. They are removable and have been discussed before. Further we see a removable singularity at $\mu=0$, that is, at $A=1$. In virtue of Eq. (4.2) the function $V(\gamma;\mu,0)$ can be made continuous at $\mu=0$ with $V(\gamma;0,0)=0$.

More interesting are the nonremovable singularities which we are going to discuss now. For this purpose it is convenient to introduce the new function U by $U(\gamma;\rho)\equiv V(\gamma;\mu,0)$ with the new variable $\rho\equiv i\mu/\gamma$. We get from Eq. (6.5)

$$U(\gamma;\rho)=(i\gamma)^{-1}\left(\frac{1+\rho}{1-\rho}\right)^{i\gamma}\left[F_{i\gamma}\left(\frac{1+\rho}{1-\rho}\right)-\frac{1}{2}\right]+\ln(-2\rho)+\psi(i\gamma)+(2i\gamma)^{-1}+C, \tag{6.13}$$

and from Eq. (6.12)

$$U(\gamma;\rho)=\int_0^\rho \frac{dt}{t}\left[1-\frac{1}{1-t^2}\left(\frac{1+t}{1-t}\right)^{i\gamma}\right]. \tag{6.14}$$

Now that we have obtained these two expressions after several manipulations, we note that a second proof of their equality is obtained by using Eq. (6.7) and the following integral representation of the digamma function[11]

$$\psi(z)=-\frac{1}{2z}-\ln 2-C+\int_0^1 \frac{dt}{t}\left[1-\frac{1}{1-t^2}\left(\frac{1-t}{1+t}\right)^{z}\right],\quad \mathrm{Re}z>0. \tag{6.15}$$

From the definition of $U(\gamma;\rho)$ and the equality $V(\gamma;\mu,\nu)=V(-\gamma;\mu,\nu)$ it follows that

$$U(\gamma;\rho)=U(-\gamma;-\rho). \tag{6.16}$$

One can derive this equality easily from Eq. (6.14). It follows also from Eq. (6.13), by using Eq. (4.3) and observing that

$$\begin{aligned}U(\gamma;\rho)-U(-\gamma;-\rho)&=\psi(i\gamma)-\psi(-i\gamma)+(i\gamma)^{-1}+\ln(-\rho)-\ln(\rho)\\&\quad+(i\gamma)^{-1}\Gamma(1+i\gamma)\Gamma(1-i\gamma)\left(\frac{1-\rho}{1+\rho}\right)^{-i\gamma}\left(\frac{\rho-1}{\rho+1}\right)^{i\gamma}\end{aligned} \tag{6.17}$$

vanishes identically for all nonreal ρ. This can be derived with the help of Eq. (3.7),

$$\psi(i\gamma)-\psi(-i\gamma)+(i\gamma)^{-1}=i\pi\coth\pi\gamma.$$

By means of analytical continuation we then find that the expression of Eq. (6.17) is identically zero and this again proves Eq. (6.16).

Considered as functions of ρ, the expressions in Eqs. (6.13) and (6.14) show several branch cuts. Taken together, they must yield the same branch cut with the same discontinuity for both expressions separately. We are going to show that this is indeed true. For the purpose of this paper the discontinuity of a function f across a branch cut in a point z on the cut can be defined by

$$\mathrm{Disc.}\,f(z)\equiv\lim_{\epsilon\downarrow 0}[f(z(1+i\epsilon))-f(z(1-i\epsilon))]. \tag{6.18}$$

For example,

$$\mathrm{Disc.}\ln(-z)=-2\pi i,\quad z>0,$$

and, with the help of Eq. (4.3) [cf. also Eq. (5.22)],

$$\mathrm{Disc.}\,(i\gamma)^{-1}z^{i\gamma}F_{i\gamma}(z)=2\pi i,\quad z>1. \tag{6.19}$$

For the discontinuity across the cut $-\infty<\rho<-1$ arising from the integrand in Eq. (6.14), we obtain

$$D_\gamma\equiv D(\gamma;\rho)=2(i\gamma)^{-1}\sinh(\pi\gamma)\times\left(\frac{\rho-1}{\rho+1}\right)^{i\gamma}\left[F_{i\gamma}\left(\frac{1-\rho}{1+\rho}\right)-\frac{1}{2}\right]. \tag{6.20}$$

We need this function only for $-\infty<\rho<-1$ and for $1<\rho<\infty$, where it is regular. Equation (6.16) implies at once that the discontinuity across the cut $1<\rho<\infty$ arising from the integrand in Eq. (6.14) is equal to $D_{-\gamma}=D(-\gamma;\rho)$. One can verify with the help of Eq. (4.3) that

$$D(\gamma;\rho)+D(-\gamma;-\rho)=-2\pi i.$$

The derivation of all the discontinuities will not be given here. We summarize the results in the following scheme.

Branch cut	Discontinuity	Arising from
$-\infty<\rho<-1$	D_γ	Eq. (6.14)
$1<\rho<\infty$	$D_{-\gamma}$	Eq. (6.14)
$-\infty<\rho<-1$	D_γ	$((1+\rho)/(1-\rho))^{i\gamma}$
$1<\rho<\infty$	$D_{-\gamma}+2\pi i$	$((1+\rho)/(1-\rho))^{i\gamma}$
$0<\rho<1$	$2\pi i$	$F_{i\gamma}((1+\rho)/(1-\rho))$
$0<\rho<\infty$	$-2\pi i$	$\ln(-2\rho)$

The first two lines concern Eq. (6.14): the last four lines concern the respective expressions of Eq. (6.13). On the third line we see the same branch cut as on the first line. Furthermore, combination of the last three lines just gives the branch cut of the second line. So we see that Eqs. (6.13) and (6.14) have indeed the same branch cut structure.

So far we have studied $W(\gamma;\mu,0)$. However, our goal is the function $W(\gamma;\mu,\nu)$. The final step now is the observation that $W(\gamma;\mu,0)$ is transformed into $W(\gamma;\mu,\nu)$ by means of the substitution

$$\mu \to (\mu+\nu)/(1-\mu\nu/\gamma^2),$$

under the restriction

$$|\mu/\gamma|<1, \quad |\nu/\gamma|<1. \tag{6.21}$$

It is easy to find that this substitution yields

$$A \to AB, \quad A^{i\gamma} \to A^{i\gamma}B^{i\gamma},$$

where $B=(1+i\nu/\gamma)/(1-i\nu/\gamma)$, see Eqs. (4.8) and (5.7). So we have

$$W(\gamma;\mu,\nu)=W\left(\gamma;\frac{\mu+\nu}{1-\mu\nu/\gamma^2},\ 0\right). \tag{6.22}$$

By defining

$$V(\gamma;\mu,\nu)\equiv V\left(\gamma;\frac{\mu+\nu}{1-\mu\nu/\gamma^2},\ 0\right), \tag{6.23}$$

we get the following expressions

$$V(\gamma;\mu,\nu)=W(\gamma;\mu,\nu)+C+\ln\frac{(-2\mu-2\nu)\operatorname{sgn}(s)}{1-\mu\nu/\gamma^2} \tag{6.24}$$

$$=\frac{1}{i\gamma}A^{i\gamma}B^{i\gamma}[F_{i\gamma}(AB)-\tfrac{1}{2}]+\psi(i\gamma)+\frac{1}{2i\gamma}+C$$

$$+\ln\left(\frac{2\mu+2\nu}{i\gamma}\,\frac{1}{1-\mu\nu/\gamma^2}\right) \tag{6.25}$$

$$=\int_0^{(\mu+\nu)/(1-\mu\nu/\gamma^2)}\frac{dt}{t}\left[1-\frac{1}{1+t^2/\gamma^2}\left(\frac{1+it/\gamma}{1-it/\gamma}\right)^{i\gamma}\right]. \tag{6.26}$$

By means of changing the variable of integration according to

$$t\to\tau, \quad t=\frac{\tau(\mu+\nu)}{1-\mu\nu\tau^2/\gamma^2},$$

and, denoting τ again by t, we obtain from Eq. (6.26)

$$V(\gamma;\mu,\nu)=\int_0^1\frac{dt}{t}\,\frac{1+\mu\nu t^2/\gamma^2}{1-\mu\nu t^2/\gamma^2}$$

$$\times\left[1-\frac{(1-\mu\nu t^2/\gamma^2)^2}{(1+\mu^2t^2/\gamma^2)(1+\nu^2t^2/\gamma^2)}\right.$$

$$\left.\times\left(\frac{1+i\mu t/\gamma}{1-i\mu t/\gamma}\right)^{i\gamma}\left(\frac{1+i\nu t/\gamma}{1-i\nu t/\gamma}\right)^{i\gamma}\right]. \tag{6.27}$$

The integrand of the integral in Eq. (6.27) is analytic in t, γ, μ, and ν on the domain defined by

$$0<|t|<\min(|\gamma/\mu|,\ |\gamma/\nu|),$$

and it can be analytically continued to $t=0$, so $V(\gamma;\mu,\nu)$ is analytic in γ, μ, and ν on the domain defined by Eq. (6.21),

$$|\mu/\gamma|<1, \quad |\nu/\gamma|<1.$$

Since the integrand in Eq. (6.27) is real if t, γ, μ, and ν are real, $V(\gamma;\mu,\nu)$ is a real-valued function for real γ, μ, and ν. Consequently, $V(\gamma;\mu,\nu)$ is *real* analytic in any one of the three variables if the other two are real. We point out that the desired analytical properties of $W(\gamma;\mu,\nu)$ follow from Eq. (6.24).

Now we shall give the proof of Eq. (5.19). For this purpose we introduce the variable $\sigma\equiv i\nu/\gamma$ in addition to $\rho=i\mu/\gamma$ used before [Eq. (6.13)]. We consider the limit of $V(\gamma;\mu,\nu)$ for $\gamma,\mu,\nu\to 0$ such that ρ and σ remain constant. In view of Eq. (6.21) we have to require $|\rho|<1$, $|\sigma|<1$. From Eq. (6.26) it easily follows that

$$\lim_{\substack{\gamma,\mu,\nu\to 0\\ |\rho|<1,\,|\sigma|<1}} V(\gamma;\mu,\nu)=\int_0^{-i(\rho+\sigma)/(1+\rho\sigma)}\frac{dt}{t}\left[1-\frac{1}{1+t^2}\right]$$

$$=\tfrac{1}{2}\ln\frac{(1-\rho^2)(1-\sigma^2)}{(1+\rho\sigma)^2}. \tag{6.28}$$

The power series expansion of this expression yields

$$\lim_{\substack{\gamma,\mu,\nu\to 0\\ |\rho|<1,\,|\sigma|<1}} v_{2n}(\mu,\nu)(i\gamma)^{-2n}=-\frac{1}{2n}\,[\rho^{2n}+\sigma^{2n}-2(-\rho\sigma)^n].$$

Considering now W and w_{2n}, we observe that the term $-\ln(1-\mu\nu/\gamma^2)$ occurring in Eq. (6.24) has the effect of cancelling the term $-\frac{1}{2}\ln(1+\rho\sigma)^2$ in Eq. (6.28). Therefore, we have

$$\lim_{\substack{\gamma,\mu,\nu\to 0\\ |\rho|<1,\,|\sigma|<1}} w_{2n}(\mu,\nu)(i\gamma)^{-2n}=-\frac{1}{2n}\,(\rho^{2n}+\sigma^{2n}), \quad n>0, \tag{6.29}$$

and this proves Eq. (5.19).

Finally we report that we have utilized Eq. (6.22) to derive (for $n=0,1,2$) $w_{2n}(\mu,\nu)$ from $w_{2n}(\mu,0)$ which is much easier to obtain. This alternative method yields a check on the derivation of $d_{2n}(\mu,\nu)$ performed in Sec. 4, see Eqs. (4.17) and (4.25). Since $w_0(\mu,0)=d_0(\mu,0)=\Gamma(0,2\mu)$ for $\mu>0$, we have to expand

$$\Gamma\left(0,\ \frac{2\mu+2\nu}{1-\mu\nu/\gamma^2}\right)$$

in powers of γ^{-2}. The expansion is carried out with the help of the addition theorem for the incomplete gamma functions (Ref. 11, p. 341):

$$\Gamma(a,x)-\Gamma(a,x+y)=\gamma(a,x+y)-\gamma(a,x)$$

$$=\exp(-x)\,x^{a-1}\sum_{n=0}^{\infty}(-x)^{-n}(1-a)_n$$

$$\times[1-\exp(-y)\,e_n(y)], \quad |y|<|x|,$$

which implies in particular

$$\Gamma(0,x+y)=\Gamma(0,x)$$

$$+\exp(-x)\left[-\frac{y}{x}+\frac{y^2}{x^2}\,\frac{1+x}{2}\right]+O(y^3), \quad y\to 0.$$

B. The functions $V(\gamma;\xi)$ and $W(\gamma;\xi)$

The function $W(\gamma;\xi)$ has been defined in Eq. (5.10) for real ξ,

$$W(\gamma;\xi)=(i\gamma)^{-1}\exp(-\xi)[F_{i\gamma}(\exp(i\xi/\gamma))-\tfrac{1}{2}]+H(\gamma).$$

Repeating the procedure of Sec. 6A, we get

$$V(\gamma;\xi)=W(\gamma;\xi)+C+\ln[-\xi\operatorname{sgn}(s)], \tag{6.30a}$$

and

$$V(\gamma;\xi) = (1/i\gamma)\exp(-\xi)[F_{i\gamma}(\exp(i\xi/\gamma)) - \tfrac{1}{2}] + \psi(i\gamma) + 1/(2i\gamma) + C + \ln(\xi/i\gamma). \qquad (6.30\text{b})$$

The latter expression can be analytically continued into the complex ξ plane. We obtain

$$\frac{d}{d\xi} V(\gamma;\xi) = \frac{1}{\xi} - \frac{1}{2\gamma} \exp(-\xi)\cot\frac{\xi}{2\gamma},$$

and

$$\lim_{\xi\to 0} V(\gamma;\xi) = 0,$$

so

$$V(\gamma;\xi) = \int_0^1 \frac{dt}{t}\left[1 - \exp(-\xi t)\frac{\xi t}{2\gamma}\cot\frac{\xi t}{2\gamma}\right]. \qquad (6.31)$$

This integral representation implies that $V(\gamma;\xi)$ is analytic in γ and in ξ on the domain defined by

$$|\xi/\gamma| < 2\pi. \qquad (6.32)$$

Let us briefly consider the singularities outside the domain (6.32). From Eq. (6.30b) we see that $\ln[\xi/(i\gamma)]$ yields a branch cut $0 < i\xi/\gamma < \infty$, and $F_{i\gamma}$ yields a branch cut $1 < \exp(i\xi/\gamma) < \infty$. It turns out that the first branch cut can be removed and that the actual branch cuts are given by

$$0 < i\xi/\gamma + 2\pi i n < \infty, \quad n = \pm 1, \pm 2, \cdots. \qquad (6.33)$$

So we have in the plane of the complex variable ξ/γ a set of branch cuts consisting of vertical lines parallel to the imaginary axis, starting from the points $2\pi n$ ($n = \pm 1, \pm 2, \cdots$) on the real axis and going downwards. Only the negative imaginary axis itself ($n = 0$) is a removable branch cut.

The series expansion in powers of $(i\gamma)^{-2}$ converges if Eq. (6.32) is satisfied,

$$V(\gamma;\xi) = \sum_{n=0}^{\infty} v_{2n}(\xi)(i\gamma)^{-2n}, \quad |\xi| < 2\pi|\gamma|. \qquad (6.34)$$

The coefficients v_{2n} are closely related to the coefficients w_{2n} of W [see Eq. (5.14)],

$$v_0(\xi) = \Gamma(0,\xi) + \ln\xi + C = -\sum_{n=1}^{\infty} \frac{(-\xi)^n}{n \cdot n!},$$

$$v_{2n}(\xi) = w_{2n}(\xi), \quad n > 0. \qquad (6.35)$$

The closed expressions for $v_{2n}(\xi)$ follow easily from Eq. (6.31) if one utilizes the expansion

$$z\cot z = \sum_{n=0}^{\infty} (-)^n (2z)^{2n} B_{2n}/(2n)!, \quad |z| < \pi. \qquad (6.36)$$

The coefficients of the expansion in powers of ξ can also be obtained in closed form. We write

$$V(\gamma;\xi) = -\sum_{n=1}^{\infty} \bar{v}_n(\gamma)(-\xi)^n/n!, \quad |\xi| < 2\pi|\gamma|. \qquad (6.37)$$

Starting from Eq. (6.31) and applying the generating function of the Bernoulli polynomials $B_n(\cdot)$, we find after some manipulations [cf. Eq. (5.22)]

$$\bar{v}_n(\gamma) = 1/(2i\gamma) + (1/n)(i\gamma)^{-n} B_n(i\gamma). \qquad (6.38)$$

We know that $V(\gamma;\xi)$ and $\bar{v}_n(\gamma)$ are functions of γ^2 rather than of γ, and we show this explicitly by recasting Eqs. (6.37) and (6.38) into the form

$$V(\gamma;\xi) = -\sum_{n=1}^{\infty} \frac{(-\xi)^n}{n \cdot n!} \sum_{m=0}^{[n/2]} \binom{n}{2m} (i\gamma)^{-2m} B_{2m}. \qquad (6.39)$$

Here $[n/2]$ means the integral part of $n/2$ and we have used the relation

$$B_n(x) = \sum_{m=0}^{n} \binom{n}{m} x^{n-m} B_m \qquad (6.40)$$

and the fact that $B_m = 0$ for $m = 3, 5, 7, \cdots$.

7. SUMMARY AND DISCUSSION

We have given in Sec. 2 closed expressions for $T_{c,l=1}(p,p';k^2)$, for $T_{cs,l=1}(p,p';k^2)$ and for a_1 and r_1, corresponding to $V_c + V_s$ with V_s the $l=1$ Yamaguchi-type potential. In Eq. (3.6) we introduced the function

$$H(\gamma) = \psi(i\gamma) + (2i\gamma)^{-1} - \ln[-i\gamma\,\mathrm{sgn}(s)],$$

which replaces the often used function $h(\gamma)$ in the definition of the effective range function if the Coulomb potential V_c is attractive. When V_c is repulsive, $H(\gamma)$ is identical to $h(\gamma)$. The effective range functions corresponding to $V_c + V_{rs}$ for several rational separable potentials V_{rs} have been discussed in Sec. 3, and the function $W(\gamma;\mu,\nu)$, which plays the central role here, has been introduced.

In Sec. 4 we investigated $F_{i\gamma}(AB)$. This section concludes with some formulas useful for the high-energy limit, $k\to\infty$. In Sec. 5 we studied $W(\gamma;\mu,\nu)$ of Eq. (3.10) and an auxiliary function $W(\gamma;\xi)$; see Eq. (5.10). This function is very useful for numerical computations, due to the relationship

$$W(\gamma;\mu,\nu) = W(\gamma;2\gamma[\arctan(\mu/\gamma) + \arctan(\nu/\gamma)]) \qquad (5.11)$$

and the fact that we found simple closed expressions for the expansion coefficients $w_{2n}(\xi)$ of $W(\gamma;\xi)$. The equality

$$w_{2n} = d_{2n} - B_{2n}/(2n)$$

holds for $w_{2n}(\xi)$ and $d_{2n}(\xi)$ as well as for $w_{2n}(\mu,\nu)$ and $d_{2n}(\mu,\nu)$. We have obtained

$$d_{2n}(\xi) = (2n)^{-1}\exp(-\xi)\,B_{2n}\,e_{2n-1}(\xi) \qquad (5.13\text{c})$$

and

$$d_{2n}(\mu,\nu) = (2n)^{-1}\exp(-2\mu - 2\nu) \times [B_{2n}e_{2n}(2\mu + 2\nu) - (\mu^{2n} + \nu^{2n}) + f_{2n}(\mu,\nu)]. \qquad (5.20)$$

In Eq. (4.20) d_0, d_2, and d_4 have been given explicitly.

In Sec. 6 we have proved that $W(\gamma;\mu,\nu)$ is a real analytic function of γ^2 for real μ and ν. The only singularities in the complex γ plane are the branch cuts

$$(-i\mu, i\mu), \quad (-i\nu, i\nu).$$

They correspond in the complex k plane to the branch cuts

$$(i\alpha, i\infty), \quad (-i\infty, -i\alpha), \quad (i\beta, i\infty), \quad (-i\infty, -i\beta).$$

The power series

$$\sum_{n=0}^{\infty} (i\gamma)^{-2n} w_{2n}(\xi),$$

where ξ has the value

$$\xi = 2\gamma \arctan(\mu/\gamma) + 2\gamma \arctan(\nu/\gamma),$$

is equal to $W(\gamma;\mu,\nu)$. It converges if

$$|\xi| < 2\pi|\gamma|. \tag{6.32}$$

The map of this region into the complex k plane gives a region of convergence which is much larger than the disk $|k| < \min(\alpha, \beta)$. [However, this disk is not wholly contained in (6.32)]. In particular the whole real k axis belongs to the region of convergence determined by Eq. (6.32).

The effective range function K_l of the examples of the potentials in Sec. 3 is a real analytic function of k^2 with the branch cut $-\infty < k^2 < \max(-\alpha^2, -\beta^2)$ and possibly with isolated poles of finite order. The position of these poles depends on the particular potential and is in general difficult to predict. In Eq. (3.11) K_l is regular at $k=0$. We have also given an example of a potential for which K_0 has a *pole* at $k=0$ [Eqs. (3.14)–(3.16)]. For a general rational separable potential with real positive β_i [Ref. 1, Eq. (97)] we conjecture that, except for the branch cut $-\infty < k^2 < \max(-\beta_1^2, \ldots, -\beta_n^2)$, the effective range function K_l is a "real-meromorphic" function of k^2 (i.e., real analytic except for a finite number of poles of finite order).

The singularities of the effective range function have thus been determined in principle. Its numerical calculation is facilitated, which is in particular due to Eq. (5.11)ff. The use of these equations is not restricted to the effective range function. They can also be applied to other quantities playing a role in the scattering by $V_c + V_{rs}$.

ACKNOWLEDGMENTS

The author is grateful to Professor R. van Wageningen for a careful reading of the manuscript and for many useful comments. This investigation forms a part of the research program of the Foundation for Fundamental Research of Matter (FOM), which is financially supported by the Netherlands Organization for Pure Scientific Research (ZWO).

[1]H. van Haeringen and R. van Wageningen, J. Math. Phys. 16, 1441 (1975).
[2]H. van Haeringen, Nucl. Phys. A 253, 355 (1975).
[3]J. Hamilton, I. Øverbö, and B. Tromborg, Nucl. Phys. B 60, 443 (1973).
[4]H. Cornille and A. Martin, Nuovo Cimento 26, 298 (1962).
[5]G. Cattapan, G. Pisent, and V. Vanzani, Nucl. Phys. A 241, 204 (1975).
[6]L. Črepinšek, C.B. Lang, H. Oberhummer, W. Plessas, and H.F.K. Zingl, Acta Phys. Austriaca 42, 139 (1975).
[7]H. van Haeringen, J. Math. Phys. 17, 995 (1976).
[8]H. van Haeringen, *Proceedings of the VII International Conference on Few Body Problems in Nuclear and Particle Physics, Delhi* (North-Holland, Amsterdam, 1976).
[9]Note that a_l has the dimension $(fm)^{2l+1}$ and that r_l has the dimension $(fm)^{1-2l}$. Sometimes a_l (or $-a_l$) is called the scattering *volume*.
[10]Here we follow the convention of Hamilton.[3] Some authors define K_l in a different way, by taking the real part of the right hand side of Eq. (3.4) for real γ. Then $\mathrm{Re}(h(\gamma))$ is often called $g(\gamma)$, or confusingly $h(\gamma)$.
[11]W. Magnus, F. Oberhettinger, and R.P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Springer-Verlag, Berlin, 1966).
[12]It may happen that the Coulomb-modified phase shifts δ_l^c vanish identically, so that the function K_l cannot be defined at all. This pathological situation must be excluded. Furthermore, a rational separable potential for general l should have normalizable form factors g, such that $p^{-l}g(p)$ is a rational function of p^2.
[13]Ž. Bajzer, Z. Physik A 278, 97 (1976).

The number of bound states of the Coulomb plus Yamaguchi potential

H. van Haeringen, C. V. M. van der Mee, and R. van Wageningen

Natuurkundig Laboratorium der Vrije Universiteit, Amsterdam, The Netherlands

(Received 15 October 1976)

It is shown that certain assertions on the number of bound states of a Coulomb plus Yamaguchi potential which Zachary [J. Math. Phys. **12**, 1379 (1971); **14**, 2018 (1973)] claims to have proved are incorrect. We prove that there are always infinitely many bound states if the Coulomb part of the potential is attractive and that, in case the Coulomb part of the potential is repulsive, there is one bound state only if the Yamaguchi potential is sufficiently attractive.

In this paper we correct some assertions which Zachary[1] claims to have proved concerning the number of bound states (in the $l=0$ partial wave projected space) for the Coulomb plus Yamaguchi potential.

We prove that the number of s wave bound states is always infinite if the Coulomb part of the potential is attractive, for a repulsive as well as for an attractive Yamaguchi potential. Zachary found (by means of numerical calculations) that the number of bound states would be 0 or 1 in this case.

In case the Coulomb part of the potential is repulsive, we prove that there is one and only one bound state if the Yamaguchi potential is sufficiently attractive, and that there is no bound state otherwise. Zachary found in this case that the number of bound states could be 0, 1, or 2. See Ref. 1, pp. 1384 and 1385.

We start with the observation that all the bound states are given by the poles of the T operator. In the notation of Ref. 2, we have $V=V_c+V_s$. Here V_s is the rank-one separable Yamaguchi potential with strength λ and range parameter β. V_s is attractive or repulsive when $\lambda>0$ or $\lambda<0$, respectively. Further, V_c is the pure Coulomb potential with strength s, V_c being attractive when $s>0$ and repulsive when $s<0$. Furthermore we shall use the variable κ which is connected to the energy by $E=-\kappa^2$, $\kappa>0$. Then

$$T=T_c+T_{cs}, \tag{1}$$

$$T_{cs}=-\frac{|g^c\rangle\langle g^c|}{\lambda^{-1}+\langle g|G_c|g\rangle}, \tag{2}$$

where g^c is the Coulomb-modified form factor. When V_c is repulsive, neither T_c nor g^c has poles. Below we shall show[3] that, when V_c is attractive, the pure Coulomb poles in T_c are cancelled by corresponding poles in T_{cs}. Then it follows that the poles of T are obtained by solving the equation

$$\lambda^{-1}+\langle g|G_c|g\rangle=0. \tag{3}$$

In the case g is the Yamaguchi form factor, the second term is known in closed form [cf. Eq. (83) of Ref. 2] and we have

$$\lambda^{-1}=\frac{1}{2\beta(\beta+\kappa)^2}\,\frac{1}{1-s/\kappa}\;{}_2F_1(1,-s/\kappa;2-s/\kappa;[(\beta-\kappa)/(\beta+\kappa)]^2). \tag{4}$$

This is essentially Eq. (31) of Zachary.

We now first consider the case that V_c is attractive, i.e., $s>0$. In that case T_c has the pure Coulomb bound-state poles at $\kappa=s/n$, $n=1, 2,\cdots$. The origin $\kappa=0$ is the limit point of these poles. However, we do not expect bound states of V_c+V_s at *these* energies. In fact, T_{cs} has poles at exactly the same points $\kappa=s/n$ and its residues cancel the residues of T_c. It follows that T_c+T_{cs} has for these values of κ "removable poles" (in the terminology of Ref. 4). This can be shown in the following way. In the neighborhood of the point $\kappa=s/n$, where we fix n for the moment, we have

$$T_c\approx\frac{G_0^{-1}|\kappa_n\rangle\langle\kappa_n|G_0^{-1}}{-\kappa^2+s^2/n^2}\quad(\kappa\approx s/n), \tag{5}$$

where $|\kappa_n\rangle$ is the pure Coulomb bound state vector. Using then

$$G_c=G_0+G_0T_cG_0\approx G_0T_cG_0,$$
$$|g^c\rangle=(1+T_cG_0)|g\rangle\approx T_cG_0|g\rangle,$$

where both approximations hold near the pure Coulomb bound state poles, we get from Eqs. (1) and (2),

$$T\approx T_c-\frac{T_cG_0|g\rangle\langle g|G_0T_c}{\langle g|G_0T_cG_0|g\rangle}. \tag{6}$$

Insertion of Eq. (5) into Eq. (6) shows that the residues of T_c and T_{cs} cancel,[3] i.e.,

$$\lim_{\kappa\to s/n}(-\kappa^2+s^2/n^2)T=0. \tag{7}$$

It is also clarifying to consider the following interesting equality, which holds without approximation,

$$\langle g|G|g\rangle^{-1}=\lambda+\langle g|G_c|g\rangle^{-1}. \tag{8}$$

Clearly, the poles of $\langle g|G_c|g\rangle$ are no poles of $\langle g|G|g\rangle$ (and vice versa) as long as $\lambda\neq0$. Furthermore, the resolvent G and the T operator have the same poles, which follows easily from

$$G=G_0+G_0TG_0.$$

We now turn to the solution of Eq. (4). All the variables in Eq. (4) are real and it follows that the whole expression is real. Due to $s>0$ we have $-s/\kappa<0$. Now it is known that ${}_2F_1(a,b;c;z)/\Gamma(c)$ is an entire analytic function of a, b, and c if z is fixed and $|z|<1$. It follows that the expression on the right-hand side of Eq. (4) has simple poles at $s/\kappa=n=1, 2, \cdots$. (These are just the pure Coulomb bound state poles.) At such a pole it behaves as[3]

$$(n-s/\kappa)^{-1}2s(\beta-s/n)^{2n-2}(\beta+s/n)^{-2n-2},$$

from which it follows that the residues have the same sign for all n. Therefore, if we vary κ from $s/(n+1)$ to s/n (i.e., between any pair of consecutive poles), that expression varies continuously from $+\infty$ to $-\infty$ and adopts every real number at least once. (Below we shall find that it adopts every real number *just* once.) This holds for every $n=1, 2,\cdots$, so Eq. (4) has infinitely many solutions for every real value of λ, i.e., there is a bound state corresponding to $\kappa=s/n$ ($n=1, 2,\cdots$) for an arbitrarily strongly repulsive or attractive Yamaguchi potential. The origin $\kappa=0$ (zero energy) is the only accumulation point of the bound state energies.

A second way to prove this, which at the same time gives more detailed information about the position of the bound state energies with respect to the pure Coulomb bound states, is to insert the completeness relation

$$\mathbb{1}=\sum_{n=1}|\kappa_n\rangle\langle\kappa_n| + \int_0^\infty dk\,k^2|k+\rangle\langle k+| \tag{9}$$

into Eq. (3). Here again $|\kappa_n\rangle$ are the bound state vectors and $|k+\rangle$ are the scattering states of the attractive pure Coulomb potential. Using then $G_c=-(\kappa^2+H_c)^{-1}$ where $H_c=H_0+V_c$, we get

$$\lambda^{-1}=\sum_{n=1}^{\infty}\frac{\langle g|\kappa_n\rangle\langle\kappa_n|g\rangle}{\kappa^2-s^2/n^2} + \int_0^\infty\frac{dk\,k^2}{\kappa^2+k^2}\langle g|k+\rangle\langle k+|g\rangle. \tag{10}$$

The integrand and each term of the infinite sum is a monotonically decreasing function of κ on each of the intervals $s/(n+1)<\kappa<s/n$, $n=0, 1,\cdots$; This can be seen either by inspection or by means of differentiation with respect to κ^2. It follows that the right-hand side of Eq. (10) is a monotonically decreasing function of κ on the above intervals. So if κ increases between any pair of adjacent poles[3] [from $s/(n+1)$ to s/n, say], the expression on the right-hand side of Eq. (10) *decreases continuously and monotonically* from $+\infty$ to $-\infty$. Therefore, Eq. (10) has for every real value of λ one and only one solution in the interval $s/(n+1)<\kappa<s/n$, for $n=1, 2,\cdots$. Furthermore, in the $n=0$ interval $s<\kappa<\infty$ there is one and only one solution for every real *positive* value of λ, since the right-hand side of Eq. (10) varies then continuously and monotonically from $+\infty$ to 0.

This means that in the case of an attractive Yamaguchi part there is just one bound state below the pure Coulomb ground state, with binding energy $E_B<-s^2$. This is the ground state of V_c+V_s. By increasing the Yamaguchi strength, $\lambda\to\infty$, we get an infinite binding energy as expected, $E_B\to\infty$. Also all other bound states of V_c+V_s, namely those with $n=2,3,\cdots$, are shifted downwards with respect to the corresponding pure Coulomb bound states. But in this case the bound states always remain above the next lower pure Coulomb bound states. On the other hand, if the Yamaguchi part is repulsive, all bound states are shifted upwards with respect to the pure Coulomb bound states, but every state remains below the next higher pure Coulomb state, no matter how strongly repulsive V_s is. This is a remarkable and quite unexpected phenomenon.

Now we consider the case that V_c is repulsive. In this case the pure Coulomb scattering states $|k+\rangle$ form a complete set in the $l=0$ space. The completeness relation now takes the form

$$1=\int_0^\infty dk\,k^2|k+\rangle\langle k+|. \tag{11}$$

Again, we insert Eq. (11) into Eq. (3) and use the fact that $G_c=-(\kappa^2+H_c)^{-1}$ with $H_c|k+\rangle=k^2|k+\rangle$. Then Eq. (3) becomes

$$\lambda^{-1}-\int_0^\infty\frac{dk\,k^2}{\kappa^2+k^2}\langle g|k+\rangle\langle k+|g\rangle=0. \tag{12}$$

The integrand is clearly real positive and it is a continuous and monotonically decreasing function of κ for $0<\kappa<\infty$. The same holds for the integral. It is maximal for $\kappa=0$. We denote the corresponding strength by λ_0,

$$\lambda_0^{-1}=\int_0^\infty dk\langle g|k+\rangle\langle k+|g\rangle. \tag{13}$$

It follows by inspection that Eq. (12) has one and only one solution if the Yamaguchi potential is sufficiently attractive, i.e., if $\lambda\geqslant\lambda_0$. When $\lambda<\lambda_0$ there is no solution and therefore no bound state. An explicit expression for λ_0 follows from

$$\beta^3/\lambda_0=\tfrac{1}{2}-2\nu\exp(4\nu)\,\Gamma(0,4\nu), \tag{14}$$

in the notation of Ref. 5 ($\nu=-s/\beta>0$). For this value of λ_0, the Coulomb-modified scattering length is infinite, $a_{cs}^{-1}=0$, see Eq. (34) of Ref. 5. We notice that

$$0<x\exp(x)\Gamma(0,x) = x\exp(x)\int_x^\infty dt\,t^{-1}\exp(-t)<1,\quad x>0,$$

so that the right-hand side of Eq. (14) is always positive and therefore $\lambda_0^{-1}>0$, cf. Eq. (13). It also follows from Eq. (14) that $\beta^3/\lambda_0<\frac{1}{2}$. This is satisfactory since it is known that for the *pure* Yamaguchi potential the bound state appears just at zero energy if the strength λ is equal to $2\beta^3$. Addition of the repulsive Coulomb potential must have the effect that $\lambda_0>2\beta^3$.

Finally we note that Eq. (12) implies that the expression on the right-hand side of Eq. (4) is a continuous and monotonically decreasing function of κ for $0<\kappa<\infty$ if $s<0$. This can be proved directly, but with considerably more effort, as follows. Starting from a well-known integral representation for the hypergeometric function, we can recast Eq. (4) into the form

$$\lambda^{-1}=2\kappa\int_0^1 dt\,t^{-s/\kappa}N^{-2}, \tag{15}$$

with

$$N=(\beta+\kappa)^2-t(\beta-\kappa)^2\geqslant 0.$$

Here we have also utilized

$${}_2F_1(1,i\gamma;2+i\gamma;z)=(1-z)\,{}_2F_1(2,1+i\gamma;2+i\gamma;z).$$

We differentiate the right-hand side of Eq. (15) with respect to κ and obtain, after a few partial integrations,

$$2\int_0^1 dt\,t^{-s/\kappa}N^{-3}\{2(\beta^2-\kappa^2)(1-t)+\ln t[(\beta+\kappa)^2+t(\beta-\kappa)^2]\}. \tag{16}$$

One easily verifies that

$$\ln t=\sum_{n=1}^{\infty}\frac{2}{2n-1}\left(\frac{t-1}{t+1}\right)^{2n-1}<2\frac{t-1}{t+1},\quad 0<t<1.$$

Substitution of this inequality shows that the integrand is dominated by

$$t^{-s/\kappa}N^{-3}\,4\kappa(t-1)[\kappa+\beta(1-t)/(1+t)],$$

which is clearly negative for $0<t<1$, so that the integral of Eq. (16) is also negative. This proves the monotonicity.

We note that almost all the assertions of this paper remain valid when the Yamaguchi potential is replaced by an arbitrary rank-one separable potential. It is not difficult to verify this. There is one important exception, however. When we discussed the solutions of Eq. (4), we assumed the energy to be negative. It can be shown, with the help of Eqs. (10) and (12), that Eq. (4) has indeed no solution for positive energy if g is the Yamaguchi form factor. That is, there is no bound state in the continuum. However, by a special choice of the form factor it is possible to construct a bound state at positive energy. Such a pathological situation will not be discussed here.

The results of this paper agree with our intuitive idea, namely that the range of V_c+V_s, being still infinite, causes an infinite number of bound states in case V_c is attractive. On the other hand, it is known that an attractive rank-one separable potential has at most one bound state at negative energy. Addition of a repulsive Coulomb potential should not change the situation.

The mistake of Zachary shows that the hypergeometric functions occurring here are complicated objects. The source of the difficulties is that the energy variable is contained in the parameters of ${}_2F_1$ as well as in its argument. In particular in the zero energy region one should be careful. Numerical calculations might fail here because the well known ordinary power series of ${}_2F_1$ converges very slowly. A method for practical calculations, in particular useful in this region, has been developed in Ref. 4.

ACKNOWLEDGMENTS

This investigation forms a part of the research program of the Foundation for Fundamental Research of Matter (FOM), which is financially supported by the Netherlands Organization for Pure Scientific Research (ZWO).

[1]W. W. Zachary, J. Math. Phys. **12**, 1379 (1971); **14**, 2018 (1973).
[2]H. van Haeringen and R. van Wageningen, J. Math. Phys. **16**, 1441 (1975).
[3]We assume that $\langle g|\kappa_n\rangle \neq 0$. When g is the Yamaguchi form factor this means that $s \neq n\beta$ for $n=2,3,\cdots$.
[4]H. van Haeringen, J. Math. Phys. **18**, 927 (1977).
[5]H. van Haeringen, Nucl. Phys. A **253**, 355 (1975).

Regge trajectories for a velocity dependent potential*

Erasmo M. Ferreira

Stanford Linear Accelerator Center, Stanford University, Stanford, California and Pontifícia Universidade Católica, Rio de Janeiro, Brazil

Ricardo Merí and Javier Sesma

Departmento de Física Teórica, Universidad de Valencia, Valencia, Spain

(Received 4 October 1976)

We study the distribution of the singularities in the complex angular momentum plane of the S matrix for a velocity-dependent potential, and note some deviations with respect to the general behavior established for static potentials. We analyze the physical implications of our results concerning the existence of bound states and resonances.

1. INTRODUCTION

The study of the analytical properties of the S matrix in potential scattering as a function of the complex angular momentum has received considerable attention since the work by Regge.[1] The book by Newton[2] provides a good review of this field. More recently, several articles studying Regge poles for local[3] and nonlocal[4] potentials have been published.

Velocity-dependent potentials were introduced and used in nuclear physics to describe the nucleon—nucleon interaction. These potentials present peculiar properties as compared with static (i.e., velocity-nondependent) potentials.[5] In a previous paper[6] we have studied the singularities of the S matrix in the complex linear momentum plane for a velocity-dependent potential. Several authors have investigated the general analytic properties of the scattering amplitude, in the complex linear and angular momentum planes, for velocity-dependent potentials. Weigel[7] has obtained the Jost functions and the equation of the Regge trajectories for a group of soluble velocity-dependent potentials as well as certain analytical properties of the S matrix and the location of its singularities at zero energy. Butera and Girardello[8] have investigated, in the case of a velocity-dependent potential, the feasibility of the Watson—Sommerfeld transformation, the validity of the Bargmann and Levinson theorems concerning bound states and a generalization of the N/D method. In order to extend our knowledge of the characteristic features of these potentials, it seemed to us interesting to investigate the distribution of singularities of the S matrix in the complex angular momentum plane for physical values of the energy. Such an analysis reveals a somewhat peculiar analytic behavior as compared to static potentials, and provides interesting information about physical aspects (bound states and resonances) of the velocity-dependent potentials.

In Sec. 2 we write down the S matrix for a simple form of velocity-dependent potential[6] and obtain some general results concerning to the location of its poles in the complex angular momentum plane. Sections 3 and 4 are devoted to the description of the Regge trajectories for low and high energies, respectively. Finally, Sec. 5 shows a numerical investigation of the poles at intermediate energies and presents some conclusions concerning bound states and resonances.

2. THE S MATRIX

Let us consider a spinless particle of mass m and energy E in a velocity-dependent potential of the form

$$V(\mathbf{r},\mathbf{p}) = A\mathbf{p}\cdot\theta(b-r)\mathbf{p}/2m, \tag{2.1}$$

which was first introduced by Razavy, Field, and Levinger[9] for the description of nuclear forces. It possesses spherical symmetry and its radial shape is a velocity-dependent square barrier ($A>0$) or well ($A<0$) of range b and intensity A. The S_l function for the l wave is given by[6]

$$S_l(k) = -\frac{kh_l^{2\prime}(kb)j_l(k'b)-(1+A)k'h_l^2(kb)j_l'(k'b)}{kh_l^{1\prime}(kb)j_l(k'b)-(1+A)k'h_l^1(kb)j_l'(k'b)}, \tag{2.2}$$

where

$$k^2 = 2mE/\hbar^2, \quad k'^2 = k^2/(1+A). \tag{2.3}$$

The general analyticity properties of the S matrix in the complex l plane have already been discussed.[7,8] $S_l(k)$ is a meromorphic function of k for fixed real l, and a meromorphic function of l for fixed k, and its poles are given by the zeros of the denominator in the right-hand side of Eq. (2.2). Let us call

$$\alpha = kb, \quad \beta = k'b, \tag{2.4}$$

$$\lambda = l + \tfrac{1}{2}, \tag{2.5}$$

and denote

$$\mathcal{H}_\lambda(s) = \alpha H^1_{\lambda+1}(\alpha)/H^1_\lambda(\alpha), \tag{2.6}$$

$$\mathcal{J}_\lambda(s') = \beta J_{\lambda+1}(\beta)/J_\lambda(\beta), \tag{2.7}$$

with

$$s = \alpha^2, \quad s' = \beta^2 = s/(1+A). \tag{2.8}$$

In our study, the adimensional parameter s, related to the energy through

$$s = 2mEb^2/\hbar^2, \tag{2.9}$$

will take only physical (i.e., real) values. The equation determining the poles of the S_l function can be written in the form

$$\mathcal{H}_\lambda(s) - (1+A)\mathcal{J}_\lambda(s') + A(\lambda - 1/2) = 0. \tag{2.10}$$

From well-known properties of the Hankel and Bessel functions, it can be seen that $\mathcal{H}_\lambda(s)$ and $\mathcal{J}_\lambda(s')$ are real functions of their arguments for real λ. Similarly, one can easily verify the relations

$$\mathcal{H}_{\bar{\lambda}}(s)=\overline{\mathcal{H}_\lambda(s)}, \quad \text{for negative } s, \tag{2.11}$$

and

$$\mathcal{J}_{\bar{\lambda}}(s')=\overline{\mathcal{J}_\lambda(s')}, \quad \text{for real } s', \tag{2.12}$$

where the bar denotes complex conjugate. These properties make evident that, for a given negative energy, the solutions of Eq. (2.10) in the complex λ plane are either real or appear in the complex conjugate pairs. This result is also obtained for static potentials.[2] Instead, there are two results, valid for static potentials, which do not apply in our case. These results concern the poles in the right λ half-plane and say: (i) At negative energies such poles may only be real, and (ii) at positive energies they can only lie above the real semiaxis. In fact, for a certain range of values of the intensity parameter A, we will find complex poles in the first and fourth quadrants for negative energies and in both the first and fourth quadrants for positive energies.

3. REGGE TRAJECTORIES AT LOW ENERGIES

In the case of low (positive or negative) energies, $|s| \ll 1$, the description of the Regge trajectories is more easily done by writing in Eq. (2.10) the expressions for $\mathcal{H}_\lambda(s)$ and $\mathcal{J}_\lambda(s')$ defined by Eqs. (2.6) and (2.7) in terms of series expansions of the Hankel and Bessel functions. We obtain

$$\begin{aligned}\mathcal{H}_\lambda(s)=2\Big\{&\sum_{n=0}^{\infty}(-s/4)^n/n!\,1(-\lambda)\cdots(-\lambda-1+n)\\ &-[sG_\lambda(s)/4(\lambda+1)]\\ &\times\sum_{n=0}^{\infty}(-s/4)^n/n!\,1(\lambda+2)\cdots(\lambda+1+n)\Big\}\\ &\times\Big\{\sum_{n=0}^{\infty}(-s/4)^n/n!\,1(-\lambda+1)\cdots(-\lambda+n)\\ &-G_\lambda(s)\sum_{n=0}^{\infty}(-s/4)^n/n!\,1(\lambda+1)\cdots(\lambda+n)\Big\}^{-1}\end{aligned} \tag{3.1}$$

and

$$\mathcal{J}_\lambda(s')=\frac{s'}{2(\lambda+1)}\,\frac{\sum_{n=0}^{\infty}(-s'/4)^n/n!\,1(\lambda+2)\cdots(\lambda+1+n)}{\sum_{n=0}^{\infty}(-s'/4)^n/n!\,1(\lambda+1)\cdots(\lambda+n)}, \tag{3.2}$$

where

$$G_\lambda(s)=(s/4)^\lambda\exp(-i\pi\lambda)\,\Gamma(1-\lambda)/\Gamma(1+\lambda). \tag{3.3}$$

Following a terminology used by Newton,[2] we shall call 0 poles the poles that tend to $\lambda=0$ and C poles those that tend to a point $\lambda_0\neq 0$, as $s\to 0$. We discuss these two cases separately.

A. 0 poles

Let us consider the possibility that Eq. (2.10) is satisfied with $\lambda\to 0$ as $s\to 0$. From Eqs. (3.1) and (3.2), it is easy to see that this possibility exists only if λ and s are related in such a way that $G_\lambda(s)$, given by Eq. (3.3), is of the form

$$G_\lambda(s)=1+O(\lambda)+O(s), \quad |\lambda|,\ |s|\ll 1. \tag{3.4}$$

This implies

$$\lambda=O([\log|s|]^{-1}), \quad |s|\ll 1. \tag{3.5}$$

Equation (2.10) then takes a form similar to that studied by Keller, Rubinow, and Goldstein.[10] Terms $O(s)$ are negligible compared to terms $O(\lambda)$, and the condition for the existence of poles becomes

$$G_\lambda(s)=1-4\lambda/A(1-2\lambda)+O(s). \tag{3.6}$$

Replacing the expression for $G_\lambda(s)$, Eq. (3.3), and taking logarithms in both sides of Eq. (3.6), we obtain

$$\begin{aligned}&\lambda[\log(s/4)-i\pi]+\log[\Gamma(1-\lambda)/\Gamma(1+\lambda)]+2n\pi i\\ &\quad\approx\log[1-4\lambda/A(1-2\lambda)], \quad n=\pm1,\ \pm2,\ \pm3,\ \cdots.\end{aligned} \tag{3.7}$$

Following a procedure similar to that used by Keller *et al.*[10] we obtain

$$\begin{aligned}\log\left(\frac{s}{4}\right)\approx&-n\frac{2\pi i}{\lambda}+\pi i-2\gamma-2\sum_{m=1}^{\infty}\frac{\zeta(2m+1)}{2m+1}\lambda^{2m}\\ &+\frac{1}{\lambda}\log\left(1-\frac{4\lambda}{A(1-2\lambda)}\right).\end{aligned} \tag{3.8}$$

Here, γ is the Euler constant and ζ is the Riemann zeta function. Retaining in explicit form terms up to the first order in λ, we have

$$\log\left(\frac{s}{4}\right)=-n\frac{2\pi i}{\lambda}+\pi i-2\gamma-\frac{4}{A}-\frac{8}{A}\left(1+\frac{1}{A}\right)\lambda+O(\lambda^2). \tag{3.9}$$

Let us call

$$\begin{aligned}&\delta(s)=-[\log|s/4|]^{-1}, \quad \theta(s)=\arg s,\\ &\phi(s)=2\gamma+4/A-(\pi-\theta)i.\end{aligned} \tag{3.10}$$

The expression for λ in terms of s, for small λ and s, can be obtained from Eq. (3.8) and is given by

$$\begin{aligned}&\lambda=n\,2\pi i\delta\{1+\phi\delta+[\phi^2+n2\pi i8(A+1)/A^2]\,\delta^2+O(\delta^3)\},\\ &\qquad |\lambda|,\ |s|\ll 1 \quad\text{and}\quad n=\pm1,\ \pm2,\ \pm3,\ \cdots.\end{aligned} \tag{3.11}$$

For negative energies ϕ is real. So, as the energy tends to zero through negative values, the poles follow trajectories given approximately by

$$\mathrm{Re}\lambda\approx-n^2\,32\pi^2[(A+1)/A^2][-\log|s/4|]^{-3}, \tag{3.12a}$$

$$\mathrm{Im}\lambda\approx n\,2\pi[-\log|s/4|]^{-1}, \tag{3.12b}$$

where

$$s\to 0^- \quad\text{and}\quad n=\pm1,\ \pm2,\ \pm3,\ \cdots.$$

For positive energies ϕ becomes complex. The threshold behavior of the trajectories is then given approximately by

$$\mathrm{Re}\lambda\approx n\,2\pi^2[-\log|s/4|]^{-2}, \tag{3.13a}$$

$$\mathrm{Im}\lambda\approx n\,2\pi[-\log|s/4|]^{-1}, \tag{3.13b}$$

where

$$s\to 0^+ \quad\text{and}\quad n=\pm1,\ \pm2,\ \pm3,\ \cdots.$$

Equations (3.12) and (3.13) allow us to describe the threshold behavior of the 0 poles. For values of the parameter A in the range $-1<A$, that is, for a barrier ($0<A$) or a "shallow" well ($-1<A<0$), as the energy approaches zero through negative values, the poles move to the origin in the λ plane coming from the second and third quadrants. Then, as the energy increases from zero towards positive values, they leave the origin moving towards the first and third quadrants. For values of A in the range $A<-1$, that is, in the case of "deep" well, the 0 poles, which lie in the first and fourth quadrants for small negative energy values, reach the origin at zero energy and enter the first and third quadrants as the energy increases. Of course, for any value of A, at negative energies, values of n with opposite sign in Eq. (3.12) correspond to complex conjugate trajectories.

We thus observe that, for $A<-1$, there occur complex poles in the right-hand side of the λ-plane for negative values of the energy. This result is peculiar of the velocity-dependent potential studied here, since it is well known that, for static potentials, poles at negative energy must be real if $\mathrm{Re}\lambda>0$.

B. *C* poles

Now we investigate the existence of poles which, as $s\to 0$, approach points $\lambda_0\neq 0$ in the λ plane. From Eqs. (3.1)–(3.3) it is clear that, except for negative integer values of λ_0, we have

$$\lim_{s'\to 0}\mathcal{G}_{\lambda_0}(s')=0, \tag{3.14}$$

$$\lim_{s\to 0}H_{\lambda_0}(s)=0, \quad \text{for } \mathrm{Re}\lambda_0<0, \tag{3.15a}$$

$$=2\lambda_0, \quad \text{for } \mathrm{Re}\lambda_0>0. \tag{3.15b}$$

In view of this we conclude that Eq. (2.10) can be satisfied only for

$$\mathrm{Re}\lambda_0>0, \quad \lambda_0=A/2(2+A). \tag{3.16}$$

So, we have one and only one C pole, which at zero energy is in a point of the plane determined by the value l_C of the angular momentum given by

$$l_0=-1/(2+A). \tag{3.17}$$

In what follows this pole will be denominated "special pole" in connection with the terminology used in Ref. 6. In view of the restriction $\mathrm{Re}\lambda_0>0$, such pole does not appear as a C pole for values of A in the range $-2<A<0$.

Let us examine the trajectory of the special pole for low energies. Following Newton,[2] we distinguish three cases:

(i) $1<\lambda_0$, i.e., $-4<A<-2$. We have in this case

$$\mathrm{Re}\lambda\approx\lambda_0[1+s/(2+A)(\lambda_0^2-1)], \tag{3.18a}$$

$$\mathrm{Im}\lambda\approx(A/(2+A)^2)|s/4|^{\lambda_0}\sin(\pi-\theta)\lambda_0 \times\Gamma(1-\lambda_0)/\Gamma(1+\lambda_0). \tag{3.18b}$$

(ii) $\lambda_0=1$, i.e., $A=-4$. Now it becomes

$$\mathrm{Re}\lambda\approx 1+(s/4)\{\log|s/4|+2\gamma-1/2\}, \tag{3.19a}$$

$$\mathrm{Im}\lambda\approx(\theta-\pi)s/4. \tag{3.19b}$$

(iii) $0<\lambda_0<1$, i.e., $A<-4$ or $0<A$. We obtain

$$\mathrm{Re}\lambda\approx\lambda_0-[A/(2+A)^2]|s/4|^{\lambda_0} \times\cos(\pi-\theta)\lambda_0\Gamma(1-\lambda_0)/\Gamma(1+\lambda_0), \tag{3.20a}$$

$$\mathrm{Im}\lambda\approx(A/(2+A)^2)|s/4|^{\lambda_0} \times\sin(\pi-\theta)\lambda_0\Gamma(1-\lambda_0)/\Gamma(1+\lambda_0). \tag{3.20b}$$

These expressions can be obtained easily from Eqs. (3.1) and (3.2), retaining only the dominant terms as $s\to 0$. The quantities θ and γ are the same as in Eq. (3.10).

The displacement of the special pole on the complex λ plane as the energy passes through zero can be described in the following way. For $A<-2$ the pole is on the real axis for a small negative value of the energy, moves to the left towards the point λ_0 as $s\to 0^-$ and then leaves the real axis, moving leftwards and downwards as the energy becomes positive. For $0<A$ the pole is moving along the real axis towards the right as $s\to 0^-$ and, at $\lambda=\lambda_0$, passes to the complex plane moving leftwards and upwards as the energy becomes positive.

It remains to examine the possible existence of poles located at negative integer values of λ for energy equal to zero. These values of λ are indeterminacy points of the S matrix,[7] where the trajectories of its zeros and of its poles intersect. At small energy values there is a pole in the neighborhood of each of the values $\lambda=-n$, $n=1,2,3,\cdots$. Let us examine the form of the Regge trajectories in the vicinity of these points.

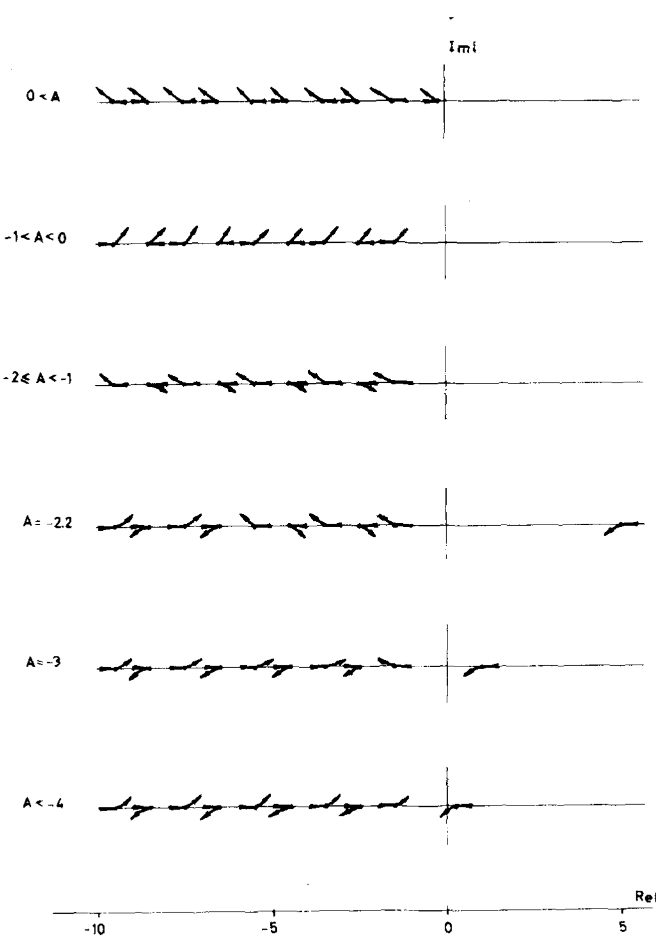

FIG. 1. Schematic threshold behavior of the C poles for a velocity-dependent potential, for different values of the intensity parameter A, as described in the text. The arrows indicate the direction of displacement of the poles as the energy goes, passing through zero, from small negative to small positive values.

For a value of λ near $-n$, retaining only the dominant terms in Eqs. (3.1)–(3.3), we obtain

$$\mathcal{H}_{-n+\epsilon}(s) \approx -2(s/4)(1-\delta_{n,1})/(n-1)+2(s/4)^n \times[\log|s/4|-i(\pi-\theta)]/[(n-1)!]^2, \qquad (3.21)$$

$$\mathcal{J}_{-n+\epsilon}(s') \approx -2n/[1-n!\,(n-1)!\,\epsilon/(s'/4)^n], \qquad (3.22)$$

where $|\epsilon|$ is an infinitesimal or order $O(s^n)$. Taking these expressions into Eq. (2.10), we obtain

$$\mathrm{Re}\epsilon \approx [(A-2n(2+A))/A\,(2n+1)n!\,(n-1)!\,]\,(s'/4)^n, \qquad (3.23a)$$

$$\mathrm{Im}\epsilon \approx (\pi-\theta)(1+A)[4/A(2n+1)((n+1)!)^2]^2(ss'/16)^n, \qquad (3.23b)$$

where s' is given by Eq. (2.8).

Equations (3.23) determine the threshold behavior of the C trajectories. In Fig. 1 we have sketched, for the different values of the intensity parameter A, the motion of the poles as the energy goes through zero, varying from small negative to small positive values.

4. REGGE TRAJECTORIES AT HIGH ENERGIES

Now let us consider the limits of infinite (positive and negative) energies, $s \to \pm\infty$. Of course, $|s'| \to \infty$. From the asymptotic expansions of the Bessel and Hankel functions for large values of the variable (see Ref. 11, p. 364) it is easy to conclude the impossibility of existence of solutions of Eq. (2.10) in the finite λ plane for infinite energies. Thus, all poles go to infinity in the λ plane as the energy increases or decreases without limit.

Equation (2.10) determining the poles can be written in the form

$$\alpha H_\lambda^{1\prime}(\alpha)/H_\lambda^1(\alpha) - (1+A)\beta J_\lambda'(\beta)/J_\lambda(\beta) + A/2 = 0. \qquad (4.1)$$

The Hankel and Bessel functions and their derivatives can be replaced by their uniform asymptotic expansions (see Ref. 11, pp. 368–69) to give respectively

$$\alpha\frac{H_\lambda'(\alpha)}{H_\lambda(\alpha)} = -\lambda^{2/3}\left(\frac{(1-z^2)}{\zeta}\right)^{1/2}\frac{\mathrm{Ai}(\rho)\lambda^{-2/3}\sum_{k=0}^\infty c_k(\zeta)/\lambda^{2k} + \exp(2\pi i/3)\,\mathrm{Ai}'(\rho)\sum_{k=0}^\infty d_k(\zeta)/\lambda^{2k}}{\mathrm{Ai}(\rho)\sum_{k=0}^\infty a_k(\zeta)/\lambda^{2k} + \exp(2\pi i/3)\mathrm{Ai}'(\rho)\sum_{k=0}^\infty b_k(\zeta)/\lambda^{2k}}, \qquad (4.2)$$

and

$$\beta\frac{J_\lambda'(\beta)}{J_\lambda(\beta)} = -\lambda^{2/3}\left(\frac{(1-z'^2)}{\xi}\right)^{1/2}\frac{\mathrm{Ai}(\rho')\lambda^{-2/3}\sum_{k=0}^\infty c_k(\xi)/\lambda^{2k} + \mathrm{Ai}'(\rho')\sum_{k=0}^\infty d_k(\xi)/\lambda^{2k}}{\mathrm{Ai}(\rho')\sum_{k=0}^\infty a_k(\xi)/\lambda^{2k} + \mathrm{Ai}'(\rho')\lambda^{-4/3}\sum_{k=0}^\infty b_k(\xi)/\lambda^{2k}}. \qquad (4.3)$$

The notation here is the same used in Ref. 11. We have denoted, for brevity,

$$\alpha = \lambda z, \quad \beta = \lambda z', \qquad (4.4)$$

$$\rho = \exp(2\pi i/3)\,\lambda^{2/3}\zeta, \quad \rho' = \lambda^{2/3}\,\xi. \qquad (4.5)$$

The relation between the variables ρ and z is given in Ref. 11, p. 368, and a similar relation holds between the variables ξ and z'. The poles move to infinity in the complex angular momentum plane as $|s| \to \infty$. However, due to the presence in Eq. (4.5) of the variables ζ and ξ which may tend to zero, the moduli of ρ and ρ' do not necessarily tend to infinity. If the modulus of ρ and/or ρ' increase without limit as $|s| \to \infty$, Eqs. (4.2) and (4.3) can be conveniently simplified. We can use the asymptotic expansions of the Airy functions (see Ref. 11, p. 448) to obtain

$$\alpha H_\lambda^{1\prime}(\alpha)/H_\lambda^1(\alpha) = -\lambda(1-z^2)^{1/2}\{1+\lambda^{-1}\zeta^{1/2} \times[-b_0(\zeta)+c_0(\zeta)\,\zeta^{-1}-\zeta^{-2}/4]+O(\lambda^{-2})\}, \qquad (4.6)$$

valid for

$$|\alpha| \to \infty, \quad |\lambda| \to \infty, \quad |\rho| \to \infty, \quad |\arg\rho| < \pi, \qquad (4.6')$$

and

$$\beta J_\lambda'(\beta)/J_\lambda(\beta) = \lambda(1-z'^2)^{1/2}\{1+\lambda^{-1}\xi^{1/2} \times[b_0(\xi)-c_0(\xi)\xi^{-1}-\xi^{-2}/4]+O(\lambda^{-2})\}, \qquad (4.7)$$

valid for

$$|\beta| \to \infty, \quad |\lambda| \to \infty, \quad |\rho'| \to \infty, \quad |\arg\rho'| < \pi. \qquad (4.7')$$

We now consider three different possibilities, namely, (i) $\lambda \approx \alpha$ in such a way that $\rho \to$ const, $|\rho'| \to \infty$, (ii) $\lambda \approx \beta$ giving $\rho' \to$ const, $|\rho| \to \infty$, and (iii) λ different from α and β, i.e., $|\rho|$, $|\rho'| \to \infty$, as $s \to \pm\infty$. We thus identify three kinds of poles which we shall call α, β, and γ poles, respectively.

A. α poles

Let us first assume that, as $|s| \to \infty$ and $|\lambda| \to \infty$, it becomes $\rho \to$ const, $|\rho'| \to \infty$. Taking Eq. (4.7) into Eq. (4.1), we obtain for the poles

$$H_\lambda^{1\prime}(\alpha)/H_\lambda^1(\alpha) - W(\lambda) = 0, \qquad (4.8)$$

where $W(\lambda)$ can be approximated by

$$W(\lambda) = z^{-1}((1+A)(1-z'^2)^{1/2} \times\{1+\lambda^{-1}[b_0(\xi)-c_0(\xi)\xi^{-1}-\xi^{-2}/4]\} - A\lambda^{-1}/2+O(\lambda^{-2})). \qquad (4.9)$$

If $W(\lambda)$ is a constant, (4.8) belongs to a type discussed by Cochran.[12] Actually in our case $W(\lambda)$ tends to a constant

$$W(\lambda) \to W = (1+A)[A/(1+A]^{1/2}, \qquad (4.10)$$

as the energy goes to infinity. Then the solutions of Eq. (4.8) are given by

$$\lambda_n = \alpha + c_n \exp(-2\pi i/3)(\alpha/2)^{1/3} + (1/60)c_n^2 \exp(-4\pi i/3)(\alpha/2)^{-1/3} + O(\alpha^{-1}), \quad (4.11)$$

where c_n denotes the nth solution of the equation

$$\mathrm{Ai}'(c_n) = (\alpha/2)^{1/3} W \exp(i\pi/3)\,\mathrm{Ai}(c_n), \quad (4.12)$$

given approximately by

$$c_n = a_n + (\alpha/2)^{-1/3}/W \exp(i\pi/3) + O(\alpha^{-2/3}), \quad (4.13)$$

where a_n stands for the nth zero of the Airy function.[13] We thus obtain that the poles tend to the points given by

$$\lambda_n = \alpha + a_n \exp(-2\pi i/3)(\alpha/2)^{1/3} - 1/W + O(\alpha^{-1/3}). \quad (4.14)$$

It is easy to see that there are poles, different from those given by Eq. (4.14), corresponding also to $\rho \to$ const, $|\rho'| \to \infty$. If we make use of the relations

$$H^1_{-\lambda}(\alpha) = \exp(\lambda\pi i) H^1_\lambda(\alpha), \quad H^{1\prime}_{-\lambda}(\alpha) = \exp(\lambda\pi i) H^{1\prime}_\lambda(\alpha), \quad (4.15)$$

in Eq. (4.8), we obtain for the pole equation

$$H^{1\prime}_{-\lambda}(\alpha)/H^1_{-\lambda}(\alpha) - W(\lambda) = 0, \quad (4.16)$$

whose solutions can be obtained similarly to those of Eq. (4.8). The values of λ so obtained are nearly the symmetrical of those given by Eq. (4.14). They are not precisely the same with opposite sign because of the fact that $W(-\lambda) \neq W(\lambda)$, and actually $W(-\lambda) = -W(\lambda) + O(\lambda^{-1})$. For negative energies, the solutions of Eq. (4.16) are exactly the complex conjugate of those of Eq. (4.8), as can be seen by writing Eq. (4.16) in the equivalent form

$$H^{1\prime}_{\bar\lambda}(\alpha)/H^1_{\bar\lambda}(\alpha) - W(\bar\lambda) = 0, \quad s < 0. \quad (4.17)$$

Summarizing, for large positive energy we have two families of trajectories which are nearly symmetrical with respect to the origin,

$$\mathrm{Re}(\pm\lambda_n) = \alpha - (a_n/2)(\alpha/2)^{1/3} \mp \mathrm{Re}(1/W) + O(\alpha^{-1/3}), \quad (4.18a)$$

$$\mathrm{Im}(\pm\lambda_n) = -(\sqrt{3}\,a_n/2)(\alpha/2)^{1/3} \mp \mathrm{Im}(1/W) + O(\alpha^{-1/3}), \quad (4.18b)$$

and pairs of complex conjugate trajectories for large negative energy

$$\mathrm{Re}\lambda_n = -\mathrm{Re}(1/W) + O(|\alpha|^{-1/3}), \quad (4.19a)$$

$$\pm\mathrm{Im}\lambda_n = |\alpha| - a_n(|\alpha|/2)^{1/3} - \mathrm{Im}(1/W) + O(|\alpha|^{-1/3}). \quad (4.19b)$$

Notice that W is pure imaginary for $-1 < A < 0$, and is real otherwise. So we find that for large negative energies the trajectories tend to rise vertically in the λ plane, approaching asymptotically the imaginary axis for $-1 < A < 0$ and the line $\mathrm{Re}\lambda = -W^{-1}$ for $A < -1$ or $0 < A$.

B. β poles

Now, let us examine the possibility that $|\rho| \to \infty$ $\rho' \to$ const, when $|s|$, $|\lambda| \to \infty$. We may substitute Eq. (4.6) in Eq. (4.1) to obtain for the pole equation

$$J'_\lambda(\beta)/J_\lambda(\beta) + V(\lambda) = 0, \quad (4.20)$$

where $V(\lambda)$ can be approximated by

$$V(\lambda) = [z'(1+A)]^{-1}((1-z^2)^{1/2} \times \{1 + \lambda^{-1}\zeta^{1/2}[-b_0(\zeta) + c_0(\zeta)\zeta^{-1} - \zeta^{-2}/4]\} - \lambda^{-1}A/2 + O(\lambda^{-2})). \quad (4.21)$$

Equation (4.20) can be treated similarly to Eq. (4.8). As in the precedent subsection, there are poles other than those given by Eq. (4.20) corresponding to $|\rho| \to \infty$, $\rho' \to$ const. Taking the complex conjugate of Eq. (4.20), we obtain

$$J'_{\bar\lambda}(\beta)/J_{\bar\lambda}(\beta) + \overline{V(\lambda)} = 0, \quad \text{for real } \beta, \quad (4.22a)$$

$$J'_{\bar\lambda}(\beta)/J_{\bar\lambda}(\beta) - \overline{V(\lambda)} = 0, \quad \text{for imaginary } \beta. \quad (4.22b)$$

At positive energies the solutions of Eq. (4.22) are not the complex conjugate of the solutions of Eq. (4.20), since $\pm\overline{V(\lambda)} \neq V(\bar\lambda)$. At negative energies, instead, Eqs. (4.22) can be written

$$J'_{\bar\lambda}(\beta)/J_\lambda(\beta) + V(\bar\lambda) = 0, \quad (4.23)$$

whose solutions are the complex conjugate of the solutions of Eq. (4.20).

Let us summarize, describing the asymptotic form of the trajectories of the β poles. Let a_n be the nth zero of the Airy function. At positive energies and for values $0 < A$, we have

$$\mathrm{Re}\lambda_n = \beta + a_n(\beta/2)^{1/3} + O(\beta^{-1/3}), \quad (4.24a)$$

$$\mathrm{Im}\lambda_n = (1+A)/A^{1/2} + O(\beta^{-1/3}); \quad (4.24b)$$

for values $-1 < A < 0$, it becomes

$$\mathrm{Re}\lambda_n = \beta + a_n(\beta/2)^{1/3} + (1+A)/(-A)^{1/2} + O(\beta^{-1/3}), \quad (4.25a)$$

$$\mathrm{Im}\lambda_n = O(\beta^{-1/3}), \quad (4.25b)$$

and for values $A < -1$, we obtain

$$\mathrm{Re}\lambda_n = (\sqrt{3}\,a_n/2)(|\beta|/2)^{1/3} \mp (1+A)/(-A)^{1/2} + O(|\beta|^{-1/3}), \quad (4.26a)$$

$$\mathrm{Im}(\pm\lambda_n) = |\beta| + (a_n/2)(|\beta|/2)^{1/3} + O(|\beta|^{-1/3}), \quad (4.26b)$$

where the signs $\mp$ in the right-hand side of Eq. (4.26a) are in correspondence with those of the left-hand side of Eq. (4.26b). At negative energies we have pairs of complex conjugate trajectories for $0 < A$,

$$\mathrm{Re}\lambda_n = (\sqrt{3}\,a_n/2)(|\beta|/2)^{1/3} + O(|\beta|^{-1/3}), \quad (4.27a)$$

$$\pm\mathrm{Im}\lambda_n = |\beta| + (a_n/2)(|\beta|/2)^{1/3} + (1+A)/A^{1/2} + O(|\beta|^{-1/3}), \quad (4.27b)$$

and for $-1 < A < 0$,

$$\mathrm{Re}\lambda_n = (\sqrt{3}\,a_n/2)(|\beta|/2)^{1/3} - (1+A)/(-A)^{1/2} + O(|\beta|^{-1/3}), \quad (4.28a)$$

$$\pm\mathrm{Im}\lambda_n = |\beta| + (a_n/2)(|\beta|/2)^{1/3} + O(|\beta|^{-1/3}), \quad (4.28b)$$

and pure real trajectories for $A < -1$,

$$\mathrm{Re}\lambda_n = \beta + a_n(\beta/2)^{1/3} + (1+A)/(-A)^{1/2} + O(\beta^{-1/3}), \quad (4.29a)$$

$$\mathrm{Im}\lambda_n = 0. \tag{4.29b}$$

C. γ poles

Finally, let us consider the possibility that $|\rho|$, $|\rho'| \to \infty$ for $|s|$, $|\lambda| \to \infty$. In this case, Eqs. (4.6) and (4.7) are simultaneously valid. The equation of the poles takes the form

$$\lambda\{(1-z^2)^{1/2}(1+\lambda^{-1}B(\xi)) + (1+A)(1-z'^2)^{1/2} \times (1+\lambda^{-1}C(\xi))\} - A/2 + O(\lambda^{-1}) = 0, \tag{4.30}$$

where we have abbreviated

$$B(\zeta) = \zeta^{1/2}[-b_0(\zeta) + c_0(\zeta)\zeta^{-1} - \zeta^{-2}/4], \tag{4.31a}$$

$$C(\xi) = \xi^{1/2}[b_0(\xi) - c_0(\xi)\xi^{-1} - \xi^{-2}/4]. \tag{4.31b}$$

Dividing by λ and grouping terms of the same order, Eq. (4.30) becomes

$$(1-z^2)^{1/2} + (1+A)(1-z'^2)^{1/2} + \lambda^{-1}[(1-z^2)^{1/2}B(\zeta) + (1+A)(1-z'^2)^{1/2}C(\xi) - A/2] + O(\lambda^{-2}) = 0. \tag{4.32}$$

It is evident that this equation can be satisfied only for $(1+A) < 0$. In this case we have

$$\lambda = \alpha/(2+A)^{1/2} - [(1+A)/A(2+A)] \times \{B(\zeta) - C(\xi) - A/2[-(1+A)]^{1/2}\} + O(\lambda^{-2}). \tag{4.33}$$

At first sight, Eq. (4.33) does not seem an explicit expression for λ in terms of α, because the right-hand side depends on λ through $B(\zeta)$ and $C(\xi)$. However, both z and z' approach constant values as α tends to infinity. In fact,

$$z^2 = 2 + A + O(\alpha^{-1}), \quad z'^2 = (2+A)/(1+A) + O(\alpha^{-1}), \tag{4.34}$$

and so, retaining only the most relevant terms, ζ and ξ can be approximated by

$$\tfrac{2}{3}\zeta^{3/2} \approx \ln(\{\{1+[-(1+A)]^{1/2}\}/(2+A)^{1/2}\}) - [-(1+A)]^{1/2}, \tag{4.35a}$$

$$\tfrac{2}{3}\xi^{3/2} \approx \ln(\{1+[-1/(1+A)]^{1/2}\}/[(2+A)/(1+A)]^{1/2}) - [-1/(1+A)]^{1/2}. \tag{4.35b}$$

Therefore, $B(\zeta)$ and $C(\xi)$ can be replaced by constants in Eq. (4.33). The values of these constants, for a given intensity parameter A, can be obtained from Eqs. (4.31) by using the expressions for ζ and ξ given by Eqs. (4.35) and the definitions of the functions b_0 and c_0 (see Ref. 11, p. 368).

5. DISCUSSION

To discuss the physical implications of the Regge trajectories and to relate 0 and C poles (at threshold) with α, β, and γ poles (asymptotically), we have investigated the location of the poles at intermediate (positive and negative) energies, using the method of "steepest descent" to obtain the numerical solutions of Eq. (2.10). $\mathcal{H}_\lambda(s)$ has been calculated through Eq. (3.1), with double precision in the sums, and $\mathcal{G}_\lambda(s')$ was evaluated using a continued fraction expansion.

In order to see the peculiarities of velocity-dependent potentials, it is interesting to compare our results with those obtained in the case of static potentials. The singularities of the S matrix for a static square potential and its physical implications have been thoroughly discussed by Nussenzveig.[3,14] Summarizing, the poles for a cutoff potential are grouped into two classes. Class-I poles are associated with the interior of the potential and can be related to bound states and reso-

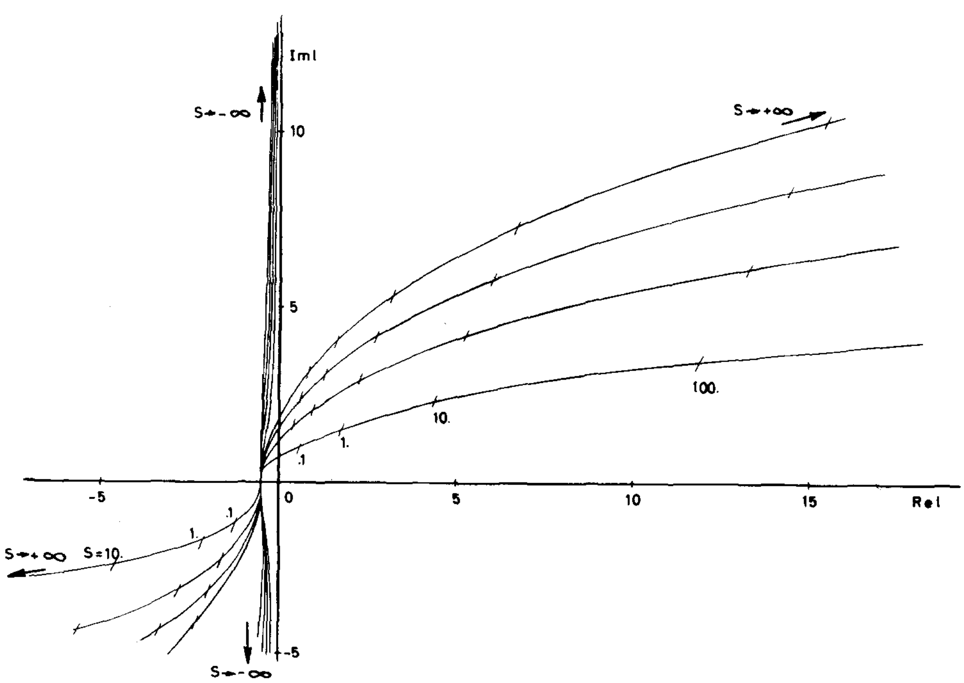

FIG. 2. Regge trajectories of the Class-II poles for a "deep" ($A = -3$) velocity-dependent square well potential. We show the first four of an infinite family of trajectories. For large negative energies the poles appear as α poles, following trajectories asymptotically tangent to the line $\mathrm{Re}\,l = -1/2 + 1/\sqrt{6}$ in both upper and lower half-planes. As the energy becomes less negative, the poles follow two complex conjugate families of trajectories, reaching the 0-point $l = -1/2$ ($\lambda = 0$) at zero energy. When the energy goes to positive values the poles leave the 0 point following two families of trajectories, one in the first quadrant and the other in the third quadrant of the λ plane. These two families are nearly symmetric with respect to the point $l = -1/2$. At large positive energies the poles appear as α poles. The numbers along the first trajectory denote the values of the energy parameter s. The transversal dashes connect poles corresponding to the same value of s.

nances. Its trajectories are C trajectories at threshold and at high energy they appear as β poles (in our terminology). Class-II poles are motivated by the sharp cutoff in the potential and correspond to surface waves in much a similar manner to those produced by a hard sphere. Its trajectories are 0 trajectories at low energy and they become α poles at high energy.

In the case of a velocity-dependent square well (or barrier) potential, we find also Class-I and Class-II poles. Class-II poles offer an aspect quite similar to those for a static well (or barrier), except for some minor differences. (For instance, for $A < -1$, at negative energies the poles lie in the first and fourth quadrants of the λ plane). As in the static case, they are not related to bound states or resonances. In Fig. 2 we show the Regge trajectories of the Class-II poles for an intensity parameter $A = -3$. The general aspect of these trajectories is about the same for other values of A.

Class-I poles follow C trajectories at low energy and become β poles at high energy, as for static potentials. However, there are some interesting differences with the static case in what concerns the bound states and resonances. Let us discuss our results.

For a velocity-dependent barrier ($0 < A$), there are no bound states and an infinite number of resonances can occur for each physical value of the angular momentum. For a low barrier ($0 < A \ll 1$), these resonances are very broad. For a high barrier ($1 \ll A$), the width of the resonances becomes roughly proportional to $AE^{1/2}$ and, in contrast with what happens for a static barrier,[14] the resonances become broader as the height A of the barrier increases. Thus a velocity-dependent barrier does not show sharp resonances, the resonance effects having its greatest intensity for $A \approx 1$.

In the case of a "shallow" velocity-dependent well ($-1 < A < 0$), there are no bound states. Each Class-I trajectory originates an infinite number of resonances, one for each physical angular momentum. These resonances become sharper as the intensity parameter A tends to -1. The trajectories of the Class-I poles for $A = -1/2$ can be seen in Fig. 3.

For a "deep" velocity-dependent well ($A < -1$), an infinite number of bound states appear for each physical angular momentum. The binding energy spectrum is unbounded. In contrast with what is found for static potentials, for a given Regge trajectory the poles with higher angular momentum have larger binding energy. There are no resonances associated to Class-I poles. In Fig. 4 we show the Regge trajectories of the Class-I poles for $A = -3$.

Besides Class-I and Class-II poles, we find in the case of our velocity-dependent potential a Class-III pole which in Ref. 6 we have denominated "special" pole. For a velocity-dependent barrier ($0 < A$), this Class-III pole has a behavior very similar to Class-I poles: it follows a C trajectory at low energy and becomes a β pole at high energy. For a "shallow" velocity dependent well ($-1 < A < 0$), it presents some properties of Class-I and Class-II poles. At low energy it appears as a Class-II pole, following a 0 trajectory; at high energy, instead, it resembles a Class-I pole, as it becomes a β pole. For a "moderately deep" velocity-dependent well ($-2 < A < -1$), the Class-III pole follows a 0 trajectory at low energy and appears as a γ pole at high energy. It originates one resonance at each physical value of l. Finally, for a "very deep" velocity-dependent well ($A < -2$), the Class-III pole has a C trajectory at low energy and behaves as a γ pole at high energy.

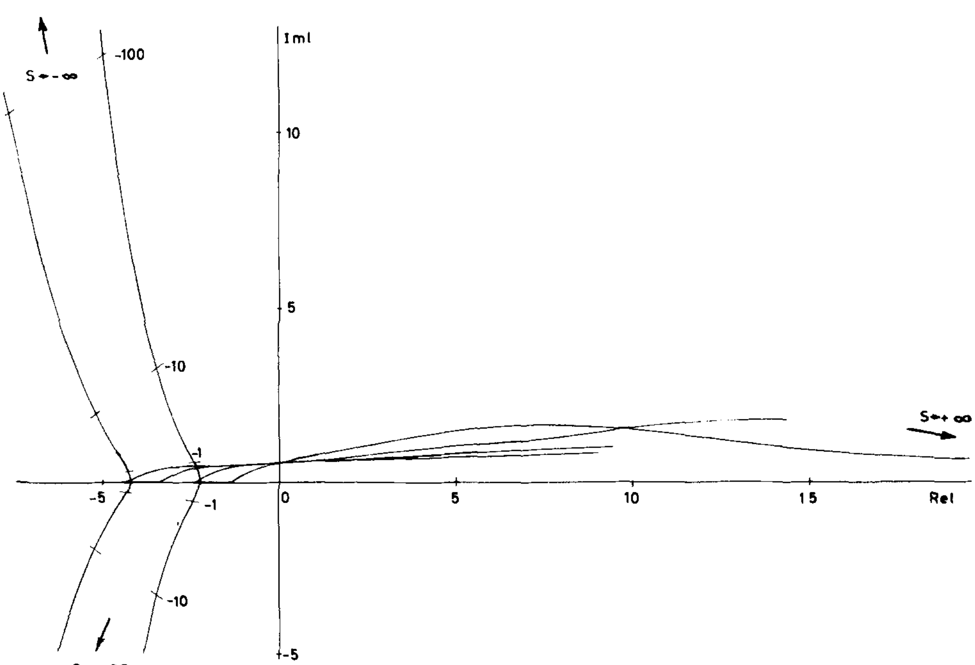

FIG. 3. Regge trajectories of the Class-I poles for a velocity-dependent "shallow" well potential ($A = -1/2$). We have drawn the first four of the infinite family of trajectories. For large negative energies the poles appear as β poles, at the infinity in the second and third quadrants. They follow two complex conjugate families of trajectories, approaching the real axis as the energy increases towards zero. The pairs of complex conjugate poles meet on the real axis for different negative values of the energy. The two poles of each pair then move in opposite directions along the real axis. At zero energy each pole reaches one of the C points, $l = -1.5, -2.5, -3.5, \cdots$. As the energy becomes positive the poles leave these C points moving rightwards and upwards. Each trajectory originates one resonance for each physical value of the angular momentum. When the energy reaches large positive values, the poles behave as β poles, following trajectories asymptotic to the real axis.

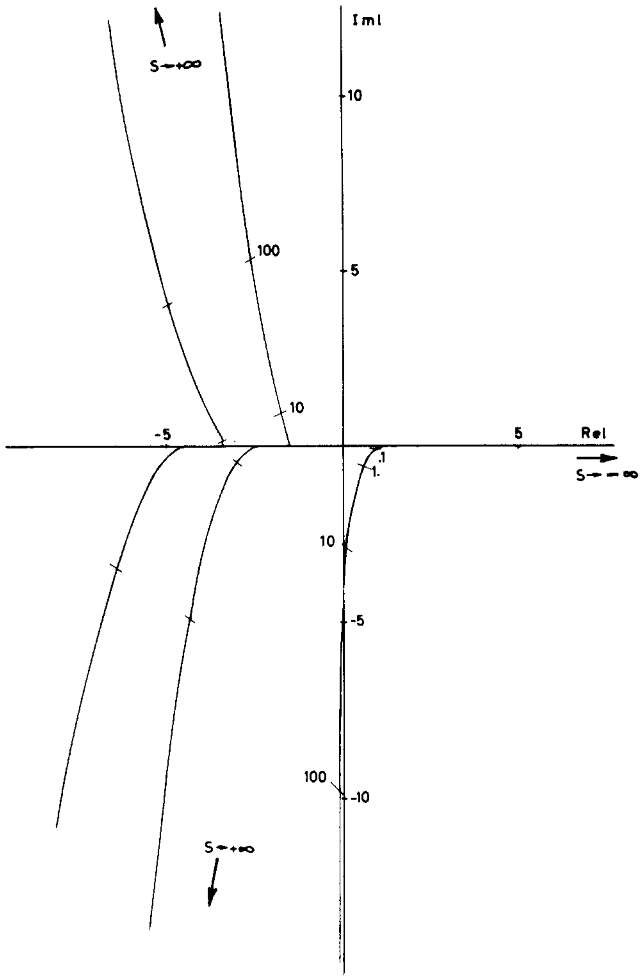

FIG. 4. Regge trajectories of the Class-I and Class-III poles for a velocity-dependent "deep" well potential ($A = -3$). We show the first four of an infinite family of Class-I trajectories and the unique Class-III trajectory. For large negative energies the Class-I poles appear, as β poles, at the infinity of the positive real semiaxis. As the energy increases towards zero, all these poles move along the real axis. Each trajectory originates one bound state of each physical value of angular momentum. At zero energy the Class-I poles reach the C points $l = -1.5, -2.5, -3.5, -4.5, \cdots$. When the energy takes positive values, the poles leave the C points in the form described qualitatively in Fig. 1. As the energy reaches large positive values, the poles move, along two nearly complex conjugate families of trajectories, towards the infinity of the second and third quadrants, where they behave as β poles. The Class-III pole is located for large negative energies also at the infinity of the positive real semiaxis. However, it behaves as a γ pole. As the energy goes to zero, the pole moves leftwards, along the real axis, causing one bound state at each physical angular momentum such that $l > 1$. At zero energy the pole reaches the point $l = 1$, thus giving a P-wave bound state of binding energy equal to zero or a zero energy resonance. As the energy takes positive values, the pole leaves the point $l = 1$ moving leftwards and downwards, and originates a broad S-wave resonance. Finally, at large positive energy, it becomes a γ pole, moving downwards along an asymptotically vertical trajectory.

This Class-III pole originates one bound state for each physical angular momentum $l > 1/(-A-2)$ and a possible resonance at each physical l such that $l < 1/(-A-2)$. These resonances become broader as l decreases. If the parameter $-A$ has a value slightly smaller than $2 + 1/l$, with l integer, there occurs a sharp resonance of angular momentum l. The Regge trajectory associated with those resonances shows a peculiar feature. As the energy increases from zero, the pole leaves the real axis moving downwards and towards the left. So, the resonances are associated with poles in the fourth quadrant of the λ plane. For these poles we have $\text{Im}\lambda < 0$ and $d(\text{Re}\lambda)/dE < 0$, giving for the resonance a positive width. In Fig. 4 we show the trajectory of the Class-III pole for $A = -3$.

To conclude, it is interesting to point out the main difference between our results for the velocity-dependent square potential and the results obtained for a static one.[14] A velocity-dependent square barrier does not originate sharp resonances, in contrast with what happens for a static barrier. A velocity-dependent square well can give an infinity of sharp resonances and no bound states (shallow well) or only some resonances and an infinity of bound states (deep well), whereas a static well has a limited number of bound states and no sharp resonances.

ACKNOWLEDGMENTS

We are grateful to the staff of Centro de Cálculo de la Universidad de Zaragoza for the facilities given to us in the use of the computational equipment. It is a pleasure to acknowledge the John Simon Guggenheim Memorial Foundation (E. M. F.) and the Ministerio de Educación y Ciencia (R. M.) for fellowships granted to the authors.

*This work was supported by Instituto de Estudios Nucleares, Spain.

[1]T. Regge, Nuovo Cimento **14**, 951 (1959).
[2]R. G. Newton, *The Complex j-Plane* (Benjamin, New York, 1964).
[3]H. M. Nussenzveig, Ann. Phys. **34**, 23 (1965); J. Math. Phys. **10**, 83, 125 (1969); F. F. K. Cheung, Phys. Lett. B **30**, 257 (1969); H. H. Aly, K. Schilcher, and H. J. W. Müller, Lett. Nuovo Cimento **1**, 707 (1969); H. H. Aly and P. Narayanaswamy, Phys. Lett. B **28**, 603 (1969); Lett. Nuovo Cimento **2**, 729 (1969); H. H. Aly and H. J. W. Müller, Lett. Nuovo Cimento **4**, 675 (1970); E. M. Ferreira and J. Sesma, J. Math. Phys. **11**, 3245 (1970); P. W. Johnson, J. Math. Phys. **12**, 1610 (1971); R. Kronenfeld, Am. J. Phys. **39**, 1056 (1971); J. B. Delos and C. E. Carlson, Phys. Rev. A **11**, 210 (1975); R. O. Mastalir, J. Math. Phys. **16**, 743, 749, 752 (1975); S. K. Bose, A. Jabs and H. J. W. Müller-Kirsten, Phys. Rev. D **13**, 1489 (1976).
[4]D. Gutkowski and A. Scalia, J. Math. Phys. **10**, 2306 (1969).
[5]E. M. Ferreira, N. Guillen, and J. Sesma, J. Math. Phys. **8**, 2243 (1967).
[6]E. M. Ferreira, N. Guillen, and J. Sesma, J. Math. Phys. **9**, 1210 (1968).
[7]M. Weigel, Z. Physik **185**, 186, 199 (1965).
[8]P. Butera and L. Girardello, Nuovo Cimento A **54**, 127, 141 (1969).
[9]M. Razavy, G. Field and J. S. Levinger, Phys. Rev. **125**, 269 (1962).
[10]J. B. Keller, S. I. Rubinow, and M. Goldstein, J. Math. Phys. **4**, 829 (1963).
[11]M. Abramowitz and I. Stegun, Eds., *Handbook of Mathematical Functions* (Dover, New York, 1965).
[12]J. A. Cochran, Numer. Math. **7**, 238 (1965).
[13]V. B. Headley and V. K. Barwell, Math. Comput. **29**, 863 (1975).
[14]H. M. Nussenzveig, *Causality and Dispersion Relations* (Academic, New York, 1972).

Phase-space approach to relativistic quantum mechanics. I. Coherent-state representation for massive scalar particles

Gerald Kaiser

Department of Mathematics, University of Toronto, Toronto, Canada M5S 1A1
(Received 23 November 1976)

We construct a family of equivalent representations U_λ ($\lambda > 0$) of the restricted Poincaré group $\mathcal{P}_+^\uparrow$ for a massive scalar particle on spaces $\mathcal{K}_\lambda$ of functions defined over "phase space" P_λ. Each P_λ is a submanifold of the forward tube, and $\mathcal{K}_\lambda$ consists of restrictions on holomorphic solutions of the Klein–Gordon equation to P_λ. Each $\mathcal{K}_\lambda$ has a resolution of the identity in terms of "coherent states" e_z, $z \in P_\lambda$, which are wavepackets characterized by an invariant extremal property.

1. INTRODUCTION

This is the first in a series of papers devoted to a phase-space formulation of relativistic quantum mechanics. In this paper we construct representations of the "coherent-state" type for a free massive scalar particle. In forthcoming papers we extend the present formalism to particles with spin, supply our "phase spaces" with natural symplectic structures, and formulate a covariant phase-space quantization. The results of this paper were announced in Ref. 1.

We begin by sketching the coherent-state representation.

In addition to the well-known configuration-space and momentum-space representations of quantum mechanics for a nonrelativistic particle, there is a class of representations on spaces of functions over classical phase space,[2-7] the most common of which is known as the "coherent-state" representation. The simplest such representation[8] is constructed as follows: let X_k and P_k be the position and momentum operators for a particle in R^n ($k=1,\ldots,n$) and form the nonnormal operators $a_k = X_k + iP_k$. These are found to have an overcomplete set of eigenvectors $e_z : a_k e_z = \bar{z}_k e_z$ (the bar denotes complex conjugation), one for each $z = x - iy \in C^n$, and each e_z is a minimum-uncertainty wave packet with $\langle X_k\rangle = x_k$ and $\langle P_k\rangle = y_k$. The coherent-state representation is then the representation of wavefunctions f by functions $f(z) \equiv \langle e_z | f\rangle$. These functions are entire and satisfy

$$\langle f|g\rangle = \pi^{-n}\int_{C^n} \overline{f(z)}\, g(z)\exp(-|z|^2/2)\, d^{2n}z, \tag{1.1}$$

where $|z|^2 = |z_1|^2 + \cdots + |z_n|^2$, $d^{2n}z$ is Lebesgue measure, and the left-hand side denotes the inner product of f and g in the given Hilbert space $\mathcal{H}$ (say, of functions over configuration space).

In spite of its usefulness and intuitive appeal, the coherent-state representation is generally regarded as something of a fluke. The formal combinations $X_k \pm iP_k$, on which it is based, cannot be justified in physical terms, and the use of non-Hermitian operators as anything other than a technical device is regarded with suspicion.

It is one of the aims of this paper to show that representations similar to the above can in fact spring from physical principles, and that the resulting formalism can, as above, be interpreted as a phase-space representation of the given quantum system. The general argument goes as follows: The positivity of the quantum Hamiltonian permits the extension of the one-parameter unitary group $\exp(-itH)$ (t real) representing dynamics in $\mathcal{H}$ to a holomorphic semigroup $\exp(-i\tau H)$ ($\tau = t - i\beta$, $\beta > 0$). On a classical level, evolution in complex time (were it possible) would result in a complexification of the configuration space (hence complex space-time). This has a counterpart at the quantum level in that wavefunctions evolved in complex time, $\exp(-i\tau H)f = \exp(-itH)\exp(-\beta H)f$, may be continued analytically from R^n (configuration space) to a subset (possible all) of C^n. In particular, if the given system is a free nonrelativistic particle, this continuation is even possible at the classical level and gives the complexified position $\mathbf{z}(\tau) = \mathbf{x}_0 + \tau(\mathbf{p}/m) = (\mathbf{x}_0 + t\mathbf{p}/m) - i\beta\mathbf{p}/m$, which is a combination of the type $\mathbf{x} - i\mathbf{p}$. Hence the complexified space can, at every complex "instant" $t - i\beta$, be interpreted as a classical phase space. Moreover, the set of analytically continued solutions carries a representation of the quantum dynamics on functions over phase space.

In Sec. 2 we develop this idea for a free scalar nonrelativistic particle and arrive at a representation which essentially coincides with the usual coherent-state representation. An analogous construction is carried out in Sec. 3 for a relativistic free scalar particle (with positive mass). The ensuing formalism appears to be new and has the general features of the coherent-state representation. The "phase spaces" P_λ of Sec. 3 are products of R^n (configuration space) with an n-dimensional hyperboloid (roughly, a mass shell). It is shown that in the nonrelativistic limit ($c \to \infty$) the formalism goes over smoothly to the formalism of Sec. 2. In Sec. 4 we study the relativistic coherent states e_z, $z \in P_\lambda \approx C^n$. We show that e_{x-iy} is a wavepacket with $\langle X_k\rangle = x_k$ and $\langle P_k\rangle = b_\lambda y_k$, where b_λ is a constant and X_k are the position operators obtained by Newton and Wigner[11] by axiomatizing the notion of "localized states". These results partly justify calling P_λ a "phase space." The e_z are shown to be characterized by an extremal property which, we suggest, is a covariant substitute for minimal uncertainty.

2. NONRELATIVISTIC PARTICLE

The wave function of a free, spinless nonrelativistic particle in R^n evolves under the Schrödinger equation

$$i\frac{\partial f}{\partial t} = Hf, \quad H = -\frac{1}{2m}\Delta. \tag{2.1}$$

The solutions are given by

$$f(\mathbf{x}, t) = [\exp(-itH)f](\mathbf{x})$$
$$= (2\pi)^{-n/2} \int_{R^n} \exp(-itp^2/2m + i\mathbf{x}\cdot\mathbf{p})\hat{f}(\mathbf{p})\, d^np, \tag{2.2}$$

where $\hat{f}(\mathbf{p})$ is the Fourier transform of the initial function $f(\mathbf{x}, 0) \in L^2(R^n)$. Now let $\mathbf{z} = \mathbf{x} - i\mathbf{y} \in C^n$ and let $\tau = t - i\beta$ be in the lower half-plane C^- $(\beta > 0)$. Then $\exp(-i\tau p^2/2m + i\mathbf{z}\cdot\mathbf{p})$ decays rapidly as $|\mathbf{p}| \to \infty$, and Eq. (2.2) defines a function $f(\mathbf{z}, \tau) \equiv [\exp(-i\tau H)f](\mathbf{z})$, holomorphic in $\mathcal{D} = C^n \times C^-$. Let $\mathcal{H} = \{f(\mathbf{z}, \tau) \mid \hat{f}(\mathbf{p}) \in L^2(R^n)\}$ be the vector space of all such functions. Then, for each $\beta > 0$, the function $f_\beta(\mathbf{z}) = f(\mathbf{z}, -i\beta) = [\exp(-\beta H)f](\mathbf{z})$ is entire in C^n. Let $\mathcal{H}_\beta$ be the space of all such functions $f_\beta(\mathbf{z})$. On $\mathcal{H}_\beta$ define the map $\exp(-itH)$ by

$$\exp(-itH)[\exp(-\beta H)f] = \exp(-\beta H)[\exp(-itH)f],$$
$$f \in L^2(R^n). \tag{2.3}$$

We shall make $\mathcal{H}_\beta$ into a Hilbert space such that $t \to \exp(-itH)$ is a unitary representation of dynamics on $\mathcal{H}_\beta$.

Thus, let $\beta > 0$ and $\mathbf{z} = \mathbf{x} - i\mathbf{y} \in C^n$. Then

$$f_\beta(\mathbf{z}) = [\exp(-\beta H)f](\mathbf{z})$$
$$= (2\pi)^{-n/2} \int_{R^n} \exp(-\beta p^2/2m + i\mathbf{z}\cdot\mathbf{p})\hat{f}(\mathbf{p})\, d^np$$
$$\equiv \langle e_{\mathbf{z}}^\beta | f\rangle, \tag{2.4}$$

where

$$\langle e_{\mathbf{z}}^\beta | \mathbf{p}\rangle = (2\pi)^{-n/2}\exp(-\beta p^2/2m + i\mathbf{z}\cdot\mathbf{p}) \tag{2.5}$$

with Fourier transform

$$\langle e_{\mathbf{z}}^\beta | \mathbf{x}'\rangle = (2\pi\beta/m)^{-n/2}\exp[-m(\mathbf{z}-\mathbf{x}')^2/2\beta]. \tag{2.6}$$

The $e_{\mathbf{z}}^\beta$ are minimum-uncertainty spherical wavepackets with $\langle X_k\rangle = x_k$, $\langle P_k\rangle = (m/\beta)y_k$, $\Delta X_k = \sqrt{\beta/2m}$ and $\Delta P_k = \sqrt{m/2\beta}$. They are eigenvectors of $a_k(\beta) = X_k + i(\beta/m)P_k$ with eigenvalue $\bar{z}_k$.

For $f_\beta \in \mathcal{H}_\beta$ define

$$\|f\|_\beta^2 = \int_{C^n} |f_\beta(\mathbf{z})|^2\, d\mu_\beta(\mathbf{z}), \tag{2.7}$$

where

$$d\mu_\beta(\mathbf{z}) = (m/\pi\beta)^{n/2}\exp(-my^2/\beta)\, d^nx\, d^ny. \tag{2.8}$$

Theorem 1: Let $\beta > 0$ and $\hat{f}(\mathbf{p}) \in L^2(R^n)$. Then

$$\|f\|_\beta = \|\hat{f}\|. \tag{2.9}$$

In particular,

(a) $\|\cdot\|_\beta$ is a norm on $\mathcal{H}_\beta$ under which $\mathcal{H}_\beta$ is a Hilbert space.

(b) The map $\exp(-itH)$ is unitary on $\mathcal{H}_\beta$.

(c) The map $\exp(-\beta H)$ is unitary from $L^2(R^n)$ onto $\mathcal{H}_\beta$ and intertwines the dynamics on $L^2(R^n)$ with the dynamics on $\mathcal{H}_\beta$.

Remark: Equation (2.9) can be polarized to give a resolution of the identity: For f, g, in $L^2(R^n)$,

$$\langle f|g\rangle_\beta \equiv \int_{C^n} \langle f|e_{\mathbf{z}}^\beta\rangle\langle e_{\mathbf{z}}^\beta|g\rangle\, d\mu_\beta(\mathbf{z})$$
$$= \langle f|g\rangle_{L^2(R^n)}. \tag{2.10}$$

Hence $f \to \langle e_{\mathbf{z}}^\beta | f\rangle$ is a "representation" of f by an entire function. The connection with the coherent-state representation is as follows: Set $m = \beta = 1$ and let $\tilde{f}(\mathbf{z}) = \pi^{n/4} \times \exp(z^2/4) f_\beta(\mathbf{z})$. Then (2.10) becomes

$$\pi^{-n}\int_{C^n} \overline{\tilde{f}(\mathbf{z})}\tilde{g}(\mathbf{z})\exp(-\tfrac{1}{2}|\mathbf{z}|^2)\, d^nx\, d^ny = \langle f|g\rangle_{L^2(R^n)}, \tag{2.11}$$

so that $f \to \tilde{f}(\mathbf{z})$ is (essentially) the ordinary coherent-state representation [in most of the literature, $\mathbf{z} = (\mathbf{x} - i\mathbf{y})/\sqrt{2}$; the weight function is then $\exp(-|\mathbf{z}|^2)$].

Proof: Let $\hat{f} \in \mathcal{S}(R^n)$. By (2.4), $f_\beta(\mathbf{x} - i\mathbf{y}) = \check{g}_{\beta,\mathbf{y}}(\mathbf{x})$ where $g_{\beta,\mathbf{y}}(\mathbf{p}) = \exp(-\beta p^2/2m + \mathbf{y}\cdot\mathbf{p})\hat{f}(\mathbf{p})$ and $\check{g}$ denotes the inverse Fourier transform of g. Thus, by Plancherel's theorem (and Fubini's),

$$\|f\|_\beta^2 = (m/\pi\beta)^{n/2}\int \exp(-my^2/\beta)\, dy$$
$$\times \int \exp(-\beta p^2/m + 2\mathbf{y}\cdot\mathbf{p})|\hat{f}(\mathbf{p})|^2\, dp$$
$$= \int |\hat{f}(\mathbf{p})|^2\, dp = \|\hat{f}\|^2,$$

which proves (2.9) for $\hat{f} \in \mathcal{S}(R^n)$, hence also for $\hat{f} \in L^2(R^n)$ by continuity. (a)—(c) are obvious. ∎

For the definition of intertwining operators, see Ref. 12.

3. RELATIVISTIC PARTICLE

In the last section we obtained unitary maps from $L^2(R^n)$ onto Hilbert spaces $\mathcal{H}_\beta$ where the role of δ functions is played by spherical wavepackets $e_{\mathbf{z}}^\beta$ in $L^2(R^n)$ [$\mathcal{H}_\beta$ is continuously imbedded in $L^2(R^n)$ by restricting $f_\beta(\mathbf{z})$ to R^n]. This formalism is nonrelativistic since the inner product for $L^2(R^n)$ is not Lorentz-invariant. In this section we define covariant counterparts of the $e_{\mathbf{z}}^\beta$ and prove the analog of Theorem 1. We begin with the relativistic version of the free-particle Schrödinger equation, namely the Klein—Gordon equation (for a free scalar particle of mass $m > 0$ in $n+1$ space—time dimensions):

$$\left(-\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \Delta - m^2\right)f(\mathbf{x}, t) = 0. \tag{3.1}$$

The positivity of the energy played an essential role in Sec. 2, and will do so again here. Hence we confine ourselves to positive-energy solutions.[13] These are given by

$$f(x) = f(\mathbf{x}, x_0) = [\exp(-ix_0H)f](\mathbf{x})$$
$$= (2\pi)^{-n/2}\int_{R^n}\exp(-ixp)\hat{f}(\mathbf{p})\, d\Omega(\mathbf{p}), \tag{3.2}$$

where

$$x_0 = ct, \quad H = +(m^2c^2 - \Delta)^{1/2}, \quad xp = x_0\omega - \mathbf{x}\cdot\mathbf{p},$$
$$\omega = (m^2c^2 + \mathbf{p}^2)^{1/2}, \quad d\Omega(\mathbf{p}) = d^np/\omega$$

is the Lorentz-invariant measure on the mass shell,[15] and $\hat{f}(\mathbf{p})/\omega$ is the ordinary Fourier transform of the initial function $f(\mathbf{x}, 0)$ (considered, say, as a tempered distribution in R^n). For every $\hat{f}(\mathbf{p}) \in L^2(\Omega)$, i.e., with

$$\|\hat{f}\|^2 \equiv \int_{R^n}|\hat{f}(\mathbf{p})|^2\, d\Omega(\mathbf{p}) < \infty, \tag{3.3}$$

the corresponding solution $f(x)$ is the boundary value of a function $f(z)$ holomorphic in the *forward tube*[14,15]

$$\mathcal{T}=R^{n+1}-iV_+=\{z=x-iy\,|\,x\in R^{n+1},\ y\in V_+\},$$

where $V_+=\{y=(y_0,\mathbf{y})\in R^{n+1}\,|\,y_0>|\mathbf{y}|\}$ is the open forward light cone in R^{n+1}. This is because $|\exp(-izp)|$ $\equiv|\exp(-iz_0\omega+i\mathbf{z}\cdot\mathbf{p})|=\exp(-y_0\omega+\mathbf{y}\cdot\mathbf{p})\equiv\exp(-yp)$ $<\exp[-(y_0-|\mathbf{y}|)|\mathbf{p}|]$; hence $|\exp(-izp)|$ decays rapidly as $|\mathbf{p}|\to\infty$ for fixed $z=x-iy\in\mathcal{T}$. $\mathcal{T}$ will replace the domain $\mathcal{D}=C^n\times C^-$ of Sec. 2, and is strictly contained in $\mathcal{D}$. Thus, for $z\in\mathcal{T}$,

$$f(z)=(2\pi)^{-n/2}\int_{R^n}\exp(-izp)\hat{f}(\mathbf{p})\,d\Omega(\mathbf{p})$$
$$\equiv\langle e_z|f\rangle,\tag{3.4}$$

where

$$\langle e_z|p\rangle=(2\pi)^{-n/2}\exp(-izp)\tag{3.5}$$

and $\langle e_z|f\rangle$ denotes the inner product in $L^2(\Omega)$. The vectors e_z are in $L^2(\Omega)$, since for $z,w\in\mathcal{T}$,

$$\langle e_z|e_w\rangle=(2\pi)^{-n}\int_{R^n}\exp[-i(z-\bar{w})p]\,d\Omega(\mathbf{p})$$
$$=(2/i)\Delta_+(z-\bar{w})$$
$$=(1/\pi)(mc/2\pi\eta)^{\nu}K_\nu(\eta mc),\tag{3.6}$$

where Δ_+ is the familiar two-point function for the free scalar field of mass m,[14] $\eta=[-(z-\bar{w})^2]^{1/2}\equiv[-(z_0-\bar{w}_0)^2+\sum_1^n(z_k-\bar{w}_k)^2]^{1/2}$ is uniquely defined by analytic continuation from $\eta=[-(z-\bar{z})^2]^{1/2}=2(y_0^2-\mathbf{y}^2)^{1/2}$ when $z=w=x-iy\in\mathcal{T}$, K_ν is a modified Bessel function,[16] and throughout the rest of this paper we set $\nu=(n-1)/2$. The analog of the space $\mathcal{H}$ of holomorphic solutions $f(\mathbf{z},\tau)$ in $\mathcal{D}$ is the space $\mathcal{K}=\{f(z)\,|\,f\in L^2(\Omega)\}$ of functions defined by (3.4). Recall now that $\mathcal{H}_\beta$ was obtained from $\mathcal{H}$ by restricting functions $f\in\mathcal{H}$ to the "phase space" $P_\beta^{\mathrm{NR}}=\{(\mathbf{z},-i\beta)\,|\,\mathbf{z}\in C^n\}\approx C^n$. This set is, however, not contained in $\mathcal{T}$. To obtain a relativistic phase space we reason as follows: $\mathcal{D}$ can be obtained (roughly) as a deformation of $\mathcal{T}$ by letting $c\to\infty$ while fixing $\tau=z_0/c$. A set in $\mathcal{T}$ which goes into P_β^{NR} under this "deformation" is

$$P_\lambda=\{(\mathbf{z},-i(\lambda^2+\mathbf{y}^2)^{1/2})\,|\,\mathbf{z}=\mathbf{x}-i\mathbf{y}\in C^n\},\tag{3.7}$$

with $\lambda=\beta c>0$. We will show that P_λ is a suitable phase space. For $\lambda=0$ Eq. (3.7) defines P_0 as a subset of the boundary of $\mathcal{T}$.

The sets P_λ are clearly not invariant under Lorentz transformations. To make the formalism manifestly Poincaré-covariant, we will also need the sets

$$P_\lambda'=\{(\mathbf{z},x_0-i(\lambda^2+\mathbf{y}^2)^{1/2})\,|\,\mathbf{z}=\mathbf{x}-i\mathbf{y}\in C^n,\quad x_0\in R\}.\tag{3.8}$$

Every function $f(z)\in\mathcal{K}$ defines a "boundary value" function on P_0' (and by restriction also on P_0) as follows: for $z=(\mathbf{x}-i\mathbf{y},x_0-i|\mathbf{y}|)\in P_0'$,

$$f(z)\equiv f_0(\mathbf{x}-i\mathbf{y},x_0)$$
$$=((1/\omega)\exp(-|\mathbf{y}|\omega+\mathbf{y}\cdot\mathbf{p}-ix_0\omega)\hat{f}(\mathbf{p}))^{\check{}}(\mathbf{x})\tag{3.9}$$

[see (3.4)]. It follows from (3.9) that, for fixed $\mathbf{y}\in R^n$ and $x_0\in R$, f_0 is in $L^2(R^n)$ as a function of $\mathbf{x}$. [Actually, as we shall see, $f_0\in L^2(C^n)$ in $\mathbf{x}-i\mathbf{y}$ and $f\to f_0$ in $L^2(C^n)$ as $\lambda\to 0$.] Thus $f(z)$ makes sense even when $z\in P_0'$ (though its pointwise values no longer have meaning).

Given $\lambda\geq 0$ and $f\in\mathcal{K}$, define

$$\|f\|_\lambda^2=\int_{P_\lambda}|f(z)|^2\,d\mu_\lambda(z),\tag{3.10}$$

where

$$d\mu_\lambda(z)=C_\lambda\,d^n x\,d^n y,\quad z=(\mathbf{x}-i\mathbf{y},z_0)\in P_\lambda\tag{3.11}$$

with

$$C_\lambda=[(2/\pi)(\pi\lambda/mc)^{\nu+1}K_{\nu+1}(2\lambda mc)]^{-1},\quad\lambda>0,\tag{3.12}$$

and $C_0\equiv\lim_{\lambda\to 0}C_\lambda=(mc)^{n+1}/\pi^\nu\Gamma(\nu+1)$. C_λ is a continuous, monotone increasing function of λ on $[0,\infty)$, with

$$C_\lambda\sim mc(mc/\pi\lambda)^{n/2}\exp(2\lambda mc)\quad\text{as}\quad\lambda mc\to\infty.\tag{3.13}$$

These facts, and others needed later, follow from certain properties of the K_ν,[16] which we summarize in Appendix A. We may regard $d\mu_\lambda$ either as a measure on P_λ or as a measure on C^n. In the latter interpretation (which will also be useful) we write (3.10) as

$$\|f\|_\lambda^2=\int_{C^n}|f_\lambda(\mathbf{z})|^2\,d\mu_\lambda(\mathbf{z}),\tag{3.14}$$

where $f_\lambda(\mathbf{z})=f(\mathbf{z},-i(\lambda^2+\mathbf{y}^2)^{1/2})$ is the restriction of $f\in\mathcal{K}$ to P_λ. Let $\mathcal{K}_\lambda=\{f_\lambda(\mathbf{z})\,|\,f\in\mathcal{K}\}$ be the space of all such restrictions (boundary values, if $\lambda=0$) and denote the map $\hat{f}(\mathbf{p})\to f_\lambda(\mathbf{z})$ from $L^2(\Omega)$ onto $\mathcal{K}_\lambda$ by D_λ. Similarly let $\mathcal{K}_\lambda'$ be the space of restrictions $f_\lambda(x,\mathbf{y})=f(\mathbf{x}-i\mathbf{y},\ x_0-i(\lambda^2+\mathbf{y}^2)^{1/2})$ to P_λ' and denote the corresponding map by D_λ'. Since each $f_\lambda(x,\mathbf{y})\in\mathcal{K}_\lambda'$ satisfies (3.1) in $x\in R^{n+1}$, $\mathcal{K}_\lambda'$ is simply the space of solutions with initial values in $\mathcal{K}_\lambda$. Notice that (3.14) is defined for $f_\lambda\in\mathcal{K}_\lambda'$ as well as for $f_\lambda\in\mathcal{K}_\lambda$.

Now $L^2(\Omega)$ carries a unitary, irreducible representation of the restricted Poincaré group $\mathcal{P}_+^\uparrow$,[14] given by

$$(U(a,\Lambda)\hat{f})(\mathbf{p})=\exp(iap)\hat{f}(\Lambda^{-1}p),\tag{3.15}$$

where $(a,\Lambda)\in\mathcal{P}_+^\uparrow$ acts on space–time according to

$$(a,\Lambda)x=\Lambda x+a,\quad x\in R^{n+1}.\tag{3.16}$$

In (3.15) $p=(\mathbf{p},\omega)$ denotes a point on the mass shell (a homogeneous space for the Lorentz group) rather than the corresponding momentum vector $\mathbf{p}$. The representation (3.15) defines a corresponding representation on $\mathcal{K}$ given by

$$(\tilde{U}(a,\Lambda)f)(z)=f(\Lambda^{-1}(z-a))\tag{3.17}$$

(where we have extended the action of $\mathcal{P}_+^\uparrow$ to $\mathcal{T}$ by linearity). Now P_λ' is a homogeneous space of $\mathcal{P}_+^\uparrow$ [in fact, $P_\lambda'\approx\mathcal{P}_+^\uparrow/\mathrm{SO}(n)$ since the stability subgroup at, say, $(0,-i\lambda)$ is $\mathrm{SO}(n)$]. Hence (3.17) gives a representation U_λ' on $\mathcal{K}_\lambda'$ by restriction (taking boundary values, if $\lambda=0$). Since extension sets up a one–one correspondence between $\mathcal{K}_\lambda$ and $\mathcal{K}_\lambda'$ we also have a representation U_λ on $\mathcal{K}_\lambda$, but this one is less direct since P_λ is not invariant under $\mathcal{P}_+^\uparrow$.

The next theorem, which is our first main result, shows that U, U_λ, and U_λ' are all unitarily equivalent.

Theorem 2: Let $\lambda\geq 0$ and $\hat{f}\in L^2(\Omega)$. Then

$$\|f\|_\lambda=\|\hat{f}\|.\tag{3.18}$$

In particular,

(a) $\|\cdot\|_\lambda$ is a norm on $\mathcal{K}_\lambda$ ($\mathcal{K}_\lambda'$) under which $\mathcal{K}_\lambda$ ($\mathcal{K}_\lambda'$) is Hilbert space.

(b) U_λ (U'_λ) is a unitary irreducible representation of $\mathcal{P}^\uparrow_+$ on $\mathcal{K}_\lambda$ ($\mathcal{K}'_\lambda$).

(c) D_λ (D'_λ) is unitary from $L^2(\Omega)$ onto $\mathcal{K}_\lambda$ ($\mathcal{K}'_\lambda$) and intertwines the representations U and U_λ (U'_λ) of $\mathcal{P}^\uparrow_+$.

Proof: The proof is completely parallel to that of Theorem 1. Let $\hat{f} \in \mathcal{S}(R^n)$ and note that

$$f(z) = (2\pi)^{-n/2} \int_{R^n} \exp(-ix_0\omega + i\mathbf{x}\cdot\mathbf{p} - yp)\hat{f}(\mathbf{p})\, d^n p/\omega$$

$$= ((1/\omega)\exp(-ix_0\omega - yp)\hat{f})\check{}\,(\mathbf{x}). \tag{3.19}$$

Hence

$$\|f\|_\lambda^2 = C_\lambda \int d^n y \int d^n x \,|f(\mathbf{x} - i\mathbf{y}, -i(\lambda^2 + \mathbf{y}^2)^{1/2})|^2$$

$$= C_\lambda \int d^n y \int d^n p\,|(1/\omega)\exp[-(\lambda^2+\mathbf{y}^2)^{1/2}\omega + \mathbf{y}\cdot\mathbf{p}] \times \hat{f}(\mathbf{p})|^2$$

$$= C_\lambda \int d^n p [|\hat{f}(\mathbf{p})|^2/\omega^2] \int d^n y \exp[-2(\lambda^2+\mathbf{y}^2)^{1/2}\omega + 2\mathbf{y}\cdot\mathbf{p}]$$

$$= \int (d^n p/\omega)\,|\hat{f}(\mathbf{p})|^2 = \|\hat{f}\|^2,$$

where we have used (A6) with $\alpha = 0$. This proves (3.18) for $\hat{f} \in \mathcal{S}(R^n)$, hence for $\hat{f} \in L^2(\Omega)$ by continuity. (a) and (b) are obvious, and the intertwining property follows from

$$(D'_\lambda U(a,\Lambda)\hat{f})(z) = (D'_\lambda(\exp(iap)\hat{f}(\Lambda^{-1}p)))(z)$$

$$= (2\pi)^{-n/2}\int \exp(-izp + iap)\hat{f}(\Lambda^{-1}p)\,d\Omega(p)$$

$$= (2\pi)^{-n/2}\int \exp[-i(z-a)\Lambda p'] \times \hat{f}(p')\,d\Omega(\Lambda p')$$

$$= (2\pi)^{-n/2}\int \exp[-i(\Lambda^{-1}(z-a))p'] \times \hat{f}(p')\,d\Omega(p')$$

$$= f(\Lambda^{-1}(z-a)) = (U'_\lambda D'_\lambda f)(z), \quad z \in P'_\lambda,$$

where we have used the invariance of $d\Omega(p)$. ∎

The norm $\|\cdot\|_\lambda$ on $\mathcal{K}_\lambda$ and $\mathcal{K}'_\lambda$ defines an inner product $\langle\cdot|\cdot\rangle_\lambda$ on these spaces by polarization. As an immediate consequence of Theorem 2 we have

Corollary 1: Let $\lambda \geq 0$ and $\hat{f}, \hat{g} \in L^2(\Omega)$. Then

$$\langle f|g\rangle_\lambda \equiv \int_{P_\lambda} \langle f|e_z\rangle\langle e_z|g\rangle\, d\mu_\lambda(z)$$

$$= \langle f|g\rangle_{L^2(\Omega)}. \tag{3.20}$$

In particular, taking $\hat{f} = e_w$ ($w \in \mathcal{T}$), we obtain

$$g(w) = \langle e_w|g\rangle = \int_{P_\lambda}\langle e_w|e_z\rangle\langle e_z|g\rangle\, d\mu_\lambda(z)$$

$$= \int_{P_\lambda}\langle e_w|e_z\rangle g(z)\, d\mu_\lambda(z). \tag{3.21}$$

Equation (3.21), restricted to $w \in P_\lambda$, states that $\langle e_w|e_z\rangle$ is a (hence *the*) reproducing kernel[17] for $\mathcal{K}_\lambda$.

In the sequel we will sometimes identify the spaces $L^2(\Omega)$, $\mathcal{K}$, $\mathcal{K}_\lambda$, and $\mathcal{K}'_\lambda$ (as Theorem 2 permits us to do). Thus f could stand for $\hat{f}(\mathbf{p})$ or $f(z)$ as an element of $\mathcal{K}$, $\mathcal{K}_\lambda$, or $\mathcal{K}'_\lambda$. We will also set $c = 1$ except in considerations involving the nonrelativistic limit.

We can now make precise the sense in which a function $f \in \mathcal{K}$ takes on its boundary value on P_0.

Corollary 2: (a) Each $\mathcal{K}_\lambda$ is a closed subspace of $L^2(C^n)$. (b) Let $0 \leq \lambda < \lambda'$ and $\hat{f} \in L^2(\Omega)$. Then

$$\|f_{\lambda'} - f_\lambda\|_{L^2(C^n)} \to 0 \quad \text{as } \lambda' \downarrow \lambda.$$

Proof: (a) follows from (3.18), and (b) follows essentially from the proof of Theorem 2:

$$\|f_{\lambda'} - f_\lambda\|^2_{L^2(C^n)} = \int dp[|\hat{f}(\mathbf{p})|^2/\omega^2]\int dy \times (\exp[-(\lambda^2+\mathbf{y}^2)^{1/2}\omega] - \exp[-(\lambda'^2+\mathbf{y}^2)^{1/2}\omega])^2\exp(2\mathbf{y}\cdot\mathbf{p})$$

$$= \int dp[|\hat{f}(\mathbf{p})|^2/\omega^2][\omega/C_\lambda + \omega/C_{\lambda'} - 2J(p,\lambda,\lambda')]$$

where

$$J = \int dy \exp\{-[(\lambda^2+\mathbf{y}^2)^{1/2} + (\lambda'^2+\mathbf{y}^2)^{1/2}]\omega + 2\mathbf{y}\cdot\mathbf{p}\}.$$

But $\omega/C_{\lambda'} \leq J \leq \omega/C_\lambda$; hence

$$\|f_{\lambda'} - f_\lambda\|^2_{L^2(C^n)} \leq (C_\lambda^{-1} - C_{\lambda'}^{-1})\|\hat{f}\|^2,$$

which implies (b). ∎

We conclude this section by showing that the e_z-representation on $\mathcal{K}_\lambda$ is indeed a relativistic version of the e_z^β-representation on $\mathcal{H}_\beta$. For given $\beta > 0$, define

$$f_\beta^{\mathrm{NR}}(\mathbf{x} - i\mathbf{y}) = \exp(-m\mathbf{y}^2/2\beta)(\exp(-\beta\mathbf{p}^2/2m + \mathbf{y}\cdot\mathbf{p})\hat{f}(\mathbf{p}))\check{}\,(\mathbf{x})$$

$$= \left[\exp\left[-\frac{\beta m}{2}\left(\frac{\mathbf{p}}{m} - \frac{\mathbf{y}}{\beta}\right)^2\right]\hat{f}(\mathbf{p})\right]\check{}\,(\mathbf{x}). \tag{3.22}$$

Theorem 3: Let $\beta > 0$ and $\hat{f}(\mathbf{p}) \in L^2(R^n)$. Then $f_\beta^{\mathrm{NR}}(\mathbf{z}) \in L^2(C^n)$ and

$$J(c) \equiv \|mc\exp(\beta mc^2)f_{\beta c} - f_\beta^{\mathrm{NR}}\|^2_{L^2(C^n)} \to 0 \quad \text{as } c \to \infty,$$

where $f_{\beta c}$ is the function in $\mathcal{K}_{\beta c}$ corresponding to $\hat{f} \in L^2(R^n) \subset L^2(\Omega)$. The proof is given as Appendix B.

4. THE WAVEPACKETS e_z

In this section we study the "relativistic coherent states" e_z. We show that they are centered about $\mathbf{x} = \mathrm{Re}(\mathbf{z})$, travel with average momentum proportional to $\mathbf{y}$, and are characterized by a property which is a covariant analog of minimal uncertainty.

To compute the position of the center of e_z, we need position operators. It was shown by Newton and Wigner[11] that certain group-theoretical postulates about (idealized) "localized states"—e.g., that any space translate of a localized state be "orthogonal" to the state[18]—uniquely determine a set of self-adjoint operators [here given on $L^2(\Omega)$]

$$X_k = i\left(\frac{\partial}{\partial p_k} - \frac{p_k}{2\omega^2}\right), \quad k = 2, \ldots, n, \tag{4.1}$$

whose (generalized) eigenvectors are the localized states. (The notion of being localized in this sense, however, depends on the frame of reference.) In a later paper, dealing with quantization, we will show that the operators (4.1) can also be obtained naturally from the formalism of Sec. 3. For the purpose of this section, we simply adopt (4.1) as the definition of position operators.

We begin by computing the expectation of X_k in e_z:

$$\langle e_z | X_z e_z\rangle = \int \frac{dp}{\omega}\langle e_z|p\rangle i\left(\frac{\partial}{\partial p_k} - \frac{p_k}{2\omega^2}\right)\langle p|e_z\rangle$$
$$= \int dp\, \frac{\langle e_z|p\rangle}{\sqrt{\omega}}\, i\frac{\partial}{\partial p_k}\left(\frac{\langle p|e_z\rangle}{\sqrt{\omega}}\right)$$
$$= \mathrm{Re}\int dp\, \frac{\langle e_z|p\rangle}{\sqrt{\omega}}\, i\frac{\partial}{\partial p_k}\left(\omega^{-1/2}\frac{\exp(i\bar{z}p)}{(2\pi)^{n/2}}\right)$$
$$= x_k\langle e_z|e_z\rangle.$$

Thus

$$\langle X_k\rangle = x_k \equiv \mathrm{Re}(z_k). \tag{4.2}$$

To find the expectation of P_α, let

$$G(m,y) = \int_{R^n}\exp(-2yp)\,d\Omega(p) = 2(\pi m/\lambda)^\nu K_\nu(2\lambda m)$$
$$= a(m)\varphi^{-\nu}K_\nu(\varphi) = b(y)\varphi^\nu K_\nu(\varphi), \tag{4.3}$$

where $a(m) = 2(2m^2\pi)^\nu$, $b(y) = 2(2\lambda^2/\pi)^{-\nu}$, $\varphi = 2\lambda m$, and $\lambda = \lambda(y) \equiv (y_\alpha y^\alpha)^{1/2}$ with all the y_α considered as independent variables. $G(m,y)$ will be a "partition function" (as in statistical mechanics) for generating expectations. Thus, using (A2),

$$\int p_\alpha \exp(-2yp)\,d\Omega(p) = -\frac{1}{2}\frac{\partial G}{\partial y_\alpha}$$
$$= -2m^2 y_\alpha a(m)\frac{1}{\varphi}\frac{\partial}{\partial\varphi}(\varphi^{-\nu}K_\nu(\varphi))$$
$$= 2m^2 y_\alpha a(m)\varphi^{-\nu-1}K_{\nu+1}(\varphi);$$

hence

$$\langle P_\alpha\rangle = -\frac{1}{2G}\frac{\partial G}{\partial y^\alpha} = \frac{K_{\nu+1}(2\lambda m)}{K_\nu(2\lambda m)}\cdot\frac{m}{\lambda}y_\alpha \tag{4.4}$$

in the state e_z $(z = x - iy)$. Similarly,

$$\int p_\alpha p_\beta \exp(-2yp)\,d\Omega(p) = \frac{1}{4}\frac{\partial^2 G}{\partial y^\alpha \partial y^\beta}$$
$$= 4m^4 y_\alpha y_\beta a(m)\varphi^{-\nu-2}K_{\nu+2}(\varphi)$$
$$- m^2 g_{\alpha\beta}a(m)\varphi^{-\nu-1}K_{\nu+1}(\varphi),$$

giving

$$\langle P_\alpha P_\beta\rangle = \frac{K_{\nu+2}(2\lambda m)}{K_\nu(2\lambda m)}\cdot\frac{m^2}{\lambda^2}y_\alpha y_\beta$$
$$-\frac{K_{\nu+1}(2\lambda m)}{K_\nu(2\lambda m)}\frac{m}{2\lambda}g_{\alpha\beta}. \tag{4.5}$$

Equations (4.4) and (4.5) give the expected momenta and their correlation matrix

$$C_{\alpha\beta} = \langle P_\alpha P_\beta\rangle - \langle P_\alpha\rangle\langle P_\beta\rangle. \tag{4.6}$$

To gain a rough idea of the behavior of $\langle P_\alpha\rangle$ and $C_{\alpha\beta}$, we consider the limiting cases $\lambda m \to \infty$. From (A3) and (A4) we obtain

$$\frac{\nu}{\lambda^2}y_\alpha \overset{0}{\leftarrow} \langle P_\alpha\rangle \overset{\infty}{\rightarrow} \frac{m}{\lambda}y_\alpha, \tag{4.7}$$

$$\frac{\nu}{\lambda^2}\left(\frac{y_\alpha y_\beta}{\lambda^2} - \tfrac{1}{2}g_{\alpha\beta}\right) \overset{0}{\leftarrow} C_{\alpha\beta}$$
$$\overset{\infty}{\rightarrow} \frac{m}{2\lambda}\left(1 + \frac{n}{2\lambda m}\right)\left(\frac{y_\alpha y_\beta}{\lambda^2} - g_{\alpha\beta}\right) + \frac{n}{8\lambda^2}g_{\alpha\beta}. \tag{4.8}$$

Hence the uncertainties in energy and momentum obey

$$\frac{n-1}{2\lambda^2}\left(\frac{1}{2} + \frac{\mathbf{y}^2}{\lambda^2}\right) \overset{0}{\leftarrow} C_{00} \overset{\infty}{\rightarrow} \frac{n}{8\lambda^2} + \frac{m\,\mathbf{y}^2}{2\lambda\,\lambda^2}, \tag{4.9}$$

$$\frac{n-1}{2\lambda^2}\left(\frac{1}{2} + \frac{y_k^2}{\lambda^2}\right) \overset{0}{\leftarrow} C_{kk} \overset{\infty}{\rightarrow} \frac{m}{2\lambda} - \frac{n}{8\lambda^2} + \frac{m}{2\lambda}\cdot\frac{y_k^2}{\lambda^2}. \tag{4.10}$$

Finally, we need an estimate on the uncertainty in position: At $x_0 = 0$,

$$\langle p|(X_k - x_k)e_z\rangle = \left[i\left(\frac{\partial}{\partial p_k} - \frac{p_k}{2\omega^2}\right) - x_k\right](2\pi)^{-n/2}$$
$$\times\exp(-y_0\omega + \mathbf{y}\cdot\mathbf{p} - i\mathbf{x}\cdot\mathbf{p})$$
$$= i\left(y_k - y_0\frac{p_k}{\omega} - \frac{p_k}{2\omega^2}\right)\langle p|e_z\rangle;$$

hence

$$\langle (X_k - x_k)^2\rangle = G^{-1}\int\left[y_k - y_0\frac{p_k}{\omega}\left(1 + \frac{1}{2y_0\omega}\right)\right]^2$$
$$\times\exp(-2yp)\,d\Omega(p). \tag{4.11}$$

The integral is difficult to evaluate, and we merely derive an upper bound in the rest frame. Setting $\mathbf{y} = 0$ and $y_0 = \lambda$,

$$\langle (X_k - x_k)^2\rangle = \lambda^2 G^{-1}\int\frac{p_k^2}{\omega^2}\left(1 + \frac{1}{2y_0\omega}\right)^2\exp(-2\lambda\omega)\,d\Omega(p)$$
$$\le \lambda^2 G^{-1}\int\left(1 + \frac{1}{2y_0\omega}\right)^2\exp(-2\lambda\omega)\,d\Omega(p)$$
$$= \lambda^2 + \lambda G^{-1}\int\left(\frac{1}{\omega} + \frac{1}{4\lambda\omega^2}\right)\exp(-2\lambda\omega)\,d\Omega(p).$$

Now

$$-\frac{1}{2\lambda m}\frac{\partial G}{\partial m} = \int\left(\frac{1}{\omega} + \frac{1}{2\lambda\omega^2}\right)\exp(-2\lambda\omega)\,d\Omega(p),$$

hence

$$\langle (X_k - x_k)^2\rangle \le \lambda^2 - \frac{1}{2mG}\frac{\partial G}{\partial m}$$
$$= \lambda^2 - \frac{2\lambda^2 b(y)}{G}\frac{1}{\varphi}\frac{\partial}{\partial\varphi}(\varphi^\nu K_\nu(\varphi))$$
$$= \lambda^2 + \frac{2\lambda^2 b(y)}{G}\varphi^{\nu-1}K_{\nu-1}(\varphi).$$

The position uncertainty therefore satisfies

$$\langle (X_k - x_k)^2\rangle \le \lambda^2 + \frac{\lambda}{m}\frac{K_{\nu-1}(2\lambda m)}{K_\nu(2\lambda m)}; \tag{4.12}$$

hence

$$\langle (X_k - x_k)^2\rangle \lesssim [\nu/(\nu-1)]\lambda^2 \quad \text{as } \lambda m \to 0, \tag{4.13}$$

$$\lesssim \lambda^2 + \lambda/m \quad \text{as } \lambda m \to \infty. \tag{4.13'}$$

For $\nu \equiv (n-1)/2 = 1$ (which is in fact the physical case), (4.13) must be replaced with

$$\langle (X_k - x_k)^2 \rangle \lesssim \lambda^2 - 2\lambda^2 \ln(2\lambda m) \quad \text{as } \lambda m \to 0. \tag{4.13'}$$

Thus $\Delta X_k \to 0$ when $\lambda \to 0$.

We can now draw consequences from the above computations. Equations (4.2) and (4.4) confirm that e_{x-iy} is a wavepacket centered about $\mathbf{x}$ with expected energy—momentum proportional to $(y_0, \mathbf{y})$. Note that

$$(\langle P_\alpha \rangle \langle P^\alpha \rangle)^{1/2} = m \frac{K_{\nu+1}(2\lambda m)}{K_\nu(2\lambda m)} \equiv m_\lambda > m. \tag{4.14}$$

We shall call m_λ the "effective mass" for the particle in P_λ. The factor $K_{\nu+1}/K_\nu$ represents a kind of renormalization which takes into effect the fluctuations in energy—momentum. m_λ has the asymptotic behavior

$$\nu/\lambda \overset{0}{\sim} m_\lambda \overset{\infty}{\sim} m. \tag{4.15}$$

Equations (4.7)—(4.10), (4.13), and (4.15) show the following pattern: When $\lambda m \to 0$, the expectations and uncertainties of physical observables in the state e_z become independent of the mass. Thus, roughly, when $\lambda \to 0$ (i.e., z approaches the boundary of $\mathcal{T}$), analyticity fails and fluctuations take over. On the other hand, we have seen that $\lambda m \equiv \lambda mc \equiv \beta mc^2 \to \infty$ gives a smooth transition to the nonrelativistic formalism (Theorem 3). Thus we expect $\langle P_k \rangle \to m y_k / \beta \equiv m y_k / \lambda$, $C_{kk} \to m/2\beta \equiv m/2\lambda$, and $\langle (X_k - x_k)^2 \rangle \to \beta/2m \equiv \lambda/2m$. The first two are born out by (4.7) and (4.10). Equation (4.13'), though consistent with this expectation, shows that in obtaining the estimate (4.12) we gave up too much ground.

The nonrelativistic wavepackets e_z^β have the attractive feature of being minimum-uncertainty states. So far we have not shown that the e_z have a similar property. Now uncertainty products do not seem to be natural measure of the optimality of relativistic states. The position operators X_k are not covariant,[11] and furthermore it is not obvious how to define an *invariant* counterpart to the uncertainty product. We conclude by proving that the e_z are characterized by a simple, invariant property which we propose as an adequate substitute for minimal uncertainty. For $z \in \mathcal{T}$ let

$$\tilde{e}_z(w) = \langle e_w | e_z \rangle / \| e_z \|, \quad w \in \mathcal{T}.$$

Theorem 4: Let $z \in \mathcal{T}$. Then $\tilde{e}_z$ is the unique (up to a constant phase factor) solution to the following problem: Find $f \in \mathcal{K}$ such that $\|f\| = 1$ and $|f(z)|$ is a maximum.

Proof: We have

$$|\langle e_z | f \rangle| \le \| e_z \| \| f \|,$$

and equality holds if and only if f is a constant multiple of e_z. ∎

Remark: Theorem 4 can be restated as a variational principle[19]: $\hat{e}_z(w) \equiv \langle e_w | e_z \rangle / \| e_z \|^2$ is the unique function f in $\mathcal{K}$ such that $f(z) = 1$ and $\|f\|$ is a minimum. The above form seems to be more appropriate for quantum mechanics. See also Ref. 20.

5. CONCLUSION

We have developed a formalism analogous to that of the coherent-state representation. By this analogy we have called P_λ a "phase space." We then showed that, at least so far as the e_z are concerned, P_λ is indeed a space parametrized by coordinates and momenta. Now in the classical notion of phase space, a central role is played by Poisson brackets and canonical transformations, i.e., by symplectic structure.[21,22] These geometrical aspects will be dealt with in a later paper, where the present formalism will be given a geometrical foundation and made manifestly covariant.

ACKNOWLEDGMENTS

I wish to thank Lon Rosen for reading the manuscript and making many helpful suggestions. I have also benefited from numerous conversations with Ivan Kupka and Peter Milman.

APPENDIX A

We collect here some properties of the modified Bessel functions K_ν and evaluate some integrals needed in Secs. 3 and 4.

The functions $K_\nu(\xi)$ are defined[16] for $\operatorname{Re}\nu > -\frac{1}{2}$ and $\operatorname{Re}\xi > 0$ by

$$K_\nu(\xi) = \frac{\sqrt{\pi}(\xi/2)^\nu}{\Gamma(\nu + \frac{1}{2})} \int_0^\infty \exp(-\xi \cosh t) \sinh^{2\nu} t \, dt. \tag{A1}$$

They satisfy

$$\begin{aligned} &\left(-\frac{1}{\xi}\frac{d}{d\xi}\right)^m (\xi^\nu K_\nu(\xi)) = \xi^{\nu-m} K_{\nu-m}(\xi), \\ &\left(-\frac{1}{\xi}\frac{d}{d\xi}\right)^m (\xi^{-\nu} K_\nu(\xi)) = \xi^{-\nu-m} K_{\nu+m}(\xi), \end{aligned} \tag{A2}$$

for $m = 1, 2, \cdots$ and

$$\begin{aligned} &K_\nu(\xi) \sim \tfrac{1}{2}\Gamma(\nu)(\xi/2)^{-\nu}, \quad \xi \to 0 \quad (\nu \neq 0), \\ &K_0(\xi) \sim -\ln(\xi/2), \quad \xi \to 0, \\ &K_\nu(\xi) \sim \sqrt{\pi/2\xi}\, e^{-\xi}, \quad \xi \to +\infty. \end{aligned} \tag{A3}$$

In Sec. 4 we use

$$\begin{aligned} &\frac{\Gamma(\nu+k)}{\Gamma(\nu)} \left(\frac{\xi}{2}\right)^{-k} \overset{0}{\sim} \frac{K_{\nu+k}(\xi)}{K_\nu(\xi)} \overset{\infty}{\sim} 1 + \frac{k^2 + 2k\nu}{2\xi}, \\ &\frac{2(n-1)}{\xi^2} \overset{0}{\sim} \frac{K_{\nu+2}(\xi)}{K_\nu(\xi)} - \left(\frac{K_{\nu+1}(\xi)}{K_\nu(\xi)}\right)^2 \overset{\infty}{\sim} \frac{1}{\xi} + \frac{n}{\xi^2}. \end{aligned} \tag{A4}$$

To evaluate

$$I(y_0, \mathbf{y}) = \int_{R^n} \frac{d^n p}{(1+\mathbf{p}^2)^{1/2}} \exp[-2y_0(1+\mathbf{p}^2)^{1/2} + 2\mathbf{y}\cdot\mathbf{p}],$$

$$\lambda \equiv (y_0^2 - \mathbf{y}^2)^{1/2} > 0,$$

note that I is Lorentz-invariant; hence

$$\begin{aligned} I(y_0, \mathbf{y}) = I(\lambda, 0) &= \int_{R^n} \frac{d^n p}{(1+\mathbf{p}^2)^{1/2}} \exp[-2\lambda(1+\mathbf{p}^2)^{1/2}] \\ &= \frac{2\pi^{n/2}}{\Gamma(n/2)} \int_0^\infty \frac{r^{n-1}\, dr}{(1+r^2)^{1/2}} \exp[-2\lambda(1+r^2)^{1/2}] \\ &= \frac{2\pi^{n/2}}{\Gamma(n/2)} \int_0^\infty \sinh^{n-1} t \exp(-2\lambda \cosh t)\, dt \\ &= 2\left(\frac{\pi}{\lambda}\right)^\nu K_\nu(2\lambda), \quad \nu \equiv \frac{n-1}{2}. \end{aligned} \tag{A5}$$

Consequently, using (A2),

$$\int_{R^n} \frac{d^n p}{(1+\mathbf{p}^2)^{1/2}} p_\alpha \exp[-2y_0(1+\mathbf{p}^2)^{1/2} + 2\mathbf{y}\cdot\mathbf{p}]$$
$$= -\frac{1}{2}\frac{\partial}{\partial y^\alpha} I(y_0, \mathbf{y})$$
$$= 4y_\alpha\left(-\frac{1}{2\lambda}\frac{\partial}{\partial(2\lambda)}\right)\left[\left(\frac{\pi}{\lambda}\right)^\nu K_\nu(2\lambda)\right]$$
$$= \frac{2}{\pi} y_\alpha \left(\frac{\pi}{\lambda}\right)^{\nu+1} K_{\nu+1}(2\lambda), \tag{A6}$$

where $p_0 = (1+\mathbf{p}^2)^{1/2}$.

APPENDIX B. PROOF OF THEOREM 3

We can set $m=\beta=1$ without loss. Note

$$f_1^{\rm NR}(\mathbf{x}-i\mathbf{y}) = \exp(-\mathbf{y}^2/2)\langle e^1_{\mathbf{x}-i\mathbf{y}}|f\rangle_{L^2(R^n)}.$$

Hence by (2.8) and (2.9),

$$\|f_1^{\rm NR}\|^2_{L^2(C^n)} = \pi^{n/2}\|f_1^{\rm NR}\|_{H_1} = \pi^{n/2}\|\hat f\|_{L^2(R^n)} < \infty.$$

Note also

$$\|ce^{c^2}f_c\|^2_{L^2(C^n)} = \frac{c^2e^{2c^2}}{C_c}\|f_c\|^2_{K_c}$$
$$\le \pi^{n/2}c\|\hat f\|^2_{L^2(\Omega)}[1+O(c^{-2})]$$
$$\le \pi^{n/2}\|\hat f\|^2_{L^2(R^n)}[1+O(c^{-2})].$$

Now

$$J = \int\int dx\,dy\,|[\{(c/\omega)\exp(c^2-yp) - \exp[-\tfrac12(\mathbf{p}-\mathbf{y})^2]\}\hat f]^\vee(\mathbf{x})|^2$$
$$= \int dp\,|\hat f(\mathbf{p})|^2\int dy((c/\omega)\exp(c^2-yp) - \exp[-\tfrac12(\mathbf{p}-\mathbf{y})^2])^2.$$

Choose α, γ such that $\frac12 < \gamma < \alpha < 1$. Then $\int_{|p|>c^{1-\alpha}} dp \times |\hat f(\mathbf{p})|^2 \to 0$ as $c\to\infty$; hence

$$J_1 \equiv \int_{|p|>c^{1-\alpha}} dp\,|\hat f(\mathbf{p})|^2\int_{R^n} dy$$
$$\times\{(c/\omega)\exp(c^2-yp) - \exp[-\tfrac12(\mathbf{p}-\mathbf{y})^2]\}^2$$
$$\le 4\pi^{n/2}\|\chi_c\hat f\|^2_{L^2(R^n)} \to 0 \quad\text{as } c\to\infty,$$

where $\chi_c(\mathbf{p})$ is the characteristic function of $\{|\mathbf{p}|>c^{1-\alpha}\}$.

Define θ and φ by $|\mathbf{y}| = c\sinh\theta$, $|\mathbf{p}| = c\sinh\varphi$. Then $y_0 = (c^2-\mathbf{y}^2)^{1/2} = c\cosh\theta$ and $\omega = c\cosh\varphi$; hence $yp = y_0\omega - \mathbf{y}\cdot\mathbf{p} \ge c^2\cosh(\theta-\varphi) \ge c^2 + (c^2/2)(\theta-\varphi)^2$. Thus for arbitrary $a\ge 0$,

$$G_a(\mathbf{p}) \equiv \int_{|y|>c\sinh a} dy\,\exp(2c^2-2yp)$$
$$\le \frac{2c^n\pi^{n/2}}{\Gamma(n/2)}\int_a^\infty \sinh^{n-1}\cosh\theta\exp[-c^2(\theta-\varphi)^2]\,d\theta$$
$$\le \frac{2^{1-n}c^n\pi^{n/2}}{\Gamma(n/2)}\int_a^\infty \exp[(n-1)\theta](e^\theta+e^{-\theta})\exp[-c^2(\theta-\varphi)^2]d\theta$$
$$\le \frac{2^{2-n}c^n\pi^{n/2}}{\Gamma(n/2)}\int_a^\infty \exp[n\theta - c^2(\theta-\varphi)^2]\,d\theta$$
$$= \frac{2^{2-n}c^{n-1}\pi^{n/2}}{\Gamma(n/2)}\exp(n\varphi + n^2/4c^2)\int_{c(a-\varphi)-n/2c}^\infty \exp(-u^2)\,du.$$

Let $a = \sinh^{-1}(c^{-\gamma})$. Then, for $|\mathbf{p}| < c^{1-\alpha}$,

$$c(a-\varphi) - n/2c \ge c[\sinh^{-1}(c^{-\gamma}) - \sinh^{-1}(c^{-\alpha})] - n/2c$$
$$\equiv g(c).$$

$g(c)$ is independent of p and $g(c)\sim c^{1-\gamma}$, $c\to\infty$. Also, $|\mathbf{p}| < c^{1-\alpha} \Rightarrow \varphi < c^{-\alpha}$. Hence

$$J_2 \equiv \int_{|p|<c^{1-\alpha}} dp\,|\hat f(\mathbf{p})|^2\int_{|y|>c^{1-\gamma}} dy\,\exp(2c^2-2yp)$$
$$\le \frac{2^{2-n}c^{n-1}\pi^{n/2}}{\Gamma(n/2)}\exp(nc^{-\alpha}+n^2/4c^2)\left(\int_{g(c)}^\infty \exp(-u^2)\,du\right)$$
$$\times\|\hat f\|^2_{L^2(R^n)} \to 0, \quad c\to\infty.$$

Now

$$2c^2 - 2yp = y^2 + p^2 - 2yp = (y-p)^2$$
$$\equiv (y_0-\omega)^2 - (\mathbf{y}-\mathbf{p})^2$$
$$\ge -(\mathbf{y}-\mathbf{p})^2.$$

Hence

$$\int_{|p|<c^{1-\alpha}} dp\,|\hat f(\mathbf{p})|^2\int_{|y|>c^{1-\gamma}} dy\,\exp[-(\mathbf{y}-\mathbf{p})^2] \le J_2$$

and

$$\int_{|p|<c^{1-\alpha}} dp\,|\hat f(\mathbf{p})|^2\int_{|y|>c^{1-\gamma}} dy\{(c/\omega)\exp(c^2-yp)$$
$$-\exp[-\tfrac12(\mathbf{y}-\mathbf{p})^2]\}^2 \le 4J_2 \to 0 \quad\text{as } c\to\infty.$$

Finally,

$$J_3 \equiv \int_{|p|<c^{1-\alpha}} dp\,|\hat f(\mathbf{p})|^2\int_{|y|<c^{1-\gamma}} dy\{(c/\omega)\exp(c^2-yp)$$
$$-\exp[-\tfrac12(\mathbf{y}-\mathbf{p})^2]\}^2$$
$$= \int_{|p|<c^{1-\alpha}} dp\,|\hat f(\mathbf{p})|^2\int_{|y|<c^{1-\gamma}} dy\,\exp[-(\mathbf{y}-\mathbf{p})^2]$$
$$\times[(c/\omega)\exp(c^2\delta^2/2) - 1]^2,$$

where

$$\delta = |(1+\mathbf{y}^2/c^2)^{1/2} - (1+\mathbf{p}^2/c^2)^{1/2}|$$
$$\le \tfrac12|\mathbf{y}^2/c^2 - \mathbf{p}^2/c^2| \le \tfrac12(c^{-2\gamma}+c^{-2\alpha}) \le c^{-2\gamma}.$$

We have used the estimate

$$|(1+u^2)^{1/2} - (1+v^2)^{1/2}|$$
$$= \left|\int_u^v \frac{x\,dx}{(1+x^2)^{1/2}}\right| \le \left|\int_u^v x\,dx\right| = \tfrac12|v^2-u^2|.$$

Hence for sufficiently large c and $|\mathbf{p}| < c^{1-\alpha}$,

$$[(c/\omega)\exp(c^2\delta^2/2) - 1]^2$$
$$\le \exp(c^2\delta^2) + 1 - (2c/\omega)\exp(c^2\delta^2/2)$$
$$\le (1+2c^2\delta^2) + 1 - 2(1-p^2/2c^2)\exp(c^2\delta^2/2)$$
$$\le 2[1-\exp(c^2\delta^2/2)] + 2c^2\delta^2 + c^{-2\alpha}\exp(c^2\delta^2/2)$$
$$\le 2c^2\delta^2 + c^{-2\alpha}(1+c^2\delta^2)$$
$$\equiv h(c) \to 0 \quad\text{as } c\to\infty.$$

Thus

$$J_3 \leq h(c) \int_{|p|<c^{1-\alpha}} dp\, |\hat{f}(p)|^2 \int_{|y|<c^{1-\gamma}} dy\, \exp[-(\mathbf{y}-\mathbf{p})^2]$$

$$\leq h(c)\pi^{n/2}\|\hat{f}\|^2_{L^2(R^n)} \to 0 \quad \text{as } c \to \infty.$$

which proves that $J \to 0$ as $c \to \infty$. ∎

*This work is part of the author's Ph.D. thesis (submitted to the University of Toronto, 1977).

[1]G. Kaiser, "Relativistic Coherent-State Representations," in *Proceedings of the Fifth International Colloquim on Group Theoretical Methods in Physics, Montreal*, 1976 (Academic, New York) (to be published).

[2]E.P. Wigner, Phys. Rev. **40**, 749 (1932).

[3]J.E. Moyal, Proc. Cambridge Phil. Soc. **45**, 99 (1945).

[4]J.R. Klauder, Ann. Phys. (N.Y.) **11**, 123 (1960).

[5]V. Bargmann, Commun. Pure Appl. Math. **14**, 187 (1961).

[6]I.E. Segal, Illinois J. Math. **6**, 500 (1962).

[7]A. Grossmann, G. Loupias, and E.M. Stein, Ann. Inst. Fourier **18**, 343 (1968).

[8]For other representations of the "coherent-state" type, see Refs. 9, 10.

[9]A.O. Barut and L. Girardello, Commun. Math. Phys. **21**, 41 (1971).

[10]A.M. Perelomov, Commun. Math. Phys. **26**, 222 (1972).

[11]T.D. Newton and E.P. Wigner, Rev. Mod. Phys. **21**, 400 (1949).

[12]I.M. Gel'fand, M.I. Graev, and N. Ya. Vilenkin, *Generalized Functions, Vol.* 5 (Academic, New York, 1966).

[13]The definiteness of the energy is necessary in order that our representation of $\mathcal{P}_+^\uparrow$ be irreducible; choosing it to be positive is also in the spirit of quantum field theory, where solutions of (3.1) enter as one-particle test functions for the field. (See Ref. 14.)

[14]R.F. Streater and A.S. Wightman, *PCT, Spin and Statistics and All That* (Benjamin, New York, 1964).

[15]M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol.* 2 (Academic, New York, 1975).

[16]M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions* (Natl. Bureau of Standards, Washington, D.C., 1964).

[17]H. Meschkowski, *Hilbertsche Räume mit Kernfunktion* (Springer-Verlag, Berlin, 1962).

[18]For a more rigorous treatment, see A.S. Wightman, Rev. Mod. Phys. **34**, 845 (1962).

[19]S. Bergman, *The Kernel Function and Conformal Mapping* (Amer. Math. Soc., Providence, R.I., 1970), 2nd ed.

[20]J.R. Klauder, J. Math. Phys. **5**, 177 (1964).

[21]S. MacLane, *Geometrical Mechanics* I, II, Univ. of Chicago lecture notes, 1968.

[22]R. Abraham and J.E. Marsden, *Foundations of Mechanics* (Benjamin, New York, 1967).

Curvature invariants and space–time singularities

John A. Thorpe*

Mathematics Department, State University of New York, Stony Brook, New York
(Received 7 September 1976)

This paper collects together in a general setting observer dependent curvature invariants for space–time and applies them to an analysis of curvature singularities. Observer dependent quantities, such as energy and momentum densities and tidal stresses, are dependent not only on the space–time point but also on the observer's 4-velocity. The properties of these invariants are discussed, and it is shown that they completely describe the behavior of curvature along timelike curves. In particular, curvature singularities can be characterized by unboundedness of these invariants.

1. INTRODUCTION

One of the major difficulties one encounters in dealing with the curvature of space—time is that of describing in a meaningful way the size of the curvature. The components R_{ijkl} of the Riemann tensor in a system of local coordinates are useless for this purpose because unless the Riemann tensor is very special (constant curvature), its components may be made arbitrarily large simply by an appropriate choice of coordinates. The scalar polynomial invariants of curvature are of some use in measuring the size of curvature since the values of these invariants are coordinate independent. But, as is well known, there exist Riemann tensors all of whose scalar polynomial invariants vanish but which are nevertheless not zero.

From the point of view of an observer in space—time, what really matters are those quantities which can be felt and measured such as tidal forces and energy density. These quantities are observer dependent. Two different observers sitting at the same point of space—time will measure different tidal forces and energy densities because these quantities depend not only on the point in space—time but also on the 4-velocity of the observer.

In this paper we shall study a collection of observer-dependent curvature invariants. These invariants are scalar functions on the bundle of unit timelike tangent vectors rather than on space—time itself. Some of these invariants have direct interpretations in terms of tidal effects, spatial curvature, and energy and momentum densities. Others are related to the matter and conformal parts of the Riemann tensor. Using these invariants, it is possible to deal meaningfully with questions relating to the size of curvature and its growth along timelike curves. In particular, curvature singularities can be characterized by unboundedness of these invariants.

Currently the most satisfactory formulation of the concept of curvature singularity is by way of parallelly propagated frames.[1] Curvature is said to be unbounded along a timelike curve γ if some component of the Riemann tensor relative to a parallel frame field along γ is unbounded. If curvature "blows up" in this sense in finite proper time, the curve γ is said to run into a parallelly propagated curvature singularity. A refinement of this idea[2] says that curvature is unbounded along γ if relative to every orthonormal frame field along γ there is some component of the Riemann tensor which is unbounded. The key idea in both of these formulations is that the growth of curvature along a timelike curve γ can be measured by curvature components in suitable orthonormal frame fields along γ. Viewed another way, curvature components can be regarded as scalar functions on the orthonormal frame bundle and the growth of curvature along a timelike curve γ can be measured by the growth of these scalar functions along suitable lifts of γ to this frame bundle.

Our approach to curvature singularities is similar, but does not require introducing orthonormal frame fields. Each timelike curve γ parametrized by proper time has a natural lift, defined by its velocity vector field, to the bundle of unit timelike vectors. By identifying γ with its natural lift, observer dependent quantities can be regarded as defined along γ. Hence we can measure the growth of curvature along timelike curves by means of observer dependent invariants. We shall show that a timelike curve of bounded acceleration runs into a parallelly propagated curvature singularity if and only if some observer dependent curvature invariant is unbounded along γ. Furthermore, the singularity can be classified as tidal, matter, conformal, etc., depending on which invariants are unbounded.

Some of the invariants constructed here for unit timelike vectors are also defined on null vectors. We shall show that unless all these invariants vanish along the limit set of an imprisoned timelike curve, such a curve (when it exists) must run into a curvature singularity.

2. INVARIANTS

Let M be a space—time. Thus M is a four-dimensional manifold with metric tensor g of signature $(-+++)$. We shall assume for convenience that M is oriented and time oriented. An *observer* ("instantaneous observer" in the terminology of Sachs and Wu[3]) is a unit timelike tangent vector v at some point p of M. Attached to each observer is a *rest space* $v^\perp$ consisting of all tangent vectors at p which are orthogonal to v. Thus each observer v at p determines a $3+1$ orthogonal decomposition of the tangent space M_p to M at p into a spacelike 3-plane ($v^\perp$) and a timelike line (all multiples of v).

We shall construct from the Riemann tensor a collection of self-adjoint linear operators on the rest space $v^\perp$, whose eigenvalues we shall call principal curvatures. The idea is to combine the invariant decomposition of the Riemann tensor with the $3+1$ decomposition of the tangent space attached to our observer to obtain linear operators which are self-adjoint and hence diagonalizable relative to the positive definite inner product on $v^\perp$.

The invariants constructed here are essentially known in special contexts. A $3+1$ decomposition of the tangent space occurs naturally in general relativity in two situations: (i) fluid flow, where the observer at each point is the 4-velocity of the fluid, and (ii) slicing of the space-time by a family of spacelike hypersurfaces, where the observer at each point is the future timelike unit vector orthogonal to the slicing. Our interest here is in studying the behavior of these invariants along arbitrary timelike curves, parametrized by proper time, where the observer at each point of the curve is the tangent vector to the curve.

The construction is as follows. Given any tensor L_{ijkl} with the symmetries of the Riemann tensor ($L_{ijkl} = -L_{jikl} = -L_{ijlk} = L_{klij}$) we can construct, using a given observer v, the symmetric 2-covariant tensor $L_{ij} = L_{ikjl}v^kv^l$. Since $L_{ij}v^j = L_{ji}v^j = 0$, we may, without any loss of information, view L_{ij} as a tensor on $v^\perp$. Since $v^\perp$ is spacelike, the metric on $v^\perp$ is positive definite so the tensor L_{ij} can be diagonalized. Thus there exists an orthonormal basis $\{e_1, e_2, e_3\}$ for $v^\perp$ and real numbers $\{\lambda_1, \lambda_2, \lambda_3\}$ such that, relative to this basis, $L_{ij} = \lambda_i\delta_{ij}$. The λ_i are the eigenvalues of the self-adjoint linear operator $L^i{}_j = g^{ik}L_{kj}$ on $v^\perp$. We shall always order the λ_i so that $\lambda_1 \leq \lambda_2 \leq \lambda_3$.

We shall apply this construction to five tensors.

(1) The Riemann tensor. Taking $L_{ijkl} = R_{ijkl}$, we obtain the symmetric tensor $L_{ij} = R_{ikjl}v^kv^l$ which is the tidal stress, or tidal force,[1] as seen by the observer v. We shall denote the corresponding eigenvalues by $\{\tau_1, \tau_2, \tau_3\}$ and call them the *principal tidal curvatures* (as measured by v). τ_1 and τ_3 represent, respectively, the minimum and maximum values of tidal stress as measured by the observer v. Geometrically, the τ_i represent (up to sign) the critical values of Riemannian sectional curvature,[4] a function with domain the manifold of non-null 2-planes at p, restricted to the (compact) set of timelike 2-planes containing v.

(2) The double dual of the Riemann tensor. Taking $L_{ijkl} = \frac{1}{4}\epsilon_{ijmn}R^{mnqr}\epsilon_{qrkl}$, we find that the associated tensor L_{ij} has components, relative to any orthonormal basis with $e_0 = v$, which are simply the spatial components R_{ijkl} $(i,j,k,l > 0)$ of the Riemann tensor. The associated eigenvalues will be denoted by $\{\kappa_1, \kappa_2, \kappa_3\}$ and will be called the *principal spatial curvatures*. Geometrically, the κ_i represent the critical values of Riemannian sectional curvature, restricted to the (compact) set of spacelike 2-planes orthogonal to v.

(3) The Weyl tensor. Taking $L_{ijkl} = C_{ijkl}$, we find that the associated tensor L_{ij} is the electric part of the Weyl tensor.[5] The associated eigenvalues will be denoted by $\{\epsilon_1, \epsilon_2, \epsilon_3\}$ and will be called the *principal conformal electric curvatures*.

(4) The dual of the Weyl tensor. Taking $L_{ijkl} = \frac{1}{2}\epsilon_{ijmn}C^{mn}{}_{kl}$, the associated tensor L_{ij} is the magnetic part of the Weyl tensor.[5] The associated eigenvalues $\{\mu_1, \mu_2, \mu_3\}$ will be called the *principal conformal magnetic curvatures*.

(5) The Ricci part of the Riemann tensor. Taking $L_{ijkl} = R_{ijkl} - C_{ijkl}$, we obtain the part of the Riemann tensor which is determined algebraically, through the field equations, by the energy—momentum tensor. The associated tensor L_{ij} is a linear combination of the spatial parts of the Ricci tensor R_{ij} and the metric tensor g_{ij} and hence has the same eigendirections as the spatial part of the Ricci tensor. The associated eigenvalues $\{\delta_1, \delta_2, \delta_3\}$ will be called the *principal matter curvatures*.

The principal curvatures are continuous real valued functions on the bundle $\mathcal{O}$ of unit timelike vectors over M. They are invariants in the sense that they are constant under the action of the orthogonal group O(3) (SO(3) for the conformal magnetic curvatures) represented as the subgroup of the Lorentz group at p leaving the observer v fixed.

These curvatures have the following properties.

(i) If the principal tidal curvatures are identically zero on $\mathcal{O}$, then M has zero curvature. Indeed, if $\tau_1 = \tau_3 = 0$, then the minimum and maximum values of sectional curvature on 2-planes containing v are both zero for all $v \in \mathcal{O}$. It follows that sectional curvature is identically zero on the set of all timelike 2-planes, hence on all 2-planes. This then implies that $R_{ijkl} = 0$.

(ii) The mean tidal curvature is equal to the value of Ricci curvature on the observer v,

$$\tau_1 + \tau_2 + \tau_3 = R_{ij}v^iv^j,$$

and the mean spatial curvature $\kappa_1 + \kappa_2 + \kappa_3$ is equal to the energy density as measured by v (assuming the Einstein equations with cosmological constant equal to zero).

Verification of the first statement is straightforward, using an orthonormal basis with $e_0 = v$ and $\{e_1, e_2, e_3\}$ in the principal tidal directions. To check the second, choose $e_0 = v$ and $\{e_1, e_2, e_3\}$ in the principal spatial directions. Then the energy density as measured by v is given by

$$\begin{aligned} T_{ij}v^iv^j &= (R_{ij} - \tfrac{1}{2}\rho g_{ij})v^iv^j \\ &= R^k{}_{ikj}v^iv^j + \tfrac{1}{2}g^{ij}R^k{}_{ikj} \\ &= R^k{}_{0k0} + \sum_{i<k} g^{ii}R^k{}_{iki} \\ &= \sum_{0<i<k} R_{kiki} = \sum_{i=1}^{3} \kappa_i. \end{aligned}$$

Here, T_{ij} denotes the energy—momentum tensor and ρ denotes scalar curvature.

(iii) The principal tidal curvatures and the principal spatial curvatures are, in general, independent invariants. However, for vacuum solutions of the Einstein equations, the principal spatial curvatures are the negatives of the principal tidal curvatures. This is because, when $R_{ij} = 0$, $R_{ijkl} = C_{ijkl}$ so that the Riemann tensor is the negative of its double dual.[6,7] Also, in a vacuum, the principal tidal directions in $v^\perp$ coincide with the principal spatial directions, for each v in $\mathcal{O}$.

(iv) The components of the Riemann tensor relative to an orthonormal basis $\{e_0, e_1, e_2, e_3\}$ for M_p (e_0 timelike) can be conveniently displayed in a 6×6 matrix

$[R_{\alpha\beta}]$, where α and β run through the set of index pairs $\{01, 02, 03, 23, 31, 12\}$. $[R_{\alpha\beta}]$ is then just the matrix for the Riemann tensor regarded as a quadratic form on the space of bivectors at p. This matrix splits naturally into 3×3 blocks

$$[R_{\alpha\beta}]=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}\end{bmatrix},$$

where A_{11} and A_{22} are symmetric matrices and A_{21} is the transpose of A_{12}. The matrices A_{11}, A_{22}, and the symmetric part of A_{12} represent, respectively, the tidal, the spatial, and the conformal magnetic parts of curvature, while the skew-symmetric part of A_{12} represents (up to a factor of 2) momentum density, all as seen by the observer e_0. The matter and conformal electric parts of curvature are represented by linear combinations $\frac{1}{2}(A_{11}\pm A_{22})\pm\frac{1}{6}[\mathrm{tr}(A_{11}-A_{22})]I$ of A_{11}, A_{22}, and the identity matrix I.

Choosing the orthonormal basis $\{e_0, e_1, e_2, e_3\}$ so that the spacelike vectors e_1, e_2, e_3 are in principal curvature directions forces certain curvature components to be zero. Thus, for example, choosing e_1, e_2, e_3 to be in the principal tidal directions diagonalizes A_{11}. Choosing e_1, e_2, e_3 in the principal spatial directions diagonalizes A_{22}. In a vacuum, A_{11} and A_{22} are simultaneously diagonalized for each choice of e_0 in $\mathcal{O}$ by choosing e_1, e_2, e_3 in the common principal tidal and spatial directions. Further, in a vacuum, the matrix A_{12} is symmetric (momentum density is zero) so $A_{12}=A_{21}$, yielding the canonical form

$$[R_{\alpha\beta}]=\left[\begin{array}{ccc|ccc} \tau_1 & 0 & 0 & & & \\ 0 & \tau_2 & 0 & & B & \\ 0 & 0 & \tau_3 & & & \\ \hline & & & -\tau_1 & 0 & 0 \\ & B & & 0 & -\tau_2 & 0 \\ & & & 0 & 0 & -\tau_3 \end{array}\right],$$

where the symmetric matrix B represents the conformal magnetic part of curvature as seen by the observer e_0.

For one special observer e_0 at p, aligned in a direction particularly well oriented in relation to the Riemann tensor, the matrix B can simultaneously be cast into a canonical form, and one obtains the Petrov canonical form[6] for $[R_{\alpha\beta}]$. The advantage of being content with diagonalizing only one block (two in a vacuum) in $[R_{\alpha\beta}]$ is, however, that this diagonalization can be done for *every* choice of observer e_0. In particular, along any timelike curve in M, e_0 can be chosen to be the unit tangent vector yielding particularly nice representations of the Riemann tensor along the curve, displayed with respect to orthonormal frame fields adapted to the curve.

(v) If all the principal curvatures and their corresponding principal directions, together with the momentum density, are known for any one observer v at $p\in M$, then the full Riemann tensor is completely determined at p. In fact, by (iv), it suffices to know the principal tidal, spatial, and conformal magnetic curvatures and directions together with momentum density (as seen by v) since each of the blocks A_{ij} in the matrix representation for R relative to any orthonormal basis with $e_0=v$ can be computed from this information.

(vi) Given any L_{ijkl} with the symmetries of the Riemann tensor and any $v\in\mathcal{O}$, we can construct another tensor $\tilde{L}_{ijkl}$ with the same symmetries by

$$\tilde{L}_{ijkl}=v_{[i}L_{j]mn[l}v_{k]}v^m v^n=v_{[i}L_{j][l}v_{k]}.$$

$\tilde{L}_{ijkl}$ represents that part of L_{ijkl} which is determined by the associated tensor $L_{ij}=L_{ikjl}v^k v^l$. In particular, the tensor $v_{[i}R_{j]mn[l}v_{k]}v^m v^n$ represents the tidal part of the Riemann tensor; it is $\neq 0$ if and only if some tidal curvature $\tau_i\neq 0$ at v. This condition, that $v_{[i}R_{j]mn[l}v_{k]}v^m v^n\neq 0$, plays an important role in the singularity theorems of Hawking and Penrose.[1,8]

Remark: Many of the invariants described here carry over also to "null observers." For v a null vector and $L_{ijkl}=-L_{jikl}=-L_{ijlk}=L_{klij}$, we can define L_{ij} on $v^\perp$ as before by

$$L_{ij}=L_{ikjl}v^k v^l.$$

The metric on $v^\perp$ is no longer positive definite, but, if we restrict to a two-dimensional spacelike plane P in $v^\perp$, L_{ij} can still be diagonalized. Moreover, the result is independent of the 2-plane selected since any spacelike vector in $v^\perp$ differs from one in P by a multiple of v and

$$L_{ij}(u^i+av^i)(w^j+bv^j)=L_{ij}u^i w^j.$$

Thus each L_{ij} defines two invariants λ_1 and λ_2 rather than three as when v is timelike. The tensor $v_{[i}L_{j]mn[l}v_{k]}v^m v^n$ still represents that part of L_{ijkl} which is determined by L_{ij}. In particular, $v_{[i}L_{j]mn[l}v_{k]}v^m v^n\neq 0$ if and only if $\lambda_i\neq 0$ at v for $i=1$ or 2. It should be noted however that, in contrast to the timelike case, these null observer invariants fail to fully determine the Riemann tensor at the given point, even with the addition of energy density $T_{ij}v^i v^j$ and magnitude $T_{ik}T_j{}^k v^i v^j$ of 4-momentum to the list of invariants. In Taub–NUT space,[1,9] for example, all these invariants are zero for v any null vector tangent to the horizon, yet $R_{ijkl}\neq 0$ there.

3. SINGULARITIES

Observer dependent invariants are especially useful for describing the behavior of curvature along timelike curves because each such curve defines an observer at each of its points and so one can study the growth of the invariants as one moves along the curve. Any invariant growing without bound in finite proper time along a curve of bounded acceleration will signal a singularity. At the other extreme, if all the invariants are zero at points along a curve, then the full Riemann tensor is zero at those points.

In order to describe the growth of curvature, it suffices to consider the following ten invariants: the principal tidal, spatial, and conformal magnetic curvatures (three each) and the magnitude $T_{ik}T_j{}^k v^i v^j+(T_{ij}v^i v^j)^2$ of momentum density. By (iv) and (v) of the previous section, if these ten invariants are bounded along a timelike curve γ, then so are all of the observer dependent curvature invariants. Further, $R_{ijkl}=0$ at $\gamma(t)$ if (and only if) these ten invariants are zero at $v=\dot\gamma(t)$, where $\dot\gamma(t)$ is the unit tangent vector to γ at $\gamma(t)$.

For γ a timelike curve parametrized by proper time, the *natural lift* of γ is the curve $\dot\gamma$ in U defined by $\dot\gamma(t)$ = the tangent vector to γ at $\gamma(t)$. We shall say that γ runs into a *curvature singularity* if

(a) γ has domain a half-open interval $[a, b)$,

(b) γ has bounded acceleration, and

(c) $\lambda\circ\dot\gamma$ is unbounded for some curvature invariant λ.

Remarks: (i) The terminology "γ runs into a singularity" does not mean that there is some point $p=\lim_{t\to b}\gamma(t)$ in M which is singular in some sense. Rather it means that the curve γ, a world line for some particle or moving observer, encounters unbounded curvature as $t\to b$ and hence cannot be extended to proper time $t=b$, not in the space—time M and not any smooth extension of M. (In order to find an actual *point* for γ to run into, it would be necessary to enlarge M as a topological space, necessarily destroying the Lorentz manifold structure in the process. Such a construction has been carried out by Schmidt.[10])

(ii) A more inclusive concept of curvature singularity would replace condition (c) by the weaker condition

(c$'$) $\lim_{t\to b}\lambda\circ\dot\gamma(t)$ fails to exist for some λ.

This condition would also prevent extension of γ to $t=b$, but the effects on the observer would be less dramatic and possibly physically insignificant.

(iii) The nature of the singularity can be specified further by noting which invariants blow up. Thus a curvature singularity is a *matter singularity* if the momentum density or one of the principal matter curvatures is unbounded. It is a *conformal singularity* if one of the principal conformal (electric or magnetic) curvatures is unbounded. Thus the Schwarzschild singularity is a conformal singularity whereas the Friedmann singularities are matter singularities. A curvature singularity is *tidal* or *spatial* if one of the tidal or spatial principal curvatures is unbounded, respectively.

The nature of the singularity encountered by an observer falling radially into a Schwarzschild black hole, for example, can be seen from the values of the curvature invariants along a radial world line. In Schwarzschild coordinates (t, r, θ, φ), if v is the unit vector in the direction of the (timelike for $r<2m$) radial vector $-\partial/\partial r$, then the principal curvatures can be read off from the curvature components relative to the normalized coordinate frame.[11] The nonzero ones are

$$\tau_1=-\kappa_3=\mu_1=-2m/r^3,$$

$$\tau_2=\tau_3=-\kappa_1=-\kappa_2=\mu_2=\mu_3=m/r^3.$$

Thus this Schwarzschild singularity is tidal, spatial, and conformal magnetic.

(iv) A timelike curve γ: $[a, b)\to M$ of bounded acceleration runs into a curvature singularity (as described above) if and only if it runs into a parallelly propagated curvature singularity (as described, e.g., in Hawking and Ellis[1]). Indeed, since γ has bounded acceleration, each parallel orthonormal frame field $\{E_i\}$ along γ differs from an orthonormal frame field $\{\tilde E_i\}$ adapted to γ (with $\tilde E_0=\dot\gamma$) by a bounded one-parameter family of boosts. Hence the Riemann tensor R will have an unbounded component relative to $\{E_i\}$ (a parallelly propagated singularity) if and only if it has an unbounded component relative to $\{\tilde E_i\}$. But the matrix for R relative to $\{\tilde E_i\}$ has the block form (relative to the observer $E_0=\dot\gamma$) as described in (iv) of the previous section so an unbounded component relative to $\{\tilde E_i\}$ corresponds to a tidal, spatial, conformal magnetic, or momentum density singularity depending on whether it appears in the A_{11} block, the A_{22} block, the symmetric part of the A_{12} block, or the skew-symmetric part of the A_{12} block.

(v) To illustrate the use of observer dependent invariants, we shall prove a refinement of a theorem of Hawking and Ellis.[12] Recall that a future inextendible timelike curve $\gamma:[a, b)\to M$ is imprisoned if there exists a compact set K in M and a sequence $\{t_\nu\}$ in $[a, b)$ converging to b such that $\gamma(t_\nu)\in K$ for all ν. The limit set $\mathcal{L}$ of a curve γ imprisoned in K is the set of all limit points of all such sequences. $\mathcal{L}$ is nonempty since K is compact. Since the space of directions on K is also compact (it is a fiber bundle over K with fiber a 3-sphere), there are at each point of $\mathcal{L}$ one or more limiting directions. These directions are limit points of the sequence of directions determined by the tangent vectors $\dot\gamma(t_\nu)$ to γ at $\gamma(t_\nu)$. Clearly, each of these directions is either timelike or null (we shall see shortly that they are in fact null). We shall refer to these limiting directions as *directions along* $\mathcal{L}$.

Theorem: Suppose γ is a future inextendible timelike curve of bounded acceleration imprisoned in the compact set K. Then either

(a) all null-observer-dependent curvature invariants vanish along the limit set $\mathcal{L}$ of γ, or

(b) γ runs into a curvature singularity.

Proof: Let $p\in\mathcal{L}$ and let $v\neq 0$ lie in a limiting direction of γ at p. Thus there exists a sequence $\{t_\nu\}$ in $[a, b)$ converging to b such that $\lim_{\nu\to\infty}\gamma(t_\nu)=p$ and $\lim_{\nu\to\infty}[\dot\gamma(t_\nu)]=[v]$, where for a nonzero vector w the notation $[w]$ means the direction of w. Since γ is future inextendible, $\lim_{t\to b}\gamma(t)$ cannot exist so there must exist a neighborhood U of p and a sequence $\{s_\nu\}$ in $[a, b)$ converging to b such that $\gamma(s_\nu)\notin U$ for all ν. By passing to subsequences if necessary we can assume that $t_\nu<s_\nu<t_{\nu+1}$ for all ν.

Now $[v]$, being a limit of timelike directions, must be timelike or null. If $[v]$ were timelike, the vector v could be chosen to be a unit timelike vector in which case we would have $v=\lim_{\nu\to\infty}\dot\gamma(t_\nu)$. Letting A be a bound on the acceleration along γ and choosing $\epsilon>0$ such that all points reachable from p in proper time $\leq\epsilon$ along timelike curves with initial velocity v and acceleration bounded by A are contained in U, it follows that, for sufficiently large ν, all points reachable from $\gamma(t_\nu)$ in proper time $\leq\epsilon$ along timelike curves with initial velocity $\dot\gamma(t_\nu)$ and acceleration bounded by A are contained in U. In particular, $\gamma(t)\in U$ for $t_\nu\leq t\leq t_\nu+\epsilon$ and hence

$$t_{\nu+1} - t_\nu > s_\nu - t_\nu > \epsilon,$$

for sufficiently large ν. But this is impossible since $\{t_\nu\}$ converges to b. Hence v must be null.

Now let λ be any principal curvature at v and let w be a unit spacelike vector in the corresponding principal direction, so that

$$L_{ikjl} v^k v^l w^i w^j = \lambda,$$

where L_{ijkl} is one of the five tensors described in the previous section. Let v_ν and w_ν be vectors tangent to M at $\gamma(t_\nu)$ such that

(i) v_ν is in the direction of $\dot\gamma(t_\nu)$ for each ν, and $\lim_{\nu\to\infty} v_\nu = v$,

(ii) w_ν is a unit vector in $v_\nu^\perp$ for each ν, and $\lim_{\nu\to\infty} w_\nu = w$.

Then

$$\begin{aligned}\lambda &= L_{ikjl} v^k v^l w^i w^j \\ &= \lim_{\nu\to\infty} L_{ikjl} v_\nu{}^k v_\nu{}^l w_\nu{}^i w_\nu{}^j \\ &= \lim_{\nu\to\infty} (g_{mn} v_\nu{}^m v_\nu{}^n)(L_{ikjl}\dot\gamma(t_\nu)^k \dot\gamma(t_\nu)^l w_\nu{}^i w_\nu{}^j).\end{aligned}$$

Since $\lim_{\nu\to\infty} g_{mn} v_\nu{}^m v_\nu{}^n = g_{mn} v^m v^n = 0$, the only way λ can be different from zero is for the sequence

$$\{L_{ikjl}\dot\gamma(t_\nu)^k \dot\gamma(t_\nu)^l w_\nu{}^i w_\nu{}^j\}$$

to be unbounded. But this implies that γ runs into a curvature singularity.

If λ represents an observer dependent curvature invariant other than a principal curvature, the proof is similar. For example, if $\lambda = T_{ij} v^i v^j$ then

$$\begin{aligned}\lambda &= \lim_{\nu\to\infty} T_{ij} v_\nu{}^i v_\nu{}^j \\ &= \lim_{\nu\to\infty} (g_{mn} v_\nu{}^m v_\nu{}^n)(T_{ij}\dot\gamma(t_\nu)^i \dot\gamma(t_\nu)^j)\end{aligned}$$

so $\lambda \neq 0$ only if energy density is unbounded along γ.

ACKNOWLEDGMENTS

I would like to thank the Mathematical Institute of Oxford University for its hospitality during the period when the bulk of this work was done. The relativity group there was most helpful to me. I am indebted to R. Penrose and P. Sommers for especially valuable discussions. In addition I am grateful to J. Ehlers and R. Geroch for their comments on an early draft of this paper.

*This research was partially supported by a grant from the Research Foundation of the State University of New York.

[1]S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space—time* (Cambridge U. P., Cambridge, 1973).
[2]G. F. R. Ellis and A. R. King, Comm. Math. Phys. **38**, 119 (1974).
[3]R. K. Sachs and H. Wu, *General Relativity for Mathematicians* (Springer-Verlag, New York and Berlin, 1976).
[4]S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry I* (Interscience, New York, 1963).
[5]S. W. Hawking, Astrophys. J. **145**, 544 (1966).
[6]A. Z. Petrov, *Einstein Spaces* (Pergamon, Elmsford, N.Y., 1969).
[7]J. A. Thorpe, J. Math. Phys. **10**, 1 (1969).
[8]S. W. Hawking and R. Penrose, Proc. Roy. Soc. (London) A **314**, 529 (1970).
[9]E. T. Newman, L. Tamburino, and T. J. Unti, J. Math. Phys. **4**, 915 (1963).
[10]B. G. Schmidt, J. Gen. Rel. Grav. **1**, 269 (1971).
[11]C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973), p. 860.
[12]Reference 1, p. 290.

Nonstandard vector connections given by nonstandard spinor connections*

C. P. Luehr

Department of Mathematics, University of Florida, Gainesville, Florida 32611

M. Rosenbaum† and M. P. Ryan, Jr.

Centro de Estudios Nucleares, Universidad Nacional Autónoma de México, México 20, D.F., Mexico

L. C. Shepley‡

Center for Relativity Theory, Physics Department, University of Texas at Austin, Austin, Texas 78712
(Received 5 August 1976)

A unique two-component spinor connection, which we call the standard connection, is determined by the requirement that it be compatible with the spinor inner product and that it give rise to the standard 4-vector connection. Here we take the most general spinor connection, presuming that the conjugate spinor connection is uniquely determined by it, and examine which 4-vector connections are thereby determined. We classify such nonstandard spinor connections and the resulting 4-vector connections and show that the most general torsion tensor can be so generated. However, it is not possible to generate in this way the most general tensor describing incompatibility of the 4-vector connection with the 4-vector inner product. These results illuminate the relationship which must exist between nonstandard theories for spinors and for vectors.

I. INTRODUCTION

The universe is best modeled by a four-dimensional, pseudo-Riemannian manifold. Of the geometric structures with physical importance, the vector fields X and the inner product $\bullet$ seem to be most fundamental. To differentiate a vector field, a connection is needed, and it is hardly surprising that customarily a connection is used which is compatible with the inner product. Most often, too, the requirement is made that the torsion vanish, so that a unique connection ∇_X is determined, which we shall call the standard connection.

Here we examine a more general situation. Since two-component spinors may be used to form vectors, we consider the most general connection on this latter set. It has been shown[1] that a unique spinor connection (which we shall call the standard connection) is determined by the requirements that it be compatible with the spinor inner product $\blacktriangle$ and that it produce the standard vector connection when spinors are combined to form vectors. In extending this result, we determine how a general spinor connection determines a vector connection, and in particular, whether the most general vector connection arises from a spinor connection.

We consider linear connections, and we presume that the connection acting on complex conjugate spinors is given by the complex conjugate of the spinor connection. The results are first, that the most general vector connection is not given. Second, we classify those vector connections which are determined in this way, and exhibit the degree to which they violate the requirements of vector inner product compatibility and of being torsionless. Perhaps the most important of our results is that the general torsion tensor does arise from this linear formalism. However, the only compatibility tensor (the measure of compatibility between the vector inner product and the connection) which arises in this fashion is proportional to the 4-vector identity. The information which originates from this tensor is lost in the usual Palatini variational method.[2]

We interpret our results as shedding light on the relationship between possible nonstandard theories for spinors and vectors. Thus if a spinor field theory results in some general spinor connection, it would then seem most natural to use this connection to generate a vector connection. In general this vector connection would have torsion and would be incompatible with the vector inner product, On the other hand, it is desirable to generate a vector connection which is incompatible with the vector inner product in the most general way conceivable for use in a Palatini-type variational principle. We are therefore currently studying modified formalisms in order to investigate purely spinoral variational principles for relativity (without hybrid terms as are often used[3]).

In Sec. II we briefly review some of the basic concepts of spinorial analysis. In Sec. III we parametrize the most general linear spin connection. In this section we prove several theorems relating the 4-vector torsion and metric compatibility tensors to these parameters. We discuss the necessary generalization we feel these results point to in the conclusion, Sec. IV. For the purpose of completeness, we have included an appendix where we state several helpful identities using permutation and contraction operators. Throughout we use the notation of Luehr and Rosenbaum.[1] As needed, we briefly sketch some of the results of their paper, but without any attempt at completeness. For a more complete and precise development, we refer the reader to it and the references cited there.

II. SPINOR SPACES AND CONNECTIONS

At each point q in the spacetime manifold $\mathcal{M}$, the vector space over the complex numbers of two-dimensional spinors is denoted by $(\mathcal{S}_2)_q$ (or briefly as $\mathcal{S}_2$, with q omitted) with elements u, v, etc. $\mathcal{S}_2$ is endowed with a bilinear, nonsingular inner product $\blacktriangle$ which is antisymmetric,

$$u \blacktriangle v = -v \blacktriangle u. \tag{1}$$

We will also use, u, v, etc. to denote spinor fields, with $u(q)$ being used for emphasis to denote the value of the spinor field u at a point q. In a given basis a spinor field is represented by a pair of complex functions on $\mathcal{M}$, and $u \blacktriangle v$ is a single complex function.

We analogously define a second two-dimensional complex vector space $\bar{\mathcal{S}}_2$ at q called conjugate spinor space, its elements being denoted $\bar{u}, \bar{v}$, etc. $\bar{\mathcal{S}}_2$ is endowed with a bilinear, nonsingular, antisymmetric inner product, also denoted by $\blacktriangle$. We will also use $\bar{u}, \bar{v}$, etc., to denote conjugate spinor fields with the value of $\bar{u}$ at q being denoted by $\bar{u}(q)$. The quantity $\bar{u} \blacktriangle \bar{v}$ is a single complex function.

At each point q, $\mathcal{S}_2$ and $\bar{\mathcal{S}}_2$ are related by maps $u \in \mathcal{S}_2 \to \bar{u} \in \bar{\mathcal{S}}_2$ and $\bar{w} \in \bar{\mathcal{S}}_2 \to \bar{\bar{w}} \in \mathcal{S}_2$, called the conjugation operation, having the following properties:

$$\overline{(\alpha u)} = \bar{\alpha}\bar{u}, \quad \overline{u+v} = \bar{u} + \bar{v}, \quad \overline{(\bar{u})} = u, \tag{2}$$

for all $u, v \in \mathcal{S}_2$ or $\bar{\mathcal{S}}_2$ and for all complex numbers α (where $\bar{\alpha}$ is the ordinary complex conjugate of α). Furthermore

$$\overline{u \blacktriangle v} = \bar{u} \blacktriangle \bar{v}. \tag{3}$$

Note that the conjugation operation can be regarded as a single map which is a bijection of the union $\mathcal{S}_2 \cup \bar{\mathcal{S}}_2$ onto itself.

The space $\bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2$ at a point q is isomorphic to the tangent space $\mathcal{M}_q$ of the manifold at q, and we will identify these two spaces. A 4-vector X thus is identified as a linear combination (over the reals) of terms of the form $\bar{u}u$. The expression $\bar{u}u$ is an abbreviated notation for the tensor product $\bar{u} \otimes u$.

The inner product of two elements A, B in $\bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2$ will be denoted $-A \overset{\blacktriangle}{\blacktriangle} B$. On the other hand, if A and B are thought of as tangent vectors of $\mathcal{M}$, their inner product [which is of signature $(+++-)$ since $\mathcal{M}$ is spacetime] is denoted $A \bullet B$. The relation between these notations is

$$-A \overset{\blacktriangle}{\blacktriangle} B = A \bullet B. \tag{4}$$

The introduction of the inner product $\bullet$ is redundant; it will be used when we wish to emphasize that the inner product of two spin tensors can be viewed as the inner product of vectors or vector tensors.

An $\mathcal{S}_2$ connection associates with each vector field X on $\mathcal{M}$ an operator on spinor fields. Here X is treated as a derivation operator, and the connection D'_X acting on $u(q)$ produces a spinor field $D'_X u(q)$. The connection obeys the axioms

$$\begin{aligned} &D'_X(fu+v) = (Xf)u + fD'_X u + D'_X v,\\ &D'_{gX+Y}u = gD'_X u + D'_Y v, \end{aligned} \tag{5}$$

where $u(q)$, $v(q)$ are any two spinor fields, X, Y are vector fields (derivation operators), $f(q)$ is any complex function, and $g(q)$ is any real function.

In analogy, an $\bar{\mathcal{S}}_2$ connection is an operator which obeys the same axioms as an $\mathcal{S}_2$ connection but which operates on conjugate spinors. We can extend D'_X to act on $\bar{\mathcal{S}}_2$ by the requirement

$$D'_X\bar{u} = \overline{D'_X u}. \tag{6}$$

In a future paper we will explore the possibility of dispensing with this relation and having an $\bar{\mathcal{S}}_2$ connection not necessarily related to a given $\mathcal{S}_2$ connection.

One $\mathcal{S}_2$ connection D_X is called "standard." D_X is uniquely determined by the requirements:

(a) When operating on $\bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2$ spin–tensor fields (i.e., 4-vector fields), the connection D_X coincides with the standard 4-vector connection. Thus it is torsionless and is compatible with the 4-vector inner product.

(b) D_X is compatible with the spinor inner product.

The most general linear $\mathcal{S}_2$ connection D'_X is uniquely given by a field $K_X(q)$ with entries in $\mathcal{S}_2 \otimes \mathcal{S}_2$ by the relation

$$D'_X u = D_X u + K_X \blacktriangle u. \tag{7}$$

Similarly the most general $\bar{\mathcal{S}}_2$ connection D''_X is uniquely given by a field $\bar{L}_X(q)$ with entries in $\bar{\mathcal{S}}_2 \otimes \bar{\mathcal{S}}_2$ by the relation

$$D''_X\bar{u} = D_X\bar{u} + \bar{L}_X \blacktriangle \bar{u}. \tag{8}$$

Notice that the map $u \in \mathcal{S}_2 \to \bar{u} \in \bar{\mathcal{S}}_2$ generates a unique map $K_X \in \mathcal{S}_2 \otimes \mathcal{S}_2 \to \bar{K}_X \in \bar{\mathcal{S}}_2 \otimes \bar{\mathcal{S}}_2$. The relation $\overline{D'_X u} = D''_X \bar{u}$ is equivalent to

$$\bar{K}_X = \bar{L}_X. \tag{9}$$

As we said, we will demand this equality, $D''_X = D'_X$.

It was previously shown[1] that D'_X generates a unique $\bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2$ connection. Let A be a $\bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2$ spin–tensor field (i.e., a 4-vector field). The $\bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2$ connection is given by requiring that the Leibnitz rule for covariant differentiation hold with respect to tensor products and that covariant differentiation commute with the projection map $\bar{\mathcal{S}}_2 \otimes \mathcal{S}_2 \to \bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2$. The result is that D'_X gives the 4-vector connection

$$D'_X A = D_X A + \bar{K}_X \blacktriangle A - A \blacktriangle \tilde{K}_X. \tag{10}$$

Since the standard connection preserves the 4-vector inner product, we have

$$X(A \bullet B) = (D_X A) \bullet B + A \bullet (D_X B). \tag{11}$$

Further, D_X is torsion free,

$$D_A B - D_B A = [A, B]. \tag{12}$$

The most general 4-vector connection is of the form

$$\nabla_X A = D_X A + C_X \bullet A. \tag{13}$$

The 4-vector connection D'_X generated by the general spinor connection produces C'_X,

$$D'_X A = D_X A + C'_X \bullet A = D_X A - C_X \overset{\blacktriangle}{\blacktriangle} A. \tag{14}$$

This C'_X is given by

$$C'_X = -(23)(\bar{K}_X I_2 + \bar{I}_2 K_X), \tag{15}$$

where I_2 and $\bar{I}_2$ are the identity operators on $\mathcal{S}_2$ and $\bar{\mathcal{S}}_2$ and (23) is a permutation on spinor files.

At this point we note that C_X is in $\mathcal{M}_q\otimes\mathcal{M}_q$ or equivalently in $(\bar{\mathcal{S}}_2\otimes_H\mathcal{S}_2)\otimes(\bar{\mathcal{S}}_2\otimes_H\mathcal{S}_2)$. Thus the general C_X is defined by 16 independent real numbers at each point $q \in \mathcal{M}$. The tensor B in $\mathcal{M}_q\otimes\mathcal{M}_q\otimes\mathcal{M}_q=\mathcal{M}_q^{\otimes 3}$ is defined by

$$X\bullet B=\tilde{C}_X. \tag{16}$$

B has 64 independent real components (at each q). The linear operator K_X has four independent complex components, or eight real ones. We also introduce the spin—tensor J defined by

$$X\bullet J=K_X. \tag{17}$$

J thus has 32 independent real components so that it is clear that the most general C_X cannot be of the form C'_X.

III. RELATIONS AMONG SPINOR AND VECTOR CONNECTIONS

The torsion tensor T associated with a 4-vector connection ∇_X is defined by

$$XY \,\vdots\, T=\nabla_X Y-\nabla_Y X-[X,Y]. \tag{18}$$

Thus

$$T=[1-(12)]B, \tag{19}$$

where (12) is a permutation of vector files. In terms of spinor spaces, at a point q, B is a member of $(\bar{\mathcal{S}}_2\otimes_H\mathcal{S}_2)^{\otimes 3}$ and

$$T=[1-(13)(24)]B. \tag{20}$$

The compatibility tensor U is defined by

$$XYY \,\vdots\, U=X(Y\bullet Z)-(\nabla_X Y)\bullet Z-Y\bullet(\nabla_X Z), \tag{21}$$

so that U vanishes if and only if ∇_X is compatible with the 4-vector inner product. In terms of B, U is given by

$$U=-[1+(23)]B=-[1+(35)(46)]B. \tag{22}$$

The tensor B' generated by the spin—tensor J is given by

$$\begin{aligned} X\bullet B'=\tilde{C}'_X &=(12)C'_X=-(13)(24)(23)(\bar{K}_X I_2+\bar{I}_2K_X)\\ &=+(13)(24)(23)\{X\blacktriangle[(12)\bar{J}I_2+(35)(46)J\bar{I}_2]\}\\ &=-X\blacktriangle[(354)(12)\bar{J}I_2-(36)J\bar{I}_2]. \end{aligned} \tag{23}$$

Thus we have

$$B'=(354)(12)\bar{J}I_2-(36)J\bar{I}_2. \tag{24}$$

Consequently the torsion tensor T' associated with B' is

$$T'=[1-(13)(24)][(354)(12)\bar{J}I_2-(36)J\bar{I}_2]. \tag{25}$$

The compatibility tensor U' associated with B' is

$$U'=-[1+(35)(46)][(354)(12)\bar{J}I_2-(36)J\bar{I}_2]. \tag{26}$$

This expression can be rewritten as

$$\begin{aligned} U'&=(12)(54)[1-(34)]\bar{J}I_2+(36)[1-(34)]J\bar{I}_2\\ &=[(12)C_{34}\bar{J}+C_{34}J]I_4, \end{aligned} \tag{27}$$

where C_{34} is contraction on the third and fourth spinor files. Note that $I_4=-(23)\bar{I}_2I_2=$ 4-vector identity operator.

We decompose J into independent parts which we will later relate to vector connections. In order to do so, we first need a lemma:

Lemma: (23) $(\nu+i\rho)I_2=[1-(24)](\nu+i\rho)I_2$ where ν,ρ are 4-vectors.

Proof: $\nu+i\rho$ is a sum of terms of the form $\bar{a}b$ in terms of spinors. We choose an arbitrary spinor v and write $v\blacktriangle(12)(23)\bar{a}bI_2=\bar{a}bv$. However, note that

$$\begin{aligned} \bar{a}bv&=\bar{a}(bv-vb)+\bar{a}vb\\ &=(v\blacktriangle b)\bar{a}I_2+\bar{a}vb\\ &=v\blacktriangle(12)\bar{a}bI_2-v\blacktriangle(12)(24)\bar{a}bI_2. \end{aligned}$$

Since v is arbitrary we have

$$(23)\bar{a}bI_2=[1-(24)]\bar{a}bI_2. \qquad \text{QED} \tag{28}$$

Theorem 1: Any J can be expressed as (here θ,ϕ,μ,σ are 4-vectors or Hermitian spin tensors)

$$J=(\theta+i\phi)I_2+(24)(\mu+i\sigma)I_2+H, \tag{29}$$

where H is completely symmetric on the last three spinor files,

$$(23)H=(34)H=(24)H=(234)H=(243)H=H.$$

Notice that $J\in(\bar{\mathcal{S}}_2\otimes_H\mathcal{S}_2)\otimes\mathcal{S}_2\otimes\bar{\mathcal{S}}_2$, but the individual terms in the right side of Eq. (29) are in general in $\bar{\mathcal{S}}_2\otimes\mathcal{S}_2\otimes\mathcal{S}_2\otimes\mathcal{S}_2$ (these spaces are, however, the same—see below).

Proof: Define the symmetrizing operator by

$$\mathcal{S}(234)=\tfrac{1}{6}[1+(234)+(243)+(23)+(34)+(24)]$$

and let

$$H=\mathcal{S}(234)J. \tag{30}$$

Similarly the antisymmetrizing operator is

$$\mathcal{A}(234)=\tfrac{1}{6}[1+(234)+(243)-(23)-(34)-(24)].$$

Now

$$\mathcal{A}(234)J=0, \tag{31}$$

since J is in the tensor product of two-dimensional spaces. We rewrite this equation as

$$J=H+\tfrac{1}{3}\{[1-(23)]+[1-(24)]+[1-(34)]\}J.$$

Since any antisymmetric spin—tensor is proportional to I_2, we have

$$\tfrac{1}{3}[1-(34)]J=(\theta'+i\phi')I_2, \tag{32a}$$

$$\tfrac{1}{3}[1-(23)]J=(24)(\mu'+i\sigma')I_2, \tag{32b}$$

$$\tfrac{1}{3}[1-(24)]J=(23)(\nu+i\rho)I_2, \tag{32c}$$

for suitable vectors $\theta',\phi',\mu',\sigma',\nu,\rho$.

By the previous lemma, Eq. (32c) is a combination of Eqs. (32a) and (32b), and the decomposition, Eq. (29), results. QED

Theorem 2: The decomposition for J is unique; that is, at a point q, $J=0$ if and only if $\theta=\phi=\mu=\sigma=0$, $H=0$.

Proof: Consider $J\blacktriangle I_2$. Since $H\blacktriangle I_2=0$ we have

$$J\blacktriangle I_2=2(\theta+i\phi)+(\mu+i\sigma).$$

Similarly

$$[(24)J]\blacktriangle I_2=\theta+i\phi+2(\mu+i\sigma).$$

Consequently

$$\theta + i\phi = \tfrac{1}{3}[2J - (24)J] \blacktriangle I_2, \tag{33}$$

$$\mu + i\sigma = \tfrac{1}{3}[2(24)J - J] \blacktriangle I_2. \tag{34}$$

Equations (29), (30), (33), and (34) show that there is a one-to-one correspondence between J and the set $(\theta, \phi, \mu, \sigma, H)$. QED

As a check we count the number of linearly independent parameters involved in J. As pointed out below Eq. (17), J has 32 real independent parameters. This number can be found by counting the dimension of $\bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2$ (four real dimensions) times the dimension of $\mathcal{S}_2 \otimes \mathcal{S}_2$ (four complex or eight real dimensions) since $J \in (\bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2) \otimes \mathcal{S}_2 \otimes \mathcal{S}_2$. Notice that in Theorems 1 and 2 we considered J as being in $\bar{\mathcal{S}}_2 \otimes \mathcal{S}_2 \otimes \mathcal{S}_2 \otimes \mathcal{S}_2$ which also has dimension 16 complex or 32 real. The space $(\bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2) \otimes \mathcal{S}_2 \otimes \mathcal{S}_2$ is in fact the same as $\bar{\mathcal{S}}_2 \otimes \mathcal{S}_2 \otimes \mathcal{S}_2 \otimes \mathcal{S}_2$. The reason for this equality comes from the facts that $\mathcal{S}_2 = \mathcal{C}\mathcal{S}_2$ and $\bar{\mathcal{S}}_2 \otimes \mathcal{S}_2 = \mathcal{C}[\bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2]$ where $\mathcal{C}$ is complexification (multiplication by complex numbers). Thus we have

$$\begin{aligned}(\bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2) &\otimes \mathcal{S}_2 \otimes \mathcal{S}_2 \\ &= (\bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2) \otimes \mathcal{C}\mathcal{S}_2 \otimes \mathcal{C}\mathcal{S}_2 = \mathcal{C}[(\bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2) \otimes \mathcal{S}_2 \otimes \mathcal{S}_2] \\ &= [\mathcal{C}(\bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2)] \otimes \mathcal{S}_2 \otimes \mathcal{S}_2 = \bar{\mathcal{S}}_2 \otimes \mathcal{S}_2 \otimes \mathcal{S}_2 \otimes \mathcal{S}_2.\end{aligned}$$

Equivalently a typical element $\bar{a}bcd$ of $\bar{\mathcal{S}}_2 \otimes \mathcal{S}_2 \otimes \mathcal{S}_2 \otimes \mathcal{S}_2$ is a sum of elements from $(\bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2) \otimes \mathcal{S}_2 \otimes \mathcal{S}_2$:

$$\bar{a}bcd = \frac{1}{2i}[\overline{(a+ib)}(a+ib) - \bar{a}a - \bar{b}b + i\overline{(a+b)}(a+b) - i\bar{a}a - i\bar{b}b]\,cd.$$

Incidently, $(\bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2) \otimes \mathcal{S}_2^{\otimes n}$ is the same as $\bar{\mathcal{S}}_2 \otimes \mathcal{S}_2^{n+1}$ provided $n \geq 1$.

We now return to the general 4-vector connection. Each J determines a B', and the decomposition of J results in

$$\begin{aligned}B' &= (354)(12)\bar{J}I_2 - (36)J\bar{I}_2 \\ &= -2(45)\,\theta\bar{I}_2 I_2 - [(13)+(24)](45)\mu\bar{I}_2 I_2 + i[(13) - (24)](45)\sigma\bar{I}_2 I_2 + (45)(12)\bar{H}I_2 - (36)H\bar{I}_2.\end{aligned} \tag{35}$$

Notice that ϕ does not appear in B'. We can also write this expression as

$$B' = \{2\theta + [(13)+(24)]\mu - i[(13)-(24)]\sigma\}I_4 + (45)(12)\bar{H}I_2 - (36)H\bar{I}_2. \tag{36}$$

The analagous expression for T' is

$$T' = [1 - (13)(24)]\{2[\theta + i(24)\sigma]I_4 + (45)(12)\bar{H}I_2 - (36)\bar{H}I_2\}. \tag{37}$$

Similarly U' is given by

$$U' = -2(2\theta + \mu)I_4. \tag{38}$$

We now state a few theorems about the values of B', T', U' at at given point $q \in \mathcal{M}$:

Theorem 3: At each point q, $T' = 0$ if and only if $\theta = \sigma = 0$, $H = 0$.

Proof: Clearly $\theta = \sigma = 0$, $H = 0$ implies $T' = 0$. To show the converse we express these parameters in terms of T',

$$-[(36)\mathcal{S}(246)T'] \blacktriangle \bar{I}_2 = H, \tag{39}$$

$$-\tfrac{1}{6}\{\{(45)[1-(15)]\}[1-(24)]T'\} \blacktriangle I_2\}\} \blacktriangle \bar{I}_2 = \theta - i\sigma. \tag{40}$$

QED

Theorem 4: $U' = 0$ if and only if $2\theta + \mu = 0$.

Proof: See Eq. (38).

Theorem 5: $B' = 0$ if and only if $\theta = \mu = \sigma = 0$, $H = 0$.

Proof: We must show that if $B' = 0$, then we have $\theta = \mu = \sigma = 0$, $H = 0$. Since $B' = 0$ implies $T' = 0$, we have $\theta = \sigma = 0$, $H = 0$. Further $B' = 0$ implies $U' = 0$ so that $\mu = 0$ also. QED

Next we show that the most general torsion tensor T is of the form T'. Notice that T has 24 real independent components. T' does also: H is symmetric in the last three spinor files, so that H has eight complex or 16 real independent components. The two vectors θ and μ add eight more real components.

Theorem 6: Any torsion tensor T is of the form given by T' [see Eq. (25)], so that T can be given by a suitable spinor connection.

Proof: Define

$$\mathcal{A}(mn) = \tfrac{1}{2}[1 - (mn)], \quad \mathcal{S}(mn) = \tfrac{1}{2}[1 + (mn)].$$

It readily follows that

$$\begin{aligned}[1 - (1\ 2)] &= [1 - (13)(24)] \\ &= 4[1 - (1\ 2)][\mathcal{A}(35)\mathcal{A}(13)\mathcal{S}(24) + \mathcal{A}(46)\mathcal{A}(24)\mathcal{S}(13)].\end{aligned} \tag{41}$$

Therefore,

$$T = [1 - (1\ 2)]B'', \tag{42}$$

where

$$B'' = 4[\mathcal{A}(35)\mathcal{A}(13)\mathcal{S}(24) + \mathcal{A}(46)\mathcal{A}(24)\mathcal{S}(13)]B. \tag{43}$$

Now define J'' by

$$-(36)J''\bar{I}_2 = 4\mathcal{A}(35)\mathcal{A}(13)\mathcal{S}(24)B \tag{44}$$

to correspond to the second term in Eq. (25). Equivalently, we have the definition

$$J'' = 2C_{56}(36)\mathcal{A}(13)\mathcal{S}(24)B. \tag{45}$$

We note that since $B \in (\bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2)^{\otimes 3}$,

$$\bar{B} = (12)(34)(56)B. \tag{46}$$

We take the complex conjugate of Eq. (44),

$$-(36)\bar{J}''I_2 = 4\mathcal{A}(35)\mathcal{A}(13)\mathcal{S}(24)(12)(34)(56)B \tag{47}$$

and multiply both sides by $-(354)(12)(36)$; and so

$$(354)(12)\bar{J}''I_2 = 4\mathcal{A}(46)\mathcal{A}(24)\mathcal{S}(13)B. \tag{48}$$

We therefore have

$$B'' = -(36)J''\bar{I}_2 + (354)(12)\bar{J}''I_2 \tag{49}$$

so that Eq. (25) does indeed hold, with the spinor connection being determined by $J = J''$ given by Eq. (45). QED

As a final note, we derive the condition on the parameters of J so that D'_X be compatible with the spinor inner product. From Eq. (29), we compute $K_X = -X \blacktriangle J$. When K_X is broken into a symmetric part and an antisymmetric part we find

$$K_X = -X\blacktriangle H - \int(12)X\blacktriangle(24)(\mu + i\sigma)I_2$$
$$-\tfrac{1}{2}[2X\blacktriangle(\theta + i\phi) + X\blacktriangle(\mu + i\sigma)]I_2. \qquad (50)$$

Theorem 7: D_X' is compatible with the spinor inner product if and only if $2\theta + \mu = 2\phi + \sigma = 0$. Such a D_X' gives rise only to a 4-vector connection which is compatible with the 4-vector inner product, so that $U' = 0$.

Proof: Recall that the condition that D_X' be compatible with the spinor inner product is

$$K_X = \widetilde{K}_X. \qquad (51)$$

Consequently the coefficient of I_2 in the last term of Eq. (50) must vanish for arbitrary X; that is,

$$2(\theta + i\phi) + (\mu + i\sigma) = 0. \qquad (52)$$

By Theorem 4, thus $U' = 0$ also. QED

IV. CONCLUSION

The most general linear 2-spinor connection is used to generate a connection on the conjugate spinors and in turn to generate a 4-vector connection. By parametrizing the spin—tensors J, which describe the spinor connections, we have exhibited those 4-vector connections which arise from them. One result can be viewed as an alternate proof of the uniqueness of the spinor connection which is compatible with the spinor inner product $\blacktriangle$, which generates a vector connection compatible with the vector inner product $\bullet$, and which results in vanishing torsion T.

The most general torsion T can be obtained from these spinor connections. This is the main result of this paper. It is not possible, however, to generate every tensor U which describes the degree to which the vector connection is incompatible with the vector inner product. The form of U which can be generated is proportional to the 4-vector identity operation I_4.

In searching for acceptable theories of spacetime geometry it is desirable to investigate nonstandard connections. Because of the undoubted importance of spinor fields,[4] it seems especially desirable to seek a spinorial formalism for these theories. Our results show that the usual spinorial formalism has to be modified in order to generate all possible nonstandard 4-vector connections from spinor connections. We are currently investigating such modifications and possible physical applications. One application, for example, is to the study of Palatini-type variational principles, where quite general variations in the connection are usually assumed to be possible.[2,5] Other applications include the study of spinor fields obeying theories with nonstandard connections in spatially homogeneous manifolds.

APPENDIX

We use an index-free language requiring facility at handling permutations and contractions. Permutations are treated in various texts,[6] but we feel it desirable to give here some helpful explicit expressions involving contractions. For definitions see Refs. 1 and 6.

In permutations such as (mn) or contractions such as C_{ab}, the integers m, n, a, b refer to files in spin—tensors. Remember that these integers refer to files counted from the left. C_{ab} decreases the total number of files by two, and this effect shows up particularly if commutation relations involving (mn) and C_{ab} are examined. In the following list, we take $m < n$ as well as $a < b$:

$$(mn)C_{ab} = C_{ab}(mn) \quad \text{if } m < n < a < b, \qquad \text{(A1a)}$$

$$(mn)C_{ab} = C_{ab}(m\ n+1) \quad \text{if } m < a \text{ and } a \le n < b-1, \qquad \text{(A1b)}$$

$$(mn)C_{ab} = C_{ab}(m\ n+2) \quad \text{if } m < a \text{ and } n \ge b-1, \qquad \text{(A1c)}$$

$$(mn)C_{ab} = C_{ab}(m+1\ n+1) \quad \text{if } a \le m < b-1 \text{ and } n < b-1, \qquad \text{(A1d)}$$

$$(mn)C_{ab} = C_{ab}(m+1\ n+2) \quad \text{if } a \le m < b-1 \text{ and } n \ge b-1, \qquad \text{(A1e)}$$

$$(mn)C_{ab} = C_{ab}(m+2\ n+2) \quad \text{if } m \ge b-1. \qquad \text{(A1f)}$$

Since C_{ab} does involve an antisymmetric inner product, as extra minus sign results if C_{ab} is preceded by a permutation using the same files

$$C_{ab}(ab) = -C_{ab}. \qquad \text{(A2)}$$

For equations in which the C_{ab} operation is preceded by a permutation involving either but not both indices, it is convenient to adopt the notation

$$(m \cdots a) = \begin{cases} (m\ m+1 \cdots a-1\ a) & \text{if } m \le a, \\ (m\ m-1 \cdots a+1\ a) & \text{if } a \le m, \end{cases} \qquad \text{(A3)}$$

for the cyclic permutation acting between the m and a files. We then have

$$C_{ab}(ma) = (m \cdots a-1)C_{mb} \quad \text{if } m < a, \qquad \text{(A4a)}$$

$$C_{ab}(ap) = (p-1 \cdots a)C_{pb} \quad \text{if } a < p < b, \qquad \text{(A4b)}$$

$$C_{ab}(an) = -(n-2 \cdots a)C_{bn} \quad \text{if } b < n, \qquad \text{(A4c)}$$

$$C_{ab}(mb) = -(m \cdots b-2)C_{ma} \quad \text{if } m < a, \qquad \text{(A4d)}$$

$$C_{ab}(pb) = (p-1 \cdots b-2)C_{ap} \quad \text{if } a < p < b, \qquad \text{(A4e)}$$

$$C_{ab}(bn) = (n-2 \cdots b-1)C_{an} \quad \text{if } b < n. \qquad \text{(A4f)}$$

Permutations can, of course, be applied to far more general functions. In the case of vector files, our convention is to emphasize the vector nature of a permutation by writing $(\mathrm{m\ n})$. In this paper we have tried to translate all vector quantities into equivalent spinorial quantities. An operation such as $(\mathrm{m\ n})$ is then equivalent to a pair of spinor operations $(ab)(a+1\ b+1)$. Each vector file corresponds to two spinor files (the left one actually a conjugate spinor file), so that $a = 2m-1$ when a vector—tensor is reinterpreted as a spin—tensor. However, mixed vector—spinor objects are common, so that $a \ne 2m-1$ in general. Further, it is not convenient to distinguish between spinor and conjugate spinor files. A permutation (mn) thus may involve changing a spinor with a conjugate spinor file.

*Supported in part by International Scientific Exchange Program Grants, National Science Foundation OIP75-09783 A01 and Consejo Nacional de Ciencia y Tecnologia No. 995.
†Asesor del Instituto Nacional de Energia Nuclear.
‡Supported in part also by National Science Foundation Grant MPS74-12498 A01.
[1]C. P. Luehr and M. Rosenbaum, J. Math. Phys. **15**, 1120 (1974).
[2]V. D. Sandberg, Phys. Rev. D **12**, 3013 (1975).
[3]See for example, A. Jamiolkowski, Rep. Math. Phys. **2**, 1 (1971) and E. Schmutzer, Z. Naturforsch. **19a**, 1027 (1964).
[4]R. Geroch, J. Math. Phys. **9**, 1739 (1968), Paper I; **11**, 343 (1970), Paper II.
[5]For a classification of types of 4-vector connections, see J. A. Schouten, *Ricci-Calculus* (Springer-Verlag, Berlin, 1954), 2nd ed.
[6]See, for example, A. Messiah, *Quantum Mechanics* (North-Holland, Amsterdam, 1961), Vol. II, Appendix D; M. Hamermesh, *Group Theory* (Addison-Wesley, Reading, Massachusetts, 1962), Chapter I.

Matrix superpropagators with derivatives

A. K. Kapoor

Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay 400005, India
(Received 28 October 1975; revised manuscript received 2 August 1976)

An earlier method of evaluation of matrix superpropagators without derivatives is extended to cover all such cases of interest. Matrix superpropagators with derivatives are reduced to superpropagators without derivatives by a straightforward application of Wick's theorem for time ordered products. A simple connection is found between the superpropagators involving the fields with derivatives at one of the points only and superpropagators obtained by replacing fields with derivatives by fields without derivatives. The results of the present paper are sufficient to allow evaluation of all superpropagators, with and without derivatives, encountered in the second order for nonlinear chiral Lagrangians.

I. INTRODUCTION

The use of suitable summation techniques for nonpolynomial Lagrangians has been shown to give supression of ultraviolet divergences.[1] The nonpolynomial Lagrangians naturally arise in gravity modified theories and also when nonlinear realizations of internal symmetry are used.[2] Nonlinear realizations of chiral symmetries were used to construct nonlinear Lagrangians which reproduced current algebra results in a simpler way when used in tree graph approximation.[3] It is therefore of interest to study these nonlinear Lagrangians beyond a tree graph approximation using nonpolynomial Lagrangian techniques. However, this requires methods for the calculation of vacuum expectation values (VEV's) of time ordered products of functions of matrix fields and their derivatives. These time ordered products, called superpropagators, have been studied extensively for the SU(2) case and methods exist for evaluation of SU(2) matrix superpropagators.[4] For gravity modified theories and theories with the SU(N) type of symmetry, a technique for calculation of matrix superpropagators,[5]

$$\langle T\mathcal{F}(\Phi(x))\mathcal{F}'(\Phi(y))\rangle_0, \tag{I.1}$$

without derivatives was developed by Ashmore and Delbourgo.[6] In Ref. 7 it was shown that the superpropagator (I.1) can be written as an integral over a set of complex matrices,

$$\langle T\mathcal{F}(\Phi(x))\mathcal{F}'(\Phi(y))\rangle_0 = N\int dU \exp(-\operatorname{Tr}(U^\dagger U))\mathcal{F}(cU)\mathcal{F}'(cU^\dagger), \tag{I.2}$$

where $\Phi(x)$ is an $n\times n$ Hermitian matrix field satisfying

$$\langle T\Phi_{ij}(x)\,\Phi_{kl}(y)\rangle_0 = \delta_{jk}\delta_{il}\Delta(x-y). \tag{I.3}$$

The integrations in (I.2) are over the set of all $n\times n$ matrices U with complex elements. The volume element dU is given by

$$dU = \prod_{i,j=1}^{n} dU_{ij}. \tag{I.4}$$

N is a normalization constant and $c^2=\Delta(x-y)$.

Assuming that the superpropagator to be calculated has the form

$$\langle TW(\Phi(x))W'(\Phi(y))\rangle_0, \tag{I.5}$$

where W and W' are scalar functions of $\Phi(x)$ and $\Phi(y)$, respectively, it was shown[7] that integrals could be easily carried out following methods developed in statistical mechanics.[8,9]

In this paper we show that it is possible to use the representation (I.2) even for those superpropagators, without derivatives, which may not have the form (I.5). Examples of such superpropagators are[10]

$$\langle T\{\exp[\lambda_1\Phi(x)]\exp[\lambda_2\Phi(y)]\}\rangle_0, \tag{I.6}$$

$$\langle T\{\exp[\lambda_1\Phi(x)]\exp[\lambda_2\Phi(y)]\}\{\exp[\lambda_3\Phi(x)]\exp[\lambda_4\Phi(y)]\}\rangle_0. \tag{I.7}$$

In Sec. II we review the method of Ref. 7 and its extension. As an illustration of the method of Sec. II we evaluate the superpropagator (I.6) and (I.7) in Sec. V. In Sec. III we show that, by a straightforward application of Wick's theorem for time ordered products, the superpropagators with derivatives can be reduced to superpropagators without derivatives. We obtain an extremely useful connection between superpropagators involving, on one hand, fields with derivatives at one of the points only and on the other hand, superpropagators obtained by removing derivatives. In Sec. IV we consider some examples of superpropagators with derivatives. Calculation of these reduces to evaluation of superpropagators of the type

$$\langle T\, e^{\mu_1\Phi(x)}_{i_1j_1}\cdots e^{\mu_m\Phi(x)}_{i_mj_m}\, e^{\mu_{m+1}\Phi(y)}_{k_1l_1}\cdots e^{\mu_{m+n}\Phi(y)}_{k_nl_n}\rangle_0. \tag{I.8}$$

In Sec. V we consider the evaluation of superpropagators of type (I.8) in detail. The last section contains some concluding remarks.

II. MATRIX SUPERPROPAGATORS WITHOUT DERIVATIVES

In this section we review the method of Ref. 7 and discuss its generalization for any matrix superpropagator without derivatives. Let $\Phi_{ij}(x)$, $i,j=1,2,\ldots,n$, be an $n\times n$ Hermitian matrix field.[11] Any such matrix field can be written as

$$\Phi=\sum_{\alpha=1}^{n^2}\frac{\lambda^\alpha}{\sqrt{2}}\phi_\alpha(x), \tag{II.1}$$

where $\phi_\alpha(x)$ $(\alpha=1,2,\ldots,n^2)$ are n^2 independent fields and λ^α are Hermitian matrices, the $n\times n$ generalizations of Gell-Mann matrices, obeying

$$\operatorname{Tr}(\lambda^\alpha\lambda^\beta)=2\delta_{\alpha\beta}. \tag{II.2}$$

We assume that all ϕ_α have the same mass and a two point function given by

$$\langle T\phi_\alpha(x)\phi_\beta(y)\rangle_0=\delta_{\alpha\beta}\Delta(x-y) \tag{II.3}$$

so that the matrix field Φ obeys (I.3).

Using the exponential shift lemma,[12,13] the VEV of a time ordered product of functions of fields ϕ_α may be written as

$$\langle T\mathcal{J}(\phi_\alpha(x),\phi_\beta(y))\rangle_0 = \frac{1}{\pi^{n^2}}\int\left(\prod_{\alpha=1}^{n^2}du_\alpha\right)\exp\left(-\sum_{\alpha=1}^{n^2}|u_\alpha|^2\right)\mathcal{J}(c_\alpha u_\alpha, c'_\beta u^*_\beta), \tag{II.4}$$

where

$c_\alpha c'_\alpha=\langle T\phi_\alpha(x)\phi_\alpha(y)\rangle_0$ (α is not summed).

If masses of all ϕ_α are equal, we can take $c_\alpha=c'_\alpha=c$, for all α.

Integrations over u_α in Eq. (II.4) run over all complex values. In matrix notation, the above representation for the superpropagator

$$\langle T\mathcal{J}(\Phi(x),\Phi(y))\rangle_0$$

takes the form[14]

$$\frac{1}{\pi^{n^2}}\int\left(\prod_{\alpha=1}^{n^2}du_\alpha\right)\exp\left(-\sum_{\alpha=1}^{n^2}|u_\alpha|^2\right)\mathcal{J}\left(c\sum_\alpha\lambda^\alpha u_{\alpha/\sqrt{2}},c\sum_\alpha\lambda^\alpha u^*_{\alpha/\sqrt{2}}\right)$$

$$=\frac{1}{\pi^{n^2}}\int dU\exp(-\operatorname{Tr}(U^\dagger U))\mathcal{J}(cU,cU^\dagger), \tag{II.5}$$

where we have defined

$$U=\sum\lambda_\alpha u_{\alpha/\sqrt{2}},\quad dU=\prod_{i,j}dU_{ij}. \tag{II.6}$$

Since u_α vary over all complex values, integration over U is on the set of all complex matrices.

In Ref. 7 we showed that the integrations in (II.5) could be performed when the function $\mathcal{J}$ factorizes as a product of two scalar functions $W(\Phi(x))$ and $W'(\Phi(y))$. For such superpropagators we obtained following formula[15] in Ref. 7,

$$\langle TW(\Phi(x))W'(\Phi(y))\rangle_0 \hat{=}\int dZ\prod_{j<k}|z_j-z_k|^2\exp\left(-\sum_{k=1}^{n}|z_k|^2\right)W(cZ)W'(cZ^\dagger), \tag{II.7}$$

where integrations in (II.7) now run over all complex diagonal matrices. We now show that the integral representation (II.5) can be used to evaluate superpropagators of the form[5,10]

$$\langle T\{\mathcal{A}(\Phi(x))\mathcal{B}(\Phi(y))\}\rangle_0, \tag{II.8a}$$

$$\langle T\{\mathcal{A}(\Phi(x))\mathcal{B}(\Phi(y))\}\{\mathcal{C}(\Phi(x))\mathcal{D}(\Phi(y))\}\rangle_0, \tag{II.8b}$$

$$\langle T\{\mathcal{A}(\Phi(x))\mathcal{B}(\Phi(y))\mathcal{C}(\Phi(x))\mathcal{D}(\Phi(y))\}\rangle_0, \tag{II.8c}$$

for which (II.7) is not applicable. The method of integration to be described is essentially the same as that of Ginibre[8] in a slightly different notation.

We will first give the details of the method of integration for superpropagator (II.8a) and the corresponding results for (II.8b) and (II.8c) will then be written. For the superpropagator (II.8a) we have[15]

$$\langle T\{\mathcal{A}(\Phi(x))\mathcal{B}(\Phi(y))\}\rangle_0 \hat{=}\int dU\exp(-\{U^\dagger U\})\{\mathcal{A}(cU)\mathcal{B}(cU^\dagger)\}. \tag{II.9}$$

As $n\times n$ matrices not having distinct eigenvalues form a set of measure zero in the space of all complex $n\times n$ matrices, we can restrict integrations in (II.9) over the set of all complex matrices with distinct eigenvalues. For any complex matrix U with distinct eigenvalues, there exists a nonsingular matrix X such that

$$U=XZX^{-1}, \tag{II.10}$$

where Z is a diagonal matrix,

$$(Z)_{ij}=z_i\delta_{ij}. \tag{II.11}$$

Changing variables from U to (X,Z) we obtain[16]

$$\int\prod_{i\neq j}(X^{-1}dX)_{ij}\prod_{i=1}^{n}dz_i\prod_{j<k}|z_j-z_k|^4 \times\exp[-\{Z^\dagger(X^\dagger X)Z(X^\dagger X)^{-1}\}] \times\{\mathcal{A}(cZ)(X^\dagger X)^{-1}\mathcal{B}(cZ^\dagger)(X^\dagger X)\} \tag{II.12}$$

for the right-hand side of (II.9).

A nonsingular matrix X can be written in the form

$$X=UYV, \tag{II.13}$$

where U is unitary, Y is upper triangular with the diagonal elements equal to 1 and V is a diagonal matrix. Performing the change of variable defined by (II.13) and defining

$$H=Y^\dagger Y, \tag{II.14}$$

the integral (II.12) can be written in the form

$$\int dZ\prod_{j<k}|z_j-z_k|^4\int dH\exp(-\{Z^\dagger HZH^{-1}\}) \times\{\mathcal{A}(cZ)H^{-1}\mathcal{B}(cZ^\dagger)H\}, \tag{II.15}$$

where

$$dH=\left(\prod_{i=1}^{n}dH_{ii}\right)\left(\prod_{i<j}dH_{ij}\right),\quad dZ=\prod_i dz_i, \tag{II.16}$$

H_{ij} $(i<j)$, z_k assume all complex values, and H_{ii} assume all real values (since H is Hermitian). For details of Jacobians of various transformations and of steps leading to (II.15) see Ginibre.[8] For later use it is sufficient to note that the matrix H is Hermitian, positive definite, and obeys

$$\det H=1,\quad \det H^{(p)}=1,\quad 1\leq p<n, \tag{II.17}$$

where $H^{(p)}$ is the matrix obtained by deleting the last p rows and p columns of the matrix H. Defining

$$(H')_{ij}=H_{ij},\quad g_i=H_{in},\quad i,j=1,2,\ldots,n-1,$$

we can write

$$H=\begin{bmatrix}H' & g\\ g^\dagger & H_{nn}\end{bmatrix}\quad (H'\equiv H^{(1)}). \tag{II.18}$$

We write H as[17]

$$H=\begin{bmatrix}H' & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}I_{n-1} & H'^{-1}g\\ g^\dagger & H_{nn}\end{bmatrix}. \tag{II.19}$$

Then the inverse of H can be easily worked out to be

$$H^{-1}=\begin{bmatrix} I_{n-1}+H'^{-1}gg^\dagger & -H'^{-1}g \\ -g^\dagger & 1 \end{bmatrix}\begin{bmatrix} H'^{-1} & 0 \\ 0 & 1\end{bmatrix}$$

$$=\begin{bmatrix} H'^{-1}+H'^{-1}gg^\dagger H'^{-1} & -H'^{-1}g \\ -g^\dagger H'^{-1} & 1\end{bmatrix}, \tag{II.20}$$

where we have made use of (II.17) to obtain

$$H_{nn}=1+g^\dagger H'^{-1}g$$

and we have eliminated H_{nn}. Using (II.18) and (II.20) we obtain

$$\{Z^\dagger HZH^{-1}\}=|z_n|^2+\{Z'^\dagger H'Z'H'^{-1}\}$$
$$+g^\dagger H'^{-1}(Z'^\dagger-z_n^*)H'(Z'-z_n)H'^{-1}g \tag{II.21}$$

and

$$\{\mathcal{A}(cZ)H^{-1}\mathcal{B}(cZ^\dagger)H\}$$
$$=a_nb_n+\{\mathcal{A}(cZ')H'^{-1}\mathcal{B}(cZ'^\dagger)H'\}$$
$$+g^\dagger H'^{-1}(\mathcal{B}(cZ'^\dagger)-b_n)H'(\mathcal{A}(cZ')-a_n)H'^{-1}g, \tag{II.22}$$

where Z' is the matrix obtained by deleting the last row and column of the matrix Z. a_n and b_n are the nth diagonal elements of the diagonal matrices $\mathcal{A}(cZ)$ and $\mathcal{B}(cZ^\dagger)$[18]:

$$a_n=(\mathcal{A}(cZ))_{nn}=\mathcal{A}(cz_n), \text{ etc.}$$

The integral (II.15) then takes the form

$$\int dZ\prod_{j<k}|z_j-z_k|^4\int dH'\exp(-|z_n|^2-\{Z'^\dagger H'Z'H'^{-1}\})$$
$$\times\int dg\exp[-g^\dagger H'^{-1}(Z'^\dagger-z_n^*)H'(Z'-z_n)H'^{-1}g]$$
$$\times[a_nb_n+\{\mathcal{A}(cZ')H'^{-1}\mathcal{B}(cZ'^\dagger)H'\}$$
$$+g^\dagger H'^{-1}(\mathcal{B}(cZ'^\dagger)-b_n)H'(\mathcal{A}(cZ')-a_n)H'^{-1}g], \tag{II.23}$$

where $dg=\prod_{i=1}^{n-1}dg_i$ and g integrations run over all complex values. The integration over g can be performed using[19]

$$\int dg\exp(-g^\dagger Qg)=\pi^{n-1}/\det Q, \tag{II.24}$$

$$\int dg\exp(-g^\dagger Qg)g^\dagger Ag=\frac{\pi^{n-1}}{\det Q}\operatorname{Tr}(AQ^{-1}), \tag{II.25}$$

where Q and A are $(n-1)\times(n-1)$ matrices independent of g. Carrying out the integration over g we obtain for (II.15)[20]

$$\int dZ\prod_{j<k}|z_j-z_k|^4\int dH\exp(-\{Z^\dagger HZH^{-1}\})$$
$$\times\{\mathcal{A}(cZ)H^{-1}\mathcal{B}(cZ^\dagger)H\}$$
$$\hat{=}\int dZ\prod_{j<k}|z_j-z_k|^4\int dH'\exp(-|z_n|^2-\{Z'^\dagger H'Z'H'^{-1}\})$$
$$\times\prod_{j=1}^{n-1}|z_j-z_n|^{-2}\Bigg[a_nb_n+\{\mathcal{A}(cZ')H'^{-1}\mathcal{B}(cZ'^\dagger)H'\}$$
$$+\left\{\left(\frac{\mathcal{A}(cZ')-a_n}{Z'-z_n}\right)H'^{-1}\left(\frac{\mathcal{B}(cZ'^\dagger)-b_n}{Z'^\dagger-z_n^*}\right)H'\right\}\Bigg]. \tag{II.26}$$

The above equation relates integrations over an $n\times n$ matrix H to integrations over an $(n-1)\times(n-1)$ matrix H'. Repeating steps from (II.18)–(II.26) $(n-1)$ times we will be left with integrations over Z alone.

Formulas similar to (II.26) for the superpropagators (II.8b) and (II.8c) can be easily obtained. For (II.8b) we have

$$\langle T\{\mathcal{A}(\Phi(x))\mathcal{B}(\Phi(y))\}\{\mathcal{C}(\Phi(x))\mathcal{D}(\Phi(y))\}\rangle_0$$
$$\hat{=}\int dZ\prod_{j<k}|z_j-z_k|^4\int dH\exp(-\{Z^\dagger HZH^{-1}\})$$
$$\times\{\mathcal{A}(cZ)H^{-1}\mathcal{B}(cZ^\dagger)H\}\{\mathcal{C}(cZ)H^{-1}\mathcal{D}(cZ^\dagger)H\}. \tag{II.27}$$

The traces $\{Z^\dagger HZH^{-1}\}$, $\{\mathcal{A}H^{-1}\mathcal{B}H\}$, and $\{\mathcal{C}H^{-1}\mathcal{D}H\}$ are again written in terms of H', g, etc, and integrations over g are performed[19] to obtain, for II.27)

$$\langle T\{\mathcal{A}(\Phi(x))\mathcal{B}(\Phi(y))\}\{\mathcal{C}(\Phi(x))\mathcal{D}(\Phi(y))\}\rangle_0$$
$$\hat{=}\int dZ\prod_{j<k}|z_j-z_k|^4\prod_{j<n}|z_j-z_n|^{-2}$$
$$\times\int dH'\exp(-|z_n|^2-\{Z'^\dagger H'Z'H'^{-1}\})$$
$$\times\Bigg(\Bigg[b_na_n+\{\mathcal{B}'H'\mathcal{A}'H'^{-1}\}+\Bigg\{\left(\frac{\mathcal{B}'-b_n}{Z'^\dagger-z_n^*}\right)H'$$
$$\times\left(\frac{\mathcal{A}'-a_n}{Z'-z_n}\right)H'^{-1}\Bigg\}\Bigg]\times\Bigg[d_nc_n+\{\mathcal{C}'H'^{-1}\mathcal{D}'H'\}$$
$$+\Bigg\{\left(\frac{\mathcal{D}'-d_n}{Z'^\dagger-z_n^*}\right)H'\left(\frac{\mathcal{C}'-c_n}{Z'-z_n}\right)H'^{-1}\Bigg\}\Bigg]$$
$$+\Bigg\{\left(\frac{\mathcal{B}'-b_n}{Z'^\dagger-z_n^*}\right)H'\left(\frac{\mathcal{A}'-a_n}{Z'-z_n}\right)H'^{-1}\left(\frac{\mathcal{D}'-d_n}{Z'^\dagger-z_n^*}\right)H'$$
$$\times\left(\frac{\mathcal{C}'-c_n}{Z'-z_n}\right)H'^{-1}\Bigg\}\Bigg), \tag{II.28}$$

where for the sake of simplicity we have omitted the arguments of $\mathcal{A}'$, $\mathcal{B}'$, $\mathcal{C}'$, and $\mathcal{D}'$.

Finally for the superpropagator (II.8c) we have

$$\langle T\{\mathcal{A}(\Phi(x))\mathcal{B}(\Phi(y))\mathcal{C}(\Phi(x))\mathcal{D}(\Phi(y))\}\rangle$$
$$\hat{=}\int dZ\prod_{j<k}|z_j-z_k|^4\int dH\exp(-\{Z^\dagger HZH^{-1}\})$$
$$\times\{\mathcal{A}(cZ)H^{-1}\mathcal{B}(cZ^\dagger)H\mathcal{C}(cZ)H^{-1}\mathcal{D}(cZ^\dagger)H\}$$
$$\hat{=}\int dZ\prod_{j<k}|z_j-z_k|^4\prod_{i<n}|z_i-z_n|^{-2}$$
$$\times\int dH'\exp(-|z_n|^2-\{Z'^\dagger H'Z'H'^{-1}\})$$
$$\times\Bigg[a_nb_nc_nd_n+\{\mathcal{A}'H'^{-1}\mathcal{B}'H'\mathcal{C}'H'^{-1}\mathcal{D}'H'\}$$
$$+a_nd_n\Bigg\{\left(\frac{\mathcal{B}'-b_n}{Z'^\dagger-z_n^*}\right)H'\left(\frac{\mathcal{C}'-c_n}{Z'-z_n}\right)H'^{-1}\Bigg\}$$
$$+b_nc_n\Bigg\{\left(\frac{\mathcal{D}'-d_n}{Z'^\dagger-z_n^*}\right)H'\left(\frac{\mathcal{A}'-a_n}{Z'-z_n}\right)H'^{-1}\Bigg\}$$

$$+\left\{\left(\frac{\mathcal{A}'-a_n}{Z'-z_n}\right)H'^{-1}\left(\frac{\mathcal{B}'-b_n}{Z'^{\dagger}-z_n^*}\right)H'\mathcal{C}'H'^{-1}\mathcal{D}'H'\right\}$$

$$+\left\{\mathcal{A}'H'^{-1}\mathcal{B}'H'\left(\frac{\mathcal{C}'-c_n}{Z'-z_n}\right)H'^{-1}\left(\frac{\mathcal{D}'-d_n}{Z'^{\dagger}-z_n^*}\right)H'\right\}$$

$$+\left\{\left(\frac{\mathcal{A}'-a_n}{Z'-z_n}\right)H'^{-1}\left(\frac{\mathcal{B}'-b_n}{Z'^{\dagger}-z_n^*}\right)H'\left(\frac{\mathcal{C}'-c_n}{Z'-z_n}\right)H'^{-1}\left(\frac{\mathcal{D}'-d_n}{Z'^{\dagger}-z_n^*}\right)H'\right\}$$

$$+\left\{\left(\frac{\mathcal{C}'-c_n}{Z'-z_n}\right)H'^{-1}\left(\frac{\mathcal{B}'-b_n}{Z'^{\dagger}-z_n^*}\right)H'\right\}\left\{\left(\frac{\mathcal{A}'-a_n}{Z'-z_n}\right)H'^{-1}$$

$$\times\left(\frac{\mathcal{D}'-d_n}{Z'^{\dagger}-z_n^*}\right)H'\right\}\Bigg]. \qquad \text{(II.29)}$$

We will now write down formula (II.26), (II.28), and (II.29) for the case $n=2$ explicitly. In this case the matrices H, $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$, and Z are two-dimensional while the primed matrices are one-dimensional. Also in view of Eq. (II.8), $H'=1$. Therefore, the H integrations in Eqs. (II.26), (II.28), and (II.29) take the form

$$\int dH \exp(-\{Z^{\dagger}HZH^{-1}\})\{\mathcal{A}(cZ)H^{-1}\mathcal{B}(cZ^{\dagger})H^{-1}\}$$

$$=\frac{\pi\exp(-|z_1|^2-|z_2|^2)}{|z_1-z_2|^2}\lfloor b_2a_2+b_1a_1+B_{21}A_{21}\rfloor, \qquad \text{(II.26')}$$

$$\int dH \exp(-\{Z^{\dagger}HZH^{-1}\})\{\mathcal{A}(cZ)H^{-1}\mathcal{B}(cZ^{\dagger})H\}$$

$$\times\{\mathcal{C}(cZ)H^{-1}\mathcal{D}(cZ^{\dagger})H\}$$

$$=\frac{\pi\exp(-|z_1|^2-|z_2|^2)}{|z_1-z_2|^2}[(b_2a_2+b_1a_1+B_{21}A_{21})$$

$$\times(d_2c_2+d_1c_1+D_{21}C_{21})+A_{21}B_{21}C_{21}D_{21}], \qquad \text{(II.28')}$$

$$\int dH \exp(-\{Z^{\dagger}HZH^{-1}\})\{\mathcal{A}H^{-1}\mathcal{B}H\mathcal{C}H^{-1}\mathcal{D}H\}$$

$$=\frac{\pi\exp(-|z_1|^2-|z_2|^2)}{|z_1-z_2|^2}(a_2b_2c_2d_2+a_1b_1c_1d_1$$

$$+a_2d_2B_{21}C_{21}+b_2c_2D_{21}A_{21}+c_1d_1A_{21}B_{21}$$

$$+a_1b_1C_{21}D_{21}+2A_{21}B_{21}C_{21}D_{21}) \qquad \text{(II.29')}$$

where we have defined

$$A_{21}=\frac{a_2-a_1}{z_2-z_1},\quad B_{21}=\frac{b_2-b_1}{z_2^*-z_1^*},$$

$$C_{21}=\frac{c_2-c_1}{z_2-z_1},\quad D_{21}=\frac{d_2-d_1}{z_2^*-z_1^*}. \qquad \text{(II.30)}$$

As an application of formulas (II.26)–(II.29) evaluation of superpropagators (I.6) and (I.7) will be discussed in detail in Sec. V.

III. REDUCTION OF SUPERPROPAGATORS WITH DERIVATIVES

In this section we will show that Wick's theorem for time ordered products can be used to reduce the superpropagators with derivatives to superpropagators without derivatives. These superpropagators without derivatives may then be evaluated using methods of Sec. II.

The steps for reducing the matrix superpropagators with derivatives to scalar superpropagators without derivatives are as follows:

(a) Suppose we are required to evaluate

$$\langle T\mathcal{F}_{ij}(\Phi(x),\partial_\mu\Phi(x))\mathcal{F}'_{kl}(\Phi(y),\partial_\nu\Phi(y))\rangle_0. \qquad \text{(III.1)}$$

The most general form for (III.1) is[21]

$$\delta_{ij}\delta_{kl}S_1+\delta_{il}\delta_{jk}S_2.$$

Then

$$9S_1+3S_2$$

$$=\langle T\{\mathcal{F}(\Phi(x),\partial_\mu\Phi(x))\}\{\mathcal{F}'(\Phi(y),\partial_\nu\Phi(y))\}\rangle_0$$

and

$$3S_1+3S_2$$

$$=\langle T\{\mathcal{F}(\Phi(x),\partial_\mu\Phi(x))\mathcal{F}'(\Phi(y),\partial_\nu\Phi(y))\}\rangle_0. \qquad \text{(III.2)}$$

This means that it is sufficient to evaluate various scalar superpropagators obtained by contracting internal indices in all possible ways.

(b) If we consider the time ordered product of a field $\phi(x)$ with a number of other operators, for example,

$$T(:\phi(x)K_1(\phi(x))::K_2(\phi(y)):), \qquad \text{(III.3)}$$

then using Wick's theorem for T products we get an expression of the form

$$:\phi(x)T(:K_1(\phi(x))::K_2(\phi(y)):):$$

$$+T:\underbrace{\phi(x)K_1(\phi(x))::K_2}(\phi(y)):, \qquad \text{(III.4)}$$

where $\underbrace{\phi(x)K_2}(\phi(y))$ stands for sum of terms obtained by pairing $\phi(x)$ chronologically with one of the $\phi(y)$ fields in all possible ways. Obviously

$$\underbrace{\phi(x)K_2}(\phi(y))=\langle T\phi(x)\phi(y)\rangle_0\frac{\partial}{\partial\phi(y)}K_2(\phi(y)).$$

Taking the VEV of (III.4) we obtain

$$\langle T:K_1(\phi(x))\phi(x)::K_2(\phi(y)):\rangle_0$$

$$=\langle T:K_1(\phi(x))::\frac{\partial}{\partial\phi(y)}K_2(\phi(y)):\rangle_0\langle T\phi(x)\phi(y)\rangle_0$$

as the VEV of the first term in Eq. (III.4) is zero.

(c) For the case when a multiplet of fields $\Phi(x)$ is present we have[13]

$$\langle T\mathcal{K}_1(\Phi(x))\Phi_{ij}(x)\mathcal{K}_2(\Phi(y))\rangle_0$$

$$=\langle T\Phi_{ij}(x)\Phi_{kl}(y)\rangle_0\left\langle T\mathcal{K}_1(\Phi(x))\frac{\partial}{\partial\Phi_{kl}(y)}\mathcal{K}_2(\Phi(y))\right\rangle_0. \qquad \text{(III.5)}$$

In view of Eq. (I.3) we obtain

$$\langle T\mathcal{K}_1(\Phi(x))\Phi_{ij}(x)\mathcal{K}_2(\Phi(y))\rangle_0$$

$$=\Delta(x-y)\left\langle T\mathcal{K}_1(\Phi(x))\frac{\partial}{\partial\Phi_{jl}}\mathcal{K}_2(\Phi(y))\right\rangle_0. \qquad \text{(III.6)}$$

Any superpropagator with derivatives can be reduced to a sum of superpropagators without derivatives by repeated use of Eq. (III.6) to pair off all fields having derivatives.

(d) Let us now assume that fields with derivatives are at the point x only. Let us further assume, for sake of definiteness, that there is only one derivative present, i.e., we are considering the superpropagators of the form

$$\langle T\mathcal{J}(\Phi(x))\partial_\mu\Phi(x)_{lm}\mathcal{J}'(\Phi(y))\rangle_0$$

$$=\partial^x_\mu\Delta(x-y)\left\langle T\mathcal{J}(\Phi(x))\,\frac{\partial}{\partial\Phi(y)_{ml}}\mathcal{J}'(\Phi(y))\right\rangle_0. \qquad \text{(III.7)}$$

If we replace $\partial_\mu\Phi(x)$ by $\Phi(x)$ in the left-hand side expression we get

$$\langle T\mathcal{J}(\Phi(x))\Phi(x)_{lm}\mathcal{J}'(\Phi(y))\rangle_0$$

$$=\Delta(x-y)\left\langle T\mathcal{J}(\Phi(x))\,\frac{\partial}{\partial\Phi(y)_{ml}}\mathcal{J}'(\Phi(y))\right\rangle_0.$$

Hence

$$\langle T\mathcal{J}(\Phi(x))\partial_\mu\Phi(x)_{lm}\mathcal{J}'(\Phi(y))\rangle_0$$

$$=(\partial^x_\mu\Delta/\Delta)\langle T\mathcal{J}(\Phi(x))\Phi(x)_{lm}\mathcal{J}'(\Phi(y))\rangle_0.$$

It is obvious that this result is a general one. For example, we have,

$$\langle T\mathcal{J}(\Phi(x))D^{(1)}_x\Phi(x)_{l_1m_1}\cdots D^{(r)}_x\Phi(x)_{l_rm_r}\mathcal{J}'(\Phi(y))\rangle_0$$

$$=[D^{(1)}_x\Delta(x-y)D^{(2)}_x\Delta(x-y)\cdots D^{(r)}_x\Delta(x-y)/\Delta^r(x-y)]$$

$$\times\langle T\mathcal{J}(\Phi(x))\Phi(x)_{l_1m_1}\cdots\Phi(x)_{l_rm_r}\mathcal{J}'(\Phi(y))\rangle_0, \qquad \text{(III.8)}$$

where $D^{(1)}_x\cdots D^{(r)}_x$ are the differential operators at the point x.

Thus superpropagators having fields with derivatives at one of the points and no derivatives at the other are related in a simple way to superpropagators obtained by replacing fields with derivatives by fields without derivatives. As an example, we will have

$$\langle T(\exp[\lambda\Phi(x)]\partial_\mu\Phi(x)\exp[-\lambda\Phi(x)])_{ij}\,e^{\lambda\Phi(y)}_{kl}\rangle_0$$

$$=(\partial_\mu\Delta/\Delta)\langle T(\exp[\lambda\Phi(x)]\Phi(x)\exp[-\lambda\Phi(x)])_{ij}$$

$$\times e^{\lambda\Phi(y)}_{kl}\rangle_0$$

$$=(\partial_\mu\Delta/\Delta)\langle T\Phi(x)_{ij}\,e^{\lambda\Phi(y)}_{kl}\rangle_0$$

$$=(\partial_\mu\Delta/\Delta)\langle T\Phi(x)_{ij}\lambda\Phi(y)_{kl}\rangle_0$$

$$=\lambda\delta_{il}\delta_{jk}(\partial_\mu\Delta/\Delta).$$

IV. EXAMPLES FROM NONLINEAR LAGRANGIANS

In this section we consider some examples of superpropagators with derivatives from nonlinear chiral Lagrangians. Using the methods of Sec. III these are related to scalar superpropagators without derivatives.

Let $P_\alpha(x)$, $\alpha=0,1,\ldots,8$ be a nonet of pseudoscalar fields. Let us define the matrix $P(x)$ by

$$P(x)=\sum_{\alpha=0}^{8}\frac{\lambda^\alpha}{\sqrt{2}}P_\alpha(x).$$

Then assuming the same masses for all fields we have

$$\langle TP(x)_{ij}P(y)_{kl}\rangle=\delta_{il}\delta_{jk}\Delta(x-y). \qquad \text{(IV.1)}$$

A typical nonlinear Lagrangian for pseudoscalar mesons has the form[22]

$$\mathcal{L}=-\frac{1}{2C}\{\partial_\mu M^\dagger\partial_\mu M\}+\{A(M+M^\dagger)\}, \qquad \text{(IV.2)}$$

where A is a numerical matrix. In exponential parametrization[23]

$$M=\exp[i\lambda P(x)] \qquad \text{(IV.3)}$$

and for rational parametrization

$$M=[1+i\lambda P(x)][1-i\lambda P(x)]^{-1}. \qquad \text{(IV.4)}$$

If other fields, for example, fermions or gauge fields are present, then $(M^\dagger\partial_\mu M)_{ij}$ and $(M\partial_\mu M^\dagger)_{ij}$ also appear in the coupling. We will work with the exponential parametrization.

Using[24]

$$\frac{d}{d\xi}\exp[\mathcal{H}(\xi)]=\int_0^1 dt\exp[(1-t)\mathcal{H}(\xi)]\,\frac{d\mathcal{H}}{d\xi}\,\exp[t\mathcal{H}(\xi)] \qquad \text{(IV.5)}$$

for a matrix function $\mathcal{H}(\xi)$, we have[10]

$$\{\partial_\mu\exp[\lambda P(x)]\partial_\mu\exp[-\lambda P(x)]\}$$

$$=\int_0^1\int_0^1 dt\,du\{\exp[\lambda(1-t)P]\partial_\mu P(x)\exp(\lambda tP)$$

$$\times\exp[-\lambda(1-u)P]\partial_\mu P(x)\exp(-\lambda uP)\}$$

$$=\int_0^1\int_0^1 dt\,du\{\partial_\mu P(x)\exp[\lambda(-1+t+u)P]\partial_\mu P$$

$$\times\exp[-\lambda(-1+t+u)P]$$

and

$$\exp[\lambda P(x)]\,\partial_\mu\exp[-\lambda P(x)]=\int_0^1\exp(\lambda tP(\,\partial_\mu P(x)$$

$$\times\exp(-\lambda tP)\,dt.$$

Thus we only need to consider superpropagators involving the following functions:

$$\partial_\mu P(x)\exp[\lambda P(x)]\partial_\nu P(x)\exp[-\lambda P(x)],$$

$$\exp[\lambda P(x)]\partial_\mu P(x)\exp[-\lambda P(x)], \qquad \text{(IV.6)}$$

$$\exp[\lambda P(x)].$$

For the superpropagators involving the fields with derivatives at one of the points only, formula (III.8) directly relates them to superpropagators without derivatives. Therefore we shall now discuss examples of superpropagators with derivatives at both the points.

When derivatives are present at both the points x and y we pair off all the fields with derivatives at one of the points. Then we are left with a superpropagator with derivatives at one of the points for which results of the previous section are applicable. To contract fields $P(x)$ with the exponential function $\exp[\lambda P(y)]$ we cast the relation[25]

$$\underbrace{P(x)_{ij}[\exp(\lambda P(y))]_{kl}}=\lambda\Delta(x-y)\int_0^1 e^{\lambda tP(y)}_{kj}\,e^{\lambda(1-t)P(y)}_{il}\,dt \qquad \text{(IV.7)}$$

in the following two convenient forms:

$$\{\mathcal{A}\underbrace{P(x)\}\{\exp}[\lambda P(y)]\mathcal{B}\}$$

$$=\lambda\Delta(x-y)\int_0^1\{\exp[\lambda tP(y)]\mathcal{A}\exp[\lambda(1-t)P(y)]\mathcal{B}\}\,dt \qquad \text{(IV.8)}$$

and

$$\{\underbrace{P(x)\mathcal{A}}\exp[\lambda P(y)]\mathcal{B}\}$$
$$=\lambda\Delta(x-y)\int_0^1\{\exp[\lambda tP(y)]\mathcal{A}\}\{\exp[\lambda(1-t)P(y)]\mathcal{B}\}, \tag{IV.9}$$

where $\mathcal{A}$ and $\mathcal{B}$ are any two operator or c-number matrices.

Example 1: Let

$$\langle T(\exp[\lambda P(x)]\partial_\mu P(x)\exp[-\lambda P(x)])_{ij}$$
$$\times(\exp[\lambda' P(y)]\partial_\nu P(y)\exp[-\lambda' P(y)])_{kl}\rangle_0$$
$$=\delta_{ij}\delta_{kl}S^{(1)}_{\mu\nu}+\delta_{il}\delta_{jk}S^{(2)}_{\mu\nu}. \tag{IV.10}$$

Then

$$9S^{(1)}_{\mu\nu}+3S^{(2)}_{\mu\nu}$$
$$=\langle T\{\exp[\lambda P(x)]\partial_\mu P(x)\exp[-\lambda P(x)]\}$$
$$\times\{\exp[-\lambda' P(y)]\,\partial_\nu P(y)\exp[\lambda' P(y)]\}\rangle_0$$
$$=\langle T\{\partial_\mu P(x)\}\{\partial_\nu P(y)\}\rangle_0$$
$$=3\partial^x_\mu\partial^y_\nu\Delta(x-y)$$

and

$$3S^{(1)}_{\mu\nu}+9S^{(2)}_{\mu\nu}=\langle T\{\exp[\lambda P(x)]\partial_\mu P(x)\exp[-\lambda P(x)]$$
$$\times\exp[\lambda' P(y)]\partial_\nu P(y)\exp[-\lambda' P(y)]\}\rangle_0.$$

We first pair off $\partial_\mu P(x)$ and then make use of the result (III.8) of the previous section to get[26]

$$3S^{(1)}+9S^{(2)}$$
$$=\partial^x_\mu\partial^y_\nu\Delta(x-y)\langle T\{\exp[-\lambda P(x)]\exp[\lambda' P(y)]\}$$
$$\times\{\exp[-\lambda' P(y)]\exp[\lambda P(x)]\}\rangle_0$$
$$+\lambda'\partial^x_\mu\Delta(x-y)\int_0^1 dt\,\langle T\{\exp[-\lambda P(x)]\exp[\lambda' tP(y)]\}$$
$$\times\{\exp[\lambda' t' P(y)]\partial_\nu P(y)\exp[-\lambda' P(y)]\exp[\lambda P(x)]\}\rangle_0$$
$$-\lambda'\partial^x_\mu\Delta(x-y)\int_0^1\langle T\{\exp[-\lambda P(x)]\exp[\lambda' P(y)]\partial_\nu P(y)$$
$$\times\exp[\lambda' t' P(y)]\}\{\exp[-\lambda' tP(y)]\exp[\lambda P(x)]\}\rangle_0$$
$$=\partial^x_\mu\partial^y_\nu\Delta(x-y)\langle T\{\exp[-\lambda P(x)]\exp[\lambda' P(y)]\}$$
$$\times\{\exp[-\lambda' P(y)]\exp[\lambda P(x)]\}\rangle_0$$
$$+(\partial^x_\mu\Delta\partial^y_\nu\Delta/\Delta)[\lambda'\int_0^1\langle T\{\exp[-\lambda P(x)]\exp[\lambda' tP(y)]\}$$
$$\times\{\exp[-\lambda' tP(y)]P(y)\exp[\lambda P(x)]\}\rangle_0\,dt+\lambda'\to-\lambda']. \tag{IV.11}$$

Example 2: Now we reduce the superpropagator

$$S_{\mu\nu\rho\sigma}=\langle T\{\exp[\lambda P(x)]\partial_\mu P(x)\exp[-\lambda P(x)]\partial_\nu P(x)\}$$
$$\times\{\exp[\lambda' P(y)]\partial_\rho P(y)\exp[-\lambda' P(y)]\partial_\sigma P(y)\}\rangle_0$$

to the sum of superpropagators without derivatives. First contract $\partial_\mu P(x)$ and $\partial_\nu P(y)$ in all possible ways by making use of Eqs. (IV.8) and (IV.9). We thus get

$$S_{\mu\nu\rho\sigma}=\partial^x_\nu\partial^y_\rho\Delta(x-y)S^{(1)}_{\mu\sigma}+\lambda'\partial^x_\nu\Delta(x-y)S^{(2)}_{\mu\rho\sigma}$$
$$+(\lambda'\to-\lambda',\ \rho\longleftrightarrow\sigma), \tag{IV.12}$$

where

$$S^{(1)}_{\mu\sigma}=\langle T\{\exp[\lambda P(x)]\partial_\mu P(x)\exp[-\lambda P(x)]$$
$$\times\exp[-\lambda' P(y)]\partial_\sigma P(y)\exp[\lambda' P(y)]\}\rangle_0$$
$$=\partial^x_\mu\partial^y_\sigma\Delta(x-y)\langle T\{\exp[-\lambda P(x)]\exp[-\lambda' P(y)]\}$$
$$\times\{\exp[\lambda' P(y)]\exp[\lambda P(x)]\}\rangle_0+\lambda'\partial^x_\mu\Delta(x-y)$$
$$\times\int_0^1\langle\{\exp[-\lambda P(x)]\exp[-\lambda' tP(y)]\}$$
$$\times\{\exp[-\lambda' t' P(y)]\partial_\sigma P(y)\exp[\lambda' P(y)]\exp[\lambda P(x)]\}\rangle_0\,dt$$
$$-\lambda'\partial^x_\mu\Delta(x-y)\int_0^1\langle\{\exp[-\lambda P(x)]\exp[-\lambda' P(y)]\partial_\sigma P(y)$$
$$\times\exp[\lambda' tP(y)]\}\{\exp[\lambda' t' P(y)]\exp[\lambda P(x)]\}\rangle_0\,dt$$
$$=\partial^x_\mu\partial^y_\sigma\Delta\langle T\{\exp[-\lambda P(x)]\exp[-\lambda' P(y)]\}$$
$$\times\{\exp[\lambda' P(y)]\exp[\lambda P(x)]\}\rangle_0$$
$$+[\lambda'(\partial^x_\mu\Delta\partial^y_\sigma\Delta/\Delta)\int_0^1 dt\langle T\{\exp[-\lambda P(x)]\exp[-\lambda' tP(y)]\}$$
$$\times\{\exp[\lambda' tP(y)]P(y)\exp[\lambda P(x)]\}\rangle_0+\lambda'\to-\lambda']. \tag{IV.13}$$

Similarly,

$$S^{(2)}_{\mu\rho\sigma}=[(\partial^x_\mu\partial^y_\rho\Delta\partial^y_\sigma\Delta/\Delta)$$
$$\times\int_0^1 dt\langle T\{\exp[-\lambda P(x)]\exp[-\lambda' tP(y)]\}$$
$$\times\{\exp[\lambda P(x)]\exp[\lambda' tP(y)]P(y)\}\rangle_0$$
$$+(\rho\longleftrightarrow\sigma,\ \lambda\to-\lambda)]+\lambda(\partial^x_\mu\Delta\partial^y_\rho\Delta\partial^y_\sigma\Delta/\Delta^2)$$
$$\times[\lambda'\iint dt\,ds\;t\langle T\{\exp[-\lambda P(x)]\exp[\lambda' tsP(y)]\}$$
$$\times\{\exp[-\lambda' tsP(y)]P^2(y)\exp[\lambda P(x)]\}\rangle_0+\lambda\to-\lambda]$$
$$-\lambda'(\partial^x_\mu\Delta\partial^y_\rho\Delta\partial^y_\sigma\Delta/\Delta)\iint dt\,ds\langle T\{\exp[-\lambda P(x)]$$
$$\times\exp[-\lambda'(t-s)P(y)]P(y)\}$$
$$\times\{\exp[\lambda P(x)]\exp[\lambda'(t-s)P(y)]P(y)\}\rangle_0. \tag{IV.14}$$

One of the aims of the above exercise was to show that for superpropagators involving any two of the functions (IV.6) it is sufficient to calculate the superpropagator

$$\langle T\{\exp[\mu_1 P(x)]\exp[\mu_3 P(y)]\}\{\exp[\mu_2 P(x)]\exp[\mu_4 P(y)]\}\rangle_0. \tag{IV.15}$$

It must be emphasized that the simplication achieved by the use of result (III.8) is tremendous. If, for instance, in the second example above we do not use (III.8) and proceed to pair off all the four fields with derivatives we will end up with the task of evaluating superpropagators of type (IV.16) with m and n ranging from 1 to 6, as compared to the present range 1 to 2 when the result (III.8) is used. The superpropagator (IV.15) is a special case of matrix superpropagators

$$\langle T\, e^{\mu_1 P(x)}_{i_1 j_1} \cdots e^{\mu_m P(x)}_{i_m j_m}\, e^{\mu_{m+1} P(y)}_{k_1 l_1} \cdots e^{\mu_{m+n} P(y)}_{k_n l_n} \rangle_0 \qquad \text{(IV.16)}$$

which will be considered in the next section. If we start with some parametrization other than the exponential one, then we will encounter superpropagators of type (IV.16) with some other functions replacing exponentials. However, all such superpropagators can be calculated once we know superpropagators of type (IV.16).

V. THE SUPERPROPAGATOR $\langle e^{\mu_1 \Phi(x)} \otimes \cdots \otimes e^{\mu_{m+n}\Phi(y)} \rangle_0$

In this section we first evaluate the superpropagators (I.6) and (I.7) as an example of the method of Sec. II. We will then show that the evaluation of superpropagators of type (IV.16) can be considerably simplified by relating various superpropagators to each other.

Example 1. Let $\Phi(x)$ be a 3×3 Hermitian matrix field. Then

$$\begin{aligned}\langle T\{\exp[\lambda_1\Phi(x)]\exp[\lambda_2\Phi(y)]\}\rangle_0 &\mathrel{\hat{=}} \int dU \exp(-\{U^\dagger U\})\{\exp(\lambda_1 c U)\exp(\lambda_2 c U^\dagger)\} \\ &= \int dZ\, dH \prod_{j<k} |z_j - z_k|^4 \exp(-\{Z^\dagger H Z H^{-1}\}) \\ &\quad \times \{\exp[\sigma_1 Z] H^{-1} \exp(\sigma_2 Z^\dagger) H\} \qquad \text{(V.1)}\end{aligned}$$

where we have used (II.15) and defined $\sigma_1 = \lambda_1 c$ and $\sigma_2 = \lambda_{2c}$. Using (II.26) we obtain

$$\begin{aligned}&\int dH \exp(-\{Z^\dagger H Z H^{-1}\})\{\exp[\sigma_1 Z]H^{-1}\exp(\sigma_2 Z^\dagger)H\} \\ &= (\pi^2/|z_1 - z_3|^2|z_2 - z_3|^2) \\ &\quad \times \int dH' \exp(-|z_3|^2 - \{Z'^\dagger H' Z' H'^{-1}\}) \\ &\quad \times \Bigg[\exp(\sigma_1 z_3)\exp(\sigma_2 z_3^*) + \{\exp(\sigma_1 Z')H'^{-1}\exp(\sigma_2 Z'^\dagger)H'\} \\ &\quad + \left\{\left(\frac{\exp(\sigma_1 Z') - \exp(\sigma_1 z_3)}{Z' - z_3}\right) H'^{-1} \right. \\ &\quad \left. \times \left(\frac{\exp(\sigma_2 Z'^\dagger) - \exp(\sigma_2 z_3^*)}{Z'^\dagger - z_3^*}\right) H' \right\}\Bigg]. \qquad \text{(V.2)}\end{aligned}$$

Using (II.26′) for the H' integration, we obtain for (V.2),

$$\begin{aligned}&\frac{\pi^3 \exp(-\sum_k |z_k|^2)}{\prod_{j<k}|z_j - z_k|^2}\Bigg[\exp(\sigma_1 z_3)\exp(\sigma_2 z_3^*) + \exp(\sigma_1 z_2)\exp(\sigma_2 z_2^*) \\ &+ \exp(\sigma_1 z_1)\exp(\sigma_2 z_2^*) + \frac{\exp(\sigma_1 z_1) - \exp(\sigma_1 z_2)}{z_1 - z_2} \\ &\times \frac{\exp(\sigma_2 z_1^*) - \exp(\sigma_2 z_2^*)}{z_1^* - z_2^*} + \frac{\exp(\sigma_1 z_2) - \exp(\sigma_1 z_3)}{z_2 - z_3} \\ &\times \frac{\exp(\sigma_2 z_2^*) - \exp(\sigma_2 z_3^*)}{z_2^* - z_3^*} + \frac{\exp(\sigma_1 z_1) - \exp(\sigma_1 z_3)}{z_1 - z_3} \\ &\times \frac{\exp(\sigma_2 z_1^*) - \exp(\sigma_2 z_3^*)}{z_1^* - z_3^*} + \frac{1}{|z_1 - z_2|^2} \\ &\times \left(\frac{\exp(\sigma_1 z_1) - \exp(\sigma_1 z_3)}{z_1 - z_3} - \frac{\exp(\sigma_1 z_2) - \exp(\sigma_1 z_3)}{z_2 - z_3}\right) \\ &\times \left(\frac{\exp(\sigma_2 z_1^*) - \exp(\sigma_2 z_3^*)}{z_1^* - z_3^*} - \frac{\exp(\sigma_2 z_2^*) - \exp(\sigma_2 z_3^*)}{z_2^* - z_3^*}\right)\Bigg]. \qquad \text{(V.3)}\end{aligned}$$

Using (V.3), we get for (V.1),

$$\begin{aligned}&\langle T\{\exp[\lambda_1\Phi(x)]\exp[\lambda_2\Phi(y)]\}\rangle_0 \\ &\mathrel{\hat{=}} \int dZ \prod_{j<k} |z_j - z_k|^2 \exp(-|z_1|^2 - |z_2|^2 - |z_3|^2) \\ &\times [\sum_k \exp(\sigma_1 z_k + \sigma_2 z_k^*) + 3E_1(\sigma_1, \mathbf{z})E_1(\sigma_2, \mathbf{z}^*) \\ &+ E_2(\sigma_1, \mathbf{z})E_2(\sigma_2, \mathbf{z}^*) + 2E_1(\sigma_1, \mathbf{z})E_2(\sigma_2, \mathbf{z}^*)\sum_k z_k^* \\ &+ 2E_2(\sigma_1, \mathbf{z})E_1(\sigma_2, \mathbf{z}^*)\sum_k z_k + E_2(\sigma_1, \mathbf{z})E_2(\sigma_2, z^*) \\ &\times (2\sum_k |z_k|^2 + \sum_{j<k}(z_j^* z_k + z_k^* z_j)], \qquad \text{(V.4)}\end{aligned}$$

where the functions E_0, E_1, and E_2 are defined in Appendix A.

The integrals in (V.4) are similar to those discussed in Ref. 7 and can be evaluated by using techniques developed in statistical mechanics.[9] We first note that $\prod_{j<k}|z_j - z_k|^2$ is the square of the absolute value of the Vandermonde determinant,

$$\prod_{j<k}|z_j - z_k|^2 = \begin{vmatrix} 1 & 1 & 1 \\ z_1 & z_2 & z_3 \\ z_1^2 & z_2^2 & z_3^2 \end{vmatrix}^2 \qquad \text{(V.5)}$$

$$= \begin{vmatrix} 3 & \sum z_k^* & \sum z_k^{*2} \\ \sum z_k & \sum z_k^* z_k & \sum z_k^{*2} z_k \\ \sum z_k^2 & \sum z_k^* z_k^2 & \sum z_k^{*2} z_k^2 \end{vmatrix}. \qquad \text{(V.6)}$$

One of the integrals in (V.4) is

$$\int dZ \prod_{j<k} |z_j - z_k|^2 \exp(-|z_1|^2 - |z_2|^2 - |z_3|^2) \times \sum_k \exp(\sigma_1 z_k + \sigma_2 z_k^*). \qquad \text{(V.7)}$$

We substitute (V.6) for $\prod_{j<k}|z_j - z_k|^2$. Since the integrand in (V.7) is completely symmetric in z_1, z_2, and z_3 we replace the first column of (V.6) by 3, $3z_1$, and $3z_1^2$. We then eliminate z_1 from the other two columns by subtracting suitable multiples of the first column. This process can be repeated once again with z_2 and z_3 to give

$$\begin{aligned}&\int dZ \prod_{j<k} |z_j - z_k|^2 \exp(-|z_1|^2 - |z_2|^2 - |z_3|^2) \\ &\times \sum_k \exp(\sigma_1 z_k + \sigma_2 z_k^*) \\ &= 6\int dz_1\, dz_2\, dz_3 \exp(-|z_1|^2 - |z_2|^2 - |z_3|^2) \\ &\times \begin{vmatrix} 1 & z_2^* & z_3^{*2} \\ z_1 & z_2^* z_2 & z_3^{*2} z_3 \\ z_1^2 & z_2^* z_2^2 & z_3^{*2} z_3^2 \end{vmatrix} \sum_k \exp(\sigma_1 z_k + \sigma_2 z_k^*). \qquad \text{(V.8)}\end{aligned}$$

Each of the three integrals in (V.8) can be evaluated by integrating each column separately and using

$$\int dz \exp(-|z|^2) z^m z^{*n} = \pi\Gamma(m+1)\delta_{mn}. \qquad (V.9)$$

The integral (V.7) then becomes

$$6(6+6\zeta+\zeta^2)e^{\zeta},$$

where $\zeta=\sigma_1\sigma_2=\lambda_1\lambda_2 c^2=\lambda_1\lambda_2\Delta(x-y)$.

In the integrals involving E_1 and E_2 we use the formulas (A6), (A7), and (V.6). For example,

$$E_2(\sigma_1,\mathbf{z})E_2(\sigma_2,\mathbf{z}^*)\prod_{j<k}|z_j-z_k|^2$$
$$=\begin{vmatrix} 3 & \sum z_k^* & \sum e^{\sigma_2 z_k^*} \\ \sum z_k & \sum z_k^* z_k & \sum e^{\sigma_2 z_k^*} z_k \\ \sum e^{\sigma_1 z_k} & \sum z_k^* e^{\sigma_1 z_k} & \sum e^{\sigma_2 z_k^*} e^{\sigma_1 z_k}\end{vmatrix}. \qquad (V.10)$$

Again arguments, similar to those leading to (V.8) from (V.7), give

$$\int dz_1\, dz_2\, dz_3 \exp(-|z_1|^2-|z_2|^2-|z_3|^2)$$
$$\times\left(\sum|z_k|^2\right)\prod_{j<k}|z_j-z_k|^2 E_2(\sigma_1,\mathbf{z})E_2(\sigma_2,\mathbf{z}^*)$$
$$=6\int dz_1\, dz_2\, dz_3 \exp(-|z_1|^2-|z_2|^2-|z_3|^2)\sum_k|z_k|^2$$
$$\times\begin{vmatrix} 1 & z_2^* & e^{\sigma_2 z_3^*} \\ z_1 & z_2^* z_2 & e^{\sigma_2 z_3^*} z_3 \\ e^{\sigma_1 z_1} & z_2^* e^{\sigma_1 z_2} & e^{\sigma_2 z_3^*} e^{\sigma_1 z_3}\end{vmatrix}$$
$$\hat{=}\, 6(\zeta e^{\zeta}+4e^{\zeta}-5\zeta-4).$$

Other integrals in (V.4) are calculated in a similar way and we finally get

$$\langle T\{\exp[\lambda_1\Phi(x)]\exp[\lambda_2\Phi(y)]\}\rangle_0$$
$$=(\zeta-3)+(6+4\zeta+\zeta^2/2)e^{\zeta}, \qquad (V.11)$$

where the normalization constant has been fixed by noting that the left-hand side has value 3 for $\zeta=0$.

Example 2. The second example we wish to consider is

$$\langle T\{\exp[\lambda_1\Phi(x)]\exp[\lambda_2\Phi(y)]\}\{\exp[\lambda_3\Phi(x)]\exp[\lambda_4\Phi(y)]\}\rangle_0$$

$$\hat{=}\int dU \exp(-\{U^\dagger U\})\{\exp(\lambda_1 cU)\exp(\lambda_2 cU^\dagger)\}$$
$$\times\{\exp(\lambda_3 cU)\exp(\lambda_4 cU^\dagger)\} \qquad (V.12)$$
$$=\int dZ \prod_{j<k}|z_j-z_k|^4\int dH \exp(-\{Z^\dagger HZH^{-1}\})$$
$$\times\{\exp(\sigma_1 Z)H^{-1}\exp(\sigma_2 Z^\dagger)H\}\{\exp(\sigma_3 Z)H^{-1}\exp(\sigma_4 Z^\dagger)H\}$$
$$\hat{=}\int dZ\prod_{j<k}|z_j-z_k|^4\Big(1/\prod_{i<3}|z_i-z_3|^2\Big)\exp(-|z_3|^2)$$
$$\times\int dH' \exp(-\{Z'^\dagger H' Z' H'^{-1}\}$$
$$\times\Bigg(\Bigg[\exp(\sigma_1 z_3+\sigma_2 z_3^*)+\{\exp(\sigma_1 Z')H'^{-1}\exp(\sigma_2 Z'^\dagger)H'\}$$
$$+\Bigg\{\left(\frac{\exp(\sigma_1 Z')-\exp(\sigma_1 z_3)}{Z'-z_3}\right)H'^{-1}$$
$$\times\left(\frac{\exp(\sigma_2 Z'^\dagger)-\exp(\sigma_2 z_3^*)}{Z'^\dagger-z_3^*}\right)H'\Bigg\}\Bigg]$$
$$\times\Bigg[\exp(\sigma_3 z_3+\sigma_4 z_3^*)+\{\exp(\sigma_3 Z')H'^{-1}\exp(\sigma_4 Z'^\dagger)H'\}$$
$$+\Bigg\{\left(\frac{\exp(\sigma_3 Z')-\exp(\sigma_3 z_3)}{Z'-z_3}\right)H'^{-1}$$
$$\times\left(\frac{\exp(\sigma_4 Z'^\dagger)-\exp(\sigma_4 z_3^*)}{Z'^\dagger-z_3^*}\right)H'\Bigg\}\Bigg]$$
$$+\Bigg\{\left(\frac{\exp(\sigma_1 Z')-\exp(\sigma_1 z_3)}{Z'-z_3}\right)H'^{-1}$$
$$\times\left(\frac{\exp(\sigma_4 Z'^\dagger)-\exp(\sigma_4 z_3^*)}{Z'-z_3}\right)H'\left(\frac{e^{\sigma_3 Z'}-e^{\sigma_3 z_3}}{Z'-z_3}\right)H'^{-1}$$
$$\times\left(\frac{\exp(\sigma_2 Z'^\dagger)-\exp(\sigma_2 z_3^*)}{Z'^\dagger-z_3^*}\right)H'\Bigg\}\Bigg), \qquad (V.13)$$

where we have again defined $\sigma_k=c\lambda_k$ and used (II.28) in the last step. The integrals, in (V.13), over the 2×2 matrix H' can be written down immediately by making use of (II.28′) and (II.29′). We define

$$p_k=\exp(\sigma_1 z_k),\quad q_k=\exp(\sigma_2 z_k^*),\quad r_k=\exp(\sigma_3 z_k),$$
$$s_k=\exp(\sigma_4 z_k^*),\quad k=1,2,3, \qquad (V.14)$$

and

$$P_{jk}=\frac{p_j-p_k}{z_j-z_k},\quad Q_{jk}=\frac{q_j-q_k}{z_j^*-z_k^*},$$
$$R_{jk}=\frac{r_j-r_k}{z_j-z_k},\quad S_{jk}=\frac{s_j-s_k}{z_j^*-z_k^*},\quad j\neq k. \qquad (V.15)$$

Using (II.28′) and (II.29′) to write the H' integration and noting relation (A9) we obtain

$$\langle T\{\exp[\lambda_1\Phi(x)]\exp[\lambda_2\Phi(y)]\}$$
$$\times\{\exp[\lambda_3\Phi(x)]\exp[\lambda_4\Phi(x)]\}\rangle_0$$
$$\hat{=}\int dZ\prod_{j<k}|z_j-z_k|^2\exp\Big(-\sum_k|z_k|^2\Big)T(\mathbf{z}), \qquad (V.16)$$

where the integrand $T(\mathbf{z})$, after some algebra, can be written in the form

$$T(\mathbf{z})=\Big[\sum_i p_i q_i+\sum_{j<k}P_{jk}Q_{jk}+E_2(\sigma_1,\mathbf{z})E_2(\sigma_2,\mathbf{z}^*)\Big]$$
$$\times\Big[\sum_i r_i s_i+\sum_{j<k}R_{jk}S_{jk}+E_2(\sigma_3,\mathbf{z})E_2(\sigma_4,\mathbf{z}^*)\Big]$$
$$+\sum_{j<k}P_{jk}Q_{jk}E_2(\sigma_3,\mathbf{z})E_2(\sigma_4,\mathbf{z})$$
$$+\sum_{j<k}R_{jk}S_{jk}E_2(\sigma_1,\mathbf{z})E_2(\sigma_2,\mathbf{z}^*)$$
$$+\sum_{j<k}P_{jk}Q_{jk}R_{jk}S_{jk}$$
$$-\frac{1}{(z_1-z_2)}(P_{13}R_{23}-P_{23}R_{13})\frac{1}{(z_1^*-z_2^*)}(Q_{13}S_{23}-Q_{23}S_{13})$$
$$+3E_2(\sigma_1,\mathbf{z})E_2(\sigma_2,\mathbf{z}^*)E_2(\sigma_3,\mathbf{z})E_2(\sigma_4,\mathbf{z}^*). \qquad (V.17)$$

The expression for T can be written as

$$T(\mathbf{z})=\sum_k T_k(\mathbf{z}), \tag{V.18}$$

$$T_1(\mathbf{z})=\left(\sum_i p_i q_i\right)\left(\sum_i r_i s_i\right),$$
$$=\sum\exp(\sigma_1 z_i+\sigma_2 z_i^*)\sum\exp(\sigma_3 z_i+\sigma_4 z_i^*), \tag{V.19}$$

$$T_2(\mathbf{z})=\sum_i p_i q_i\sum_{j<k}R_{jk}S_{jk}+\sum_i r_i s_i\sum_{j<k}P_{jk}Q_{jk}$$
$$=\sum_{i=1}^{3}\sum_{j<k}\left[\exp(\sigma_1 z_i+\sigma_2 z_i^*)\right.$$
$$\times\left(\frac{\exp(\sigma_3 z_j)-\exp(\sigma_3 z_k)}{z_j-z_k}\right)\left(\frac{\exp(\sigma_4 z_j^*)-\exp(\sigma_4 z_k^*)}{z_j^*-z_k^*}\right)$$
$$\left.+\begin{pmatrix}\sigma_1\to\sigma_3\\ \sigma_2\to\sigma_4\end{pmatrix}\right], \tag{V.20}$$

$$T_3(\mathbf{z})=\sum_i p_i q_i E_2(\sigma_3,\mathbf{z})E_2(\sigma_4,\mathbf{z})+(\sigma_1\to\sigma_3,\sigma_2\to\sigma_4), \tag{V.21}$$

$$T_4(\mathbf{z})=4E_2(\sigma_1,\mathbf{z})E_2(\sigma_2,\mathbf{z}^*)E_2(\sigma_3,\mathbf{z})E_2(\sigma_4,\mathbf{z}^*), \tag{V.22}$$

$$T_5(\mathbf{z})=2E_2(\sigma_1,\mathbf{z})E_2(\sigma_2,\mathbf{z}^*)\sum_{j<k}R_{jk}S_{jk}, \tag{V.23}$$

$$T_6(\mathbf{z})=2E_2(\sigma_3,\mathbf{z})\,E_2(\sigma_4,\mathbf{z}^*)\sum_{j<k}P_{jk}Q_{jk}, \tag{V.24}$$

$$T_7(\mathbf{z})=\left(\sum_{j<k}P_{jk}Q_{jk}\right)\left(\sum_{j<k}R_{jk}S_{jk}\right), \tag{V.25}$$

$$T_8(\mathbf{z})=\sum_{j<k}P_{jk}Q_{jk}R_{jk}S_{jk}, \tag{V.26}$$

$$T_9(\mathbf{z})=\frac{-1}{|z_1-z_2|^2}(P_{13}Q_{23}-P_{23}Q_{13})(R_{13}S_{23}-R_{23}S_{13}). \tag{V.27}$$

Integrals involving T_1–T_3 can be evaluated by methods similar to those used in the previous example. For evaluation of the integrals of T_4–T_9 see Appendix B.

All superpropagators of type

$$\langle T\, e_{i_1 j_1}^{\mu_1\Phi(x)}\cdots e_{i_m j_m}^{\mu_m\Phi(x)}\cdot e_{k_1 l_1}^{\mu_{m+1}\Phi(y)}\cdots e_{k_n l_n}^{\mu_{m+n}\Phi(y)}\rangle_0 \tag{V.28}$$

can be evaluated by contracting various indices in all possible ways and then by applying the method of Sec. II to the resulting scalar superpropagators. However, we shall now give a simple method of getting some identities between the various scalar superpropagators so obtained. These identities simplify the calculation of a superpropagator of type (V.28) to a considerable extent. We illustrate this by means of examples of (V.28) with $m,n\leq 2$.

(a) $m=1, n=1$

The matrix superpropagator

$$\langle T\, e_{ij}^{\mu_1\Phi(x)}\, e_{kl}^{\mu_2\Phi(y)}\rangle_0 \tag{V.29}$$

can be written as

$$\delta_{ij}\delta_{kl}S_1+\delta_{il}\delta_{jk}S_2, \tag{V.30}$$

where

$$9S_1+3S_2=\langle T\{\exp[\mu_1\Phi(x)]\}\{\exp[\mu_2\Phi(y)]\}\rangle_0$$
$$\equiv S_{1,1}(\mu_1;\mu_2) \tag{V.31}$$

and

$$3S_1+9S_2=\langle T\{\exp[\mu_1\Phi(x)]\exp[\mu_2\Phi(y)]\}\rangle$$
$$\equiv S_{1,1}(\mu_1,\mu_2). \tag{V.32}$$

The superpropagator (V.29) was calculated by Ashmore and Delbourgo by first calculating $S_{1,1}(\mu_1;\mu_2)$ and then relating $S_{1,1}(\mu_1,\mu_2)$ by solving a set of recurrence relations. Here we give another method of obtaining a relation between (V.31) and (V.32). Consider

$$\frac{d}{d\mu_1}\langle T\{\exp[\mu_1\Phi(x)]\}\{\exp[\mu_2\Phi(y)]\}\rangle_0$$
$$=\langle T\{\Phi(x)\exp[\mu_1\Phi(x)]\}\{\exp[\mu_2\Phi(y)]\}\rangle_0$$
$$=\mu_2\Delta(x-y)\langle T\{\exp[\mu_1\Phi(x)]\exp[\mu_2\Phi(y)]\}\rangle_0, \tag{V.33}$$

where in the last step we have paired off the field $\Phi(x)$ using Eq. (IV.8). Equation (V.33) gives a simpler way to evaluate $S_{1,1}(\mu_1,\mu_2)$.

(b) $m=2, n=1$

In the case of superpropagator

$$\langle T\, e_{i_1 j_1}^{\mu_1\Phi(x)}\, e_{i_2 j_2}^{\mu_2\Phi(x)}\, e_{kl}^{\mu_3\Phi(y)}\rangle_0 \tag{V.34}$$

we must know the following scalar propagators:

$$S_{2,1}(\mu_1;\mu_2;\mu_3)$$
$$=\langle T\{\exp[\mu_1\Phi(x)]\}\{\exp[\mu_2\Phi(x)]\}\{\exp[\mu_3\Phi(y)]\}\rangle_0, \tag{V.35a}$$

$$S_{2,1}(\mu_1,\mu_2;\mu_3)$$
$$=\langle T\{\exp[\mu_1\Phi(x)]\exp[\mu_2\Phi(x)]\}\{\exp[\mu_3\Phi(y)]\}\rangle_0, \tag{V.35b}$$

$$S_{2,1}(\mu_2,\mu_3;\mu_1)$$
$$=\langle T\{\exp[\mu_2\Phi(x)]\exp[\mu_3\Phi(y)]\}\{\exp[\mu_1\Phi(x)]\}\rangle_0, \tag{V.35c}$$

$$S_{2,1}(\mu_3,\mu_1;\mu_2)$$
$$=\langle T\{\exp[\mu_3\Phi(y)]\exp[\mu_1\Phi(x)]\}\{\exp[\mu_2\Phi(x)]\}\rangle_0, \tag{V.35d}$$

$$S_{2,1}(\mu_1,\mu_2,\mu_3)$$
$$=\langle T\{\exp[\mu_1\Phi(x)]\exp[\mu_2\Phi(x)]\exp[\mu_3\Phi(y)]\}\rangle_0. \tag{V.35e}$$

Of the five superpropagators listed above, $S_{2,1}(\mu_1,\mu_2;\mu_3)$ can be calculated by using (II.7). $S_{2,1}(\mu_1,\mu_2;\mu_3)$ and $S_{2,1}(\mu_1,\mu_2,\mu_3)$ are of the same form as $S_{1,1}(\mu_1;\mu_2)$ and $S_{1,1}(\mu_1,\mu_2)$, respectively. Finally, derivatives of $S_{2,1}(\mu_1;\mu_2;\mu_3)$ w.r.t. μ_1 and μ_2 are seen to be linear combinations of $S_{2,1}(\mu_2,\mu_3;\mu_1)$ and $S_{2,1}(\mu_3,\mu_1;\mu_2)$, by an argument similar to that used in derivation of (V.33). Hence one can solve for $S_{2,1}(\mu_2,\mu_3;\mu_1)$ and $S_{2,1}(\mu_3,\mu_1;\mu_2)$.

(c) $m=2, n=2$

For the matrix superpropagator

$$\langle T\, e_{i_1 j_1}^{\mu_1\Phi(x)}\, e_{i_2 j_2}^{\mu_2\Phi(x)}\, e_{k_1 l_1}^{\mu_3\Phi(y)}\, e_{k_2 l_2}^{\mu_4\Phi(y)}\rangle_0, \tag{V.36}$$

one can again write down the scalar superpropagators which must be evaluated. Of these some can be evaluated using representation (II.7) and some of the others can be related to derivatives of these. The only cases which are not covered in this fashion are

$$S_{2,2}(\mu_1,\mu_3;\mu_2,\mu_4)$$
$$=\langle T\{\exp[\mu_1\Phi(x)]\exp[\mu_3\Phi(y)]\}\{\exp[\mu_2\Phi(x)]\exp[\mu_4\Phi(y)]\}\rangle_0 \tag{V.37}$$

and

$$S_{2,2}(\mu_1,\mu_3,\mu_2,\mu_4)$$
$$=\langle T\{\exp[\mu_1\Phi(x)]\exp[\mu_3\Phi(y)]\exp[\mu_2\Phi(x)]\exp[\mu_4\Phi(y)]\}\rangle_0. \tag{V.38}$$

Evaluation of (V.37) has already been discussed and (V.38) can be calculated in a similar way.

Therefore all superpropagators of type (V.28) can be evaluated. Of these only superpropagators with $m,n\leq 2$ are needed for evaluation of superpropagators with derivatives.

VI. CONCLUSION

All the matrix superpropagators with or without derivatives needed for second order graphs can be evaluated by the methods presented here.

In the above discussion we have worked with exponential parametrization. However, one need not restrict oneself to the exponential parametrization; the whole discussion can be carried over for any other parametrization. The case when $\Phi(x)$ is a symmetric Hermitian matrix is also of interest for its application to gravity modified theories. In this case result (II.5) is valid except that integrations will now be over all complex symmetric matrices and the volume element will be $dU=\Pi_{i\leq j}dU_{ij}$. A method of integration over complex symmetric matrices is therefore needed.

ACKNOWLEDGMENTS

I wish to thank Mr. M.S. Sri Ram, with whom this work was originally started, for his comments on an earlier version of the manuscript and many useful discussions. I would like to thank Dr. H.S. Mani for his interest in the problem and encouragement while this work was in progress.

APPENDIX A

In this appendix we derive some useful results. Let U denote a 3×3 matrix. The matrix U will satisfy an equation of the form

$$U^3+p_1U^2+p_2U+p_3I=0 \tag{A1}$$

where p_1, p_2, and p_3 are scalar functions of the matrix U and can be expressed in terms of $\{U^n\}$, $n=1,2,3$. Hence any matrix function of U, such as $\exp(\lambda U)$, can be written as

$$\exp(\lambda U)=E_0+UE_1+U^2E_2, \tag{A2}$$

where E_0, E_1, and E_2 are functions of p_1, p_2, and p_3. Thus, for a diagonal matrix Z ($Z_{ij}=z_i\delta_{ij}$),

$$\exp(\lambda Z)=E_0+ZE_1+Z^2E_2, \tag{A3}$$

E_0, E_1, and E_2 will be functions of λ and $\mathbf{z}$. If the diagonal elements of Z are distinct, then the three equations

$$\begin{aligned}\exp(\lambda z_1)&=E_0+z_1E_1+z_1^2E_2,\\ \exp(\lambda z_2)&=E_0+z_2E_1+z_2^2E_2,\\ \exp(\lambda z_3)&=E_0+z_3E_1+z_3^2E_3,\end{aligned} \tag{A4}$$

obtained by equating diagonal elements of both sides of (A3), can be solved for E_k to obtain

$$E_0(\lambda,\mathbf{z})=\begin{vmatrix}e^{\lambda z_1} & z_1 & z_1^2\\ e^{\lambda z_2} & z_2 & z_2^2\\ e^{\lambda z_3} & z_3 & z_3^2\end{vmatrix}\Bigg/\begin{vmatrix}1 & z_1 & z_1^2\\ 1 & z_2 & z_2^2\\ 1 & z_3 & z_3^2\end{vmatrix}, \tag{A5}$$

$$E_1(\lambda,\mathbf{z})=\begin{vmatrix}1 & e^{\lambda z_1} & z_1^2\\ 1 & e^{\lambda z_2} & z_2^2\\ 1 & e^{\lambda z_3} & z_3^2\end{vmatrix}\Bigg/\begin{vmatrix}1 & z_1 & z_1^2\\ 1 & z_2 & z_2^2\\ 1 & z_3 & z_3^2\end{vmatrix}, \tag{A6}$$

and

$$E_2(\lambda,\mathbf{z})=\begin{vmatrix}1 & z_1 & e^{\lambda z_1}\\ 1 & z_2 & e^{\lambda z_2}\\ 1 & z_3 & e^{\lambda z_3}\end{vmatrix}\Bigg/\begin{vmatrix}1 & z_1 & z_1^2\\ 1 & z_2 & z_2^2\\ 1 & z_3 & z_3^2\end{vmatrix}. \tag{A7}$$

It must be noted that E_k are completely symmetric in z_1, z_2, and z_3.

For use in Sec. V we record the following formulas for the E's:

$$\frac{\exp(\lambda z_j)-\exp(\lambda z_k)}{z_j-z_k}=E_1+(z_j+z_k)E_2,\quad j\neq k, \tag{A8}$$

and

$$\frac{1}{z_1-z_2}\left(\frac{\exp(\lambda z_1)-\exp(\lambda z_3)}{z_1-z_3}-\frac{\exp(\lambda z_2)-\exp(\lambda z_3)}{z_2-z_3}\right)=E_2. \tag{A9}$$

These results follow directly from Eqs. (A4).

The left-hand side of (A8) can be written as (assuming $j=1$, $k=2$),

$$\begin{aligned}\frac{\exp(\lambda z_1)-\exp(\lambda z_2)}{z_1-z_2}&=\exp(\lambda z_2)\,\frac{\exp[\lambda(z_1-z_2)-1]}{z_1-z_2}\\ &=\lambda\exp(\lambda z_2)\int_0^1\exp[\lambda\alpha(z_1-z_2)]\,d\alpha\\ &=\lambda\int_0^1\exp[\lambda\alpha z_1+\lambda(1-\alpha)z_2]\,d\alpha.\end{aligned} \tag{A10}$$

Equation (A10) can be put in the form

$$\begin{aligned}\frac{\exp(\lambda z_1)-\exp(\lambda z_2)}{z_1-z_2}&=\lambda\int d\alpha_1\,d\alpha_2\,\theta(\alpha_1)\theta(\alpha_2)\delta(\alpha_1+\alpha_2-1)\\ &\quad\times\exp(\alpha_1z_1+\alpha_2z_2)\\ &=\lambda\int d\alpha\,\delta(\alpha_1+\alpha_2+\alpha_3-1)\theta(\alpha_1)\theta(\alpha_2)\delta(\alpha_3)\\ &\quad\times\exp(\lambda\boldsymbol{\alpha}\cdot\mathbf{z}),\end{aligned} \tag{A11}$$

where $\boldsymbol{\alpha}\cdot\mathbf{z}=\sum_{k=1}^{3}\alpha_kz_k$ and $d\alpha=\Pi_k\,d\alpha_k$. Similarly,

$$E_2(\lambda,\mathbf{z})=\lambda^2\int d\boldsymbol{\alpha}\,\delta\left(\sum_{i=1}^{3}\alpha_i-1\right)\prod_{i=1}^{3}\theta(\alpha_i)\times\exp(\lambda\boldsymbol{\alpha}\cdot\mathbf{z}). \tag{A12}$$

If we expand $\exp(\lambda\boldsymbol{\alpha}\cdot\mathbf{z})$ in powers of $\boldsymbol{\alpha}$ and integrate term by term we obtain

$$E_2(\lambda,\mathbf{z})=\sum_{m_i\geq 0}\frac{\lambda^{m+2}}{(m+2)!}z_1^{m_1}z_2^{m_2}z_3^{m_3},\quad m=\sum m_i. \tag{A13}$$

This power series expansion for E_2 can also be obtained from Eq. (A7) directly. One can write down similar power series for E_0 and E_1,

$$E_0(\lambda,\mathbf{z})=1+\sum_{m_i\geq 0}\frac{\lambda^{m+3}}{(m+3)}z_1^{m_1}z_2^{m_2}z_3^{m_3} \tag{A14}$$

and

$$E_1(\lambda,\mathbf{z})=\sum_{m_i\geq 0}\frac{\lambda^{m+1}}{(m+1)!}z_1^{m_1}z_2^{m_2}z_3^{m_3}-(z_1+z_2+z_3)\sum_{m_i\geq 0}\frac{\lambda^{m+2}}{(m+2)!}z_1^{m_1}z_2^{m_2}z_3^{m_3}, \tag{A15}$$

where

$$m=m_1+m_2+m_3.$$

Formula (A14) and (A15) can be proved directly from (A5) and (A6).

In place of $E_k(\lambda,\mathbf{z})$ we will sometimes write $E_k(\lambda,z_1,z_2,z_3)$.

APPENDIX B

In this appendix we discuss the details of the evaluation of integral (V.16) where T is given by Eqs. (V.18)–(V.27). As we have already remarked, T_1, T_2, and T_3 can be easily evaluated using methods of Example 1 of Sec. V. Therefore, we first consider integrals of T_4, T_5, and T_6. In all these integrands there is a factor of type

$$\prod_{j<k}|z_j-z_k|^2E_2(\sigma_1,\mathbf{z})E_2(\sigma_2',\mathbf{z}^*) \tag{B1}$$

which can be replaced by

$$6\begin{vmatrix}1 & z_2^* & e^{\sigma_2 z_3^*}\\ z_1 & z_2^*z_2 & e^{\sigma_2 z_3^*}z_3\\ e^{\sigma_1 z_1} & z_2^*e^{\sigma_1 z_2} & e^{\sigma_2 z_3^*}e^{\sigma_1 z_3}\end{vmatrix} \tag{B2}$$

inside the integral, by arguments similar to those used in writing Eq. (V.8) from (V.7). In view of Eqs. (V.15), (V.22)–(V.24), (A11), and (A12) we first consider integrals of the form

$$\int dz_1\,dz_2\,dz_3\exp\left(-\sum|z_k|^2\right)\begin{vmatrix}1 & z_2^* & e^{\sigma_2 z_3^*}\\ z_1 & z_2^*z_2 & e^{\sigma_2 z_3^*}z_3\\ e^{\sigma_1 z_1} & e^{\sigma_1 z_2^*} & e^{\sigma_1 z_3+\sigma_2 z_3^*}\end{vmatrix}\times\exp(\sigma_3\boldsymbol{\alpha}\cdot\mathbf{z}+\sigma_4\boldsymbol{\beta}\cdot\mathbf{z}), \tag{B3}$$

where $\boldsymbol{\alpha}=(\alpha_1,\alpha_2,\alpha_3)$ and $\boldsymbol{\alpha}\cdot\mathbf{z}=\sum_k\alpha_kz_k$, etc. As each column of the determinant in (B3) depends on only one of the integration variables, each column can be integrated separately to give

$$\int dz_1\,dz_2\,dz_3\exp(-|z_1|^2-|z_2|^2-|z_3|^2)\times\begin{vmatrix}1 & z_2^* & e^{\sigma_2 z_3^*}\\ z_1 & z_2^*z_2 & e^{\sigma_2 z_3^*}z_3\\ e^{\sigma_1 z_1} & z_2^*e^{\sigma_1 z_2} & e^{\sigma_2 z_3^*+\sigma_1 z_3}\end{vmatrix}\exp(\sigma_3\boldsymbol{\alpha}\cdot\mathbf{z}+\sigma_4\boldsymbol{\beta}\cdot\mathbf{z}^*)$$

$$=\frac{1}{\sigma_4}\frac{\partial}{\partial\beta_2}\int dz_1\,dz_2\,dz_3\times\exp(-|z_1|^2-|z_2|^2-|z_3|^2+\sigma_3\boldsymbol{\alpha}\cdot\mathbf{z}+\sigma_4\boldsymbol{\beta}\cdot\mathbf{z}^*)\times\begin{vmatrix}1 & 1 & e^{\sigma_2 z_3^*}\\ z_1 & z_2 & e^{\sigma_2 z_3^*}z_3\\ e^{\sigma_1 z_1} & e^{\sigma_1 z_2} & e^{\sigma_2 z_3^*+\sigma_1 z_3}\end{vmatrix}$$

$$=\frac{1}{\sigma_4}\frac{\partial}{\partial\beta_2}\left\{\begin{vmatrix}1 & 1 & 1\\ \sigma_4\beta_1 & \sigma_4\beta_2 & (\sigma_2+\sigma_4\beta_3)\\ e^{\sigma_1\sigma_4\beta_1} & e^{\sigma_1\sigma_4\beta_2} & e^{\sigma_1(\sigma_2+\sigma_4\beta_3)}\end{vmatrix}\times\exp(\sigma_3\sigma_4\boldsymbol{\alpha}\cdot\boldsymbol{\beta}+\sigma_3\sigma_2\alpha_3\beta_3)\right\} \tag{B4}$$

$$=\frac{1}{\sigma_4}\frac{\partial}{\partial\beta_2}\{D(\sigma_1,\sigma_2,\sigma_4,\beta)\exp(\sigma_3\sigma_4\boldsymbol{\alpha}\cdot\boldsymbol{\beta}+\sigma_3\sigma_2\alpha_3\beta_3)\}, \tag{B5}$$

where

$$D(\sigma_1,\sigma_2,\sigma_4\beta)=\begin{vmatrix}1 & 1 & 1\\ \sigma_4\beta_1 & \sigma_4\beta_2 & (\sigma_2+\sigma_4\beta_3)\\ e^{\sigma_1\sigma_4\beta_1} & e^{\sigma_1\sigma_4\beta_2} & e^{\sigma_1(\sigma_2+\sigma_4\beta_3)}\end{vmatrix}. \tag{B6}$$

Hence using (A12) and (B1)–(B5) we get

$$\int dZ\prod_{j<k}|z_j-z_k|^2\exp\left(-\sum|z_k|^2\right)\times E_2(\sigma_1,\mathbf{z})E_2(\sigma_2,\mathbf{z}^*)E_2(\sigma_3,\mathbf{z})E_2(\sigma_4,\mathbf{z}^*)$$

$$=(\sigma_3\sigma_4)^2\int d\boldsymbol{\beta}\,\delta\left(\sum\beta_k-1\right)\prod_k\theta(\beta_k)\times\int d\boldsymbol{\alpha}\,\delta\left(\sum\alpha_k-1\right)\prod_k\theta(\alpha_k)\times\frac{1}{\sigma_4}\frac{\partial}{\partial\beta_2}[\exp(\sigma_3\sigma_4\boldsymbol{\alpha}\cdot\boldsymbol{\beta}+\sigma_2\sigma_3\alpha_3\beta_3)D(\sigma_1,\sigma_2,\sigma_4,\boldsymbol{\beta})]. \tag{B7}$$

Integrations over $\boldsymbol{\alpha}$ can be immediately written down by making use of (A12). So the right-hand side of (B7) becomes

$$\sigma_4\int d\boldsymbol{\beta}\,\delta\left(\sum\beta_k-1\right)\theta(\beta_1)\theta(\beta_2)\theta(\beta_3)\times\frac{\partial}{\partial\beta_2}[D(\sigma_1,\sigma_2,\sigma_4,\boldsymbol{\beta})E_2(\sigma_3,\sigma_4\beta_1,\sigma_4\beta_2,\sigma_2+\sigma_4\beta_3)]. \tag{B8}$$

The derivative $\partial/\partial\beta_2$ in (B8) can be transferred to the δ and θ functions to give, for (B8),

$$-\sigma_4\int d\boldsymbol{\beta}\,D(\sigma_1,\sigma_2,\sigma_4,\boldsymbol{\beta})E_2(\sigma_3,\sigma_4\beta_1,\sigma_4\beta_2,\sigma_2+\sigma_4\beta_3)\times\left[\delta'\left(\sum\beta_k-1\right)\theta(\beta_1)\theta(\beta_2)\theta(\beta_3)+\delta\sum\beta_k-1)\theta(\beta_1)\delta(\beta_2)\theta(\beta_3)\right]$$

$$=-\sigma_4\int d\boldsymbol{\beta}\,\delta\left(\sum\beta_k-1\right)\delta(\beta_2)\theta(\beta_1)\theta(\beta_3)\times D(\sigma_1,\sigma_2,\sigma_4,\boldsymbol{\beta})E_2(\sigma_3,\sigma_4\beta_1,\sigma_4\beta_2,\sigma_2+\sigma_4\beta_3) \tag{B9}$$

where we have used the fact that the integrand in the term involving δ' is antisymmetric under exchange of β_1 and β_2 and hence the integral of this term must be zero. The two δ functions in (B9) can be utilized to carry out integrations over β_2 and β_3, and (B9) can be written as

$$-\sigma_4\int_0^1 d\beta_1 \begin{vmatrix} 1 & 1 & 1 \\ \sigma_4\beta_1 & 0 & \sigma_2+\sigma_4(1-\beta_1) \\ e^{\sigma_1\sigma_4\beta_1} & 1 & \exp[\sigma_1(\sigma_2+\sigma_4(1-\beta_1))] \end{vmatrix}$$

$$\times E_2(\sigma_3, \sigma_4\beta_1, 0, \sigma_2+\sigma_4(1-\beta_1)). \tag{B10}$$

Using the notation $\xi=\sigma_4\beta_1$ and $\eta=\sigma_2+\sigma_4(1-\beta_1)$ the integrand of (B10) can be written as[27]

$$\begin{vmatrix} 1 & 1 & 1 \\ \xi & 0 & \eta \\ e^{\sigma_1\xi} & 1 & e^{\sigma_1\eta} \end{vmatrix} E_2(\sigma_3,\xi,0,\eta)$$

$$=\{-\xi[\exp(\sigma_1\eta)-1]+\eta[\exp(\sigma_1\xi)-1]\}E_2(\sigma_3,\xi,0,\eta)$$

$$=\exp(\sigma_1\eta)\xi E_2(\sigma_3,\xi,0,\eta)+\exp(\sigma_1\xi)\eta E_2(\sigma_3,\xi,0,\eta)$$

$$-(\eta-\xi)E_2(\sigma_3,\xi,0,\eta)$$

$$=[\exp(\sigma_1\eta)-1][\exp(\sigma_3\eta)-1]/\eta-[\exp(\sigma_1\xi)-\exp(\sigma_1\eta)]$$

$$\times[\exp(\sigma_3\xi)-\exp(\sigma_3\eta)]/(\xi-\eta)$$

$$+[\exp(\sigma_1\xi)-1][\exp(\sigma_3\xi)-1]/\xi. \tag{B11}$$

Defining $\zeta=(\xi-\eta)/2$ and noting $\xi+\eta=\sigma_2+\sigma_4$ the above expression can be put in the form

$$[\exp(\sigma_1\eta)-1][\exp(\sigma_3\eta)-1]/\eta+[\exp(\sigma_1\xi)-1][\exp(\sigma_3\xi)-1]/\xi$$

$$-\ 4\exp[(\sigma_1+\sigma_3)(\sigma_2+\sigma_4)/2]\sinh\sigma_1\zeta\,\sinh\sigma_3\zeta/\zeta$$

$$=[\exp(\sigma_1\xi)-1][\exp(\sigma_3\xi)-1]/\xi$$

$$+[\exp(\sigma_1\eta-1][\exp(\sigma_3\eta)-1]/\eta$$

$$-2\exp[(\sigma_1+\sigma_3)(\sigma_2+\sigma_4)/2][\cosh(\sigma_1+\sigma_3)\zeta$$

$$-\cosh(\sigma_1-\sigma_3)\zeta]/\zeta. \tag{B12}$$

Each of the three terms in the above expression can be integrated by choosing ξ,η,ζ as integration variables and expanding the integrand in a power series.

Other integrals, T_5—T_9 are evaluated by essentially similar procedure as the above after making use of the steps mentioned below.

(1) An integral involving T_5 contains $\sum_{j<k}R_{jk}S_{jk}$, which can be written as

$$\sum_{j<k}R_{jk}S_{jk}=\sum_{j<k}\frac{\exp(\sigma_3 z_j)-\exp(\sigma_3 z_k)}{z_j-z_k}$$

$$\times\frac{\exp(\sigma_4 z_j^*)-\exp(\sigma_4 z_k^*)}{z_j^*-z_k^*}. \tag{B13}$$

Next we use Eq. (A11) to substitute

$$R_{jk}=\sigma_3\int d\alpha\,\delta(\textstyle\sum\alpha_i-1)\theta(\alpha_j)\theta(\alpha_k)\delta(\alpha_i)\exp(\sigma_3\boldsymbol{\alpha}\cdot\mathbf{z}),\quad i\neq j,k \tag{B14}$$

and use similar expressions for S_{jk}. The rest of the steps are similar to those for T_4. The integral of T_6 is exactly similar.

(2) For T_7 we substitute for P_{jk} and Q_{jk} from Eqs. (V.15) and (A8) to get

$$\sum_{j<k}P_{jk}Q_{jk}$$

$$=\sum_{j<k}[E_1(\sigma_1,\mathbf{z})+(z_j+z_k)E_2(\sigma_1,\mathbf{z})]$$

$$\times[E_1(\sigma_2,\mathbf{z}^*)+(z_j^*+z_k^*)E_2(\sigma_2,\mathbf{z}^*)]$$

$$=3E_1(\sigma_1,\mathbf{z})E_1(\sigma_2,\mathbf{z}^*)$$

$$+2(\textstyle\sum z_k)E_2(\sigma_1,\mathbf{z})E_1(\sigma_2,\mathbf{z}^*)$$

$$+2(\textstyle\sum z_k^*)E_2(\sigma_2,\mathbf{z}^*)E_1(\sigma_1,\mathbf{z})$$

$$+2\left[\sum_k|z_k|^2+\sum_{j<k}(z_jz_k^*+z_j^*z_k)\right]E_2(\sigma_1,\mathbf{z})E_2(\sigma_2,\mathbf{z}^*). \tag{B15}$$

Next we use (A6) and (A7) for products of E's. The expression $\sum_{j<k}R_{jk}S_{jk}$ is simplified in the same way as outlined for T_5 above.

The evaluation of an integral of T_8 is exactly parallel to T_7.

(3) Finally, for T_9 we use (A8)

$$T_9=\frac{1}{(z_1-z_2)}(P_{13}R_{23}-P_{23}R_{13})\frac{1}{(z_1^*-z_2^*)}(Q_{13}S_{23}-Q_{23}S_{13})$$

$$=[E_1(\sigma_1,\mathbf{z})E_2(\sigma_3,\mathbf{z}^*)+\sigma_1\longleftrightarrow\sigma_3]$$

$$\times[E_1(\sigma_2,\mathbf{z})E_2(\sigma_4,\mathbf{z}^*)+\sigma_2\longleftrightarrow\sigma_4]$$

$$=E_1(\sigma_1,\mathbf{z})E_2(\sigma_3,\mathbf{z})E_1(\sigma_2,\mathbf{z}^*)E_2(\sigma_4,\mathbf{z}^*)+\cdots \tag{B16}$$

where dots stand for other terms obtained by interchanges $\sigma_1\longleftrightarrow\sigma_3$ and $\sigma_2\longleftrightarrow\sigma_4$. We now use Eq. (A6) for E_1 and Eq. (A12) for E_2 and retrace the steps of the evaluation of T_4.

[1]S. Okubo, Progr. Theor. Phys. (Kyoto) **11**, 80 (1954); R. Arnowitt and S. Deser, Phys. Rev. **100**, 349 (1955); E.S. Fradkin, Nucl. Phys. **49**, 624 (1963); G.V. Effimov, Zh. Eksp. Teor. Fiz. **44**, 2107—17 (1963) [Sov. Phys. JETP **17**, 1417 (1963)]; G. Feinberg and A. Pais, Phys. Rev. **131**, 2724 (1963); W. Guttinger, Fortschr. Phys. **14**, 483 (1966); H.M. Fried, Nuovo Cimento A **52**, 1333 (1967); M.K. Volkov, Ann. Phys. (N.Y.) 49, 202 (1968). For a recent work on infinity suppression in gravity modified Yang—Mills theories see, M.S. Sri Ram, "Infinity Suppression in Gravity Modified Yang—Mills Theories," I.I.T., Kanpur, preprint. For a study of nonlinear conformal invariant QED see, Tulsi Dass and Radhey Shyam, "Conformal Invariant Finite Quantum Electrodynamics," I.I.T., Kanpur, preprint.

[2]For a review of nonpolynomial Lagrangian theories, see *Proceedings of Coral Gables Conference on Fundamental Interactions at High Energies* 1971, edited by A. Salam (Gordon and Breach, New York, 1971).

[3]S. Gasiorowicz and D.A. Geffen, Rev. Mod. Phys. **41**, 531 (1969).

[4]A. Hunt, K. Koller, and Q. Shafi, Phys. Rev. D **3**, 1327 (1971); R. Delbourgo, J. Math. Phys. **13**, 464 (1972); P.T. Davies, J. Phys. A: Gen. Phys. **5**, 1479, 1698 (1972).

[5] We use script symbols to denote matrix functions. Ordinary symbols denote scalar functions with no free indices.

[6] J. Ashmore and R. Delbourgo, J. Math. Phys. **14**, 176, 569 (1973).

[7] A. K. Kapoor, J. Math. Phys. **17**, 61 (1976).

[8] J. Ginibre, J. Math. Phys. **6**, 440 (1965).

[9] M. L. Mehta, *Random Matrices* (Academic, New York, 1972).

[10] Curly brackets here, and, throughout the paper, denote the trace over internal indices.

[11] We assume that components of Φ are linearly independent. If they satisfy some condition, for example $\mathrm{Tr}\Phi=0$ or $\Phi_{ji}=\Phi_{ij}$, some of the considerations given here will require modification.

[12] R. Delbourgo, A. Salam, and J. Strathdee, Phys. Rev. **187**, 1999 (1969).

[13] Throughhout this paper we assume normal ordering for fields at the same point.

[14] In Eq. (5) of Ref. 7, U^* should be replaced by $U^\dagger$ [compare with Eq. (II. 5) here]. Similarly $u^*_{\alpha\beta}$ in Eq. (4) and Z^* in Eqs. (7) and (8) of Ref. 7 should be replaced by $u^*_{\beta\alpha}$ and $Z^\dagger$, respectively.

[15] The symbol $\hat{=}$ indicates that an over all normalization factor has been omitted on the right-hand side.

[16] For evaluation of Jacobians of matrix transformations see, W. L. Deemer and I. Olkin, Biometrika 38, 345 (1951). For the transformations considered in this paper all details can be found in Ref. 8 and Appendix A. 24 of Ref. 9.

[17] I_{n-1} denotes an $(n-1)\times(n-1)$ unit matrix.

[18] If, for example, $\mathcal{A}(Z)=\exp(Z)$, then $a_k=\exp(z_k)$.

[19] If g is an N component column, A, B, Q are $N\times N$ matrices independent of g, and Q is positive definite Hermitian, then the following results hold:

$$\int dg\exp(-g^\dagger Qg)=\frac{\pi^N}{\det Q},$$

$$\int dg\exp(-g^\dagger Qg)g^\dagger Ag=\frac{\pi^N}{\det Q}\mathrm{Tr}(AQ^{-1}),$$

and

$$\int dg\exp(-g^\dagger Qg)(g^\dagger Ag)(g^\dagger Bg)$$

$$=\frac{\pi^N}{\det Q}[\mathrm{Tr}(AQ^{-1})\mathrm{Tr}(BQ^{-1})+\mathrm{Tr}(AQ^{-1}BQ^{-1})],$$

where $dg=\Pi_{k=1}^N dg_k$ and the integrals are over all complex values for each g_k. These results are easily proved by first changing variables to $h=Q^{1/2}g$ and then reducing the resulting integrals to Gaussian integrations. For example, the third integral above becomes, on a change of variables,

$$(\det Q)^{-1}(Q^{-1/2}AQ^{-1/2})_{ij}(Q^{-1/2}BQ^{-1/2})_{kl}\int dh\exp(-h^\dagger h)h_i^*h_jh_k^*h_l.$$

The most general form for the integral over h is given by

$$\int dh\exp(-h^\dagger h)h_i^*h_jh_k^*h_l=C(\delta_{ij}\delta_{kl}+\delta_{il}\delta_{jk}).$$

Contracting the indices we get

$$\begin{aligned}CN(N+1)&=\int dh\exp(-h^\dagger h)(h^\dagger h)^2\\&=\int(\Pi_k dh_k)\exp(-\Sigma_k|h_k|^2)(\Sigma|h_k|^2)^2\\&=\int(\Pi_k dh_k)\exp(-\Sigma_k|h_k|^2)\\&\quad\times[\Sigma|h_k|^4+2\Sigma_{j<k}|h_j|^2|h_k|^2]\\&=\pi^N[2N+2N(N-1)/2].\end{aligned}$$

Hence $C=\pi^N$. Any other integral involving $\exp(-g^\dagger g)$ and any number of factors of g and g^* can be evaluated in a similar fashion.

[20] The expressions such as $[\mathcal{A}(Z')-a_n]/(Z'-z_n)$ stand for a $(n-1)\times(n-1)$ diagonal matrix with kth element, on the diagonal, given by $(a_k-a_n)/(z_k-z_n)$.

[21] This follows immediately from invariance of the two point function (I. 3) under SU(n) transformations $\Phi\to V\Phi V^{-1}$, where V is any $n\times n$ unitary matrix. The arguments will fail if all components of Φ do not have the same mass.

[22] The constant C is fixed by demanding that the fields $P(x)$ have the correct kinetic energy term.

[23] See, for example, J. A. Cronin, Phys. Rev. **161**, 1483 (1967).

[24] See, for example, H. M. Fried, *Functional Methods and Models in Quantum Field Theory* (M.I.T. Press, Cambridge, Mass., 1972), Appendix A.

[25] Equation (IV. 7) is obtained by taking $\xi=P(y)_{ji}$ in Eq. (IV. 5).

[26] Here $t'=1-t$. In the rest of the paper we use this notation for integration variables like t, s, u, etc.

[27] In arriving at (B11) we have used the following three alternative expressions for E_2:

$$E_2(\sigma,\mathbf{z})=\frac{1}{z_1-z_2}\left[\frac{e^{\sigma z_1}-e^{\sigma z_3}}{z_1-z_3}-\frac{e^{\sigma z_2}-e^{\sigma z_3}}{z_2-z_3}\right],$$

$$E_2(\sigma,\mathbf{z})=\frac{1}{z_2-z_3}\left[\frac{e^{\sigma z_2}-e^{\sigma z_1}}{z_2-z_1}-\frac{e^{\sigma z_3}-e^{\sigma z_1}}{z_3-z_1}\right],$$

and

$$E_2(\sigma,\mathbf{z})=\frac{1}{z_3-z_1}\left[\frac{e^{\sigma z_3}-e^{\sigma z_2}}{z_3-z_2}-\frac{e^{\sigma z_1}-e^{\sigma z_2}}{z_1-z_2}\right],$$

which follow from (A9) and the fact that $E_k(\sigma,\mathbf{z})$ are completely symmetric in z_1, z_2, and z_3.

Existence, uniqueness and properties of the solutions of the Boltzmann kinetic equation for a weakly ionized gas. I

Frédéric A. Molinet*

Université de Paris (VII), Laboratoire de Physique Théorique et Mathématique, Tour 33/43, 2 Place Jussieu 75 221, Paris, France
(Received 1 December 1975)

The Boltzmann kinetic equation for a weakly ionized gas in the presence of a time dependent exterior electric field and a static exterior magnetic field has been transformed into an integral equation. Existence and uniqueness theorems have been proved for inverse power-law potentials of the form A/r^s with $s>3$ and for a large class of initial distribution functions. For soft potentials ($3<s\leq 5$), these theorems have been derived from the general properties of the integral operator. For hard potentials, $5<s\leq+\infty$, where no general properties of the integral operator can be directly proved, an iteration procedure which constitutes the main part of the present work has been developed. In each case, some important properties of the solution have been established.

I. INTRODUCTION

The theory of ionized gases in the presence of electric and magnetic exterior fields,[1] leads to a system of integrodifferential equations describing the evolution of the distribution function of each type of particles. Generally, these equations which are similar to the usual Boltzmann equation in kinetic theory, are not independent. However, when the ionization is weak,[2] it is possible to neglect in each equation the terms involving the collisions between charged particles. It remains then, for each charged component a reduced equation, the general form of which is given by

$$\frac{\partial f}{\partial t}+\mathbf{v}\cdot\nabla_{\mathbf{r}}f+(\mathbf{\Gamma}+\mathbf{v}\times\mathbf{\Omega})\cdot\nabla_{\mathbf{v}}f=\mathcal{J}(f,f_n),\qquad \mathbf{\Gamma}=\frac{q\mathbf{E}}{m},\quad \mathbf{\Omega}=\frac{q\mathbf{B}}{m},\tag{1.1}$$

where $f(\mathbf{r},\mathbf{v},t)$, q, and m are, respectively, the distribution function, the elementary charge, and the elementary mass of the type of charged particles considered, and where $f_n(\mathbf{r},\mathbf{v}_n,t)$ is the distribution function of the neutral particles. The vectors $\mathbf{E}$ and $\mathbf{B}$ denote, respectively, the exterior electric and magnetic fields.

The present work is devoted to the resolution of Eq. (1.1) under the following assumptions:

(i) the magnetic field $\mathbf{B}$ is uniform and independent of time,

(ii) the electric field $\mathbf{E}(t)$ depends only on the time t,

(iii) the velocity distribution function of the neutral particles is a Maxwellian distribution function given by

$$f_n(\mathbf{r},\mathbf{v}_n,t)=f_n(v_n)=a\exp(-\beta_n v_n^2),\qquad a=N\left(\frac{M}{2\pi kT}\right)^{3/2},\quad \beta_n=\frac{M}{2kT},\tag{1.2}$$

where N, M, T are, respectively, the density, the elementary mass, and the temperature of the neutral particles, and where k is the Boltzmann constant.

The two-body interaction between a charged and a neutral particle is described by inverse power force laws A/r^s, where $3<s\leq+\infty$, the limiting case $s=+\infty$ corresponding to the interaction of rigid spheres.

Moreover, for power-law potentials with $s<+\infty$, an angular cutoff has been introduced which excludes the grazing collisions. This mathematical trick, which was first used by Grad,[3] allows us to split the collision operator K, defined by $Kf=\mathcal{J}(f,f_n)$, into two parts and write

$$Kf=-\nu f+Hf,\tag{1.3}$$

where $f\to\nu f$ is a multiplication operator by a function and where H is an integral operator the kernel of which is defined over $\Delta\times\Delta$, where Δ is the whole three-dimensional velocity space. The Eq. (1.1) reduces then to

$$\frac{\partial f}{\partial t}+(\mathbf{\Gamma}+\mathbf{v}\times\mathbf{\Omega})\cdot\frac{\partial f}{\partial\mathbf{v}}=-\nu f+Hf.\tag{1.4}$$

This last equation differs, however basically, from the linearized Boltzmann equation in the kinetic theory of neutral gases. Indeed, the integral operator defined by the collision operator of the latter one, which has been studied by Hilbert[4] and Hecke[5] in the case of rigid spheres and by Grad[3,6] and Cercignani[7] in the general case of power-law potentials with angular cutoff, is completely continuous in the Lebesgue function space $L_2(\Delta)$, whereas for hard potentials ($5<s\leq+\infty$) the integral operator H of Eq. (1.4) verifies much weaker properties.

In Part I of the present work, we investigated separately the cases $3<s\leq 5$ (soft potentials) and $5<s\leq+\infty$ (hard potentials).

After a transformation of Eq. (1.4) into an integral equation, we establish in Sec. II the general properties of the corresponding integral operator. For soft potentials, these properties allow us to prove in Sec. III, the existence and uniqueness of the solution of equation (1.4) for a wide class of initial conditions.

On the other hand, for hard potentials, where no general properties of the integral operator can be directly proved, an iteration procedure similar to that given in Ref. 8, has been developed. The proof of the convergence of this procedure which is based on a careful construction of majorants, constitutes the main part of the present work. The details of our method are given in Sec. IV for the particular case of rigid spheres. In this case, the basic theorems on the existence and uniuniqueness of the solution of the integral equation form of the Boltzmann equation (1.4) as well as some properties of this solution in connection with those of the resolvent kernel, have been established.

The general case of hard potentials ($5 < s < +\infty$) together with other important topics such as the proof of the existence of the moments of the solution, will be treated in Part II.

II. INTEGRAL FORM OF THE BOLTZMANN EQUATION AND PROPERTIES OF THE KERNEL

A. Integral form of the Boltzmann equation

In order to set up solution procedures for a wide class of initial conditions as well as for general laws of interaction between electrons and neutral particles or ions, it is useful to transform the Boltzmann equation from an integrodifferential to a purely integral form. The simplest procedure to achieve this transformation is to consider Eq. (1.4) and integrate both sides along the characteristics of the differential operator D,

$$D=\frac{\partial}{\partial t}+(\mathbf{\Gamma}+\mathbf{v}\times\mathbf{\Omega})\cdot\frac{\partial}{\partial \mathbf{v}}, \tag{2.1}$$

while taking into account the proper boundary conditions.

The equations of the characteristics of (2.1) are given by the general solution of the differential vector equation,

$$\frac{d\mathbf{v}}{dt}=\mathbf{\Gamma}(t)+\mathbf{v}\times\mathbf{\Omega}. \tag{2.2}$$

One particular solution of (2.2) is given by

$$\mathbf{v}_1(t)=\int_{t_0}^{t}\mathcal{R}_{\Omega(t-t')}\mathbf{\Gamma}(t')\,dt', \tag{2.3}$$

where $\mathcal{R}_{\Omega(t-t')}$ is the rotation operator of angle $\Omega(t-t')$ around the vector $\mathbf{\Omega}$. The general solution of (2.2) is therefore

$$\mathbf{v}=\mathcal{R}_{\Omega(t-t_0)}\mathbf{u}+\int_{t_0}^{t}\mathcal{R}_{\Omega(t-t')}\mathbf{\Gamma}(t')\,dt', \tag{2.4}$$

where $\mathbf{u}$ is the value of $\mathbf{v}$ at $t=t_0$.

The transformation $S_{t,t_0}:\mathbf{u}\to\mathbf{v}=S_{t,t_0}\mathbf{u}$ defined by (2.4) has an inverse. Indeed by applying the operator $\mathcal{R}_{\Omega(t_0-t)}$ to both terms of (2.4), we obtain

$$\mathbf{u}=\mathcal{R}_{\Omega(t_0-t)}\mathbf{v}-\int_{t_0}^{t}\mathcal{R}_{\Omega}(t_0-t')\mathbf{\Gamma}(t')\,dt'\equiv S_{t_0,t}\mathbf{v} \tag{2.5}$$

and hence $S_{t,t_0}^{-1}=S_{t_0,t}$.

By performing the variable change defined by (2.5), the Boltzmann equation can be written in the form

$$\frac{\partial}{\partial t}f(S_{t,t_0}\mathbf{u},t)=-\nu(|S_{t,t_0}\mathbf{u}|)f(S_{t,t_0}\mathbf{u},t)+\mathcal{S}(S_{t,t_0}\mathbf{u},t), \tag{2.6}$$

where we have put

$$[Hf](\mathbf{v},t)=\mathcal{S}(\mathbf{v},t). \tag{2.7}$$

The integration of (2.6) with respect to t yields

$$f(S_{t,t_0}\mathbf{u},t)=f(\mathbf{u},t_0)\exp\left[-\int_{t_0}^{t}\nu(|S_{t',t_0}\mathbf{u}|)\,dt'\right]+\int_{t_0}^{t}\mathcal{S}(S_{t',t_0}\mathbf{u},t')\exp\left[-\int_{t'}^{t}\nu(|S_{t'',t_0}\mathbf{u}|)dt''\right]dt'. \tag{2.8}$$

By coming back, to (2.8), to the variables $\mathbf{v}$ and t and according to the definition of $\mathcal{S}(\mathbf{v},t)$, we finally obtain the following linear integral equation verified by $f(\mathbf{v},t)$:

$$f(\mathbf{v},t)=f(S_{t_0,t}\mathbf{v},t_0)\exp\left[-\int_{t_0}^{t}\nu(|S_{t',t}\mathbf{v}|)\,dt'\right]+\int_{t_0}^{t}\int_{\Delta}k(\mathbf{v},t,\mathbf{v}',t')f(\mathbf{v}',t')\,d\mathbf{v}'\,dt', \tag{2.9}$$

where Δ is the three-dimensional velocity space and where the kernel $k(\mathbf{v},t,\mathbf{v}',t')$ is given by

$$k(\mathbf{v},t,\mathbf{v}',t')=H(S_{t',t}\mathbf{v},\mathbf{v}')\exp\left[-\int_{t'}^{t}\nu(|S_{t'',t}\mathbf{v}|)\,dt''\right]. \tag{2.10}$$

Conversely, now let $F(\mathbf{v},t)$ be a solution of (2.9). According to the properties of the transformation defined by (2.5), $F(\mathbf{v},t)$ also verifies Eq. (2.8). This last equation may be written in the form

$$F(S_{t,t_0}\mathbf{u},t)=F(\mathbf{u},t_0)\exp\left[-\int_{t_0}^{t}\nu(|S_{t',t_0}\mathbf{u}|)\,dt'\right]+\exp\left[-\int_{t_0}^{t}\nu(|S_{t'',t_0}\mathbf{u}|)\,dt''\right]\int_{t_0}^{t}G(\mathbf{u},t')\,dt', \tag{2.11}$$

where $G(\mathbf{u},t')$ is given by

$$G(\mathbf{u},t')=\exp\left[-\int_{t'}^{t_0}\nu(|S_{t'',t_0}\mathbf{u}|)\,dt''\right]\mathcal{S}(S_{t',t_0}\mathbf{u},t'). \tag{2.12}$$

According to a theorem due to Lebesgue, the derivative of the function $t\to\int_{t_0}^{t}g(t')\,dt'$, where $g(t')$ is a locally integrable function is equal to $g(t)$ almost everywhere. Hence, the time derivative of $F(S_{t,t_0}\mathbf{u},t)$ exists almost everywhere and is given by

$$\frac{\partial}{\partial t}F(S_{t,t_0}\mathbf{u},t)=-\nu(|S_{t,t_0}\mathbf{u}|)F(S_{t,t_0}\mathbf{u},t)+\mathcal{S}(S_{t,t_0}\mathbf{u},t). \tag{2.13}$$

In view of (2.4) and (2.2), Eq. (2.13) may also be written in the form

$$\frac{\partial}{\partial t}F(\mathbf{v},t)+(\mathbf{\Gamma}+\mathbf{v}\times\mathbf{\Omega})\cdot\frac{\partial F}{\partial\mathbf{v}}(\mathbf{v},t)=-\nu(v)F(\mathbf{v},t)+\mathcal{S}(\mathbf{v},t). \tag{2.14}$$

This last equation is just the original Boltzmann equation. It must be noted however that this result does not imply that the partial derivatives of F on the left-hand side of (2.14) exist and are finite separately. Only the sum of the two terms on the left-hand side of (2.14) exists almost everywhere. The original integrodifferential equation is therefore satisfied in a generalized sense.

In order to set up solution procedures for the integral equation (2.9), we will study in the next section, the basic properties of the kernel of this equation.

B. General properties of the kernel $k(\mathbf{v}, t, \mathbf{v}', t')$

The properties of the function $\nu(v)$ and of the kernel $H(\mathbf{v}, \mathbf{v}')$ which have been established in Appendix B imply related properties of the kernel $k(\mathbf{v}, t, \mathbf{v}', t')$. In the Lebesgue function spaces $L_p(\Delta)$, where Δ is the whole velocity space, a highly interesting property is given by the following theorem.

Theorem 2.1: For soft potentials $(3 < s \leq 5)$, the linear operator $K_{t,t'}$, generated by the kernel $k(\mathbf{v}, t, \mathbf{v}', t')$, $(t_0 \leq t' \leq t \leq t_1)$ is uniformly bounded on the function spaces $L_p(\Delta)$, $1 \leq p \leq +\infty$ and $\mathcal{B}_{\mathcal{R}}(\Delta)$. In order to prove this theorem, we consider the inequalities (B29) and (B37) of Appendix B. Using the fact that $\nu(v)$ is a positive function, we obtain

$$\int_\Delta k(\mathbf{v}, t, \mathbf{v}', t')\, d\mathbf{v} \leq \nu(0), \tag{2.15}$$

$$\int_\Delta k(\mathbf{v}, t, \mathbf{v}', t')\, d\mathbf{v}' < C_2 \nu(0), \tag{2.16}$$

where C_2 is a positive constant.

Let $\tilde{f} \in L_p(\Delta)$. For $1 \leq p \leq +\infty$, we have

$$\|K_{t,t'}\tilde{f}\|_p = \left(\int_\Delta d\mathbf{v} \left| \int_\Delta k(\mathbf{v}, t, \mathbf{v}', t') f(\mathbf{v}')\, d\mathbf{v}' \right|^p\right)^{1/p}.$$

Making use of the identity

$$k(\mathbf{v}, t, \mathbf{v}', t') f(\mathbf{v}) = [k(\mathbf{v}, t, \mathbf{v}', t')]^{(p-1)/p} \{[k(\mathbf{v}, t, \mathbf{v}', t')]^{1/p} f(\mathbf{v})\}$$

and of the fact that $k(\cdot)$ is a positive function, then, by Hölder's inequality, we see that

$$\begin{aligned} |K_{t,t'}\tilde{f}|^p &= \left| \int_\Delta k(\mathbf{v}, t, \mathbf{v}', t') f(\mathbf{v}')\, d\mathbf{v}' \right|^p \\ &\leq \left(\int_\Delta k(\mathbf{v}, t, \mathbf{v}', t')\, d\mathbf{v}'\right)^{p-1} \\ &\quad \times \int_\Delta k(\mathbf{v}, t, \mathbf{v}', t') |f(\mathbf{v}')|^p\, d\mathbf{v}'. \end{aligned}$$

Inserting the preceding inequality into Eq. (2.15), we obtain

$$|K_{t,t'}\tilde{f}|^p < [C_2 \nu(0)]^{p-1} \int_\Delta k(\mathbf{v}, t, \mathbf{v}', t') |f(\mathbf{v}')|^p\, d\mathbf{v}'$$

and finally, according to Fubini's theorem, we can write

$$\|K_{t,t'}\tilde{f}\|_p < C_2^{(p-1)/p} \nu(0) \|\tilde{f}\|_p. \tag{2.17}$$

Similarly, according to the definition of the norm in $L_\infty(\Delta)$, we obtain for $p = +\infty$,

$$\|K_{t,t'}\tilde{f}\|_\infty < C_2 \nu(0) \|\tilde{f}\|_\infty. \tag{2.18}$$

Inequalities (2.17) and (2.18) hold for all values of t and t' belonging to a finite interval $[t_0, t_1]$. Thus, since C_2 and $\nu(0)$ are constants, independent of t and t', it results that the linear operator $K_{t,t'}$ is uniformly bounded on the function spaces $L_p(\Delta)$, $1 \leq p \leq +\infty$.

If the function f belongs to the set $\mathcal{B}_{\mathcal{R}}(\Delta)$ of all bounded transformations on $\mathcal{R}$ to Δ, then, by (2.16) and the definition of the norm in $\mathcal{B}_{\mathcal{R}}(\Delta)$, we have

$$\begin{aligned} \|K_{t,t'}f\| &= \sup_{\mathbf{v} \in \Delta} \left| \int_\Delta k(\mathbf{v}, t, \mathbf{v}', t') f(\mathbf{v}')\, d\mathbf{v}' \right| \\ &\leq \|f\| \sup_{\mathbf{v} \in \Delta} \int_\Delta k(\mathbf{v}, t, \mathbf{v}', t')\, d\mathbf{v}' < C_2 \nu(0) \|f\|. \end{aligned} \tag{2.19}$$

Inequality (2.19) completes the proof of Theorem 2.1.

For hard potentials $(5 < s \leq +\infty)$, $\nu(v)$ is a monotonically increasing function of v. Consequently, Eqs. (B29) and (B37) as well as Eqs. (2.15) and (2.16) no longer hold. The procedure used above to prove Theorem 2.1 can therefore not be applied to the present case. Moreover, for $s = +\infty$ (rigid spheres), by using Eqs. (B14) and (B19), it is easy to disprove that $K_{t,t'}$ is uniformly bounded on $L_p(\Delta)$ $(1 \leq p \leq +\infty)$. Nevertheless, as it will be seen in Sec. IV, the resolvent kernel of the kernel $k(\mathbf{v}, t, \mathbf{v}', t')$, verify, in the case of hard potentials, some highly interesting properties, the proof of which constitutes the main part of the present work.

III. METHODS OF SOLUTION OF THE INTEGRAL EQUATION FOR SOFT POTENTIALS $(3 < s \leq 5)$

A. Existence and uniqueness of the solution

Let $\mathcal{B}_F(A)$ be the set at all bounded transformations on $A = [t_0, t_1]$ to $F = L_p(\Delta)$, $1 \leq p \leq +\infty$ and let $\|\tilde{f}\| = \sup_{t \in [t_0, t_1]} \|\tilde{f}_t\|_p$ be the norm defined on $\mathcal{B}_F(A)$ which is a Banach space. Furthermore, let $f_{t_0} : \mathbf{v} \to f_{t_0}(\mathbf{v}) = f(\mathbf{v}, t_0)$ be the initial distribution function (at $t = t_0$). We now present a proof of the following:

Theorem 3.1: If the class $\tilde{f}_{t_0}$ defined by f_{t_0} belongs to the Lebesgue function space $L_p(\Delta)$, $1 \leq p \leq +\infty$, the integral equation form of the Boltzmann equation, Eq. (2.9), has a unique solution in the function space $\mathcal{B}_F(A)$, $A = [t_0, t_1]$.

Proof: Applying the method of successive approximations to the integral equation (2.9), we obtain the following sequence:

$$h_n(\mathbf{v}, t) = \varphi(\mathbf{v}, t) + \int_{t_0}^t \int_\Delta k(\mathbf{v}, t, \mathbf{v}', t') h_{n-1}(\mathbf{v}', t')\, d\mathbf{v}'\, dt',$$

$$h_0(\mathbf{v}, t) = 0, \tag{3.1}$$

where

$$\varphi(\mathbf{v}, t) = f(S_{t_0,t}\mathbf{v}, t_0) \exp\left[-\int_{t_0}^t \nu(|S_{t',t}\mathbf{v}|)\, dt'\right]. \tag{3.2}$$

We will first prove that

$$\|\tilde{h}_{t,n+1} - \tilde{h}_{t,n}\|_p \leq (N_p/2^n) \exp[2\alpha(t - t_0)], \tag{3.3}$$

where $N_p = \|\tilde{f}_{t_0}\|_p$ and where α is a positive constant.

In view of (3.2), we have $\|\tilde{\varphi}\|_p \leq \|\tilde{f}_{t_0}\|_p = N_p$. Thus, inequality (3.3) holds for $n = 0$. We will assume now that (3.3) is satisfied at the order $n - 1$. Then, according to Eq. (3.1) and to the theorem of Sec. II.B, we have

$$\begin{aligned} \|\tilde{h}_{t,n+1} - \tilde{h}_{t,n}\|_p &\leq \int_{t_0}^t dt' \|K_{t,t'}(\tilde{h}_{t,n} - \tilde{h}_{t,n-1})\|_p \\ &\leq \int_{t_0}^t \alpha \|\tilde{h}_{t,n} - \tilde{h}_{t,n-1}\|_p\, dt' \\ &\leq (N_p/2^n) \exp[2\alpha(t - t_0)], \end{aligned}$$

and therefore inequality (3.3) holds.

If we denote the transformation $t \to \tilde{h}_{t,n}$ by $\tilde{h}_n$, it follows from the inequality (3.3) and from the definition of the norm on the Banach space $\mathcal{B}_F(A)$ that the series $(\|\tilde{h}_{n+1} - \tilde{h}_n\|)$ are convergent. Hence, the sequence $\tilde{h}_n$ converges on $\mathcal{B}_F(A)$ to a limit $\tilde{f} \in \mathcal{B}_F(A)$ which is a solution of Eq. (2.9).

Let us suppose now that $\tilde{f}$ and $\tilde{f}'$ are two different solutions of Eq. (2.9), belonging to $\mathcal{B}_F(A)$. Then, if we apply the method of successive approximations to the homogeneous integral equation verified by the difference $\tilde{g}=\tilde{f}-\tilde{f}'$, we obtain

$$\|\tilde{g}_{t,n+1}\|_p \leq \int_{t_0}^{t}\|K_{t,t'}\tilde{g}_{t,n}\|_p\,dt' \leq \alpha\int_{t_0}^{t}\|\tilde{g}_{t,n}\|_p\,dt'$$

and, by repeating the same process,

$$\|\tilde{g}_{t,n+1}\|_p \leq \frac{\alpha^n}{(n-1)!}\int_{t_0}^{t}(t-\theta)^{n-1}\|\tilde{g}_{t,1}\|_p\,d\theta.$$

Furthermore, $\|\tilde{g}_{t,1}\|_p=\|\tilde{g}_t\|_p \leq \|\tilde{g}\| < +\infty$. Hence

$$\|\tilde{g}_{n+1}\| \leq \alpha^n\|g\|(t_1-t_0)^n/n!,$$

which implies that $\lim\|\tilde{g}_{n+1}\|=0$ as $n\to\infty$ and therefore $\tilde{f}=\tilde{f}'$.

Corollary: If the function f_{t_0} belongs to $G=\mathcal{B}_{\mathcal{R}}(\Delta)$, the integral equation form of the Boltzmann equation has a unique solution in the function space $\mathcal{B}_G(A)$, $A=[t_0,t_1]$.

Proof: Since $f_{t_0}\in\mathcal{B}_{\mathcal{R}}(\Delta)$ and $K_{t,t'}$ is uniformly bounded on $\mathcal{B}_{\mathcal{R}}(\Delta)$ (See Sec. II. B) we get in a way similar to the above that

$$\|h_{t,n+1}-h_{t,n}\| \leq N/2^n \exp[\alpha(t-t_0)], \tag{3.4}$$

where $N=\|f_{t_0}\|$ is the norm of f_{t_0} in $\mathcal{B}_{\mathcal{R}}(\Delta)$. Hence, if we denote the transformation $t\to h_{t,n}$ by h_n, it follows from (3.4) that the sequence h_n converges on $\mathcal{B}_G(A)$ to a unique limit f belonging to $\mathcal{B}_G(A)$ and verifying the integral equation (2.9).

B. Properties of the solution

In order to establish some interesting properties of the solution, we start with the proof of two important lemmas.

Let us consider the series of the iterated kernels of $k(\mathbf{v},t,\mathbf{v}',t')$. If we denote the sum of the n first terms of this series by $R^{(n)}(\mathbf{v},t,\mathbf{v}',t')$ we have

$$R^{(n)}(\mathbf{v},t,\mathbf{v}',t')=\sum_{p=1}^{n}k^{(p)}(\mathbf{v},t,\mathbf{v}',t'), \tag{3.5}$$

where $k^{(p)}$ is the iterated kernel of order p.

Lemma 3.1: The sequence $R^{(n)}(\mathbf{v},t,\mathbf{v}',t')$ converges for all $(\mathbf{v},t)\in\Delta\times[t_0,t_1]$ and for almost all $(\mathbf{v}',t')\in\Delta\times[t_0,t_1]$, to a limit $R(\mathbf{v},t,\mathbf{v}',t')$. For all $(\mathbf{v},t)\in\Delta\times[t_0,t_1]$, this limit is integrable on the set of all $(\mathbf{v}',t')\in\Delta\times[t_0,t_1]$. The limit is also integrable on the set of all $\mathbf{v}\in\Delta$ for almost all $(v',t')\in\Delta\times[t_0,t_1]$.

Proof: According to the results of the preceding section, the sequence $h_n(\mathbf{v},t)$ defined by (3.1) is uniformly convergent on Δ, provided that φ belongs to $\mathcal{B}_R(\Delta)$. In addition, by the definition of the iterated kernels, we have

$$h_{n+1}(\mathbf{v},t)=\varphi(\mathbf{v},t)+\int_{t_0}^{t}\int_\Delta R^{(n)}(\mathbf{v},t,\mathbf{v}',t')\varphi(\mathbf{v}',t')\,d\mathbf{v}'\,dt'. \tag{3.6}$$

Thus, since $R^{(n)}>0$, if we choose $\varphi(\mathbf{v},t)=C$, where C is a constant, we find that

$$\lim_{n\to\infty}\int_{t_0}^{t}\int_\Delta R^{(n)}(\mathbf{v},t,\mathbf{v}',t')\,d\mathbf{v}'\,dt' < A_1(t_0,t_1) < +\infty. \tag{3.7}$$

Hence, according to the theorem of Beppo Levi,[9] the ascending sequence $R^{(n)}(\mathbf{v},t,\mathbf{v}',t')$ converges for all $(\mathbf{v},t)\in\Delta\times[t_0,t_1]$ and for almost all $(\mathbf{v}',t')\in\Delta\times[t_0,t_1]$ to a limit $R(\mathbf{v},t,\mathbf{v}',t')$ which is integrable on the set of all $(\mathbf{v}',t')\in\Delta\times[t_0,t_1]$.

In order to prove the last statement of Lemma 3.1, we choose $\varphi(\mathbf{v},t)>0$ and integrable. By Theorem 3.1, the sequence $\|\tilde{h}_{t,n+1}\|_1$ is uniformly convergent on the space $\mathcal{R}$. Hence,

$$\int_\Delta d\mathbf{v}\int_{t_0}^{t}\int_\Delta R^{(n)}(\mathbf{v},t,\mathbf{v}',t')\varphi(\mathbf{v}',t')\,d\mathbf{v}'\,dt' < A_2(t_0,t_1) < +\infty.$$

Furthermore, according to the first part of Lemma 3.1, we have

$$\lim_{n\to\infty}\int_{t_0}^{t}\int_\Delta R^{(n)}(\mathbf{v},t,\mathbf{v}',t')\varphi(\mathbf{v}',t')\,d\mathbf{v}'\,dt' = \int_{t_0}^{t}\int_\Delta R(\mathbf{v},t,\mathbf{v}',t')\varphi(\mathbf{v}',t')\,d\mathbf{v}'\,dt'.$$

Hence,

$$\int_\Delta d\mathbf{v}\int_{t_0}^{t}\int_\Delta R(\mathbf{v},t,\mathbf{v}',t')\varphi(\mathbf{v}',t')\,d\mathbf{v}'\,dt' = \int_{t_0}^{t}\int_\Delta \varphi(\mathbf{v}',t')\,d\mathbf{v}'\,dt'\int_\Delta R(\mathbf{v},t,\mathbf{v}',t')\,d\mathbf{v} < A_2(t_0,t_1).$$

The last inequality implies that

$$\int_\Delta R(\mathbf{v},t,\mathbf{v}',t')\,d\mathbf{v} < A_3(t_0,t_1) < +\infty$$

for almost all $(\mathbf{v}',t')\in\Delta\times[t_0,t_1]$.

Lemma 3.2: For almost all $(\mathbf{v}',t')\in\Delta\times[t_0,t_1]$, the resolvent kernel $R(\mathbf{v},t,\mathbf{v}',t')$ verifies

$$\int_\Delta R(\mathbf{v},t,\mathbf{v}',t')\,d\mathbf{v}=\nu(v').$$

Proof: By Lemma 3.1, the following equality holds for almost all $(\mathbf{u},\theta)\in\Delta\times[t_0,t_1]$,

$$R(\mathbf{v},t,\mathbf{u},\theta)=\sum_{p\geq 1}k^{(p)}(\mathbf{v},t,\mathbf{u},\theta). \tag{3.8}$$

Multiplying both sides of (3.8) by $k(\mathbf{u},\theta,\mathbf{v}',t')$ and integrating with respect to $(\mathbf{u},\theta)$, we obtain

$$\int_{t'}^{t}\int_\Delta R(\mathbf{v},t,\mathbf{u},\theta)k(\mathbf{u},\theta,\mathbf{v}',t')\,d\mathbf{u}\,d\theta = \sum_{p\geq 1}\int_{t'}^{t}\int_\Delta k^{(p)}(\mathbf{v},t,\mathbf{u},\theta)k(\mathbf{u},\theta,\mathbf{v}',t')\,d\mathbf{u}\,d\theta. \tag{3.9}$$

Furthermore, according to the definition of the iterated kernels, we can write

$$\begin{aligned}\int_{t'}^{t}\int_\Delta k^{(p)}(\mathbf{v},t,\mathbf{u},\theta)k(\mathbf{u},\theta,\mathbf{v}',t')\,d\mathbf{u}\,d\theta &= \int_{t'}^{t}\int_\Delta k(\mathbf{v},t,\mathbf{u},\theta)k^{(p)}(\mathbf{u},\theta,\mathbf{v}',t')\,d\mathbf{u}\,d\theta\\ &= k^{(p+1)}(\mathbf{v},t,\mathbf{v}',t').\end{aligned} \tag{3.10}$$

Substituting (3.10) into (3.9) and using (3.8), we obtain the following integral equation, verified by the resolvent kernel for almost all $(v', t') \in \Delta \times [t_0, t_1]$:

$$R(\mathbf{v}, t, \mathbf{v}', t')$$

$$= k(\mathbf{v}, t, \mathbf{v}', t') + \int_{t'}^{t} \int_{\Delta} k(\mathbf{u}, \theta, \mathbf{v}', t') R(\mathbf{v}, t, \mathbf{u}, \theta)\, d\mathbf{u}\, d\theta. \tag{3.11}$$

By Lemma 3.1, $R(\mathbf{v}, t, \mathbf{v}', t')$ is integrable on the set of variables $\mathbf{v} \in \Delta$ for almost all $(\mathbf{v}', t')$. If, therefore, we integrate both sides of (3.11) with respect to $\mathbf{v}$ and apply Fubini's theorem, we obtain

$$\mathcal{G}(t, \mathbf{v}', t') = \varphi(t, \mathbf{v}', t') + \int_{t'}^{t} \int_{\Delta} k(\mathbf{u}, \theta, \mathbf{v}', t') \times \mathcal{G}(t, \mathbf{u}, \theta)\, d\mathbf{u}\, d\theta, \tag{3.12}$$

where

$$\mathcal{G}(t, \mathbf{v}', t') = \int_{\Delta} R(\mathbf{v}, t, \mathbf{v}', t')\, d\mathbf{v},$$

$$\varphi(t, \mathbf{v}', t') = \int_{\Delta} k(\mathbf{v}, t, \mathbf{v}', t')\, d\mathbf{v}.$$

According to Theorem 3.1, the integral equation (3.12) has a unique solution $\mathcal{G}(t, v', t')$ verifying $\mathcal{G}(t, \mathbf{v}', t') < A_3$ for almost all $(\mathbf{v}', t') \in \Delta \times [t_0, t_1]$ since, by Eq. (2.15), $\varphi \leqslant \nu(0)$. We will prove now that this unique solution is $\nu(v')$. For this purpose, substituting $\nu(v')$ for $\mathcal{G}(t, \mathbf{v}', t')$ in the second term on the right-hand side of (3.12) which we denote by Q, and using (2.10), we obtain

$$Q = \int_{t'}^{t} \int_{\Delta} \nu(u) H(S_{t',\theta}\mathbf{u}, \mathbf{v}') \times \exp\left[-\int_{t'}^{\theta} \nu(|S_{\tau,\theta}\mathbf{u}|)\, d\tau\right] d\mathbf{u}\, d\theta. \tag{3.13}$$

Let $\mathbf{w} = S_{t',\theta}\mathbf{u}$. According to (2.4), we have

$$\mathbf{u} = S_{\theta,t'}\mathbf{w} = \mathcal{R}_{\Omega(\theta - t')}\mathbf{w} + \int_{t'}^{\theta} \mathcal{R}_{\Omega(\theta - t'')}\mathbf{\Gamma}(t'')\, dt''$$

and

$$|\mathbf{u}| = |\mathcal{R}_{\Omega(t'-\theta)}\mathbf{u}| = \left|\mathbf{w} - \int_{\theta}^{t'} \mathcal{R}_{\Omega(t'-t'')}\mathbf{\Gamma}(t'')\, dt''\right|, \tag{3.14}$$

$$|S_{\tau,\theta}\mathbf{u}| = |\mathcal{R}_{\Omega(t'-\tau)} S_{\tau,\theta}\mathbf{u}| = \left|\mathbf{w} - \int_{\tau}^{t'} \mathcal{R}_{\Omega(t'-t'')}\mathbf{\Gamma}(t'')\, dt''\right|. \tag{3.15}$$

Hence,

$$Q = \int_{\Delta} H(\mathbf{w}, \mathbf{v}')\, d\mathbf{w} \int_{t'}^{t} \nu\left(\left|\mathbf{w} - \int_{\theta}^{t'} \mathcal{R}_{\Omega(t'-t'')}\mathbf{\Gamma}(t'')\, dt''\right|\right) \times \exp\left[-\int_{t'}^{\theta} \nu\left(\left|\mathbf{w} - \int_{\tau}^{t'} \mathcal{R}_{\Omega(t'-t'')}\mathbf{\Gamma}(t'')\, dt''\right|\right) d\tau\right] d\theta. \tag{3.16}$$

Performing in (3.16) the integration with respect to θ, and using (B14), (3.15) and (2.10), we finally obtain

$$Q = \nu(v') - \int_{\Delta} k(\mathbf{w}, t, \mathbf{v}', t')\, d\mathbf{w}.$$

This shows that $\nu(v')$ is a solution of (3.12) which implies that

$$\int_{\Delta} R(\mathbf{v}, t, \mathbf{v}', t')\, d\mathbf{v} = \nu(\mathbf{v}')$$

for almost all $(\mathbf{v}', t') \in \Delta \times [t_0, t_1]$.

The last result completes the statement of Lemma 3.2.

Now, we want to prove the following:

Theorem 3.2: If f_{t_0} is a nonnegative function on Δ, belonging to the function space $\mathcal{L}_1(\Delta)$, then for all $t \in [t_0, t_1]$

$$\int_{\Delta} f(\mathbf{v}, t)\, d\mathbf{v} = \int_{\Delta} f(\mathbf{v}, t_0)\, d\mathbf{v}.$$

Proof: By (3.2), $f_{t_0} \geqslant 0$ implies $\varphi \geqslant 0$. Hence, according to Theorem 3.1, if $f_{t_0} \in \mathcal{L}_1(\Delta)$, $f(\mathbf{v}, t)$ is the limit for almost all $\mathbf{v} \in \Delta$, of the ascending sequence

$$h_n(\mathbf{v}, t) = \varphi(\mathbf{v}, t) + \int_{t_0}^{t} \int_{\Delta} R^{(n-1)}(\mathbf{v}, t, \mathbf{v}', t')\varphi(\mathbf{v}', t')\, d\mathbf{v}' dt'. \tag{3.17}$$

This implies that for almost all $\mathbf{v} \in \Delta$,

$$\int_{t_0}^{t} \int_{\Delta} R^{(n)}(\mathbf{v}, t, \mathbf{v}', t')\varphi(\mathbf{v}', t')\, d\mathbf{v}'\, dt' < f(\mathbf{v}, t) - \varphi(\mathbf{v}, t) < B(t_0, t_1) < +\infty. \tag{3.18}$$

Consequently, according to the theorem of Beppo Levi, the ascending sequence $R^{(n)}\varphi$ converges to a limit $R\varphi$ and for almost all $\mathbf{v} \in \Delta$,

$$\lim_{n\to\infty} \int_{t_0}^{t} \int_{\Delta} R^{(n)}(\mathbf{v}, t, \mathbf{v}', t')\varphi(\mathbf{v}', t')\, d\mathbf{v}'\, dt' = \int_{t_0}^{t} \int_{\Delta} R(\mathbf{v}, t, \mathbf{v}', t')\varphi(\mathbf{v}', t')\, d\mathbf{v}'\, dt'$$

and

$$f(\mathbf{v}, t) = \varphi(\mathbf{v}, t) + \int_{t_0}^{t} \int_{\Delta} R(\mathbf{v}, t, \mathbf{v}', t')\varphi(\mathbf{v}', t')\, d\mathbf{v}'\, dt'. \tag{3.19}$$

Furthermore, since $f(\mathbf{v}, t)$ and $\varphi(\mathbf{v}, t)$ are integrable with respect to $\mathbf{v}$, the second term on the right-hand side of (3.19) is also integrable on Δ. Hence, using Fubini's theorem and Lemma 3.2, we obtain

$$\int_{\Delta} f(\mathbf{v}, t)\, d\mathbf{v} = \int_{\Delta} \varphi(\mathbf{v}, t)\, d\mathbf{v} + \int_{t_0}^{t} \int_{\Delta} \nu(v')\varphi(\mathbf{v}', t')\, d\mathbf{v}'\, dt'. \tag{3.20}$$

Now, in view of (3.2), we can write

$$\int_{t_0}^{t} \int_{\Delta} \nu(v')\varphi(\mathbf{v}', t')\, d\mathbf{v}'\, dt' = \int_{t_0}^{t} \int_{\Delta} \nu(v') f(S_{t_0,t'}\mathbf{v}', t_0) \exp\left[-\int_{t_0}^{t'} \nu(|S_{\theta,t'}\mathbf{v}'|)\, d\theta\right] d\mathbf{v}'\, dt' \tag{3.21}$$

Finally, using the same variable change $S_{t_0,t'}\mathbf{v}' = \mathbf{w}'$ as above (see Eq. 3.13), we have

$$\int_{t_0}^{t} \int_{\Delta} \nu(v')\varphi(\mathbf{v}', t')\, d\mathbf{v}'\, dt' = \int_{\Delta} f(\mathbf{w}, t_0)\, d\mathbf{w} - \int_{\Delta} \varphi(\mathbf{w}, t)\, d\mathbf{w}, \tag{3.22}$$

which implies

$$\int_{\Delta} f(\mathbf{v}, t)\, d\mathbf{v} = \int_{\Delta} f(\mathbf{v}, t_0)\, d\mathbf{v}. \tag{3.23}$$

Corollary: If f_{t_0} is a nonnegative function on Δ, belonging to $\mathcal{L}_1(\Delta)$, then, for almost all $\mathbf{v} \in \Delta$, $f(\mathbf{v}, t)$ has a finite upper bound for $t \in [t_0, +\infty[$.

Proof: Follows immediately from Theorem 3.2.

IV. METHODS OF SOLUTION OF THE INTEGRAL EQUATION FOR RIGID SPHERES ($s = \infty$)

It has been shown in Appendix B that for $s > 5$, the collision frequency is a monotonically increasing function of v. As a consequence, Theorem 2.1 no longer holds. Thus, the procedure of the previous section cannot be used in this case. However, it is still possible to prove the existence and uniqueness of the solution of Eq. (2.9) for a large class of initial distribution functions. Our method, which is based on the construction of "majorant" series of the series obtained by applying the iteration procedure to Eq. (2.9), will be illustrated here in the particular case of rigid spheres. The general case $5 < s < \infty$ will be treated in Part II.

A. Properties of the resolvent kernel

In order to derive the basic properties of the resolvent kernel for rigid spheres, we first prove some lemmas.

Let $k_d(\mathbf{v}, t, \mathbf{v}', t')$ be the kernel of the integral equation (2.9), for rigid spheres. In view of (B20) and (B21), it can be easily shown by applying the Schwarz inequality to the iterated kernel defined by

$$k_d^{(p)}(\mathbf{v}, t, \mathbf{v}', t') = \int_{t_0}^{t} \int_{\Delta} k_d(\mathbf{v}, t, \mathbf{u}, \theta) k_d^{(p-1)}(\mathbf{u}, \theta, \mathbf{v}', t')\, d\mathbf{u}\, d\theta \tag{4.1}$$

so that for $p \geq 2$,

$$k_d^{(p)}(\mathbf{v}, t, \mathbf{v}', t') < \left(\frac{M+m}{M-m}\right)^{2(p-2)} a_5. \tag{4.2}$$

However, the majorant series defined by (4.2) diverges. In order to get a more useful majorant, we consider the sequence of integrals

$$I_p(\mathbf{v}, t) = \int_{t_0}^{t} \int_{\Delta} k_d^{(p)}(\mathbf{v}, t, \mathbf{v}', t') \exp(-\alpha v'^2)\, d\mathbf{v}'\, dt', \tag{4.3}$$

where α is a positive constant.

If we replace β_n by $(M/m)\alpha$ in the expressions of $H_d(\mathbf{v}, \mathbf{v}')$, $\nu_d(v)$, $Y(v)$ without changing the other coefficients, then, according to (B10), (B14), (B19), and (B22), we have

$$H_d(\mathbf{v}, \mathbf{v}'; \alpha) \exp(-\alpha v'^2) = H_d(\mathbf{v}', \mathbf{v}; \alpha) \exp(-\alpha v^2), \tag{4.4}$$

$$\int_{\Delta} H_d(\mathbf{v}', \mathbf{v}; \alpha)\, d\mathbf{v}' = \nu(v; \alpha), \tag{4.5}$$

$$\int_{\Delta} H_d(\mathbf{v}, \mathbf{v}'; \alpha)\, dv' = \left(\frac{M+m}{M-m}\right)^2 \nu(v; \alpha), \tag{4.6}$$

$$\frac{1}{l_\alpha} v < \nu(v; \alpha) \leq \frac{1}{l_\alpha}(v + \bar{v}_\alpha), \tag{4.7}$$

where

$$l_\alpha = l\left(\frac{\alpha}{\beta}\right)^{1/2}, \quad \bar{v}_\alpha = \bar{v}_n \left(\frac{\beta}{\alpha}\right)^{1/2}$$

$$\nu(v; \alpha) = 4\pi a_d \left(\frac{M}{M+m}\right)^2 Y(v; \alpha).$$

Now let Ψ_p, $p \geq 0$, be the sequence of functions defined by

$$\Psi_p(v; \alpha) = 2^{-p}(v + M_p + |M_p - v| + 2\bar{v}_\alpha)^p \exp[-\alpha/4(v + M_p + |M_p - v| - 2\gamma_1)^2], \tag{4.8}$$

where

$$\gamma_1 = \Gamma_M(t_1 - t_0), \quad \Gamma_M = \sup_{t \in [t_0, t_1]} |\Gamma(t)|,$$

and where

$$M_p = \sup_{v \geq 0}[h_p(v)] = \tfrac{1}{2}(\gamma_1 - \bar{v}_\alpha) + [\tfrac{1}{4}(\gamma_1 + \bar{v}_\alpha)^2 + p/2\alpha]^{1/2} > \gamma_1 \tag{4.9}$$

with

$$h_p(v) = (v + \bar{v}_\alpha)^p \exp[-\alpha(v - \gamma_1)^2].$$

Now, we want to prove the following:

Lemma 4.1: For $\alpha \leq (2\Gamma_M l)^{-1}$, a constant $C_\alpha > 0$ exists such that

$$\int_{\Delta} k_d(\mathbf{v}, t, \mathbf{v}', t'; \alpha) \Psi_p(v'; \alpha)\, d\mathbf{v}' < C_\alpha \Psi_{p+1}(v; \alpha), \quad p \geq 0, \tag{4.10}$$

where $k_d(\mathbf{v}, t, \mathbf{v}', t'; \alpha)$ is defined by

$$k_d(\mathbf{v}, t, \mathbf{v}', t'; \alpha) = H_d(S_{t',t}\mathbf{v}, \mathbf{v}'; \alpha) \exp[-\int_{t'}^{t} \nu_d(|S_{\theta,t}\mathbf{v}|)\, d\theta]. \tag{4.11}$$

Proof: We consider separately the cases $v \leq M_{p+1}$ and $v > M_{p+1}$.

(a) $v \leq M_{p+1}$

According to (4.8), we have

$$\Psi_{p+1}(v; \alpha) = (M_{p+1} + \bar{v}_\alpha)^{p+1} \exp[-\alpha(M_{p+1} - \gamma_1)^2], \quad 0 \leq v \leq M_{p+1}$$

and

$$\Psi_p(v'; \alpha) \leq (M_p + \bar{v}_\alpha)^p \exp[-\alpha(M_p - \gamma_1)^2], \quad v' \in [0, +\infty[.$$

Then, by (4.6),

$$\int_{\Delta} H_d(S_{t',t}\mathbf{v}, \mathbf{v}'; \alpha) \Psi_p(v'; \alpha)\, d\mathbf{v}' \leq \frac{M+m}{M-m}^{\;2} \nu_d(|S_{t',t}\mathbf{v}|; \alpha)(M_p + \bar{v}_\alpha)^p \exp[-\alpha(M_p - \gamma_1)^2]. \tag{4.12}$$

Furthermore, according to (2.5) and (4.7), we may write

$$|S_{t',t}\mathbf{v}| \leq v + \Gamma_M(t - t'), \quad \nu_d(|S_{t',t}\mathbf{v}|; \alpha) < 1/l_\alpha[l_\alpha \nu_d(v; \alpha) + \Gamma_M(t - t') + \bar{v}_\alpha]$$

and since

$$\nu_d(v; \alpha) \geq \nu_d(0; \alpha) = \bar{v}_\alpha / l_\alpha, \quad \nu_d(|S_{\theta,t}\mathbf{v}|) > \nu_d(0) = \bar{v}_n / l,$$

we have

$$\nu_d(|S_{t',t}\mathbf{v}|; \alpha) < \nu(v; \alpha)[2 + (\Gamma_M / \bar{v}_\alpha)(t - t')], \tag{4.13}$$

$$\exp[-\textstyle\int_{t'}^{t}\nu_d(|S_{\theta,t}\mathbf{v}|)\,d\theta]\leqslant\exp[-(\bar{v}_n/l)(t-t')]. \qquad (4.14)$$

Let $\mathcal{Y}_{\max}=\sup_{x\geqslant 0}[\mathcal{Y}(x)]$, where $\mathcal{Y}(x)=[2+(\Gamma_M/\bar{v}_\alpha)x]\times\exp[-(\bar{v}_n/l)x]$. It is easy to verify that

$$\mathcal{Y}_{\max}=\begin{cases}2, & \Gamma_M\leqslant\dfrac{2\bar{v}_n\bar{v}_\alpha}{l},\\[2mm] \dfrac{\Gamma_M l}{\bar{v}_n\bar{v}_\alpha}\exp\left(\dfrac{2\bar{v}_n\bar{v}_\alpha}{\Gamma_M l}-1\right)\equiv b_\alpha, & \Gamma_M>\dfrac{2\bar{v}_n\bar{v}_\alpha}{l}.\end{cases}$$

Hence, according to (4.13) and (4.14), we have for $v\geqslant 0$,

$$\exp[-\textstyle\int_{t'}^{t}\nu_d(|S_{\theta,t}\mathbf{v}|)\,d\theta]\nu_d(|S_{t',t}\mathbf{v}|;\alpha)<a_\alpha\nu_d(v;\alpha), \qquad (4.15)$$

where $a_\alpha=\max[2,b_\alpha]$. Furthermore, since $v\leqslant M_{p+1}$, we have $\nu_d(v;\alpha)\leqslant(1/l_\alpha)(M_{p+1}+\bar{v}_\alpha)$ and, by (4.9) $M_{p+1}+\bar{v}_\alpha<K_\alpha(M_p+\bar{v}_\alpha)$ with $K_\alpha=(M_1+\bar{v}_\alpha)/(M_0+\bar{v}_\alpha)$. Hence $\nu_d(v;\alpha)<(K_\alpha/l_\alpha)(M_p+\bar{v}_\alpha)$ and since $M_{p+1}>M_p>\gamma_1$,

$$\begin{aligned}&\nu_d(v;\alpha)(M_p+\bar{v}_\alpha)^p\exp[-\alpha(M_p-\gamma_1)^2\\&\quad<\frac{K_\alpha}{l_\alpha}(M_p+\bar{v}_\alpha)^{p+1}\exp[-\alpha(M_p-\gamma_1)^2]\\&\quad<\frac{2}{l_\alpha}(M_{p+1}+\bar{v}_\alpha)^{p+1}\exp[-\alpha(M_{p+1}-\gamma_1)^2]=\frac{K_\alpha}{l_\alpha}\Psi_{p+1}(v;\alpha).\end{aligned}$$

Inserting this result into (4.12), we obtain the inequality

$$\int_\Delta k_d(\mathbf{v},t,\mathbf{v}',t';\alpha)\Psi_p(v';\alpha)\,d\mathbf{v}'<\frac{K_\alpha}{l_\alpha}\left(\frac{M+m}{M-m}\right)^2a_\alpha\Psi_{p+1}(v;\alpha) \qquad (4.16)$$

which is of the form given by (4.10).

(b) $v>M_{p+1}$

According to (4.8), we now have $\Psi_p(v;\alpha)=(v+\bar{v}_\alpha)^p\times\exp[-\alpha(v-\gamma_1)^2]$. For $v'>v$, this implies that $\Psi_p(v';\alpha)<\Psi_p(v;\alpha)$. Hence, according to (4.6),

$$\int_S d\hat{\mathbf{u}}\int_v^\infty H_d(S_{t',t}\mathbf{v},\mathbf{v}';\alpha)\Psi_p(v';\alpha)v'^2\,dv'<\left(\frac{M+m}{M-m}\right)^2\nu(|S_{t',t}\mathbf{v}|;\alpha)(v+\bar{v}_\alpha)^p\exp[-\alpha(v-\gamma_1)^2], \qquad (4.17)$$

where the integral of $\hat{\mathbf{u}}$ is taken over the unit sphere S.

For $M_p<v'\leqslant v$, we may write $\Psi_p(v';\alpha)<\exp(-\alpha\gamma_1^2)\times(v+\bar{v}_\alpha)^p\exp[-\alpha(v'^2-2\gamma_1 v)]$. Then, by (4.4) and (4.5), we have

$$\begin{aligned}&\int_S d\hat{\mathbf{u}}\int_{M_p}^{v}H_d(S_{t',t}\mathbf{v},\mathbf{v}';\alpha)\Psi_p(v';\alpha)v'^2\,dv'\\&\quad<\exp(-\alpha\gamma_1^2)(v+\bar{v}_\alpha)^p\nu_d(|S_{t',t}\mathbf{v}|;\alpha)\\&\quad\times\exp[-\alpha(|S_{t',t}\mathbf{v}|^2-2\gamma_1 v)].\end{aligned} \qquad (4.18)$$

For $v'\leqslant M_p$, we have

$$\Psi_p(v';\alpha)=(M_p+\bar{v}_\alpha)^p\exp[-\alpha(M_p-\gamma_1)^2],$$
$$(M_p+\bar{v}_\alpha)^p<(v+\bar{v}_\alpha)^p,$$

and

$$\begin{aligned}&\exp(-\alpha M_p^2)\int_S d\hat{\mathbf{u}}\int_0^{M_p}H_d(S_{t',t}\mathbf{v},\mathbf{v}';\alpha)v'^2\,dv'\\&\quad<\int_S d\hat{\mathbf{u}}\int_0^{M_p}\exp(-\alpha v'^2)H_d(S_{t',t}\mathbf{v},\mathbf{v}';\alpha)v'^2\,dv'\\&\quad<\nu_d(|S_{t',t}\mathbf{v}|;\alpha)\exp(-\alpha|S_{t',t}\mathbf{v}|^2),\end{aligned}$$

which implies

$$\begin{aligned}&\int_S d\hat{\mathbf{u}}\int_0^{M_p}H_d(S_{t',t}\mathbf{v},\mathbf{v}';\alpha)\Psi_p(v';\alpha)v'^2\,dv'\\&\quad<\exp(-\alpha\gamma_1^2)(v+\bar{v}_\alpha)^p\nu_d(|S_{t',t}\mathbf{v}|;\alpha)\\&\quad\times\exp[-\alpha(|S_{t',t}\mathbf{v}|^2-2\gamma_1 v)].\end{aligned} \qquad (4.19)$$

Inequalities (4.17), (4.18), and (4.19), together with (4.15), then give

$$\begin{aligned}&\int_\Delta k_d(\mathbf{v},t,\mathbf{v}',t')\Psi_p(v';\alpha)\,d\mathbf{v}'\\&\quad<\frac{1}{l_\alpha}\left(\frac{M+m}{M-m}\right)^2a_\alpha\Psi_{p+1}(v;\alpha)+2\exp[-\alpha(\gamma_1^2-2\gamma_1 v)]\\&\quad\times(v+\bar{v}_\alpha)^p\nu_d(|S_{t',t}\mathbf{v}|;\alpha)\exp[-\alpha|S_{t',t}\mathbf{v}|^2\\&\quad-\int_{t'}^{t}\nu_d(|S_{\theta,t}\mathbf{v}|)\,d\theta].\end{aligned} \qquad (4.20)$$

Furthermore, for $|S_{t',t}\mathbf{v}|\geqslant v$ and in view of (4.15), we have

$$\nu_d(|S_{t',t}\mathbf{v}|;\alpha)\exp[-\textstyle\int_{t'}^{t}\nu_d(|S_{\theta,t}\mathbf{v}|)\,d\theta-\alpha|S_{t',t}\mathbf{v}|^2]\leqslant a_\alpha\nu_d(v;\alpha)\exp(-\alpha v^2).$$

If $|S_{t',t}\mathbf{v}|<v$, then $\nu_d(|S_{t',t}\mathbf{v}|;\alpha)<\nu_d(v;\alpha)$ and since $v>M_{p+1}>\gamma_1$, it follows immediately from (2.5) and (B22) that

$$\begin{aligned}&|S_{t',t}\mathbf{v}|^2=|\mathcal{R}_{\Omega(t-t')}S_{t',t}\mathbf{v}|^2\\&\quad>(v-|\textstyle\int_{t'}^{t}\mathcal{R}_{\Omega(t-\theta)}\Gamma(\theta)\,d\theta|)^2>[v-\Gamma_M(t-t')]^2,\\&\nu_d(|S_{\theta,t}\mathbf{v}|)>(1/l)|S_{\theta,t}\mathbf{v}|>(1/l)(v-\gamma_1).\end{aligned}$$

Hence

$$\begin{aligned}&\exp[-\textstyle\int_{t'}^{t}\nu_d(|S_{\theta,t}\mathbf{v}|)\,d\theta-\alpha|S_{t',t}\mathbf{v}|^2]\\&\quad<\exp[(\gamma_1/l)(t_1-t_0)]\exp(-\alpha v^2)\exp[(2\alpha\Gamma_M-1/l)v(t-t')].\end{aligned} \qquad (4.21)$$

Now, for $\alpha\leqslant(2\Gamma_M l)^{-1}$, the term on the right-hand side of inequality (4.21 is lower than $\exp[(\gamma_1/l)(t_1-t_0)-\alpha v^2]$. Hence, by (4.20) and (4.7),

$$\int_\Delta k_d(\mathbf{v},t,\mathbf{v}',t';\alpha)\Psi_p(v';\alpha)<\frac{1}{l_\alpha}\left[a_\alpha\left(\frac{M+m}{M-m}\right)^2+2d_\alpha\right]\Psi_{p+1}(v;\alpha),\quad p\geqslant 0, \qquad (4.22)$$

where $d_\alpha=\max\{a_\alpha,\exp[(\gamma_1/l)(t_1-t_0)]\}$. This result completes the proof of Lemma 4.1.

Let us now consider the sequence of integrals I_p defined by (4.3).

Lemma 4.2: For any positive α, the series $\sum_{p\geqslant 1}I_p$ converges uniformly on $\Delta\times[t_0,t_1]$.

Proof: We consider separately the cases $\alpha > \beta$ and $\alpha \leq \beta$ $(\beta = m/kT)$.

(a) $\alpha > \beta$

If $\alpha > \beta$, $\exp(-\alpha v'^2) < \exp(-\beta v'^2)$. Then, for $p \geq 1$, it follows that

$$I_p(\mathbf{v}, t) < \int_{t_0}^{t} \int_\Delta k_d^{(p)}(\mathbf{v}, t, \mathbf{v}', t') \exp(-\beta v'^2)\, d\mathbf{v}'\, dt' = i_p(\mathbf{v}, t). \tag{4.23}$$

We will first prove that for $\beta \leq (2\Gamma_M l)^{-1}$ and $p \geq 1$, we have

$$i_p(\mathbf{v}, t) < \exp(-\beta\gamma_1^2) C_\beta^p \frac{(t-t_0)^p}{p!} \Psi_p(v;\beta) = U_p^{(\beta)}. \tag{4.24}$$

According to Lemma 4.1, we can write

$$i_1(\mathbf{v}, t) = \int_{t_0}^{t} \int_\Delta k_d(\mathbf{v}, t, \mathbf{v}', t'; \beta) \exp(-\beta v'^2)\, d\mathbf{v}'\, dt'$$

$$< \exp(\beta\gamma_1^2) \int_{t_0}^{t} \int_\Delta k_d(\mathbf{v}, t, \mathbf{v}', t'; \beta) \Psi_0(v';\beta)\, d\mathbf{v}'\, dt'$$
$$< \exp(\beta\gamma_1^2) C_\beta \Psi_1(v;\beta)(t-t_0).$$

Thus, inequality (4.24) holds for $p = 1$. We will assume now that (4.24) is satisfied at the order $p-1$. Then, according to Eq. (4.1) and Lemma 4.1, we have

$$i_p(\mathbf{v}, t) = \int_{t_0}^{t} \int_\Delta k_d(\mathbf{v}, t, \mathbf{v}', t'; \beta) i_{p-1}(\mathbf{v}', t')\, d\mathbf{v}'\, dt'$$
$$< \exp(\beta\gamma_1^2) C_\beta^{p-1} \int_{t_0}^{t} \frac{(t-t_0)^{p-1}}{(p-1)!} dt'$$
$$\times \int_\Delta k_d(\mathbf{v}, t, \mathbf{v}', t'; \beta) \Psi_{p-1}(v';\beta)\, d\mathbf{v}' < U_p^{(\beta)}$$

and therefore, inequality (4.24) holds.

Consequently, if $\beta \leq (2\Gamma_M l)^{-1}$, we obtain from (4.23) and (4.24), for $p \geq 1$

$$I_p(\mathbf{v}, t) < U_p^{(\beta)}. \tag{4.25}$$

Similarly, if $\beta > (2\Gamma_M l)^{-1} = \delta$, we have

$$i_1(\mathbf{v}, t) < \int_{t_0}^{t} \int_\Delta k_d(\mathbf{v}, t, \mathbf{v}', t'; \delta) \exp(-\delta v'^2)\, d\mathbf{v}'\, dt',$$

and by repeating the same process,

$$I_p(\mathbf{v}, t) < U_p^{(\delta)}. \tag{4.26}$$

(b) $\alpha \leq \beta$

In this case, we can write

$$I_p(v, t) < \int_{t_0}^{t} \int_\Delta k_d^{(p)}(\mathbf{v}, t, \mathbf{v}', t'; \alpha) \exp(-\alpha v'^2)\, d\mathbf{v}'\, dt'$$

which yields for $p \geq 1$,

$$I_p(\mathbf{v}, t) < U_p^{(\nu)}, \quad \nu = \min(\alpha, \delta). \tag{4.27}$$

Furthermore, according to (4.8) and (4.9), the sequence $U_p^{(\mu)}$ where $\mu = \min(\alpha, \beta, \delta)$, verifies

$$U_p^{(\mu)} < \exp(\mu\gamma_1^2) C_\mu^p \frac{(t_1-t_0)^p}{p!} (M_p + \bar{v}_\mu)^p \exp[-\mu(M_p - \gamma_1)^2]$$
$$\equiv V_p^{(\mu)}.$$

This shows that the infinite series $\sum_{p\geq1} I_p$ has the majorant $\sum_{p\geq1} V_p^{(\mu)}$, where the convergence of the last series can be easily proved by applying d'Alembert's criterium. Indeed, by (4.9), we have

$$\frac{M_{p+1} + \bar{v}_\mu}{M_p + \bar{v}_\mu} < \frac{p+1}{p}$$

and since $M_{p+1} > M_p$, $\exp\{-\mu[(M_{p+1} - \gamma_1)^2 - (M_p - \gamma_1)^2]\} < 1$. Then, obviously

$$\lim_{p\to\infty} \left(\frac{V_{p+1}}{V_p}\right) = 0.$$

Consequently, since $I_p > 0$, the series $\sum_{p\geq1} I_p$ converges uniformly on $\Delta\times[t_0, t_1]$, for any positive α.

The previous results allow us to establish some properties of the resolvent kernel which are similar, but not identical to those given by Lemmas 3.1 and 3.2.

Theorem 4.1: The sequence $R_d^{(n)}(\mathbf{v}, t, \mathbf{v}', t')$ converges for all $(\mathbf{v}, t) \in \Delta\times[t_0, t_1]$ and for almost all $(\mathbf{v}', t') \in \Delta \times[t_0, t_1]$ to a limit $R_d(\mathbf{v}, t, \mathbf{v}', t')$. For almost all $(\mathbf{v}, t) \in \Delta\times[t_0, t_1]$, this limit is integrable on the set of all $\mathbf{v} \in \Delta$.

Proof: By Lemma 4.2 and the definition of $R_d^{(n)}$, we have, for any positive α,

$$\lim_{n\to\infty} \int_{t_0}^{t} \int_\Delta R_d^{(n)}(\mathbf{v}, t, \mathbf{v}', t') \exp(-\alpha v'^2)\, d\mathbf{v}'\, dt'$$
$$< C(t_0, t_1) < +\infty.$$

Hence, according to the theorem of Beppo Levi,[9] the ascending sequence $R_d^{(n)}(\mathbf{v}, t, \mathbf{v}', t')$ converges for all $(\mathbf{v}, t) \in \Delta\times[t_0, t_1]$ and for almost all $(\mathbf{v}', t') \in \Delta\times[t_0, t_1]$ to a limit $R_d(\mathbf{v}, t, \mathbf{v}', t')$ which verifies

$$\int_{t_0}^{t} \int_\Delta R_d(\mathbf{v}, t, \mathbf{v}', t') \exp(-\alpha v'^2)\, d\mathbf{v}'\, dt' < +\infty. \tag{4.28}$$

In order to prove the second part of Theorem 4.1, we use the porperties of the majorant $S_\mu(v, t) = \sum_{p\geq1} U_p^{(\mu)}$, $\mu = \min(\alpha, \beta, \delta)$ of the series $\sum_{p\geq1} I_p$. It is obvious that the infinite series $\sum_{p\geq1} U_p^{(\mu)}$, which has the majorant $\sum_{p\geq1} V_p^{(\mu)}$, is absolutely and uniformly convergent. This allows the integration of the series $U_p^{(\mu)}$ term by term. Then, according to (4.8) and (4.24), we have

$$\int_\Delta S_\mu(v, t)\, d\mathbf{v} = 4\pi \exp(-\mu\gamma_1^2)(A_1 + A_2),$$

where

$$A_1 = \sum_{p\geq1} C_\mu^p \frac{(t-t_0)^p}{p!} \int_0^{M_p} v^2 \Psi_p(v;\alpha)\, dv$$
$$= \sum_{p\geq1} C_\mu^p \frac{(t-t_0)^p}{p!} \frac{M_p^3}{3} (M_p + \bar{v}_\mu)^p \exp[-\mu(M_p - \gamma_1)^2],$$
$$A_2 = \sum_{p\geq1} C_\mu^p \frac{(t-t_0)^p}{p!} \int_{M_p}^{\infty} (v + \bar{v}_\mu)^p \exp[-\mu(v-\gamma_1)^2] v^2\, dv$$
$$< \int_0^\infty \exp[-\mu(v-\gamma_1)^2 \exp[C_\mu(t-t_0)(v+v_\mu)]\, v^2\, dv.$$

The proof of the uniform convergence of the series A_1 is similar to that of the series $V_p^{(\mu)}$. Consequently, the majorant $S_\mu(v, t)$ as well as the series $\sum_{p\geq1} I_p(\mathbf{v}, t)$, are integrable with respect to $\mathbf{v}$. Hence, according to the first part of Theorem 4.1, we have

$$\int_\Delta \sum_{p\geq1} I_p(\mathbf{v}, t)\, d\mathbf{v}$$
$$= \int_\Delta d\mathbf{v} \int_{t_0}^{t} \int_\Delta R_d(\mathbf{v}, t, \mathbf{v}', t') \exp(-\alpha v'^2)\, d\mathbf{v}'\, dt' < +\infty.$$

Furthermore, since R_d is a positive function, it follows from Fubini's theorem that

$$\int_{t_0}^{t}\int_\Delta \exp(-\alpha v'^2)\,d\mathbf{v}'\,dt' \int_\Delta R_d(\mathbf{v},t,\mathbf{v}',t')\,d\mathbf{v} < +\infty, \tag{4.29}$$

which implies that $R_d(\mathbf{v},t,\mathbf{v}',t')$ is integrable on the set of all $\mathbf{v}\in\Delta$, for almost all $(\mathbf{v}',t')\in\Delta\times[t_0,t_1]$.

We note however, that $R_d(\mathbf{v},t,\mathbf{v}',t')$ is not integrable with respect to $\mathbf{v}'$.

Theorem 4.2: For almost all $(\mathbf{v}',t')\in\Delta\times[t_0,t_1]$, the resolvent kernel $R_d(\mathbf{v},t,\mathbf{v}',t')$ verifies

$$\int_\Delta R_d(\mathbf{v},t,\mathbf{v}',t')\,dv = \nu(v').$$

Proof: Applying again the method used to prove Lemma 3.2, it can be shown in a similar way that for almost all $(\mathbf{v}',t')$, $R_d(\mathbf{v},t,\mathbf{v}',t')$ verifies

$$\begin{aligned}\int_\Delta R_d(\mathbf{v},t,\mathbf{v}',t')\,d\mathbf{v} &= \int_\Delta k_d(\mathbf{v},t,\mathbf{v}',t')\,d\mathbf{v} + \int_{t'}^{t}\int_\Delta k_d(\mathbf{u},\theta,\mathbf{v}',t')\,d\mathbf{u}\,d\theta\\ &\quad\times\int_\Delta R_d(\mathbf{v},t,\mathbf{u},\theta)\,d\mathbf{v}\end{aligned} \tag{4.30}$$

and that $\nu(v')$ is a particular solution of this integral equation, satisfying (4.29).

We will prove now that for almost all $(\mathbf{v}',t')$, $\nu(v')=\int_\Delta R_d(\mathbf{v},t,\mathbf{v}',t')\,d\mathbf{v}$. For this, it is sufficient to prove that if $j(t,\mathbf{v}',t')$ denotes another positive solution of (4.30), which verifies (4.29) for any positive α, then $j(t,\mathbf{v},t')=\int_\Delta R_d(\mathbf{v},t,\mathbf{v}',t')\,d\mathbf{v}$ for almost all $(\mathbf{v}',t')$.

According to our assumptions, we have

$$j(t,\mathbf{v}',t') = \int_\Delta k_d(\mathbf{v},t,\mathbf{v}',t')\,d\mathbf{v} + \int_{t'}^{t}\int_\Delta k_d(\mathbf{u},\theta,\mathbf{v}',t')j(t,\mathbf{u},\theta)\,d\mathbf{u}\,d\theta \tag{4.31}$$

and

$$\int_{t_0}^{t}\int_\Delta j(t,\mathbf{v}',t')\exp(-\alpha v'^2)\,d\mathbf{v}'\,dt' < +\infty, \quad \forall\ \alpha>0. \tag{4.32}$$

Next, we apply the method of successive approximations to (4.31) which yields

$$j(t,\mathbf{v}',t') = \int_\Delta R_d^{(n)}(\mathbf{v},t,\mathbf{v}',t')\,d\mathbf{v} + \int_{t_0}^{t}\int_\Delta k^{(n)}(\mathbf{u},\theta,\mathbf{v}',t')j(t,\mathbf{u},\theta)\,d\mathbf{u}\,d\theta. \tag{4.33}$$

We then multiply both sides of (4.33) by $\exp(-\alpha v'^2)$ and integrate them with respect to $(\mathbf{v}',t')$. By applying Fubini's theorem and using (4.3), (4.23), (4.24), and (4.8), we obtain

$$\begin{aligned}&\int_{t_0}^{t}\int_\Delta \left[j(t,\mathbf{v}',t') - \int_\Delta R^{(n)}(\mathbf{v},t,\mathbf{v}',t')\,d\mathbf{v}\right]\exp(-\alpha v'^2)\,d\mathbf{v}'\,dt'\\ &\quad< \int_{t_0}^{t}\int_\Delta j(t,\mathbf{u},\theta)i_n(u,\theta)\,d\mathbf{u}\,d\theta < \int_{t_0}^{t}\int_\Delta j(t,\mathbf{u},\theta)U_n(u;\mu)\,d\mathbf{u}\,d\theta\\ &\quad< \exp(-\mu\gamma_1^2)(C_n+D_n),\end{aligned} \tag{4.34}$$

where

$$C_n = C_\mu^n\frac{(t-t_0)^n}{n!}(M_n+\bar v_\mu)^n\int_{t_0}^{t}\int_{\Delta_n} j(t,\mathbf{u},\theta)\exp[-\mu(M_n-\gamma_1)^2]\,d\mathbf{u},a$$

$$D_n = C_\mu^n\frac{(t-t_0)^n}{n!}\int_{t_0}^{t}\int_{\Delta-\Delta_n} j(t,\mathbf{u},\theta)(u+\bar v_\mu)^n \times\exp[-\mu(u-\gamma_1)^2]\,d\mathbf{u}\,d\theta,$$

and where $\mu=\min(\alpha,\beta,\delta)$ and $\Delta_n=\{\mathbf{u}\mid 0\le u\le M_n\}$.

Using (4.32) and the fact that $j(t,\mathbf{u},\theta)$ is a positive function, we have

$$\begin{aligned}&\int_{t_0}^{t}\int_{\Delta_n} j(t,\mathbf{u},\theta)\exp[-\mu(M_n-\gamma_1)^2]\,d\mathbf{u}\,d\theta\\ &\quad< \int_{t_0}^{t}\int_\Delta j(t,\mathbf{u},\theta)\exp[-\mu(u-\gamma_1)^2]\,d\mathbf{u}\,d\theta\\ &\quad= A(t) < +\infty.\end{aligned} \tag{4.35}$$

Thus, the term C_n is analogous to the general term of the series V_n^μ and consequently $\lim_{n\to\infty}C_n=0$.

The same result holds for D_n. Indeed, since $j(t,\mathbf{u},\theta)>0$, we have $D_n<E_n$, where

$$E_n = C_\mu^n\frac{(t-t_0)^n}{n!}\int_{t_0}^{t}\int_\Delta (u+v_\mu)^n j(t,\mathbf{u},\theta)\exp[-\mu(u-\gamma_1)^2]\,d\mathbf{u}\,d\theta.$$

Moreover, the ascending sequence $B_n=\sum_{\nu=1}^{n}E_\nu$ has the majorant

$$M = \int_{t_0}^{t}\int_\Delta j(t,\mathbf{u},\theta)\exp\{C_\mu(u+\bar v_\mu)(t-t_0)-\mu(u-\gamma_1)^2\}\,d\mu\,d\theta$$

and, in virtue of (4.32), $M<+\infty$. Hence B_n is convergent and consequently $\lim_{n\to\infty}E_n=0$, which implies that $\lim_{n\to\infty}D_n=0$. Finally, according to Theorem 4.1, if we take the limit $n\to\infty$ of both sides of (4.34), we obtain

$$\int_{t_0}^{t}\int_\Delta \left|j(t,\mathbf{v}',t') - \int_\Delta R(\mathbf{v},t,\mathbf{v}',t')\,d\mathbf{v}\right|\exp(-\alpha v'^2)\,d\mathbf{v}'\,dt' = 0,$$

which implies that $j(t,\mathbf{v}',t')=\int_\Delta R(\mathbf{v},t,\mathbf{v}',t')\,d\mathbf{v}$ for almost all $(\mathbf{v}',t')$.

B. Existence and uniqueness of the solution

The properties of the resolvent kernel given in Sec. IV.A. (Theorems 4.1 and 4.2) allow us to prove the following existence and uniqueness theorem:

Theorem 4.3: If the class $\tilde f_{t_0}$ of initial distribution functions $\tilde f_{t_0}$ belongs to the Lebesgue function space $L_1(\Delta)$ and if $f_{t_0}(\mathbf{v})>0$ on Δ, the integral equation form of the Boltzman equation for rigid spheres has a unique solution f_t in $L_1(\Delta)$, having the properties $f_t(\mathbf{v})>0$ and $\int_\Delta f_t(\mathbf{v})\,d\mathbf{v}=\int_\Delta f_{t_0}(\mathbf{v})\,d\mathbf{v}$.

Proof: Applying the method of successive approximations to the integral equation (2.9), we obtain the following sequence:

$$\varphi_{n+1}(\mathbf{v},t) = \varphi(\mathbf{v},t) + \int_{t_0}^{t}\int_\Delta R_d^{(n)}(\mathbf{v},t,\mathbf{v}',t')\varphi(\mathbf{v}',t')\,d\mathbf{v}'\,dt', \tag{4.36}$$

where $\varphi_1(\mathbf{v},t)=\varphi(\mathbf{v},t)$ and

$$\varphi(\mathbf{v},t) = f(S_{t_0,t}\mathbf{v},t_0)\exp\left[-\int_{t_0}^{t}\nu_d(|S_{\theta,t}\mathbf{v}|)\,d\theta\right]. \tag{4.37}$$

Obviously, $\|\tilde{\varphi}\|_1 = \|\tilde{f}_{t_0}\|_1$ and since $f_{t_0}(\mathbf{v}) = f(\mathbf{v}, t_0) > 0$, $\varphi(\mathbf{v}, t) > 0$. Consequently, since $R_d^{(n)}(\mathbf{v}, t, \mathbf{v}', t')$ is bounded for all finite n (see Eq. 4.2), $R_d^{(n)}(\mathbf{v}, t, \mathbf{v}', t')\varphi(\mathbf{v}', t')$ is integrable with respect to $(\mathbf{v}, t') \in \Delta\times[t_0, t_1]$.

Furthermore, according to Fubini's theorem and Theorem 4.2 (Sec. IV.A), we have

$$\int_\Delta d\mathbf{v} \int_{t_0}^t \int_\Delta R_d^{(n)}(\mathbf{v}, t, \mathbf{v}', t')\, \varphi(\mathbf{v}', t')\, d\mathbf{v}'\, dt'$$

$$= \int_{t_0}^t \int_\Delta \varphi(\mathbf{v}', t')\, d\mathbf{v}'\, dt' \int_\Delta R_d^{(n)}(\mathbf{v}, t, \mathbf{v}', t')\, d\mathbf{v}$$

$$< \int_{t_0}^t \int_\Delta \nu(v')\, \varphi(\mathbf{v}', t')\, d\mathbf{v}'\, dt'$$

$$= \int_\Delta f(\mathbf{v}, t_0)\, d\mathbf{v} - \int_\Delta \varphi(\mathbf{v}, t)\, d\mathbf{v} < +\infty.$$

Hence, by virtue of the theorem of Beppo Levi, the ascending sequence $\int_{t_0}^t \int_\Delta R_d^{(n)}(\mathbf{v}, t, \mathbf{v}', t')\, d\mathbf{v}'\, dt'$ converges for almost all $\mathbf{v} \in \Delta$, to a limit which is integrable with respect to $\mathbf{v}$. Moreover, since this limit is a majorant for the different terms of the preceding sequence, for almost all $\mathbf{v}$, it follows from a further application of Beppo Levi's theorem, that for almost all $\mathbf{v} \in \Delta$,

$$\lim_{n\to\infty} \int_{t_0}^t \int_\Delta R_d^{(n)}(\mathbf{v}, t, \mathbf{v}', t')\varphi(\mathbf{v}', t')\, d\mathbf{v}'\, dt'$$

$$= \int_{t_0}^t \int_\Delta R_d(\mathbf{v}, t, \mathbf{v}', t')\varphi(\mathbf{v}', t')\, d\mathbf{v}'\, dt'$$

which implies that $\lim_{n\to\infty}\varphi_n(\mathbf{v}, t) = f(\mathbf{v}, t)$, where

$$f(\mathbf{v}, t) = \varphi(\mathbf{v}, t) + \int_{t_0}^t \int_\Delta R_d(\mathbf{v}, t, \mathbf{v}', t')\varphi(\mathbf{v}', t')\, d\mathbf{v}'\, dt'. \qquad (4.38)$$

Obviously, $f(\mathbf{v}, t) > 0$. Moreover, $f(\mathbf{v}, t)$ is a solution of the integral equation (2.9) for rigid spheres. Indeed, according to Theorem 4.1, it can be shown in a similar way as in Sec. III.B, that R_d is a solution of the integral equation (3.11) where k_d must be substituted for k. Hence, if we multiply both sides of (4.38) by k_d and integrate them with respect to $(\mathbf{v}', t')$, we obtain

$$\int_{t_0}^t \int_\Delta k_d(\mathbf{v}, t, \mathbf{v}', t') f(\mathbf{v}', t')\, d\mathbf{v}'\, dt'$$

$$= \int_{t_0}^t \int_\Delta R_d(\mathbf{v}, t, \mathbf{v}', t')\varphi(\mathbf{v}', t')\, d\mathbf{v}'\, dt' = f(\mathbf{v}, t) - \varphi(\mathbf{v}, t)$$

which proves our assertion. Furthermore, if we integrate both sides of (4.38) with respect to $\mathbf{v}$ and apply Theorem 4.2, we have

$$\int_\Delta f(\mathbf{v}, t)\, d\mathbf{v}$$

$$= \int_\Delta \varphi(\mathbf{v}, t)\, d\mathbf{v} + \int_{t_0}^t \int_\Delta \nu_d(v')\varphi(\mathbf{v}', t')\, d\mathbf{v}'\, dt'$$

$$= \int_\Delta f(\mathbf{v}, t_0)\, d\mathbf{v}. \qquad (4.39)$$

Let us now assume that $g_t \in \mathcal{L}_1(\Delta)$ is another solution of (2.9) for rigid spheres, having the properties $g(v, t) > 0$ and $\int_\Delta g(\mathbf{v}, t)\, dv = \int_\Delta f(\mathbf{v}, t_0)\, d\mathbf{v}$. Then, for any positive n, we have

$$g(v, t) = \varphi(v, t) + \int_{t_0}^t \int_\Delta R_d^{(n)}(v, t, v', t')\varphi(v', t')\, dv'\, dt'$$

$$+ \int_{t_0}^t \int_\Delta k_d^{(n+1)}(\mathbf{v}, t, \mathbf{v}', t') g(\mathbf{v}' t')\, d\mathbf{v}'\, dt' \qquad (4.40)$$

and, by (4.36),

$$\int_\Delta g(\mathbf{v}, t)\, d\mathbf{v} = \int_\Delta \varphi_{n+1}(\mathbf{v}, t)\, d\mathbf{v} + \int_{t_0}^t \int_\Delta g(\mathbf{v}', t')\, d\mathbf{v}'\, dt'$$

$$\times \int_\Delta k_d^{(n+1)}(\mathbf{v}, t, \mathbf{v}', t')\, d\mathbf{v}. \qquad (4.41)$$

As stated before, the ascending sequence $\int_\Delta \varphi_{n+1}(\mathbf{v}, t)\, d\mathbf{v}$ has the limit $\int_\Delta f(\mathbf{v}, t_0)\, d\mathbf{v}$. Hence, from (4.41), it follows that

$$\lim_{n\to\infty}\left\{\int_\Delta |g(\mathbf{v}, t) - \varphi_{n+1}(\mathbf{v}, t)|\, d\mathbf{v}\right\} = \int_\Delta |g(\mathbf{v}, t) - f(\mathbf{v}, t)|\, dv = 0$$

which implies that for all $t \in [t_0, t_1]$, $g_t = f_t$.

C. Properties of the solution

The existence and uniqueness theorem given above, shows that a solution does exist for any value of the exterior electric and magnetic fields, provided that the initial distribution function is positive and summable over Δ. We will prove now, under similar initial conditions, some important properties of this solution which are directly connected with the properties of the resolvent kernel. Other properties, such as the existence of the moments of the solution, which need new assumptions on the initial distribution function, will be investigated in paper II.

Theorem 4.4: If f_{t_0} is a positive function, belonging to $\mathcal{L}_1(\Delta)$, the solution $f(\mathbf{v}, t)$ given by Eq. (4.38) is bounded on $\Delta\times[t_0, +\infty[$ for almost all $(\mathbf{v}, t)$.

Proof: We first prove the following inequality:

$$R_d(\mathbf{v}, t, \mathbf{v}', t') < k_d^{(1)}(\mathbf{v}, t, \mathbf{v}', t') + k_d^{(2)}(\mathbf{v}, t, \mathbf{v}', t') + A_1\nu(v'), \qquad (4.42)$$

where A_1 is a positive constant.

As it has been noted above, $R_d(\mathbf{v}, t, \mathbf{v}', t')$ is a solution of the integral equation (3.11), with k_d substituted to k. Hence, if we write this equation in the form

$$R_d(\mathbf{v}, t, \mathbf{v}', t')$$

$$= k_d^{(1)}(\mathbf{v}, t, \mathbf{v}', t') + k_d^{(2)}(\mathbf{v}, t, \mathbf{v}', t')$$

$$+ \int_{t'}^t \int_\Delta k_d^{(2)}(\mathbf{u}, \theta, \mathbf{v}', t') R_d(\mathbf{v}, t, \mathbf{u}, \theta)\, d\mathbf{u}\, d\theta, \qquad (4.43)$$

we see that in order to prove inequality (4.42) we have to find a majorant for the last term on the right-hand side of Eq. (4.43). Using the classical transformation

$$\int_{t'}^t \int_\Delta k_d^{(2)}(\mathbf{u}, \theta, \mathbf{v}', t') R_d(\mathbf{v}, t, \mathbf{u}, \theta)\, d\mathbf{u}\, d\theta$$

$$= \int_{t'}^t \int_\Delta k_d^{(2)}(\mathbf{v}, t, \mathbf{u}, \theta) R_d(\mathbf{u}, \theta, \mathbf{v}', t')\, d\mathbf{u}\, d\theta \qquad (4.44)$$

and applying Theorem 4.2 we obtain for almost all $(\mathbf{v}', t')$ the following majorant:

$$\int_{t'}^t \int_\Delta k_d^{(2)}(\mathbf{u}, \theta, \mathbf{v}', t') R_d(\mathbf{v}, t, \mathbf{u}, \theta)\, d\mathbf{u}\, d\theta$$

$$\leqslant \nu(v') \int_{t'}^t \left[\sup_{(\mathbf{v},\mathbf{u})\in\Delta\times\Delta} (k^{(2)}(\mathbf{v}, t, \mathbf{u}, \theta))\right] d\theta. \qquad (4.45)$$

Furthermore, according to the definition of the iterated kernels, we have

$$
\begin{aligned}
&k_d^{(2)}(\mathbf{v}, t, \mathbf{v}', t') \\
&= \int_{t'}^{t} \int_\Delta k_d(\mathbf{v}, t, \mathbf{u}, \theta) k_d(\mathbf{u}, \theta, \mathbf{v}', t')\, d\mathbf{u}\, d\theta \\
&= \int_{t'}^{t} d\theta \int_\Delta H_d(S_{\theta,t}\mathbf{v}, \mathbf{u}) H_d(S_{t',\theta}\mathbf{u}, \mathbf{v}') \\
&\quad\times \exp\left[-\int_\theta^t \nu_d(|S_{\tau,t}\mathbf{v}|)\, d\tau - \int_{t'}^{\theta} \nu_d(|S_{\tau,\theta}\mathbf{u}|)\, d\tau\right] \\
&\le \exp[-\nu(0)(t-t')] \int_{t'}^{t} d\theta \int_\Delta H_d(S_{\theta,t}\mathbf{v}, \mathbf{u}) \\
&\quad\times H_d(S_{t',\theta}\mathbf{u}, \mathbf{v}')\, d\mathbf{u}. \qquad (4.46)
\end{aligned}
$$

Using Schwarz's inequality and taking into account Eqs. (B20) and (B21), gives

$$\int_\Delta H_d(S_{\theta,t}\mathbf{v}, \mathbf{u}) H_d(S_{t',\theta}\mathbf{u}, \mathbf{v}')\, d\mathbf{u} < a_1 \left(\frac{M}{M-m}\right)^{1/2}. \qquad (4.47)$$

Hence

$$\sup_{(\mathbf{v},\mathbf{u})\in\Delta\times\Delta} [k^{(2)}(\mathbf{v}, t, \mathbf{u}, \theta)] \le a_1 \left(\frac{M}{M-m}\right)^{1/2} (t-\theta) \times \exp[-\nu(0)(t-\theta)]. \qquad (4.48)$$

Now we consider the general form of the solution $f(\mathbf{v}, t)$ given by Eq. (4.38). Since $f(\mathbf{v}, t_0)$ is positive and summable, we have $\varphi(\mathbf{v}, t) > 0$ and $\int_\Delta \varphi(\mathbf{v}, t)\, d\mathbf{v} < n$, where $n = \int_\Delta f(\mathbf{v}, t_0)\, d\mathbf{v}$. This implies that for almost all $(\mathbf{v}, t)$, $\varphi(\mathbf{v}, t) < C$, where C is a positive constant. Hence, according to Eqs. (4.42) and (B19), we have

$$f(\mathbf{v}, t) < \varphi(\mathbf{v}, t) + \frac{C}{\nu_d(0)} \left(\frac{M+m}{M-m}\right)^2 + \frac{2a_1 n}{[\nu_d(0)]^2} \left(\frac{M}{M-m}\right)^{1/2}, \qquad (4.49)$$

which implies that $f(\mathbf{v}, t)$ is bounded on $\Delta \times [t_0, +\infty[$ for almost all $(\mathbf{v}, t)$, as was to be shown.

Corollary: If f_{t_0} is a positive function belonging to $\mathcal{B}_R(\Delta) \cap \mathcal{L}_1(\Delta)$, $f(\mathbf{v}, t)$ is bounded on $\Delta \times [t_0, +\infty[$.

Proof: Follows immediately from the inequality (4.49).

BRIEF CONCLUSION TO PAPER I

The main results which have been obtained in this article are the existence and uniqueness theorems for inverse power-law potentials of the form A/r^s with $3 < s \le 5$ and for rigid spheres. Other useful results concerning the properties of the resolvent kernel of the integral equation form of the Boltzmann equation and the properties of the solutions for a wide class of initial conditions, have also been derived.

The method for dealing with rigid spheres will be extended to the case $5 < s < +\infty$, in Part II, where a complete discussion of the various results obtained and some indications on the possibilities of further developments will be given. An application to the perfectly Lorentzian case will also be studied.

ACKNOWLEDGMENTS

This work was done in the Laboratoire de Physique Théorique et Mathématique, Université de Paris VII. The author wishes to express his thanks to Professor R. Jancel for his helpful comments and his critical reading of a first draft manuscript and Professor Th. Kahan for many useful discussions.

APPENDIX A: GENERAL FORM OF THE COLLISION INTEGRAL

The collision integral on the right-hand side of (1.1) may be written in the form (see Ref. 1, pp. 258 and 565)

$$\mathcal{J}(f, f_n) = \int_\Delta \int_0^{2\pi} \int_0^{\pi/2} [f(\mathbf{r}, \mathbf{v}', t) f_n(v_n') - f(\mathbf{r}, \mathbf{v}, t) f_n(v_n)] \times B(\theta, g)\, d\mathbf{v}_n\, d\epsilon\, d\theta, \qquad (A1)$$

where the integration with respect to $\mathbf{v}_n$ is performed over the whole three-dimensional velocity space Δ and where the polar coordinates (θ, ϵ) are defined in Ref. 1, p. 565.

The velocities $\mathbf{v}'$ and $\mathbf{v}_n'$ of the particles after the interaction are respectively related to the velocities $\mathbf{v}$ and $\mathbf{v}_n$ of the same particles before the interaction by the following equations:

$$
\begin{aligned}
\mathbf{v}' &= \mathbf{v} - 2\frac{M}{M+m} \hat{k}(\hat{k} \cdot \mathbf{g}), \\
\mathbf{v}_n' &= \mathbf{v}_n + 2\frac{m}{M+m} \hat{k}(\hat{k} \cdot \mathbf{g}),
\end{aligned}
\qquad (A2)
$$

where $\hat{k}$ is a unit vector in the direction of $\mathbf{g} - \mathbf{g}'$ with $\mathbf{g} = \mathbf{v} - \mathbf{v}_n$, $\mathbf{g}' = \mathbf{v}' - \mathbf{v}_n'$, and $\mathbf{g} \cdot \hat{k} = g \cos\theta$.

APPENDIX B

1. Construction of the kernel of the integral operator *H*

By performing the variable change $\mathbf{g} = \mathbf{v} - \mathbf{v}_n$, the first term on the right-hand side of (A1) can be written in the form

$$[Hf](\mathbf{r}, \mathbf{v}, t) = -a \int_\Delta \int_0^{2\pi} \int_0^{\pi/2} \exp(-\beta_n v_n'^2) f(\mathbf{r}, \mathbf{v}', t) B(\theta, g)\, d\mathbf{g}\, d\epsilon\, d\theta, \qquad (B1)$$

where $f_n(v_n')$ has been replaced by the Maxwellian distribution $f_n(v_n') = a \exp(-\beta_n v_n'^2)$.

The transformation $\hat{k} \to -\hat{k}$ does not change the Eqs. (A2). Hence, if we state $B(\pi - \theta, g) = B(\theta, g)$, we have

$$[Hf](\mathbf{r}, \mathbf{v}, t) = -a/2 \int \exp(-\beta_n v_n'^2) f(\mathbf{r}, \mathbf{v}', t) Q(\theta, g)\, d\hat{k}\, dg, \qquad (B2)$$

where $Q(\theta, g) = B(\theta, g) \csc\theta$.

To find the kernel of H requires some manipulation. We will follow a procedure proposed by Grad[6] and applied by Cercignani[7a] (p. 70–1) to the linearized collision operator of the Boltzmann equation of neutral gases.

Let $\mathbf{g} = \mathbf{g}_1 + \mathbf{g}_2$, where $\mathbf{g}_1$ and $\mathbf{g}_2$ are respectively parallel and perpendicular to $\hat{k}$. After integrating over the plane $\mathbf{g}_2$ which is perpendicular to $\hat{k}$, we may combine the one-dimensional g_1 integration in the direction $\hat{k}$ with the integral of $\hat{k}$ over a unit sphere to give a three-dimensional integration over the components of the vector $\mathbf{g}_1 = g_1\hat{k}$. Then, Eq. (B2) can be written in the form

$$[Hf](\mathbf{r}, \mathbf{v}, t) = -a/g_1^2 \int \exp(-\beta_n v_n'^2) f(\mathbf{r}, \mathbf{v}', t)\, Q_1(\mathbf{g}_1, \mathbf{g}_2)\, d\mathbf{g}_1\, d\mathbf{g}_2, \quad \text{(B3)}$$

where $Q_1(\mathbf{g}_1, \mathbf{g}_2) = Q(\theta, g)$.

Furthermore, since $\mathbf{g}_1 = \hat{k}(\hat{k}\cdot\mathbf{g})$, the first equation (A2) allows us to introduce in (B3) the integration variable $\mathbf{v}'$ instead of $\mathbf{v}$. It follows immediately that

$$H(\mathbf{v}, \mathbf{v}') = a\frac{M+m}{2M}\frac{1}{|\mathbf{v}-\mathbf{v}'|^2}\int_{\mathrm{P}} \exp(-\beta_n v_n'^2) \times Q_1\left(\frac{M+m}{2M}|\mathbf{v}-\mathbf{v}'|, g_2\right) d\mathbf{g}_2, \quad \text{(B4)}$$

where the domain P is the plane perpendicular to $\mathbf{v}-\mathbf{v}'$ and where $\mathbf{v}_n'$ is given by

$$\mathbf{v}_n' = \mathbf{v}' + \frac{M+m}{2M}(\mathbf{v}-\mathbf{v}') - \mathbf{g}_2. \quad \text{(B5)}$$

In addition, since $\mathbf{g}_2\cdot(\mathbf{v}-\mathbf{v}') = 0$, $v_n'^2$ can be written as

$$v_n'^2 = \frac{1}{2}\left(\frac{m^2}{2M^2}|\mathbf{v}-\mathbf{v}'|^2 + 2|\mathbf{g}_2-\mathbf{w}| + \frac{m}{M}(v^2 - v'^2)\right), \quad \text{(B6)}$$

where $\mathbf{w} = \frac{1}{2}(\mathbf{v}+\mathbf{v}')$. If we state $\mathbf{w} = \mathbf{w}_1 + \mathbf{w}_2$, where $\mathbf{w}_1$ is parallel to $\mathbf{v}-\mathbf{v}'$ and where $\mathbf{w}_2$ is in the plane P, we have

$$w_1^2 = \left(\mathbf{w}\cdot\frac{(\mathbf{v}-\mathbf{v}')}{|\mathbf{v}-\mathbf{v}'|}\right)^2 = \frac{1}{4}\frac{(v^2-v'^2)^2}{|\mathbf{v}-\mathbf{v}'|^2} \quad \text{(B7)}$$

and therefore

$$H(\mathbf{v}, \mathbf{v}') = k_s(\mathbf{v}, \mathbf{v}')\exp\left(-\frac{\beta}{2}(v^2 - v'^2)\right), \quad \beta = \frac{m}{2kT}, \quad \text{(B8)}$$

where $k_s(\mathbf{v}, \mathbf{v}')$ is a symmetric kernel given by

$$k_s(\mathbf{v}, \mathbf{v}') = a\frac{M+m}{2M}\frac{1}{|\mathbf{v}-\mathbf{v}'|^2}\exp\left[-\frac{\beta_n}{4}\left(\frac{m^2}{M^2}|\mathbf{v}-\mathbf{v}'|^2 + \frac{(v^2-v'^2)}{|\mathbf{v}-\mathbf{v}'|^2}\right)\right] \times \int_{\mathrm{P}} \exp(-\beta_n|\mathbf{g}_2-\mathbf{w}_2|^2)\, Q_1\left(\frac{M+m}{2M}|\mathbf{v}-\mathbf{v}'|, \mathbf{g}_2\right) d\mathbf{g}_2. \quad \text{(B9)}$$

When $M = m$, $k_s(\mathbf{v}, \mathbf{v}')$ is equal to the kernel k_2 introduced by Grad.[6]

2. General properties of the kernel $H(\mathbf{v}, \mathbf{v}')$

Interchange of $\mathbf{v}$ and $\mathbf{v}'$ in Eq. (B8) shows that

$$H(\mathbf{v}', \mathbf{v}) = H(\mathbf{v}, \mathbf{v}')\exp[\beta(v^2 - v'^2)]. \quad \text{(B10)}$$

Furthermore, if we consider the second term on the right-hand side of (A1) and proceed in a similar way as before, we obtain

$$\nu(v) = \int f_n(v_n)\, B(\theta, g)\, d\mathbf{v}_n\, d\epsilon\, d\theta \quad \text{(B11)}$$
$$= a\frac{M+m}{2M}\int_\Delta \frac{1}{|\mathbf{v}-\mathbf{v}'|^2}d\mathbf{v}'\int_{\mathrm{P}} \exp(-\beta_n v_n^2) \times Q_1\left(\frac{M+m}{2M}(\mathbf{v}-\mathbf{v}'), \mathbf{g}_2\right) dg_2,$$

where

$$\mathbf{v}_n = \mathbf{v} - \frac{M+m}{2M}(\mathbf{v}-\mathbf{v}') - \mathbf{g}_2 \quad \text{(B12)}$$

and

$$v_n^2 = \frac{1}{2}\left(\frac{m^2}{2M^2}|\mathbf{v}-\mathbf{v}'|^2 + 2|\mathbf{g}_2-\mathbf{w}|^2 - \frac{m}{M}(v^2 - v'^2)\right). \quad \text{(B13)}$$

Comparison of (B11) and (B13) to (B4) and (B6) yields

$$\nu(v) = \int_\Delta H(\mathbf{v}', \mathbf{v})\, d\mathbf{v}'. \quad \text{(B14)}$$

3. Special cases

A. Rigid spheres

If we denote the sum of the radii of the interacting particles by D, the expression of Q_1, in the particular case of rigid spheres, is given by $Q_1 = D^2 g_1$. Hence, Eqs. (B8) and (B9) yield

$$H_d(\mathbf{v}, \mathbf{v}') = a_d\frac{1}{|\mathbf{v}-\mathbf{v}'|}\exp\left[-\frac{\beta_n}{4}\left(\frac{v^2-v'^2}{|\mathbf{v}-\mathbf{v}'|} + \frac{m}{M}|\mathbf{v}-\mathbf{v}'|^2\right)\right], \quad \text{(B15)}$$

where

$$a_d = \left(\frac{M}{2\pi kT}\right)^{1/2}\left(\frac{m+M}{M}\right)^2\sigma N, \quad \sigma = \frac{D^2}{4}.$$

Furthermore, by performing the variable change $\mathbf{u} = \mathbf{v}' - \mathbf{v}$, Eqs. (B14) and (B15) give

$$\nu_d(v) = 4\pi a_d\left(\frac{M}{M+m}\right)^2 Y(v), \quad \text{(B16)}$$

where

$$Y(v) = \left(v + \frac{1}{2\beta_n v}\right)\int_0^{2v}\exp\left(-\frac{\beta_n}{4}x^2\right)dx + \frac{1}{\beta_n}\exp(-\beta_n v^2). \quad \text{(B17)}$$

Integration of the two terms of (B15) with respect to $\mathbf{v}'$ yields in a similar way,

$$\int_\Delta H_d(\mathbf{v}, \mathbf{v}')\, d\mathbf{v}' = 4\pi a_d\left(\frac{M}{M-m}\right)^2 Y(v). \quad \text{(B18)}$$

Hence, according to (B16),

$$\int_\Delta H_d(\mathbf{v}, \mathbf{v}')\, dv' = \left(\frac{M+m}{M-m}\right)^2 \nu_d(v). \quad \text{(B19)}$$

Similarly, if we integrate the square of the term on the right-hand side of (B15), we obtain

$$\int_\Delta [H_d(\mathbf{v}', \mathbf{v})]^2 d\mathbf{v}' < \frac{1}{8\pi l^2}\left(\frac{M+m}{M}\right)^4\left(\frac{M}{\pi kT}\right)^{1/2} \equiv a_1, \quad \text{(B20)}$$

$$\int_\Delta [H_d(\mathbf{v}, \mathbf{v}')]^2 d\mathbf{v}' < \frac{M}{M-m}a_1, \quad \text{(B21)}$$

where $l = (4\pi N\sigma)^{-1}$.

Furthermore, by (B17), it can be shown that the collision frequency is monotonically increasing and that

$$\frac{1}{l}v < \nu(v) \leq \frac{1}{l}(v + \bar{v}_n), \quad \text{(B22)}$$

where $\bar{v}_n = l\nu(0) = (8kT/\pi M)^{1/2}$.

B. Power-law potentials

If we consider force laws of the type A/r^s, $s > 0$, we have

$$B(\theta, g) = g^{\gamma}\beta(\theta), \quad \gamma = (s-5)/(s-1). \tag{B23}$$

It can be shown by introducing an angular cutoff in (B23) which excludes the grazing collisions (see Grad[6]) that for $s > 3$ and for all $(\mathbf{v}', t') \in \Delta \times \Delta$, the following inequality holds:

$$k_s(\mathbf{v}, \mathbf{v}') < b k_s^{(d)}(\mathbf{v}, \mathbf{v}'), \tag{B24}$$

where $k_s^{(d)}$ denotes the symmetric kernel defined by (B9), for rigid spheres, and where b is a positive constant. According to (B8), inequality (B24) yields

$$H(\mathbf{v}, \mathbf{v}') < b H_d(\mathbf{v}, \mathbf{v}') \tag{B25}$$

and, in view of (B19), (B20), and (B21),

$$\int_\Delta H(\mathbf{v}, \mathbf{v}')\, d\mathbf{v}' < b\left(\frac{M+m}{M-m}\right)^2 \nu_d(v), \tag{B26}$$

$$\int_\Delta [H(\mathbf{v}', \mathbf{v})]^2\, d\mathbf{v}' < a_1 b^2, \tag{B27}$$

$$\int_\Delta [H(\mathbf{v}, \mathbf{v}')]^2\, d\mathbf{v}' < \frac{M}{M-m} a_1 b^2. \tag{B28}$$

Another property, which is very easy to prove (see Grad[6]), is that the collision frequency $\nu(v)$ is a monotonic function of v which is decreasing for $s < 5$ and increasing for $s > 5$. For $s = 5$, $\nu(v)$ is a constant. As a consequence, for $s \leqslant 5$,

$$\int_\Delta H(\mathbf{v}', \mathbf{v}) = d\mathbf{v}' = \nu(v) \leqslant \nu(0). \tag{B29}$$

Furthermore, according to the angular cutoff, we may write

$$a_2 < \frac{\beta(\theta)}{\sin\theta} < a_3, \tag{B30}$$

where a_2 and a_3 are positive constants. This implies

$$a_2 g_1^{\gamma} < Q_1(g_1, g_2) < a_4 g_1^{\gamma}, \quad a_4 = C^t \tag{B31}$$

and, by (B9) and $g_1 = (M+m)/(2M)(\mathbf{v} - \mathbf{v}')$,

$$a_2 I(\mathbf{v}, \mathbf{v}') < H(\mathbf{v}, \mathbf{v}') < a_4 I(\mathbf{v}, \mathbf{v}'), \tag{B32}$$

where

$$I(\mathbf{v}, \mathbf{v}') = \left(\frac{M+m}{2M}\right)^{\gamma+1} \frac{\pi a}{\beta_n} \frac{1}{|\mathbf{v} - \mathbf{v}'|^{2-\gamma}} \times \exp\left[-\frac{\beta_n}{4} \frac{(v^2 - v'^2)}{|\mathbf{v} - \mathbf{v}'|} + \frac{m}{M}|\mathbf{v} - \mathbf{v}'|^2\right]. \tag{B33}$$

In addition, according to (B14), (B32), and (B33),

$$a_2 \int_\Delta I(\mathbf{v}, \mathbf{v}')\, d\mathbf{v}' < \nu(v) < a_4 \int_\Delta I(\mathbf{v}', \mathbf{v})\, d\mathbf{v}', \tag{B34}$$

$$\int_\Delta I(\mathbf{v}, \mathbf{v}')\, d\mathbf{v}' = \left(\frac{M+m}{M-m}\right)^{\gamma+1} \int_\Delta I(\mathbf{v}', \mathbf{v})\, d\mathbf{v}'. \tag{B35}$$

Hence

$$C_1 \nu(v) < \int_\Delta H(\mathbf{v}, \mathbf{v}')\, d\mathbf{v}' < C_2 \nu(v), \tag{B36}$$

where

$$C_1 = \frac{a_2}{a_1}\left(\frac{M+m}{M-m}\right)^{\gamma+1}, \quad C_2 = \frac{a_4}{a_2}\left(\frac{M+m}{M-m}\right)^{\gamma+1}.$$

In the case $s \leqslant 5$, (B36) yields

$$\int_\Delta H(\mathbf{v}, \mathbf{v}')\, d\mathbf{v}' < C_2 \nu(0). \tag{B37}$$

*Present address: Laboratoire Central de Télécommunications, 18-20 Rue Grande Dame Rose, 78 140 Velizy-Villacoublay, France.

[1] R. Jancel and T. Kahan, *Electrodynamique des Plasmas* (Dunod, Paris, 1963).

[2] S. Chapman and T.G. Cowling, *The Mathematical Theory of Non-Uniform Gases* (Cambridge University Press, Cambridge, 1952).

[3] H. Grad, (a) *Handbuch der Physik*, edited by S. Flügge (Springer, Berlin, 1958), Vol. XII, p. 205; (b) *Rarefied Gas Dynamics*, edited by F.M. Devienne (Pergamon, New York, 1960), p. 100; (c) Phys. Fluids 6, 147 (1963); (d) Commun. Pure Appl. Math. 18, 345—54 (1965).

[4] D. Hilbert, Math. Ann. 72, 562 (1912).

[5] E. Hecke, Math. Z. 12, 274 (1918).

[6] H. Grad, *Rarefied Gas Dynamics*, edited by J.A. Laurmann (Academic, New York, 1962), Vol. 1, p. 26.

[7] C. Cercignani, (a) *Mathematical Methods in Kinetic Theory* (Plenum, New York, 1969); (b) Phys. Fluids, 10, 2097 (1967).

[8] F. Molinet, C.R. Acad. Sci. Paris, Ser. B, 267, 1257, 1305 (1968); 269, 237 (1969).

[9] F. Riesz and B. Nagy, *Leçons d'analyse Fonctionelle* (Gauthier-Villars, Paris, 1968), 5th ed.

Paraboson uniqueness for infinitely many degrees of freedom

Steven Robbins

Division of Mathematics, Computer Science and Systems Design, The University of Texas at San Antonio, San Antonio, Texas 78285
(Received 30 July 1976)

As with the boson field, the creation and annihilation operators for the paraboson field are unbounded, and the usual paraboson relations, therefore, cannot be rigorously satisfied. There are no Weyl relations for parabosons, and so the paraboson relations must be treated in unbounded form. A new approach for treating such unbounded relations is developed, and it is found that the paraboson field is determined by the requirement that the appropriate operator is the one-dimensional number operator. A similar result is shown to hold for the boson field.

1. INTRODUCTION AND STATEMENT OF MAIN RESULTS

The free boson field which is sometimes referred to as the zero-interaction boson field, the positive energy boson field, the Fock boson field or the Fock–Cook boson field, was first described by Fock[1] in 1932. The first mathematically rigorous treatment was later given by Cook[2] in 1953. The boson field is characterized by a pair of commutation relations, (1.1) and (1.2), involving the creation operator $C(z)$, and its adjoint, $C^*(z)$, the corresponding annihilation operator. The range of z is a Hilbert space H, the single particle (state) space.

$$[C^*(z), C(y)] = \langle y, z\rangle, \tag{1.1}$$

$$[C(z), C(y)] = 0. \tag{1.2}$$

The free boson field has additional structure given by a representation, Γ, of the unitary group of H by operators on K and a distinguished vector $v \in K$ which represents the vacuum or zero-particle state. The important relationships among C, Γ, and v are described in the hypotheses of Theorem 1.

Relations (1.1) and (1.2) are satisfied only in a formal sense since operators satisfying (1.1) must be unbounded. Because of the difficulty in dealing with unbounded operators,[3] uniqueness results (especially when H is infinite dimensional) are usually stated in terms of the operators[4] $W(z) = \exp[iR(z)]$, where $R(z)$ is the closure of $\sqrt{1/2}[C(z) + C^*(z)]$. The relations (1.1) and (1.2) are formally equivalent to the following relation which was introduced by Weyl[5,6] in 1927:

$$W(z)W(y) = \exp[\tfrac{1}{2}i\,\mathrm{Im}(\langle z, y\rangle)]W(z+y). \tag{1.3}$$

The operators $W(z)$ are unitary and so one need not talk about unbounded operators at all.

In 1953, Green[7] introduced the paraboson relations

$$[[C^*(z), C(y)]_+, C(x)] = 2\langle x, z\rangle C(y), \tag{1.4}$$

$$[[C(z), C(y)]_+, C(x)] = 0, \tag{1.5}$$

which generalize the relations (1.1) and (1.2) in the sense that operators formally satisfying (1.1) and (1.2) also satisfy (1.4) and (1.5). As in the boson case, these relations cannot be satisfied by bounded operators[8] and therefore cannot be rigorously satisfied. Unfortunately, there are no known formally equivalent relations which involve only bounded operators as with the Weyl relation (1.3) for bosons. The relations (1.4) and (1.5) are usually too difficult to work with mathematically, but a partial simplification can be made.

It will be shown that the operator

$$n(z) = \tfrac{1}{2}[C(z)C^*(z) + C^*(z)C(z)] \tag{1.6}$$

is self-adjoint and can be interpreted as the operator representing the "number of particles in the state z," up to an additive constant. (1.4) implies that

$$[n(z), C(y)] = \langle y, z\rangle C(z), \tag{1.7}$$

which is a necessary condition for this interpretation. [Other conditions are also necessary, such as that $n(z)$ is bounded from below and has integral spectrum when z is normalized.] For $\|z\| = 1$, Eq. (1.7) is formally equivalent to

$$\exp[itn(z)]C(y)\exp[-itn(z)] = C[\exp(itP_z)y], \tag{1.8}$$

where P_z is the projection on the one-dimensional space spanned by z. Since $\exp[itn(z)]$ is unitary for each real t, this relation is much easier to handle than (1.4). The relation (1.5) cannot be so treated, but it will be handled in a simpler manner. It will be part of the conclusion of our uniqueness result, rather than the hypothesis.

Our goal is a uniqueness result for parabosons similar to the result of Segal[9] for ordinary bosons in which the Weyl relation is replaced by a relation similar to (1.8). We first look at a new boson result along these lines. Of course, (1.8) will have to be supplemented by another relation which tells us that we actually have bosons and not some other form of parabosons. This is done by including a particular form of (1.1). For this theorem we use a slightly simpler form for $n(z)$, which is the same as the definition given in (1.6) (up to an additive constant) when (1.1) is satisfied.

Theorem 1: Suppose H and K are complex Hilbert spaces; C is a map from H into the set of closed, densely defined operators on K; Γ is a continuous unitary representation of the full unitary group of H on K; and v is a unit vector in K such that for all vectors y and z in H, all unitary operators U on H and all nonzero complex numbers γ,

$$C(z+y) \supset C(z) + C(y), \tag{1.9}$$
$$C(\gamma z) = \gamma C(z),$$

$$\Gamma(U)v = v, \tag{1.10}$$

$$\Gamma(U)C(z)\Gamma(U)^{-1} = C(Uz), \tag{1.11}$$

$$C^*(z)C(z) = C(z)C^*(z) + \|z\|^2, \tag{1.12}$$

$$d\Gamma \geq 0,$$

v is cyclic for the algebra generated by the $C(z)$ and $C^*(z)$ as z ranges through H.

Suppose further that the operator $n(z)$ defined by

$$n(z) = C(z)C^*(z) \tag{1.13}$$

satisfies

$$\exp[itn(z)]C(y)\exp[-itn(z)] = C[\exp(itP_z)y] \tag{1.14}$$

when z is a unit vector and that v is an analytic vector for each $n(z)$. Then $\{H, C, K, \Gamma, v\}$ is unitarily equivalent to the free boson field.

$d\Gamma \geq 0$ is interpreted as follows. If A is a (possibly unbounded) self-adjoint operator on H, then $\exp(itA)$ is a continuous one-parameter unitary group on H and since Γ is a continuous unitary representation, $\Gamma[\exp(itA)]$ is a continuous one-parameter unitary group on K. By Stone's theorem, there exists a self-adjoint operator on K, denoted by $d\Gamma(A)$, such that $\Gamma[\exp(itA)] = \exp[itd\Gamma(A)]$. We write $d\Gamma \geq 0$ when $A \geq 0$ implies $d\Gamma(A) \geq 0$. "v is cyclic for the algebra generated by the $C(z)$ and $C^*(z)$" means that if $\mathcal{A}$ is the algebra with unit generated by all the operators in the form $C(z)$ and $C^*(z)$ as z ranges through H, then $\mathcal{A}v$ is a dense subset of K. It will be shown that the other hypotheses of the theorem imply that v is in fact in the domain of all operators in $\mathcal{A}$. Analytic vectors were first introduced by Nelson[10] in 1959. v is an analytic vector for $n(z)$ means that the series expansion of $\exp[itn(z)]$ when applied to v term by term converges absolutely for sufficiently small values of t. The relation (1.12) states that $C^*(z)C(z)$ and $C(z)C^*(z)$ have the same domain and that on this domain their difference is the scalar $\|z\|^2$. Notice that it is not necessary to assume any form of the relation (1.2).

There are only three differences in the hypotheses of the boson and paraboson theorems we will prove. Two of these are minor. The difference in the definition of $n(z)$ has already been noted. The paraboson theorem needs the extra hypothesis that the single particle space H is infinite dimensional. This additional hypothesis is actually necessary since even in the one-dimensional case spurious representations are possible.[11] It is interesting that the infinite dimensional case is in this sense more regular than the finite dimensional case.

The third difference, and the only one of consequence, is that the relation (1.12) is replaced by the condition that $C^*(z)C(z)$ and $C(z)C^*(z)$ commute, a fact which follows from (1.12). Thus the major difference in the hypotheses of the two theorems is the weakening of condition (1.12).

Since there is not one free paraboson field, but a countable collection of them, the conclusion of Theorem 2 is not one of uniqueness, but that the given field can be expressed in terms of the different free paraboson fields.

Theorem 2: Suppose H and K are complex Hilbert spaces with H infinite dimensional; C is a map from H into the set of closed, densely defined operators on K; Γ is a continuous unitary representation of the full unitary group of H on K; and v is a unit vector in K such that for all vectors y and z in H, all unitary operators U on H, and all nonzero complex numbers γ,

$$C(z+y) \supset C(z) + C(y),$$

$$C(\gamma z) = \gamma C(z),$$

$$\Gamma(U)v = v, \tag{1.15}$$

$$\Gamma(U)C(z)\Gamma(U)^{-1} = C(Uz) \tag{1.16}$$

$$C^*(z)C(z) \text{ and } C(z)C^*(z) \text{ commute}, \tag{1.17}$$

$$d\Gamma \geq 0,$$

v is cyclic for the algebra generated by the $C(z)$ and $C^*(z)$ as z ranges through H.

Suppose further that the operator $n(z)$ defined by

$$n(z) = \tfrac{1}{2}[C^*(z)C(z) + C(z)C^*(z)] \tag{1.18}$$

satisfies

$$\exp[itn(z)]C(y)\exp[-itn(z)] = C[\exp(itP_z)y] \tag{1.19}$$

when z is a unit vector, and that v is an analytic vector for each $n(z)$. Then $\{H, C, K, \Gamma, v\}$ is unitarily equivalent to a direct sum of free paraboson fields of distinct orders.

In Sec. 2 we give a proof of Theorem 1. Section 3 gives the definition of the free paraboson field of order p and here we show that the free paraboson fields satisfy the hypotheses of Theorem 2. Section 4 is devoted to the proof of Theorem 2. Slightly stronger results are discussed in Sec. 5.

2. THE FREE BOSON FIELD

That the free boson field satisfies the hypotheses of Theorem 1 was first proved by Cook.[12] v is analytic for $n(z)$ since $n(z)v = 0$. We will now prove Theorem 1. $\{H, C, K, \Gamma, v\}$ denotes the collection satisfying the hypotheses of Theorem 1 and $\{H, C_0, K_0, \Gamma_0, v_0\}$ denotes the free boson field over H. We will construct a Hilbert space isomorphism $\phi: K \to K_0$ such that

$$\phi v = v_0,$$

$$\phi\Gamma(U)\phi^{-1} = \Gamma_0(U),$$

and

$$\phi C(z)\phi^{-1} = C_0(z). \tag{2.1}$$

Let $\mathcal{A}$ be the algebra (with unit) generated by the operators in the form $C(x)$ or $C^*(y)$ where x and y are arbitrary elements of H, and for $z \in H$, let $\mathcal{A}_z$ be the subset of $\mathcal{A}$ consisting of those elements for which the x and y are restricted to be either parallel or orthogonal to z. $\mathcal{A}'$ will represent the subset of $\mathcal{A}$ whose elements do not involve any annihilation operators.

$v \in \mathrm{Dom}(n(z))$ so $v \in \mathrm{Dom}(C^*(z))$.

$$\exp[it\, d\Gamma(P_z)]C^*(z)v = C^*[\exp(itP_z)z]v = \exp(-it)C^*(z)v$$

so

$$d\Gamma(P_z)C^*(z)v = -C^*(z)v,$$

and since $d\Gamma(P_z) \geq 0$,

$$C^*(z)v = 0. \tag{2.2}$$

Thus, $n(z)v = 0$. Since this implies that v is (trivially) an analytic vector for $n(z)$, the hypothesis that v is an analytic vector for each $n(z)$ can be replaced by the weaker condition that for each z, v is in the domain of $C^*(z)$. In fact, (1.10) and (1.11) imply that it is sufficient to require that v is in the domain of $C^*(z)$ for some nonzero z.

We next show that each element of $\mathcal{A}$ has v in its domain. It is sufficient to show this for $\mathcal{A}_z$ since by (1.9), each element of $\mathcal{A}$ extends an element of $\mathcal{A}_z$.

In fact, if A_z is a monomial in $\mathcal{A}_z$, using (1.14) it follows by induction on the length of A_z that $A_z v$ is an eigenvector of $n(z)$. It is therefore in the domain of $C^*(z)$ and by (1.12) also in the domain of $C(z)$.

Let $\mathcal{D} = \mathcal{A}v$. $\mathcal{D}$ is invariant under each $C(z)$ and $C^*(z)$ and is dense since v is cyclic. For each z, every element of $\mathcal{D}$ is a finite linear combination of vectors in the form $A_z v$ with $A_z \in \mathcal{A}_z$. Thus, each member of $\mathcal{D}$ is a finite linear combination of eigenvectors of $n(z)$ and therefore is an analytic vector for $n(z)$. If $w \in \mathcal{D}$, then polarization of (1.12) gives

$$C^*(z)C(y)w = C(y)C^*(z)w + \langle y, z\rangle w. \tag{2.3}$$

If $A \in \mathcal{A}$, then (2.2) and (2.3) can be used to eliminate all annihilation operators from Av and express Av as a vector in the form $A'v$ with $A' \in \mathcal{A}'$. If $A' \in \mathcal{A}'$ and A' is a monomial, (1.10) and (1.11) imply that $A'v$ (if it is not zero) is an eigenvector of $d\Gamma(I)$ with eigenvalue equal to the number of terms in the product A'. Thus, $A'v$ is orthogonal to v unless A' is just a multiple of the identity. This shows that $\langle Av, v\rangle$ is determined by (2.2) and (2.3) whenever $A \in \mathcal{A}$. If A_1 and A_2 are in $\mathcal{A}$, and $A_2^\dagger$ represents the "formal" adjoint of A_2 (the product of the adjoints of the terms of A_2 in reverse order), then

$$\langle A_1 v, A_2 v\rangle = \langle A_2^\dagger A_1 v, v\rangle.$$

Since $A_2^\dagger A_1 \in \mathcal{A}$, $\langle A_1 v, A_2 v\rangle$ is determined. Thus, if w_1 and w_2 are elements of $\mathcal{D}$, $\langle w_1, w_2\rangle$ is uniquely determined by (2.2) and (2.3).

We are now ready to define ϕ. For $A \in \mathcal{A}$, let A_0 be the operator on K_0 obtained from A by replacing each $C(z)$ by $C_0(z)$ and each $C^*(z)$ by $C_0^*(z)$.

We then define

$$\phi(Av) = A_0 v_0$$

and extend ϕ to $\mathcal{D}$ by linearity. ϕ is well defined because it preserves inner products, as the discussion above has shown. Thus, ϕ has a unique extension to $K = \overline{\mathcal{D}}$ (the closure of $\mathcal{D}$), giving a Hilbert space isomorphism. Clearly, $\phi v = v_0$ and $\phi\Gamma(U) = \Gamma_0(U)\phi$ since $\phi\Gamma(U)$ and $\Gamma_0(U)\phi$ agree on the dense set $\mathcal{D}$. Instead of establishing (2.1) directly we show that

$$C^*(z) = \phi^{-1} C_0^*(z)\phi.$$

Here we must be careful since $C^*(z)$ is an unbounded operator so it is not necessarily sufficient to show that these two operators agree on a dense set. However, the set $\mathcal{D}$ is quite special since it is a dense set of analytic vectors for each $n(z)$ and thus also for $\sqrt{n(z)}$. $\mathcal{D}$ is therefore a core for $\sqrt{n(z)}$. [A core of a closed, densely defined operator, T, is a dense subset of $\mathrm{Dom}(T)$ such that the restriction of T to the core has closure equal to T.] $\phi^{-1}n_0(z)\phi$ and $n(z)$ agree on $\mathcal{D}$ which is a dense set of analytic vectors for $n(z)$ so

$$\phi^{-1}n_0(z)\phi = n(z).$$

Thus,

$$\phi^{-1}\sqrt{n_0(z)}\phi = \sqrt{n(z)}.$$

If $C^*(z) = VS$ and $C_0^*(z) = V_0 S_0$ are the polar decompositions, then

$$S = [C(z)C^*(z)]^{1/2} = \sqrt{n(z)}$$

and

$$S_0 = [C_0(z)C_0^*(z)]^{1/2} = \sqrt{n_0(z)}.$$

$\phi^{-1}C_0^*(z)\phi$ has polar decomposition

$$(\phi^{-1}V_0\phi)(\phi^{-1}S_0\phi) = (\phi^{-1}V_0\phi)S.$$

$C^*(z)$ and $\phi^{-1}C_0^*(z)\phi$ agree on $\mathcal{D}$ which is a core for S. $\mathcal{D}$ is a common core for $C^*(z)$ and $\phi^{-1}C_0^*(z)\phi$ by the following lemma. Thus,

$$C^*(z) = \phi^{-1}C_0^*(z)\phi.$$

The proof of the lemma completes the proof of Theorem 1.

Lemma 2.1: Suppose S is self-adjoint, V is bounded and VS is closed. If $\mathcal{D}$ is a core for S, then $\mathcal{D}$ is also a core for VS.

Proof: Let $S_1 = S|_{\mathcal{D}}$ so $(VS)|_{\mathcal{D}} = VS_1$. $VS_1 \subset VS$ so $\overline{VS_1} \subset VS$.

$$(VS_1)^* = S_1^* V^*$$

since V is bounded, and

$$(S_1^* V^*)^* \supset V^{**}S_1^{**} = V\overline{S}_1 = VS,$$

$$\overline{VS_1} = (VS_1)^{**} = (S_1^* V^*)^* \supset V^{**}S_1^{**} = V\overline{S}_1 = VS.$$

Thus $\overline{VS_1} = VS$ so $\mathcal{D}$ is a core for VS. ∎

3. THE FREE PARABOSON FIELD

We now construct the free paraboson field of order p and show that it satisfies the hypotheses of Theorem 2. Let $\{H, B, \tilde{K}, \tilde{\Gamma}, \tilde{v}\}$ be the free boson field over H. We use $B(z)$ to represent the creation operators which were represented by $C(z)$ in Sec. 2. Let K_0 be a finite dimensional Hilbert space with bounded self-adjoint operators A_α, $1 \leq \alpha \leq p$, satisfying

$$[A_\alpha, A_\beta]_+ = 2\delta_{\alpha\beta},$$

as in the Clifford algebra. Let v_0 be a fixed unit vector in K_0. Define

$$\tilde{K}' = K_0 \otimes \left(\bigotimes_{\alpha=1}^{p} \tilde{K}\right),$$

$$\Gamma'(U) = I \otimes \left(\bigotimes_{\alpha=1}^{p} \tilde{\Gamma}(U)\right),$$

$$v = v_0 \otimes \left(\bigotimes_{\alpha=1}^{p} \tilde{v} \right),$$

$$B_\alpha(z) = A_\alpha \otimes (I \otimes \cdots \otimes I \otimes B(z) \otimes I \otimes \cdots \otimes I),$$

where there are p terms in parentheses and $B(z)$ occurs in the α position.[13]

v is in the domain of each polynomial in the operators $B_\alpha(z)$ and $B_\alpha^*(z)$. Let $\mathcal{D}'$ be the set of all vectors in $\tilde{K}'$ which are in the form of some such polynomial applied to v. If $\alpha \neq \beta$, then on $\mathcal{D}'$,

$$[B_\alpha(z), B_\alpha(y)] = 0, \tag{3.1}$$

$$[B_\alpha^*(z), B_\alpha(y)] = \langle y, z \rangle, \tag{3.2}$$

$$[B_\alpha(z), B_\beta(y)]_+ = 0, \tag{3.3}$$

$$[B_\alpha^*(z), B_\beta(y)]_+ = 0, \tag{3.4}$$

and

$$B_\alpha^*(z) v = 0. \tag{3.5}$$

Let $K' = \bar{\mathcal{D}}'$. Since

$$\Gamma'(U) B_\alpha(z) \Gamma'(U)^{-1} = B_\alpha(Uz)$$

and

$$\Gamma'(U) v = v,$$

$\mathcal{D}'$ is invariant under $\Gamma'(U)$ and so is K'. We will now consider the operators $B_\alpha(z)$ and $\Gamma'(U)$ as operators on K'.

Let

$$C'(z) = \overline{B_1(z) + B_2(z) + \cdots + B_p(z)}.$$

Relations (3.1)–(3.4) imply that on D',

$$[[C'^*(x), C'(y)]_+, C'(z)] = 2\langle z, x \rangle C'(y), \tag{3.6}$$

$$[[C'(x), C'(y)]_+, C'(z)] = 0.$$

These are Green's paraboson relations. We also have

$$C'^*(z) v = 0, \tag{3.7}$$

$$C'^*(y) C'(z) v = p \langle z, y \rangle v. \tag{3.8}$$

Let $\mathcal{D}$ be the set of all vectors in K' which have the form of some polynomial in the operators $C'(z)$ and $C'^*(z)$ applied to v and let K be the closure of $\mathcal{D}$. From (3.6), (3.7), and (3.8) it follows that vectors in $\mathcal{D}$ can be expressed as the image of v under polynomials not involving any $C'^*(z)$.

Since

$$\Gamma'(U) C'(z) \Gamma'(U)^{-1} = C'(Uz),$$

$\mathcal{D}$, and thus K, are invariant under $\Gamma'(U)$. Let $\Gamma(U)$ be the restriction of $\Gamma'(U)$ to K. The restriction of $C'(z)$ to $\mathcal{D}$ (as an operator on K) has $\mathcal{D}$ in the domain of its adjoint, and is thus closable. Let $C(z)$ be its closure. $\{H, C, K, \Gamma, v\}$ is the free paraboson field of order p.

Define

$$n'(z) = \tfrac{1}{2}[C'^*(z), C'(z)]_+,$$

$$n(z) = \tfrac{1}{2}[C^*(z), C(z)]_+.$$

On $\mathcal{D}'$,

$$[n'(z), B_\alpha(y)] = \langle y, z \rangle B_\alpha(z),$$

$$[n'(z), B_\alpha^*(y)] = -\langle z, y \rangle B_\alpha^*(z),$$

and

$$n'(z) v = \tfrac{1}{2} p v.$$

If $\|z\| = 1$, then also on $\mathcal{D}'$,

$$[d\Gamma'(P_z), B_\alpha(y)] = \langle y, z \rangle B_\alpha(z),$$

$$[d\Gamma'(P_z), B_\alpha^*(y)] = -\langle z, y \rangle B_\alpha^*(z),$$

and

$$d\Gamma'(P_z) v = 0.$$

Thus, on $\mathcal{D}'$, $d\Gamma'(P_z)$ and $n'(z) - \frac{1}{2}p$ agree. $\mathcal{D}'$ is a set of finite linear combinations of eigenvectors of $d\Gamma'(P_z)$ and so $\mathcal{D}'$ is a dense set of analytic vectors of $d\Gamma'(P_z)$, and thus of $n'(z)$. Therefore, $n'(z)$ is essentially self-adjoint on $\mathcal{D}'$ and the closure of $n'(z)$ is $d\Gamma'(P_z) + \frac{1}{2}p$. Similarly, $n(z)$ is essentially self-adjoint on $\mathcal{D}$ and the closure of $n(z)$ is $d\Gamma(P_z) + \frac{1}{2}p$. In fact, $n(z)$ and $n'(z)$ are already closed as we shall see.

On $\mathcal{D}'$,

$$[d\Gamma'(I), B_\alpha(z)] = B_\alpha(z).$$

Thus, the spectrum of $d\Gamma'(I)$ is contained in the set of nonnegative integers. This is similar for $d\Gamma(I)$. Let

$$\mathcal{D}'_k = \{w \in K : d\Gamma'(I) w = kw\},$$

$$\mathcal{D}_k = \{w \in K : d\Gamma(I) w = kw\},$$

$$\mathcal{D}'_{k,0} = \mathcal{D}'_k \cap \mathcal{D}',$$

$$\mathcal{D}_{k,0} = \mathcal{D}_k \cap \mathcal{D}.$$

$\mathcal{D}'_{k,0}$ is a dense subset of $\mathcal{D}'_k$ and the linear span of the union of the sets $\mathcal{D}'_{k,0}$ is $\mathcal{D}'$. This is similar for $\mathcal{D}_{k,0}$, $\mathcal{D}_k$, and $\mathcal{D}$.

In the following, z is a fixed unit vector. Let $C' = C'(z)$ and $C = C(z)$. We note that C'^*C' and $C'C'^*$ are bounded operators from $\mathcal{D}'_k$ into $\mathcal{D}'_k$ while C^*C and CC^* are bounded operators from $\mathcal{D}_k$ into $\mathcal{D}_k$. We exhibit the proof of this for C^*C.

It is sufficient to show that $(C^*C)^{1/2}$ is bounded on $\mathcal{D}_k$. Assume $w \in \mathcal{D}_{k,0}$. Since $w \in \mathcal{D}$,

$$2d\Gamma(P_z) w + pw = C^*Cw + CC^*w,$$

$$2\langle d\Gamma(P_z) w, w \rangle + p\|w\|^2 = \langle C^*Cw, w \rangle + \langle CC^*w, w \rangle$$

$$2\langle d\Gamma(P_z) w, w \rangle + p\|w\|^2 \geq \langle C^*Cw, w \rangle.$$

Since $d\Gamma(P_z) \leq d\Gamma(I)$,

$$(2k + p)\|w\|^2 \geq \langle C^*Cw, w \rangle = \|(C^*C)^{1/2} w\|^2.$$

Therefore, $(C^*C)^{1/2}$ is bounded by $\sqrt{2k+p}$ on $\mathcal{D}_{k,0}$ and since $\mathcal{D}_{k,0}$ is dense in $\mathcal{D}_k$, $(C^*C)^{1/2}$ is bounded on $\mathcal{D}_k$. Thus, each vector in $\mathcal{D}_k$, and so each vector in $\mathcal{D}$, is an analytic vector for C^*C and CC^*. It follows from (3.6) that for vectors in $\mathcal{D}$, C^*C and CC^* commute. Thus they commute on a dense, invariant set of analytic vectors. This implies that they are commuting self-adjoint operators. Since C^*C and CC^* commute and are positive, their sum is self-adjoint. Thus $n(z)$ is closed and

$$n(z) - \tfrac{1}{2}p = d\Gamma(P_z).$$

Similarly,

$n'(z) - \frac{1}{2}p = d\Gamma'(P_z)$.

Let P_k' be the projection on $\mathcal{D}_k'$ and P_k the projection on $\mathcal{D}_k$. For $w \in K'$, let $w_k = P_k'w$. If $w \in K$, then $w_k = P_k w$. On $\mathcal{D}'$,

$$[d\Gamma'(I), C'(z)] = C'(z),$$

$$[d\Gamma'(I), C'^*(z)] = -C'^*(z).$$

If $u \in K'$,

$$P_{k+1}'C'P_k'u = C'P_k'u, \quad P_k'C'^*P_{k+1}'u = C'^*P_{k+1}'u.$$

From these it follows that

$$\langle P_{k+1}'C'w, u\rangle = \langle C'P_k'w, u\rangle$$

so

$$P_{k+1}'C'w = C'P_k'w = C'w_k.$$

If $w \in \mathrm{Dom}(C')$, $C'w = \sum_k C'w_k$. If w is also in K, then $w_k \in \mathcal{D}_k$ and since C and C' agree on $\mathcal{D}_k$, $Cw_k = C'w_k$. Thus, $\sum_k Cw_k$ converges so $w \in \mathrm{Dom}(C)$. This shows that if $w \in \mathrm{Dom}(C') \cap K$, then $w \in \mathrm{Dom}(C)$ and C is the restriction of C' to $\mathrm{Dom}(C') \cap K$.

Relations (3.1), (3.2), and (3.5) imply that for each fixed α, $\{B_\alpha, K'\}$ is isomorphic to a direct sum of free boson fields and the relations

$$B_\alpha(z+y) \supset B_\alpha(z) + B_\alpha(y), \quad B_\alpha(\gamma z) = \gamma B_\alpha(z) \quad \text{if} \quad \gamma \neq 0$$

follow from the corresponding properties of free boson fields.[14] Thus

$$C'(z+y) \supset C'(z) + C'(y), \quad C'(\gamma z) = \gamma C'(z) \quad \text{if} \quad \gamma \neq 0$$

so

$$C(z+y) \supset C(z) + C(y), \quad C(\gamma z) = \gamma C(z) \quad \text{if } \gamma \neq 0.$$

$n(z)$ satisfies (1.19) because $d\Gamma(P_z)$ does and v is analytic for $n(z)$ because $n(z)v = \frac{1}{2}pv$.

We have now shown that the free paraboson field of order p satisfies the hypotheses of Theorem 2.

4. PROOF OF THEOREM 2

Suppose $\{H, C, K, \Gamma, v\}$ satisfies the hypotheses of Theorem 2. $n(z)$ is self-adjoint because $C^*(z)C(z)$ and $C(z)C^*(z)$ are commuting positive self-adjoint operators. Let $\mathcal{A}$ be the algebra of all polynomials in the creation and annihilation operators. Our first goal is to show that the vectors in $\mathcal{A}v$ are analytic vectors for each $n(z)$. Suppose z is a fixed unit vector in H. Let $C = C(z)$ and $n = n(z)$. Let $n = \int \lambda E(d\lambda)$ be the spectral resolution of n.

Lemma 4.1: Suppose Δ is a bounded interval of real numbers and $E(\Delta)w = w$. Then for each nonnegative integer k, $Cw \in \mathrm{Dom}(n^k)$, $C^*w \in \mathrm{Dom}(n^k)$ and we have

$$n^kCw = C(n+1)^kw, \tag{4.1}$$

$$E(\Delta+1)Cw = Cw, \tag{4.2}$$

$$n^kC^*w = C^*(n-1)^kw, \tag{4.3}$$

$$E(\Delta-1)C^*w = C^*w. \tag{4.4}$$

Proof: Since $E(\Delta)w = w$, $w \in \mathrm{Dom}(n)$ so $w \in \mathrm{Dom}(C)$. C is bounded on $E(\Delta)K$ since

$$\|Cw\|^2 = \langle C^*Cw, w\rangle \leq 2\langle nw, w\rangle \leq 2(\sup\Delta)\|w\|^2,$$

$$\exp(itn)Cw = \exp(itn)C\exp(-itn)\exp(itn)w$$
$$= \exp(it)C\exp(itn)w,$$

$$\frac{\exp(itn)-1}{it}Cw = \exp(it)C\left(\frac{\exp(itn)-1}{it}\right)w + \left(\frac{\exp(it)-1}{it}\right)Cw.$$

Since

$$\left(\frac{\exp(itn)-1}{it}\right)w \in E(\Delta)K$$

and C is bounded on $E(\Delta)K$,

$$\lim_{t\to 0} C\left(\frac{\exp(itn)-1}{it}\right)w = C\lim_{t\to 0}\left(\frac{\exp(itn)-1}{it}\right)w = Cnw.$$

Thus,

$$\lim_{t\to 0}\left(\frac{\exp(itn)-1}{it}\right)Cw = Cnw + Cw = C(n+1)w,$$

so $Cw \in \mathrm{Dom}(n)$ and $nCw = C(n+1)w$.

Let $K_0 = \{u \in K : E(\Delta')u = u$ for some bounded interval $\Delta'\}$. If $u \in K_0$,

$$\langle nu, C^*w\rangle = \langle Cnu, w\rangle = \langle (n-1)Cu, w\rangle$$
$$= \langle u, C^*(n-1)w\rangle.$$

Since n is essentially self-adjoint on K_0, $C^*w \in \mathrm{Dom}(n)$ and $nC^*w = C^*(n-1)w$. This establishes (4.1) and (4.3) for $k=1$. The proof for general k is accomplished by induction; the induction step to prove (4.1) is almost identical to the proof for $k=1$. The important points here are that $n^kw \in E(\Delta)K$ and n^kC is bounded on $E(\Delta)K$.

Let K_d be the closed subspace of K generated by the eigenvectors of n. (4.1) and (4.3) imply that if $w \in K_d$, then (4.2) and (4.4) hold and K_d is invariant under C and C^*. Now suppose $w \in K_d^\perp$. Under this condition, (4.2) and (4.4) have been proved[15] using slightly different hypotheses. We give a simpler proof here based on the following fact about self-adjoint operators.

Lemma 4.1a: Let T be a self-adjoint operator with spectral resolution $T = \int \lambda E(d\lambda)$. Suppose that $E([0,\infty))x = x$ and for each positive integer k, $x \in \mathrm{Dom}(T^k)$. Let

$$b = \lim_{k\to\infty}\|A^kx\|^{1/k} \quad \text{and} \quad a = b - \lim_{k\to\infty}\|(b-A)^kx\|^{1/k}.$$

Then $[a,b]$ is the smallest closed interval such that $E([a,b])x = x$.

The proof of this is straightforward and is similar to the proof that on a finite measure space the L_p norms of a function approach the L_∞ norm.

Suppose $w \in K_d^\perp$, $\Delta = (a,b)$, and $E(\Delta)w = w$. Then

$$\|n^kCw\| = \|C(n+1)^kw\|$$
$$= \|(C^*C)^{1/2}(n+1)^kw\|$$
$$= \|(n+1)^k(C^*C)^{1/2}w\|.$$

Since $(C^*C)^{1/2}$ commutes with $n+1$, Lemma 4.1a implies that

$$\lim\|(n+1)^k(C^*C)^{1/2}w\|^{1/k} \le \lim\|(n+1)^k w\|^{1/k} \le b+1.$$

Thus, for some $\delta \ge 0$,

$$\lim\|n^k Cw\|^{1/k} = b+1-\delta \le b+1.$$

Similarly,

$$\|(b+1-\delta-n)^k Cw\| = \|(b-\delta-n)^k (C^*C)^{1/2}w\|$$

and

$$\begin{aligned}\lim\|(b-\delta-n)^k(C^*C)^{1/2}w\|^{1/k} &= \lim\|(b-n)^k \\ &\quad\times (C^*C)^{1/2}w\|^{1/k} - \delta \\ &\le \lim\|(b-n)^k w\|^{1/k} - \delta \\ &\le b-a-\delta.\end{aligned}$$

Thus,

$$\begin{aligned}&b+1-\delta-\lim\|(b+1-\delta-n)^k Cw\|^{1/k} \\ &\quad \ge b-\delta+1-(b-a-\delta) = a+1,\end{aligned}$$

so $E([a+1, b+1])Cw = Cw$. Since $Cw \in K_a^\perp$,

$$E((a+1, b+1))Cw = Cw.$$

Thus, (4.2) holds and a similar argument yields (4.4). ∎

Lemma 4.2: If $w \in \mathrm{Dom}(n^{k+1/2})$, then $Cw \in \mathrm{Dom}(n^k)$, $C^*w \in \mathrm{Dom}(n^k)$,

$$n^k Cw = C(n+1)^k w,$$

and

$$n^k C^* w = C^*(n-1)^k w.$$

Proof: Let $Q_j = E([j-1, j))$ and $w_j = Q_j w$. We first show that $\sum_j C(n+1)^k w_j$ converges.

$$\langle C(n+1)^k w_i, C(n+1)^k w_j\rangle = \langle C^*C(n+1)^k w_i, (n+1)^k w_j\rangle,$$

which is 0, if $i \ne j$, since C^*C commutes with n.

$$\begin{aligned}\|C(n+1)^k w_j\|^2 &= \langle C^*C(n+1)^k w_j, (n+1)^k w_j\rangle \\ &\le 2(j+1)^{2k+1}\|w_j\|^2.\end{aligned}$$

Since $w \in \mathrm{Dom}(n^{k+1/2})$, $Q_j n^{k+1/2} w = n^{k+1/2} w_j$, and so $\sum_j \|n^{k+1/2} w_j\|^2$ converges.

$$\|n^{k+1/2} w_j\|^2 \ge (j-1)^{2k+1}\|w_j\|^2,$$

and so $\sum_j (j-1)^{2k+1}\|w_j\|^2$ converges, which implies that $\sum_j (j+1)^{2k+1}\|w_j\|^2$ converges. Since $\sum_j (n+1)^k w_j$ converges (by a similar argument),

$$(n+1)^k w = \sum_j (n+1)^k w_j$$

and

$$C(n+1)^k w = \sum_j C(n+1)^k w_j.$$

By Lemma 4.1,

$$C(n+1)^k w_j = n^k C w_j$$

and so $\sum_j n^k C w_j$ converges. Since $\sum_j C w_j$ converges, $Cw = \sum_j C w_j$, $Cw \in \mathrm{Dom}(n^k)$ and

$$n^k Cw = \sum_j n^k C w_j.$$

Thus, $n^k Cw = C(n+1)^k w$. A similar argument works for C^*. ∎

Lemma 4.3: If w is an analytic vector for n, then so are Cw and C^*w.

Proof: Assume w is an analytic vector for n. Since $w \in \mathrm{Dom}(n^k)$ for every k, by Lemma 4.2,

$$\begin{aligned}n^k Cw &= C(n+1)^k w, \\ \|n^k Cw\| &= \|C(n+1)^k w\| \\ &= \|(C^*C)^{1/2}(n+1)^k w\| \\ &\le \sqrt{2}\|(n+1)^{k+1/2} w\| \\ &\le \sqrt{2}\|(n+1)^{k+1} w\|.\end{aligned}$$

Thus,

$$\sum_k \frac{\|n^k Cw\| t^k}{k!}$$

converges since

$$\sum_k \frac{\sqrt{2}\,\|(n+1)^{k+1} w\| t^k}{k!}$$

converges. The same reasoning can be applied to show that C^*w is an analytic vector for n. ∎

Our next goal is to show that a certain subset of the analytic vectors of $n(z)$ is invariant under $C(y)$ and $C^*(y)$ when y is orthogonal to z. Assume now that y and z are orthogonal unit vectors in H and let $C = C(z)$, $n = n(z)$, $D = C(y)$, and $m = n(y)$. Let $n = \int \lambda E(d\lambda)$ and $m = \int \lambda F(d\lambda)$ be the spectral resolutions.

Lemma 4.4: If w is analytic for n and $w \in \mathrm{Dom}(m)$, then Dw and D^*w are analytic for n.

Proof: (1.19) implies that n, m, and D^*D are commuting self-adjoint operators. Let

$$\begin{aligned}Q_j &= E([j-1, j)) \\ R_j &= F([j-1, j)), \\ P_{ij} &= Q_i R_j.\end{aligned}$$

Q_i and R_j commute, so P_{ij} is a projection. Let $w_{ij} = P_{ij} w$, so that $w = \sum_{ij} w_{ij}$. We first show that $\sum_{ij} D w_{ij}$ converges. The terms in this summation are orthogonal since

$$\langle D w_{ij}, D w_{kl}\rangle = \langle w_{ij}, D^* D w_{kl}\rangle$$

and D^*D commutes with P_{ij}. As before

$$\|D w_{ij}\|^2 \le 2j\|w_{ij}\|^2$$

and $\sum j\|w_{ij}\|^2$ converges since $w \in \mathrm{Dom}(\sqrt{m})$. Thus $Dw = \sum D w_{ij}$.

We now show that $Q_i D w_{ij} = D w_{ij}$. DP_{ij} is a bounded operator with bound $\sqrt{2j}$.

$$\exp(itn) DP_{ij} \exp(-itn) = \exp(itn) D \exp(-itn) P_{ij} = DP_{ij}.$$

DP_{ij} commutes with n and so it commutes with Q_i,

$$D w_{ij} = DP_{ij} w = DP_{ij} Q_i w = Q_i DP_{ij} w = Q_i D w_{ij}.$$

Our next step is to show that for each k, $\sum_{ij} n^k D w_{ij}$ converges. Since $Q_i D w_{ij} = D w_{ij}$, $D w_{ij} \in \mathrm{Dom}(n^k)$. The terms in the summation are orthogonal and

$$\|n^k D w_{ij}\|^2 \le i^{2k}(2j)\|w_{ij}\|^2.$$

If N is the set of natural numbers and μ is counting measure on $N \times N$, then since $w \in \mathrm{Dom}((n+1)^{2k})$, the double sequence $\{i^{2k}\|w_{ij}\|\}$ is in $L_2(N \times N, \mu)$. Here, k is fixed,

i and j are the indices of the sequence. Since $w \in \mathrm{Dom}(m+1)$, $\{j\|w_{ij}\|\}$ is also in $L_2(N\times N, \mu)$. Therefore, the product of these two double sequences, $\{i^{2k}j\|w_{ij}\|^2\}$ is in $L_1(N\times N, \mu)$ so $\sum_{ij}\|n^k Dw_{ij}\|^2$ converges. This shows that $Dw \in \mathrm{Dom}(n^k)$, and

$$\|n^k Dw\|^2 \le \Bigl(\sum_{ij} j^2\|w_{ij}\|^2\Bigr)^{1/2}\Bigl(\sum_{ij} i^{4k}\|w_{ij}\|^2\Bigr)^{1/2}$$

$$\|n^k Dw\| \le \alpha\Bigl(\sum_{ij} i^{4k}\|w_{ij}\|^2\Bigr)^{1/4}$$

$$\le \alpha\|(n+1)^{2k}w\|^{1/2}$$

for the appropriate constant, α. Thus, the radius of convergence of

$$\sum_k \frac{\|n^k Dw\|t^k}{k!}$$

is at least as great as that of

$$\sum_k \frac{\|(n+1)^{2k}w\|^{1/2}t^k}{k!}. \tag{4.5}$$

We now use the following elementary fact about power series: If $\sum_k(a_k t^k/k!)$ converges for $|t|<\beta$, then $\sum_k[(a_{2k})^{1/2}t^k/k!]$ converges for $|t|<\frac{1}{2}\beta$.

This can easily be proved using Stirling's formula. Applying this result, we get that the radius of convergence of (4.5) is at least half as great as that of

$$\sum_k \frac{\|(n+1)^k w\|t^k}{k!},$$

which is positive since w is analytic for $n+1$. A similar argument for D^* completes the proof of Lemma 4.4. ∎

We are now ready to prove that $\mathcal{A}v$ is a set of analytic vectors for each $n(z)$. We first note that since

$$C(x+y)\supset C(x)+C(y), \tag{4.6}$$

if z is fixed, we may assume that the operators in $\mathcal{A}$ involve only vectors $y\in H$ which are either parallel or orthogonal to z. Since v is analytic for $n(z)$, Lemmas 4.3 and 4.4 and a simple inductive argument give the result.

Our next step is to show that the paraboson relationship, (1.4) is satisfied on $\mathcal{A}v$. We first note that if $w \in \mathcal{A}v$, we have shown in Lemma 4.4 that $nDP_{ij}w = DP_{ij}nw$, so $nDw_{ij} = Dnw_{ij}$. A now familiar argument shows that all relevant summations converge, so $nDw = Dnw$. We now have that if z and y are orthogonal and $w\in\mathcal{A}v$,

$$[n(z), C(z)]w = \|z\|^2 C(z)w$$

and

$$[n(z), C(y)]w = 0.$$

If y is now arbitrary, we can write y as a sum of two terms, one parallel to z and one orthogonal to z and use (4.6) to get

$$[n(z), C(y)]w = \langle y, z\rangle C(z)w.$$

Polarization of this (in z) then gives (1.4).

Let z be a unit vector in H. For the free paraboson field of order p, $n(z)$ and $d\Gamma(P_z)$ differ by $\frac{1}{2}p$, a scalar which is independent of z. We now show that we have a similar result here. $n(z)$ and $d\Gamma(P_z)$ are commuting self-adjoint operators and so

$$B(z) = \overline{2[n(z) - d\Gamma(P_z)]}$$

is a well-defined self-adjoint operator which has $\mathcal{A}v$ as an invariant set of analytic vectors. We will show that $B(z)$ is independent of z. It is sufficient to show that $B(z)w$ is independent of z for each $w\in\mathcal{A}v$ since $\mathcal{A}v$ is a core for $B(z)$. Since

$$\exp[itd\Gamma(P_z)]C(y)\exp[-it\,d\Gamma(P_z)] = C[\exp(itP_z)y],$$

$$\exp[itB(z)]C(y)\exp[-itB(z)] = C(y),$$

and

$$\exp[itB(z)]C^*(y)\exp[-itB(z)] = C^*(y),$$

so if $A\in\mathcal{A}$,

$$\exp[itB(z)]A\exp[-itB(z)] = A, \tag{4.7}$$

$$\exp[itB(z)]Av = A\exp[itB(z)]v,$$

and

$$\frac{\exp[itB(z)]-1}{it}Av = A\left(\frac{\exp[itB(z)]-1}{it}\right)v.$$

As t approaches 0, the left side approaches $B(z)Av$ and so the right side converges.

$$\frac{1}{it}[\exp(itB(z))-1]v$$

approaches $B(z)v$ and since A is closable ($\mathcal{A}v$ is in the domain of A^*), the right side approaches $AB(z)v$. Thus, $B(z)Av = AB(z)v$, and therefore it is sufficient to show that $B(z)v$ is independent of z. The proof of this fact is identical to a proof previously given.[16] We sketch it here. If y is orthogonal to z, then

$$d\Gamma(P_z)C^*(z)C(y)v = -C^*(z)C(y)v$$

so $C^*(z)C(y)v = 0$ since $d\Gamma(P_z)\ge 0$. Thus, if y and z are (not necessarily orthogonal) unit vectors,

$$C^*(z)C(y)v = C^*(z)C(P_z y)v$$

$$= \langle y, z\rangle C^*(z)C(z)v$$

$$= \langle y, z\rangle B(z)v$$

and similarly,

$$C^*(z)C(y)v = \langle y, z\rangle B(y)v,$$

thus, $B(y)v = B(z)v$ if $\langle y, z\rangle \ne 0$. Therefore, $B(z)v$ is independent of z. Let $B = B(z)$.

The next step is to show that the spectrum of B contains only nonnegative integers. B commutes with each $\Gamma(U)$ since

$$\Gamma(U)B(z)\Gamma(U)^{-1} = B(Uz) = B(z),$$

and so B commutes with each $d\Gamma(P_z)$. We will use the following lemma:

Lemma 4.5: Suppose $\mathcal{D}$ is a subset of K which is invariant under each $C(z)$ and $C^*(z)$, and (1.4) holds on $\mathcal{D}$. Suppose $w\in\mathcal{D}$ and $\{z_1, z_2, \ldots, z_k\}$ is an orthonormal set of vectors in H such that for $j = 1, 2, \ldots, k$, $d\Gamma(P_{z_j})w = 0$. Let S_j be the symmetric group on j elements and define

$$\Phi_j = \sum_{\tau\in S_j}(-1)^\tau C(z_{\tau j})C(z_{\tau(j-1)})\cdots C(z_{\tau 1})w.$$

Then

$$\|\Phi_j\|^2=(j!)^2\langle B(B-1)(B-2)\cdots(B-j+1)w, w\rangle.$$

Proof: We set $C_j=C(z_j)$.

$$\|\Phi_j\|^2=\sum_{\pi}\langle C^*_{\pi j}\Phi_j, (-1)^{\pi}C_{\pi(j-1)}\cdots C_{\pi 1}w\rangle,$$

but each term in the summation is independent of π since Φ_j is already antisymmetric, so

$$\|\Phi_j\|^2=j!\langle C_j^*\Phi_j, C_{j-1}\cdots C_1 w\rangle.$$

Similarly,

$$\langle C_j^*\Phi_j, (-1)^{\pi}C_{\pi(j-1)}\cdots C_{\pi 1}w\rangle$$

is independent of $\pi\in S_{j-1}$, so

$$\|\Phi_j\|^2=j\langle C_j^*\Phi_j, \Phi_{j-1}\rangle.$$

Since

$$C_i^*C_j w=\delta_{ij}Bw,$$

if we use (1.4) to eliminate C_j^* and C_j from

$$C_j^*\Phi_j=\sum(-1)^{\pi}C_j^*C_{\pi j}\cdots C_{\pi 1}w,$$

the result will consist of terms linear in B and terms independent of B. Since the result must be antisymmetric in $C_1, C_2, \ldots, C_{j-1}$, there are scalars α_j and β_j such that

$$C_j^*\Phi_j=(\alpha_j+B\beta_j)\Phi_{j-1}.$$

Thus,

$$\|\Phi_j\|^2=j\langle(\alpha_j+B\beta_j)\Phi_{j-1}, \Phi_{j-1}\rangle$$

and similarly,

$$\|\Phi_k\|^2=k!\left\langle\prod_{j=2}^{k}(\alpha_j+B\beta_j)Bw, w\right\rangle$$

since $\Phi_1=C_1w$.

The scalars α_j and β_j have been determined[17] for the case in which B was a scalar. The result and proof are the same here:

$$\alpha_j+B\beta_j=j(B-j+1).$$

This completes the proof of Lemma 4.5. ■

Since H is infinite dimensional, for any $w\in Av$ and positive integer k, there exists an orthonormal set $\{z_1, z_2, \ldots, z_k\}$ such that $d\Gamma(P_{z_j})w=0$. Thus, $\mathcal{D}=\mathcal{A}v$ and w satisfy the hypotheses of Lemma 4.5. If P is any spectral projection of B, then P commutes with each $d\Gamma(P_z)$ and by (4.7), for each $A\in\mathcal{A}$, $PA\subset AP$. Thus, if we take $\mathcal{D}=P\mathcal{A}v$ and any $w\in P\mathcal{A}v$, these also satsify the hypotheses of the lemma. Therefore, for all $w\in PAv$ and all nonnegative integers, k,

$$\langle B(B-1)\cdots(B-k)w, w\rangle\geq 0.$$

The union of the sets $P\mathcal{A}v$ as P ranges over the spectral projections of B corresponding to bounded sets is a dense set on which each polynomial in B is essentially self-adjoint. Thus,

$$B(B-1)(B-2)\cdots(B-k)\geq 0$$

and so the spectrum of B is contained in the set of nonnegative integers.

Let $K_p=\{w\in K: Bw=pw\}$ and let P_p be the projection onto K_p. $K_p=\overline{P_p\mathcal{A}v}=\overline{\mathcal{A}P_pv}$ so that if $P_pv\neq 0$,

$$v_p=P_pv/\|P_pv\|$$

is cyclic for K_p.

$$C^*(z)v_p=0$$

and since $\Gamma(U)$ commutes with B, K_p is invariant under $\Gamma(U)$. Therefore,

$$\Gamma(U)v_p=v_p$$

and

$$C^*(y)C(z)v_p=\langle z, y\rangle Bv_p=p\langle z, y\rangle v_p.$$

Each K_p reduces $C(z)$ and (1.4) holds on $\mathcal{A}v_p$. It is now a simple matter to construct an isomorphism (as in Sec. 2) between $\{H, C, K_p, \Gamma, v_p\}$ and the free paraboson field of order p.

5. EXTENSIONS

We have already remarked that the hypothesis of Theorem 1 that v is analytic for each $n(z)$ can be replaced by the condition that v is in the domain of $C(z)$ for some nonzero z in H. Given the other hypotheses of Theorems 1 or 2, the condition

$$\Gamma(U)v=v \tag{5.1}$$

for all unitaries, U, on H can be replaced by the weaker condition that there exists a nonnegative self-adjoint operator, A, on H whose discrete spectrum does not contain 0 such that

$$\Gamma[\exp(itA)]v=v.$$

That this condition implies (5.1) has been proved in a previous paper.[18] The proof relies heavily on the positivity of $d\Gamma$.

In Theorems 1 and 2, it is also possible to weaken the positivity condition, $d\Gamma\geq 0$, if the full invariance condition (5.1) is kept. In Theorem 1, $d\Gamma\geq 0$ is used only to prove that

$$C^*(z)v=0 \tag{5.2}$$

for all $z\in H$. In Theorem 2 it is used to prove (5.2) and

$$C^*(y)C(z)v=0 \tag{5.3}$$

when y is orthogonal to z. (5.2) and (5.3) are implied by a weaker positivity condition. (5.2) holds if $d\Gamma(A)\geq 0$ for some nonnegative self-adjoint operator, A, other than 0. (5.3) also follows if this operator is not a multiple of the identity. These statements have been proved[19] for parafermions using the condition that $C(z)$ is a bounded operator.

We now modify the proof given there to include the unbounded case. First we remark that because of (5.1) and

$$\Gamma(U)C^*(z)\Gamma(U)^{-1}=C^*(Uz),$$

$\|C^*(z)v\|$ depends only on $\|z\|$ and so there is a constant α, such that

$$\|C^*(z)v\|=\alpha\|z\|.$$

Similarly, there exist constants β and γ such that

$$\|C^*(y)C(z)v\| = \beta\|y\| \, \|z\|$$

when y is orthogonal to z and

$$\|C^*(z)C(z)v\| = \gamma\|z\|^2.$$

Thus, there exists a constant δ such that for arbitrary y and z,

$$\|C^*(y)C(z)v\| \leq \delta\|y\| \, \|z\|.$$

In particular, $C^*(z)v$ is a continuous function of z and $C^*(y)C(z)v$ is continuous in both y and z.

Suppose A is a self-adjoint operator on H and $z \in \mathrm{Dom}(A)$,

$$\exp[it\, d\Gamma(A)]C^*(z)v = C^*[\exp(itA)z]v,$$

$$\left(\frac{\exp[it\, d\Gamma(A)] - 1}{it}\right)C^*(z)v = -C^*\left(\frac{\exp(itA) - 1}{it}z\right)v. \tag{5.4}$$

As t approaches zero,

$$\frac{\exp(itA) - 1}{it}z \to Az,$$

so by the continuity of $C^*(z)v$ noted above, the right side of (5.4) approaches $-C^*(Az)v$. Thus, $C^*(z)v$ is in the domain of $d\Gamma(A)$ and

$$d\Gamma(A)C^*(z)v = -C^*(Az)v. \tag{5.5}$$

Suppose $A \geq 0$, $r \neq 0$ is in the spectrum of A, and $d\Gamma(A) \geq 0$. For each $\epsilon > 0$ there is a unit vector $z \in \mathrm{Dom}(A)$ such that $\|(A - r)z\| < \epsilon$,

$$0 \leq \langle d\Gamma(A)C^*(z)v, C^*(z)v\rangle,$$

$$0 \leq -\langle C^*((A - r)z)v, \; C^*(z)v\rangle - r\langle C^*(z)v, C^*(z)v\rangle,$$

$$0 \leq \alpha^2\epsilon - r\alpha^2 = \alpha^2(\epsilon - r).$$

Since ϵ is arbitrary, $\alpha = 0$. This gives (5.2).

If y and z are in the domain of A, then in a manner similar to the one used to derive (5.5), we get that $C^*(y)C(z)v$ is in the domain of $d\Gamma(A)$ and

$$d\Gamma(A)C^*(y)C(z)v = C^*(y)C(Az)v - C^*(Ay)C(z)v. \tag{5.6}$$

Now, if $A \geq 0$, $A \neq 0$, A is not a multiple of the identity and $d\Gamma(A) \geq 0$, we choose two numbers r and s in the spectrum of A with $r < s$. For each $\epsilon > 0$ there are orthogonal unit vectors y and z such that $\|(A - s)z\|$ and $\|(A - r)y\|$ are less than ϵ. Using (5.6) we obtain

$$0 \leq \langle C^*(y)C(Az)v, C^*(y)C(z)v\rangle - \langle C^*(Ay)C(z)v, C^*(y)C(z)v\rangle$$

$$0 \leq 2\beta^2\epsilon + (r - s)\beta^2.$$

Again, since ϵ is arbitrary and $r < s$, $\beta = 0$. This gives (5.3).

[1] V. Fock, Z. Phys. 75, 622 (1932).
[2] J.M. Cook, Trans. Am. Math. Soc. 74, 222 (1953).
[3] For a good discussion of unbounded operators, see Sec. viii.5 of M. Reed and B. Simon, *Functional Analysis*, *Methods of Modern Mathematical Physics* (Academic, New York, 1972), Vol. 1.
[4] For an exception to this, see Theorem 4.14.1 of C.R. Putnam, *Commutation Properties of Hilbert Space Operators and Related Topics* (Springer, New York, 1967).
[5] H. Weyl, Z. Phys. 46, 1 (1927).
[6] I.E. Segal, *Mathematical Problems of Relativistic Physics* (Am. Math. Soc., Providence, 1963), p. 17.
[7] H.S. Green, Phys. Rev. 90, 270 (1953).
[8] S. Robbins, Trans. Am. Math. Soc. 209, 389 (1975).
[9] I.E. Segal, Ill. J. Math. 6, 500 (1962).
[10] E. Nelson, Ann. Math. (N.Y.) 70, 572 (1959). See also M. Reed and B. Simon, *Fourier Analysis*, *Self Adjointness*, *Methods of Modern Mathematical Physics* (Academic, New York, 1976).
[11] See Ref. 8.
[12] See Ref. 2.
[13] For a definition of the tensor product of unbounded operators, see Definition 2 of Ref. 2.
[14] See p. 225 of Ref. 2.
[15] S. Robbins, Commun. Math. Phys. 38, 111 (1974).
[16] See p. 114 of Ref. 15.
[17] O.W. Greenberg and A.M.L. Messiah, Phys. Rev. 138, B1155 (1965).
[18] See Theorem 2 of Ref. 15.
[19] See Theorem 3 of Ref. 15.

Covariant canonical formulation and centers of mass and motion for a relativistic two-body problem

Donald E. Fahnline*

The Pennsylvania State University, Altoona, Pennsylvania 16601
(Received 27 August 1976)

Fokker has given a multitime action principle yielding the equations of motion for a relativistic two-body system in which the first particle responds to the retarded Liénard–Wiechert field of the second, while the second responds to the advanced field of the first. The present paper exhibits a single time parameter Lagrangian which avoids the no-interaction theorems and leads to a manifestly covariant and canonical description of this system via Dirac's generalized dynamics. Within this formulation the system possesses a manifestly covariant and canonical center of mass and a separate manifestly covariant but noncanonical "center of motion" which moves with the constant 4-velocity corresponding to the conserved total 4-momentum. In the one-dimensional case the two centers coincide, and the differential equation for the internal motion reduces to the classical Kepler problem. In the general case the motion lies in a plane, and the center of mass moves in a circle about the center of motion.

INTRODUCTION

This paper presents a single-parameter Lagrangian which avoids the no-interaction theorems[1–5] and leads to a manifestly covariant and canonical formulation, a manifestly covariant and canonical center of mass, and a manifestly covariant "center of motion" for a relativistic action-at-a-distance two-body problem due to Fokker[6]: One spinless electrically charged particle responds without self-action to the retarded Liénard–Wiechert field of a second, while the second responds similarly to the advanced field of the first. Although this problem is time asymmetric, the fact that it has ordinary differential equations of motion has led authors to examine it as a close approximation to the difference–differential equations of the time-symmetric Wheeler–Feynman problem,[7,8] or in the hope that it will illuminate more physical problems.[9–12]

Fokker[6] derives the equations of motion and the conserved total 4-momentum vector P^μ of the two-particle system from a multitime action principle and the externally imposed constraint that the particles maintain null separation r^μ. Due to the constraint, the coordinates of the particles are not independent. The resulting equations of motion express the proper acceleration of each particle in terms of the proper acceleration of the other, the proper velocities, and the coordinates. Bruhns,[10] proceeding in the same manner, finds the conserved total angular momentum tensor $J^{\mu\nu}$ and simplifies the equations of motion to express the proper accelerations in terms of the proper velocities and the coordinates.

Staruszkiewicz[13] finds a single-parameter canonical formulation for the theory by introducing the spatial components $\mathbf{r}$ of the separation vector and the components of a center of mass as independent generalized coordinates. Neither the particle coordinates nor the time component r^0 of the separation vector are canonical. The center of mass has a complicated equation of motion and is a canonical position variable only if P^μ is replaced by its numerical value.

Künzle[11] gives a differential geometric multitime formulation of the general problem of two interacting relativistic particles and finds that the spatial components of the particle coordinates can be independent canonical variables if the particles have null separation. He studies Fokker's problem as an approximation to an interaction he derives within his formulation from the Liénard–Wiechert fields with the acceleration terms ignored.[12] After finding P^μ and $J^{\mu\nu}$ and their Poisson brackets via Noether's theorem and their properties as generators of the Poincaré group, he uses them to construct, following Synge,[14] a center of mass defined up to an arbitrary term parallel to P^μ. Although this center of mass has constant velocity and the appropriate brackets with the spatial components of $J^{\mu\nu}$ and P^μ, it is not a canonical position variable because its spatial components do not have zero Poisson brackets with each other.

Working within the frame defined by his center of mass, Künzle[12] reduces the problem to quadratures and gives numerical results for the case of equal rest masses and for the case of circular motion for arbitrary mass ratios. Rudd and Hill,[9] Staruszkiewicz,[7] Bruhns,[10] and Künzle[12] give solutions for the special case where the motion is confined to one dimension by the initial conditions. Bruhns also gives circular motion solutions.

The first section of the present paper exhibits a single-parameter Lagrangian with a Lagrange multiplier term which yields simultaneously the light-cone constraint and Fokker's and Bruhns' equations of motion in forms containing derivatives with respect to the single scalar parameter s instead of the two proper times of the particles. It also yields definitions for the particle momenta and, via Noether's theorem, the conserved momentum P^μ and angular momentum $J^{\mu\nu}$; each of these contains interaction contributions in accordance with the general theorem of Van Dam and Wigner.[15] The Lagrangian guarantees the existence of a canonical formulation of Fokker's problem in which the Hamiltonian is a scalar and the temporal components as well as the spatial components of the particle positions and momenta are independent canonical variables in the sense of Dirac's theory of generalized dynamics.[16,17] This guarantee justifies the definition of a manifestly covariant Poisson bracket, from which results the expected basic brackets and the Lie algebra of the Poincaré group. The world-line condition brackets fundamental to the no-interaction theorems[1-5] do not arise here because of the independence of the components.

The "center of motion" 4-vector defined in the second section is the Synge–Künzle center of mass[12] modified to include a specification of the time component. Although its does not lie on the line joining the particles and is not canonical, its derivative is the constant 4-velocity corresponding to P^μ. It thus provides a natural frame, the "center of motion frame," for describing the motion of the particles, and its proper time provides a natural choice for the parameter s. The name "center of motion" distinguishes it from the separate center of mass and suggests its relation to the motion of the system.

The third section defines a center of mass for the system. It lies on the line joining the particles and is a covariant position variable canonically conjugate to P^μ. Although it is functionally more complicated than Staruszkiewicz's center of mass,[13] its motion is simpler. This is most apparent in the center of motion frame, as discussed in the fourth section. If $J^{\mu\nu}$ is not zero, the particles move in a plane, the center of mass moves in this plane in a circle about the center of motion, and $\mathbf{r}$ remains tangent to this circle at the center of mass. These geometric relations and the specification that s be the center of motion proper time simplify the reduction of the problem to quadratures. If $J^{\mu\nu}$ is zero, the centers of mass and motion coincide, the motion is (excluding an exceptional case) one-dimensional through the center of motion, and the reduced equation for the internal motion has the simple inverse square form of the classical Kepler problem.

LAGRANGIAN FORMULATION

The symbol $x_n{}^\mu$ represents the Minkowski space coordinates of particle n, where Greek indices ranging from 0 to 3 denote the components of 4-vectors and $x_n{}^0 \equiv ct_n$. The particular subscripts n, $f = 1, 2$ always refer to the two particles and are never equal when they appear in the same equation. All other small Latin indices range from 1 to 3 and denote the components of the spatial part $\mathbf{a}$ of a 4-vector a^μ. The metric tensor is $g^{ii} = -g^{00} = 1$, $g^{\mu\nu} = 0$ for $\mu \neq \nu$. The form $a \cdot b$ denotes the scalar product $a^\mu b_\mu$ of two 4-vectors. A dot above a variable indicates differentiation with respect to a single scalar parameter s subject only to the nonholonomic constraints that the velocities $\dot{x}_n{}^\mu$ be timelike and future pointing:

$$\dot{x}_n \cdot \dot{x}_n < 0, \quad \dot{x}_n^0 > 0. \tag{1}$$

The first object of this section is to show that the equations of motion for Fokker's time-asymmetric two-body problem result from the single-parameter Lagrangian

$$L \equiv -\sum_n (m_n c \omega_n + g\xi/2\sigma_n) + \lambda r \cdot r, \tag{2}$$

where $g \equiv e_1 e_2 / c$ is the coupling constant in Gaussian units, m_n and e_n are the constant rest mass and electric charge of particle n, the scalar λ is a Lagrange multiplier, and

$$r^\mu \equiv x_1{}^\mu - x_2{}^\mu, \quad \xi \equi -\dot{x}_1 \cdot \dot{x}_2 > 0,$$

$$\omega_n \equiv (-\dot{x}_n \cdot \dot{x}_n)^{1/2} > 0, \quad \sigma_n \equiv -\dot{x}_n \cdot r.$$

By the Lagrangian equations the particle momenta are[18]

$$p_n{}^\mu \equiv \partial L / \partial \dot{x}_{n\mu} = m_n c \dot{x}_n{}^\mu / \omega_n + g \dot{x}_f{}^\mu (\sigma_1^{-1} + \sigma_2^{-1})/2 - g\xi r^\mu / 2\sigma_n{}^2, \tag{3}$$

and the momentum conjugate to the Lagrange multiplier is

$$\pi \equiv \partial L / \partial \dot{\lambda} = 0; \tag{4}$$

their derivatives are

$$\dot{p}_n{}^\mu \equiv \partial L / \partial x_{n\mu} = (-1)^n [g\xi(\sigma_1^{-2} \dot{x}_1{}^\mu + \sigma_2^{-2} \dot{x}_2{}^\mu)/2 - 2\lambda r^\mu] \tag{5}$$

and

$$\dot{\pi} \equiv \partial L / \partial \lambda = r \cdot r. \tag{6}$$

Fokker's externally applied constraint

$$r \cdot r = 0 \tag{7}$$

now appears as a consequence of (4) and (6). With the assumption that $r^0 > 0$ for definiteness, the derivatives of (7) imply

$$\sigma \equiv \sigma_1 = \sigma_2 > 0 \tag{8}$$

and

$$\ddot{r} \cdot r = -\dot{r} \cdot \dot{r} \leq 0. \tag{9}$$

Equation (8) simplifies (3) to

$$p_n{}^\mu = m_n c \dot{x}_n{}^\mu / \omega_n + g \dot{x}_f{}^\mu / \sigma - g\psi r^\mu / 2, \tag{10}$$

where $\psi \equiv \xi / \sigma_1 \sigma_2 > 0$.

Equating the derivative of (3) with (5) and taking the scalar product with $\dot{x}_{n\mu}$ lead to an expression for the Lagrange multiplier:

$$2\lambda / g\sigma = \psi^2 + (\ddot{x}_1 \cdot \dot{x}_2 - \dot{x}_1 \cdot \ddot{x}_2)/2\sigma^3. \tag{11}$$

The elimination of λ now reduces (3) and (5) to the single-parameter equivalent of Fokker's equations of motion[6,10]:

$$\frac{m_n c}{g\sigma} \frac{d}{ds} \frac{\dot{x}_n{}^\mu}{\omega_n} = -\frac{\ddot{x}_f{}^\mu}{\sigma^2} - \frac{\dot{x}_f{}^\mu \ddot{x}_f \cdot r}{\sigma^3} + r^\mu \left(\frac{\psi \ddot{x}_f \cdot r}{\sigma^2} - \frac{\dot{x}_n \cdot \ddot{x}_f}{\sigma^3} \right) + (-1)^n \frac{\omega_f^2}{\sigma^2} \left(\frac{\dot{x}_f}{\sigma} - \psi r^\mu \right). \tag{12}$$

Equation (12) implies

$$\frac{m_n c r_\mu}{g\sigma} \frac{d}{ds} \frac{\dot{x}_n{}^\mu}{\omega_n} = -\frac{(-1)^n}{\rho_f{}^2} \tag{13}$$

and

$$\frac{m_n c \rho_f \dot{x}_{f\mu}}{g\sigma^2} \frac{d}{ds} \frac{\dot{x}_n{}^\mu}{\omega_n} = \frac{\dot{x}_{n\mu}}{\sigma^2} \left(\frac{d}{ds} \frac{\dot{x}_f{}^\mu}{\omega_f} \right) + (-1)^n \left(\frac{g}{m_f c \rho_n{}^2} - \frac{1}{\rho_f} \right) \left(\frac{1}{\rho_f{}^2} - \psi \right), \tag{14}$$

where $\rho_n \equiv \sigma_n/\omega_n > 0$. Using (13) and (14) in (12) yields

$$(-1)^n \frac{\eta}{\omega_n}\frac{d}{ds}\frac{\mathring{x}_n^{\,\mu}}{\omega_n} = g^2\left(\frac{g}{m_n c\rho_f^{\,2}} - \frac{1}{\rho_n}\right)\left(\frac{r^\mu}{\rho_n} - \frac{\mathring{x}_n^{\,\mu}}{\omega_n}\right) + \frac{g m_f c \rho_n^{\,2}}{\rho_f}\left(\frac{g}{m_f c\rho_n^{\,2}} - \frac{1}{\rho_f}\right)\left(\psi\rho_f r^\mu - \frac{\mathring{x}_f^{\,\mu}}{\omega_f}\right), \tag{15}$$

where $\eta \neq m_1 m_2 c^2 \rho_1\rho_2 - g^2$. This is the single-parameter equivalent of Bruhns' equations of motion,[10] which are expressed in terms of the proper times τ_n, the proper velocities $v_n^{\,\mu} \equiv dx_n^{\,\mu}/d\tau_n$ satisfying $v_n \cdot v_n = -c^2$, and the proper accelerations $a_n^{\,\mu} \equiv dv_n^{\,\mu}/d\tau_n$. The translation to the proper time form follows from the identities

$$\omega_n\, ds = c\, d\tau_n, \quad \mathring{x}_n^{\,\mu}/\omega_n = v_n^{\,\mu}/c,$$

$$\psi = -v_1 \cdot v_2/c^2\rho_1\rho_2, \quad \rho_n = -v_n \circ r/c.$$

Since both (12) and (15) satisfy

$$\mathring{x}_{n\mu}\, d(\mathring{x}_n^{\,\mu}/\omega_n)/ds = 0$$

identically, each of them contains only six independent equations for the eight accelerations $\ddot{x}_n^{\,\mu}$. Equation (9) provides a seventh equation, leaving one acceleration arbitrary in accordance with the arbitrariness of the parameter s. The initial conditions must satisfy (7) and (8).

Within this single-parameter formalism the invariance of L under the space—time translations and rotations of the Poincaré group implies, via Noether's theorem, that the energy-momentum 4-vector

$$P^\mu \equiv p_1^{\,\mu} + p_2^{\,\mu} = \sum_n (m_n c + g/\rho_n)\,\mathring{x}_n^{\,\mu}/\omega_n - g\psi r^\mu \tag{16}$$

and the angular momentum tensor

$$J^{\mu\nu} \equiv \sum_n (x_n^{\,\mu} p_n^{\,\nu} - x_n^{\,\nu} p_n^{\,\mu}) \tag{17}$$

are conserved. Equations (3), (7), and (8) and the proper time notation reduce P^μ and $J^{\mu\nu}$ to the expressions given by other authors.[6,10,12,13] The interpretation of P^μ as the 4-momentum of a composite particle and the definitions of the centers of motion and mass given in the next sections require that the initial conditions determine P^μ to be timelike and future pointing. The scalar $m > 0$ defined by

$$m^2c^2 \equiv -P \circ P = \sum_n (m_n c + g/\rho_n)^2 + 2\eta\psi \tag{18}$$

is conserved and represents the rest mass of the composite particle. Simultaneously, mc^2 represents the total energy of the system in the frame where $P^i = 0$.

The following incomplete discussion of the single-parameter canonical formulation of Fokker's problem suffices as justification of the Poisson bracket used in the next sections to study the centers of mass and motion. The definitions of the particle momenta in (3) yield

$$\Gamma_n \equiv -p_n \circ r = m_n c\rho_n + g(\zeta+1)\,\zeta^{1-n}/2 + g\psi\zeta^{3-2n}\, r\circ r/2, \tag{19}$$

$$\Lambda_n \equiv -p_n \circ p_n = m_n^{\,2}c^2 + [g(\zeta+1)\,\zeta^{1-n}/2\rho_f]^2 - (g\psi\zeta^{3-2n}/2)^2\, r\circ r + g m_n c\rho_n\psi - g^2\psi(\zeta+1)\,\zeta^{4-3n}/2, \tag{20}$$

and

$$\Xi \equiv -p_1 \cdot p_2 = m_1 m_2 c^2 \rho_1\rho_2\psi + (m_1/\rho_1\zeta + m_2/\rho_2)\, g(\zeta+1)\,c/2 - (m_1\rho_1/\zeta + m_2\rho_2\zeta)\,g\psi c/2 - (g\psi/2)^2\, r\cdot r, \tag{21}$$

where $\zeta \equiv \sigma_2/\sigma_1$. [At this point in the Hamiltonian analysis (7) and (8) are unavailable.] The elimination of the four velocity dependent variables ρ_n, ψ, and ζ from these five equations yields one scalar constraint $\phi_1(p_n^{\,\mu}, x_n^{\,\mu}) = 0$. Alternatively, the existence of ϕ_1 is guaranteed by the zero-degree homogeneity of the $p_n^{\,\mu}$ of (3) under simultaneous variation of the $\mathring{x}_n^{\,\mu}$. Equation (4) is a second scalar constraint: $\phi_2 \equiv \pi = 0$. Hence the Lagrangian is nonstandard in the sense that the elimination of the velocities $\mathring{x}_n^{\,\mu}$ and $\mathring{\lambda}$ in terms of $x_n^{\,\mu}$, λ, $p_n^{\,\mu}$, and π and the usual transition to the Hamiltonian are impossible. Nevertheless, the transition is possible within Dirac's theory of generalized dynamics,[16,17] which in the present case prescribes the scalar Hamiltonian

$$H \equiv v_1\phi_1 + v_2\phi_2 - L + \sum_n p_n \circ \mathring{x}_n = v_1\phi_1 + v_2\phi_2 - \lambda r\circ r, \tag{22}$$

where v_1 and v_2 are scalar functions of $\mathring{x}_n^{\,\mu}$ and $\mathring{\lambda}$ only.

The Poisson bracket of two variables A and B in this formulation is

$$[A,B] \equiv \frac{\partial A}{\partial\lambda}\frac{\partial B}{\partial\pi} - \frac{\partial A}{\partial\pi}\frac{\partial B}{\partial\lambda} + \sum_n\left(\frac{\partial A}{\partial x_n^{\,\alpha}}\frac{\partial B}{\partial p_{n\alpha}} - \frac{\partial A}{\partial p_n^{\,\alpha}}\frac{\partial B}{\partial x_{n\alpha}}\right), \tag{23}$$

where $x_n^{\,\alpha}$, $p_n^{\,\alpha}$, λ, and π are independent canonical variables , and the constraints $\phi_1 = 0$ and $\phi_2 = 0$ apply only after the derivatives have been taken. The only nonzero basic brackets are

$$[x_n^{\,\mu}, p_n^{\,\nu}] = g^{\mu\nu} \tag{24}$$

and

$$[\lambda, \pi] = 1. \tag{25}$$

The brackets of P^μ and $J^{\mu\nu}$ produce the Poincaré algebra:

$$[P^\mu, P^\nu] = 0, \tag{26}$$

$$[J^{\mu\nu}, P^\sigma] = g^{\mu\sigma}P^\nu - g^{\nu\sigma}P^\mu, \tag{27}$$

and

$$[J^{\mu\nu}, J^{\rho\sigma}] = g^{\mu\rho}J^{\nu\sigma} - g^{\nu\rho}J^{\mu\sigma} + g^{\nu\sigma}J^{\mu\rho} - g^{\mu\sigma}J^{\nu\rho}. \tag{28}$$

A CENTER OF MOTION

The "center of motion" discussed in this section is

$$x^\mu \equiv (J^{\mu\nu}P_\nu - bP^\mu)/P\circ P, \tag{29}$$

where

$$b \equiv -p_1 \circ x_1 - p_2 \circ x_2. \tag{30}$$

It is a modification of the center of mass defined by

Synge[14] and applied to Fokker's problem by Künzle[12]; they specify only $b \equiv -x\cdot P$ so that their center is defined only up to an arbitrary term parallel to P^μ, and their time component x^0 is undefined in the frame where $P^i=0$. The definition (30) does not fit the general approach exemplified by Lorente and Roman,[19] because b is constructed from the individual particle variables rather than the generators P^μ and $J^{\mu\nu}$. The new name emphasizes that the role of x^μ is distinct from that of a center of mass.

Since P^μ and $J^{\mu\nu}$ are conserved, the derivative of the center of motion is

$$\mathring{x}^\mu = \mathring{b}\, P^\mu/m^2c^2, \tag{31}$$

where

$$\mathring{b} = m_1c\omega_1 + m_2c\,\omega_2 > 0 \tag{32}$$

follows from (5), (7), (8), and (10). Hence $\mathring{x}^\mu$ is, like P^μ, timelike and future pointing. The equation

$$m\omega_x \equiv m(-\mathring{x}\cdot\mathring{x})^{1/2} = m_1\omega_1 + m_2\omega_2 = \mathring{b}/c, \tag{33}$$

resulting from (31) and (32), implies that the proper velocity of the center of motion is constant:

$$v_x{}^\mu \equiv dx^\mu/d\tau_x = c\mathring{x}^\mu/\omega_x = P^\mu/m, \tag{34}$$

where the definition of the center of motion proper time τ_x requires $v_x\cdot v_x = -c^2$. Using $\omega_n = \sigma/\rho_n$ in (33) leads to

$$\sigma = m\omega_x/(m_1\rho_1^{-1} + m_2\rho_2^{-1}). \tag{35}$$

Since $\omega_x = c\mathring{\tau}_x$, etc., (33) with an appropriate choice of the initial values of the proper times integrates to

$$m\tau_x = m_1\tau_1 + m_2\tau_2 = b/c^2. \tag{36}$$

Equations (16), (19), and (34) give

$$\Gamma = mc\rho_x = \Gamma_1 + \Gamma_2 = m_1c\rho_1 + m_2c\rho_2 + 2g, \tag{37}$$

where $\rho_x \equiv \sigma_x/\omega_x = -\mathring{x}\cdot r/\omega_x$ and $\Gamma \equiv -P\cdot r$.

The conservation of P^μ and $J^{\mu\nu}$ implies the conservation of the Pauli–Lubanski vector

$$W_\mu \equiv J^*_{\mu\nu}P^\nu, \tag{38}$$

where $J^*_{\mu\nu} \equiv \epsilon_{\mu\nu\alpha\beta}J^{\alpha\beta}/2$ and $\epsilon_{\mu\nu\alpha\beta}$ is the completely antisymmetric tensor with $\epsilon_{0123} = -\epsilon^{0123} = 1$. Equations (10), (16), and (17) give

$$W_\mu = \epsilon_{\mu\nu\alpha\beta}r^\nu p_1{}^\alpha p_2{}^\beta = \eta\epsilon_{\mu\nu\alpha\beta}r^\nu\mathring{x}_1{}^\alpha\mathring{x}_2{}^\beta/\sigma^2. \tag{39}$$

Numerous orthogonality relations result from (38) and (39), including

$$W\cdot P = 0$$

and

$$(x_n{}^\mu - x^\mu)\,W_\mu = 0. \tag{40}$$

Since P^μ is timelike, W^μ must be spacelike or zero. Equation (39) with the elimination of ψ from (18) and

$$\mathring{r}\cdot\mathring{r}/\sigma^2 = 2\psi - \rho_1^{-2} - \rho_2^{-2} \tag{41}$$

yields

$$W\cdot W/\eta = \eta\mathring{r}\cdot\mathring{r}/\sigma^2 = m^2c^2 - mc\rho_x(m_1c\rho_1^{-1} + m_2c\rho_2^{-1}). \tag{42}$$

Equation (35) then yields

$$\sigma = m^2c^2\omega_x\rho_x/(m^2c^2 - W\cdot W/\eta). \tag{43}$$

The spin angular momentum tensor associated with x^μ is

$$S_x^{\mu\nu} \equiv J^{\mu\nu} - (x^\mu P^\nu - x^\nu P^\mu). \tag{44}$$

By (29) and (38), $S_x^{\mu\nu}$ is conserved and satisfies

$$S_x^{\mu\nu}P_\nu = 0 \tag{45}$$

and

$$W_\mu = S_x{}^*{}_{\mu\nu}P^\nu.$$

The Poisson brackets of x^μ are

$$[x^\mu, P^\nu] = g^{\mu\nu}, \tag{46}$$

$$[J^{\mu\nu}, x^\sigma] = g^{\mu\sigma}x^\nu - g^{\nu\sigma}x^\mu, \tag{47}$$

and

$$[x^\mu, x^\nu] = -S_x^{\mu\nu}/P\cdot P. \tag{48}$$

Hence, although x^μ transforms properly under the Poincaré group, its components are not independent canonical variables.

A CENTER OF MASS

According to (23), any form $z^\mu \equiv d_1x_1{}^\mu + d_2x_2{}^\mu$, where d_1 and d_2 are numbers such that $d_1 + d_2 = 1$, is a canonical position variable conjugate to P^μ. The corresponding internal canonical variables are r^μ and $p_r^\mu \equiv d_2p_1{}^\mu - d_1p_2{}^\mu$. The Newtonian specification $d_n \equiv m_n/(m_1+m_2)$ provides an example. Staruszkiewicz's center of mass[13] with

$$d_n \equiv [1 + (-1)^n(m_2^2 - m_1^2)/m^2]/2$$

also fits this prescription, under the condition that $m^2 \equiv -P\cdot P/c^2$ be replaced by its numerical value. However, the motions of these centers of mass are complicated.

Some simplification results from choosing the following definition with nonnumerical weighting factors:

$$z^\mu \equiv \sum_n \Gamma_n x_n{}^\mu/\Gamma, \tag{49}$$

where $\Gamma_n \equiv -p_n\cdot r = m_nc\rho_n + g$ by (7), (8), and (19). Equations (7), (16), and (17) yield an expression for z^μ in the general form proposed by Lorente and Roman for position operators in quantum mechanics[19]:

$$z^\mu = (J^{\mu\nu}r_\nu + x_1\cdot rP^\mu)/P\cdot r. \tag{50}$$

But this definition does not fit their approach, because its is constructed from the individual particle coordinates and because r^μ is null, not timelike. (Since P^μ is timelike by assumption, the denominator cannot be zero.)

This center of mass resembles the Newtonian center of mass in that it lies on the line joining the two particles, yields the inverse relations

$$x_n{}^\mu = z^\mu - (-1)^n\Gamma_f r^\mu/\Gamma, \tag{51}$$

and has the Poisson brackets of a manifestly covariant canonical position variable conjugate to P^μ:

$$[z^\mu, z^\nu]=0, \tag{52}$$

$$[z^\mu, P^\nu]=g^{\mu\nu}, \tag{53}$$

and

$$[J^{\mu\nu}, z^\sigma]=g^{\mu\sigma}z^\nu - g^{\nu\sigma}z^\mu. \tag{54}$$

In the limit of small velocities and coupling constant, both $\mathbf{z}$ and $\mathbf{x}$ reduce to the Newtonian center of mass, while t_z and t_x reduce to the instantaneous Newtonian time.

The spin angular momentum tensor corresponding to z^μ is

$$S_z^{\mu\nu}\equiv J^{\mu\nu} - (z^\mu P^\nu - z^\nu P^\mu) \tag{55}$$

$$=\eta(r^\mu \mathring{r}^\nu - r^\nu \mathring{r}^\mu)/\sigma\Gamma, \tag{56}$$

where the equality follows from (16), (17), and (49). Hence it obeys

$$S_z{}^{\mu\nu}r_\nu=0, \tag{57}$$

$$S_z{}^{\mu\nu}S_{z\,\mu\nu}=0, \tag{58}$$

and

$$W_\mu=S_z{}^*{}_{\mu\nu}P^\nu. \tag{59}$$

Its Poisson brackets are

$$[S_z{}^{\mu\nu}, z^\sigma]=0, \tag{60}$$

$$[S_z{}^{\mu\nu}, P^\sigma]=0, \tag{61}$$

and

$$[S_z{}^{\mu\nu}, S_z{}^{\rho\sigma}]=[J^{\mu\nu}, S_z{}^{\rho\sigma}]$$
$$=g^{\mu\rho}S_z{}^{\nu\sigma}-g^{\nu\rho}S_z{}^{\mu\sigma}+g^{\nu\sigma}S_z{}^{\mu\rho}-g^{\mu\sigma}S_z{}^{\nu\rho}. \tag{62}$$

Since

$$\mathring{\Gamma}_n=m_n c\mathring{\rho}_n=(-1)^n\sigma(m_n c\rho_n\psi - m_n c\rho_n^{-1}+g\rho_f^{-2}) \tag{63}$$

by (5), (8), (10), and (19), the derivative of z^μ is

$$\mathring{z}^\mu=\sigma(P^\mu - W\circ W r^\mu/mc\rho_x\eta)/mc\rho_x. \tag{64}$$

Hence, $\mathring{z}^\mu$ is not constant nor even necessarily timelike: (42) provides $W\circ W/\eta < m^2c^2$ but is insufficient for determining the sign of

$$\mathring{z}\circ\mathring{z}=(\sigma/mc\rho_x)^2(2W\circ W/\eta - m^2c^2). \tag{65}$$

Nevertheless, the motion of z^μ is simple in the center of motion frame discussed in the next section. The basis of this simplicity is the general relation between the centers of mass and motion, which follows from their definitions:

$$z^\mu=x^\mu+J^{\mu\nu}r_\nu/P\circ r - J^{\mu\nu}P_\nu/P\circ P$$
$$-J^{\alpha\beta}P_\alpha r_\beta P^\mu/(P\circ r)(P\circ P). \tag{66}$$

This equation implies

$$(z^\mu - x^\mu)(z_\mu - x_\mu)=W\circ W/(P\circ P)^2\geqslant 0 \tag{67}$$

and the orthogonality relations

$$(z^\mu - x^\mu)P_\mu=0, \tag{68}$$

$$(z^\mu - x^\mu)r_\mu=0, \tag{69}$$

$$(\mathring{z}^\mu - \mathring{x}^\mu)P_\mu=0, \tag{70}$$

and

$$(\mathring{z}^\mu - \mathring{x}^\mu)(z_\mu - x_\mu)=0. \tag{71}$$

CENTER OF MOTION FRAME

Since P^μ is timelike and future pointing by assumption, (34) and the Poisson brackets of x^μ and P^μ guarantee that a Lorentz transformation followed by a space translation suffices for reaching the "center of motion" frame, where $P^0=mc$, $\mathbf{P}=0$, $\mathring{\mathbf{x}}=0$, $x^0=c\tau_x$, and $\mathbf{x}=0$. This frame differs from Künzle's center of mass frame[12] in that x^0 is specified in terms of the particle variables.

In the center of motion frame the definition of ρ_x and (7) yield

$$\rho_x=r^0=r\equiv|\mathbf{r}|; \tag{72}$$

(29) yields

$$J^{\mu 0}=0; \tag{73}$$

and (38), (44), and (55) yield

$$S_x{}^{0i}=0, \quad S_z{}^{0i}=mcz^i, \tag{74}$$

$$W^0=0, \ \mathbf{S}_x=\mathbf{S}_z=\mathbf{J}=-\mathbf{W}/mc, \tag{75}$$

where $J_i\equiv\frac{1}{2}\epsilon_{ijk}J_{jk}$, etc. Hence, (66) implies

$$z^0=x^0, \quad \mathbf{z}=\mathbf{J}\times\mathbf{r}/mcr. \tag{76}$$

These equations describe the geometry of the system in the center of motion frame. If $J\equiv|\mathbf{J}|\neq 0$, (40) and (75) show that the particles move in the plane perpendicular to the conserved spin $\mathbf{S}_z=\mathbf{J}$, and (76) shows that the center of mass moves in this plane in a circle of radius J/mc about the origin, while $\mathbf{r}$ moves so as to always pass through the center of mass perpendicular to $\mathbf{z}$. If $J=0$, (76) implies that the center of mass and the center of motion coincide.

The choice $s=\tau_x$ is natural and convenient, especially in the center of motion frame. Although this specification simplifies the equations of motion to forms which may be interesting in themselves, in combination with the conservation laws it already provides a formal reduction to quadratures simpler than that obtained by Künzle[12]: Eqs. (43) and (75) with $\omega_x=c$ yield

$$\sigma=c\rho_x\eta/(\eta - J^2). \tag{77}$$

Solving (37) and (42) simultaneously for the ρ_n as functions of ρ_x also provides expressions for η, ψ, and σ as functions of ρ_x via the definition of η, (18) or (41), and (77). Employing these expressions in (37) and (63) yields $\mathring{\rho}_x=f(\rho_x)$, which integrates to give ρ_x and the above scalars as functions of τ_x.

Equations (56), (75), and (77) yield the expression

$$\mathbf{\Omega}=mc^2\mathbf{J}/(\eta - J^2) \tag{78}$$

for the angular velocity of $\mathbf{r}$. Since η is already known as a function of τ_x, this equation integrates to provide the direction of $\mathbf{r}$. Finally, (51) and (76) provide the particle coordinates.

If $J\neq 0$, (8) and (77) guarantee $\eta\neq 0$ and $\eta - J^2\neq 0$, so that the demoninator in (78) cannot be zero. The case $J=0$, $\eta=0$ is discussed by Künzle.[12]

If $J=0$ and $\eta\neq 0$, (78) implies that the direction of r^μ is constant. Since $\mathbf{z}=0$ for $J=0$, the motion is one-dimensional along a line through the origin in the center of motion frame. Although solutions for this case have been given by several authors,[7,9,10,12] a particularly simple method of solution results from the present formalism: In any frame (41) and (42) yield

$$2\psi=\rho_1^{-2}+\rho_2^{-2} \tag{79}$$

and

$$m\rho_x^{-1}=m_1\rho_1^{-1}+m_2\rho_2^{-1}, \tag{80}$$

while (77) yields

$$\sigma=c\rho_x. \tag{81}$$

Employing these equations with (37) and (63) gives

$$\frac{m_1m_2}{m}\,\frac{d^2}{d\tau_x^2}\,\frac{\rho_1\rho_2}{\rho_x}=gc\left(\frac{\rho_1\rho_2}{\rho_x}\right)^{-2}, \tag{82}$$

which is simply the inverse square law differential equation of the classical Kepler problem. Its well-known solution, (37), (51), and (80) yield the particle coordinates as explicit functions of τ_x.

DISCUSSION

The distinguishing feature of the present treatment of Fokker's time-asymmetric two-body problem is that the temporal components of all 4-vectors have equal status with the spatial components as independent functions of a single scalar parameter, despite the constraints $v_n\circ v_n=-c^2$ on the proper velocities and the constraint that the particles have null separation. This feature is based on the Langrange multiplier term in the Lagrangian and on the generalized theory of Hamiltonian dynamics, which Dirac devised for such constrained systems.[16] One consequence is the avoidance of the no-interaction theorems,[1–5] which is possible because the Hamiltonian is a scalar rather than a component of a 4-vector, and because it is the generator of the system motion with respect to the scalar parameter rather than with respect to the proper times or the time components of the particle coordinates.

A second consequence of the single-parameter Lagrangian formulation is the availability of the manifestly covariant canonical particle coordinates and momenta for the construction of the centers of mass and motion.

The light-cone constraint causes the separation of the centers of motion and mass; with the constraint in force the separation persists, and the particles and the center of mass have nonzero accelerations with respect to the center of motion proper time even for zero coupling constant. Constraints can have similar effects in Newtonian mechanics. The center of mass appears to act as an information center: it receives information at the speed of light from the two particles about their advanced or retarded positions and momenta to correlate with the center of motion proper time.

Although the centers of motion and mass and the conservation laws already provide a reduction of the problem to quadratures, the simple form of the reduced equation for the one-dimensional case suggests that there is still more structure within this system remaining to be discovered.

*Supported in part by the Commonwealth Campus Scholarly Activities Fund of The Pennsylvania State University.
[1]D. G. Currie, T. F. Jordan, and E. C. G. Sudarshan, Rev. Mod. Phys. **35**, 350 (1963).
[2]D. G. Currie, J. Math. Phys. **4**, 1470 (1963).
[3]J. T. Cannon and T. F. Jordan, J. Math. Phys. **5**, 299 (1964).
[4]H. Leutwyler, Nuovo Cimento **37**, 556 (1965).
[5]For discussion and a collection of articles on the no-interaction theorems and on progress in action-at-a-distance particle dynamics, see *The Theory of Action-at-a-Distance in Particle Dynamics*, edited by E. H. Kerner (Gordon and Breach, New York, 1973).
[6]A. D. Fokker, Physica **9**, 33 (1929).
[7]A. Staruszkiewicz, Ann. Physik **25**, 362 (1970).
[8]J. A. Wheeler and R. P. Feynman, Rev. Mod. Phys. **21**, 425 (1949).
[9]R. A. Rudd and R. N. Hill, J. Math. Phys. **11**, 2704 (1970).
[10]B. Bruhns, Phys. Rev. D **8**, 2370 (1973).
[11]H. P. Künzle, J. Math. Phys. **15**, 1033 (1974).
[12]H. P. Künzle, Int. J. Theor. Phys. **11**, 395 (1974).
[13]A. Staruszkiewicz, Ann. Inst. H. Poincaré **14**, 69 (1971).
[14]J. L. Synge, *Relativity: The Special Theory* (North-Holland, Amsterdam, 1965), 2nd ed., pp. 219–22.
[15]H. Van Dam and E. P. Wigner, Phys. Rev. **142**, 838 (1966).
[16]P. A. M. Dirac, Can. J. Math. **2**, 129 (1950).
[17]Dirac's generalized Hamiltonian dynamics is discussed in detail in E. C. G. Sudarshan and N. Mukunda, *Classical Dynamics: A Modern Perspective* (Wiley, New York, 1974).
[18]When $n=1$, referring to particle 1, the subscript f assumes the value 2, and vice versa.
[19]M. Lorente and P. Roman, J. Math. Phys. **15**, 70 (1974).

Asymptotic behavior of Feynman integrals with zero mass particles

Edward B. Manoukian

Centre de Recherches Mathématiques, Université de Montréal, Montréal, Canada
(Received 29 November 1976)

The asymptotic behavior of Feynman integrals when some (or all) of their external momenta *become small and/or some (or all) become large* in Euclidean space nonexceptionally (and the power counting theorem) are derived for integrals involving, in general, *zero* mass particles as well as a subset of their masses.

We derive the asymptotic behavior of Feynman integrals when some (or all) of their external momenta become *small and/or* some (or all) become *large* in Euclidean space nonexceptionally (and the power counting theorem in Minkowski space as well) for integrals involving *zero* mass particles as well. This analysis is a generalization, in a very general way, of the previous classic work[1] (and Ref. 2) dealing with the *large* momentum behavior of Feynman integrals with only *nonzero* mass particles.[3]

We consider integrals of the form

$$\int_{-\infty}^{\infty} dk_1 \int_{-\infty}^{\infty} dk_2 \cdots \int_{-\infty}^{\infty} dk_t f(p_1, p_2, \ldots; k_1, k_2, \cdots), \tag{1}$$

where $k_i \in (k_1, k_2, \cdots)$ for $i = 1, 2, \ldots, t$ and f is a function of n real variables considered as the components of an n-vector in an n-dimensional Euclidean space R^n. Let I be an arbitrarily chosen subspace of R^n associated with the integration variables. Choose E to be any subspace of R^n such that E and I are disjoint and $R^n = I + E$. Let $\Lambda(I)$ be the projection operation along the subspace I. The integral (1) may be conveniently rewritten as

$$f_I(\mathbf{P}) = \int_I d\mathbf{P}' f(\mathbf{P} + \mathbf{P}'), \tag{2}$$

and the absolute convergence of the integral (2) implies that $f_I(\mathbf{P})$ depends only on the projection of $\mathbf{P}$ along the subspace I. Let $\mathbf{P}$ be a vector in R^n of the form

$$\mathbf{P} = \sum_{i=1}^{j} \mathbf{L}_i \eta_i \eta_{i+1} \cdots \eta_j + \mathbf{L}_{j+1} + \sum_{i=j+2}^{m} \mathbf{L}_i \lambda_{j+2} \lambda_{j+3} \cdots \lambda_i, \tag{3}$$

where $m \leq n+1$ and $\mathbf{L}_1, \ldots, \mathbf{L}_j, \mathbf{L}_{j+2}, \ldots, \mathbf{L}_m$ are independent vectors and the subsets $\mathbf{L}_1, \mathbf{L}_2, \ldots, \mathbf{L}_r; \cdots$ span subspaces $S \equiv \{\mathbf{L}_1, \mathbf{L}_2, \ldots, \mathbf{L}_r\}; \cdots$. $\mathbf{L}_{j+1}$ is a vector confined to a finite region W in R^n. The parameters $\eta_1, \ldots, \eta_j, \lambda_{j+2}, \ldots, \lambda_m$ are real and nonnegative.

Definition: A function $f(\mathbf{P})$ is said to belong to a class $\mathbf{A}_n$ if and only if, for any nonempty subspaces $S, S' \subseteq R^n$ there exists two pairs of asymptotic coefficients $\alpha(S)$, $\alpha^0(S')$, $\beta(S)$, $\beta^0(S')$, with the latter two nonnegative, such that

$$f\left(\sum_{i=1}^{j} \mathbf{L}_i \eta_i \eta_{i+1} \cdots \eta_j + \mathbf{L}_{j+1} + \sum_{i=j+2}^{m} \mathbf{L}_i \lambda_{j+2} \lambda_{j+3} \cdots \lambda_i\right)$$
$$= O\Big\{\eta_1^{\alpha(\{\mathbf{L}_1\})} \cdots \eta_j^{\alpha(\{\mathbf{L}_1, \ldots, \mathbf{L}_j\})} \lambda_{j+2}^{\alpha^0(\{\mathbf{L}_{j+2}, \ldots, \mathbf{L}_m\})} \cdots \lambda_m^{\alpha^0(\{\mathbf{L}_m\})}$$
$$\times \sum_{\gamma_1, \ldots, \gamma_j} (\ln \eta_{\pi_1})^{\gamma_1} \cdots (\ln \eta_{\pi_j})^{\gamma_j}$$
$$\times \sum_{\gamma_{j+2}^0, \ldots, \gamma_m^0} (\ln 1/\lambda_{\pi_{j+2}})^{\gamma_{j+2}^0} \cdots (\ln 1/\lambda_{\pi_m})^{\gamma_m^0}\Big\}$$

for $\eta_1, \eta_2, \ldots, \eta_j \to \infty$ and $\lambda_{j+2}, \lambda_{j+3}, \ldots, \lambda_m \to 0$, independently, where the sums are over all nonnegative integers $\gamma_1, \ldots, \gamma_j$; $\gamma_{j+2}^0, \ldots, \gamma_m^0$ such that

$$\sum_{r=1}^{i} \gamma_r \leq \beta(\{\mathbf{L}_1, \ldots, \mathbf{L}_{\pi_i}\}), \quad 1 \leq i \leq j,$$

$$\sum_{r=j+2}^{i} \gamma_r^0 \leq \beta^0(\{\mathbf{L}_{\pi_i}, \ldots, \mathbf{L}_m\}), \quad j+2 \leq i \leq m,$$

with the asymptotic coefficients $\beta(\{\mathbf{L}_1\}), \ldots, \beta(\{\mathbf{L}_1, \ldots, \mathbf{L}_j\})$; $\beta^0(\{\mathbf{L}_{j+2}, \ldots, \mathbf{L}_m\}), \ldots, \beta^0(\{\mathbf{L}_m\})$ arranged in increasing order,

$$\beta(\{\mathbf{L}_1, \ldots, \mathbf{L}_{\pi_1}\}) \leq \beta(\{\mathbf{L}_1, \ldots, \mathbf{L}_{\pi_2}\}) \leq \cdots \leq \beta(\{\mathbf{L}_1, \ldots, \mathbf{L}_{\pi_j}\}),$$

$$\beta^0(\{\mathbf{L}_{\pi_{j+2}}, \ldots, \mathbf{L}_m\}) \leq \beta^0(\{\mathbf{L}_{\pi_{j+3}}, \ldots, \mathbf{L}_m\}) \leq \cdots$$
$$\leq \beta^0(\{\mathbf{L}_{\pi_m}, \ldots, \mathbf{L}_m\}),$$

and $(\pi_1, \ldots, \pi_j)$; $(\pi_{j+2}, \ldots, \pi_m)$ are a permutation of $(1, \ldots, j)$; $(j+2, \ldots, m)$, respectively.

Theorem: If a function $f(\mathbf{P})$ belongs to class $\mathbf{A}_n$, and for *any* given $I \subseteq R^n$, is integrable over *any* finite region of I or of any of its, arbitrarily decomposed, subspaces excluding the origin and simultaneously $\alpha^0(S'') + \dim S'' > 0$, $\alpha(S') + \dim S' < 0$ for all S', $S'' \subseteq I$, then $f_I(\mathbf{P})$ exists and belongs to class $\mathbf{A}_{n-t}$ with asymptotic coefficients

$$\alpha_I(\bar{S}') = \max_{\Lambda(I)\tilde{S} = \bar{S}'} [\alpha(\tilde{S}) + \dim \tilde{S} - \dim \bar{S}'], \tag{4}$$

$$\alpha_I^0(\bar{S}'') = \min_{\Lambda(I)\tilde{S} = \bar{S}''} [\alpha^0(\tilde{S}) + \dim \tilde{S} - \dim \bar{S}''], \tag{5}$$

for all $\bar{S}'$ and $\bar{S}'' \subseteq E$ (with $\bar{S}'$ and $\bar{S}''$ defined with respect to the vector $\mathbf{P}$ in $f_I(\mathbf{P})$, etc.) with Eqs. (4) and (5) defined in reference to the parameters η and λ, respectively. And

$$\beta_I(\bar{S}') = \max_{\tilde{S}_{\text{maximizing}}} \beta(\tilde{S}) + \sum_{i=1}^{t} p_i, \tag{6}$$

$$\beta_I^0(\bar{S}'') = \max_{\tilde{S}_{\text{minimizing}}} \beta^0(\tilde{S}) + \sum_{i=1}^{t} p_i^0, \tag{7}$$

where in (6) max runs over all $\tilde{S}$ maximizing the expression (4) and in (7) max runs over all $\tilde{S}$ minimizing the expression (5). The parameters p_i, p_i^0 take on the values 0 or 1. Write I as a decomposition into t, arbitrarily chosen, one-dimensional disjoint subspaces $I_1, I_2, \ldots, I_t : I = I_1 + I_2 + \cdots + I_t$. If all the minimizing subspaces for an I_i integration, after performing the $I_{i+1}+I_{i+2}+\cdots+I_t$ integration, relative to any one of the minimizing subspaces for the I_{i-1} integration after performing the $I_i+I_{i+1}+\cdots+I_t$ integration have the same dimension then $p_i^0=0$, and 1 otherwise, for $i=2,\ldots,t$. If all the minimizing subspaces for the I_1 integration relative to $\bar{S}''$ in (7) after performing the $I_2+I_3+\cdots+I_t$ integration have the same dimension then $p_1^0=0$, and 1 otherwise. Similarly p_i is defined by replacing minimizing by maximizing and $\bar{S}''$ by $\bar{S}'$ in the definition just given for p_i^0.

To prove the theorem we proceed by induction[1] and suppose that f when integrated out over *any* chosen $1\leqslant j<t$ of its variables, or any of their linear recombinations belongs to class $\mathbf{A}_{n-j}$ and together with all its subintegrals satisfy the criteria of the theorem. Denote such a j-fold integral by F. We then show that an integral of the form $\int_{-\infty}^{\infty} dk_1\, F(\mathbf{P}+\mathbf{L}k_1)$ enjoys the same properties and belongs to class $\mathbf{A}_{n-j-1}$, where k_1 is arbitrarily chosen. The asymptotic coefficients of F will be denoted by α', β', $\alpha^{0\prime}$, $\beta^{0\prime}$. $\mathbf{P}$ is a vector of the form in (3) with $0\leqslant m\leqslant n-j-1$. By the Heine–Borel theorem such an infinite interval $\{k_1:-\infty\leqslant k_1\leqslant\infty\}$ may be covered[1] by a finite number of subintervals of the forms:

$$J_1=\{k_1:k_1=z\eta_1\eta_2\cdots\eta_j;\ |z|\geqslant b_z,$$

$$\text{for some number } b_z>1,\ z=\pm|z|\},$$

$$J_>^{i_1\cdots i_r}=\Big\{k_1:k_1=\sum_{s=1}^{r}A_{i_1i_2\cdots i_s}\eta_s\eta_{s+1}\cdots\eta_j+z\eta_{r+1}\cdots\eta_j;$$

$$\omega_{i_1\cdots i_r}\eta_r\geqslant|z|\geqslant b_{i_1\cdots i_r},$$

$$\text{for some numbers } b_{i_1\cdots i_r}>1 \text{ and}$$

$$0<\omega_{i_1\cdots i_r}<1;\ A_{i_1},A_{i_1i_2},\ldots,A_{i_1\cdots i_r}\in R^1;\ z=\pm|z|\Big\},\quad 1\leqslant r<j,$$

$$J_>^{i_1\cdots i_j}=\Big\{k_1:k_1=\sum_{s=1}^{j}A_{i_1i_2\cdots i_s}\eta_s\eta_{s+1}\cdots\eta_j+z;$$

$$\omega_{i_1\cdots i_j}\eta_j\geqslant|z|\geqslant b_{i_1\cdots i_j},$$

$$\text{for some } b_{i_1\cdots i_j}>1 \text{ and } 0<\omega_{i_1\cdots i_j}<1;$$

$$A_{i_1},A_{i_1i_2},\ldots,A_{i_1\cdots i_j}\in R^1,\ z=\pm|z|\Big\},$$

$$J_2^{i_1\cdots i_j}=\Big\{k_1:k_1=\sum_{s=1}^{j}A_{i_1\cdots i_s}\eta_s\cdots\eta_j+z;$$

$$b_{i_1\cdots i_j}\geqslant|z|\geqslant C_{i_1\cdots i_{j+1}},$$

$$\text{for some } b_{i_1\cdots i_j}>1 \text{ and } 0<C_{i_1\cdots i_{j+1}}<1,$$

$$A_{i_1},\ldots,A_{i_1\cdots i_j}\in R^1,\quad z=\pm|z|\Big\},$$

$$J_<^{i_1\cdots i_r}=\Big\{k_1:k_1=\sum_{s=1}^{j}A_{i_1\cdots i_s}\eta_s\cdots\eta_j$$

$$+\sum_{s=j+1}^{r}A_{i_1\cdots i_s}\lambda_{j+1}\lambda_{j+2}\cdots\lambda_s$$

$$+\lambda_{j+1}\lambda_{j+2}\cdots\lambda_r z;\ \lambda_{j+1}\equiv 1,$$

$$C_{i_1\cdots i_n}\geqslant|z|\geqslant\tilde{\omega}_{i_1\cdots i_r}\lambda_{r+1},$$

$$\text{for some } 0<C_{i_1\cdots i_r}<1 \text{ and } \tilde{\omega}_{i_1\cdots i_r}>1,$$

$$A_{i_1},\ldots,A_{i_1\cdots i_r}\in R^1,\quad z=\pm|z|\Big\},\quad j+1\leqslant r<m,$$

$$J_<^{i_1\cdots i_m}=\Big\{k_1:k_1=\sum_{s=1}^{j}A_{i_1\cdots i_s}\eta_s\cdots\eta_j+A_{i_1\cdots i_{j+1}}$$

$$+\sum_{s=j+2}^{m}A_{i_1\cdots i_s}\lambda_{j+2}\cdots\lambda_m+\lambda_{j+2}\cdots\lambda_m z;$$

$$C_{i_1\cdots i_m}\geqslant|z|\geqslant 0 \text{ for some } 0<C_{i_1\cdots i_m}<1,$$

$$A_{i_1},\ldots,A_{i_1\cdots i_m}\in R^1,\quad z=\pm|z|\Big\}.$$

The properties of the parameters in each subinterval are dictated by the covering process of the Heine–Borel theorem, and more information on them than the one given above is not essential here. The integration over k_1 then reduces to ones belonging successively to each of the subintervals above with the positive integers $i_1,\ldots,i_m$ taking on a finite set of values. Integrations over subintervals such as $J_1, J_>^{i_1\cdots i_r}, J_>^{i_1\cdots i_j}, J_2^{i_1\cdots i_j}$ have been treated in Ref. 1, and, if $\alpha'(\{\mathrm{L}\})+1<0$, then the integration over J_1 is absolutely convegent for $\infty\geqslant|z|\geqslant b_z$. It is easily seen that if $\alpha^{0\prime}(\{\mathrm{L}\})+1>0$, then the integral over the subinterval $J_<^{i_1\cdots i_m}$ is absolutely convergent with $C_{i_1\cdots i_m}\geqslant|z|\geqslant 0$, $0<C_{i_1\cdots i_m}<1$. The integrations over the subintervals $J_<^{i_1\cdots i_r}$ are carried out essentially in the same way as the ones over the $J_>^{i_1\cdots i_r}$ subintervals (by making the formal replacements λ_{r+1}, $|z|\to 1/\lambda_{r+1}$, $1/|z|$) with $|z|\leqslant C_{i_1\cdots i_r}$ for some $0<C_{i_1\cdots i_r}<1$, $\lambda_{r+1}/|z|\leqslant C_{z,i_1\cdots i_r}$ for some $0<C_{z,i_1\cdots i_r}<1$ and $\tilde{\omega}_{i_1\cdots i_r}\geqslant C_{z,i_1\cdots i_r}^{-1}$, with the range: $C_{i_1\cdots i_r}\geqslant|z|\geqslant\tilde{\omega}_{i_1\cdots i_r}\lambda_{r+1}$.

Let I' be written as the disjoint union of two subspaces I_1' and I_2', where I_1' is a one-dimensional subspace associated with the integration over k_1 and where I' is associated with $j+1$ integration variables. According to the induction hypothesis $\alpha'(\tilde{S}')+\dim\tilde{S}'<0$ and $\alpha^{0\prime}(\tilde{S}'')+\dim\tilde{S}''>0$ for $\tilde{S}'$, $\tilde{S}''\subseteq I_2'$; and

$$\alpha_{I_2'}(S')\equiv\alpha'(S')=\max_{\Lambda(I_2')\tilde{S}=S'}[\alpha(\tilde{S})+\dim\tilde{S}-\dim S'],$$

$$\alpha_{I_2'}^0(S'')\equiv\alpha^{0\prime}(S'')=\min_{\Lambda(I_2')\tilde{S}=S''}[\alpha^0(\tilde{S})+\dim\tilde{S}-\dim S''].$$

From the above, the infrared convergence criterion $\alpha^{0\prime}(\{\mathrm{L}\})+1>0$ through $J_<^{i_1\cdots i_m}$ and the ultraviolet convergence one $\alpha'(\{\mathrm{L}\})+1<0$ through J_1 we readily obtain the result stated in the theorem, by induction, in a standard manner by the same reasoning as in the earlier case[1] since, in particular the j variables, the associated integration subspaces are arbitrary and the joint $(j+1)$-dimensional integral is absolutely convergent with the easily derivable results

$$\min_{S\subseteq I'}[\alpha^0(S)+\dim S]>0$$

and[1]

$$\max_{S\subseteq I'}[\alpha(S)+\dim S]<0,$$

in particular, and the iterated integrals all yield a unique result by Fubini's theorem; and[1] $\Lambda(I_1')S'=S$,

$\Lambda(I_2')S''=S'$ are equivalent to $\Lambda(I')S''=S$ with I_1', I_2' and S disjoint. Proof of (6) and (7) is identical to the one given in Ref. 2 with elementary modifications, notably in Lemma 1, and will not be repeated.

The above analysis is to be applied to Feynman integrals with integrands of the form

$$I_0=P/\prod_i (Q_i^2+\mu_i^2)^{\gamma_i}, \quad \mu_i^2\geq 0,\ \gamma_i>0 \tag{8}$$

with

$$P=\sum A^i_{m_{i1},\ldots,n_{i1},\ldots,\tau_{i1}}\prod_j (p_j)^{m_{ij}}\prod_j (k_j)^{n_{ij}}\prod_s (\mu_s)^{\tau_{is}}, \tag{9}$$

where the sum is over nonnegative integers m_{ij}, n_{ij}, τ_{is}, i; and

$$Q_i=\sum_j a_j^i p_j+\sum_{j=1}^{t} b_j^i k_j.$$

For every term $(Q_i^2+\mu_i^2)^{\gamma_i}$ we introduce a vector $\mathbf{V}_j$ in R^n such that $\mathbf{V}_j\cdot\mathbf{P}=Q_j$ (4-vector indices are suppressed) and $\mathbf{P}$ is a vector in the form in (3) with j replaced by i. Similarly we introduce vectors $\mathbf{V}_j$ to define p_j and k_j in (9). We readily see that

$$\begin{aligned}
I_0 &= 0\{\eta_1^{\alpha(\{\mathbf{L}_1\})}\cdots\eta_i^{\alpha(\{\mathbf{L}_1,\ldots,\mathbf{L}_i\})}\\
&\quad\times\lambda_{i+2}^{\alpha^0(\{\mathbf{L}_{i+2},\ldots,\mathbf{L}_m\})}\cdots\lambda_m^{\alpha^0(\{\mathbf{L}_m\})}\},\\
\alpha(\{\mathbf{L}_1,\ldots,\mathbf{L}_r\}) &= -2\sum_j\gamma_j+\max_i\Big(\sum_j m_{ij}+\sum_j n_{ij}\Big),\\
&\quad 1\leq r\leq i,
\end{aligned}\tag{10}$$

where the sums are over all j's such that $\mathbf{V}_j$ is not orthogonal to $\{\mathbf{L}_1,\ldots,\mathbf{L}_r\}$ and max is taken for i in the expression (9). And

$$\begin{aligned}
\alpha^0(\{\mathbf{L}_r,\ldots,\mathbf{L}_m\}) &= -2\sum_j{}'\gamma_j+\min_i\Big(\sum_j m_{ij}+\sum_j n_{ij}\Big),\\
&\quad i+2\leq r\leq m,
\end{aligned}\tag{11}$$

where the sumes are over all j's such that the $\mathbf{V}_j$'s are *orthogonal* to $\{\mathbf{L}_1,\ldots,\mathbf{L}_{r-1}\}$ *and* $\mathbf{L}_{i+1}$ (if $\mathbf{L}_{i+1}$ is not the zero vector) but not orthogonal to the space $\{\mathbf{L}_r,\ldots,\mathbf{L}_m\}$, $i+2\leq r\leq m$ with $\mathbf{L}_{i+1}$ excluded in $\{\mathbf{L}_1,\ldots,\mathbf{L}_{r-1}\}$; min is taken for i in the expression (9) and the first sum is restricted only to those j's with $\mu_j^2\equiv 0$ in $\prod_i (Q_i^2+\mu_i^2)^{\gamma_i}$. The vector $\mathbf{L}_{i+1}$ (for $\mathbf{L}_{i+1}\neq 0$) is a characteristic of the momenta which are not asymptotic (i.e., neither $p\to\infty$ nor $p\to 0$). As the parameters $\eta_1,\ldots,\eta_i$ are taken, independently, large and the parameters $\lambda_{i+2},\ldots,\lambda_m$ are taken, independently, small, no partial sums of the asymptotic momenta p_{i1}, p_{i2}, $\cdots$, some of which becoming small ($\neq 0$) and/or some becoming large can vanish, i.e., the asymptotic momenta, are nonexceptional.

Application of the above work and especially to renormalized Feynman amplitudes will be given in subsequent work.

[1]S. Weinberg, Phys. Rev. **118**, 838 (1960).
[2]J. P. Fink, J. Math. Phys. **9**, 1389 (1968).

[3]For a different approach to the power counting theorem, alone, in general see: W. Zimmermann, Commun. Math. Phys. **11**, 1 (1968); Z.H. Lowenstein and W. Zimmermann, Commun. Math. Phys. **44**, 73 (1975) [also J.H. Lowenstein and E. Speer, *ibid.* **47**, 43 (1976)] the conclusion of which agrees with the part of our work restricted to the power counting theorem. By writing a propagator carrying a momentum Q_i by a polynomial in Q_i times $[Q_i^2+\mu_i^2-i\epsilon(Q_i^2+\mu_i^2)]^{-1}$ for $\mu_i^2\geq 0$ the limit $\epsilon\to 0$, in Minkowski space, as a covariant distribution may be then also carried out.

Quantum theory and measures on Hilbert space

H. Krips

History and Philosophy of Science Department, Melbourne University, Parkville, Victoria 3052, Australia
(Received 28 September 1976)

It is shown that merely the decision to represent the fundamental events of quantum theory by vectors in a Hilbert space uniquely determines the form of any probability measure over the set of events. The proof proceeds by showing that any nonnegative measure on the vectors of a Hilbert space, for which the sum of the measures on the vectors of any complete orthonormal set is 1, takes the usual form demanded by quantum theory. This result represents a considerable strengthening of one of the consequences of Gleason's theorem.

I. INTRODUCTION

The proof which will be presented here incorporates an unpublished suggestion for a proof by Dorling.[1] Indeed several moves made in our proof correspond to moves sketched at points by Dorling, as well as by other authors in the literature (see later).[2]

The central theorem which we shall prove is:

Theorem T: Let m be any measure on the unit norm vectors of a Hilbert space H, separable, and of more than two dimensions, for which

(1) $0 \le m(\phi) \le 1$ for any ϕ in H of unit norm,

(2) if $\{\phi_i\}$ is any complete orthonormal (c.o.n.) set of vectors in H, then $\sum m(\phi_i) = 1$,

(3) there is a Ψ in H for which $m(\Psi) = 1$, Ψ of unit norm.

Then $m(\phi) = |\langle \Psi, \phi\rangle|^2$, for any ϕ in H of unit norm.

Before proving this theorem—which will occupy us in Secs. II—III—we shall discuss the physical significance of the theorem for quantum theory (Q.T.).[3]

In Q.T. we associate with each system S at time t a separable Hilbert space H. And with each nondegenerate variable A for S at t we associate a c.o.n. set of vectors $\{\phi_i\}$, the "eigenvectors" of A, one for each of the values of A. We shall let a_i be the value of A associated with the eigenvector ϕ_i of A. We associate a probability with each value a_i of A for S at t. We denote it by $P[A, i]$ (suppressing S, t indices). How this probability is interpreted does, of course, vary from one treatment of Q.T. to the others—sometimes it is the conditional probability of measuring A to have value a_i in S at t, conditional on A being measured in S at t; and sometimes it is the probability of A *having* the value a_i in S at t. But, in whatever way it is interpreted, the following hold:

(1)′ $1 \ge P[A, i] \ge 0$,

(2)′ $\sum_i P[A, i] = 1$.

Moreover, if S at t is in a "pure state", then, by definition,

(3)′ $P[B, j] = 1$ for some B, j.

It is important to note that (1)′ and (2)′ hold just as a matter of *probability theory*—and not of Q.T. This is because we so define a "measurement" that the events of measuring A to have value a_i in S at t, for various i, are mutually exclusive and exhaustive—as are the events of A having value a_i in S at t, for various i.

Now we introduce the central tenets of Q.T.:

(I) $P[A, i] = m(\phi_i)$ for some function m [4],

(II) every c.o.n. set of vectors in H is the set of the eigenvectors of some nondegenerate variable.

What (I) says should not be underestimated—it says that not only is there an eigenvector for each $\langle A, a_i\rangle$, but *also*, if $\langle A, a_i\rangle$ and $\langle B, b_j\rangle$ share the same eigenvector—despite A and B being *different* variables—then $P[A, i] = P[B, j]$.[5] In other words dependence on $\langle A, i\rangle$ can be taken as dependence on the representative eigenvector ϕ_i.

(II) says that eigenvectors provide a 1 : 1 representation of the variables and their values—it says that not only is there a c.o.n. set of vectors for each variable and its values, but visa versa. Thus (I) and (II) indicate that the Hilbert space formalism is to be taken seriously as providing representatives for the physical variables of a system and their values.

With the help of (I), (II), and (1)′—(3)′, we immediately see that the function m satisfies the conditions (1)—(3); and hence , from theorem T, we derive the Born interpretation:

If S at t is in the pure state Ψ, then $P[A, i] = |\langle \phi_i, \Psi\rangle|^2$. Thus, from merely adopting probability theory and (I) and (II) (which specify how the Hilbert space formalism is to map the "real world"), we get out the fundamental law of Q.T.—a law which gives rise to all the physically significant "quantum effects," such as superposition interference, etc.

Other attempts have of course been made to obtain a similar result.[6,7] Of these, the strongest result is that of Gleason.[6] Gleason considers a measure m on *projection operators* onto H—rather than on the vectors of H directly. It satisfies:

(1)″ $m(E) \ge 0$,

(2)″ $m(I) = 1$, where I is the identity operator on H,

(3)″ if $EF = 0$, then $m(E+F) = m(E) + m(F)$.

He then derives that there is a positive definite Hermitian unit trace (p.h.u.t.) operator W for which $m(E) = \mathrm{Tr}\, WE$, for any E. He can then introduce, as central tenets of Q.T.:

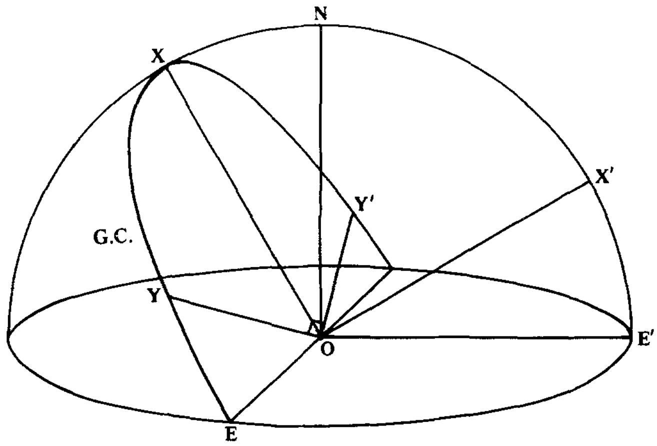

FIG. 1.

(I)′ $P[A, i] = m'(E_i)$ for some function m', where $E_i = P[\phi_i]$ (the projection operator into ϕ_i),

(II)′ as for II,

(III)^r $m'(E+F) = m'(E) + m'(F)$ if $EF = 0$.

From (2)′, we can then see that $m'(I) = 1$, and hence that m' satisfies (1)″ – (3)″, and hence we derive:

For any S at t there is a "density operator" W for which $P[A, i] = \mathrm{Tr}\, W E_i$.

We can then derive, as a special case, the "Born interpretation," via the trivial intermediate theorem:

If $P[B, j] = 1$ for some B, j, then $W = E_j'$, where E_j' is the projector onto the eigenvector of B for value b_j.

(This follows trivially from the W being p.h.u.t.)

What are the advantages/disadvantages of the approach to the Born interpretation via (I), (II) and Theorem T, as compared to the approach via (I)′–(III)′ and Gleason's theorem (let us call it G)?

The G approach has the obvious advantage that we derive that there are only two sorts of states in Q.T.—viz. "mixed states," for which there is a density operator which is *not* a projection operator; *and* "pure states." In the T approach, more axioms have to be added to derive this restriction. We show how this can plausibly be done in Ref. 8 [where we also incidentally show how the restriction to dimension of $H > 2$ can be removed, as well as arguing for (II)]. On the other hand, the *disadvantage* with the G approach, is that the assumption (III)′ is intuitively less plausible than (2)′ (its analog in the T approach), because (2)′ follows just from probability theory. Because of this we would argue that the G approach is over-all less preferable, as a basis for axiomatizing Q.T.—even though its axioms may be logically more integrated when we come to derive the existence of density operators for all states in Q.T. As a second point in favor of considering the T approach, we note that the G approach uses logically stronger axioms [this is indicated by the fact that from (I)′–(III)′, but not from (I), (II) alone, can we derive that states in Q.T. are of the linear "density operator form"]. Thus, even were the G approach to be preferred over-all as a basis for axiomatizing Q.T., the T approach is of interest in that it shows that the Born interpretation can be derived from the weaker (I), (II).

Finally we note that in a previous paper,[8] we also presented a derivation of the Born interpretation, but we needed an extra assumption of rotational symmetry for m, in order to establish the continuity of m. Thus one of the interesting consequences of Theorem T is that rotational symmetry of m in H (and hence its continuity too) is derived instead of assumed.

II. PROOF OF THEOREM 1

The proof of Theorem T will be presented in several steps. First we will present a series of lemmata, which will lead up to a proof of:

Theorem 1: Let m be a map of the points of a three-dimensional real Euclidian sphere S onto $[0, 1]$, such that

(a) the north pole N has value 1 $[m(N) = 1]$,

(b) $\sum m(P_i) = 1$ for any orthotriad of points P_1, P_2, P_3.[9]

Then $m(P) = (\sin\theta)^2$, where θ is the latitude of P (as measured from the equator).

We will later generalize Theorem 1 to a hypersphere in an N-dimensional *complex* Hilbert space, in order to prove Theorem T. The lemmata and theorems in this section will all be taken to refer to points on the sphere S, for which there is a mapping m satisfying (a), (b) above.

Lemma 1: Let X be the northmost point (the "apex") of a great circle GC through Y. Then $m(X) \geq m(Y)$.[10] Also $m(E) = 0$ for any equatorial E.

Proof: There is a orthotriad of points X, X', E where E is on the equator, and X' is orthogonal to GC. (see Fig. 1).

$\therefore$ $m(X) + m(X') + m(E) = 1$.

But there is also an orthotriad E, E', N, where E' is also equatorial.

$\therefore$ $m(E) + m(E') + m(N) = 1$

But $m(N) = 1$, and hence $m(E) = 0$. $\therefore$

(i) $m(X) + m(X') = 1$.

But since X' is orthogonal to GC, there is an orthotriad X', Y, Y', where Y' is also on GC. Hence

(ii) $m(Y) + m(Y') + m(X') = 1$.

From (i) and (ii), we see that

$m(X) \geq m(Y)$ since $m(Y') \geq 0$. QED

Lemma 2: Suppose X and Y are both northerly points, and suppose Y is to the south of, but at the same longitude as X. Then $m(X) \geq m(Y)$.

Proof: Construct the great circle GC with X as apex—let its equatorial diameter be AB. Construct the set of all great circles through Y,[11] and let L be the locus of their apexes. Obviously Y itself will be one such apex, corresponding to the great circle ABY, and N will be another such apex corresponding to the great circle

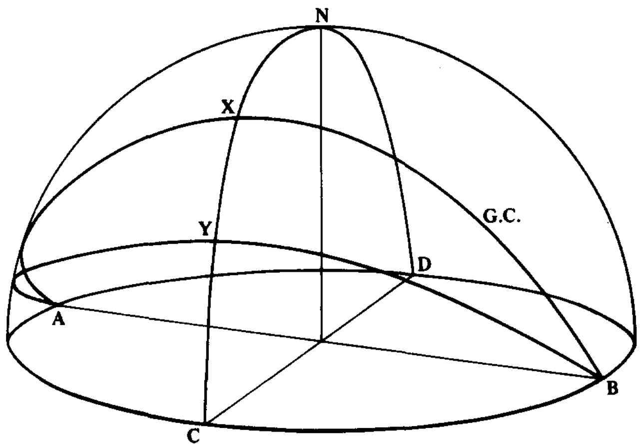

FIG. 2.

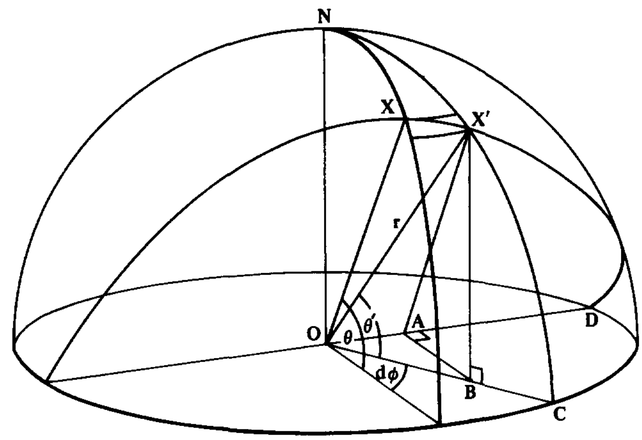

FIG. 3.

CND (see Fig. 2.) Hence L will be a continuous line joining Y and N (traced out as AB turns into CD—see Ref. 11). Obviously L must intersect GC. (The only way a line can join N and Y is by intersecting GC.) Let the point of intersection be P.

Then, by Lemma 1, and since P is an apex of a great circle through Y, we have that

(i) $m(P) \geqslant m(Y)$.

But, by Lemma 1, and since P is on GC, the apex of which is X, we have that

(ii) $m(X) \geqslant m(P)$.

Hence by (i) and (ii),

$m(X) \geqslant m(Y)$. QED

Lemma 3: Let X be the apex of a great circle which also passes through X', let $d\phi$ be the difference in longitude between X and X', and let X have latitude θ and X' have latitude θ'. Then $\tan\theta'/\tan\theta = \cos d\phi$.

Proof: Let the meridian of longitude for X' intersect the equator at C. Join OC. Project X' down to intersect OC orthogonally at B. Project a perpendicular from B across to intersect OD at A. (See Fig. 3.) $\angle X'BA$ is obviously $90°$ since $X'B$ is normal to the equatorial plane (i.e., $X'B \perp OC$ by construction, and $X'B$ is in a meridianal plane). Join $X'A$. Let OX' be r in length. Now $OB = r\cos\theta'$ (since $\angle X'BA = 90$), and $X'B = r\sin\theta'$. Also $\angle X'AB = \theta$ (since a great circle is inclined at the same angle to the equatorial plane, at all points along its equatorial diameter, that angle being the latitude of its apex); and hence

$$X'B = AB\tan\theta \quad (\text{since } \angle X'BA = 90).$$

Hence $AB = r\sin\theta'/\tan\theta$. But $\angle ABO = d\phi$ (alternate angles); and hence

$$\cos d\phi = AB/OB = r\sin\theta'/(\tan\theta \cdot r\cos\theta')$$
$$= \tan\theta'/\tan\theta \qquad \text{QED}$$

Lemma 4: Let X and X' be two northerly points where X' is to the south of X, but on a different longitude. Then $m(X') \leqslant m(X)$.

Proof: Construct a great circle GC_1 with X as apex. Let X_1 be a point on GC_1, which has a longitude $d\phi$ closer to X' than X has. (See Fig. 4.) Then, by Lemma 3, $\tan\theta_1/\tan\theta = \cos d\phi$, where θ_1 and θ are the latitudes of X_1 and X respectively. Construct a great circle GC_2 with X_1 as apex. Let X_2 be a point on GC_2 which has a longitude $d\phi$ closer to X' then X_1 has. Then $\tan\theta_2/\tan\theta_1 = \cos d\phi$, by Lemma 3. Construct a set of N such points $X_1, X_2, \ldots, X_N$, and let X_N have the same longitude as X'; i.e., $N\,d\phi = \Delta$ where Δ is the difference in longitude between X and X'. Then $\tan\theta_n/\tan\theta_{n-1} = \cos d\phi$ for all $n = 1, \ldots, N$, by Lemma 3, and hence

$$\frac{\tan\theta_1}{\tan\theta}\,\frac{\tan\theta_2}{\tan\theta_1}\,\frac{\tan\theta_N}{\tan_{N-1}} = \prod_{n=1}^{N}\cos d\phi,$$

i.e., $\tan\theta_N/\tan\theta = (\cos\Delta/N)^N$.

But $(\cos\Delta/N)^N \to 1$ as $N \to \infty$. Hence, by letting N be large enough, θ_N can be made as near to θ as one likes. In particular, since $\theta > \theta'$, we see that for large enough N, $\theta_N > \theta'$ also; i.e., we see that X_N is at the same longitude as X', but to *its north*.

Now by Lemma 1, we see that $m(X_n) \leqslant m(X_{n-1})$ for all $n = 2, \ldots, N$, and that $m(X) \geqslant m(X_1)$. Hence $m(X) \geqslant m(X_N)$. But, from Lemma 2, we see that $m(X_N) \geqslant m(X')$, since X_N is on the same longitude but to the north of X'. Hence $m(X) \geqslant m(X')$. QED

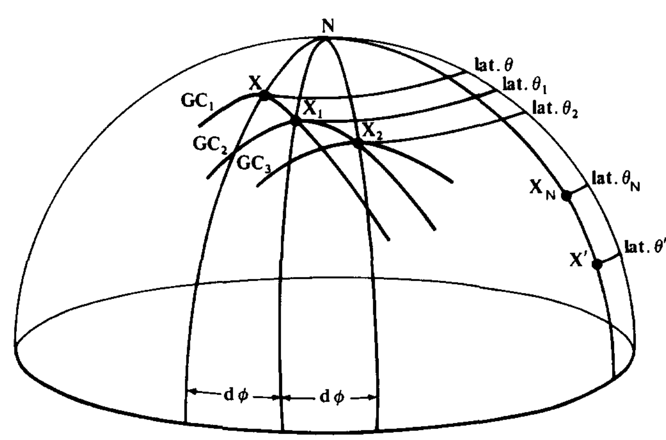

FIG. 4.

Lemma 5: m is a constant of latitude and a continuous function of latitude, at *at least* all points in the northern hemisphere other than equatorial and polar points.

Proof: (All points will be assumed to be northerly.) Let $m_{\max}(\theta)$ be the g.l.b. (greatest lower bound) of all the $m(X)$ for X at latitudes $> \theta$. Since $m(x_1) < m(X)$ for any x_1 at latitude θ and any X at latitudes $> \theta$ (by Lemma 4), it follows that $m(x_1)$ is a l.b. on the set of $m(X)$; and hence

(i) $m(x_1) \leq m_{\max}(\theta)$ for any x_1 at latitude θ.

Let $m_{\min}(\theta)$ be the l.u.b. (least upper bound) on all the $m(X)$ for X at latitudes $< \theta$. Similarly to (i), we can prove:

(ii) $m(x_1) \geq m_{\min}(\theta)$ for any x_1 at latitude θ.

Then let $\{\theta_n\}$ be any denumerable set of latitudes for which $(m_{\max}(\theta_n) - m_{\min}(\theta_n)) \geq \epsilon$, for some $\epsilon > 0$, where $\theta_n > \theta_{n'}$ for any $n > n'$. Also let $\{X_n\}$ be any set of points such that X_n is above latitude θ_n and below latitude θ_{n+1} (if there is one). Then

$$
\begin{aligned}
m(X_1) &\geq m_{\max}(\theta_1) \quad \text{(by definition of g.l.b.)}\\
&\geq \epsilon + m_{\min}(\theta_1) \quad (\textit{ex hypothesi}),\\
&\geq \epsilon \quad (\text{since } m_{\min}(\theta_1) \geq 0).
\end{aligned}
$$

Also

$$
\begin{aligned}
m(X_2) &\geq m_{\max}(\theta_2)\\
&\geq m_{\max}(\theta_2) - (m_{\min}(\theta_2) - m(X_1))
\end{aligned}
$$

[since $m_{\min}(\theta_2) \geq m(X_1)$, X_1 being below θ_2 *ex hypothesi*]

$$
\begin{aligned}
&\geq (m_{\max}(\theta_2) - m_{\min}(\theta_2)) + m(X_1)\\
&\geq \epsilon + \epsilon\\
&\geq 2\epsilon.
\end{aligned}
$$

By induction, we easily prove that

$$m(X_n) \geq n\epsilon;$$

and hence, since $m(X_n) \leq 1$,

$$n \leq 1/\epsilon.$$

Hence there is a *finite* upper bound on the *number* of latitudes θ at which $(m_{\max}(\theta) - m_{\min}(\theta)) \geq \epsilon$. Let the set of those latitudes be the set C_ϵ.

Then what we have shown is that the cardinality of $C_\epsilon \leq 1/\epsilon$, for any $\epsilon > 0$.

Part 2: Here we show that if C_ϵ has even *one* member for $\epsilon > 0$, then $C_{\epsilon/2}$ has *indefinitely* many members, which contradicts the final result of Part 1. Hence we show that C_ϵ is empty for any $\epsilon > 0$.

Now let X be at a latitude which belongs to C_ϵ. Obviously $X \neq N$, since for θ to be a latitude in C_ϵ, $m_{\max}(\theta)$ must exist—and obviously $m_{\max}(90)$ does *not* exist.

Let P be the plane orthogonal to X. The intersection of the sphere S with P will then contain all pairs of points $\langle Y_1, Z_1\rangle$, $\langle Y_2, Z_2\rangle, \ldots, \langle Y_n, Z_n\rangle, \cdots$, which form orthotriads that include X; *and* any point on the intersection of S with P is a member of such a pair. We shall now show that any such Y_n is in $C_{\epsilon/2}$.

Let θ_n be the latitude of Y_n. Then for any $\epsilon' > 0$, there is a Y' north of Y_n for which $m_{\max}(\theta_n) + \epsilon' > m(Y')$—otherwise it would be $m_{\max}(\theta_n) + \epsilon'$ which would be the g.l.b. on all the $m(Y)$ for Y north of Y_n. Now let Y_n' be normal to Z_n, *and* north of Y_n but south of Y' (see Fig. 5). Then, by Lemma 4, $m(Y_n') \leq m(Y')$; and hence

(i) $m(Y_n') \leq m_{\max}(\theta_n) + \epsilon'$.

Let X' be normal to Z_n and to Y_n'. Then obviously X' is *south* of X. (This follows trivially from Lemma 6—see later.) Similarly for any $\epsilon'' > 0$ there is a Y_n'' normal to Z_n and south of Y_n for which

(ii) $m(Y_n'') \geq m_{\min}(\theta_n) - \epsilon''$.

Let X'' be normal to Z_n and Y_n''. Obviously X'' is north of X. Now because $\langle X'', Y_n'', Z_n\rangle$ and $\langle X', Y_n', Z_n\rangle$ are orthotriads,

$$m(X'') + m(Y_n'') + m(Z_n) = 1,$$

$$m(X') + m(Y_n') + m(Z_n) = 1;$$

and hence

$$m(X'') - m(X') = m(Y_n') - m(Y_n'').$$

But, by definition of l.u.b. and g.l.b.,

$$m(X'') - m(X') \geq m_{\max}(\theta) - m_{\min}(\theta),$$

where, since the latitude of X is in C_ϵ,

$$m_{\max}(\theta) - m_{\min}(\theta) \geq \epsilon.$$

Hence

$$m(Y_n') - m(Y_n'') \geq \epsilon;$$

so that from (i) and (ii), for *any* $\epsilon, \epsilon' > 0$, $m_{\max}(\theta_n) - m_{\min}(\theta_n) + \epsilon' + \epsilon'' \geq \epsilon$.

Hence, letting $\epsilon' = \epsilon'' = \epsilon/4$,

$$m_{\max}(\theta_n) - m_{\min}(\theta_n) \geq \epsilon/2,$$

which means that θ_n is in $C_{\epsilon/2}$.

But, the plane P on which all the Y_n lie, is *not* a latitude plane (since X cannot be N—see second paragraph of part 2). Hence there are *indefinitely* many points on the intersection of the sphere S with P, which have different latitudes. Hence there are indefinitely many Y_n at different latitudes; and hence indefinitely many members of $C_{\epsilon/2}$.

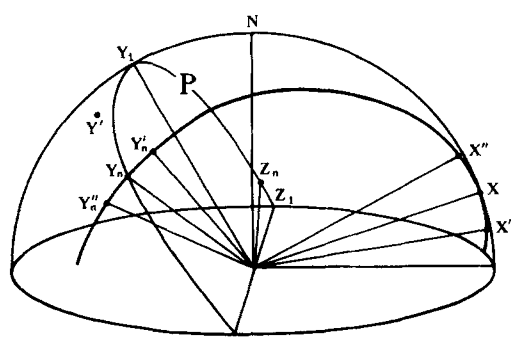

FIG. 5.

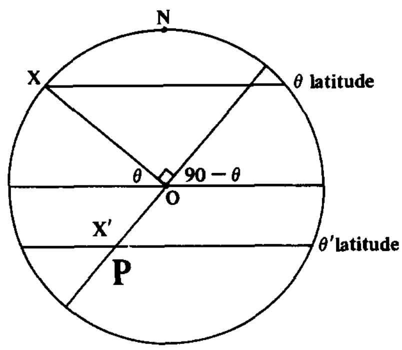

FIG. 6.

But this contradicts the result derived in part 1 of this proof; and hence, by *reductio ad absurdum*, we see that C_ϵ cannot even have *one member*.

Part 3: Here we show m is a constant of latitude and hence a continuous function of latitude. First we trivially see, from Part 2, that at *any latitude* θ, for which $m_{\max}(\theta)$ and $m_{\min}(\theta)$ exist,

$$m_{\max}(\theta)=m_{\min}(\theta)$$

Hence, by (i) and (ii) (from part 1) $m(x_1)=m_{\max}(\theta)=m_{\min}(\theta)$ for any x_1 at (northern) latitude θ, $\theta\neq 0$ or 90. But $m(X)=0$ for any X at latitude θ, and $m(N)=1$ for N at latitude 90; and hence we see that m is a constant of latitude. We can therefore introduce a function m of latitude, so that $m(\theta)$ is the value of $m(X)$ for any X at latitude θ (where we define a function as a *single valued* mapping).

Second, for any $\epsilon>0$ and $\theta>0$, there must be a $\theta'<\theta$ for which $m(\theta')>m_{\min}(\theta)-\epsilon$ because, if this were not so, then $(m_{\min}(\theta)-\epsilon)$ would be the l.u.b. of the $m(\theta')$ for $\theta'<\theta$. Hence, since we have just proved that $m_{\min}(\theta)=m(\theta)$, for $0<\theta<90$, we see that for any $\epsilon>0$, and $0<\theta<90$, there is a $\theta'<\theta$ such that $m(\theta')>m(\theta)-\epsilon$.

But also $m_{\min}(\theta)\geqslant m(\theta')$ for any $\theta'<\theta$ (by definition of l.u.b.); and hence for $0<\theta<90$, $m(\theta)\geqslant m(\theta')$.

Hence, for any $\epsilon>0$ and $0<\theta<90$, there is a $\theta'<\theta$ for which $m(\theta)\geqslant m(\theta')>m(\theta)-\epsilon$. But this means that m is a continuous function of θ from below, for $90>\theta>0$. Similarly m can be shown to be a continuous function of θ from above, for $0<\theta<90$.

Hence we have shown m is continuous at all northern latitudes, except 0 and 90. QED

Lemma 6: Iff $\sin^2\theta+\sin^2\theta'+\sin^2\theta''=1$ is there an orthotriad with members at latitudes θ, θ', θ''.

Proof: At any latitude, including θ, there are members of various orthotriads—let one such member be X. Construct the plane P through O, normal to OX. Obviously P intersects the sphere at all latitudes between $90-\theta$ and $\theta-90$. (see Fig. 6.) One of these latitudes is θ', because, from

$$\sin^2\theta+\sin^2\theta'+\sin^2\theta''=1$$

it follows that

$$\sin^2\theta+\sin^2\theta'\leqslant 1,$$

i.e.,

$$\sin^2\theta'\leqslant\cos^2\theta,\ \text{i.e.,}\ \sin^2\theta'\leqslant\sin^2(90-\theta),$$

and hence $\theta-90\leqslant\theta'\leqslant 90-\theta$.

Let one of the points at which P intersects the sphere at latitude θ' be X'. Now there are *two* points on the sphere, orthogonal to both X and X'—both of them along the same line normal to OXX'. Let them be X_1'' and X_2'' Moreover, we know that, for *any* orthotriad X_1, X_2, X_3 with members at latitudes θ_1, θ_2, θ_3 respectively, we have $\sin^2\theta_1+\sin^2\theta_2+\sin^2\theta_3=1$.[12] Hence X_1'' and X_2'' must be at the two latitudes, for which the $\sin^2$ is $1-(\sin^2\theta)-(\sin^2\theta')$. Hence there is an orthotriad (viz., X, X', X_1'') at latitudes θ, θ', θ'', merely in virtue of $\sin^2\theta+\sin^2\theta'+\sin^2\theta''=1$. QED

We are now finally in a position to prove Theorem 1 (of the first paragraph of Sec. II above).

Proof (*of Theorem* 1): Because m exists it follows, by Lemma 5, that there is an m'', which maps values of $\sin^2\theta$, $0\leqslant\theta\leqslant 90$, onto the closed interval $[0, 1]$, such that:

$m''(\sin^2\theta)=m(P)$ for any P which is at latitude θ.

Note that the rotational symmetry of m (from Lemma 5) is necessary to ensure the single-valuedness of m''. Moreover,

(a)′ $m''(1)=1$ [since $m(N)=1$], and $m''(0)=0$ (by Lemma 1),

and, from Lemma 5,

(b)′ m'' is continuous everywhere on the open interval (0, 1) and bounded.

And, from Lemma 6, we see that

(c)′ if $\sum x_i=1$, then $\sum m''(x_i)=1$.

(The continuity of m'' follows trivially from the continuity of m.)

But (c)′ is easily shown to be equivalent to the "Cauchy equation" on interval (0, 1), q.v.[13] let $\bar{x}_1, \bar{x}_2$, $(\bar{x}_1+\bar{x}_2)$ all be in (0, 1); and set $x_1=\bar{x}_1+\bar{x}_2$, $x_2=1-(\bar{x}_1+\bar{x}_2)$, and $x_3=0$. Then, from (c)′ and (a)′,

$$m''(\bar{x}_1+\bar{x}_2)+m''(1-(\bar{x}_1+\bar{x}_2))=1.$$

But now set $x_1=\bar{x}_1$, $x_2=\bar{x}_2$, and $x_3=1-(\bar{x}_1+\bar{x}_2)$. Then, from (c)′,

$$m''(\bar{x}_1)+m''(\bar{x}_2)+m''(1-(\bar{x}_1+\bar{x}_2))=1.$$

Hence $m''(\bar{x}_1)+m''(\bar{x}_2)=m''(\bar{x}_1+\bar{x}_2)$, for any $\bar{x}_1$, $\bar{x}_2$, $(\bar{x}_1+\bar{x}_2)$ in (0, 1) [since if $(\bar{x}_1+\bar{x}_2)$ is in (0, 1), so is $1-(\bar{x}_1+\bar{x}_2)$].[14]

It is well known, however, that the Cauchy equation for function m'', which is continuous and bounded everywhere on some interval, has the solution $m''(x)=x$, for all x on that interval (see p. 187 of Ref. 5). Hence, by (b)′, $m''(\sin^2\theta)=\sin^2\theta$, for all θ such that $0>\theta>90$.

Now we can easily see that $m(N')=1$, where N' is the south pole (since N' forms an orthotriad with any two equatorial points). And hence any result derived for northern points can be obtained for southern points too (the same relations hold for southern points as for northern points). In particular we can derive $m'(\sin^2\theta)$

$=\sin^2\theta$, for $-90<\theta<0$, where $m'(\sin^2\theta)=m(P)$ for any P at latitude $\theta<0$.

Hence we see that, for any P on the whole of S, $m(P)=\sin^2\theta$ for any P at latitude θ, $\theta\neq 0$ or 90 or -90. But $m(N)=1=\sin^2 90$; and, for any equatorial E, $m(E)=0=\sin^2 0$. Also, $m(N')=1=\sin^2(-90)$. Hence, for any P on S, $m(P)=\sin^2\theta$, where θ is the latitude of S. (In particular, m is continuous at the polar and equatorial points.) QED

III. PROOF OF THEOREM T

We shall now prove Theorem T.

In what follows, m and Ψ are as defined in (1)–(3) of the statement of Theorem T.

Lemma 7: For any o.n. set of vectors $\{\phi_j\}$ which spans Ψ, we have that $\sum m(\phi_j)=1$.

Proof: Let $\{\phi_j\}$ span the closed subspace V, which, *ex hypothesi*, includes Ψ. Construct a c.o.n. set of vectors in H consisting of $\{\Psi;\phi_1'\cdots\phi_k'\cdots;\phi_1''\cdots\phi_l''\cdots\}$, where $\{\Psi;\phi_k'\}$ span V. (Such a set can always be constructed by the "Gram–Schmidt" process.) We have $m(\Psi)+\sum m(\phi_k')+\sum m(\phi_l'')=1$ [by (2) of Theorem T], and hence, since $m(\Psi)=1$ and $m(\phi)\geq 0$, we have that $m(\phi_k')=0$ and $m(\phi_l'')=0$ for all k, l. But since $\{\phi_j\}$ spans V (as well as $\{\Psi,\phi_k'\}$) we must have that

$$\{\phi_1\cdots\phi_j\cdots;\phi_1''\cdots\phi_l''\cdots\}$$

is c.o.n. in H too. Hence

$$\sum m(\phi_j)+\sum m(\phi_l'')=1.$$

Since $m(\phi_l'')=0$ for all l, we have

$$\sum m(\phi_j)=1. \qquad \text{QED}$$

Now we can finally prove the Theorem T:

Proof: Let ϕ be any vector of unit norm in H. Then either $\langle\phi,\Psi\rangle$ is real or complex (where $m(\Psi)=1$). *First* suppose that $\langle\phi,\Psi\rangle$ is real. There are two subcases:

(a) $\phi\perp\Psi$. Then let ϕ_1 be a vector orthogonal to Ψ and ϕ, and V be the *real* subspace spanned by Ψ, ϕ, and ϕ_1, i.e., V is the closed set of real linear combinations of Ψ, ϕ, and ϕ_1 (of unit norm).

(b) ϕ not $\perp\Psi$. Then there is a ϕ_1, for which $\phi\perp\phi_1$ and Ψ is a linear combination of ϕ_1 and ϕ—i.e., $\Psi=\alpha\phi_1+\beta\phi$. Moreover $\beta=\langle\phi,\Psi\rangle$ and hence is real; and $\alpha=\langle\phi_1,\Psi\rangle$ which we can choose real (by adjusting the phase of ϕ_1). We then let ϕ_2 be a vector orthogonal to both ϕ_1 and ϕ, and let V be the *real* subspace spanned by ϕ, ϕ_1, and ϕ_2—which will include Ψ.

In either of these two cases, we therefore have a three-dimensional closed linear subspace V which contains Ψ and ϕ. V obviously forms a three-dimensional Hilbert space[15] with the same operations of addition, scalar multiplication, and scalar product as defined on the broader H. It is crucial to note, however, that V is *real*—i.e., all scalar products of vectors in V are real. In particular for any vector ϕ' in V, there is a unique triple of real numbers $\{C_1,C_2,C_3\}$, where $\phi'=\sum_{i=1}^3 C_i\phi_i$, for any given base set $\{\phi_i\}$ in V.

Now consider the real Euclidean space E_3. We can coordinate it so that for any X_1 in E_3 there is a representative triple of real numbers—the "coordinates of X_1." Moreover, we can define a Hilbert space on E_3, by defining a scalar product $(X,Y)=$ cosine of $\angle XOY$ (O being the point with coordinates $\{0,0,0\}$).

It is obvious that there is an isomorphism i between the spaces E_3 and V—which preserves scalar product, vector sums, and multiplication by a constant, viz., $\mathrm{i}(\phi)$ is that vector in E_3 with coordinates $\{C_1,C_2,C_3\}$, where $\phi=\sum_{i=1}^3 C_i\phi_i$. We can then define a measure m' on those vectors of E_3 which are of *unit norm*—viz. on those vectors which are on the sphere of unit radius about O, q.v.,

$$m'(X)=m(\mathrm{i}^{-1}(X)).$$

This measure m' is defined over the set of points on a sphere in E_3—let it be S—and is easily seen to satisfy the conditions (a), (b) imposed on the m of Theorem 1, q.v.: Let X_1, X_2, X_3 be any o.n. triad in S. Then, since i preserves scalar product, we see that $\mathrm{i}(X_1)$, $\mathrm{i}(X_2)$, $\mathrm{i}(X_3)$ is an o.n. triad in V. But Ψ is spanned by any o.n. triad in V, since V is, *ex hypothesi*, a three-dimensional Hilbert space in its own right. Hence, by Lemma 7, $\sum m(\mathrm{i}(X_i))=1$; so that $\sum m'(X_i)=1$.[16]

Therefore, we finally get that, for any X on S, $m'(X)=\sin^2\theta$, where θ is the latitude of X (by Theorem 1). But $\sin^2\theta=(N,X)^2=(S,X)^2$, where S and N are the north and south poles respectively; and it is a corollary of Theorem 1 that these are the only points for which m' has value 1—and hence that $\mathrm{i}(\Psi)$ is N or S [since $m(\Psi)=1$]. Thus we see that $m(\mathrm{i}^{-1}(X))=\langle\mathrm{i}^{-1}(N),\mathrm{i}^{-1}(X)^2\rangle=\langle\mathrm{i}^{-1}(S),\mathrm{i}^{-1}(X)\rangle^2$ and hence that $m(\phi)=\langle\Psi,\phi\rangle^2$ [since we have just seen that Ψ is either $\mathrm{i}^{-1}(N)$ or $\mathrm{i}^{-1}(S)$, and since some X is $i(\phi)$, for $\|\phi\|=1$].

Second, suppose $\langle\phi,\Psi\rangle$ is *not* real. Now, if $m(\Psi)=1$, then it follows that there is a set $\{\phi_i\}$ for which $\{\Psi;\phi_i\}$ is c.o.n. in H, and $m(\phi_i)=0$ for all i. (For proof see Lemma 7). But if $\{\Psi;\phi_i\}$ is c.o.n. in H, so is $\{\Psi\exp(i\alpha);\Phi_i\}$. Hence, by condition (2) of Theorem T, $m(\Phi\exp(i\alpha))+\sum m(\phi_i)=1$. Hence, since $m(\phi_i)=0$, we have that $m(\Psi\exp(i\alpha))=1$. Moreover, α may be so chosen that $\langle\phi,\Psi\exp(i\alpha)\rangle$ is real. Hence the condition (3) in Theorem T guarantees that there is a Ψ' for which $m(\Psi')=1$ *and* $\langle\phi,\Psi'\rangle$ is real for any given ϕ. The whole series of proofs up to the proof of Theorem T may then be repeated with Ψ' instead of Ψ; and we then derive (applying the first part of this theorem) that $m(\phi)=\langle\phi,\Psi'\rangle^2$ where $\langle\phi,\Psi'\rangle^2=|\langle\phi,\Psi\rangle|^2$ since $\Psi'=\Psi\exp(i\alpha)$. Hence, in general, $m(\phi)=|\langle\phi,\Psi\rangle|^2$. QED

This completes the proof of the required Theorem T.[17]

[1]"A proof of the non-existence of partial hidden variable theories," by J. Dorling, Chelsea College, London.

[2]Dorling does not provide suggestions for the following proofs which appear in our paper: Lemmata 2, 3, and most of 4, 5, and Theorem 1.

[3]The title of Ref. 1 notwithstanding, our proof has little relevance to the "hidden variables question." Condition I, in what follows, *need* not be satisfied by hidden variable theories; and is in fact not satisfied by the theory of Bohm and Bub, Rev. Mod. Phys. **38**, 453 (1966), for example.

[4]Again m may depend on S and t but the dependence is suppressed—in fact if S and t is in a pure state, $m(\ ) = |\langle\ ,\Psi\rangle|^2$, where Ψ is the eigenvector of B for value j [see (3)'].

[5]In an earlier article, Found. Phys. **4**, 181 (1974), this point was not sufficiently well emphasized by us, but we correct this here.

[6]A. Gleason, J. Math. Mech. **6**, 885 (1957).

[7]J. von Neuman, *Mathematical Foundations of Quantum Mechanics* (Princeton U. P., Princeton, N.J., 1968), p. 297.

[8]H. Krips, Found. Phys. **4**, 381 (1974).

[9]P_i and P_j are "orthogonal," i.e., $P_i \perp P_j$, if the line from P_i to the center O of the sphere is at right angles to the line from P_j to O.

[10]The lemma comes straight from Ref. 1, although, in vector notation, the same proof occurs on p. 450 of J. Bell, Rev. Mod. Phys. **38**, 447 (1966).

[11]The most convenient way to imagine this is to imagine turning AB through 360°, and, at each angle it passes through as it turns, construct a great circle through Y on AB as diameter.

[12]This is most easily seen, by realizing that the components of ON along OX_1, OX_2, OX_3 respectively are $\cos(90-\theta_1)$, $\cos(90-\theta_2)$, and $\cos(90-\theta_3)$ if we set $\|ON\| = 1$.

[13]This is just the proof called "step 1," p. 187 of Ref. 5.

[14]Note that here we essentially use the three-dimensionality of the sphere. For dimension less than 3, the proof breaks down here.

[15]The relevant theorem here is proven as in M. Naimark, *Normed Rings* (Noordhoff, Gröningen, 1964), p. 86.

[16]That Ψ is spanned by $\{i(X_1), i(X_2), i(X_3)\}$ as far as V is concerned is not, strictly speaking, sufficient to derive this conclusion. We also need to make the point that the same operations of addition and scalar multiplication which apply in H also apply in V—cf. p. 85 of Ref. 15.

[17]The author would like to mention the substantial encouragement and help received from talks with Dr. H.A. Cohen.

Heavens and their integral manifolds

Charles P. Boyer

Instituto de Investigación en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, México 20, D. F., Mexico

Jerzy F. Plebański*

Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional 14—740, México 14, D. F.
(Received 30 June 1976)

By using the apparatus of exterior forms, a new spinorial notation and Cartan's theory of integral manifolds, some new results concerning complex strong heavenly metrics are established. In particular, the study of a subfamily of two-variable heavens (types $G\otimes[—]$, $D\otimes[—]$, and $N\otimes[—]$) is reduced to linear equations, a prolongation process related to first and second heavenly equations is studied leading to a (presumably) infinite hierarchy of 1-forms, and finally, the symmetries of the studied structure are investigated from the point of view of its description by Pfaffian forms, elucidating in this way previous results concerning Killing vectors.

1. INTRODUCTION

This paper is the fifth in a series of articles dedicated to the study of the analytic continuation of general relativity, with special emphasis on the solutions of the complex Einstein equations characterized by the self-dual conformal curvature. (These spaces have been called heavens by Newman[1] and Penrose.[2] The first article[3] of the series outlined the formalism of complex tetrads, forms, and spinors used subsequently and established heavens as the integral manifolds of a certain partial differential equation of order and degree 2. Actually, two equivalent partial differential equations for a single function were given—the first and second heavenly equations. In the second article,[4] a generalization of the Goldberg—Sachs theorem to complex Riemannian spaces was given, elucidating the important role of complex null strings. Then in[5] many explicit heavens of various algebraic type were studied and the problem of finding the conformal Killing vectors for an arbitrary heaven was reduced to a single equation while the general theory of Killing spinors in both real and complex Riemannian spaces was studied in Ref. 6.

The purpose of this article is to study the general integral manifolds of heaven from a geometric point of view. While the general form of a regular integral manifold is only given implicitly, many of its properties are obtained and studied form the point of view of both a concise spinorial language and closed Pfaffian 1-forms. The outline of the paper is as follows: in order to make the succeeding sections more palatable we show first in Sec. 2 how a special subcase of two-variable heavens can be solved completely by the systematic use of elementary exterior differential calculus which reduces the problem to the linear two-dimensional complex Laplace equation. For this case we give a complete classification of the self-dual conformal curvature types.

In Sec. 3 we organize some of the basic results of the previous papers in the concise spinorial language, while Sec. 4 treats the integral manifolds in terms of pairs of Pfaffian 1-forms. The idea of prolongations first introduced by Cartan is used to study further properties,[7–9] in particular the relation between the first and second heavenly equations and establishing a hierarchy (presumably infinite) of 1-forms. Then in Sec. 5 we investigate the structure of the regular integral manifolds of heavens from the point of view of the general Cartan theory.[8,10] Furthermore, some explicit results concerning subcases when the problem can be reduced to linear structures are presented. Finally, in the last section the symmetry group which maps the heavenly integral manifolds into each other is computed and the relation to Killing vectors is discussed.

2. TWO-VARIABLE HEAVENS

In Ref. 5 the general solution of the reduced two variable problem [Eq. (2.1) below] was solved using the method of first integrals. However, since the computations involved were quite complicated, the classification of the algebraic degeneracy of the conformal curvature was not given. In this section we present this classification as well as the general form of the metric, connections, and curvature using fairly simple computations. Our techniques illustrate the ease in which differential forms can be used to solve concrete problems which at first sight appear formidable. The reduced two variable equation of Ref. 5 is

$$\Theta_{xx}\Theta_{yy}-\Theta_{xy}\Theta_{xy}=1. \tag{2.1}$$

Here we have transformed the constant $-k^2$ in Eq. (2.34) of Ref. 5 to 1 by a complex dilatation. The case when $k^2=0$ was completely solved in Ref. 5 and yields algebraically special metrics. We will also briefly discuss this case in the present context. In both cases we succeed in linearizing the theory.

To write (2.1) in differential form language we first write the contact 1-form

$$d\Theta-u\,dx-v\,dy=0, \tag{2.2}$$

which implies $u=\Theta_x$, $v=\Theta_y$. Then it is easy to see that (2.1) becomes

$$du\wedge dv-dx\wedge dy=0. \tag{2.3a}$$

Taking the exterior derivative of (2.2) we find

$$du\wedge dx+dv\wedge dy=0. \tag{2.3b}$$

Now we can consider (2.1) to be equivalent to (2.3) with the added condition that the 2-forms (2.3) be in involution[8] with respect to the variables x and y, i.e., $dx \wedge dy \neq 0$. Now we integrate (2.3b) considering x and v as underlying variables. This is equivalent to a half Legendre transformation and can be understood as a certain discrete transformation in the linear symplectic group Sp(4, $\mathbb{C}$), upon integration we find

$$d\phi = u\,dx + y\,dv = 0 \tag{2.4}$$

subject to the constraint $dy \wedge dx = \phi_{vv} dv \wedge dx \neq 0$. Substituting (2.4) into (2.3a) we find

$$\phi_{xx} + \phi_{vv} = 0. \tag{2.5}$$

Thus Eq. (2.1) is equivalent to the complex Laplace equation as long as $dv \wedge dx \neq 0$ and $\phi_{vv} \neq 0$. The only solution of (2.1) which is not equivalent to (2.5) is the case $\Theta_{yy} = 0$ which yields the solutions

$$\Theta = \pm ixy + ay + \Theta^0(x), \tag{2.6}$$

where a is constant and Θ^0 is an arbitrary holomorphic function of x. This is a special case of a class solved in Ref. 3 of type $[N] \otimes [-]$. We mention also that equivalently we could integrate (2.3a) with independent variables x and v which gives another Laplace equation when substituted into (2.3b).

Now the general solution of (2.5) is well known and can be written as

$$\phi = f(z) + \bar{f}(\bar{z}) \tag{2.7}$$

where $z = x + iv$, $\bar{z} = x - iv$ (bar does not denote complex conjugate), f and $\bar{f}$ are arbitrary holomorphic functions. It is a straightforward calculation to express the metric in terms of the new quantities. Putting $F := f_{zz}$, $\bar{F} := \bar{f}_{\bar{z}\bar{z}}$ we find

$$ds^2 = dp\left[dz + d\bar{z} - \frac{2dp}{F+\bar{F}} + 4i\frac{F-\bar{F}}{F+\bar{F}}dq\right] + dq\left[-iF\,dz + i\bar{F}\,d\bar{z} - \frac{8F\bar{F}}{F+\bar{F}}dq\right], \tag{2.8}$$

where the condition $\phi_{vv} \neq 0$ implies $F + \bar{F} \neq 0$. Similarly we can compute the heavenly connections. Using the notation of Ref. 3,

$$\begin{aligned}\Gamma_{12} &= \Gamma_{34} = -K\,dp + L\,dq,\\ \Gamma_{42} &= -L\,dp + M\,dq,\\ \Gamma_{31} &= -N\,dp + K\,dq,\end{aligned} \tag{2.9}$$

we have

$$\begin{aligned} N &= -i\frac{(\dot{F} - \dot{\bar{F}})}{(F+\bar{F})^3}, \quad K = \frac{\dot{F}^2 + \dot{\bar{F}}^2}{(F+\bar{F})^3},\\ L &= 4i\frac{(\bar{F}^2\dot{F} - F^2\dot{\bar{F}})}{(F+\bar{F})^3}, \quad M = 8\frac{(F^3\dot{\bar{F}} + \bar{F}^3\dot{F})}{(F+\bar{F})^3}. \end{aligned} \tag{2.10}$$

After a little straightforward algebra we obtain the components of the conformal curvature,

$$\begin{aligned}
\tfrac{1}{2}C^{(1)} &= (F+\bar{F})^{-4}\left(\ddot{F} + \ddot{\bar{F}} - \frac{3(\dot{F} - \dot{\bar{F}})^2}{F+\bar{F}}\right),\\
\tfrac{1}{2}C^{(2)} &= 2i(F+\bar{F})^{-4}\left(F\ddot{\bar{F}} - \bar{F}\ddot{\bar{F}} + \frac{3(\dot{F} - \dot{\bar{F}})}{(F+\bar{F})}(\bar{F}\dot{F} + F\dot{\bar{F}})\right),\\
\tfrac{1}{2}C^{(3)} &= -4(F+\bar{F})^{-4}\left(\bar{F}^2\ddot{F} + F^2\ddot{\bar{F}} + \frac{\dot{F}\dot{\bar{F}}(F^2 + \bar{F}^2) - 3(\bar{F}^2\dot{F}^2 + F^2\dot{\bar{F}}^2) - 4\dot{F}\dot{\bar{F}}F\bar{F}}{F+\bar{F}}\right),\\
\tfrac{1}{2}C^{(4)} &= 8i(F+\bar{F})^{-4}\left(\ddot{F}\bar{F}^3 - \ddot{\bar{F}}F^3 - \frac{3(\dot{F}\bar{F} + \dot{\bar{F}}F)(F\bar{F}^2 - \dot{\bar{F}}F^2)}{F+\bar{F}}\right),\\
\tfrac{1}{2}C^{(5)} &= 16(F+\bar{F})^{-4}\left(\ddot{F}\bar{F}^4 + \ddot{\bar{F}}F^4 - \frac{3(\dot{F}\bar{F}^2 - \dot{\bar{F}}F^2)(\dot{F}\bar{F}^2 - \dot{\bar{F}}F^2)}{F+\bar{F}}\right).
\end{aligned} \tag{2.11}$$

The degenerate case $\Theta_{xx}\Theta_{yy} - \Theta_{xy}^2 = 0$ described in Ref. 5 can be treated similarly. Indeed in this case (2.3a) is replaced by

$$du \wedge dv = 0 \tag{2.12a}$$

and using (2.4) we have

$$\phi_{xx} = 0 \tag{2.12b}$$

whose solutions are immediate, viz.,

$$\phi = \phi^1(v)x + \phi^0(v). \tag{2.12c}$$

This case was completely classified in Ref. 5 so we do not repeat it here.

Returning to the curvature components (2.11) we can construct Penrose's fourth order equation for the spinor $K^A = (K^1, K^2)$ to determine the Penrose–Petrov classification.[3,11] In our case this equation can be written in the biquadratic form

$$\left(\ddot{F} - 3\frac{\dot{F}^2}{F+\bar{F}}\right)(K^1 - 2i\bar{F}K^2)^4 + \left(\ddot{\bar{F}} - 3\frac{\dot{\bar{F}}^2}{F+\bar{F}}\right)(K^1 + 2iFK^2)^4 + 2\cdot\frac{3\dot{F}\dot{\bar{F}}}{F+\bar{F}}(K^1 + 2iFK^2)^2(K^1 - 2i\bar{F}K^2)^2 = 0. \tag{2.13}$$

Owing to the biquadratic nature of (2.13), the roots can be obtained fairly easily. Introducing

$$A := \ddot{F} - 3\frac{\dot{F}^2}{F+\bar{F}}, \quad B := 3\frac{\dot{F}\dot{\bar{F}}}{F+\bar{F}}, \quad C := \ddot{\bar{F}} - 3\frac{\dot{\bar{F}}^2}{F+\bar{F}} \tag{2.14}$$

we write (2.13) in the factorized form

$$(\alpha^+_A K^A)(\alpha^-_B K^B)(\beta^+_C K^C)(\beta^-_D K^D) = 0, \tag{2.15}$$

where the components of the spinors $\alpha^\pm_A$, $\beta^\pm_A$ are

$$\begin{aligned}
\alpha^\pm_1 &= A^{1/2} \pm i[B + (B^2 - AC)^{1/2}]^{1/2},\\
\alpha^\pm_2 &= -2i\bar{F}A^{1/2} \mp 2F[B + (B^2 - AC)^{1/2}]^{1/2},\\
\beta^\pm_1 &= A^{1/2} \pm i[B - (B^2 - AC)^{1/2}]^{1/2},\\
\beta^\pm_2 &= -2i\bar{F}A^{1/2} \mp 2F[B - (B^2 - AC)^{1/2}]^{1/2}.
\end{aligned} \tag{2.16}$$

The coincidences are, of course, obtained by the vanishing of any spinor scalar products between $\alpha^\pm_A$, $\beta^\pm_A$. We find the following classification:

(1) F and $\bar{F}$ constant, or $\bar{F}$ constant and $F = (\alpha z + \beta)^{-1/2} + k$ or vice versa;

Flat

(2) $\bar{F}$ constant, F arbitrary or vice versa, $[N] \otimes [-]$;

(3) $F = \alpha_1 z + \beta_1$, $\bar{F} = \alpha_2\bar{z} + \beta_2$ or $F = (\alpha_1 z + \beta_1)^{-1/2} + k$, $\bar{F} = (\alpha_2\bar{z} + \beta_2)^{-1/2} - k$, $[D] \otimes [-]$;

(4) F and $\bar{F}$ otherwise $[G] \otimes [-]$,

where $\alpha_i \neq 0$, β_i and k are constants in the above classification. It can also be mentioned that the corresponding Killing vectors can be worked out from the results of Ref. 5.

3. STRONG HEAVENS IN THE SPINORIAL NOTATION

The results derived in Ref. 3 concerning strong heavens were obtained by working with the spinorial formalism; however they were stated in a notation which did not make explicit use of the spinorial character of the various quantities concerned. We can now considerably improve the condensed presentation of these results by the simple device of introducing in place of the variables $\{xypq\}$ and $\{pqrs\}$ which were used in Ref. 3, the new variables defined by

$$x := -p^1, \quad y := -p^2, \quad p := q_1, \quad q := q_2, \quad r := i\bar q_{\dot 1}, \quad s := i\bar q_{\dot 2} \tag{3.1}$$

which will be interpreted as *formal spinors* $(p^A, q_A, \bar q_{\dot A})$. The spinorial indices should be then manipulated according to the standard rules

$$\psi_A = \epsilon_{AB}\psi^B, \quad \bar\psi_{\dot A} = \epsilon_{\dot A\dot B}\bar\psi^{\dot B}, \quad \psi^B = \psi_A\epsilon^{AB}, \quad \bar\psi^{\dot B} = \bar\psi_{\dot A}\epsilon^{\dot A\dot B}, \tag{3.2}$$

where

$$(\epsilon_{AB}) = (\epsilon_{\dot A\dot B}) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = (\epsilon^{\dot A\dot B}) = (\epsilon^{AB}). \tag{3.3}$$

The key functions[3] depend now on their respective variables written in spinor notation, viz.

$$\Omega = \Omega(q_A, \bar q_{\dot A}), \quad \Theta = \Theta(p^A, q_A). \tag{3.4}$$

The corresponding heavenly equations assume the form

$$\frac{1}{2}\frac{\partial^2\Omega}{\partial q_A\partial\bar q_{\dot B}}\frac{\partial^2\Omega}{\partial q^A\partial\bar q^{\dot B}} + 1 = 0, \qquad \frac{1}{2}\frac{\partial^2\Theta}{\partial p^A\partial p_B}\frac{\partial^2\Theta}{\partial p_A\partial p^B} + \frac{\partial^2\Theta}{\partial p^A\partial q_A} = 0, \tag{3.5}$$

and the (strongly) heavenly metric takes the form of

$$H: ds^2 = 2e^1\otimes e^2 + 2e^3\otimes e^4 = 2\frac{\partial^2\Omega}{\partial q_A\partial\bar q_{\dot B}}\, dq_A\otimes dq_B = 2\,dq^A\otimes\left(dp_A - \frac{\partial^2\Theta}{\partial p^A\partial p^B}\, dq^B\right). \tag{3.6}$$

The heavenly tetrad and its inverse are then given by

$$(g^{A\dot B}) = \sqrt{2}\begin{pmatrix} e^4 & e^2 \\ e^1 & -e^3 \end{pmatrix} = \sqrt{2}\begin{pmatrix} dq^1, & \frac{\partial^2\Omega}{\partial q_1\partial\bar q_{\dot B}}\, d\bar q_{\dot B} \\ dq^2, & \frac{\partial^2\Omega}{\partial q_2\partial\bar q_{\dot B}}\, d\bar q_{\dot B} \end{pmatrix} = \sqrt{2}\begin{pmatrix} dq^1, & -dp^1 - \frac{\partial^2\Theta}{\partial p_1\partial p_A}\, dq_A \\ dq^2, & -dp^2 - \frac{\partial^2\Theta}{\partial p_2\partial p_A}\, dq_A \end{pmatrix}, \tag{3.7}$$

$$(\partial_{A\dot B}) := (g^a_{A\dot B}\partial_a) = \sqrt{2}\begin{pmatrix} \partial_4, & \partial_2 \\ \partial_1, & -\partial_3 \end{pmatrix} = -\sqrt{2}\begin{pmatrix} \frac{\partial}{\partial q^1}, & \frac{\partial^2\Omega}{\partial q^1\partial\bar q_{\dot A}}\frac{\partial}{\partial\bar q^{\dot A}} \\ \frac{\partial}{\partial q^2}, & \frac{\partial^2\Omega}{\partial q^2\partial\bar q_{\dot A}}\frac{\partial}{\partial\bar q^{\dot A}} \end{pmatrix} = -\sqrt{2}\begin{pmatrix} \frac{\partial}{\partial q^1} + \frac{\partial^2\Theta}{\partial q^1\partial p^A}\frac{\partial}{\partial p_A}, & -\frac{\partial}{\partial p^1} \\ \frac{\partial}{\partial q^2} + \frac{\partial^2\Theta}{\partial p^2\partial p^A}\frac{\partial}{\partial p_A}, & -\frac{\partial}{\partial p^2} \end{pmatrix}. \tag{3.8}$$

For the invariant d'Alembertian in the strong heavens we find

$$\Box\Phi := \nabla_\alpha\nabla^\alpha\Phi = 2(\partial_2\partial_1 + \partial_3\partial_4)\Phi = -2\frac{\partial^2\Omega}{\partial q_A\partial\bar q_B}\frac{\partial^2}{\partial q^A\partial\bar q^{\dot B}}\Phi = 2\frac{\partial}{\partial p_A}\left(\frac{\partial}{\partial q^A} + \frac{\partial^2\Theta}{\partial p^A\partial p^B}\frac{\partial}{\partial p_B}\right)\Phi. \tag{3.9}$$

Now, the basic relation which establishes the bridge between the Ω and Θ formalisms is

$$p^A = -\frac{\partial\Omega}{\partial q_A}. \tag{3.10}$$

It will be useful to consider a parallel object

$$\bar p^{\dot A} = -\frac{\partial\Omega}{\partial\bar q_{\dot A}}. \tag{3.11}$$

Then, for the base of the (closed anti-self-dual 2-forms we have

$$(\mathcal{S}^{\dot A\dot B}) = \begin{pmatrix} 2e^4\wedge e^1, & -e^1\wedge e^2 + e^3\wedge e^4 \\ -e^1\wedge e^2 + e^3\wedge e^4, & 2e^3\wedge e^2 \end{pmatrix} = \begin{pmatrix} dq^A\wedge dq_A, & -dq^A\wedge dp_A \\ d\bar q^{\dot A}\wedge d\bar p_{\dot A}, & -d\bar q^{\dot A}\wedge d\bar q_{\dot A} \end{pmatrix}. \tag{3.12}$$

Writing this, we notice that

$$\mathcal{S}^{\dot 1\dot 2} = -dq^A\wedge dp_A = d\bar q^{\dot A}\wedge d\bar p_{\dot A}, \qquad \mathcal{S}^{\dot 2\dot 2} = -d\bar q^{\dot A}\wedge d\bar q_{\dot A} = dp^A\wedge dp_A + 2d\left(\frac{\partial\Theta}{\partial p_A}\right)\wedge dq_A. \tag{3.13}$$

At the same time, for the base of the self-dual 2-forms we have

$$(\mathcal{S}^{AB}) = \begin{pmatrix} 2e^4\wedge e^2, & e^1\wedge e^2 + e^3\wedge e^4 \\ e^1\wedge e^2 + e^3\wedge e^4, & 2e^3\wedge e^1 \end{pmatrix} = \left(2dq^{(A}\wedge\frac{\partial^2\Omega}{\partial q_{B)}\partial\bar q_{\dot C}}\, d\bar q_{\dot C}\right) = -\left[2dq^{(A}\wedge dp^{B)} + \frac{\partial^2\Theta}{\partial p_A\partial p_B}\, dq^C\wedge dq_C\right]. \tag{3.14}$$

The invariant volume along V_4 is given by

$$-dV := *1 = e^1\wedge e^2\wedge e^3\wedge e^4 = \tfrac{1}{2}\, dq^A\wedge dq_A\wedge dp^B\wedge dp_B = -\tfrac{1}{4}\, dq^A\wedge dq_A\wedge d\bar q^{\dot B}\wedge d\bar q_{\dot B}. \tag{3.15}$$

Now, our $g^{A\dot B}$ from (3.7) induces the spinorial connection [for the definition, see Ref. 3, Eq. (1.16)]

$$\Gamma_{AB} = dq^R\frac{\partial^2\Omega}{\partial q^R\partial\bar q^{\dot S}}\frac{\partial^3\Omega}{\partial\bar q_{\dot S}\partial q^A\partial q^B} = -dq^C\frac{\partial^3\Theta}{\partial p^A\partial p^B\partial p^C}, \quad \bar\Gamma_{\dot A\dot B} = 0. \tag{3.16}$$

The curvature form

$$\mathcal{R}_{AB} := d\Gamma_{AB} + \Gamma_{AS} \wedge \Gamma^{S}{}_{B}$$
$$= \frac{1}{2} \frac{\partial^4 \Theta}{\partial p^A \partial p^B \partial p^C \partial p^D} \cdot 2dq^C \wedge \left(dp^D + \frac{\partial^2 \Theta}{\partial p_D \partial p^E} dq^E\right)$$
$$= -\tfrac{1}{2} C_{ABCD} \mathcal{S}^{CD}, \tag{3.17}$$

determines the only nontrivial spinorial curvature quantity

$$C_{ABCD} = \frac{\partial^4 \Theta}{\partial p^A \partial p^B \partial p^C \partial p^D}. \tag{3.18}$$

Of course, because H with the null tetrad oriented as in (3.7) is a strong heaven, $\bar{C}_{\dot{A}\dot{B}\dot{C}\dot{D}} = 0 = R_{ab}$. The heavenly conformal curvature can also be expressed in terms of the first key function

$$C_{ABCD} = \frac{\partial^2 \Omega}{\partial q^{(A} \partial \bar{q}^{\dot{R}}} \frac{\partial}{\partial \bar{q}_{\dot{R}}} \left\{ \frac{\partial^2 \Omega}{\partial q^B \partial \bar{q}^{\dot{S}}} \frac{\partial}{\partial \bar{q}_{\dot{S}}} \frac{\partial}{\partial q^C} \frac{\partial}{\partial q^{D)}} \Omega \right\}. \tag{3.19}$$

(The symmetrization affects here only the undotted indices $ABCD$.)

We will mention that from the two expressions for ds^2 in (3.6) one directly infers that

$$\frac{\partial^2 \Omega}{\partial q_A \partial q_B} + \frac{\partial^2 \Theta}{\partial p_A \partial p_B} = 0. \tag{3.20}$$

Now, a basic advantage of the present notation is that if we restrict the heavenly factor of the gauge group $SL(2,\mathbb{C})$ in $\mathcal{G} = SL(2,\mathbb{C}) \times \overline{SL}(2,\mathbb{C})$ to *constant* transformations, it coincides directly with the freedom of $SL(2,\mathbb{C})$ transformations of our formal spinors. These transformations (which maintain the simple expression for Γ_{AB} in the terms of Θ) represent the *ambiguity group* of the present spinorial description of strong heavens in the Θ formalism. Notice that according to (3.12) the "hellish" 2-forms $\mathcal{S}^{\dot{A}\dot{B}}$ are invariants of this group, as it should be. It should be observed, however, that working with the Ω formalism, we then have two independent ambiguity transformations of our formal spinors, $SL(2,\mathbb{C})$ and $\overline{SL}(2,\mathbb{C})$, with *constant* coefficients, where $\overline{SL}(2,\mathbb{C})$ *does not* coincide with the hellish factor in $\mathcal{G}$ restricted to constant transformations. This fact should be remembered when working with the Ω formalism if one wants to avoid confusions.

4. PROLONGATIONS AND THE DERIVATION OF THE HEAVENLY STRUCTURE

In this section we will apply Cartan's idea of prolongation[7-9] to study the heavenly integral manifolds and their relation to an hierarchy of key functions. Since the procedure of prolongation of an ideal of differential forms is probably new to the reader we will proceed rather cautiously. We will also employ the spinorial formalism developed in the proceding section. Our treatment here is local; however, the formalism used is readily adaptable to a global treatment. The standard mathematical device for patching the local information together to obtain a global theory is to use the theory of algebraic sheaves.[2] Then the problem of the existence of exact global 1-forms along some integral manifold is a problem involving sheaf cohomology theory. We only mention this as a future road to the global theory and all our integrations here will be in star shaped regions.

Now consider two independent 4-forms given in local coordinates by

$$\alpha := (dp^1 \wedge dq_1 + dp^2 \wedge dq_2) \wedge d\bar{q}_{\dot{1}} \wedge d\bar{q}_{\dot{2}}$$
$$\equiv \tfrac{1}{2}(dp_A \wedge dq^A) \wedge (d\bar{q}_{\dot{B}} \wedge d\bar{q}^{\dot{B}}), \tag{4.1a}$$
$$\beta := (dp^1 \wedge dp^2 + d\bar{q}_{\dot{1}} \wedge d\bar{q}_{\dot{2}}) \wedge dq_1 \wedge dq_2$$
$$\equiv \tfrac{1}{4}(dp_A \wedge dp^A + d\bar{q}_{\dot{A}} \wedge d\bar{q}^{\dot{A}}) \wedge (dq_B \wedge dq^B), \tag{4.1b}$$

involving six complex variables $\{p_A, q_A, \bar{q}_{\dot{A}}\}$. We can take the global manifold here as $\mathbb{C}^6$. Now an integral manifold I of the ideal of differential forms generated by (4.1) is a pair (N, i) where N is an analytic manifold and $i: N \to \mathbb{C}^6$ is an immersion (locally 1–1) such that the pullback $i^*\omega := \omega(i(p)) = 0$ for $p \in N$ and ω in the ideal. Hereafter, we will take poetic license and simply write an integral manifold I as any subspace which satisfies

$$\alpha = 0, \quad \beta = 0. \tag{4.2}$$

By external multiplication of α and β by the basis 1-forms $\{dp_A, dq_A, d\bar{q}_{\dot{A}}\}$ one easily finds that all $6 = \binom{6}{5}$ of the possible external products of five differentials of these variables vanish on I as a consequence of (4.2). Therefore, for any integral manifold we have

$$\dim I \leq 4. \tag{4.3}$$

We are interested in exactly four-dimensional integral manifolds along which we explicitly assume

$$0 \neq dV = -\tfrac{1}{4} dq^A \wedge dq_A \wedge dp^B \wedge dp_B$$
$$= \tfrac{1}{4} dq^A \wedge dq_A \wedge d\bar{q}^{\dot{B}} \wedge d\bar{q}_{\dot{B}}, \tag{4.4}$$

i.e., the ideal generated by α and β satisfying (4.2) is an involution[8,10] with respect to either set of variables $\{\bar{q}_{\dot{A}}, q_A\}$ or $\{q_A, p_A\}$. The last equality in (4.4) follows from $\beta = 0$.

Now with $dV \neq 0$ we can select in particular, $\{\bar{q}_{\dot{A}}, q_A\}$ as local independent coordinates for I, then having

$$p_A = p_A(q_B, \bar{q}_{\dot{B}}) \tag{4.5}$$

so that Eqs. (4.1) and (4.2) become

$$\alpha = \left(-\frac{\partial p^1}{\partial q_2} + \frac{\partial p^2}{\partial q_1}\right) dV = 0, \tag{4.6a}$$

$$\beta = \left(\frac{\partial p^1}{\partial \bar{q}_{\dot{1}}} \frac{\partial p^2}{\partial \bar{q}_{\dot{1}}} - \frac{\partial p^2}{\partial \bar{q}_{\dot{1}}} \frac{\partial p^2}{\partial \bar{q}_{\dot{1}}} + 1\right) dV = 0. \tag{4.6b}$$

As a consequence of the local inversion of the Poincaré lemma, (4.6a) implies the existence of a function $\Omega = \Omega(q_A, \bar{q}_{\dot{A}})$ such that

$$p^A = -\frac{\partial \Omega}{\partial q_A}. \tag{4.7}$$

Plugging (4.7) into (4.6b) we recover the first heavenly equation (3.5a). Thus our integral manifolds I of (4.2) are, at least locally, in 1–1 correspondence with the solutions of the first heavenly equation. We will now analyze various consequences of Eq. (4.2).

In order to proceed systematically with the program, we will now state an elementary lemma[11] (which will heretofore be referred to as **L**):

Lemma **L**: Let $\mathcal{D}_n$ denote a star-shaped region of a complex analytic manifold M_n of complex dimension n, and let $x^1, \ldots, x^k$ be k independent local coordinates in $\mathcal{D}_n$, i.e., $dx^1 \wedge dx^2 \wedge \cdots \wedge dx^k \neq 0$, $k \leq n$. Let e be a 1-form ($e \in \Lambda^1$) such that in $\mathcal{D}_n : de \wedge dx^1 \wedge \cdots dx^k = 0$. Then there exists in $\mathcal{D}_n$ 0-forms $x, y_1, \ldots, y_k \in \Lambda^0$ such that

$$e = dx + \sum_{i=1}^{k} y_i dx^i. \tag{4.8}$$

We can now rewrite (4.1a) and (4.1b) in the form

$$\alpha = d(p^A dq_A) \wedge d\bar{q}_{\dot{1}} \wedge d\bar{q}_{\dot{2}}, \tag{4.9a}$$

$$\beta = \tfrac{1}{2} d(p^A dp_A + \bar{q}^{\dot{A}} d\bar{q}_{\dot{A}}) \wedge dq_1 \wedge dq_2, \tag{4.9b}$$

and applying L to the integral manifolds for which $\alpha = \beta = 0$, i.e., (4.2) is satisfied, we infer the existence of functions Ω, Σ and s^A such that

$$p^A dq_A + \bar{p}^{\dot{A}} d\bar{q}_{\dot{A}} = -d\Omega, \tag{4.10a}$$

$$s^A dq_A + \tfrac{1}{2} p^A dp_A + \tfrac{1}{2} q^{\dot{A}} d\bar{q}_{\dot{A}} = d\Sigma. \tag{4.10b}$$

Differentiating these relations we have of course

$$dp^A \wedge dq_A = -d\bar{p}^{\dot{A}} \wedge d\bar{q}_{\dot{A}}, \tag{4.11a}$$

$$\tfrac{1}{2}(dp^A \wedge dp_A + d\bar{q}^{\dot{A}} \wedge d\bar{q}_{\dot{A}}) = -ds^A \wedge dq_A. \tag{4.11b}$$

Now, the new spinors which have appeared in these relations are $\bar{p}_{\dot{A}}$ and s_A. We can now observe that the pairs of spinors $\{p_A, q_A\}$ and $\{\bar{p}_{\dot{A}}, \bar{q}_{\dot{A}}\}$ play symmetric roles. Indeed, multiplying externally (4.11a) by $dq_1 \wedge dq_2$ we deduce the equation

$$\begin{aligned} \bar{\alpha} &:= (d\bar{p}^{\dot{1}} \wedge d\bar{q}_{\dot{1}} + d\bar{p}^{\dot{2}} \wedge d\bar{q}_{\dot{2}}) \wedge dq_1 \wedge dq_2 \\ &\equiv \tfrac{1}{2}(d\bar{p}_{\dot{A}} \wedge d\bar{q}^{\dot{A}}) \wedge (dq_B \wedge dq^B) = 0 \end{aligned} \tag{4.12}$$

on an integral manifold. Now take the external "squares" of both sides of (4.11a); this gives

$$\begin{aligned} d\bar{q}_{\dot{1}} \wedge d\bar{q}_{\dot{2}} \wedge d\bar{p}^{\dot{1}} \wedge d\bar{p}^{\dot{2}} &\equiv \tfrac{1}{4} d\bar{q}^{\dot{A}} \wedge d\bar{q}_{\dot{A}} \wedge d\bar{p}^{\dot{B}} \wedge d\bar{p}_{\dot{B}} \\ &= dq_1 \wedge dq_2 \wedge dp^1 \wedge dp^2 \\ &\equiv \tfrac{1}{4} dq^A \wedge dq_A \wedge dp^B \wedge dp_B \equiv dV \neq 0. \end{aligned} \tag{4.13}$$

Consequently, the functions $\{\bar{p}_{\dot{A}}, \bar{q}_{\dot{B}}\}$ are independent. Now, eliminating $dp^B \wedge dp_B$ in equality (4.13) by the use of (4.11b) leads to

$$\begin{aligned} 0 = \bar{\beta} &:= (d\bar{p}^{\dot{1}} \wedge d\bar{p}^{\dot{2}} + dq_1 \wedge dq_2) \wedge d\bar{q}_{\dot{1}} \wedge d\bar{q}_{\dot{2}} \\ &\equiv \tfrac{1}{4}(d\bar{p}_{\dot{A}} \wedge d\bar{p}^{\dot{A}} + dq_A \wedge dq^A) \wedge (d\bar{q}_{\dot{B}} \wedge d\bar{q}^{\dot{B}}). \end{aligned} \tag{4.14}$$

Thus, equations $\alpha = \beta = 0$ and $\bar{\alpha} = \bar{\beta} = 0$ imply each other and are related by the formal transformation $\{p_A, q_A, \bar{q}_{\dot{A}}\} \to \{\bar{p}_{\dot{A}}, \bar{q}_{\dot{A}}, q_A\}$.

Now, $\bar{\alpha} = 0$ by the application of L and again gives us (4.10a). From $\bar{\beta} = 0$, however, by the application of L we obtain the new information that there exist functions $\bar{s}^{\dot{A}}$ and $\bar{\Sigma}$ such that

$$\bar{s}^{\dot{A}} d\bar{q}_{\dot{A}} + \tfrac{1}{2} \bar{p}^{\dot{A}} d\bar{p}_{\dot{A}} + \tfrac{1}{2} q^A dq_A = d\bar{\Sigma}. \tag{4.15}$$

This relation differentiated gives, of course,

$$\tfrac{1}{2}(d\bar{p}^{\dot{A}} \wedge d\bar{p}_{\dot{A}} + dq^A \wedge dq_A) = -d\bar{s}^{\dot{A}} \wedge d\bar{q}_{\dot{A}}. \tag{4.16}$$

It is now clear that the equations $\bar{\alpha} = \bar{\beta} = 0$ again lead, through the elimination of $\bar{p}_{\dot{A}}$ in the form of $\bar{p}_{\dot{A}} = \partial\Omega/\partial\bar{q}^{\dot{A}}$, to the first heavenly equation, (3.5a). Therefore, we can now equivalently state the problem of the integral manifold as follows: Postulating simultaneously any of the two pairs of equations in 1-forms,

$$\begin{aligned} \mathcal{A} \begin{cases} d\Sigma = \tfrac{1}{2} p^A dp_A + s^A d\bar{q}_A + \tfrac{1}{2} q^{\dot{A}} d\bar{q}_{\dot{A}}, \\ -d\Omega = p^A dq_A + \bar{p}^{\dot{A}} d\bar{q}_{\dot{A}}, \end{cases} & \\ \left.\begin{aligned} &\phantom{-d\Omega = p^A dq_A + \bar{p}^{\dot{A}} d\bar{q}_{\dot{A}},} \\ &d\bar{\Sigma} = \tfrac{1}{2} q^A dq_A + \bar{s}^{\dot{A}} d\bar{q}_{\dot{A}} + \tfrac{1}{2} \bar{p}^{\dot{A}} d\bar{p}_{\dot{A}} \end{aligned}\right\} \bar{\mathcal{A}}, & \end{aligned} \tag{4.17}$$

one is led to the first heavenly equation. Thus, it is reasonable, instead of considering separately the pair $\mathcal{A}$ or the pair $\bar{\mathcal{A}}$, to consider the three equations in (4.17) as 1-forms where there enter six spinors $(p_A, q_A, s_A, \bar{p}_{\dot{A}}, \bar{q}_{\dot{A}}, \bar{s}_{\dot{A}})$ and the three key functions $(\Sigma, \Omega, \bar{\Sigma})$ *together* as equations which determine an integral manifold in the corresponding multidimensional space.

We shall thus call the three relations in (4.17) the *nucleus* $\mathcal{N}$ of the heavenly structure of 1-forms. The integrability conditions of $\mathcal{N}$ have of course the shape of three equations in 2-forms,

$$\partial\mathcal{N}: \quad \begin{aligned} &\tfrac{1}{2} dp^A \wedge dp_A + ds^A \wedge dq_A + \tfrac{1}{2} d\bar{q}^{\dot{A}} \wedge d\bar{q}_{\dot{A}} = 0 \\ &dp^A \wedge dq_A + d\bar{p}^{\dot{A}} \wedge d\bar{q}_{\dot{A}} = 0 \\ &\tfrac{1}{2} dq^A \wedge dq_A + d\bar{s}^{\dot{A}} \wedge d\bar{q}_{\dot{A}} + \tfrac{1}{2} d\bar{p}^{\dot{A}} \wedge d\bar{p}_{\dot{A}} = 0 \end{aligned} \quad \begin{aligned} &\partial\mathcal{A}, \\ &\partial\bar{\mathcal{A}}. \end{aligned} \tag{4.18}$$

(Of course, it is enough to postulate $\partial\mathcal{A}$ in order to deduce $\partial\bar{\mathcal{A}}$ and vice versa.)

Now, we are going to show that $\mathcal{N}$ forms a natural part of some much wider structure of 1-forms, which, among other things, also describes the integral manifold of the second heavenly equation. For this purpose, we first respectively eliminate in the expressions for α and $\bar{\alpha}$ [the formulas (4.1) and (4.12)] $d\bar{q}_{\dot{B}} \wedge d\bar{q}^{\dot{B}}$ by using (4.11b), and $dq_B \wedge dq^B$ by using (4.16); this leads to the equations

$$\alpha^{alt} := (dp_A \wedge dq^A) \wedge (ds_B \wedge dq^B) = 0, \tag{4.19a}$$

$$\bar{\alpha}^{alt} := (d\bar{p}_{\dot{A}} \wedge d\bar{q}^{\dot{A}}) \wedge (d\bar{s}_{\dot{B}} \wedge d\bar{q}^{\dot{B}}) = 0. \tag{4.19b}$$

At this point it is convenient to observe that the numerical identity

$$\epsilon_{AB}\epsilon_{CD} + \epsilon_{BC}\epsilon_{AD} + \epsilon_{CA}\epsilon_{BD} = 0 \tag{4.20}$$

(any object skew in the three indices in two dimensions vanishes), when contracted with $dk^A \wedge dl^B \wedge dm^C \wedge dn^D$ provides a general Λ^4 identity

$$\begin{aligned} G: dk^A \wedge dl_A \wedge dm^B \wedge dn_B &+ dk^A \wedge dn_A \wedge dl^B \wedge dm_B \\ &+ dk^A \wedge dm_A \wedge dl^B \wedge dn_B = 0. \end{aligned} \tag{4.21}$$

In particular, identifying here $m^A = n^A$ we obtain a special identity

$$S: dk^A \wedge dm_A \wedge dl^B \wedge dm_B = -\tfrac{1}{2} dk^A \wedge dl_A \wedge dm^B \wedge dm_B. \tag{4.22}$$

Now, by using S, we can rewrite (4.19a)–(4.19b) in the form

$$\alpha^{alt} = -\tfrac{1}{2} ds^A \wedge dp_A \wedge dq^B \wedge dq_B = 0, \tag{4.23a}$$

$$\bar{\alpha}^{alt} = -\tfrac{1}{2} d\bar{s}^{\dot{A}} \wedge d\bar{p}_{\dot{A}} \wedge d\bar{q}^{\dot{B}} \wedge d\bar{q}_{\dot{B}} = 0. \tag{4.23b}$$

Now, rewrite (4.11b) and (4.16) in the form

$$-\tfrac{1}{2} d\bar{q}^{\dot{A}} \wedge d\bar{q}_{\dot{A}} = \tfrac{1}{2} dp^A \wedge dp_A + ds^A \wedge dq_A, \tag{4.24a}$$

$$-\tfrac{1}{2} dq^A \wedge dq_A = \tfrac{1}{2} d\bar{p}^{\dot{A}} \wedge d\bar{p}_{\dot{A}} + d\bar{s}^{\dot{A}} \wedge d\bar{q}_{\dot{A}}. \tag{4.24b}$$

By taking the external "squares" of the both sides of

(4.24a) and (4.24b), and by applying in doing so (4.22), we obtain the Λ^4 equations

$$\gamma := dp^A \wedge dp_A \wedge ds^B \wedge dq_B - \tfrac{1}{2} ds^A \wedge ds_A \wedge dq^B \wedge dq_B = 0, \tag{4.25a}$$

$$\gamma := d\bar{p}^{\dot{A}} \wedge d\bar{p}_{\dot{A}} \wedge d\bar{s}^{\dot{B}} \wedge d\bar{q}_{\dot{B}} - \tfrac{1}{2} d\bar{s}^{\dot{A}} \wedge d\bar{s}_{\dot{A}} \wedge d\bar{q}^{\dot{B}} \wedge d\bar{q}_{\dot{B}} = 0. \tag{4.25b}$$

Now, by application of **L**, we infer from (4.23a) and (4.23b) the existence of the functions such that

$$s^A\, dp_A + r^A\, dq_A = d\Theta, \tag{4.26a}$$

$$\bar{s}^{\dot{A}}\, d\bar{p}_{\dot{A}} + \bar{r}^{\dot{A}}\, d\bar{q}_{\dot{A}} = d\bar{\Theta}. \tag{4.26b}$$

But according to (4.13), the variables $\{p_A, q_A\}$ and $\{\bar{p}_{\dot{A}}, \bar{q}_{\dot{A}}\}$ are respectively independent. Therefore, we have

$$\Theta = \Theta(p_A, q_A), \quad \bar{\Theta} = \bar{\Theta}(\bar{p}_{\dot{A}}, \bar{q}_{\dot{A}}), \tag{4.27a}$$

$$s^A = \frac{\partial \Theta}{\partial p_A}, \quad \bar{s}^{\dot{A}} = \frac{\partial \bar{\Theta}}{\partial \bar{p}_{\dot{A}}}. \tag{4.27b}$$

Now, putting (4.27) into (4.25a) and (4.25b) one easily obtains

$$\gamma \equiv 2\left(\frac{1}{2}\frac{\partial^2 \Theta}{\partial p^A \partial p_B} \cdot \frac{\partial^2 \Theta}{\partial p_A \partial p^B} + \frac{\partial^2 \Theta}{\partial p^A\, \partial_A}\right) dV = 0, \tag{4.28a}$$

$$\bar{\gamma} \equiv 2\left(\frac{1}{2}\frac{\partial^2 \bar{\Theta}}{\partial \bar{p}^{\dot{A}} \partial \bar{p}_{\dot{B}}} \cdot \frac{\partial^2 \bar{\Theta}}{\partial \bar{p}_{\dot{A}} \partial \bar{p}^{\dot{B}}} + \frac{\partial^2 \bar{\Theta}}{\partial \bar{p}^{\dot{A}} \partial \bar{q}_{\dot{A}}}\right) dV = 0. \tag{4.28b}$$

It follows that Θ must fulfill the second heavenly equation, (3.5b) and $\bar{\Theta}$ fulfills a copy of the same equation in the variables $\bar{q}_{\dot{A}}$, $\bar{p}_{\dot{A}}$:

$$\frac{1}{2}\frac{\partial^2 \bar{\Theta}}{\partial \bar{p}^{\dot{A}} \partial \bar{p}_{\dot{B}}} \cdot \frac{\partial^2 \bar{\Theta}}{\partial \bar{p}_{\dot{A}} \partial \bar{p}^{\dot{B}}} + \frac{\partial^2 \bar{\Theta}}{\partial \bar{p}^{\dot{A}} \partial \bar{q}_{\dot{A}}} = 0. \tag{4.29}$$

We should like to observe that in Ref. 3 the fact that the first and the second heavenly equations are equivalent was described in an implicit manner only. In the present study, the above derived implication $(\alpha = \beta = 0) \to (\alpha^{alt} = \gamma = 0)$ explains a part of the mechanism of this equivalence. The inverse implication $(\alpha^{alt} = \gamma = 0) \to (\alpha = \beta = 0)$ can be also easily derived, indeed, $\gamma = 0$ is equivalent to the statement that the external square of the closed 2-form, $\frac{1}{2} dp^A \wedge dp_A + ds^A \wedge dq_A$, vanishes. Thus, this 2-form is simple, and hence by the application of the Darboux theorem, can be represented as $-d(\bar{q}_{\dot{1}}\, d\bar{q}_{\dot{1}} + d\tau) = -\frac{1}{2} d\bar{q}^{\dot{A}} \wedge d\bar{q}_{\dot{A}}$. Therefore, $\gamma = 0$ assures (4.24a). On the other hand, α^{alt} is equivalent to (4.19a) and by use of (4.24a) reduces to $\alpha = 0$. Moreover, (4.24a) multiplied externally by $dq^B \wedge dq_B$ clearly gives $\beta = 0$.

Now, in Ref. 5 it was found that in the study of the Killing vectors in strong heavens in the Θ function formalism, an important role is played by a new function, Λ. We will now be able to show that this function can be interpreted as a natural member of a structure of 1-forms, which naturally emerges by a further extension of the procedures applied in this section.

Indeed, by closing (4.26a) and (4.26b) we have

$$ds^A \wedge dp_A + dr^A \wedge dq_A = 0, \tag{4.30a}$$

$$d\bar{s}^{\dot{A}} \wedge d\bar{p}_{\dot{A}} + d\bar{r}^{\dot{A}} \wedge d\bar{q}_{\dot{A}} = 0. \tag{4.30b}$$

Now, externally multiplying (4.30a) and (4.30b) by $dp^A \wedge dq_A$ and $d\bar{p}^{\dot{A}} \wedge d\bar{q}_{\dot{A}}$, respectively, and by applying S, (4.22), we infer that

$$ds^A \wedge dq_A \wedge dp^B \wedge dp_B + dr^A \wedge dp_A \wedge dq^B \wedge dq_B = 0, \tag{4.31a}$$

$$d\bar{s}^{\dot{A}} \wedge d\bar{q}_{\dot{A}} \wedge d\bar{p}^{\dot{B}} \wedge d\bar{p}_{\dot{B}} + d\bar{r}^{\dot{A}} \wedge d\bar{p}_{\dot{A}} \wedge d\bar{q}^{\dot{B}} \wedge d\bar{q}_{\dot{B}} = 0. \tag{4.31b}$$

These relations can now be used in (4.25a) and (4.25b) transforming these equations to the form

$$\gamma = -(\tfrac{1}{2} ds^A \wedge ds_A + dr^A \wedge dp_A) \wedge dq^B \wedge dq_B = 0, \tag{4.32a}$$

$$\bar{\gamma} = -(\tfrac{1}{2} d\bar{s}^{\dot{A}} \wedge d\bar{s}_{\dot{A}} + d\bar{r}^{\dot{A}} \wedge d\bar{p}_{\dot{A}}) \wedge d\bar{q}^{\dot{B}} \wedge d\bar{q}_{\dot{B}} = 0. \tag{4.32b}$$

Consequently, by applying **L**, we infer the existence of functions such that

$$\tfrac{1}{2} s^A ds_A + r^A\, dp_A + t^A\, dq_A = d\Lambda, \tag{4.33a}$$

$$\tfrac{1}{2} \bar{s}^{\dot{A}}\, d\bar{s}_{\dot{A}} + \bar{r}^{\dot{A}}\, d\bar{p}_{\dot{A}} + \bar{t}^{\dot{A}}\, d\bar{q}_{\dot{A}} = d\Lambda. \tag{4.33b}$$

By closing these relations we have of course,

$$\tfrac{1}{2} ds^A \wedge ds_A + dr^A \wedge dp_A + dt^A \wedge dq_A = 0, \tag{4.34a}$$

$$\tfrac{1}{2} d\bar{s}^{\dot{A}} \wedge d\bar{s}_{\dot{A}} + d\bar{r}^{\dot{A}} \wedge d\bar{p}_{\dot{A}} + d\bar{t}^{\dot{A}} \wedge d\bar{q}_{\dot{A}} = 0. \tag{4.34b}$$

Now, if we understand Λ as $\Lambda = \Lambda(p_A, q_A)$, then from (4.33a) we have

$$\frac{\partial \Lambda}{\partial p_A} = r^A + \tfrac{1}{2} s^B \frac{\partial s_B}{\partial p_A}$$

so that (4.26a) implies

$$\frac{\partial \Theta}{\partial q_A} + \frac{1}{2}\frac{\partial \Theta}{\partial p_B}\frac{\partial^2 \Theta}{\partial p^B \partial p_A} = \frac{\partial \Lambda}{\partial p^A}. \tag{4.35}$$

Now due to the identity $(\partial/\partial p_A)(\partial/\partial p^A) = 0$, one easily sees that Eqs. (4.35) imply and are implied by the second heavenly equation (3.5b). Notice that (4.35) is just the spinorial version of the last two equations of Eqs. (2.33) of Ref. 5, which appear in the master equation for determining the Killing vectors in heaven.

Now our procedure of prolongations to obtain new Pfaffian 1-forms can be continued presumably indefinitely. However, there is one important difference. From the 1-forms we have constructed up to now, that is Eqs. (4.17), (4.26), and (4.33), we can choose any neighboring pair (4.33a) and (4.26a), (4.26a) and (4.10b), or (4.10b) and (4.10a), or the corresponding barred pairs to reconstruct the entire heavenly structure. This, however, appears not to be the case as we continue further up the ladder. That is, if we construct the next 1-form and its barred associate by applying the same techniques as previously we can not use this 1-form in conjunction with (4.33a) to derive the second heavenly equation or its associated 1-forms (4.26a) or (4.10b). It appears that above the Λ 1-form (4.33a) infinitely many 1-forms appear and that possibly all are needed to regain the entire structure. We will now write our *heavenly hierarchy* of 1-forms in a much more concise notation and also assign complex dilatation weights to the variables which appear.

We begin by noticing that all the 1-forms constructed so far enjoy a scale invariance of the following type:

$$\begin{aligned}&q_A \to \lambda \exp(-i\mu/2) q_A, \quad p_A \to \lambda \exp(i\mu/2) p_A,\\ &s_A \to \lambda \exp(i3\mu/2) s_A, \quad r_A \to \lambda \exp(5i\mu/2) r_A,\\ &t_A \to \lambda \exp(7i\mu/2) t_A, \quad \Omega \to \lambda^2 \Omega, \quad \Sigma \to \lambda^2 \exp(i\mu)\Sigma,\\ &\Theta \to \lambda^2 \exp(2i\mu)\Theta, \quad \Lambda \to \lambda^2 \exp(3i\mu)\Lambda,\end{aligned} \tag{4.36}$$

where $\lambda, \mu \in \mathbb{C}$. The corresponding transformations for the barred quantities can be obtained from (4.36) by putting a bar on the corresponding variables and changing $\mu \to -\mu$.

This invariance exhibits the fact that all our 1-forms can be characterized by their weights with respect to $\exp(i\mu)$. This suggests the following change of notation. We introduce spinors $\psi_A(j)$ and scalars $\Phi(l)$ defined by

$$\begin{aligned}&\psi_A(-\tfrac{1}{2}) = q_A, \quad \bar{\psi}_{\dot A}(\tfrac{1}{2}) = \bar{q}_{\dot A}, \quad \Phi(0) = -\Omega,\\ &\psi_A(\tfrac{1}{2}) = p_A, \quad \bar{\psi}_{\dot A}(-\tfrac{1}{2}) = \bar{p}_{\dot A}, \quad \Phi(1) = \Sigma, \quad \Phi(-1) = \bar{\Sigma},\\ &\psi_A(\tfrac{3}{2}) = s_A, \quad \bar{\psi}_{\dot A}(-\tfrac{3}{2}) = \bar{s}_{\dot A}, \quad \Phi(2) = \Theta, \quad \Phi(-2) = \bar{\Theta},\\ &\psi_A(\tfrac{5}{2}) = r_A, \quad \bar{\psi}_{\dot A}(-\tfrac{5}{2}) = \bar{r}_{\dot A}, \quad \Phi(3) = \Lambda, \quad \Phi(-3) = \bar{\Lambda},\\ &\psi_A(\tfrac{7}{2}) = t_A, \quad \bar{\psi}_{\dot A}(-\tfrac{7}{2}) = \bar{t}_{\dot A}.\end{aligned} \tag{4.37}$$

It is understood here that j is a half-odd integer, while l is an integer. Then we can extend $\psi_A(j)$ to all negative half-odd integers and $\bar{\psi}_{\dot A}(j)$ to all positive half-odd integers by

$$\psi_A(j) = 0 \leftarrow j < -\tfrac{1}{2}, \tag{4.38a}$$

$$\bar{\psi}_{\dot A}(j) = 0 \leftarrow j > \tfrac{1}{2}. \tag{4.38b}$$

Now the important point is that we can apparently also extend $\psi_A(j)$, $\bar{\psi}_{\dot A}(j)$ to the remaining half-odd integers and $\Phi(l)$ to all integers by the prolongation process. Indeed using (4.37) and (4.38), we can write all our previous Pfaffian 1-forms (4.17), (4.26), and (4.33) succinctly as

$$d\Phi(l) = \tfrac{1}{2} \sum_j \{\psi^A(l-j)\, d\psi_A(j) + \bar{\psi}^{\dot A}(l-j)\, d\bar{\psi}_{\dot A}(j)\}, \tag{4.39}$$

where j runs over all half-odd integers. Now the previously obtained 1-forms are given by the range $l = -3, \ldots, 3$. However, we have checked the validity of (4.39) for the larger range $l = -7, \ldots, 7$. Indeed it appears that (4.39) is valid for *all* integers l. Hence, we conjecture that the *heavenly hierarchy* given by (4.39) is, in fact, infinite. We have not been able to prove our conjecture, however. One might think that an inductive proof would work, but a closer examination shows that one must invoke the induction hypothesis at each stage of the prolongation process, i.e., it is necessary to alternate invoking the induction hypothesis with implementing Lemma **L**. In spite of this we see no reason why the prolongation process should break down for higher values of l.

Now the closure relations for (4.39) are given by

$$\omega(l) \equiv \tfrac{1}{2} \sum_j \{d\psi^A(l-j) \wedge d\psi_A(j) + d\bar{\psi}^{\dot A}(l-j) \wedge d\bar{\psi}_{\dot A}(j)\} = 0 \tag{4.40}$$

It is clear that the relations (4.40) split in a natural fashion into three subfamilies: the *pure heavenly* subfamily (no dotted spinors)

$$\omega(l) = \tfrac{1}{2} \sum_j d\psi^A(l-j) \wedge d\psi_A(j) = 0, \quad l \geq 2; \tag{4.41}$$

the *pure hellish* subfamily (no undotted spinors)

$$\omega(l) = \tfrac{1}{2} \sum_j d\bar{\psi}^{\dot A}(l-j) \wedge d\bar{\psi}_{\dot A} = 0, \quad l \leq -2; \tag{4.42}$$

and the subfamily where the three forms, $d\Phi(1)$, $d\Phi(0)$, and $d\Phi(-1)$ necessarily mix the undotted spinors. The corresponding equations are of course

A_+
A_-

$$\begin{aligned}\omega(1) &= \tfrac{1}{2} d\psi^A(\tfrac{1}{2}) \wedge d\psi_A(\tfrac{1}{2}) + d\psi^A(\tfrac{3}{2}) \wedge d\psi_A(-\tfrac{1}{2})\\ &\quad + \tfrac{1}{2} d\bar{\psi}^{\dot A}(\tfrac{1}{2}) \wedge d\bar{\psi}_{\dot A}(\tfrac{1}{2}) = 0,\\ \omega(0) &= d\psi^A(\tfrac{1}{2}) \wedge d\psi_A(-\tfrac{1}{2}) + d\bar{\psi}^{\dot A}(-\tfrac{1}{2}) \wedge d\bar{\psi}_{\dot A}(\tfrac{1}{2}),\\ \omega(-1) &= \tfrac{1}{2} d\psi^A(-\tfrac{1}{2}) \wedge d\psi_A(-\tfrac{1}{2}) + d\bar{\psi}^{\dot A}(-\tfrac{3}{2}) \wedge d\bar{\psi}_{\dot A}(\tfrac{1}{2})\\ &\quad + \tfrac{1}{2} d\bar{\psi}^{\dot A}(-\tfrac{1}{2}) \wedge d\bar{\psi}_{\dot A}(-\tfrac{1}{2}) = 0.\end{aligned} \tag{4.43}$$

It is quite clear that A^+ or A^- assumed is enough to reproduce *all* the structure considered.

Finally we mention that from (3.12) one easily sees that equalities (4.43) amount to the description of the forms $s^{\dot A \dot B}$ through the alternative formulas

$$\begin{aligned}s^{\dot 1 \dot 1} &= d\psi^A(-\tfrac{1}{2}) \wedge d\psi_A(-\tfrac{1}{2})\\ &= -2\, d\bar{\psi}^{\dot A}(-\tfrac{3}{2}) \wedge d\bar{\psi}_{\dot A}(\tfrac{1}{2}) - d\bar{\psi}^{\dot A}(-\tfrac{1}{2}) \wedge d\bar{\psi}_{\dot A}(-\tfrac{1}{2}),\\ s^{12} &= -d\psi^A(\tfrac{1}{2}) \wedge d\psi_A(-\tfrac{1}{2}) = +d\bar{\psi}^{\dot A}(-\tfrac{1}{2}) \wedge d\bar{\psi}_{\dot A}(\tfrac{1}{2}),\\ s^{\dot 2 \dot 2} &= -d\psi^{\dot A}(\tfrac{1}{2}) \wedge d\bar{\psi}_{\dot A}(\tfrac{1}{2}) = 2 d\psi^A(\tfrac{3}{2}) \wedge d\psi_A(-\tfrac{1}{2})\\ &\quad + d\psi^A(\tfrac{1}{2}) \wedge d\psi_A(\tfrac{1}{2}).\end{aligned} \tag{4.44}$$

5. GENERAL PROPERTIES OF THE INTEGRAL MANIFOLDS

In this section we consider some important properties of the heavenly integral manifolds. While we have not been able to find an explicit expression for the general solution, we can use Cartan's theory[8,10] to construct stepwise the regular integral manifolds in terms of their tangent spaces. This will allow us, for example, to determine at each step the arbitrariness of the integral manifolds, i.e., on how many arbitrary functions of how many variables the general manifold depends. To do this we can begin with any of the equivalent forms of the heavenly manifolds. It seems best to use the already partially integrated description given in terms of the two Pfaffian 1-forms (4.10b) and (4.26a). Here we write them in component form

$$\begin{aligned}\omega_1 &= d\Theta + s_2\, dp_1 - s_1\, dp_2 + r_2\, dq_1 - r_1\, dq_2,\\ \omega_2 &= d\Sigma + \tfrac{1}{2} s_2\, dq_1 - \tfrac{1}{2} s_1\, dq_2 + \tfrac{1}{2} \bar{q}_{\dot 2}\, d\bar{q}_{\dot 1} - \tfrac{1}{2} \bar{q}_{\dot 1}\, d\bar{q}_{\dot 2}\\ &\quad + \tfrac{1}{2} p_2\, dp_1 - \tfrac{1}{2} p_1\, dp_2.\end{aligned} \tag{5.1}$$

These forms are, of course, zero on an integral manifold. The closure easily gives

$$\begin{aligned}d\omega_1 &= ds_2 \wedge dp_1 - ds_1 \wedge dp_2 + dr_2 \wedge dq_1 - dr_1 \wedge dq_2,\\ d\omega_2 &= ds_2 \wedge dq_1 - ds_1 \wedge dq_2 + d\bar{q}_{\dot 2} \wedge d\bar{q}_{\dot 1} + dp_2 \wedge dp_1.\end{aligned} \tag{5.2}$$

Now we are working on complex Euclidean n-space where $n = 12$. We wish to find the regular integral manifolds by successive applications of the Cauchy–Kowalewski theorem. Now at a regular point in $\mathbb{C}^{12}$, the rank of the system (5.1) is 2, so the Cartan character $s_0 = 2$. A vector in the tangent plane to a solution must satisfy

$$X \lrcorner\, \omega_1 = X \lrcorner\, \omega_2 = 0 \tag{5.3}$$

where $\lrcorner$ denotes the inner product between differential forms and vector fields. The polar system is obtained by adjoining to (5.1) the 1-forms

$$X_1 \lrcorner d\omega_1, \quad X_1 \lrcorner d\omega_2, \tag{5.4}$$

where X satisfies (5.3). This system has rank 4 (i.e., $s_0 + s_1 = 4$), so $s_1 = 2$. A two-dimensional integral manifold is obtained by constructing X_2 to satisfy (5.3) and

$$X \lrcorner (X_1 \lrcorner d\omega_1) = X \lrcorner (X_1 \lrcorner d\omega_2) = 0. \tag{5.5}$$

Its polar system is obtained by adding to (5.1) and (5.4) the 1-forms $X_2 \lrcorner d\omega_1$ and $X_2 \lrcorner d\omega_2$ which has rank 6 and thus $s_2 = 2$. Continuing in this way we obtain a three-dimensional integral manifold with tangent vectors $\{X_1, X_2, X_3\}$, where X_3 satisfies (5.3), (5.5), and (5.5) with X_1 replaced by X_2. The polar system is obtained by adding $X_3 \lrcorner d\omega_1$ and $X_3 \lrcorner d\omega_2$ to the previous polar system. Its rank is 8, thus $s_3 = 2$. The four-dimensional integral manifolds $\{X_1, X_2, X_3, X_4\}$ are constructed as before with X_4 orthogonal to the last constructed polar system. However, if we add $X_4 \lrcorner d\omega_1$ and $X_4 \lrcorner d\omega_2$ to this polar system, its rank remains 8, since we are in $\mathbb{C}^{12}$ and $12 - 4 = 8$. Thus the genus $g = 4$ and the maximal regular integral manifolds have complex dimension four, which of course we already knew. We also have $s_4 = 0$. Now we can use Cartan's criteria (Ref. 8, p. 75) to state for example, that the general solution for heavens depends on *two arbitrary functions of three complex variables*. To sum up, the regular maximal integral manifolds of heaven are determined by

$$\begin{aligned} &X_i \lrcorner \omega_1 = X_i \lrcorner \omega_2 = 0, \\ &X_i \lrcorner (X_j \lrcorner d\omega_1) = X_i \lrcorner (X_j \lrcorner d\omega_2) = 0, \quad i \neq j, \end{aligned} \tag{5.6}$$

$i, j = 1, \ldots, 4$. The only qualification that we must add is that (4.4) be satisfied, i.e., that the system (5.1) be in involution with respect to the spinors q_A, p_A. Equations (5.6) give, at least implicitly, the general integral manifolds of heaven in terms of the tangent spaces at each point.

In order to find explicit integral manifolds for heaven, we deal with the second heavenly equation in the form given by (4.25a) and the 2-form $d\omega$, given by (5.2). In component form, (4.25a) reads

$$\begin{aligned} &ds_1 \wedge ds_2 \wedge dq_1 \wedge dq_2 + ds_2 \wedge dq_1 \wedge dp_1 \wedge dp_2 - ds_1 \wedge dq_2 \\ &\quad \wedge dp_1 \wedge dp_2 = 0 \end{aligned} \tag{5.7}$$

on an integral manifold. In fact we would like to be able to linearize, at least partially, the differential equations for an integral manifold. Indeed we will see that the case treated in detail in Sec. 2 is a special case of the linearization that follows. Again as in Sec. 2, the trick is to integrate the equation $d\omega_1 = 0$, treating q_1, q_2, p_1, and s_1 as independent variables. We have the existence of a function $\Psi(p_1, s_1, q_1, q_2)$ with

$$d\Psi - s_2\, dp_1 - p_2\, ds_1 - r_2\, dq_1 + r_1\, dq_2 = 0 \tag{5.8}$$

on an integral manifold. Plugging (5.8) back into (5.7) we obtain the differential equation

$$\Psi_{p_1 s_1} \Psi_{s_1 q_2} - \Psi_{p_1 q_2} \Psi_{s_1 s_1} + \Psi_{s_1 q_1} - \Psi_{p_1 p_1} = 0. \tag{5.9}$$

The advantage of this form is the following: As seen from (2.7), the derivatives of Θ with respect to q_1 do not enter into the calculation of the metric, and thus it is also this way with Ψ. Moreover, the nonlinear terms in (5.9) contain derivatives with respect to the q_A's. Thus (5.9) is susceptible to linearization involving nontrivial metrics. We mention that the condition that the original variables p_A, q_A be independent (i.e., $dV \neq 0$) imply that $\Psi_{s_1 s_1} \neq 0$, and that only those heavenly manifolds such that $\Theta_{p_2 p_2} \neq 0$ are amenable to the above treatment. The case $\Theta_{p_2 p_2} = 0$ is easily handled, however, as shown in Ref. 3.

Now it is straightforward to determine the metric in terms of the function Ψ. Indeed, the necessary derivatives are

$$\begin{aligned} &\Theta_{p_2 p_2} = \Psi^{-1}_{s_1 s_1}, \quad \Theta_{p_1 p_2} = -\Psi^{-1}_{s_1 s_1} \Psi_{s_1 p_1}, \\ &\Theta_{p_1 p_1} = \Psi^{-1}_{s_1 s_1} \Psi^2_{s_1 p_1} - \Psi_{p_1 p_1}. \end{aligned} \tag{5.10}$$

Simple substitution of (5.10) into (3.7) then gives the spinorial components of the metric in the Ψ formalism. Similarly the connection and curvature components can be computed; however, we do not give these explicitly. The important point is that as in the Θ formalism, the metric, connections, and curvature do not involve derivatives with respect to q_A.

With this in mind we look for solutions of (5.9) with $\Psi_{s_1 q_2} = 0$. [The counterpart in the Θ formalism is $\Theta_{p_2 q_2} = \Theta_{p_1 q_2} = 0$ which does not enter the second heavenly equation (3.5b) explicitly.] Thus Ψ can be written as $\Psi = F(p_1, s_1, q_1) + \psi(p_1, q_1, q_2)$ and the analysis of (5.9) splits into two cases depending on whether $F_{s_1 s_1 s_1}$ vanishes or not.

Case 1: $F_{s_1 s_1 s_1} \neq 0$.

This case reduces to the three dimensional *complex Laplace equation* after some gauging and changes of variables,

$$\Psi_{p_1 p_1} + \Psi_{p_3 p_3} + \Psi_{p_4 p_4} = 0, \tag{5.11}$$

where now Ψ is a function of p_1, p_3, p_4, and

$$\begin{aligned} &p_3 = s_1 + \zeta, \quad p_4 = iq_1 + i\zeta, \\ &\zeta := s_1 + \int \alpha(q_1)\, dq_1, \quad \alpha(q_1) := \psi_{p_1 q_2}. \end{aligned}$$

There are many ways, of course, to solve (5.11) depending on different domains of holomorphy.[12] The general solution can be given explicitly and depends on two holomorphic functions of two complex variables. We mention in addition, regarding solution techniques for (5.11), Ref. 13 where group theoretical techniques are used and Ref. 14 where an operational calculus approach is used. We also mention that the special case when $\alpha(q_1) = 1$ and Ψ is independent of q_1 reduces to the two-dimensional Laplace equation treated in Sec. 2.

Case 2: $F_{s_1 s_1 s_1} = 0$.

This case has two vanishing conformal curvature components, i.e., $C^{(5)} = C^{(4)} = 0$. The quantities necessary to compute the metric are

$$\begin{aligned} &\Psi_{s_1 s_1} = g^1(q_1) p_1 + g^0(q_1), \\ &\Psi_{s_1 p_1} = g^1(q_1) s_1 + \tfrac{1}{2} g^1_{q_1} p_1^2 + g^0_{q_1} p_1 + f^1(q_1), \\ &\Psi_{p_1 p_1} = \Psi_{s_1 s_1 q_1} s_1 + \tfrac{1}{6} g^1_{q_1 q_1} p_1^3 + \tfrac{1}{2} g^0_{q_1 q_1} p_1^2 + f^1_{q_1} p_1 \\ &\qquad + f^0_{q_1}(q_1) + \Psi_{s_1 s_1} h(\xi, q_1), \end{aligned} \tag{5.12}$$

where g^0, g^1, f^0, f^1 are arbitrary functions of q_1, h is an arbitrary function of its arguments, and $\xi := q_2 - \frac{1}{2}g^1(p_1 + g^0/g^1)^2$. This case has some overlap with the case treated beginning with (4.12a) in Ref. 5, but in general they are not equivalent.

It is clear from the above analysis that many classes of metrics appear and can be given explicitly and hence studied in much more detail along the lines of Sec. 2. We will not do this here, however. Finally, it is mentioned that a similar linearization yielding nontrivial metrics is obtained by setting $\Psi_{s_1 q_2} = 0$ in (5.9).

6. SYMMETRIES OF THE SECOND HEAVENLY EQUATION

In this section we describe the symmetries of the second heavenly equation. In fact, we show that essentially the infinitesimal symmetries coincide with the Killing vectors obtained in Ref. 5 aside from the function Λ, which we have already seen arises from the prolongation process described in Sec. 4. Generally it would be of interest to study the symmetry of the complete heavenly hierarchy or at least the system of 1-forms which begin with and end with Λ (i.e., $l = -3, \ldots, 3$). This could shed light on the meaning of the hierarchy. However, we content ourselves here with finding the infinitesimal symmetries of the pair of Pfaffian 1-forms (4.26a) and (4.33a). The reason for choosing here the 1-form (4.33a) instead of (4.10b) to represent the heavenly integral manifolds is that it allows for the dependence of the symmetries on Λ which is of interest from the point of view of the Killing vectors.[5] We will show, however, that this dependence is not allowed as transformations on the space spanned in a local chart by $(q_A, p_A, \circledcirc)$.

Now let M be a differential manifold and let $\mathcal{I}$ be an ideal in the Grassmann algebra $\Lambda(M)$. Also let $\mathcal{I}$ be closed under exterior differentiation. Suppose that $\mathcal{I}$ is generated by ω_i and $d\omega_i$ and let (N, i) be an immersed submanifold which annuls $\mathcal{I}$, i.e., an integral manifold. Then the local symmetry group for $\mathcal{I}$ is given by the set of all local diffeomorphisms $\phi: M \to M$ such that

$$\phi^*\omega \in \mathcal{I} \quad \text{for all } \omega \in \mathcal{I}, \tag{6.1}$$

where ϕ^* denotes the pullback of ϕ. Hence ϕ is a mapping on M such that the integral manifolds of $\mathcal{I}$ are mapped into themselves. Infinitesimally (6.1) reads

$$\pounds_X \omega = \lambda^i \omega_i \tag{6.2}$$

for all $\omega \in \mathcal{I}$ and where $\omega_i \in \mathcal{I}$, λ^i are locally holomorphic functions on M, and X is the vector field describing a local one-parameter trajectory ϕ_t.

Now let us apply (6.2) to the ideal $\mathcal{I}$ of differential forms generated by the two 1-forms (4.26a), (4.33a), and their closures (4.30a), (4.34a). Explicitly, we write

$$\omega_1 = d\circledcirc - s^A\, dp_A - r^A\, dq_A, \tag{6.3a}$$

$$\omega_2 = d\Lambda - \tfrac{1}{2}s^A\, ds_A - r^A\, dp_A - t^A\, dq_A. \tag{6.3b}$$

Then applying (6.2) we have

$$\pounds_X \omega_1 = \lambda_1^1 \omega_1 + \lambda_1^2 \omega_2, \tag{6.4a}$$

$$\pounds_X \omega_2 = \lambda_2^1 \omega_1 + \lambda_2^2 \omega_2. \tag{6.4b}$$

We mentioned that the commutivity of the Lie derivative and the exterior derivative applied to (6.4) guarantees that the 2-forms $d\omega_1$ and $d\omega_2$ are back in $\mathcal{I}$ after an infinitesimal transformation. Now in order to solve (6.4) we write the ω's out explicitly and make use of the identity[15]

$$\pounds_X \omega = d(X \lrcorner\, \omega) + X \lrcorner\, d\omega. \tag{6.5}$$

Then equating the coefficients of the independent 1-forms on the space $\mathbb{C}^{12}$ we obtain a system of first order coupled partial differential equations for the vector fields X. In order to facilitate matters it is convenient to define functions F and G by

$$F = X \lrcorner\, \omega_1, \quad G = X \lrcorner\, \omega_2. \tag{6.6}$$

Then upon equating coefficients in (6.4) we obtain the equations

$$\begin{aligned}
&F_{t_A} = 0, \quad X^{q^A} = F_{r_A} = G_{t_A},\\
&X^{p^A} = F_{s_A} + \tfrac{1}{2}s^A F_\Lambda = G_{r_A},\\
&X^{s^A} = F_{p_A} + s^A F_{\circledcirc} + r^A F_\Lambda = G_{s_A} + \tfrac{1}{2}s^A G_\Lambda,\\
&X^{r^A} = F_{q_A} + r^A F_{\circledcirc} + t^A F_\Lambda = G_{p_A} + s^A G_{\circledcirc} + r^A G_\Lambda,
\end{aligned} \tag{6.7}$$

where X^{q^A} denotes the component of X multiplying ∂_{q^A}, etc. The first three of Eqs. (6.7) can be integrated immediately to give

$$\begin{aligned}
&F = F^A r_A + F^0,\\
&G = F^A t_A + (F^0_{s_A} + \tfrac{1}{2}s^A F^0_\Lambda) r_A + G^{00},
\end{aligned} \tag{6.8}$$

where F^A, F^0, and G^{00} are arbitrary functions of the spinors q_A, p_A, s_A and scalars $\circledcirc, \Lambda$. Plugging (6.8) into the last two of Eqs. (6.7) and doing some algebra, we have

$$\begin{aligned}
&F^0_\Lambda = 0, \quad F^0 = H^A s_A + H^0, \quad F^B_\Lambda = F^B_{s_A} = F^B_{\circledcirc} = F^B_{p_A} = 0,\\
&F^0_{p_A} + s^A F^0_{\circledcirc} = G^{00}_{s_A} + \tfrac{1}{2}s^A G^{00}_\Lambda,\\
&G^{00}_{p_A} + s^A G^{00}_{\circledcirc} = H^B_{q_A} s_B + H^0_{q_A},\\
&F^B_{q_A} + \epsilon^{BA} s_C H^C + \epsilon^{BA} H^0_{\circledcirc} = H^B_{p_A} + \tfrac{1}{2}s^A H^B_{\circledcirc} + \epsilon^{BA} G^{00}_\Lambda,
\end{aligned} \tag{6.9}$$

where H^A and H^0 are arbitrary functions of q_A, p_A, $\circledcirc$, and Λ. The integration of (6.9) is straightforward but rather tedious. First we notice that F^B is a function only of q_A and after some algebra we find that H^B, H^0, and G^{00} must have the forms

$$\begin{aligned}
&H^B = H^{BC}(q) p_C + H^{B0}(q),\\
&H^0 = h^1(q)\circledcirc + h^0(q, p),\\
&G^{00} = g^2(q)\Lambda + g^1(q, p)\circledcirc + g^0(q, p, s),
\end{aligned} \tag{6.10}$$

and we are left with the constraint equations

$$\begin{aligned}
&g^0 = \tfrac{1}{2}F^A_{q_B} s_A s_B + h^0_{p_A} s_A + \bar{g}^{00}(q, p),\\
&H^{BA} = F^B_{q_A} + \epsilon^{BA}(h^1 - g^2),\\
&g^0_{p_A} = (H^{BC}_{q_A} - \epsilon^{BA} h^1_{q_C}) p_C s_B + (H^{B0}_{q_A} - \alpha^1 \epsilon^{BA}) s_B + h^0_{q_A},
\end{aligned} \tag{6.11}$$

where h^0 and α^1 are functions of q_A only. Upon further integrations of (6.11) we find that both h^1 and g^2 must be constants and

$$F^A = \phi_{q_A} + C_1 \epsilon^{AB} q_B,$$

$$
\begin{aligned}
H^{AB} &= \phi_{q_A q_B} + C_2 \epsilon^{BA}, \\
H^{A0} &= \psi_{q_A} - a_0 \epsilon^{AB} q_B, \qquad\qquad (6.12)\\
h^0 &= \tfrac{1}{6}\phi_{q_A q_B q_C} p_A p_B p_C + \tfrac{1}{2}\psi_{q_A q_B} p_A p_B + \chi_{q_A} p_A + h^{00}(q), \\
g^0 &= \tfrac{1}{24}\phi_{q_A q_B q_C q_D} p_A p_B p_C p_D + \tfrac{1}{6}\psi_{q_A q_B q_C} p_A p_B p_C \\
&\quad + \tfrac{1}{2}\phi_{q_A q_B q_C} p_A p_B s_C + \psi_{q_A q_B} p_A s_B + \tfrac{1}{2}\chi_{q_A q_B} p_A p_B \\
&\quad + h^{00}_{q_A}(q) p_A + g^{00}(q).
\end{aligned}
$$

This ends the computation of the infinite dimensional Lie algebra $\mathcal{L}$ of infinitesimal symmetries of the heavenly manifolds. It is not difficult to see that the only symmetry in (6.12) which is not a projection onto transformations of the space with local coordinates $(q_A, p_A, \circledast)$ is that symmetry generated by the function $g^{00}(q)$. This function generates q_A dependent translations of Λ and leaves (6.3b) invariant since the spinor t_A is essentially arbitrary. The Lie algebra $\mathcal{L}_0$ generated by these translations is therefore less interesting. Indeed, it can be seen that $\mathcal{L}_0$ is an ideal in $\mathcal{L}$ and we thus consider the factor algebra $\mathcal{L}/\mathcal{L}_0$. The projections of these onto the base space spanned by $(q_A, p_A, \circledast)$ are given by the vector fields

$$
\begin{aligned}
X^{q^A} &= \phi_{q_A} + C_1 \epsilon^{AB} q_B, \\
X^{p^A} &= \phi_{q_A q_B} p_B + C_2 \epsilon^{BA} p_B + \psi_{q_A} - a_0 \epsilon^{AB} q_B, \qquad (6.13)\\
X^{\circledast} &= (C_1 + 3C_2)\circledast + h^0.
\end{aligned}
$$

It is now easy to see that the vector fields (6.13) are precisely the Killing vectors in spinorial notation given by Eq. (2.33) of Ref. 5, with $\alpha_0 = 0$. On the other hand, we have understood the α_0 term in terms of the prolongation variable Λ in Sec. 4.

Finally, we mention the possible use of the symmetries to obtain solutions of the second heavenly equation. Given a symmetry vector field we can find relative invariants which essentially reduces the number of variables of the original partial differential equation by one. Indeed if we know three independent symmetries we can reduce the problem to quadratures. Moreover, given any solution we can obtain other solutions by group transformations.

Closing this paper, we should like to conclude that we believe that its results, although not as complete as one might desire, seem to justify our belief that (i) it is profitable to use an abbreviated spinorial notation as introduced in Sec. 3, and (ii) that the apparatus of the canonical Cartan's theory of integral manifolds is suitable when striving towards better understanding of the nature of heavens.

ACKNOWLEDGMENTS

We would like to acknowledge an illuminating discussion with Dr. M. Flato during the International Symposium on Mathematical Physics in Mexico City (January 1976). The kind interest in this work of Dr. J.D. Finley, III is also gratefully appreciated.

*On leave of absence from the University of Warsaw, Warsaw, Poland.

[1]E.T. Newman, Tel Aviv GR7 Lecture (1974); Riddle of Gravity Symposium in honor of P. Bergmann, Syracuse, N.Y. (1975); to Proceedings of International Symposium on Mathematical Physics, México D.F., pp. 747–56 (1976).

[2]R. Penrose, First Award Gravity Foundation Essay, "The Nonlinear Graviton (1975); "Nonlinear Graviton and Curved Twistor Theory," Oxford Preprint.

[3]J.F. Plebański, J. Math. Phys. **16**, 2395 (1975).

[4]J.F. Plebański and S. Hacyan, J. Math. Phys. **16**, 2403 (1975).

[5]J.D. Finley, III and J.F. Plebański, J. Math. Phys. **17**, 585 (1976).

[6]S. Hacyan and J.F. Plebański, "Some Basic Properties of Killing Spinors," C.I.E.A. preprint (1976).

[7]E. Cartan, *Oeuvres Completes* (Gauthier-Villars, Paris, 1955), Partie 2, Vol. 2, pp. 571–714.

[8]E. Cartan, *Les systèmes differentiels extérieurs et leurs applications géométriques* (Hermann, Paris, 1945).

[9]H.D. Wahlquist and F.B. Estabrook, J. Math. Phys. **16**, 1 (1975).

[10]W. Slebodzinski, *Exterior Forms and Their Applications* (Monografie Matematyczne, Warszawa, 1970).

[11]J.F. Plebański, *Spinors, Tetrads and Forms* (Monograph of Centro de Invest. y Estudios Avanzados del I.P.N., Mexico City, 1974).

[12]R. Gilbert, *Function Theoretic Methods in Partial Differential Equations* (Academic, New York, 1969).

[13]C.P. Boyer, E.G. Kalnins, and W. Miller, Jr. Nagoya Math. J. **60**, 35 (1976).

[14]L. Infeld and J.F. Plebański, Bull. Acad. Pol. Sci. III 4, 215 (1956).

[15]R. Hermann, *Differential Geometry and the Calculus of Variations* (Academic, New York, 1968).

Symmetries of the Hamilton–Jacobi equation

C. P. Boyer

Instituto de Investigaciónes en Mathemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, Mexico 20, D. F.

E. G. Kalnins

Mathematics Department, University of Waikato, Hamilton, New Zealand
(Received 12 July 1976)

We present a detailed discussion of the infinitesimal symmetries of the Hamilton–Jacobi equation (an arbitrary first order partial differential equation). Our presentation elucidates the role played by the characteristic system in determining the symmetries. We then specialize to the case of a free particle in one space and one time dimension, and study the local Lie group of point transformations locally isomorphic to O(3,2). We show that the separation of variables of the corresponding Hamilton–Jacobi equation in the form of a sum is related to orbits in the Schrödinger subalgebra of o(3,2). The remaining orbits of o(3,2) yield symmetry related solutions which separate in more complicated product forms. Finally some connections with the primordial equation of hydrodynamics (without force terms) are made.

INTRODUCTION

One of the most important techniques in finding explicit solutions of partial differential equations is that of Lie group theory. This is said while keeping in mind the recent developments which illustrate the intimate connection of the time honored method of separation of variables with the theory of Lie groups.[1–5] Up to now most of this development has treated only second order linear partial differential equations, although the first and perhaps best understood example of separation of variables occurred for the nonlinear Hamilton–Jacobi equation.[6–9] Indeed there is a close connection between the separation of variables for second order linear partial differential equations of hyperbolic–elliptic type and the corresponding quadratic Hamilton–Jacobi equation which describe the characteristic surfaces of the former. This connection is usually described in the dual formulation in terms of a covariant Riemannian metric[10] $ds^2 = g_{ij}\,dx^i\,dx^j$. However, even for parabolic equations like the time dependent Schrödinger and heat equations we will see that the connection with a Hamilton–Jacobi equation of first degree in the temporal derivative remains, in the sense that they both admit the same type of separable coordinates. This is no doubt related to the fact that such coordinates are projectively related to quadratic surfaces in a higher dimensional pseudo-Riemannian space. However, we will show shortly how the elliptic Hamilton–Jacobi equation (sums of quadratics) is related by a simple point transformation to the parabolic Hamilton–Jacobi equation (first order derivative in time). It is also emphasized that the separation of the parabolic type presents a unified picture of four types[2,3] of potentials V, the free particle ($V=0$), the linear potential ($V=ax$), and the attractive and repulsive harmonic oscillators ($V=\pm\,\omega x^2$).

Now generally any first order partial differential equation can be cast by the process of embedding in a space of one higher dimension, into the Hamilton–Jacobi form

$$S_t + H(x^i, p_i, t) = 0,\qquad p_i = S_{x^i} \tag{0.1}$$

(subindices with respect to variables denote differentiation). The importance of this equation in geometrical optics, the calculus of variations, and obtaining explicit solutions of Hamilton's equations of classical mechanics is well known. (For the classical treatment see Chap. 2 of Ref. 11; for modern treatments see Chap. 13 of Ref. 12 and Chap. 4 of Ref. 13.) There is also a close connection with the theory of canonical transformations which we mention briefly here since the treatment in the sequel is complementary to this in the sense that it relates to contact transformations. Indeed consider a manifold (Hamiltonian manifold) with local coordinates (x^i, p_i, t) which has a closed 2-form ω and a function H such that

$$\omega = dp_i \wedge dx^i - dH \wedge dt. \tag{0.2}$$

Now each submanifold such that $\omega = 0$ implies the existence of a function $S(x^i, t)$ which is a solution of the Hamilton–Jacobi equation (0.1) (for more details see, e.g., Chap. 13 of Ref. 12). On the other hand, if we consider $-H$ as a coordinate, then the transformations which leave ω invariant form the pseudogroup of canonical transformations over a $(2n+2)$-dimensional manifold. Then restricting H to be a function will give a subpseudogroup which depends upon H, of course.

We now consider the special case of a free particle in a Riemannian (or pseudo-Riemannian) n-space with contravariant metric g^{ij}. Then (0.1) becomes

$$S_t + g^{ij} S_{x^i} S_{x^j} = 0\,. \tag{0.3}$$

If we introduce a change of variables $T = t + S$, $z = t - S$, an easy calculation shows that (0.3) is equivalent to

$$T_z^2 + g^{ij} T_{x^i} T_{x^j} = 1 \tag{0.4}$$

as long as both sets (x^i, t) and (x^i, z) can be treated as independent variables. Two comments are in order: First, the local symmetry group of point transformations of (0.3) and (0.4) are isomorphic. It was shown in Ref. 14 that when g^{ij} is the flat Euclidean metric, the local symmetry group of point transformations of (0.3) is a factor group of order 2 of $O(n+2, 2)$. Second, the above change of coordinates involving the dependent variable shows that (0.3) is equivalent to a Riemannian (pseudo-Riemannian) metric.

In this paper we study in detail the symmetries and separable coordinates of the equation

$$S_t + S_x^2 = q + p^2 = 0. \qquad (*)$$

This equation can be obtained from (0.3) by partial separation, at least in the case when g^{ij} admits a Killing vector. Thus from the point of view of separable coordinates we only study here subgroup coordinates. In fact the more general point transformation symmetries of (*) will yield coordinates not associated with the usual separation of variables. From this point of view the similarity solutions or complete integrals we obtain are more general than ordinary R-separation; however, we do not study here the usual quadratic orthogonal separation involving quadratic forms. Those, of course, do not appear in (*), but they will appear in the analog of (0.4), i.e.,

$$T_z^2 + T_x^2 = 1. \qquad (**)$$

We plan, to treat these in a subsequent work. Recently[15] it was shown that in a Riemannian or pseudo-Riemannian metric space there are two types of separation, those coming from local symmetry groups and those coming from the usual orthogonal separation, and that the latter are described by contravariant quadratic symmetric forms (Killing tensors).

The outline of the paper is as follows: In Sec. 1 we compute the Lie algebra of vector fields depending on both coordinates and momenta which are infinitesimal symmetries for an arbitrary first-order partial-differential equation. This computation elucidates the role played by the characteristic system in determining the symmetries. We discuss some of the underlying structure of this infinite-dimensional Lie algebra. Then we specialize to the subalgebra of point transformation symmetries of (*). These generate a finite-dimensional local Lie group-conformal transformations in R^3, locally isomorphic to O(3, 2). We then classify the orbits in the Lie algebra O(3, 2) under conjugacy with respect to the group. In Sec. 2 we obtain all R-separable coordinates systems for (*). In Sec. 3 we present a similarity solution[16] for each of the orbit representatives found in Sec. 1 and discuss the connection with the separation of variables of Sec. 2. Some remarks concerning the general solution and characteristic vector fields are also made.

Finally, in Sec. 4 we present a discussion of symmetries which derives from the fact that the x derivative of (*) yields the primordial equation of hydrodynamics without force terms[11,16]

$$p_t + 2pp_x = 0 \qquad (***)$$

(the connection holds for n spatial dimensions). This allows one to relate a subalgebra of symmetries of (***) to a subalgebra of symmetries of (*). Moreover, even symmetries of (***) which are not symmetries of (*) can be used to determine complete integrals of the latter, or vice-versa.

1. THE INFINITESIMAL SYMMETRIES OF THE HAMILTON-JACOBI EQUATION

Consider an n dimensional manifold M with local coordinates[17] x^i and an arbitrary first order differential equation on M

$$G(x^i, u_{x^i}, u) = 0. \qquad (1.1a)$$

We wish to determine the infinitesimal symmetries of such an equation which depend on all the variables present. To do this we consider the cotangent bundle $T^*(M)$ over M with local coordinates (x^i, p_i), and construct the product manifold $T^*(M)\times R$. Now $T^*(M)$ has a canonical 1-form $p_i\,dx^i$ which provides the contact 1-form

$$\alpha = du - p_i\,dx^i$$

on $T^*(M)\times R$. Solutions of (1.1a) will be surfaces in $T^*(M)\times R$

$$G(x^i, p_i, u) = 0,$$

which also annul the 1-form α. Now, following Cartan,[18] we construct the closed ideal (closed refers to exterior differentiation) I defined by

$$G(x^i, p_i, u), \qquad (1.1b)$$

$$\alpha = du - p_i\,dx^i, \qquad (1.1c)$$

$$dG = G_{x^i}\,dx^i + G_{p_i}\,dp_i + G_u\,du, \qquad (1.1d)$$

$$d\alpha = dx^i \wedge dp_i. \qquad (1.1e)$$

The surfaces in $T^*(M)\times R$ which annul I will be the solutions of the differential equation (1.1a). Stated more precisely we look for immersed submanifolds whose pullback annuls I.

Now the symmetries of the differential equation (1.1a) will be those local C^2 diffeomorphisms on $T^*(M)\times R$ whose pullback maps I into I. Stated infinitesimally this reads[12,19]

$$\underset{X}{\mathcal{L}} G = \xi G, \qquad (1.2a)$$

$$\underset{X}{\mathcal{L}} \alpha = \lambda\alpha + \eta\, dG + (A_i\,dx^i + B^i\,dp_i)G, \qquad (1.2b)$$

where $\underset{X}{\mathcal{L}}$ denotes the Lie derivative with respect to the vector field X, and $\xi, \lambda, \eta, A_i, B^i$ are functions on $T^*(R^n)\times R$, where ξ, A_i, B^i must be nonsingular in a neighborhood of $G=0$ but are other wise arbitrary. We have replaced M by the Euclidean manifold R^n. It should be mentioned here that the commutivity of the exterior derivative and the Lie derivative guarantee that dG and $d\alpha$ are back in I when an infinitesimal transformation is applied, and so Eqs. (1.2) suffice to define the symmetry condition for all of I. Notice that the Lie algebra $\mathcal{G}$ of symmetries is more general than just contact transformations since it is not necessary that the contact 1-form α be preserved. The contact transformations which are symmetries of (1.1a) form a Lie subalgebra $\mathcal{G}_c \subset \mathcal{G}$ given by the special case $\eta = A_i = B^i = 0$. To determine the Lie algebra $\mathcal{G}$, we use the expressions[12] valid for a 0-form f and any form ω

$$\underset{X}{\pounds} f = X \lrcorner\, df,$$
$$\underset{X}{\pounds} \omega = d(X \lrcorner\, \omega) + X \lrcorner\, d\omega, \tag{1.3}$$

where $\lrcorner$ denotes the natural inner product between vector fields and exterior differential forms. Applying (1.3) to (1.2a) and (1.2b) and defining the function on $T^*(R^n)\times R$, $F = X \lrcorner\, \alpha$, we equate coefficients of the independent 1-forms in (1.2b) to obtain

$$X^{x^i} = -F_{p_i} + \eta G_{p_i} + B^i G, \tag{1.4a}$$

$$X^{p_i} = F_{x^i} + p_i F_u - \eta(G_{x^i} + p_i G_u) - A_i G, \tag{1.4b}$$

$$X^u = F + p_i X^{x^i} = F - p_i F_{p_i} + \eta p_i G_{p_i} + B^i p_i G, \tag{1.4c}$$

and from (1.2a) we find

$$X \lrcorner\, dG = G_{x^i} X^{x^i} + G_{p_i} X^{p_i} + G_u X^u = \xi G, \tag{1.4d}$$

where the superscripts on the vector field X denote its component, i.e.,

$$X = X^{x^i}\partial_{x^i} + X^{p_i}\partial_{p_i} + X^u \partial_u. \tag{1.4e}$$

Now, inserting (1.4a)–(1.4c) into (1.4d), we obtain a linear first-order partial-differential equation for the function F which immediately yields the system

$$\frac{dx^i}{d\tau} = G_{p_i}, \quad \frac{du}{d\tau} = p_i G_{p_i}, \quad \frac{dp_i}{d\tau} = -(G_{x^i} + p_i G_u), \tag{1.5a}$$

$$\frac{dF}{d\tau} = (\xi - G_{x^i} B^i + G_{p_i} A_i - p_i B^i G_u) G + G_u F. \tag{1.5b}$$

We recognize that Eqs. (1.5a) describe nothing more than the *characteristic system*[11,12] of Eq. (1.1a). Thus the function F has two parts; one determined by Eqs. (1.5b), plus an arbitrary function which depends only on the characteristic curves of (1.1a).

Now the characteristic vector fields $\mathcal{J}$ in $\mathcal{G}$ are those which satisfy $X \lrcorner\, \omega \in I$ for all ω in I. By using the identity

$$\underset{X}{\pounds}(Y \lrcorner\, \omega) = [X, Y] \lrcorner\, \omega + Y \lrcorner\, \underset{X}{\pounds}\omega, \tag{1.6}$$

it is easy to show[12] that $\mathcal{J}$ is in fact an ideal in $\mathcal{G}$ for any ideal of forms I.

However, in some sense the terms in Eqs. (1.4) proportional to G are trivial, e.g., A_i and B^i, since if we restrict the vector fields to the surface in $T^*(R^n)\times R$ defined by (1.1a), these parts vanish. Indeed we can consider all vector fields in (1.4) which satisfy

$$Y \lrcorner\, \alpha = EG, \quad Y \lrcorner\, d\alpha = \beta G, \tag{1.7}$$

where E and β are arbitrary 0- and 1-forms, respectively, on $T^*(R^n)\times R$ which are nonsingular near $G=0$. Clearly all such vector fields are characteristic. Moreover, by using (1.3), (1.4d), and (1.6b), it is not difficult to show that they form an ideal $\tilde{\mathcal{J}}$ in $\mathcal{G}$. Thus it is often convenient to consider the factor algebra $\mathcal{G}/\tilde{\mathcal{J}}$. We can always choose A_i and B^i such that the term multiplying G in (1.5b) vanishes in which case we have

$$\frac{dF}{d\tau} = G_u F. \tag{1.8}$$

In general when $G_u = 0$, (1.1a) takes the standard Hamilton–Jacobi form (0.1) and the symmetries are determined by an arbitrary function of the characteristic strips. In this case the first two of Eqs. (1.5a) are just Hamilton's equation of classical mechanics. For example, for the free particle in Euclidean space, the function F takes the form

$$F = F(x^i - 2p_i t, S - 2p^2 t - qt, p_i, q). \tag{1.9a}$$

The point transformation symmetries are locally isomorphic to $\mathrm{O}(n+2,2)$ as shown in Ref. 14. Now $\mathcal{G}/\tilde{\mathcal{J}}$ admits a Lie algebra semidirect sum

$$\mathcal{G}/\tilde{\mathcal{J}} = \mathcal{G}_C/\tilde{\mathcal{J}} \,\oplus\!\!\!\!\!\supset\, \mathcal{J}/\tilde{\mathcal{J}} \tag{1.9b}$$

where $\mathcal{G}_C/\tilde{\mathcal{J}}$ is generated by the contact symmetries given by the function F which satisfies the characteristic system (1.5a) and (1.8), and $\mathcal{J}/\tilde{\mathcal{J}}$ describes the characteristics given by the function η.

For the remainder of this section we will discuss only point transformation symmetries $\mathcal{G}_p \subset \mathcal{G}_C/\tilde{\mathcal{J}}$ for (*). To find them from (1.4), we set $A^i = B_i = \eta = 0$ and impose the condition

$$X^{x^i}_{p_j} = 0,$$

i.e., the transformations on the base space are independent of p_j. Doing this explicitly for the case when $G=0$ is given by (*) and using (1.9c) will determine the point transformation symmetries of (*). From this analysis one can find that the vector fields span the finite dimensional Lie algebra $\mathrm{o}(3,2)$. [In n space and 1 time dimensions, $\mathrm{o}(n+2,2)$.] However, to understand better the appearance of the Lie algebra $\mathrm{o}(3,2)$ of the conformal group, we introduced in Ref. 14 the graph $W(t,x,S)=0$ of solutions of (*). Then upon computing the derivatives $W_x + W_S S_x = W_t + W_S S_t = 0$ and introducing the Minkowski variables

$$x^0 = 2^{-1/2}(t+2S), \quad x^2 = 2^{-1/2}(t-2S), \quad x^1 = x, \tag{1.10a}$$

we find that W satisfies

$$(W_{x^0})^2 - (W_{x^1})^2 - (W_{x^2})^2 = 0. \tag{1.10b}$$

Thus the point transformation symmetries of (*) are precisely the conformal transformations of the cone (1.10b).

This global approach[20] has distinct advantages over the infinitesimal method: (i) Without much work we have reduced the problem to known results; (ii) the geometry elucidates the meaning of the symmetries; (iii) we obtain certain symmetries which are not connected to the identity component of the group and thus are not obtainable through infinitesimal methods. However, it should also be mentioned that in general it is not always so easy to find such a nice geometrical situation in which case infinitesimal methods provide the most straightforward approach.

The symmetries of a cone in a pseudo-Euclidean space of three dimensions with signature $(+,-,-)$ form the conformal group $C^{1,2}$ which is a certain factor group of the pseudo-orthogonal group $\mathrm{O}(3,2)$. More precisely we can consider the group $\mathrm{O}(3,2)$ as a group of transformations in a five-dimensional pseudo-Euclidean space with signature $(+,-,-,-,+)$ which leaves the quadratic form $\eta_0^2-\eta_1^2-\eta_2^2-\eta_3^2+\eta_4^2$ invariant. We now consider the 5-cone $\eta_\alpha\eta^\alpha=0$ and define homogeneous coordinates

$$x^\mu=\eta^\mu/(\eta^3+\eta^4), \tag{1.11}$$

where $\mu=0,1,2$. The linear action of O(3, 2) on the 5-cone given by

$$\eta'_a=\Lambda_a{}^b\eta_b,$$

with $\Lambda_a{}^b\in O(3,2)$, then induces through (1.11) a nonlinear action on the Minkowski space $M=\{x^u\}$, which we will give shortly. However, it is seen that the action of O(3, 2) on M is not effective. Indeed, there are two members of O(3, 2) which act as the identity transformation on M, namely the subgroup $Z_2=\{\Lambda\in O(3,2)$: $\Lambda\eta=\pm\eta\}$. Hence, the conformal group $C^{1,2}\sim O(3,2)/Z_2$.

Now the group O(3, 2) consists of four components, where the component connected to the identity is $SO_0(3,2)\equiv\{\Lambda\in O(3,2)$: $\det\Lambda=1$, $\Lambda_0{}^0\Lambda_4{}^4-\Lambda_0{}^4\Lambda_4{}^0>1\}$. The other three components are obtained by reversing the signs of $\det\Lambda$ and $\Lambda_0{}^0\Lambda_4{}^4-\Lambda_0{}^4\Lambda_4{}^0$. Notice that $SO_0(3,2)\subset C^{1,2}$. The whole O(3, 2) can be obtained by extending $SO_0(3,2)$ by two discrete operations P-parity and T-covariant time reversal given by

$$P=\{x^0\to x^0, x^1\to -x^1, x^2\to x^2\},$$

$$T=\{x^0\to -x^0, x^1\to x^1, x^2\to x\},$$

respectively. In terms of (t,x,S), we have

$$P=\{t\to t, x\to -x, S\to S\}, \tag{1.12a}$$

$$T=\{t\to -2S, x\to x, S\to -\tfrac{1}{2}t\}. \tag{1.12b}$$

We will also be interested in a discrete symmetry R obtained by combining P with a certain member of $SO_0(2,1)\subset SO_0(3,2)$, namely

$$R=\{t\to S, x\to x, S\to t\}. \tag{1.12c}$$

Finally we mention the well-known inversion symmetry

$$I=\{(t,x,S)\to(t,x,S)/(4tS-x^2)\}. \tag{1.12d}$$

It is emphasized that the symmetries (1.12b)–(1.12d) are nontrivial symmetries of the Hamilton–Jacobi equation (*). Indeed, (1.12b) and (1.12c) imply that, given a solution $S(x,t)$ of (*), we can use the implicit function theorem and solve for $t=t(x,S)$, which again satisfies

$$t_x^2+t_S=0,$$

i.e., is another solution of (*).

We now give the group transformations of $SO_0(3,2)$ in terms of the original Hamilton–Jacobi variables (t,x,S):

(1) *O(2, 1) transformations*:

$$x'=\Lambda^1{}_1x+\frac{\Lambda^1{}_0+\Lambda^1{}_2}{\sqrt{2}}\,t+\sqrt{2}(\Lambda^1{}_0-\Lambda^1{}_2)S,$$

$$t'=\frac{(\Lambda^0{}_1+\Lambda^2{}_1)}{\sqrt{2}}x+\frac{(\Lambda^0{}_0+\Lambda^0{}_2+\Lambda^2{}_0+\Lambda^2{}_2)}{2}t+(\Lambda^0{}_0+\Lambda^2{}_0-\Lambda^0{}_2-\Lambda^2{}_2)S,$$

$$S'=\frac{(\Lambda^0{}_1-\Lambda^2{}_1)}{2\sqrt{2}}x+\frac{(\Lambda^0{}_0+\Lambda^0{}_2-\Lambda^2{}_0-\Lambda^2{}_2)}{4}t+\frac{(\Lambda^0{}_0-\Lambda^0{}_2-\Lambda^2{}_0+\Lambda^2{}_2)}{2}S, \tag{1.13a}$$

where $\Lambda^i{}_j\in O(2,1)$, $i,j=0,1,2$.

(2) *Translations*:

$$x'=x+a,\quad t'=t+\tau,\quad S'+S+\sigma, \tag{1.13b}$$

with $a,\ \tau,\sigma\in R$.

(3) *Dilatations*:

$$x'=\rho x,\quad t'=\rho t,\quad S'=\rho S, \tag{1.13c}$$

with $\rho>0$.

(4) *Special conformal transformations*:

$$\begin{aligned}x'&=\sigma^{-1}(x,t,S)[x+C_1(x^2-4tS)],\\ t'&=\sigma^{-1}(x,t,S)[t+C_+(x^2-4tS)],\\ S'&=\sigma^{-1}(x,t,S)[S+C_-(x^2-4tS)],\end{aligned} \tag{1.13d}$$

where

$$\sigma(x,t,S)=1-2C_+t-4C_-S+2C_1x+(C_+C_--C_1^2)(4tS-x^2)$$

and $C_1, C_\pm\in R$. It is mentioned that the special conformal transformations can be generated by a translation, an inversion, and another translation.

Now the group action (1.13) is really only a local group since the points where $\sigma(x,t,S)$ vanishes map finite points to infinity. Nevertheless, a global Lie group can be defined if we consider the "cone" compactification of R^3, making the manifold homeomorphic with the sphere S^3. Although this is necessary for a global Lie group, for our purposes it is more convenient to work with the local coordinates (x,t,S), keeping in mind that under finite group transformations singularities can occur. Hence, what we are really dealing with is a finite pseudogroup. Although the study of such singularities is of interest, we will not consider them further here. We only mention that Sard's theorem[13] guarantees that they form a set of measure zero.

In what follows we will be interested in two different formulations of the Lie algebra o(3,2). The first is the covariant formulation with a basis given by M_{ab} with $a,b=0,\ldots,4$, which satisfy the Lie brackets

$$[M_{ab},M_{cd}]=g_{ad}M_{bc}+g_{bc}M_{ad}-g_{ac}M_{bd}-g_{bd}M_{ac}. \tag{1.14}$$

On the η-space realization used previously the M_{ab} can be realized as $\eta_a\partial_b-\eta_b\partial_a$. However, on R^3 it is more convenient to consider the realization[14] [these are the point transformations of 1.4 for (*) projected onto R^3]

$$\begin{aligned}&X_1=\partial_t,\quad X_2=t\partial_t+\tfrac{1}{2}x\partial_x,\quad X_3=t^2\partial_t+tx\partial_x+\tfrac{1}{4}x^2\partial_S,\\ &X_4=\partial_x,\quad X_5=t\partial_x+\tfrac{1}{2}x\partial_S,\quad X_6=\partial_S,\\ &X_7=\tfrac{1}{2}x\partial_x+S\partial_S,\\ &X_8=\tfrac{1}{2}xt\partial_t+(tS+\tfrac{1}{4}x^2)\partial_x+\tfrac{1}{2}xS\partial_S,\\ &X_9=\tfrac{1}{4}x^2\partial_t+Sx\partial_x+S^2\partial_S,\quad X_{10}=\tfrac{1}{2}x\partial_t+S\partial_x.\end{aligned} \tag{1.15}$$

It is not difficult to see that the generators $X_1,\ldots,X_7$ form a subalgebra of o(3,2). In fact this subalgebra is maximal and generates the subgroup of $so_0(3,2)$ which leaves a lightlike two-plane invariant. It has the structure $\mathrm{gl}(2,R) \ni w$, i.e., the general linear algebra with the Heisenberg–Weyl subalgebra as an ideal. However, we will be more interested in the subalgebra formed by the generators $X_1, \ldots, X_6$ whose structure is $s_1 \sim \mathrm{sl}(2,R) \ni w$. This algebra generates a group known as the Schrödinger group $\mathcal{S}_1$ since it is the group which leaves invariant the Schrodinger equation for a free particle in one space and one time dimension.[14,21,22] The existence of the Schrödinger group $\mathcal{S}_1$ as a subgroup of $\mathrm{O}(3,2)$, or more generally[14,23] $\mathcal{S}_n \subset \mathrm{O}(n+2,2)$, emphasizes the close connection between the Schrödinger and heat equations on one hand and the Hamilton–Jacobi equation on the other. The subgroup $\mathcal{S}_1$ will play an important role in what follows. It is also seen that the discrete symmetry R given by (1.8c) provides us with another Schrödinger subgroup $\mathcal{S}_1$, conjugate to $\mathcal{S}_1$, through the mappings $X_1 \leftrightarrow X_6$, $X_2 \leftrightarrow X_7$, $X_3 \leftrightarrow X_9$, $X_4 \leftrightarrow X_4$, $X_5 \leftrightarrow X_{10}$, $X_8 \leftrightarrow X_8$.

Rather than write down the commutation relations explicitly for the generators (1.15), it is more convenient to express them in terms of generators M_{ab} satisfying (1.14), viz.,

$$\begin{aligned}
M_{10} &= -(1/\sqrt{2})(X_5 + 2X_{10}),\\
M_{20} &= X_2 - X_7,\\
M_{30} &= (1/2\sqrt{2})(X_1 + \tfrac{1}{2}X_6 - 2X_3 - 4X_9),\\
M_{21} &= -(1/\sqrt{2})(X_5 - 2X_{10}),\\
M_{31} &= \tfrac{1}{2}(X_4 + 4X_8),\\
M_{32} &= (1/2\sqrt{2})(X_1 - \tfrac{1}{2}X_6 + 2X_3 - 4X_9),\\
M_{40} &= (1/2\sqrt{2})(X_1 + \tfrac{1}{2}X_6 + 2X_3 + 4X_9),\\
M_{41} &= \tfrac{1}{2}(X_4 - 4X_8),\\
M_{42} &= (1/2\sqrt{2})(X_1 - \tfrac{1}{2}X_6 - 2X_3 + 4X_9),\\
M_{43} &= X_2 + X_7.
\end{aligned} \tag{1.16}$$

Now we are interested in the orbit structure of $\mathrm{O}(3,2)$ under the adjoint action of the conformal group $C^{1,2}$. In fact this problem has been solved in several places[4,24,25]; however, in none of these are the results in a form particularly suited for our needs. As will be seen in the next section, for the purpose of separation of variables the subgroup $\mathcal{S}_1$ plays a distinguished role. Therefore, we want to pick orbit representatives which are members of the Lie algebra s_1 of $\mathcal{S}_1$ if possible. The procedure we use to do this is to notice that every member of $\mathrm{O}(3,2)$ stabilizes a timelike, spacelike, or lightlike vector. Of course, specific elements may stabilize more than one type of vector. We then study each case separately by looking at the adjoint action of the stability subgroup and picking orbit representatives in s_1 when possible. When this is done, we must then check for conjugacy under the full $C^{1,2}$ group, again picking member of s_1 when possible. In this way we obtain a complete set of orbit representatives emphasizing which are conjugate to members of s_1 and which are not.

We begin by classifying the orbits of s_1. Now in the case of the linear Schrödinger equation treated in Refs. 2 and 3, the orbits of the factor algebra of s_1 by the central element X_6 were considered. The reason for this is that for all linear equations it is convenient to think in terms of diagonalizing operators and from this point of view X_6 is irrelevant. However, in the case of nonlinear equations one cannot always diagonalize operators in this sense. Instead, we can construct relative invariants,[26] i.e., if the infinitesimal generator X is a symmetry of the differential equation (1.1a), we can construct the graph $f(x^i, u_{x^i}, u) = 0$ of a solution u which satisfies $X \lrcorner\, df = Xf = 0$. In the special case when (1.1a) is a linear equation, this is equivalent to diagonalization of operators in the factor algebra as long as we consider general orbit representatives which include the central operators. The case at hand should illustrate the point. Thus we are interested in classifying orbits in s_1 under three particular groups: (i) the Galilei group G_1 extended by dilatations, $D \otimes G_1$; (ii) the Schrödinger group $\mathcal{S}_1$; (iii) the full conformal group $C^{1,2}$. The first group $D \otimes G_1$ is of particular interest since this is the geometrical group closely associated with the separation of variables. That is, two coordinate system which differ by dilatations of (x,t), or by Galilei transformations, essentially look the same. In Refs. 2, 3 there are some inconsistencies concerning this point. Conjugacy under $\mathcal{S}_1$ and $C^{1,2}$ are of interest for obvious reasons.

The orbits of s_1 under $D \otimes G_1$ are

$$\begin{gathered}
X_1 \pm X_6,\quad X_2 + aX_6,\quad X_3 \pm X_6,\quad X_1 + X_3 + aX_6,\\
X_1 - X_3 + aX_6,\quad X_1 \pm X_5,\quad X_3 \pm X_4,\\
X_1,\quad X_3,\quad X_4,\quad X_5,\quad X_6,
\end{gathered} \tag{1.17}$$

where $-\infty < a < \infty$. We will discuss the connection of these orbits with the separation of variables of (*) in the next section.

Under $\mathcal{S}_1$ we gain the type of equivalences discussed in Refs. 2, 3, viz.,

$$\begin{gathered}
X_1 \pm X_6,\quad X_2 + aX_6,\quad X_1 + X_3 + aX_6,\\
X_1 + X_5,\quad X_1,\quad X_4,\quad X_6,
\end{gathered} \tag{1.18}$$

where again $-\infty < a < \infty$. In both the above cases $a = 0$ is a degenerate orbit.

Under the full conformal group $C^{1,2}$ the orbits of s_1 become

$$\begin{gathered}
X_1 \pm X_6,\quad X_2 + X_6,\quad X_1 + X_3 \pm X_6,\quad X_1 + X_5,\\
X_1 + X_3,\quad X_2,\quad X_1.
\end{gathered} \tag{1.19}$$

Thus under $C^{1,2}$ we can dilate a to ± 1 using X_7, and interestingly enough we find that X_4 is on the same orbit as $X_1 - X_6$ through a rotation generated by $X_5 - X_{10}$. Again the last three entries in (1.15) correspond to degenerate orbits.

Now we wish to clasify the orbit structure of $\mathrm{o}(3,2)$ under the conformal group. As mentioned previously we first classify the one-parameter subalgebras of the stability subgroups and then later take into account conjugacy under the full $C^{1,2}$.

A. Timelike

We take the vector $(1,0,0,0,0)$ for which the stability subgroup is $\mathrm{O}(3,1)$ generated by the rotations $\{M_{21}, M_{31},$

$M_{32}\}$ and the boosts $\{M_{41}, M_{42}, M_{43}\}$. The one-parameter subalgebras are well known,[27,28] and, using (1.16), we have the orbits

$$M_{21} \sim (X_5 - 2X_{10}), \quad M_{43} = X_2 + X_7,$$
$$M_{32} + M_{42} \sim (X_1 - \tfrac{1}{2}X_6), \tag{1.20}$$
$$M_{21} + aM_{43} \sim (X_5 - 2X_{10}) + a(X_2 + X_7),$$

where here $0 < a < \infty$. Our conjugacy is under $O(3,1)$ and not just the connected component $SO_0(3,1)$.

B. Spacelike

We choose the vector $(0,1,0,0,0)$ for which the stability subgroup is $O(2,2)$ generated by $\{M_{20}, M_{30}, M_{40}, M_{32}, M_{42}, M_{43}\}$. Here it is convenient to employ the well-known Lie algebra isomorphism $o(2,2) \sim o(2,1) \oplus o(2,1)$, where $\oplus$ is a Lie algebra direct sum. Explicitly, we construct

$$J_3^{\pm} = \tfrac{1}{2}(M_{40} \pm M_{32}),$$
$$K_1^{\pm} = \tfrac{1}{2}(M_{43} \pm M_{20}), \tag{1.21a}$$
$$K_2^{\pm} = \tfrac{1}{2}(M_{30} \pm M_{42}),$$

which can be seen to generate a commuting pair of $O(2,1)$ algebras which satisfy

$$[J_3, K_2] = K_2, \quad [J_3, K_2] = -K_1, \quad [K_1, K_2] = -J_3. \tag{1.21b}$$

To find the orbits of this $o(2,1)^+ \oplus o(2,1)^-$ under $O(2,2)$, we first notice that the $o(2,1)^-$ is conjugate to $o(2,1)^+$ by a discrete transformation in $O(2,2)$ [explicitly in terms of our model this is the transformation R given by (1.8c) combined with certain dilatations in $SO_0(2,2)$]. Thus we have the usual one-parameter subalgebras of $o(2,1)^+$. Then we must find the nontrivial extensions of these orbits by the orbits of $o(2,1)^-$. This is done by the method of the Goursat twist as discussed for example in Ref. 28. Finally one checks for conjugacy of the extensions under $O(2,2)$. Accordingly, we find the orbits

$$J_3^+ + aJ_3^- \sim X_1 + X_3 + a(X_6 + X_9),$$
$$-1 < a < 1, \quad a \neq 0,$$
$$J_3^+ + aK_1^- \sim X_1 + X_3 + aX_7, \quad 0 < a < \infty,$$
$$J_3^+ \pm (J_3^- + K_2^-) \sim X_1 + X_3 \pm X_6,$$
$$K_1^+ + (J_3^- + K_2^-) \sim X_2 + X_6, \tag{1.22}$$
$$J_3^+ + K_2^+ \pm (J_3^- + K_2^-) \sim X_1 \pm X_6,$$
$$K_1^+ + aK_1^- \sim X_2 + aX_7, \quad -1 \leq a \leq 1, \quad a \neq 0,$$
$$J_3^+ \sim X_1 + X_3, \quad K_1^+ \sim X_2, \quad J_3^+ + K_2^+ \sim X_1.$$

In arriving at (1.22) we have taken full advantage of the dilatations in $O(2,2)$ generated by X_2 and X_7 to remove some of the annoying constants which multiply the various X's in the expression in (1.16).

C. Lightlike

We choose the vector $(0,0,0,1,1)$ for which the stability subgroup $D \otimes E(2,1)$ is generated by the $o(2,1)$ subalgebra $\{M_{21}, M_{10}, M_{20}\}$, the translations $\{M_{31} + M_{41}, M_{32} + M_{42}, M_{30} + M_{40}\}$, and the dilatation M_{43}. Again we use a modified Goursat twist[28] method to find the nontrivial extensions of the $o(2,1)$ subalgebra (modified since the ideal is solvable rather than Abelian). In order to simplify the notation, we introduce $J_3 = M_{21}$, $K_1 = M_{20}$, $K_2 = M_{10}$ which satisfy (1.17b), for the Abelian subalgebra $P_1 = M_{31} + M_{41}$, $P_2 = M_{32} + M_{42}$, $P_3 = M_{30} + M_{40}$, which transform as the designated components of an $O(2,1)$ vector, and $D = M_{43}$ which commutes with $o(2,1)$ and satisfies $[D, P_i] = -P_i$. We thus obtain the following orbit representatives:

$$J_3 + aD \sim X_5 - X_{10} + a(X_2 + X_7), \quad 0 < a < \infty,$$
$$J_3 + P_3 \sim X_1 + X_5 - X_{10},$$
$$K_1 + bD \sim X_2 + aX_7, \quad -1 \leq a \leq 1,$$
$$K_1 + P_2 \sim X_2 - X_7, \tag{1.23}$$
$$J_3 + K_2 + D \sim X_2 + X_5 + X_7,$$
$$J_3 + K_2 + P_2 \sim X_1 + X_5,$$
$$P_2 \sim X_1 - X_6, \quad P_3 \sim X_1 + X_6,$$
$$P_2 + P_3 \sim X_1, \quad D \sim X_2 + X_7.$$

Again we have made use of the dilatations in $D \otimes E(2,1)$ to simplify the operators in terms of the X's.

Now in order to obtain all orbits of $o(3,2)$, we only have to check the above results for conjugacy under the full $C^{1,2}$. Since we have already done this for the s_1 subalgebra, we can restrict our attention to the remaining cases. Indeed for the timelike case we can use dilatations to adjust some of the constants appearing in (1.16), and we see that all of the orbits (1.20) also appear as orbits in the other two cases. In fact, there are no further simplifications due to conjugacy other than identifying those orbits which appear in both cases. We have collected our results in Table I, indicating in which of the three cases the various orbits appear as well as which are members of the Schrödinger subalgebra s_1 as well as its maximal proper extension $gl(2,R) \otimes w$ in $o(3,2)$.

2. SEPARATION OF VARIABLES

For the purpose of separating variables in (*) it is more convenient to use the equivalent homogeneous equation

$$W_x^2 + WW_t = 0 \tag{2.1}$$

obtained from (*) by the substitution $S = \ln W$. We are in general interested in R-separability, that is, we look for a transformation of coordinates

$$x = F(v_1, v_2), \quad t = G(v_1, v_2), \tag{2.2}$$

$v_1, v_2 \in R$, where F and G are once differentiable real functions, such that the solution of (2.1) takes the form

$$W = \exp[Q(v_1, v_2)]A(v_1)B(v_2), \tag{2.3}$$

where Q can not be written as the sum of functions of the single variables unless it vanishes. It is clear that a solution of (2.1) of the form (2.3) implies a solution of (*) of the form

$$S = Q(v_1, v_2) + \ln A(v_1) + \ln B(v_2).$$

We proceed by considering the cases $Q = 0$ and $Q \neq 0$ separately. First, it is convenient to introduce a notion of equivalence. Two coordinates will be said to be equivalent if they can be related by a member of the group

TABLE I. Orbits in o(3, 2) classified under $C^{1,2}$. t, s, l denote respectively timelike, spacelike, and lightlike.

Orbit		Representative	Type	Remarks
$\mathrm{gl}(2,R)\oplus w$	s_1	$X_1+\epsilon X_6$	$t(\epsilon=-1), s, l$	$\epsilon=\pm 1, 0$
		X_2+X_6	s	
		X_2	s, l	
		X_1+X_5	l	
		$X_1+X_3+\epsilon X_6$	s	$\epsilon=\pm 1, 0$
		X_2+aX_7	$t(a=1), s, l$	$-1\le a\le 1$
		$X_1+X_3+aX_7$	s	$0<a<\infty$
		$X_2+X_5+X_7$	l	
		$X_1+X_5-X_{10}$	l	0
		$X_5-X_{10}+a(X_2+X_7)$	t, l	$-0\le a<\infty$
		$X_1+X_3+a(X_6+X_9)$	s	$-1\le a\le 1,\ a\ne 0$

$D\otimes G_1$ discussed previously. We also consider any two systems to be equivalent if they differ by a constant multiple, i.e., $(v_1, v_2)\sim(v_1', v_2')$ if $(v_1', v_2')=\alpha(v_1, v_2)$, α constant.

A. Pure separability, $Q = 0$

Rewriting (2.1) in terms of the coordinates v_1 and v_2, we obtain

$$a_{11}W_1^2+a_{12}W_1W_2+a_{22}W_2^2+a_1WW_1+a_2WW_2=0, \tag{2.4}$$

where $a_{11}=(G_2/D)^2$, $a_{12}=-2G_1G_2/D^2$, $a_{22}=(G_1/D)$, $a_1=-F_2/D$, $a_2=F_1/D$, $D=F_1G_2-F_2G_1$, and the subscripts on W, G, F indicate differentiation with the respective variable. The conditions for separability can be further subdivided into two cases:

(i) $a_{12}\ne 0$: This is only possible if W is an exponential in one variable, say v_2, and the coefficients depend only on the remaining variable v_1. Upon redefining the variable v_1 this gives rise to coordinates of the form $t=v_2+h(v_1)$, $x=v_1$, where h is an arbitrary function of v_1. These coordinates describe nonorthogonal coordinate axes and always give rise to exponential solutions. We will not consider these any further in this article.

(ii) $a_{12}=0$: Without loss of generality we can take $G_1=0$ and hence $t=v_2$. By multiplying (2.4) by F_1^2 we can take the coefficients as $a_{11}=2$, $a_1=-F_1F_2$, and $a_2=F_1^2$. The conditions for separability are then

$$F_1^2=f(v_1)g(v_2),\quad F_1F_2=h(v_1), \tag{2.5}$$

with f, g, and h arbitrary functions of their respective variables. By redefining the variable v_1, the conditions (2.5) imply

$$F=v_1p(v_2)+q(v_2),\quad pp_2=\alpha,\quad pq_2=\beta,$$

where α and β are constants. Without loss of generality we can put $q=0$, and we find two cases:

(1) $\alpha=0$, $x=v_1$, $t=v_2$,

(2) $\alpha\ne 0$, $x=v_1v_2^{1/2}$, $t=v_2$.

B. R-separability, $Q\neq 0$

We now wish to classify all coordinate systems for which (2.1) admits solutions of the form (2.3) for nontrivial real function Q. The appearance of the Q will give rise to a factor a_0W^2 added to Eq. (2.4). We now only consider the case $a_{12}=0$ and we obtain, preceeding as before, the nonzero coefficients

$$a_{11}=1,\quad a_1=2Q_1-F_1F_2,\quad a_2=F_1^2,$$
$$a_0=Q_1^2+F_1(F_1Q_2-F_2Q_1).$$

The condition for separability then gives

$$F_1^2=f(v_1)g(v_2),\quad 2Q_1-F_1F_2=h(v_1),$$
$$Q_1^2+F_1(F_1Q_2-F_2Q_1)=f(v_1)g(v_2)+p(v_1), \tag{2.6}$$

where again f, g, h, p, q are arbitrary functions of their denoted variables. By suitably redefining the variable v_1, we have from the first of Eqs. (2.6)

$$F=v_1u(v_2)+W(v_2)$$

and from the second

$$Q=\tfrac{1}{4}v_1^2uu_2+\tfrac{1}{2}v_1uW_2.$$

Then from the third equation in (2.6), a straightforward computation yields

$$u^3u_{22}=A,$$
$$u^3W_{22}=B, \tag{2.7}$$

where A and B are constants. Now we can integrate the first of these equations to give $u=(av_2^2+b)^{1/2}$. We consider the following cases:

(1) $a=0$: We can take $u=1$. Then by using equivalence under space translations, Galilei transformations, and dilatations, the coordinates can be brought to the form

$$x=v_1\pm v_2^2,\quad v_2=t\quad\text{with}\quad Q=\pm v_1v_2.$$

(2) $b=0$: We may take $u=v_2$ and similarly bring the coordinates to one of the forms

$$x=v_1v_2\pm 1/v_2,\quad t=v_2\quad\text{with}\ Q=\tfrac{1}{4}v_1^2v_2v_1/2v_2,$$
$$x=v_1v_2,\qquad t=v_2,\quad Q=\tfrac{1}{4}v_1^2v_2.$$

(3) $a/b>0$, $a, b\ne 0$: Using dilatation, we can take $u=(v_2^2+1)^{1/2}$. Again using Galilei and space translation, we find

$$x=v_1(v_2^2+1)^{1/2},\quad t=v_2,\quad Q=\tfrac{1}{4}v_1^2v_2.$$

(4) a, $b\ne 0$, $a/b<0$: Similarly we find

$$x=v_1|1-v_2^2|^{1/2},\quad t=v_2,\quad Q=(\epsilon/4)v_1^2v_2,$$

where $\epsilon=\mathrm{sgn}\,(1-v_2^2)$. Thus we have shown that up to equivalence under the group $D\otimes G_1$, there are precisely seven coordinate systems such that (2.1) and hence the Hamilton–Jacobi equation (*) is separable. Moreover, these coordinates coincide with the separable coordinate system[2] for the Schrödinger equation $U_{xx}+iU_t=0$ and the heat equation $U_{xx}+U_t=0$. The list of separable coordinates is presented in Table II, where equivalences under the full Schrödinger group is also noted. It is also mentioned here that the separation of variables for (*) also implies the equivalence of the four types of potentials; i.e., free particle, linear potential, and attractive and repulsive harmonic oscillator. Indeed it is not difficult to give explicitly the transformations which map the time dependent Hamilton–Jacobi equation with a linear potential, attractive, or repulsive harmonic oscillator potential onto (*). Thus it follows also that their local symmetry groups of point transformations are all isomorphic to O(3, 2). A closer connection will be seen explicitly in the next section.

3. SIMILARITY SOLUTIONS

In this section we give a systematic treatment of similarity solutions of (*) by giving the solution which corresponds to each of the orbit representatives in Table I. We can then say that any similarity solution obtainable from point transformations must be related to one of our representative solutions by at most a transformation in $C^{1,2}$. Moreover, we will show how the orbits of the subalgebra s_1 relate to the method of separation of variables presented in Sec. 2, or more specifically that to each system of separable coordinates (ξ, τ), there corresponds an orbit representative of s_1 such that the similarity variable is ξ and the similarity solution is the solution obtained by the separation of variables of (*). In this way we will obtain complete integrals of (*). Any arbitrary parameter which has been transformed away by our orbit analysis can, of course, always be reinstated. As is well known,[11] then, the general solution can always be obtained by forming the envelope of any complete integral. It seems likely that all known explicit complete integrals of (*) can be obtained by group theoretical methods.

More generally let $f(x^i, u) = 0$ be the graph of a solution u of (1.1a) and $X \in \mathcal{G}_p$; then f is called a *relative invariant* with respect to X if

$$\underset{X}{\pounds} f = X \lrcorner\, df = Xf = 0. \tag{3.1}$$

For every such f which satisfies (3.1), we can solve implicitly for u which when combined with the original differential equation (1.1a) reduces (1.1) to a differential equation with one less variable. Any solution u obtained in this way is called a *similarity solution.*[16] It is clear in general that in order to specify a unique solution for an equation in n independent variables, we must demand that f be a relative invariant for $n-1$ members X_α of g, $\alpha = 1, \ldots, n-1$. The X_α's need not commute, but owing to (3.1) they must form a subalgebra of g. Thus, the problem of finding complete similarity solutions relates to the problem of classifying all subalgebras of a given Lie algebra.[28] The preceeding discussion of similarity solutions has a simple geometric interpretation. We restrict ourselves here to R^3. Indeed (3.1) says that for any vector field X we construct surfaces in R^3 such that X lies in its tangent plane at each point. The tangent planes to all integral surfaces at a point intersect along X, i.e., X defines the characteristics of (3.1). If in addition X is a symmetry of a differential equation as given by (1.1a), $\underset{X}{\pounds}$ describes the infinitesimal dragging of the tangent plane to an integral surface of the equation (1.1a) in such a way that the tangent plane lines up with the tangent plane of another solution. For a general first order equation the possible tangent planes form a one-parameter family which envelops the Monge cone at a given point. Now, choosing a tangent plane defined by a generator of the Monge cone and X, we are guaranteed that, by moving along the curve generated by X, there will be a generator of the Monge cone which lies in the tangent plane at each point. In this way we describe an integral surface which satisfies both (1.1a) and (3.1). There are two qualifications to be made: First X cannot be collinear to the generator of the Monge cone; second X must not imply a relationship between the independent variables for (1.1a).

Now in the practical computation of relative invariants one uses the characteristic equations of a given vector field, viz.,

$$X = \xi^i(\mathbf{x}, u)\partial_{x^i} + \eta(\mathbf{x}, u)\partial_u, \tag{3.2}$$

then $u(\mathbf{x})$ can be obtained by solving

$$\frac{dx^1}{\xi^1(\mathbf{x}, u)} = \cdots = \frac{dx^n}{\xi^n(\mathbf{x}, u)} = \frac{du}{\eta(\mathbf{x}, u)}. \tag{3.3}$$

In our case any $X \in \mathrm{o}(3,2)$ takes the form

$$X = a(x, t)\partial_t + b(x, t, S)\partial_x + c(x, S)\partial_S, \tag{3.4}$$

where the coefficients can be read off from (1.11). The characteristic equations for (3.4) are then

$$\frac{dt}{a(x, t)} = \frac{dx}{b(x, t, S)} = \frac{dS}{c(x, S)}. \tag{3.5}$$

Solving any two of the Eqs. (3.5) when combined with (*) will then give the similarity solution corresponding to the vector field (3.4).

We now proceed to discuss the similarity solutions for the subgroup $\mathcal{S}_1$ and their relation to the separation of variables of the previous section. For the s_1 subalgebra we see from (1.11) that both b and c are independent

TABLE II. Separable coordinates (*) classified under $D \otimes G_1$. Subgroupings indicate equivalence under $\mathcal{S}_1$.

Coordinates	Multiplier	Operator	Remarks	
$x = v_1, \quad t = v_2$	$Q = 0$	$X_1 + \epsilon X_6$	$\epsilon = \pm 1,\ 0$	
$x = v_1 v_2, \quad t = v_2$	$Q = \frac{1}{4} v_1^2 v_2$	$X_3 + \epsilon X_6$	"	
$x = v_1 + \epsilon v_2^2, \quad t = v_2$	$Q = \epsilon v_1 v_2$	$X_1 + \epsilon X_5$	"	
$x = v_1 v_2 + \epsilon / v_2, \quad t = v_2$	$Q = v_1^2 v_2/4 - \epsilon v_1/2v_2$	$X_3 + \epsilon X_4$	"	
$x = v_1 v_2^{1/2}, \quad t = v_2$	$Q = 0$	$2X_2 + aX_6$		$-\infty < a < \infty$
$x = v_1 \lvert 1 - v_2^2 \rvert^{1/2}, \quad t = v_2$	$Q = \frac{1}{4}\epsilon v_1^2 v_2$	$X_1 - X_3 + aX_6$	$\epsilon = \mathrm{sgn}(1 - v_2^2)$	"
$x = v_1 \lvert 1 + v_2^2 \rvert^{1/2}, \quad t = v_2$	$Q = v_1^2 v_2/4$	$X_1 + X_3 + aX_6$		"

of S (transformations which act linearly on $w=e^S$), and a is independent of x. Thus we can integrate the first two of Eqs. (3.5) to give the similarity variable $\xi=\xi(x,t)$. Then expressing x as a function of ξ and t, we have

$$dS=\frac{C(x(\tau,t))}{a(x)}dt. \tag{3.6}$$

Integrating along the characteristic ξ, we obtain S as

$$S=\int_{\xi=\text{const}}\frac{C(x(\xi,t))}{a(t)}dt+F(\xi). \tag{3.7}$$

Substituting (3.7) back into (*) yields a first order ordinary differential equation for F which can then be integrated to give the explicit similarity solution.

As mentioned previously, it is the geometric subgroup $D\otimes\mathcal{G}_1$, which is relevant for the separation of variables; therefore, we consider the orbit representatives given by (1.13) for the similarity solutions. We will see that for each orbit in (1.13) the similarity variable ξ will correspond precisely to the variable v_1 for one of the separable coordinate systems listed in Table II, although there are degenerate cases. The separation constant corresponds to the parameter a in (1.13), i.e., to the one-parameter extensions by the central element X_6. In some cases the separation constant can be transformed to ± 1 or 0 by a member of $D\otimes\mathcal{G}_1$ which alters only slightly the functional form of the solution. We also group together those orbits (1.14) and separable systems which are inequivalent under the Schrödinger group $\mathcal{S}_1$. As in Refs. 2, 3, these systems are denoted by the appelations, harmonic oscillator, repulsive harmonic oscillator, free particle, and linear potential, since they reduce (*) to the time-independent Hamilton—Jacobi equation with the corresponding type of potential. Within this grouping we label by 1 and 2 coordinates which are equivalent under $\mathcal{S}_1$ but inequivalent the subgroup $D\otimes\mathcal{G}_1$ since they appear differently from a geometric point of view. We will give the details for the first case only.

A. Harmonic oscillator

The separable coordinates are

$$\xi=v_1=x/(1+t^2)^{1/2},\quad \tau=v_2=t. \tag{3.8}$$

Substituting these into (*) and using the ansatz

$$S=\tfrac{1}{4}\xi^2\tau+F(\xi)+G(\tau), \tag{3.9}$$

we obtain

$$F_\xi^2+\tfrac{1}{4}\xi^2+(1+\tau^2)G_\tau=0. \tag{3.10}$$

Separation implies

$$(1+\tau^2)G_\tau=a, \tag{3.11a}$$

which reduces (3.9) to the time-independent Hamilton—Jacobi equation with a harmonic oscillator potential

$$F_\xi^2+\tfrac{1}{4}\xi^2+a=0. \tag{3.11b}$$

Integrating Eqs. (3.11) and placing into (3.9), we find

$$S=\tfrac{1}{4}\xi^2\tau+a\tan^{-1}\tau-a\sin^{-1}(\xi/2\sqrt{-a})+\tfrac{1}{2}\sqrt{-a}\,\xi(1+\xi^2/4a)^{1/2}. \tag{3.12}$$

From the point of view of similarity solutions it is easy to see that the coordinates (3.8) correspond to the orbit $X_1+X_3+aX_6$ of (1.13) for which Eq. (3.5) is

$$\frac{dt}{1+t^2}=\frac{dx}{tx}=\frac{dS}{\frac{1}{4}x^2+a}. \tag{3.13}$$

The first two of these equations give precisely the variable ξ of (3.8), while the first and third [or what amounts to (3.7)] gives, integrating along the characteristic ξ, (3.9) with $G=a\tan^{-1}\tau$. Then substituting (3.9) back into (*) gives (3.11b) and hence the similarity solution (3.12). We point out that the case $a=0$ is degenerate.

B. Repulsive harmonic oscillator

(1) The separable coordinates are

$$\xi=v_1=x/|t|^{1/2},\quad \tau=v_2=t, \tag{3.14}$$

which correspond to the orbit $2X_2+aX_6$ whose subsidiary conditions are

$$\frac{dt}{2t}=\frac{dx}{x}=\frac{dS}{a}. \tag{3.15}$$

Integrating (3.15) gives

$$S=\tfrac{1}{2}a\ln\tau+F(\xi), \tag{3.16a}$$

which upon substituting into (*) gives

$$F_\xi^{\,2}-\tfrac{1}{2}\xi F_\xi+a=0, \tag{3.16b}$$

yielding the solutions

$$\begin{aligned}S&=\tfrac{1}{2}a\ln\tau+\tfrac{1}{8}\xi^2-a\cosh^{-1}\left(\frac{\xi}{\sqrt{8a}}\right)+\left(\frac{a}{8}\right)^{1/2}\xi\left(\frac{\xi^2}{8a}-1\right)^{1/2},\quad \tau>0\\ S&=\tfrac{1}{2}a\ln\tau-\tfrac{1}{8}\xi^2-a\sinh^{-1}\left(\frac{\xi}{\sqrt{8a}}\right)+\left(\frac{a}{8}\right)^{1/2}\xi\left(\frac{\xi^2}{8a}+1\right)^{1/2},\quad \tau<0.\end{aligned} \tag{3.17}$$

Again the case $a=0$ is degenerate.

(2) The separable coordinates are

$$\xi=v_1=x/|t^2-1|^{1/2},\quad \tau=v_2=t, \tag{3.18}$$

corresponding to the orbit $X_1-X_3+aX_6$ in (1.13) whose equations are

$$\frac{dt}{t^2-1}=\frac{dx}{tx}=\frac{dS}{\frac{1}{4}x^2-a}. \tag{3.19}$$

Integrating, we obtain

$$\begin{aligned}S&=\tfrac{1}{4}\xi^2\tau+a\coth^{-1}\tau+F(\xi),\qquad \tau^2>1,\\ S&=\tfrac{1}{4}\xi^2\tau+a\tanh^{-1}\tau+F(\xi),\qquad \tau^2<1,\end{aligned} \tag{3.20a}$$

where $F(\xi)$ satisfies

$$F_\xi^{\,2}-\tfrac{1}{4}\xi^2-\operatorname{sgn}(\tau^2-1)a=0, \tag{3.20b}$$

leading to the solutions

$$S=\tfrac{1}{4}\xi^2\tau+a\coth^{-1}\tau+a\sinh^{-1}\left(\frac{\xi}{2\sqrt{a}}\right)+\tfrac{1}{2}\sqrt{a}\,\xi\left(\frac{\xi^2}{4a}+1\right)^{1/2},\quad \tau^2>1,$$

$$S=-\tfrac{1}{4}\xi^2\tau+a\tanh^{-1}\tau-a\cosh^{-1}\left(\frac{\xi}{2\sqrt{a}}\right)+\tfrac{1}{2}\sqrt{a}\,\xi\left(\frac{\xi^2}{4a}-1\right)^{1/2},\quad \tau^2<1. \tag{3.21}$$

As mentioned previously cases (1) and (2) are related by a transformation in S_1. The transformation which takes (3.21) into (3.17) is given by

$$t'=\frac{1+t}{1-t},\quad x'=\frac{2^{1/2}x}{(1-t)},\quad S'=S+\frac{x^2}{4(1-t)}. \tag{3.22}$$

It is also mentioned that Eq. (3.16b) can be cast into the form of a repulsive harmonic oscillator by replacing F by $F\pm\xi^2/8$. Again in both cases (1) and (2), $a=0$ is degenerate.

C. Free particle

(1) The separable coordinates are

$$\xi=v_1=x/t,\quad \tau=v_2=t, \tag{3.23a}$$

corresponding to the orbit $X_3+\epsilon X_6$ in (1.13).

The subsidiary conditions (3.5) are

$$\frac{dt}{t^2}=\frac{dx}{tx}=\frac{dS}{\frac{1}{4}x^2+\epsilon}, \tag{3.23b}$$

giving rise to

$$S=\tfrac{1}{4}\xi^2\tau-\epsilon/\tau+F(\xi), \tag{3.24a}$$

where

$$F_\xi^{\,2}+\epsilon=0. \tag{3.24b}$$

Thus we have the solution

$$S=\tfrac{1}{4}\xi^2\tau-\epsilon/\tau\pm\sqrt{-\epsilon}\,\xi. \tag{3.25}$$

(2) The coordinates are simply the usual Cartesian ones $\xi=x$, $\tau=t$, corresponding to the orbit representative $X_1+\epsilon X_6$ whose equations are

$$\frac{dt}{1}=\frac{dx}{0}=\frac{dS}{\epsilon}, \tag{3.26}$$

giving rise to

$$S=\epsilon t+F(x) \tag{3.27a}$$

with

$$F_x^{\,2}+\epsilon=0. \tag{3.27b}$$

Hence, the similarity solution is simply

$$S=\epsilon t\pm\sqrt{-\epsilon}\,x+c. \tag{3.28}$$

Here we allow $\epsilon=0$ as well as $\epsilon=\pm1$ so as to include the degenerate orbits X_3 and X_1.

D. Linear potential

(1) The separable coordinates are

$$\xi=v_1=\frac{x}{t}+\frac{\epsilon}{2t^2},\quad \tau=v_2=t, \tag{3.29a}$$

corresponding to the orbit $X_3+\epsilon X_4$ with the subsidiary conditions

$$\frac{dt}{t^2}=\frac{dx}{tx+\epsilon}=\frac{dS}{\frac{1}{4}x^2}, \tag{3.29b}$$

which gives rise to

$$S=\frac{\xi^2\tau}{4}+\frac{\epsilon\xi}{4\tau}-\frac{\epsilon^2}{48\tau^3}+F(\xi) \tag{3.30a}$$

with

$$F_\xi^{\,2}-\tfrac{1}{2}\epsilon\xi=0. \tag{3.30b}$$

Integrating (3.41b), we find the solution

$$S=\frac{\xi^2\tau}{4}+\frac{\epsilon\xi}{4\tau}-\frac{\epsilon^2}{48\tau^3}+\tfrac{1}{3}\sqrt{2\epsilon}\,\xi^{3/2}. \tag{3.31}$$

(2) The separable coordinates are

$$\xi=v_1=x-\tfrac{1}{2}\epsilon t^2,\quad \tau=v_2=t, \tag{3.32}$$

corresponding to $X_1+\epsilon X_5$ with the equations

$$\frac{dt}{1}=\frac{dx}{\epsilon t}=\frac{dS}{\frac{1}{2}\epsilon x}. \tag{3.33}$$

Integrating, we find

$$S=\frac{\epsilon\xi\tau}{2}+\frac{\epsilon^2\tau^3}{12}+F(\xi), \tag{3.34a}$$

with

$$F_\xi^{\,2}+\tfrac{1}{2}\epsilon\xi=0, \tag{3.34b}$$

giving rise to the solution

$$S=\tfrac{1}{2}\epsilon\xi\tau+\tfrac{1}{12}\epsilon^2\tau^3+\tfrac{1}{3}\sqrt{-2\epsilon}\,\xi^{3/2}. \tag{3.35}$$

Again we allow $\epsilon=0$ as well as ±1 in order to include the degenerate cases. The group transformation which takes (3.35) to (3.31) and (3.28) to (3.25) is

$$t'=-1/t,\quad x'=x/t,\quad S'=S-x^2/4t. \tag{3.36}$$

It can be seen that this is the square of the transformation (3.22).

The remaining orbits in (1.13) and (1.14) are degenerate in the sense that they give rise to special cases. X_4 gives the usual cartesian separation and the special solution $S=\text{const}$, where as X_5 which is equivalent to X_4 under S_1, gives the degenerate solution $a=0$ in (3.17). A relative invariant of X_6 violates the condition that x and t be independent (in involution). However, we should notice that it does not violate the independence of x and z in (**) and thus gives rise to a nontrivial solution. It is interesting that under the full conformal group these cases are equaivalent to those already discussed. In fact under $C^{1,2}$ we have only the four types given by the potentials and their degenerate cases as noted in (1.15). We can always set the separation constant equal to ±1 or 0.

As mentioned previously the subalgebra of o(3,2) generated by $X_1,\ldots,X_7$ is maximal and contains s_1. Moreover, its structure is $\mathrm{gl}(2,R)\otimes w_1$, but now X_6 is not in the center. However, we notice from (1.11) that for this subalgebra the coefficient b given by (3.4) still has no S dependence; hence, we should obtain a similarity variable $\xi(x,t)$ upon integrating the first two of Eqs. (3.4). Indeed this suggests that there may be some type of separation of variables not considered in Sec. 2 which lead to these solutions. We will now show that this is indeed the case. We will only consider orbits inequivalent under the full conformal group $C^{1,2}$; however, we expect that again classifying the subalgebra $gl(2,R) \oplus w$, under its subgroup $D \otimes G_1$, will lead to a more geometric picture compatible with the separation of variables. From Table I we pick out the following orbit representatives of $gl(2,R) \otimes w_1$ which are not in s_1:

(1) $X_2 + aX_7$, $-1 \le a \le 1$, $a \ne 0$: The subsidiary equations are

$$\frac{dt}{t} = \frac{dx}{\frac{1}{2}(a+1)x} = \frac{dS}{aS}. \tag{3.37}$$

The similarity variable is

$$\xi = t^{-(a+1)/2}x, \tag{3.38a}$$

giving the form

$$S = \xi^a F(\xi). \tag{3.38b}$$

Plugging (3.48b) back into (*), we find that $F(\xi)$ satisfies

$$F_\xi{}^2 - \tfrac{1}{2}(a+1)\xi F_\xi + aF = 0. \tag{3.38c}$$

Thus we see that we have the separation of (*) in the form of a product instead of a sum. If we look into the separation process in some detail, we will see that the conditions for separation involve a coupling between the coordinate functions (2.2) and the separable solution in the variable $v_2 = t$. For this reason this type of separation is much more complicated and usually not considered for equations of the kind of (*). However, here we are led to these naturally by considering similarity solutions. Now Eq. (3.38c) is a special case of Chrystal's equation[29] whose solution is given implicitly by

$$F = \left[\frac{(a+1)^2}{4} - u^2\right]\frac{\xi^2}{4a},$$
$$\xi[u \mp (a+1)/2]^{(a+1)/2}[u \mp (a-1)/2]^{(1-a)/2} = C \tag{3.39a}$$

with $a \ne \pm 1$ and C a constant. For the degenerate cases $a = \pm 1$ we have the regular solutions

$$F = \tfrac{1}{2}(\xi + C)^2, \quad a = -1, \tag{3.39b}$$

$$F = -\tfrac{1}{4}C^2 \mp \tfrac{1}{2}C\xi, \quad a = 1, \tag{3.39c}$$

and, for $a = 1$, the singular solution[29]

$$F = \tfrac{1}{4}\xi^2 + C, \quad a = 1. \tag{3.39d}$$

(2) $X_1 + X_3 + aX_7$, $0 < a < \infty$: The Pfaffian equations are

$$\frac{dt}{1+t^2} = \frac{dx}{tx + \frac{1}{2}ax} = \frac{dS}{\frac{1}{4}x^2 + aS}, \tag{3.40}$$

which upon integrating the first two of these equations gives the similarity variable

$$\xi = \frac{x}{(1+t^2)^{1/2}}\left(\frac{1+it}{1-it}\right)^{ia/4}, \tag{3.41a}$$

while the first and third gives

$$S = \frac{\xi^2\tau}{4}\left(\frac{1+i\tau}{1-i\tau}\right)^{-ia/2} + \left(\frac{1+i\tau}{1-i\tau}\right)^{-ia/2} F(\xi), \tag{3.41b}$$

where F satisfies

$$F_\xi{}^2 - \tfrac{1}{2}a\xi F_\xi + aF + \tfrac{1}{4}\xi^2 = 0. \tag{3.42c}$$

This equation has the form of the general Chrystal's equation.[29] Its general solution is given by

$$\xi^2(u - 2i \mp \tfrac{1}{2}a)^{1 \pm a/2i}(u + 2i \mp \tfrac{1}{2}a)^{-1 \pm a/2i} = C, \tag{3.43a}$$

where C is an arbitrary constant and

$$F = (\xi^2/4a)(\tfrac{1}{4}a^2 - 1 - u^2). \tag{3.43b}$$

We mention that (3.42c) has a singular solution which we ignor since it occurs when a is pure imaginary.

(3) $X_2 + X_5 + X_7$: The subsidiary equations are

$$\frac{dt}{t} = \frac{dx}{t+x} = \frac{dS}{S + \frac{1}{2}x}, \tag{3.44}$$

which yields the similarity variable

$$\xi = x/t - \ln t \tag{3.45a}$$

and the form

$$S = \tfrac{1}{2}\xi\tau\ln\tau + \tfrac{1}{4}\tau\ln^2\tau + \tau F(\xi), \tag{3.45b}$$

where F satisfies

$$F_\xi{}^2 - (\xi+1)F_\xi + \tfrac{1}{2}\xi + F = 0. \tag{3.45c}$$

The general solution of this equation is given by

$$(\pm u - 1)\exp(\pm u - 1) = C^{e^\xi}, \tag{3.46a}$$

where C is an arbitrary constant and

$$F = \tfrac{1}{4}(1 + \xi^2 - u^2). \tag{3.46b}$$

Thus it is seen that the remaining two cases [(2) and (3) above] separate in the product form with an additional multiplier term $Q(\xi, \tau)$ present.

There now remains from Table I only three cases of orbit representatives of $o(3,2)$ which are not in $gl(2,R) \otimes w_1$. Of these the first two to be considered are in fact closer related to (**).

(4) $X_5 - X_{10} + X_1$: The Pfaffian subsidiary equations are

$$\frac{dt}{1 - \frac{1}{2}x} = \frac{dx}{t - S} = \frac{dS}{\frac{1}{2}x} \tag{3.47a}$$

or alternatively in terms of $z = t - S$, $T = t + S$, we have

$$\frac{dz}{1-x} = \frac{dx}{z} = \frac{dT}{1} \tag{3.47b}$$

from which we find the similarity variable

$$\xi^2 = z^2 + (x-1)^2 \tag{3.48a}$$

and the solution

$$T = \sin^{-1}[(x-1)/\xi] + F(\xi), \tag{3.48b}$$

where F satisfies

$$\xi^2 F_\xi^2 + 1 - \xi^2 = 0, \tag{3.48c}$$

giving rise to the general solution

$$T = \sin^{-1}[(x-1)/\xi] + \xi^2 - 1 - \tan^{-1}\xi^2 - 1 + C. \tag{3.49}$$

Clearly this case is related to the separation of (**) in polar coordinates.

(5) $X_5 - X_{10} + a(X_2 + X_7)$: The subsidiary equations are

$$\frac{dt}{at - \frac{1}{2}x} = \frac{dx}{t - S + ax} = \frac{dS}{\frac{1}{2}x + aS} \tag{3.50a}$$

or in terms of z and T

$$\frac{dz}{az-x}=\frac{dx}{z+ax}=\frac{dT}{aT}. \tag{3.50b}$$

From the first two equations the similarity variable is

$$\xi=(x^2+z^2)^{1/2}\left(\frac{z-ix}{z+ix}\right)^{a/2i} \tag{3.51a}$$

while the second two equations give the form

$$T=\left(\frac{z+ix}{z-ix}\right)^{a/2i}F(\xi), \tag{3.51b}$$

where $F(\xi)$ satisfies

$$(a^2+1)\xi^2F_\xi{}^2-2a^2\xi F_\xi F+a^2F^2-\xi^2=0. \tag{3.51c}$$

The general solution of this equation is given implicitly by

$$\frac{a}{a^2+1}\operatorname{sgn}\xi=\left(\frac{aF+i(a^2+1)\xi^2-a^2F^2}{aF-i(a^2+1)\xi^2-a^2F^2}\right)^{\pm a/2i} \times\left|\frac{F}{F\pm(a^2+1)\xi^2-a^2F^2}\right|. \tag{3.52a}$$

The case $a=0$ is degenerate and leads to

$$T=\pm x^2+z^2+C, \tag{3.52b}$$

which in terms of S gives a certain translation in S and t of the fundamental solution $x^2/4t$.

(6) $X_1+X_3+a(X_6+X_9)$: The Pfaffian equations are

$$\frac{dt}{1+t^2+\frac{1}{4}ax^2}=\frac{dx}{x(t+aS)}=\frac{dS}{a(1+S^2)+\frac{1}{4}x^2}. \tag{3.53}$$

We have not been able to find a simple way to integrate these equations explicitly. This ends the list of similarity solutions for (*). We mention also that it would be interesting to see if there is any relation (perhaps of a projective nature) between the solutions presented here and the semisubgroup separation of variables for the graph equation (1.10b) and hence the wave equation in 3-space.[4]

Before ending this section we briefly comment on one other solution generated by a symmetry, namely the general solution generated by the characteristcis. However, since (*) is not quasilinear, this solution cannot be written as a similarity solution. The characteristics for any first order equation are determined from the Eqs. (1.5a) or equivalently from the characteristic vector fields (1.7). The relative invariant[30] obtained from the vector fields in $\mathcal{J}/\tilde{\mathcal{J}}$ given by

$$Y=\eta(x,t,S,p,q)(2p\partial_x+\partial_t+p^2\partial_S) \tag{3.54a}$$

is determined by the equations

$$\frac{dx}{2p}=\frac{dt}{1}=\frac{dS}{p^2}=\frac{dp}{0}=\frac{dq}{0}, \tag{3.54b}$$

giving rise to the general solution of (*) in terms of the characteristic strips[11]

$$S=p^2t+F(x-2pt,p), \tag{3.54c}$$

where F is an arbitrary function of its arguments. Indeed the above analysis can be made much simpler if we use the characteristic $\xi=x-2pt$ as an underlying variable. We can consider (*) to be generated by the ideal

$$r=p^2+q=0,\quad dr=0, \tag{3.55a}$$

$$dx\wedge dp+dt\wedge dq=0. \tag{3.55b}$$

Then clearly $dx=d\xi+2p\,dt+2t\,dp$, so (3.55b) becomes

$$d\xi\wedge dp+dt\wedge dr=0, \tag{3.55c}$$

which implies the existence of a function $V(\xi,t)$ with $p=V_\xi$ and $r=V_t$. Then (3.55a) implies that it is independent of t, and thus the general solution is given by

$$p=V_\xi(\xi), \tag{3.55d}$$

which is equivalent to (3.54c) as long as $d\xi\wedge dt\neq 0$. In fact, it can easily be seen that $V(\xi)$ is equal to F in (3.54c), modulo an additive constant. In the next section we will see that (3.55d) is closely related to prolongations of (*).

4. PROLONGATIONS

The concept of prolongation was first introduced by Cartan[18,31] in his study of what has since been called infinite pseudogroups. His idea was to obtain and classify certain pseudogroups (infinite groups in Cartan's language) by taking successively higher derivatives of Lie's differential equations for finite Lie groups. Indeed a classification of certain types of pseudogroups has by now been rigorously established, using essentially this idea.[13,32] However, here we wish only to apply the first prolongation of (*), that is we take the derivative with respect to x of (*) and notice that it gives precisely (***). The question that is raised is then what is the connection between the symmetries of (*) and (***)? We do not intend to give here a full analysis of this question but only to point out some interesting relationships.

Since (***) is a first order quasilinear partial differential equation, the analysis performed in the beginning of Sec. 1 applies. We are only interested in the point transformation symmetries of (***) since only these can be projected to symmetries of $R^2\times R^1$ with local coordinates (x,t,p). Then, using (1.4) and (1.5), we find the pseudogroup of point transformations of (***) to be generated by the vector fields (projections onto $R^2\times R^1$)

$$\begin{aligned}X=&\,2pF^1(x,t,p)\partial_x-2pF^0(x-2pt,p)\partial_p\\&+\Big(F^1(x,t,p)+\frac{x+2pt}{2p}F^0(x-2pt,p)\\&+g(x-2pt,p)\Big)\partial_t,\end{aligned} \tag{4.1}$$

where F^0, F^1, g are arbitrary function of their arguments. It is easy to see that the ideal I of characteristic vector fields of (***) is generated by $F^1(x,t,p)$.

We now look for those members of the symmetry algebra $\mathcal{G}_*$ of * given by (1.4) and (1.5b) which can be related to a subalgebra of (4.1) whose vector fields when prolonged[33] to act on the variables S and q can be identified with a subalgebra of $\mathcal{G}_*$. This prolongation can be accomplished through the use of (1.4c) and (1.4d) and give precisely those vector fields in $\mathcal{G}_*$ for which X^x, X^t, and X^p are independent of S and q. A straightforward computation gives constraints on the vector fields (4.1) which imply the existence of a function $H(x-2pt,p)$

such that

$$F^0=\frac{1}{2p}H_x=\frac{1}{2p'}H_\xi,$$
$$g=\frac{1}{2p}\left(H_{p'}+\frac{\xi}{2p'}H_\xi+c\xi\right), \tag{4.2}$$

where c is a constant and we use the change of variables $p'=p$, $\xi=x-2pt$. The prolongation to the S and q components of the vector fields now proceeds via (1.4c) and (1.4d) respectively. These prolonged vector fields can be written

$$\begin{aligned}X^x&=2p\widetilde{F}^1(x,t,p)+cx,\\ X^t&=\widetilde{F}^1-(1/2p)H_p+ct,\\ X^p&=-H_x,\\ X^q&=-2pH_x+E^0(x,t,S,p,q)(p^2+q)\\ X^S&=p^2\widetilde{F}^1(x,t,p)+\tfrac{1}{2}pH_p-H+cS\\ &\quad+E^1(x,t,S,p,q)(p^2+q),\end{aligned} \tag{4.3}$$

where $\widetilde{F}^1$ is an arbitrary function of its arguments and E^1 and E^0 are arbitrary except for being nonsingular, at $p^2+q=0$. Again as in Sec. 1 it is convenient to factor these terms out and use (4.3) modulo E^0 and E^1. We now consider some explicit examples.

The first example to be considered is the characteristic collineation given by the arbitrary functions F^1 in (4.1). For (***) this gives rise to the general solution

$$p=f(x-2pt). \tag{4.4}$$

Now the prolonged vector fields given by $\widetilde{F}^1$ in (4.3) will generate the general solution of (*) given by (3.54c) or (3.55d). In fact we can easily identify f in (4.4) with $V_\xi=V_x$ in (3.55d).

As another example we consider those point transformation symmetries of (***) which can be prolonged to point transformation symmetries of (*) or vice-versa. These can be found by simply demanding the condition that the x, t, and S components of the vector fields in (4.3) be independent of p and q. Through a straightforward calculation we arrive at a finite-dimensional subalgebra spanned by the vector fields[34]

$$\begin{aligned}&Y_1=\partial_t,\quad Y_4=\partial_x,\\ &Y_2=t\partial_t+\tfrac{1}{2}x\partial_x-\tfrac{1}{2}p\partial_p,\quad Y_5=t\partial_x+\tfrac{1}{2}\partial_p,\\ &Y_3=t^2\partial_t+tx\partial_x+(\tfrac{1}{2}x-tp)\partial_p,\\ &Y_7=\tfrac{1}{2}x\partial_x+\tfrac{1}{2}p\partial_p.\end{aligned} \tag{4.5}$$

We have used a notation suggested by (1.15); the prolongation of (4.5) by adding the q and S components via (4.3) gives precisely the corresponding Y's in (1.15). Conversely, we can obtain the above vector fields from the corresponding ones in (1.15) [the subalgebra $\mathrm{gl}(2,R)\otimes w$] by lifting the latter to $T^*(R^2)\times R^1$ and projecting onto a surface with S and q constant. We notice that Y_6 is missing from (4.5) since X_6 projects to the identity for constant S, i.e., $Y_6=0$. The structure of the generators (4.5) is $\mathrm{gl}(2,R)\oplus a_2$, where a_2 is a two-dimensional Abelian ideal generated by Y_4 and Y_5. Hence the prolongation process *does not* conserve Lie brackets. However, X_6 generates a one-dimensional ideal of $\mathrm{gl}(2,R)\otimes w$, and thus there is a Lie algebra isomorphism between the factor algebra $[\mathrm{gl}(2,R)\otimes w]/\{X_6\}$ in (1.15) and $\mathrm{gl}(2,R)\otimes a_2$ given by (4.5). Of course, the subalgebra $s_1\sim\mathrm{sl}(2,R)\otimes w$ of $\mathcal{G}_*$ obtained by removing X_7 is a central extension of the subalgebra $\mathrm{sl}(2,R)\otimes a_2$ of $\mathcal{G}_{***}$ obtained by removing Y_7 from (4.5).

Now there is an interesting connection between the similarity solutions of (4.5) and the corresponding ones for (1.15) given in Sec. 3. Indeed the orbit representatives of $\mathrm{gl}(2,R)\otimes a_2$ under the adjoint action of the group are

$$\begin{aligned}&Y_2+aY_7\quad(a\geq 0),\quad Y_1+Y_3+aY_7\quad(-\infty<a>\infty)\\ &Y_1+Y_5,\quad Y_2+Y_5+Y_7,\quad Y_1+Y_7,\quad Y_1,\quad Y_7.\end{aligned} \tag{4.6}$$

Comparing (4.6) with the orbit representatives of $\mathrm{gl}(2,R)\otimes w$ in Table I and considering the factor algebra $[\mathrm{gl}(2,R)\otimes w]/\{X_6\}$, we see that the only difference is the appearance of Y_1+Y_7 and Y_7 and the ranges of a in (4.6). This is so since X_7 and X_1+X_7 are conformally equivalent to X_2 and X_2+X_6 respectively. Similarly, the differences in the ranges of a are explained by conformal equivalence. Now the connection of the corresponding similarity solutions of (*) and (***) is this: Take the x derivative of one of the similarity solutions in $\mathrm{gl}(2,R)\otimes w$ obtained in Sec. 3 and put $p=S_x$; then this solution is precisely the similarity solution obtained from the corresponding orbit representative in $\mathrm{gl}(2,R)\otimes a_2$ for (***). It should be added that the multiple of X_6 for a similarity solution of (*) becomes an integration constant for the corresponding similarity solution of (***). A simple example should illustrate the point. Consider the similarity solution for $2X_2+X_6$ given by (3.17). Considering only $t>0$, we find

$$p=S_x=\frac{1}{2t^{1/2}}\left(\frac{\xi}{2}+\sqrt{\frac{\xi^2}{4}-4a}\right),\quad \xi=\frac{x}{t^{1/2}}, \tag{4.7a}$$

which is the similarity solution of (***) obtained from $2Y_2$ with proper identification of the integration constant. Indeed

$$\frac{dt}{2t}=\frac{dx}{x}=\frac{-dp}{p} \tag{4.7b}$$

gives the similarity variable $\xi=x/t^{1/2}$ and

$$p=t^{-1/2}f(\xi). \tag{4.7c}$$

If we call $f=F_\xi$ and substitute (4.7c) into (***), we get

$$2F_\xi F_{\xi\xi}-\tfrac{1}{2}\xi F_{\xi\xi}-\tfrac{1}{2}F_\xi=0, \tag{4.7d}$$

which is precisely the x derivative of (3.16b).

More generally, we can consider the entire subalgebra $\mathcal{H}\subset\mathcal{G}_{***}$ determined by (4.2). Now looking at (4.3) we see as before that we must not only factor out E^0 and E^1 but also the constant part in the function H, i.e., the generator X_6 in (1.15). That this can be done follows readily from the form of the generators in (4.3), namely, that the only S dependence of the vector fields in (4.3), mod (E^0,E^1), is of the type $S\partial_S$. Then the prolongation process defines an isomorphism of $\mathcal{H}$ onto the subalgebra of $\mathcal{G}_*$ given by (4.3) modulo the above equivalences. We mention that one can find nontrivial similarity solutions for (***) which upon integration give solutions of (*) and that the prolonged vector field corresponding to such solutions are vector fields on $T^*(R^2)\times R$ which are not the lifts of vector fields on R^3.

Finally we make a few comments on the members of $\mathcal{G}_*$ and $\mathcal{G}_{***}$ which are not related by prolongations. For example, looking at (1.15), we seen that all the members of o(3, 2) which cannot be prolonged to members of $\mathcal{G}_{***}$ are those vector fields whose components involve the variable S. Nevertheless, they yield similarity solutions of (*) for which we can determine, in principal, S and hence $p=S$ at nonsingular points, and are guaranteed that p will satisfy (***). Conversely, from those members of $\mathcal{G}_{***}$ that cannot be prolonged to symmetries of (*), we can also determine a p through the similarity methods which upon integration with respect to x provides a solution of (*). The problem is from the group theoretical standpoint that the prolongation process discussed above no longer gives a symmetry. However, they can be interpreted as generalized symmetries since they are a symmetry of one equation and give rise to solutions of both. In this connection it would be interesting to study further the symmetries of the complete prolonged ideal of differential forms which contains both (*) and (***) and possibly any further prolongations in the spirit of Ref. 35.

ACKNOWLEDGMENT

We wish to thank I. Kupka, W. Miller Jr., and P. Winternitz for helpful discussions on this work. The second author would like to thank the Consejo Nacional de Ciencia y Tecnología (CONACYT of México for financial support and the members of the Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas for their hospitality during the period in which this work was done.

[1]For a review see W. Miller, Jr., in *Proceedings of the Advanced Seminar on Special Functions* (Academic, New York, 1975). There is also a forthcoming book; W. Miller, Jr., *Symmetry and Separation of Variables for Linear Differential Equations* (Addison-Wesley, Reading, Mass., to appear). Also for a selective choice see Refs. 2—5.

[2]E.G. Kalnins and W. Miller, Jr., J. Math. Phys. 15, 1728 (1974).

[3]C.P. Boyer, E.G. Kalnins, and W. Miller, J. Math. Phys. 16, 499 (1975).

[4]E.G. Kalnins and W. Miller, Jr., J. Math. Phys. 16, 2507 (1975).

[5]C.P. Boyer, E.G. Kalnins, and W. Miller, Jr., Nagoya Math. J. 60, 35 (1976).

[6]J. Liouville, J. de Math. 11, 345 (1846).

[7]P. Stackel, Habilitationschrift Halle, 1891; Math. Ann. 42, 537 (1893).

[8]T. Levi-Civita, Math. Ann. 59, 383 (1904); F. Dall'Acqua, Math. Ann. 66, 398 (1908); P. Burgatti, R.C. Acad. Lincei 20, 108 (1911).

[9]M.S. Iarov-Iarovoi, J. Appl. Math. Mech. 27, 1499 (1964); P. Havas, J. Math. Phys. 16, 1461 (1975); B. Carter, Commun. Math. Phys. 10, 280 (1968).

[10]H.P. Robertson, Math. Ann. 98, 749 (1927); L.P. Eisenhart, Ann. Math. 35, 284 (1934).

[11]R. Courant and D. Hilbert, *Methods of Mathematical Physics, Vol. II* (Interscience, New York, 1962).

[12]R. Herman, *Differential Geometry and the Calculus of Variations* (Academic, New York, 1968).

[13]S. Sternberg, *Lectures in Differential Geometry* (Prentice-Hall, Englewood Cliffs, N.J., 1964).

[14]C.P. Boyer and M. Peñafiel, Nuovo Cimento B 31, 195 (1976).

[15]N.M.J. Woodhouse, Commun. Math. Phys. 44, 9 (1975).

[16]G.W. Blumen and J.D. Cole, *Similarity Methods for Differential Equations* (Springer, New York, 1974): W.F. Ames, *Nonlinear Partial Differential Equations in Engineering, Vol. 2* (Academic, New York, 1972).

[17]Although the derivation of the infinitesimal symmetries of a first-order partial-differential equation seems not to have appeared explicitly before, we do not claim that it is new here. Some indications can be found in Hermann's book,[12] while the technique used here was shown to us by I. Kupka.

[18]E. Cartan, *Les systêmes differentials extêrieurs et leurs applications geometriques* (Hermann, Paris, 1945).

[19]B.K. Harrison and F.B. Estabrook, J. Math. Phys. 12, 653 (1971).

[20]Although the treatment here is global, the symmetries do not form a global Lie group as will be seen shortly.

[21]U. Niederer, Helv. Phys. Acta. 45, 802 (1972).

[22]C.P. Boyer, R.T. Sharp, and P. Winternitz, J. Math. Phys. 17, 1438 (1976).

[23]G. Burdet, M. Perrin, and P. Sorba, Commun. Math. Phys. 34, 85 (1973).

[24]H. Zassenhaus, Can. Math. Bull. 1, 31, 101, 183 (1958).

[25]G. Burdet and M. Perrin, J. Math. 16, 2172 (1975).

[26]L.P. Eisenhart, *Continuous Groups of Transformations* (Dover, New York, 1961).

[27]D. Finkelstein, Phys. Rev. 100, 924 (1955); P. Winternitz and I. Fris, Sov. J. Nucl. Phys. 1, 636 (1965).

[28]J. Patera, P. Winternitz, and H. Zassenhaus, J. Math. Phys. 16, 1597 (1975).

[29]E.L. Ince, *Ordinary Differential Equations* (Dover, New York, 1956).

[30]Of course here, we must consider a relative invariant on $T^*(R^2)\times R$.

[31]E. Cartan, *Oeuvres completes* (Gauthier-Villars, Paris, 1953), Partie 2, Vol. 2.

[32]S. Kobayashi, *Transformation Groups in Differential Geometry* (Springer, New York, 1972).

[33]The term prolongation appears in different contexts in the literature, Cartan's prolongation[18,31] means taking the total or partial derivatives in a differential ideal and add those as new variables. It is reasonable to extend this definition to include any process by which new variables can be added to an ideal of differential forms (see, e.g., Ref. 35 below). Thus we can start with an ideal of forms for (***) and by integration find a new variable S which satisfied (*). This prolongation is a kind of inverse of Cartan's prolongation. This prolongation can then induce prolongations on vector fields which are symmetries of the ideal. For example the lift of a vector field on M to one on $T^*(M)$ is a special type of prolongation.

[34]A finite subalgebra of $\mathcal{G}_{***}$ given by (4.1) which contains the vector fields (4.5) was given in G. Rosen and G.W. Ullrich, SIAM J. Appl. Math. 24, 286 (1973). Apparently, there are two vector fields here which cannot be prolonged to symmetries of (*).

[35]H.D. Wahlquist and F.B. Estabrook, J. Math. Phys. 16, 1 (1975).

On self-reciprocal functions under a class of integral transforms

Kurt Bernardo Wolf

Instituto de Investigación en Matemáticas, Aplicadas y en Sistemas (IIMAS), Universidad Nacional Autónoma de México, México D. F., Mexico

(Received 28 July 1976)

We use the fact that a rather general class of integral transforms—complex linear and radial canonical transforms—are equivalent to hyperdifferential operators, to formulate the problem of self-reciprocal functions under these transforms as an eigenvalue problem for (second-order) differential operators. We thus find the solution for Fourier, Hankel, bilateral Laplace, Bargmann, Weierstrass–Gauss and Barut–Girardello transforms. These involve the Schrödinger attractive and repulsive harmonic oscillator and/or centrifugal potential wavefunctions. A general concept of "self-reproducing" functions is introduced which includes all of the above plus linear potential wavefunctions. In particular, two new generalized bases for Bargmann's Hilbert space of analytic functions are found.

I. INTRODUCTION

A function $f(x)$ will be said to be self-reciprocal under an integral transform $\mathcal{T}_M$ [defined through integration over an interval $\mathbb{I}\subseteq\mathbb{R}$ with a kernel $A_M(x,x')$] when

$$(\mathcal{T}_M f)(x)=\int_{\mathbb{I}} dx' A_M(x,x')f(x')=\lambda f(x), \quad \lambda\in\mathbb{C}. \tag{1.1}$$

This corresponds to the eigenfunction problem for the operator $\mathcal{T}_M$. The cases we are interested in include the well-known cases of the Fourier[1] and Hankel[2] transforms, as well as the bilateral Laplace,[3] Bargmann,[4] Weierstrass–Gauss[5] (which represents the time evolution of the solutions of the heat equation), and Barut–Girardello[6] transforms. These constitute special cases of a class of integral transforms termed canonical transforms[7,8] which will be described in Sec. II.

The functions which are self-reciprocal under the Fourier transform are well known,[9] while some properties of functions self-reciprocal under Hankel transforms have been analyzed in the work of Hardy and Titchmarsh.[10] Further results on the Hankel self-reciprocal functions and the consideration of the (unilateral) Laplace transform and some of its variants has been presented in a series of papers by Indian mathematicians.[11] The solution we present to the problem (1.1) makes use of the observation that, for the class of canonical transforms and $f\in\mathcal{C}_M^\infty$ [the intersection of the space $\mathcal{C}^\infty$ and the space of functions for which the integral (1.1) exists at least in a generalized sense], one can realize $\mathcal{T}_M$ as a hyperdifferential operator

$$(\mathcal{T}_{M(\tau)}f)(x)=\exp(i\tau H^\omega)f(x), \tag{1.2}$$

where τ is a continuous parameter which for certain values yields the particular transforms mentioned above, and H^ω is a second-order differential operator, self-adjoint in $\mathcal{L}^2(\mathbb{I})$. Clearly, the solution of the eigenvalue equation

$$H^\omega\Phi_\mu^\omega(x)=\mu\Phi_\mu^\omega(x) \tag{1.3}$$

solves (1.1) with $\lambda=\exp(i\tau\mu)$.

When τ is a rational multiple of π and the spectrum $\Lambda^\omega=\{\mu\}$ of H^ω is discrete and integer-spaced (cases of Fourier and Hankel), the spectrum $\Lambda_M=\{\lambda\}$ of $\mathcal{T}_M$ will consist of a finite number N_M of values of unit modulus, and will divide the space $\mathcal{C}_M^\infty$ into subspaces $\mathcal{S}_\lambda$ of self-reciprocal functions labeled by the eigenvalue λ. The functions spanning each of these spaces will be the subsets of eigenfunctions Φ_μ^ω with $\mu\equiv\mu \bmod N_M$. When the usual closing procedure is implemented, the $\mathcal{S}_\lambda$ become Hilbert spaces. This is applied to the Fourier and Hankel transforms in Secs. III and IV. When the spectrum Λ^ω and hence Λ_M are continuous, the generalized eigenfunctions of H^ω will still provide the self-reciprocal functions of (1.1). This is the case of the bilateral Laplace and Bargmann transforms analyzed in Secs. V and VI respectively. The case of the Bargmann transform is particularly interesting since its generalized self-reciprocal functions (the repulsive quantum oscillator wavefunctions), being orthonormal (in the sense of Dirac) and complete in $\mathcal{L}^2(\mathbb{R})$ are so too, in the same sense, in the Bargmann–Hilbert space[4] $\mathcal{F}_B$ [of entire analytic functions of growth (2, 1/2)]. This generalized basis is new and adds to the known harmonic oscillator and coherent-state[4,12] bases of $\mathcal{F}_B$. The results for the Weierstrass–Gauss and Barut–Girardello transforms are sketched in Secs. VII and VIII. As for the general case of the complex linear[7] and radial[8] canonical transforms, we explore the generalized "self-reproducing" functions in Sec. IX. These include, beside the functions studied before, the Airy functions.

II. CANONICAL TRANSFORMS AND THEIR HYPERDIFFERENTIAL REALIZATION

To every complex unimodular 2×2 matrix $M=\begin{pmatrix}a&b\\c&d\end{pmatrix}$ we associate the integral transform $\mathcal{T}_M$ given as in (1.1) with the kernel

$$A_M(x,x')=(2\pi|b|)^{-1/2}\varphi_b\exp\left(\frac{i}{2b}(ax'^2-2x'x+dx^2)\right), \tag{2.1a}$$

$$\varphi_b=\exp[-(i/2)(\pi/2+\arg b)], \tag{2.1b}$$

which we call the canonical transform.[7]

When a,b,c,d are real,[13] $\mathcal{T}_M$ can be seen to be a unitary mapping from $\mathcal{L}^2(\mathbb{R})$ onto $\mathcal{L}^2(\mathbb{R})$, while if these parameters are complex, the resulting transform [when bounded: for $\mathrm{Im}(a/b)>0$ and b real if $a=0$] is a unitary mapping between $\mathcal{L}^2(\mathbb{R})$ and Hilbert spaces $\mathcal{F}_M$ of analytic functions defined through the scalar product over the complex plane, for $\bar f=\mathcal{T}_M f$ and $\bar g=\mathcal{T}_M g$,

$$(\bar f,\bar g)_M=\int_{\mathbb{C}} d\mu_M(x)\,\bar f(x)^*\,\bar g(x), \tag{2.2a}$$

$$d\mu_M(x)=2(2\pi v)^{-1/2}\exp\left[\frac{1}{2v}(ux^2-2xx^*+u^*x^{*2})\right]d\,\mathrm{Re}\,x\;d\,\mathrm{Im}\,x, \qquad (2.2b)$$

$$u=a^*d-b^*c, \quad v=2\,\mathrm{Im}\,b^*a. \qquad (2.2c)$$

The functions $\tilde{f}(x)$ in the spaces $\mathcal{J}_M$ are characterized as $\tilde{f}(x)=\exp(-ux^2/2v)\tilde{f}_B(x)$, where $\tilde{f}_B(x)$ are elements in the Bargmann–Hilbert space.[4] The transform inverse to (1.1) is given by

$$f(x)=(\mathcal{T}_M^{-1}\tilde{f})(x)=\int_{\mathbb{C}} d\mu_M(x')A_M(x',x)^*\tilde{f}(x'). \qquad (2.3)$$

For the transforms we are interested in, we can identify the following:

(Fourier) $\mathcal{J}=\exp(i\pi/4)\,\mathcal{T}_F,\quad F=\begin{pmatrix}0&1\\-1&0\end{pmatrix},$

(Bilateral Laplace) $\mathcal{L}=i\sqrt{2\pi}\,\mathcal{T}_L,\quad L=\begin{pmatrix}0&i\\i&0\end{pmatrix},\qquad (2.4b)$

(Bargmann) $\mathcal{B}=(2\pi)^{1/4}\mathcal{T}_B,\quad B=\frac{1}{\sqrt{2}}\begin{pmatrix}1&-i\\-i&0\end{pmatrix},\qquad (2.4c)$

[with the specific choice of phase $\pm i=\exp(\pm i\pi/2)$]

(Weierstrass–Gauss) $\mathcal{W}_t=\mathcal{T}_{W_t},\quad W_t=\begin{pmatrix}1&-2it\\0&1\end{pmatrix}.\qquad (2.4d)$

We will find it useful to define the geometric transform as

$$(\mathcal{T}_{G(\beta,c)}f)(x)=e^{\beta/2}\exp[\tfrac{1}{2}ie^{\beta}cx^2]f(e^{\beta}x),$$
$$G(\beta,c)=\begin{pmatrix}e^{-\beta}&0\\c&e^{\beta}\end{pmatrix}, \qquad (2.4e)$$

which can be obtained from the general case (2.1) letting[7] $b\to 0$.

These results and their derivation are found in Ref. 7, where we also analyze the behavior of the measure (2.2) when $v\to 0$. We should point out the novelty that in our treatment the Weierstrass–Gauss transform[5] becomes a unitary transformation between Hilbert spaces with a conserved scalar product and a proper inversion. One more result which we can extract from Ref. 7 is the fact that, when two transforms $\mathcal{T}_{M_1}$ and $\mathcal{T}_{M_2}$ are bounded, their composition (through integration over $\mathbb{R}$) follows the product of the matrices $M_1M_2=M_3$, so that $\mathcal{T}_{M_1}\circ\mathcal{T}_{M_2}=\varphi\mathcal{T}_{M_3}$, where φ is a phase (± 1) depending on 1–2 matrix elements of the M_i's. Notice, however, that it is not necessary that a transform $\mathcal{J}_M$ be bounded in order to have a nonvanishing domain $\mathcal{C}_M^\infty$ dense in $\mathcal{L}^2(\mathbb{I})$. This remark applies to the case of the bilateral Laplace transform, which is unbounded. Finally, we should stress that, while all (bounded) transforms are unitary mappings between $\mathcal{L}^2(\mathbb{I})$ and $\mathcal{J}_M$, when seen as mappings between $\mathcal{L}^2(\mathbb{I})$ and $\mathcal{L}^2(\mathbb{I})$, the transforms $\mathcal{T}_M$ with complex M are nonunitary.

In Ref. 7 it was shown that for $f\in\mathcal{C}_M^\infty$, the integral transforms (1.1)–(2.1) are equivalent to the action of the hyperdifferential operators (1.2). Specifically, for the one-parameter subsets which contain the transforms (2.4),

$$H^h=\tfrac{1}{2}(-\Delta+x^2)\ \text{generating}\begin{pmatrix}\cos\tau&-\sin\tau\\\sin\tau&\cos\tau\end{pmatrix}, \qquad (2.5a)$$

$$H^r=\tfrac{1}{2}(-\Delta-x^2)\ \text{generating}\begin{pmatrix}\cosh\tau&-\sinh\tau\\-\sinh\tau&\cosh\tau\end{pmatrix}, \qquad (2.5b)$$

$$H^f=-\tfrac{1}{2}\Delta\ \text{generating}\begin{pmatrix}1&-\tau\\0&1\end{pmatrix}, \qquad (2.5c)$$

where

$$\Delta=\frac{d^2}{dx^2}. \qquad (2.5d)$$

Thus, the Fourier transform can be written in hyperdifferential form (1.2) with H^h as given by (2.4a) and $\tau=-\pi/2$, the Laplace and Bargmann transforms with H^r as (2.5b) and $\tau=-i\pi/2$ and $i\pi/4$, respectively, while the Weierstrass–Gauss transform appears with H^f as (2.5c) and $\tau=2it$. This last case is commonly known.[5] Finally, the generators of geometric transforms (2.4d) are first-order differential operators which can be seen to be, for the β parameter,

$$H^d=-i\left(x\frac{d}{dx}+\frac{1}{2}\right)\ \text{generating}\begin{pmatrix}e^{-\tau}&0\\0&e^{\tau}\end{pmatrix}, \qquad (2.6a)$$

while for the γ parameter it is simply

$$\tfrac{1}{2}x^2\ \text{generating}\begin{pmatrix}1&0\\\tau&1\end{pmatrix}. \qquad (2.6b)$$

In Ref. 8 we considered the "radial part" of a special n-dimensional version of the transform (2.1). This class of integral transforms have the form (1.1) over the interval $\mathbb{I}=\mathbb{R}^+$ (the positive half-axis), while the kernel, instead of (2.1), turns out to be

$$A_M^{[k]}(x,x')=b^{-1}\exp(-ik\pi)(xx')^{1/2}\exp\left[\frac{i}{2b}(ax'^2+dx^2)\right]J_{2k-1}(xx'/b). \qquad (2.7)$$

As before, for a,b,c,d real,[14] (1.1)–(2.7) is a unitary mapping from $\mathcal{L}^2(\mathbb{R}^+)$ onto $\mathcal{L}^2(\mathbb{R}^+)$. As particular cases we have the following transforms:

(Hankel) $\mathcal{H}^{[k]}=\exp(ik\pi)\mathcal{T}_F^{[k]},\quad F=\begin{pmatrix}0&1\\-1&0\end{pmatrix},\qquad (2.8a)$

(Barut–Girardello[15]) $\mathcal{G}^{[k]}=\mathcal{T}_B^{[k]},\quad B=\frac{1}{\sqrt{2}}\begin{pmatrix}1&-i\\-i&1\end{pmatrix}.\qquad (2.8b)$

These transforms can also be put in hyperdifferential form (1.2) for functions in $\mathcal{C}_{M[k]}^\infty$. For the Hankel transform, the operator $H^{h[k]}$ has the form (2.5a) with $\tau=-\pi/2$ and the Barut–Girardello transform, (2.5b) with $\tau=i\pi/4$ but, instead of (2.6), the "radial" operator Δ is in these cases

$$\Delta^{[k]}=\frac{d^2}{dx^2}-\frac{(2k-1)^2-1/4}{x^2}. \qquad (2.9)$$

The parameters of the transform kernel (2.7), when extended to complex values, define a unitary map from $\mathcal{L}^2(\mathbb{R}^+)$ to spaces $\mathcal{J}_M^{[k]}$ with a scalar product which corresponds basically to the radial part of the kth spherical harmonic part of (2.2). Thus, the "radial part" of an n-dimensional Bargmann transform is the Barut–Girardello transform and similar "radial Weierstrass–Gauss" transforms, for example, can be constructed.

III. SELF-RECIPROCAL FUNCTIONS UNDER THE FOURIER TRANSFORM

The results presented in the last section allow us to state that the eigenfunctions of H^h in (2.5a), namely the quantum harmonic oscillator wavefunctions

$$\Phi_n^h(x) = (\pi^{1/2}2^n n!)^{-1/2}\exp(-x^2/2)H_n(x), \quad n=0,1,2,\cdots \tag{3.1}$$

will be self-reciprocal under the Fourier transform (2.4a). This is a well-known fact[9] which will clearly illustrate our method. Since the spectrum of H^h is $\mu = n+\frac{1}{2}$, $n=0,1,2,\cdots$, then

$$(\mathcal{F}\Phi_n^h)(x) = \exp(-i\pi n/2)\,\Phi_n^h(x). \tag{3.2}$$

Equation (3.2) allows us to split $\mathcal{C}_F^\infty$ [and its closure $\mathcal{L}^2(\mathbb{R})$] into four subspaces $\mathcal{S}_\lambda$, for $\lambda = 1, i, -1$, or $-i$ (recall that $\mathcal{F}^4=1$). Each of these subspaces is generated by the set Φ_n^h with $n \equiv (0,1,2$, or $3)$ mod4 respectively. Clearly, the intersection of two different $\mathcal{S}_\lambda$'s is empty, while the union of the four is $\mathcal{L}^2(\mathbb{R})$. The raising and lowering operators

$$2^{-1/2}\left[x - \frac{d}{dx}\right]\Phi_n^h(x) = (n+1)^{1/2}\Phi_{n+1}^h(x), \tag{3.3a}$$

$$2^{-1/2}\left[x + \frac{d}{dx}\right]\Phi_n^h(x) = n^{1/2}\Phi_{n-1}^h(x), \tag{3.3b}$$

are n-independent and will thus map the $\mathcal{S}_\lambda$ to $\mathcal{S}_{\lambda'}$ rotating the λ plane counterclockwise and clockwise by $\pi/2$.

IV. SELF-RECIPROCAL FUNCTIONS UNDER THE HANKEL TRANSFORM

Here we follow the general procedure as in the last section. It is well known that the normalized eigenfunctions of (2.5a) with (2.9) are

$$\Phi_n^{h(k)}(x) = \left(\frac{2n!}{\Gamma(n+2k)}\right)^{1/2}\exp(-x^2/2)\,x^{2k-1/2}L_n^{(2k-1)}(x^2),$$
$$n=0,1,2,\cdots, \quad x\in\mathbb{R}^+. \tag{4.1}$$

The "radial" Schrödinger harmonic oscillator with centrifugal potential wavefunctions, for $k \geq 1$, (4.1) is the only set of eigenfunctions, while for $0<k<1$ we have more than one self-adjoint extension of (2.5a)—(2.9) in $\mathbb{R}^+$, one of which still has the eigenfunctions (4.1). The spectrum is $\mu = 2(n+k)$ with $n=0,1,2\cdots$ and thus

$$(\mathcal{H}_k\Phi_n^{h(k)})(x) = \exp(-in\pi)\,\Phi_n^{h(k)}(x), \tag{4.2}$$

exactly as in (3.2) and splitting the space $\mathcal{S}^{(k)}$ in $\mathcal{S}_\lambda^{(k)}$ ($\lambda = 1$ or -1) generated by $\Phi_n^{h(k)}$ with $n \equiv (0$ or $1)$ mod2, respectively, as before.

The n-independent differential operators which raise and lower the index n in (4.1) will similarly rotate the λ plane. From the raising and lowering operators for the upper index of the Laguerre polynomials, we find

$$\left(x + \frac{2k-1/2}{x} - \frac{d}{dx}\right)\Phi_n^{h(k)}(x) = 2(n+2k)^{1/2}\,\Phi_n^{h(k+2)}(x), \tag{4.3a}$$

$$\left(x + \frac{2k-3/2}{x} + \frac{d}{dx}\right)\Phi_n^{h(k)}(x) = 2(n+2k+1)^{1/2}\Phi_n^{h(k-2)}(x), \tag{4.3b}$$

which will thus map $\mathcal{S}_\lambda^{(k)}$ onto $\mathcal{S}_\lambda^{(k\pm2)}$, i.e., self-reciprocal functions of the Hankel transform of index k to index $k\pm2$.

The characterization of the self-reciprocal functions under the Hankel transform with either eigenvalue λ thus seems almost trivial and certainly simpler than that presented in Refs. 10 and 11.

V. SELF-RECIPROCAL FUNCTIONS UNDER THE BILATERAL LAPLACE TRANSFORM

The bilateral Laplace transform can be realized through the hyperdifferential operator (1.2) with exponent (2.5b), the quantum repulsive oscillator Hamiltonian, and $\tau = -i\pi$, on a suitable function space. The eigenfunctions of (2.5b) are the repulsive oscillator quantum eigenfunctions

$$\Phi_\mu^{r\pm}(x) = C_\mu D_{i\mu-1/2}[\pm\sqrt{2}\exp(3i\pi/4)x], \quad \mu\in\mathbb{R}, \quad x\in\mathbb{R}, \tag{5.1a}$$

$$C_\mu = 2^{-3/4}\pi^{-1}\Gamma(-i\mu+\tfrac{1}{2})\exp[-\tfrac{1}{4}i\pi(i\mu+\tfrac{1}{2})], \tag{5.1b}$$

where D_ν is the parabolic cylinder function. [See Ref. 16, Eq. (2.22) for their computation; the method closely follows that of Miller *et al.*, Ref. 17.] The spectrum of H^r covers twice the real line, so $\mu\in\mathbb{R}$ and Φ_μ^{r+} and Φ_μ^{r-} are mutually orthogonal and a generalized basis for $\mathcal{L}^2(\mathbb{R})$. Our statement is now that (5.1) are self-reciprocal under the Laplace transform and that, due to (2.4b),

$$(\mathcal{L}\Phi_\mu^{r\pm})(x) = i\sqrt{2\pi}\exp(\pi\mu/2)\,\Phi_\mu^{r\pm}(x). \tag{5.2}$$

The statements (5.1)—(5.2) can be made more transparent with the use of the technique of Ref. 17, of transforming them to simpler operators on the same orbit under the group as that generated by (2.4e). For this it is sufficient to notice that $\mathcal{T}_L$ in (2.4b) [as well as the Bargmann transform $\mathcal{T}_B$ in (2.4c), next section] is on the same orbit as the dilatation transformation (2.4e), that is

$$\mathcal{T}_{M_0}\mathcal{T}_L\mathcal{T}_{M_0}^{-1} = \mathcal{T}_{G(-i\pi/2,0)}, \quad M_0 = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix} \tag{5.3}$$

(where M_0 represents the square root of the inverse Fourier transform, as $M_0^2 = F^{-1}$). Equation (5.3) can be verified by simply multiplying the corresponding 2×2 matrices. Now, the generalized eigenfunctions of H^d in (2.6) are

$$\Phi_\mu^{d\pm}(x) = (2\pi)^{-1/2}x_\pm^{i\mu-1/2}, \quad x_\pm \equiv \begin{cases}\pm x, & x \gtrless 0,\\ 0, & x \lessgtr 0,\end{cases} \tag{5.4}$$

with eigenvalue $\mu\in\mathbb{R}$. As (5.3) holds, we have that

$$\Phi_\mu^{r\pm}(x) = (\mathcal{T}_{M_0}^{-1}\Phi_\mu^{d\pm})(x). \tag{5.5}$$

The generalized orthonormality of the set (5.4) and its completeness for $\mathcal{L}^2(\mathbb{R})$ is known from the theory of Mellin transforms,[3] and from here the same statement follows for their unitary transforms (5.1) whose direct verification is far less straightforward than for (5.4). The action of transform (5.3) is that of dilatation by a factor $e^\beta = \exp(-i\pi/2)$. On the basis (5.4) this is clearly seen to be

$$\begin{aligned}(\mathcal{T}_{G(-i\pi/2,0)}\Phi_\mu^{d\pm})(x) &= \exp(-i\pi/4)\,\Phi_\mu^{d\pm}[\exp(-i\pi/2)x]\\ &= \exp(\pi\mu/2)\,\Phi_\mu^{d\pm}(x).\end{aligned} \tag{5.6}$$

From here and (2.4b), Eq. (5.2) follows.

One point which merits brief consideration is the following fact: $(\Phi_\mu^{r\sigma}, \Phi_\mu^{r\sigma'}) = \delta(\mu - \mu')\delta_{\sigma\sigma'}$ $(\sigma = \pm)$, under the ordinary scalar product on the real line. This implies, through the parseval identity for the Laplace transform, that for the scalar product

$$(\bar{f}, \bar{g})_L = \frac{1}{2\pi} \int_{-i\infty}^{i\infty} dx\, \bar{f}(x)^* \bar{g}(x) \tag{5.7a}$$

the same generalized orthogonality relation holds,

$$(\Phi_\mu^{r\sigma}, \Phi_{\mu'}^{r\sigma'})_L = \exp(\pi\mu)\delta(\mu - \mu')\,\delta_{\sigma\sigma'}. \tag{5.7b}$$

Completeness does not hold, however, as the transform is only isometric.

VI. SELF-RECIPROCAL FUNCTIONS UNDER THE BARGMANN TRANSFORM

The Bargmann transform[4] is closely related to the bilateral Laplace transform, as it has the same generating operator (1.2), namely in (2.5b). The value of the parameter τ is here $i\pi/4$. Indeed, we can point to the fact that $B^2 = L^{-1}$ through (2.4b) and (2.4c). This can be verified easily through integration. The self-reciprocal functions (5.1) of the Laplace transform will thus also be self-reciprocal under the Bargmann transform. With the proper constants from (2.4c) we have that

$$(\mathcal{B}\Phi_\mu^{r\pm})(x) = (2\pi)^{1/4}\exp(-\pi\mu/4)\Phi_\mu^{r\pm}(x). \tag{6.1}$$

This result can also be derived by noting that $\mathcal{T}_{M_0}\mathcal{T}_B\mathcal{T}_{M_0}^{-1} = \mathcal{T}_{G(i\pi/4,0)}$ where M_0 is given by (5.3), and repeating the argument of the last section. Finally, direct verification through integration[18] is also possible. As $B^{-2} = L$ this proves the result for Laplace transforms as well.

Now, use of the Parseval identity for Bargmann transforms and the unitarity of the transform, informs us that the set $\Phi_\mu^{r\pm}$ is a complete, orthonormal generalized basis for the Bargmann space $\mathcal{F}_B$ of analytic functions. Recall that the better-known bases for Bargmann's Hilbert space are the denumerable ("harmonic oscillator") monomials, i.e., powers of x, and the overcomplete coherent-state basis.[4,12] The repulsive oscillator basis can now be added to the list. The orthogonality relation (5.7) amounts to the same statement for the $\mathcal{F}_L$ space. On similar grounds, at the end of Sec. IX we will show that the Airy functions can be used to construct another such generalized basis.

VII. SELF-RECIPROCAL FUNCTIONS UNDER THE WEIERSTRASS-GAUSS TRANSFORM

For the Weiserstrass–Gauss transform we can use the eigenfunctions of (2.5c) and (2.5d) (recalling that the spectrum of this operator covers twice the positive half-axis), and choose

$$\Phi_\mu^f(x) = (2\pi)^{-1/2}\exp(i\mu x), \quad x \in \mathbb{R}, \quad \mu \in \mathbb{R} \tag{7.1}$$

which has eigenvalues $\frac{1}{2}\mu^2$ as an appropriate basis. The exponentiation (1.2) with $\tau = 2it$ then yields

$$(\mathcal{W}_t\Phi_\mu^f)(x) = \exp(-\mu^2 t)\,\Phi_\mu^f(x) \tag{7.2}$$

which is a well-known property of the heat equation Green's function.[5] Thus the functions (7.1) also provide a generalized orthogonal basis for the corresponding spaces $\mathcal{F}_{W_t}$.

A "radial" Weierstrass–Gauss transform for the matrix W_t in (2.4) with a modified Bessel function in the kernel (2.7) would have its self-reciprocal functions given by the eigenfunctions of $-\frac{1}{2}\Delta^{\{k\}}$ in (2.9), viz.,

$$\Phi_\mu^{f\{k\}}(x) = (\mu x)^{1/2} J_{2k-1}(\mu x), \quad x \in \mathbb{R}^+, \quad \mu \in \mathbb{R}^+ \tag{7.3}$$

with eigenvalue $\frac{1}{2}\mu^2$. An equation identical to (7.2) for these transforms follows. These functions will be used below.

VIII. SELF-RECIPROCAL FUNCTIONS UNDER THE BARUT-GIRARDELLO TRANSFORM

The Barut–Girardello transform[6] has the integral kernel (2.7) which stems from (2.8b) generated by (2.5b) with Δ given by (2.9). The generator is the quantum repulsive oscillator Hamiltonian with a centrifugal potential. Its eigenfunctions are

$$\Phi_\mu^{r\{k\}}(x) = C_\mu' x^{-1/2} M_{i\mu/2,k-1/2}(-ix^2), \quad x \in \mathbb{R}^+, \quad \mu \in \mathbb{R}, \tag{8.1a}$$

$$C_\mu' = 2^{(i\mu-1)/2}\pi^{-1/2}\Gamma(k + i\mu/2)\exp[i\pi(-i\tfrac{1}{4}\mu + k)]/\Gamma(2k), \tag{8.1b}$$

with eigenvalue μ, where $M_{\alpha\beta}$ is the Whittaker function. These functions can be found through the technique of Ref. 17, parallel to that of (5.5) through the use of the appropriate kernel (2.7). Since $\tau = i\pi/4$, as stated before, it follows that

$$(\mathcal{G}_k\Phi_\mu^{r\{k\}})(x) = \exp(-\pi\mu/4)\,\Phi_\mu^{r\{k\}}(x). \tag{8.2}$$

Remarks similar to those made in Sec. VI can be made to point out that (8.1) constitutes a new generalized basis for the Barut–Girardello space[8] $\mathcal{F}_B^{\{k\}}$.

IX. SELF-REPRODUCING FUNCTIONS UNDER CANONICAL TRANSFORMS

A useful generalization to the concept of self-reciprocal functions (1.1) to the class of canonical transforms $\mathcal{T}_M$, $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ given by (2.1) or (2.7) is to ask for functions $\phi_\mu^\omega(x)$ such that

$$[\mathcal{T}_M\phi_\mu^\omega](x) = C_\mu^M \exp i(\alpha_M x^2 + \beta_M x)\,\phi_\mu^\omega(\gamma_M x + \delta_M), \tag{9.1}$$

where C_μ^M, α_M, ..., δ_M are constants. We can call such functions "self-reproducing" under $\mathcal{T}_M$. This has been used in Ref. 19 in order to find the irreducible representation matrix elements of SL(2,R) for all subgroup reduction chains as well as for the nonsubgroup Airy function basis.[20] In Ref. 16 we were able to state some general results on separation of variables for a class of two-variable parabolic differential equations through exploring relations of the type (9.1), where $\mathcal{T}_{M(t)}$ represented the time evolution of a system governed by such an equation. It also allows us, via the unitarity of the transform to find new generalized function bases for spaces of analytic functions á la Bargmann, as we have done with the parabolic cylinder functions in Sec. VI.

The basic step is to write

$$\mathcal{T}_M = \mathcal{T}_{G(\beta,c)}\,\mathcal{H}_t^\omega, \tag{9.2}$$

where $\mathcal{T}_{G(\beta,c)}$ is the geometric transform (2.4e) and

TABLE I. Self-reciprocating functions under canonical transforms of the linear (2.1) ($\mathbb{I}=\mathbb{R}$) and radial (2.7) ($\mathbb{I}=\mathbb{R}^+$) types. The table headings refer to the constants appearing in Eq. (9.1).

Function ϕ_μ^ω	C_μ^M	α_M	β_M	γ_M	δ_M
Harmonic oscillator (3.1) (Hermite × Gaussian) $\mu=n+\frac{1}{2}$, $n=0, 1, 2, \cdots$ Radial [k] Harm. Oscill. (4.1) (Laguerre × Gaussian) $\mu=2(n+k)$, $n=0, 1, \cdots$	$(a^2+b^2)^{-1/4} \times \exp\left(-i\mu \tan^{-1}\frac{b}{a}\right)$	$\frac{ac+bd}{2(a^2+b^2)}$	0	$\frac{1}{(a^2+b^2)^{1/2}}$	0
Repulsive oscillator (5.1) (Parabolic cylinder) $\mu\in\mathbb{R}$ (twice) Radial [k] Rep. Oscill. (8.1) (Whittaker function)	$(a^2-b^2)^{-1/4} \times \exp\left(-i\mu \tanh^{-1}\frac{b}{a}\right)$	$\frac{ac-bd}{2(a^2-b^2)}$	0	$\frac{1}{(a^2-b^2)^{1/2}}$	0
Schrödinger free particle (7.1) (imaginary exponential) $\mu\in\mathbb{R}$ Radial [k] centrifugal pot. (7.3) (Bessel function) $\mu\in\mathbb{R}^+$	$a^{-1/2} \times \exp\left(-\frac{i}{2}\mu^2\frac{b}{a}\right)$	$\frac{c}{2a}$	0	$\frac{1}{a}$	0
Linear potential (9.4) (Airy function) $\mu\in\mathbb{R}$	$a^{-1/2}\exp i\left(-\mu\frac{b}{a}-\frac{5b^3}{12a^3}\right)$	$\frac{c}{2a}$	$-\frac{b}{a^2}$	$\frac{1}{a}$	$\frac{b^2}{2a^2}$

$\mathcal{H}_t^\omega=\exp(itH^\omega)$, H^ω being any of the operators considered in (2.5) or any other operator in the SL(2,C) orbit of one of these. By writing the corresponding 2×2 matrices for the transforms in (9.2), we can easily find β, c, and t in terms of the matrix elements of M through a set of coupled algebraic equations. Thus, when ϕ_μ^ω is an eigenfunction of H^ω, the action of $\mathcal{H}_t^\omega$ is to multiply ϕ_μ^ω by $\exp(i\mu t)$ and that of $\mathcal{T}_{G(\beta,c)}$ is given by (2.4e), yielding the form (9.1) for the transform function.

A simple example will illustrate the procedure for the harmonic oscillator functions $\Phi_n^h(x)$ in (3.1),

$$\left[\mathcal{T}_{\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)}\Phi_n^h\ (x)\right]=\left[\mathcal{T}_{\left(\begin{smallmatrix}\alpha&0\\\gamma&\alpha^{-1}\end{smallmatrix}\right)}\mathcal{T}_{\left(\begin{smallmatrix}\cos t&-\sin t\\\sin t&\cos t\end{smallmatrix}\right)}\Phi_n^h\right](x), \tag{9.3a}$$

$$\alpha=(a^2+b^2)^{1/2},\quad \gamma=(ac+bd)/\alpha,\quad \tan t=-b/a. \tag{9.3b}$$

Now, the right-factor matrix is generated by (2.5a), (see Sec. III), hence

$$\begin{aligned}\left[\mathcal{T}_{\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)}\Phi_n^h\right](x)&=\int_{-\infty}^{\infty}dx'\,A_M(x,x')\,\Phi_n^h(x')\\&=\exp[i(n+1/2)t]\left[\mathcal{T}_{\left(\begin{smallmatrix}\alpha&0\\\gamma&\alpha^{-1}\end{smallmatrix}\right)}\Phi_n^h\right](x)\\&=\alpha^{-1/2}\exp[i(n+1/2)t]\exp i\left(\frac{\gamma}{2\alpha}x^2\right)\Phi_n^h\left(\frac{x}{\alpha}\right).\end{aligned} \tag{9.3c}$$

Equation (9.3) can of course be verified directly using integral tables. The fact that Φ_n^h appears in the integrand and in the right-hand side of this equation, has thus been given a group-theoretical interpretation. The list of self-reproducing functions can then be drawn from the eigenfunctions of the operators (2.5) or the "radial" ones with (2.9). There is one further extension which for economy we have not mentioned at all in this article, but which appears in full detail in Ref. 16: The extension of the sl(2,R) algebra through a semidirect sum with a Heisenberg–Weyl algebra w with generators $x, -id/dx$, and 1 to an algebra $w \oplus \mathrm{sl}(2,R)$. When exponentiated to the group $W\otimes \mathrm{SL}(2,R)$, this brings in one new interesting orbit generated by $H^l=\frac{1}{2}P^2+Q$, i.e., the quantum free-fall or linear potential Schrödinger Hamiltonian. Its generalized eigenfunctions can be found to be

$$\Phi_\mu^l(x)=2^{1/3}Ai(2^{1/3}[x-\mu]),\quad x,\mu\in\mathbb{R}. \tag{9.4}$$

Now, in following Ref. 16 to deal with $W\otimes \mathrm{SL}(2,R)$ we can add (9.4) to the list of self-reproducing functions. In Table I, we summarize the results for the harmonic oscillator (3.1), repulsive oscillator (5.1), free (7.1), and linear potential (9.4), as well as the radial harmonic oscillator (4.1), radial repulsive oscillator (8.1), and pure centrifugal (7.3) Schrödinger eigenfunctions. It should be noted that the choice of these eigenfunctions (rather than the most general eigenfunction of a linear combination of these Schrödinger Hamiltonians) is no restriction at all, since assume we wish to ascertain the self-reproducing formula for a function $\mathcal{T}_{M_1}\phi_\mu^\omega$ where ϕ_μ^ω is one of the functions above. Now $\mathcal{T}_{M_1}\phi_\mu^\omega$ has the form in the right-hand side of (9.1). Write (9.2) as

$$\mathcal{T}_M\mathcal{T}_{M_1}=\mathcal{T}_{M_2}=\mathcal{T}_{M_1}\mathcal{T}_{G(\beta',c')}\mathcal{H}_{t'}^\omega, \tag{9.5}$$

where it is as easy to find t' and $G(\beta',c')$ in terms of M as it was before. Thus, the most general self-reproducing functions under canonical transforms are given by $\mathcal{T}_{M_1}\phi_\mu^\omega$ where ϕ_μ^ω appear in the table and have the structure (9.1). Table I can be used for all values such that the entries are nonsingular.[21] In particular, the table gives the results on self-reciprocal functions found in Secs. III–VIII as can be checked by replacing the proper matrix elements (2.4) and (2.6) into the six first entries of the table. As far as the last entry on Airy functions is concerned, it is of interest to point out that the Bargmann transform (2.4c) of (9.4), namely

$$\begin{aligned}[\mathcal{B}\Phi_\mu^l](x)=2^{1/2}\pi^{1/4}\exp[\tfrac{1}{2}x^2-\sqrt{2}x+\tfrac{5}{12}-\mu]\\ \times Ai(2^{5/6}x-2^{-2/3}-2^{1/3}\mu),\end{aligned} \tag{9.6}$$

constitutes a generalized orthonormal complete set of functions for Bargmann's Hilbert space $\mathcal{F}_B$. The argument follows that of the $\Phi_\mu^{r\pm}$ basis in Sec. VI.

[1]J. Arsac, *Fourier Transforms and the Theory of Distributions* (Prentice-Hall, Englewood Cliffs, N.J., 1966); S. Bochner, *Fourier Transforms* (Princeton U.P., Princeton, N.J., 1949); H. Dym and H.P. McKean, *Fourier Series and Integrals* (Academic, New York, 1972); A. Erdelyi *et al.*, Bate-

man manuscript project *Tables of Integral Transforms* (McGraw-Hill, New York, 1954), Vol. I; F. Oberhettingen, *Fourier Transforms of Distributions and Their Inverses* (Academic, New York, 1973); A. Papoulis, *The Fourier Integral and its Applications* (McGraw-Hill, New York, 1962); L. Schwartz, *Methods for the Physical Sciences* (Hermann, Paris, 1966); I. N. Sneddon, *Fourier Transforms* (McGraw-Hill, New York, 1951); E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals* (Clarendon, Oxford, 1967).

[2]Several of the references in Ref. 1, mainly E. C. Titchmarsh, contain ample material on Hankel or Bessel transforms; A. Erdelyi *et al.* Ref. 1, Vol. II, Chap. VIII; F. Oberhettingen, *Tables of Bessel Transforms* (Springer, Berlin, 1972); G. N. Watson, *A Treatise on the Theory of Bessel Functions* (Cambridge U. P., New York, 1922).

[3]V. A. Ditkin and A. P. Prudinov, *Integral Transforms and Operational Calculus* (Fitzmatgiz, Moscow, 1961); G. Doetsch, *Introduction to the Theory and Application of the Laplace Transform* (Birkhauser, Basel, 1959); M. G. Smith, *Laplace Transform Theory* (Van Nostrand, London, 1966); S. Colombo and J. Lavoine, *Transformations de Laplace et de Mellin* (Gauthier-Villars, Paris, 1972).

[4]V. Bargmann, Commun. Pure Appl. Math. **14**, 187 (1961); **20**, 1 (1967); in *Analytical Methods in Mathematical Physics*, edited by R. P. Gilbert and R. G. Newton (Gordon and Breach, New York, 1970); V. Bargmann, P. Butera, L. Girardello, and J. R. Klauder, Rep. Math. Phys. **2**, 221 (1971); P. Kramer, M. Moshinsky, and T. H. Seligman, in *Group Theory and its Applications*, edited by E. M. Loebl (Academic, New York, 1975), Vol. III.

[5]F. Tricomi, Math. Z. **40**, 720 (1936); G. Doetsch, Math. Z. **41**, 283 (1936); F. Tricomi, Ann. Inst. Henri Poincaré **8**, 111 (1938); J. Blackman, Duke Math. J. **19**, 671 (1952); D. V. Widder, Ann. Mat. Pura Appl. **42**, 279 (1956); P. G. Rooney, Canad. J. Math. **9**, 459 (1957); **10**, 613 (1958); G. G. Bilodeau, Canad. J. Math. **13**, 593 (1961); P. G. Rooney, Canad. Math. Bull. **6**, 45 (1953); D. V. Widder, J. Aust. Math. Soc. **4**, 1 (1964).

[6]A. O. Barut and L. Girardello, Commun. Math. Phys. **21**, 41 (1971); P. Kramer *et al.*, Ref. 4.

[7]K. B. Wolf, J. Math. Phys. **15**, 1295 (1974).

[8]K. B. Wolf, J. Math. Phys. **15** 2102 (1974).

[9]See E. C. Titchmarsh, Ref. 1, Sec. 3.8.

[10]G. H. Hardy and E. C. Titchmarsh, Quart. J. Math. **1**, 196 (1930), this material basically constitues Chapter IX of the book by E. C. Titchmarsh in Ref. 1. There, the case $\lambda = 1$ [in Eq. (1.1)] is studied and the case $\lambda = -1$ rather briefly commented upon. The self-reciprocal property is related with the parity under inversions through a line in the complex plane of a function closely related to the Mellin and Laplace transforms of the self-reciprocal function.

[11]B. Mohan, Proc. London Math. Soc. **34**, 231 (1932); Proc. Edinburgh Math. Soc. **4**, 53 (1934); Quart. J. Math. **10**, 252 (1939); S. K. Bose, Ganita **5**, 25 (1954); K. P. Bhatnagar, Bull. Calcutta Math. Soc. **46**, 179 (1954); G. Krishna, Proc. Natl. Acad. Sci. India **26**, 343 (1960); S. Masood, Proc. Natl. Acad. Sci. India **29**, 372 (1960); V. V. L. N. Rao, Proc. Cambridge Philos. Soc. **57**, 561 (1961) and references therein.

[12]J. Glauber, Phys. Rev. **130**, 2529 (1963); **131**, 2766 (1963); C. Mandel and E. Wolf, Rev. Mod. Phys. **37**, 231 (1965).

[13]M. Moshinsky, Proceedings of the XV Solvay Conference in Physics (Brussels, 1970); M. Moshinsky and C. Quesne, J. Math. Phys. **12**, 1772, 1780 (1971); M. Moshinsky, SIAM J. Appl. Math. **25**, 193 (1973).

[14]M. Moshinsky, T. H. Seligman, and K. B. Wolf, J. Math. Phys. **13**, 901 (1972).

[15]As the use of the normalization constant is not yet fixed (see Kramer *et al.*, Ref. 4 and Ref. 6) we opt for setting it to unity in our context.

[16]J. B. Wolf, J. Math. Phys. **17**, 601 (1976).

[17]E. G. Kalnins and W. Miller, Jr., J. Math. Phys. **15**, 1728 (1974).

[18]After completing squares in the exponent, use I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products* (Academic, New York, 1965), Eq. 7.724; in the integration, the contour can be shifted back to the real line as the integrand behaves properly in the second and fourth quadrant.

[19]C. P. Boyer and K. B. Wolf, "Finite SL(2, R) Representation Matrices of the D_k^+ Series for all Subgroup Reductions," Comumicaciones Técnicas CIMAS 6B, No. 104 (1975); to appear in Rev. Mex. Fís. (Spring 1977).

[20]K. B. Wolf, "The Airy Function as a Non-subgroup Basis for the Oscillator Representation $(D^+_{1/4} \dot{+} D^+_{3/4})$ of SL(2, R)," Comunicaciones Técnicas IIMAS GB No. 118 (1975), to appear in Rev. Mex. Fís. (Spring, 1977).

[21]Singularities occur (notably in the repulsive oscillator case) when the decomposition (9.2) fails for a given M and ω. This is the case of the Fourier transform (2.4a). In this case, the transforms of the three last entries are $(2\pi)^{-1/2}x^{-i\lambda-1/2}$, $\delta(x-\lambda)$, and $(2\pi)^{-1/2}\exp i(-\lambda q + q^3/6)$, so that self-reproduction fails correspondingly.

Spatial coherence of acoustic signals in randomly inhomogeneous waveguides—A multiple-scatter theory

G. R. Sutton* and J. J. McCoy

The Catholic University of America, Washington, D. C. 20064
(Received 11 May 1976)

A normal mode theory is formulated for estimating the loss of spatial coherence of acoustic signals propagating through a randomly inhomogeneous waveguide. The theory is applicable for experiments in which multiple-scatter effects are important provided the rate at which energy is scattered with increasing range is limited. Other limitations on the theory require that intermodal scattering between propagating and nonpropagating modes be negligible and that the intramodal scatter of energy within a single mode be narrow-angled, when measured in the waveguide plane. The scattering of acoustic signals, of significant frequency content ranging upwards from tens of Hertz, by large scale interval wave fields, i.e., wavelengths of the order of 1 kilometer or greater, satisfy the conditions for applicability of the theory.

INTRODUCTION

The propagation of scalar radiation fields in randomly inhomogeneous media has been the subject of an intense research effort over the past two decades. One problem of fundamental interest is an infinite, statistically homogeneous medium characterized by a weakly fluctuating refractive index field superimposed on a constant background value, i. e., $n^2(\mathbf{x}) = n_0^2[1+\epsilon\mu(\mathbf{x})]$, where n_0^2 is a constant, $\mu(\mathbf{x})$ is everywhere of the order of unity and $\epsilon \ll 1$. A radiation field is introduced by the statement that the intensity across a given plane, say the $z=0$ plane, is independent of position in the plane and is directed normal to the plane in the positive z direction. In the absence of refractive index fluctuations the resulting radiation field is given by a plane wave. The effect of the refractive index fluctuations can be investigated by considering the intensity across a second z plane, $z>0$. Since the refractive index fluctuations have been assumed to be statistically homogeneous and the intensity at the $z=0$ plane is independent of absolute position, the intensity will be independent of position at any plane for which $z>0$. However, the intensity will no longer be unidirectional, as given by a plane wave. To describe the intensity at a plane $z>0$ a directional intensity spectrum is introduced, which may be interpreted as a spatial analog of the power spectrum of random signal theory. The spectrum resolves the intensity at the z plane into propagation directions. The prediction problem is for a descriptor of this intensity spectrum as a function of range, radiation wavelength, and the statistics of the refractive index fluctuations.

The problem of scalar radiation fields in randomly inhomogeneous waveguides has received considerably less attention than has the infinite medium problem. One can view the infinite waveguide as a two-dimensional random medium in which there is an infinity of modes, some of which can propagate energy and others that are evanescent. In the absence of the random inhomogeneities the energy in each propagating mode remains in that mode. The presence of the inhomogeneities introduces a modal coupling, which results in a continual transfer of energy between the modes of propagation. The analog of the infinite medium problem described above is a radiation field at $z=0$ that has energy in only one mode and is unidirectional. The prediction problem for the intensity spectrum at $z>0$ must account for the resolutions of the intensity into modes of propagation and the further resolution of the intensity, in a given mode, into directions of flow in the plane of the waveguide. In our study we consider this problem.

Previous research on randomly inhomogeneous waveguides[1–8] appears to be limited to the resolution of the energy into the several waveguide modes. Furthermore, the majority of these efforts have also been limited to single scatter treatments of the problem. Thus, their results are limited in range as well as in the strength of the inhomogeneities. Of the papers that discuss multiple-scatter solutions, the present study is most closely related to the work of Marcuse.[7] Like Marcuse's theory, our development of a theory that is suitable in the multiscatter region is based on a range incrementing procedure. This technique, which appears to have been first used in discussing stochastic scattering by Beran,[9] is now used quite commonly. Its applicability is clearly limited to problems in which the scattering can be described as a forward scattering. Beran used it to investigate the infinite medium problem discussed at the outset of this paper for a scattering experiment in which $\bar{\lambda}/l_m \ll 1$, where $\bar{\lambda}$ denotes the radiation field wavelength and l_m denotes the smallest correlation length of the fluctuating refractive index field. We refer to this problem as the optics problem since it models the propagation of a laser beam through atmospheric turbulence. More recently, Beran and McCoy[10–12] used the same technique to investigate propagation through a highly anisotropic infinite medium for which $\bar{\lambda}/l_{Hm} \ll 1$ but $\bar{\lambda}/l_{VM} \gtrsim 1$. Here l_{Hm} is the minimum horizontal correlation length and l_{VM} is the maximum vertical correlation length. We refer to this problem as the acoustics problem since it models the long range propagation of a sound beam through an oceanic internal wave field.

The point of departure of our studies from that discussed in the preceding paragraph is founded in the motivation. Marcuse's work is directed to propagation in fiber optic waveguides, for which the principal concern is to estimate the loss of energy to the waveguide due to boundary interaction and to estimate the transfer of energy among the waveguide modes. The degree of spatial coherence of the energy within a single waveguide mode is of little interest. Our principal interests, like those of Beran and McCoy, are motivated by ocean acoustics experiments in which the spatial coherence of the signal is of primary importance for ascertaining receiver array gain or receiver array resolution capabilities. Our formulation differs from that of Beran and McCoy in being a normal mode formulation, whereas theirs is a parabolic formulation. Thus, the two formulations are complementary with the normal mode theory becoming more suitable with decreasing frequency content for the acoustic signal.

1. DEVELOPMENT OF GOVERNING EQUATIONS

The waveguide is referred to by Cartesian coordinates with the y axis directed normal to the waveguide plane, which is taken to be horizontal. The acoustic medium is described by a weakly-random sound speed field superimposed on a deterministic background field. The statistics of the fluctuating field are assumed to be homogeneous and isotropic for measurements taken in any given horizontal plane. The possibility of statistical inhomogeneity for measurements taken over the depth coordinate is retained, as is the possibility of statistical anisotropy for measurements taken over the depth coordinate when compared to those taken in a waveguide plane. The background sound speed field can vary with y but not with position in the waveguide plane. The acoustic field is taken to be harmonic in time with circular frequency, ω.

The waveguide with the background medium is described by the normal mode functions, $Y_i(y)$, defined by the eigenvalue problem

$$\frac{d^2 Y_i}{dy^2} + [\bar{k}^2(y) - \beta_i^2]\, Y_i = 0, \tag{1}$$

together with appropriate conditions at the waveguide face(s). The depth dependent mean wavenumber is denoted by $\bar{k}^2 = \omega^2/\bar{c}^2$, where $\bar{c}/(y)$ is the background sound speed field; the eigenvalue corresponding to the ith modal function is denoted by β_i. In general, the eigenvalue problem will give both a discrete and a continuous spectrum, the continuous portion being required to describe any energy leakage out of a waveguide formed by a sound speed well. The continuous spectrum is understood in terms of lateral waves, the discrete spectrum in terms of trapped modes.

The acoustic pressure field in the random waveguide problem, $\hat{p}(\mathbf{x})$, can be formally represented by

$$\hat{p}(\mathbf{x}) = \sum_i \hat{p}_i(\mathbf{r})\, Y_i(y), \tag{2}$$

where the summation is over both discrete and continuous modes. The $\hat{p}_i(\mathbf{r})$ are termed the modal amplitudes and vary with position in the waveguide plane, located by the two-dimensional position vector, $\mathbf{r}$. The modal amplitudes are governed by the set of differential equations,

$$\nabla_\perp^2 \hat{p}_i + \beta_i^2 \hat{p}_i = -\epsilon \sum_j \mu_{ij}(\mathbf{r})\, \hat{p}_j, \tag{3}$$

where

$$\mu_{ij}(\mathbf{r}) = \int \bar{k}^2(y)\, \mu(\mathbf{x})\, Y_i(y)\, Y_j(y)\, dy. \tag{4}$$

The two-dimensional Laplacian is denoted by $\nabla_\perp^2$, the randomly varying wavenumber field by $\epsilon \bar{k}^2(y)\, \mu(\mathbf{x})$. Here μ is a stochastic function of position of unit variance. Hence, ϵ provides a measure of the "strength" of the variations. We assume $\epsilon \ll 1$.

The theory is to be formulated in terms of modal coherence functions defined according to

$$\{\hat{\Gamma}_{ii}(x_1, x_2, z)\} = \{\hat{p}_i(x_1, z)\, \hat{p}_i^*(x_2, z)\}, \tag{5}$$

where the braces indicate an ensemble averaging. Thus, the coherence function is a spatial correlation function taken at two points in the same z plane, where z is taken to correspond to the principal propagation, or range, direction. For the problem specified in the Introduction, i.e., an initial plane wave directed along the z axis and homogeneous statistics measured in the x, z plane, the modal coherence functions vary with $x_{12} = x_1 - x_2$ being independent of absolute position along the x axis. The x_{12} transform $\{\hat{\Gamma}_{ii}(x_{12}, z)\}$ is termed the modal intensity directional spectral density. This function provides the intensity resolutions discussed in the Introduction.

The modal coherence function for $x_{12} = 0$, i.e.,

$$\{\hat{I}_i(z)\} = \{\hat{\Gamma}_{ii}(0, z)\} \tag{6}$$

is termed the modal intensity function. It provides a measure, somewhat imprecise, of the modal energy flux per unit area that passes through the point $x_1 = x_2, z$. (For the problem posed, $\{\hat{I}_i\}$ is independent of the absolute location of $x_1 = x_2$.) A more precise measure of modal energy flux per unit area would weight the energy flux in a direction making an angle θ with the axis, by $\cos\theta$, to incorporate the projection of the unit area normal to z onto a plane normal to the propagation direction. For the narrow-angled spectra of interest, $\theta \ll 1$ and the modal intensity function provides a useful estimate of energy flux per unit area.

The more familiar mutual coherence function for two points in the same z plane is given by

$$\{\hat{\Gamma}(x_1, y_1, x_2, y_2, z)\} = \{\hat{p}(x_1, y_1, z)\, \hat{p}^*(x_2, y_2, z)\}. \tag{7}$$

By substituting Eq. (2) into Eq. (7), we can write

$$\{\hat{\Gamma}(x_1, y_1, x_2, y_2, z)\} = \sum_i \sum_j \{\hat{\Gamma}_{ij}(x_1, x_2, z)\}\, Y_i(y_1)\, Y_j(y_2), \tag{8}$$

where in writing Eq. (8) we were required to introduce cross-modal coherence functions, which are obvious generalizations of the $\{\hat{\Gamma}_{ii}\}$. Upon setting $y_1 = y_2$ in Eq. (8) and integrating over the waveguide depth, we obtain, upon making use of the orthonormality of the Y_i,

$$\int \{\hat{\Gamma}(x_1, y, x_2, y, z)\}\, dy = \sum_i \{\hat{\Gamma}_{ii}(x_1, x_2, z)\}. \tag{9}$$

Equation (9) states that the modal coherence functions provide the modal decomposition of the averaged (taken over the waveguide depth) coherence for two points located along the same horizontal line positioned orthogonal to z.

In the range incrementing derivation procedure, we divide the waveguide by a series of range planes separated by a distance, Δz. Since the problems to which the theory is to be applied involve only forward propagation, we then calculate the modal coherence functions at one range plane in terms of those of the plane immediately preceding it. The increment Δz is taken to be small enough to enable use of a single-scatter theory for this calculation. Also, in making this calculation, we assume the statistics of the $\hat{p}_i$ measured on any range plane to be independent of the statistics of the sound speed fluctuations in the interval "in front of" the range plane. This is clearly consistent with the forward propagation assumption over most of the interval if we take Δz to be large relative to $l_{\rm HM}$, the maximum correlation length along a line in the horizontal plane. Finally, we make a series of approximations to simplify the single-scatter solution to a point at which computationally useful expressions are obtained. Conditions justifying the approximations are drawn when the approximations are made and are summarized by Eqs. (37).

Making use of a perturbation solution of Eqs. (3), we write the following expression for the cross-modal coherence function for two points on the $(j+1)\Delta z$ plane:

$$\begin{aligned}\{\hat{\Gamma}_{ij}(x_{12},(j+1)\Delta z)\} &= \{\hat{\Gamma}_{ij}^{(0)}(x_{12},(j+1)\Delta z)\} + \epsilon^2 \sum_k \sum_l \Big[\int_A \int_A G_i(\mathbf{r}_1,\mathbf{r}')G_j^*(\mathbf{r}_2,\mathbf{r}'') \\ &\quad \times \sigma_{ikjl}(\mathbf{r}'-\mathbf{r}'')\{\hat{\Gamma}_{kl}^{(0)}(\mathbf{r}',\mathbf{r}'')\}d\mathbf{r}'\,d\mathbf{r}'' \\ &\quad + \int_A \int_{A_R} G_i(\mathbf{r}_1,\mathbf{r}')G_k(\mathbf{r}',\mathbf{r}'')\sigma_{ikkl}(\mathbf{r}'-\mathbf{r}'') \\ &\quad \times \{\hat{\Gamma}_{lj}^{(0)}(\mathbf{r}'',\mathbf{r}_2)\}d\mathbf{r}''\,d\mathbf{r}' + \int_A \int_{A_R} G_j^*(\mathbf{r}_2,\mathbf{r}')G_k^*(\mathbf{r}',\mathbf{r}'') \\ &\quad \times \sigma_{jkkl}(\mathbf{r}'-\mathbf{r}'')\{\hat{\Gamma}_{il}^{(0)}(\mathbf{r}_1,\mathbf{r}'')\}d\mathbf{r}''\,d\mathbf{r}' \Big], \end{aligned} \tag{10}$$

where

$$\begin{aligned}\sigma_{ijkl}(\mathbf{r}'-\mathbf{r}'') &= \{\mu_{ij}(\mathbf{r}')\mu_{kl}(\mathbf{r}'')\} \\ &= \int\int \bar{k}^2(y')\bar{k}^2(y'')\sigma(\mathbf{x}',\mathbf{x}'') \\ &\quad \times Y_i(y')Y_j(y')Y_k(y'')Y_l(y'')\,dy'\,dy'', \end{aligned} \tag{11}$$

where

$$\sigma(\mathbf{x}',\mathbf{x}'') = \{\mu(\mathbf{x}')\mu(\mathbf{x}'')\}. \tag{12}$$

In Eqs. (10), the area A is the region of the waveguide plane between the jth and $(j+1)$th range planes; the area A_R is the region between a generic intermediate range plane located by z' and the $(j+1)$th plane. The term $\{\hat{\Gamma}_{ij}^{(0)}(\mathbf{r}',\mathbf{r}'')\}$ is the cross-modal coherence function, for a pair of points not necessarily in the same range plane, for the waveguide with no random sound speed variations beyond the jth range plane. Thus, $\{\hat{\Gamma}_{ij}^{(0)}\}$ is determined from the cross-modal coherence functions on the jth plane by solving a homogeneous waveguide problem. Finally, the modal Green's functions, $G_i(\mathbf{r},\mathbf{r}')$, are taken to be free space Green's functions, given in terms of Hankel functions of the first kind of order zero, $H_0^{(1)}$, for β_i^2 real and positive, and by the continuation of $H_0^{(1)}$ for other β_i^2. Real positive β_i^2 define the propagating modes of the waveguide. Use of the free space Green's function as well as the choice of the domains of integration in Eqs. (10) is consistent with the restriction to forward-directed scattering.

In Eq. (12), σ is the spatial correlation function of the randomly varying sound-speed field. The matrix σ_{ijkl} is a modal correlation matrix defined by this correlation function and the modal eigenfunctions. For homogeneous and isotropic statistics for measurements taken in a waveguide plane,

$$\sigma_{ijkl}(\mathbf{r}',\mathbf{r}'') = \sigma_{ijkl}(|\mathbf{r}'-\mathbf{r}''|). \tag{13}$$

As indicated previously, we next reduce the integrals in Eq. (10) to more manageable expressions. The first simplifications to be introduced are (i) to neglect any coupling in Eqs. (10) between propagating and nonpropagating wave modes, and (ii) to approximate the Hankel functions in the G's by their large agreement asymptotic representation. We shall see that the strength of the coupling between propagation modes depends on the magnitude of $\beta_i - \beta_n$, the difference of the modal eigenvalues; the larger $\beta_i - \beta_n$, the weaker the coupling. Thus, coupling will be strongest between propagating modes, with real values for β_i^2, that may be termed neighbors. Coupling between propagating and nonpropagating wave modes will be much weaker. It is to be noted, however, that this coupling, no matter how small, will act continuously and is to be expected to have significant effects over long enough ranges. Our neglect restricts the theory to experiments in which these latter ranges are very long when compared to those over which the effects of coupling between the propagating modes and the loss of coherence of the energy within a given wave mode become significant. The justification for approximating the Hankel functions by their asymptotic representations is that we can neglect the region of integration within which the arguments are not large. These regions have a linear extent of the order of β_i^{-1}. Of the remaining length scales encountered in the integral the shortest is of order $l_{\rm Hm}$. Thus, the approximation is valid if $\beta_i l_{\rm Hm} \gg 1$, a restriction that is imposed on the theory.

Introducing these approximations into Eq. (10), the first integral term is written

$$\begin{aligned} I_{ij}^{(1)}(\mathbf{r}_1,\mathbf{r}_2) &= \frac{\epsilon^2}{8\pi}\beta_i^{1/2}\beta_j^{1/2} \\ &\quad \times \int_A \int_A \frac{\exp[i(\beta_i|\mathbf{r}_1-\mathbf{r}'| - \beta_j|\mathbf{r}_2-\mathbf{r}''|)]}{|\mathbf{r}_1-\mathbf{r}'|^{1/2}|\mathbf{r}_2-\mathbf{r}''|^{1/2}} \\ &\quad \times \sigma_{ikjl}(\mathbf{r}'-\mathbf{r}'')\{\hat{\Gamma}_{kl}^{(0)}(\mathbf{r}',\mathbf{r}'')\}d\mathbf{r}'\,d\mathbf{r}''. \end{aligned} \tag{14}$$

We consider the terms $|\mathbf{r}_1-\mathbf{r}'|$ and $|\mathbf{r}_2-\mathbf{r}''|$, write the second as

$$|\mathbf{r}_2-\mathbf{r}''| = z_2 - z'' + \frac{(x_2-x'')^2}{2(z_2-z'')} + O\left[\frac{(x_2-x'')^4}{(z_2-z'')^3}\right], \tag{15}$$

and note that we can truncate this expression, as it appears in the exponent, after two terms provided, we can show that

$$\frac{\beta_j(x_2-x'')^4}{(z_2-z'')^3} \ll 1 \tag{16}$$

over the integration region that contributes significantly. Thus, if the directional intensity spectrum of the signal is bounded by an angle θ, the restriction will be satisfied provided

$$\beta_j\theta^4\Delta z \ll 1. \tag{17}$$

This condition is to be interpreted as an upper limitation on the size of Δz. In a previous study[8] we showed that the angular spread due to a single intramodal scattering is of the order of $(\beta_j l_{\rm Hm})^{-1}$, when $\beta_j l_{\rm Hm} \gg 1$. Use of this in Eq. (17) together with the requirement that $\Delta z \gg l_{\rm HM}$ leads to the following condition on the relative sizes of the signal wavelength and the sound speed fluctuations correlation length scales:

$$\frac{l_{\rm HM}}{\beta_j^3 l_{\rm Hm}^4} \ll 1. \tag{18}$$

This restriction is less severe than one to be ultimately accepted on the theory.

The truncation of the expression for $|\mathbf{r}_2 - \mathbf{r}''|$ as it appears in the denominator, after a single term, is justified if $\theta^2 \ll 1$, a condition that is consistent with the others to be accepted.

With these approximations we write Eq. (14) as

$$\begin{aligned} I_{ij}^{(1)}(\mathbf{r}_1, \mathbf{r}_2) = {} & \frac{\epsilon^2}{8\pi\beta_i^{1/2}\beta_j^{1/2}} \exp[i(\beta_i - \beta_j)z] \\ & \times \iint \exp[i(\beta_j z'' - \beta_i z')] \\ & \times \exp(i[\beta_i(x_1 - x')^2/2(z - z') \\ & - \beta_j(x_2 - x'')^2/2(z - z'')]) \\ & \times (z - z')^{-1/2}(z - z'')^{-1/2} \\ & \times \sigma_{ikjl}(\mathbf{r}' - \mathbf{r}'')\{\hat{\Gamma}_{kl}^{(0)}(\mathbf{r}', \mathbf{r}'')\} d\mathbf{r}' \, d\mathbf{r}''. \end{aligned} \tag{19}$$

In writing Eq. (19) we have set $z_1 = z_2 = z$. The integration over x'' and x' are now accomplished in the manner of the saddle point method of integration. (This approximation amounts to use of a geometric theory over a range $l_{\rm HM}$.) The lowest order term gives

$$\begin{aligned} I_{ij}^{(1)}(\mathbf{r}_1, \mathbf{r}_2) = \frac{\epsilon^2}{4\beta_i\beta_j} \Bigl[& \iint \sigma_{ikjl}(x_2, z'' - z')\{\hat{\Gamma}_{kl}^{(0)}(x_{12}, z', z'')\} \\ & \times \exp[i(\beta_j z'' - \beta_i z')] \, dz'' \, dz' \Bigr] \exp[i(\beta_i - \beta_j)z]. \end{aligned} \tag{20}$$

The first order correction, when compared to the coefficient of the exponential in Eq. (20) is seen to be of the order of

$$\frac{1}{L^2}\left(\frac{z - z''}{\beta_j} - \frac{z - z'}{\beta_i}\right), \tag{21}$$

where L^{-1} is a measure of the curvature of $\sigma_{ikjl}\{\hat{\Gamma}_{kl}^{(0)}\}$ with respect to x_{12}. For $\beta_i = \beta_j$, which we shall find to be the important case, the above is written

$$(z' - z'')/\beta_i L^2. \tag{22}$$

Noting that $L \gtrsim l_{\rm Hm}$ and that $z' - z'' < l_{\rm HM}$, neglect of the first order correction is justified provided

$$l_{\rm HM}/\beta_i l_{\rm Hm}^2 \ll 1. \tag{23}$$

The condition is recognized as a statement that $l_{\rm HM}$ falls within the geometric theory range. This condition is somewhat weaker than that ultimately accepted on the theory.

To simplify Eq. (20) further requires the approximation that

$$\begin{aligned} & \{\hat{\Gamma}_{kl}^{(0)}(x_{12}, z', z'')\} \\ & = \{\hat{\Gamma}_{kl}(x_{12}, j\Delta z)\} \exp(i[\beta_k(z' - j\Delta z) - \beta_l(z'' - j\Delta z)]). \end{aligned} \tag{24}$$

Equation (24) essentially neglects diffraction effects over a Δz range. The approximation is somewhat stronger than Eq. (23) in that it requires

$$\beta_k \theta^2 \Delta z \ll 1. \tag{25}$$

Introducing Eq. (24) into Eq. (20) gives

$$\begin{aligned} & I_{ij}^{(1)}(x_1, x_2, (j+1)\Delta z) \\ & = \frac{\epsilon^2}{4\beta_i\beta_j}\{\hat{\Gamma}_{kl}(x_{12}, j\Delta z)\} \exp[i(\beta_i - \beta_j)(j+1)\Delta z \\ & \quad + i(\beta_l - \beta_k)j\Delta z] \iint \sigma_{ikjl}(x_{12}, z' - z'') \\ & \quad \times \exp[i(\beta_k - \beta_i)z' + i(\beta_j - \beta_l)z''] \, dz'' \, dz'. \end{aligned} \tag{26}$$

Introducing sum and difference coordinates we write

$$\begin{aligned} & I_{ij}^{(1)}(x_{12}, (j+1)\Delta z) \\ & = \frac{\epsilon^2}{4\beta_i\beta_j}\{\hat{\Gamma}_{kl}(x_{12}, j\Delta z)\} \exp[i(\beta_i - \beta_j)(j+1)\Delta z \\ & \quad + i(\beta_l - \beta_k)j\Delta z] \iint \sigma_{ikjl}(x_{12}, s_z) \exp[i(\beta_k - \beta_i + \beta_j - \beta_l)p_z \\ & \quad + \frac{i}{2}(\beta_k - \beta_i - \beta_j + \beta_l)s_z] \, dp_z \, ds_z. \end{aligned} \tag{27}$$

Again we make use of the assumption that $\Delta z \gg l_{\rm HM}$, which enables our uncoupling the regions over which the two integrals are to be carried out. The integral over p_z is seen to be equal to Δz, provided

$$\beta_k - \beta_i - \beta_j - \beta_l = 0,$$

and it is seen to be negligibly small if this equation is not satisfied. In particular, we note that $i = j$ requires that $k = l$ and we write

$$I_{ii}^{(1)} = \frac{\epsilon^2 \delta_{kl}}{\beta_i} \beta_k \bar{\sigma}_{ik}(x_{12})\{\hat{\Gamma}_{kk}(x_{12}, j\Delta z)\}\Delta z, \tag{28}$$

where

$$\bar{\sigma}_{ik}(x_{12}) = \frac{1}{2\beta_i\beta_k}\int_0^\infty \sigma_{ikik}(x_{12}, s_z)\cos[(\beta_k - \beta_i)s_z] \, ds_z. \tag{29}$$

Equation (28) indicates that the first series on the rhs of Eq. (10) does not result in coupling the diagonal, $\{\hat{\Gamma}_{ii}\}$, to the off-diagonal, $\{\hat{\Gamma}_{ij}\}$ terms. Coupling of the diagonal terms is controlled by $\bar{\sigma}_{ik}(x_{12})$, which is seen to be a spectral representation of the correlation matrix, $\sigma_{ikik}(x_{12}, s_z)$. Thus, the coupling between diagonal terms will fall off very rapidly with increasing $|\beta_i - \beta_i|$ beyond a value of $l_{\rm Hm}^{-1}$. Since the theory requires that $\beta_i l_{\rm Hm} \gg 1$, this conclusion justifies the statement that modal coupling will be strongest between neighboring propagation modes.

The reduction of the second and third series proceeds in a manner similar to the above. Taken together we obtain for the generic diagonal term

$$I_{ii}^{(2)} = -\epsilon^2 \delta_{kl}\{\hat{\Gamma}_{ii}(x_{12}, j\Delta z)\}\bar{\sigma}_{ik}(0)\Delta z. \tag{30}$$

Thus, in the reduced formulation there is no coupling between diagonal and off-diagonal terms. We write for the diagonal terms the following system of difference equations:

$$\begin{aligned} & \{\hat{\Gamma}_{ii}(x_{12}, (j+1)\Delta z)\} \\ & = \{\hat{\Gamma}_{ii}(x_{12}, j\Delta z)\} - \epsilon^2[\{\hat{\Gamma}_{ii}(x_{12}, j\Delta z)\}\sum_k \bar{\sigma}_{ik}(0) \\ & \quad - \frac{1}{\beta_i}\sum_k \beta_k \bar{\sigma}_{ik}(x_{12})\{\hat{\Gamma}_{kk}(x_{12}, j\Delta z)\}]\Delta z. \end{aligned} \tag{31}$$

This set of difference equations can be approximated by a set of differential equations, which we write,

$$\frac{\partial\{\hat{\Gamma}_{ii}(x_{12},z)\}}{\partial z}=-\epsilon^2\Big(\sum_k\bar{\sigma}_{ik}(0)\Big)\{\hat{\Gamma}_{ii}(x_{12},z)\}+\frac{\epsilon^2}{\beta_i}\sum_k\beta_k\bar{\sigma}_{ik}(x_{12})\{\Gamma_{kk}(x_{12},z)\}. \tag{32}$$

The differential equations follow exactly from the difference equations in the limit of $\Delta z\to 0$. However, our derivation procedure requires $\Delta z\gg l_{\mathrm{HM}}$, making this final step an approximation. It is similar to the approximation taken in formulating a continuum theory to predict the response of a discrete system.

Equations (32) is the sought-after scattering theory for predicting the modal coherence functions. We can immediately use these equations to obtain a formulism on $\{\hat{I}_i(z)\}$, the modal intensity functions, and on $\{\hat{I}_i^c(z)\}$, the coherent modal intensity function given by $\{\hat{\Gamma}_{ii}(\infty,z)\}$. We write

$$\frac{d\{\hat{I}_i(z)\}}{dz}=\epsilon^2\left[\frac{1}{\beta_i}\sum_k\beta_k\bar{\sigma}_{ik}(0)\{\hat{I}_k(z)\}-\Big(\sum_k\bar{\sigma}_{ik}(0)\Big)\{\hat{I}_i(z)\}\right] \tag{33}$$

and

$$\frac{d\{I_i^c(z)\}}{dz}=-\epsilon^2\Big(\sum_k\bar{\sigma}_{ik}(0)\Big)\{I_i^c(z)\}. \tag{34}$$

It is to be noted that a complete formulation is obtained in terms of modal intensity functions only for the incident plane wave and for statistics that are homogeneous in the waveguide plane. For beamed signals or for inhomogeneous statistics, $\{\hat{\Gamma}_{ii}\}$ will depend on absolute position of the observation point along x. For such problems the task of determining the energy distribution among the wave modes is coupled to that of determining the modal coherence measured at two points along the x axis in an inseparable manner.[10,12]

Equation (33) provides a statement of energy conservation,

$$\sum_i\beta_i\{\hat{I}_i(z)\}=\text{const}, \tag{35}$$

as expected of a theory incorporating only forward scattering and neglecting any energy losses. The equations on the coherent modal intensity functions are uncoupled as expected, since the transfer of energy between modes is accomplished by scattering, which will always result in a loss of coherence.

Equations (34) allow an estimate of the rate at which energy is being scattered. Thus, the limit of the single-scatter region, i.e., the limiting range over which multiple-scatter effects can be neglected, can be estimated by

$$z^*=\min\Big(\epsilon^2\sum_k\bar{\sigma}_{ik}(0)\Big)^{-1}. \tag{36}$$

Equation (36) provides the last restriction on the theory, $\Delta z\ll z^*$; this restriction is needed to justify use of the single-scatter theory over the range increment. Summarizing, the derivation procedure requires a range increment, Δz such that

$$l_{\mathrm{HM}}\ll\Delta z\ll\min\Big(\epsilon^2\sum_k\bar{\sigma}_{ik}(0)\Big)^{-1} \tag{37a}$$

and

$$\beta_k\theta^2\Delta z\ll 1. \tag{37b}$$

2. SOLUTION OF GOVERNING EQUATIONS

The separation distance, x_{12}, appears in Eqs. (32) as a parameter. The set of equations constitute a set of constant coefficient ordinary differential equations, the theory of which is well established. The general solution of the set of equations is given by a linear combination of solutions of the form

$$\{\hat{\Gamma}_{ii}(z;x_{12})\}=\frac{1}{\beta_i}\gamma_i(x_{12})\exp[-\epsilon^2 0(x_{12})z], \tag{38}$$

which upon substitution in Eqs. (32) gives an eigenvalue problem for a. We write the characteristic equation as

$$\left|\left[a(x_{12})-\hat{\sigma}_i-\Big(\sum_{l\neq i}\bar{\sigma}_{il}(0)\Big)\right]\delta_{ik}+(1-\delta_{ik})\bar{\sigma}_{ik}(x_{12})\right|=0, \tag{39}$$

where

$$\hat{\sigma}_i(x_{12})=\bar{\sigma}_{ii}(0)-\bar{\sigma}_{ii}(x_{12}). \tag{40}$$

In general, Eq. (39) has N roots, N being equal to the number of propagating modes. We denote the roots by $a^{(u)}(x_{12})$ and assume them to be distinct. Nondistinct roots introduce no conceptual difficulties. Associated with each characteristic value, $a^{(u)}(x_{12})$, is a characteristic vector $\gamma_i^{(u)}(x_{12})$ obtained in the usual manner. The symmetry of the $\bar{\sigma}_{ik}(x_{12})$ provides an orthogonality relationship for the characteristic vectors. We make their definition unique by a normalization prescription.

Two limiting propagation experiments can be identified, a coherence dominated experiment, defined by the condition that $(\hat{\sigma}_i-\hat{\sigma}_j)\gg\bar{\sigma}_{ij}^2$, and a modal intensity distribution dominated experiment, defined by the reverse condition that $(\hat{\sigma}_i-\hat{\sigma}_j)^2\ll\bar{\sigma}_{ij}^2$. For a coherence dominated experiment, the determinant of Eq. (39) is approximated by one that is diagonal, leading to characteristic values

$$a^{(u)}(x_{12})=\bar{\sigma}_{uu}(0)-\bar{\sigma}_{uu}(x_{12})+\sum_{l\neq u}\bar{\sigma}_{ul}(0), \tag{41}$$

and characteristic vectors

$$\gamma_i^{(u)}(x_{12})=\delta_{iu}. \tag{42}$$

The first term on the rhs of Eq. (41) gives the rate of intramodal scatter, scatter from the u mode; the second gives the separation distance over which this intramodal scatter correlates; the third gives the rate of intermodal scatter out of the u mode. The first two terms are obtained for an uncoupled mode theory. The third term can be interpreted in terms of an apparent dissipation mechanism. For a modal intensity distribution dominated experiment, the characteristic equation can be approximated by

$$\left|\left[a(x_{12})-\sum_{l\neq i}\bar{\sigma}_{il}(0)\right]\delta_{ik}+(1-\delta_{ik})\bar{\sigma}_{ik}(x_{12})\right|=0. \tag{43}$$

All experiments are modal intensity distribution dominated experiments for small enough x_{12}. For $x_{12}=0$, we have

$$\left|\left[a(0)-\sum_{l\neq i}\bar{\sigma}_{il}(0)\right]\delta_{ik}+(1-\delta_{ik})\bar{\sigma}_{ik}(0)\right|=0. \tag{44}$$

Equation (44) can be directly obtained from the equations governing the modal intensity functions, Eqs. (33).

Equation (44) has one root given by $a^{(1)}(0)=0$. This is consistent with the energy conserving nature of the theory. The components of the characteristic vector associated with this root are equal. This leads to the conclusion that the continual scattering of energy will tend to an equal distribution of acoustic intensity among all the propagating wave modes. While this conclusion is intuitively satisfying it would appear to be of little physical significance since the range required for the uniform distribution of energy is expected to be far greater than those in practical ocean acoustic experiments. Indeed, we suspect that such ranges are comparable to those at which coupling between propagating and nonpropagating wave modes is significant. The theory will fail before such ranges are reached.

3. APPLICATION OF THEORY TO PROPAGATION EXPERIMENTS

In discussing the applicability of the theory to a realistic propagation experiment it is natural to first look at the several conditions introduced in the derivation. For example, Eqs. (17), (18), (23), (25), and (37). Appearing in some of these equations are a number of environmental parameters and the question of estimating these parameters for a particular medium and acoustic experiment arises. Further, the tremendous range of size scales for measuring sound speed fluctuations in the ocean, from meters or less up to several tens of kilometers, would almost invariably indicate that the anticipated experiment is beyond the range of validity of this as well as any other coherence theory. It is important to realize, however, that experience has shown the presence of a selection rule in most coherence experiments. Thus, although the range of environmental size scales may be tremendous, usually a much narrower range is actually operative in a given experiment. Determining which of the size scales is of importance in a given experiment is a task requiring some experience with the problem. Using a scattering theory such as the one developed here in a "self-consistent" manner can be useful in this regard. In this manner, for example, the theory of Beran and McCoy would indicate that the most important horizontal size scales for controlling the horizontal coherence of moderately low frequency acoustic experiments (of the order of a few hundred Hertz) over ranges of the order of several hundreds of kilometers are of the order of a kilometer, with the important range of sizes covering somewhere between one and two orders of magnitude.[13] This can be compared to the four or five orders of magnitude that one might observe from environmental data.

The theory is equally applicable to the optics problem and to the acoustics problem discussed in the Introduction. This is in opposition to the parabolic formulations, for which the highly anisotropic theory developed for ocean acoustic experiments is distinctly different from the isotropic theory developed for atmospheric laser experiments. Some explanation of this opposition is, therefore, in order. The explanation is that the manifest differences of the parabolic formulations would exhibit themselves in our normal mode formulation in the next step in applying the formulation, namely, reducing it to a computational algorithm. Then, for efficiency, it would be important to consider the sizes of the various off-diagonal terms of the determinant in Eq. (43). As noted in deriving the theory, the $\bar{\sigma}_{ik}$ terms will, in general, be greatest for $i=k\pm 1$ falling off with increasing separation of the integers i and k. That is, the intermodal scattering by the large size scale fluctuations is confined to neighboring modes. Thus, the nonzero elements of the determinant are banded about the main diagonal. In this the isotropic and anisotropic theories will be the same. When they differ is that, for an isotropic medium $\bar{\sigma}_{ik}$ is relatively insensitive to the absolute values of i and k. That is, the rate of energy transfer between adjacent modes is relatively insensitive to the particular pair of modes. For the highly anisotropic medium, on the other hand, $\bar{\sigma}_{ik}$ is strongly sensitive to the absolute values of i and k, as can be clearly demonstrated, from a single scatter treatment of the local nature of the scattering.[8] The rate of transfer between adjacent modes is very much greater for the lower order wave modes. Thus, for the anisotropic medium the width of the band of nonzero elements about the main diagonal will exhibit a very sharp shrinkage for $i\approx k$ values above some cut-off value. As a consequence the energy in the lower order wave modes will rapidly be scattered into other lower order wave modes with the continual scatter up the spectrum, to higher order wave modes, occurring but only at a much reduced rate. Indeed, the rate of decrease was shown in Ref. 8 to be so great that to an observer attuned to the more rapid rate of transfer among the lower order wave modes the continued transfer would appear to be blocked. A saturation of sorts would appear to occur with intermodal scatter rapidly spreading to all of the lower order wave modes and then stopping. This saturation is consistent with the highly anisotropic parabolic formulation.

ACKNOWLEDGMENTS

The work was carried out during the tenure of G. R. Sutton as a Royal Australian Navy Postgraduate Scholar at the Catholic University of America. The contributions of J. J. McCoy were supported by an Army Research Office grant numbered DAHC04-75-G-0069. These supports are greatly appreciated.

*Present address: School of Mathematical Sciences, The New South Wales Institute of Technology, N.S.W. 2007, Australia.

[1]M. Isakovich, Sov. Phys. Acoust. 3, 35 (1957).
[2]J. Samuels, J. Acoust. Soc. Am. 31, 319 (1959).
[3]A. Lapin, Sov. Phys. Acoust. 15, 490 (1970).
[4]V. Bezrodnyi and J. Fuks, Sov. Phys. Acoust. 17, 444 (1972).
[5]A. Nayfeh, J. Acoust. Soc. Am. 56, 3 (1974).
[6]I. Tolstoy and C. Clay, *Ocean Acoustics* (McGraw-Hill, New York, 1966).
[7]D. Marcuse, *Theory of Dielectric Optical Waveguides* (Academic, New York, 1974).
[8]G.R. Sutton and J.J. McCoy, J. Acoust. Soc. Am. 60, 833–9 (1976).
[9]M.J. Beran, J. Opt. Soc. Am. 56, 1475 (1966).
[10]M.J. Beran and J.J. McCoy, J. Math. Phys. 15, 11, 1901 (1974).
[11]M.J. Beran and J.J. McCoy, J. Acoust. Soc. Am. 56, 6, 1667 (1974).
[12]M.J. Beran and J.J. McCoy, Abstr. J. Acoust. Soc. Am. 58, S1, 551 (1975); J. Math. Phys. 17, 7, 1186 (1976).
[13]M.J. Beran, J.J. McCoy, and B.B. Adams, Naval Research Laboratory Report No. 7809 (1975).

Painlevé functions of the third kind*

Barry M. McCoy and Craig A. Tracy

The Institute for Theoretical Physics, State University of New York at Stony Brook, Stony Brook, New York 11794

Tai Tsun Wu

Gordon McKay Laboratory, Harvard University, Cambridge, Massachusetts 02138
(Received 4 October 1976)

We explicitly construct the one-parameter family of solutions, $\eta(\theta;\nu,\lambda)$, that remain bounded as $\theta\to\infty$ along the positive real θ axis for the Painlevé equation of third kind

$$ww''=(w')^2-\theta^{-1}ww'+2\nu\theta^{-1}(w^3-w)+w^4-1,$$

where, as $\theta\to\infty$, $\eta\sim 1-\lambda\Gamma(\nu+1/2)2^{-2\nu}\theta^{-\nu-1/2}e^{-2\theta}$. We further construct a representation for $\psi(t;\nu,\lambda)=-\ln[\eta(t/2;\nu,\lambda)]$, where $\psi(t;\nu,\lambda)$ satisfies the differential equation

$$\psi''+t^{-1}\psi'=(1/2)\sinh(2\psi)+2\nu t^{-1}\sinh(\psi).$$

The small-θ behavior of $\eta(\theta;\nu,\lambda)$ is described for $|\lambda|<\pi^{-1}$ by

$$\eta(\theta;\nu,\lambda)\sim 2^{\sigma}B\theta^{\sigma}.$$

The parameters σ and B are given as explicit functions of λ and ν. Finally an identity involving the Painlevé transcendent $\eta(\theta;\nu,\lambda)$ is proved. These results for the special case $\nu=0$ and $\lambda=\pi^{-1}$ make rigorous the analysis of the scaling limit of the spin–spin correlation function of the two-dimensional Ising model.

I. INTRODUCTION

The Painlevé equation of the third kind is

$$w''=\frac{1}{w}(w')^2-\frac{1}{\theta}w'+\frac{1}{\theta}(\alpha w^2+\beta)+\gamma w^3+\frac{\delta}{w}, \tag{1.1}$$

where prime denotes differentiation with respect to the variable θ and α, β, γ, and δ are constants. The importance of (1.1) in the theory of ordinary differential equations was first discussed by Painlevé[1] and later by Gambier.[2]

In this paper we develop the theory for the one-parameter, bounded (as $\theta\to\infty$ along the positive real axis), solutions of (1.1) when the constants α, β, γ, and δ satisfy

$$\alpha(-\delta)^{1/2}+\beta(\gamma)^{1/2}=0. \tag{1.2}$$

Under the assumption (1.2) there is no loss in generality if we consider in place of (1.1) the equation

$$w''=\frac{1}{w}(w')^2-\frac{1}{\theta}w'+\frac{2\nu}{\theta}(w^2-1)+w^3-\frac{1}{w}, \tag{1.3}$$

where ν is a constant.

If we denote by $\eta(\theta;\nu,\lambda)$ the one-parameter family of solutions of (1.3) that remain bounded as θ approaches infinity along the positive real axis, we shall prove

Theorem 1: The function $\eta(\theta;\nu,\lambda)$ satisfies (1.3) and for sufficiently large, positive θ and $\mathrm{Re}\nu>-\frac{1}{2}$, $\eta(\theta;\nu,\lambda)$ has the representation

$$\frac{1-\eta(\theta;\nu,\lambda)}{1+\eta(\theta;\nu,\lambda)}=G(t;\nu,\lambda), \tag{1.4a}$$

$$t=2\theta, \tag{1.4b}$$

where

$$G(t;\nu,\lambda)=\sum_{n=0}^{\infty}\lambda^{2n+1}g_{2n+1}(t;\nu), \tag{1.5}$$

$$g_1(t;\nu)=\int_1^{\infty}dy\,\frac{\exp(-ty)}{(y^2-1)^{1/2}}\left(\frac{y-1}{y+1}\right)^{\nu}, \tag{1.6a}$$

and for $n\geq 1$

$$g_{2n+1}(t;\nu)=(-1)^n\int_1^{\infty}dy_1\cdots\int_1^{\infty}dy_{2n+1}\left[\prod_{j=1}^{2n+1}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^{\nu}\right]\times\left[\prod_{j=1}^{2n}(y_j+y_{j+1})^{-1}\right]\left[\prod_{j=1}^{n}(y_{2j}^2-1)\right]. \tag{1.6b}$$

The parameter λ is subject only to the condition $|\lambda|<R(t)$ where $R(t)$ is the radius of convergence of (1.5) viewed in the complex λ plane. Simple bounds on $R(t)$ follow from the inequalities of Ref. 3, Eqs. (3.156)–(3.159). The restriction $\mathrm{Re}\nu>-\frac{1}{2}$ can be lifted in (1.6) by first changing the contour of integration to the contour $\mathcal{C}$ which is the contour beginning at infinity and looping around the branch point at $y=1$. The additional factor $\sin\pi(\nu-\frac{1}{2})$ can be incorporated into λ.

It is an important feature concerning the theory of the function $\eta(\theta;\nu,\lambda)$ that if we define $\psi(t;\nu,\lambda)$ by

$$\text{(i)}\quad \eta(\theta;\nu,\lambda)=\exp[-\psi(t;\nu,\lambda)],\quad t=2\theta,$$
$$\text{(ii)}\quad \psi(t;\nu,\lambda)\to 0\quad\text{as}\quad t\to+\infty, \tag{1.7}$$

then we have

Theorem 2: For t sufficiently large and $\mathrm{Re}\nu>-\frac{1}{2}$

$$\psi(t;\nu,\lambda)=\sum_{n=0}^{\infty}\lambda^{2n+1}\psi_{2n+1}(t;\nu), \tag{1.8}$$

where we have

$$\psi_1(t;\nu)=2g_1(t;\nu) \tag{1.9a}$$

and for $n\geq 1$

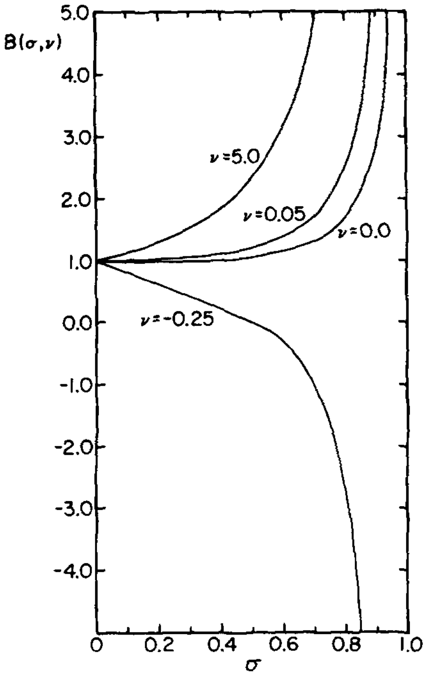

FIG. 1. Quantity $B(\sigma,\nu)$ as a function of σ for various values of ν. The slope of $B(\sigma,\nu)$ at $\sigma=0$ is $2\gamma+\ln 2+\psi(\frac{1}{2}+\nu)$. For $\nu=\nu^*\simeq 0.0245$ the slope is zero. For $\nu>\nu^*$ ($<\nu^*$) the slope at the origin is positive (negative). For $\nu=0$ the minimum of $B(\sigma,\nu)$ occurs at $\sigma\simeq 0.23$. The scale of the figure is too large to see the $B(\sigma,\nu)<1$ behavior for $\nu\geq 0$. For $\nu<0$ $B(\sigma,\nu)$ vanishes at $\sigma=1+2\nu$.

$$\psi_{2n+1}(t;\nu)=\frac{2}{2n+1}\int_1^\infty dy_1\cdots\int_1^\infty dy_{2n+1}\left[\prod_{j=1}^{2n+1}\frac{\exp(-ty_j)}{y_j+y_{j+1}}\right]\times\left[\prod_{j=1}^{2n+1}\left(\frac{y_j-1}{y_j+1}\right)^{\nu-1/2}+\prod_{j=1}^{2n+1}\left(\frac{y_j-1}{y_j+1}\right)^{\nu+1/2}\right] \tag{1.9b}$$

with $y_{2n+2}\equiv y_1$ in (1.9b). Again the restriction $\mathrm{Re}(\nu)>-\frac{1}{2}$ can be lifted by using the contour C. To examine the analytic properties of $\eta(\theta;\nu,\lambda)$ and $\psi(t;\nu,\lambda)$ in the complex λ plane, representation (3.38) is useful.

As emphasized by Painlevé[1] the point $\theta=0$ plays a unique role in the theory of the third Painlevé transcendent. It is the only point in the finite θ plane for which a branch point or an essential singular point of a solution of (1.1) can occur. Furthermore it has been shown[1,2] that if $\theta=0$ is an analytic point, then the solution is a meromorphic function. Thus it is important to examine the behavior of a solution of (1.3) in the neighborhood of $\theta=0$. It is easy to demonstrate that for $t\to 0$ ($t=2\theta$) a formal solution of (1.3) is

$$w(t/2)=Bt^\sigma\{1-\nu B^{-1}(1-\sigma)^{-2}t^{1-\sigma}+B\nu(1+\sigma)^{-2}t^{1+\sigma}+[\tfrac{1}{4}\nu^2B^{-2}(1-\sigma)^{-4}-\tfrac{1}{16}B^{-2}(1-\sigma)^{-2}]t^{2-2\sigma}+O(t^2)\}, \tag{1.10}$$

where $-1<\mathrm{Re}\sigma<1$ but otherwise σ and B are arbitrary.

In general a solution that behaves as (1.10) for $t\to 0$ will not remain bounded as $t\to+\infty$. When $0\leq\lambda<\pi^{-1}$ the bounded solution $\eta(t/2;\nu,\lambda)$ behaves as (1.10) for $t\to 0$ but the coefficients σ and B are now functions of λ and ν. Using Theorem 2 we shall prove

Theorem 3: The solution $\eta(t/2;\nu,\lambda)$ has the small-t expansion (1.10) for $0\leq\lambda<1/\pi$ where

$$\sigma=\sigma(\lambda)=(2/\pi)\arcsin(\pi\lambda) \tag{1.11}$$

and

$$B=B(\sigma,\nu)=2^{-3\sigma}\frac{\Gamma^2((1-\sigma)/2)}{\Gamma^2((1+\sigma)/2)}\frac{\Gamma((1+\sigma)/2+\nu)}{\Gamma((1-\sigma)/2+\nu)}, \tag{1.12}$$

where $\Gamma(x)$ is the gamma function.

In Fig. 1 the function $B(\sigma,\nu)$ is graphed. Using Theorem 3 we can determine the small t behavior of $\eta(t/2;\nu,\lambda)$ for $\lambda\geq\pi^{-1}$ (see Sec. IV.I, also the case $\lambda<0$ is discussed).

We conclude our presentation of the theory of the Painlevé transcendent $\eta(\theta;\nu,\lambda)$ by proving a useful identity ($0\leq\lambda\leq\pi^{-1}$).

Theorem 4: If we define the functions

$$f_{2n}(t;\nu)=\frac{(-1)^n}{n}\int_1^\infty dy_1\cdots\int_1^\infty dy_{2n}\left[\prod_{j=1}^{2n}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\times\left(\frac{y_j-1}{y_j+1}\right)^\nu\frac{1}{y_j+y_{j+1}}\right]\prod_{j=1}^{n}(y_{2j}^2-1) \tag{1.13}$$

with $y_{2n+1}\equiv y_1$, then we have

$$\tfrac{1}{2}[1+\eta(\theta;\nu,\lambda)]\eta^{-1/2}(\theta;\nu,\lambda)\exp\left[\int_\theta^\infty dx\left\{\tfrac{1}{4}x\eta^{-2}(x;\nu,\lambda)\times[(1-\eta^2(x;\nu,\lambda))^2-(\eta'(x;\nu,\lambda))^2]+\frac{\nu}{2\eta(x;\nu,\lambda)}(1-\eta(x;\nu,\lambda))^2\right\}\right]=\exp\left[-\sum_{n=1}^\infty\lambda^{2n}f_{2n}(2\theta;\nu)\right], \tag{1.14a}$$

where prime denotes differentiation with respect to x. Using definition (1.7) of the function $\psi(t;\nu,\lambda)$ the above identity becomes

$$\cosh\tfrac{1}{2}\psi(t;\nu,\lambda)\exp\left\{\tfrac{1}{4}\int_t^\infty ds\,s\left[-\left(\frac{d\psi}{ds}\right)^2+\sinh^2\psi+\frac{4\nu}{s}\sinh^2\tfrac{1}{2}\psi\right]\right\}=\exp\left[-\sum_{n=1}^\infty\lambda^{2n}f_{2n}(t;\nu)\right], \tag{1.14b}$$

where all ψ functions appearing under the integral sign are functions of s, ν, and λ.

Theorems 1, 2, 3, and 4 are proved and discussed in Secs. II, III, IV, and V, respectively.

For the special case $\nu=0$ and $\lambda=\pi^{-1}$ these four theorems make rigorous the analysis of the scaling limit of the spin–spin correlation function of the two-dimensional Ising model carried out by Wu, McCoy, Tracy, and Barouch.[3] It is perhaps not inappropriate to describe in some detail how the above theorems fit into the work of Ref. 3. However it should be stressed that the remainder of this section is irrelevant for the mathematical discussion that follows in Secs. II—V.

If we denote by ξ the correlation length $[\xi=\xi(T)$, $T=$ temperature, and $\xi\to\infty$ as $T\to T_c^{\pm}$ where T_c is the critical temperature] and by $\langle\sigma_{0,0}\sigma_{M,N}\rangle$ the spin—spin correlation function for the two-dimensional Ising model on a square lattice, and if we further assume for simplicity of presentation that the vertical and horizontal interaction energies are equal, then by *scaling limit* we mean that limit

$$\xi\to\infty,\quad R\equiv(M^2+N^2)^{1/2}\to\infty \tag{1.15a}$$

such that

$$t=R/\xi \text{ is fixed.} \tag{1.15b}$$

In this limit the correlation function $\langle\sigma_{0,0}\sigma_{M,N}\rangle$ becomes[3]

$$\langle\sigma_{0,0}\sigma_{M,N}\rangle=R^{-1/4}F_{\pm}(t)+R^{-5/4}F_{1\pm}(t)+o(R^{-5/4}), \tag{1.16}$$

where $F_{\pm}(t)$ and $F_{1\pm}(t)$ (these are commonly called *scaling functions*) are functions of the single variable t defined by (1.15b).

In Secs. III and IV of Ref. 3 an expansion valid for large t was developed [these results are summarized by Eqs. (2.26)—(2.30) of Ref. 3]. The expansion for $F_-(t)$ is the right-hand side of (1.14a) of Theorem 4 (for $\lambda=\pi^{-1}$ and $\nu=0$) times the factor $(2t)^{1/4}(\sinh 2\beta_c E_1+\sinh 2\beta_c E_2)^{1/8}$. The expansion for $F_+(t)/F_-(t)$ is the right-hand side of (1.5) of Theorem 1 (for $\nu=0$ and $\lambda=\pi^{-1}$). These infinite series expansions are only useful for large t. For small t the functions $g_{2n+1}(t;\nu)$ of Theorem 1 behave as

$$g_{2n+1}(t;\nu)=c_{2n+1}(\ln t)^{2n+1}+c_{2n}(\ln t)^{2n}+\cdots+c_1(\ln t)+c_0+o(1)\quad(t\to 0) \tag{1.17}$$

and similarly for the functions $f_{2n}(t;\nu)$ of Theorem 4.

Therefore, to study the small-t behavior of $F_{\pm}(t)$ the representation of $F_{\pm}(t)$ as an infinite series of multiple integrals is not directly the most convenient representation. This representation of $F_{\pm}(t)$ as an infinite series of multiple integrals can be thought of as the coordinate space analog of the dispersion integral representation of the two-point function. What is needed is a way to sum up this dispersion integral representation.

In Ref. 3 this was accomplished in two ways. One way (that of Sec. V) was to develop a separate perturbation scheme valid for small t. The other method (that of Sec. VI of Ref. 3) was to introduce an integral equation that could be solved in terms of Painlevé functions. This approach led to the representation of $F_{\pm}(t)$ in terms of Painlevé functions [these results are summarized by Eq. (2.39) of Ref. 3]. In terms of Theorem 1 of this paper $F_+(t)/F_-(t)$ was shown to be the left-hand side of (1.4a) for $\nu=0$ and $\lambda=\pi^{-1}$ and in terms of Theorem 4 $F_-(t)$ was shown to be the left-hand side of (1.14a) for $\lambda=\pi^{-1}$ and $\nu=0$ times the factor $(2t)^{1/4}(\sinh 2\beta_c E_1+\sinh 2\beta_c E_2)^{1/8}$. The methods used in Sec. VI of Ref. 3 though correct are not rigorous. Theorem 1 of this paper *rigorously* proves that the infinite series representation of $F_+(t)/F_-(t)$ is simply related to Painlevé functions, and likewise Theorem 4 rigorously relates the infinite series representation of $F_-(t)$ to Painlevé functions. Stated somewhat crudely, the Painlevé transcendents $\eta(\theta;\nu,\lambda)$ are the functions that sum the dispersion integral representation of the two-point functions $F_{\pm}(t)$.

In light of Theorems 1 and 4 the small-t behavior of $F_{\pm}(t)$ follows once the small-t behavior of $\eta(t/2;0,\pi^{-1})$ is known. To determine this behavior the analysis of Ref. 3 had to make crucial use of the unpublished thesis of Myers[4] where Painlevé functions of the third kind arose in the study of scattering from a strip. Though Myers' analysis is rigorous it gives only the small-t behavior of $\eta(t/2;\nu,\lambda)$ for the case $\nu=0$ and $\lambda=\pi^{-1}$. Theorem 3 gives a direct proof (that is, the scattering problem is avoided) of the small-t behavior of $\eta(t/2;\nu,\lambda)$. Theorem 2 is essential to prove Theorem 3.

II. THEOREM 1 AND THE FUNCTION $G(t;\nu,\lambda)$

A. Restricted Painlevé equation of third kind

The most general Painlevé equation of the third kind is given by (1.1) where the constants α, β, γ, and δ are arbitrary. If we assume that the constants α, β, γ, and δ are restricted so that (1.2) is satisfied, then (1.1) can be reduced to (1.3). To demonstrate this we let

$$w(z)=A\eta(\theta),\quad \theta=Bz, \tag{2.1}$$

where z denotes the independent variable in (1.1), and A and B are constants that are to be determined. Substituting (2.1) into (1.1) it follows that

$$\frac{d^2\eta}{d\theta^2}=\frac{1}{\eta}\left(\frac{d\eta}{d\theta}\right)^2-\frac{1}{\theta}\frac{d\eta}{d\theta}+\frac{\alpha A}{B}\frac{1}{\theta}\eta^2+\frac{\beta}{AB}\frac{1}{\theta}+\frac{\gamma A^2}{B^2}\eta^3+\frac{\delta}{A^2B^2}\frac{1}{\eta}. \tag{2.2}$$

This equation is of the form (1.3) if we have

$$\frac{\alpha A}{B}=-\frac{\beta}{AB}=2\nu \tag{2.3a}$$

and

$$\frac{\gamma A^2}{B^2}=-\frac{\delta}{A^2B^2}=1. \tag{2.3b}$$

From (2.3b) we see A and B are given by

$$A^2=(-\delta/\gamma)^{1/2},\quad B^2=(-\delta\gamma)^{1/2}. \tag{2.4}$$

In order that (2.3a) is satisfied we demand

$$2\nu=\alpha/(\gamma)^{1/2}=-\beta/(-\delta)^{1/2} \tag{2.5}$$

which is just (1.2).

The condition (1.2) arises naturally in the following context. In general if $w(\theta)$ is a solution to (1.1), then $[Aw(\theta)]^{-1}$ is a solution to (1.1) with *different* α, β, γ, and δ (here A is a constant). If we demand that $[Aw(\theta)]^{-1}$ is a solution for the *same* α, β, γ, and δ, then A is fixed and the parameters α, β, γ, and δ must satisfy (1.2). From now on we discuss only (1.3).

B. Function $G(t;\nu,\lambda)$

As stated in the Introduction we denote by $\eta(\theta;\nu,\lambda)$ the one-parameter family of bounded (as $\theta\to\infty$ along the positive real θ axis) solutions to (1.3). We associate with $\eta(\theta;\nu,\lambda)$ the function $G(t;\nu,\lambda)$ where

$$G(t;\nu,\lambda)=\frac{1-\eta(\theta;\nu,\lambda)}{1+\eta(\theta;\nu,\lambda)} \tag{2.6a}$$

and

$$t=2\theta. \tag{2.6b}$$

From (1.3) and (2.6) it follows that $G(t;\nu,\lambda)$ satisfies the differential equation

$$G''+\frac{1}{t}G'-\left(1+\frac{2\nu}{t}\right)G=G''G^2-2(G')^2G+\frac{1}{t}G'G^2+G^3-\frac{2\nu}{t}G^3, \tag{2.7}$$

where the prime denotes differentiation with respect to the variable t.

Theorem 1 states that the one-parameter bounded solutions to (2.7) are given by (1.5) and (1.6). It is the goal of this section to prove Theorem 1. The method of proof is to substitute (1.5)–(1.6) into (2.7) and explicitly demonstrate that this is indeed a solution.

We begin the proof of Theorem 1 by establishing some useful identities which we state as lemmas.

C. Preliminary lemmas

Lemma 2.1: A necessary and sufficient condition that $G(t;\nu,\lambda)$ as defined by (1.5)–(1.6) satisfy (2.7) is for $k=0,1,2,\cdots$,

$$\begin{aligned}g''_{2k+1}&+\frac{1}{t}g'_{2k+1}-\left(1+\frac{2\nu}{t}\right)g_{2k+1}\\&=\sum_{l=0}^{k-1}\sum_{m=0}^{k-l-1}\Big\{g_{2l+1}\Big[g''_{2(k-l-m)-1}+\frac{1}{t}g'_{2(k-l-m)-1}\\&\quad-\left(1+\frac{2\nu}{t}\right)g_{2(k-l-m)-1}\Big]g_{2m+1}+2g_{2(k-l-m)-1}[g_{2l+1}g_{2m+1}\\&\quad-g'_{2l+1}g'_{2m+1}]\Big\}\end{aligned} \tag{2.8}$$

where $g_{2n+1}(t;\nu)$ are defined by (1.6) and for $k=0$ the right-hand side of (2.8) is defined to be zero.

Proof: Since for $t>0$ $G(t;\nu,\lambda)$ has a finite radius of convergence in the λ plane we are allowed to equate equal powers of λ when (1.5) is substituted into (2.7). The precise form of the right-hand side of (2.8) follows by simple manipulations of power series. Clearly if (2.8) is true, then multiplication of this equation by λ^{2k+1} and summing over k reproduces (2.7).

If we define $g_{2n+1}(t;\nu)$, $n=0,1,2,\cdots$ by (1.6), then an alternate representation of these functions for $n=1,2,\cdots$ is

Lemma 2.2:

$$\begin{aligned}g_{2n+1}(t;\nu)&=(-1)^n\int_1^\infty dy_1\cdots\int_1^\infty dy_{2n+1}\\&\times\left[\prod_{j=1}^{2n+1}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^\nu\right]\\&\times\left[\prod_{j=1}^{2n}(y_j+y_{j+1})^{-1}\right](y_1y_{2n+1}-1)\prod_{j=2}^{n}(y_{2j-1}^2-1),\end{aligned} \tag{2.9}$$

where for $n=1$ the last product is replaced by unity.

Proof: (i) $n=1$ case

From (1.6) we have

$$\begin{aligned}g_3(t;\nu)=(-1)\int_1^\infty dy_1\int_1^\infty dy_2\int_1^\infty dy_3\left[\prod_{j=1}^{3}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\right.\\\left.\times\left(\frac{y_j-1}{y_j-1}\right)^\nu\right]\frac{y_2^2-1}{(y_1+y_2)(y_2+y_3)}.\end{aligned} \tag{2.10}$$

If we cyclically permute the integration variable labels in (2.10), then we can write $g_3(t;\nu)$ as

$$\begin{aligned}g_3(t;\nu)=-\tfrac{1}{3}\int_1^\infty dy_1\int_1^\infty dy_2\int_1^\infty dy_3\left[\prod_{j=1}^{3}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\right.\\\left.\times\left(\frac{y_j-1}{y_j+1}\right)^\nu\frac{1}{y_j+y_{j+1}}\right][(y_2^2-1)(y_3+y_1)\\+(y_3^2-1)(y_1+y_2)+(y_1^2-1)(y_2+y_3)]\end{aligned} \tag{2.11}$$

with $y_4\equiv y_1$.

The quantity in the second square brackets in (2.11) can be written as

$$\begin{aligned}&(y_2^2-1)(y_3+y_1)+(y_3^2-1)(y_1+y_2)+(y_1^2-1)(y_2+y_3)\\&\quad=(y_1+y_2)(y_1y_2-1)+(y_1+y_3)(y_1y_3-1)\\&\qquad+(y_2+y_3)(y_2y_3-1).\end{aligned} \tag{2.12}$$

Using this in (2.11) and writing the three resulting terms as one term (again by cyclically permuting the labels of the integration variables) we obtain

$$\begin{aligned}g_3(t;\nu)=-\int_1^\infty dy_1\int_1^\infty dy_2\int_1^\infty dy_3\left[\prod_{j=1}^{3}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\right.\\\left.\times\left(\frac{y_j-1}{y_j+1}\right)^\nu\right]\frac{y_1y_3-1}{(y_1+y_2)(y_2+y_3)}\end{aligned} \tag{2.13}$$

which is (2.9) for $n=1$.

(ii) $n=2$ case

For $n=2$ the part of the integrand in (1.6b) that is not invariant under cyclic permutations of the integration variable labels is

$$(y_2^2-1)(y_4^2-1)(y_5+y_1). \tag{2.14}$$

Under the five cyclic permutations of the labels (1,2,3,4,5) the quantity (2.14) becomes the sum of five terms, viz.

$$(y_2^2-1)(y_4^2-1)(y_5+y_1)+(y_3^2-1)(y_5^2-1)(y_1+y_2)$$
$$+(y_4^2-1)(y_1^2-1)(y_2+y_3)+(y_5^2-1)(y_2^2-1)(y_3+y_4)$$
$$+(y_1^2-1)(y_3^2-1)(y_4+y_5)$$
$$=(y_3^2-1)(y_5y_1-1)(y_5+y_1)+(y_4^2-1)(y_1y_2-1)(y_1+y_2)$$
$$+(y_5^2-1)(y_2y_3-1)(y_2+y_3)+(y_1^2-1)(y_3y_4-1)(y_3+y_4)$$
$$+(y_2^2-1)(y_4y_5-1)(y_4+y_5). \tag{2.15}$$

This can be written more compactly as

$$(y_2^2-1)(y_4^2-1)(y_5+y_1)+\text{cyclic permutations}$$
$$=(y_3^2-1)(y_5y_1-1)(y_5+y_1)+\text{cyclic permutations}. \tag{2.16}$$

If (2.16) is used in (1.6b) for $n=2$ we obtain (2.9) for $n=2$.

(iii) General case

We write integrand of (1.6b) as

$$\left[\prod_{j=1}^{2n+1}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^{\nu}(y_j+y_{j+1})^{-1}\right]$$
$$\times(y_{2n+1}+y_1)\prod_{j=1}^{n}(y_{2j}^2-1), \tag{2.17}$$

where $y_{2n+2}\equiv y_1$. The quantity in square brackets in (2.17) is invariant under cyclic permutations of the integration variable labels. We claim that

$$\prod_{j=1}^{n}(y_{2j}^2-1)(y_{2n+1}+y_1)+\text{cyclic perm.}$$
$$=\prod_{j=2}^{n}(y_{2j-1}^2-1)(y_1y_{2n+1}-1)(y_{2n+1}+y_1)+\text{cyclic perm.} \tag{2.18}$$

From (2.18) the result (2.9) follows. To demonstrate (2.18) we first examine that piece of the left-hand side of (2.18) which is of degree $(2n+1)$. There are $2(2n+1)$ such terms and they are of the form

$$y_2^2y_4^2\cdots y_{2n}^2y_{2n+1}+y_2^2y_4^2\cdots y_{2n}^2y_1+\text{cyclic perm.} \tag{2.19}$$

This can be rewritten as

$$y_3^2y_5^2\cdots y_{2n+1}^2y_1+y_1^2y_3^2\cdots y_{2n-1}^2y_{2n+1}+\text{cyclic perm.}$$
$$=y_3^2y_5^2\cdots y_{2n-1}^2y_1y_{2n+1}(y_1+y_{2n+1})+\text{cyclic perm.} \tag{2.20}$$

Now consider the terms of (2.18) that are of degree $2n-1$. These terms arise by replacing some y_{2j}^2 in (2.19) by -1 or y_1y_{2n+1} by -1. This can be done at n places. Thus the second term of the left-hand side of (2.18) is

$$-[y_3^2y_5^2\cdots y_{2n-3}^2y_{2n+1}y_1+y_3^2y_5^2\cdots y_{2n-5}^2y_{2n-1}^2y_{2n+1}y_1+\cdots$$
$$+y_5^2y_7^2\cdots y_{2n-1}^2+y_3^2y_5^2\cdots y_{2n-1}^2](y_{2n+1}+y_1)$$
$$+\text{cyclic perm.} \tag{2.21}$$

If one compares (2.20) and (2.21) with the right-hand side of (2.18), then one sees both of these terms are present. The third term comes from leaving out an additional y_{2j-1}^2 or y_1y_{2n+1}, a term which is again clearly present on the right-hand side of (2.18). Continuing so, we see that the lemma is proved.

Our final lemma is

Lemma 2.3:

$$g''_{2k+1}(t;\nu)+\frac{1}{t}g'_{2k+1}(t;\nu)-\left(1+\frac{2\nu}{t}\right)g_{2k+1}(t;\nu)$$
$$=2(-1)^k\int_1^\infty dy_1\cdots\int_1^\infty dy_{2k+1}$$
$$\times\left[\prod_{j=1}^{2k+1}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^{\nu}\right]$$
$$\times\left[\prod_{j=1}^{2k}(y_j+y_{j+1})^{-1}\right](y_1y_{2k+1}-1)\prod_{j=2}^{k}(y_{2j-1}^2-1)$$
$$\times\sum_{l=0}^{k-1}\sum_{m=0}^{k-l-1}(y_{2l+1}+y_{2l+2})(y_{2k-2m}+y_{2k+1-2m}) \tag{2.22}$$

with $k=1,2,3,\cdots$ and $\prod_{j=2}^{k}(y_{2j-1}^2-1)$ is defined to be unity for $k=1$.

Proof: For notational convenience we denote by L_ν the differential operator

$$L_\nu=\frac{d^2}{dt^2}+\frac{1}{t}\frac{d}{dt}-\left(1+\frac{2\nu}{t}\right). \tag{2.23}$$

From Lemma 2.2 we have

$$L_\nu g_{2k+1}(t;\nu)$$
$$=(-1)^k\int_1^\infty dy_1\cdots\int_1^\infty dy_{2k+1}$$
$$\times\prod_{j=1}^{2k+1}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^{\nu}\prod_{j=1}^{2k}(y_j+y_{j+1})^{-1}(y_1y_{2k+1}-1)$$
$$\times\prod_{j=2}^{k}(y_{2j-1}^2-1)\left[(y_1+y_2+\cdots+y_{2k+1})^2\right.$$
$$\left.-\frac{1}{t}(y_1+y_2+\cdots+y_{2k+1})-\left(1+\frac{2\nu}{t}\right)\right]. \tag{2.24}$$

We now proceed to integrate by parts the $1/t$ terms in (2.24). We first note the following identities:

$$\frac{y_{2j-1}+2\nu}{(y_{2j-1}^2-1)^{1/2}}\left(\frac{y_{2j-1}-1}{y_{2j-1}+1}\right)^{\nu}dy_{2j-1}$$
$$=d\left[\frac{1}{(y_{2j-1}^2-1)^{1/2}}\left(\frac{y_{2j-1}-1}{y_{2j-1}+1}\right)^{\nu}(y_{2j-1}^2-1)\right] \tag{2.25a}$$

and

$$\frac{y_{2j}-2\nu}{(y_{2j}^2-1)^{3/2}}\left(\frac{y_{2j}-1}{y_{2j}+1}\right)^{\nu}dy_{2j}$$
$$=-d\left[\frac{1}{(y_{2j}^2-1)^{1/2}}\left(\frac{y_{2j}-1}{y_{2j}+1}\right)^{\nu}\right]. \tag{2.25b}$$

We write the $1/t$ terms in the integrand of (2.24) as

$$-\frac{1}{t}(y_1+2\nu)-\frac{1}{t}(y_2-2\nu)-\frac{1}{t}(y_3+2\nu)-\cdots-\frac{1}{t}(y_{2k+1}+2\nu). \tag{2.26}$$

The $1/t$ part of (2.24) in view of (2.26) is a sum of $2k+1$ terms. Each term is a $(2k+1)$-dimensional integral. We integrate by parts a single integral of each of these multidimensional integrals. The term we choose to integrate by parts is the term with the structure of (2.25). We integrate the factors according to (2.25) and differentiate the remaining multiplicative factors. The differentiation creates terms of two classes. One class of terms will not contain a $1/t$ factor (these terms come from differentiating the exponential factor which brings down a t factor canceling the $1/t$ factor in front) and the other class will contain an overall $1/t$ factor. We denote by $[L_\nu g_{2k+1}(t;\nu)]_1$ that part of (2.24) which upon integration by parts in the above described manner contains no $1/t$ factors, and by $[L_\nu g_{2k+1}(t;\nu)]_2$ the part that contains the $1/t$ factor. Thus we have

$$L_\nu g_{2k+1}(t;\nu)=[L_\nu g_{2k+1}(t;\nu)]_1+[L_\nu g_{2k+1}(t;\nu)]_2. \tag{2.27}$$

We have, carrying out this integration by parts (all boundary terms vanish),

$$\begin{aligned}&[L_\nu g_{2k+1}(t;\nu)]_1\\&=(-1)^k\int_1^\infty dy_1\cdots\int_1^\infty dy_{2k+1}\\&\times\prod_{j=1}^{2k+1}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^\nu\prod_{j=1}^{2k}(y_j+y_{j+1})^{-1}(y_1y_{2k+1}-1)\\&\times\{(y_1+y_2+\cdots+y_{2k+1})^2-(y_1^2-1)-(y_3^2-1)-\cdots\\&-(y_{2k+1}^2-1)+(y_2^2-1)+(y_4^2-1)+\cdots+(y_{2k}^2-1)-1\}.\end{aligned}\tag{2.28}$$

The last factor in (2.28) can be combined to obtain

$$\begin{aligned}&[L_\nu g_{2k+1}(t;\nu)]_1\\&=2(-1)^k\int_1^\infty dy_1\cdots\int_1^\infty dy_{2k+1}\\&\times\prod_{j=1}^{2k+1}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^\nu\prod_{j=1}^{2k}(y_j+y_{j+1})^{-1}(y_1y_{2k+1}-1)\\&\times\sum_{l=0}^{k-1}\sum_{m=0}^{k-l-1}(y_{2k-2m}+y_{2k-2m+1})(y_{2l+1}+y_{2l+2}).\end{aligned}\tag{2.29}$$

Comparing (2.27) and (2.29) with (2.22) we see that to prove this lemma we must establish

$$[L_\nu g_{2k+1}(t;\nu)]_2=0. \tag{2.30}$$

We have from the integration by parts

$$\begin{aligned}&[L_\nu g_{2k+1}(t;\nu)]_2\\&=\frac{1}{t}(-1)^k\int_1^\infty dy_1\cdots\int_1^\infty dy_{2k+1}\\&\times\prod_{j=1}^{2k+1}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^\nu\left[\sum_{j=0}^{k}(y_{2j+1}^2-1)\frac{d}{dy_{2j+1}}\right.\\&\left.-\sum_{j=1}^{k}\frac{d}{dy_{2j}}(y_{2j}^2-1)\right]\left[\prod_{j=1}^{2k}(y_j+y_{j+1})^{-1}\right.\\&\left.\times(y_{2k+1}y_1-1)\prod_{j=2}^{k}(y_{2j-1}^2-1)\right].\end{aligned}\tag{2.31}$$

Performing the indicated differentiations $[L_\nu g_{2k+1}(t;\nu)]_2$ becomes

$$\begin{aligned}&[L_\nu g_{2k+1}(t;\nu)]_2\\&=\frac{1}{t}(-1)^k\int_1^\infty dy_1\cdots\int_1^\infty dy_{2k+1}\\&\times\prod_{j=1}^{2k+1}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^\nu\prod_{j=1}^{2k}(y_j+y_{j+1})^{-1}\\&\times\prod_{j=2}^{k}(y_{2j-1}^2-1)(y_{2k+1}y_1-1)\left\{\sum_{j=1}^{k-1}(y_{2j+1}^2-1)\right.\\&\times\left[-\frac{1}{y_{2j+1}+y_{2j+2}}-\frac{1}{y_{2j}+y_{2j+1}}+\frac{2y_{2j+1}}{y_{2j+1}^2-1}\right]\\&+(y_1^2-1)\left[-\frac{1}{y_1+y_2}+\frac{y_{2k+1}}{y_{2k+1}y_1-1}\right]+(y_{2k+1}^2-1)\\&\times\left[-\frac{1}{y_{2k}+y_{2k+1}}+\frac{y_1}{y_{2k+1}y_1-1}\right]-\sum_{j=1}^{k}\left[2y_{2j}-(y_{2j}^2-1)\right.\\&\left.\left.\times\left(\frac{1}{y_{2j-1}+y_{2j}}+\frac{1}{y_{2j}+y_{2j+1}}\right)\right]\right\}.\end{aligned}\tag{2.32}$$

We now claim that the term inside the curly brackets in (2.32) is zero. To see this we group the terms in (2.32) with common denominators. Thus the sum of terms that have the denominator (y_1+y_2) is

$$\frac{1}{y_1+y_2}\left[-(y_1^2-1)+(y_2^2-1)\right]=y_2-y_1 \tag{2.33a}$$

and similarly for the other denominator factors:

$$\frac{1}{y_{2j}+y_{2j+1}}\left[-(y_{2j+1}^2-1)+y_{2j}^2-1\right]=y_{2j}-y_{2j+1}, \tag{2.33b}$$

$$\frac{1}{y_{2j+1}+y_{2j+2}}\left[-(y_{2j+1}^2-1)+y_{2j+2}^2-1\right]=y_{2j+2}-y_{2j+1}, \tag{2.33c}$$

and

$$\frac{1}{y_{2k}+y_{2k+1}}\left[-(y_{2k+1}^2-1)+y_{2k}^2-1\right]=y_{2k}-y_{2k+1}. \tag{2.33d}$$

As a result of this combination we see that the term in curly brackets in (2.32) becomes

$$\begin{aligned}&y_2-y_1+\sum_{j=1}^{k-1}(y_{2j}-y_{2j+1})+\sum_{j=1}^{k-1}(y_{2j+2}-y_{2j+1})+y_{2k}-y_{2k+1}\\&+2\sum_{j=1}^{k-1}y_{2j+1}-2\sum_{j=1}^{k}y_{2j}+\left[\frac{y_{2k+1}(y_1^2-1)+y_1(y_{2k+1}^2-1)}{y_1y_{2k+1}-1}\right],\end{aligned}$$

a quantity which is identically zero. Hence (2.30) follows and thus the lemma is proved.

D. Cases $k=0$, $k=1$, $k=2$

The problem is to show that (2.8) holds for all k. For $k=0$ (2.8) reduces to showing

$$L_\nu g_1(t;\nu)=0, \tag{2.34}$$

where L_ν is given by (2.23). That is we want to demonstrate

$$\int_1^\infty dy\, \frac{\exp(-ty)}{(y^2-1)^{1/2}} \left(\frac{y-1}{y+1}\right)^\nu \left[y^2 - \frac{1}{t}y - \left(1+\frac{2\nu}{t}\right)\right] = 0. \tag{2.35}$$

This clearly follows by using (2.25a) in the integration by parts of the $1/t$ term. This result is well known.

For $k=1$ (2.8) reduces to showing

$$L_\nu g_3(t;\nu) = 2g_1(g_1 - g_1'^2). \tag{2.36}$$

From Lemma 2.3 we see that

$$L_\nu g_3(t;\nu) = -2\int_1^\infty dy_1 \int_1^\infty dy_2 \int_1^\infty dy_3 \times\left[\prod_{j=1}^{3} \frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^\nu\right](y_1y_3-1). \tag{2.37}$$

Using the definition (1.6a) of $g_1(t;\nu)$ we see that the right-hand side of (2.36) is precisely (2.37). Hence (2.8) is true for $k=1$.

The case $k=2$ is somewhat more involved. This case along with $k=3$ must have separate proofs from the case of arbitrary k ($\geqslant 4$) as for $k \leqslant 3$ the structure of (2.8) is lacking certain complexities that are present in the general case. This will become apparent as we proceed into the proof.

However certain general comments concerning (2.8) can be made at this point. To prove (2.8) we have found it necessary to put the integrands of the integral representations of the terms appearing in (2.8) into such a form that the integrands contain the same number of denominator factors. By use of Lemma 2.3 we see that $L_\nu g_{2k+1}(t;\nu)$ has $2k-2$ denominator factors in the integrand of its integral representation. This same number of denominator factors occurs in the term

$$2\sum_{l=0}^{k-1}\sum_{m=0}^{k-l-1} g_{2(k-m-l)-1}[g_{2l+1}g_{2m+1} - g'_{2l+1}g'_{2m+1}]$$

which appears in (2.8). However the term

$$\sum_{l=0}^{k-1}\sum_{m=0}^{k-l-1} g_{2l+1}g_{2m+1}L_\nu g_{2(k-m-l)-1}$$

which also appears in (2.8) has only $2k-4$ denominator factors in its integral representation (apply Lemma 2.3 to $L_\nu g_{2(k-m-l)-1}$ and use the definitions of g_{2l+1} and g_{2m+1}). Thus instead of (2.8) we will prove the equivalent identity

$$L_\nu g_{2k+1}(t;\nu) - 2\sum_{l=0}^{k-1}\sum_{m=0}^{k-l-1} g_{2(k-m-l)-1}[g_{2l+1}g_{2m+1} - g'_{2l+1}g'_{2m+1}] = \sum_{l=0}^{k-1}\sum_{m=0}^{k-l-1} g_{2l+1}g_{2m+1}L_\nu g_{2(k-m-l)-1}. \tag{2.38}$$

The key to proving (2.38) will be to write the left-hand side of (2.38) in a form that contains only $2k-4$ denominator factors of the type (y_j+y_{j+1}). Once this is done the two sides of (2.38) can be successfully compared. The remainder of this section is the proof of (2.38) for $k=2$.

Using Lemma 2.2 for $g_{2l+1}, g_{2m+1}, g'_{2l+1}$, and g'_{2m+1}, Lemma 2.3 for $L_\nu g_{2k+1}$, and the definition (1.6) for $g_{2(k-m-l)-1}$ we can write the left-hand side of (2.38) for $k=2$ as

$$L_\nu g_5 - 2\sum_{l=0}^{1}\sum_{m=0}^{1-l} g_{2(2-m-l)-1}[g_{2l+1}g_{2m+1} - g'_{2l+1}g'_{2m+1}] = 2\int_1^\infty dy_1\cdots\int_1^\infty dy_5 \prod_{j=1}^{5}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^\nu \times\prod_{j=1}^{4}(y_j+y_{j+1})^{-1}\mathcal{L}_5(y_1,\ldots,y_5), \tag{2.39}$$

where

$$\begin{aligned}\mathcal{L}_5(y_1,\ldots,y_5) &= (y_1+y_2)(y_2+y_3)[(y_1y_5-1)(y_3^2-1) - (y_5y_3-1)\\&\quad\times[y_1(y_3+y_4+y_5)-1]] + (y_3+y_4)(y_4+y_5)\\&\quad\times[(y_1y_5-1)(y_3^2-1) - (y_3y_1-1)[(y_1+y_2+y_3)y_5-1]],\end{aligned} \tag{2.40}$$

where we used the labeling $1,2,\ldots,2l+1$ for g_{2l+1}; $2l+2, 2l+3,\ldots,2(k-m)$ for $g_{2(k-l-m)-1}$; and $2(k-m)+1,\ldots,2k+1$ for g_{2m+1}. We note that the $l=0$, $m=0$ term is zero. In this expression for $\mathcal{L}_5$ we use the identities

$$\begin{aligned}(y_1y_5-1)(y_3^2-1) - (y_1y_3-1)[(y_1+y_2+y_3)y_5-1] &= (y_1y_5-1)y_3[(y_2+y_3)-(y_1+y_2)]\\&\quad - (y_1y_3-1)y_5(y_2+y_3)\end{aligned} \tag{2.41a}$$

and

$$\begin{aligned}(y_1y_5-1)(y_3^2-1) - (y_5y_3-1)[(y_3+y_4+y_5)y_1-1] &= (y_1y_5-1)y_3[(y_3+y_4)-(y_4+y_5)]\\&\quad - (y_3y_5-1)y_1(y_3+y_4)\end{aligned} \tag{2.41b}$$

to rewrite $\mathcal{L}_5$ so that (2.39) becomes [note that by (2.41) we have factored out one denominator term]

$$\begin{aligned}&L_\nu g_5 - 2\sum_{l=0}^{1}\sum_{m=0}^{1-l} g_{2(2-m-l)-1}[g_{2l+1}g_{2m+1} - g'_{2l+1}g'_{2m+1}]\\&= 2\int_1^\infty dy_1\cdots\int_1^\infty dy_5\prod_{j=1}^{5}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^\nu\\&\quad\times\left\{(y_1y_5-1)\left[\frac{y_3}{y_4+y_5} - \frac{y_3}{y_3+y_4} + \frac{y_3}{y_1+y_2} - \frac{y_3}{y_2+y_3}\right] - \frac{y_1(y_3y_5-1)}{y_4+y_5} - \frac{y_5(y_1y_3-1)}{y_1+y_2}\right\}.\end{aligned} \tag{2.42}$$

In the last two terms we make the change of variables $y_1 \leftrightarrow y_3$ and $y_3 \leftrightarrow y_5$, respectively. Then (2.42) becomes

$$L_\nu g_5 - 2\sum_{l=0}^{1}\sum_{m=0}^{1-l} g_{2(2-m-l)-1}[g_{2l+1}g_{2m+1} - g'_{2l+1}g'_{2m+1}] = 2\int_1^\infty dy_1\cdots\int_1^\infty dy_5\prod_{j=1}^{5}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^\nu$$

$$\times(y_1y_5-1)\left[-\frac{y_3}{y_3+y_4}-\frac{y_3}{y_2+y_3}\right]$$
$$=-2\int_1^\infty dy_1\cdots\int_1^\infty dy_5\prod_{j=1}^5\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}$$
$$\times\left(\frac{y_j-1}{y_j+1}\right)^\nu(y_1y_5-1). \tag{2.43}$$

A few words are in order to explain this last step. Suppose we have

$$h=\int_1^\infty dy_1\int_1^\infty dy_2\prod_{j=1}^2\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^\nu\left(\frac{y_2}{y_1+y_2}\right). \tag{2.44}$$

Making the change of variables $y_1\leftrightarrow y_2$ in (2.44), adding this to (2.44) and dividing by two we find

$$h=\tfrac12\int_1^\infty dy_1\int_1^\infty dy_2\prod_{j=1}^2\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^\nu. \tag{2.45}$$

The result (2.45) was used in the last step of (2.43). We now compare (2.43) with the right-hand side of (2.38). We have [recall (2.34)]

$$\sum_{l=0}^{1}\sum_{m=0}^{1-l}g_{2l+1}g_{2m+1}L_\nu g_{3-2m-2l}=g_1^2L_\nu g_3. \tag{2.46}$$

Using (2.37) for $L_\nu g_3$ and (1.6a) for g_1 we conclude that (2.46) is exactly (2.43). Thus (2.8) is true for $k=2$.

E. Integral representation of (2.38) for general k

Before we proceed to the case $k=3$, we derive an integral representation for the left-hand side of (2.38) for general k. If we use Lemma 2.2 for g_{2l+1}, g_{2m+1}, g'_{2l+1}, and g'_{2m+1}, Lemma 2.3 for $L_\nu g_{2k+1}$, and definition (1.6b) for $g_{2(k-m-l)-1}$ and use the labeling $1,2,\ldots,2l+1$ for g_{2l+1}; $2l+2, 2l+3,\ldots,2(k-m)$ for $g_{2(k-l-m)-1}$, and $2(k-m)+1,\ldots,2k+1$ for g_{2m+1}, we find that the left-hand side of (2.38) can be written as

$$L_\nu g_{2k+1}-2\sum_{l=0}^{k-1}\sum_{m=0}^{k-l-1}g_{2(k-m-l)-1}[g_{2m+1}g_{2l+1}-g'_{2m+1}g'_{2l+1}]$$
$$=2(-1)^k\int_1^\infty dy_1\cdots\int_1^\infty dy_{2k+1}\prod_{j=1}^{2k+1}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}$$
$$\times\left(\frac{y_j-1}{y_j+1}\right)^\nu\prod_{j=1}^{2k}(y_j+y_{j+1})^{-1}\mathcal{L}_{2k+1}(y_1,\ldots,y_{2k+1}) \tag{2.47}$$

with $\mathcal{L}_{2k+1}$ given by

$$\mathcal{L}_{2k+1}=\sum_{m=1}^{k-1}(y_1+y_2)(y_{2k-2m}+y_{2k-2m+1})\prod_{\substack{n_1=2\\n_1\neq k-m+1}}^{k}(y_{2n_1-1}^2-1)$$
$$\times\Big\{(y_1y_{2k+1}-1)(y_{2k-2m+1}^2-1)-(y_{2k+1}y_{2k-2m+1}-1)$$
$$\times\Big(y_1\sum_{n_2=0}^{2m}y_{2k+1-n_2}-1\Big)\Big\}+\sum_{l=1}^{k-1}(y_{2l+1}+y_{2l+2})(y_{2k}+y_{2k+1})$$
$$\times\prod_{\substack{n_1=2\\n_1\neq l+1}}^{k}(y_{2n_1-1}^2-1)\Big\{(y_1y_{2k+1}-1)(y_{2l+1}^2-1)$$
$$-(y_1y_{2l+1}-1)\Big(y_{2k+1}\sum_{n_3=1}^{2l+1}y_{n_3}-1\Big)\Big\}$$
$$+\sum_{l=1}^{k-1}\sum_{m=1}^{k-l-1}(y_{2l+1}+y_{2l+2})(y_{2k-2m}+y_{2k-2m+1})$$
$$\prod_{\substack{n_1=2\\n_1\neq l+1,\,k-m+1}}^{k}(y_{2n_1-1}^2-1)\Big\{(y_1y_{2k+1}-1)(y_{2l+1}^2-1)$$
$$\times(y_{2k-2m+1}^2-1)-(y_1y_{2l+1}-1)(y_{2k+1}y_{2k+1-2m}-1)$$
$$\times\Big(\sum_{n_3=1}^{2l+1}y_{n_3}\sum_{n_2=0}^{2m}y_{2k+1-n_2}-1\Big)\Big\}. \tag{2.48}$$

As in the $k=2$ case, the $l=0$, $m=0$ term canceled. The first term in (2.48) (the term involving the sum $\sum_{m=1}^{k-1}$) is the $l=0$, $m\neq0$ terms of (2.38); the second term in (2.48) (the term involving the sum $\sum_{l=1}^{k-1}$) is the $l\neq0$, $m=0$ terms of (2.38); and the third term which involves the double sum is the $l\neq0$, $m\neq0$ terms of (2.38). We write the first term in curly brackets in (2.48) as

$$(y_1y_{2k+1}-1)(y_{2k-2m+1}^2-1)-(y_{2k+1}y_{2k-2m+1}-1)$$
$$\times\Big(y_1\sum_{n_2=0}^{2m}y_{2k+1-n_2}-1\Big)$$
$$=(y_1y_{2k+1}-1)y_{2k-2m+1}(y_{2k-2m+1}-y_{2k+1})$$
$$-y_1\sum_{n_2=1}^{2m}y_{2k+1-n_2}(y_{2k-2m+1}y_{2k+1}-1)$$
$$=(y_1y_{2k+1}-1)y_{2k-2m+1}\Big(\sum_{n_2=1}^{2m}y_{2k+1-n_2}-\sum_{n_2=0}^{2m-1}y_{2k+1-n_2}\Big)$$
$$-y_1(y_{2k-2m+1}y_{2k+1}-1)\sum_{n_2=1}^{2m}y_{2k+1-n_2}$$
$$=-y_{2k-2m+1}(y_1y_{2k+1}-1)\sum_{n_2=0}^{2m-1}y_{2k+1-n_2}$$
$$+(y_1-y_{2k-2m+1})\sum_{n_2=1}^{2m}y_{2k+1-n_2}, \tag{2.49}$$

the second term in curly brackets as

$$(y_1y_{2k+1}-1)(y_{2l+1}^2-1)-(y_1y_{2l+1}-1)\Big(y_{2k+1}\sum_{n_3=1}^{2l+1}y_{n_3}-1\Big)$$
$$=y_{2l+1}(y_1y_{2k+1}-1)(y_{2l+1}-y_1)-y_{2k+1}\Big(\sum_{n_3=2}^{2l+1}y_{n_3}\Big)(y_1y_{2l+1}-1)$$
$$=-y_{2l+1}(y_1y_{2k+1}-1)\sum_{n_3=1}^{2l}y_{n_3}+(y_{2k+1}-y_{2l+1})\sum_{n_3=2}^{2l+1}y_{n_3}, \tag{2.50}$$

and the third term in curly brackets as

$$(y_1y_{2k+1}-1)(y_{2l+1}^2-1)(y_{2k-2m+1}^2-1)-(y_1y_{2l+1}-1)$$
$$\times(y_{2k+1}y_{2k+1-2m}-1)\sum_{n_3=1}^{2l+1}y_{n_3}\sum_{n_2=0}^{2m}y_{2k+1-n_2}-1)$$
$$=-(y_1y_{2k+1}-1)[y_{2l+1}(y_{2k-2m+1}^2-1)(y_1-y_{2l+1})$$
$$+y_{2k-2m+1}(y_{2l+1}^2-1)(y_{2k+1}-y_{2k-2m+1})$$
$$+y_{2l+1}y_{2k-2m+1}(y_1-y_{2l+1})(y_{2k+1}-y_{2k-2m+1})]$$

$$
\begin{aligned}
&-(y_1y_{2l+1}-1)(y_{2k+1}y_{2k-2m+1}-1)\\
&\times\left[y_1\sum_{n_2=1}^{2m} y_{2k+1-n_2}+y_{2k+1}\sum_{n_3=2}^{2l+1} y_{n_3}\right.\\
&\left.+\sum_{n_3=2}^{2l+1} y_{n_3}\sum_{n_2=1}^{2m} y_{2k+1-n_2}\right],
\end{aligned}
$$

so that $\mathcal{L}_{2k+1}$ becomes

$$
\begin{aligned}
\mathcal{L}_{2k+1}&\\
=&\sum_{m=1}^{k-1}(y_1+y_2)(y_{2k-2m}+y_{2k-2m+1})\prod_{\substack{n_1=2\\ n_1\neq k-m+1}}^{k}(y_{2n_1-1}^2-1)\\
&\times\Big[-y_{2k-2m+1}(y_1y_{2k+1}-1)\\
&\times\sum_{n_2=0}^{2m-1} y_{2k+1-n_2}+(y_1-y_{2k-2m+1})\sum_{n_2=1}^{2m} y_{2k+1-n_2}\Big]\\
&+\sum_{l=1}^{k-1}(y_{2l+1}+y_{2l+2})(y_{2k}+y_{2k+1})\prod_{\substack{n_1=2\\ n_1\neq l+1}}^{k}(y_{2n_1-1}^2-1)\\
&\times\left[-y_{2l+1}(y_1y_{2k+1}-1)\sum_{n_3=1}^{2l} y_{n_3}+(y_{2k+1}-y_{2l+1})\sum_{n_3=2}^{2l+1} y_{n_3}\right]\\
&-\sum_{l=1}^{k-1}\sum_{m=1}^{k-l-1}(y_{2l+1}+y_{2l+2})(y_{2k-2m}+y_{2k-2m+1})\\
&\times\prod_{\substack{n_1=2\\ n_1\neq k-m+1,\,l+1}}^{k}(y_{2n_1-1}^2-1)\Big\{(y_1y_{2k+1}-1)\\
&\times[y_{2l+1}(y_{2k-2m+1}^2-1)(y_1-y_{2l+1})+y_{2k-2m+1}(y_{2l+1}^2-1)\\
&\times(y_{2k+1}-y_{2k-2m+1})+y_{2l+1}y_{2k-2m+1}(y_1-y_{2l+1})\\
&\times(y_{2k+1}-y_{2k+1-2m})]+(y_1y_{2l+1}-1)(y_{2k+1}y_{2k-2m+1}-1)\\
&\times\Big[y_1\sum_{n_2=1}^{2m} y_{2k+1-n_2}+y_{2k+1}\sum_{n_3=2}^{2l+1} y_{n_3}\\
&+\sum_{n_3=2}^{2l+1} y_{n_3}\sum_{n_2=1}^{2m} y_{2k+1-n_2}\Big]\Big\}.
\end{aligned}
\qquad (2.51)
$$

Frequently when working with the quantity $\mathcal{L}_{2k+1}$ we will perform operations upon $\mathcal{L}_{2k+1}$ (for instance, symmetrizing the integration variable labels) that leave the value of the right-hand side of (2.47) unchanged. Under these circumstances we will use the symbol "=" to mean that $\mathcal{L}_{2k+1}$ as given above and the right-hand side of the equation have identical values when substituted into (2.47). From the context of the equation it will be clear when we are using this meaning of "=".

F. Graphs and $\mathcal{L}_{2k+1}$

It is convenient to develop a graphical representation of the various terms that occur in $\mathcal{L}_{2k+1}$. The basic factor appearing in (2.47) is the quantity

$$\mathcal{L}_{2k+1}(y_1,\ldots,y_{2k+1})\prod_{j=1}^{2k}(y_j+y_{j+1})^{-1}. \qquad (2.52)$$

We can represent all such terms by the following rules:

(1) $\Pi_{j=1}^{2k}(y_j+y_{j+1})^{-1}$ is represented by a straight line with $2k+1$ points [see Fig. 2(a)].

(2) $(y_j+y_{j+1})\,\Pi_{l=1}^{2k}(y_l+y_{l+1})^{-1}$ is represented by a straight line with $2k+1$ points and one additional line connecting the points j and $j+1$ [see Fig. 2(b)].

(3) $(y_j^2-1)\,\Pi_{l=1}^{2k}(y_l+y_{l+1})^{-1}$ is represented by a straight line with $2k+1$ points and a circle centered about the jth point [see Fig. 2(c)].

(4) $y_j\,\Pi_{l=1}^{2k}(y_l+y_{l+1})^{-1}$ is represented by a straight line with $2k+1$ points and a "×" through the jth point [see Fig. 2(d)].

(5) Suppose we have a term $\mathcal{L}'$ which is a part of $\mathcal{L}_{2k+1}$. The order of $\mathcal{L}'$ is $2k+1$ and by the graph of $\mathcal{L}'$ we mean the graph of the integrand

$$\mathcal{L}'\prod_{j=1}^{2k}(y_j+y_{j+1})^{-1}$$

as constructed in accordance with rules (1)–(4).

Sometimes we wish to multiply some integrand factor $\mathcal{L}'$ by the factor $\mathcal{L}''$. If $\mathcal{L}'$ is a single graph, the product will be in general many graphs. To illustrate this multiplication of the graph $\mathcal{L}'$ by some other factor $\mathcal{L}''$ we draw the graph of $\mathcal{L}'$ and merely place $\mathcal{L}''$ to the extreme left. Of course, we may also explicitly draw all the graphs corresponding to $\mathcal{L}'\mathcal{L}''$ in accordance with rules (1)–(4).

G. Case $k=3$

For $k=3$ we write

$$\mathcal{L}_7=\sum_{l=0}^{2}\sum_{m=0}^{2-l}\mathcal{L}_7(l,m), \qquad (2.53)$$

where from (2.51) it follows that

$$
\begin{aligned}
\mathcal{L}_7(0,0)&=0,\\
\mathcal{L}_7(0,1)&=(y_1+y_2)(y_4+y_5)(y_3^2-1)\\
&\quad\times[-y_5(y_1y_7-1)(y_6+y_7)+(y_1-y_5)(y_6+y_5)],\\
\mathcal{L}_7(0,2)&=(y_1+y_2)(y_2+y_3)(y_5^2-1)[-y_3(y_1y_7-1)\\
&\quad\times(y_7+y_6+y_5+y_4)+(y_1-y_3)(y_6+y_5+y_4+y_3)],\\
\mathcal{L}_7(1,0)&=(y_3+y_4)(y_6+y_7)(y_5^2-1)\\
&\quad\times[-y_3(y_1y_7-1)(y_1+y_2)+(y_7-y_3)(y_2+y_3)],\\
\mathcal{L}_7(2,0)&=(y_5+y_6)(y_6+y_7)(y_3^2-1)[-y_5(y_1y_7-1)\\
&\quad\times(y_1+y_2+y_3+y_4)+(y_7-y_5)(y_2+y_3+y_4+y_5)],
\end{aligned}
$$

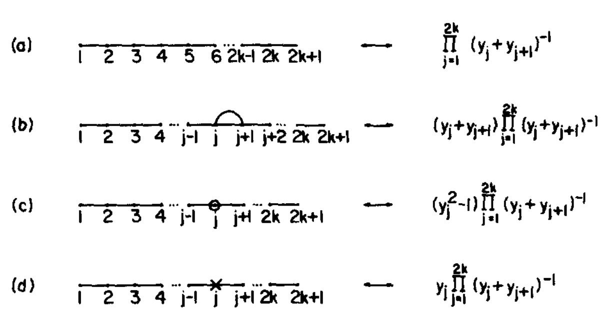

FIG. 2. (a) Graphical representation of $\Pi_{l=1}^{2k}(y_l+y_{l+1})^{-1}$. (b) Graphical representation of $(y_j+y_{j+1})\Pi_{l=1}^{2k}(y_l+y_{l+1})^{-1}$. (c) Graphical representation of $(y_j^2-1)\Pi_{l=1}^{2k}(y_l+y_{l+1})^{-1}$. (d) Graphical representation of $y_j\Pi_{l=1}^{2k}(y_l+y_{l+1})^{-1}$.

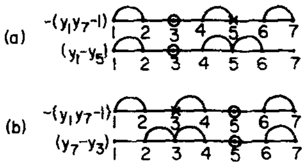

FIG. 3. (a) Graph of $\mathcal{L}_7(0,1)$ as defined by (2.54). (b) Graph of $\mathcal{L}_7(1,0)$ as defined by (2.54).

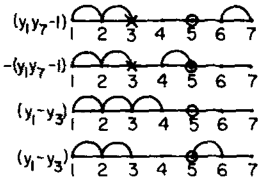

FIG. 5. Graph of $\mathcal{L}_7(0,2)$ as defined by (2.54).

$$\mathcal{L}_7(1,1) = -(y_3+y_4)(y_4+y_5)\{(y_1y_7-1)[y_3(y_5^2-1)(y_1-y_3)+y_5(y_3^2-1)(y_7-y_5)+y_3y_5(y_1-y_3)(y_7-y_5)]+(y_1y_3-1)(y_7y_5-1)[y_1(y_6+y_5)+y_7(y_2+y_3)+(y_2+y_3)(y_6+y_5)]\}. \tag{2.54}$$

The graphs of $\mathcal{L}_7(0,1)$ and $\mathcal{L}_7(1,0)$ are given in Fig. 3(a) and Fig. 3(b), respectively. From the graphs it is clear that $\mathcal{L}_7(0,1)$ and $\mathcal{L}_7(1,0)$ are equal [in the sense of "=" following (2.51)].

$\mathcal{L}_7(0,1)$ consists of two terms as illustrated in Fig. 3(a). If we let $1 \leftrightarrow 5$ in the second term, the integrand is antisymmetric and thus when integrated gives zero. Hence

$$\mathcal{L}_7(0,1) = -(y_1+y_2)(y_4+y_5)(y_3^2-1)y_5(y_1y_7-1)(y_6+y_7) \tag{2.55}$$

and similarly $(1 \leftrightarrow 3)$

$$\mathcal{L}_7(1,0) = -(y_3+y_4)(y_6+y_7)(y_5^2-1)y_3(y_1y_7-1)(y_1+y_2). \tag{2.56}$$

Both (2.55) and (2.56) can be reduced further. This reduction is essentially the same as that of (2.44) and (2.45) [in (2.55) symmetrize $5 \leftrightarrow 6$ and in (2.56) symmetrize $2 \leftrightarrow 3$]. Thus $\mathcal{L}_7(0,1)$ and $\mathcal{L}_7(1,0)$ become

$$\mathcal{L}_7(0,1) = -\tfrac{1}{2}(y_1y_7-1)(y_1+y_2)(y_4+y_5)(y_5+y_6)\times(y_6+y_7)(y_3^2-1) \tag{2.57}$$

and

$$\mathcal{L}_7(1,0) = -\tfrac{1}{2}(y_1y_7-1)(y_1+y_2)(y_2+y_3)(y_3+y_4)\times(y_6+y_7)(y_5^2-1), \tag{2.58}$$

respectively. The graph of $\mathcal{L}_7(0,1)$ is displayed in Fig. 4.

We now examine the term $\mathcal{L}_7(0,2)$. There are four basic terms in $\mathcal{L}_7(0,2)$ and these are displayed in Fig. 5.

The second graph has the reduction

$$-(y_1y_7-1)(y_1+y_2)(y_2+y_3)y_3(y_4+y_5)(y_5^2-1) \to -\tfrac{1}{2}(y_1y_7-1)(y_1+y_2)(y_2+y_3)(y_3+y_4)(y_4+y_5)(y_5^2-1) \tag{2.59}$$

which is obtained by symmetrizing the y_3 variable $(3 \leftrightarrow 4)$. This reduction always occurs when the graph is of the type Fig. 5(b). The general structure required for this reduction is shown in Fig. 6. The third graph of $\mathcal{L}_7(0,2)$ [Fig. 5(c)] gives zero weight to the integral (2.47) since the integrand is antisymmetric under the interchange $1 \leftrightarrow 3$. Thus $\mathcal{L}_7(0,2)$ becomes

$$\mathcal{L}_7(0,2) = -(y_1y_7-1)(y_1+y_2)(y_2+y_3)y_3(y_5^2-1)(y_6+y_7) - \tfrac{1}{2}(y_1y_7-1)(y_1+y_2)(y_2+y_3)(y_3+y_4)\times(y_4+y_5)(y_5^2-1)+(y_1-y_3)(y_1+y_2)\times(y_2+y_3)(y_5^2-1)(y_5+y_6). \tag{2.60}$$

This reduced form for $\mathcal{L}_7(0,2)$ is shown in Fig. 7. The second term in (2.60) [Fig. 7(b)] has the correct number of factored denominators (in a graph this always corresponds to four loops).

A similar reduction for $\mathcal{L}_7(2,0)$ gives

$$\mathcal{L}_7(2,0) = (y_5+y_6)(y_6+y_7)(y_3^2-1)[-y_5(y_1y_7-1)(y_1+y_2) - \tfrac{1}{2}(y_1y_7-1)(y_4+y_5)(y_3+y_4)+(y_7-y_5)(y_2+y_3)]. \tag{2.61}$$

The graph of (2.61) is shown in Fig. 8 and should be compared with Fig. 7.

From (2.54) we can write $\mathcal{L}_7(1,1)$ as

$$\mathcal{L}_7(1,1) = -(y_3+y_4)(y_4+y_5)\{(y_1y_7-1)[y_3y_5(y_1-y_3)(y_7-y_5)+(y_5^2-1)y_3(y_1+y_2)+(y_3^2-1)y_5(y_7+y_6)]+(y_2+y_3)[y_7(y_1y_3-1)(y_5y_7-1)-y_3(y_1y_7-1)(y_5^2-1)]+(y_5+y_6)\times[y_1(y_1y_3-1)(y_5y_7-1)-y_5(y_1y_7-1)(y_3^2-1)]+(y_1y_3-1)(y_5y_7-1)(y_2+y_3)(y_5+y_6)\}. \tag{2.62}$$

We now use the identities

$$y_{2k+1}(y_1y_{2l+1}-1)(y_{2k-2m+1}y_{2k+1}-1) - y_{2l+1}(y_{2k-2m+1}^2-1)(y_1y_{2k+1}-1) = (y_{2k+1-2m}^2-1)(y_{2l+1}-y_{2k+1})+y_{2k+1}y_{2k-2m+1}\times(y_1y_{2l+1}-1)(y_{2k+1}-y_{2k-2m+1}) \tag{2.63}$$

and

$$y_1(y_1y_{2l+1}-1)(y_{2k-2m+1}y_{2k+1}-1) - y_{2k-2m+1}(y_{2l+1}^2-1)(y_1y_{2k+1}-1)$$

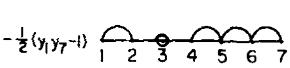

FIG. 4. Quantity (2.57).

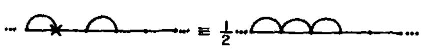

FIG. 6. General reduction formula. See discussion following (2.59).

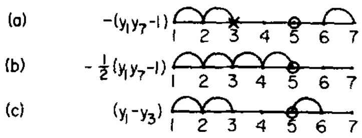

FIG. 7. Quantity (2.60).

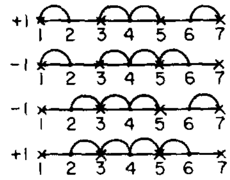

FIG. 9. Quantity (2.65).

$$= (y_{2l+1}^2 - 1)(y_{2k-2m+1} - y_1) + y_1 y_{2l+1}(y_1 - y_{2l+1})(y_{2k-2m+1} y_{2k+1} - 1)$$

for $k=3$, $l=1$, and $m=1$ in (2.62). The second and third terms multiplying the factor $(y_1 y_7 - 1)$ can be further reduced (these terms are of the general structure of Fig. 6). Carrying this out we can write $\mathcal{L}_7(1,1)$ as

$$\begin{aligned}\mathcal{L}_7(1,1) = &-(y_3+y_4)(y_4+y_5)[(y_1y_7-1)[y_3y_5(y_1-y_3)(y_7-y_5)\\ &+\tfrac{1}{2}(y_1+y_2)(y_2+y_3)(y_5^2-1)+\tfrac{1}{2}(y_5+y_6)(y_6+y_7)\\ &\times(y_3^2-1)]+(y_2+y_3)[(y_5^2-1)(y_3-y_7)\\ &+y_5y_7(y_1y_3-1)(y_7-y_5)]+(y_5+y_6)\\ &\times[(y_3^2-1)(y_5-y_1)+y_1y_3(y_1-y_3)(y_5y_7-1)]\\ &+(y_1y_3-1)(y_5y_7-1)(y_2+y_3)(y_5+y_6)].\end{aligned} \tag{2.64}$$

We now examine the term

$$\begin{aligned}&y_1y_7(y_3+y_4)(y_4+y_5)\,y_3y_5(y_1-y_3)(y_7-y_5)\\ &\quad=(y_3+y_4)(y_4+y_5)\,y_1y_3y_5y_7[(y_1+y_2)\\ &\qquad-(y_2+y_3)][(y_7+y_6)-(y_6+y_5)]\end{aligned} \tag{2.65}$$

occurring in $\mathcal{L}_7(1,1)$. We draw the graph of (2.65) in Fig. 9. The first term cancels the second and third term, the fourth (let $5\to7$ and $7\to5$ in the first and third graphs).

We now combine the terms $\mathcal{L}_7(0,2)$, $\mathcal{L}_7(2,0)$, and $\mathcal{L}_7(1,1)$. One way to create denominator factors from a term like $(y_j - y_k)$ is to write this as $(y_j + y_{j+1}) - (y_{j+1} + y_{j+2}) + \cdots - (y_{k-1} + y_k)$. This identity has been extensively used already. However there are terms where this is of no use. For instance in (2.60) for $\mathcal{L}_7(0,2)$ there occurs the term $(y_1 - y_3)$. If we were to rewrite this as $(y_1+y_2)-(y_2+y_3)$ we would introduce the factors $(y_1+y_2)^2$ and $(y_2+y_3)^2$. We do not want terms of this form. Such a problem term occurs in (2.61) for $\mathcal{L}_7(2,0)$ [the $(y_7 - y_5)$ term] and two such terms in (2.64). We combine these terms:

$$\begin{aligned}J_7 \equiv\ &(y_1-y_3)(y_1+y_2)(y_2+y_3)(y_5+y_6)(y_5^2-1)\\ &+(y_7-y_5)(y_2+y_3)(y_5+y_6)(y_6+y_7)(y_3^2-1)\\ &-(y_3-y_7)(y_2+y_3)(y_3+y_4)(y_4+y_5)(y_5^2-1)\\ &-(y_5-y_1)(y_3+y_4)(y_4+y_5)(y_5+y_6)(y_3^2-1).\end{aligned} \tag{2.66}$$

The graph of J_7 is shown in Fig. 10. The first term cancels the third term and the second term is canceled by the fourth term. This can be seen by the change of variables $1\to3$, $2\to4$, $3\to7$, $4\to6$, $5\to5$, $6\to2$, and $7\to1$ (this relabeling is seen most easily by comparing the first graph with the third graph of Fig. 10).

Thus we have

$$\begin{aligned}&\prod_{j=1}^{6}(y_j+y_{j+1})^{-1}[(y_1-y_3)(y_1+y_2)(y_2+y_3)(y_5+y_6)(y_5^2-1)]\\ &\quad=\frac{(y_1-y_3)(y_5^2-1)}{(y_3+y_4)(y_4+y_5)(y_6+y_7)}\to\frac{(y_3-y_7)(y_5^2-1)}{(y_7+y_6)(y_6+y_5)(y_1+y_2)}\\ &\quad=\prod_{j=1}^{6}(y_j+y_{j+1})^{-1}[(y_3-y_7)\\ &\qquad\times(y_5^2-1)(y_2+y_3)(y_3+y_4)(y_4+y_5)]\end{aligned}$$

which is the third graph. Hence we have demonstrated

$$J_7 = 0 \tag{2.67}$$

where we use the sense of "=" as discussed after (2.51).

We now examine the term

$$-(y_3+y_4)(y_4+y_5)\,y_1y_3(y_1-y_3)(y_5y_7-1)(y_5+y_6) \tag{2.68}$$

in $\mathcal{L}_7(1,1)$ [see (2.64)]. This clearly gives zero contribution since the above integrand [multiplied as always by $\prod_{j=1}^{6}(y_j+y_{j+1})^{-1}$] is antisymmetric under the interchange $1\leftrightarrow3$. The same is true for the term

$$(y_3+y_4)(y_4+y_5)(y_5+y_6)\,y_5y_7(y_1y_3-1)(y_7-y_5) \tag{2.69}$$

occurring in $\mathcal{L}_7(1,1)$. Collecting these results we have

$$\begin{aligned}&\mathcal{L}_7(0,2)+\mathcal{L}_7(2,0)+\mathcal{L}_7(1,1)\\ &\quad=-(y_1y_7-1)(y_1+y_2)(y_2+y_3)(y_5^2-1)[y_3(y_6+y_7)\\ &\qquad+\tfrac{1}{2}(y_3+y_4)(y_4+y_5)]-(y_1y_7-1)(y_5+y_6)(y_6+y_7)\\ &\qquad\times(y_3^2-1)[y_5(y_1+y_2)+\tfrac{1}{2}(y_4+y_5)(y_3+y_4)]\\ &\qquad-(y_1y_7-1)(y_3+y_4)(y_4+y_5)[\tfrac{1}{2}(y_2+y_3)(y_1+y_2)\\ &\qquad\times(y_5^2-1)+\tfrac{1}{2}(y_5+y_6)(y_6+y_7)(y_3^2-1)]\\ &\qquad+(y_3+y_4)(y_4+y_5)\,y_3y_5(y_1-y_3)(y_7-y_5)\\ &\qquad-(y_3+y_4)(y_4+y_5)(y_1y_3-1)(y_5y_7-1)(y_2+y_3)(y_5+y_6).\end{aligned} \tag{2.70}$$

Though the last term in (2.70) contains four denomina-

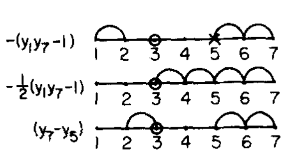

FIG. 8. Quantity (2.61).

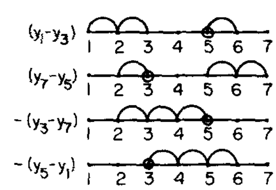

FIG. 10. Quantity (2.66).

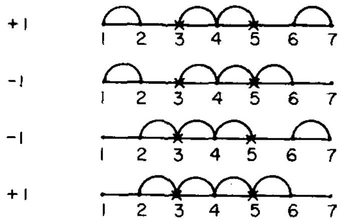

FIG. 11. Quantity $y_3y_5(y_3+y_4)(y_4+y_5)(y_1-y_3)(y_7-y_5)$.

tor type factors, the presence of the two terms $(y_1y_3-1)(y_5y_7-1)$ is not desired. Thus we expect further reductions of this term along with the other terms in (2.70) that do not have four denominator factors.

We examine the combination

$$(y_3+y_4)(y_4+y_5)[y_3y_5(y_1-y_3)(y_7-y_5) - (y_1y_3-1)(y_5y_7-1)(y_2+y_3)(y_5+y_6)]. \tag{2.71}$$

Now

$$\begin{aligned} y_3y_5(y_1-y_3)(y_7-y_5) &= y_3y_5(y_1+y_2)(y_6+y_7) - y_3y_5(y_1+y_2)(y_5+y_6) \\ &\quad - y_3y_5(y_2+y_3)(y_6+y_7) + y_3y_5(y_2+y_3)(y_5+y_6) \end{aligned} \tag{2.72}$$

so that the term $y_3y_5(y_3+y_4)(y_4+y_5)(y_1-y_3)(y_7-y_5)$ can be viewed as a sum of four terms. These terms are displayed in Fig. 11. In the first graph we let $1 \leftrightarrow 3$ and $5 \leftrightarrow 7$, in the second graph $1 \leftrightarrow 3$, and $5 \leftrightarrow 7$ in the third graph to obtain

$$y_3y_5(y_1-y_3)(y_7-y_5) = (y_1-y_3)(y_7-y_5)(y_2+y_3)(y_5+y_6). \tag{2.73}$$

Using (2.73) the expression (2.71) becomes

$$\begin{aligned} &(y_3+y_4)(y_4+y_5)(y_2+y_3)(y_5+y_6) \\ &\qquad\times[(y_1-y_3)(y_7-y_5)-(y_1y_3-1)(y_5y_7-1)] \\ &\quad= -(y_2+y_3)(y_3+y_4)(y_4+y_5)(y_5+y_6) \\ &\qquad\times[(y_1y_7-1)(y_3y_5-1)+y_1(y_5-y_3)+y_7(y_3-y_5)]. \end{aligned} \tag{2.74}$$

By letting $5 \leftrightarrow 3$ in the second and third terms in (2.74) we see that the integrand obtained from (2.74) [that is, multiply (2.74) by $\prod_{j=1}^{6}(y_j+y_{j+1})^{-1}$] is antisymmetric. Hence (2.74) is equivalent to

$$-(y_2+y_3)(y_3+y_4)(y_4+y_5)(y_5+y_6)(y_1y_7-1)(y_3y_5-1). \tag{2.75}$$

Multiplying (2.75) by $\prod_{j=1}^{6}(y_j+y_{j+1})^{-1}$ we have

$$-\frac{(y_1y_7-1)(y_3y_5-1)}{(y_1+y_2)(y_6+y_7)} \tag{2.76}$$

which will be integrated over in (2.47). We relabel the variables by $1\to 3$, $3\to 7$, $5\to 1$, and $7\to 5$ (keeping the even labels fixed) so that (2.76) is equivalent to

$$-\frac{(y_1y_7-1)(y_3y_5-1)}{(y_2+y_3)(y_5+y_6)} \tag{2.77}$$

which implies (2.75) is equivalent to

$$-(y_1+y_2)(y_6+y_7)(y_3+y_4)(y_4+y_5)(y_1y_7-1)(y_3y_5-1). \tag{2.78}$$

Using these results (2.70) becomes

$$\begin{aligned} &\mathcal{L}_7(0,2)+\mathcal{L}_7(2,0)+\mathcal{L}_7(1,1) \\ &\quad= -(y_1y_7-1)\{(y_1+y_2)(y_2+y_3)(y_5^2-1)[y_3(y_6+y_7) \\ &\qquad+\tfrac{1}{2}(y_3+y_4)(y_4+y_5)]+(y_5+y_6)(y_6+y_7)(y_3^2-1) \\ &\qquad\times[y_5(y_1+y_2)+\tfrac{1}{2}(y_3+y_4)(y_4+y_5)] \\ &\qquad+(y_3+y_4)[\tfrac{1}{2}(y_1+y_2)(y_2+y_3)(y_5^2-1) \\ &\qquad+\tfrac{1}{2}(y_5+y_6)(y_6+y_7)(y_3^2-1) \\ &\qquad+(y_3y_5-1)(y_1+y_2)(y_6+y_7)]\}. \end{aligned} \tag{2.79}$$

Making use of the identity

$$\begin{aligned} y_3(y_5^2-1) &= (y_3+y_4)(y_5^2-1)-(y_4+y_5)(y_4y_5-1) \\ &\quad+(y_4^2-1)(y_5+y_6)-y_6(y_4^2-1) \end{aligned} \tag{2.80}$$

we see that by a relabeling of the integration variable labels the quantity

$$\frac{y_3(y_5^2-1)}{(y_3+y_4)(y_4+y_5)(y_5+y_6)}$$

can be replaced by

$$\frac{(y_3+y_4)(y_5^2-1)}{(y_3+y_4)(y_4+y_5)(y_5+y_6)} - \frac{1}{2}\,\frac{(y_4+y_5)(y_4y_5-1)}{(y_3+y_4)(y_4+y_5)(y_5+y_6)}$$

in (2.79). A similar transformation on the term $y_5(y_1+y_2)$ in (2.79) results in the equivalent expression for (2.79),

$$\begin{aligned} &\mathcal{L}_7(0,2)+\mathcal{L}_7(2,0)+\mathcal{L}_7(1,1) \\ &\quad= -(y_1y_7-1)\{(y_1+y_2)(y_2+y_3)(y_5^2-1)[(y_3+y_4)(y_6+y_7) \\ &\qquad+\tfrac{1}{2}(y_3+y_4)(y_4+y_5)]+(y_5+y_6)(y_6+y_7)(y_3^2-1) \\ &\qquad\times[(y_4+y_5)(y_1+y_2)+\tfrac{1}{2}(y_3+y_4)(y_4+y_5)] \\ &\qquad+(y_3+y_4)(y_4+y_5)[\tfrac{1}{2}(y_1+y_2)(y_2+y_3)(y_5^2-1) \\ &\qquad+\tfrac{1}{2}(y_5+y_6)(y_6+y_7)(y_3^2-1)]-\tfrac{1}{2}(y_4+y_5)(y_4y_5-1) \\ &\qquad\times(y_1+y_2)(y_2+y_3)(y_6+y_7)-\tfrac{1}{2}(y_3+y_4)(y_3y_4-1) \\ &\qquad\times(y_1+y_2)(y_5+y_6)(y_6+y_7)+(y_1+y_2)(y_3+y_4) \\ &\qquad\times(y_4+y_5)(y_6+y_7)(y_3y_5-1)\}. \end{aligned} \tag{2.81}$$

The last three terms of (2.81) cancel. To see this we multiply these terms by the factor $\prod_{j=1}^{6}(y_j+y_{j+1})^{-1}$ to obtain

$$\begin{aligned} &-\frac{1}{2}\,\frac{(y_1y_7-1)(y_4y_5-1)}{(y_3+y_4)(y_5+y_6)} - \frac{1}{2}\,\frac{(y_1y_7-1)(y_3y_4-1)}{(y_2+y_3)(y_4+y_5)} \\ &\quad+\frac{(y_1y_7-1)(y_3y_5-1)}{(y_2+y_3)(y_5+y_6)}. \end{aligned} \tag{2.82}$$

Letting $3\to 2$, $4\to 3$ in the first term and $4\to 5$, $5\to 6$ in the second term we see that (2.82) is zero. Hence using this in (2.81) and adding the result to $\mathcal{L}_7(0,1)+\mathcal{L}_7(1,0)$ [see (2.57) and (2.58)] we find that (2.47) for the case $k=3$ becomes

$$2(-1)^3\int_1^\infty dy_1\cdots\int_1^\infty dy_7\prod_{j=1}^7\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^\nu$$
$$\times\prod_{j=1}^6(y_j+y_{j+1})^{-1}\mathcal{L}_7$$
$$=2\int_1^\infty dy_1\cdots\int_1^\infty dy_7\prod_{j=1}^7\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^\nu$$
$$\times(y_1y_7-1)\left(\frac{3}{2}\frac{(y_3^2-1)}{(y_2+y_3)(y_3+y_4)}+\frac{3}{2}\frac{(y_5^2-1)}{(y_4+y_5)(y_5+y_6)}\right.$$
$$\left.+\frac{(y_5^2-1)}{(y_5+y_6)(y_6+y_7)}+\frac{(y_3^2-1)}{(y_1+y_2)(y_2+y_3)}\right). \tag{2.83}$$

We now compare the result (2.83) with the right-hand side of (2.38). From (2.38) and (2.34)

$$\sum_{l=0}^{2}\sum_{m=0}^{2-l}g_{2l+1}g_{2m+1}L_\nu g_{2(3-m-l)-1}$$
$$=g_1^2L_\nu g_5+2g_1g_3L_\nu g_3. \tag{2.84}$$

Using the definition of the function g_1 and g_3 and Lemma 2.3 for $L_\nu g_3$ and $L_\nu g_5$ we can write (2.84) as

$$\sum_{l=0}^{2}\sum_{m=0}^{2-l}g_{2l+1}g_{2m+1}L_\nu g_{2(3-m-l)-1}$$
$$=2\int_1^\infty dy_1\cdots\int_1^\infty dy_7\prod_{j=1}^7\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^\nu$$
$$\times\left(\frac{(y_1y_5-1)(y_3^2-1)}{(y_2+y_3)(y_3+y_4)}+\frac{(y_1y_5-1)(y_3^2-1)}{(y_3+y_4)(y_4+y_5)}\right.$$
$$\left.+\frac{(y_1y_5-1)(y_3^2-1)}{(y_1+y_2)(y_2+y_3)}+2\frac{(y_1y_3-1)(y_5^2-1)}{(y_4+y_5)(y_5+y_6)}\right). \tag{2.85}$$

By relabeling the integration variable subscripts we see that (2.85) and (2.83) are identical. Hence we have proved identity (2.8) for $k=3$.

H. Cases $k\geqslant 4$

We have proved (2.38) for $k=1$, 2, and 3. To prove Theorem 1 we must prove (2.38) [and hence (2.8)] for $k\geqslant 4$. In the preceding section the $k=3$ case of (2.38) was presented. Rather than give the most direct proof possible for $k=3$, we presented a proof that parallels as much as possible the general proof of this section. Even so the general proof is involved and at places special cases are presented to help see the cancellation that is taking place.

1. Alternative form for $\mathcal{L}_{2k+1}$

We start with $\mathcal{L}_{2k+1}$ as given by (2.51) and write

$$\mathcal{L}_{2k+1}=\sum_{l=1}^{k-1}\sum_{m=0}^{k-l-1}\mathcal{L}_{2k+1}(l,m) \tag{2.86}$$

with $\mathcal{L}_{2k+1}(0,0)\equiv 0$. Equation (2.51) can be rewritten (by adding and subtracting terms) as

$$\mathcal{L}_{2k+1}=\sum_{m=1}^{k-1}(y_1+y_2)(y_{2k-2m}+y_{2k-2m+1})\prod_{\substack{n_1=2\\ n_1\neq k-m+1}}^{k}(y^2_{2n_1-1}-1)$$
$$\times\left(-y_{2k-2m+1}(y_1y_{2k+1}-1)\sum_{n_2=0}^{2m-1}y_{2k+1-n_2}\right.$$
$$\left.+(y_1-y_{2k-2m+1})\sum_{n_2=1}^{2m}y_{2k+1-n_2}\right)+\sum_{l=1}^{k-1}(y_{2l+1}+y_{2l+2})$$
$$\times(y_{2k}+y_{2k+1})\prod_{\substack{n_1=2\\ n_1\neq l+1}}^{k}(y^2_{2n_1-1}-1)\left(-y_{2l+1}(y_1y_{2k+1}-1)\right.$$
$$\left.\times\sum_{n_3=1}^{2l}y_{n_3}+(y_{2k+1}-y_{2l+1})\sum_{n_3=2}^{2l+1}y_{n_3}\right)$$
$$-\sum_{l=1}^{k-1}\sum_{m=1}^{k-l-1}(y_{2l+1}+y_{2l+2})(y_{2k-2m}+y_{2k-2m+1})$$
$$\times\prod_{\substack{n_1=2\\ n_1\neq k-m+1,\,l+1}}^{k}(y^2_{2n_1-1}-1)\left((y_1y_{2k+1}-1)\right.$$
$$\times y_{2l+1}y_{2k-2m+1}(y_1-y_{2l+1})(y_{2k+1}-y_{2k-2m+1})$$
$$+(y_1y_{2l+1}-1)(y_{2k+1}y_{2k-2m+1}-1)$$
$$\times\sum_{n_3=2}^{2l+1}y_{n_3}\sum_{n_2=1}^{2m}y_{2k+1-n_2}+(y_1y_{2k+1}-1)(y^2_{2l+1}-1)y_{2k-2m+1}$$
$$\times\sum_{n_2=0}^{2m-1}y_{2k+1-n_2}+(y_1y_{2k+1}-1)(y^2_{2k-2m+1}-1)y_{2l+1}$$
$$\times\sum_{n_3=1}^{2l}y_{n_3}+\sum_{n_2=1}^{2m}y_{2k+1-n_2}[y_1(y_1y_{2l+1}-1)$$
$$\times(y_{2k+1}y_{2k-2m+1}-1)-y_{2k-2m+1}(y^2_{2l+1}-1)(y_1y_{2k+1}-1)]$$
$$+\sum_{n_3=2}^{2l+1}y_{n_3}[y_{2k+1}(y_1y_{2l+1}-1)(y_{2k-2m+1}y_{2k+1}-1)$$
$$\left.-y_{2l+1}(y^2_{2k-2m+1}-1)(y_1y_{2k+1}-1)]\right). \tag{2.87}$$

We first examine the $l=0$, $m=1$ term of (2.87), i.e.,

$$\mathcal{L}_{2k+1}(0,1)=\prod_{n_1=2}^{k-1}(y^2_{2n_1-1}-1)(y_1+y_2)(y_{2k-2}+y_{2k-1})$$
$$\times[-y_{2k-1}(y_1y_{2k+1}-1)(y_{2k}+y_{2k+1})$$
$$+(y_1-y_{2k-1})(y_{2k-1}+y_{2k})]. \tag{2.88}$$

The term containing (y_1-y_{2k-1}) in (2.88) gives zero contribution to $\Pi_{j=1}^{2k}(y_j+y_{j+1})^{-1}\mathcal{L}_{2k+1}(0,1)$ (let $1\leftrightarrow 2k-1$). Hence we have

$$\mathcal{L}_{2k+1}(0,1)=-\prod_{n_1=2}^{k-1}(y^2_{2n_1-1}-1)(y_1+y_2)(y_{2k-2}+y_{2k-1})y_{2k-1}$$
$$\times(y_1y_{2k+1}-1)(y_{2k}+y_{2k+1}). \tag{2.89}$$

Furthermore, symmetrizing the y_{2k-1} variable (recall argument associated with Fig. 6) we have

$$\mathcal{L}_{2k+1}(0,1)=-\tfrac{1}{2}\prod_{n_1=2}^{k-1}(y^2_{2n_1-1}-1)(y_1+y_2)(y_{2k-2}+y_{2k-1})$$
$$\times(y_{2k-1}+y_{2k})(y_{2k}+y_{2k+1})(y_1y_{2k+1}-1). \tag{2.90}$$

Similar transformations result in

$$\mathcal{L}_{2k+1}(1,0)=-\tfrac{1}{2}\prod_{n_1=3}^{k}(y_{2n_1-1}^2-1)(y_1+y_2)(y_2+y_3)$$
$$\times(y_3+y_4)(y_{2k}+y_{2k+1})(y_1y_{2k+1}-1). \qquad (2.91)$$

Both $\mathcal{L}_{2k+1}(0,1)$ and $\mathcal{L}_{2k+1}(1,0)$ have the required four denominator type factors and a single $(y_1y_{2k+1}-1)$ factor.

We now analyze

$$\sum_{m=2}^{k-1}\mathcal{L}_{2k+1}(0,m) \text{ and } \sum_{l=2}^{k-1}\mathcal{L}_{2k+1}(l,0).$$

From (2.87) we have

$$\sum_{m=2}^{k-1}\mathcal{L}_{2k+1}(0,m)$$
$$=\sum_{m=2}^{k-1}(y_1+y_2)(y_{2k-2m}+y_{2k-2m+1})\prod_{\substack{n_1=2\\ n_1\neq k-m+1}}^{k}(y_{2n_1-1}^2-1)$$
$$\times\Big[-y_{2k-2m+1}(y_1y_{2k+1}-1)\sum_{n_2=0}^{2m-3}y_{2k+1-n_2}-y_{2k-2m+1}$$
$$\times(y_1y_{2k+1}-1)(y_{2k-2m+2}+y_{2k-2m+3})+(y_1-y_{2k-2m+1})$$
$$\times(y_{2k-2m+1}+y_{2k-2m+2})+(y_1-y_{2k-2m+1})\sum_{n_2=1}^{2m-2}y_{2k+1-n_2}\Big]. \qquad (2.92)$$

Symmetrizing the $y_{2k-2m+1}$ variable in the second term in square brackets in (2.92) $(y_{2k-2m+1}\leftrightarrow y_{2k-2m+2})$ and observing that the term $(y_{2k-2m}+y_{2k-2m+1})(y_1-y_{2k-2m+1})(y_{2k-2m+1}+y_{2k-2m+2})(y_1+y_2)$ is equivalent to zero $(1\leftrightarrow 2k-2m+1)$ the quantity (2.92) becomes

$$\sum_{m=2}^{k-1}\mathcal{L}_{2k+1}(0,m)$$
$$=\sum_{m=2}^{k-1}(y_1+y_2)(y_{2k-2m}+y_{2k-2m+1})\prod_{\substack{n_1=2\\ n_1\neq k-m+1}}^{k}(y_{2n_1-1}^2-1)$$
$$\times\Big[-y_{2k-2m+1}(y_1y_{2k+1}-1)$$
$$\times\sum_{n_2=0}^{2m-3}y_{2k+1-n_2}-\tfrac{1}{2}(y_{2k-2m+1}+y_{2k-2m+2})(y_{2k-2m+2}$$
$$+y_{2k-2m+3})(y_1y_{2k+1}-1)+(y_1-y_{2k-2m+1})\sum_{n_2=1}^{2m-2}y_{2k+1-n_2}\Big]. \qquad (2.93)$$

Similarly for $\sum_{l=2}^{k-1}\mathcal{L}_{2k+1}(l,0)$ we have

$$\sum_{l=2}^{k-1}\mathcal{L}_{2k+1}(l,0)$$
$$=\sum_{l=2}^{k-1}(y_{2l+1}+y_{2l+2})(y_{2k}+y_{2k+1})\prod_{\substack{n_1=2\\ n_1\neq l+1}}^{k}(y_{2n_1-1}^2-1)$$
$$\times\Big[-y_{2l+1}(y_1y_{2k+1}-1)\sum_{n_3=1}^{2l-2}y_{n_3}-\tfrac{1}{2}(y_{2l-1}+y_{2l})$$
$$\times(y_{2l}+y_{2l+1})(y_1y_{2k+1}-1)+(y_{2k+1}-y_{2l+1})\sum_{n_3=2}^{2l-1}y_{n_3}\Big]. \qquad (2.94)$$

We now claim

$$\sum_{l=1}^{k-1}\sum_{m=1}^{k-l-1}\prod_{\substack{n_1=2\\ n_1\neq l+1,\,k-m+1}}^{k}(y_{2n_1-1}^2-1)(y_{2l+1}+y_{2l+2})$$
$$\times(y_{2k-2m}+y_{2k-2m+1})y_1y_{2l+1}y_{2k-2m+1}y_{2k+1}$$
$$\times(y_1-y_{2l+1})(y_{2k+1}-y_{2k-2m+1})=0 \qquad (2.95)$$

which is the generalization of (2.65). To demonstrate this we write

$$y_1-y_{2l+1}=(y_1+y_2)-(y_2+y_3)+\cdots$$
$$+(y_{2l-1}+y_{2l})-(y_{2l}+y_{2l+1}) \qquad (2.96)$$
$$y_{2k+1}-y_{2k-2m+1}=(y_{2k+1}+y_{2k})-(y_{2k}+y_{2k-1})$$
$$+\cdots-(y_{2k-2m}+y_{2k-2m+1}).$$

Then for a fixed l and m each term in (2.95) can be written as a sum of $4l(m+1)$ terms. A typical term is of the form

$$(-1)^{p+q}\prod_{\substack{n_1=2\\ n_1\neq l+1,\,k-m+1}}^{k}(y_{2n_1-1}^2-1)(y_{2l+1}+y_{2l+2})(y_{2k-2m}+y_{2k-2m+1})$$
$$\times y_1y_{2l+1}y_{2k-2m+1}y_{2k+1}(y_p+y_{p+1})(y_q+y_{q+1}), \qquad (2.97)$$

where $p=1,2,\ldots,2l$ and $q=2k+1,\ 2k,\ldots,2k-2m$.

Keeping q fixed we examine the term with p replaced by $2l+1-p$. It is

$$(-1)^{p+q}\prod_{\substack{n_1=2\\ n_1\neq l+1,\,k-m+1}}^{k}(y_{2n_1-1}^2-1)(y_{2l+1}+y_{2l+2})(y_{2k-2m}+y_{2k-2m+1})$$
$$\times y_1y_{2l+1}y_{2k-2m+1}y_{2k+1}(y_{2l+1-p}+y_{2l+2-p})(y_q+y_{q+1}). \qquad (2.98)$$

These two terms [(2.97) and (2.98)] are equivalent as can be seen from their graphs (see Fig. 12). They differ by an overall minus sign and thus add to give zero. Since this is true for fixed l, m, and q, we have pairwise cancellation as the index p runs through $1,2,\ldots,2l$. Hence it follows that (2.95) is true.

The term

$$(y_{2l+1}+y_{2l+2})(y_{2k-2m}+y_{2k-2m+1})(y_1y_{2k+1}-1)$$
$$\times(y_{2l+1}^2-1)\,y_{2k-2m+1}\sum_{n_2=0}^{2m-1}y_{2k+1-n_2} \qquad (2.99)$$

occurring in (2.87) is equivalent to

$$(y_{2l+1}+y_{2l+2})(y_{2k-2m}+y_{2k-2m+1})(y_1y_{2k+1}-1)(y_{2l+1}^2-1)$$
$$\times\tfrac{1}{2}(y_{2k-2m+1}+y_{2k-2m+2})(y_{2k-2m+2}+y_{2k-2m+3})$$
$$+(y_{2l+1}+y_{2l+2})(y_{2k-2m}+y_{2k-2m+1})(y_1y_{2k+1}-1)$$
$$\times y_{2k-2m+1}\sum_{n_2=0}^{2m-3}y_{2k+1-n_2} \qquad (2.100)$$

as can be seen by symmetrizing y_{2k-m+1} when it multiplies

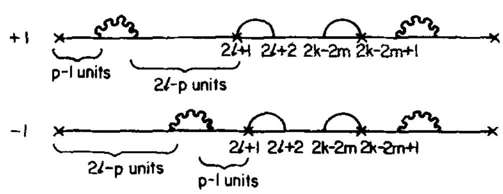

FIG. 12. Quantities (2.97) and (2.98). The loops that are moved are with wavy lines.

(a)

(b)

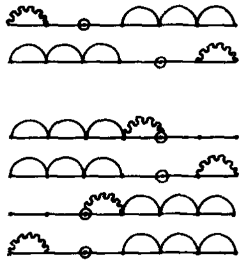

FIG. 13. (a) $l+m=1$ terms for $k=3$. (b) $l+m=2$ terms for $k=3$.

the last two terms of the sum $\sum_{n_2=0}^{2m-1} y_{2k+1-n_2}$. Likewise we can symmetrize the variable y_{2l+1} occurring in (2.87) when it multiplies the last two terms of the sum

$$\sum_{n_3=1}^{2l} y_{n_3}.$$

Collecting all these results and using the identities given by (2.63) we find that $\mathcal{L}_{2k+1}$ as given by (2.87) can be written in the equivalent form

$$\begin{aligned}
\mathcal{L}_{2k+1}
&= -\tfrac{1}{2}\prod_{n_1=2}^{k-1}(y_{2n_1-1}^2-1)(y_1+y_2)(y_2+y_3)(y_3+y_4)(y_{2k}+y_{2k+1}) \\
&\quad\times(y_1y_{2k+1}-1)-\tfrac{1}{2}\prod_{n_1=2}^{k-1}(y_{2n_1-1}^2-1)(y_1+y_2)(y_{2k-2}+y_{2k-1}) \\
&\quad\times(y_{2k-1}+y_{2k})(y_{2k}+y_{2k+1})+\sum_{m=2}^{k-1}\prod_{\substack{n_1=2\\ n_1\neq k-m+1}}^{k}(y_{2n_1-1}^2-1) \\
&\quad\times(y_1+y_2)(y_{2k-2m}+y_{2k-2m+1})\Big[-\tfrac{1}{2}(y_1y_{2k+1}-1) \\
&\quad\times(y_{2k-2m+1}+y_{2k-2m+2})(y_{2k-2m+2}+y_{2k-2m+3}) \\
&\quad+(y_1-y_{2k-2m+1})\sum_{n_2=1}^{2m-2}y_{2k+1-n_2}-y_{2k-2m+1}(y_1y_{2k+1}-1) \\
&\quad\times\sum_{n_2=0}^{2m-3}y_{2k+1-n_2}\Big]+\sum_{l=2}^{k-1}(y_{2l+1}+y_{2l+2})(y_{2k}+y_{2k+1}) \\
&\quad\times\prod_{\substack{n_1=2\\ n_1\neq l+1}}^{k}(y_{2n_1-1}^2-1)\Big[-\tfrac{1}{2}(y_1y_{2k+1}-1)(y_{2l-1}+y_{2l}) \\
&\quad\times(y_{2l}+y_{2l+1})+(y_{2k+1}-y_{2l+1})\sum_{n_3=2}^{2l-1}y_{n_3}-y_{2l+1} \\
&\quad\times(y_1y_{2k+1}-1)\sum_{n_3=1}^{2l-2}y_{n_3}\Big]-\sum_{l=1}^{k-1}\sum_{m=1}^{k-l-1}\prod_{\substack{n_1=2\\ n_1\neq k-m+1,\,l+1}}^{k}(y_{2n_1-1}^2-1) \\
&\quad\times(y_{2l+1}+y_{2l+2})(y_{2k-2m}+y_{2k-2m+1})\Big\{\tfrac{1}{2}(y_1y_{2k+1}-1) \\
&\quad\times(y_{2l+1}^2-1)(y_{2k-2m+1}+y_{2k-2m+2})(y_{2k-2m+2}+y_{2k-2m+3}) \\
&\quad+\tfrac{1}{2}(y_1y_{2k+1}-1)(y_{2k-2m+1}^2-1)(y_{2l}+y_{2l-1})(y_{2l}+y_{2l+1}) \\
&\quad-y_{2l+1}y_{2k-2m+1}(y_1-y_{2l+1})(y_{2k+1}-y_{2k-2m+1}) \\
&\quad+(y_1y_{2k+1}-1)(y_{2l+1}^2-1)\,y_{2k-2m+1}\sum_{n_2=0}^{2m-3}y_{2k+1-n_3} \\
&\quad+(y_1y_{2k+1}-1)(y_{2k-2m+1}^2-1)\,y_{2l+1}\sum_{n_3=1}^{2l-2}y_{n_3} \\
&\quad+\sum_{n_2=1}^{2m}y_{2k+1-n_2}\Big[(y_{2l+1}^2-1)(y_{2k-2m+1}-y_1)+y_1y_{2l+1} \\
&\quad\times(y_1-y_{2l+1})(y_{2k-2m+1}y_{2k+1}-1)\Big]+\sum_{n_3=2}^{2l+1}y_{n_3} \\
&\quad\times\Big[(y_{2k-2m+1}^2-1)(y_{2l+1}-y_{2k+1})+y_{2k+1}y_{2k-2m+1}(y_1y_{2l+1}-1) \\
&\quad\times(y_{2k+1}-y_{2k-2m+1})]+(y_1y_{2l+1}-1)\Big] \\
&\quad\times(y_{2k-2m+1}y_{2k+1}-1)\sum_{n_3=2}^{2l+1}y_{n_3}\sum_{n_2=1}^{2m}y_{2k+1-n_2}\Big\}.
\end{aligned} \tag{2.101}$$

The advantage of the representation (2.101) for $\mathcal{L}_{2k+1}$ as opposed to the previous representations [as for example (2.51)] is, for one, the separation of the "end effects" and the "bulk effects" of the integrand. Also the splitting (2.63) has been introduced into (2.101).

2. Summing $\mathcal{L}_{2k+1}(l, m)$ for $m+l=k-1$

For the case $k=3$ the graphs that appear in $\mathcal{L}_7(0,1)$ and $\mathcal{L}_7(1,0)$ when all reductions have been completed [see Fig. 13(a)] can be obtained from the set of graphs for $\mathcal{L}_7(2,0)+\mathcal{L}_7(0,2)+\mathcal{L}_7(1,1)$ [see Fig. 13(b)]. We claim that this is a general result. That is to say, if we sum all $\mathcal{L}_{2k+1}(l,m)$ such that $l+m=k-1$, then from the final reduced form for this sum there is a simple prescription to obtain the remaining terms. Therefore, we examine the sum

$$S_{2k+1}=\sum_{\substack{l=0\\ l+m=k-1}}^{k-1}{}'\sum_{m=0}^{k-l-1}{}'\,\mathcal{L}_{2k+1}(l,m) \tag{2.102}$$

and proceed to reduce this to the desired form [four loops in all graphs and a single $(y_1y_{2k+1}-1)$ factor].

From (2.101) and the definition of S_{2k+1} we have

$$\begin{aligned}
S_{2k+1}
&=(y_1+y_2)(y_2+y_3)\prod_{n_1=3}^{k}(y_{2n_1-1}^2-1)\Big\{-\tfrac{1}{2}(y_1y_{2k+1}-1)(y_3+y_4) \\
&\quad\times(y_4+y_5)+(y_1-y_3)\sum_{n_2=1}^{2k-4}y_{2k+1-n_2}-y_3(y_1y_{2k+1}-1) \\
&\quad\times\sum_{n_2=0}^{2k-5}y_{2k+1-n_2}\Big\}\ +(y_{2k-1}+y_{2k})(y_{2k}+y_{2k+1}) \\
&\quad\times\prod_{n_1=2}^{k-1}(y_{2n_1-1}^2-1)\Big\{-\tfrac{1}{2}(y_1y_{2k+1}-1)(y_{2k-2}+y_{2k-1}) \\
&\quad\times(y_{2k-3}+y_{2k-2})+(y_{2k+1}-y_{2k-1})\sum_{n_3=2}^{2k-3}y_{n_3} \\
&\quad-y_{2k-1}(y_1y_{2k+1}-1)\sum_{n_3=1}^{2k-4}y_{n_3}\Big\}-\sum_{l=1}^{k-2}(y_{2l+1}+y_{2l+2}) \\
&\quad\times(y_{2l+2}+y_{2l+3})\prod_{\substack{n_1=2\\ n_1\neq l+1,\,l+2}}^{k}(y_{2n_1-1}^2-1)\Big\{\tfrac{1}{2}(y_1y_{2k+1}-1) \\
&\quad\times(y_{2l+1}^2-1)(y_{2l+3}+y_{2l+4})(y_{2l+4}+y_{2l+5}) \\
&\quad+\tfrac{1}{2}(y_1y_{2k+1}-1)(y_{2l+3}^2-1)(y_{2l}+y_{2l+1})(y_{2l-1}+y_{2l}) \\
&\quad-y_{2l+1}y_{2l+3}(y_1-y_{2l+1})(y_{2k+1}-y_{2l+3})+(y_1y_{2k+1}-1) \\
&\quad\times(y_{2l+1}^2-1)\,y_{2l+3}\sum_{n_2=0}^{2k-2l-5}y_{2k+1-n_2}+(y_1y_{2k+1}-1)
\end{aligned}$$

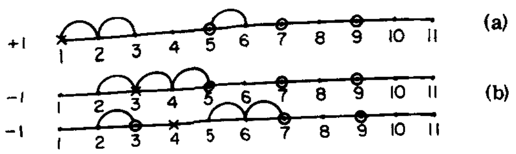

FIG. 14. Some typical terms contributing to J_{2k+1} for $k=5$. (a) Typical graph from $J_{11}^{(1)}$. (b) Two graphs from $J_{11}^{(3)}$.

$$\times(y_{2l+3}^2-1)\,y_{2l+1}\sum_{n_3=1}^{2l-2} y_{n_3}+\sum_{n_2=1}^{2k-2l-2} y_{2k+1-n_2}$$
$$\times[(y_{2l+1}^2-1)(y_{2l+3}-y_1)+y_1y_{2l+1}(y_1-y_{2l+1})$$
$$\times(y_{2l+3}y_{2k+1}-1)]+\sum_{n_3=2}^{2l+1} y_{n_3}[(y_{2l+3}^2-1)(y_{2l+1}-y_{2k+1})$$
$$+y_{2k+1}y_{2l+3}(y_1y_{2l+1}-1)(y_{2k+1}-y_{2l+3})]+(y_1y_{2l+1}-1)$$
$$\times(y_{2l+3}y_{2k+1}-1)\sum_{n_3=2}^{2l+1} y_{n_3}\sum_{n_2=1}^{2k-2l-2} y_{2k+1-n_2}\Bigg\}. \tag{2.103}$$

3. $J_{2k+1}=0$

Recalling the discussion that resulted in the definition of the quantity J_7 [just before (2.66)], we see that an analogous argument for the terms appearing in (2.103) leads to the definition

$$J_{2k+1}=J_{2k+1}^{(1)}+J_{2k+1}^{(2)}+J_{2k+1}^{(3)}+J_{2k+1}^{(4)} \tag{2.104}$$

with

$$J_{2k+1}^{(1)}=(y_1+y_2)(y_2+y_3)(y_1-y_3)\times\prod_{n_1=3}^{k}(y_{2n_1-1}^2-1)\sum_{n_2=1}^{2k-4} y_{2k+1-n_2}, \tag{2.105a}$$

$$J_{2k+1}^{(2)}=(y_{2k-1}+y_{2k})(y_{2k}+y_{2k+1})(y_{2k+1}-y_{2k-1})\times\prod_{n_1=2}^{k-1}(y_{2n_1-1}^2-1)\sum_{n_3=2}^{2k-3} y_{n_3}, \tag{2.105b}$$

$$J_{2k+1}^{(3)}=-\sum_{l=1}^{k-2}(y_{2l+1}+y_{2l+2})(y_{2l+2}+y_{2l+3})\prod_{\substack{n_1=2\\ n_1\neq l+1,\,l+2}}^{k}(y_{2n_1-1}^2-1)\times\sum_{n_2=2}^{2l+1} y_{n_2}(y_{2l+3}^2-1)(y_{2l+1}-y_{2k+1}), \tag{2.105c}$$

and

$$J_{2k+1}^{(4)}=-\sum_{l=1}^{k-2}(y_{2l+1}+y_{2l+2})(y_{2l+2}+y_{2l+3})\times\prod_{\substack{n_1=2\\ n_1\neq l+1,\,l+2}}^{k}(y_{2n_1-1}^2-1)\sum_{n_2=2l+3}^{2k} y_{n_2}(y_{2l+1}^2-1)(y_{2l+3}-y_1) \tag{2.105d}$$

where we changed the labeling in the last sums appearing in (2.105c) and (2.105d). Furthermore the result $J_7=0$ of the previous section leads us to conjecture that

$$J_{2k+1}=0 \tag{2.106}$$

where "=" is interpreted in the generalized sense.

We now prove that (2.106) is true. From an examination of the graphs associated with $J_{2k+1}^{(\alpha)}$, $\alpha=1,2,3$, and 4, it is clear that

$$J_{2k+1}^{(1)}=J_{2k+1}^{(2)} \quad\text{and}\quad J_{2k+1}^{(3)}=J_{2k+1}^{(4)}. \tag{2.107}$$

A typical graph coming from the set of graphs associated with $J_{11}^{(1)}$ is shown in Fig. 14(a) and two types of graphs appearing in $J_{11}^{(3)}$ are displayed in Fig. 14(b). The important point to emphasize is that all graphs associated with $J_{2k+1}^{(1)}$ are such that the three loops appearing in the graph divide the line connecting "1" to "$2k+1$" into two disjoint lines [in Fig. 14(a) the disjoint lines are from 3 to 5 and from 6 to 11]. The graphs associated with $J_{2k+1}^{(3)}$ are of two basic types. There are the graphs that divide the line "1" to "$2k+1$" into two disjoint lines [the first graph in Fig. 14(b) is of this type] and there are the graphs that divide the line into three disjoint lines [the second graph of Fig. 14(b) is this type and the disjoint lines are 1 to 2, 3 to 5, and 7 to 11].

We now claim that the subset of graphs of $J_{2k+1}^{(3)}$ with two disjoint lines exactly cancels all the graphs of $J_{2k+1}^{(1)}$. The remaining graphs of $J_{2k+1}^{(3)}$ that are of the three-line type cancel amongst themselves to give zero. Once these two statements are demonstrated we will have proved (2.106).

We first count the number of two-line graphs in $J_{2k+1}^{(1)}$ and $J_{2k+1}^{(3)}$. In $J_{2k+1}^{(1)}$ there are clearly $2(k-2)$ terms [factor "2" comes from (y_1-y_k)]. The two-line graphs of $J_{2k+1}^{(3)}$ come from the last two terms of the sum $\sum_{n_2=2}^{2l+1} y_{n_2}$ in (2.105c). Thus for fixed l there are two two-line graphs and hence $2(k-2)$ graphs in all. The three-line graphs of $J_{2k+1}^{(3)}$ result from the $\sum_{n_2=2}^{2l-1} y_{n_2}$ terms. Fix the integer l ($\leq k-2$) and let q be one of the values $1,2,\ldots,l-1$. Then one term in (2.105c) can be written as

$$-(y_{2l+1}+y_{2l+2})(y_{2l+2}+y_{2l+3})\times\prod_{\substack{n_1=2\\ n_1\neq l+1,\,l+2}}^{k}(y_{2n_1-1}^2-1)(y_{2q}+y_{2q+1})(y_{2l+3}^2-1)(y_{2l+1}-y_{2k+1}). \tag{2.108}$$

The graphs corresponding to (2.108) are shown in Fig. 15. To these two graphs we consider the complement graphs as shown in Fig. 16 [these are obtained from (2.108) by letting $q\to q$ and $l\to k-1-l+q$]. From Figs. 15 and 16 it is clear that the sum of the diagram and its complement gives zero (the first three-line graph in Fig. 15 is canceled by the second three-line graph in Fig. 16). Hence the sum of all three-line graphs in $J_{2k+1}^{(3)}$ gives zero.

The two-line graphs of $J_{2k+1}^{(3)}$ are of the form

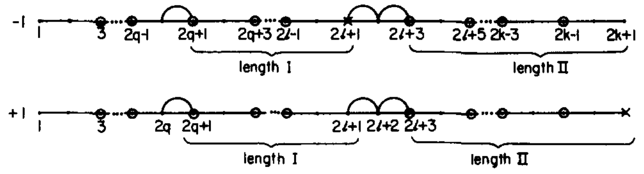

FIG. 15. Three-line graphs of $J_{2k+1}^{(3)}$ for q and l fixed.

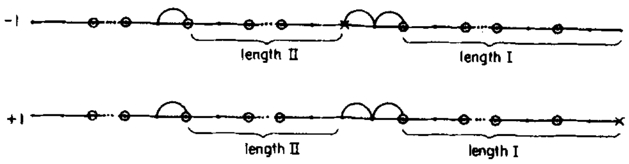

FIG. 16. Complement graph of Fig. 15. The second graph cancels the first graph of Fig. 15.

$$-\sum_{l=1}^{k-2}(y_{2l}+y_{2l+1})(y_{2l+1}+y_{2l+2})(y_{2l+2}+y_{2l+3})$$
$$\times \prod_{\substack{n_1=2\\ n_1\neq l+1}}^{k}(y_{2n_1-1}^2-1)(y_{2l+1}-y_{2k+1}) \qquad (2.109)$$

and are shown in Fig. 17. In Fig. 18 we draw a two-line graph associated with $J_{2k+1}^{(1)}$. If we choose the lengths as shown (which is always possible), then we conclude from a comparison of Figs. 17 and 18 that the two-line graphs of $J_{2k+1}^{(1)}$ and $J_{2k+1}^{(3)}$ cancel to give zero. Thus we have established (2.106).

4. Further cancellation in (2.103)

We examine the terms

$$A_1=\sum_{l=1}^{k-2}(y_{2l+1}+y_{2l+2})(y_{2l+2}+y_{2l+3})\prod_{\substack{n_1=2\\ n_1\neq l+1,\, l+2}}^{k}(y_{2n_1-1}^2-1)$$
$$\times\sum_{n_2=2l+3}^{2k} y_{n_2}\,y_1y_{2l+1}(y_1-y_{2l+1})(y_{2l+3}y_{2k+1}-1) \qquad (2.110)$$

and

$$A_2=\sum_{l=1}^{k-2}(y_{2l+1}+y_{2l+2})(y_{2l+2}+y_{2l+3})\prod_{\substack{n_1=2\\ n_1\neq l+1,\, l+2}}^{k}(y_{2n_1-1}^2-1)$$
$$\times\sum_{n_2=2}^{2l+1} y_{n_2}\,y_{2k+1}y_{2l+3}(y_{2k+1}-y_{2l+1})(y_1y_{2l+1}-1) \qquad (2.111)$$

that appear in (2.103) (note that $\sum_{n_2=1}^{2k-2l-2} y_{2k+1-n_2}=\sum_{n_2=2l+3}^{2k} y_{n_2}$). We now demonstrate that when A_1 and A_2 are used in (2.47) (A_1 and A_2 are parts of $\mathcal{L}_{2k+1}$) and the integration is performed the result is zero. That is to say, we show

$$A_1 = A_2 = 0, \qquad (2.112)$$

where "$=$" is used in the generalized sense.

For fixed l we examine one term in (2.110). If we relabel the integration variables $1\to 2l+1$, $2\to 2l,\ldots,2l\to 2$, and $2l+1\to 1$ while the remaining labels are fixed, then the integrand is antisymmetric in y_1 and y_{2l+1} [recall that we are always implicitly multiplying the factors A_1 and A_2 by $\prod_{j=1}^{2k}(y_j+y_{j+1})^{-1}$] and hence zero. For the term A_2 we relabel the variables $y_{2l+3}\to y_{2k+1}$, $y_{2l+4}\to y_{2k},\ldots,y_{2k}\to y_{2l+4}$, and $y_{2k+1}\to y_{2l+3}$ and note that each term in A_2 is antisymmetric in y_{2k+1} and y_{2l+3}. Hence (2.112) is proved.

5. Final form for S_{2k+1}

Summarizing the results so far we have demonstrated that S_{2k+1} of (2.103) can be written as [this is the generalization of (2.70)]

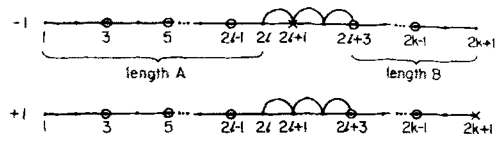

FIG. 17. Two-line graphs of $J_{2k+1}^{(3)}$.

$$S_{2k+1}=(y_1+y_2)(y_2+y_3)\prod_{n_1=3}^{k}(y_{2n_1-1}^2-1)\Big(-\tfrac12(y_1y_{2k+1}-1)$$
$$\times(y_3+y_4)(y_4+y_5)-y_3(y_1y_{2k+1}-1)\sum_{n_2=6}^{2k+1}y_{n_2}\Big)$$
$$+(y_{2k-1}+y_{2k})(y_{2k}+y_{2k+1})\prod_{n_1=2}^{k-1}(y_{2n_1-1}^2-1)$$
$$\times\Big(-\tfrac12(y_1y_{2k+1}-1)(y_{2k-2}+y_{2k-1})(y_{2k-3}+y_{2k-2})$$
$$-y_{2k-1}(y_1y_{2k+1}-1)\sum_{n_2=1}^{2k-4}y_{n_2}\Big)-\sum_{l=1}^{k-2}(y_{2l+1}+y_{2l+2})$$
$$\times(y_{2l+2}+y_{2l+3})\prod_{\substack{n_1=2\\ n_1\neq l+1,\, l+2}}^{k}(y_{2n_1-1}^2-1)$$
$$\times\Big(\tfrac12(y_1y_{2k+1}-1)(y_{2l+1}^2-1)(y_{2l+3}+y_{2l+4})(y_{2l+4}+y_{2l+5})$$
$$+\tfrac12(y_1y_{2k+1}-1)(y_{2l+3}^2-1)(y_{2l}+y_{2l+1})(y_{2l-1}+y_{2l})$$
$$-y_{2l+1}y_{2l+3}(y_1-y_{2l+1})(y_{2k+1}-y_{2l+3})$$
$$+(y_1y_{2k+1}-1)(y_{2l+1}^2-1)y_{2l+3}$$
$$\times\sum_{n_2=2l+6}^{2k+1}y_{n_2}+(y_1y_{2k+1}-1)(y_{2l+3}^2-1)y_{2l+1}$$
$$\times\sum_{n_2=1}^{2l-2}y_{n_2}+(y_1y_{2l+1}-1)(y_{2l+3}y_{2k+1}-1)$$
$$\times\sum_{n_3=2}^{2l+1}y_{n_3}\sum_{n_2=2l+3}^{2k}y_{n_2}\Big). \qquad (2.113)$$

The generalization of the term in (2.71) is

$$\sum_{l=1}^{k-2}(y_{2l+1}+y_{2l+2})(y_{2l+2}+y_{2l+3})\prod_{\substack{n_1=2\\ n_1\neq l+1,\, l+2}}^{k}(y_{2n_1-1}^2-1)$$
$$\times\Big[y_{2l+1}y_{2l+3}(y_1-y_{2l+1})(y_{2k+1}-y_{2l+3})$$
$$-(y_1y_{2l+1}-1)(y_{2l+3}y_{2k+1}-1)\sum_{n_2=2}^{2l+1}y_{n_2}\sum_{n_3=2l+3}^{2k}y_{n_3}\Big]. \qquad (2.114)$$

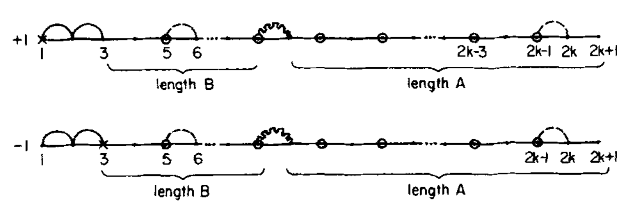

FIG. 18. Two-line graphs of $J_{2k+1}^{(1)}$. Moving loop (denoted by wavy line) starts at 5 and goes to $2k-1$ (shown in dotted lines).

As was done in going from (2.72) to (2.74), we want to write the first term in square brackets in (2.114) in a form so that the two sum-terms appearing in the second term in (2.114) become a common factor. To do this we write

$$(y_1 - y_{2l+1})(y_{2k+1} - y_{2l+3}) = \left(\sum_{n_2=1}^{2l} y_{n_2} - \sum_{n_2=2}^{2l+1} y_{n_2}\right)\left(\sum_{n_3=2l+4}^{2k+1} y_{n_3} - \sum_{n_3=2l+3}^{2k} y_{n_3}\right) \tag{2.115}$$

and examine the graphs of

$$y_{2l+1}y_{2l+3}(y_{2l+1}+y_{2l+2})(y_{2l+2}+y_{2l+3}) \times \prod_{\substack{n_1=2\\ n_1\neq l+1,\, l+2}}^{k} (y_{2n_1-1}^2 - 1)(y_1 - y_{2l+1})(y_{2k+1}-y_{2l+3}). \tag{2.116}$$

It is clear that we can relabel the integration variable subscripts so that the following is true [when used in (2.116), which in turn will be used in (2.47)]:

$$y_{2l+1}y_{2l+3}(y_1 - y_{2l+1})(y_{2k+1}-y_{2l+3}) = (y_1 - y_{2l+1})(y_{2k+1}-y_{2l+3}) \sum_{n_2=2}^{2l+1} y_{n_2} \sum_{n_3=2l+3}^{2k} y_{n_3}. \tag{2.117}$$

Using (2.117) in (2.114) we obtain

$$\sum_{l=1}^{k-2} (y_{2l+1}+y_{2l+2})(y_{2l+2}+y_{2l+3}) \prod_{\substack{n_1=2\\ n_1\neq l+1,\, l+2}}^{k} (y_{2n_1-1}^2-1) \sum_{n_2=2}^{2l+1} y_{n_2} \times \sum_{n_3=2l+3}^{2k} y_{n_3}[-(y_1y_{2k+1}-1)(y_{2l+1}y_{2l+3}-1) - y_1(y_{2l+3}-y_{2l+1}) - y_{2k+1}(y_{2l+1}-y_{2l+3})], \tag{2.118}$$

where we used the algebraic identity

$$-(y_1 - y_{2l+1})(y_{2k+1}-y_{2l+3}) + (y_1y_{2l+1}-1)(y_{2l+3}y_{2k+1}-1) = (y_1y_{2k+1}-1)(y_{2l+1}y_{2l+3}-1) + y_1(y_{2l+3}-y_{2l+1}) + y_{2k+1}(y_{2l+1}-y_{2l+3}). \tag{2.119}$$

It is clear that the last two terms in (2.118) give zero as the second and third terms will lead to an integrand that is antisymmetric in y_{2l+3} and y_{2l+1}. Also by relabeling we can let

$$\sum_{n_2=2}^{2l+1} y_{n_2} \sum_{n_3=2l+3}^{2k} y_{n_3} \to \sum_{n_2=1}^{2l} y_{n_2} \sum_{n_3=2l+4}^{2k+1} y_{n_3}.$$

Hence S_{2k+1} of (2.113) becomes

$$\begin{aligned} S_{2k+1} = &-(y_1y_{2k+1}-1)(y_1+y_2)(y_2+y_3)\prod_{n_1=3}^{k}(y_{2n_1-1}^2-1) \\ &\times \left[\tfrac{1}{2}(y_3+y_4)(y_4+y_5) + y_3 \sum_{n_2=6}^{2k+1} y_{n_2}\right] - (y_1y_{2k+1}-1) \\ &\times (y_{2k-1}+y_{2k})(y_{2k}+y_{2k+1}) \prod_{n_1=2}^{k-1}(y_{2n_1-1}^2-1) \\ &\times \left[\tfrac{1}{2}(y_{2k-3}+y_{2k-2})(y_{2k-2}+y_{2k-1}) + y_{2k-1}\sum_{n_2=1}^{2k-4} y_{n_2}\right] \\ &- (y_1y_{2k+1}-1)\sum_{l=1}^{k-2}(y_{2l+1}+y_{2l+2})(y_{2l+2}+y_{2l+3}) \\ &\times \prod_{\substack{n_1=2\\ n_1\neq l+1,\, l+2}}^{k}(y_{2n_1-1}^2-1)\Big[\tfrac{1}{2}(y_{2l+1}^2-1)(y_{2l+3}+y_{2l+4}) \\ &\times (y_{2l+4}+y_{2l+5}) + \tfrac{1}{2}(y_{2l+3}^2-1)(y_{2l-1}+y_{2l})(y_{2l}+y_{2l+1}) \\ &+ (y_{2l+1}^2-1)y_{2l+3}\sum_{n_2=2l+6}^{2k+1} y_{n_2} + (y_{2l+3}^2-1)y_{2l+1} \\ &\times \sum_{n_2=1}^{2l-2} y_{n_2} + (y_{2l+1}y_{2l+3}-1)\sum_{n_2=1}^{2l} y_{n_2}\sum_{n_3=2l+4}^{2k+1} y_{n_3}\Big]. \end{aligned} \tag{2.120}$$

This can be written more compactly by combining the first two terms in (2.120) into the $l=0$ and $l=k-1$ terms of the third term. Doing this (2.120) becomes

$$\begin{aligned} S_{2k+1} = &-(y_1y_{2k+1}-1)\sum_{l=0}^{k-1}(y_{2l+1}+y_{2l+2})(y_{2l+2}+y_{2l+3}) \\ &\times \Bigg[\prod_{\substack{n_1=2\\ n_1\neq l+2}}^{k}(y_{2n_1-1}^2-1)\tfrac{1}{2}(y_{2l+3}+y_{2l+4})(y_{2l+4}+y_{2l+5}) \\ &+ \prod_{\substack{n_1=2\\ n_1\neq l+1}}^{k}(y_{2n_1-1}^2-1)\tfrac{1}{2}(y_{2l-1}+y_{2l})(y_{2l}+y_{2l+1}) \\ &+ \prod_{\substack{n_1=2\\ n_1\neq l+2}}^{k}(y_{2n_1-1}^2-1)y_{2l+3}\sum_{n_2=2l+6}^{2k+1}y_{n_2} + \prod_{\substack{n_1=2\\ n_1\neq l+1}}^{k}(y_{2n_1-1}^2-1) \\ &\times y_{2l+1}\sum_{n_2=1}^{2l-2}y_{n_2} + \prod_{\substack{n_1=2\\ n_1\neq l+1,\, l+2}}^{k}(y_{2n_1-1}^2-1)(y_{2l+1}y_{2l+3}-1) \\ &\times \sum_{n_2=1}^{2l}y_{n_2}\sum_{n_3=2l+4}^{2k+1}y_{n_3}\Bigg], \end{aligned} \tag{2.121}$$

where we must have the convention that any product term

$$\prod_{\substack{n_1=2\\ n_1\neq l+1}}^{k}(y_{2n_1-1}^2-1)$$

is zero for $l=0$,

$$\prod_{\substack{n_1=2\\ n_1\neq l+2}}^{k}(y_{2n_1-1}^2-1)$$

is zero for $l=k-1$, and

$$\prod_{\substack{n_1=2\\ n_1\neq l+1,\, l+2}}^{k}(y_{2n_1-1}^2-1)$$

is zero for either $l=0$ or $l=k-1$.

Consider the term

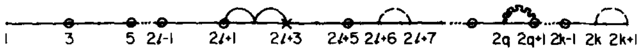

FIG. 19. Graph of (2.122) for a particular value of q. The range over which q varies is indicated by dotted loops.

$$(y_{2l+1}+y_{2l+2})(y_{2l+2}+y_{2l+3}) \prod_{\substack{n_1=2\\ n_1\neq l+2}}^{k} (y_{2n_1-1}^2-1)\, y_{2l+3} \times \sum_{q=l+3}^{k}(y_{2q}+y_{2q+1}) \tag{2.122}$$

which is the third term in (2.121). For a particular q the graph of (2.122) is shown in Fig. 19. From the figure it is clear that the integrand is divided into three parts. We examine the integrand associated with the graph between the double loop and the single loop. It is

$$\left[y_{2l+3}\prod_{n=l+3}^{q}(y_{2n-1}^2-1)\right]\prod_{j=2l+3}^{2q-1}(y_j+y_{j+1})^{-1}. \tag{2.123}$$

We claim that this integrand factor can be replaced by

$$\begin{aligned}\tfrac{1}{2}\Bigg[&\sum_{p=0}^{q-l-2}(y_{2l+3+2p}+y_{2l+4+2p})\prod_{n=l+2}^{l+1+p}(y_{2n}^2-1)\prod_{n=l+3+p}^{q}(y_{2n-1}^2-1)\\ &-\sum_{p=0}^{q-l-3}(y_{2l+4+2p}+y_{2l+5+2p})\prod_{\substack{n=l+3\\ n\neq l+3+p}}^{q}(y_{2n-1}^2-1)\\ &\times(y_{2l+3}y_{2l+5+2p}-1)\Bigg]\prod_{j=2l+3}^{2q-1}(y_j+y_{j+1})^{-1},\end{aligned} \tag{2.124}$$

where the product symbol is to be interpreted as unity if the upper index is less than the lower index. To prove (2.124) we start with the algebraic identity

$$\begin{aligned}y_{2l+3}\prod_{n=l+3}^{q}(y_{2n-1}^2-1)&=\sum_{p=0}^{q-l-2}(y_{2l+3+2p}+y_{2l+4+2p})\prod_{n=l+2}^{l+1+p}(y_{2n}^2-1)\prod_{n=l+3+p}^{q}(y_{2n-1}^2-1)\\ &-\sum_{p=0}^{q-l-3}(y_{2l+4+2p}+y_{2l+5+2p})(y_{2l+4+2p}\,y_{2l+5+2p}-1)\\ &\times\prod_{n=l+2}^{l+1+p}(y_{2n}^2-1)\prod_{n=l+p+4}^{q}(y_{2n-1}^2-1)-y_{2q}\prod_{n=l+2}^{q}(y_{2n}^2-1)\end{aligned} \tag{2.125}$$

and note that when used as an integrand [multiplied by $\prod_{j=2l+3}^{2q-1}(y_j+y_{k+1})^{-1}$] the last term in (2.125) is equivalent to the term on the left-hand side of (2.125). Hence (2.123) is equivalent to

$$\begin{aligned}\tfrac{1}{2}\Bigg[&\sum_{p=0}^{q-l-2}(y_{2l+3+3p}+y_{2l+4+2p})\prod_{n=l+2}^{l+1+p}(y_{2n}^2-1)\prod_{n=l+3+p}^{q}(y_{2n-1}^2-1)\\ &-\sum_{p=0}^{q-l-3}(y_{2l+4+2p}+y_{2l+5+2p})(y_{2l+4+2p}\,y_{2l+5+2p}-1)\\ &\times\prod_{n=l+2}^{l+1+p}(y_{2n}^2-1)\prod_{n=l+p+4}^{q}(y_{2n-1}^2-1)\Bigg]\prod_{j=2l+3}^{2q-1}(y_j+y_{j+1})^{-1}.\end{aligned} \tag{2.126}$$

By the change of variables $2l+4+2p\to 2l+3$, $2l+3+2p\to 2l+4,\ldots, 2l+4\to 2l+3+2p$, and $2l+3\to 2l+4+2p$ in the last term in (2.126) and remembering that the quantity $\prod_{j=2l+3}^{2q-1}(y_j+y_{j+1})^{-1}(y_{2l+4+2p}+y_{2l+5+2p})$ remains invariant under this change of labels we obtain the equivalent expression (2.124).

Similarly we have that the quantity

$$\left[y_{2l+1}\prod_{n=q+1}^{l}(y_{2n-1}^2-1)\right]\prod_{j=2q}^{2l}(y_j+y_{j+1})^{-1} \tag{2.127}$$

occurring in (2.121) can be replaced by

$$\begin{aligned}\tfrac{1}{2}\Bigg[&\sum_{p=0}^{l-q}(y_{2q+2p}+y_{2q+2p+1})\prod_{n=p+q+1}^{l}(y_{2n}^2-1)\prod_{n=q+1}^{q+p}(y_{2n-1}^2-1)\\ &-\sum_{p=0}^{l-q-1}(y_{2q+1+2p}+y_{2q+2+2p})(y_{2q+1+2p}\,y_{2l+1}-1)\\ &\times\prod_{\substack{n=q+1\\ n\neq p+q+1}}^{l}(y_{2n-1}^2-1)\Bigg]\prod_{j=2q}^{2l}(y_j+y_{j+1})^{-1}.\end{aligned} \tag{2.128}$$

Using these results S_{2k+1} of (2.121) becomes (see Fig. 20)

$$\begin{aligned}S_{2k+1}=&-(y_1y_{2k+1}-1)\sum_{l=0}^{k-1}(y_{2l+1}+y_{2l+2})(y_{2l+2}+y_{2l+3})\\ &\times\Bigg\{\prod_{\substack{n_1=2\\ n_1\neq l+2}}^{k}(y_{2n_1-1}^2-1)\tfrac{1}{2}(y_{2l+3}+y_{2l+4})(y_{2l+4}+y_{2l+5})\\ &+\prod_{\substack{n_1=2\\ n_1\neq l+1}}^{k}(y_{2n_1-1}^2-1)\tfrac{1}{2}(y_{2l-1}+y_{2l})(y_{2l}+y_{2l+1})\\ &+\prod_{n=2}^{l+1}(y_{2n-1}^2-1)\sum_{q=l+3}^{k}(y_{2q}+y_{2q+1})\tfrac{1}{2}\\ &\times\Bigg[\sum_{p=0}^{q-l-2}(y_{2l+3+2p}+y_{2l+4+2p})\prod_{n=l+2}^{l+1+p}(y_{2n}^2-1)\\ &\times\prod_{n=l+3+p}^{k}(y_{2n-1}^2-1)-\sum_{p=0}^{q-l-3}(y_{2l+4+2p}+y_{2l+5+2p})\\ &\times\prod_{\substack{n=l+3\\ n\neq l+3+p}}^{k}(y_{2n-1}^2-1)(y_{2l+3}y_{2l+5+2p}-1)\Bigg]\\ &+\prod_{n=l+2}^{k}(y_{2n-1}^2-1)\sum_{q=1}^{l-1}(y_{2q-1}+y_{2q})\tfrac{1}{2}\\ &\times\Bigg[\sum_{p=0}^{l-q}(y_{2q+2p}+y_{2q+2p+1})\prod_{n=2}^{q+p}(y_{2n-1}^2-1)\\ &\times\prod_{n=p+q+1}^{l}(y_{2n}^2-1)-\sum_{p=0}^{l-q-1}(y_{2q+1+2p}+y_{2q+2+2p})\\ &\times(y_{2q+1+2p}\,y_{2l+1}-1)\prod_{\substack{n=2\\ n\neq p+q+1}}^{l}(y_{2n-1}^2-1)\Bigg]\\ &+\prod_{\substack{n=2\\ n\neq l+1,\,l+2}}^{k}(y_{2n-1}^2-1)(y_{2l+1}y_{2l+3}-1)\sum_{n=1}^{2l}y_n\sum_{n=2l+4}^{2k+1}y_n\Bigg\}.\end{aligned} \tag{2.129}$$

We now consider the term

$$I_{2k+1}=\sum_{l=0}^{k-1}(y_{2l+1}+y_{2l+2})(y_{2l+2}+y_{2l+3})\Bigg\{-\tfrac{1}{2}\prod_{n=2}^{l+1}(y_{2n-1}^2-1)$$

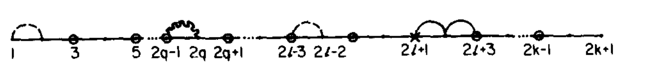

FIG. 20. Graph of one term of (2.129). Dotted loops indicate range of the wavy loop.

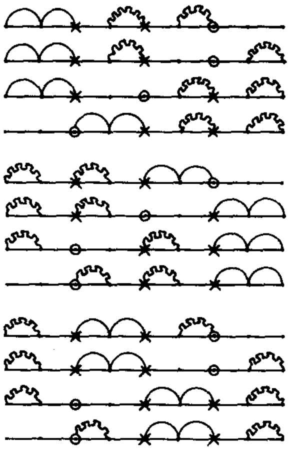

FIG. 21. Graphs associated with I_9. The first eight graphs are multiplied by $-\frac{1}{2}$ and the last four graphs by $+1$.

$$\times \sum_{q=l+3}^{k} (y_{2q}+y_{2q+1}) \sum_{p=0}^{q-l-3} (y_{2l+4+2p}+y_{2l+5+2p})$$

$$\times \prod_{\substack{n=l+3\\ n\neq l+3+p}}^{k} (y_{2n-1}^2-1)(y_{2l+3}y_{2l+5+2p}-1) - \tfrac{1}{2}\prod_{n=l+2}^{k}(y_{2n-1}^2-1)$$

$$\times \sum_{q=1}^{l-1}(y_{2q-1}+y_{2q})\sum_{p=0}^{l-q-1}(y_{2q+1+2p}+y_{2q+2+2p})$$

$$\times \prod_{\substack{n=2\\ n\neq p+q+1}}^{l}(y_{2n-1}^2-1)(y_{2q+1+2p}y_{2l+1}-1) + \prod_{\substack{n=2\\ n\neq l+1,\, l+2}}^{k}(y_{2n-1}^2-1)$$

$$\times (y_{2l+1}y_{2l+3}-1)\sum_{n=1}^{2l} y_n \sum_{n=2l+4}^{2k+1} y_n \Bigg\} \tag{2.130}$$

which is part of S_{2k+1}. We claim that

$$I_{2k+1}=0, \tag{2.131}$$

where equality is in the generalized sense. We first examine a special case. Consider $k=4$, then there are twelve terms in (2.130) and the graphs of these terms are shown in Fig. 21.

Concerning the terms with the structure $(y_\alpha y_\beta - 1)$ we indicate by "×" the presence of the y_α and y_β terms. From an examination of Fig. 21 it is clear that graphs 5–8 are just a reversed labeling of the first four graphs. Hence we need only consider the first four graphs with weight -1 and the last four graphs. However, it is clear from Fig. 21 that the last four graphs have the same structure as do the first four graphs. Hence they add to give zero, i.e., $I_9=0$.

The general case proceeds along similar lines. Some typical graphs are shown in Fig. 22. As in the $k=4$ case, the terms arising from the second term in (2.130) can be combined with the first term as they are

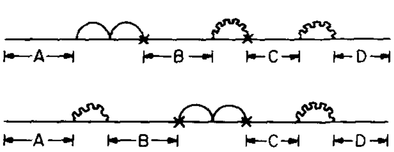

FIG. 22. Typical graphs of I_{2k+1}. The first graph comes from the first set of terms and the second graph from the third set of terms.

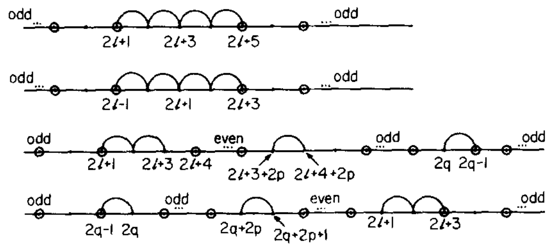

FIG. 23. Graphs of terms appearing in S_{2k+1} as defined by (2.132). The labels "odd" and "even" refer to whether (y_α^2-1) occurs at an odd site or an even site, respectively.

of the same structure. Furthermore, if the distances $A, B, C,$ and D as depicted in Fig. 22 are made equal, then the two graphs cancel. One need only check that the (y_α^2-1) type terms are in the correct place. An examination of (2.130) convinces oneself that they are in the correct places for cancellation. Hence (2.131) follows, and incorporating this result into (2.129) results in our final form for S_{2k+1},

$$S_{2k+1} = -\tfrac{1}{2}(y_1y_{2k+1}-1)\sum_{l=0}^{k-1}(y_{2l+1}+y_{2l+2})(y_{2l+2}+y_{2l+3})$$

$$\times \Bigg\{ \prod_{\substack{n=2\\ n\neq l+2}}^{k}(y_{2n-1}^2-1)(y_{2l+3}+y_{2l+4})(y_{2l+4}+y_{2l+5})$$

$$+ \prod_{\substack{n=2\\ n\neq l+1}}^{k}(y_{2n-1}^2-1)(y_{2l-1}+y_{2l})(y_{2l}+y_{2l+1})$$

$$+ \prod_{n=2}^{l+1}(y_{2n-1}^2-1)\sum_{q=l+3}^{k}(y_{2q}+y_{2q+1})\sum_{p=0}^{q-l-2}(y_{2l+3+2p}+y_{2l+4+2p})$$

$$\times \prod_{n=l+2}^{l+1+p}(y_{2n}^2-1)\prod_{n=l+3+p}^{k}(y_{2n-1}^2-1) + \prod_{n=l+2}^{k}(y_{2n-1}^2-1)$$

$$\times \sum_{q=1}^{l-1}(y_{2q-1}+y_{2q})\sum_{p=0}^{l-q}(y_{2q+2p}+y_{2q+1+2p})$$

$$\times \prod_{n=2}^{p+q}(y_{2n-1}^2-1)\prod_{n=p+q+1}^{l}(y_{2n}^2-1)\Bigg\}. \tag{2.132}$$

In Fig. 23 we display a graph of a typical term from each of the four basic terms in (2.132). In the last two graphs the "even" and "odd" structure of (y_α^2-1) should be noted.

6. Final form for $\mathcal{L}_{2k+1}$

Equation (2.132) is the result of summing $\mathcal{L}_{2k+1}(l,m)$ subject to the restriction $l+m=k-1$. We now claim that $\mathcal{L}_{2k+1}$ [(defined by (2.86)] is in fact

$$\mathcal{L}_{2k+1} = -\tfrac{1}{2}(y_1y_{2k+1}-1)\sum_{l=0}^{k-1}(y_{2l+1}+y_{2l+2})\sum_{m=0}^{k-1-l}(y_{2l+2+2m}+y_{2l+3+2m})$$

$$\times \prod_{n=l+2}^{l+m+1}(y_{2n-1}^2-1)\Bigg\{\prod_{\substack{n=2\\ n\neq l+2,\ldots,l+m+2}}^{k}(y_{2n-1}^2-1)(y_{2l+3+2m}$$

$$+y_{2l+4+2m})(y_{2l+4+2m}+y_{2l+5+2m}) + \prod_{\substack{n=2\\ n\neq l+1,l+2,\ldots,l+m+1}}^{k}$$

$$(y_{2n-1}^2-1)(y_{2l-1}+y_{2l})(y_{2l}+y_{2l+1}) + \prod_{n=2}^{l+1}(y_{2n-1}^2-1)$$

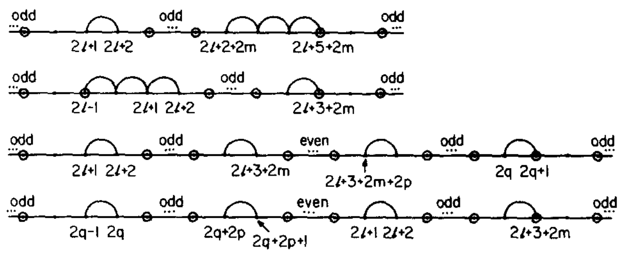

FIG. 24. Graphs of terms appearing in L_{2k+1} as given by (2.133).

$$\times \sum_{q=l+3}^{k} (y_{2q}+y_{2q+1}) \sum_{p=0}^{q-l-2} (y_{2l+3+2m+2p}+y_{2l+4+2m+2p})$$
$$\times \prod_{n=l+m+2}^{l+m+1+p} (y_{2n}^2-1) \prod_{n=l+m+3+p}^{k} (y_{2n-1}^2-1) + \prod_{n=l+2+m}^{k} (y_{2n-1}^2-1)$$
$$\times \sum_{q=1}^{l-1} (y_{2q-1}+y_{2q}) \sum_{p=0}^{l-q} (y_{2q+2p}+y_{2q+1+2p}) \prod_{n=2}^{q+p} (y_{2n-1}^2-1)$$
$$\times \prod_{n=p+q+1}^{l} (y_{2n}^2-1)\Bigg\}. \qquad (2.133)$$

A graph of a typical term from each of the four basic terms of (2.133) is shown in Fig. 24. Figure 24 should be compared with Fig. 23, the $m=0$ case of Fig. 24.

To demonstrate (2.133) one can proceed in two ways. The first method is to repeat the analysis starting at (2.101) leading to (2.132) where now $l+m$ is fixed to be less than $k-1$. This will result in (2.133). Alternatively one can study special cases and note that the general case (2.133) is obtained from the specific case (2.132) by letting $l \to l+m$ in certain terms containing the index l. These special cases indicate the transition from Fig. 23 to Fig. 24.

7. Proof of (2.38)

Using (2.133) in (2.47) we obtain an integral representation of the left-hand side of (2.38). We now compare this with the integral representation of the right-hand side of (2.38) and demonstrate that the two representations are identical. This will establish (2.38) as an identity which in view of Lemma 2.1 proves Theorem 1.

Consider the right-hand side of (2.38). A graph of a typical term is shown in Fig. 25 (the labeling is first g_{2l+1}, then g_{2m+1}, and finally $L_\nu g_{2(k-m-l)-1}$ where we use Lemma 2.3 for this last term). This graph can be made equivalent to the last graph of Fig. 24 by rearranging the graph in the order 1—2—3—4—5 as indicated in the figure. The factor "2" on the right-hand side comes about since there are two graphs in Fig. 24 to each graph in Fig. 25. The first two graphs are a degenerate form of Fig. 25 graphs.

Thus Theorem 1 is proved.

III. THEOREM 2 AND THE FUNCTION $\psi(t;\nu,\lambda)$

A. Differential equation and the functions $\psi_{2n+1}(t;\nu)$

We define $\psi(t;\nu,\lambda)$ by Eq. (1.7). In terms of the function $G(t;\nu,\lambda)$ the definition of $\psi(t;\nu,\lambda)$ is

$$G(t;\nu,\lambda)=\tanh[\tfrac{1}{2}\psi(t;\nu,\lambda)]. \qquad (3.1)$$

From either (1.7) or (3.1) and either (1.3) or (2.7) it follows that $\psi(t;\nu,\lambda)$ satisfies the differential equation

$$\psi''+\frac{1}{t}\psi'=\tfrac{1}{2}\sinh(2\psi)+\frac{2\nu}{t}\sinh(\psi) \qquad (3.2)$$

with

$$\psi(t;\nu,\lambda)\sim 2g_1(t;\nu)\lambda \qquad (3.3)$$

as t approaches infinity along the positive t axis.

The λ expansion of the function $G(t;\nu,\lambda)$ [see Eq. (1.5)] induces a corresponding λ expansion for the function $\psi(t;\nu,\lambda)$,

$$\psi(t;\nu,\lambda)=\sum_{n=0}^{\infty}\lambda^{2n+1}\psi_{2n+1}(t;\nu). \qquad (3.4)$$

The defining relation (3.1) in conjunction with (1.5) requires that

$$g_1(t;\nu)=\tfrac{1}{2}\psi_1(t;\nu), \qquad (3.5a)$$
$$g_3(t;\nu)=\tfrac{1}{2}\psi_3(t;\nu)-\tfrac{1}{3}[\tfrac{1}{2}\psi_1(t;\nu)]^3, \qquad (3.5b)$$
$$g_5(t;\nu)=\tfrac{1}{2}\psi_5(t;\nu)+[\tfrac{1}{2}\psi_1(t;\nu)]^2[\tfrac{1}{2}\psi_3(t;\nu)]+\tfrac{2}{15}[\tfrac{1}{2}\psi_1(t;\nu)]^5, \qquad (3.5c)$$

etc.

The content of Theorem 2 is the assertion that the functions $\psi_{2n+1}(t;\nu)$ as defined by (3.1)—(3.5) possess the representation (1.9). To prove Theorem 2 we define $\psi(t;\nu,\lambda)$ by (1.8) and (1.9) and demonstrate that either (3.1) or (3.2) is true. We choose to demonstrate (3.1).

If (3.1) is true, then it certainly follows that

$$\frac{\partial G}{\partial\lambda}=\tfrac{1}{2}\operatorname{sech}^2[\tfrac{1}{2}\psi]\frac{\partial\psi}{\partial\lambda}$$
$$=\tfrac{1}{2}[1-\tanh^2(\tfrac{1}{2}\psi)]\frac{\partial\psi}{\partial\lambda}$$
$$=\tfrac{1}{2}[1-G^2(t;\nu,\lambda)]\frac{\partial\psi}{\partial\lambda}. \qquad (3.6)$$

With the boundary condition

$$G(t;\nu,0)=1 \qquad (3.7)$$

and the assumption that (3.6) is true, it follows that (3.1) is true. Equation (3.6) can be written in the equivalent form

$$\tfrac{1}{2}(2k+1)\psi_{2k+1}(t;\nu)$$
$$=(2k+1)g_{2k+1}(t;\nu)+\sum_{m=0}^{k-1}\tfrac{1}{2}[2(k-m)-1]\psi_{2(k-m)-1}(t;\nu)$$
$$\times\sum_{l=0}^{m} g_{2l+1}(t;\nu)g_{2(m-l)+1}(t;\nu). \qquad (3.8)$$

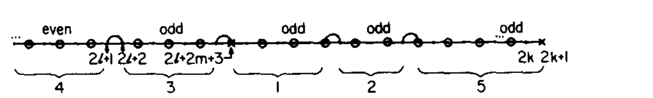

FIG. 25. Graph of a typical term from the right-hand side of (2.38). The numbers beneath the graph represent the ordering to be followed to show equivalence with the graphs of Fig. 24 (in this case the last graph).

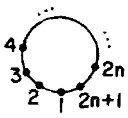

FIG. 26. Quantity (3.11).

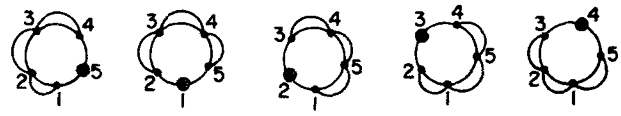

FIG. 28. Quantity $G_5(3)$.

B. Graphs and a lemma

The defining equation (1.9b) can be written in a slightly different form

$$\psi_{2n+1}(t;\nu)=\frac{2}{2n+1}\int_1^\infty dy_1\cdots\int_1^\infty dy_{2n+1}\times\prod_{j=1}^{2n+1}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^\nu\frac{1}{y_j+y_{j+1}}\times\left[\prod_{j=1}^{2n+1}(y_j+1)+\prod_{j=1}^{2n+1}(y_j-1)\right]. \quad (3.9)$$

To prove (3.8) it will be useful to rewrite the term

$$\prod_{j=1}^{2n+1}(y_j+1)+\prod_{j=1}^{2n+1}(y_j-1) \quad (3.10)$$

appearing in (3.9) in a different form. To help visualize the structure of these terms a graphical representation will now be introduced.

Since the factor $(y_{2n+1}+y_1)$ appears in the denominator of the integrand of (3.9), the linear graphs introduced in Sec. II are not the most convenient. We use circular graphs to emphasize the cyclic nature of the integrand in (3.9). Thus the factor

$$\prod_{j=1}^{2n+1}(y_j+y_{j+1})^{-1} \quad (3.11)$$

is represented by a circular graph of $2n+1$ points (see Fig. 26). We adopt the same rules as in Sec. II concerning "loops" and "circles."

Thus the integrand factor

$$(y_2^2-1)(y_5^2-1)(y_3+y_4)\prod_{j=1}^{5}(y_j+y_{j+1})^{-1} \quad (3.12)$$

has the graph shown in Fig. 27. As in Sec. II we omit the term $\Pi_{j=1}^{2n+1}(y_j+y_{j+1})^{-1}$ that multiplies the various factors in (3.9). Thus, for example, when we speak of the graph of the factor

$$(y_2^2-1)(y_5^2-1)(y_3+y_4) \quad (3.13)$$

that appears in an integrand with five variables we always mean (3.12).

Furthermore for the graphs considered in this section we make the additional restrictions:

(i) All graphs have an odd number of points.

(ii) All graphs have an odd number of loops.

(iii) Following any loop there immediately follows another loop or a circle. If a circle follows, then either a loop or a point must follow this circle. If a point follows, then a circle must follow this point.

(iv) The sign of a graph is defined to be $(-1)^{N_c}$ where N_c is the number of circles appearing in the graph. The integrand associated with the graph carries this sign. As a result of the above rule, $N_c=\frac{1}{2}(K-L)$ where K is the total number of points of the graph and L is the number of loops in the graph.

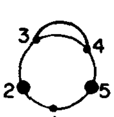

FIG. 27. Quantity (3.12).

With these restrictions in mind we make the following definition:

$G_K(L)=$ the sum of all labeled graphs of K points with L loops. (3.14)

As an example the set of graphs $G_5(3)$ is shown in Fig. 28.

We use the word graph and the integrand associated with such a graph interchangeably. With this understanding we now prove

Lemma 3.1:

$$\prod_{j=1}^{2k+1}(y_j+1)+\prod_{j=1}^{2k+1}(y_j-1)=\sum_{j=0}^{k}G_{2k+1}(2j+1). \quad (3.15)$$

Proof: At any site in a graph of $2k+1$ points there are five different configurations at this site (see Fig. 29). We represent each possible configuration at a site by a vector:

$$|LL\rangle=\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix},\quad |L_R\rangle=\begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix},\quad |L_l\rangle=\begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix},$$
$$|P\rangle=\begin{pmatrix}0\\0\\0\\1\\0\end{pmatrix},\quad |C\rangle=\begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix}. \quad (3.16)$$

Now consider the points j and $j+1$ in a graph. We assume j has the configuration $|\alpha\rangle$ and $j+1$ has the con-

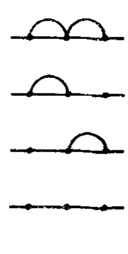

FIG. 29. Five distinct configurations at a site.

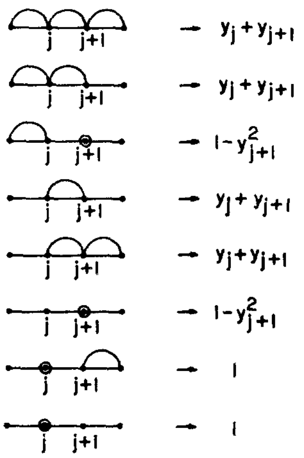

FIG. 30. All possible configurations at sites j and $j+1$ with their respective weights. From this Fig. (3.18) follows.

figuration $|\beta\rangle$ ($\alpha, \beta = LL$, L_R, L_l, C, or P). To this part of the graph we assign in accordance with the above rules an integrand factor. We denote this factor by

$$\langle\alpha|M(j,j+1)|\beta\rangle. \tag{3.17}$$

Using the graphical rules we have (see Fig. 30)

$$M(j,j+1)=\begin{bmatrix} y_j+y_{j+1} & 0 & y_j+y_{j+1} & 0 & 0 \\ y_j+y_{j+1} & 0 & y_j+y_{j+1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-y_{j+1}^2 \\ 0 & 0 & 0 & 0 & 1-y_{j+1}^2 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}. \tag{3.18}$$

To avoid double counting the circles we assign the weight one when they occur at site j and the weight $1-y_{j+1}^2$ when they occur at site $j+1$. Then we have

$$\sum_{j=0}^{k} G_{2k+1}(2j+1) = \mathrm{Tr}\left\{\prod_{j=1}^{2k+1} M(j,j+1)\right\}, \tag{3.19}$$

where $y_{2k+2} \equiv y_1$.

If we make the similarity transformation

$$\tilde{M}(j,j+1) = UM(j,j+1)U^{-1}, \tag{3.20}$$

where

$$U=\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{3.21}$$

then

$$\tilde{M}(j,j+1)=\begin{bmatrix} y_j+y_{j+1} & y_j+y_{j+1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2(1-y_{j+1}^2) \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \end{bmatrix}. \tag{3.22}$$

Thus (3.19) can be written as

$$\sum_{j=0}^{k} G_{2k+1}(2j+1) = \mathrm{Tr}\left\{\prod_{j=1}^{2k+1} V(j,j+1)\right\} \tag{3.23}$$

with

$$V(j,j+1)=\begin{pmatrix} y_j+y_{j+1} & y_j+y_{j+1} & 0 \\ 0 & 0 & 2(1-y_{j+1}^2) \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}. \tag{3.24}$$

We write (3.23) as

$$\mathrm{Tr}\left\{\prod_{j=1}^{2k+1} V(j,j+1)\right\} = \mathrm{Tr}\{[B(1)\,V(1,2)\,B^{-1}(2)][B(2)\,V(2,3)\,B^{-1}(3)]\cdots \times [B(2k+1)\,V(2k+1,1)\,B^{-1}(1)]\}, \tag{3.25}$$

where we define

$$B(j)=\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -y_j+1 \\ \frac{1}{2} & -\frac{1}{2} & -y_j \\ \frac{1}{2} & \frac{1}{2} & -y_j-1 \end{bmatrix}. \tag{3.26}$$

From (3.24) and (3.26) it follows that

$$B(j)\,V(j,j+1)\,B^{-1}(j+1) = \begin{bmatrix} y_{j+1}+1 & 0 & 0 \\ y_{j+1}^2+\frac{1}{2}y_{j+1}-\frac{1}{2} & 0 & -y_{j+1}^2+\frac{1}{2}y_{j+1}+\frac{1}{2} \\ 0 & 0 & y_{j+1}-1 \end{bmatrix}. \tag{3.27}$$

Since the second column of the matrix in (3.27) consists of all zeros, the matrix elements $(y_{j+1}^2+\frac{1}{2}y_{j+1}-\frac{1}{2})$ and $(-y_{j+1}^2+\frac{1}{2}y_{j+1}+\frac{1}{2})$ do not affect the value of (3.25). Hence these terms can be set equal to zero when evaluating the trace in (3.25). Doing this we see that (3.27), (3.25), and (3.23) imply that (3.15) is true.

C. Proof of (3.8)

If we use Lemma 3.1 and let $n \to k-m-1$ in (3.9) we have

$$\psi_{2(k-m)-1}(t;\nu) = \frac{2}{2k-2m-1}\int_1^\infty dy_1 \cdots \int_1^\infty dy_{2k-2m-1}\left[\prod_{j=1}^{2k-2m-1}\frac{\exp(-ty_j)}{(y_j-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^{\nu}(y_j+y_{j+1})^{-1}\right]\left[\sum_{j=0}^{k} G_{2k+1}(2j+1)\right]. \tag{3.28}$$

Each term in $G_{2k+1}(2j+1)$ contains at least one loop. Since the first term in square brackets in (3.28) is invariant under cyclic permutations of the integration variable labels, each term in $G_{2k+1}(2j+1)$ may be cyclically permuted (by cyclically permuting the labels on the graph) so that one of the loops occurring in $G_{2k+1}(2j+1)$ connects the points "1" and "$2k-2m-1$." We denote this permuted version of (3.28) by placing a prime on $G_{2k+1}(2j+1)$.

Now consider the right-hand side of (3.8). If we use the definition of the functions g_{2l+1} and $g_{2(m-l)+1}$ (see

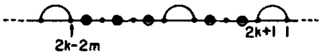

FIG. 31. Typical graph occurring in argument following (3.32).

Theorem 1), the permuted version of (3.28) just described, and the identity

$$\sum_{m=0}^{k-1}\sum_{j=0}^{k-1-m}=\sum_{j=0}^{k-1}\sum_{m=0}^{k-1-j}, \tag{3.29}$$

we have upon letting $j\to j+1$ in the first sum of the right-hand side of (3.29)

$$\begin{aligned}\sum_{m=0}^{k-1}\sum_{l=0}^{m}&\frac{2(k-m)-1}{2}\psi_{2(k-m)-1}g_{2l+1}g_{2(m-l)+1}\\ &=\int_1^\infty dy_1\cdots\int_1^\infty dy_{2k+1}\left[\prod_{j=1}^{2k+1}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^\nu\right.\\ &\quad\left.\times(y_j+y_{j+1})^{-1}\right]\sum_{j=1}^{k}\left\{\sum_{m=0}^{k-j}(-1)^m\sum_{l=0}^{m}\frac{G'_{2k-2m-1}(2j-1)}{y_{2k-2m-1}+y_1}\right.\\ &\quad\times(y_{2k+1}+y_1)(y_{2k-2m-1}+y_{2k-2m})(y_{2k-2m+2}+y_{2k-2m+2l+1})\\ &\quad\left.\times\prod_{j=k-m}^{k-m+l-1}(y_{2j+1}^2-1)\prod_{j=k-m+l+1}^{k}(y_{2j}^2-1)\right\}.\end{aligned} \tag{3.30}$$

The primed graphs in $G'_{2k-2m-1}(2j-1)$ all contain a factor $(y_{2k-2m-1}+y_1)$ which is canceled in (3.30) by the same term that appears in the denominator. Hence to give a graphical representation of the term [valid when used in (3.30)]

$$\frac{G'_{2k-2m-1}(2j-1)}{y_{2k-2m-1}+y_1} \tag{3.31}$$

we imagine starting with the sum of graphs of $2k-2m-1$ points and $2j-1$ loops. Each graph's labels are cyclically permuted so that a term $(y_1+y_{2k-2m-1})$ appears (that is, a loop from "1" to "$2k-2m-1$"). When this loop is removed from each graph the result is (3.31).

We now claim that

$$\begin{aligned}&G_{2k+1}(2j+1)\\ &=\sum_{m=0}^{k-j}(-1)^m\sum_{l=0}^{m}\left[\frac{G'_{2k-2m-1}(2j-1)}{y_{2k-2m-1}+y_1}(y_{2k+1}+y_1)\right.\\ &\quad\times(y_{2k-2m-1}+y_{2k-2m})(y_{2k-2m+2l}+y_{2k-2m+2l+1})\\ &\quad\left.\times\prod_{j=k-m}^{k-m+l-1}(y_{2j+1}^2-1)\prod_{j=k-m+l+1}^{k}(y_{2j}^2-1)\right],\end{aligned} \tag{3.32}$$

where equality in (3.32) is used in the sense that the left-hand side and the right-hand side lead to identical results when used in (3.30).

Consider the set of graphs $G_{2k+1}(2j+1)$. We cyclically permute the labels of the graphs such that a loop connects the points "1" and "$2k+1$." Imagine proceeding from $2k+1$ until two loops are encountered. This second loop must start at an even label which we denote by $2k-2m$ (see Fig. 31). Clearly the smallest m can be is zero which corresponds to the three loops together (see Fig. 32). The largest m can be is $k-j$ since the graph must contain $2j+1$ loops in all. Thus the set of permuted graphs with the points $2k-2m+1$ to $2k$ omitted with just one loop between these points and $2j+1$ loops in all is

$$\frac{G'_{2k-2m-1}(2j-1)}{y_{2k-2m-1}+y_1}(y_{2k+1}+y_1)(y_{2k-2m-1}+y_{2k-2m}). \tag{3.33}$$

The remaining terms in (3.32), i.e.,

$$\sum_{l=0}^{m}\prod_{j=k-m}^{k-m+l-1}(y_{2j+1}^2-1)(y_{2k-2m+2l}+y_{2k-2m+2l+1})\prod_{j=k-m+l+1}^{k}(y_{2j}^2-1)$$

are just all ways of putting in the final loop. The factor $(-1)^m$ gives the correct sign for the m inserted circles between $2k-2m$ and $2k+1$. Summing this from $m=0$ to $k-j$ gives all possible (permuted) graphs in $G_{2k+1}(2j+1)$. Hence (3.32) is true.

Using (3.32) in (3.30) we have that (3.8) can be written as

$$\begin{aligned}(2k+1)g_{2k+1}&+\sum_{m=0}^{k-1}\frac{2(k-m)-1}{2}\psi_{2(k-m)-1}\sum_{l=0}^{m}g_{2l+1}g_{2(m-l)+1}\\ &=\int_1^\infty dy_1\cdots\int_1^\infty dy_{2k+1}\left[\prod_{j=1}^{2k+1}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\right.\\ &\quad\left.\times\left(\frac{y_j-1}{y_j+1}\right)^\nu(y_j+y_{j+1})^{-1}\right]\sum_{j=0}^{k}G_{2k+1}(2j+1),\end{aligned} \tag{3.34}$$

where we identified $(2k+1)g_{2k+1}$ as $G_{2k+1}(1)$. Using Lemma 3.1 we see that (3.34) is just $\frac{1}{2}(2k+1)\psi_{2k+1}$. Thus we have proved (3.8) and hence Theorem 2.

From Theorem 2 we can prove that underlying the nonlinear differential equation (3.2) and hence the Painlevé equation (1.1) with the restriction (1.2) there is an associated *linear* integral equation.

Consider the integral operator $\mathcal{K}$ defined on $\mathcal{L}^2(1,\infty,d\sigma_\pm)$ by

$$(\mathcal{K}f)(x)=\int_1^\infty d\sigma_\pm(y)\exp[-\theta(x+y)](x+y)^{-1}f(y), \tag{3.35a}$$

where the measure $d\sigma_\pm$ is

$$d\sigma_\pm=d\sigma_\pm(y)=\left(\frac{y-1}{y+1}\right)^{\nu\pm1/2}dy. \tag{3.35b}$$

The scalar product is

$$(g,f)_\pm=\int_1^\infty d\sigma_\pm(y)\overline{g(y)}f(y). \tag{3.36}$$

The operator $\mathcal{K}$ is Hilbert–Schmidt for all real $\theta>0$. As $\theta\to0$ the Hilbert–Schmidt norm of $\mathcal{K}$ approaches infinity (the approach is $\sim\ln\theta^{-1}$). We denote by $\lambda_j^\pm(\theta,\nu)$

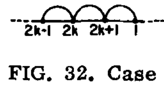

FIG. 32. Case $m=0$ in argument following (3.32).

the eigenvalues and by $\phi_j^{\pm}(x;\theta,\nu)$ the orthonormal eigenfunctions of K. For brevity we sometimes write $\lambda_j^{\pm}$ for $\lambda_j^{\pm}(\theta,\nu)$. Thus we have

$$(K\phi_j^{\pm})(x)=\lambda_j^{\pm}(\theta,\nu)\,\phi_j^{\pm}(x;\theta,\nu), \tag{3.37}$$

where + (−) refers to the measure $d\sigma_+$ ($d\sigma_-$).

We now prove

Corollary: The Painlevé transcendents $\eta(\theta;\nu,\lambda)$ of Theorem 1 possess the representation

$$\eta(\theta;\nu,\lambda)=\prod_{j=1}^{\infty}\left(\frac{1-\lambda_j^+\lambda}{1+\lambda_j^+\lambda}\right)^{a_j^+}\prod_{j=1}^{\infty}\left(\frac{1-\lambda_j^-\lambda}{1+\lambda_j^-\lambda}\right)^{a_j^-}, \tag{3.38}$$

where

$$a_j^{\pm}=a_j^{\pm}(\theta,\nu)=(\lambda_j^{\pm})^{-1}\int_\theta^\infty d\zeta\left|\int_1^\infty d\sigma_{\pm}(y)\right.$$
$$\left.\times\exp(-\zeta y)\,\phi_j^{\pm}(y;\theta,\nu)\right|^2. \tag{3.39}$$

Proof: Using

$$(y_{2n+1}+y_1)^{-1}=\int_0^\infty d\zeta\exp[-\zeta(y_1+y_{2n+1})]$$

in the representation (1.9b) of the functions $\psi_{2n+1}(t;\nu)$ we see we can write $\psi_{2n+1}(t;\nu)$ as

$$\psi_{2n+1}(t;\nu)=\frac{2}{2n+1}\int_0^\infty d\zeta\,[(e,K^{2n}e)_+ +(e,K^{2n}e)_-], \tag{3.40}$$

where

$$e(y)=\exp[-(\zeta+\theta)y]. \tag{3.41}$$

Using Mercer's theorem we can write (3.40) as

$$\psi_{2n+1}(t;\nu)=\frac{2}{2n+1}\int_0^\infty d\zeta\left\{\sum_{j=1}^{\infty}(\lambda_j^+)^{2n}|(e,\phi_j^+)_+|^2\right.$$
$$\left.+\sum_{j=1}^{\infty}(\lambda_j^-)^{2n}|(e,\phi_j^-)_-|^2\right\}. \tag{3.42}$$

Recalling the elementary relation (valid for $|x|<1$)

$$\sum_{n=1}^{\infty}\frac{2}{2n+1}x^{2n+1}=\ln\left(\frac{1+x}{1-x}\right),$$

we can conclude from (3.42) and (1.8) that for $|\lambda| < \min[(\lambda_1^+)^{-1},(\lambda_1^-)^{-1}]$ (where $\lambda_1^+\geq\lambda_2^+\geq\cdots$ and $\lambda_1^-\geq\lambda_2^-\geq\cdots$) $\psi(t;\nu,\lambda)$ has the representation

$$\psi(t;\nu,\lambda)=\int_0^\infty d\zeta\left\{\sum_{j=1}^{\infty}(\lambda_j^+)^{-1}|(e,\phi_j^+)_+|^2\ln\left(\frac{1+\lambda_j^+\lambda}{1-\lambda_j^+\lambda}\right)\right.$$
$$\left.+\sum_{j=1}^{\infty}(\lambda_j^-)^{-1}|(e,\phi_j^-)_-|^2\ln\left(\frac{1+\lambda_j^-\lambda}{1-\lambda_j^-\lambda}\right)\right\}. \tag{3.43}$$

Defining $a_j^{\pm}$ by (3.39) and recalling (1.7) we conclude that the Painlevé transcendent $\eta(\theta;\nu,\lambda)$ is given by (3.38).

From (3.38) we see that the closest singularity in the complex λ plane occurs at $\min[(\lambda_1^+)^{-1},(\lambda_1^-)^{-1}]$. This gives the radius of convergence of (1.7) in the complex λ plane. The restriction $|\lambda|<\min[(\lambda_1^+)^{-1},(\lambda_1^-)^{-1}]$ can be lifted in (3.38). From the theory of analytic continuation we know that, for fixed θ and ν, $\eta(\theta;\nu,\lambda)$ is given by the right-hand side of (3.38) whenever the infinite products converge. A necessary condition that (3.38) converge in the complex λ plane is $\lambda\neq\pm(\lambda_j^+)^{-1}$ and $\lambda\neq\pm(\lambda_j^-)^{-1}$ for all j. We conjecture this is also sufficient.

It is an open problem to compute the quantities $\lambda_j^{\pm}$ and $a_j^{\pm}$ appearing in (3.38).

IV. THEOREM 3 AND COROLLARIES

A. Formal small-t expansion

A formal small-t expansion of the differential equation (3.2) is

$$\psi(t)\sim-\sigma\ln t-\ln B+\sum_{j=1}^{\infty}\sum_{k=1}^{j+1}a_{j,k}t^{j-\sigma(j+2-2k)}. \tag{4.1}$$

The coefficients $a_{j,k}$ are determined from (3.2) by equating like powers of t and are unique functions of σ and B (and ν). The requirement that (4.1) be asymptotic as $t\to0$ requires that

$$-1<\mathrm{Re}\,\sigma<1, \tag{4.2}$$

but otherwise the coefficients σ and B are arbitrary. If we define

$$w(t)=\exp[-\psi(t)], \tag{4.3}$$

then

$$w(t)\sim Bt^{\sigma}\left\{1+\sum_{j=1}^{\infty}\sum_{k=1}^{j+1}b_{j,k}t^{j-\sigma(j+2-2k)}\right\} \tag{4.4}$$

is a formal small-t expansion of (1.3) where we again assume (4.2). The coefficients $b_{j,k}$ can be determined from either (4.1) and (4.3) (assuming $a_{j,k}$ are known) or directly from the differential equation (1.3). The first few coefficients are

$$\begin{aligned}
b_{1,1}&=-\nu B^{-1}(1-\sigma)^{-2},\\
b_{1,2}&=B\nu(1+\sigma)^{-2},\\
b_{2,1}&=\tfrac14\nu^2B^{-2}(1-\sigma)^{-4}-\tfrac1{16}B^{-2}(1-\sigma)^{-2},\\
b_{2,2}&=-\nu^2(1+\sigma)^{-1}(1-\sigma)^{-2},\\
b_{2,3}&=\tfrac1{16}B^2(1+\sigma)^{-2}+\tfrac34\nu^2B^2(1+\sigma)^{-4},
\end{aligned} \tag{4.5}$$

etc.

Computation of the coefficients of the terms $t^{3-3\sigma}$ and $t^{4-4\sigma}$ ($b_{3,1}$ and $b_{4,1}$, respectively) in the expansion (4.4) shows that these terms are zero. This is a general result, i.e.,

$$b_{n,1}=0,\quad n=3,4,5,\cdots. \tag{4.6}$$

To prove (4.6) we can proceed by induction. Since the argument is straightforward we omit the proof. Thus for $n\geq3$ there are no terms of the form $t^{n-n\sigma}$ in (4.4).

When $\sigma=0$ (4.4) becomes a formal power series expansion in the variable t about the point $t=0$. This formal power series can be shown to converge. The result that there exists a one-parameter family of solutions to (1.3) such that the point $t=0$ is an analytic point is known.[1,2,5] Furthermore when $t=0$ ($\theta=0$) is an analytic point, the solution to (1.3) is known to be a meromorphic function.[1,2,5]

B. $\psi_{2n+1}(t;\nu)$ and $\psi(t;\nu,\lambda)$ as $t\to 0$

We define for $n\geq 2$

$$\psi_n(t;\nu)=\frac{2}{n}\int_1^\infty dy_1\cdots\int_1^\infty dy_n\prod_{j=1}^n\frac{\exp(-ty_j)}{y_j+y_{j+1}}\times\left[\prod_{j=1}^n\left(\frac{y_j-1}{y_j+1}\right)^{\nu-1/2}+\prod_{j=1}^n\left(\frac{y_j-1}{y_j+1}\right)^{\nu+1/2}\right] \quad (4.7)$$

with $y_{n+1}\equiv y_1$ (this merely defines ψ_n for even integers and coincides with Theorem 2 for odd integers).

Lemma 4.1: As $t\to 0$ along the positive real axis

$$\psi_n(t;\nu)=\sigma_n\ln\left(\frac{1}{t}\right)+B_n+o(1) \quad (4.8)$$

where

$$\sigma_n=\frac{4}{n}\int_0^1 dx_1\cdots\int_0^1 dx_n\prod_{j=1}^n(x_j+x_{j+1})^{-1}\delta\left(1-\sum_{j=1}^n x_j\right), \quad (4.9)$$

$x_{n+1}\equiv x_1$, and

$$B_n=B_n^{(1)\prime}+B_n^{(1)\prime\prime}, \quad (4.10)$$

with

$$B_n^{(1)\prime}=\frac{4}{n}\int_0^1 dx_1\cdots\int_0^1 dx_n\prod_{j=1}^n(x_j+x_{j+1})^{-1}\times\ln x_1\,\delta\left(1-\sum_{j=1}^n x_j\right) \quad (4.11)$$

and

$$B_n^{(1)\prime\prime}=\frac{2}{n}\lim_{t\to 0}\left\{\int_1^\infty dy_1\cdots\int_1^\infty dy_n\prod_{j=1}^n\exp(-ty_j)(y_j+y_{j+1})^{-1}\times\left[\prod_{j=1}^n\left(\frac{y_j-1}{y_j+1}\right)^{\nu+1/2}+\prod_{j=1}^n\left(\frac{y_j-1}{y_j+1}\right)^{\nu-1/2}\right]-2\int_0^\infty dy_1\cdots\int_0^\infty dy_n\prod_{n=1}^n\exp(-ty_j)(y_j+y_{j+1})^{-1}\times[1-\exp(-y_1)]\right\}, \quad (4.12)$$

with $\delta(x)$ denoting the Dirac delta function.

Proof: Let $F(y)$ be such that $F(y)/y$ is bounded for all $y>0$ and $F(y)\to\infty$. Define

$$\Theta=\Theta(y_1,y_2,\ldots,y_n)=\prod_{j=1}^n\theta(y_j-1), \quad (4.13)$$

where

$$\theta(x)=\begin{cases}1 & \text{if } x>0,\\ 0 & \text{if } x\leq 0.\end{cases} \quad (4.14)$$

We write

$$\psi_n=\frac{2}{n}\int_0^\infty dy_1\cdots\int_0^\infty dy_n\prod_{j=1}^n\exp(-ty_j)(y_j+y_{j+1})^{-1}\times\left[\Theta(y_1,\ldots,y_n)\prod_{j=1}^n\left(\frac{y_j-1}{y_j+1}\right)^{\nu+1/2}+\Theta(y_1,\ldots,y_n)\prod_{j=1}^n\left(\frac{y_j-1}{y_j+1}\right)^{\nu-1/2}-2F(y_1)\right]+I_n, \quad (4.15)$$

where

$$I_n=\frac{4}{n}\int_0^\infty dy_1\cdots\int_0^\infty dy_nF(y_1)\prod_{j=1}^n\exp(-ty_j)(y_j+y_{j+1})^{-1}. \quad (4.16)$$

The limit $t\to 0$ exists for the first term in (4.15) (that is for the quantity ψ_n-I_n).

We choose

$$F(y)=1-\exp(-y) \quad (4.17)$$

which clearly satisfies the above two requirements of $F(y)$. Thus $B_n^{(1)\prime\prime}$ is just the $t\to 0$ limit of ψ_n-I_n with the choice (4.17) for $F(y)$.

We make the change of variables

$$\rho=\sum_{j=1}^n x_j,\quad x_j=\rho^{-1}y_j,\quad j=1,2,\ldots,n, \quad (4.18)$$

in (4.16) with the choice (4.17).

Then

$$I_n=\int_0^\infty\frac{d\rho}{\rho}\exp(-t\rho)\int_0^1 dx_1\cdots\int_0^1 dx_n\,\delta\left(1-\sum_{j=1}^n x_j\right)\times[1-\exp(-\rho x_1)]\prod_{j=1}^n(x_j+x_{j+1})^{-1}. \quad (4.19)$$

If we make use of the identity

$$\ln\left(\frac{x}{y}\right)=\int_0^\infty\frac{d\xi}{\xi}[\exp(-\xi y)-\exp(-\xi x)], \quad (4.20)$$

then I_n becomes

$$\begin{aligned}I_n&=\int_0^1 dx_1\cdots\int_0^1 dx_n\delta\left(1-\sum_{j=1}^n x_j\right)\times\prod_{j=1}^n(x_j+x_{j+1})^{-1}\ln\left(\frac{t+x_1}{t}\right)\\&=\ln\left(\frac{1}{t}\right)\int_0^1 dx_1\cdots\int_0^1 dx_n\delta\left(1-\sum_{j=1}^n x_j\right)\prod_{j=1}^n(x_j+x_{j+1})^{-1}\\&\quad+\int_0^1 dx_1\cdots\int_0^1 dx_n\delta\left(1-\sum_{j=1}^n x_j\right)\prod_{j=1}^n(x_j+x_{j+1})^{-1}\ln x_1\\&\quad+o(1)\quad(\text{as } t\to 0).\end{aligned} \quad (4.21)$$

This proves the Lemma.

From (4.8) and the fact that

$$\psi(t;\nu,\lambda)=\sum_{n=0}^\infty\lambda^{2n+1}\psi_{2n+1}(t;\nu), \quad (4.22)$$

we conclude for $|\lambda|<1/\pi$

$$\psi(t;\nu,\lambda)=+\sigma\ln t^{-1}-\ln B+o(1) \quad (4.23)$$

as $t\to 0^+$ where

$$\sigma=\sum_{n=0}^\infty\lambda^{2n+1}\sigma_{2n+1} \quad (4.24)$$

and

$$-\ln B=\sum_{n=0}^\infty\lambda^{2n+1}B_{2n+1}, \quad (4.25)$$

where σ_{2n+1} and B_{2n+1} are given by Lemma 4.1. For the steps (4.23)–(4.25) to be completely rigorous we must ensure that the error estimate in (4.8) remains $o(1)$

when summed over n in (4.22). We do not present a rigorous proof of this point. Heuristically, if one sums the *leading* term in the $o(1)$ term in (4.8), then the result is still $o(1)$. Also the function $B(\sigma, \nu=0)$ was computed numerically by a procedure independent of the steps (4.23)–(4.25) and to within numerical accuracy (five to six decimal places) the result agrees with that given by (1.12).

C. Computation of σ_n

Lemma 4.2: If we denote by σ_n the quantity defined in (4.9) then

$$\sigma_n = (2/n)\, \pi^{n-2} B(\tfrac{1}{2}, n/2), \qquad (4.26)$$

where $B(x,y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$ is the beta function.

Proof: Consider the integral

$$J_n = \int_0^\infty dx_1 \cdots \int_0^\infty dx_n \prod_{j=1}^{n-1} (x_j + x_{j+1})^{-1} \exp(-x_1 - x_n). \qquad (4.27)$$

Let

$$\lambda = \sum_{j=1}^{n} x_j, \quad \alpha_j = \frac{x_j}{\lambda}, \quad j = 1, 2, \ldots, n-1, \qquad (4.28)$$

then the Jacobian is

$$\frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(\alpha_1, \alpha_2, \ldots, \alpha_{n-1}, \lambda)} = \lambda^{n-1}. \qquad (4.29)$$

Since there are $(n-1)$ factors in the denominator of J_n, we get

$$\begin{aligned} J_n &= \int_0^1 d\alpha_1 \int_0^{1-\alpha_1} d\alpha_2 \cdots \int_0^{1-\alpha_1 - \cdots - \alpha_{n-2}} d\alpha_{n-1} \int_0^\infty d\lambda \\ &\quad \times \exp\left[-\lambda\alpha_1 - \lambda\left(1 - \sum_{j=1}^{n-1} \alpha_j\right)\right] \\ &\quad \times \left\{\prod_{j=1}^{n-2} (\alpha_j + \alpha_{j+1}) \left[\alpha_{n-1} + \left(1 - \sum_{j=1}^{n-1} \alpha_j\right)\right]\right\}^{-1} \\ &= \int_0^1 d\alpha_1 \int_0^1 d\alpha_2 \cdots \int_0^1 d\alpha_n \delta\left(1 - \sum_{j=1}^{n} \alpha_j\right) \prod_{j=1}^{n} (\alpha_j + \alpha_{j+1})^{-1}, \end{aligned} \qquad (4.30)$$

where $\alpha_{n+1} \equiv \alpha_1$. Hence in view of (4.9) we have shown that

$$\sigma_n = (4/n) J_n. \qquad (4.31)$$

To evaluate (4.27) we use the method of Mellin transforms. If we define

$$F(\xi) = \int_0^\infty x^{-\xi} f(x)\, dx, \qquad (4.32)$$

then for f and $g \in L^2$ we have the Mellin convolution formula

$$\int_0^\infty f(x) g(x)\, dx = \frac{1}{2\pi i} \int_{1/2 - i\infty}^{1/2 + i\infty} F(\xi) G(1-\xi)\, d\xi. \qquad (4.33)$$

Now the Mellin transform of $(x+y)^{-1}$ is

$$\int_0^\infty x^{-\xi} \frac{1}{x+y} dx = \frac{\pi}{\sin \pi\xi} y^{-\xi}, \qquad (4.34)$$

where $0 < \mathrm{Re}\,\xi < 1$ while that of $\exp(-x)$ is of course

$$\int_0^\infty x^{-\xi} \exp(-x)\, dx = \Gamma(1-\xi). \qquad (4.35)$$

Therefore,

$$\begin{aligned} J_n &= \frac{1}{2\pi i} \int_{1/2 - i\infty}^{1/2 + i\infty} d\xi \left(\frac{\pi}{\sin \pi\xi}\right)^{n-1} \Gamma(1-\xi)\Gamma(\xi) \\ &= \frac{1}{2\pi i} \int_{1/2 - i\infty}^{1/2 + i\infty} d\xi \left(\frac{\pi}{\sin \pi\xi}\right)^{n} \\ &= \tfrac{1}{2}\pi^{n-1} \int_{-\infty}^{\infty} d\xi \frac{1}{(\cosh \pi\xi)^n} = \tfrac{1}{2}\pi^{n-2} B(\tfrac{1}{2}, n/2), \end{aligned} \qquad (4.36)$$

Therefore, (4.31) and (4.36) prove the Lemma.

The result (1.11) of Theorem 3 now follows from Lemma 4.2. We note that for n odd, (4.26) can be written as

$$\sigma_{2n+1} = \frac{1}{2n+1} \frac{\pi^{2n}}{2^{n-1}} \frac{(2n-1)!!}{n!} \qquad (4.37)$$

D. Relating B_n to integral equations

From (4.11) it follows that

$$\begin{aligned} B_n^{(1)\prime} &= \frac{4}{n} \int_0^\infty d\lambda \int_0^1 dx_1 \cdots \int_0^1 dx_n\, \delta\left(1 - \sum_{j=1}^{n} x_j\right) \\ &\quad \times \ln x_1 \prod_{j=1}^{n-1} (x_j + x_{j+1})^{-1} \exp[-\lambda(x_1 + x_n)]. \end{aligned} \qquad (4.38)$$

Reversing the steps that went from (4.27) to (4.30) we see that (4.38) can be written as

$$\begin{aligned} B_n^{(1)\prime} &= \frac{4}{n} \int_0^\infty d\lambda \int_0^\infty dy_1 \cdots \int_0^\infty dy_n\, \delta(\lambda - \Sigma y_j) \\ &\quad \times \ln(y_1 \lambda^{-1}) \prod_{j=1}^{n-1} (y_j + y_{j+1})^{-1} \exp(-y_1 - y_n) \\ &= \frac{4}{n} \int_0^\infty dy_1 \cdots \int_0^\infty dy_n \ln\left(\frac{y_1}{y_1 + \cdots + y_n}\right) \\ &\quad \times \prod_{j=1}^{n-1} (y_j + y_{j+1})^{-1} \exp(-y_1 - y_n). \end{aligned} \qquad (4.39)$$

Using identity (4.20) for the logarithm term in (4.39) we conclude

$$\begin{aligned} B_n^{(1)\prime} &= \frac{4}{n} \int_0^\infty dy_1 \cdots \int_0^\infty dy_n \int_0^\infty \frac{d\xi}{\xi} \\ &\quad \times \{\exp[-\xi(y_1 + \cdots + y_n)] - \exp(-\xi y_1)\} \\ &\quad \times \prod_{j=1}^{n-1} (y_j + y_{j+1})^{-1} \exp(-y_1 - y_n). \end{aligned} \qquad (4.40)$$

We now wish to split the above integral into two parts. However as it stands, the integrals taken separately are divergent. Thus we write

$$B_n^{(1)\prime} = \lim_{\epsilon \to 0} [B_n^{(1)}(\epsilon) + B_n^{(2)}(\epsilon)], \qquad (4.41)$$

where

$$\begin{aligned} B_n^{(1)}(\epsilon) &= -\frac{4}{n} \int_\epsilon^{1/\epsilon} \frac{d\xi}{\xi} \int_0^\infty dy_1 \cdots \int_0^\infty dy_n \\ &\quad \times \exp[-(1+\xi) y_1] \prod_{j=1}^{n-1} (y_j + y_{j+1})^{-1} \exp(-y_n) \end{aligned} \qquad (4.42)$$

and

$$B_n^{(2)}(\epsilon)=\frac{4}{n}\int_\epsilon^\infty \frac{d\xi}{\xi}\int_0^\infty dy_1\cdots\int_0^\infty dy_n$$
$$\times\exp(-y_1)\prod_{j=1}^{n-1}\exp(-\xi y_j)(y_j+y_{j+1})^{-1}$$
$$\times\exp[-y_n(1+\xi)]. \tag{4.43}$$

We let $x_j=\xi y_j$ in this last expression (and then followed by $\zeta=1/\xi$) to obtain

$$B_n^{(2)}(\epsilon)=\frac{4}{n}\int_0^{1/\epsilon} d\zeta\int_0^\infty dx_1\cdots\int_0^\infty dx_n$$
$$\times\exp(-\zeta x_1)\prod_{j=1}^{n-1}\frac{\exp(-x_j)}{x_j+x_{j+1}}\exp[-(1+\zeta)x_n]. \tag{4.44}$$

Since we are interested only in the $t\to 0$ limit of the integrals occuring in (4.12), we may write $B_n^{(1)\prime\prime}$ as

$$B_n^{(1)\prime\prime}=\frac{2}{n}\lim_{t\to 0}\Bigg\{\int_1^\infty dy_1\cdots\int_1^\infty dy_n$$
$$\times\exp[-t(y_1+y_n)]\prod_{j=1}^{n}(y_j+y_{j+1})^{-1}\left[\prod_{j=1}^{n}\left(\frac{y_j-1}{y_j+1}\right)^{\nu+1/2}\right.$$
$$\left.+\prod_{j=1}^{n}\left(\frac{y_j-1}{y_j+1}\right)^{\nu-1/2}\right]-2\int_0^\infty dy_1\cdots\int_0^\infty dy_n$$
$$\times\exp[-t(y_1+y_n)]\prod_{j=1}^{n}(y_j+y_{j+1})^{-1}[1-\exp(-y_1)]\Bigg\}. \tag{4.45}$$

That is, we do not change the value of (4.12) if we set $t=0$ in $\exp(-ty_2),\ldots,\exp(-ty_{n-1})$ and leave only the factor $\exp[-t(y_1+y_n)]$. As we did for $B_n^{(1)\prime}$, we break $B_n^{(1)\prime\prime}$ into a sum of terms. As (4.45) stands, the individual integrals are divergent. First we use

$$\frac{\exp[-t(y_1+y_n)]}{y_1+y_n}=\int_t^\infty d\xi\exp[-\xi(y_1+y_n)] \tag{4.46}$$

in (4.45) and then write (also let $t\to\epsilon$)

$$B_n^{(1)\prime\prime}=\lim_{\epsilon\to 0}[B_n^{(3)}+B_n^{(4)}(\epsilon)+B_n^{(5)}(\epsilon)], \tag{4.47}$$

where

$$B_n^{(3)}(\epsilon)=-\frac{4}{n}\int_\epsilon^{\epsilon^{-1}} d\xi\int_0^\infty dy_1\cdots\int_0^\infty dy_n[1-\exp(-y_1)]$$
$$\times\exp[-\xi(y_1+y_n)]\prod_{j=1}^{n-1}(y_j+y_{j+1})^{-1}$$
$$=-\frac{4}{n}\int_\epsilon^{\epsilon^{-1}}\frac{d\xi}{\xi}\int_0^\infty dx_1\cdots\int_0^\infty dx_n$$
$$\times\{\exp(-x_1)-\exp[-(1+\zeta)x_1]\}$$
$$\times\prod_{j=1}^{n-1}(x_j+x_{j+1})^{-1}\exp(-x_n), \tag{4.48}$$

$$B_n^{(4)}(\epsilon)=\frac{2}{n}\int_\epsilon^\infty d\xi\int_1^\infty dy_1\cdots\int_1^\infty dy_n$$
$$\times\exp[-\xi(y_1+y_n)]\prod_{j=1}^{n-1}(y_j+y_{j+1})^{-1}\prod_{j=1}^{n}\left(\frac{y_j-1}{y_j+1}\right)^{\nu-1/2}, \tag{4.49}$$

and $B_n^{(5)}(\epsilon)$ is just $B_n^{(4)}(\epsilon)$ with $\nu-\frac12$ replaced by $\nu+\frac12$ in (4.49).

Comparing (4.42) and (4.48) we see that

$$B_n^{(1)}(\epsilon)+B_n^{(3)}(\epsilon)=-\frac{4}{n}\int_\epsilon^{\epsilon^{-1}}\frac{d\xi}{\xi}\int_0^\infty dx_1\cdots\int_0^\infty dx_n$$
$$\times\exp(-x_1)\prod_{j=1}^{n-1}(x_j+x_{j+1})^{-1}\exp(-x_n). \tag{4.50}$$

The ξ integration is now decoupled from the x_j variables so that (4.50) is simply

$$B_n^{(1)}(\epsilon)+B_n^{(3)}(\epsilon)=-2\sigma_n\ln(1/\epsilon), \tag{4.51}$$

where we used the results of Lemmas 4.1 and 4.2.

The quantities $B_n^{(2)}(\epsilon)$, $B_n^{(4)}(\epsilon)$, and $B_n^{(5)}(\epsilon)$ remain to be computed. Equations (4.44) and (4.49) are in the form of an iterated kernel. Therefore, we now examine the integral equations associated with these kernels.

E. Integral equations

Lemma 4.3:

$$\int_1^\infty dy\left(\frac{y-1}{y+1}\right)^\nu\frac{\phi_{p,\nu}(y)}{x+y}=\lambda_p\phi_{p,\nu}(x) \tag{4.52}$$

with

$$\lambda_p=\pi\operatorname{sech}\pi p,\quad 0\le p<\infty, \tag{4.53}$$
$$\phi_{p,\nu}(x)=C_{p,\nu}F(\tfrac12+ip,\tfrac12-ip;1+\nu;\tfrac12-\tfrac12x), \tag{4.54}$$

where $F(a,b;c;x)$ is the hypergeometric function and

$$C_{p,\nu}=[\pi^{-1}\Gamma^{-2}(\nu+1)p\sinh\pi p\,\Gamma(\tfrac12+\nu+ip)\Gamma(\tfrac12+\nu-ip)]^{1/2}. \tag{4.55}$$

Furthermore the $\phi_{p,\nu}(x)$ are orthogonal, i.e.,

$$\int_1^\infty dy\left(\frac{y-1}{y+1}\right)^\nu\phi_{p,\nu}(y)\phi_{p',\nu}(y)=\delta(p-p'), \tag{4.56}$$

where $\delta(x)$ is the Dirac delta function.

The functions $\phi_{p,\nu}(x)$ can alternatively be expressed in terms of Legendre functions[6]

$$\phi_{p,\nu}(x)=C_{p,\nu}\Gamma(1+\nu)\left(\frac{x+1}{x-1}\right)^{\nu/2}P^{-\nu}_{-1/2+ip}(x). \tag{4.57}$$

Lemma 4.3 is a special case of the inversion formulas for the generalized Mehler–Fock transform.[7,8]

To compute $B_n^{(2)}(\epsilon)$ we need

Lemma 4.4:

$$\int_0^\infty\frac{\exp(-y)}{x+y}\chi_p(y)\,dy=\lambda_p\chi_p(x), \tag{4.58}$$

where

$$\chi_p(x)=(2\lambda_p)^{-1/2}\int_1^\infty d\xi\exp[-(\xi-1)x/2]\phi_{p,0}(\xi) \tag{4.59}$$

and $\phi_{p,0}(\xi)$ is the $\nu=0$ case of (4.54) and λ_p is given by (4.53).

Proof: Consider the integral equation

$$\int_0^\infty \frac{\exp[-(u+v)]}{u+v} g(v)\,dv = \lambda g(u). \tag{4.60}$$

We can write this as

$$\lambda g(u) = \int_0^\infty dv\, g(v) \int_1^\infty d\xi' \exp[-\xi'(u+v)]. \tag{4.61}$$

If we multiply both sides of this equation by $\exp(-\xi u)$, and if we integrate the result over u from zero to infinity, then (4.61) becomes

$$\lambda G(\xi) = \int_1^\infty d\xi' \frac{1}{\xi+\xi'} G(\xi'), \tag{4.62}$$

where $G(\xi)$ is the Laplace transform of $g(u)$, i.e.,

$$G(\xi) = \int_0^\infty \exp(-\xi u)\, g(u)\, du. \tag{4.63}$$

From Lemma 4.3,

$$G(\xi) = \phi_{p,0}(\xi)$$

and

$$\lambda = \lambda_p = \pi \operatorname{sech} \pi p. \tag{4.64}$$

From (4.61) and (4.63),

$$\begin{aligned}\lambda g(u) &= \int_1^\infty d\xi' \exp(-\xi' u) G(\xi')\\ &= \int_1^\infty d\xi' \exp(-\xi' u)\phi_{p,0}(\xi'). \end{aligned}\tag{4.65}$$

Letting $f(x) = \exp(x/2) g(u)$, $x = 2u$ we see that $f(x)$ satisfies (4.58). The overall constant in (4.59) has been chosen so that

$$\int_0^\infty \exp(-x)\chi_p(x)\chi_{p'}(x)\,dx = \delta(p-p'). \tag{4.66}$$

From Lemma 4.4 and (4.4) it follows that

$$B_n^{(2)}(\epsilon) = \frac{4}{n}\int_0^{\epsilon^{-1}} d\zeta \int_0^\infty dp\, \lambda_p^{n-1} \left|(\exp(-\zeta x), \chi_p(x))\right|^2, \tag{4.67}$$

where

$$\begin{aligned}(\exp(-\zeta x), \chi_p(x)) &= \int_0^\infty \exp[-(\zeta+1)x]\,\chi_p(x)\,dx\\ &= (2\lambda_p)^{-1/2}\int_0^\infty d\xi\, \phi_{p,0}(\xi)\\ &\quad\times \int_0^\infty dx \exp[-(\zeta+\xi/2+\tfrac12)x]\\ &= (2\lambda_p)^{-1/2}\int_1^\infty d\xi \frac{\phi_{p,0}(\xi)}{\zeta+\xi/2+\frac12}\\ &= (2\lambda_p)^{1/2}\phi_{p,0}(2\zeta+1). \end{aligned}\tag{4.68}$$

Thus

$$B_n^{(2)}(\epsilon) = \frac{8}{n}\int_0^{\epsilon^{-1}} d\zeta \int_0^\infty dp \left(\frac{\pi}{\cosh \pi p}\right)^n \left|\phi_{p,0}(2\zeta+1)\right|^2. \tag{4.69}$$

Using (4.57) for $\nu=0$ we have

$$B_n^{(2)}(\epsilon) = \frac{8}{n}\int_0^{\epsilon^{-1}} d\zeta \int_0^\infty dp \left(\frac{\pi}{\cosh \pi p}\right)^n \times p \tanh p\pi [P_{-1/2+ip}(1+2\zeta)]^2. \tag{4.70}$$

From Lemma 4.3 and (4.49) it follows that

$$B_n^{(4)}(\epsilon) = \frac{2}{n}\int_\epsilon^\infty d\xi \int_0^\infty dp \left(\frac{\pi}{\cosh \pi p}\right)^{n-1} \times \left|(\exp(-\xi x), \phi_{p,\nu-1/2}(x))\right|^2 \tag{4.71}$$

and

$$B_n^{(5)}(\epsilon) = \frac{2}{n}\int_\epsilon^\infty d\xi \int_0^\infty dp \left(\frac{\pi}{\cosh \pi p}\right)^{n-1} \times \left|(\exp(-\xi x), \phi_{p,\nu+1/2}(x))\right|^2, \tag{4.72}$$

where the scalar product in (4.71) and (4.72) is

$$(\exp(-\xi x), \phi_{p,\mu}(x)) = \int_1^\infty dx \left(\frac{x-1}{x+1}\right)^\mu \exp(-\xi x)\phi_{p,\mu}(x). \tag{4.73}$$

Thus to prove Theorem 3 we need to compute the integrals (4.70)–(4.72).

F. $B_n^{(2)}(\epsilon)$

In (4.70) we do the ζ integration first. Now

$$\int_0^{\epsilon^{-1}} [P_{-1/2+ip}(1+2\zeta)]^2 d\zeta = \tfrac12 \int_1^\Lambda [P_{-1/2+ip}(z)]^2 dz, \tag{4.74}$$

where $z = 1+2\zeta$ and $\Lambda = 1+2\epsilon^{-1}$. We are interested in computing (4.74) in the limit $\Lambda\to\infty$.

For any two Legendre functions w_ν and w_σ on the cut, we have[9]

$$\int_a^b w_\nu(Z) w_\sigma(Z)\,dZ = [(\nu-\sigma)(\nu+\sigma+1)]^{-1} \times \left[(1-Z^2)\left(w_\nu \frac{d}{dZ} w_\sigma - w_\sigma \frac{d}{dZ} w_\nu\right)\right]_a^b. \tag{4.75}$$

Letting $\nu = -\frac12 + ip$, $\sigma = -\frac12 + ip'$, then we have in particular

$$\begin{aligned}&\int_1^\Lambda dz\, P_{-1/2+ip}(z) P_{-1/2+ip'}(z)\\ &= (p^2 - p'^2)^{-1}(\Lambda^2-1)[P_{-1/2+ip}(\Lambda)P'_{-1/2+ip'}(\Lambda)\\ &\quad - P_{-1/2+ip'}(\Lambda)P'_{-1/2+ip}(\Lambda)]. \end{aligned}\tag{4.76}$$

Writing

$$P_{-1/2+ip'}(z)(p-p')^{-1} = [P_{-1/2+ip'}(z) - P_{-1/2+ip}(z)](p-p')^{-1} + P_{-1/2+ip}(z)(p-p')^{-1}$$

in (4.76) we obtain

$$\int_1^\Lambda dz\, P_{-1/2+ip}(z) P_{-1/2+ip'}(z) = \frac{\Lambda^2-1}{p+p'}\left\{P_{-1/2+ip}(\Lambda)\frac{\partial}{\partial\Lambda}\frac{P_{-1/2+ip'}(\Lambda) - P_{-1/2+ip}(\Lambda)}{p-p'}\right.$$

$$-\frac{\partial P_{-1/2+ip}(\Lambda)}{\partial\Lambda}\left[\frac{P_{-1/2+ip'}(\Lambda)-P_{-1/2+ip}(\Lambda)}{p-p'}\right]\Bigg\}. \tag{4.77}$$

We now let $p'\to p$ in (4.77),

$$\int_1^\Lambda dz\,[P_{-1/2+ip}(z)]^2=\frac{\Lambda^2-1}{2p}\left(\frac{\partial P_{-1/2+ip}(\Lambda)}{\partial\Lambda}\frac{\partial P_{-1/2+ip}(\Lambda)}{\partial p}-P_{-1/2+ip}(\Lambda)\frac{\partial^2}{\partial p\partial\Lambda}P_{-1/2+ip}(\Lambda)\right). \tag{4.78}$$

We need to compute the right-hand side of (4.78) to order $o(1)$ as $\Lambda\to\infty$. Using Eq. (3.2.9) of Ref. 6, one can show for $\Lambda\to\infty$ (along the positive real axis)

$$P_{-1/2+ip}(\Lambda)=(2/\pi)^{1/2}\,\mathrm{Re}\left(\frac{2^{ip}\Gamma(ip)}{\Gamma(\frac12+ip)}\Lambda^{-1/2+ip}\right)+O\left(\frac{1}{\Lambda}\right). \tag{4.79}$$

We have also

$$\frac{\partial P_{-1/2+ip}(\Lambda)}{\partial\Lambda}=\frac{2^{-1/2+ip}\Gamma(ip)}{\Gamma(\frac12+ip)}\pi^{-1/2}(-\tfrac12+ip)\Lambda^{-3/2+ip}+\text{complex conj.}+O(\Lambda^{-2}), \tag{4.80}$$

$$\frac{\partial P_{-1/2+ip}(\Lambda)}{\partial p}=[i\ln2+i\psi(ip)-i\psi(\tfrac12+ip)+i\ln\Lambda]\times\frac{2^{-1/2+ip}\Gamma(ip)}{\pi^{1/2}\Gamma(\frac12+ip)}\Lambda^{-1/2+ip}+\text{complex conj.}+O(\Lambda^{-1}), \tag{4.81}$$

and

$$\frac{\partial^2 P_{-1/2+ip}(\Lambda)}{\partial p\partial\Lambda}=i\frac{2^{-1/2+ip}\Gamma(ip)}{\pi^{1/2}\Gamma(\frac12+ip)}[(\ln2+\psi(ip)-\psi(\tfrac12+ip)+\ln\Lambda)(-\tfrac12+ip)+1]\Lambda^{-3/2+ip}+\text{complex conj.}+O(\Lambda^{-2}), \tag{4.82}$$

where $\psi(x)=(d/dx)\ln\Gamma(x)$ is the psi function.

Substituting (4.79)–(4.82) into (4.78) and using the relations $\Gamma(ip)\Gamma(-ip)=\pi p^{-1}\sinh^{-1}(\pi p)$ and $\Gamma(\frac12+ip)\times\Gamma(\frac12-ip)=\pi\,\mathrm{sech}(\pi p)$ the result

$$\int_1^\Lambda dz\,[P_{-1/2+ip}(z)]^2=(\pi p\tanh\pi p)^{-1}[\ln2+\ln\Lambda+\mathrm{Re}\psi(ip)-\mathrm{Re}\psi(\tfrac12+ip)]+\frac{1}{2\pi p}\mathrm{Im}\left(\frac{\Gamma^2(ip)}{\Gamma^2(\frac12+ip)}(2\Lambda)^{2ip}\right)+o(1) \tag{4.83}$$

follows.

Using (4.83) and (4.74) in (4.70) we obtain

$$B_n^{(2)}(\epsilon)=\sigma_n\ln\left(\frac1\epsilon\right)+\frac{4}{\pi n}\int_0^\infty dp\left(\frac{\pi}{\cosh\pi p}\right)^n\times[\ln4+\mathrm{Re}\psi(ip)-\mathrm{Re}\psi(\tfrac12+ip)]-\frac1n\pi^n, \tag{4.84}$$

where we used the result

$$\lim_{\Lambda\to\infty}\int_0^\infty dp\left(\frac{\pi}{\cosh\pi p}\right)^n\tanh\pi p\,\mathrm{Im}\left(\frac{\Gamma^2(ip)}{\Gamma^2(\frac12+ip)}(2\Lambda)^{2ip}\right)=-\tfrac12\pi^{n+1}. \tag{4.85}$$

G. $B_n^{(4)}(\epsilon)$, $B_n^{(5)}(\epsilon)$, and B_n

The matrix element needed in (4.71) and (4.72) is [recall (4.57) and (4.73)]

$$(\exp(-\xi x),\phi_{p,\nu}(x))=C_{p,\nu}\Gamma(1+\nu)\int_1^\infty dx\left(\frac{x-1}{x+1}\right)^{\nu/2}\exp(-\xi x)P^{-\nu}_{-1/2+ip}(x). \tag{4.86}$$

The integral (4.86) is known[7,10] and the result is

$$(\exp(-\xi x),\phi_{p,\nu}(x))=C_{p,\nu}\Gamma(1+\nu)\xi^{-1}W_{-\nu,ip}(2\xi), \tag{4.87}$$

where $W_{k,\mu}(x)$ is a Whittaker function.[11] Using (4.87) in (4.71) and (4.72) we have

$$B_n^{(4)}(\epsilon)+B_n^{(5)}(\epsilon)=\frac2n\int_\epsilon^\infty d\xi\int_0^\infty dp\left(\frac{\pi}{\cosh\pi p}\right)^{n-1}\frac{p\sinh\pi p}{\pi}\times\xi^{-2}[\Gamma(\nu+1+ip)\Gamma(\nu+1-ip)(W_{-1/2-\nu,ip}(2\xi))^2+\Gamma(\nu+ip)\Gamma(\nu-ip)(W_{1/2-\nu,ip}(2\xi))^2]. \tag{4.88}$$

We first examine the ξ integration. Define

$$F_1(\epsilon)=\int_\epsilon^\infty d\xi\,\xi^{-2}[W_{1/2-\nu,ip}(2\xi)]^2 \tag{4.89}$$

and

$$F_2(\epsilon)=\int_\epsilon^\infty d\xi\,\xi^{-2}[W_{-1/2-\nu,ip}(2\xi)]^2. \tag{4.90}$$

Let $\hat F_{1,2}(Z)$ be the respective Mellin transforms, i.e.,

$$\hat F_{1,2}(Z)=\int_0^\infty\epsilon^{Z-1}F_{1,2}(\epsilon)\,d\epsilon,\quad \mathrm{Re}Z>0. \tag{4.91}$$

Using (4.89) and (4.90) in (4.91) and interchanging the orders of integration so that the ϵ integration can be trivially performed, we obtain

$$\hat F_1(Z)=Z^{-1}\int_0^\infty d\xi\,\xi^{Z-2}[W_{1/2-\nu,ip}(2\xi)]^2=Z^{-1}2^{-Z+1}\int_0^\infty d\xi\,\xi^{Z-2}[W_{1/2-\nu,ip}(\xi)]^2 \tag{4.92}$$

and

$$\hat F_2(Z)=Z^{-1}2^{-Z+1}\int_0^\infty d\xi\,\xi^{Z-2}[W_{-1/2-\nu,ip}(\xi)]^2. \tag{4.93}$$

The integrals appearing in (4.92) and (4.93) are known[12] and we have for $\mathrm{Re}Z>0$

$$\int_0^\infty\xi^{Z-2}[W_{1/2-\nu,ip}(\xi)]^2d\xi=\frac{\Gamma(Z+2ip)\Gamma(Z)\Gamma(-2ip)}{\Gamma(\nu-ip)\Gamma(\nu+ip+Z)}\,{}_3F_2(2ip+Z,Z,\nu+ip;1+2ip,\nu+ip+Z;1)+\frac{\Gamma(Z-2ip)\Gamma(Z)\Gamma(2ip)}{\Gamma(\nu+ip)\Gamma(\nu-ip+Z)}\times{}_3F_2(Z,Z-2ip,\nu-ip;1-2ip,\nu-ip+Z;1), \tag{4.94}$$

where ${}_3F_2(a_1,a_2,a_3;b_1,b_2;Z)$ is a generalized hypergeo-

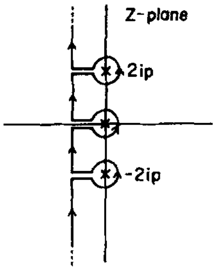

FIG. 33. Contour C used in (4.95b).

metric function.[13] From (4.92)–(4.94) we see that $\hat{F}_{1,2}(Z)$ has poles on the line $\mathrm{Re}Z=0$ at $Z=\pm 2ip$ and $Z=0$. To compute the small-ϵ behavior of $F_{1,2}(\epsilon)$ to order $o(1)$ it is sufficient to study the behavior of $\hat{F}_{1,2}(Z)$ on the line $\mathrm{Re}Z=0$ since

$$F_{1,2}(\epsilon)=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\epsilon^{-Z}\hat{F}_{1,2}(Z)\,dZ \tag{4.95a}$$

$$=\frac{1}{2\pi i}\int_C \epsilon^{-Z}\hat{F}_{1,2}(Z)\,dZ, \tag{4.95b}$$

where (4.95a) is the Mellin inversion formula and the countour C in (4.95b) is shown in Fig. 33. The integral along the straight line lying in the $\mathrm{Re}Z<0$ plane is $o(1)$ as $\epsilon\to 0$.

We now examine $\hat{F}_{1,2}(Z)$ at $Z=\pm 2ip$ and $Z=0$. We first expand (4.94) about $Z=0$. For $p>0$

$$\frac{\Gamma(Z+2ip)\Gamma(Z)\Gamma(-2ip)}{\Gamma(\nu-ip)\Gamma(\nu+ip+Z)}=Z^{-1}\frac{\Gamma(2ip)\Gamma(-2ip)}{\Gamma(\nu+ip)\Gamma(\nu-ip)}\{1+Z[\psi(2ip)-\gamma-\psi(\nu+ip)]+O(Z^2)\}, \tag{4.96}$$

where $\gamma=0.5772\cdots$ in Euler's constant. By definition

$$ {}_3F_2(2ip+Z,Z,\nu+ip;1+2ip,\nu+ip+Z;1)=\sum_{n=0}^{\infty}\frac{(2ip+Z)_n(Z)_n(\nu+ip)_n}{(1+2ip)_n(\nu+ip+Z)_n n!}, \tag{4.97}$$

where $(a)_n\equiv\Gamma(a+n)/\Gamma(a)$. Expanding (4.97) about $Z=0$ we have for $p>0$

$$ {}_3F_2(2ip+Z,Z,\nu+ip;1+2ip,\nu+ip+Z;1)=1+Z\sum_{n=1}^{\infty}\frac{2ip}{n(2ip+n)}+O(Z^2)$$

$$=1+Z[\psi(1+2ip)+\gamma]+O(Z^2). \tag{4.98}$$

Thus we have shown that for $Z\to 0$

$$\int_0^\infty \xi^{Z-2}[W_{1/2-\nu,ip}(\xi)]^2\,d\xi=2Z^{-1}\frac{\Gamma(-2ip)\Gamma(2ip)}{\Gamma(\nu-ip)\Gamma(\nu+ip)}\{1+Z[\mathrm{Re}\psi(2ip)-\mathrm{Re}\psi(\nu+ip)+\mathrm{Re}\psi(1+2ip)]+O(Z^2)\}, \tag{4.99}$$

and hence for $Z\to 0$

$$\hat{F}_1(Z)=f_{-2}Z^{-2}+f_{-1}Z^{-1}+O(1), \tag{4.100a}$$

where

$$f_{-2}=\pi[\Gamma(\nu+ip)\Gamma(\nu-ip)p\cosh(\pi p)\sinh(\pi p)]^{-1} \tag{4.100b}$$

and

$$f_{-1}=f_{-2}[2\,\mathrm{Re}\psi(2ip)-\ln 2-\mathrm{Re}\psi(\nu+ip)]. \tag{4.100c}$$

To obtain the Laurent expansion of $\hat{F}_2(Z)$ about $Z=0$, one replaces ν by $\nu+1$ in (4.100).

From (4.92), (4.94), and (4.97) we have for $Z\to\pm 2ip$

$$\hat{F}_1(Z)=\pm\frac{1}{2ip}2^{\mp 2ip+1}\left[\frac{\Gamma(\pm 2ip)}{\Gamma(\nu\pm ip)}\right]^2\frac{1}{Z\mp 2ip}+O(1). \tag{4.101}$$

Using (4.100) and (4.101) in (4.95b) and recalling (4.88)–(4.90) we have for $\epsilon\to 0^+$

$$\begin{aligned}B_n^{(4)}(\epsilon)+B_n^{(5)}(\epsilon)&=\sigma_n\ln\left(\frac{1}{\epsilon}\right)-\sigma_n\ln 2+\frac{4}{n\pi}\int_0^\infty dp\left(\frac{\pi}{\cosh\pi p}\right)^n\\&\quad\times[2\,\mathrm{Re}\psi(2ip)-\tfrac12\mathrm{Re}\psi(\nu+ip)-\tfrac12\mathrm{Re}\psi(\nu+1+ip)]\\&\quad+\lim_{\epsilon\to 0}\frac{4}{n\pi}\int_0^\infty dp\left(\frac{\pi}{\cosh\pi p}\right)^{n-1}\sinh\pi p\left\{\mathrm{Im}\left[2^{-2ip}\left(\frac{\Gamma(2ip)}{\Gamma(\nu+ip)}\right)^2\exp(-2ip\ln\epsilon)\right]\Gamma(\nu+ip)\Gamma(\nu-ip)\right.\\&\quad\left.+\mathrm{Im}\left[2^{-2ip}\left(\frac{\Gamma(2ip)}{\Gamma(\nu+1+ip)}\right)^2\exp(-2ip\ln\epsilon)\right]\Gamma(\nu+1+ip)\Gamma(\nu+1-ip)\right\}.\end{aligned} \tag{4.102}$$

In deriving (4.102) we made the identification [see (4.31) and (4.36)]

$$\sigma_n=\frac{4}{\pi n}\int_0^\infty dp\left(\frac{\pi}{\cosh\pi p}\right)^n. \tag{4.103}$$

The only nonzero contribution in the limit $\epsilon\to 0$ to the last integral in (4.102) is in the region $p\sim 0$. A computation shows this integral is $-(1/n)\pi^n$.

We now use (4.10), (4.41), (4.51), (4.84), and (4.102) to obtain [note that the $\ln(1/\epsilon)$ terms cancel]

$$B_n=3\ln 2\sigma_n-\frac{2}{n}\pi^n+\frac{4}{n\pi}\int_0^\infty dp\left(\frac{\pi}{\cosh\pi p}\right)^n\times[2\,\mathrm{Re}\psi(ip)-\tfrac12\mathrm{Re}\psi(\nu+ip)-\tfrac12\mathrm{Re}\psi(\nu+1+ip)], \tag{4.104}$$

where we used the functional equation[14]

$$\psi(2\psi)=\tfrac12\psi(x)+\tfrac12\psi(x+\tfrac12)+\ln 2.$$

H. $B(\sigma,\nu)$

To complete the proof of Theorem 3 we must compute the sum (4.25) where we have shown that the coefficients are given by (4.104). Rather than regard B as a function of λ and ν, it will prove more natural to think of B as a function of σ and ν where $\sigma=\sigma(\lambda)=2\pi^{-1}\arcsin(\pi\lambda)$.

Then for $\sigma<1$ ($\lambda<1/\pi$) it follows from (4.104) that

$$-\ln B=\sum_{n=0}^{\infty}B_{2n+1}\lambda^{2n+1}$$

$$=3\sigma\ln 2-\ln\left(\frac{1+\sin\pi\sigma/2}{1-\sin\pi\sigma/2}\right)+I_1+I_2+I_3, \qquad (4.105\text{a})$$

where

$$I_1=\frac{4}{\pi}\int_0^\infty dp\ \ln\left(\frac{\cosh\pi p+\sin\pi\sigma/2}{\cosh\pi p-\sin\pi\sigma/2}\right)\operatorname{Re}\psi(ip), \qquad (4.105\text{b})$$

$$I_2=-\frac{2}{\pi}\int_0^\infty dp\ \ln\left(\frac{\cosh\pi p+\sin\pi\sigma/2}{\cosh\pi p-\sin\pi\sigma/2}\right)\operatorname{Re}\psi(\nu+ip), \qquad (4.105\text{c})$$

and

$$I_3=-\frac{\nu}{\pi}\int_0^\infty dp\ \ln\left(\frac{\cosh\pi p+\sin\pi\sigma/2}{\cosh\pi p-\sin\pi\sigma/2}\right)(\nu^2+p^2)^{-1}. \qquad (4.105\text{d})$$

From (4.105b)

$$\frac{\partial I_1}{\partial\sigma}=4\cos\pi\sigma/2\int_0^\infty dp\frac{\cosh\pi p}{\cosh^2\pi p-\sin^2\pi\sigma/2}\operatorname{Re}\psi(ip)$$

$$=2\cos\pi\sigma/2\int_{-\infty}^\infty dp\frac{\cosh\pi p}{\cosh^2\pi p-\sin^2\pi\sigma/2}\operatorname{Re}\psi(-ip), \qquad (4.106)$$

where the second equality follows from the fact that $\operatorname{Re}\psi(ip)$ is an even function of p with no singularities on the real p axis. $\operatorname{Im}\psi(-ip)$ is an odd function of p with a pole with residue -1 at $p=0$. Hence

$$2\cos\pi\sigma/2\int_\Omega dp\ \frac{\cosh\pi p}{\cosh^2\pi p-\sin^2\pi\sigma/2}\operatorname{Im}\psi(-ip)$$

$$=2\pi i\cos\pi\sigma/2(1-\sin^2\pi\sigma/2)^{-1}, \qquad (4.107)$$

where the contour of the integration Ω is the real p axis from $-\infty$ to $-\epsilon$, a semicircle lying in the upper half-plane centered at the origin with radius ϵ, and the real axis from $+\epsilon$ to $+\infty$. The limit $\epsilon\to 0^+$ is then understood. Multiplying (4.107) by $+i$ and adding the result to (4.106) we have

$$2\cos\pi\sigma/2\int_\Omega dp\ \psi(-ip)\ \frac{\cosh\pi p}{\cosh^2\pi p-\sin^2\pi\sigma/2}$$

$$=\frac{\partial I_1}{\partial\sigma}-2\frac{d}{d\sigma}\ln\left(\frac{1+\sin\pi\sigma/2}{1-\sin\pi\sigma/2}\right). \qquad (4.108)$$

The integral

$$J_1\equiv 2\cos\pi\sigma/2\int_\Omega dp\ \psi(-ip)\ \frac{\cosh\pi p}{\cosh^2\pi p-\sin^2\pi\sigma/2}, \qquad (4.109)$$

can be evaluated by applying Cauchy's theorem to

$$2\cos\pi\sigma/2\int_{C_R}dz\ \psi(-iz)\ \frac{\cosh\pi z}{\cosh^2\pi z-\sin^2\pi\sigma/2}, \qquad (4.110)$$

where the contour C_R is shown in Fig. 34. Letting $R\to\infty$ in (4.110) results in

$$J_1=\psi\left(\frac{1+\sigma}{2}\right)+\psi\left(\frac{1-\sigma}{2}\right)-\pi\cos\pi\sigma/2(1-\sin^2\pi\sigma/2). \qquad (4.111)$$

Hence from (4.108), (4.109), and (4.111)

$$\frac{\partial I_1}{\partial\sigma}=\psi\left(\frac{1+\sigma}{2}\right)+\psi\left(\frac{1-\sigma}{2}\right)+\frac{d}{d\sigma}\ln\left(\frac{1+\sin^2\pi\sigma/2}{1-\sin^2\pi\sigma/2}\right). \qquad (4.112)$$

Since I_1 $(\sigma=0)=0$ it follows from (4.112) that

$$I_1=2\ln\Gamma\left(\frac{1+\sigma}{2}\right)-2\ln\Gamma\left(\frac{1-\sigma}{2}\right)+\ln\left(\frac{1+\sin\pi\sigma/2}{1-\sin\pi\sigma/2}\right). \qquad (4.113)$$

The evaluation of I_2 is similar. We have for $\nu>0$

$$\frac{\partial I_2}{\partial\sigma}=-\cos\pi\sigma/2\int_{-\infty}^\infty dp\ \psi(\nu-ip)\ \frac{\cosh\pi p}{\cosh^2\pi p-\sin^2\pi\sigma/2}, \qquad (4.114)$$

since $\operatorname{Im}\psi(\nu-ip)$ is an odd function of p with no singularities on the real axis. We again use the contour of Fig. 34 (the semicircles are no longer necessary) with the result

$$\frac{\partial I_2}{\partial\sigma}=-\tfrac{1}{2}\psi\left(\frac{1+\sigma}{2}+\nu\right)-\tfrac{1}{2}\psi\left(\nu+\frac{1-\sigma}{2}\right)+\tfrac{1}{2}\cos\pi\sigma/2$$

$$\times\int_{-\infty}^\infty\frac{dp}{\nu+ip}\ \frac{\cosh\pi p}{\cosh^2\pi p-\sin^2\pi\sigma/2}. \qquad (4.115)$$

The integral appearing in (4.115) is unchanged if we replace $(\nu+ip)^{-1}$ by $(\nu-ip)^{-1}$. Hence

$$\frac{\partial I_2}{\partial\sigma}=-\tfrac{1}{2}\psi\left(\frac{1+\sigma}{2}+\nu\right)-\tfrac{1}{2}\psi\left(\nu+\frac{1-\sigma}{2}\right)+\frac{\nu}{2}\cos\frac{\pi}{2}\sigma$$

$$\times\int_{-\infty}^\infty dp\ (\nu^2+p^2)^{-1}\frac{\cosh\pi p}{\cosh^2\pi p-\sin^2\pi\sigma/2}. \qquad (4.116)$$

Integrating (4.116) $[I_2(\sigma=0)=0]$ we have

$$I_2=-\ln\left(\frac{1+\sigma}{2}+\nu\right)+\ln\Gamma\left(\frac{1-\sigma}{2}+\nu\right)-I_3. \qquad (4.117)$$

It follows from (4.105a), (4.113), and (4.117) that $B(\sigma,\nu)$ is given by (1.12). The small-t behavior (1.17) of the functions $g_{2n+1}(t;\nu)$ now follows from Theorem 3 and Eq. (3.5).

I. Small-t behavior of $\eta(t/2;\nu,\lambda)$ for $\lambda\geqslant\pi^{-1}$

As $\sigma\to 1$ ($\lambda\to\pi^{-1}$) we have from (1.12)

$$B(\sigma,\nu)=b_{-2}(1-\sigma)^{-2}+b_{-1}(1-\sigma)^{-1}+b_0+O(1-\sigma) \qquad (4.118\text{a})$$

with

$$b_{-2}=\tfrac{1}{2}\nu, \qquad (4.118\text{b})$$

$$b_{-1}=\tfrac{3}{2}\nu\ln 2-\tfrac{1}{2}\nu\psi(\nu+1)-\gamma\nu+\tfrac{1}{4}, \qquad (4.118\text{c})$$

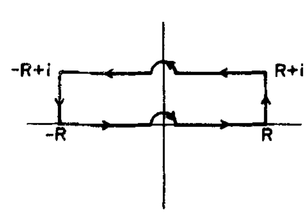

FIG. 34. Contour C_R used in (4.110) and (4.114).

and

$$b_0=\tfrac{9}{4}(\ln 2)^2\nu-\tfrac{3}{4}\nu\ln 2[\psi(\nu)+\psi(\nu+1)+4\gamma] + \tfrac{1}{16}\nu\{[\psi(\nu)+\psi(\nu+1)]^2+\psi'(\nu+1)-\psi'(\nu)\} - \tfrac{1}{2}\gamma+\gamma\nu\psi(\nu+1)+\gamma^2\nu. \quad (4.118d)$$

We now use (4.118) in (4.4) [also use (4.5)] and recall that there are no terms of the form $t^{n-n\sigma}$ for $n\geq 3$. The limit $\sigma\to 1$ exists with the result that for $t\to 0$ (along positive real axis)

$$\eta(t/2;\nu,\pi^{-1})\sim\tfrac{1}{2}t\{\nu\ln^2 t-C(\nu)\ln t+1/(4\nu)[C^2(\nu)-1]\}, \quad (4.119)$$

where

$$C(\nu)=1+2\nu[3\ln 2-2\gamma-\psi(\nu+1)]. \quad (4.120)$$

We note that $\lim_{\nu\to 0}(4\nu)^{-1}[C^2(\nu)-1]=3\ln 2-\gamma$.

The correction terms to (4.119) are most easily determined by using the differential equation (1.3). For example, for the special case $\nu=0$ we find[3,4]

$$\eta(\theta;0,\pi^{-1})=-\theta\Omega-\frac{\theta^5}{128}(8\Omega^3-8\Omega^2+4\Omega-1)+O(\theta^9\Omega^5), \quad (4.121)$$

where $\Omega\equiv\ln(\theta/4)+\gamma$.

The case $\lambda>1/\pi$ can be similarly examined. We write for real positive μ

$$\lambda=(1/\pi)\cosh(\pi\mu), \quad (4.122)$$

so that [see (1.10)]

$$\sigma=1+2i\mu. \quad (4.123)$$

We examine here the case $\nu=0$. Then using (4.123) in (1.11) for $\nu=0$ we see that (4.4) becomes for $\mu>0$, $t\to 0^+$

$$\eta(t/2;0,\lambda)\sim\frac{1}{4\pi}t\sinh(\pi\mu)\,\mathrm{Im}\{\Gamma^2(-i\mu)\exp[2i\mu\ln(t/8)]\}. \quad (4.124)$$

If we write

$$\Gamma(iy)=|\Gamma(iy)|\exp[i\phi(y)] =\left[\frac{\pi}{y\sinh\pi y}\right]^{1/2}\exp[i\phi(y)], \quad (4.125)$$

then (4.124) becomes ($t\to 0$, $\mu>0$)

$$\eta(t/2;0,\lambda)\sim-\frac{1}{4\mu}t\sin[2\mu\ln(t/8)+2\phi(\mu)]. \quad (4.126)$$

Thus for $\lambda>\pi^{-1}$ there are an infinite number of zeros of the function $\eta(t/2;0,\lambda)$ lying on the positive t axis with $t=0$ being a limit point of these zeros. The asymptotic spacing of these zeros follows from (4.126). The correction terms to (4.124) [or (4.126)] can be found from the differential equation.

The case $\lambda<0$ can also be studied. From (1.4) and (1.5) it follows that

$$\eta(t/2;\nu,-\lambda)=\frac{1}{\eta(t/2;\nu,\lambda)}. \quad (4.127)$$

Hence we see that for $\lambda<-\pi^{-1}$, $\eta(t/2;\nu,\lambda)$ has an infinite number of poles clustering to zero on the positive t axis.

V. THEOREM 4

We commence the proof of Theorem 4 by using (1.4a) to rewrite the left-hand side of (1.14a) in terms of G as

$$[1-G^2(t)]^{-1/2}\exp\int_t^\infty dt'\left\{t'\frac{[G^2(t')-G'^2(t')]}{[1-G^2(t')]^2}+\frac{2\nu}{1-G^2(t')}\right\}, \quad (5.1)$$

where $G'(t)=(d/dt)G(t)$. The first factor may be written in the form

$$[1-G^2(t)]^{-1/2}=\exp\left(-\tfrac{1}{2}\int_t^\infty dt\,\frac{2G(t')G'(t')}{[1-G^2(t')]}\right) \quad (5.2)$$

and therefore Theorem 4 is established if we can demonstrate

$$-\sum_{n=1}^\infty\lambda^{2n}f_{2n}(t;\nu)=\int_t^\infty dt'\left\{\frac{t'[G^2(t')-G'^2(t')]}{[1-G^2(t')]^2}-\frac{G(t)G'(t')}{[1-G^2(t')]}+\frac{2\nu}{1-G^2(t')}\right\}. \quad (5.3)$$

Furthermore, because $f_{2n}(t;\nu)$ and $G^2(t)$ vanish exponentially rapidly as $t\to\infty$ (5.3) will be demonstrated if we can show

$$\sum_{n=1}^\infty\lambda^{2n}f'_{2n}(t;\nu)=\frac{t[G^2(t)-G'^2(t)]}{[1-G^2(t)]^2}-\frac{G(t)G'(t)}{[1-G^2(t)]}+\frac{2\nu}{1-G^2(t)} \quad (5.4)$$

or, using the differential equation for $G(t)$ (2.7)

$$(1-G^2)^2\sum_{n=1}^\infty\lambda^{2n}f''_{2n}(t;\nu)=G^2(1-G'^2)-GG''(1-G^2). \quad (5.5)$$

Here, all factors of t have been removed by use of the differential equation.

To demonstrate (5.5) we first define in analogy to g_{2k+1} (1.6b)

$$h_{2k+1}=(-1)^{k+1}\int_1^\infty dy_1\cdots\int_1^\infty dy_{2k+1}\prod_{j=1}^{2k+1}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}} \times\left(\frac{y_j-1}{y_j+1}\right)^\nu\prod_{j=1}^{2k}\frac{1}{y_j+y_{j+1}}\prod_{j=1}^{k+1}(y_{2j-1}^2-1). \quad (5.6)$$

Then it is seen from the definition (1.13) that

$$f''_{2n}=\sum_{l=1}^n g_{2(n-l)+1}h_{2l-1}. \quad (5.7)$$

Thus, if we define

$$H(t;\nu,\lambda)=\sum_{n=1}^\infty\lambda^{2n+1}h_{2n+1}, \quad (5.8)$$

(5.5) reduces to

$$(1-G^2)^2H=G(1-G'^2)-G''(1-G^2) \quad (5.9)$$

and therefore our theorem will be proven if we can demonstrate that

$$H-(G-G'')=G^2G''-GG'^2+2G^2H-G^4H. \quad (5.10)$$

The coefficient of λ^{2n+1} of the left-hand side of (5.10) is

$$\begin{aligned}
&h_{2n+1}-(g_{2n+1}-g''_{2n+1})\\
&\quad=(-1)^{n+1}\int_1^\infty dy_1\cdots\int_1^\infty dy_{2n+1}\prod_{j=1}^{2n+1}\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\\
&\quad\times\left(\frac{y_j-1}{y_j+1}\right)^{\nu}\prod_{j=1}^{2n}\frac{1}{y_j+y_{j+1}}\left\{\prod_{j=1}^{n+1}(y_{2j-1}^2-1)\right.\\
&\quad\left.+\prod_{j=1}^{n}(y_{2j}^2-1)\left[1-\left(\sum_{j=1}^{2n+1}y_j\right)^2\right]\right\}.
\end{aligned}\tag{5.11}$$

Rewrite the term involving $(\sum y_j)^2$, using

$$\begin{aligned}
&\sum_{j=1}^{n}(y_{2j-1}+y_{2j})\sum_{k=j}^{n}(y_{2k}+y_{2k+1})\\
&\quad=\frac{1}{2}\left\{\left(\sum_{j=1}^{2n+1}y_j\right)^2+\sum_{l=1}^{n}y_{2l}^2-\sum_{l=1}^{n+1}y_{2l-1}^2\right\}
\end{aligned}\tag{5.12}$$

to obtain for the term in brackets

$$\begin{aligned}
&\prod_{j=1}^{n+1}(y_{2j-1}^2-1)+\prod_{j=1}^{n}(y_{2j}^2-1)\left[1-\left(\sum_{j=1}^{2n+1}y_j\right)^2\right]\\
&\quad=\prod_{j=1}^{n+1}(y_{2j-1}^2-1)+\prod_{j=1}^{n}(y_{2j}^2-1)\left[\sum_{l=1}^{n}(y_{2l}^2-1)\right.\\
&\quad\left.-\sum_{l=1}^{n+1}(y_{2l-1}^2-1)-2\sum_{j=1}^{n}(y_{2j-1}+y_{2j})\sum_{k=j}^{n}(y_{2k}+y_{2k+1})\right].
\end{aligned}\tag{5.13}$$

Then use the identity

$$\begin{aligned}
&X_1X_3\cdots X_{2n+1}+X_2X_4\cdots X_{2n}\left[\sum_{l=1}^{n}X_{2l}-\sum_{l=1}^{n+1}X_{2l-1}\right]\\
&\quad=-\sum_{j=1}^{n}(X_{2j-1}-X_{2j})\sum_{k=j}^{n}(X_{2k}-X_{2k+1})\\
&\quad\times\prod_{l_1=1}^{j-1}X_{2l_1}\prod_{l_2=j}^{k-1}X_{2l_2+1}\prod_{l_3=k+1}^{n}X_{2l_3}
\end{aligned}\tag{5.14}$$

with

$$X_l=y_l^2-1\tag{5.15}$$

to obtain

$$\begin{aligned}
&h_{2n+1}-(g_{2n+1}-g''_{2n+1})=(-)^{n+1}\int_1^\infty dy_1\cdots dy_{2n+1}\\
&\quad\times\prod_{j=1}^{2n+1}\left[\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^{\nu}\right]\prod_{j=1}^{2n}\frac{1}{y_j+y_{j+1}}\\
&\quad\times\sum_{j=1}^{n}(y_{2j-1}+y_{2j})\sum_{k=j}^{n}(y_{2k}+y_{2k+1})\left\{-2\prod_{l=1}^{n}(y_{2l}^2-1)\right.\\
&\quad-(y_{2j-1}-y_{2j})(y_{2k}-y_{2k+1})\\
&\quad\left.\times\prod_{l_1=1}^{j-1}(y_{2l_1}^2-1)\prod_{l_2=j}^{k-1}(y_{2l_2+1}^2-1)\prod_{l_3=k+1}^{n}(y_{2l_3}^2-1)\right\}.
\end{aligned}\tag{5.16}$$

The first term in brackets in (5.16) gives the term $2G^2H$ in (5.10) and we note that after we expand the product

$$\begin{aligned}
&(y_{2j-1}-y_{2j})(y_{2k}-y_{2k+1})\\
&\quad=y_{2j-1}(y_{2k}-y_{2k+1})-y_{2j}y_{2k}+y_{2j}y_{2k+1}
\end{aligned}\tag{5.17}$$

that the first term on the right-hand side vanishes when used in the double sum over j and k in (5.16). Moreover we may rewrite $y_{2j}y_{2k}$ using

$$\begin{aligned}
&\left[\sum_{l=2j}^{2k}y_l\right]^2\\
&\quad=y_{2j}y_{2k}+y_{2j}\sum_{l=j}^{k-1}(y_{2l}+y_{2l+1})+y_{2k}\sum_{l=j+1}^{k}(y_{2l-1}+y_{2l})\\
&\quad+\sum_{l_1=j}^{n}(y_{2l_1}+y_{2l_1+1})\sum_{l_2=l_1+1}^{k}(y_{2l_2-1}+y_{2l_2})\\
&\quad+\sum_{l_1=j}^{n}(y_{2l_1}+y_{2l_1+1})\sum_{l_2=j}^{l_1}(y_{2l_2-1}+y_{2l_2})
\end{aligned}\tag{5.18}$$

and obtain

$$\begin{aligned}
&h_{2n+1}-(g_{2n+1}-g''_{2n+1})=(-1)^{n+1}\int_1^\infty dy_1\cdots dy_{2n+1}\\
&\quad\times\prod_{j=1}^{2n+1}\left[\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^{\nu}\right]\prod_{j=1}^{2n}\frac{1}{y_j+y_{j+1}}\\
&\quad\times\sum_{j=1}^{n}(y_{2j-1}+y_{2j})\sum_{k=j}^{n}(y_{2k}+y_{2k+1})\left\{-2\prod_{l=1}^{n}(y_{2l}^2-1)\right.\\
&\quad+\prod_{l_1=1}^{j-1}(y_{2l_1}^2-1)\prod_{l_2=j}^{k-1}(y_{2l_2+1}^2-1)\prod_{l_3=k+1}^{n}(y_{2l_3}^2-1)\\
&\quad\times\left[-y_{2j}y_{2k+1}+\left(\sum_{l=2j}^{2k}y_l\right)^2-y_{2j}\sum_{l=j}^{k-1}(y_{2l}+y_{2l+1})-y_{2k}\right.\\
&\quad\times\sum_{l=j+1}^{k}(y_{2l-1}+y_{2l})-\sum_{l_1=j}^{n}(y_{2l_1}+y_{2l_1+1})\sum_{l_2=l_1+1}^{k}(y_{2l_2-1}+y_{2l_2})\\
&\quad\left.\left.-\sum_{l_1=j}^{n}(y_{2l_1}+y_{2l_1+1})\sum_{l_2=j}^{l_1}(y_{2l_2-1}+y_{2l_2})\right]\right\}.
\end{aligned}\tag{5.19}$$

We note that the terms involving $y_{2j}\sum_{l=j}^{k-1}(y_{2l}+y_{2l+1})$ and $y_{2k}\sum_{j=1}^{k}(y_{2l-1}+y_{2l})$ are equal and we eliminate y_{2j} and y_{2k+1} using

$$y_{2j}=\sum_{l=2j}^{2k}y_l-\sum_{l=j}^{k}(y_{2l+1}+y_{2l}).\tag{5.20a}$$

and

$$y_{2k+1}=\sum_{l=2k+1}^{2n+1}y_l-\sum_{l=k+1}^{n}(y_{2l}+y_{2l+1}).\tag{5.20b}$$

Therefore, upon combining terms we find that

$$\begin{aligned}
&h_{2n+1}-(g_{2n+1}-g''_{2n+1})=\int_1^\infty dy\cdots dy_{2n+1}\\
&\quad\times\prod_{j=1}^{2n+1}\left[\frac{\exp(-ty_j)}{(y_j^2-1)^{1/2}}\left(\frac{y_j-1}{y_j+1}\right)^{\nu}\right]\prod_{j=1}^{2n}\frac{1}{y_j+y_{j+1}}\\
&\quad\times\sum_{j=1}^{n}(-1)^{j-1}(y_{2j-1}+y_{2j})\sum_{k=j}^{n}(-1)^{n-k}(y_{2k}+y_{2k+1})\\
&\quad\times\left\{2(-1)^{k-j+1}\prod_{l=1}^{n}(y_{2l}^2-1)+\prod_{l_1=1}^{j-1}(y_{2l_1}^2-1)\prod_{l_2=j}^{k-1}(y_{2l_2+1}^2-1)\right.\\
&\quad\times\prod_{l_3=k+1}^{n}(y_{2l_3}^2-1)\left[(-1)^{k-j}\left(\sum_{l=2j}^{2k}y_l\right)^2\quad-(-1)^{k-j}\sum_{l_1=2j}^{2k}y_{l_1}\right.\\
&\quad\times\sum_{l_2=2k+1}^{2n}y_{l_2}-\sum_{l_1=j}^{n}(-1)^{l_1-j}(y_{2l_1}+y_{2l_1+1})
\end{aligned}$$

$$\times \sum_{l_2=l_1+1}^{k} (-1)^{l_2-l_1}(y_{2l_2-1}+y_{2l_2})(-1)^{k-l_2} \Bigg]\Bigg\}, \qquad (5.21)$$

from which (5.10) follows. Hence, Theorem 4 is established.

ACKNOWLEDGMENTS

The authors express their gratitude to Professor C.N. Yang, Professor R. Roskies, and Mr. J. Hamm for useful discussions. One of us (C.T.) would like to thank Professor M. Blume for the hospitality extended at Brookhaven National Laboratory where some numerical computations were performed. We wish to particularly thank Professor L. Mittag for showing us the solution of (4.52) in the two special cases $\nu=\pm\frac{1}{2}$ when the eigenfunctions are elementary.

*Work supported in part by the National Science Foundation Grant No. MPS-74-132-08-A01, and by Energy Research and Development Administration Contract No. AT(11-1)-3227.

[1]P. Painlevé, Acta Math. **25**, 1 (1902).
[2]B. Gambier, Acta Math. **33**, 1 (1910).
[3]T.T. Wu, B.M. McCoy, C.A. Tracy, and E. Barouch, Phys. Rev. B **13**, 316 (1976).
[4]J. Myers, Ph.D. thesis Harvard University, 1962 (unpublished).
[5]N.P. Erugin, Differ. Urav. 3, 1821 (1967) [Differ. Eq. **3**, 943 (1967)].
[6]See, for example, A Erdelyi, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. I, p. 122, Eq. (3).
[7]J.S. Lowndes, Proc. Cambridge Philos. Soc. **60**, 57 (1964).
[8]See also I.N. Sneddon, *The Use of Integral Transforms* (McGraw-Hill, New York, 1972), Chapter 7.
[9]See Ref. 6, p. 170.
[10]See, for example, *Tables of Integrals, Series, and Products*, edited by I.S. Gradshteyn and I.M. Ryzhik (Academic, New York, 1965), Eq. (7.142).
[11]See Ref. 6, p. 264.
[12]See for example, A. Erdelyi, *Tables of Integral Transforms* (McGraw-Hill, New York, 1954), Vol. II, Eq. (20.342).
[13]See, for example, Ref. 6, p. 182.
[14]See, for example, Ref. 6, p. 16.

Variation method and nonlinear stability problems

Din-Yu Hsieh

Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912

(Received 4 August 1976)

The direct variational method, developed for studying the asymptotic behavior of a wide class of nonlinear oscillation and wave problems, is extended to the study of the nonlinear stability problems. For systems which are unstable against small disturbances but stable against finite amplitude disturbances, the variational method can yield significant results in regimes even far away from the critical region. The procedure of the variational method is illustrated by applications to various physical problems: the Duffing oscillation, a model wave equation for an unstable mechanical system, the two-stream instability in plasma, and the nonlinear Kelvin–Helmholtz stability problem.

1. INTRODUCTION

An approximate, direct variational method has been developed to deal with a class of nonlinear oscillation and wave problems.[1-3] The method starts with reformulating the problems by equivalent variational problems, then some judiciously chosen asymptotic trial solutions with adjustable parameters are directly substituted in the variational formulation to be varied. The basic idea underlying the method is to make use of as much prior information and expectation as possible and incorporate them into the form of the trial solution. Thus it is expected that the system of equations governing the adjustable unknown parameters would be much simpler than the original problem. In dealing with nonlinear oscillation and wave problems, we have been mainly interested in the asymptotic oscillatory solution. Therefore, approximate solutions for the amplitudes and phases can be obtained by singling out the secular terms. In other perturbation methods which deal with similar problems, the lowest order solutions usually turn out also to be oscillatory or sinusoidal.[4] It is thus not surprising that results obtained from the variational method are essentially in agreement with those obtained from other methods.

Now when the system is stable against small perturbations, small disturbances will usually manifest themselves as oscillations or waves. As the amplitude increases beyond the linear regime, nonlinear interactions will enter to modify the solution. However, if the nonlinearity is relatively weak, it is expected that the basic oscillatory behavior would still persist, while the amplitudes and phases of the basic oscillation may evolve and become related to each other. But the process of this evolution would be slow in comparison with the basic oscillation. Thus the nonlinear solutions are essentially a perturbation of the basic linear solution. The situation is quite different if the system is unstable against small perturbations. In that case, a basic linear solution does not exist to begin with. Therefore, the perturbation methods are usually limited to dealing with the evolutionary process in the neighborhood of the critical region which is the border between linear stability and instability.

In many physical problems, although the system is unstable against some small perturbation, it may be stable against finite perturbations. In other words, when the small amplitude oscillations or waves grow due to instability to certain finite amplitude, the oscillations or waves would settle down to some oscillations and waves with relatively slowly varying finite amplitudes. Thus an asymptotic oscillatory solution does exist even though the linear problem is unstable. It is clear that the variational method we have developed is also capable of dealing with this class of nonlinear stability problems. As in previous studies, we shall again use specific examples to illustrate the procedure of the scheme. It is to be noted that the method is still being developed, while the problem of nonlinear stability is extremely intriguing. Therefore, the results we obtained from the various studies may raise more questions than answers. It is hoped that more and more can be learned from these practical applications to concrete problems.

In the following, we shall first treat a Duffing type oscillation problem, then a model wave equation for an unstable mechanical system to illustrate in detail some basic features of the nonlinear stability problems. Then we shall apply the same method to two-stream stability problems in plasma and the nonlinear Kelvin–Helmholtz stability problem.

2. DUFFING STABILITY

Consider the Duffing equation

$$\frac{d^2u}{dt^2} - au + ru^3 = 0, \tag{2.1}$$

where a and r are real constants. When $r=0$, the solutions of the linear equation are given by $\exp\{\pm\sqrt{a}\,t\}$. Thus the linear system is stable if $a<0$ and it is unstable if $a \geq 0$. However, it may be readily seen that the solution of the nonlinear equation is always bounded for $r>0$.

Equation (2.1) can actually be solved exactly. Multiply (2.1) by du/dt and then integrate; we obtain

$$\left(\frac{du}{dt}\right)^2 = F(u), \tag{2.2}$$

where

$$F(u) = C + au^2 - (r/2)u^4, \tag{2.3}$$

and C is an integration constant which is the value of $(du/dt)^2$ when $u=0$.

Solutions are permitted only for $F(u) \geq 0$. Since $F(u) \to \pm\infty$, as $|u| \to \infty$ for $r \lessgtr 0$, thus the system is stable,

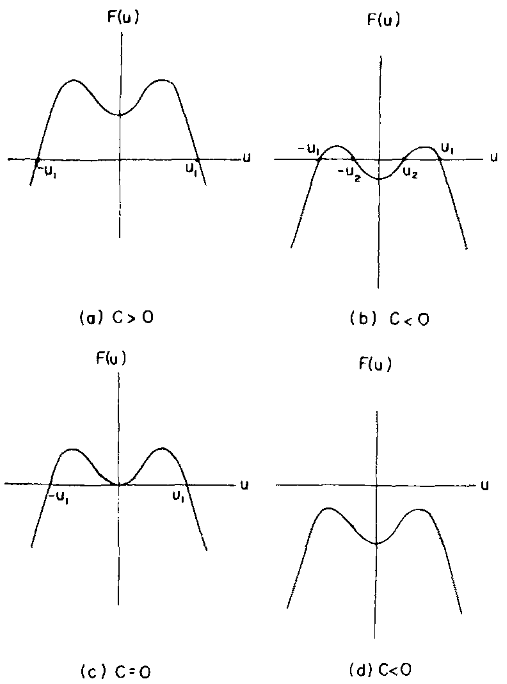

FIG. 1.

i.e., u is bounded if and only if $r > 0$. Therefore, even if $a > 0$, i.e., the linear system is unstable, the nonlinear system is stable so long as $r > 0$.

Take the case that $a > 0$. $F(u)$ can be schematically represented by Fig. 1. Fig. 1(a) represents the case $C > 0$, Fig. 1(b) and 1(d) the cases $C < 0$, and Fig. 1(c) the case $C = 0$. Note that the maxima of $F(u)$ are located at $u = \pm\sqrt{a/r}$.

When $C > 0$, the solution is oscillatory. The maximum amplitude u_1 is given by the zero of $F(u)$. The system oscillates between u_1 and $(-u_1)$. The period of the oscillation T is given by the relation

$$T = 2\int_{-u_1}^{u_1} \frac{du}{\sqrt{F(u)}}\,. \tag{2.4}$$

Thus the amplitude and frequency of the oscillation are in general related to each other, and the maximum amplitude $u_1 > \sqrt{a/r}$, which cannot be made as small as we please.

For $C = 0$ [Fig. 1(c)], it represents the degenerate case that the period of the oscillation becomes infinite. Thus the oscillatory solution degenerates to a single pulse.

When $C < 0$, then $u(t)$ never changes sign. If $|C|$ is very large [Fig. 1(d)], then there is no solution. When $|C|$ is small enough, two possible oscillatory states are possible, either $u(t)$ is always positive or always negative.

The detailed motion of the system for this particular problem can be expressed in terms of elliptic functions.

This problem is equivalent to the variational problem that the functional

$$J = \int_0^t \left[\frac{1}{2}\left(\frac{du}{dt}\right)^2 + \frac{a}{2}u^2 - \frac{r}{4}u^4\right]dt, \tag{2.5}$$

is stationary. Now let us take a trial solution of the form

$$u(t) = A(t)\sin\omega t + B(t), \tag{2.6}$$

where $A(t)$ and $B(t)$ are both slowly varying functions of time. Let us substitute (2.6) in (2.5), making use of the slowly varying nature of A and B, then as we have done in previous studies, and particularly in the previous study of the Duffing problem, we obtain

$$J = \int_0^t \left[\frac{1}{4}\left(\frac{dA}{dt}\right)^2 + \frac{1}{4}(\omega^2 + a)A^2 - \frac{3}{32}rA^4 - \frac{3}{4}rA^2B^2 + \frac{1}{2}\left(\frac{dB}{dt}\right)^2 + \frac{a}{2}B^2 - \frac{r}{4}B^4\right]dt. \tag{2.7}$$

The Euler equations obtained from variations of A and B are

$$\frac{d^2A}{dt^2} - (\omega^2 + a)A + \tfrac{3}{4}rA^3 + 3rB^2A = 0, \tag{2.8}$$

and

$$\frac{d^2B}{dt^2} - aB + rB^3 + \tfrac{3}{2}rA^2B = 0. \tag{2.9}$$

Since $A(t)$ and $B(t)$ are slowly varying functions of time, thus as a first approximation, the terms d^2A/dt^2 and d^2B/dt^2 can be neglected, and we have

$$A[3r(A^2/4 + B^2) - (\omega^2 + a)] = 0 \tag{2.10}$$

and

$$B[r(\tfrac{3}{2}A^2 + B^2) - a] = 0. \tag{2.11}$$

One solution is $A = 0$, then $B = 0$, or $B^2 = a/r$. These are also the equilibrium solutions from the original equation (2.1). The former corresponds to the degenerate case depicted in Fig. 1(c), while the latter corresponds to another degenerate case when the maxima of $F(u)$ touch the u axis. From Eq. (2.9) or (2.1), we can infer that the former is an unstable equilibrium while the latter case is stable.

Now ω has been implicitly assumed to be real. Otherwise, the averaging scheme cannot be carried out. Thus for $A \neq 0$, it is necessary that

$$3r(A^2/4 + B^2) \geq a. \tag{2.12}$$

This is the criterion of nonlinear stability. Then the nonlinear frequency relation is given by

$$\omega^2 = 3r(A^2/4 + B^2) - a. \tag{2.13}$$

We can again distinguish two cases.

(i) $B = 0$

Then (2.12) and (2.13) become

$$A^2 \geq 4a/3r, \tag{2.14}$$

and

$$\omega^2 = 3rA^2/4 - a. \tag{2.15}$$

The solution corresponds to the case depicted in Fig. 1(a). A is now identified with the maximum amplitude u_1, the stability criterion (2.14) is consistent with our previous estimate of the lower bound of u_1, while the frequency relation (2.15) is an approximate representation of (2.4).

(ii) $B \neq 0$

Then (2.11) yields

$$a = r(\tfrac{3}{2}A^2 + B^2),$$

and (2.13) becomes

$$B^2 \geq a/5r \geq \tfrac{3}{8}A^2, \quad \text{and} \quad \omega^2 = r(2B^2 - \tfrac{3}{4}A^2).$$

This solution corresponds to the case depicted in Fig. 1(b).

3. HARMONICS AND SUBHARMONICS IN WAVES

Let us now take a trial solution of the form

$$u(t) = A(t)\sin(\omega t) + B(t)\sin(p\omega t), \tag{3.1}$$

where $p \neq 1$ is some positive real number. It is hoped that this trial solution could better represent the solution as depicted in Fig. 1(a). A more realistic representation would be

$$u(t) = A(t)\sin(\omega t + \delta(t)) + B(t)\sin(p\omega t + \nu(t)), \tag{3.2}$$

as in a previous study.[1] A completely analogous development can be followed for the present stability analysis. In the previous study of the problem of forced oscillation, it is found that if the system is dissipative only the subharmonics and harmonics of order three can persist. Furthermore, δ and ν can be taken to be zero if the dissipation is small. For the present problem, oscillation cannot persist indefinitely for this free system if any dissipation is present. Therefore, it is not feasible to consider the corresponding problem with dissipation. However, it is also expected valid that the dominant interactions are among the subharmonics and harmonics of order three. Therefore, we shall only consider the cases $p = \frac{1}{3}$ and $p = 3$.

Take the case $p = 3$. The Euler's equations obtained from variations of A and B are then[4]:

$$\frac{d^2A}{dt^2} - (\omega^2 + a)A + \tfrac{3}{4}rA^3 + \tfrac{3}{2}rB^2A = \tfrac{3}{4}rA^2B, \tag{3.3}$$

$$\frac{d^2B}{dt^2} - (9\omega^2 + a)B + \tfrac{3}{4}rB^3 + \tfrac{3}{2}rA^2B = \tfrac{1}{4}rA^3. \tag{3.4}$$

For the case $p = \frac{1}{3}$, we have

$$\frac{d^2A}{dt^2} - (\omega^2 + a)A + \tfrac{3}{4}rA^3 + \tfrac{3}{2}rB^2A = \tfrac{1}{4}rB^3, \tag{3.3$'$}$$

$$\frac{d^2B}{dt^2} - \left(\frac{\omega^2}{9} + a\right)B\tfrac{3}{4}rB^3 + \tfrac{3}{2}rA^2B = \tfrac{3}{4}rAB^2. \tag{3.4$'$}$$

Thus if $\{A, B; \omega\}$ characterizes the equations (3.3) and (3.4), then the equations (3.3′) and (3.4′) are characterized by $\{B, A; \omega/3\}$. Therefore, these two sets of equations are completely equivalent, and it is sufficient to treat equations (3.3) and (3.4) only. Since $A(t)$ and $B(t)$ are again slowly varying functions of t, the terms d^2A/dt^2 and d^2B/dt^2 are neglected, and we obtain

$$A[(\omega^2 + a) - \tfrac{3}{4}r(A^2 + 2B^2 - AB)] = 0, \tag{3.5}$$

$$(9\omega^2 + a)B - (r/4)(3B^3 + 6A^2B - A^3) = 0. \tag{3.6}$$

Let us now discuss the various possible solutions.

(i) $A = 0$

Then only the B mode exists. If we write $\nu = 3\omega$, Eq. (3.6) becomes

$$(\nu^2 + a)B - \tfrac{3}{4}rB^3 = 0, \tag{3.7}$$

which is exactly the same as the case (i) treated in Sec. 2. The stability criterion is, as before, given by

$$B^2 \geq 4a/3r. \tag{3.8}$$

The fact that $A = 0$ is a permissible solution implies that subharmonic generation is not necessarily required.

(ii) If $B = 0$, then it is necessary that $A = 0$. In other words, if $A \neq 0$, then $B \neq 0$. Thus a harmonic of order three will always accompany the fundamental mode.

(iii) $A \neq 0$

Then we have

$$\omega^2 = \tfrac{3}{4}r(A^2 + 2B^2 - AB) - a$$
$$= (r/36)(3B^2 + 6A^2 - A^3/B) - a/9. \tag{3.9}$$

Since ω^2 has to be positive, it is necessary for stability that

$$A^2 + 2B^2 - AB \geq 4a/3r, \tag{3.10}$$

and

$$B^2 + 2A^2 - A^3/3B \geq 4a/3r. \tag{3.11}$$

We can take $A > 0$ without loss of generality, then it may be seen that it is favorable to stability for $B < 0$.

Now Eqs. (3.9) can be rewritten as

$$\omega^2 = (rB^2/32)\,G(A/B), \tag{3.12}$$

and

$$a = (rB^2/32)\,F(A/B), \tag{3.13}$$

where

$$F(x) = x^3 + 21x^2 - 27x + 51, \tag{3.14}$$

and

$$G(x) = -(x^3 - 3x^2 - 3x + 3). \tag{3.15}$$

Thus the stability criterion can be represented by

$$G(A/B) > 0, \tag{3.16}$$

while the amplitudes A and B are to be determined by the equation (3.13). Thus once any B is chosen, then equation (3.13) will determine A. The number of possible solutions are either one or three. Whether any of these solutions are permissible is to be checked by relation (3.16), and the frequency ω is in turn determined by Eq. (3.12). Figure 2 represents the curves $F(x)$ and $G(x)$. Thus $G(x) > 0$ only if

$$0.66 < x < 3.6 \quad \text{or} \quad x < -1.26.$$

So the stability criterion requires that

$$0.66 < A/B < 3.6 \quad \text{or} \quad A/B < -1.26. \tag{3.17}$$

In order that Eq. (3.13) be satisfied, we can see that there is a solution only if

$$32a/rB^2 < 1809. \tag{3.18}$$

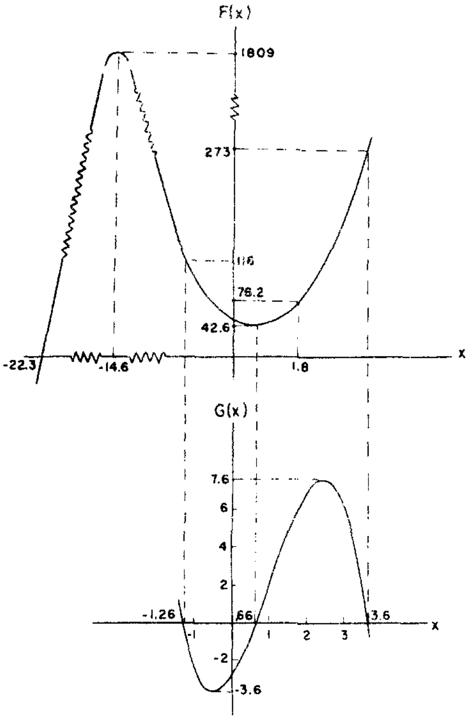

FIG. 2.

Then there is at least one solution. In fact it may be seen that, there are two solutions if $273 < 32a/rB^2 < 1809$, or $42.6 < 32a/rB^2 < 116$, and three solutions if $116 < 32a/rB^2 < 273$, and one solution if $32a/rB^2 < 42.6$.

These results can be interpreted in the manner of harmonic and subharmonic generation. When $|A/B| \gg 1$, we can interpret ω as the fundamental frequency and regard it as harmonic generation. On the other hand, when $|A/B| = O(1)$, we may interpret 3ω as the fundamental frequency and regard it as subharmonic generation, while relation (3.18) gives the threshold amplitude of the fundamental mode for stability. However, viewed with our knowledge about the exact solution of the problem, these interpretations are rather artifical. Indeed, given a and r of the system, the solution is determined completely by the initial conditions, or alternatively determined by the maximum amplitude, or some other parameter which characterizes the average amplitude, e.g., average energy. The values of A, B, and ω, given by Eqs. (3.12) and (3.13), are then the best values if the exact solution is to be approximated by the two term representation (3.1). The case (i), for which $A=0$, is the best result for the one term representation.

In general, the more terms we take in the representation, the better is the approximation. However, with the exact solution as the guide, it becomes clear also that when $|A/B| \gg 1$, the two term representation will differ little from the one term representation. Therefore good approximation can already be achieved with a one term approximation. It is only when $|A/B| = O(1)$, that the two term representation is significantly better than the one term representation. Only then may it be necessary to consider the representation with more than two terms. This may indeed be the case in the neighborhood of the critical region when $\omega \approx 0$, which corresponds to the case as depicted in Fig. 1(c). The exact solution for that case is a periodic succession of intermittant pulses.

One way to rule out some of the multiplicity of possible solutions is to go back to Eqs. (3.3) and (3.4), and study the stability of these differential equations about these possible solutions with constant amplitudes. Then it can be found, the solution for case (i) is always stable, i.e., the representation by a single mode is always good in some sense. On the other hand, for case (iii), the constant amplitude solution is stable only for

$$1.8 < A/B < 3.61.$$

Therefore, only one solution is permissible when

$$76.2 < 32a/rB^2 < 273.$$

4. A WAVE EQUATION

Let us consider a model wave equation in the following form:

$$u_{tt} + 2vu_{xt} + (v^2 - c^2)u_{xx} - au + ru^3 = 0, \tag{4.1}$$

where v, a, c, and r are all real constants. The linear system with $r=0$ has been discussed by Sturrock.[5] It is evident that this equation is a generalization of the Duffing problem of oscillation to wave propagation. Some of the results we obtained in the previous two sections may thus be applied to this system with slight modifications.

Let us first consider the linear case when $r=0$. If we assume a solution of the form $\exp[i(kx+\omega t)]$, then the dispersion relation is readily found to be

$$(\omega + kv)^2 = c^2k^2 - a. \tag{4.2}$$

Thus the system is stable for $a \leq 0$. When $a > 0$, one of the modes is unstable in the interval,

$$|k| < \sqrt{a}/c, \tag{4.3}$$

where c and $\sqrt{a}$ are taken to be positive.

It may also be seen that[5] the system exhibits absolute instability when $c > |v|$ and convective instability when $c < |v|$, since the characteristic velocities of the system are $-v \pm c$. Thus they are in different directions when $c > |v|$ and in the same direction when $c < |v|$. An initial pulse of compact spatial support will spread and grow in amplitude in both cases. However, for the case $c < |v|$ the amplitude at any fixed x will die out for large t, i.e., the instability is convective. For periodic wave trains, the distinction between absolute and convective instabilities is not as noteworthy.

When $r > 0$, we expect that the instability of the linear system would be stabilized by the nonlinear effect, and waves with finite and slowly varying amplitude would persist. We shall again apply the variational method to this problem.

The functional corresponding to Eq. (4.1) is

$$J = \int_0^t dt \int_{-\infty}^{\infty} \tfrac{1}{2}[u_t^2 + 2vu_xu_t + (v^2 - c^2)u_x^2 + au^2 - (r/2)u^4]dx. \tag{4.4}$$

As in the Duffing problem, let us consider the trial solution first in the form

$$u(x, t) = A(x, t)\sin(S(x, t)) + B(x, t), \tag{4.5}$$

where A and B are slowly varying functions of x and t. We shall also denote

$$\omega = S_t \quad \text{and} \quad k = S_x, \tag{4.6}$$

and they are also slowly varying functions of x and t. Following the procedure we used before, we obtain

$$J \cong \int_0^t dt \int_{-\infty}^{\infty} \tfrac{1}{2}[\tfrac{1}{2}A_t^2 + \tfrac{1}{2}\omega^2 A^2 + B_t^2$$
$$+ v(A_t A_x + \omega k A^2 + 2B_t B_x) + (v^2 - c^2)(\tfrac{1}{2}A_x^2 + \tfrac{1}{2}k^2 A^2 A^2 + B_x^2)$$
$$+ a(\tfrac{1}{2}A^2 + B^2) - (r/2)(\tfrac{3}{8}A^4 + 3A^2B^2 + B^4)]\,dx. \tag{4.7}$$

The Euler equations from variations of A and B are

$$A_{tt} + 2vA_{xt} + (v^2 - c^2)A_{xx} - [\omega^2 + 2v\omega k + (v^2 - c^2)k^2 + a]A$$
$$+ \tfrac{3}{4}rA^3 + 3rAB^2 = 0, \tag{4.8}$$

and

$$B_{tt} + 2vB_{xt} + (v^2 - c^2)B_{xx} - aB + rB^3 + \tfrac{3}{2}rA^2B = 0, \tag{4.9}$$

while the variation with respect to S leads to

$$(\omega A^2)_t + v[(\omega A^2)_x + (kA^2)_t] + (v^2 - c^2)(kA^2)_x = 0. \tag{4.10}$$

Since A and B are slowly varying functions of (x, t), the second derivatives of A and B will be neglected in (4.8) and (4.9), and we have

$$A[(\omega + kv)^2 + a - c^2k^2 - 3r(A^2/4 + B^2)] = 0, \tag{4.11}$$

and

$$B[r(\tfrac{3}{2}A^2 + B^2) - a] = 0. \tag{4.12}$$

These equations are similar to Eqs. (2.10) and (2.11). As before, one permissible solution is $A = 0$. Then we have $B = 0$, or $B^2 = a/r$, which represent constant states.

If $A \neq 0$, then we have

$$(\omega + kv)^2 = c^2k^2 + 3r(A^2/4 + B^2) - a, \tag{4.13}$$

thus the system is stable for all k when

$$3r(A^2/4 + B^2) \geq a. \tag{4.14}$$

The last relation implies that the system can be completely stabilized when the amplitude is large enough. Now for any particular mode with finite k, the stability condition need not be so restrictive. It could be replaced by

$$3r(A^2/4 + B^2) \geq a^2 - c^2k^2. \tag{4.15}$$

When the stability condition is satisfied, then Eq. (4.13) gives the nonlinear dispersion relation for that mode of the wave.

As before, we again have two cases to consider. Let us first consider the case $B = 0$. Then Eqs. (4.13) and (4.15) become

$$\omega = -kv \pm [c^2k^2 + \tfrac{3}{4}rA^2 - a]^{1/2}, \tag{4.16}$$

and

$$\tfrac{3}{4}rA^2 \geq a - c^2k^2. \tag{4.17}$$

Thus for those k such that $k < \sqrt{a}/c$, the amplitude of the wave has to exceed certain values as given by relation (4.17). From the dispersion relation (4.16), we can also obtain the group velocity

$$\frac{d\omega}{dk} = -v \pm \frac{c^2k + \frac{3}{4}rA\,dA/dk}{[c^2k^2 + \frac{3}{4}rA^2 - a]^{1/2}}. \tag{4.18}$$

It is of interest to note that, if we take ω and k as constants, Eq. (10) leads to

$$(\omega + kv)A_t + [\omega v + k(v^2 - c^2)]A_x = 0$$

or

$$A_t - c_g A_x = 0, \tag{4.19}$$

where

$$c_g = -v \pm \frac{c^2k}{[c^2k^2 + \frac{3}{4}rA^2 - a]^{1/2}}, \tag{4.20}$$

after using Eq. (4.16). The last result is consistent with (4.18), if we set $dA/dk = 0$. For $\frac{3}{4}rA^2 > a - c^2k^2$, A can indeed be considered as independent of k. However, at the critical amplitude $\frac{3}{4}rA^2 = a - c^2k^2$, the term dA/dk has to be retained in order to avoid the singular behavior of the group velocity. At this critical amplitude we have

$$\omega = -kv \quad \text{and} \quad \frac{d\omega}{dk} = -v. \tag{4.21}$$

Let us also note that for the linear case, for $k < \sqrt{a}/c$, the growth rate of the instability as given by Eq. (4.2) is

$$\omega_i = (a - c^2k^2)^{1/2}, \tag{4.22}$$

and ω_i has a maximum at $k = 0$. It is interesting to see that, if we vary with respect to ω and k the functional given by Eq. (4.7), we obtain

$$\omega + kv = 0 \tag{4.23}$$

and

$$\omega v + k(v^2 - c^2) = 0. \tag{4.24}$$

In conjunction with (4.16), Eq. (4.23) will imply that the amplitude be the critical amplitude, while Eq. (4.24) then gives $k = 0$, i.e., the mode that corresponds to maximum growth from the linear instability.

The physical picture which emerges from the above analysis may be put in the following manner. For an initial small pulse given initially by

$$u(x, 0) = \int_0^{\infty} a_k \exp(ikx)\,dk, \tag{4.25}$$

with $u_t(x, 0)$ such that only one branch (say the + branch) of the solutions as given by the dispersion relation is present, we expect that for large t, the solution is approximately given by

$$u(x, t) \sim \int_0^{\infty} A(k)\exp[i(kx + \omega(k)t)]\,dk, \tag{4.26}$$

where

$$A(k) = a_k, \quad \omega(k) \cong -kv + [c^2k^2 - a], \quad \text{for } k \geq \frac{\sqrt{a}}{c}, \tag{4.27}$$

and

$$A(k) = \frac{4}{3r}(a - c^2k^2), \quad \omega(k) = -kv, \quad \text{for } k \leq \frac{\sqrt{a}}{c}. \tag{4.28}$$

On the other hand, if we attempt to represent the asymptotic solution by a single mode, then the mode with $k = 0$ would give the best representation.

The case with $B \neq 0$ can again be analyzed in a similar manner. Using (4.12), relation (4.15) gives

$$B^2 \geq \frac{a}{5r} - \frac{2}{5r} c^2k^2, \quad B^2 \geq \frac{3}{8} A^2 - \frac{1}{2r} c^2k^2,$$

$$A^2 \leq \frac{4}{15r} (2a + c^2k^2).$$

The value of B is more or less determined by the initial condition, and hence can be considered as given. In order that all k are included, it is necessary that

$$B^2 \geq a/5r.$$

Then A is given by Eq. (4.12), and we can proceed as in the previous case.

5. HARMONIC AND SUBHARMONICS IN WAVES

As in Sec. 3, let us now use the trial solution of the form

$$u(x, t) = A(x, t) \sin(S(x, t)) + B(x, t) \sin(pS(x, t)). \quad (5.1)$$

We shall again take $p = 3$ or $p = \frac{1}{3}$, and it is sufficient as before to consider only the case $p = 3$. Substituting (5.1) in (4.4), and carrying out the usual procedure, we obtain

$$\begin{aligned} J \cong \int_0^t dt \int_{-\infty}^{+\infty} \tfrac{1}{2}[\tfrac{1}{2}(A_t^2 + \omega^2 A^2 + B_t^2 + 9\omega^2 B^2) + v(A_t A_x + \omega k A^2 \\ + B_t B_x + 9\omega k B^2) + \tfrac{1}{2}(v^2 - c^2)(A_x^2 + k^2 A^2 + B_x^2 + 9k^2 B^2) \\ + (a/2)(A^2 + B^2) - (r/2)(\tfrac{3}{8} A^4 + \tfrac{3}{2} A^2 B^2 + \tfrac{3}{8} B^2 - \tfrac{1}{2} A^3 B)]\, dx. \end{aligned} \quad (5.2)$$

The Euler's equations from variation with respect to A, B, and S are

$$\begin{aligned} A_{tt} + 2vA_{tx} + (v^2 - c^2) A_{xx} - [\omega^2 + 2v\omega k \\ + (v^2 - c^2) k^2 + a] A + \frac{3r}{4} (A^3 + 2AB^2 - A^2 B) = 0, \end{aligned} \quad (5.3)$$

$$\begin{aligned} B_{tt} + 2vB_{tx} + (v^2 - c^2) B_{xx} - 9 \bigg(\omega^2 + 2v\omega k \\ + (v^2 - c^2) k^2 + \frac{a}{9}\bigg) B + \frac{r}{4} (3B^3 + 6A^2 B - A^3) = 0, \end{aligned} \quad (5.4)$$

and

$$[(A^2 + 9B^2)(\omega + kv)]_t + [(A^2 + 9B^2)(\omega v + kv^2 - kc^2)]_x = 0. \quad (5.5)$$

If the second derivatives of A and B and the first derivatives of ω and k are neglected, then we obtain:

$$A\{[(\omega + kv)^2 + a - c^2k^2] - (3r/4)(A^2 + 2B^2 - AB)\} = 0, \quad (5.6)$$

$$[(\omega + kv)^2 + a/9 - c^2k^2] B - (r/36)(3B^3 + 6A^2 B - A^3) = 0, \quad (5.7)$$

and

$$(A^2 + 9B^2)_t + \frac{(\omega + kv) v - kc^2}{\omega + kv} (A^2 + 9B^2)_x = 0. \quad (5.8)$$

Like the case treated in Sec. 3, among the permissible solutions, if $A = 0$, then it reduces to the case treated in Sec. 4; if $B = 0$ then it is necessary that $A = 0$ also. The more interesting case is that in which both A and B are not vanishing, and we have

$$\begin{aligned} (\omega + kv)^2 &= (3r/4)(A^2 + 2B^2 - AB) + c^2k^2 - a \\ &= (r/36B)(3B^3 + 6A^2 B - A^3) + c^2k^2 - a/9, \end{aligned} \quad (5.9)$$

or

$$(\omega + kv)^2 = c^2k^2 + \frac{rB^2}{32} G\left(\frac{A}{B}\right), \quad (5.10)$$

and

$$a = \frac{rB^2}{32} F\left(\frac{A}{B}\right), \quad (5.11)$$

where $F(x)$ and $G(x)$ are the same function given in Sec. 3.

Again, Eq. (5.8) can be written as

$$(A^2 + 9B^2)_t - c_g A_x = 0,$$

where

$$c_g = \frac{d\omega}{dk} = -v \pm c^2 k \bigg/ \left[c^2k^2 + \frac{rB^2}{32} G\left(\frac{A}{B}\right)\right]^{1/2}. \quad (5.12)$$

As in Sec. 3, the permissible values of A and B are nonunique. Here we may interpret this intriguing feature in the following manner. We are trying to approximate the true solution by "best" representations in terms of Fourier modes. More than one representation can be "best" in comparison with its neighboring representations in the sense of the variational method.

These harmonic and subharmonic generations may serve as a mechanism to transfer energy from one part of the spectrum to others. It is evident from the possibility of these harmonic and subharmonic generations, that the details of the asymptotic solution for an initial pulse are indeed extremely complex. Conceivably, three or more term representations of the solution will yield even more complex results. However, in the spirit of the approximate variational method, the results obtained so far have enough information on the gross features of the problem for us to pause and analyze the implications in conjunction with observations in various physical problems.

6. TWO-STREAM INSTABILITY

Let us consider a multicomponent plasma, consisting of electrons and ions, coupled only through the self-consistent fields. Then the basic equations are as follows:

$$\frac{\partial \mathbf{v}_i}{\partial t} + (\mathbf{v}_i \cdot \nabla) \mathbf{v}_i = \frac{e_i}{m_i} \left(\mathbf{E} + \frac{1}{c} \mathbf{v}_i \times \mathbf{B}\right) - \frac{1}{m_i n_i} \nabla p_i, \quad (6.1)$$

$$\frac{\partial n_i}{\partial t} + \nabla \cdot (n_i \mathbf{v}_i) = 0, \quad (6.2)$$

$$\nabla \times \mathbf{B} = \frac{4\pi}{c} \mathbf{J} + \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t}, \quad (6.3)$$

$$\nabla \times \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}, \quad (6.4)$$

$$\mathbf{J} = \sum_i e_i n_i \mathbf{v}_i, \quad (6.5)$$

where e_i denotes the charge, m_i the mass, $\mathbf{v}_i$ the macroscopic velocity, n_i the density, p_i the partial pressure of the ith constituent, and $\mathbf{E}$, $\mathbf{B}$, and $\mathbf{J}$ are the electric field, magnetic field, and electric current, respectively. An equation of state to define p_i will also be needed, which we shall take as

$$p_i = m_i n_i^2 \frac{\partial U_i}{\partial n_i}, \quad (6.6)$$

where $U_i(n_i, s_i)$ is the internal energy of the ith species. We shall consider the case that the heating of plasma is negligible, and that the entropies s_i are constants.

If we introduce the scalar and vector potentials ϕ and $\mathbf{A}$ such that

$$\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}, \quad \mathbf{B} = \nabla \times \mathbf{A}, \tag{6.7}$$

then Maxwell's equations (6.3)—(6.5) become

$$\nabla^2 \mathbf{A} - \frac{1}{c^2}\frac{\partial^2 A}{\partial t^2} = -\frac{4\pi}{c}\mathbf{J}, \tag{6.8}$$

$$\nabla^2 \phi - \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} = -4\pi \sum_i e_i n_i, \tag{6.9}$$

when the following gauge condition is used:

$$\nabla \cdot \mathbf{A} + \frac{1}{c}\frac{\partial \phi}{\partial t} = 0. \tag{6.10}$$

It can be readily verified that an equivalent variational problem to what is formulated above is that the functional

$$I = \int_{t_i}^{t_2} dt \int_v d^3x \left\{ \frac{1}{8\pi}\left(\nabla\phi + \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}\right)^2 - \frac{1}{8\pi}(\nabla \times \mathbf{A})^2 + \sum_i \left[\frac{m_i n_i}{2}\mathbf{v}_i^2 - m_i n_i U_i - e_i n_i \phi + \frac{e_i n_i}{c}\mathbf{v}_i \cdot \mathbf{A} + \alpha_i\left(\frac{\partial n_i}{\partial t} + \nabla \cdot (n_i \mathbf{v}_i)\right)\right]\right\} \tag{6.11}$$

is an extremum, subject to the subsidiary conditions

$$\frac{\partial n_i}{\partial t} + \nabla \cdot (n_i \mathbf{v}_i) = 0. \tag{6.2}$$

The variables to be varied are n_i, $\mathbf{v}_i$, ϕ, and $\mathbf{A}$, and α_i are the Lagrange multipliers. Equations (6.8) and (6.9) can also be obtained, if condition (6.10) is imposed.

Let us now consider a plasma which consists of electrons and an ion background which is uniform and stationary. The electrons are, however, divided into two beams in the primary state. Therefore, we can consider the plasma as composed of three species, with species 1 and 2 to designate the two electron beams and the species 0, the ion background. Thus, we have

$$e_1 = e_2 = -e_0 = -e, \quad m_1 = m_2 = m, \qquad v_0 = 0, \quad m_0 \to \infty, \quad \langle n_1 + n_2\rangle = n_0 = N_0, \tag{6.12}$$

where m is the electronic mass, e the absolute value of the electronic charge, N_0 the background ion density, and the mass of the ion is taken to be infinitely large to make the ion background stationary. The last equation, in which $\langle\ \rangle$ denotes the spatial average, is the condition for the over-all neutrality of the plasma.

We shall consider only the one-dimensional pattern. Therefore, everything will depend on x and t only. Then the magnetic field does not play a role at all and we can take $\mathbf{A} = 0$. Thus the functional I becomes

$$I = \int_{t_1}^{t_2} dt \int_{-\infty}^{+\infty} \left\{ \frac{1}{8\pi}\left(\frac{\partial\phi}{\partial x}\right)^2 - e(n_1 + n_2 - N_0)\phi + \sum_{i=1}^{2}\left[\frac{m}{2}n_i \mathbf{v}_i^2 - m n_i U_i + \alpha_i\left(\frac{\partial n_i}{\partial t} + \nabla \cdot (n_i \mathbf{v}_i)\right)\right]\right\} dx. \tag{6.13}$$

The primary state which describes two electron beams moving uniformly with velocities V_1 and V_2 relative to the ion background is given by the following solution:

$$v_1 = V_1, \quad v_2 = V_2, \quad n_1 = N_1, \quad n_2 = N_2, \quad N_1 + N_2 = N_0,$$
$$\phi = 0, \quad \alpha_1 = mV_1 x - \frac{m}{2}V_1^2 t \equiv \alpha_{10}, \quad \alpha_2 = mV_2 x - \frac{m}{2}V_2^2 t \equiv \alpha_{20}. \tag{6.14}$$

It may be pointed out that the Lagrangian multipliers $\alpha_{1,2}$ play the role of the velocity potential as revealed from the equation obtained from the variation of I with respect to $v_{1,2}$.

The internal energy functions U_i are of the form

$$U_i = \frac{KT_i}{m_i(\gamma - 1)}\left(\frac{n_i}{N_i}\right)^{\gamma-1}, \tag{6.15}$$

where K is the Boltzmann's constant, T_i is the equilibrium temperature of the ith species in the primary state, and γ is the ratio of the specific heats. For the electron plasma under consideration, γ is taken to be 3.[6]

Now let us take the trial solutions:

$$v_i = V_i + v_{i1}(x,t)\sin(S(x,t)), \tag{6.16}$$

$$n_i = N_i + n_{i1}(x,t)\sin(S(x,t)), \tag{6.17}$$

$$\phi = \phi_1(x,t)\sin(S), \tag{6.18}$$

$$\alpha_i = \alpha_{i0} + \alpha_{i1}\cos(S). \tag{6.19}$$

With $\gamma = 3$, we can rewrite (6.15) as

$$m_i n_i U_i = \frac{K}{2}N_i T_i\left(1 + \frac{n_{i1}}{N_i}\sin(S)\right)^3. \tag{6.20}$$

Substitute (6.16)—(6.20) into (6.13), and carrying out the averaging procedure as before, we obtain

$$\begin{aligned} I \cong I_1 = \int_{t_1}^{t_2} dt \int_v dx \Big\{ &\frac{1}{16\pi}(\phi_{1,x}^2 + \phi_1^2 S_x^2) \\ &+ \sum_{i=1}^{2}\Big[\frac{m}{2}\left(N_i V_i^2 + V_i v_{i1} n_{i1} + \frac{N_i}{2}v_{i1}^2\right) \\ &- \frac{K}{2}N_i T_i\left(1 + \frac{3}{2}\frac{n_{i1}^2}{N_i^2}\right) - \frac{e}{2}n_{i1}\phi_1 \\ &+ \frac{\alpha_{i0}}{2}(n_{i1}v_{i1,x} + v_{i1}n_{i1,x}) \\ &+ \frac{\alpha_{i1}}{2}(n_{i1}S_t + N_i V_{i1} S_x + V_i n_{i1} S_x)\Big]. \end{aligned} \tag{6.21}$$

Now let us vary Eq. (6.21) with respect to n_{i1}, v_{i1}, ϕ_1, and α_{i1}, and then neglect terms containing their derivatives with respect to x and t. We thus obtain the following equations:

$$\delta n_{i1}: \ -e\phi_1 + \alpha_{i1}(\omega + kV_i) - 3KT_i\left(\frac{n_{i1}}{N_i}\right) = 0, \tag{6.22}$$

$$\delta v_{i1}: \ m v_{i1} + \alpha_{i1} k = 0, \tag{6.23}$$

$$\delta \phi_1: \ k^2\phi_1 - 4\pi e(n_{11} + n_{21}) = 0, \tag{6.24}$$

$$\delta\alpha_{i1}:\ (\omega + kV_i) n_{i1} + kN_i v_{i1} = 0, \tag{6.25}$$

where we have again written

$$S_x = k \ \text{ and } \ S_t = \omega.$$

Denote the plasma frequencies ω_{pi}, and the sound speeds c_i by

$$\omega_{pi}^2 = 4\pi N_i e^2/m, \tag{6.26}$$

and

$$c_i^2 = 3KT_i/m. \tag{6.27}$$

The elimination of n_{i1}, v_{i1}, ω_1, and α_{i1} then leads to the dispersion relation

$$1 = \frac{\omega_{p1}^2}{(\omega + kV_1)^2 - c_1^2 k^2} + \frac{\omega_{p2}^2}{(\omega + kV_2)^2 - c_2^3 k^2} \tag{6.28}$$

The last dispersion is the same as the dispersion relation for the linear problem. This is not surprising, because when the trial solution took the form of (6.10)—(6.19), the resulting approximate functional turned out to consist of only linearly interacting terms if γ is exactly 3. This may indicate that the nonlinear effects may only have marginal effect on the basic instability of the two stream problem.

The detailed analysis of the dispersion relation (6.28) is well known.[7] We may only mention that for small c_1 and c_2. The system is unstable at least for some range of k. However, when

$$c_1 + c_2 > |V_1 - V_2| \tag{6.29}$$

the instability is inhibited for all k.

It may be of interest if $\gamma \neq 3$, then (6.20) is replaced by

$$m_i n_i U_i = \frac{K}{\gamma - 1} N_i T_i \left(1 + \frac{n_{i1}}{N_i} \sin(S)\right)^\gamma. \tag{6.30}$$

Then to the order of $(n_i/N_i)^3$, Eq. (6.22) is to be replaced by

$$-e\phi_1 + \alpha_{i1}(\omega + kV_i) - \gamma KT_i\left(\frac{n_{i1}}{N_i}\right) - \tfrac{1}{8}\gamma(\gamma-2)(\gamma-3)KT_i\left(\frac{n_{i1}}{N_i}\right)^3 = 0. \tag{6.31}$$

If we set consistently the sound speed c_i by

$$c_i^2 = \gamma KT_i/m,$$

then the dispersion relation is corrected to become

$$1 = \omega_{p1}^2 \Big/ \left\{(\omega + kV_1)^2 - k^2 c_1^2\left[1 + \frac{(\gamma-2)(\gamma-3)}{8}\left(\frac{n_{11}}{N_1}\right)^2\right]\right\} + \omega_{p2}^2 \Big/ \left\{(\omega + kV_2)^2 - k^2 c_2^2\left[1 + \frac{(\gamma-2)(\gamma-3)}{8}\left(\frac{n_{21}}{N_1}\right)^2\right]\right\}. \tag{6.32}$$

Thus if $\gamma > 3$, the instability is inhibited by the nonlinearity; while if $\gamma < 3$, the instability of the system is enhanced by the nonlinearity.

If we vary Eq. (6.21) with respect to S, we obtain

$$\frac{1}{8\pi}\frac{\partial}{\partial x}(k\phi_1^2) + \frac{1}{2}\sum_{i=1}^{2}\left(\frac{\partial}{\partial t}(\alpha_{i1}n_{i1}) + \frac{\partial}{\partial x}(N_i\alpha_{i1}v_{i1} + V_i\alpha_{i1}n_{i1})\right) = 0. \tag{6.33}$$

When ω and k are taken as constants, the last equation can again be written as

$$\frac{\partial\phi_1}{\partial t} - c_g\frac{\partial\phi_1}{\partial x} = 0, \tag{6.34}$$

where $c_g = d\omega/dk$ as given by the dispersion relation (6.28).

When $\gamma = 3$, the nonlinear effect on instability can be brought out explicitly if we include harmonics in the trial function. To this end, let us take the following trial solution:

$$v_i = V_i + v_{i1}\sin(S) + v_{i2}\cos(2S), \tag{6.35}$$

$$n_i = N_i + n_{i1}\sin(S) + n_{i2}\cos(2S), \tag{6.36}$$

$$\phi = \phi_1\sin(S) + \phi_2\cos(2S), \tag{6.37}$$

$$\alpha_i = \alpha_{i0} + \alpha_{i1}\cos(S) + \alpha_{i2}\sin(2S). \tag{6.38}$$

Then the approximate functional becomes

$$\begin{aligned} I \cong I_1 + \int_{t_1}^{t_2} dt \int_v d^3x \Bigg\{ & \frac{1}{16\pi}(\phi_{2,x}^2 + 4\phi_2^2 S_x^2) \\ & + \sum_{i=1}^{2}\Bigg[\frac{m}{2}\left(V_i v_{i2} n_{i2} + \frac{N_i}{2}v_{i2}^2 - \frac{1}{4}n_{i2}v_{i1}^2\right) \\ & - \frac{K}{2}N_i T_i\left[\frac{3}{2}\left(\frac{n_{i2}}{N_i}\right)^2 - \frac{3}{4}\frac{n_{i1}^2 n_{i2}}{N_i^3}\right] - \frac{e}{2}n_{i2}\phi_2 \\ & + \frac{\alpha_{i0}}{2}(n_{i2}v_{i2,x} + v_{i2}n_{i2,x}) - \frac{\alpha_{i1}}{4}S_x(n_{i1}v_{i2} + v_{i1}n_{i2}) \\ & - \alpha_{i2}(n_{i2}S_t + N_i v_{i2}S_x + V_i n_{i2}S_x - \tfrac{1}{2}n_{i1}v_{i1}S_x)\Bigg]\Bigg\}. \end{aligned} \tag{6.39}$$

The variations with respect to n_{i1}, v_{i1}, ϕ_1, α_{i1}, n_{i2}, v_{i2}, ϕ_2, and α_{i2} then lead to

$$\delta n_{i1}:\quad -e\phi_1 + \alpha_{i1}(\omega + kV_i) - 3KT_i\left(\frac{n_{i1}}{N_i}\right) = -\frac{3}{2}KT_i\left(\frac{n_{i1}n_{i2}}{N_i^2}\right) + \frac{\alpha_{i1}}{2}kv_{i2} - \alpha_{i2}kv_{i1}, \tag{6.40}$$

$$\delta v_{i1}:\quad N_i(mv_{i1} + \alpha_{i1}k) = \frac{m}{2}n_{i2}v_{i1} + \frac{\alpha_{i1}}{2}kn_{i2} - \alpha_{i2}kn_{i1}, \tag{6.41}$$

$$\delta\phi_1:\quad k^2\phi_1 - 4\pi e(n_{11} + n_{21}) = 0, \tag{6.42}$$

$$\delta\alpha_{i1}:\quad (\omega + kV_i)n_{i1} + kN_i v_{i1} = \frac{k}{2}(n_{i1}v_{i2} + v_{i1}n_{i2}), \tag{6.43}$$

$$\delta n_{i2}:\quad e\phi_2 - 2\alpha_{i2}(\omega + kV_i) - 3kT_i\left(\frac{n_{i2}}{N_i}\right) = -\frac{3}{4}KT_i\left(\frac{n_{i1}}{N_i}\right)^2 + \frac{m}{4}v_{i1}^2 + \frac{\alpha_{i1}}{2}kv_{i1}, \tag{6.44}$$

$$\delta v_{i2}:\quad N_i(mv_{i2} - 2k\alpha_{i2}) = \frac{k}{2}\alpha_{i1}n_{i1}, \tag{6.45}$$

$$\delta\phi_2:\quad k^2\phi_2 - \pi e(n_{12} + n_{22}) = 0, \tag{6.46}$$

$$\delta\phi_{i2}:\quad (\omega + kV_i)n_{i2} + kN_i v_{i2} = \frac{k}{2}(n_{i1}v_{i1}). \tag{6.47}$$

In these expressions, the derivatives with respect to x and t of the varied quantities have again been neglected.

It may be noted that in comparison with equations

(6.22)–(6.25) the terms on the right-hand side of the equations (6.40)–(6.43) are those due to the harmonic generations. Moreover, the structures of the Eqs. (6.44)–(6.47) are quite similar to those of (6.40)–(6.43). The nature of the mutual interaction is also reminiscent to what we have discussed in Sec. 5. Again a permissible set of solutions is $\{u_{i,1}\}=0$, where u stands for any of the variables of concern. On the other hand, $\{u_{i,2}\}=0$ will imply $\{u_{u,1}\}=0$. Thus harmonics are always generated from a fundamental mode, due to any nonlinear interaction; whereas the subharmonics can only be generated when a certain threshold amplitude is reached for the fundamental mode. A more detailed analysis of Eq. (6.40)–(6.47) will be presented elsewhere.

7. THE KELVIN-HELMHOLTZ STABILITY

Consider two incompressible, inviscid fluids separated by an interface

$$F(\mathbf{r},t)=0. \tag{7.1}$$

We shall use subscripts 1 and 2 to denote variables in these two fluids. Let us assume that the flow of the fluids is irrotational. Thus we have in the ith region that

$$\mathbf{v}_i=\nabla\phi_i, \tag{7.2}$$

$$\nabla^2\phi_i=0, \tag{7.3}$$

and

$$\frac{p_i}{\rho_i}+\tfrac{1}{2}(\nabla\phi_i)^2+\Omega_i+\frac{\partial\phi_i}{\partial t}=f_i(t), \tag{7.4}$$

where $\mathbf{v}_i$ is the fluid velocity, ϕ_i the velocity potential, ρ_i the density, Ω_i the external force potential, and $f_i(t)$ an integration constant. The interfacial conditions are

$$\frac{\partial F}{\partial t}+(\mathbf{v}\cdot\nabla)F=0 \quad \text{on} \quad F=0, \tag{7.5}$$

and

$$p_1-p_2=\sigma\left(\frac{1}{R_a}+\frac{1}{R_b}\right) \quad \text{on } F=0, \tag{7.6}$$

where σ is the surface tension coefficient, and R_a and R_b are the principal radii of curvature at the point under consideration on $F=0$. R_a (R_b) is to be positive if its center of curvature lies on the side of region 1, and negative otherwise.

The problem formulated above is equivalent to the variational principle that the following functional J is stationary[8]:

$$J=\int_{t_1}^{t_2}dt\,(\mathcal{J}_1+\mathcal{J}_2+\mathcal{J}_a), \tag{7.7}$$

where

$$\mathcal{J}_i=\rho_i\int_{V_i}d^3r\left(\frac{\partial\phi_i}{\partial t}+\frac{1}{2}(\nabla\phi_i)^2+\Omega_i\right),\quad i=1,2 \tag{7.8}$$

and

$$\mathcal{J}_a=\sigma\int_F da, \tag{7.9}$$

where the integral in (7.9) is the total surface area of the interface $F=0$. The variations with respect to ϕ_i lead to Eqs. (7.3) and (7.5), while the variation with respect to the interface leads to Eq. (7.6).

Let us consider two-dimensional flow problems and let the gravity be the only external force field which applies in the direction of $(-y)$. Thus

$$\Omega_i=gy. \tag{7.10}$$

Now take the trial solution of the interface relation (7.1) of the form

$$y=\eta(x,t)\equiv A(x,t)\sin(S(x,t)), \tag{7.11}$$

and let $y>\eta$ be region 1, $y<\eta$ be region 2.

Thus

$$\mathcal{J}_1=\rho_1\int_{-\infty}^{\infty}dx\int_{\eta}^{\infty}\left[\frac{\partial\phi_1}{\partial t}+\frac{1}{2}\left(\frac{\partial\phi_1}{\partial x}\right)^2+\frac{1}{2}\left(\frac{\partial\phi_1}{\partial y}\right)^2+gy\right]dy, \tag{7.12}$$

$$\mathcal{J}_2=\rho_2\int_{-\infty}^{+\infty}dx\int_{-\infty}^{\eta}\left[\frac{\partial\phi_2}{\partial t}+\frac{1}{2}\left(\frac{\partial\phi_2}{\partial x}\right)^2+\frac{1}{2}\left(\frac{\partial\phi_2}{\partial y}\right)^2+gy\right]dy, \tag{7.13}$$

and

$$\mathcal{J}_a=\sigma\int_{-\infty}^{+\infty}dx\left[1+\left(\frac{\partial\eta}{\partial x}\right)^2\right]^{1/2}. \tag{7.14}$$

For the problem of the Kelvin–Helmoholtz stability, the primary flow state is as follows: The upper fluid (1) and the lower fluid (2) are divided by a horizontal interface ($y=0$). Both fluids are moving uniformly along the x direction with different velocities (U_1 and U_2). Thus we shall take the trial solution for ϕ_1 and ϕ_2 as

$$\phi_1=U_1x+C(x,t)\cos(S)\exp(-ky), \tag{7.15}$$

$$\phi_2=U_2x+B(x,t)\cos(S)\exp(ky), \tag{7.16}$$

where $k=\partial S/\partial x>0$. The forms of the trial solutions are suggested by the linear theory.[9]

In a more complete theory, indeed, the spatial and temporal variations of A, B, C, and S should be taken into account in order to study the modulation of the nonlinear waves. We shall here however only present the results that A, B, C, k, and $\omega=\partial S/\partial t$ are constants. Then after carrying out some straightforward calculations, we find up to the order of $O(A^4)$

$$\mathcal{J}_1=\rho_1\int_{-\infty}^{+\infty}dx\,\{\tfrac{1}{2}(\omega+kU_1)CA[1+\tfrac{1}{8}(kA)^2]+\frac{k}{4}C^2[1+(kA)^2]-\frac{g}{4}A^2\}, \tag{7.17}$$

$$\mathcal{J}_2=\rho_2\int_{-\infty}^{+\infty}dx\,\{-\tfrac{1}{2}(\omega+kU_2)BA[1+\tfrac{1}{8}(kA)^2]+\frac{k}{4}B^2[1+(kA)^2]+\frac{g}{4}A^2\}, \tag{7.18}$$

$$\mathcal{J}_a=\sigma\int_{-\infty}^{+\infty}dx\,[1+\tfrac{1}{4}(kA)^2-\tfrac{3}{64}(kA)^4]. \tag{7.19}$$

The variations with respect to A, B, and C thus lead to

$$\delta A:\ \rho_1\left(\tfrac{1}{2}(\omega+kU_1)C[1+\tfrac{3}{8}(kA)^2]+\frac{k^3}{2}AC^2-\frac{g}{2}A\right)$$
$$+\rho_2\left(-\tfrac{1}{2}(\omega+kU_2)B[1+\tfrac{3}{8}(kA)^2]+\frac{k^3}{2}AB^2+\frac{g}{2}A\right)$$
$$+\frac{\sigma k^2}{2}A[1-\tfrac{3}{8}(kA)^2]=0, \tag{7.20}$$

$$\delta B:\ -\tfrac{1}{2}(\omega+kU_2)A[1+\tfrac{1}{8}(kA)^2]+\frac{k}{2}B[1+(kA)^2]=0, \tag{7.21}$$

$$\delta C:\ \tfrac{1}{2}(\omega+kU_1)A[1+\tfrac{1}{8}(kA)^2]+\frac{k}{2}C[1+(kA)^2]=0. \tag{7.22}$$

Denote

$$G=(1+\tfrac{1}{8}k^2A^2)/(1+k^2A^2),$$

then we have

$$B=\frac{(\omega+kU_2)}{k}GA, \tag{7.23}$$

$$C=-\frac{(\omega+kU_1)}{k}GA. \tag{7.24}$$

Substituting them in (7.20), we have

$$[(\rho_1+\rho_2)\omega^2+2(\rho_1U_1+\rho_2U_2)k\omega+(\rho_1U_1^2+\rho_2U_2^2)k^2]$$
$$\times G[1+(\tfrac{3}{8}-G)k^2A^2]-[gk(\rho_2-\rho_1)+\sigma k^3(1-\tfrac{3}{8}k^2A^2)]=0.$$

Thus

$$\omega=\frac{1}{(\rho_1+\rho_2)}\{-(\rho_1U_1+\rho_2U_2)k\pm\sqrt{W(k,[U_1-U_2],A)}\}, \tag{7.25}$$

where

$$W(k,[U_1-U_2],A)=\frac{(\rho_1+\rho_2)[gk(\rho_2-\rho_1)+\sigma k^3(1-\tfrac{3}{8}k^2A^2)]}{G[1+(\tfrac{3}{8}-G)k^2A^2]}$$
$$-\rho_1\rho_2k^2(U_1-U_2)^2. \tag{7.26}$$

The system is stable when ω is real or when $W\geqslant 0$. As $A\to 0$, we recover the result of the linear theory,[9] and the system is stable for all k if

$$(U_1-U_2)^2\leqslant\frac{2(\rho_1+\rho_2)}{\rho_1\rho_2}[\sigma g(\rho_2-\rho_1)]^{1/2}\equiv U_c^2. \tag{7.27}$$

For $(U_1-U_2)^2>U_c^2$, there exists a certain range of k, for which the system is linearly unstable. However, as can be seen from the expression of W, the system can be stabilized by the nonlinear effects. The expression of W in (7.26) is valid only up to $O(A^2)$, thus the extent of stabilization that can be inferred from this formula is still somewhat limited. This restriction can be relaxed and a more detailed analysis can be carried out but we shall not pursue it further here.

It is of interest that if we also vary Eqs. (7.17)–(7.19) with respect to ω and k, then we obtain after using (7.21) and (7.22)

$$\delta\omega:\ \omega=-\frac{(\rho_1U_1+\rho_2U_2)k}{(\rho_1+\rho_2)}, \tag{7.28}$$

which by (7.25) implies

$$W=0. \tag{7.29}$$

Making use of (7.28) and (7.29), we obtain

$$\delta k:\ \frac{1}{2}\frac{\rho_1\rho_2(U_1-U_2)^2G}{(\rho_1+\rho_2)}[2(1+\tfrac{3}{8}k^2A^2)-G(1+3k^2A^2)]$$
$$-\sigma k(1-\tfrac{3}{8}k^2A^2)=0. \tag{7.30}$$

It is of interest to note that as $A\to 0$, Eq. (7.30) is exactly the same as

$$dW/dk=0. \tag{7.31}$$

Analogous to the discussions at the end of the Sec. 4, we can infer that for those Fourier components which are linearly unstable, the asymptotic amplitude will reach that given by $W=0$, and the dispersion relation of propagation is given by (7.28). If we attempt to represent the asymptotic solution by a single mode, the mode, i.e., the wavenumber k, determined by Eq. (7.30), which corresponds to maximum growth from the linear instability, would give the best representation.

The previous analysis can be extended to include harmonics and subharmonics. We shall limit our study to include second harmonics only. Thus the trial solutions adopted will be

$$\eta(x,t)=A\sin(S)+A_2\cos(2S), \tag{7.32}$$

$$\phi_1=U_1x+C\cos(S)\exp(-ky)+C_2\sin(2S)\exp(-2ky), \tag{7.33}$$

and

$$\phi_2=U_2x+B\cos(S)\exp(ky)+B_2\sin(2S)\exp(2ky). \tag{7.34}$$

Then after somewhat lengthy yet straightforward computations, we obtain, up to $O(A^4+A_2^4)$,

$$\mathcal{J}_1/\rho_1=\tfrac{1}{2}(\omega+kU_1)\{CA[1+\tfrac{1}{2}kA_2+(k^2/8)(A^2+2A_2^2)]$$
$$-2C_2[A_2+\tfrac{1}{2}kA^2+k^2A_2(A^2+\tfrac{1}{2}A_2^2)]\}$$
$$+(k/4)C^2[1+k^2(A^2+A_2^2)]+(k/2)C_2^2[1+4k^2(A^2+A_2^2)]$$
$$-kCC_2[kA+\tfrac{3}{2}k^2AA_2+\tfrac{9}{8}k^3A(A^2+2A_2^2)]-(g/4)(A^2+A_2^2). \tag{7.35}$$

Let us denote

$$\mathcal{J}_1=\Psi(\omega,k,\rho_1,U_1,g,A,A_2,C,C_2),$$

then

$$\mathcal{J}_2=\Psi(\omega,k,\rho_2,U_2,-g,-A,-A_2,B,B_2), \tag{7.36}$$

and

$$\mathcal{J}_a=\sigma[1+\tfrac{1}{4}(k^2A^2+4k^2A_2^2)$$
$$-\tfrac{3}{64}(k^4A^4+16k^4A^2A_2^2+16k^4A_2^4)]. \tag{7.37}$$

It is clear that when we vary the above equations with respect to A, B, and C, they will be the set of equations (7.20)–(7.22) with additional terms involving A_2, B_2, and C_2. A similar set of equations with the roles of $\{A,B,C\}$ and $\{A_2,B_2,C_2\}$ reversed will be obtained when we vary with the above with respect to A_2, B_2, and C_2. The situation is analogous to what we have discussed in Sec. 5. We shall not go into the detailed study of this nonlinear interaction. Let us limit ourselves to the case that $(kA)^2$ is small, and get all the correction terms up to $O(A^2)$. Then we can take $A_2, B_2, C_2=O(A^2)$. Thus the Euler equations become

$$\delta A:\ \rho_1\{\tfrac{1}{2}(\omega+kU_1)C[1+\tfrac{3}{8}(kA)^2+\tfrac{1}{2}(kA_2)]-(\omega+kU_1)kAC_2 + (k^3/2)AC^2-k^2CC_2-(g/2)A\}+\rho_2\{-\tfrac{1}{2}(\omega+kU_2) \times B[1+\tfrac{3}{8}(kA)^2-\tfrac{1}{2}(kA_2)]-(\omega+kU_2)kAB_2+(k^3/2)AB^2 + k^2BB_2+(g/2)A\}+(\sigma k^2/2)A[1-\tfrac{3}{8}(kA)^2]=0, \tag{7.38}$$

$$\delta B:\ -\tfrac{1}{2}(\omega+kU_2)A[1+\tfrac{1}{8}(kA)^2-\tfrac{1}{2}kA_2] + (k/2)B[1+(kA)^2]+k^2B_2A=0, \tag{7.39}$$

$$\delta C:\ \tfrac{1}{2}(\omega+kU_1)A[1+\tfrac{1}{8}(kA)^2+\tfrac{1}{2}kA_2] + (k/2)C[1+(kA)^2]-k^2C_2A=0, \tag{7.40}$$

$$\delta A_2:\ \rho_1[(\omega+kU_1)(\tfrac{1}{4}kCA-C_2)-(g/2)A_2] + \rho_2[(\omega+kU_2)(\tfrac{1}{4}kBA+B_2)+(g/2)A_2]+2\sigma k^2A_2=0, \tag{7.41}$$

$$\delta B_2:\ (\omega+kU_2)[A_2-\tfrac{1}{2}kA^2]+kB_2+k^2BA=0, \tag{7.42}$$

$$\delta C_2:\ -(\omega+kU_1)[A_2+\tfrac{1}{2}kA^2]+kC_2-k^2CA=0. \tag{7.43}$$

Using Eqs. (7.23) and (7.24), we thus obtain

$$A_2=\frac{A^2}{2[2\sigma k^2-g(\rho_2-\rho_1)]}[\rho_2(\omega+kU_2)^2-\rho_1(\omega+kU_1)^2], \tag{7.44}$$

$$B_2=-(\omega+kU_2)\left(\frac{A_2}{k}+\frac{1}{2}A^2\right), \tag{7.45}$$

and

$$C_2=(\omega+kU_1)\left(\frac{A_2}{k}-\frac{1}{2}A^2\right). \tag{7.46}$$

From (7.39) and (7.40), we have

$$B=\left(\frac{\omega+kU_2}{k}\right)A[1+\tfrac{1}{8}(kA)^2+\tfrac{3}{2}kA_2], \tag{7.47}$$

$$C=-\left(\frac{\omega+kU_1}{k}\right)A[1+\tfrac{1}{8}(kA)^2-\tfrac{3}{2}kA_2]. \tag{7.48}$$

Substituting (7.44)–(7.48) into (7.38), we obtain

$$(\rho_1+\rho_2)\omega^2+2(\rho_1U_1+\rho_2U_2)k\omega+(\rho_1U_1^2+\rho_2U_2^2)k^2 - gk(\rho_2-\rho_1)-\sigma k^3-\alpha(kA)^2=0, \tag{7.49}$$

where

$$\alpha=\tfrac{1}{2}gk(\rho_2-\rho_1)+\tfrac{1}{8}\sigma k^3+\frac{[\rho_2(\omega+kU_2)^2-\rho_1(\omega+kU_1)^2]^2}{2k[g(\rho_2-\rho_1)-2\sigma k^2]}. \tag{7.50}$$

Thus

$$\omega=\frac{1}{(\rho_1+\rho_2)}\{-(\rho_1U_1+\rho_2U_2)k \pm(W[k,(U_1-U_2),0]+\alpha(\rho_1+\rho_2)(kA)^2)^{1/2}\}. \tag{7.51}$$

In expression (7.50), the values of ω are taken to be those given by the linear relation, i.e., expression (7.51) with $A=0$. Therefore, the nonlinearity will inhibit the linear instability if $\alpha>0$, and enhance the instability if $\alpha<0$. When the linear instability is arrested by the nonlinearity, the asymptotic amplitudes will be given by

$$W[k,(U_1-U_2),0]+\alpha(\rho_1+\rho_2)(kA)^2=0$$

or

$$(\rho_1+\rho_2)[gk(\rho_2-\rho_1)+\sigma k^3] -\rho_1\rho_2k^2(U_1-U_2)^2+\alpha(\rho_1+\rho_2)(kA)^2=0. \tag{7.52}$$

Then ω is again given by Eq. (28), and we have

$$\alpha=\tfrac{1}{2}gk(\rho_2-\rho_1)+\tfrac{1}{8}\sigma k^3+\frac{k(\rho_2-\rho_1)^2[g(\rho_2-\rho_1)+\sigma k^2]^2}{2(\rho_1+\rho_2)^2[g(\rho_2-\rho_1)-2\sigma k^2]} \tag{7.53}$$

or

$$\alpha=-\tfrac{3}{8}\sigma k^3+\frac{1}{2}\frac{\rho_1\rho_2}{\rho_1+\rho_2}k^2(U_1-U_2)^2 +\frac{1}{2}\frac{(\rho_2-\rho_1)^2[\rho_1\rho_2k^2(U_1-U_2)^2]^2}{[gk(\rho_2-\rho_1)-2\sigma k^3](\rho_1+\rho_2)^4} \tag{7.54}$$

If we denote

$$\Gamma=-\rho_2gk\alpha,$$

then this is the same Γ as that given by Nayfih and Saric.[10] Thus when a perturbation expansion approach is adopted, the results of our analysis agree with those obtained by other established methods.

It may be remarked that when a perturbation expansion approach is adopted, and if instead of treating the amplitudes as constants, the amplitudes are considered as slowly varying functions of space and time, then a nonlinear Schrödinger equation can also be derived. For the problem of nonlinear water wave, i.e., when the upper fluid is absent, one version of derivation of the nonlinear Schrödinger equation using the variational method was given by Yen and Lake.[11]

8. DISCUSSION

We have presented a diverse set of problems to illustrate how the variational method developed earlier for nonlinear oscillation and wave problems can be extended to the study of nonlinear stability problems. If the conventional perturbation expansion approach is used, the variational method can lead to the same results as obtained by other more established methods. However, the variational method apparently can yield valuable information far away from the critical region, which the ordinary perturbation methods may not be able to deal with. The first example of Duffing stability is particularly illuminating on this point. It is also clear from the preceding study that we are far from fully understanding the whole story about the variational method as applied to the nonlinear stability problems. Some immediate questions raised from the previous study are, among others, the apparent nonuniqueness of permissible solutions when harmonics are included; the problem of successive harmonic generations; and the modulation of waves far away from the critical region. The nonlinear stability problems are difficult and challenging. The variational method presented above at least offered a new and hopeful perspective on this class of intriguing problems.

ACKNOWLEDGMENTS

This study is supported by NSF and ARPA. Computational assistance from Paul Hsieh, an eighth grader, is appreciated. Part of the work was carried out when the author was visiting the Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England, in 1975.

[1]D. Y. Hsieh, J. Math. Phys. **16**, 275—80 (1975).
[2]D. Y. Hsieh, J. Math. Phys. **16**, 1630—36 (1975).
[3]D. Y. Hsieh, J. Acoust. Soc. Am. **58**, 977—82 (1975).
[4]A. H. Nayfeh, *Perturbation Methods* (Wiley, New York, 1973).
[5]P. Sturrock, Phys. Rev. **112**, 1488—503 (1958).
[6]I. B. Bernstein and S. K. Trehan, Nucl. Fusion **1**, 3—41 (1960).
[7]P. C. Clemmow and J. P. Dougherty, *Electrodynamics of Particles and Plasmas* (Addison-Wesley, Reading, Massachusetts, 1969).
[8]D. Y. Hsieh, in *Finite-Amplitude Wave Effects in Fluids*, edited by L. Bjørnø (IPC Science and Technology Press, Surrey, England, 1974).
[9]S. Chanchasekhar, *Hydrodynamics and Hydromagnetic Stability* (Oxford U. P., New York, 1961).
[10]A. H. Nayfeh and W. S. Saric, J. Fluid Mech. **55**, 311—27 (1972).
[11]H. C. Yuen and B. M. Lake, Phys. Fluids **18**, 956—60 (1975).

Rederivation of the Jin-Martin lower bound

Seichi Naito

Department of Physics, Osaka City University, Sumiyoshiku, Osaka, Japan
(Received 28 June 1976)

With the help of the Phragmén–Lindelöf theorem, we can rederive the full Jin–Martin's results on the total cross section lower bound.

Jin and Martin,[1] using fairly involved Herglotz-function arguments, have obtained the lower bound on total cross sections. Their bound in a slightly weakened form has been rederived simply[2] by using the Phragmén–Lindelöf theorem.[3] In view of the importance of the lower bound, it will be interesting to show how we can derive full results of Jin and Martin. This problem is investigated by using the technique which is more complicated than Simon's technique.

For simplicity, we consider the spinless elastic scattering $A+B\to A+B$ (s channel) coupled by crossing to $\bar{A}+B\to\bar{A}+B$ (u channel), and the scattering amplitude is denoted by $F(s,t)$. Then it is easily shown that the full results[1] by Jin and Martin can be obtained from the fact that *there never occurs the case* when both

$$\lim_{s\to+\infty} s^2\,\mathrm{Re}[F(s,t)]=0 \tag{1}$$

and

$$\lim_{s\to+\infty} s^4\,\mathrm{Im}[F(s,t)]=0 \quad [\text{for } 0\le t<\min(4M_A^2,4M_B^2)] \tag{2}$$

hold simultaneously. Therefore, we shall hereafter prove the above fact.

The analyticity in s and the polynomial upper boundedness of $F(s,t)$ make it possible to write the dispersion relation with two subtractions[4]:

$$F(s,t)=A(t)+B(t)\omega+\frac{\omega^2}{\pi}\left(\int_{\rho+t/2}^{\infty}d\omega'\,\frac{\mathrm{Im}[F_{\mathrm{I}}(\omega'+\sigma-\tfrac{1}{2}t,t)]}{\omega'^2(\omega'-\omega)}+\int_{\rho+t/2}^{\infty}d\omega'\,\frac{\mathrm{Im}[F_{\mathrm{II}}(\omega'+\sigma-\tfrac{1}{2}t,t]}{\omega'^2(\omega'+\omega)}\right), \tag{3}$$

where the unitarity of the S matrix gives

$$\mathrm{Im}[F_{\mathrm{I}}(\omega'+\sigma-\tfrac{1}{2}t,t)]\ge 0 \tag{4}$$

and

$$\mathrm{Im}[F_{\mathrm{II}}(\omega'+\sigma-\tfrac{1}{2}t,t)]\ge 0 \quad [\text{for } 0\le t<\min(4M_A^2,4M_B^2)]. \tag{5}$$

In (3), we have used the following notations:

$$\omega\equiv s-\sigma+\tfrac{1}{2}t,$$

$$\sigma\equiv M_A^2+M_B^2,$$

and

$$\rho\equiv 2M_AM_B.$$

With the help of (2), Eq. (3) can be rewritten as

$$F(s,t)=[B(t)+d_1]\omega+[A(t)+d_2]+d_3/\omega+d_4/\omega^2+d_5(\omega)+d_6(\omega), \tag{6}$$

where

$$d_i\equiv-\frac{1}{\pi}\int_{\rho+t/2}^{\infty}d\omega'\,\omega'^{i-3}\,\mathrm{Im}[F_{\mathrm{I}}(\omega'+\sigma-\tfrac{1}{2}t,t)] \quad (i=1,\cdots,4),$$

$$d_5(\omega)\equiv\frac{\omega^2}{\pi}\int_{\rho+t/2}^{\infty}d\omega'\,\frac{\mathrm{Im}[F_{\mathrm{II}}(\omega'+\sigma-\tfrac{1}{2}t,t)]}{\omega'^2(\omega'+\omega)}, \tag{7}$$

and

$$d_6(\omega)\equiv\frac{1}{\omega^2}\,\frac{1}{\pi}\int_{\rho+t/2}^{\infty}d\omega'\,\frac{\omega'^2\,\mathrm{Im}[F_{\mathrm{I}}(\omega'+\sigma-\tfrac{1}{2}t,t)]}{\omega'-\omega}.$$

In the following discussion, it is essential that $d_6(\omega)$ defined by (7) has the upper bound

$$\lim_{\omega\to+\infty}|\omega^2 d_6(\omega)|\le\beta \quad (\beta \text{ is some constant}). \tag{8}$$

Inequality (8) can be derived from (1), (2), and the polynomial upper boundedness

$$|F(s,t)|\le|s|^N \text{ as } |s|\to\infty, \tag{9}$$

by using the same technique as in the Appendix of our previous paper.[5] Then (1), (2), (6), and (8) lead to

$$\lim_{\omega\to+\infty}\omega\{(B+d_1)\omega+(A+d_2)+d_3/\omega+d_5(\omega)\}=0. \tag{10}$$

With the help of (10), we find that

$$\delta_1\equiv(1/\pi)\int_{\rho+t/2}^{\infty}d\omega'\,\omega'^{-2}\,\mathrm{Im}[F_{\mathrm{II}}(\omega'+\sigma-\tfrac{1}{2}t,t)] \tag{11}$$

should be finite; if $\delta_1=+\infty$, we obtain

$$\lim_{\omega\to+\infty}\omega^{-1}d_5(\omega)=+\infty,$$

which contradicts (10). QED Then we find

$$d_5(\omega)=\delta_1\omega-\frac{\omega}{\pi}\int_{\rho+t/2}^{\infty}d\omega'\,\frac{\mathrm{Im}[F_{\mathrm{II}}(\omega'+\sigma-\tfrac{1}{2}t,t)]}{\omega'(\omega'+\omega)}.$$

and

$$\delta_1=-B-d_1 \quad [\text{from } (10)]. \tag{12}$$

By carrying out the above procedure successively, (6) is finally reduced to

$$F(s,t)=d_4/\omega^2+\tilde{d}_5(\omega)+d_6(\omega), \tag{13}$$

with

$$\tilde{d}_5(\omega)\equiv-\frac{1}{\omega}\,\frac{1}{\pi}\int_{\rho+t/2}^{\infty}d\omega'\,\frac{\omega'\,\mathrm{Im}[F_{\mathrm{II}}(\omega'+\sigma-\tfrac{1}{2}t,t)]}{\omega'+\omega}. \tag{14}$$

As in the case of δ_1, the assumption $\delta_5=-\infty$, where

$$\delta_5 \equiv -\frac{1}{\pi}\int_{\rho+t/2}^{\infty} d\omega'\,\omega'\,\mathrm{Im}[F_{II}(\omega'+\sigma-\tfrac{1}{2}t,t)], \tag{15}$$

is easily shown to lead to the contradiction with (1), (2), (8), and (13), so that δ_5 should be finite. Then the integral for $\omega^2\tilde{d}_5(\omega)$ is uniformly convergent[6] on account of (5), so that

$$\lim_{\omega\to+\infty} \omega^2\tilde{d}_5(\omega) = \delta_5. \tag{16}$$

After all, (1), (2), (13), and (16) give

$$\lim_{\omega\to+\infty} \omega^2 d_6(\omega) = -d_4 - \delta_5. \tag{17}$$

On the other hand, the integral (7) for $\omega^2 d_6(\omega)$ gives

$$\lim_{|\omega|\to\infty,\ \theta=\pi/2} \omega^2 d_6(\omega) = 0, \tag{18}$$

with

$$\omega = |\omega| \exp(i\theta).$$

In (18), we have used the fact that the integral is uniformly convergent[6] on account of (2) and (4). Since $\omega^2 d_6(\omega)$ is analytic and polynomially upper bounded in the region $R=\{\omega \mid |\omega| \geqslant 1,\ 0\leqslant \theta \leqslant \pi/2\}$ [as easily seen from (9) and (13)], we can apply the Phragmén–Lindelöf theorem[3] to $\omega^2 d_6(\omega)$, and consequently we obtain from (17) and (18)

$$d_4 + \delta_5 = 0. \tag{19}$$

However, (4), (5), (7), (15), and (19) lead to the physically unrealizable condition

$$\mathrm{Im}[F_I(\omega'+\sigma-\tfrac{1}{2}t,t)] \equiv \mathrm{Im}[F_I(\omega'+\sigma-\tfrac{1}{2}t,t)] \equiv 0. \tag{20}$$

Thus we have proved that *it never occurs that* (1) *and* (2) *hold simulaneously*.

[1] Y. S. Jin and A. Martin, Phys. Rev. **135**, B1369 (1964), Ref. 11 of T. Kinoshita, Phys. Rev. **154**, 1438 (1967), M. Sugawara, Phys. Rev. Lett. **14**, 336 (1965).
[2] B. Simon, Phys. Rev. D **1**, 1240 (1970).
[3] E. C. Titchmarsh, *The Theory of Functions* (Oxford U.P., New York, 1939), 2nd ed., p. 177 (5.61) and p. 179 (5.64).
[4] Y. S. Jin and A. Martin, Phys. Rev. **135**, B1375 (1964).
[5] S. Naito, Phys. Rev. D 13, 2884 (1976).
[6] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Cambridge U.P., New York, 1952), p. 70.

Dilations and interaction

Detlev Buchholz and Klaus Fredenhagen

II. Institut für Theoretische Physik der Universität Hamburg, Germany
(Received 30 August 1976)

As a consequence of the geometrical features of dilations massless particles do not interact in a local, dilationally invariant quantum theory. This result also holds in models in which dilations are only a symmetry of the S matrix.

1. INTRODUCTION AND MAIN RESULTS

The conventional argument showing that massless particles do not interact in a local, dilationally invariant quantum theory is in the simplest case the following one (see, e.g., Ref. 1): suppose ϕ is a scalar Wightman field transforming under dilations according to

$$D(\lambda)\phi(\lambda^{-1}x)D(\lambda)^{-1} = \lambda^d \cdot \phi(x). \tag{1}$$

If ϕ has a nonvanishing matrix element between the vacuum and a massless one-particle state, d can only be one. Then ϕ has canonical dimension and this implies that it is a free field. This reasoning is quite correct. However, since the argument depends upon the existence of a field ϕ with the special properties mentioned above, the conclusion appears to us to be rather premature. First, there is no physical reason to rule out *ab initio* all models in which the basic fields do not transform like a finite-dimensional representation under dilations. And secondly, even if the fields transform in this way, it could happen that they do not interpolate between the vacuum and the massless one-particle states. In general one should only expect that suitable polynomials in the fields have this property. It is the aim of the present note to close these apparent loopholes. Using only the geometrical features of dilations and the basic properties of local field theory, we give a fairly general argument confirming the above no-go theorem.

The setting used for the analysis may be sketched as follows: We deal with an irreducible field algebra $\mathfrak{F}$ of bounded operators acting on a Hilbert space $\mathcal{H}$. $\mathfrak{F}$ is generated by a net $\mathcal{O} \to \mathfrak{F}(\mathcal{O})$ of local algebras attached to the regions $\mathcal{O}$ of Minkowski space. We may forego here a formal specification of the usual structural assumptions on the theory like locality, covariance, spectrum condition and uniqueness of the vacuum. (For a detailed discussion see, for example, Ref. 2). In addition to these familiar properties we require that there be a continuous, unitary representation $\lambda \to D(\lambda)$ of the multiplicative group of the positive reals in $\mathcal{H}$. The operators $D(\lambda)$, the dilations, satisfy

$$D(\lambda)U(x) = U(\lambda x)D(\lambda) \quad \text{and} \quad D(\lambda)U(\Lambda) = U(\Lambda)D(\lambda), \tag{2a}$$

where $x = (x_0, \mathbf{x}) \to U(x)$ are the translations and $\Lambda \to U(\Lambda)$ the Lorentz transformations. Moreover, the dilations $D(\lambda)$ induce automorphisms of the field algebra $\mathfrak{F}$ with appropriate geometric properties:

$$D(\lambda)\mathfrak{F}(\mathcal{O})D(\lambda)^{-1} = \mathfrak{F}(\lambda \cdot \mathcal{O}). \tag{2b}$$

These rather general assumptions suffice to prove the following statement:

If there exist massless particles in the model, (i.e., a family of subspaces $\mathcal{H}_1^{(i)} \subset \mathcal{H}$ on which the unitaries $U(x)$, $U(\Lambda)$ act like an irreducible representation of the Poincaré group with mass zero and helicity s_i), then the S matrix for these particles is trivial.

Our interest in this problem arose in discussions with Haag on supersymmetric field theories. In a recent article Haag, Lopuszanski, and Sohnius have analyzed the structure of all possible supersymmetries of the S matrix.[3] They found out that in a pure S-matrix formalism there is essentially only one way of a complete fusion between internal and geometrical symmetries, including dilations. Since such a structure looks very promising from the point of view of physics, one may ask whether it can be embedded into a conventional field theoretical setting. As a consequence of our analysis the answer to this question is negative: If the theory is to describe collisions of massless particles and if dilations are to be a proper, unbroken symmetry, one has to abandon some of the usual field theoretical assumptions. At present it is unclear how the assumptions have to be modified and we refrain from speculations. However, we want to emphasize that even in a modified scheme the local observables (the currents, etc.) should have a structure similar to that of $\mathfrak{F}$ given above. What may then be learned from our analysis is that massless particles in the vacuum sector of the observable algebra do not interact. It is therefore unlikely that particles like the photon and the η'-meson (both of which carry the charge quantum numbers of the vacuum) can be incorporated into such a scheme. This apparently restricts the possible range of application of these models to weak interaction physics.

2. THE PROOF

The central idea of the proof is very simple: we derive an asymptotic expansion for the function $\lambda \to D(\lambda)AD(\lambda)^{-1}$ at $\lambda = 0$, where A is a suitable local operator taken from $\mathfrak{F}$. It turns out that

$$D(\lambda)AD(\lambda)^{-1} = (\Omega, A\Omega) \cdot 1 + \lambda \cdot \phi + o(\lambda), \tag{3}$$

where this expansion is understood in the sense of operator valued distributions; Ω denotes the vector representing the vacuum and ϕ is some local field. Now the crucial point is that if ϕ is not zero, it creates a massless particle from the vacuum. It then follows from

Huyghens' principle (i.e., the timelike commutation relations between local and asymptotic fields given in Ref. 4) that the S matrix of this particle can only be trivial.

Unfortunately, there are models in which, for kinematical reasons, all local operators A give rise to a vanishing ϕ. However, this defect can be cured by a slight modification of the above expansion: dilating and boosting the operator A simultaneously, one arrives at an expression similar to (3), but with a nontrivial ϕ. To abbreviate the argument, we confine our attention to models involving only one type of massless particles with helicity $s=0$. But we shall give a brief outline of how to proceed in more complicated situations.

Now let A be any operator from $\mathfrak{F}$ which is localized in a bounded region $\mathcal{O}\subset\mathbb{R}^4$. We regularize A according to

$$A_\varphi=\int dt\,\varphi(t)\,U(t)AU(t)^{-1}, \tag{4}$$

where $t\to U(t)$ are the time translations. $\varphi(t)$ is a test function with compact support which has a Fourier transform $\tilde{\varphi}(\omega)$ with a twofold zero at $\omega=0$. The smoothed operator A_φ is still local and we get the following bound on its two-point function:

Lemma 1: Let $\Delta\to E(\Delta)$ be the spectral projections of the mass operator $M=(P^2)^{1/2}$ where $\Delta\subseteq\mathbb{R}^+$ is any Borel set of mass values. Then

$$|(A_\varphi\Omega,\,E(\Delta)U(\mathbf{x})A_\varphi\Omega)|\le c\cdot(1+|\mathbf{x}|^4)^{-1}$$
$$\cdot\{\|E(\Delta)A\Omega\|^2+\|E(\Delta)A^*\Omega\|^2\}$$

where the constant c depends neither on $\mathbf{x}$ nor on Δ.

Proof: Using the methods of the Jost–Lehmann–Dyson representation, one can show that the function

$$h_\Delta(x)=(A\Omega,\,E(\Delta)U(x)A\Omega)-(A^*\Omega,\,E(\Delta)U(-x)A^*\Omega)$$

vanishes in the spacelike complement of some bounded region $\mathcal{O}_1$ which depends only on the localization region $\mathcal{O}$ of A (see, e.g., Ref. 5, Lemma 6.2). Now, if one puts $\psi(t)=\int ds\,\bar{\varphi}(s)\varphi(s+t)$, one gets, owing to the spectrum condition,

$$\begin{aligned}(A_\varphi\Omega,\,&E(\Delta)U(\mathbf{x})A_\varphi\Omega)\\&=\int dt\,\psi(t)(A\Omega,\,E(\Delta)U(t,\mathbf{x})A\Omega)\\&=\int dt\,\psi^+(t)(A\Omega,\,E(\Delta)U(t,\mathbf{x})A\Omega)\\&=\int dt\psi^+(t)\{(A\Omega,\,E(\Delta)U(t,\mathbf{x})A\Omega)\\&\quad-(A^*\Omega,\,E(\Delta)U(-t,-\mathbf{x})A^*\Omega)\}\\&=\int dt\,\psi^+(t)h_\Delta(t,\mathbf{x}),\end{aligned}$$

where

$$\begin{aligned}\psi^+(t)&=(2\pi)^{-1/2}\int_0^\infty d\omega\tilde{\psi}(\omega)\exp(-i\omega t)\\&=\int_0^\infty d\omega\,|\tilde{\varphi}(\omega)|^2\exp(-i\omega t).\end{aligned}$$

Since $|\tilde{\varphi}(\omega)|^2$ is a test function with a fourfold zero at $\omega=0$, it is easy to verify that $\psi^+(t)$ is continuous and $|\psi^+(t)|\le c\cdot(1+|t|^5)^{-1}$. Taking the support properties of $h_\Delta(x)$ into account, one arrives at

$$\begin{aligned}|(A_\varphi\Omega,\,&E(\Delta)U(\mathbf{x})A_\varphi\Omega)|\\&\le\int_{|t|\ge|\mathbf{x}|-R}dt\,|\psi^+(t)|\cdot|h_\Delta(t,\mathbf{x})|\\&\le c\cdot\int_{|t|\ge|\mathbf{x}|-R}dt(1+|t|^5)^{-1}\cdot\{\|E(\Delta)A\Omega\|^2+\|E(\Delta)A^*\Omega\|^2\},\end{aligned}$$

where R is some length which depends only on A. From this inequality the statement of the lemma follows at once. ∎

We take now the operator A_φ and carry out the following manipulations: First we dilate it, then we boost it in a fixed direction, and finally we smear it in the two remaining spatial directions. For the boosts we take those in the x_1 direction:

$$K_\lambda=\begin{pmatrix}\frac{1}{2}(\lambda+\lambda^{-1}) & \frac{1}{2}(\lambda-\lambda^{-1}) & & \\ \frac{1}{2}(\lambda-\lambda^{-1}) & \frac{1}{2}(\lambda+\lambda^{-1}) & & \\ & & 1 & \\ & & & 1\end{pmatrix},\quad\lambda>0. \tag{5}$$

Then if $\mathbf{x}^\perp=(0,x_2,x_3)$ denotes the projection of $\mathbf{x}$ onto the (x_2,x_3)-plane and if $d^2x^\perp=dx_2\,dx_3$, we set

$$\begin{aligned}B_\lambda=\lambda^{-1}\cdot\int d^2x^\perp f(\mathbf{x}^\perp)U(\mathbf{x}^\perp)U(K_\lambda)D(\lambda)A_\varphi\\\times D(\lambda)^{-1}U(K_\lambda)^{-1}U(\mathbf{x}^\perp)^{-1},\end{aligned} \tag{6}$$

where $f(\mathbf{x}^\perp)$ is any test function with compact support. To begin with, we examine the localization properties of B_λ: Since A_φ is localized in some bounded region $\mathcal{O}$, it follows from (6) that B_λ is localized in $\{\lambda\cdot K_\lambda\mathcal{O}+\operatorname{supp}f\}$. Now $\lim_{\lambda\to0}\lambda\cdot K_\lambda=P$, where P is the projection onto the ray $(a,-a,0,0)$, $a\in\mathbb{R}$. Therefore, the operators B_λ are, for sufficiently small λ, localized in a fixed bounded region $\mathcal{O}_1$. The next step is to show that the sequence $B_\lambda\Omega$ converges to a (possibly zero) one-particle state in the limit of small λ:

Proposition 2: Let B_λ be the operator defined in relation (6). Then the weak limit $\text{w-}\lim_{\lambda\to0}B_\lambda\Omega$ exists and is an element of $\mathcal{H}_1$.

Proof: The proof of this assertion is based on Lemma 1. Since $U(K_\lambda)$ commutes with $U(\mathbf{x}^\perp)$ and $E(\Delta)$, we may write

$$\begin{aligned}\|E(\Delta)B_\lambda\Omega\|^2=\lambda^{-2}\cdot\int d^2x^\perp\int d^2y^\perp\bar{f}(\mathbf{x}^\perp)f(\mathbf{y}^\perp)\\\times(A_\varphi\Omega,\,E(\lambda\Delta)U(\lambda^{-1}[\mathbf{y}^\perp-\mathbf{x}^\perp])A_\varphi\Omega),\end{aligned}$$

where we have made use of relation (2a). If we set $g(\mathbf{x}^\perp)=\int d^2y^\perp\bar{f}(\mathbf{y}^\perp)f(\mathbf{x}^\perp+\mathbf{y}^\perp)$, we get, using Lemma 1,

$$\begin{aligned}\|E(\Delta)&B_\lambda\Omega\|^2\\&=\lambda^{-2}\cdot\int d^2x^\perp g(\mathbf{x}^\perp)(A_\varphi\Omega,\,E(\lambda\Delta)U(\lambda^{-1}\mathbf{x}^\perp)A_\varphi\Omega)\\&\le\sup_{\mathbf{y}^\perp}|g(\mathbf{y}^\perp)|\cdot c\int d^2x^\perp(1+|\mathbf{x}^\perp|^4)^{-1}\\&\quad\cdot\{\|E(\lambda\Delta)A\Omega\|^2+\|E(\lambda\Delta)A^*\Omega\|^2\}.\end{aligned}$$

Putting $\Delta=\mathbb{R}^+$, it follows that the sequence $B_\lambda\Omega$ is uniformly bounded in λ. Putting $\Delta_c=[a,b]$, where $0<a\le b<\infty$, it follows that $\lim_{\lambda\to0}\|E(\Delta_c)B_\lambda\Omega\|=0$ because the continuity properties of the spectral resolution imply $\lim_{\lambda\to 0}\|E(\lambda\Delta_c)\Phi\| = 0$ for every vector $\Phi \in \mathcal{H}$. Thus the sequence $B_\lambda\Omega$ converges weakly to zero on the orthogonal complement of the one-particle space $\mathcal{H}_1$. It remains to establish its convergence on $\mathcal{H}_1$. Since we are dealing with only one type of massless particles with helicity $s = 0$, we may identify the one-particle states $\Psi \in \mathcal{H}_1$ with their momentum space wavefunctions $\Psi(\mathbf{p})$ in $L^2(\mathbf{R}^3, d^3p/2|\mathbf{p}|)$. The dilations and Poincaré transformations act on these functions as follows:

$$(D(\lambda)\Psi)(\mathbf{p}) = \lambda \cdot \Psi(\lambda\mathbf{p}),$$
$$(U(t,\mathbf{x})\Psi)(\mathbf{p}) = \exp[i(t|\mathbf{p}| - \mathbf{xp})]\Psi(\mathbf{p}), \tag{7a}$$

and

$$(U(\Lambda)\Psi)(\mathbf{p}) = \Psi(\Lambda^{-1}\circ\mathbf{p}), \tag{7b}$$

where $\Lambda^{-1}\circ\mathbf{p}$ denotes the spatial components of the 4-vector $\Lambda^{-1}(|\mathbf{p}|,\mathbf{p})$. What is crucial now is that the wavefunction $(A_\varphi\Omega)(\mathbf{p})$ of the one-particle state $E(\{0\})A_\varphi\Omega$ is continuous in $\mathbf{p}$ if A_φ is the operator defined in relation (4). To verify this, we fix a set $\mathcal{L}$ of Lorentz transformations Λ which are close to the identity I, e.g., $\mathcal{L} = \{\Lambda : \|\Lambda - I\| \le \frac{1}{2}\}$. Since A is local it is obvious that all operators $(U(\Lambda)^{-1}AU(\Lambda) - A)$, $\Lambda \in \mathcal{L}$, are localized in a bounded region $\mathcal{O}$ of configuration space. Therefore, we get the estimate, using relation (7) and Lemma 1,

$$\begin{aligned}(\pi/|\mathbf{p}|)|\tilde{\varphi}(|\mathbf{p}|)|^2 \cdot |(A\Omega)(\Lambda\circ\mathbf{p}) - (A\Omega)(\mathbf{p})|^2 &= (1/2|\mathbf{p}|)\cdot|((U(\Lambda)^{-1}AU(\Lambda) - A)_\varphi\Omega)(\mathbf{p})|^2\\ &= (2\pi)^{-3}\cdot\int d^3x \exp(i\mathbf{xp})\cdot((U(\Lambda)^{-1}AU(\Lambda)-A)_\varphi\Omega,\\ &\qquad E(\{0\})U(\mathbf{x})(U(\Lambda)^{-1}AU(\Lambda)-A)_\varphi\Omega)\\ &\le c\cdot\{\|(U(\Lambda)-1)E(\{0\})A\Omega\|^2 + \|(U(\Lambda)-1)E(\{0\})A^*\Omega\|^2\},\end{aligned}$$

and this inequality holds for all $\Lambda \in \mathcal{L}$ and $\mathbf{p} \in \mathbf{R}^3$. Since we may take for $\tilde{\varphi}$ a test function which has a zero only at the origin, it is evident that $\lim_{\Lambda\to I}(A\Omega)(\Lambda\circ\mathbf{p}) = (A\Omega)(\mathbf{p})$ for $\mathbf{p} \ne 0$. But this shows that $(A\Omega)(\mathbf{p})$ and therefore also $(A_\varphi\Omega)(\mathbf{p}) = (2\pi)^{1/2}\tilde{\varphi}(|\mathbf{p}|)(A\Omega)(\mathbf{p})$ are continuous at $\mathbf{p} \ne 0$ because for every sequence $\mathbf{p}_n$ converging to $\mathbf{p}$ we can specify a sequence of Lorentz transformations Λ_n such that, for sufficiently large n, $\Lambda_n\circ\mathbf{p} = \mathbf{p}_n$ and $\lim_n\Lambda_n = I$. In order to establish the continuity of $(A_\varphi\Omega)(\mathbf{p})$ at $\mathbf{p} = 0$, we estimate

$$\begin{aligned}(1/2|\mathbf{p}|)|(A_\varphi\Omega)(\mathbf{p})|^2 &= (2\pi)^{-3}\int d^3x \exp(i\mathbf{xp})\cdot(A_\varphi\Omega, E(\{0\})U(\mathbf{x})A_\varphi\Omega) \le c.\end{aligned}$$

This bound holds uniformly for all $\mathbf{p} \in \mathbf{R}^3$ and implies $\lim_{\mathbf{p}\to 0}(A_\varphi\Omega)(\mathbf{p}) = 0$. Now we are almost finished: Using relation (7), we get for the wavefunction $(B_\lambda\Omega)(\mathbf{p})$ of the one-particle state $E(\{0\})B_\lambda\Omega$

$$(B_\lambda\Omega)(\mathbf{p}) = 2\pi\cdot\tilde{f}(\mathbf{p}^\perp)\cdot(A_\varphi\Omega)(\lambda\cdot K_\lambda^{-1}\circ\mathbf{p}),$$

where

$$\tilde{f}(\mathbf{p}^\perp) = (2\pi)^{-1}\int d^2x^\perp \exp(-i\mathbf{x}^\perp\mathbf{p}^\perp)f(\mathbf{x}^\perp) \quad \text{with } \mathbf{p}^\perp = (0, p_2, p_3).$$

An easy calculation shows that $\lim_{\lambda\to 0}\lambda\cdot K_\lambda^{-1}\circ\mathbf{p} = \frac{1}{2}(|\mathbf{p}| + p_1)\mathbf{e}_1$, where $\mathbf{e}_1 = (1,0,0)$. Taking into account the continuity of $\mathbf{p} \to (A_\varphi\Omega)(\mathbf{p})$, we get

$$\lim_{\lambda\to 0}(B_\lambda\Omega)(\mathbf{p}) = 2\pi\cdot\tilde{f}(\mathbf{p}^\perp)\cdot(A_\varphi\Omega)(\tfrac{1}{2}[|\mathbf{p}| + p_1]\mathbf{e}_1). \tag{8}$$

It then follows from the bounded convergence theorem that the limit $\lim_{\lambda\to 0}\int(d^3p/2|\mathbf{p}|)\Psi(\mathbf{p})(B_\lambda\Omega)(\mathbf{p})$ exists for all test functions $\Psi(\mathbf{p})$ with compact support. These functions are dense in $L^2(\mathbf{R}^3, d^3p/2|\mathbf{p}|)$ and since the vectors $B_\lambda\Omega$ are uniformly bounded in λ we conclude that the weak limit $w\text{-}\lim_{\lambda\to 0}E(\{0\})B_\lambda\Omega$ exists. This finishes the proof of the statement. ∎

Remark: Using the above proposition and the localization properties of the operators B_λ, one can show that $\lim_{\lambda\to 0}B_\lambda$ also exists on a dense set of vectors in $\mathcal{H}$.

The wavefunction of the one-particle state $\text{w-}\lim_{\lambda\to 0} B_\lambda\Omega$ is given by the right hand side of equation (8). It is therefore easy to specify a local operator A for which this vector is nontrivial: Pick, for example, a one-particle state $\Phi \in \mathcal{H}_1$ which is invariant under spatial rotations $R \to U(R)$. Since $\mathfrak{F}$ is irreducible, there exists a local operator $A_1 \in \mathfrak{F}$ such that the matrix element $(\Phi, A_1\Omega)$ is not zero. The operator $A = \int d\mu(R)U(R) \times A_1U(R)^{-1}$, where $d\mu(R)$ is the Haar measure on the group of rotations, then has the desired property. If one takes A_1 Hermitian and the functions φ, f real, one can even arrange for the approximating operators B_λ to be Hermitian.

In the remainder of this section we shall show that the existence of an operator sequence B_λ with properties mentioned above implies that the massless particles do not scatter. The argument is based on results recently derived in Ref 4 in the context of collision theory for massless particles. We recapitulate the main facts briefly: As in the massive case, there are collision states

$$\Phi_1 \overset{\text{in}}{\times}\cdots\overset{\text{in}}{\times}\Phi_n \text{ and } \Phi_1 \overset{\text{out}}{\times}\cdots\overset{\text{out}}{\times}\Phi_n$$

in $\mathcal{H}$ corresponding to incoming and outgoing configurations $\Phi_1, \ldots, \Phi_n \in \mathcal{H}_1$ of massless particles. These vectors have the familiar Fock structure known from a free theory. They can be generated from the vacuum Ω with the aid of asymptotic fields A^{in} and A^{out}. The bounded functions of the fields which are localized in a region $\mathcal{O}$ constitute the local asymptotic field algebras $\mathfrak{F}^{\text{in}}(\mathcal{O})$ and $\mathfrak{F}^{\text{out}}(\mathcal{O})$ respectively. They have commutation relations with the basic fields which may be interpreted as the field theoretical version of Huyghens' principle: If $\mathcal{O}$ is any bounded region and if $\mathcal{O}_+$, $\mathcal{O}_-$ are two regions which have a positive and negative timelike distance from $\mathcal{O}$, then

$$[F, F_+^{\text{in}}] = 0 \quad\text{and}\quad [F, F_-^{\text{out}}] = 0 \tag{9}$$

for arbitrary $F \in \mathfrak{F}(\mathcal{O})$, $F_+^{\text{in}} \in \mathfrak{F}^{\text{in}}(\mathcal{O}_+)$, and $F_-^{\text{out}} \in \mathfrak{F}^{\text{out}}(\mathcal{O}_-)$ This relation is the key to the proof of the following statement.

Proposition 3: If there exists a bounded region $\mathcal{O} \subset \mathbb{R}^4$ and a sequence of Hermitian operators $B_\lambda \in \mathfrak{F}(\mathcal{O})$ which converges weakly on the vacuum to some nonzero vector in $\mathcal{H}_1$, then the collision states

$$\Phi_1 \overset{\text{in}}{\times}\cdots\overset{\text{in}}{\times}\Phi_n \text{ and } \Phi_1 \overset{\text{out}}{\times}\cdots\overset{\text{out}}{\times}\Phi_n$$

coincide for arbitrary configurations $\Phi_1, \ldots, \Phi_n \in \mathcal{H}_1$. Consequently, the S matrix is trivial.

Proof: We define $B_\lambda^L = U(L)B_\lambda U(L)^{-1}$, where $L=(\Lambda, x)$ is an arbitrary Poincaré transformation and $U(L) = U(x)U(\Lambda)$ is the corresponding unitary in $\mathcal{H}$. Since B_λ converges weakly on the vacuum to some nontrivial one-particle state $\Phi \in \mathcal{H}_1$ we get $w\text{-}\lim_{\lambda\to 0} B_\lambda^L \Omega = U(L)\Phi = \Phi_L$. These vectors form a total set in $\mathcal{H}_1$ because the Poincaré transformations are irreducibly represented in $\mathcal{H}_1$. Now the operators B_λ^L are localized in the region $L\mathcal{O}$. Using relation (9), we get therefore

$$(F_+^{\text{in}}\Phi_L, F_-^{\text{out}}\Omega) = \lim_{\lambda\to 0}(F_+^{\text{in}}B_\lambda^L\Omega, F_-^{\text{out}}\Omega) = \lim_{\lambda\to 0}(F_+^{\text{in}}\Omega, F_-^{\text{out}}B_\lambda^L\Omega) = (F_+^{\text{in}}\Omega, F_-^{\text{out}}\Phi_L), \quad (10)$$

provided $F_+^{\text{in}} \in \mathfrak{F}^{\text{in}}(L\mathcal{O}_+)$ and $F_-^{\text{out}} \in \mathfrak{F}^{\text{out}}(L\mathcal{O}_-)$. Since the operators A^{in} and A^{out} are free fields, it is straightforward to verify that the bounded operators F_+^{in} and F_-^{out} in this relation may be replaced by products of smeared field operators $A_1^{\text{in}}, \ldots, A_m^{\text{in}}$ and $A_{m+1}^{\text{out}}, \ldots, A_n^{\text{out}}$, which are localized in $L\mathcal{O}_+$ and $L\mathcal{O}_-$ respectively. Thus we arrive at

$$(A_1^{\text{in}}\cdots A_m^{\text{in}}\Phi_L, A_{m+1}^{\text{out}}\cdots A_n^{\text{out}}\Omega) = (A_1^{\text{in}}\cdots A_m^{\text{in}}\Omega, A_{m+1}^{\text{out}}\cdots A_n^{\text{out}}\Phi_L). \quad (11)$$

Now we can prove the proposition by induction. For a one-particle state there is nothing to show, so let us assume that

$$\Phi_1 \overset{\text{in}}{\times} \cdots \overset{\text{in}}{\times} \Phi_m = \Phi_1 \overset{\text{out}}{\times} \cdots \overset{\text{out}}{\times} \Phi_m$$

for arbitrary configurations $\Phi_1, \ldots, \Phi_m \in \mathcal{H}_1$. This implies in particular that $A_1^{\text{in}}\cdots A_m^{\text{in}}\Omega = A_1^{\text{out}}\cdots A_m^{\text{out}}\Omega$ and, using relation (11), we get

$$\begin{aligned}(A_1^{\text{in}}\cdots A_m^{\text{in}}\Phi_L, A_{m+1}^{\text{out}}\cdots A_n^{\text{out}}\Omega) &= (A_1^{\text{in}}\cdots A_m^{\text{in}}\Omega, A_{m+1}^{\text{out}}\cdots A_n^{\text{out}}\Phi_L)\\ &= (A_1^{\text{out}}\cdots A_m^{\text{out}}\Omega, A_{m+1}^{\text{out}}\cdots A_n^{\text{out}}\Phi_L)\\ &= (A_1^{\text{out}}\cdots A_m^{\text{out}}\Phi_L, A_{m+1}^{\text{out}}\cdots A_n^{\text{out}}\Omega),\end{aligned}$$

where the last equality sign follows from an explicit calculation of the scalar products. If we set $\Phi_1 = A_1^{\text{in}}\Omega, \ldots, \Phi_n = A_n^{\text{out}}\Omega$, we can reexpress this equation in terms of the collision states,

$$(\Phi_1 \overset{\text{in}}{\times} \cdots \overset{\text{in}}{\times} \Phi_m \overset{\text{in}}{\times} \Phi_L, \Phi_{m+1} \overset{\text{out}}{\times} \cdots \overset{\text{out}}{\times} \Phi_n) = (\Phi_1 \overset{\text{out}}{\times} \cdots \overset{\text{out}}{\times} \Phi_m \overset{\text{out}}{\times} \Phi_L, \Phi_{m+1} \overset{\text{out}}{\times} \cdots \overset{\text{out}}{\times} \Phi_n),$$

provided $A_1^{\text{in}}, \ldots, A_n^{\text{out}}$ are operators with the special localization properties mentioned above. However, keeping in mind that the vectors Φ_L form a total set in $\mathcal{H}_1$, one can extend this equation by continuity to arbitrary configurations $\Phi_1, \ldots, \Phi_n, \Phi_L \in \mathcal{H}_1$[4] and it is then obvious that

$$\Phi_1 \overset{\text{in}}{\times} \cdots \overset{\text{in}}{\times} \Phi_{m+1} = \Phi_1 \overset{\text{out}}{\times} \cdots \overset{\text{out}}{\times} \Phi_{m+1}. \qquad \blacksquare$$

Combining the two propositions it follows that the massless particles in $\mathcal{H}_1$ do not interact if the dilations are a true symmetry. We have established this result only for one type of massless particles with helicity $s=0$. In the presence of a family of one-particle spaces $\mathcal{H}_1^{(i)} \subset \mathcal{H}$ on which the unitaries $U(x)$, $U(\Lambda)$ act like an irreducible representation of the Poincaré group with mass zero and helicity s_i, the main modifications are in the second part of the proof of Proposition 2: For vectors $\Psi \in \mathcal{H}_1^{(k)}$ relation (7b) changes according to

$$(U(\Lambda)\Psi)_k(\mathbf{p}) = \exp[is_k\alpha(\Lambda, \mathbf{p})](\Psi)_k(\Lambda^{-1}\circ\mathbf{p}),$$

where the index k refers to the space $\mathcal{H}_1^{(k)}$. The functions $\alpha(\Lambda, \mathbf{p})$ are the Wigner phases.[6] They are not completely fixed by the structural relations imposed by the Lorentz group. As a matter of fact we may choose a convention such that the functions $\alpha(\Lambda, \mathbf{p})$ are simultaneously continuous in Λ and $\mathbf{p}$ except at $\mathbf{p}=0$; moreover, we may require that $\alpha(K_\lambda, \mathbf{p}) = 0$, where K_λ are the boosts in the x_1 direction introduced in relation (5). It is then easy to verify that the functions $(B_\lambda\Omega)_k(\mathbf{p})$ are continuous and that the analog of relation (8) holds. The proof of Proposition 3 carries over almost literally, and we may therefore omit the details.

Finally we want to point out a further generalization of our main result. In an asymptotically complete theory of massless particles there always exist two representations $D^{\text{in}}(\lambda)$ and $D^{\text{out}}(\lambda)$ of the group of dilations which act on the asymptotic fields A^{in} and A^{out}, respectively, as in a free field theory. Their commutation relations with the translations $U(x)$ and Lorentz transformations $U(\Lambda)$ are again given by (2a). However, they do not, in general, act on the basic fields according to relation (2b). In order that the dilations are an asymptotically visible symmetry, it would be sufficient to require

$$D^{\text{in}}(\lambda) = D^{\text{out}}(\lambda) = D(\lambda) \quad (12)$$

and relation (2b) could be dropped. But this assumption still implies that the S matrix is trivial! To verify this, one has only to realize that Propositions 2 and 3 still hold in this case with obvious modifications. The proof of Proposition 2 depends on the clustering properties of the vacuum and relation (2a) and therefore applies. Of course, the operators B_λ are in general not local. However, relation (10) which was crucial for the proof of Proposition 3 can still be established. This follows simply from the fact that the asymptotic nets $\mathcal{O} \to \mathfrak{F}^{\text{in}}(\mathcal{O})$ and $\mathcal{O} \to \mathfrak{F}^{\text{out}}(\mathcal{O})$ transform under the dilations $D(\lambda) = D^{\text{in}}(\lambda) = D^{\text{out}}(\lambda)$ according to relation (2b). Hence, if, for example, $A \in \mathfrak{F}(\mathcal{O})$, where $\mathcal{O}$ is any bounded region which contains the origin and if $F_+^{\text{in}} \in \mathfrak{F}^{\text{in}}(\mathcal{O}_+)$, where $\mathcal{O}_+$ has a positive timelike separation from $\mathcal{O}$, one gets for $\lambda < 1$

$$[D(\lambda)AD(\lambda)^{-1}, F_+^{\text{in}}] = D(\lambda)[A, D(\lambda^{-1})F_+^{\text{in}}D(\lambda^{-1})^{-1}]D(\lambda)^{-1} = 0$$

by Huyghens' principle. A similar relation holds for $F_-^{\text{out}} \in \mathfrak{F}^{\text{out}}(\mathcal{O}_-)$. It is then easy to verify that the operators B_λ commute for small λ with the operators in $\mathfrak{F}^{\text{in}}(\mathcal{O}_+)$ and $\mathfrak{F}^{\text{out}}(\mathcal{O}_-)$ where the regions $\mathcal{O}_+$, $\mathcal{O}_-$ depend only on the localization properties of f and A_φ. The rest of the argument can then be carried over.

ACKNOWLEDGMENTS

We are indebted to R. Haag for stimulating discussions and to G. Mack and G. Roepstorff for useful remarks on helicity representations.

[1]G. F. Dell-Antonio, Nuovo Cimento A 12, 756 (1972).
[2]S. Doplicher, R. Haag, and J.E. Roberts, Commun. Math. Phys. 13, 1 (1969).
[3]R. Haag, J.T. Lopuszanski, and M. Sohnius, Nucl. Phys. B 88, 257 (1975).
[4]D. Buchholz, Commun. Math. Phys. 42, 269 (1975); "Collision Theory for Massless Bosons," DESY Report 76/22 (1976) (to be published in Commun. Math. Phys.).
[5]S. Doplicher, R. Haag, and J.E. Roberts, Commun. Math. Phys. 35, 49 (1974).
[6]E. P. Wigner, Ann. Math. 40, 149 (1939); A.S. Wightman, "L'invariance dans la mécanique quantique relativiste," in *Dispersion Relations and Elementary Particles* (Hermann, Paris, 1960).

Renormalized off-energy-shell Coulomb scattering

J. Zorbas*

Department of Mathematics, University of British Columbia, Vancouver, B.C. Canada V6T 1W5
(Received 28 July 1976)

The complex-energy distorted plane waves for scattering via a general class of two-body Coulomb-like potentials are shown to satisfy off-energy-shell Lippmann–Schwinger equations in the case of two particle scattering and appropriately iterated Weinberg–Van Winter and Faddeev equations in the case of three particle scattering. A renormalized off-energy-shell formalism is defined for scattering involving more than one charged fragment in either the incoming or outgoing channel. The existence of the limit to real energies of the renormalized off-energy-shell formalism is verified.

I. INTRODUCTION

A knowledge of the various possible off-energy-shell two-body T matrices allows one to construct the kernel of the Faddeev equations.[1] An important case for which integral representations of the off-shell two-body T matrix are known is that of the pure Coulomb potential.[2–4] These integral representations have been applied to elastic and rearrangement scattering of a charged particle by an uncharged fragment (see Ref. 5 for references to previous results). Unfortunately, when there is more than one charged fragment in either the incoming or outgoing channel, the solutions to the off-shell equations will not have physical on-energy-shell limits. In this paper we apply the recently developed stationary Hilbert space scattering formalism[6,7] together with various results concerning the existence and integrability of the Green's functions[8] to define a renormalized off-energy-shell scattering formalism for general Coulomb-like scattering.

In order to define a renormalized off-shell formalism, the usual off-shell scattering theory must be related to the operators $W_{\pm\epsilon}^{(\alpha)}$ defined as follows:

$$W_{\pm\epsilon}^{(\alpha)} = (\pm)\int_0^{\pm\infty} du \exp(\mp u)\Omega^{(\alpha)}(u/\epsilon),$$
$$\Omega^{(\alpha)}(t) = \exp(iHt)\exp(-iH_\alpha t)P^{(\alpha)}, \tag{1.1}$$

where the self-adjoint operators H and H_α denote the full Hamiltonian and α-channel Hamiltonian respectively and $P^{(\alpha)}$ denotes the projector onto the channel subspace $\mathcal{H}^{(\alpha)}$. The operators $W_{\pm\epsilon}^{(\alpha)}$ provide the link between the time-dependent and stationary short-range scattering theory[9–11] and can be shown[10,11] to have the following strong Riemann–Stieltjes integral representations:

$$W_{\pm\epsilon}^{(\alpha)} = \int_{-\infty}^{+\infty} \frac{\pm i\epsilon}{H-\lambda \pm i\epsilon}\, d_\lambda E_\lambda^{H_\alpha} P^{(\alpha)}, \tag{1.2}$$

where $E_\lambda^{H_\alpha}$ denotes the spectral function of H_α. If the Green's functions corresponding to $(H-\lambda\pm i\epsilon)^{-1}$ satisfy various conditions [see Theorem (4.1)] the complex-energy distorted plane waves can be defined in terms of the Fourier transform of the Green's functions and the operators $W_{\pm\epsilon}^{(\alpha)}$ can be expanded in terms of these complex-energy distorted plane waves. Thus there exists a mathematically rigorous relation between the time-dependent theory based on the operators $\Omega^{(\alpha)}(t)$ appearing in (1.1) and the time-independent formalism based on the complex-energy distorted plane waves defined in terms of the Fourier transform of the full Green's functions.

In the case of N-body Coulomb-like scattering Dollard[12–14] has shown that one can formulate a time-dependent scattering theory via the α-channel modified or renormalized wave operators, $\Omega_\pm^{(\alpha)}$, defined as follows:

$$\Omega_\pm^{(\alpha)} = s\text{-}\lim_{t\to\pm\infty}\hat{\Omega}^{(\alpha)}(t), \quad \hat{\Omega}^{(\alpha)}(t) = \Omega^{(\alpha)}(t)\exp[-iG^{(\alpha)}(t)], \tag{1.3}$$

where the time-dependent renormalization term $G^{(\alpha)}(t)$ is given by

$$G^{(\alpha)}(t) = \epsilon(t)\sum_{j<k}\frac{e_j e_k m_j m_k}{|m_j\hat{\mathbf{p}}_k - m_k\hat{\mathbf{p}}_j|}\log\left[\frac{2|t|\,|m_j\hat{\mathbf{p}}_k - m_k\hat{\mathbf{p}}_j|^2}{m_j m_k(m_j+m_k)}\right],$$
$$\epsilon(t) = \begin{cases} 1, & t>0,\\ -1, & t<0,\end{cases} \tag{1.4}$$

where e_j, m_j, and $\hat{\mathbf{p}}_j$ denote respectively the total charge, mass, and momentum of the jth fragment. The existence of the renormalized wave operators (1.3) has been shown[12,13] for N-body scattering via a general class of two-body Coulomb-like potentials [in particular for two-body potentials which satisfy condition (β) of Sec. IIB].

It is clear that the time-dependent renormalization of the operators $\Omega^{(\alpha)}(t)$ needed to obtain a satisfactory time-dependent Coulomb scattering theory will induce a modification of the off-energy-shell formalism based on the complex-energy distorted plane waves. This modification was derived in a Hilbert space context[7] and will be briefly reviewed in Sec. II together with various technical results which will be required later.

In Sec. IV sufficient conditions are given in order that the operators $W_{\pm\epsilon}^{(\alpha)}$ can be expanded in terms of complex-energy distorted plane waves. The complex-energy distorted plane waves are shown to satisfy off-energy-shell Lippmann–Schwinger equations in Sec. V and appropriately iterated Weinberg–Van Winter and Faddeev equations in Sec. VI. A renormalized off-energy-shell formalism for two- and three-particle Coulomb scattering is defined in Secs. V and VII respectively. The results of Ref. 7 are used to show the existence of the limit to real energies of the renormalized complex-energy distorted plane waves and renormalized off-energy-shell T matrices for two- and three-particle scattering via a general class of two-body Coulomb-like potentials.

II. PRELIMINARIES

A. Coordinates and notation

Corresponding to the channel α the N particles making up the scattering system will be grouped into n_α fragments. By a suitable transformation the coordinates corresponding to the particles making up a fragment j can be transformed to a center-of-mass coordinate $\mathbf{Z}_j$ together with a set of "internal" coordinates (if j is composite) denoted collectively by x^j. Since only two-body forces are considered in this paper it is convenient to separate the total center-of-mass coordinate. Thus the n_α center-of-mass coordinates $(\mathbf{Z}_1, \ldots, \mathbf{Z}_{n_\alpha})$ are transformed into a total center-of-mass coordinate $\bar{\mathbf{x}}$ together with $n_\alpha - 1$ "relative" center-of-mass coordinates $(\mathbf{x}_1, \ldots, \mathbf{x}_{n_\alpha-1})$. Such a choice of coordinates induces a decomposition of the Hilbert space $\mathcal{H} = \mathcal{L}^2(R^{3N})$ as follows: $\mathcal{H} = \mathcal{H}_{\text{c.m.}} \otimes \mathcal{H}_{\text{int}} = \mathcal{L}^2(R^3) \otimes \mathcal{L}^2(R^{3(N-1)})$. Furthermore, the wave and scattering operators factor as $\Omega_\pm^{(\alpha)} = \Pi \otimes \Omega_\pm^{(\alpha)}$ and $S_{\alpha\beta} = \Pi \otimes S_{\alpha\beta}$, where Π denotes the identity on $\mathcal{H}_{\text{c.m.}}$ and the operators acting in $\mathcal{H}_{\text{int}}$ have been denoted by the same symbols as the operators acting in $\mathcal{H}$. All results contained in Ref. 7 can now be expressed as results in $\mathcal{H}_{\text{int}}$ without reference to $\mathcal{H}_{\text{c.m.}}$.

In the following the coordinates after removing the total center-of-mass will be collectively denoted by x, where $x = (x_{n_\alpha-1}, x^{n_\alpha}) = (\mathbf{x}_1, \ldots, \mathbf{x}_{n_\alpha-1}, x^1, \ldots, x^{n_\alpha})$. The conjugate momentum corresponding to $x_{n_\alpha-1}$ will be collectively denoted by $p_{n_\alpha-1}$, where $p_{n_\alpha-1} = (\mathbf{p}_1, \ldots, \mathbf{p}_{n_\alpha-1})$. Furthermore, the coordinates $x_{n_\alpha-1}$ will be chosen so that the total energy $E^{(\alpha)}$ associated with H_α takes the following form:

$$E^{(\alpha)} = \sum_{j=1}^{n_\alpha-1} \mathbf{p}_j^2 + E_{\text{int}}^{(\alpha)}, \tag{2.1}$$

where $E_{\text{int}}^{(\alpha)}$ denotes the bound state energy of the composite fragments making up the channel α.

In order to derive the existence of an off-shell formalism, various Hilbert space relations will be expanded in terms of free plane waves,

$$\phi_{p_{n_\alpha-1}}(x_{n_\alpha-1}) = (2\pi)^{-3(n_\alpha-1)/2} \exp(ip_{n_\alpha-1} \cdot x_{n_\alpha-1}),$$

where

$$p_{n_\alpha-1} \cdot x_{n_\alpha-1} = \sum_{j=1}^{n_\alpha-1} \mathbf{p}_j \cdot \mathbf{x}_j,$$

via the Fourier transform. The Fourier transform is a unitary map from the position representation $\mathcal{L}^2(R^{3(n_\alpha-1)})$ to the momentum representation $\mathcal{L}^2(R^{3(n_\alpha-1)}, dp_{n_\alpha-1})$, which is defined as follows:

$$\hat{f}(p_{n_\alpha-1}) = \text{l.i.m.} \int_{R^{3(n_\alpha-1)}} dx_{n_\alpha-1} \overline{\phi_{p_{n_\alpha-1}}(x_{n_\alpha-1})} f(x_{n_\alpha-1}), \tag{2.2}$$

where $f \in \mathcal{L}^2(R^{3(n_\alpha-1)})$ and l.i.m. means the $\mathcal{L}^2$-limit $D \to R^{3(n_\alpha-1)}$, where D is a sequence of bounded sets. Furthermore,

$$f(x_{n_\alpha-1}) = \text{l.i.m.} \int_{R^{3(n_\alpha-1)}} dp_{n_\alpha-1} \phi_{p_{n_\alpha-1}}(x_{n_\alpha-1}) \hat{f}(p_{n_\alpha-1}) \tag{2.3}$$

for each $f \in \mathcal{L}^2(R^{3(n_\alpha-1)})$.

B. The Hamiltonian

In this paper we will assume that the scattering system consists of N distinguishable spinless particles interacting via two-body Coulomb-like potentials and described by the self-adjoint Hamiltonian H acting in $\mathcal{L}^2(R^{3(N-1)})$ having the form

$$H = H_0 + \sum_{i<i'} V_{ii'}, \quad H_0 = -\sum_{j=1}^{N-1} \nabla_j^2, \tag{2.4}$$

where H_0 is the unique self-adjoint extension of the formal sum of Laplace operators ∇_j^2 [Ref. 15, Chap. V, Sec. (5.2)]. In particular, it will be convenient to require that each $V_{ii'}$ satisfies the following condition:

$$\begin{aligned}&(\mathcal{B})\ V = V_c + V_s, \quad V_s \in \mathcal{L}^2(R^3) \cap \mathcal{L}^1(R^3),\\ &V_c(\mathbf{x}) = Z|\mathbf{x}|^{-1}, \quad \text{where } Z \text{ is a constant.}\end{aligned} \tag{2.5}$$

If each $V_{ii'}$ satisfies $(\mathcal{B})$, then[16] $V_{ii'}$ is a Kato potential, i.e., $\mathcal{D}(V_{ii'}) \supset \mathcal{D}(H_0)$ and for any $a > 0$ there exists a $b > 0$ such that

$$\|V_{ii'}\psi\| \leq a\|H_0\psi\| + b\|\psi\|, \quad \psi \in \mathcal{D}(H_0), \tag{2.6}$$

and thus[15] H and the α-channel Hamiltonian H_α are self-adjoint with $\mathcal{D}(H) = \mathcal{D}(H_\alpha) = \mathcal{D}(H_0)$.

C. Stationary Coulomb scattering

The existence of stationary renormalization terms denoted by $F_{\pm\epsilon}^{(\alpha)*}$ for N-particle Coulomb scattering was shown in Refs. 6 and 7. In particular for $G^{(\alpha)}(t)$ given by (1.4) the $F_{\pm\epsilon}^{(\alpha)*}$ can be computed and are given by

$$\begin{aligned}F_{\pm\epsilon}^{(\alpha)*}(p_{n_\alpha-1}) &= \left\{(\pm)\int_0^{\pm\infty} du \exp[\mp u + iG^{(\alpha)}(u/\epsilon)]\right\}^{-1}\\ &= \Gamma\left(1 \pm i \sum_{j<k}^{n_\alpha} \frac{m_j m_k e_j e_k}{|\mathbf{p}_{jk}|}\right)^{-1}\\ &\quad\times \exp\left[\pm i \sum_{j<k}^{n_\alpha} \frac{m_j m_k e_j e_k}{|\mathbf{p}_{jk}|} \log \frac{\epsilon m_j m_k (m_j + m_k)}{2|\mathbf{p}_{jk}|^2}\right],\end{aligned} \tag{2.7}$$

where $\mathbf{p}_{jk} = \sum_{l=1}^{n_\alpha-1} C_{jk}^l \mathbf{p}_l = m_j \hat{\mathbf{p}}_k - m_k \hat{\mathbf{p}}_j$ with the constants C_{jk}^l depending on the particular choice of relative center-of-mass coordinates. The expressions (2.7) are well-defined for each $\epsilon > 0$ on $\mathcal{D}^{(\alpha)}$ the set of functions dense in $\mathcal{H}^{(\alpha)}$ having the form $\phi = \phi_1 \Pi_{j=1}^{n_\alpha} \chi_j$, where χ_j, $j = 1, \ldots, n_\alpha$ denote the bound states making up the channel α and $\phi_1 \in \mathcal{L}^2(R^{3(n_\alpha-1)}; dp_{n_\alpha-1})$ satisfies

$$\Gamma\left(1 \pm i \sum_{j<k}^{n_\alpha} \frac{m_j m_k e_j e_k}{|\mathbf{p}_{jk}|}\right)^{-1} \phi_1 \in \mathcal{L}^2(R^{3(n_\alpha-1)}; dp_{n_\alpha-1}). \tag{2.8}$$

It has been shown[7] that if the renormalized wave operators $\Omega_\pm^{(\alpha)}$ for Coulomb-like scattering exist then they are related to $W_{\pm\epsilon}^{(\alpha)}$ as follows:

$$\Omega_\pm^{(\alpha)}\psi = \text{s-}\lim_{\epsilon \to +0} W_{\pm\epsilon}^{(\alpha)} F_{\pm\epsilon}^{(\alpha)*}\psi \tag{2.9}$$

for $\psi \in \mathcal{D}^{(\alpha)}$. Furthermore, the S operator $S_{\alpha\beta} = (2\pi i)^{-1} \times \Omega_+^{(\beta)*}\Omega_-^{(\alpha)}$ has the following stationary representations:

$$\begin{aligned}\langle\phi|S_{\alpha\beta}\psi\rangle &= \lim_{\epsilon \to +0} (-\pi)^{-1}\\ &\quad\times \left\langle F_{+\epsilon}^{(\beta)*}\phi \middle| \int_{-\infty}^{+\infty} d_\lambda E_\lambda^{H_\beta} V^{(\beta)} \Omega_-^{(\alpha)} \frac{\epsilon}{(H_\alpha - \lambda)^2 + \epsilon^2}\psi\right\rangle\end{aligned} \tag{2.10}$$

for $\phi \in \mathcal{D}^{(\beta)}$, $\psi \in \mathcal{H}^{(\alpha)}$ with $V^{(\beta)} = H - H_\beta$ and

$$\langle\phi|S_{\alpha\beta}\psi\rangle=\lim_{\epsilon_1\to+0}\lim_{\epsilon_2\to+0}(-\pi)^{-1}\times\left\langle F^{(\beta)*}_{+\epsilon_1}\phi\right|\int_{-\infty}^{+\infty}d_\lambda E^{H_\beta}_\lambda V^{(\beta)}W^{(\alpha)}_{-\epsilon_2}F^{(\alpha)*}_{-\epsilon_2}\times\frac{\epsilon_1}{(H_\alpha-\lambda)^2+\epsilon_1^2}\psi\Big\rangle \tag{2.11}$$

for $\phi\in\mathcal{D}^{(\beta)}$, $\psi\in\hat{\mathcal{D}}^{(\alpha)}$ with $\hat{\mathcal{D}}^{(\alpha)}$ defined as follows

$$\hat{\mathcal{D}}^{(\alpha)}=\left\{\psi\in\mathcal{D}^{(\alpha)}\left|\Gamma\left(1\pm i\sum_{j<k}^{n_\alpha}\frac{m_jm_ke_je_k}{|\mathbf{p}_{jk}|}\right)^{-1}\psi\in\mathcal{D}(H_\alpha)\right.\right\}. \tag{2.12}$$

III. SCREENED OFF-SHELL T-OPERATOR

In the case of two particle scattering via short-range potentials the on-energy-shell T matrix takes the form

$$(\mathbf{p}|T|\mathbf{p}')_{E^{(0)}(p)=E^{(0)}(p')}=C\int_{R^3}d\mathbf{x}\overline{\phi_{\mathbf{p}}(\mathbf{x})}V(\mathbf{x})\phi^{+}_{\mathbf{p}'}(\mathbf{x})\big|_{E^{(0)}(p)=E^{(0)}(p')}, \tag{3.1}$$

where $\phi^{\pm}_{\mathbf{p}}(\mathbf{x})$ denote the distorted plane waves and C is a constant. It is clear from the form of $\phi^{\pm}_{\mathbf{p}}(\mathbf{x})$ for large $|\mathbf{x}|$ that the above integral will not exist in the Lebesgue sense unless $|V(\mathbf{x})|=O(|\mathbf{x}|^{-3-\eta})$, $\eta>0$, as $|\mathbf{x}|\to\infty$. Thus for potentials which decrease slower than $|\mathbf{x}|^{-3-\eta}$, $\eta>0$, for large $|\mathbf{x}|$ a convergence factor must be introduced within the T matrix and the physical T matrix understood as an appropriate limit as the convergence factor is removed. In this section we will introduce the convergence factor within the stationary Hilbert space formalism and relate the resulting screened off-shell T operator to the S operator for Coulomb scattering.

It will be convenient to consider cutoff functions $g^{(\alpha)}_R(x_{n_\alpha-1})$ which satisfy the following general requirements denoted collectively by (C): $g^{(\alpha)}_R(x_{n_\alpha-1})$ is a real measureable function of $(\mathbf{x}_1,\ldots,\mathbf{x}_{n_\alpha-1})$ which satisfies for some constant D

$$|g^{(\alpha)}_R(x_{n_\alpha-1})|\le D,\quad\text{for all } R>0,$$
$$\lim_{R\to\infty}g^{(\alpha)}_R(x_{n_\alpha-1})=1, \tag{3.2}$$

and is such that the corresponding cutoff Hamiltonian

$$H_R=H_\alpha+V^{(\alpha)}_R,\quad V^{(\alpha)}_R=g^{(\alpha)}_R V^{(\alpha)} \tag{3.3}$$

is self-adjoint with domain $\mathcal{D}(H_0)$ for each $R>0$, where

$$V^{(\alpha)}_R\in\mathcal{L}^1(R^{3(n_\alpha-1)})\cap\mathcal{L}^2(R^{3(n_\alpha-1)})\quad\text{for } 0<R<\infty.$$

If the two-body potentials making up H are Kato potentials, then $V^{(\alpha)}_R$ is a Kato potential for $R>0$ and thus H_R having the form (3.3) is self-adjoint with domain $\mathcal{D}(H_0)$ for each $R>0$.

The following theorem provides cutoff dependent generalizations of the stationary representations (2.10) and (2.11) of the S operator.

Theorem 3.1: Suppose that $H=H_0+\sum_{i<j}V_{ij}$, where each V_{ij} is symmetric on $\mathcal{D}(V_{ij})\supset\mathcal{D}(H_0)$. In addition assume that the renormalized wave operators for Coulomb-like scattering exist and the cutoff function $g^{(\beta)}_R$ satisfies (C). Then the following stationary representations of the S operator $S_{\alpha\beta}$ are valid:

$$\langle\phi|S_{\alpha\beta}\psi\rangle=\lim_{\epsilon\to+0}\lim_{R\to\infty}(-\pi)^{-1}\times\left\langle F^{(\beta)*}_{+\epsilon}\phi\right|\int_{-\infty}^{+\infty}d_\lambda E^{H_\beta}_\lambda V^{(\beta)}_R\Omega^{(\alpha)}_{-}\frac{\epsilon}{(H_\alpha-\lambda)^2+\epsilon^2}\psi\Big\rangle \tag{3.4}$$

for $\phi\in\mathcal{D}^{(\beta)}$, $\psi\in\mathcal{D}(H_\alpha)\cap\mathcal{H}^{(\alpha)}$ and

$$\langle\phi|S_{\alpha\beta}\psi\rangle=\lim_{\epsilon_1\to+0}\lim_{\epsilon_2\to+0}\lim_{R\to+\infty}(-\pi)^{-1}\times\left\langle F^{(\beta)*}_{+\epsilon_1}\phi\right|\int_{-\infty}^{+\infty}d_\lambda E^{H_\beta}_\lambda V^{(\beta)}_R W^{(\alpha)}_{-\epsilon_2}F^{(\alpha)*}_{-\epsilon_2}\times\frac{\epsilon_1}{(H_\alpha-\lambda)^2+\epsilon_1^2}\psi\Big\rangle \tag{3.5}$$

for $\phi\in\mathcal{D}^{(\beta)}$ and $\psi\in\hat{\mathcal{D}}^{(\alpha)}$.

Proof: We will only give the proof of (3.5) since (3.4) can be shown in an analogous manner.

The relation (3.5) will follow from (2.11) if the following equality is valid:

$$\int_{-\infty}^{+\infty}d_\lambda E^{H_\beta}_\lambda\,\text{s-}\lim_{R\to\infty}V^{(\beta)}_R W^{(\alpha)}_{-\epsilon_2}F^{(\alpha)*}_{-\epsilon_2}\frac{\epsilon_1}{(H_\alpha-\lambda)^2+\epsilon_1^2}\psi=\text{s-}\lim_{R\to\infty}\int_{-\infty}^{+\infty}d_\lambda E^{H_\beta}_\lambda V^{(\beta)}_R W^{(\alpha)}_{-\epsilon_2}F^{(\alpha)*}_{-\epsilon_2}\frac{\epsilon_1}{(H_\alpha-\lambda)^2+\epsilon_1^2}\psi \tag{3.6}$$

for each $\epsilon_1>0$, $\epsilon_2>0$, and $\psi\in\hat{\mathcal{D}}^{(\alpha)}$.

Due to Lemma 1 of Ref. 10 there exists constants a and b such that

$$\|V^{(\beta)}_R W^{(\alpha)}_{-\epsilon}\psi\|\le a\|HW^{(\alpha)}_{-\epsilon}\psi\|+b\|\psi\| \tag{3.7}$$

for $\psi\in\mathcal{D}(H_\alpha)$. A similar argument as given in the proof of Theorem 1, Ref. 17, applied to the first term on the right side of (3.7) yields the existence of constants A and B such that

$$\|V^{(\beta)}_R W^{(\alpha)}_{-\epsilon}\psi\|\le A\|H_\alpha\psi\|+B\|\psi\|$$

for $\psi\in\mathcal{D}(H_\alpha)$. Thus the following Bochner integrals exist (Theorem 3.7.4, Ref. 18):

$$\int_0^{\pm\infty}dt\exp[\mp\epsilon_1 t+iH_\beta t]V^{(\beta)}_R W^{(\alpha)}_{-\epsilon_2}\exp[-iH_\alpha t]F^{(\alpha)*}_{-\epsilon_2}\psi \tag{3.8}$$

for $\psi\in\hat{\mathcal{D}}^{(\alpha)}$ and $\epsilon_1>0$, $\epsilon_2>0$. Furthermore, Theorem 3′ of Ref. 10 is applicable and allows the interchange of the spectral integral and Bochner integrals which leads to the following equality:

$$\pi^{-1}\int_{-\infty}^{+\infty}d_\lambda E^{H_\beta}_\lambda V^{(\beta)}_R W^{(\alpha)}_{-\epsilon_2}\frac{\epsilon_1}{(H_\alpha-\lambda)^2+\epsilon_1^2}F^{(\alpha)*}_{-\epsilon_2}\psi=(-2\pi)^{-1}\int_0^{-\infty}dt\exp(\epsilon_1 t+iH_\beta t)V^{(\beta)}_R W^{(\alpha)}_{-\epsilon_2}\exp(-H_\alpha t)F^{(\alpha)*}_{-\epsilon_2}\psi+(2\pi)^{-1}\int_0^{\infty}dt\exp(-\epsilon_1 t+iH_\beta t)V^{(\beta)}_R W^{(\alpha)}_{-\epsilon_2}\times\exp(-iH_\alpha t)F^{(\alpha)*}_{-\epsilon_2}\psi \tag{3.9}$$

for $\epsilon_1>0$, $\epsilon_2>0$, and $\psi\in\hat{\mathcal{D}}^{(\alpha)}$. By the Lebesgue dominated convergence theorem for Bochner integrals[18] the strong limit $R\to+\infty$ of (3.9) can be taken explicitly, and it yields the Bochner integrals in (3.9) with $V^{(\beta)}_R$ replaced by $V^{(\beta)}$. Applying Theorem 3′ of Ref. 10 allows us to conclude that these Bochner integrals are equal to the first term in (3.6), which concludes the proof of (3.5).

IV. COMPLEX-ENERGY DISTORTED PLANE WAVES

In order to define renormalized complex-energy distorted plane waves corresponding to the left side of (2.9), we require the relationship between the operators $W_{\pm\epsilon}^{(\alpha)}$ and the complex-energy distorted plane waves defined by (4.3). The derivation of this relationship is given in this section.

The following theorem provides sufficient conditions for the expansion of the operators $W_{\pm\epsilon}^{(\alpha)}$ via complex-energy distorted plane waves.

Theorem 4.1: Suppose that the resolvent $(H-\zeta)^{-1}$, $\operatorname{Im}\zeta\neq 0$, has the form of an integral operator in the position representation, i.e.,

$$[(H-\zeta)^{-1}](x)=\int_{R^{3(N-1)}}dx'\,G(x,x';\zeta)\psi(x') \tag{4.1}$$

for almost all $x\in R^{3(N-1)}$ and $\psi\in\mathcal{L}^2(R^{3(N-1)})$. In addition, assume that the kernel $G(x,x';\zeta)$ satisfies for $|\operatorname{Re}\zeta|<M<\infty$

$$\int_{R^{3(N-1)}}dx\int_{R^{3(N-1)}}dx'\,|\phi(x)G(x,x';\zeta)|\leq C(\phi,M,\operatorname{Im}\zeta) \tag{4.2}$$

for each $\phi\in\mathcal{L}^1(R^{3(N-1)})\cap\mathcal{L}^2(R^{3(N-1)})$, where the finite constant $C(\phi,M,\operatorname{Im}\zeta)$ depends on ϕ, M, and $\operatorname{Im}\zeta$. Furthermore, assume that for a given channel α the bound states $\chi_j(x^j)$ making up the channel α are square integrable and $\mathcal{L}^\infty$ functions of x^j for $j=1,\ldots,n_\alpha$. Then the complex-energy distorted plane waves $\phi_{p_{n_\alpha-1}}^{(\alpha)\pm\epsilon}(x)$ defined by

$$\phi_{p_{n_\alpha-1}}^{(\alpha)\pm\epsilon}(x)=(\mp i\epsilon)\int_{R^{3(N-1)}}dx'\,G(x,x';E^{(\alpha)}\pm i\epsilon)\times\prod_{j=1}^{n_\alpha}\chi_j(x^{j'})\phi_{p_{n_\alpha-1}}(x'_{n_\alpha-1}) \tag{4.3}$$

exist for almost all $x\in R^{3(N-1)}$ and each $\epsilon>0$. Furthermore, for each $\psi\in\mathcal{H}^{(\alpha)}$ of the form $\psi=\psi_1\prod_{j=1}^{n_\alpha}x_j$ with $\hat{\psi}_1(p_{n_\alpha-1})$ a bounded function with compact support the operators $W_{\pm\epsilon}^{(\alpha)}$ can be expanded via $\phi_{p_{n_\alpha-1}}^{(\alpha)\mp\epsilon}(x)$ as follows:

$$(W_{\pm\epsilon}^{(\alpha)}\psi)(x)=\int_{R^{3(n_\alpha-1)}}dp_{n_\alpha-1}\phi_{p_{n_\alpha-1}}^{(\alpha)\mp\epsilon}(x)\hat{\psi}_1(p_{n_\alpha-1}) \tag{4.4}$$

for almost all $x\in R^{3(N-1)}$ and each $\epsilon>0$.

Proof: The requirement that $\chi_j\in\mathcal{L}^\infty$, $j=1,\ldots,n_\alpha$, and (4.2) show that $G(x,x';E^{(\alpha)}\pm i\epsilon)\prod_{j=1}^{n_\alpha}\chi_j(x^{j'})\times\phi_{p_{n_\alpha-1}}(x'_{n_\alpha-1})$ is integrable in x' for almost all $x\in R^{3(N-1)}$ and thus $\phi_{p_{n_\alpha-1}}^{(\alpha)\pm\epsilon}(x)$ defined by (4.3) exists.

In order to derive (4.4), we introduce the function $\chi_\Delta(x_{n_\alpha-1})$ as follows:

$$\chi_\Delta(x_{n_\alpha-1})=\begin{cases}1, & x_{n_\alpha-1}\in\Delta,\\ 0, & x_{n_\alpha-1}\notin\Delta,\end{cases}$$

where Δ is a compact subset of $R^{3(n_\alpha-1)}$. In analogy with the derivation of (1.2) from (1.1) the following strong Riemann–Stieltjes integral representations can be verified:

$$W_{\pm\epsilon}^{(\alpha)}=\underset{\Delta\to R^{3(n_\alpha-1)}}{\text{s-lim}}\int_{-\infty}^{+\infty}\frac{\pm i\epsilon}{H-\lambda\pm\epsilon}\chi_\Delta d_\lambda E_\lambda^{H\,\alpha}P^{(\alpha)}. \tag{4.5}$$

From the definition of the strong Riemann–Stietjes integral[10] (4.5) yields the following equality:

$$\langle\phi\,|\,W_{\pm\epsilon}^{(\alpha)}\psi\rangle=\lim_{\Delta\to R^{3(n_\alpha-1)}}\lim_{N\to\infty}\lim_{|\pi_N|\to 0}\sum_{l=1}^{N}\left\langle\phi\,\middle|\,\frac{\pm i\epsilon}{H-\lambda'_l\pm i\epsilon}\chi_\Delta E^{H_\alpha}_{(\lambda_{l-1},\lambda_l]}\psi\right\rangle \tag{4.6}$$

for each $\epsilon>0$, $\phi\in\mathcal{H}_{\text{int}}$, $\psi=\psi_1\prod_{j=1}^{n_\alpha}\chi_j\in\mathcal{H}^{(\alpha)}$, where $-N=\lambda_0<\lambda_1<\cdots<\lambda_N=N$, $\lambda'_l\in(\lambda_{l-1},\lambda_l]$, $|\pi_N|=\sup_{1\leq l\leq N}\times(\lambda_l-\lambda_{l-1})$. Applying (4.1), (4.2), and the assumption that $\chi_j\in\mathcal{L}^\infty$, $j=1,\ldots,n_\alpha$, allows us to rewrite (4.6) as follows:

$$\begin{aligned}\langle\phi\,|\,W_{\pm\epsilon}^{\alpha}\psi\rangle=&\lim_{\Delta\to R^{3(n_\alpha-1)}}\lim_{N\to\infty}\lim_{|\pi_N|\to 0}\sum_{l=1}^{N}\\&\int_{R^{3(n_\alpha-1)}}dp_{n_\alpha-1}\hat{\psi}_1(p_{n_\alpha-1})\chi_{(\lambda_{l-1},\lambda_l]}(E^{(\alpha)})\\&\times\int_{R^{3(N-1)}}dx\,\overline{\phi(x)}(\pm i\epsilon)\int_{R^{3(N-1)}}dx'\\&\times G(x,x';\lambda'_l\mp i\epsilon)\chi_\Delta(x'_{n_\alpha-1})\\&\times\prod_{j=1}^{n_\alpha}\chi_j(x^{j'})\phi_{p_{n_\alpha-1}}(x'_{n_\alpha-1})\end{aligned} \tag{4.7}$$

for $\phi\in\mathcal{L}^1(R^{3(N-1)})\cap\mathcal{L}^2(R^{3(N-1)})$, $\psi\in\mathcal{H}^{(\alpha)}$, $\hat{\psi}_1\in\mathcal{L}^1(R^{3(n_\alpha-1)})\cap\mathcal{L}^2(R^{3(n_\alpha-1)})$ and

$$\chi_{(\lambda_{l-1},\lambda_l]}(E^{(\alpha)})=\begin{cases}1, & \lambda_{l-1}<E^{(\alpha)}\leq\lambda_l,\\ 0, & E^{(\alpha)}\notin(\lambda_{l-1},\lambda_l].\end{cases}$$

In order to perform the sums and limits in (4.7) λ'_l must be replaced by $E^{(\alpha)}$. The first resolvent identity together with

$$\|(H-\zeta)^{-1}\psi\|_2\leq|\operatorname{Im}\zeta|^{-1}\|\psi\|_2 \tag{4.8}$$

yield the following inequality:

$$\begin{aligned}\sum_{l=1}^{N}\Bigg|&\int_{R^{3(n_\alpha-1)}}dp_{n_\alpha-1}\hat{\psi}_1(p_{n_\alpha-1})\chi_{(\lambda_{l-1},\lambda_l]}(E^{(\alpha)})\\&\times\int_{R^{3(N-1)}}dx\,\overline{\phi(x)}\Big\{[(H-\lambda'_l\pm i\epsilon)^{-1}\\&-(H-E^{(\alpha)}\pm i\epsilon)^{-1}]\chi_\Delta\prod_{j=1}^{n_\alpha}\chi_j\phi_{p_{n_\alpha-1}}\Big\}(x)\Bigg|\\&\leq|\pi_N|\,\epsilon^{-2}\|\hat{\psi}_1\|_1\|\phi\|_2\left\|\chi_\Delta\prod_{j=1}^{n_\alpha}\chi_j\right\|_2.\end{aligned} \tag{4.9}$$

Since the limit $|\pi_N|\to 0$ of (4.9) is zero the relations (4.7) may be replaced by the following equality:

$$\begin{aligned}\langle\phi\,|\,W_{\pm\epsilon}^{\alpha}\psi\rangle=&\lim_{\Delta\to R^{3(n_\alpha-1)}}\lim_{N\to\infty}\lim_{|\pi_N|\to 0}\sum_{l=1}^{N}\int_{R^{3(n_\alpha-1)}}dp_{n_\alpha-1}\\&\times\hat{\psi}_1(p_{n_\alpha-1})\chi_{(\lambda_{l-1},\lambda_l]}(E^{(\alpha)})\int_{R^{3(N-1)}}dx\,\overline{\phi(x)}(\pm i\epsilon)\\&\times\int_{R^{3(N-1)}}dx'\,G(x,x';E^{(\alpha)}\mp i\epsilon)\chi_\Delta(x'_{n_\alpha-1})\\&\times\prod_{j=1}^{n_\alpha}\chi_j(x^{j'})\phi_{p_{n_\alpha-1}}(x'_{n_\alpha-1}).\end{aligned} \tag{4.10}$$

By requiring that $\hat{\psi}_1(p_{n_\alpha-1})$ be a bounded function of compact support there exists an M such that $|E^{(\alpha)}(p_{n_\alpha-1})|\leq M$ for all $p_{n_\alpha-1}$ contained in the support of $\hat{\psi}_1$. Thus the following inequality is valid:

$$\begin{aligned}\int_{R^{3(n_\alpha-1)}}&dp_{n_\alpha-1}\int_{R^{3(N-1)}}dx\int_{R^{3(N-1)}}dx'\,\Bigg|\hat{\psi}_1(p_{n_\alpha-1})\\&\times\chi_{(\lambda_{l-1},\lambda_l]}(E^{(\alpha)})\overline{\phi(x)}G(x,x';E^{(\alpha)}\mp i\epsilon)\\&\times\chi_\Delta(x'_{n_\alpha-1})\prod_{j=1}^{n_\alpha}\chi_j(x^{j'})\phi_{p_{n_\alpha-1}}(x'_{n_\alpha-1})\Bigg|\\&\leq C(\phi,M,\epsilon)\prod_{j=1}^{n_\alpha}\|\chi_j\|_\infty\|\chi_{(\lambda_{l-1},\lambda_l]}\hat{\psi}_1\|_1.\end{aligned} \tag{4.11}$$

The above inequality justifies taking the limits and sum in (4.10) which then reduces to

$$\langle\phi\,|\,W^{\alpha}_{\pm\epsilon}\psi\rangle=\int_{R^{3(N-1)}}dx\,\overline{\phi(x)}\int_{R^{3(n_\alpha-1)}}dp_{n_\alpha-1}\times\phi^{(\alpha)\mp\epsilon}_{p_{n_\alpha-1}}(x)\hat{\psi}_1(p_{n_\alpha-1}). \tag{4.12}$$

The relations (4.4) follow immediately from (4.12) which concludes the proof of the theorem.

The requirements (4.1) and (4.2) can be verified for a general class of Coulomb-like potentials. This is the main content of Ref. 8 where the Green's function $G(x,x';\zeta)$ is shown to exist and satisfy (4.1) and (4.2) for two- and three-particle scattering via two-body Coulomb-like potentials which satisfy (β).

The following proposition shows the inadequacy of the complex-energy distorted plane waves defined by (4.3) when Coulomb forces are present.

Proposition 4.2: Suppose that the α-channel renormalized wave operators $\Omega^{(\alpha)}_{\pm}$ for Coulomb scattering exist and $\oplus_\alpha R^{(\alpha)}_{\pm}\mathcal{H}_{\rm int}\oplus\mathcal{H}_b=\mathcal{H}_{\rm int}$ where $R^{(\alpha)}_{\pm}\mathcal{H}_{\rm int}$ denote the ranges of $\Omega^{(\alpha)}_{\pm}$ and $\mathcal{H}_b$ denotes the subspace of $\mathcal{H}_{\rm int}$ spanned by the eigenvectors of H. Then

$$\operatorname*{w-lim}_{t\to\pm\infty}\exp(iHt)\exp(-iH_\alpha t)P^{(\alpha)}=0 \tag{4.13}$$

and

$$\operatorname*{w-lim}_{\epsilon\to+0}W^{(\alpha)}_{\pm\epsilon}=0. \tag{4.14}$$

If in addition $W^{(\alpha)}_{\pm\epsilon}$ has the expansion (4.4) then

$$\lim_{\epsilon\to+0}\int_{R^{3(N-1)}}dx\,\overline{\phi(x)}\int_{R^{3(n_\alpha-1)}}dp_{n_\alpha-1}\phi^{(\alpha)\pm\epsilon}_{p_{n_\alpha-1}}(x)\hat{\psi}_1(p_{n_\alpha-1})=0 \tag{4.15}$$

for $\phi\in\mathcal{L}^2\cap\mathcal{L}^1$ and $\psi=\psi_1\Pi^{n_\alpha}_{j=1}\chi_j$, $\chi_j\in\mathcal{L}^\infty\cap\mathcal{L}^2$, $j=1,\ldots,n_\alpha$, where $\hat{\psi}_1$ is a bounded function with compact support.

Proof: If ϕ is an eigenvector of H then $H\phi=E\phi$ where E is a constant and

$$|\langle\phi\,|\,\exp(iHt)\exp(-iH_\alpha t)P^{(\alpha)}\psi\rangle|=|\langle\phi\,|\,\exp(-iH_\alpha t)P^{(\alpha)}\psi\rangle|$$

which vanishes in the limit $t\to\pm\infty$ since $\exp(-iH_\alpha t)P^{(\alpha)}$ converges weakly to zero. Thus in order to show (4.13) we must verify the following equality:

$$\lim_{t\to\pm\infty}\left\langle\sum_{\beta=0}^{\infty}\phi^{(\beta)}_{\pm}\,|\exp(iHt)\exp(-iH_\alpha t)\psi\right\rangle=0 \tag{4.16}$$

for $\phi^{(\beta)}_{\pm}=R^{(\beta)}_{\pm}\phi$ and $\psi\in\mathcal{H}^{(\alpha)}$. It is straightforward to obtain the following inequality

$$\begin{aligned}\left|\left\langle\sum_{\beta=0}^{\infty}\phi^{(\beta)}_{\pm}\,|\exp(iHt)\exp(-iH_\alpha t)\psi\right\rangle\right| &\le\sum_{\beta=0}^{\alpha}\|\{\exp[iG^{(\beta)}(t)+iH_\beta t]\exp(-iHt)-\Omega^{(\beta)*}_{\pm}\}\phi^{(\beta)}_{\pm}\|\,\|\psi\|\\ &+\left|\sum_{\beta=0}^{\alpha}\langle\Omega^{(\beta)*}_{\pm}\phi^{(\beta)}_{\pm}\,|\exp[iG^{(\beta)}(t)+iH_\beta t-iH_\alpha t]\psi\rangle\right|\\ &+\left\|\sum_{\beta=\alpha+1}^{\infty}\phi^{(\beta)}_{\pm}\right\|\,\|\psi\|.\end{aligned} \tag{4.17}$$

By choosing α large enough the last term in (4.17) can be made as small as we like. Furthermore, for each fixed α the time-dependent terms on the right side of (4.17) vanish in the limit $t\to\pm\infty$. Thus (4.16) follows from the inequality (4.17).

The relations (4.14) follow from

$$\langle\phi\,|\,W^{(\alpha)}_{\pm\epsilon}\psi\rangle=(\pm)\int_0^{\pm\infty}du\exp(\mp u)\langle\phi\,|\,\Omega^{(\alpha)}(u/\epsilon)\psi\rangle \tag{4.18}$$

together with (4.13) and the Lebesgue dominated convergence theorem.

The equality (4.15) is an immediate consequence of (4.14).

The inadequacy of the short-range form of the wave operators (4.13) was first shown by Dollard[19] for two-particle Coulomb scattering.

If $W^{(\alpha)}_{\pm\epsilon}\psi$, $\psi=\psi_1\Pi^{n_\alpha}_{j=1}\chi_j\in\mathcal{D}^{(\alpha)}$, has the expansion (4.4) then the right side of (2.9) with $\epsilon>0$ takes the following form:

$$(W^{(\alpha)}_{\pm\epsilon}F^{(\alpha)*}_{\pm\epsilon}\psi)(x)=\int_{R^{3(n_\alpha-1)}}dp_{n_\alpha-1}\phi^{(\alpha)\mp\epsilon}_{p_{n_\alpha-1}}(x)F^{(\alpha)*}_{\pm\epsilon}(p_{n_\alpha-1})\hat{\psi}_1(p_{n_\alpha-1}) \tag{4.19}$$

for almost all $x\in R^{3(N-1)}$. Since (4.19) converges strongly to the renormalized wave operators it is appropriate to replace the short-range off-energy-shell formalism based on the complex-energy distorted plane waves $\phi^{(\alpha)\mp\epsilon}_{p_{n_\alpha-1}}(x)$ by an off-energy-shell formalism for Coulomb scattering based on the "renormalized" complex-energy distorted plane waves $\phi^{(\alpha)\mp\epsilon}_{p_{n_\alpha-1}}(x)F^{(\alpha)*}_{\pm\epsilon}(p_{n_\alpha-1})$. Thus the usual prescription for relating the Green's functions to the physical distorted plane waves must be modified by the stationary renormalization terms $F^{(\alpha)*}_{\pm\epsilon}(p_{n_\alpha-1})$.

V. TWO-PARTICLE OFF-SHELL COULOMB SCATTERING

The Hilbert space theory derived in Ref. 7 and outlined in Sec. II C will be applied in this section to define a renormalized off-shell formalism and to relate this formalism to the physical on-shell distorted plane waves and S-matrix for two-particle scattering via a general class of Coulomb-like potentials.

In order to derive the relationship between the renormalized off-shell formalism and the physical on-shell formalism we require that the physical distorted plane waves exist. That is in the case of general N-body scattering for each $\psi\in\mathcal{H}^{(\alpha)}$ of the form $\psi=\psi_1\Pi^{n_\alpha}_{j=1}\chi_j$ there exists physical α-channel distorted plane waves denoted by $\phi^{(\alpha)\mp}_{p_{n_\alpha-1}}(x)$ such that the following expansion is valid:

$$(\Omega^{(\alpha)}_{\pm}\psi)(x)=\text{l.i.m.}\int_{R^{3(n_\alpha-1)}}dp_{n_\alpha-1}\phi^{(\alpha)\mp}_{p_{n_\alpha-1}}(x)\hat{\psi}_1(p_{n_\alpha-1}), \tag{5.1}$$

where $\hat{\psi}_1(p_{n_\alpha-1})$ is defined by (2.2). In addition it will be convenient to require that the α-channel distorted plane waves satisfy

$$\int_{R^{3(n_\alpha-1)}}dp_{n_\alpha-1}|\phi^{(\alpha)\mp}_{p_{n_\alpha-1}}(x)\hat{\psi}_1(p_{n_\alpha-1})|\le C(\hat{\psi}_1), \tag{5.2}$$

where $C(\hat{\psi}_1)$ is a constant depending on $\hat{\psi}_1\in C^\infty_0(R^{3(n_\alpha-1)}\backslash\{0\})$ the C^∞ functions with compact support contained in $R^{3(n_\alpha-1)}\backslash\{0\}$. The expansion (5.1) has been verified[13,20] for two-particle scattering via a general class of Coulomb-like potentials. The inequality (5.2) can be shown for two-particle pure Coulomb scattering.

The following theorem verifies the off-energy-shell Lippmann–Schwinger equations and provides the relationship between the renormalized complex-energy distorted plane waves $\phi_{\mathbf{p}}^{\pm\epsilon}(\mathbf{x})F_{\mp\epsilon}^{*}(p)$ and the physical distorted plane waves $\phi_{\mathbf{p}}^{\pm}(\mathbf{x})$.

Theorem 5.1: Suppose that $H=H_0+V$, where V satisfies (β). Then the complex-energy distorted plane waves $\phi_{\mathbf{p}}^{\pm\epsilon}(\mathbf{x})$ satisfy the off-energy-shell Lippmann–Schwinger equations, i.e.,

$$\phi_{\mathbf{p}}^{\pm\epsilon}(\mathbf{x})=\phi_{\mathbf{p}}(\mathbf{x})-\int_{R^3}d\mathbf{x}'\,G_0(\mathbf{x},\mathbf{x}';E^{(0)}\pm i\epsilon)V(\mathbf{x}')\phi_{\mathbf{p}}^{\pm\epsilon}(\mathbf{x}') \tag{5.3}$$

for almost all $\mathbf{p},\mathbf{x}\in R^3$ and each $\epsilon>0$ where $G_0(\mathbf{x},\mathbf{x}';\zeta)$ is the kernel corresponding to $(H_0-\zeta)^{-1}$, $\operatorname{Im}\zeta>0$. Furthermore, if the expansion (5.1) is valid then the physical distorted plane waves $\phi_{\mathbf{p}}^{\pm}(\mathbf{x})$ are related to the renormalized complex-energy distorted plane waves as follows

$$\lim_{\epsilon\to+0}\int_{R^3}d\mathbf{x}\,\overline{\phi(\mathbf{x})}\int_{R^3}d\mathbf{p}\hat{\psi}(\mathbf{p})[\phi_{\mathbf{p}}^{\pm}(\mathbf{x})-\phi_{\mathbf{p}}^{\pm\epsilon}(\mathbf{x})F_{\mp\epsilon}^{*}(p)]=0, \tag{5.4}$$

for all $\phi\in\mathcal{L}^2\cap\mathcal{L}^1$ and $\hat{\psi}\in\mathcal{D}^{(0)}$ with $\hat{\psi}$ a bounded function with compact support.

Proof: The Green's functions $G(\mathbf{x},\mathbf{x}';\zeta)$ corresponding to $(H-\zeta)^{-1}$ have been shown[8] to satisfy the following equation:

$$G(\mathbf{x},\mathbf{x}';\zeta)=G_0(\mathbf{x},\mathbf{x}';\zeta)-\int_{R^3}d\mathbf{y}\,G_0(\mathbf{x},\mathbf{y};\zeta)\,V(\mathbf{y})G(\mathbf{y},\mathbf{x}';\zeta) \tag{5.5}$$

for almost all $\mathbf{x},\mathbf{x}'\in R^3$. Multiplying (5.5) with $\zeta=E^{(0)}\pm i\epsilon$ by $(\mp i\epsilon)$ and taking the Fourier transform of the result yields (5.3) after interchanging the $\mathbf{x}'$ and $\mathbf{y}$ integrations which is justified by the integrability of $G(\mathbf{y},\mathbf{x}';\zeta)$ with respect to $\mathbf{x}'$ for all $\mathbf{y}$ (Theorem 2.2, Ref. 8).

Due to the existence and integrability of $G(\mathbf{x},\mathbf{x}';\zeta)$ (see Sec. II, Ref. 8) the expansion (4.4) is valid. Thus the equality (5.4) follows immediately from (2.9), (5.1), and (4.4).

The following theorem is stated without proof since the proof is analogous to the corresponding results for three-particle scattering given in Sec. VII.

Theorem 5.2: Suppose that $H=H_0+V$ where V satisfies (β). Furthermore, assume that the cutoff function g_R satisfies (C). Then the S operator has the following expansion:

$$\begin{aligned}\langle\phi\,|S\psi\rangle=&\lim_{\epsilon_1\to+0}\lim_{\epsilon_2\to+0}\lim_{R\to\infty}(-\pi)^{-1}\int_{R^3}d\mathbf{p}\int_{R^3}d\mathbf{p}'\,\overline{\hat{\phi}(\mathbf{p})}\\&\times\hat{\psi}(\mathbf{p}')\overline{F_{+\epsilon_1}^{*}(p)}(\mathbf{p}\,|\,V_RW_{-\epsilon_2}\,|\mathbf{p}')F_{-\epsilon_2}^{*}(p')\\&\times\frac{\epsilon_1}{(E^{(0)}(p)-E^{(0)}(p'))^2+\epsilon_1^2}\end{aligned} \tag{5.6}$$

for $\hat{\phi}\in\mathcal{L}^1(R^3)\cap\mathcal{D}^{(0)}$ and $\hat{\psi}\in\hat{\mathcal{D}}^{(0)}$ where $\hat{\psi}$ is a continuous function having compact support and the off-energy-shell "T matrix" $(\mathbf{p}|\,V_RW_{-\epsilon_2}|\mathbf{p}')$ is given by

$$(\mathbf{p}\,|\,V_RW_{-\epsilon_2}\,|\mathbf{p}')=\int_{R^3}d\mathbf{x}\,\overline{\phi_{\mathbf{p}}(\mathbf{x})}V_R(\mathbf{x})\phi_{\mathbf{p}'}^{+\epsilon}(\mathbf{x}),\quad 0<R<\infty, \tag{5.7}$$

where $V_R=g_RV$.

If, in addition, the expansion (5.1) is valid with the distorted plane waves satisfying (5.2) then the S operator has the following expansion:

$$\begin{aligned}\langle\phi\,|S\psi\rangle=&\lim_{\epsilon_1\to+0}\lim_{R\to\infty}(-\pi)^{-1}\int_{R^3}d\mathbf{p}\int_{R^3}d\mathbf{p}'\,\overline{\hat{\phi}(\mathbf{p})}\hat{\psi}(\mathbf{p}')\overline{F_{+\epsilon_1}^{*}(p)}\\&\times(\mathbf{p}\,|\,V_R\Omega_-\,|\mathbf{p}')\frac{\epsilon_1}{(E^{(0)}(p)-E^{(0)}(p'))^2+\epsilon_1^2},\end{aligned} \tag{5.8}$$

for $\hat{\phi}\in\mathcal{L}^1(R^3)\cap\mathcal{D}^{(0)}$ and $\hat{\psi}\in C_0^\infty(R^3\backslash\{0\})$ where the half-off-shell "T matrix" $(\mathbf{p}|\,V_R\Omega_-|\mathbf{p}')$ is given by

$$(\mathbf{p}\,|\,V_R\Omega_-\,|\mathbf{p}')=\int_{R^3}d\mathbf{x}\,\overline{\phi_{\mathbf{p}}(\mathbf{x})}V_R(\mathbf{x})\phi_{\mathbf{p}'}^{+}(\mathbf{x}), \tag{5.9}$$

for $0<R<\infty$.

Theorem (5.2) shows that it is necessary to replace the off-energy-shell "T matrices" $(\mathbf{p}|\,V_R\Omega_-|\mathbf{p}')$ and $(\mathbf{p}|\,V_RW_{-\epsilon_2}|\mathbf{p}')$ by the following respective renormalized off-energy-shell T matrices:

$$\overline{F_{+\epsilon_1}^{*}(p)}(\mathbf{p}\,|\,V_R\Omega_-\,|\mathbf{p}') \tag{5.10}$$

and

$$\overline{F_{+\epsilon_1}^{*}(p)}(\mathbf{p}\,|\,V_RW_{-\epsilon_2}\,|\mathbf{p}')F_{-\epsilon_2}^{*}(p'). \tag{5.11}$$

The renormalized off-energy-shell T matrix (5.10) can be shown to yield the physical S matrix for the pure Coulomb potential (see the Appendix of Ref. 21).

The existence of the physical S matrix for two-particle scattering has been verified for the pure Coulomb potential[22,23,21] and for a general class of Coulomb-like potentials under the technical assumption (5.2).[21] Thus the existence of the on-energy-shell Coulomb S matrix together with Theorem (5.2) show that the renormalized T matrices (5.10) and (5.11) lead to the physical S matrix.

VI. THREE PARTICLE OFF-ENERGY-SHELL EQUATIONS

The complex-energy distorted plane waves corresponding to three-particle scattering via two-body potentials which satisfy (β) are shown to satisfy the iterated Weinberg–Van Winter equations and twice iterated Faddeev equations.

The following theorem provides the existence of solutions to the iterated off-energy-shell Weinberg–Van Winter equations.

Theorem 6.1: Suppose that $H=H_0+\sum_{i<j}V_{ij}$ where each V_{ij} satisfies (β). Furthermore, assume that the α-channel bound state $\chi^{(\alpha)}(\mathbf{x})$ satisfies $\chi^{(\alpha)}\in\mathcal{L}^\infty(R^3)$. Then the α-channel complex-energy distorted plane waves $\phi_{p_{n_\alpha-1}}^{(\alpha)\pm\epsilon}(x)$ satisfy the iterated Weinberg–Van Winter equations,

$$\begin{aligned}\phi_{p_{n_\alpha-1}}^{(\alpha)\pm\epsilon}(x)=&\tilde{\phi}_{p_{n_\alpha-1}}^{(\alpha)\pm\epsilon}(x)+\int_{R^6}dy\,G_{cl}(x,y;E^{(\alpha)}\pm i\epsilon)\tilde{\phi}_{p_{n_\alpha-1}}^{(\alpha)\pm\epsilon}(y)\\&+\int_{R^6\times R^6}dy\,dz\,G_{cl}(x,y;E^{(\alpha)}\pm i\epsilon)\\&\times G_{cl}(y,z;E^{(\alpha)}\pm i\epsilon)\phi_{p_{n_\alpha-1}}^{(\alpha)\pm\epsilon}(z),\end{aligned} \tag{6.1}$$

for almost all $x\in R^6$, $p_{n_\alpha-1}\in R^{3(n_\alpha-1)}$ and each $\epsilon>0$, where

$$\tilde{\phi}_{p_{n_\alpha-1}}^{(\alpha)\pm\epsilon}(x)=(\mp i\epsilon)\int_{R^6}dy\,H(x,y;E^{(\alpha)}\pm i\epsilon)\phi_{p_{n_\alpha-1}}(y)\chi^{(\alpha)}(y), \tag{6.2}$$

$$H(x,y;E^{(\alpha)}\pm i\epsilon)=\sum_{i<j}G_{ij}(x,y;E^{(\alpha)}\pm i\epsilon)-2G_0(x,y;E^{(\alpha)}\pm i\epsilon), \qquad (6.3)$$

and

$$G_{cl}(x,y;E^{(\alpha)}\pm i\epsilon)=\sum_{\substack{i<j\\ i\neq k,\, j\neq k}}[G_0(x,y;E^{(\alpha)}\pm i\epsilon)-G_{ij}(x,y;E^{(\alpha)}\pm i\epsilon)]\times[V_{ik}(y)+V_{jk}(y)], \qquad (6.4)$$

with $G_{ij}(x,y;\zeta)$ and $G_0(x,y;\zeta)$ the Green's functions corresponding to $(H_{ij}-\zeta)^{-1}$, $H_{ij}=H_0+V_{ij}$, and $(H_0-\zeta)^{-1}$, respectively.

Proof: The Green's functions satisfy (Theorem 5.2, Ref. 8) the iterated Weinberg–Van Winter equations

$$\begin{aligned}G(x,y;E^{(\alpha)}\pm i\epsilon)&=H(x,y;E^{(\alpha)}\pm i\epsilon)+\int_{R^6}dz\,G_{cl}(x,z;E^{(\alpha)}\pm i\epsilon)\\&\quad\times H(z,y;E^{(\alpha)}\pm i\epsilon)+\int_{R^6\times R^6}dz\,dw\,G_{cl}(x,z;E^{(\alpha)}\pm i\epsilon)\\&\quad\times G_{cl}(z,w;E^{(\alpha)}\pm i\epsilon)G(w,y;E^{(\alpha)}\pm i\epsilon).\end{aligned} \qquad (6.5)$$

Multiplying (6.5) by $(\mp i\epsilon)\phi_{p_{n_\alpha-1}}(y)\chi^{(\alpha)}(y)$ and integrating the result with respect to y yields (6.1) if the appropriate integrals can be interchanged. The interchange of the y and z integrals corresponding to the second term on the right side of (6.5) can be justified by Proposition 3.2 of Ref. 8. The interchange of the y and z,w integrals corresponding to the last term in (6.5) can be justified by Theorem 5.6 of Ref. 8 since

$$\int_{R^6}dy\,|G_{cl}(x,y;E^{(\alpha)}\mp i\epsilon)G_{cl}(y,\cdot\,;E^{(\alpha)}\mp i\epsilon)|\in\mathcal{L}^1(R^6)\cap\mathcal{L}^2(R^6), \qquad (6.6)$$

which can be verified via Proposition 3.2 and Theorem 5.2 of Ref. 8.

It has been shown (Theorem 5.3, Ref. 8) for two-body Coulomb-like potentials which satisfy (β) that there exist functions $H_{ij}(x,y;\zeta)$ and $K_{ij}(x,y;\zeta)$, $i<j$, $i,j=1,2,3$, $\epsilon>0$, which are related to the full Green's functions $G(x,y;\zeta)$ as follows:

$$G(x,y;\zeta)=G_0(x,y;\zeta)-\sum_{i<j}H_{ij}(x,y;\zeta)=-\sum_{i<j}K_{ij}(x,y;\zeta) \qquad (6.7)$$

for almost all $x,y\in R^6$ and $\zeta\notin S=\sigma(H)\cup_{i<j}\sigma(H_{ij})$ where $\sigma(K)$ denotes the spectrum of K. In addition the kernels $H_{ij}(x,y;\zeta)$ and $K_{ij}(x,y;\zeta)$ satisfy the following twice iterated Faddeev equations written in matrix form:

$$\begin{pmatrix}H_{12}(x,y;\zeta)\\H_{13}(x,y;\zeta)\\H_{23}(x,y;\zeta)\end{pmatrix}=[\Pi-A+A^2-A^3]\begin{pmatrix}G_0(\cdot,y;\zeta)-G_{12}(\cdot,y;\zeta)\\G_0(\cdot,y;\zeta)-G_{13}(\cdot,y;\zeta)\\G_0(\cdot,y;\zeta)-G_{23}(\cdot,y;\zeta)\end{pmatrix}+A^4\begin{pmatrix}H_{12}(\cdot,y;\zeta)\\H_{13}(\cdot,y;\zeta)\\H_{23}(\cdot,y;\zeta)\end{pmatrix} \qquad (6.8)$$

and

$$\begin{pmatrix}K_{12}(x,y;\zeta)\\K_{13}(x,y;\zeta)\\K_{23}(x,y;\zeta)\end{pmatrix}=[\Pi-A+A^2-A^3]\begin{pmatrix}0\\0\\-G_{23}(\cdot,y;\zeta)\end{pmatrix}+A^4\begin{pmatrix}K_{12}(\cdot,y;\zeta)\\K_{13}(\cdot,y;\zeta)\\K_{23}(\cdot,y;\zeta)\end{pmatrix}, \qquad (6.9)$$

where Π is the 3×3 identity matrix and A is a matrix integral operator which acts on column vectors as follows:

$$A\begin{pmatrix}\psi_1(\cdot)\\\psi_2(\cdot)\\\psi_3(\cdot)\end{pmatrix}=\begin{pmatrix}\int_{R^6}dx'\,G_{12}(x,x';\zeta)V_{12}(x')[\psi_2(x')+\psi_3(x')]\\\int_{R^6}dx'\,G_{13}(x,x';\zeta)V_{13}(x')[\psi_1(x')+\psi_3(x')]\\\int_{R^6}dx'\,G_{23}(x,x';\zeta)V_{23}(x')[\psi_1(x')+\psi_2(x')]\end{pmatrix}. \qquad (6.10)$$

We now define the following functions:

$$\tilde{\phi}^{(0)\pm\epsilon}_{p,ij}(x)=(\mp i\epsilon)\int_{R^6}dy\,H_{ij}(x,y;E^{(0)}\pm i\epsilon)\phi_p(y),\quad i<j,\quad i,j=1,2,3, \qquad (6.11)$$

and for the channel α involving a composite particle

$$\tilde{\phi}^{(\alpha)\pm\epsilon}_{p_{n_\alpha-1},ij}(x)=(\mp i\epsilon)\int_{R^6}dy\,K_{ij}(x,y;E^{(\alpha)}\pm i\epsilon)\times\phi_{p_{n_\alpha-1}}(y)\chi^{(\alpha)}(y),\quad i<j,\quad i,j=1,2,3. \qquad (6.12)$$

The functions $\tilde{\phi}^{(\alpha)\pm\epsilon}_{p_{n_\alpha-1},ij}(x)$ exist under the assumption $\chi^{(\alpha)}\in\mathcal{L}^\infty$ since $H_{ij}(x,y,\zeta)$ and $K_{ij}(x,y;\zeta)$ are integrable (Theorem 5.7, Ref. 8). Multiplying $H_{ij}(x,y;E^{(0)}\pm i\epsilon)$ by $(\mp i\epsilon)\phi_p(y)$ and $K_{ij}(x,y;E^{(\alpha)}\pm i\epsilon)$ by $(\mp i\epsilon)\phi_{p_{n_\alpha-1}}(y)\times\chi^{(\alpha)}(y)$ in (6.8) and (6.9) respectively and integrating the result with respect to y it is straightforward to show via Proposition 3.2 and Theorem 5.7 of Ref. 8 that the various integrals can be interchanged and thus (6.8) and (6.9) lead to equations for the functions (6.11) and (6.12), respectively.

The above results are summarized in the following theorem.

Theorem 6.2: Suppose that $H=H_0+\sum_{i<j}V_{ij}$ where each V_{ij} satisfies (β). Then:

(a) There exists functions $\tilde{\phi}^{(0)\pm\epsilon}_{p,ij}(x)$, $i>j$, $i,j=1,2,3$, defined by (6.11) which are related to the complex-energy distorted plane waves $\phi^{(0)\pm\epsilon}_p(x)$ as follows

$$\phi^{(0)\pm\epsilon}_p(x)=\phi_p(x)-\sum_{i<j}\tilde{\phi}^{(0)\pm\epsilon}_{p,ij}(x) \qquad (6.13)$$

for almost all $x\in R^6$, $p\in R^6$ and each $\epsilon>0$ and which satisfy the twice iterated Faddeev equations,

$$\begin{pmatrix}\tilde{\phi}^{(0)\pm\epsilon}_{p,12}(x)\\\tilde{\phi}^{(0)\pm\epsilon}_{p,13}(x)\\\tilde{\phi}^{(0)\pm\epsilon}_{p,23}(x)\end{pmatrix}=[\Pi-A+A^2-A^3]\times\begin{pmatrix}\int_{R^6}dy\,G_{12}(\cdot,y;E^{(0)}\pm i\epsilon)V_{12}(y)\phi_p(y)\\\int_{R^6}dy\,G_{13}(\cdot,y;E^{(0)}\pm i\epsilon)V_{13}(y)\phi_p(y)\\\int_{R^6}dy\,G_{23}(\cdot,y;E^{(0)}\pm i\epsilon)V_{23}(y)\phi_p(y)\end{pmatrix}+A^4\begin{pmatrix}\tilde{\phi}^{(0)}_{p,12}(\cdot)\\\tilde{\phi}^{(0)}_{p,13}(\cdot)\\\tilde{\phi}^{(0)}_{p,23}(\cdot)\end{pmatrix} \qquad (6.14)$$

for almost all $x \in R^6$, $p \in R^6$, and each $\epsilon > 0$.

(b) If in addition $\chi^{(\alpha)} \in \mathcal{L}^\infty$ then there exists functions $\tilde{\phi}^{(\alpha)\pm\epsilon}_{p_{n_\alpha-1},ij}(x)$, $i<j$, $i,j=1,2,3$, defined by (6.12) which are related to the complex-energy distorted plane waves $\phi^{(\alpha)\pm\epsilon}_{p_{n_\alpha-1}}(x)$ as follows:

$$\phi^{(\alpha)\pm\epsilon}_{p_{n_\alpha-1}}(x) = -\sum_{i<j} \tilde{\phi}^{(\alpha)\pm\epsilon}_{p_{n_\alpha-1},ij}(x) \tag{6.15}$$

for almost all $x \in R^6$, $p_{n_\alpha-1} \in R^{3(n_\alpha-1)}$ and each $\epsilon > 0$ and which satisfy the twice iterated Faddeev equations,

$$\begin{pmatrix} \tilde{\phi}^{(\alpha)\pm\epsilon}_{p_{n_\alpha-1},12}(x) \\ \tilde{\phi}^{(\alpha)\pm\epsilon}_{p_{n_\alpha-1},13}(x) \\ \tilde{\phi}^{(\alpha)\pm\epsilon}_{p_{n_\alpha-1},23}(x) \end{pmatrix} = [\Pi - A + A^3 - A^3] \begin{pmatrix} 0 \\ 0 \\ -\phi_{p_{n_\alpha-1}}(\cdot)\chi^{(\alpha)}(\cdot) \end{pmatrix} + A^4 \begin{pmatrix} \tilde{\phi}^{(\alpha)\pm\epsilon}_{p_{n_\alpha-1},12}(\cdot) \\ \tilde{\phi}^{(\alpha)\pm\epsilon}_{p_{n_\alpha-1},13}(\cdot) \\ \tilde{\phi}^{(\alpha)\pm\epsilon}_{p_{n_\alpha-1},23}(\cdot) \end{pmatrix} \tag{6.16}$$

for almost all $x \in R^6$, $p_{n_\alpha-1} \in R^{3(n_\alpha-1)}$ and each $\epsilon > 0$.

Remark: To obtain an explicit expression for the kernel of the three-particle equations given in Theorems 6.1 and 6.2 requires a knowledge of the separable Green's functions $G_{ij}(x,y;\zeta)$. In the case of the pure Coulomb potential integral representations of the separable Green's functions can be derived. The existence of such integral representations follows from Hostler's representation[24] of the two-particle Green's functions corresponding to the pure Coulomb potential together with the definition (see Sec. 1.5.2, Ref. 25 and Sec. 3, Ref. 8) of the separable Green's function.

VII. THREE PARTICLE OFF-SHELL COULOMB SCATTERING

The existence of the limit to real energies of the renormalized complex-energy distorted plane waves and renormalized off-energy-shell T matrices for three-particle scattering via a general class of two-body Coulomb-like potentials is verified.

The following theorem relates the α-channel renormalized complex-energy distorted plane waves $\phi^{(\alpha)\pm\epsilon}_{p_{n_\alpha-1}}(x)F^{(\alpha)*}_{\mp\epsilon}(p_{n_\alpha-1})$ to the renormalized wave operators.

Theorem 7.1: Suppose that $H = H_0 + \sum_{i<j} V_{ij}$ where each V_{ij} satisfies $(\mathcal{B})$. Then the renormalized complex-energy distorted plane waves $\phi^{(\alpha)\mp\epsilon}_{p_{n_\alpha-1}}(x)F^{(\alpha)*}_{\pm\epsilon}(p_{n_\alpha-1})$ are related to the renormalized wave operators $\Omega^{(\alpha)}_\pm$ as follows:

$$\begin{aligned}\int_{R^6} dx\, \overline{\phi(x)}(\Omega^{(\alpha)}_\pm \psi)(x) &= \lim_{\epsilon\to+0} \int_{R^6} dx\, \overline{\phi(x)} \int_{R^{3(n_\alpha-1)}} dp_{n_\alpha-1} \\ &\quad \times \phi^{(\alpha)\mp\epsilon}_{p_{n_\alpha-1}}(x) F^{(\alpha)*}_{\pm\epsilon}(p_{n_\alpha-1}) \hat{\psi}_1(p_{n_\alpha-1}),\end{aligned} \tag{7.1}$$

where $\phi \in \mathcal{L}^2 \cap \mathcal{L}^1$ and $\psi = \psi_1 \chi^{(\alpha)} \in \mathcal{D}^{(\alpha)}$, $\chi^{(\alpha)} \in \mathcal{L}^\infty(R^3)$, and $\hat{\psi}_1$ is a bounded function of compact support.

Proof: The expansion (4.4) is valid since each V_{ij} satisfies $(\mathcal{B})$. Thus (7.1) follows from (2.9).

The following theorem relates the "renormalized half-off-shell T matrix" for three-particle scattering to the S operator.

Theorem 7.2: Suppose that $H = H_0 + \sum_{i<j} V_{ij}$, where each V_{ij} satisfies $(\mathcal{B})$. Furthermore, assume that the renormalized wave operators have the expansion (5.1) with the physical distorted plane wave $\phi^{(\alpha)+}_{p_{n_\alpha-1}}(x)$ satisfying (5.2). In addition assume $g^{(\beta)}_R$ satisfies (C) and $\chi^{(\beta)} \in \mathcal{L}^\infty$. Then the S operator $S_{\alpha\beta}$ has the following expansion:

$$\begin{aligned}\langle \phi | S_{\alpha\beta}\psi\rangle &= \lim_{\epsilon\to+0} \lim_{R\to+\infty} (-\pi)^{-1} \int_{R^{3(n_\beta-1)}\times R^{3(n_\alpha-1)}} dp_{n_\beta-1} \\ &\quad \times dp'_{n_\alpha-1}\, \overline{\hat{\phi}(p_{n_\beta-1})}\hat{\psi}_1(p'_{n_\alpha-1}) \overline{F^{(\beta)*}_{+\epsilon}(p_{n_\beta-1})} \\ &\quad \times (p_{n_\beta-1}; E^{(\beta)}_{\text{int}} | V^{(\beta)}_R \Omega^{(\alpha)}_- | p'_{n_\alpha-1}; E^{(\alpha)}_{\text{int}}) \\ &\quad \times \frac{\epsilon}{(E^{(\alpha)}(p'_{n_\alpha-1}) - E^{(\beta)}(p_{n_\beta-1}))^2 + \epsilon^2}\end{aligned} \tag{7.2}$$

for $\phi = \phi_1\chi^{(\beta)} \in \mathcal{D}^{(\beta)}$, $F^{(\beta)*}_{+\epsilon}\hat{\phi} \in \mathcal{L}^1(R^{3(n_\beta-1)})$, $\epsilon > 0$, $\psi = \psi_1\chi^{(\alpha)} \in \mathcal{D}(\mathcal{H}_\alpha) \cap \mathcal{H}^{(\alpha)}$, $\hat{\psi}_1 \in C_0^\infty(R^{3(n_\alpha-1)}\setminus\{0\})$ where the "half-of-shell T matrix" $(p_{n_\beta-1}; E^{(\beta)}_{\text{int}} | V^{(\beta)}_R \Omega^{(\alpha)}_- | p'_{n_\alpha-1}; E^{(\alpha)}_{\text{int}})$ is given by

$$\begin{aligned}&(p_{n_\beta-1}; E^{(\beta)}_{\text{int}} | V^{(\beta)}_R \Omega^{(\alpha)}_- | p'_{n_\alpha-1}; E^{(\alpha)}_{\text{int}}) \\ &\quad = \int_{R^6} dx\, \overline{\phi_{p_{n_\beta-1}}(x)\chi^{(\beta)}(x)} V^{(\beta)}_R(x) \phi^{(\alpha)+}_{p'_{n_\alpha-1}}(x),\end{aligned} \tag{7.3}$$

for each $0 < R < \infty$.

Proof: The representation (3.4) together with (5.1) yields after an appropriate interchange of integrals

$$\begin{aligned}\langle \phi | S_{\alpha\beta}\psi\rangle &= \lim_{\epsilon\to+0} \lim_{R\to+\infty} (-\pi)^{-1} \int_{-\infty}^{+\infty} d_\lambda \int_{R^{3(n_\beta-1)}} dp_{n_\beta-1} \\ &\quad \times \chi_{(-\infty,\lambda]}(E^{(\beta)}) \int_{R^{3(n_\alpha-1)}} dp'_{n_\alpha-1}\, \overline{\hat{\phi}_1(p_{n_\beta-1})}\hat{\psi}_1(p'_{n_\alpha-1}) \\ &\quad \times \overline{F^{(\beta)*}_{+\epsilon}(p_{n_\beta-1})} (p_{n_\beta-1}; E^{(\beta)}_{\text{int}} | V^{(\beta)}_R \Omega^{(\alpha)}_- | p'_{n_\alpha-1}; E^{(\alpha)}_{\text{int}}) \\ &\quad \times \frac{\epsilon}{(E^{(\alpha)}(p'_{n_\alpha-1}) - \lambda)^2 + \epsilon^2},\end{aligned} \tag{7.4}$$

for ϕ and ψ specified in the theorem. It is straightforward to verify that the expression

$$\frac{\epsilon}{(E^{(\alpha)}(p'_{n_\alpha-1}) - \lambda)^2 + \epsilon^2}$$

in (7.4) can be replaced by

$$\frac{\epsilon}{(E^{(\alpha)}(p'_{n_\alpha-1}) - E^{(\beta)}(p_{n_\beta-1}))^2 + \epsilon^2}$$

and the spectral integral can be performed to obtain (7.2).

The following theorem relates the "renormalized off-energy-shell T matrix" for three-particle scattering to the S operator.

Theorem 7.3: Suppose that $H = H_0 + \sum_{i<j} V_{ij}$ where each V_{ij} satisfies $(\mathcal{B})$. Furthermore, assume $g^{(\beta)}_R$ satisfies (C) and $\chi^{(\beta)} \in \mathcal{L}^1(R^3) \cap \mathcal{L}^2(R^3)$, $x^{(\alpha)} \in \mathcal{L}^\infty(R^3) \cap \mathcal{L}^2(R^3)$. Then the S operator $S_{\alpha\beta}$ has the following expansion:

$$\begin{aligned}\langle \phi | S_{\alpha\beta}\psi\rangle &= \lim_{\epsilon_1\to+0} \lim_{\epsilon_2\to+0} \lim_{R\to+\infty} (-\pi)^{-1} \int_{R^{3(n_\beta-1)}\times R^{3(n_\alpha-1)}} dp_{n_\beta-1} \\ &\quad \times dp'_{n_\alpha-1}\, \overline{\hat{\phi}_1(p_{n_\beta-1})}\hat{\psi}_1(p'_{n_\alpha-1}) \overline{F^{(\beta)*}_{+\epsilon_1}(p_{n_\beta-1})} \\ &\quad \times (p_{n_\beta-1}; E^{(\beta)}_{\text{int}} | V^{(\beta)}_R W^{(\alpha)}_{-\epsilon_2} | p'_{n_\alpha-1}; E^{(\alpha)}_{\text{int}}) \overline{F^{(\alpha)*}_{-\epsilon_2}(p'_{n_\alpha-1})} \\ &\quad \times \frac{\epsilon_1}{(E^{(\alpha)}(p'_{n_\alpha-1}) - E^{(\beta)}(p_{n_\beta-1}))^2 + \epsilon_1^2},\end{aligned} \tag{7.5}$$

for $\phi=\phi_1\chi^{(\beta)}\in\mathcal{D}^{(\beta)}$, $F^{(\beta)*}_{+\epsilon_1}\hat{\phi}_1\in\mathcal{L}^1(R^{3(n_\beta-1)})$, for each $\epsilon_1>0$, $\psi=\psi_1\chi^{(\alpha)}\in\hat{\mathcal{D}}^{(\alpha)}$, $F^{(\alpha)*}_{-\epsilon_2}\hat{\psi}_1\in\mathcal{L}^1(R^{3(n_\alpha-1)})$, for each $\epsilon_2>0$, and $\hat{\psi}_1$ a continuous function of compact support where the "off-energy-shell T matrix" $(p_{n_\beta-1};E^{(\beta)}_{\rm int}\times|V^{(\beta)}_R W^{(\alpha)}_{-\epsilon_2}|p'_{n_\alpha-1};E^{(\alpha)}_{\rm int})$ is given by

$$(p_{n_\beta-1};E^{(\beta)}_{\rm int}|V^{(\beta)}_R W^{(\alpha)}_{-\epsilon_2}|p'_{n_\alpha-1};E^{(\alpha)}_{\rm int})=\int_{R^6}dx\,\overline{\phi_{p_{n_\beta-1}}(x)\chi^{(\beta)}}(x)V^{(\beta)}_R(x)\,\phi^{(\alpha)+\epsilon_2}_{p'_{n_\alpha-1}}(x) \tag{7.6}$$

for $0<R<\infty$.

Proof: In analogy with the proof of Theorem 7.2 the expansion (7.5) follows from the representation (3.5) together with (4.4) if the following equality is valid:

$$\begin{aligned}&\int_{R^6}dx\int_{R^{3(n_\beta-1)}}dp_{n_\beta-1}\chi_{(-\infty,\lambda]}(E^{(\beta)})\\&\times\overline{\phi_{p_{n_\beta-1}}(x)\chi^{(\beta)}(x)F^{(\beta)*}_{+\epsilon_1}(p_{n_\beta-1})\hat{\phi}_1(p_{n_\beta-1})}\\&\times V^{(\beta)}_R(x)\int_{R^{3(n_\alpha-1)}}dp'_{n_\alpha-1}\phi^{(\alpha)+\epsilon_2}_{p'_{n_\alpha-1}}(x)F^{(\alpha)*}_{-\epsilon_2}(p'_{n_\alpha-1})\\&\times\frac{\epsilon_1}{(E^{(\alpha)}(p'_{n_\alpha-1})-\lambda)^2+\epsilon_1^2}\hat{\psi}_1(p'_{n_\alpha-1})\\&=\int_{R^{3(n_\beta-1)}\times R^{3(n_\alpha-1)}}dp_{n_\beta-1}\,dp'_{n_\alpha-1}\,\overline{\hat{\phi}_1(p_{n_\beta-1})}\hat{\psi}_1(p'_{n_\alpha-1})\\&\times\overline{F^{(\beta)*}_{+\epsilon_1}(p_{n_\beta-1})}F^{(\alpha)*}_{-\epsilon_2}(p'_{n_\alpha-1})\chi_{(-\infty,\lambda]}(E^{(\beta)})\\&\times\frac{\epsilon_1}{(E^{(\alpha)}(p'_{n_\alpha-1})-\lambda)^2+\epsilon_1^2}\\&\times(p_{n_\beta-1};E^{(\beta)}_{\rm int}|V^{(\beta)}_R W^{(\alpha)}_{-\epsilon_2}|p'_{n_\alpha-1};E^{(\alpha)}_{\rm int})\end{aligned} \tag{7.7}$$

for ϕ and ψ of the form specified in the theorem and $\epsilon_1>0$, $\epsilon_2>0$, $0<R<\infty$. The interchange of integrals required to verify (7.7) can be justified via the following inequality (see Theorem 5.6 of Ref. 8):

$$\int_{R^6}dx\,|\chi^{(\beta)}(x)V^{(\beta)}_R(x)\phi^{(\alpha)+\epsilon_2}_{p'_{n_\alpha-1}}(x)|\le D(R,\operatorname{supp}\hat{\psi}_1,\epsilon_2), \tag{7.8}$$

where $D(R,\operatorname{supp}\hat{\psi}_1,\epsilon_2)$ is a constant depending on R, support of $\hat{\psi}_1$ and ϵ_2.

VIII. CONCLUDING REMARKS

In recent years several stationary Coulomb scattering theories have been derived[26–29] each involving a different off-energy-shell formalism. These formalisms lead to modified Lippmann–Schwinger equations in which the pure Coulomb potential has been replaced by momentum dependent potentials which decrease faster for large particle separation than the Coulomb potential. Since the "effective potentials" appearing in these modified Lippmann–Schwinger equations do not have the symmetries of the pure Coulomb potential it does not seem possible to obtain closed form solutions as in the case of the two-particle off-energy-shell Lippmann–Schwinger equations. In contrast the renormalized off-energy-shell formalism derived in this paper is based on the solutions of the off-energy-shell equations of short-range scattering theory. Thus, as discussed at the end of Sec. VI, explicit expressions for the kernels of the Weinberg–Van Winter and Faddeev equations for three-particle scattering via pure two-body Coulomb potentials can be derived.

The results of this paper suggest that the solutions of the Lippmann–Schwinger, Weinberg–Van Winter and Faddeev equations must possess, in the limit to real energies, an oscillatory behavior which cancels the stationary renormalization terms appearing in the definition of the renormalized off-energy-shell formalism. Thus any approximation procedure which is based on the off-energy-shell Lippmann–Schwinger and Faddeev equations must take into account the existence of such divergent phase factors. For example, the usual perturbation series based on the off-energy-shell equations must be replaced[30] by "renormalized" perturbation series which take into account the stationary renormalization terms.

*Supported by the National Research Council of Canada.

[1] L. D. Faddeev, "Mathematical aspects of the three-body problem in the quantum scattering theory," Israel Program for Scientific Translations, Jerusalem (1965).
[2] S. Okubo and D. Feldman, Phys. Rev. 117, 292 (1960).
[3] J. Schwinger, J. Math. Phys. 5, 1606 (1964).
[4] J. C. Y. Chen and A. C. Chen, *Advances in Atomic and Molecular Physics*, edited by D. R. Bates and I. Estermann (Academic, New York, 1972), Vol. 8, p. 71.
[5] J. C. Y. Chen, Case Stud. At. Phys. 3, 305 (1973).
[6] J. Zorbas, Nuovo Cimento Lett. 10, 121 (1974). Note that the signs in the gamma and exponential functions of Eq. 16 should be interchanged.
[7] J. Zorbas, J. Math. Phys. 17, 498 (1976).
[8] J. Zorbas, "Two and Three Particle Coulomb Green's Functions," preprint, U. of British Columbia (1976).
[9] E. Prugovečki, *Quantum Mechanics in Hilbert Space* (Academic, New York, 1971).
[10] W. O. Amrein, V. Georgescu, and J. M. Jauch, Helv. Phys. Acta 43, 313 (1970).
[11] D. B. Pearson, Nuovo Cimento 2A, 853 (1971).
[12] J. D. Dollard, J. Math. Phys. 5, 729 (1964).
[13] J. D. Dollard, "Non-relativistic time-dependent scattering theory and the Coulomb interaction," thesis, Princeton (1963).
[14] J. D. Dollard, Rocky Mount. J. Math. 1, 5 (1971).
[15] T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1966).
[16] E. Nelson, J. Math. Phys. 5, 332 (1964).
[17] C. Chandler and A. G. Gibson, Helv. Phys. Acta 45, 734 (1972).
[18] E. Hille and R. S. Phillips, "Functional analysis and semi-groups," Am. Math. Soc. Coll. Publ. 31 (1957).
[19] J. D. Dollard, J. Math. Phys. 7, 802 (1966).
[20] J. C. Guillot and K. Zizi, in *Scattering Theory in Mathematical Physics*, edited by J. A. Lavita and J. P. Marchand (Reidel, Dordrecht-Holland, 1974).
[21] J. Zorbas, Rep. Math. Phys. 9, 309 (1976).
[22] I. W. Herbst, Commun. Math. Phys. 35, 181 (1974).
[23] A. G. Gibson and C. Chandler, J. Math. Phys. 15, 1366 (1974).
[24] L. Hostler, J. Math. Phys. 5, 591 (1964).
[25] C. Van Winter, Mat. Fys. Skr. Dan. Vid. Selsk. 2, No. 8 (1964).
[26] D. F. Freeman and J. Nuttall, J. Math. Phys. 14, 1883 (1973).
[27] D. Masson and E. Prugovečki, J. Math. Phys. 17, 297 (1976).
[28] E. Prugovečki and J. Zorbas, J. Math. Phys. 14, 1398 (1973); 15, 268 (1974) (erratum).
[29] E. Prugovečki and J. Zorbas, Nucl. Phys. A 213, 268 (1973).
[30] J. Zorbas, "Perturbation theory for Coulomb scattering," preprint U. of British Columbia (1976).

Schrödinger equation with Yukawa potential, a differential equation with two singular points

Wolfgang Bühring

Physikalisches Institut der Universität Heidelberg, 6900 Heidelberg, Germany
(Received 22 September 1976)

The Schrödinger radial equation with Yukawa potential is treated analytically by means of a double contour integral representation for the solution. Standard solutions are defined relative to each of the singular points of the differential equation. Convergent expressions are obtained for the connection coefficients which occur in the linear relations persisting between any three of the standard solutions. These expressions are double series the terms of which are hypergeometric functions multiplied by factors which can be calculated recursively. As an application, the expression for the S matrix, which is simply related to the connection coefficients, is considered with regard to its convergence properties.

1. INTRODUCTION

An old mathematical problem of applied quantum mechanics is the Schrödinger radial equation with Yukawa potential, as has been recently pointed out again by Danos[1] on the occasion of an invited SIAM conference talk. While of the two singular points of this linear differential equation the regular singular point at the origin is unproblematic, it is the singular point at infinity which causes the difficulty. In this context a recent paper by Shere[2] on multiple asymptotic series is very important, for it gives new insight into the behavior near infinity of the solutions of our differential equation. This information stimulates us to find a kernel which is suitable for a double integral representation of the solution. By means of the integral representation we are able to solve the connection problem, i.e., to determine the coefficients in the linear relations persisting between the solutions defined relative to the different singular points of the differential equation. The method is in principle similar to but much more complicated than the classical method of finding the relation between Bessel and Hankel functions by means of a Laplace-type contour integral representation for the solution of Bessel's differential equation.

The present paper is divided into eleven main sections. By a simple transformation in Sec. 2, a modified differential equation is introduced which contains only three parameters rather than four. For this equation, standard solutions relative to the origin are defined in Sec. 3 in analogy with Bessel functions (of the first kind). The general integral representation for the solution is derived in Sec. 4. By means of the integral representation we are able, in Sec. 5, to define standard solutions relative to the singular point at infinity in analogy with Hankel functions. The asymptotic expansions[2] of these solutions are verified. The connection problem is solved in Sec. 6. The desired connection coefficients are obtained in the form of convergent double series, the terms of which are hypergeometric functions multiplied by factors which can be calculated recursively. Wronskian and circuit relations for the solutions are considered in Sec. 7, and Neumann-type solutions are introduced in Sec. 8. Section 9 is devoted to the exceptional case where one parameter of the differential equation, the angular momentum quantum number, is equal to an integer or to half an odd integer. Another exceptional case occurring when one of the other parameters of the differential equation, the exponent of the Yukawa term, happens to have any value of a particular discrete set is treated in Sec. 10. As an application, the significance of our results for the potential scattering problem is shortly explained in Sec. 11, in particular the convergence properties of the expression for the S matrix are considered. Also, the first Born approximation value of the S matrix is verified.

2. REDUCTION OF THE NUMBER OF PARAMETERS OF THE DIFFERENTIAL EQUATION

The Schrödinger radial equation with Yukawa potential may be written

$$r^2 f'' + 2rf' + [k^2 r^2 - l(l+1) - gr\exp(-\mu r)]f(r) = 0. \tag{2.1}$$

This differential equation, which depends on the four parameters k, l, g, μ, may be reduced to an equation with only three independent parameters l, G, β. For by introducing a new independent variable

$$z = kr, \tag{2.2a}$$

one obtains, with

$$f(r) = y(z), \tag{2.2b}$$

$$g/k = G, \tag{2.2c}$$

$$\mu/k = \beta, \tag{2.2d}$$

the differential equation

$$z^2 y'' + 2zy' + [z^2 - l(l+1) - Gz\exp(-\beta z)]y(z) = 0. \tag{2.3}$$

It is this z equation which we will treat in the following sections.

3. SOLUTIONS RELATIVE TO THE ORIGIN: BESSEL-TYPE SOLUTIONS OR SOLUTIONS OF THE FIRST KIND

Relative to the regular singular point at the origin we may immediately define, in analogy with Bessel functions (of the first kind), one type of standard solutions $jy(L; z)$ by

$$C(L)jy(L; z) = [\Gamma(\tfrac{3}{2})/\Gamma(\tfrac{3}{2}+L)]2^{-L}z^L \sum_{n=0}^{\infty} w_n(L)z^n, \tag{3.1}$$

where

$$L = l \tag{3.2a}$$

or

$$L = -l - 1 \tag{3.2b}$$

and [with $w_{-1}(L) = 0$]

$$w_0(L) = 1, \tag{3.3a}$$

$$w_n(L) = [-w_{n-2}(L) + G\sum_{m=0}^{n-1}(1/m!)(-\beta)^m w_{n-m-1}(L)]/ [n(n+2L+1)]. \tag{3.3b}$$

The normalization has been so chosen that the right-hand side of Eq. (3.1) reduces to the appropriate spherical Bessel function when $G \to 0$. The constant $C(L)$ is another normalization factor, depending on G, but such that $C(L) \to 1$ when $G \to 0$. It will be completely defined later by Eq. (7.1) in Sec. 7. Apart from the normalization factors, solutions of this type have been used earlier.[3]

If l is neither an integer nor half an odd integer, the solutions $jy(l; z)$ and $jy(-l-1; z)$ are well defined and linearly independent. The case when $2l$ is an integer requires special attention and will be considered later in Sec. 9.

4. INTEGRAL REPRESENTATION OF THE SOLUTION

A. Type of integral representation

Standard solutions relative to the singular point at infinity may be defined by their asymptotic expansions in a suitable sectorial neighborhood of infinity, but in order to find the connection between these solutions with known asymptotic behavior and the solutions (3.1) we need a convergent representation, i.e., an appropriate integral representation. From the work of Shere[2] we know that the asymptotic expansion is an expansion with respect to the sequence

$$z^{-n}\exp(-m\beta z), \quad n = 0, 1, 2, \cdots, \quad m = 0, 1, 2, \cdots.$$

From this fact we get an idea as to the integral kernel needed and are led to consider the double integral representation

$$y(z) = z^{-\lambda-1}(2\pi i)^{-2}\int_{C_t}\int_{C_s} K(z; s, t)v(s, t)\,ds\,dt \tag{4.1}$$

with the kernel

$$K(z; s, t) = \exp[zt + \exp(-\beta z)s]. \tag{4.2}$$

The contours C_s and C_t in the s and t plane, respectively, are not to depend on z. The factor $z^{-\lambda-1}$ has been extracted in order that we may gain some freedom. The still arbitrary parameter λ will be specified later according to our convenience.

Proceeding in a similar way as Ince[4] we substitute the double integral (4.1) into the differential equation (2.3), interchange the z differentiation with the s and t integrations, and obtain the condition

$$\int_{C_t}\int_{C_s}\sum_{n=0}^{2}\sum_{m=0}^{2} u_{mn}(s, t)\exp(-m\beta z)\, z^n \times K(z; s, t)\, v(s, t)\,ds\,dt = 0 \tag{4.3}$$

with

$$u_{00}(s, t) = \lambda(\lambda+1) - l(l+1), \tag{4.4a}$$

$$u_{01}(s, t) = -2\lambda t, \tag{4.4b}$$

$$u_{02}(s, t) = t^2 + 1, \tag{4.4c}$$

$$u_{11}(s, t) = 2\lambda\beta s - G, \tag{4.4d}$$

$$u_{12}(s, t) = \beta^2 s - 2\beta st, \tag{4.4e}$$

$$u_{22}(s, t) = \beta^2 s^2, \tag{4.4f}$$

$$u_{10}(s, t) = u_{20}(s, t) = u_{21}(s, t) = 0. \tag{4.4g}$$

Now integration of $K(z; s, t)$ with respect to s yields $\exp(\beta z)K(z; s, t)$, and integration of $K(z; s, t)$ with respect to t yields $(1/z)K(z; s, t)$. It is, therefore, possible to decrease the exponents m or n of the terms in Eq. (4.3) by partial integrations with respect to s or t according to

$$\begin{aligned}\int\int &\exp(-m\beta z)\, z^n K(z; s, t)\, u_{mn}(s, t)\, v(s, t)\,ds\,dt\\ &= \int(\exp[-(m-1)\beta z]\, z^n K(z; s, t)\, u_{mn}(s, t)\, v(s, t))_s\,dt\\ &\quad - \int\int \exp[-(m-1)\beta z]\, z^n K(z; s, t) \times \partial[u_{mn}(s, t)v(s, t)]/\partial s\,ds\,dt\\ &= \int(\exp(-m\beta z)\, z^{n-1} K(z; s, t)\, u_{mn}(s, t)\, v(s, t))_t\,ds\\ &\quad - \int\int \exp(-m\beta z)\, z^{n-1} K(z; s, t) \times \partial[u_{mn}(s, t)v(s, t)]/\partial t\,ds\,dt.\end{aligned} \tag{4.5}$$

Here the parentheses with subscript s or t denote the difference between the final and initial value after the contour has been described. The single integrals containing these parentheses are referred to as the semi-integrated terms. Repeated application of this reduction, to each term in Eq. (4.3) m times with respect to s, and n times with respect to t, causes all the powers of $\exp(-\beta z)$ and of z to disappear in the remaining double integral. The condition (4.3) then finally appears in the form

$$\int_{C_t}\int_{C_s} K(z; s, t)\sum_{n=0}^{2}\sum_{m=0}^{2}\partial^{m+n}[u_{mn}(s, t)v(s, t)]/ (\partial s^m \partial t^n)\,ds\,dt + [R] = 0, \tag{4.6}$$

where $[R]$ is an aggregate of semi-integrated terms. While the semi-integrated terms may look different, dependent on the order in which the integrations with respect to s and t have been performed, each of them contains the factors $K(z; s, t)$ and $v(s, t)$ or a partial derivative of $v(s, t)$ and may contain powers of z, $\exp(-\beta z)$, s, and t.

In order that the double integral representation (4.1) may be a solution of the differential equation (2.3) it is, therefore, necessary that the weight function $v(s, t)$ satisfy the partial differential equation

$$\sum_{n=0}^{2}\sum_{m=0}^{2}\partial^{m+n}[u_{mn}(s, t)v(s, t)]/(\partial s^m \partial t^n) = 0 \tag{4.7}$$

or, explicitly,

$$\begin{aligned}\beta^2 s^2 v_{sstt} &+ (3\beta + 2t)\beta s\, v_{stt} + [2(\lambda+2)\beta s - G]v_{st}\\ &+ [(t+\beta)^2 + 1]\, v_{tt} + 2(\lambda+2)(t+\beta)\, v_t\\ &+ (\lambda + 2 + l)(\lambda + 1 - l)\, v(s, t) = 0,\end{aligned} \tag{4.8}$$

and that the contours C_s and C_t in the s and t plane be so chosen that the semi-integrated terms $[R]$ vanish identically in z.

An appropriate solution of the st equation (4.8) is

$$v(s,t)=\sum_{m=0}^{\infty} G^m\, m!\, b_m(t)\, s^{-1-m}, \tag{4.9}$$

where the coefficients $b_m(t)$ obey, with $b_{-1}(t)\equiv 0$, the recursive system of ordinary differential equations

$$[(t-m\beta)^2+1]b_m''+2(\lambda+2)(t-m\beta)b_m' + (\lambda+2+l)(\lambda+1-l)b_m(t)=-b_{m-1}'(t). \tag{4.10}$$

The factors G^m and $m!$ in Eq. (4.9) have been extracted in view of later convenience.

Since $v(s,t)$ according to Eq. (4.9) is single-valued with respect to s, a suitable contour for the s integration in Eq. (4.1) is a closed circle around the origin traversed once in the positive sense. Of the semi-integrated terms $[R]$ then only the single integrals over s survive while all the single integrals over t disappear because of vanishing integrands. Inserting the expansion (4.9) into the integral representation (4.1) and assuming that (if the radius of the circle in the s plane is chosen sufficiently large) the series may be integrated term by term with respect to s, we obtain

$$y(z)=z^{-\lambda-1}(2\pi i)^{-1}\int_{C_t}\exp(zt) \times\sum_{m=0}^{\infty} G^m b_m(t)\exp(-m\beta z)\,dt, \tag{4.11}$$

where the result of the s integration has just produced the factor $\exp(-m\beta z)$ and has cancelled the factorial function and one factor $2\pi i$. Assuming furthermore, that now the t integration of the series may be performed term by term, we are finally led to consider solutions in the form

$$y(z)=z^{-\lambda-1}\sum_{m=0}^{\infty} G^m \exp(-m\beta z)(2\pi i)^{-1}\int_C \exp(zt)\, b_m(t)\,dt, \tag{4.12}$$

where the contour (we now write simply C rather than C_t) has to be so chosen that the remaining semi-integrated terms vanish identically, while the functions $b_m(t)$ satisfy the ordinary differential equations (4.10).

B. General solution of the t equations

In order to obtain a simple solution of the t equations (4.10) we now dispose of the parameter λ by choosing

$$\lambda=L. \tag{4.13}$$

This means, since L is given by Eqs. (3.2), that we consider only the two possibilities $\lambda=l$ or $\lambda=-l-1$. We then have to solve the equations

$$[(t-m\beta)^2+1]b_m''+2(L+2)(t-m\beta)b_m'+2(L+1)b_m(t) =-b_{m-1}'(t), \tag{4.14}$$

the solution of which is found by standard techniques to be

$$b_m(t)=[(t-m\beta)^2+1]^{-L-1}\Big(A_m+\int_{m\beta}^{t}[B_m-b_{m-1}(T)] \times[(T-m\beta)^2+1]^L\,dT\Big). \tag{4.15}$$

Since we want to have a solution $b_m(t)$ which (in case of real parameters β and L) is real when t is, the constants of integration A_m and B_m have to be real. This is a convenient but not necessary agreement.

We now want to investigate in detail the behavior of the solution (4.15) in the vicinity of its singular points. Part of this task can be done conveniently by means of the differential equations of which the $b_m(t)$ are the solution. It is, therefore, important to note that the $b_m(t)$ also satisfy a linear system of first order differential equations,

$$[(t-m\beta)^2+1]\,b_m'(t)+2(L+1)(t-m\beta)\,b_m(t)=B_m-b_{m-1}(t), \tag{4.16}$$

which may be used in place of the second order equations (4.14).

This fact is not so surprising as it might seem at first sight. For if we had restricted the parameter λ to the values of l or $-l-1$ from the beginning, it would have been possible to find a third-order equation for $v(s,t)$ in place of the fourth-order equation (4.8) and, as a consequence, a system of first-order equations for $b_m(t)$, namely Eq. (4.16) with $B_m=0$. For some parts of the present investigation, however, the solution with all the $B_m=0$ turns out to be too much restricted. The situation is quite similar here as in the case of the (spherical) Bessel functions, where some interesting information is lost by the usual treatment implying $B_0=0$ from the beginning.[5]

C. Power series expansions for the solution of the t equations

The system of differential equations (4.14) has regular singular points at $t=M\beta\pm i$ where $M=0,1,2,\cdots$ and at $t=\infty$. Relative to the points $t=M\beta\pm i$ the characteristic exponents are 0 and $-L-1$. Provided that neither L is an integer nor $(m-M)\beta=\pm 2i$, we have

$$b_m(t)=2^{-L-1}\exp[-i\pi(L+1)/2]\sum_{n=0}^{\infty} a_n^{mM}(L) \times(t-M\beta-i)^{-L-1+n}+\sum_{n=0}^{\infty} c_n^{mM}(L)(t-M\beta-i)^n$$

for

$$|t-M\beta-i|<\min(|\beta|,2), \tag{4.17a}$$

$$b_m(t)=2^{-L-1}\exp[i\pi(L+1)/2]\sum_{n=0}^{\infty} a_n^{mM\,*}(L) \times(t-M\beta+i)^{-L-1+n}+\sum_{n=0}^{\infty} c_n^{mM*}(L)\,(t-M\beta+i)^n$$

for

$$|t-M\beta+i|<\min(|\beta|,2). \tag{4.17b}$$

Relative to the point $t=\infty$ the characteristic exponents are 1 and $2L+2$. Provided that $2L$ is not an integer greater than -2, we have the expansion

$$b_m(t)=\sum_{n=0}^{\infty} d_n^{mM}(L)(t-M\beta)^{-2L-2-n}+\sum_{n=0}^{\infty} e_n^{mM}(L)(t-M\beta)^{-1-n} \tag{4.18}$$

for

$$|t-M\beta|>([(m-M)\,\mathrm{Re}(\beta)]^2+[1+|(m-M)\,\mathrm{Im}(\beta)|]^2)^{1/2},$$

if $M\leq m/2$,

or

$$|t - M\beta| > ([M\,\mathrm{Re}(\beta)]^2 + [1 + |M\,\mathrm{Im}(\beta)|]^2)^{1/2} \quad \text{if } M \geqslant m/2.$$

Convenient normalization factors have been introduced in Eqs. (4.17). The coefficients $a_n^{mM*}(L)$ and $c_n^{mM*}(L)$ are conjugate complex to $a_n^{mM}(L)$ and $c_n^{mM}(L)$, respectively, if β and L are real. Otherwise the conjugate complex has to be taken except for β and L which have to be retained unchanged. It can be seen immediately from the differential equations (4.14) or (4.16) or from the solution (4.15) that, for $n=0,1,2,\ldots$,

$$a_n^{mM}(L)=0 \quad \text{if } M>m \text{ or } M<0. \tag{4.19a}$$

As soon as the initial coefficients with $n=0$, which depend on the constants of integration A_m and B_m, have been specified, the other coefficients can be calculated by means of recurrence relations, which can conveniently be obtained if the expansions (4.17) or (4.18), respectively, are inserted into the differential equations (4.16). The recurrence relations are

$$a_n^{mm}(L)=i(2n)^{-1}\,(n+L)\,a_{n-1}^{mm}(L), \tag{4.19b}$$

$$\begin{aligned} a_n^{mM}(L) = {} & ((m-M)\beta([m-M]\beta-2i)(n-L-1))^{-1} \\ & \times(2([m-M]\beta-i)(n-1)\,a_{n-1}^{mM}(L) \\ & -(n+L-1)a_{n-2}^{mM}(L) - a_{n-1}^{m-1\,M}(L)), \quad \text{if } M<m, \end{aligned} \tag{4.19c}$$

$$\begin{aligned} c_n^{mM}(L) = {} & ((m-M)\beta([m-M]\beta-2i)n)^{-1} \\ & \times(2([m-M]\beta-i)(n+L)\,c_{n-1}^{mM}(L) \\ & -(n+2L)\,c_{n-2}^{mM}(L) - c_{n-1}^{m-1\,M}(L) + B_m\delta_{n1}), \quad \text{if } M\neq m, \end{aligned} \tag{4.20a}$$

$$\begin{aligned} c_n^{mm}(L) = {} & i[2(n+L+1)]^{-1}[(n+2L+1)\,c_{n-1}^{mm}(L) \\ & + c_n^{m-1\,m}(L) - B_m\delta_{n0}], \end{aligned} \tag{4.20b}$$

$$\begin{aligned} d_n^{mM}(L) = {} & n^{-1}(2(L+n)(m-M)\beta d_{n-1}^{mM}(L) \\ & -(2L+n)([m-M]^2\beta^2+1)\,d_{n-2}^{mM}(L) + d_{n-1}^{m-1\,M}(L)), \end{aligned} \tag{4.21}$$

$$\begin{aligned} e_n^{mM}(L) = {} & (n-2L-1)^{-1}\,(2(n-L-1)(m-M)\beta\, e_{n-1}^{mM}(L) \\ & -(n-1)([m-M]^2\beta^2+1)\,e_{n-2}^{mM}(L) + e_{n-1}^{m-1\,M}(L)). \end{aligned} \tag{4.22}$$

All the recurrence relations have been written down under the convention that the coefficients are zero whenever one of their indexes (M excepted) becomes negative. The symbol δ_{nk} is equal to 1 or 0 according as n is equal to or different from k.

As to the initial coefficients, we have

$$a_0^{mM}(L)=0 \quad \text{if } m\neq M, \tag{4.23}$$

$$c_0^{mm}(L)=i(2L+2)^{-1}\,(c_0^{m-1\,m}(L)-B_m), \tag{4.24}$$

$$e_0^{mM}(L)=(2L+1)^{-1}B_m. \tag{4.25}$$

The other initial coefficients can be presented in convenient form only if the values of L are so restricted that the required integrals exist. We then have

$$a_0^{mm}(L)=A_m+\int_{m\beta}^{m\beta+i}(B_m-b_{m-1}(T))((T-m\beta)^2+1)^L\,dT \tag{4.23$'$}$$

if $\mathrm{Re}(L)>-1$,

$$\begin{aligned} c_0^{mM}(L) = {} & ((\{M-m\}\beta+2i)(M-m)\beta)^{-L-1} \\ & \times\Big(A_m+\int_{m\beta}^{M\beta+i}[B_m-b_{m-1}(T)][(T-m\beta)^2+1]^L\,dT\Big) \end{aligned} \tag{4.24$'$}$$

for $M<m$ if $\mathrm{Re}(L)<0$ or for $M>m$ with L unrestricted,

$$d_0^{mM}(L)=A_m+\int_{m\beta}^{\infty}(B_m-b_{m-1}(T))((T-m\beta)^2+1)^L\,dT \tag{4.26}$$

if $\mathrm{Re}(L)<-\frac{1}{2}$ (or with L unrestricted if $B_m=0$ for $m=0,1,2,\ldots$).

The restrictions imposed here with respect to L will not be prohibitive. For, by starting from the integrals, we will later obtain expressions which do not need such restrictions and therefore are, by analytic continuation, valid for other values of L too. In order to demonstrate this fact in more detail let us consider $a_0^{mm}(L)$, for instance: The integral in Eq. (4.15) is equal to the sum of two terms, one being constant and the other singular of the type $(t-m\beta-i)^{L+1}$ times a function which is regular at $t=m\beta+i$. By definition of $a_0^{mm}(L)$ it is the constant term which is needed. Now if $\mathrm{Re}(L)>-1$ the singular part vanishes at $t=m\beta+i$ and the integral in Eq. (4.23$'$) is equal to the constant term. The analytic continuation of this integral with respect to L is therefore equal to the analytic continuation of the constant term, just as required.

If Eq. (4.19c) is rewritten with $m=M+N$ where $N=1,2,3,\ldots$, the resulting equation has coefficients independent of M. Therefore, and because of Eq. (4.19a) the $a_n^{M+N\,M}(L)$ are proportional to $a_0^{MM}(L)$ but otherwise independent of M. It then follows that

$$a_n^{M+N\,M}(L)=a_0^{MM}(L)\,a_n^{N0}(L)/a_0^{00}(L). \tag{4.27}$$

Consequently the recurrence relation (4.19c) is needed only with $M=0$.

As to the coefficients $d_n^{mM}(L)$ and $e_n^{mM}(L)$, the situation is more complicated since here the initial coefficients $d_0^{mM}(L)$ and $e_0^{mM}(L)$ generally do not vanish when $M\neq m$. It is therefore necessary to indicate the dependence on the set of initial coefficients by using a more detailed notation. Let us introduce another index P such that $d_n^{mMP}(L)$ denotes the special coefficients generated from the set ($d_0^{mMP}(L)=d_0^{mM}(L)=0$ for $m\neq P$, $d_0^{PMP}(L)=d_0^{PM}(L)\neq 0$). Similarly, let $e_n^{mMP}(L)$ denote the special coefficients generated from the initial coefficients ($e_0^{mMP}(L)=e_0^{mM}(L)=0$ for $m\neq P$, $e_0^{PMP}(L)=e_0^{PM}(L)\neq 0$). The special coefficients obey the same recurrence relations (4.21)—(4.22) as the general coefficients, but, for each P, all the coefficients with $m<P$ vanish. Then for $m\geqslant P$ the same set of coefficients is generated as for $P=0$ apart from a shift of the indexes m and M by P and a normalization factor independent of m and M. For since the recurrence relations (4.21)—(4.22) depend on m and M via $m-M$ only, the shift in m can be compensated by an equal shift in M. Consequently we have, writing shortly d_n^{mMP} in place of $d_n^{mMP}(L)$, etc.,

$$d_n^{P+m\;P+M\;P}=d_n^{mM0}\,d_0^{PP}/d_0^{00}, \tag{4.28a}$$

$$e_n^{P+m\;P+M\;P}=e_n^{mM0}\,e_0^{PP}/e_0^{00}. \tag{4.28b}$$

The general coefficients for $n>0$ then may be expressed in terms of the special coefficients by

$$d_n^{mM}=\sum_{P=0}^{m} d_n^{mMP}=\sum_{P=0}^{m} d_n^{m-P\,M-P\,0}\, d_0^{PP}/d_0^{00}, \tag{4.29a}$$

$$e_n^{mM}=\sum_{P=0}^{m} e_n^{mMP}=\sum_{P=0}^{m} e_n^{m-P\,M-P\,0}\, e_0^{PP}/e_0^{00}. \tag{4.29b}$$

Later in Sec. 6, we will need the sums $\sum_{m=0}^{\infty} G^m d_n^{mm}$ and $\sum_{m=0}^{\infty} G^m e_n^{mm}$. Inserting Eqs. (4.29), interchanging the summations, and introducing a new index of summation $r=m-P$ in place of m, which finally is called m again, we obtain

$$\sum_{m=0}^{\infty} G^m d_n^{mm}=\left(\sum_{m=0}^{n} G^m d_n^{mm0}\right)\left(\sum_{P=0}^{\infty} G^P d_0^{PP}\right)/d_0^{00}, \tag{4.30a}$$

$$\sum_{m=0}^{\infty} G^m e_n^{mm}=\left(\sum_{m=0}^{n} G^m e_n^{mm0}\right)\left(\sum_{P=0}^{\infty} G^P e_0^{PP}\right)/e_0^{00}. \tag{4.30b}$$

Here, use has been made of the fact that

$$d_n^{mm0}=e_n^{mm0}=0 \quad \text{for } m>n. \tag{4.31}$$

Equation (4.31) follows immediately from the recurrence relations (4.21)–(4.22) since, by definition of the case $P=0$, the initial coefficients with $n=0$ are different from zero for $m=0$ only.

This is what had to be said about the coefficients d_n^{mM} and e_n^{mM}. Even more complicated in the situation for the coefficients c_n^{mM}, but for our purpose a detailed discussion of this matter is not required.

D. A special solution of the t equations

The functions $b_m(t)$ in general, as presented in the preceding subsections, still depend on the various constants of integration A_m and B_m. We will now specify these constants in a manner suitable for several purposes.

$$B_m=0 \text{ for } m=0,1,2,\ldots, \tag{4.32a}$$

$$A_0=2^{L+1}\Gamma(L+1), \tag{4.32b}$$

$$A_m=\int_{m\beta}^{\infty} b_{m-1}(T)\,((T-m\beta)^2+1)^L\,dT \quad \text{for } m=1,2,3,\ldots . \tag{4.32c}$$

The corresponding set of functions $b_m(t)$ will be referred to as set I and denoted by $b_m^I(t)$.

5. SOLUTIONS RELATIVE TO THE SINGULAR POINT AT INFINITY: HANKEL-TYPE SOLUTIONS OR SOLUTIONS OF THE THIRD KIND

A. Integral representation

We are now prepared to define standard solutions of the original differential equation (2.3) relative to the singular point at infinity by using the integral representation (4.12) with the set I of functions $b_m(t)$ and one or the other of the contours C_0^+ and C_0^- which, starting from and returning to infinity, enclose one cingular point $t=i$ or $t=-i$, respectively, as shown in Fig. 1a. Provided that $|\arg(\beta)|\leq \pi/2$, $|\beta|>0$, $P\beta\neq 2i$ for all the integers $P\neq 0$, and L is not a negative integer, our standard solutions, defined in analogy with (spherical) Hankel functions, are

$$hy^{(1)}(L;z)=z^{-L-1}\sum_{m=0}^{\infty} G^m \exp(-m\beta z)(2\pi i)^{-1} \times\int_{C_0^+}\exp(zt)\,b_m^I(t)\,dt, \tag{5.1a}$$

$$hy^{(2)}(L;z)=z^{-L-1}\sum_{m=0}^{\infty} G^m \exp(-m\beta z)(2\pi i)^{-1} \times\int_{C_0^-}\exp(zt)\,b_m^I(t)\,dt, \tag{5.1b}$$

where

$$b_0^I(t)=2^{L+1}\Gamma(L+1)(t^2+1)^{-L-1}, \tag{5.2a}$$

$$b_m^I(t)=((t-m\beta)^2+1)^{-L-1}\int_t^{\infty} b_{m-1}^I(T)\,((T-m\beta)^2+1)^L\,dT \tag{5.2b}$$

for $m=1,2,3,\ldots$.

Here, we may agree, the powers appearing in $b_0^I(t)$ are defined according to $\arg(t-i)=-\pi/2$ and $\arg(t+i)=\pi/2$ at that point of each contour where t is on the imaginary axis while $|t|<1$. If the contours start from and return to infinity in the direction parallel to the negative real axis, as shown in Fig. 1a, the integral representations (5.1) are valid for $|\arg(z)|<\pi/2$. More generally, if each of the contours is rotated around the corresponding singular point by an angle α, the integrals exists for $|\alpha+\arg(z)|<\pi/2$. The possible angles of rotation are restricted by the presence of the other singular points, so that $-\pi+\arg(\beta)<\alpha<\pi/2$ in case of $hy^{(1)}(L;z)$ and $-\pi/2<\alpha<\pi+\arg(\beta)$ in case of $hy^{(2)}(L;z)$. Consequently, by rotation of the contour, $hy^{(1)}(L;z)$ may be continued analytically and defined in the larger sector $-\pi<\arg(z)<3\pi/2-\arg(\beta)$ and $hy^{(2)}(L;z)$ in the sector $-3\pi/2-\arg(\beta)<\arg(z)<\pi$.

B. Asymptotic expansion

For each of the integrals in Eqs. (5.1) we may now obtain an asymptotic expansion for $z\to\infty$ on the basis of Watson's lemma.[6] Inserting the power series for $b_m^I(t)$ around the relevant singular point according to Eq. (4.17a) with $M=0$ and integrating the series term by term, we have, in case of $hy^{(1)}(L;z)$,

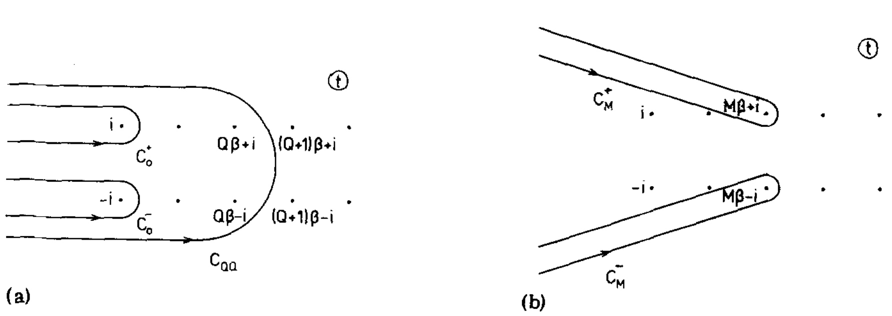

FIG. 1. Contours in the t plane suitable for the integral representation. All the contours shown start somewhere at infinity, enclose at least one of the singular points of the integrand, and return to infinity. The figures have been drawn for the case when β is real and positive. The contour C_{QQ} shown in Fig. 1a corresponds to $Q=2$, the contours in Fig. 1b to $M=2$.

$$z^{-L-1}(2\pi i)^{-1}\int_{C_0^+}\exp(zt)\,b_m^l(t)\,dt$$
$$\sim z^{-1}\exp(iz)\exp[-i\pi(L+1)/2]\sum_{n=0}^{\infty} p_{mn}z^{-n}, \tag{5.3a}$$
$$z\to\infty, \quad -\pi<\arg(z)<3\pi/2-\arg(\beta).$$

Here we have used the fact that, according to Eqs. (4.17a) and (5.2a),

$$a_0^{00}(L)=2^{L+1}\Gamma(L+1), \tag{5.4}$$

and we have introduced the coefficients

$$p_{mn}=[\Gamma(L+1)/\Gamma(L+1-n)]\,a_n^{m0}(L)/a_0^{00}(L). \tag{5.5}$$

Inserting this result into Eq. (5.1a) and treating Eq. (5.1b) in a similar way we obtain, in accordance with Shere,[2] the asymptotic expansions of the Hankel-type solutions,

$$hy^{(1)}(L;z)\sim z^{-1}\exp(iz)\exp[-i\pi(L+1)/2]$$
$$\times\sum_{m=0}^{\infty}G^m\exp(-m\beta z)\sum_{n=0}^{\infty}p_{mn}z^{-n},$$
$$z\to\infty, \quad -\pi<\arg(z)<3\pi/2-\arg(\beta), \quad |\arg(\beta)|\leqslant\pi/2, \tag{5.6a}$$

$$hy^{(2)}(L;z)\sim z^{-1}\exp(-iz)\exp[i\pi(L+1)/2]$$
$$\times\sum_{m=0}^{\infty}G^m\exp(-m\beta z)\sum_{n=0}^{\infty}p_{mn}^{*}z^{-n},$$
$$z\to\infty, \quad -3\pi/2-\arg(\beta)<\arg(z)<\pi, \quad |\arg(\beta)|\leqslant\pi/2. \tag{5.6b}$$

Here the coefficients p_{mn} and p_{mn}^*, which do not depend on L, may be computed from the recurrence relations

$$p_{mn}=(m\beta(m\beta-2i))^{-1}[-2(n-1)(m\beta-i)\,p_{m\,n-1}$$
$$-(n-1+l)(n-2-l)\,p_{m\,n-2}+p_{m-1\,n-1}], \tag{5.7a}$$

$$p_{mn}^*=(m\beta(m\beta+2i))^{-1}[-2(n-1)(m\beta+i)\,p_{m\,n-1}^*$$
$$-(n-1+l)(n-2-l)\,p_{m\,n-2}^*+p_{m-1\,n-1}^*], \tag{5.7b}$$

for $m=1,2,3,\ldots$, and

$$p_{0n}=-i(2n)^{-1}(n+l)(n-l-1)\,p_{0\,n-1}, \tag{5.7c}$$

$$p_{0n}^*=(-1)^n p_{0n}, \tag{5.7d}$$

with

$$p_{00}=p_{00}^*=1 \tag{5.7e}$$

and the coefficients equal to zero whenever $n<m$. These recurrence relations may be obtained either from the definition (5.5) of the p_{mn} in terms of the $a_n^{m0}(L)$ and the recurrence relations satisfied by the $a_n^{m0}(L)$ or directly by substituting the asymptotic expansions (5.6) into the original differential equation (2.3).

Since $L=l$ or $L=-l-1$ according to Eqs. (3.2), we have two sets of solutions, which however differ merely by constant factors,

$$hy^{(1)}(-l-1;z)=\exp[i\pi(2l+1)/2]\,hy^{(1)}(l;z), \tag{5.8a}$$

$$hy^{(2)}(-l-1;z)=\exp[-i\pi(2l+1)/2]\,hy^{(2)}(l;z), \tag{5.8b}$$

in analogy to the behavior of Hankel functions. Equations (5.8) may be used to complete the definition of $hy^{(1)}(L;z)$ and $hy^{(2)}(L;z)$ for L equal to a negative integer. Further comments on the case when L is an integer will be given in Sec. 9. The exceptional case when the parameter β has a value such that $P\beta=2i$ for some integer $P\neq 0$ will be treated in Sec. 10.

If Eqs. (5.6) are to be interpreted as multiple asymptotic expansions, with respect to the sequence $z^{-n}\exp(-m\beta z)$ for $z\to\infty$, in the sense of Shere,[2] it is necessary to impose in addition the restriction that $|\arg(\beta z)|<\pi/2$. Consequently both the asymptotic expansions (5.6) then are valid in the same, smaller sector $-\pi/2-\arg(\beta)<\arg(z)<\pi/2-\arg(\beta)$. In fact, that the integral representations (4.1) or (4.12) may be solutions of the differential equation (2.3) is evident from Sec. 4.A provided that z is restricted either to any finite domain or to any domain which is infinite but such that $|\arg(\beta z)|<\pi/2$ when $z\to\infty$. For otherwise the kernel $K(z;s,t)$ and the powers of $\exp(-\beta z)$ in the semi-integrated terms cause trouble when $z\to\infty$, and therefore the limit $z\to\infty$ of the solutions may be considered without special care only if $|\arg(\beta z)|<\pi/2$. Nevertheless, the Eqs. (5.6) seem to be meaningful in the extended sectors: The expressions (5.1) for the Hankel-type solutions are without doubt solutions of the differential equation (2.3) as long as z is finite, and the sums over m are therefore expected to converge irrespective of the phase of βz. In going from Eqs. (5.1) to Eqs. (5.6) we have replaced the coefficients of the series in powers of $G\exp(-\beta z)$ by their asymptotic expansions. The Eqs. (5.6) are therefore suitable to represent the solutions when $|z|$ is sufficiently large but finite, even if the real part of βz is not positive, just because of the convergence with respect to m. That finally, when $z\to\infty$, all the terms of the sum over m become infinite if $\mathrm{Re}(\beta z)<0$ does not matter, for it reflects the highly singular nature of the solutions at infinity, which cannot simply be accounted for by one multiplicative singular factor times an ordinary asymptotic expansion.

C. Modified integral representation

More generally, let us consider, in view of later application, the solution

$$y_m^+(z)=z^{-L-1}\sum_{m=0}^{\infty}G^m\exp(-m\beta z)$$
$$\times(2\pi i)^{-1}\int_{C_M^+}\exp(zt)\,b_m(t)\,dt, \tag{5.9}$$

where C_M^+ denotes a contour which, starting from and returning to infinity, encloses one singular point $t=M\beta+i$ as shown in Fig. 1b. Again we may agree that the powers appearing in the integrands are defined according to $\arg(t-M\beta-i)=-\pi/2$ and $\arg(t-M\beta+i)=\pi/2$ at that point of the contour where $t-M\beta$ is imaginary while $|t-M\beta|<1$. The $b_m(t)$ are still the general ones with the constants of integration not yet specified. For $m<M$ they are analytic at $t=M\beta+i$ and do not contribute to the integrals. It then remains to consider the integrals with $b_{M+N}(t)$ for $N=0,1,2,\ldots$. The relevant, at $t=M\beta+i$, singular part of $b_{M+N}(t)$ is proportional to

$$\sum_{n=0}^{\infty}a_n^{M+N\,M}(L)(t-M\beta-i)^{-L-1+n},$$

which, because of Eq. (4.27), is equal to

$$(a_0^{MM}(L)/a_0^{00}(L))\sum_{n=0}^{\infty}a_n^{N0}(L)\,(T-i)^{-L-1+n}.$$

This is, apart from the constant factor in front of the sum, essentially the singular part at $T=i$ of $b_N(T)$, where $T=t-M\beta$. For the contour integral we therefore obtain

$$\exp(-M\beta z)\int_{C_M^+}\exp(zt)\,b_{M+N}(t)\,dt$$
$$=(a_0^{MM}(L)/a_0^{00}(L))\int_{C_0^+}\exp(zT)\,b_N(T)\,dT. \qquad (5.10)$$

It then follows that

$$y_M^+(z)=G^M(a_0^{MM}(L)/a_0^{00}(L))\,y_0^+(z). \qquad (5.11)$$

On the other hand we have, by comparison of Eqs. (5.1a) and (5.9),

$$y_0^+(z)=a_0^{00}(L)\,(2^{L+1}\Gamma(L+1))^{-1}hy^{(1)}(L;z), \qquad (5.12)$$

for the general functions $b_m(t)$ and the special functions $b_m^I(t)$ can differ merely by additive terms which are analytic at $t=i$ and, if A_0 and B_0 are not the same in both cases, by the normalization $a_0^{00}(L)$ of the terms which are singular at $t=i$. From Eqs. (5.11)–(5.12) we obtain

$$y_M^+(z)=G^M a_0^{MM}(L)\,(2^{L+1}\Gamma(L+1))^{-1}hy^{(1)}(L;z). \qquad (5.13\text{a})$$

Similarly, if we consider the contour C_M^- which, starting from and returning to infinity, encloses one singular point $t=M\beta-i$ as shown in Fig. 1b, and if $y_M^-(z)$ is the solution of the type (5.9) but with the contour C_M^- rather than C_M^+, we have

$$y_M^-(z)=G^M a_0^{MM*}(L)\,(2^{L+1}\Gamma(L+1))^{-1}hy^{(2)}(L;z). \qquad (5.13\text{b})$$

The integral representation (5.9) with the contours C_M^+ or C_M^-, respectively, yields therefore, apart from normalization factors, again the Hankel-type solutions even when $M\neq 0$.

6. THE CONNECTION PROBLEM

A. General linear relation between the solutions of different kind

Let us consider the integral representation

$$y_{QQ}(z)=z^{-L-1}\sum_{m=0}^{\infty}G^m\exp(-m\beta z)$$
$$\times(2\pi i)^{-1}\int_{C_{QQ}}\exp(zt)\,b_m(t)\,dt \qquad (6.1)$$

with a contour C_{QQ} which, starting from and returning to infinity in a suitable direction as shown in Fig. 1a, encloses all the singular points $t=m\beta+i$ and $t=m\beta-i$ from $m=0$ up to $m=Q$. And let us choose a special set of functions $b_m(t)$ such that all the $b_m(t)$, in particular for $m>Q$, become analytic at all the singular points, $t=(Q+R)\beta+i$ and $t=(Q+R)\beta-i$ with $R=1,2,3,\ldots$, to the right of the contour. This means that for $m>Q$ we need to have $a_0^{mm}(L)=a_0^{mm*}(L)=0$. According to Eq. (4.23′) this condition can be satisfied if the constants of integration, in case of $-1<\mathrm{Re}(L)$, are chosen to be

$$A_m=\tfrac{1}{2}\Big[\int_{m\beta}^{m\beta+i}b_{m-1}(T)\,((T-m\beta)^2+1)^L\,dT$$
$$+\int_{m\beta}^{m\beta-i}b_{m-1}(T)\,((T-m\beta)^2+1)^L\,dT\Big], \qquad (6.2\text{a})$$

$$B_m=E\int_{m\beta-i}^{m\beta+i}b_{m-1}(T)\,((T-m\beta)^2+1)^L\,dT, \qquad (6.2\text{b})$$

with

$$E^{-1}=\int_{m\beta-i}^{m\beta+i}((T-m\beta)^2+1)^L\,dT \qquad (6.2\text{c})$$

or

$$E=-i\,\Gamma(L+\tfrac{3}{2})/(\Gamma(\tfrac{1}{2})\,\Gamma(L+1)). \qquad (6.2\text{d})$$

To specify the constants of integration for $m\leq Q$ is not necessary at this stage. Using the expansion in powers of $(t-m\beta)^{-1}$ according to Eq. (4.18) for the $b_m(t)$, we then have to consider

$$y_{QQ}(z)=z^{-L-1}\sum_{m=0}^{\infty}G^m\exp(-m\beta z)\,(2\pi i)^{-1}$$
$$\times\int_{C_{QQ}}\exp(zt)\Big(\sum_{n=0}^{\infty}d_n^{mm}(L)\,(t-m\beta)^{-2L-2-n}$$
$$+\sum_{n=0}^{\infty}e_n^{mm}(L)\,(t-m\beta)^{-1-n}\Big)dt, \qquad (6.3)$$

where $\arg(t-m\beta)=0$ when $t-m\beta$ is real and positive.

With our choice (6.2) of the constants of integration, all the $b_m(t)$, even those with $m>Q$, are analytic at the singular points of the differential equation situated to the right of the contour C_{QQ}. For each finite m, in particular $m>Q$, the contour may therefore be deformed so as to lie wholly in the region where the series in powers of $(t-m\beta)^{-1}$ converge uniformly. Then it is legitimate to integrate the series term by term, and we obtain

$$y_{QQ}(z)=z^L\sum_{m=0}^{\infty}G^m\sum_{n=0}^{\infty}(\Gamma(2L+2+n))^{-1}d_n^{mm}(L)\,z^n$$
$$+z^{-L-1}\sum_{m=0}^{\infty}G^m\sum_{n=0}^{\infty}(n!)^{-1}e_n^{mm}(L)\,z^n. \qquad (6.4)$$

Interchanging the summations over m and n and expressing, by means of Eqs. (4.30), the general coefficients $d_n^{mm}(L)$ and $e_n^{mm}(L)$ in terms of the special coefficients $d_n^{mm0}(L)$ and $e_n^{mm0}(L)$, we obtain

$$y_{QQ}(z)=(\Gamma(2L+2))^{-1}\Big(\sum_{P=0}^{\infty}G^P d_0^{PP}(L)\Big)z^L\sum_{n=0}^{\infty}w_n(L)\,z^n$$
$$+\Big(\sum_{P=0}^{\infty}G^P e_0^{PP}(L)\Big)z^{-L-1}\sum_{n=0}^{\infty}w_n(-L-1)\,z^n, \qquad (6.5)$$

where

$$w_n(L)=[\Gamma(2L+2)/\Gamma(2L+2+n)]$$
$$\times\sum_{m=0}^{n}G^m d_n^{mm0}(L)/d_0^{00}(L), \qquad (6.6\text{a})$$

$$w_n(-L-1)=(n!)^{-1}\sum_{m=0}^{n}G^m e_n^{mm0}(L)/e_0^{00}(L). \qquad (6.6\text{b})$$

That the right-hand sides of Eqs. (6.6) are equal to the coefficients $w_n(L)$ and $w_n(-L-1)$ of Sec. 3 can be verified by explicit computation of the first few coefficients in both cases, but it seems difficult to show generally for arbitrary n. Comparing Eq. (6.5) with the Bessel-type solutions (3.1) we have

$$y_{QQ}(z)=[2^L\Gamma(L+1)]^{-1}\Big(\sum_{P=0}^{\infty}G^P d_0^{PP}(L)\Big)$$
$$\times C(L)\,jy(L;z)+2^{-L}(\Gamma(\tfrac{1}{2}-L)/\Gamma(\tfrac{1}{2}))$$
$$\times\Big(\sum_{P=0}^{\infty}G^P e_0^{PP}(L)\Big)C(-L-1)\,jy(-L-1;z). \qquad (6.7)$$

This result could have been obtained from Eq. (6.4) even without the detailed discussion of the properties of the coefficients: For we know that $y_{QQ}(z)$ is a solution of the differential equation (2.3) and therefore should be some linear combination of the solutions $jy(L;z)$ and

$jy(-L-1;z)$. In fact, Eq. (6.4) shows the expected analytical structure, and therefore it suffices to consider only the terms with $n=0$, which yield the normalization factors. By comparison with the normalization factors of $jy(L;z)$ and $jy(-L-1;z)$, Eq. (6.7) then follows immediately. More generally, the corresponding coefficients for any n should agree, and this fact provides an indirect proof of Eqs. (6.6).

We have succeeded in expressing the solution $y_{QQ}(z)$ as a linear combination of the Bessel-type solutions. It remains to express $y_{QQ}(z)$ as a linear combination of the Hankel-type solutions. To do so we again start from the integral representation (6.1) and observe that the contour C_{QQ} is, after suitable deformation and apart from some parts at infinity which do not contribute to the integral, equal to the sum of all the contours C_M^+ and C_M^- where $M=0,1,2,\ldots,Q-1,Q$. We therefore have

$$y_{QQ}(z)=z^{-L-1}\sum_{m=0}^{\infty}G^m\exp(-m\beta z)(2\pi i)^{-1}\times\sum_{M=0}^{Q}\left[\int_{C_M^+}\exp(zt)\,b_m(t)\,dt+\int_{C_M^-}\exp(zt)\,b_m(t)\,dt\right] \tag{6.8}$$

or, by means of Eqs. (5.9) and (5.13),

$$y_{QQ}(z)=\sum_{M=0}^{Q}G^Ma_0^{MM}(L)][2^{L+1}\Gamma(L+1)]^{-1}hy^{(1)}(L;z)+\sum_{M=0}^{Q}G^Ma_0^{MM*}(L)][2^{L+1}\Gamma(L+1)]^{-1}hy^{(2)}(L;z). \tag{6.9}$$

By equating the two different expressions (6.7) and (6.9) for $y_{QQ}(z)$ we have the general linear relation between the solutions of different kind, valid for arbitrary values of the integer $Q=0,1,2,\cdots$. The constants of integration A_m and B_m, on which all the initial coefficients $a_0^{mm}(L)$, $a_0^{mm*}(L)$, $d_0^{mm}(L)$, and $e_0^{mm}(L)$ depend, have not yet been specified for $m\leq Q$. It is by choosing Q and the A_m and B_m for $m\leq Q$ in a suitable way, that we will obtain the desired special relations between our standard solutions.

B. Special relation between the solutions of the first and third kind

1. The connection coefficients

Our first choice of the constants of integration is such that we have, for $m=0,1,2,\ldots,Q$,

$$b_m(t)=b_m^I(t). \tag{6.10a}$$

Consequently

$$d_0^{00}(L)=2^{L+1}\Gamma(L+1), \tag{6.10b}$$

$$d_0^{mm}(L)=0 \quad \text{for } m=1,2,3,\ldots,Q, \tag{6.10c}$$

$$e_0^{mm}(L)=0 \quad \text{for } m=0,1,2,\ldots,Q, \tag{6.10d}$$

$$a_0^{00}(L)=2^{L+1}\Gamma(L+1), \tag{6.10e}$$

and, if $\mathrm{Re}(L)>-1$, $a_0^{mm}(L)$ may simply be written as an integral

$$a_0^{mm}(L)=\int_{m\beta+i}^{\infty}b_{m-1}^I(T)\left[(T-m\beta)^2+1\right]^L dT \tag{6.10f}$$

for $m=1,2,3,\ldots,Q$. Equation (6.7) then reduces to

$$y_{QQ}(z)=\left(2+[2^L\Gamma(L+1)]^{-1}\sum_{P=Q+1}^{\infty}d_0^{PP}(L)\right)\times C(L)jy(L;z)+2^{-L}(\Gamma(\tfrac12-L)/\Gamma(\tfrac12))\times\left(\sum_{P=Q+1}^{\infty}e_0^{PP}(L)\right)C(-L-1)jy(-L-1;z), \tag{6.11}$$

and this expression must be equal to the right-hand side of Eq. (6.9). Provided that the sums over the initial coefficients in Eq. (6.9) will converge, we may now take the limit $Q\to\infty$ and obtain

$$2C(L)jy(L;z)=D(L)hy^{(1)}(L;z)+D^*(L)hy^{(2)}(L;z), \tag{6.12}$$

where we have introduced the connection coefficients

$$D(L)=1+\sum_{M=1}^{\infty}G^Ma_0^{MM}(L)/a_0^{00}(L) \tag{6.13}$$

and $D^*(L)$ which is the corresponding conjugate complex function of the parameters G, β, L.

2. Convergence of the expressions for the connection coefficients

Since the left-hand side of Eq. (6.12) is well-defined, at least if $2L\neq-2,-3,-4,\cdots$, the expressions (6.13) for the connection coefficients on the right-hand side are expected to converge. Nevertheless a proof of this fact seems to be desirable. For this purpose we observe that, since the integrand in Eqs. (5.2b) and (6.10f) vanishes when $T\to\infty$ at least as fast as T^{-2}, the upper limit ∞ of the integrals may be replaced by $\epsilon+m\beta+i+\infty\exp[i\arg(\beta)]$. It then suffices to know the behavior of each of the $b_m(t)$ on the straight line defined by

$$t(r)=\epsilon+m\beta+i+r\exp[i\arg(\beta)],\quad 0\leq r\leq\infty, \tag{6.14a}$$

where $\epsilon=0$. More generally let us consider the functions $b_m(t)$ on a straight line parallel to the original one, corresponding to some real ϵ with $0\leq\epsilon<\min(2,|\beta|)$. In case of $m=1$ we then have, using the parameter r, which is real and positive, as a new variable of integration,

$$K(L)\,b_1^I(t)\,((t-\beta)^2+1)^{L+1}=\int_R^{\infty}(H(r))^L(F(r))^{-1}E\,dr \tag{6.14b}$$

with

$$K(L)=[2^{L+1}\Gamma(L+1)]^{-1}, \tag{6.14c}$$

$$R=(t-\epsilon-\beta-i)/E, \tag{6.14d}$$

$$E=\exp[i\arg(\beta)], \tag{6.14e}$$

$$H(r)=\{(rE+\epsilon)/[(r+|\beta|)E+\epsilon]\}\{(rE+\epsilon+2i)/[(r+|\beta|)E+\epsilon+2i]\}, \tag{6.14f}$$

$$F(r)=[(r+|\beta|)E+\epsilon][(r+|\beta|)E+\epsilon+2i]. \tag{6.14g}$$

By a suitable choice of ϵ, to be discussed below, $H(r)$ in case of $\mathrm{Re}(L)\geq0$ or $1/H(r)$ in case of $\mathrm{Re}(L)<0$ can be made to be bounded. A positive constant $I_1(L)$ then exists such that $|(H(r))^L|<I_1(L)$. Also, with the same choice of ϵ, there is a positive constant $I_2>0$ such that

$$|F(r)| > I_2(r+|\beta|)^2. \tag{6.14h}$$

For example, if β is real and $\mathrm{Re}(L) \geq 0$ we may simply take $I_1(L) \equiv 1$ and $I_2 = 1$, and this is true even with $\epsilon = 0$. Since the integral over the r-dependent factors of the estimate is

$$\int_R^\infty (r+|\beta|)^{-2}\,dr = (R+|\beta|)^{-1}, \tag{6.14i}$$

it follows, with $I_1(L)/I_2 = I(L) > 0$, that

$$|K(L)\,b_1^I(t)\,((t-\beta)^2+1)^{L+1}| < I(L)((t-\epsilon-i)/E)^{-1}. \tag{6.14j}$$

Assuming, on the basis of this result, that

$$|K(L)\,b_m^I(t)\,[(t-m\beta)^2+1]^{L+1}| < (I(L))^m\,[(t-\epsilon-\{m-1\}\beta-i)/E]^{-m}/m! \tag{6.14k}$$

we obtain, using again the positive real parameter r of Eq. (6.14a) with m replaced by $m+1$ as a new variable of integration,

$$|K(L)\,b_{m+1}^I(t)\,[(t-\{m+1\}\beta)^2+1]^{L+1}| < (I(L))^m \left| \int_{(t-\epsilon-\{m+1\}\beta-i)/E}^\infty (H(r))^L (F(r))^{-1} \times (r+|\beta|)^{-m} E\,dr \right| / m! \tag{6.14l}$$

or

$$|K(L)\,b_{m+1}^I(t)\,[(t-\{m+1\}\beta)^2+1]^{L+1}| < (I(L))^{m+1}[(t-\epsilon-m\beta-i)/E]^{-m-1}/(m+1)!. \tag{6.14m}$$

This is the same equation as Eq. (6.14k) apart from the replacement of m by $m+1$. And therefore, since Eq. (6.14k) is valid for $m=1$ according to Eq. (6.14j), it holds for any $m=1,2,3,\cdots$. As to the choice of ϵ, we prefer $\epsilon=0$ if the required bounds exist for $\epsilon=0$. Since $a_0^{mm}(L)$ then is just the value at $t=m\beta+i$ of the integral estimated by Eq. (6.14k), it follows immediately that

$$|K(L)\,a_0^{mm}(L)| < [I(L)/|\beta|]^m/m!. \tag{6.14n}$$

The conditions implying that $\epsilon=0$ is a possible choice, which may depend on whether we consider $a_0^{mm}(L)$ or $a_0^{mm*}(L)$ and will be stated so as to apply to both simultaneously, are that $\mathrm{Re}(L) \geq 0$ and the interval $[-2i, 2i]$ of the imaginary axis is excluded for β. We now want to show that the estimate (6.14n) remains valid even if these conditions are violated. We then have to choose an $\epsilon \neq 0$ such that the required bounds exist. As a consequence the estimate (6.14k), with a different constant $I(L)$, is valid on a path which no longer contains the point of interest $t=m\beta+i$, but starts some distance ϵ from it apart. On the other hand, the integral is known to be the sum of two terms, one singular and the other regular at $t=m\beta+i$. While the regular term tends to a constant when $t\to m\beta+i$, the singular one either diverges or vanishes according as $\mathrm{Re}(L) \leq -1$ or $\mathrm{Re}(L) > -1$, respectively. If ϵ is sufficiently small, then these terms cannot cancel because of their different order of magnitude, and the estimate (6.14k) is also approximately valid for the regular part alone. But the regular part does not change significantly when $\epsilon \to 0$, and therefore the estimate (6.14k) essentially remains true at $t=m\beta+i$ for the regular part, in particular as far as the factor $1/m!$ is concerned. Since the value at $t=m\beta+i$ of the regular part is just $a_0^{mm}(L)$, it follows that Eq. (6.14n), may be with a different constant $I(L)$, applies even when an $\epsilon\neq 0$ is needed.

We therefore have shown that the expression (6.13) for the connection coefficients, defined for $|\arg(\beta)| \leq \pi/2$, converges provided that $\beta\neq 0$, for arbitrary values of G. This means that we have solved the connection problem for arbitrary values of G and β while $|\arg(\beta)| \leq \pi/2$, except for $\beta=0$ while $G\neq 0$, a case which although our method breaks down is much simpler and can be reduced to Kummer's differential equation. It is another question, to be considered in the next subsection, whether or not we are able to obtain in all cases a sufficiently simple explicit expression for the $a_0^{mm}(L)$.

3. Explicit expressions for the connection coefficients

In order to evaluate $a_0^{mm}(L)$ as given by Eq. (6.10f) we insert the expansion of $b_{m-1}^I(t)$ around infinity,

$$b_{m-1}^I(t) = \sum_{n=0}^\infty d_n^{m-1\,00}(L)\,t^{-2L-2-n}. \tag{6.15}$$

It may be seen from the domain of convergence for Eq. (4.18) that this expansion converges uniformly on the path of integration, provided that $|\beta| > 2|\sin(\arg(\beta))|$, a condition which has been stated so as to apply simultaneously to both the $a_0^{mm}(L)$ and $a_0^{mm*}(L)$. Assuming that the parameter β obeys the restriction just mentioned we may integrate the series term by term. We then have to evaluate the integrals

$$2^{L+1}\Gamma(L+1)H(L,n) = \int_{m\beta+i}^\infty T^{-2L-2-n}((T-m\beta)^2+1)^L\,dT, \tag{6.16}$$

which may be expressed in terms of hypergeometric functions [using, for example, Ref. 7, Eq. (3.197,2)],

$$H(L,n) = 2^{-L-1}[\Gamma(L+2+n)]^{-1} n!\,(m\beta+i)^{-n-1} \times {}_2F_1(-L, n+1; L+2+n; (m\beta-i)/(m\beta+i)). \tag{6.17a}$$

By means of the transformation formulas of the hypergeometric function this can be written

$$H(L,n) = 2^{-L-1}[\Gamma(L+2+n)]^{-1} n!\,(2m\beta)^{-n-1} \times {}_2F_1(\tfrac12+n/2,\ 1+n/2;\ L+2+n;\ (1+m^2\beta^2)/(m^2\beta^2)) \tag{6.17b}$$

or

$$H(L,n) = 2^{-L-1}[\Gamma(L+2+n)]^{-1} n!\,(2i)^{-n-1} \times {}_2F_1(\tfrac12+n/2,\ L+1+n/2;\ L+2+n;\ 1+m^2\beta^2), \tag{6.17c}$$

and finally

$$\begin{aligned} H(L,n) = {} & n!\,2^{-L-2-n}\exp[-i\pi(n+1)/2]\,(1+m^2\beta^2)^{-(n+1)/2} \\ & \times[\Gamma(\tfrac12)\,(\Gamma(1+n/2)\,\Gamma(L+\tfrac32+n/2))^{-1} \\ & \times {}_2F_1(\tfrac12+n/2,\ -L-\tfrac12-n/2;\ \tfrac12;\ m^2\beta^2/(1+m^2\beta^2)) \\ & +\Gamma(-\tfrac12)\,(\Gamma(\tfrac12+n/2)\,\Gamma(L+1+n/2))^{-1} \\ & \times\exp(-i\pi/2)\,(m^2\beta^2/(1+m^2\beta^2))^{1/2} \\ & \times {}_2F_1(1+n/2,\ -L-n/2;\ \tfrac32;\ m^2\beta^2/(1+m^2\beta^2))] \end{aligned} \tag{6.17d}$$

(L not a negative integer)

or

$$H(L,n)=n!\ 2^{-L-2-n}(m^2\beta^2+1)^{-(n+1)/2}$$
$$\times\{\Gamma(L+\tfrac{1}{2})(\Gamma(L+1+n/2)\Gamma(L+\tfrac{3}{2}+n/2))^{-1}$$
$$\times{}_2F_1(\tfrac{1}{2}+n/2,\ -L-\tfrac{1}{2}-n/2;\ -L+\tfrac{1}{2};\ 1/(m^2\beta^2+1))$$
$$+\Gamma(-L-\tfrac{1}{2})(\Gamma(\tfrac{1}{2}+n/2)\Gamma(1+n/2))^{-1}$$
$$\times\exp[i\pi(L+\tfrac{1}{2})](m^2\beta^2+1)^{-L-1/2}$$
$$\times{}_2F_1(L+1+n/2,\ -n/2;\ L+\tfrac{3}{2};\ 1/(m^2\beta^2+1))\} \tag{6.17e}$$

(L not half an odd integer).

The application of the transformation formulas of the hypergeometric function requires some care in order that the appropriate branches may be taken. The fractional powers in Eqs. (6.17d)–(6.17e) have been defined so as to become real and positive if β is (and L is real). We have chosen for presentation the expressions (6.17d)–(6.17e) with the important case in mind that β is real. Equations (6.17d) or (6.17e) then are preferable according as $m^2\beta^2\leqslant 1$ or $m^2\beta^2\geqslant 1$, respectively. Other expressions for $H(L,n)$, which may be more suitable when β is complex, can be obtained from Eq. (6.17a) by means of the transformation formulas of the hypergeometric function.

For the connection coefficients we now have, from Eqs. (6.10f), (6.13), (6.15), and (6.16),

$$D(L)=1+\sum_{m=1}^{\infty}G^m\sum_{n=0}^{\infty}d_n^{m-1\,0\,0}(L)H(L,n), \tag{6.18a}$$

$$D^*(L)=1+\sum_{m=1}^{\infty}G^m\sum_{n=0}^{\infty}d_n^{m-1\,0\,0}(L)H^*(L,n), \tag{6.18b}$$

where the coefficients $d_n^{m-1\,0\,0}(L)$ can be calculated recursively from Eq. (4.21) with $M=0$ and the initial coefficients (6.10b)–(6.10c). The $H^*(L,n)$ are the corresponding conjugate complex functions of the parameters β and L. The sums over n have different domains of convergence with respect to β if $|\beta|\leqslant 2$ while $\mathrm{Im}(\beta)\neq 0$. The condition for convergence is $|\beta|>-2\sin[\arg(\beta)]$ in case of $D(L)$ and $|\beta|>2\sin[\arg(\beta)]$ in case of $D^*(L)$. The sums over m, we may recall, converge under less stringent conditions on β for arbitrary (not necessarily real) values of G. From a computational point of view, Eqs. (6.18) are useful only for a more restricted range of the parameters, since otherwise the rate of convergence may become prohibitively slow. We should keep in mind that the Hankel-type solutions and the connection coefficients have been defined in the preceding sections for $|\arg(\beta)|\leqslant\pi/2$ (and $\beta\neq 0$) only. The analytic continuation with respect to β will not be considered in the present paper.

7. RELATIONS BETWEEN DIFFERENT SOLUTIONS

A. Normalization of the solutions of the first kind

We now dispose of the normalization factor $C(L)$ of the Bessel-type solutions by choosing

$$C(L)=\{D(L)D^*(L)\}^{1/2}, \tag{7.1}$$

where $C(L)>0$ for real values of L, G, β. With $\delta(L)$ defined $\mathrm{mod}(2\pi)$ by

$$\exp\{i\delta(L)\}=D(L)/C(L) \tag{7.2}$$

we then have

$$2jy(L;z)=\exp[i\delta(L)]hy^{(1)}(L;z)+\exp[-i\delta(L)]hy^{(2)}(L;z) \tag{7.3a}$$

or

$$2jy(L;z)=\cos[\delta(L)][hy^{(1)}(L;z)+hy^{(2)}(L;z)]+i\sin[\delta(L)][hy^{(1)}(L;z)-hy^{(2)}(L;z)]. \tag{7.3b}$$

B. Wronskian relations

Let us consider the Wronskian

$$W(z)=W(y_1(z),y_2(z))=y_1(z)y_2'(z)-y_1'(z)y_2(z) \tag{7.4}$$

of two solutions $y_1(z)$ and $y_2(z)$. By standard techniques it can easily be seen from the differential equation (2.3) that $z^2W(z)$ is identically constant. Evaluating $z^2W(z)$ at infinity or at the origin, respectively, we obtain

$$z^2W(hy^{(1)}(L;z),hy^{(2)}(L;z))=-2i, \tag{7.5}$$

$$z^2[W(jy(L;z),jy(-L-1;z))]=-\sin[\pi(L+\tfrac{1}{2})]/[C(L)C(-L-1)]. \tag{7.6}$$

The last Wronskian, when evaluated by means of Eqs. (7.3a) and (7.5), gives

$$z^2W(jy(L;z),jy(-L-1;z))=-\sin[\delta(-L-1)-\delta(L)+\pi(L+\tfrac{1}{2})]. \tag{7.7}$$

By comparison with Eq. (7.6) we then have

$$C(L)C(-L-1)\sin[\delta(-L-1)-\delta(L)+\pi(L+\tfrac{1}{2})]=\sin[\pi(L+\tfrac{1}{2})]. \tag{7.8}$$

C. Circuit relations

By inspection of Eq. (3.1), which is valid for arbitrary values of $\arg(z)$, we find immediately, for any integer M, the circuit relations of the Bessel-type solutions

$$jy(L;z\exp(2M\pi i))=\exp(2ML\pi i)jy(L;z). \tag{7.9}$$

The circuit relations of the Hankel-type solutions, which then can be found by means of the connection formula (7.3a) in the usual way, are

$$hy^{(1)}(L;z\exp(2M\pi i))=\{\cos[\delta(-L-1)-\delta(L)-(2M-1)L\pi]\,hy^{(1)}(L;z)-\exp[-i(\delta(-L-1)+\delta(L)+(L+1/2)\pi)]\times\sin(2ML\pi)hy^{(2)}(L;z)\}/\cos[\delta(-L-1)-\delta(L)+L\pi], \tag{7.10a}$$

$$hy^{(2)}(L;z\exp(2M\pi i))=\{\exp[i(\delta(-L-1)+\delta(L)+(L+\tfrac{1}{2})\pi)]\times\sin(2ML\pi)hy^{(1)}(L;z)+\cos[\delta(-L-1)-\delta(L)+(2M+1)L\pi]hy^{(2)}(L;z)\}/\cos[\delta(-L-1)-\delta(L)+L\pi]. \tag{7.10b}$$

These equations can be used to extend the definition of the Hankel-type solutions and to obtain their asymptotic expansions for values of $\arg(z)$ other then those considered in Sec. 5.

8. NEUMANN-TYPE SOLUTIONS OR SOLUTIONS OF THE SECOND KIND

Solutions of the second kind may be defined by the requirement that they are to be "orthogonal" near infinity to the corresponding Bessel-type solutions, i.e.,

$$2iny(L;z)=\exp[i\delta(L)]hy^{(1)}(L;z)-\exp[-i\delta(L)]hy^{(2)}(L;z) \tag{8.1a}$$

$$2ny(L;z)=-i\cos[\delta(L)](hy^{(1)}(L;z)-hy^{(2)}(L;z))+\sin[\delta(L)](hy^{(1)}(L;z)+hy^{(2)}(L;z)). \tag{8.1b}$$

We then have

$$hy^{(1)}(L;z)=\exp[-i\delta(L)](jy(L;z)+iny(L;z)), \tag{8.2a}$$

$$hy^{(2)}(L;z)=\exp[i\delta(L)](jy(L;z)-iny(L;z)), \tag{8.2b}$$

and

$$ny(L;z)=\{\cos[\delta(-L-1)-\delta(L)+\pi(L+\tfrac{1}{2})]jy(L;z)-jy(-L-1;z)\}/\sin[\delta(-L-1)-\delta(L)+\pi(L+\tfrac{1}{2})] \tag{8.3a}$$

or, because of Eq. (7.8),

$$ny(L;z)=\{\cos[\delta(-L-1)-\delta(L)+\pi(L+\tfrac{1}{2})]jy(L;z)-jy(-L-1;z)\}C(L)C(-L-1)/\sin[\pi(L+\tfrac{1}{2})]. \tag{8.3b}$$

Equations (8.3) show the behavior near the origin of the Neumann-type solutions, provided that L is neither an integer nor half an odd integer.

The Neumann-type solutions have been defined so as to be always linearly independent of the corresponding Bessel-type solutions. In fact, from Eqs. (7.6) and (8.3b) follows the Wronskian relation

$$z^2W(jy(L;z),ny(L;z))=1. \tag{8.4}$$

9. THE CASE WHEN 2L IS AN INTEGER

A. Hankel-type solutions for L equal to an integer

While the definition of the Hankel-type solutions by the contour integrals (5.1) includes the case when $L=0,1,2,\cdots$, their asymptotic expansions (5.6)–(5.7) have been derived in Sec. 5.B under the tacit assumption that L be not equal to an integer, for otherwise the expansions (4.17) of the integrands break down and have to be replaced by more complicated expressions containing logarithmic terms. Inserting these expressions into the integrals and integrating the series term by term we find, nevertheless, the same asymptotic expansions (5.6)–(5.7). This is not an unexpected result, for the asymptotic expansions are well-defined even when L becomes equal to an integer and therefore should remain valid, by analytic continuation with respect to L.

B. General properties of the Bessel- and Neumann-type solutions

If $L=-\frac{1}{2}+N$ is half an odd integer, we may conclude from Eq. (7.8) that

$$\delta(-\tfrac{1}{2}-N)=\delta(-\tfrac{1}{2}+N). \tag{9.1}$$

It then follows from Eqs. (5.8), (7.3), and (8.1) that

$$jy(-\tfrac{1}{2}-N;z)=(-1)^N jy(-\tfrac{1}{2}+N;z), \tag{9.2}$$

$$ny(-\tfrac{1}{2}-N;z)=(-1)^N ny(-\tfrac{1}{2}+N;z). \tag{9.3}$$

It therefore suffices to investigate $jy(-\frac{1}{2}+N;z)$, $ny(-\frac{1}{2}+N;z)$, and, as will be explained in Sec. 9.D, $jy(N;z)$ and $ny(N;z)$ for $N=0,1,2,\cdots$. As to $jy(-\frac{1}{2}+N;z)$ and $jy(N;z)$, the connection problem has already been solved in Sec. 6, apart from the fact that in case of $jy(-\frac{1}{2}+N;z)$ the expression (6.17e) for $H(L,n)$ has to be replaced. The appropriate expression can be obtained either from Eq. (6.17e) by a limiting process or more simply from Eq. (6.17c) with $L=-\frac{1}{2}+N$ by means of that continuation formula of the hypergeometric function which applies in this exceptional case [Ref. 8, Eq. (15.3.13) or (15.3.14)].

If $2L$ is an integer greater than -2, then any solution $y(L;z)$ which is linearly independent of $jy(L;z)$ can, by standard techniques, be seen to have the form

$$y(L;z)=\sum_{n=0}^{\infty} g_n(L)\,z^{-L-1+n}+\alpha(L)\sum_{n=0}^{\infty} w_n(L)\,z^{L+n}\ln(z), \tag{9.4}$$

where $g_0(L)\neq 0$ and $g_{2L+1}(L)$ if $L\neq -\frac{1}{2}$ or $g_0(-\frac{1}{2})$ and $\alpha(-\frac{1}{2})\neq 0$ are the constants of integration. The coefficients $g_n(L)$ and $\alpha(L)$ can be evaluated by means of the recurrence relations [with $g_{-1}(L)=0$]

$$g_n(L)=\Big(-g_{n-2}(L)+G\sum_{m=0}^{n-1}(1/m!)(-\beta)^m g_{n-m-1}(L)\Big)\Big/[n(n-2L-1)] \quad \text{for } 0<n<2L+1 \tag{9.5a}$$

with $L=\frac{1}{2},1,\frac{3}{2},\cdots$,

$$\alpha(L)=\Big(-g_{2L-1}(L)+G\sum_{m=0}^{2L}(1/m!)(-\beta)^m g_{2L-m}(L)\Big)\Big/(2L+1) \tag{9.5b}$$

with $L=0,\frac{1}{2},1,\cdots$,

$$g_n(L)=\Big(-g_{n-2}(L)+G\sum_{m=0}^{n-1}(1/m!)(-\beta)^m g_{n-m-1}(L)-(2n-2L-1)\,\alpha(L)\,w_{n-2L-1}(L)\Big)\Big/[n(n-2L-1)] \tag{9.5c}$$

for $n>2L+1$ with $L=-\frac{1}{2},0,\frac{1}{2},\cdots$.

The coefficients $w_n(L)$ are those defined by Eqs. (3.3). By comparison with Eqs. (3.3) we have, with $L=0,\frac{1}{2},1,\cdots$,

$$g_n(L)=g_0(L)\,w_n(-L-1) \quad \text{for } n<2L+1 \tag{9.5d}$$

and

$$\alpha(L)=2g_0(L)\,[(s+L+1)\,w_{2L+1}(s)]_{s=-L-1}. \tag{9.5e}$$

Changing the constant $g_{2L+1}(L)$ is equivalent to adding a multiple of $jy(L;z)$.

If $2L=1,2,3,\cdots$, it may happen for particular pairs of values of (G,β) that Eq. (9.5b) yields $\alpha(L)=0$ and the logarithmic terms in Eq. (9.4) do disappear. Such cases are related to the indeterminacy points of the S matrix studied by Ahmadzadeh, Burke, and Tate.[3]

The problem now under consideration is to determine the constants of integration such that the solution (9.4) just represents $ny(-\frac{1}{2}+N; z)$ or $ny(N; z)$, respectively.

C. Neumann-type solutions for L equal to half an odd integer

If L is half an odd integer, the expression (8.3b) for the Neumann-type solution in terms of the Bessel-type solutions becomes indeterminate because of Eqs. (9.1) and (9.2), and the appropriate limit has to be taken. Inserting the expansion (3.1) for the Bessel-type solutions and performing the limiting process one finds that the result has the expected analytical structure (9.4). Comparison of the initial coefficients of the different series with the corresponding ones in Eq. (9.4) then yields, for $N=0$,

$$g_0(-\tfrac{1}{2}) = (2/\pi)^{1/2} C(-\tfrac{1}{2})\{\psi(1) + \ln(2) + [d\ln\{C(L)\}/dL]_{L=-1/2}\}, \tag{9.6a}$$

$$\alpha(-\tfrac{1}{2}) = (2/\pi)^{1/2} C(-\tfrac{1}{2}), \tag{9.6b}$$

and, for $N=1, 2, 3, \cdots$,

$$g_0(-\tfrac{1}{2}+N) = -(2/\pi)^{1/2} 2^{N-1}(N-1)!\, C(-\tfrac{1}{2}+N), \tag{9.6c}$$

$$\alpha(-\tfrac{1}{2}+N) = (2/\pi)^{1/2} 2^{-N}(1/N!)\, C(-\tfrac{1}{2}-N), \tag{9.6d}$$

$$g_{2N}(-\tfrac{1}{2}+N) = \pi^{-1/2}[d\{[2^{-L-1}/\Gamma(\tfrac{3}{2}+L)]\, C(-L-1) + (-1)^{N+1}[2^{L}/\Gamma(\tfrac{1}{2}-L)] \times w_{2N}(-L-1)\, C(L)\}/dL]_{L=-1/2+N}. \tag{9.6e}$$

To obtain Eq. (9.6d) use has been made of the relations

$$C(-\tfrac{1}{2}-N) = (-1)^N 2^{2N} N!\, C(-\tfrac{1}{2}+N) \times (w_{2N}(-L-1)/\Gamma(\tfrac{1}{2}-L))_{L=-1/2+N} \tag{9.7a}$$

or

$$C(-\tfrac{1}{2}-N) = -2^{2N-2}\{(N-1)!\}^2 C(-\tfrac{1}{2}+N) \times \left(-w_{2N-2}(-\tfrac{1}{2}-N) + G\sum_{m=0}^{2N-1}(1/m!) \times (-\beta)^m w_{2N-1-m}(-\tfrac{1}{2}-N)\right) \tag{9.7b}$$

which follows from Eqs. (3.1) and (9.2). Because of Eq. (9.7b), the result (9.6d) for $\alpha(-\frac{1}{2}+N)$ is consistent with Eq. (9.5b).

While the expressions for $\alpha(-\frac{1}{2})$, $g_0(-\frac{1}{2}+N)$, and $\alpha(-\frac{1}{2}+N)$ are quite simple, the expressions obtained for $g_0(-\frac{1}{2})$ and $g_{2N}(-\frac{1}{2}+N)$ are complicated because of the required derivatives of $C(L)$ and $C(-L-1)$ which are inconvenient to evaluate. And although Eq. (9.6e) can be cast into a more suitable form, we do not show this expression since it still contains the derivatives of $C(L)$.

On the basis of the analytical structure (9.4) of $ny(-\frac{1}{2}+N)$ and the expression (9.6d) for $\alpha(-\frac{1}{2}+N)$ we may now obtain the circuit relations for the Neumann-type solutions

$$ny[-\tfrac{1}{2}+N; z\exp(2M\pi i)] = (-1)^M[ny(-\tfrac{1}{2}+N; z) + 4MiC(-\tfrac{1}{2}-N) \times C(-\tfrac{1}{2}+N)\, jy(-\tfrac{1}{2}+N; z)]. \tag{9.8}$$

And then, by means of Eqs. (7.9), (8.2), and (9.8), we are able to derive the circuit relations for the Hankel-type solutions

$$hy^{(1)}(-\tfrac{1}{2}+N; z\exp(2M\pi i)) = (-1)^M\{[1-2MC(-\tfrac{1}{2}-N) \times C(-\tfrac{1}{2}+N)]\, hy^{(1)}(-\tfrac{1}{2}+N; z) - 2MC(-\tfrac{1}{2}-N) \times C(-\tfrac{1}{2}+N)\, hy^{(2)}(-\tfrac{1}{2}+N; z)\}, \tag{9.9a}$$

$$hy^{(2)}(-\tfrac{1}{2}+N; z\exp(2M\pi i)) = (-1)^M\{2MC(-\tfrac{1}{2}-N) \times C(-\tfrac{1}{2}+N)\, hy^{(1)}(-\tfrac{1}{2}+N; z) + [1+2MC(-\tfrac{1}{2}-N)\, C(-\tfrac{1}{2}+N)]\, hy^{(2)}(-\tfrac{1}{2}+N; z)]. \tag{9.9b}$$

Alternatively, Eqs. (9.9) can be obtained from Eqs. (7.10) by a limiting process.

D. Neumann-type solutions for L equal to an integer

If L becomes equal to an integer, some complications arise due to the fact that the connection coefficients $D(-L-1)$ and $D^*(-L-1)$ have simple poles at $L=N$ for $N=0,1,2,\cdots$, but while $D(-L-1)/(\Gamma(-L)$ remains analytic there, this is neither true for $C(-L-1)/\Gamma(-L)$ which is discontinuous in changing its sign nor for $\delta(-L-1)$ which is discontinuous in jumping by π. The reason for this behavior lies in our definition (7.1) of $C(-L-1)$ which for obvious reasons has been chosen so that $C(-L-1)\to +1$ if $G\to 0$, at least for $L\neq N$.

From Eq. (7.8) we find, mod(π),

$$\delta(-N-1) - \delta(N) + \pi/2 = 0, \tag{9.10}$$

and, by means of Eqs. (5.8), (7.3), and (8.2),

$$jy(-N-1; z) = \pm(-1)^N jy(N; z), \tag{9.11}$$

$$ny(-N-1; z) = \pm(-1)^N ny(N; z). \tag{9.12}$$

The sign ambiguities in Eqs. (9.11)–(9.12) are due to the discontinuities of $C(-N-1)$ and $\delta(-N-1)$ which do affect $jy(-N-1; z)$ and $ny(-N-1; z)$.

In order to find the behavior at the origin of $ny(N; z)$ it is again necessary to consider the limit of Eq. (8.3b). In a similar way as in Sec. 9.C but taking account of the complications mentioned above we obtain

$$g_0(N) = (-1)^{N+1} 2^N[\Gamma(\tfrac{1}{2})/\Gamma(\tfrac{1}{2}-N)]\, C(N), \tag{9.13a}$$

$$\alpha(N) = 2^N[\Gamma(\tfrac{1}{2})/\Gamma(\tfrac{1}{2}-N)]\, C(N)\left(-w_{2N-1}(-N-1) + G\sum_{m=0}^{2N}(1/m!)(-\beta)^m w_{2N-m}(-N-1)\right)\Big/(2N+1), \tag{9.13b}$$

$$g_{2N+1}(N) = (1/\pi)\,[d\{\cos[\delta(-L-1) - \delta(L) + \pi(L+\tfrac{1}{2})] \times C(-L-1)\sin(\pi L)\, 2^{-L}[\Gamma(\tfrac{3}{2})/\Gamma(\tfrac{3}{2}+L)] - C(L)\, 2^{L}[\Gamma(\tfrac{1}{2})/\Gamma(\tfrac{1}{2}-L)]\sin(\pi L) \times w_{2N+1}(-L-1)\}/dL]_{L=N}. \tag{9.13c}$$

Again, while the expressions for $g_0(N)$ and $\alpha(N)$ are simple, the expression for $g_{2N+1}(N)$ is complicated, and casting it into a more convenient form would be of little use because of the remaining derivatives of $C(L)$.

From Eqs. (3.1) and (9.4) the circuit relations for the Neumann-type solutions are found to be

$$ny(N; z\exp(2M\pi i)) = ny(N; z) + 2M\pi i\,\gamma(N)\,jy(N; z), \tag{9.14}$$

where

$$\gamma(N) = \alpha(N)\,C(N)\,2^N\,\Gamma(\tfrac{3}{2}+N)/\Gamma(\tfrac{3}{2}) \tag{9.15a}$$

or, because of Eq. (9.13b),

$$\gamma(N) = (-1)^N\left[C(N)\,2^N\Gamma(\tfrac{1}{2}+N)/\Gamma(\tfrac{1}{2})\right]^2 \times\left(-w_{2N-1}(-N-1) + G\sum_{m=0}^{2N}(1/m!) \times(-\beta)^m\,w_{2N-m}(-N-1)\right). \tag{9.15b}$$

By means of Eqs. (8.2) we may now obtain the circuit relations for the Hankel-type solutions

$$hy^{(1)}(N; z\exp(2M\pi i)) = [1 - M\pi\gamma(N)]\,hy^{(1)}(N; z) - M\pi\gamma(N)\exp[-2i\,\delta(N)]\,hy^{(2)}(N; z), \tag{9.16a}$$

$$hy^{(2)}(N; z\exp(2M\pi i)) = M\pi\gamma(N)\exp[2i\,\delta(N)]\,hy^{(1)}(N; z) + [1 + M\pi\gamma(N)]\,hy^{(2)}(n; z). \tag{9.16b}$$

10. THE CASE WHEN $\beta = 2i/Q$ FOR SOME INTEGER Q

A. Q equal to a positive integer

1. General aspects

If the parameter β happens to be such that $Q\beta = 2i$ for some positive integer Q, complications arise at the singular points $t = M\beta + i$ as is shown by the break down of the recurrence relations (4.19c) and (4.20a) for $m - M = Q$. The reason is that then, if $m \geq Q$, the inhomogeneous term of the differential equation for $b_m(t)$ has one of its singular points coinciding with one of the singular points of the homogeneous equation.

The necessary modifications do not touch upon the definition (5.1b) of $hy^{(2)}(L; z)$ and its asymptotic expansion, the coefficients of which remain, in fact, well-defined. The integral representation (5.1a), however, cannot be expected to still represent $hy^{(1)}(L; z)$ after the limit $\beta \to 2i/Q$ has been taken in the integrand. The reason is that the contour of the integral passes in between the two singular points $t = i$ and $t = -i + Q\beta$, which will coalesce when $\beta \to 2i/Q$. The right-hand side of Eq. (5.1a) with $\beta = 2i/Q$ therefore is obtained as the analytical continuation, with respect to β, of the integral with a contour which encloses both of these two singular points. It therefore represents, according to Eqs. (5.12)–(5.13), the solution

$$hy^{(3)}(L; z) = hy^{(1)}(L; z) + G^Q[a_0^{QQ*}(L)/a_0^{Q0}(L)]\,hy^{(2)}(L; z) \tag{10.1}$$

rather than the solution $hy^{(1)}(L; z)$.

Since the modifications begin to occur at $m = Q$, let us consider $b_Q^I(t)$, which according to Eq. (5.2b) with $Q\beta = 2i$ is

$$b_Q^I(t) = (t-3i)^{-L-1}(t-i)^{-L-1} \times\int_t^\infty b_{Q-1}^I(T)\,(T-3i)^L\,(T-i)^L\,dT. \tag{10.2}$$

The singular points of $b_{Q-1}^I(T)$ are at $T = \pm i + 2ni/Q$ where $n = 0, 1, 2, \ldots, Q-1$, and therefore $T = i$ is among them while $T = 3i$ is not. The expansion of $b_{Q-1}^I(T)$ around $T = i$ has singular terms proportional to $a_n^{Q-1\,0}(L)(T-i)^{-L-1+n}$, so that the integrand of Eq. (10.2) has a term with $a_0^{Q-1\,0}(L)(T-i)^{-1}$. Since $a_0^{Q-1\,0}(L) = 0$ unless $Q = 1$ according to Eq. (4.23), the analytical structure of $b_Q^I(t)$ remains unchanged for $Q = 2, 3, 4, \cdots$, but logarithmic terms do occur for $Q = 1$. As a further consequence,

$$a_0^{QQ*}(L) = \int_i^\infty b_{Q-1}^I(T)(T-3i)^L(T-i)^L\,dT \tag{10.3}$$

remains well defined for $Q = 2, 3, 4, \cdots$, but does not exist for $Q = 1$. More generally, the same statements are true for $b_{Q+N}^I(t)$ and $a_0^{Q+N\;Q+N*}(N)$, respectively $(N = 1, 2, 3, \cdots)$. We will therefore treat the different cases $Q = 2, 3, 4, \cdots$ or $Q = 1$ in separate subsections.

2. $Q = 2, 3, 4, \ldots$

The modifications in this case are not so drastic, since the analytical structure of the $b_{Q+N}(t)$ remains the same as in Sec. 4.C. The recurrence relation (4.19c) for $m = Q + N$ and $M = N$ with $N = 0, 1, 2, \cdots$ has simply to be replaced by

$$a_n^{Q+N\;N}(L) = (2ni)^{-1}[(n+L)\,a_{n-1}^{Q+N\;N}(L) + a_n^{Q+N-1\;N}(L)], \tag{10.4}$$

so that the original recurrence relation, from which Eq. (4.19c) followed after division by a now vanishing factor, is satisfied for arbitrary $a_n^{Q+N\;N}(L)$, and this is true even for $n = 1$ since $a_0^{Q+N-1\;N}(L) = 0$ for $Q = 2, 3, 4, \cdots$. Similar modifications occur for the $c_n^{Q+N\;N}(L)$. As to the initial coefficients, Eq. (4.23) no longer holds when $m = Q + N$ and $M = N$, but we now have

$$a_0^{Q+N\;N}(L) = -\exp(i\pi L)\,a_0^{Q+N\;Q+N\,*}(L) = -\exp(i\pi L)\int_{N\beta+i}^\infty b_{Q+N-1}^I(T) \times(T-N\beta-i)^L(T-N\beta-3i)^L\,dT. \tag{10.5}$$

This equation is a direct consequence of the fact that the singular points $t = N\beta + i$ and $t = (N+Q)\beta - i$ do coincide. Since for $Q = 2, 3, 4, \cdots$ the integral (10.5) exists, $a_0^{QQ*}(L)$ is well defined and we may solve Eq. (10.1) for $hy^{(1)}(L; z)$, which therefore is also well defined. The asymptotic expansions of $hy^{(1)}(L; z)$ or $hy^{(3)}(L; z)$ have the same structure as in Sec. 5.B. It is only the coefficients p_{Qn} which have to be modified as a consequence of Eqs. (10.4) and (10.5) with $N = 0$. In place of Eq. (5.7a) with $m = Q$ we now have

$$p_{Qn} = (2ni)^{-1}[-(n+l)(n-l-1)\,p_{Q\,n-1} + p_{Q-1\,n}], \tag{10.6}$$

with the initial coefficient

$$p_{Q0} = 0 \tag{10.7a}$$

as in Sec. 5.B in case of $hy^{(1)}(L; z)$ or

$$p_{Q0} = -\exp(i\pi L)\, a_0^{QQ*}(L)/a_0^{00}(L) \tag{10.7b}$$

in case of $hy^{(3)}(L;z)$. Changing the initial coefficient p_{Q0} is equivalent to adding a multiple of $hy^{(2)}(L;z)$. It should be noted that if a solution with $p_{Q0}\neq 0$ like $hy^{(3)}(L;z)$ is considered and its asymptotic expansion written in the form of Eq. (5.6a), the sum over n has to start at $n=0$ rather than $n=m$.

3. $Q=1$

In order to see what happens in this case we consider the Eq. (5.2b) for $b_m^I(t)$ with $m=Q=1$ and split the path of integration into three parts so that

$$J = \int_t^{\beta-i} + \int_{\beta-i}^{\beta+i} + \int_{\beta+i}^{\infty} = J_1 + J_2 + J_3 \tag{10.8a}$$

with

$$J = \int_t^{\infty} (T-i)^{-L-1}(T+i)^{-L-1}(T-\beta-i)^L (T-\beta+i)^L\, dT. \tag{10.8b}$$

By means of well-known formulas [Ref. 7, Eqs. (3.211) and (9.182.1) or (3.197.1) and (9.131.1)] there are several ways, none particularly convenient, to show that

$$J_2 = (i/2)(4/\beta^2)^{L+1}[\Gamma(L+1)\Gamma(L+1)/\Gamma(2L+2)] \times {}_2F_1(L+1,\ L+1,\ 2L+2;\ -4/\beta^2). \tag{10.9}$$

The other two integrals J_1 and J_3 are more difficult to evaluate, but it suffices to obtain those terms which do not vanish (or even are singular) when $\beta\to 2i$, for $t-i$ different from but in the vicinity of zero. We therefore put

$$\beta = 2i + \epsilon \tag{10.10a}$$

$$t = i + \eta. \tag{10.10b}$$

It suffices, and is appropriate in view of the branches required, to consider values of ϵ which are real and positive. The integral J_3 remains well defined when $\epsilon\to 0$ and is found [Ref. 7, Eq. (3.197.1); Ref. 8, Eqs. (15.1.23) or (15.1.28)] to be

$$J_3 \sim (4i)^{-1}[\psi(L/2+1) - \psi(L/2+\tfrac{1}{2})], \quad \epsilon\to 0. \tag{10.11}$$

The behavior when $\epsilon\to 0$ of the integral J_2 can be seen from Eq. (10.9) by means of a suitable transformation formula of the hypergeometric function [Ref. 8, Eq. (15.3.10)],

$$J_2 \sim \exp[-i\pi(L+\tfrac{1}{2})][\psi(1) - \psi(L+1) + i\pi/4 - (\tfrac{1}{2})\ln(\epsilon)]. \tag{10.12}$$

Introducing for J_1 a new variable of integration $s = T - i$ we have

$$J_1 = (\tfrac{1}{2})\exp[-i\pi(L+\tfrac{1}{2})] \times \int_\eta^\epsilon s^{-L-1}[1+s/(2i)]^{-L-1} \times [1+(\epsilon-s)/(2i)]^L (s-\epsilon)^L\, ds. \tag{10.13}$$

Since we need the behavior of J_1 for $\epsilon\to 0$ while $|\eta|<2$, it is reasonable to expand the second and third factor of the integrand by means of the binomial theorem. We then have to consider, for $m,n=0,1,2,\cdots$, the integrals

$$J_4(m,n) = \int_\eta^\epsilon s^{-L-1+n}(s-\epsilon)^{L+m}\, ds. \tag{10.14}$$

For $m+n=1,2,3,\cdots$ we immediately have

$$J_4(m,n) \sim -\eta^{m+n}/(m+n), \quad \epsilon\to 0, \tag{10.15}$$

but for $m=n=0$ we obtain

$$J_4(0,0) = -(L+1)^{-1}(1-\epsilon/\eta)^{L+1} \times {}_2F_1(1,\ L+1;\ L+2;\ 1-\epsilon/\eta) \tag{10.16}$$

or, by means of the appropriate transformation formula of the hypergeometric function [Ref. 8, Eq. (15.3.10)],

$$J_4(0,0) \sim \psi(L+1) - \psi(1) + \ln(\epsilon) - \ln(\eta), \quad \epsilon\to 0. \tag{10.17}$$

Collecting the results we see that the $\ln(\epsilon)$ terms cancel so that we may take the limit $\epsilon=0$ and obtain, with certain coefficients γ_n,

$$J = (\tfrac{1}{2})\exp[-i\pi(L+\tfrac{1}{2})]\Big(\psi(1) - \psi(L+1) + i\pi/2 - \ln(\eta) + \sum_{n=1}^{\infty}\gamma_n\eta^n\Big) + (4i)^{-1}[\psi(L/2+1) - \psi(L/2+\tfrac{1}{2})]. \tag{10.18}$$

On the basis of this result we see that at the singular point $t=i$ the analytical structure of $b_1^I(t)$, or more generally of the $b_{1+N}^I(t)$ for $N=0,1,2,\cdots$, is given by

$$b_{1+N}^I(t) = 2^{-L-1}\exp[-i\pi(L+1)/2]\Big(\sum_{n=0}^{\infty} f_n^{1+N\,0}(L)(t-i)^{-L-1+n} + \sum_{n=0}^{\infty} a_n^{1+N\,1*}(L)(t-i)^{-L-1+n}\ln(t-i)\Big) + \sum_{n=0}^{\infty} c_n^{1+N\,1*}(L)(t-i)^n \tag{10.19}$$

with the initial coefficients

$$f_0^{10}(L) = a_0^{00}(L)(4i)^{-1}\{2\psi(L+1) - 2\psi(1) - i\pi - \exp(i\pi L)[\psi(L/2+1) - \psi(L/2+\tfrac{1}{2})]\}, \tag{10.20a}$$

$$a_0^{11*}(L) = -(i/2)\,a_0^{00*}(L) = -(i/2)\,a_0^{00}(L), \tag{10.20b}$$

$$f_0^{1+N\,0}(L) = a_0^{1+N\,1*}(L) = 0 \quad \text{for } N=1,2,3,\cdots, \tag{10.20c}$$

$$c_n^{11*}(L) = 0 \quad \text{for } n=0,1,2,\cdots, \tag{10.20d}$$

and $c_0^{1+N\,1*}(L)$ according to Sec. 4.C for $N=1,2,3,\cdots$. Inserting the expansion (10.19) into the differential equation (4.16) we find that, while the coefficients $c_n^{1+N\,1*}(L)$ and $a_n^{1+N\,1*}(L)$ obey the same recurrence relations as in Sec. 4.C, the coefficients $f_n^{1+N\,0}(L)$ can be obtained from the recurrence relations

$$f_n^{10}(L) = (2ni)^{-1}[(n+L)f_{n-1}^{10}(L) - 2ia_n^{11*}(L) + a_{n-1}^{11*}(L) + f_n^{00}(L)], \tag{10.21a}$$

$$f_n^{1+N\,0}(L) = [-4N(N+1)(n-L-1)]^{-1}[2i(2N+1) \times (n-1)f_{n-1}^{1+N\,0}(L) - (n+L-1)f_{n-2}^{1+N\,0}(L) + 4N(N+1)a_n^{1+N\,1*}(L) + 2i(2N+1)a_{n-1}^{1+N\,1*}(L) - a_{n-2}^{1+N\,1*}(L) - f_{n-1}^{N0}(L)], \tag{10.21b}$$

where

$$f_n^{00}(L) = a_n^{00}(L). \tag{10.21c}$$

Because of Eqs. (4.23) and (10.20b) we have

$$a_n^{1+N\,1*}(L) = -(i/2)\,a_n^{N0*}(L). \tag{10.21d}$$

We are now prepared to obtain the asymptotic expansion of $hy^{(3)}(L;z)$ in a similar way as in Sec. 5.B. The integrals with the logarithmic terms

$$\int_{C_0^+} \exp(zt)\,(t-i)^{-L-1+n}\ln(t-i)\,dt$$
$$= -(d/dL)\int_{C_0^+} \exp(zt)\,(t-i)^{-L-1+n}\,dt$$
$$= -(d/dL)[\exp(iz)\,z^{L-n}(2\pi i)/\Gamma(L+1-n)]$$

yield a factor $\ln(1/z)+\psi(L+1-n)$ as compared with the usual integrals not containing the factor $\ln(t-i)$. Introducing, for $N=0,1,2,\cdots$, the new coefficients

$$q_{1+N\,n} = [\Gamma(L+1)/\Gamma(L+1-n)]\,[f_n^{1+N\,0}(L) - (i/2)\,\psi(L+1-n)\,a_n^{N0*}(L)]/a_0^{00}(L), \tag{10.22}$$

we obtain the asymptotic expansion

$$hy^{(3)}(L;z) \sim z^{-1}\exp(iz)\exp[-i\pi(L+1)/2]\sum_{m=0}^{\infty} G^m \exp(-2miz) \times \sum_{n=0}^{\infty} q_{mn} z^{-n} - (i/2)\,G z^{-1}\exp(-iz)\exp[i\pi(L+1)/2] \times \sum_{m=0}^{\infty} G^m \exp(-2miz)\sum_{n=m}^{\infty} p_{mn}^* z^{-n}\ln(1/z), \tag{10.23}$$

$$-\pi < \arg(z) < \pi, \quad z\to\infty,$$

with the initial coefficients

$$q_{00} = p_{00}^* = 1, \tag{10.24a}$$

$$q_{10} = (4i)^{-1}\{4\psi(L+1) - 2\psi(1) - i\pi - \exp(i\pi L)\,[\psi(L/2+1) - \psi(L/2+\tfrac{1}{2})]\}. \tag{10.24b}$$

While the coefficients

$$q_{0n} = p_{0n} \tag{10.25a}$$

and p_{mn}^* are those of Sec. 5.B with $\beta = 2i$, Eq. (5.7a) does not apply when $m = 1+N$, but because of Eq. (10.22) we have

$$q_{1n} = -(2ni)^{-1}\,[(n+l)(n-l-1)\,q_{1\,n-1} - q_{0n} + i(n-\tfrac{1}{2})\,p_{0\,n-1}^* + p_{0n}^*)] \tag{10.25b}$$

and

$$q_{1+N\,n} = [4N(1+N)]^{-1}\,[2i(2N+1)(n-1)\,q_{1+N\,n-1} + (n-1-l)(n-2-l)\,q_{1+N\,n-2} - q_{N\,n-1} - i(n-\tfrac{3}{2})\,p_{N\,n-2}^* + (2N+1)\,p_{N\,n-1}^*] \tag{10.25c}$$

for $N=1,2,3,\cdots$, with the coefficients equal to zero whenever an index becomes negative. Again, these recurrence relations can be derived in two different ways, as mentioned in Sec. 5.B.

The terms multiplied by $\ln(1/z)$ in Eq. (10.23) represent the asymptotic expansion of $-(i/2)\,G\,hy^{(2)}(L;z)$. Similarly, changing the initial coefficient q_{10} would correspond to adding a multiple of $hy^{(2)}(L;z)$.

Since $a_0^{11*}(L)$ does not exist when $\beta = 2i$, we cannot solve Eq. (10.1) for $hy^{(1)}(L;z)$, which also does not exist. In fact, the asymptotic expansion of $hy^{(1)}(L;z)$ does not exist for $\beta \to 2i$ or, when multiplied by an appropriate factor such that all the terms remain finite for $\beta \to 2i$, becomes proportional to the asymptotic expansion of $hy^{(2)}(L;z)$.

B. Q equal to a negative integer

1. General aspects

If $Q = -R$ is a negative integer, the complications arise at the singular points $t = M\beta - i$. Consequently the necessary modifications do not touch upon the definition (5.1a) of $hy^{(1)}(L;z)$ and its asymptotic expansion. The integral representation (5.1b), however, with the limit $\beta \to -2i/R$ taken in the integrand, represents the solution

$$hy^{(4)}(L;z) = hy^{(2)}(L;z) + G^R[a_0^{RR}(L)/a_0^{00}(L)]\,hy^{(1)}(L;z) \tag{10.26}$$

rather than the solution $hy^{(2)}(L;z)$.

2. $R = 2, 3, 4, \ldots$

For $R = 2,3,4,\cdots$ the asymptotic expansions of $hy^{(2)}(L;z)$ or $hy^{(4)}(L;z)$ have the same structure as in Sec. 5.B. It is only the coefficients p_{Rn}^* which have to be modified and are now given by

$$p_{Rn}^* = (-2ni)^{-1}\,[-(n+l)(n-l-1)\,p_{R\,n-1}^* + p_{R-1\,n}^*] \tag{10.27}$$

with the initial coefficient

$$p_{R0}^* = 0 \tag{10.28a}$$

as in Sec. 5.B in case of $hy^{(2)}(L;z)$ or

$$p_{R0}^* = -\exp(-i\pi L)\,a_0^{RR}(L)/a_0^{00}(L) \tag{10.28b}$$

in case of $hy^{(4)}(L;z)$.

3. $R = 1$

For $R = 1$, $hy^{(2)}(L;z)$ does not exist. The asymptotic expansion of $hy^{(4)}(L;z)$ then is

$$hy^{(4)}(L;z) = z^{-1}\exp(-iz)\exp[i\pi(L+1)/2] \times \sum_{m=0}^{\infty} G^m \exp(2miz)\sum_{n=0}^{\infty} q_{mn}^* z^{-n} + (i/2) G\, z^{-1} \times \exp(iz)\exp[-i\pi(L+1)/2] \times \sum_{m=0}^{\infty} G^m \exp(2miz)\sum_{n=m}^{\infty} p_{mn} z^{-n}\ln(1/z), \tag{10.29}$$

$$-\pi < \arg(z) < \pi, \quad z\to\infty,$$

with the initial coefficients

$$q_{00}^* = p_{00} = 1, \tag{10.30a}$$

$$q_{10}^* = (-4i)^{-1}\{4\psi(L+1) - 2\psi(1) + i\pi - \exp(-i\pi L)\,[\psi(L/2+1) - \psi(L/2+\tfrac{1}{2})]\}. \tag{10.30b}$$

While the coefficients

$$q_{0n}^* = p_{0n}^* \tag{10.31a}$$

and p_{mn} are those of Sec. 5.B with $\beta = -2i$, we have for the new coefficients

$$q_{1n}^* = (2ni)^{-1}\,[(n+l)\,(n-1-l)\,q_{1\,n-1}^* - q_{0n}^* - i(n-\tfrac{1}{2})\,p_{0\,n-1} + p_{0n}] \tag{10.31b}$$

and

$$q_{1+N\,n}^* = [4N(1+N)]^{-1}\,[-2i(2N+1)(n-1)\,q_{1+N\,n-1}^* + (n-1+l)(n-2-l)\,q_{1+N\,n-2}^* - q_{N\,n-1}^* + i(n-\tfrac{3}{2})\,p_{N\,n-2} + (2N+1)\,p_{N\,n-1}] \tag{10.31c}$$

for $N=1,2,3,\cdots$.

The more detailed comments missing here are quite analogous to those of Sec. 10.A.

11. APPLICATION TO POTENTIAL SCATTERING

A. Convergence of the expression for the S matrix

The main results of this paper may immediately be applied to the scattering of particles by a Yukawa potential, for it is simply the ratio of the connection coefficients $D(l)$ and $D^*(l)$ which determines the S matrix

$$S=D(l)/D^*(l)=\exp[2i\delta(l)] \tag{11.1}$$

and the scattering phase shift which is equal to our quantity $\delta(l)$ of Eq. (7.2). Since in case of the scattering problem $\beta=\mu/k$ is real and positive, the expressions (6.18) for the connection coefficients converge without any further restriction on β. That they do not converge for $\beta=0$ is related to the well-known pathology of the Coulomb scattering problem. Furthermore we note that the expressions for the connection coefficients, which appear as expansions in powers of $G=g/k$, converge for arbitrarily large coupling constants g (if $|g|<\infty$) or arbitrarily small momenta $k>0$. Since we have $I(L)=1$ if β is real and $\mathrm{Re}(L)\geq 0$, it follows from the estimate (6.14n) and Eq. (6.13) that, for $\mathrm{Re}(l)\geq 0$, the rate of (absolute) convergence is comparable to the rate of (absolute) convergence of the expansion of the exponential function

$$\exp(G/\beta)=\exp(g/\mu)=1+g/\mu+(g/\mu)^2/2!+\cdots, \tag{11.2}$$

irrespective of the value of k.

B. First Born approximation

It might be of interest to see how the well-known first Born approximation value of the S matrix is reproduced by our method. Assuming that G is sufficiently small we may expand S in powers of G and obtain to first order in G, using Eqs. (6.13) and (11.1),

$$S_{\mathrm{Born}}-1=G(a_0^{11}(l)-a_0^{11*}(l))/a_0^{00}(l) \tag{11.3}$$

or, because of Eqs. (5.2a) and (6.10e)–(6.10f),

$$S_{\mathrm{Born}}-1=-G\int_{\beta-i}^{\beta+i}(T^2+1)^{-l-1}[(T-\beta)^2+1]^l\,dT. \tag{11.4}$$

Since this integral appeared already in Sec. 10, we obtain immediately, by means of Eq. (10.9),

$$S_{\mathrm{Born}}-1=-G(i/2)(4/\beta^2)^{l+1}[\Gamma(l+1)\,\Gamma(l+1)/\Gamma(2l+2)]\times{}_2F_1(l+1,\ l+1;\ 2l+2;\ -4/\beta^2). \tag{11.5}$$

Using a suitable transformation formula for the hypergeometric function we may express this result as a Legendre function of the second kind,

$$S_{\mathrm{Born}}-1=-iGQ_l(1+\beta^2/2), \tag{11.6}$$

which, by means of a table of integrals of Bessel functions,[7] can be represented by an integral

$$S_{\mathrm{Born}}-1=-2iG\int_0^\infty (j_l(z))^2\, z\exp(-\beta z)\,dz, \tag{11.7}$$

so that finally the familiar first Born approximation result

$$S_{\mathrm{Born}}-1=2ik\int_0^\infty (j_l(kr))^2[-gr\exp(-\mu r)]\,dr \tag{11.8}$$

is obtained.

According to Eqs. (11.5) or (11.6), the first Born approximation value of the S matrix diverges when $\beta\to\pm 2i$, that is when $k\to\pm\mu/(2i)$. This fact is related to the appearance of logarithmic terms in the asymptotic behavior of the wavefunction as explained in Sec. 10.

A more extended discussion of potential scattering or of the bound state problem is beyond the scope of the present paper.

ACKNOWLEDGMENT

I am grateful to Dr. F. Naundorf for calling my attention to Ref. 2, which stimulated me to find a solution of this old problem.

[1]M. Danos, in *Studies in Applied Mathematics* 6, edited by D. Ludwig and F. W. J. Olver (Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1970).
[2]K. D. Shere, J. Math. Phys. **12**, 78–82 (1971).
[3]A. Ahmadzadeh, P. G. Burke, and C. Tate, Phys. Rev. **131**, 1315–9 (1963).
[4]E. L. Ince, *Ordinary Differential Equations* (Dover, New York, 1956).
[5]W. Bühring, ZAMM **56**, T245–T246 (1976).
[6]F. W. J. Olver, *Asymptotics and Special Functions* (Academic, New York, 1974).
[7]I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Academic, New York, 1965).
[8]M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965).

An infinite number of conservation laws for coupled nonlinear evolution equations

Richard Haberman

Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08903
(Received 12 July 1976; revised manuscript received 27 August 1976)

The n-dimensional Zakharov–Shabat eigenvalue problem and the corresponding time dependency of the vector eigenfunctions are considered. It is known that certain coupled systems of nonlinear partial differential equations are equivalent to the time invariance of the spectrum. Here, for any such coupled system, an infinite sequence of conservation laws is explicitly derived. As an example, this result is applied to the equations describing three resonantly interacting nonlinear wave envelopes.

1. INTRODUCTION

The inverse-scattering transform method was first discovered by Gardner, Greene, Kruskal, and Miura[1,2] in their now classic investigation of the Korteweg–de Vries equation. The method involves showing that a nonlinear evolution equation is equivalent to the consistency of a time-invariant linear eigenvalue problem (which, for the Korteweg–de Vries equation, was the Schrödinger eigenvalue problem) and a linear equation for the time evolution of the eigenfunction. Zakharov and Shabat[3] (based on some ideas of Lax[4]) considered an eigenvalue problem consisting of two coupled first order differential equations, enabling them to analyze the cubic nonlinear Schrödinger equation. Subsequently, by utilizing the Zakharov–Shabat eigenvalue problem, Ablowitz, Kaup, Newell, and Segur[5,6] generated a wide class of nonlinear partial differential equations. The coupled equations of a nonlinearly interacting resonant triad of wave envelopes were analyzed by Zakharov and Manakov[7] using the inverse-scattering method. This motivated Ablowitz and Haberman[8] to consider the eigenvalues ζ of the following system of differential equations:

$$\frac{\partial \mathbf{V}}{\partial x}=i\zeta\mathbf{D}\mathbf{V}+\mathbf{N}\mathbf{V}, \tag{1.1}$$

where appropriate homogeneous boundary conditions are prescribed (for example, $\mathbf{V}$ is bounded as $x\to\pm\infty$). The time dependency of the n-dimensional eigenfunction $\mathbf{V}$ is chosen such that

$$\frac{\partial \mathbf{V}}{\partial t}=\mathbf{Q}\mathbf{V}, \tag{1.2}$$

where $\mathbf{D}$, $\mathbf{N}$, and $\mathbf{Q}$ are $n\times n$ matrices (we assume $N_{ii}=0$ and $\mathbf{D}$ is constant and diagonal). The consistency of (1.1) and (1.2) and the time invariance of the eigenvalues imply that the generalized potentials $\mathbf{N}$ evolve according to a system of nonlinear partial differential equations. This incorporates the previously obtained results.

One of the fundamental properties of the nonlinear partial differential equations which have been analyzed by this inverse-scattering method is that they have an infinite sequence of conservation laws. This has been shown by a variety of different methods for all the equations which correspond to a second order spectral problem. Here we will show this for any system of nonlinear partial differential equations which correspond to (1.1) and (1.2).

2. CONSERVATION LAWS

We show that (for each i) an infinite number of conservation laws all follow from the rather trivial conservation law,

$$\frac{\partial}{\partial x}\left[\frac{\partial}{\partial t}(\ln V_{ii})\right]=\frac{\partial}{\partial t}\left[\frac{\partial}{\partial x}(\ln V_{ii})\right]$$

or

$$\frac{\partial}{\partial x}\left(\frac{\partial V_{ii}/\partial t}{V_{ii}}\right)=\frac{\partial}{\partial t}\left(\frac{\partial V_{ii}/\partial x}{V_{ii}}\right), \tag{2.1}$$

where V_{ki} is the kth component of the ith linearly independent eigenvector $\mathbf{V}^{(i)}$ ($i=1,\dots,n$).

By using (1.1) and (1.2), the latter form of (2.1) becomes

$$\frac{\partial}{\partial x}\left(\frac{\sum_k Q_{ik}V_{ki}}{V_{ii}}\right)=\frac{\partial}{\partial t}\left(\frac{\sum_k N_{ik}V_{ki}}{V_{ii}}\right).$$

This suggests that the matrix Γ be introduced, equal to the ratio of the components of *an* eigenvector

$$\Gamma_{ki}\equiv V_{ki}/V_{ii}\quad(\Gamma_{ii}\equiv 1).$$

Thus

$$\frac{\partial}{\partial x}\left(\sum_k Q_{ik}\Gamma_{ki}\right)=\frac{\partial}{\partial t}\left(\sum_k N_{ik}\Gamma_{ki}\right). \tag{2.2}$$

We obtain an infinite number of conservation laws by considering the asymptotic expansion of the eigenfunctions as $\zeta\to\infty$. This idea was first developed by Gardner, Greene, Kruskal, and Miura.[2,9] However, our approach here generalizes to $n\times n$ systems the method used for 2×2 systems by Konno, Sanuki, Ichikawa, and Wadati.[10–12] Using the Liouville–Green expansion for large ζ of (1.1) (see Appendix), the leading order asymptotic behavior as $\zeta\to\infty$ of V_{ki} is

$$V_{ki}=A_i(\zeta,t)\exp(i\zeta d_i x)[\delta_{ik}+O(1/\zeta)],$$

if the ith eigenvector is chosen in an appropriate manner. In this way we see that

$$\Gamma_{ki}=O(1/\zeta)\quad\text{for}\quad k\neq i. \tag{2.3}$$

Although the asymptotic expansion of V_{ki} is often easier, involving linear equations (see the Appendix), the asymptotic expansion of Γ_{ki} is needed and will be calculated shortly. In fact, since $\mathbf{Q}$ is assumed to be known as a function of ζ and $\mathbf{N}$, *the asymptotic expansion of* Γ_{ki}

yields, directly from (2.2), *an infinite sequence of conservation laws.*

We now obtain the asymptotic expansion of Γ_{ki} without dividing two asymptotic expansions. Using (1.1) and the definition of Γ_{ki} yields

$$\frac{\partial}{\partial x}\Gamma_{ki}=i\zeta(d_k-d_i)\Gamma_{ki}+\sum_m(N_{km}\Gamma_{mi}-\Gamma_{ki}N_{im}\Gamma_{mi}). \quad (2.4)$$

The differential equation for Γ_{ki} contains quadratic nonlinearities of the Riccati type (rather than being linear as are the equations for V_{ki}). For $k\neq i$,

$$\Gamma_{ki}=\frac{(\partial/\partial x)\Gamma_{ki}-\sum_m(N_{km}\Gamma_{mi}-\Gamma_{ki}N_{im}\Gamma_{mi})}{i\zeta(d_k-d_i)}. \quad (2.5)$$

The asymptotic expansion of Γ_{ki} for $k\neq i$,

$$\Gamma_{ki}\sim\frac{\Gamma_{ki}^{(1)}}{i\zeta(d_k-d_i)}+\frac{\Gamma_{ki}^{(2)}}{[i\zeta(d_k-d_i)]^2}+\cdots=\sum_{p=1}^{\infty}\frac{\Gamma_{ki}^{(p)}}{[i\zeta(d_k-d_i)]^p}, \quad (2.6)$$

can be explicitly determined after substituting (2.6) into (2.5). Recalling (2.3), we see

$$\Gamma_{ki}^{(1)}=-N_{ki},$$
$$\Gamma_{ki}^{(2)}=-\frac{\partial}{\partial x}N_{ki}+\sum_m N_{km}N_{mi}. \quad (2.7)$$

A recursive formula is easily obtained;

$$\Gamma_{ki}^{(p+1)}=\frac{\partial}{\partial x}\Gamma_{ki}^{(p)}-\sum_{m\neq i}N_{km}\Gamma_{mi}^{(p)}+\sum_m N_{im}\left(\sum_{s=1}^{p-1}\Gamma_{ki}^{(p-s)}\Gamma_{mi}^{(s)}\right)$$

for $p\geq 1$, (2.8)

where the last term is to be ignored for $p=1$. One notes that part of the conserved quantity, $\sum_{k\neq i}N_{ik}\Gamma_{ki}^{(p)}/[i(d_k-d_i)]^p$, is a $(p+1)$-degree polynomial in the generalized potentials **N** (the rest, of lower degree, depends also on derivatives of **N**).

3. THREE-WAVE INTERACTIONS

As an example, we will determine the infinite sequences of conservation laws for the three-wave interaction equations. By cross differentiation of (1.1) and (1.2), Ablowitz and Haberman[8] showed that

$$Q_{ik}=\alpha_{ik}N_{ik}+c_k\delta_{ik}\zeta \quad (3.1)$$

implies that the generalized potentials evolve according to

$$\frac{\partial}{\partial t}N_{ij}=\alpha_{ij}\frac{\partial}{\partial x}N_{ij}+\sum_{k\neq i,j}(\alpha_{ik}-\alpha_{kj})N_{ik}N_{kj}, \quad (3.2)$$

where, for $i\neq k$, $\alpha_{ik}=(c_i-c_k)/i(d_i-d_k)$. [These equations, (3.2), represent a resonant triad of interacting waves if we discuss 3×3 matrices and if $N_{jk}=\sigma_{jk}N_{kj}^*$ for $j>k$ with $\sigma_{ij}\sigma_{jk}=-\sigma_{ik}$ for $i>j>k$. This result is equivalent to the work of Zakharov and Manakov.[7] The conservation laws can be derived without these limitations, however.]

By substituting (3.1) into (2.2), using asymptotic expansion (2.6), we determine that the leading order nonzero term is $O(1/\zeta)$ and

$p\geq 1$, $O(1/\zeta^p)$:

$$\frac{\partial}{\partial t}\left[\sum_{k\neq i}N_{ik}\frac{\Gamma_{ki}^{(p)}}{(d_k-d_i)^p}\right]=\frac{\partial}{\partial x}\left[\sum_{k\neq i}\alpha_{ik}N_{ik}\frac{\Gamma_{ki}^{(p)}}{(d_k-d_i)^p}\right]. \quad (3.3)$$

For each i (for three-wave interactions $i=1,2,3$), (3.3) is the prescription for an infinite sequence of conservation laws ($p=1,2,\ldots,\infty$). The quantity $\sum_{k\neq i}N_{ik}\Gamma_{ki}^{(p)}/(d_k-d_i)^p$ is conserved and its flux is $\sum_{k\neq i}\alpha_{ik}N_{ik}\Gamma_{ki}^{(p)}/(d_k-d_i)^p$. Taking $p=1$ in (3.3) and using (2.7) yields

$$\frac{\partial}{\partial t}\left[\sum_{k\neq i}\frac{N_{ik}N_{ki}}{d_k-d_i}\right]=\frac{\partial}{\partial x}\left[\sum_{k\neq i}\alpha_{ik}\frac{N_{ik}N_{ki}}{(d_k-d_i)}\right]. \quad (3.4)$$

If we discuss triad interacting wave envelopes and thus let

$$\mathbf{N}=\begin{bmatrix}0 & A_1 & A_2\\ \sigma_{21}A_1^* & 0 & A_3\\ \sigma_{31}A_2^* & \sigma_{32}A_3^* & 0\end{bmatrix} \quad \text{with } \sigma_{21}\sigma_{32}=-\sigma_{31},$$

then (3.4) becomes

$$i=1:\ \frac{\partial}{\partial t}\left[\frac{\sigma_{21}}{d_2-d_1}|A_1|^2+\frac{\sigma_{31}}{d_3-d_1}|A_2|^2\right]=\frac{\partial}{\partial x}\left[\frac{\alpha_{21}\sigma_{21}}{d_2-d_1}|A_1|^2+\frac{\alpha_{31}\sigma_{31}}{d_3-d_1}|A_2|^2\right]. \quad (3.5)$$

A similar expression follows for $i=2$ and $i=3$. However, the result for $i=3$ is just a linear combination of the conservation laws which occur for $i=1$ and $i=2$. These "energy-sharing" conservation laws are well known, and can be verified (with some care) directly from (3.2). Further conservation laws can be derived using the result of Sec. 2.

APPENDIX: LIOUVILLE-GREEN ASYMPTOTIC EXPANSION FOR LINEAR SYSTEMS WITH A LARGE PARAMETER

Let us consider the asymptotic expansions for large ζ of the eigenvectors **V** given by (1.1). Although the Liouville–Green method is well known (for example, see Coddington and Levinson[13] or Nayfeh[14]), we rederive the result so as to be clear. Define $\mathbf{V}^{(i)}$ as the eigenvector whose kth component, V_{ki}, is asymptotically proportional to $\exp(i\zeta d_i x)$ for large ζ. Thus let

$$V_{ki}=\exp(i\zeta d_i x)\,W_{ki}.$$

Then from (1.1), W_{ki} satisfies

$$i\zeta(d_k-d_i)\,W_{ki}=\frac{\partial}{\partial x}W_{ki}-\sum_m N_{km}W_{mi}.$$

The asymptotic expansion of W_{ki} follows by the substitution of

$$W_{ki}\sim W_{ki}^{(0)}+\frac{W_{ki}^{(1)}}{\zeta}+\frac{W_{ki}^{(2)}}{\zeta^2}+\cdots.$$

The $O(\zeta)$ equation

$$i(d_k-d_i)W_{ki}^{(0)}=0$$

implies that

$$W_{ki}^{(0)} = A_i \delta_{ki},$$

where usually we think of A_i depending only on x. Here A_i also depends on both ζ and t (for example, the time dependence determined via (1.2) could imply $A_2 = \exp(i\zeta^3 t)$). By considering the next order equation, $O(1)$, the x dependence of A_i may be obtained:

$$i(d_k - d_i) W_{ki}^{(1)} = \frac{\partial}{\partial x} W_{ki}^{(0)} - \sum_m N_{km} W_{mi}^{(0)} = \frac{\partial A_i}{\partial x} \delta_{ki} - A_i N_{ki}.$$

Since $N_{ii} = 0$, $\partial A_i / \partial x = 0$ and

$$W_{ki}^{(1)} = \begin{cases} 0, & k = i, \\ -N_{ki} A_i / i(d_k - d_i), & k \neq i, \end{cases}$$

where $W_{ii}^{(1)} = 0$, because $W_{ii}^{(1)} \neq 0$ is equivalent to the introduction in A_i of $O(1/\zeta)$ term. In a similar way, the higher order terms may be directly calculated. Note that the equations for $W_{ki}^{(p)}$ will be linear. From these results we see that

$$V_{ki} = A_i(\zeta, t) \exp(i\zeta d_i x)[\delta_{ki} + O(1/\zeta)].$$

By considering the variables $\Gamma_{ki} = V_{ki}/V_{ii}$, the as yet unknown expression $A_i(\zeta, t)$ will be of no importance.

[1]C. S. Gardner, J. M. Greene, M. D. Kruskal, and R. M. Miura, Phys. Rev. Lett. 19, 1095 (1967).
[2]C. S. Gardner, J. M. Greene, M. D. Kruskal, and R. M. Miura, Comm. Pure Appl. Math. 27, 97 (1974).
[3]V. E. Zakharov and A. B. Shabat, Zh. Eksp. Teor. Fiz. 61, 118 (1971) [Sov. Phys. JETP 34, 62 (1972)].
[4]P. D. Lax, Comm. Pure Appl. Math. 21, 467 (1968).
[5]M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur, Phys. Rev. Lett. 31, 125 (1973).
[6]M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur, Stud. Appl. Math. 53, 249 (1974).
[7]V. E. Zakharov and S. V. Manakov, Zh. Eksp. Teor. Fiz. Pis. Red. 18, 413 (1973) [Sov. Phys. JETP Lett. 18, 243 (1973)].
[8]M. J. Ablowitz and R. Haberman, J. Math. Phys. 16, 2301 (1975).
[9]R. M. Miura, C. S. Gardner, and M. D. Kruskal, J. Math. Phys. 9, 1204 (1968).
[10]K. Konno, H. Sanuki, and Y. H. Ichikawa, Prog. Theor. Phys. 52, 886 (1974).
[11]H. Sanuki and K. Konno, Phys. Lett. A 48, 221 (1974).
[12]M. Wadati, H. Sanuki, and K. Konno, Prog. Theor. Phys. 53, 419 (1975).
[13]E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations* (McGraw-Hill, New York, 1955).
[14]A. H. Nayfeh, *Perturbation Methods* (Wiley, New York, 1973).

Addendum: A classical perturbation theory [J. Math. Phys. 18, 110 (1977)]

Charles Schwartz

Department of Physics, University of California, Berkeley, California 94720

(Received 13 February 1977)

[3]J. K. Percus, Comm. Pure Appl. Math. 17, 137 (1964).

implies that

$$W_{ki}^{(0)} = A_i \delta_{ki},$$

where usually we think of A_i depending only on x. Here A_i also depends on both ζ and t (for example, the time dependence determined via (1.2) could imply $A_2 = \exp(i\zeta^3 t)$). By considering the next order equation, $O(1)$, the x dependence of A_i may be obtained:

$$i(d_k - d_i) W_{ki}^{(1)} = \frac{\partial}{\partial x} W_{ki}^{(0)} - \sum_m N_{km} W_{mi}^{(0)} = \frac{\partial A_i}{\partial x} \delta_{ki} - A_i N_{ki}.$$

Since $N_{ii} = 0$, $\partial A_i / \partial x = 0$ and

$$W_{ki}^{(1)} = \begin{cases} 0, & k = i, \\ -N_{ki} A_i / i(d_k - d_i), & k \neq i, \end{cases}$$

where $W_{ii}^{(1)} = 0$, because $W_{ii}^{(1)} \neq 0$ is equivalent to the introduction in A_i of $O(1/\zeta)$ term. In a similar way, the higher order terms may be directly calculated. Note that the equations for $W_{ki}^{(p)}$ will be linear. From these results we see that

$$V_{ki} = A_i(\zeta, t) \exp(i\zeta d_i x)[\delta_{ki} + O(1/\zeta)].$$

By considering the variables $\Gamma_{ki} = V_{ki}/V_{ii}$, the as yet unknown expression $A_i(\zeta, t)$ will be of no importance.

[1]C. S. Gardner, J. M. Greene, M. D. Kruskal, and R. M. Miura, Phys. Rev. Lett. 19, 1095 (1967).
[2]C. S. Gardner, J. M. Greene, M. D. Kruskal, and R. M. Miura, Comm. Pure Appl. Math. 27, 97 (1974).
[3]V. E. Zakharov and A. B. Shabat, Zh. Eksp. Teor. Fiz. 61, 118 (1971) [Sov. Phys. JETP 34, 62 (1972)].
[4]P. D. Lax, Comm. Pure Appl. Math. 21, 467 (1968).
[5]M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur, Phys. Rev. Lett. 31, 125 (1973).
[6]M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur, Stud. Appl. Math. 53, 249 (1974).
[7]V. E. Zakharov and S. V. Manakov, Zh. Eksp. Teor. Fiz. Pis. Red. 18, 413 (1973) [Sov. Phys. JETP Lett. 18, 243 (1973)].
[8]M. J. Ablowitz and R. Haberman, J. Math. Phys. 16, 2301 (1975).
[9]R. M. Miura, C. S. Gardner, and M. D. Kruskal, J. Math. Phys. 9, 1204 (1968).
[10]K. Konno, H. Sanuki, and Y. H. Ichikawa, Prog. Theor. Phys. 52, 886 (1974).
[11]H. Sanuki and K. Konno, Phys. Lett. A 48, 221 (1974).
[12]M. Wadati, H. Sanuki, and K. Konno, Prog. Theor. Phys. 53, 419 (1975).
[13]E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations* (McGraw-Hill, New York, 1955).
[14]A. H. Nayfeh, *Perturbation Methods* (Wiley, New York, 1973).

Addendum: A classical perturbation theory [J. Math. Phys. 18, 110 (1977)]

Charles Schwartz

Department of Physics, University of California, Berkeley, California 94720

(Received 13 February 1977)

[3]J. K. Percus, Comm. Pure Appl. Math. 17, 137 (1964).