Scalar waves in the exterior of a Schwarzschild black hole*

S. Persides

University of Thessaloniki, Thessaloniki, Greece
(Received 8 November 1973)

Using the method of separation of variables, we study rigorously scalar waves due to a point source in the exterior of a Schwarzschild black hole. First, a Fourier analysis gives general formulas for the interior and exterior radial wave functions and their relations to solutions of special cases, the Green's function, and the frequency spectrum. Three special cases are examined. Second, Laplace transforms of the field are obtained and their properties are studied. Using the results of the Laplace transformation and some general properties of timelike curves, we prove the following theorem: The time-dependent scalar field of a point source goes to zero outside the horizon as the source falls into the black hole.

1. INTRODUCTION

It has been conjectured that a black hole has no externally measurable asymmetry. This is described colloquially by Wheeler's statement[1] that "a black hole has no hair". To establish such a property we have to prove that (a) a black hole is born without hair and (b) no hair transplant is possible on a black hole. Stating these objectives in a more precise language we must prove that (a) gravitational collapse of a slightly asymmetrical object will form a black hole with no externally measurable asymmetries (axial or spherical depending on whether we have rotation or not) and (b) no physical process can destroy this symmetry (without destroying the black hole). A process that could alter the symmetry is the fall of a test particle into the black hole. The purpose of this paper is (a) to set the necessary mathematical background for an analytically exact study of the waves emitted by a scalar particle moving in the exterior of a Schwarzschild black hole and (b) to prove rigorously that as the particle goes into the black hole, the time-dependent scalar field goes to zero everywhere outside the horizon.

Since this paper differs in attitude from all the previous work published up to now by other authors, it is worthwhile to describe briefly its relation to what has been already done towards proving the "baldness" theorem for black holes. To study the creation of a black hole gravitational collapse with nonspherical perturbations has to start from a specific model for the star.[2,3] Thus sooner or later the complexity of a realistic model forces the use of numerical methods and a computer. A seminumerical treatment by Price[4,5] has shown that scalar, electromagnetic or gravitational test fields produced by a source anchored on the collapsing star radiate away their higher ($l \geqslant s$) multipole moments. This result has been extended to a charged collapsing star by Bicak.[6] According to Price the eventual radiation of all higher multipoles is due to the curvature of space—time represented by an "effective potential" localized around $r \approx 4r_s/3$ (r_s is the Schwarzschild radius). Effective-potential methods have been used also to study the gravitational radiation emitted by a particle falling into a black hole.[7–9] Even phenomena taking place totally in the exterior of the black hole have been studied with an effective potential and numerical integration of a differential equation.[10–13]

In general the numerical techniques have become necessary, because after separation of variables all the above problems lead to a radial wave equation which cannot be solved explicitly.[14] Expansion methods[15–20] have been found appropriate and helpful for qualitative and quantitative studies but they cannot be considered as proof of a certain property. Also because of the inability to solve the radial wave equation, exact studies have been limited essentially to static situations.[21–24] The scalar static field has been studied by the author in detail in a previous paper[25] called hereafter paper I.

In the present paper we limit our treatment to scalar waves generated by a weak point source in the exterior of a Schwarzschild black hole. We neglect the contribution of the source and its waves to the curvature of the space—time and, consequently, we can superimpose solutions and derive fields due to more complicated sources. Since our objectives are to be reached as rigorously as possible we have to rule out the use of an effective potential and numerical techniques. After the separation of the angular dependence by expanding the field in spherical harmonics, we use Fourier or Laplace transforms to eliminate the time variable. Fourier transforms are needed to study the frequency spectrum of the radiation which is directly measurable by instruments. Laplace transforms are needed to relate the evolution of the field with its final value. Both transforms lead to a radial wave equation which has been studied in a previous paper[26] called hereafter paper II. We rely heavily on the properties of the solutions derived in that paper and its notation is used without repetition of definitions and formulas. In Sec. 2 we present in general the Fourier-transform analysis and a few fundamental special situations. In Sec. 3 we present the Laplace-transform method and the most important contribution of this paper, namely, the theorem concerning the radiation of multipole moments when the source falls into the black hole. Finally in Sec. 4 we comment on the achieved results and take a glimpse at future work.

2. FOURIER TRANSFORMS OF SCALAR WAVES

A. General formulation

In this section our objective is to give some explicit formulas for the scalar field $\Psi(t, r, \theta, \varphi)$ which satisfy the second order partial differential equation

$$\Box\Psi \equiv g^{\mu\nu}\Psi_{;\mu\nu} = 4\pi f(t, r, \theta, \varphi), \tag{1}$$

where $f(t, r, \theta, \varphi)$ represents a point source (Greek letters take the values 0, 1, 2, 3). This means that for a given t we have $f(t, r, \theta, \varphi) \neq 0$ only at a single point r', θ', φ'. In Eq. (1) $g^{\mu\nu}$ is the contravariant form of the Schwarzschild metric

$$g_{\mu\nu}=\mathrm{diag}\left[\left(1-\frac{r_s}{r}\right)c^2,\ -\left(1-\frac{r_s}{r}\right)^{-1},\ -r^2,\ -r^2\sin^2\theta\right] \tag{2}$$

and we have neglected the contribution of Ψ to the curvature of space–time (semicolon denotes covariant differentiation with respect to $g_{\mu\nu}$).

It has been shown in paper I that the explicit expression of the source $f(t,r,\theta,\varphi)$ affects the final behavior of the solution on the horizon. Following Misner *et al.*[11] we choose f to be invariant under coordinate transformations

$$f(t,r,\theta,\varphi)=q(u^0c)^{-1}\delta(\mathbf{r}-\mathbf{r}'), \tag{3}$$

where $u^0=dt/dS$ along the particle's trajectory $\mathbf{r}(t)=[r'(t),\ \theta'(t),\ \varphi'(t)]$. The strength q of the source can be a function of time.

If we analyze the field in spherical harmonics ($l=0,1,\cdots,\ m=0,\ \pm 1,\cdots,\ \pm l$)

$$\Psi(t,r,\theta,\varphi)=\sum_{l,m}P_{lm}(t,r)Y_{lm}(\theta,\varphi), \tag{4}$$

where

$$P_{lm}(t,r)=\int Y^*_{lm}(\theta,\varphi)\Psi(t,r,\theta,\varphi)\,d\Omega, \tag{5}$$

then Eq. (1) reduces to

$$\frac{r^2}{c^2}\left(1-\frac{r_s}{r}\right)^{-1}P_{lm,00}-\left[\left(1-\frac{r_s}{r}\right)r^2P_{lm,1}\right]_{,1}+l(l+1)P_{lm}=4\pi q(u^0c)^{-1}\delta(r-r')Y^*_{lm}(\theta',\varphi') \tag{6}$$

(commas denote partial differentiation with x^μ equal to t,r,θ,φ for $\mu=0,1,2,3$). To eliminate the time dependence we introduce the Fourier transform ($k=\omega/c$)[27,28]

$$R_{lm}(r,r_s;k)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}P_{lm}(t,r)\exp(i\omega t)\,dt \tag{7}$$

with inverse

$$P_{lm}(t,r)=\int_{-\infty}^{\infty}R_{lm}(r,r_s;k)\exp(-i\omega t)\,d\omega. \tag{8}$$

We have assumed that $P_{lm}(t,r)$ is a generalized function[29] and, consequently, the Fourier transform exists and its usual properties are preserved. From Eq. (6) we have

$$\frac{d}{dr}\left(r(r-r_s)\frac{dR_{lm}}{dr}\right)+\left(\frac{k^2r^3}{r-r_s}-l(l+1)\right)R_{lm}=f_{lm}(r,r_s;k) \tag{9}$$

where

$$f_{lm}(r,r_s;k)=(-2)\int_{-\infty}^{+\infty}q(u^0c)^{-1}\delta(r-r')Y^*_{lm}(\theta',\varphi')e^{i\omega t}\,dt. \tag{10}$$

If the right-hand side of Eq. (9) is set equal to zero we have the homogeneous radial wave equation. In paper II we have studied in detail six solutions $\mathcal{R}_i (i=1,2,3,4,5,6)$. We have given formulas for the constants K_{ij} defined from the Wronskian and relating any three solutions according to the equation

$$K_{ij}\,\mathcal{R}_k+K_{jk}\,\mathcal{R}_i+K_{ki}\,\mathcal{R}_j=0. \tag{11}$$

We have also given expansions of $\mathcal{R}_i$ in the form of power series near 0 for $\mathcal{R}_1$ and $\mathcal{R}_2$, near r_s for $\mathcal{R}_3$ and $\mathcal{R}_4$ and near $+\infty$ for $\mathcal{R}_5$ and $\mathcal{R}_6$. The expansions found for the solutions $\mathcal{R}_5$ and $\mathcal{R}_6$ are only asymptotic expressions (see Table I below). Near the point $r=r_s$ two other expansions for $\mathcal{R}_3$ and $\mathcal{R}_4$ have been given [see Eqs. (40) and (41) of paper II] as power series of $r^{-1}-r_s^{-1}$. These expansions converge for $r>r_s/2$. Hence, for given r, r_s, and k we can directly evaluate $\mathcal{R}_3$ and $\mathcal{R}_4$ and any other $\mathcal{R}_i$ using Eq. (11). This method is particularly helpful for the evaluation of $\mathcal{R}_5$ and $\mathcal{R}_6$ [also for $R_l^{(i)}$ and $R_l^{(e)}$ defined below by Eqs. (12) and (13)] since no use of asymptotic expansions is made.

TABLE I. Solutions of the homogeneous radial wave equation.

Singular point	Solutions	General term	Region of convergence
$r=0$	$\mathcal{R}_1, \mathcal{R}_2$,	r^n	$r< r_s$
$r=r_s$	$\mathcal{R}_3, \mathcal{R}_4$	$(r-r_s)^n$	$\vert r-r_s\vert < r_s$
$r=+\infty$	$\mathcal{R}_5, \mathcal{R}_6$	r^{-n}	$r=+\infty$

B. Interior and exterior solutions

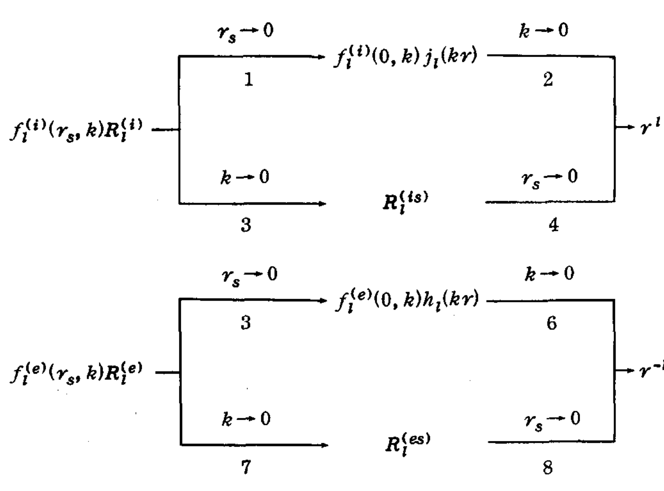
In flat space–time Eq. (9) reduces to the Bessel equation and the interior and exterior solutions are chosen[28] to be $j_l(kr)$ and $h_l(kr)$. In Schwarzschild's space–time let $R_l^{(i)}$ and $R_l^{(e)}$ be the interior and exterior solutions, respectively, for the *time-dependent* field and $R_l^{(is)}$ and $R_l^{(es)}$ for the *static* field ($R_l^{(is)}$ and $R_l^{(es)}$ are explicitly given in terms of Legendre functions and appropriately normalized in paper I). The choice of $R_l^{(i)}$ and $R_l^{(e)}$ is based on physical considerations. The most important requirement is that $R_l^{(i)}$ represent purely ingoing waves as $r\to r_s+$ and $R_l^{(e)}$ represent purely outgoing waves as $r\to+\infty$. Since $\mathcal{R}_4$ behaves as $\exp[-ik(r+r_s\ln|r-r_s|)]$ near $r=r_s$ and $\mathcal{R}_5$ as $\exp[ik(r+r_s\ln|r-r_s|)]$ near $r=+\infty$, we will satisfy these requirements if we choose

$$R_l^{(i)}=-\frac{K_{56}}{K_{45}}\mathcal{R}_4=\frac{1}{2}\left(-\frac{K_{46}}{K_{45}}\mathcal{R}_5+\mathcal{R}_6\right) \tag{12}$$

and

$$R_l^{(e)}=\mathcal{R}_5. \tag{13}$$

DIAGRAM 1. Relations among the solutions of special cases.



Furthermore, we ask that two functions $f_l^{(i)}(r_s, k)$ and $f_l^{(e)}(r_s, k)$ exist such that $f_l^{(i)} R_l^{(i)}$ and $f_l^{(e)} R_l^{(e)}$ go to the flat-space solutions when $r_s \to 0$ and to the static solutions when $k \to 0$. Schematically this is presented in the diagram below. As numbered in the diagram the limits 2 and 6 are easily verified if $f_l^{(i)}(0, k)$ and $f_l^{(e)}(0, k)$ behave as $k^{-l}(2l+1)!!$ and $ik^{l+1}[(2l-1)!!]^{-1}$ as $k \to 0$. The limits 4 and 8 have been given in paper I. The limit 5 has been shown in paper II. The limit 1 would be satisfied if $R_l^{(i)}$ had been chosen proportional to $\mathcal{R}_5 + \mathcal{R}_6$. Our choice [see Eq. (12)] will give the same limit if $K_{45}/K_{46} \to -1$ as $r_s \to 0$. Preliminary considerations have shown this to be true although rigorous proof has not yet been given. The remaining limits 3 and 7 have been established indirectly using the Laplace transform (see Sec. 3A). Obviously $f_l^{(i)}$ and $f_l^{(e)}$ are not unique. Property 2 of Sec. 3A gives a pair of such functions, but the simplest choice has still to be found. Finally, the behavior of $R_l^{(i)}$ near $r=0$ has not been examined analytically but quite probably $R_l^{(i)}$ diverges there as $\ln r$ (it is a linear combination of $\mathcal{R}_1$ and $\mathcal{R}_2$ of paper II).

In what follows we will need the Wronskian of $R_l^{(i)}$ and $R_l^{(e)}$ which is easily found to be (r is the independent variable)

$$W[R_l^{(i)}, R_l^{(e)}] = -\tfrac{1}{2} W[\mathcal{R}_5, \mathcal{R}_6] = \frac{i}{kr(r-r_s)}. \tag{14}$$

As we see the ratio K_{45}/K_{46} does not enter into the Wronskian and will not be needed in what follows.

C. Green's function and special cases

To express the solution of Eq. (9) for some given $f_{lm}(r, r_s; k)$ we are going to use Green's function $G_l(r, \rho, r_s; k)$. The source $f_{lm}(r, r_s; k)$ is different from zero only for $r > r_s$. Thus the integrals expressing the solution in terms of G_l involve integrations only from r_s to $+\infty$. The function $G(r, \rho, r_s; k)$ is the solution of Eq. (9) with right-hand side $\delta(r-\rho)$ and behaving as $R_l^{(i)}$ and $R_l^{(e)}$ at r_s+ and $+\infty$, respectively ($\rho > r_s$). The continuity of G_l at $r=\rho$ gives

$$G_l = A R_l^{(i)}(r_<, r_s; k) R_l^{(e)}(r_>, r_s; k) \tag{15}$$

where A is a constant and $r_<(r_>)$ the smaller (larger) of ρ and r. The discontinuity of dG_l/dr at $r=\rho$ determines A through the relation[28,30]

$$\left.\frac{dG_l}{dr}\right|_{r=\rho+} - \left.\frac{dG_l}{dr}\right|_{r=\rho-} = \frac{1}{\rho(\rho - r_s)}. \tag{16}$$

Using the Wronskian from Eq. (14) we find A and

$$G_l = -ik R_l^{(i)}(r_<, r_s; k) R_l^{(e)}(r_>, r_s; k). \tag{17}$$

The solution of the inhomogeneous Eq. (9) (uniquely determined by the boundary conditions) is now

$$R_{lm}(r, r_s; k) = -ik \int_{r_s}^{\infty} R_l^{(i)}(r_<, r_s; k) R_l^{(e)}(r_>, r_s; k) f_{lm}(\rho, r_s; k) d\rho. \tag{18}$$

We will give explicitly this expression for the following three cases: (i) a particle of constant charge created at $t=0$ and staying at $\mathbf{r}_0 = (r_0, \theta_0, \varphi_0)$ thereafter; (ii) a particle of constant charge created at $t=0$ at the point r_0, θ_0, φ_0 and subsequently falling freely into the black hole; (iii) a particle of constant charge moving permanently on a circular orbit $r' = r_0$, $\theta' = \pi/2$, $\varphi' = \omega_0 t$. Case (i) will be compared later with its Laplace counterpart. Case (ii) is important because we are interested in the properties of the emitted burst of radiation. Case (iii) has been studied lately[11–13] in connection with the possibility of the existence of a mechanism for gravitational synchrotron radiation.

In case (i) we have $r' = r_0$, $\theta' = \theta_0$, $\varphi' = \varphi_0$ and

$$u^0 c = \left(1 - \frac{r_s}{r_0}\right)^{-1/2}, \quad q = q_0 U(t), \tag{19}$$

where $U(t)$ is the unit step function equal to 0 for $t<0$ and 1 for $t>0$. From Eqs. (10) and (18) we have

$$f_{lm}(r, r_s; k) = -2q_0\left(1 - \frac{r_s}{r_0}\right)^{1/2} Y_{lm}^*(\theta_0, \varphi_0)\delta(r - r_0) \times\left(\pi\delta(\omega) + \frac{i}{\omega}\right), \tag{20}$$

$$R_{lm}(r, r_s; k) = 2iq_0 k\left(1 - \frac{r_s}{r_0}\right)^{1/2} Y_{lm}^*(\theta_0, \varphi_0) R_l^{(i)}(r_<, r_s; k) R_l^{(e)}(r_>, r_s; k) \times\left(\pi\delta(\omega) + \frac{i}{\omega}\right), \tag{21}$$

where $r_<(r_>)$ is the smaller (larger) of r, r_0.

In case (ii) we have $\theta' = \theta_0$, $\varphi' = \varphi_0$

$$u^0 c = \left(1 - \frac{r_s}{r_0}\right)^{1/2}\left(1 - \frac{r_s}{r'}\right)^{-1}, \quad q = q_0 U(t). \tag{22}$$

For $r > r_0$ we find $f_{lm} = 0$, while for $r \le r_0$

$$f_{lm}(r, r_s; k) = -2q_0 c^{-1} Y_{lm}^*(\theta_0, \varphi_0)\left(\frac{r_s}{r} - \frac{r_s}{r_0}\right)^{-1/2} \exp(i\omega t_r), \tag{23}$$

where t_r is the unique solution of $r - r'(t) = 0$ corresponding to a given $r \le r_0$. The Fourier transform of $P_{lm}(t, r)$ is

$$R_{lm}(r, r_s; k) = 2iq_0 c^{-1} k Y_{lm}^*(\theta_0, \varphi_0) \int_{r_s}^{r_0} R_l^{(i)}(r_<, r_s; k) R_l^{(e)}(r_>, r_s; k) \times\left(\frac{r_s}{\rho} - \frac{r_s}{r_0}\right)^{-1/2} \exp(it_\rho)\, d\rho. \tag{24}$$

In case (iii) we have $r' = r_0$, $\theta' = \pi/2$, $\varphi' = \omega_0 t$ and

$$u^0 c = [1 - r_s r_0^{-1} - r_0^2 \omega_0^2 c^{-2}]^{-1/2}, \quad q = q_0. \tag{25}$$

Hence

$$f_{lm}(r, r_s; k) = -2q_0(1 - r_s r_0^{-1} - r_0^2\omega_0^2 c^{-2})^{1/2}\, Y_{lm}(\pi/2, 0) \times \delta(r - r_0)\,\delta(\omega - m\omega_0), \tag{26}$$

and

$$R_{lm}(r, r_s; k) = 4\pi i q_0 k(1 - r_s r_0^{-1} - r_0^2\omega_0^2 c^{-2})^{1/2}\, Y_{lm}(\pi/2, 0) \times R_l^{(i)}(r_<, r_s; k) R_l^{(e)}(r_>, r_s; k)\,\delta(\omega - m\omega_0), \tag{27}$$

with $r_<(r_>)$ equal to the smaller (larger) of r and r_0. Equation (27) is obviously the Fourier transform of a periodic function.[11]

Superimposing appropriately these three cases we can calculate the field for other situations. For example, the field of a scalar particle with constant charge q_0 stationary at a given point from $t=-\infty$ to $t=0$ and disappearing at $t=0$ will be the superposition of a static field and of case (i) with charge $-q_0$ instead of q_0. If we add also case (ii) we have the field of a charge q_0 stationary for $t<0$ and starting to fall freely at $t=0$. In all cases a computer program can be written to give the frequency spectrum or the field itself at any point with input the trajectory and the charge of the particle. The computer should be asked to calculate $R_l^{(i)}$ and $R_l^{(e)}$ as described in Sec. 2A, then R_{lm} from Eqs. (21), (24), and (27). This would give us the frequency spectrum. The field can be found from Eqs. (8) and (4).

For a physicist cases (i) and (ii) obviously have a limit as $t\to+\infty$ while the contrary happens for case (iii). Moreover, it can be argued that when t becomes large only the "small ω" components of the spectrum will contribute to the integral (8), because of the factor $\exp(-i\omega t)$ in the integrand. But in the theory of Fourier transforms there is no general mathematical theorem connecting the time-dependent field and its limit as $t\to+\infty$. Such a theorem is available in the theory of Laplace transforms. In Sec. 3 we will exploit that theorem to prove that no lines of force can be planted on a black hole by a falling inwards scalar particle.

3. LAPLACE TRANSFORMS OF SCALAR WAVES

A. Laplace transforms

Our objective in Sec. 3D will be to prove rigorously and under very general conditions that the static field left in the exterior of a Schwarzschild black hole by a scalar charge falling inwards is the zero field. We develop here the necessary Laplace transform of the field starting from Eq. (6). Any field changing only for $t>0$ can be written as superposition of fields which are zero for $t<0$ and static fields. Consequently, we assume that $P_{lm}(t,r)=0$ for all $t\leq 0$. Assuming that $P_{lm}(t,r)$ is piecewise continuous and of exponential order as $t\to+\infty$ so that its Laplace transform exists,[31,32] we set

$$R_{lm}(r,r_s;\lambda)=\int_0^\infty P_{lm}(t,r)\exp(-st)\,dt. \tag{28}$$

We have from Eq. (6)

$$\frac{d}{dr}\left(r(r-r_s)\frac{dR_{lm}}{dr}\right)+\left(\frac{\lambda^2r^3}{r-r_s}-l(l+1)\right)R_{lm}=f_{lm}^*(r,r_s;s) \tag{29}$$

with $\lambda=\pm is/c$ and

$$f_{lm}^*(r,r_s;s)=-4\pi\int_0^\infty q\,\delta(r-r')\,Y_{lm}^*(\theta',\varphi')(u^0c)^{-1}\times\exp(-st)\,dt. \tag{30}$$

The left-hand side of Eq. (29) is similar to that of Eq. (9) with λ appearing instead of k. Consequently, the appropriate interior and exterior solutions are $R_l^{(i)}(r,r_s;\lambda)$ and $R_l^{(e)}(r,r_s;\lambda)$ derived from the Fourier case by simply replacing k by λ. Similarly, we have the Wronkian of the Laplace transforms from Eq. (14) and Green's function from Eq. (17):

$$G_l(r,\rho,r_s;\lambda)=-i\lambda R_l^{(i)}(r_<,r_s;\lambda)R_l^{(e)}(r_>,r_s;\lambda). \tag{31}$$

Hence the unique solution of the nonhomogeneous Eq. (29) is

$$R_{lm}(r,r_s;\lambda)=-i\lambda\int_r^\infty R_l^{(i)}(r_<,r_s;\lambda)R_l^{(e)}(r_>,r_s;\lambda)f_{lm}^*(\rho,r_s;s)d\rho. \tag{32}$$

Using the Laplace transformation we will prove that $sR_l^{(i)}R_l^{(e)}$ is an analytic function of s for $\operatorname{Re}s>0$ and has a limit as $s\to 0$. These properties are obviously not directly connected with the Laplace transformation. They are properties of the solutions of Eq. (29) and presumably can be derived directly from that equation or from the expressions (12) and (13) of the solutions $R^{(i)}$ and $R^{(e)}$. However, attempts to that direction have failed for the moment.

Our basic assumption in the indirect derivation of these properties can be stated as follows: *The field of a constant point charge created at $t=0$ and staying thereafter at a fixed point goes to a static field as* $t\to+\infty$. This assumption is quite reasonable on physical grounds and we will not elaborate on it. We will comment in Sec. 4 on the possibility of eliminating it.

We prove now the following two properties:

Property 1: *The quantity $R_l^{(i)}(r_<,r_s;\lambda)R_l^{(e)}(r_>,r_s;\lambda)$ is an analytic function of the complex variable s for* $\operatorname{Re}s>0$.

Proof: For a point charge at r_0, θ_0, φ_0 we find from Eq. (30)

$$f_{lm}^*(r,r_s;s)=-4\pi q_0\left(1-\frac{r_s}{r_0}\right)^{1/2}Y_{lm}^*(\theta_0,\varphi_0)\,\delta(r-r_0)\,\frac{1}{s} \tag{33}$$

and from Eq. (32) (with $\lambda=is/c$)

$$R_{lm}(r,r_s;\lambda)=-\frac{4\pi q_0}{c}\left(1-\frac{r_s}{r_0}\right)^{1/2}Y_{lm}^*(\theta_0,\varphi_0)R_l^{(i)}(r_<,r_s;\lambda)\times R_l^{(e)}(r_>,r_s;\lambda). \tag{34}$$

Because of the assumption $P_{lm}(t,r)$ has a finite limit as $t\to+\infty$. Hence, its Laplace transform $R_{lm}(r,r_s;\lambda)$ converges absolutely[32] and is an analytic function of s for $\operatorname{Re}s>0$. From Eq. (34) we conclude the property for the quantity $R_l^{(i)}R_l^{(e)}$.

Property 2: *As $s\to 0$ with $|\arg s|\leq\alpha$, $0<\alpha<\pi/2$ we have*

$$\lim[sR_l^{(i)}(r_<,r_s;\lambda)R_l^{(e)}(r_>,r_s;\lambda)]=\frac{2c}{r_s}P_l\left(1-\frac{2r_<}{r_s}\right)Q_l\left(1-\frac{2r_>}{r_s}\right). \tag{35}$$

Proof: The resulting static field as $t\to+\infty$ has been found in paper I. Consequently,

$$\lim_{t\to+\infty}P_{lm}(t,r)=-\frac{8\pi q_0}{r_s}\left(1-\frac{r_s}{r_0}\right)^{1/2}Y_{lm}^*(\theta_0,\varphi_0)P_l\left(1-\frac{2r_<}{r_s}\right)$$

$$\times Q_l\left(1 - \frac{2r_<}{r_s}\right). \tag{36}$$

But[32]

$$\lim_{t\to+\infty} P_{lm}(t, r) = \lim_{s\to 0}\,[s\,R_{lm}(r, r_s;\lambda)]. \tag{37}$$

Combining this equation with Eq. (36) we have property 2.

Equation (35) is related to the limits 3 and 7 in the diagram of Sec. 2B. In fact it establishes the existence of two functions $f_l^{(i)}(r_s, k)$ and $f_l^{(e)}(r_s, k)$ satisfying those limits $[R_l^{(is)}$ and $R_l^{(es)}$ are proportional to $P_l(1-2r_</r_s)$ and $Q_l(1-2r_>/r_s)]$. However, we must note that in this approach the restriction on s does not allow k to go to zero through real values. Generalizing our results we conjecture that $s=0$ is a first order pole of $R_l^{(i)}R_l^{(e)}$. This is easily verified in flat space—time but remains to be proved rigorously for $r_s \neq 0$.

B. Particle falling into the black hole

In the next section our objective will be to prove the "baldness theorem" with as few as possible restrictions on the timelike curve followed by the particle in its fall. To show that some integrals converge we present here a few properties which are common to all timelike curves representing the fall of a particle into a Schwarzschild black hole. We assume only that the curve is timelike and that $dr/d\tau$, $d\theta/d\tau$, $d\varphi/d\tau$ are continuous functions of the proper time τ (even at the horizon) and that $dr/d\tau < -\epsilon$ for $r_s \leq r < r_s + \delta$ (ϵ and δ are small positive quantities). This last assumption excludes a "pathological" timelike curve along which the particle stops or reverses itself arbitrarily close to the horizon. From these assumptions we conclude that $d\theta/dr$ and $d\varphi/dr$ exist and are finite at $r = r_s$.

The properties are as follows:

Property 1: *For any particle falling into a Schwarzschild black hole we have as* $r \to r_s+$

$$\lim\left[\left(1-\frac{r_s}{r}\right)\frac{dt}{d\tau}\right] = c\,\lim\left[\left(1-\frac{r_s}{r}\right)u^0\right] = -\frac{1}{c}\frac{dr}{d\tau}\bigg|_{r=r_s}, \tag{38}$$

$$\lim\left[\left(1-\frac{r_s}{r}\right)^{-1}\frac{dx^\alpha}{dt}\right] = -c\,\frac{dx^\alpha}{dr}\bigg|_{r=r_s} \tag{39}$$

where x^α *is equal to* r, θ, φ *for* $\alpha = 1, 2, 3$.

Proof: From the Schwarzschild line element we have

$$\left(1-\frac{r_s}{r}\right)c\,\frac{dt}{d\tau} = \left\{\left(\frac{dr}{dt}\right)^2 + r^2\left[\left(\frac{d\theta}{d\tau}\right)^2 + \sin^2\theta\left(\frac{d\varphi}{d\tau}\right)^2\right]\right.$$
$$\left.\times\left(1-\frac{r_s}{r}\right) - c^2\left(1-\frac{r_s}{r}\right)\right\}^{1/2} \tag{40}$$

and in the limit $r \to r_s+$ we have Eq. (38). Using Eq. (38) we have as $r \to r_s+$

$$\lim\left[\left(1-\frac{r_s}{r}\right)^{-1}\frac{dx^\alpha}{dt}\right] = \lim\left[\left(1-\frac{r_s}{r}\right)^{-1}\frac{d\tau}{dt}\,\frac{dx^\alpha}{d\tau}\right]$$

$$= -c\,\frac{dx^\alpha}{dr}\bigg|_{r=r_s}. \tag{41}$$

Property 2: *As* $r \to r_s+$,

$$\lim\left[c\,\frac{dt}{dr} + \left(1-\frac{r_s}{r}\right)^{-1}\right] = C \tag{42}$$

where C is a negative constant.

Proof: From the Schwarzschild line element we have

$$\left(1-\frac{r_s}{r}\right)c^2\left(\frac{dt}{dr}\right)^2 - \left(1-\frac{r_s}{r}\right)^{-1} = c^2\left(\frac{d\tau}{dr}\right)^2 + r^2\left[\frac{d\theta}{dr}^{\,2}\right.$$
$$\left. + \sin^2\theta\left(\frac{d\varphi}{dr}\right)^2\right] \tag{43}$$

or

$$c\,\frac{dt}{dr} + \left(1-\frac{r_s}{r}\right)^{-1} = \left[c\,\frac{dt}{dr}\left(1-\frac{r_s}{r}\right) - 1\right]^{-1}$$
$$\times\left[c^2\left(\frac{d\tau}{dr}\right)^2 + r^2\left(\frac{d\theta}{dr}\right)^2 + r^2\sin^2\theta\left(\frac{d\varphi}{dr}\right)^2\right]. \tag{44}$$

Taking the limit $r \to r_s+$ and using Property 1 [Eq. (39) for $\alpha = 1$], we have

$$\lim\left[c\,\frac{dt}{dr} + \left(1-\frac{r_s}{r}\right)^{-1}\right] = -\frac{1}{2}\left[c^2\left(\frac{d\tau}{dr}\right)^2 + r^2\left(\frac{d\theta}{dr}\right)^2\right.$$
$$\left. + r^2\sin^2\theta\left(\frac{d\varphi}{dr}\right)^2\right]_{r=r_s}. \tag{45}$$

Property 3: *For a particle falling into the black hole the quantity*

$$X(r) = ct + r_s\ln(r - r_s) \tag{46}$$

is continuous and finite for $r \geq r_s$.

Proof: Obviously X is continuous and finite at any point $r > r_s$. Furthermore, we can write for any $r_s < r \leq r_1$

$$X(r) = \int_{r_1}^{r}\left[c\,\frac{dt}{dr} + \left(1-\frac{r_s}{r}\right)^{-1}\right]dr + r_1 - r + ct_1 + r_s\ln(r_1 - r_s). \tag{47}$$

This relation holds because it is true for $r = r_1$ and its derivative with respect to r coincides with that obtained from Eq. (46). In Eq. (47) the integrand goes to a finite limit as $r \to r_s+$. Hence, X remains finite as $r \to r_s+$ and the property follows.

C. The "baldness" theorem

We are ready now to prove that as a scalar particle falls into a Schwarzschild black hole, the scalar field outside the horizon goes to zero. With the properties proven in Sec. 3B we are able to prove the theorem for any timelike curve, whether or not this is a radial curve or a spiral, a geodesic or the trajectory of a propelled source. This generalization is important since the particle can be subjected to nongravitational forces during its fall. Again a basic assumption is needed here as in Sec. 3A. We assume that as the particle falls into the black hole the field outside the horizon goes to a final value. The necessity of such an assumption is due to the fact that in the corresponding mathematical theorem[32,33] properties of the Laplace transform only are not sufficient to ensure the existence of the limit of the primitive function as $t \to +\infty$.

One more assumption must be made concerning the history of the field. We will assume that the field was static (not necessarily zero) before some time in the past. Without loss of generality we can assume that the field starts changing at $t=0$. It seems that this assumption can be weakened but not completely ignored. For example, physical intuition suggests that the "baldness" theorem still holds for a particle falling from infinity freely from $t=-\infty$ but does not apply to the field of a particle circling the black hole from $t=-\infty$ and falling freely after $t=0$. Perhaps the crucial point is that the radiated energy from $-\infty$ to $+\infty$ must be finite, but we will not consider this question here.

We shall presently prove the following theorem:

Theorem: *Let* $\Psi(t, r, \theta, \varphi)$ *be the scalar field due to a point scalar particle of finite strength* $q(t)$ *in the exterior of a Schwarzschild black hole and let* Ψ *be static* (independent of t) *for* $t<0$. *If the particle falls into the black hole, then*

$$\lim_{t\to+\infty} \Psi = 0 \tag{48}$$

provided the limit exists.

Proof: The existence of $\lim \Psi$ as $t \to +\infty$ guaranties the existence of $\lim P_{lm}(t, r)$ as $t\to+\infty$. If $R_{lm}(r, r_s;\lambda)$ is the Laplace transform of $P_{lm}(t, r)$, then

$$\lim_{t\to+\infty} P_{lm}(t, r) = \lim_{s\to 0}[sR_{lm}(r, r_s;\lambda)] \tag{49}$$

with $|\arg s| \le \alpha$, $0<\alpha<\pi/2$. Consequently, it is enough to show that sR_{lm} goes to zero when $s\to 0$ staying in the above sector.

Let $q(0)$ at r_0, θ_0, φ_0 be the charge responsible for the original static field. Then for $t>0$ the total field can be considered as superposition of the following three fields: (i) the original static field due to $q(0)$; (ii) a time-dependent field due to a charge $-q(0)U(t)$ stationary at the fixed point r_0, φ_0, θ_0; (iii) a time dependent field due to $q(t)\ U(t)$ which moves along a timelike curve and eventually falls into the black hole. The fields (i) and (ii) will cancel each other in the limit $t\to+\infty$ (see Property 2 of Sec. 3A) and we have to prove that $sR_{lm}\to 0$ for the Laplace transform of the field (iii) only.

If $r_{\max}$ is the maximum of $r'(t)$, then for $r>r_{\max}$ $f^*_{lm}=0$. For $r\le r_{\max}$ $(dr'/dt\ne 0)$

$$f^*_{lm}(r, r_s;s) = \sum_{t_r} (-4\pi) \left[\frac{qY^*_{lm}(\theta', \varphi')\exp(-st)}{u^0c\,|dr'/dt|}\right]_{t=t_r}, \tag{50}$$

where the summation is taken over all possible roots t_r of $r-r'(t)=0$ for a given r. Each term of the sum will contribute to R_{lm} [see Eq. (32)] a term

$$-4\pi\lambda \int_{r_s}^{r_{\max}} \left[R_l^{(i)}(r_<, r_s;\lambda)R_l^{(e)}(r_>, r_s;\lambda)\,\frac{qY^*_{lm}(\theta', \varphi')e^{-st}}{u^0c\,|dr'/dt|}\right]_{t=t_\rho} \times d\rho, \tag{51}$$

where $r_<\ (r_>)$ is the smaller (larger) of r and ρ. Since

$$u^0c\left|\frac{dr'}{dt}\right| = \frac{dr'}{d\tau} < -\epsilon \tag{52}$$

the integrand in Eq. (51) is finite at any $r>r_s$. Moreover, only one of $R_l^{(i)}$ and $R_l^{(e)}$ depends on ρ and, consequently, the product $R_l^{(i)}R_l^{(e)}$ contributes a factor $\exp(-sr_sc^{-1}\ln|\rho-r_s|)$ to the integrand. This factor combined with $\exp(-st_\rho)$ gives $\exp(-sX/c)$ which remains finite as $\rho\to r_s+$. Since the integrand remains finite in the *closed* interval $r_s\le\rho\le r_{\max}$ it goes to zero *uniformly* after multiplication by s as $s\to 0$. Hence sR_{lm} tends to zero as $s\to 0$ remaining in the stated sector. This completes the theorem.

To make the proof simpler we have deliberately neglected contributions to the field from points where $dr'/d\tau=0$. It is clear that only formal complications arise from such points without altering the result. For example, if $r'=\text{const}$ for some time interval, then $\delta(r-r')$ can be taken out of the integral in Eq. (30) and integrated easily with respect to r to give a finite term for R_{lm}. Multiplication by s will make the product sR_{lm} to go to zero as $s\to 0$.

4. REMARKS

The Fourier and Laplace transforms presented in Secs. 2A, 2C, and 3A contain the basic formulas for attacking wave problems in the exterior of a Schwarzschild black hole. It is clear that the whole procedure can be used to study weak electromagnetic[34] and gravitational fields. The vector or tensor character of these fields introduces more than one dependent variables representing the field and can lead to new or different physical phenomena but the basic method of study will remain the same.

However, all these studies are limited by our incomplete knowledge of the radial wave functions R_l, namely the solutions of Eq. (9) and its counterparts for electromagnetic and gravitational fields. We expect that a study of the analytical properties of R_l will answer many questions raised in the course of the present work, such as "Does $K_{45}/_{46}$ go to -1 as $r_s\to 0$?," "What are the simplest $f_l^{(i)}$ and $f_l^{(e)}$?," "Is $k=0$ a first order pole of $R_l^{(i)}(r_<, r_s;k)R_l^{(e)}(r_>, r_s;k)$?," "How does $R_l^{(i)}$ behave at $r=0$?," etc.

In addition to these immediate results it is possible that using properties of R_l and asymptotic expansions for the Fourier transform[29] we will be able to prove the "baldness" theorem without assuming the existence of $\lim\Psi$ as $t\to+\infty$. Also we will be able to see whether Price's results for the behavior of the field near $t=+\infty$ apply to the falling particle or are limited to the case of a source anchored on the collapsing star. This can be accomplished by examining the limit of sd^nR_l/ds^n as $s\to 0$.

That this paper represents only a small step towards understanding wave phenomena in a black-hole space becomes obvious from the fact that it raises more problems than it solves. As related subjects for study we mention the tails of the waves, the wavefront, the "collision of wavefronts" after the wave has engulfed the black hole, the initial conditions on the horizon which determine the evolution of the field inside the black hole, etc. To these problems analytical methods and numerical techniques are viewed as complimentary to each other.

With respect to the "baldness" theorem it would be interesting to examine whether or not the conditions can be still weakened. Most probably the restriction to have a static field for $t<0$ can be replaced by a weaker one. From our point of view the most undesirable condition is the assumed existence of a limit for the field as $t\to+\infty$. As we stated before there is a hope to eliminate the need of this condition using the asymptotic—expansion theory of Fourier transforms. Finally, another possibility of generalization of the "baldness" theorem must be mentioned here, namely, to more complicated black holes. However, in such cases the radial wave equation (when separation of variables is possible) is much more involved.

*Supported in part by the Greek Ministry of Culture and Science (Office of Scientific Research and Development).

[1]K.S. Thorne, *Nonspherical Gravitational Collapse* (Orange Aid Preprint Series, California Institute of Technology, 1971).
[2]V. de la Cruz, J.E. Chase, and W. Israel, Phys. Rev. Lett. **24**, 423 (1970).
[3]A.Z. Patashinskiĭ and A.A. Khar'kov, Zh. Eksp. Teor. Fiz. 59, 574 (1970) [Sov. Phys. JETP **32**, 313 (1971).
[4]R.H. Price, Phys. Rev. D 5, 2419 (1972).
[5]R.H. Price, Phys. Rev. D 5, 2439 (1972).
[6]J. Bicak, General Relativity and Gravitation **3**, 331 (1972).
[7]F.J. Zerilli, Phys. Rev. D **2**, 2141 (1970).
[8]M. Davis and R. Ruffini, Nuovo Cimento Lett. **2**, 1165 (1971).
[9]M. Davis, R. Ruffini, W.H. Press and R.H. Price, Phys. Rev. Lett. **27**, 1466 (1971).
[10]R.A. Matzner, J. Math. Phys. **9**, 163 (1968).
[11]C.W. Misner, R.A. Breuer, D.R. Brill, P.L. Chrzanowski, H.G. Hughes III, and C.M. Pereira, Phys. Rev. Lett. **28**, 998 (1972).
[12]D.M. Chitre and R.H. Price, Phys. Rev. Lett. **29**, 185 (1972).
[13]M. Davis, R. Ruffini, J. Tiomno, and F. Zerilli, Phys. Rev. Lett. **28**, 1352 (1972).
[14]W. Kundt and E.T. Newman, J. Math. Phys. **9**, 2193 (1968).
[15]E.D. Fackerell, Astrophys. J. **166**, 197 (1971).
[16]T.C. Mo and C.H. Papas, Phys. Rev. D 3, 1708 (1971).
[17]S. Persides, J. Math. Phys. **12**, 2355 (1971).
[18]I. Piir, Toim. Eesti NSV Tead, Acad. (USSR) **20**, 253 (1971).
[19]V. Unt and P. Keres, Toim. Eesti NSV Tead, Acad. (USSR) **21**, 17 (1972).
[20]J.M. Bardeen and W.H. Press, J. Math. Phys. **14**, 7 (1973).
[21]W. Israel, Commun. Math. Phys. **8**, 245 (1968).
[22]J.L. Anderson and J.M. Cohen, Astrophys. Space Sci. **9**, 146 (1970).
[23]J.M. Cohen and R.M. Wald, J. Math. Phys. **12**, 1845 (1971).
[24]R.M. Wald, Phys. Rev. D **6**, 1476 (1972).
[25]S. Persides, J. Math. Anal. Appl. **43**, 571 (1973).
[26]S. Persides, J. Math. Phys. **14**, 1017 (1973).
[27]A. Papoulis, *The Fourier Integral and its Applications* (McGraw-Hill, New York, 1962).
[28]J.D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1962).
[29]M.J. Lighthill, *Introduction to Fourier Analysis and Generalized Functions* (Cambridge U.P., Cambridge, 1970).
[30]A.K. Korn and T.M. Korn, *Mathematical Handbook for Scientists and Engineers* (McGraw-Hill, New York, 1968).
[31]B.A. Fuchs and V.I. Levin, *Functions of a Complex Variable* (Pergamon, Oxford, 1961).
[32]W. Kaplan, *Operational Methods for Linear Systems* (Addison-Wesley, Reading, 1962).
[33]G. Doetch, *Theorie und Anwendung der Laplace Transformation* (Springer, Berlin, 1937).
[34]S. Persides and B. Xanthopoulos, "The Electromagnetic Field of a Point Source in Schwarzschild's Space—Time" (to be published).

Singular perturbation method in neutron transport theory

Janusz Mika

Institute of Nuclear Research, Swierk, Otwock, Poland

(Received 14 February 1972; revised manuscript received 16 August 1972)

We consider equations of evolution with a small parameter in a Banach space. The method of singular perturbations is applied to derive inner and outer asymptotic solutions. It is shown that the neutron transport equation coupled with the equation for the concentration of delayed neutron precursors, if considered in the Hilbert space of square integrable functions, satisfies all the requirements set up in the paper.

INTRODUCTION

In a series of papers Hendry and Bell[1] and Hendry[2,3] have introduced into neutron transport theory the method of singular perturbations originally developed in fluid mechanics[4–6] as the method of matched asymptotic expansions. The numerical tests show that the method is exceptionally advantageous in evaluating the response of a subcritical reactor system to an instantaneous pulse of neutrons.

The method of singular perturbations is based upon the observation that the derivative with respect to time is multiplied by a coefficient smaller than coefficients standing by remaining terms in the neutron transport equation by several orders of magnitude. Thus the properly defined small dimensionless parameter $\epsilon > 0$ is introduced into the equation and the solution expanded into a power series in ϵ.

The analysis given in Refs. 1–3 has only a formal character and may not serve as a rigorous justification of the results. In this paper the method of singular perturbations is applied to the equations of evolution with a small parameter in a Banach space to show that in the zeroth-order perturbation the asymptotic solution tends to the exact one in a properly defined sense. The results are also generalized to the systems of equations such that only one of the equations contains a small parameter.

The question of higher-order perturbations is not dealt with in the paper. The analysis is much more involved, especially for the inner asymptotic solution, and at the moment it is not clear to the author if it can be performed within the formalism used in the paper.

In the last section it is shown that the neutron transport equation coupled with the equation for the concentration of delayed neutron precursors actually meets all the requirements set up in previous chapters.

The results known from the literature are quoted without proofs. For the details see Ref. 7–10.

EQUATIONS AND SYSTEMS OF EQUATIONS OF EVOLUTION IN A BANACH SPACE

In this section a complex Banach space E with the norm $\|\cdot\|$ will be considered.

Definition 1: A family $\{G(t); 0 \le t\}$ of bounded operators in E will be called a strongly continuous semigroup if

(i) $G(t+s) = G(t)G(s), \quad 0 \le t, \; 0 \le s,$

(ii) $G(0) = I,$

(iii) for each $x \in E$ the function $t \to G(t)x$ is strongly continuous on $[0, \infty)$.

Definition 2: The operator

$$Ax = \lim_{t \to 0} t^{-1}[G(t)x - x]$$

with the domain $D(A)$ consisting of all $x \in E$ such that the limit exists in the norm in E, is called a generator of a strongly continuous semigroup $G(t)$. It has the properties:

(i) It is a closed linear operator and its domain $D(A)$ is dense in E;

(ii) if $x \in D(A)$, then for every $t \in [0, \infty)$, $G(t)x \in D(A)$ and the function $t \to G(t)x$ is strongly continuously differentiable on $[0, \infty)$ such that

$$\frac{dG(t)x}{dt} = AG(t)x = G(t)Ax.$$

Theorem 1 (*Hille–Yosida*): A necessary and sufficient condition that a closed linear operator A with the domain $D(A)$ dense in E generates a strongly continuous semigroup $\{G(t); 0 \le t\}$ of bounded operators such that for each $t \in [0, \infty)$

$$\|G(t)\| \le \exp(\omega t)$$

for some real ω, is that the resolvent of A

$$R(\lambda, A) = (\lambda I - A)^{-1}$$

for every $\lambda > \omega$ exists and is an operator defined on the whole E, such that

$$\|R(\lambda, A)\| \le (\lambda - \omega)^{-1}.$$

Lemma 1: If the operator A satisfies the requirements of Theorem 1 and the function $t \to q(t)$ with the values from E is strongly continuously differentiable on $[0, T]$, where T is a fixed positive number, then the equation of evolution

$$\frac{dx(t)}{dt} = Ax(t) + q(t), \quad x(0) = x_0 \in D(A)$$

has for $t \in [0, T]$ a unique strongly continuously differentiable solution

$$x(t) = G(t)x_0 + \int_0^t ds\, G(t-s)q(s).$$

The integral is understood as a strong limit of Riemann sums. The same meaning will be attached to other integrals of abstract functions appearing in this paper.

Definition 3: A family $\{U(t,s); 0 \le s \le t \le T\}$ of bounded operators in E will be called a strongly continuous quasi-semigroup[11] if

(i) $\quad U(t,s) = U(t,\tau)U(\tau,s), \quad 0 \le s \le \tau \le t \le T,$

(ii) $U(t, t)=I, \quad 0 \leqslant t \leqslant T$,

(iii) for each $x \in E$ the function $t, s \to U(t, s)x$ is strongly continuous on the triangle $0 \leqslant s \leqslant t \leqslant T$.

Theorem 2: Let a family $\{A(t); 0 \leqslant t \leqslant T\}$ of closed operators in E is such that

(i) for each $t \in [0, T]$ the domain $D(A(t)) = D(A)$ is independent of t and $D(A)$ is dense in E,

(ii) for each $\tau \in [0, T]$ the operator $A(\tau)$ is a generator of a strongly continuous semigroup $\{G_\tau(t); 0 \leqslant t\}$ such that for each $t \in [0, \infty)$

$$\|G_\tau(t)\| \leqslant \exp(\alpha_\tau t)$$

and

$$\alpha = \sup_{\tau \in [0, T]} \alpha_\tau < 0,$$

(iii) for each $x \in D(A)$ the function $t \to A(t)x$ is strongly continuously differentiable on $[0, T]$.

Then the family $\{A(t); 0 \leqslant t \leqslant T\}$ generates a strongly continuous quasisemigroup $\{U(t, s); 0 \leqslant s \leqslant t \leqslant T\}$ such that for each $x \in D(A)$ and $0 \leqslant s \leqslant t \leqslant T$ the elements $U(t, s)x \in D(A)$ and the function $t, s \to U(t, s)x$ is strongly continuously differentiable on the triangle $0 \leqslant s \leqslant t \leqslant T$. The partial derivatives of $U(t, s)$ satisfy the identities

$$\frac{\partial U(t, s)}{\partial t} x = A(t)\, U(t, s)x$$

and

$$\frac{\partial U(t, s)}{\partial s} x = -\, U(t, s)A(s)x.$$

Lemma 2 (*Kato*): The quasi-semigroup $U(t, s)$ generated by the family $A(t)$ can be expressed with the semigroups $G_\tau(t)$ by the multiplicative integral

$$U(t, s) = \lim \prod_{i=1}^{n} G_{\tau_i}(t_i - t_{i-1}),$$

where $s = t_0 < t_1 < \cdots < t_n = t$ are the points of divisions of the interval $[s, t]$ and $t_{i-1} < \tau_i < t_i$. The limit is understood in the strong sense for $\max_i(t_i - t_{i-1}) \to 0$.

Lemma 3: For $0 \leqslant s \leqslant t \leqslant T$

$$\|U(t, s)\| \leqslant \exp[\alpha(t - s)],$$

where α is defined in assumption (ii) of Theorem 2.

Lemma 4: If the family $\{A(t); 0 \leqslant t \leqslant T\}$ satisfies the requirements of Theorem 2 and generates the strongly continuous quasi-semigroup $\{U(t, s); 0 \leqslant s \leqslant t \leqslant T\}$ and the function $t \to q(t)$ is strongly continuously differentiable on $[0, T]$ then the equation of evolution

$$\frac{dx(t)}{dt} = A(t)x(t) + q(t), \quad x(0) = x_0 \in D(A)$$

has for $t \in [0, T]$ a unique strongly continuously differentiable solution

$$x(t) = U(t, 0)x_0 + \int_0^t ds\, U(t, s)q(s).$$

Lemma 5: If the family $\{A(t); 0 \leqslant t \leqslant T\}$ generates the quasisemigroup $\{U(t, s); 0 \leqslant s \leqslant t \leqslant T\}$, then for $\epsilon > 0$ the family $\{\epsilon^{-1}A(t); 0 \leqslant t \leqslant T\}$ generates the quasisemigroup $\{U_\epsilon(t, s); 0 \leqslant s \leqslant t \leqslant T\}$ such that for $0 \leqslant s \leqslant t \leqslant T$

$$\|U_\epsilon(t, s)\| \leqslant \exp[(\alpha/\epsilon)(t - s)],$$

where α is defined in Lemma 3.

Lemma 6: Let the family $\{A(t); 0 \leqslant t \leqslant T\}$ satisfies the requirements of Theorem 2 and generates a strongly continuous quasisemigroup $\{U(t, s); 0 \leqslant s \leqslant t \leqslant T\}$. Let the family $\{B(t); 0 \leqslant t \leqslant T\}$ satisfies also the requirements of Theorem 2 and generates a strongly continuous quasisemigroup $\{V(t, s); 0 \leqslant s \leqslant t \leqslant T\}$. Let $\{P(t); 0 \leqslant t \leqslant T\}$ and $\{Q(t); 0 \leqslant t \leqslant T\}$ be the families of bounded operators such that the functions $t \to P(t)$ and $t \to Q(t)$ are uniformly continuously differentiable on $[0, T]$. Let finally the functions $t \to q(t)$ and $t \to r(t)$ be strongly continuously differentiable on $[0, T]$. Then the system of equations of evolution

$$\frac{dx(t)}{dt} = A(t)x(t) + P(t)y(t) + q(t), \quad x(0) = x_0 \in D(A),$$

$$\frac{dy(t)}{dt} = Q(t)x(t) + B(t)y(t) + r(t), \quad y(0) = y_0 \in D(B)$$

has for $t \in [0, T]$ a unique strongly differentiable solution

Proof of the lemma follows by the application of the perturbation theorem (see VIII. 1.22 of Ref. 7).

Lemma 7 (*Kisyński*): Let the function $t \to m(t)$ be strongly continuously differentiable on $[0, T]$, the family of operators $\{W(t, s)\}$ uniformly bounded, and the function $t, s \to W(t, s)g$ strongly continuous for each $g \in E$ on the triangle $0 \leqslant s \leqslant t \leqslant T$. Then the Volterra-type equation

$$z(t) = m(t) + \int_0^t ds\, W(t, s)z(s)$$

has a unique solution for $t \in [0, T]$. This solution may be obtained by the method of successive approximations in the form

$$z(t) = \sum_{k=0}^{\infty} m^{(k)}(t),$$

$$m^{(0)}(t) = m(t), \quad m^{(k)}(t) = \int_0^t ds\, W(t, s)\, m^{(k-1)}(s), \quad 1 \leqslant k.$$

The series is strongly convergent uniformly on $[0, T]$.

Lemma 8: Let

$$m(t) = U(t, 0)x_0 + \int_0^t ds\, U(t, s)\, q(s) + Z(t, 0)y_0 + \int_0^t ds\, Z(t, s)\, r(s),$$

where $U(t, s)$, $q(s)$, and $r(s)$ are defined in Lemma 6 and for each $g \in E$ and $0 \leqslant s \leqslant t \leqslant T$

$$Z(t, s)g = \int_0^t ds'\, U(t, s')\, P(s')\, V(s', s)g.$$

Let additionally

$$W(t, s) = Z(t, s)\, Q(s).$$

Then the solution $z(t)$ to the Volterra equation from Lemma 7 is identical with $x(t)$ from Lemma 6. Moreover, $y(t)$ can be expressed by $x(t)$ as follows:

$$y(t) = V(t, 0)y_0 + \int_0^t ds\, V(t, s)\, r(s) + \int_0^t ds\, V(t, s)\, Q(s)x(s).$$

Proof: From the assumption of Lemma 6 it follows that the above defined $m(t)$ and $W(t, s)$ satisfy the requirements of Lemma 7 and the corresponding Volterra equation has the unique solution $z(t)$. To show that $z(t)$

is identical with $x(t)$ first apply Lemma 4 to the second equation from Lemma 6 to obtain the formula giving $y(t)$ in terms of $x(t)$. Substituting the resulting equation into the first equation and again applying Lemma 4 with the integral involving $x(t)$ treated as a source term, one obtains the Volterra equation with $m(t)$ and $W(t,s)$ defined as above.

EQUATIONS OF EVOLUTION WITH A SMALL PARAMETER

Definition 4 (*Krein*): The family of functions $\{\varphi_\epsilon(t); 0<\epsilon\le\epsilon_0, 0\le t\le T\}$ tends almost uniformly on the interval $(0,T]$ to the function $\varphi(t)$ if for each $\delta>0$ there exist $t_1(\delta)>0$ and $\epsilon_1(\delta)>0$ such that

$$\|\varphi_\epsilon(t)-\varphi(t)\|<\delta,$$

for $\epsilon<\epsilon_1(\delta)$ and $\epsilon t_1(\delta)\le t\le T$.

A simple example of the family tending almost uniformly to zero on $(0,T]$ is $\varphi_\epsilon(t)=e^{-at/\epsilon}$ for $a>0$. The almost uniform convergence on the interval $(0,T]$ implies the uniform convergence on any interval $[t_0,T]$ for $t_0>0$. The inverse is not always true as it is seen for the family $\varphi_\epsilon(t)=\epsilon^{-1}e^{-at/\epsilon}$.

Lemma 9 (*Krein*): Let the family $A(t)$ satisfies the requirements of Theorem 2 and the function $t\to q(t)$ is strongly continuously differentiable on $[0,T]$ and let additionally the family of operators $dA(t)/dt$ be defined on the whole E and bounded uniformly on $[0,T]$. Then the solution to the equation of evolution

$$\epsilon\frac{dx_\epsilon(t)}{dt}=A(t)\,x_\epsilon(t)+q(t),\qquad x_\epsilon(0)=x_0\in D(A)$$

tends almost uniformly on $(0,T]$ to the function

$$\bar{x}(t)=-A^{-1}(t)q(t),$$

which is referred to as the outer asymptotic solution.

Proof: Define the function

$$\bar{z}_\epsilon(t)=x_\epsilon(t)-\bar{x}(t).$$

It satisfies the equation of evolution

$$\epsilon\frac{d\bar{z}_\epsilon(t)}{dt}=A(t)\,\bar{z}_\epsilon(t)-\epsilon\frac{d\bar{x}(t)}{dt}$$

with the initial condition

$$\bar{z}_\epsilon(0)=x_0-\bar{x}(0)\in D(A).$$

The unique solution to the above equation is, according to Lemmas 4 and 5, given by

$$\bar{z}_\epsilon(t)=U_\epsilon(t,0)\,\bar{z}_\epsilon(0)-\int_0^t ds\,U_\epsilon(t,s)\,\frac{d\bar{x}(s)}{ds}.$$

Since the family $dA(t)/dt$ is uniformly bounded and the function $t\to q(t)$ strongly continuously differentiable the term

$$\frac{dx(t)}{dt}=A^{-1}(t)\,\frac{dA(t)}{dt}\,A^{-1}(t)\,q(t)-A^{-1}(t)\,\frac{dq(t)}{dt}$$

is uniformly bounded on $[0,T]$. Thus the integral term in the expression for $\bar{z}_\epsilon(t)$ can be estimated from Lemma 5 as follows:

$$\left\|\int_0^t ds\,U_\epsilon(t,s)\,\frac{d\bar{x}(s)}{ds}\right\|\le C_0\int_0^t ds\,\exp[-\alpha(t-s)/\epsilon\le\frac{C_0\epsilon}{|\alpha|},$$

where

$$C_0=\sup_{t\in[0,T]}\left\|\frac{dx(t)}{dt}\right\|.$$

Thus

$$\|\bar{z}_\epsilon(t)\|\le\exp(-\alpha t/\epsilon)\|\bar{z}_\epsilon(0)\|+C_0\epsilon/|\alpha|$$

from which it follows that $\bar{z}_\epsilon(t)$ tends to zero almost uniformly on $(0,T]$. If $\bar{z}_\epsilon(0)=x_0-\bar{x}(0)=0$, then $x_\epsilon(t)$ tends to $\bar{x}(t)$ uniformly on $[0,T]$.

Lemma 10 (*Krein*): Under the requirements of Lemma 9 the equation of evolution

$$\epsilon\frac{d\tilde{x}_\epsilon(t)}{dt}=A(0)\tilde{x}_\epsilon(t)+q(0),\qquad\tilde{x}_\epsilon(0)=x_0\in D(A)$$

has the unique strongly differentiable solution given by

$$\tilde{x}_\epsilon(t)=G_0(t/\epsilon)x_0-G_0(t/\epsilon)\,\bar{x}(0)+\bar{x}(0),$$

where $G_0(t)$ is the semigroup generated by the operator $A(0)$. The function $\tilde{x}_\epsilon(t)$ is referred to as the inner asymptotic solution.

The family $\tilde{x}_\epsilon(t)$ is uniformly bounded on $[0,T]$ and tends almost uniformly on $(0,T]$ to

$$\tilde{\bar{x}}(t)=\bar{x}(0)$$

which is called the intermediate asymptotic solution. For any $t_2>0$ and $\delta>0$ there exists $\epsilon_2(t_2,\delta)$ such that if $\epsilon<\epsilon_2(t_2,\delta)$ then for $0\le t\le\epsilon t_2$

$$\|x_\epsilon(t)-\tilde{x}_\epsilon(t)\|<\delta.$$

Proof: From the assumed properties of $A(t)$ it follows that the operator $A(0)$ generates the semigroup $G_0(t)$. Thus from Lemma 1 the solution $\tilde{x}_\epsilon(t)$ has the form

$$\tilde{x}_\epsilon(t)=G_0\Big(\frac{t}{\epsilon}\Big)\,x_0+\frac{1}{\epsilon}\int_0^t ds\,G_0\Big(\frac{t-s}{\epsilon}\Big)\,q(0).$$

The integral can be written as

$$\frac{1}{\epsilon}\int ds\,G_0\Big(\frac{t-s}{\epsilon}\Big)\,q(0)=\int_0^{t/\epsilon}ds\,G_0(s)A(0)A^{-1}(0)\,q(0)$$

$$=\int_0^{t/\epsilon}ds\,\frac{dG_0(s)}{ds}\,A^{-1}(0)\,q(0)=G_0\Big(\frac{t}{\epsilon}\Big)\,A^{-1}(0)q(0)-A^{-1}(0)q(0),$$

from which the final expression for $\tilde{x}_\epsilon(t)$ follows.

Since $G_0(t/\epsilon)$ tends in the norm on $(0,T]$ almost uniformly to zero it follows that $\tilde{x}_\epsilon(t)$ tends almost uniformly to $\tilde{\bar{x}}(t)$.

To prove the last statement of the lemma, consider the family

$$\tilde{Z}_\epsilon(t)=x_\epsilon(t)-\tilde{x}_\epsilon(t).$$

It satisfies the equation of evolution

895

$$\epsilon \frac{d\tilde{z}_\epsilon(t)}{dt} = A(t)\tilde{z}_\epsilon(t) + [A(t) - A(0)]\tilde{x}_\epsilon(t) + q(t) - q(0)$$

with the initial condition $\tilde{z}_\epsilon(0) = 0$. The solution of the above equation is, according to Lemmas 4 and 5,

$$\tilde{z}_\epsilon(t) = \frac{1}{\epsilon}\int_0^t ds\, U_\epsilon(t,s)[A(s) - A(0)]\tilde{x}_\epsilon(s) + \frac{1}{\epsilon}\int_0^t ds\, U_\epsilon(t,s)[q(s) - q(0)].$$

Since the function $t \to d\tilde{x}_\epsilon(t)/dt$ is strongly continuously differentiable it is also uniformly bounded on $[0, T]$. This shows, on account of the evolution equation satisfied by $\tilde{x}_\epsilon(t)$, that $A(0)x_\epsilon(t)$ is also uniformly bounded in ϵ and t. Now from the fact that the domain $D(A(t)) = D(A)$ is independent of t and the function $t \to A(t)x$ is strongly continuous for any $x \in D(A)$ if follows that the function $t \to A(t)A^{-1}(0)$ is uniformly continuous. Thus the quantity

$$\|[A(s) - A(0)]A^{-1}(0)A(0)\tilde{x}_\epsilon(s)\|$$

tends to zero with $s \to 0$ uniformly in ϵ.

Now putting $t = \epsilon t_2$, changing the variables $s = \epsilon s'$ in the integrals expressing $\tilde{z}_\epsilon(t)$, and using Lemma 5 one gets

$$\|\tilde{z}_\epsilon(\epsilon t_2)\| \leq \int_0^{t_2} ds'\, \|[A(\epsilon s') - A(0)]A^{-1}(0)A(0)\tilde{x}_\epsilon(\epsilon s')\| + \int_0^{t_2} ds'\, \|q(\epsilon s') - q(0)\|.$$

This shows that for any $t_2 > 0$ and $\delta > 0$ there exists $\epsilon_2(t_2, \delta)$ such that if $\epsilon < \epsilon_2(t_2, \delta)$, then for $0 \leq t \leq \epsilon t_2$

$$\|x_\epsilon(t) - \tilde{x}_\epsilon(t)\| < \delta.$$

Theorem 3 (*Krein*): Under the assumptions of Lemma 9 for sufficiently small ϵ the solution $x_\epsilon(t)$ tends with $\epsilon \to 0$ to the asymptotic solution

$$x_\epsilon^{(\mathrm{as})}(t) = \bar{x}(t) + \tilde{x}_\epsilon(t) - \tilde{\bar{x}}(t)$$

uniformly on $[0, T]$.

Proof: It follows from Lemmas 9 and 10 that $x_\epsilon(t)$ tends to $\bar{x}(t)$ and $\tilde{x}_\epsilon(t)$ to $\tilde{\bar{x}}(t)$ almost uniformly on $(0, T]$. Thus the family

$$v_\epsilon(t) = x_\epsilon(t) - x_\epsilon^{(\mathrm{as})}(t)$$

tends to zero almost uniformly on $(0, T]$. This means that for any $\delta > 0$ there exists $t_0(\delta) > 0$ and $\epsilon_1(\delta) > 0$ such that for $\epsilon < \epsilon_1(\delta)$ and $\epsilon t_0(\delta) \leq t \leq T$

$$\|v_\epsilon(t)\| < \delta.$$

Next from Lemma 10 it follows that for $\epsilon < \epsilon_2(t_0(\delta), \frac{1}{2}\delta)$ and $0 \leq t \leq \epsilon t_0(\delta)$

$$\|x_\epsilon(t) - \tilde{x}_\epsilon(t)\| < \tfrac{1}{2}\delta.$$

Finally, from the continuity of the function $t \to A^{-1}(t)q(t)$ it is seen that for $\epsilon < \epsilon_3(\delta)$ and $0 \leq t \leq \epsilon t_0(\delta)$

$$\|\bar{x}(t) - \tilde{\bar{x}}(t)\| < \tfrac{1}{2}\delta.$$

From the above inequalities it follows that for $\epsilon < \min(\epsilon_1, \epsilon_2, \epsilon_3)$ and $0 \leq t \leq T$

$$\|v_\epsilon(t)\| < \delta.$$

This shows that the family $v_\epsilon(t)$ tends to zero uniformly on $[0, T]$ which proves the theorem.

SYSTEMS OF EQUATIONS OF EVOLUTION WITH A SMALL PARAMETER

Lemma 11: Let the families $\{A(t); 0 \leq t \leq T\}$, $\{B(t); 0 \leq t \leq T\}$, $\{P(t); 0 \leq t \leq T\}$, $\{Q(t); 0 \leq t \leq T\}$ and the functions $t \to q(t)$ and $t \to r(t)$ satisfy the requirements of Lemma 6. Let additionally the family $dA(t)/dt$ be defined on the whole E and bounded uniformly on $[0, T]$. Then the system of equations

$$A(t)\bar{x}(t) + P(t)\bar{y}(t) + q(t) = 0,$$

$$\frac{d\bar{y}(t)}{dt} = Q(t)\bar{x}(t) + B(t)\bar{y}(t) + r(t),$$

with the initial condition $\bar{y}(0) = y_0 \in D(B)$, has a unique strongly differentiable solution $\{\bar{x}(t), \bar{y}(t)\}$ (the outer asymptotic solution).

Proof: Solving the first equation for $\bar{x}(t)$ and substituting to the second equation one gets the equation of evolution

$$\frac{d\bar{y}(t)}{dt} = [B(t) - Q(t)A^{-1}(t)P(t)]\bar{y}(t) + r(t) - Q(t)A^{-1}(t)q(t), \quad \bar{y}(0) = y_0 \in D(B).$$

This equation has a unique strongly differentiable solution obtained by the perturbation theorem since $B(t)$ generates a quasisemigroup $V(t,s)$ and $Q(t)A^{-1}(t)P(t)$ is a family of bounded operators. From the properties of $A(t)$, $P(t)$, and $q(t)$ it follows that the function $t \to \bar{x}(t)$ is also strongly differentiable.

Lemma 12: With the assumptions of Lemma 11 the solution $\{x_\epsilon(t), y_\epsilon(t)\}$ to the system of equations of evolution

$$\epsilon \frac{dx_\epsilon(t)}{dt} = A(t)x_\epsilon(t) + P(t)y_\epsilon(t) + q(t),$$

$$\frac{dy_\epsilon(t)}{dt} = Q(t)x_\epsilon(t) + B(t)y_\epsilon(t) + r(t),$$

with the initial condition

$$x_\epsilon(0) = x_0 \in D(A), \quad y_\epsilon(0) = y_0 \in D(B)$$

tends to the outer asymptotic solution $\{\bar{x}(t), \bar{y}(t)\}$ almost uniformly on $(0, T]$.

Proof: Let

$$\bar{v}_\epsilon(t) = x_\epsilon(t) - \bar{x}(t), \quad \bar{w}_\epsilon(t) = y_\epsilon(t) - \bar{y}(t).$$

The families $\bar{v}_\epsilon(t)$ and $\bar{w}_\epsilon(t)$ satisfy the following system of equations

$$\epsilon \frac{d\bar{v}_\epsilon(t)}{dt} = A(t)\bar{v}_\epsilon(t) + P(t)\bar{w}_\epsilon(t) - \epsilon \frac{d\bar{x}(t)}{dt},$$

$$\frac{d\bar{w}_\epsilon(t)}{dt} = Q(t)\bar{v}_\epsilon(t) + B(t)\bar{w}_\epsilon(t),$$

with the initial conditions

$$\bar{v}_\epsilon(0) = x_0 - A^{-1}(0)[P(0)y_0 + q(0)] \in D(A),$$

$$\bar{w}_\epsilon(0) = 0$$

From Lemmas 6, 7, and 8 it follows that the above system has a unique solution and is equivalent to a single Volterra type equation

$$\bar{v}_\epsilon(t) = \bar{m}_\epsilon(t) + \epsilon^{-1}\int_0^t ds\, W_\epsilon(t,s) v_\epsilon(s),$$

where

$$\bar{m}_\epsilon(t) = U_\epsilon(t,0)\{x_0 - A^{-1}(0)[P(0)y_0 + q(0)]\}, \\ - \int_0^t ds\, V_\epsilon(t,s)\frac{d\bar{x}(s)}{ds}$$

and

$$W_\epsilon(t,s)g = Z_\epsilon(t,s)Q(s)g,$$

and

$$Z_\epsilon(t,s)g = \int_s^t ds\, U_\epsilon(t,s')P(s')V(s',s)g,$$

for any $g \in E$.

The function $\bar{W}_\epsilon(t)$ is expressed in the following form:

$$\bar{w}_\epsilon(t) = \int_0^t ds'\, V(t,s')Q(s')\bar{v}_\epsilon(s').$$

From Lemma 9 it follows that the family $d\bar{x}(t)/dt$ is uniformly bounded on $[0,T]$ and Lemma 5 shows that the integral term in the expression for $\bar{m}_\epsilon(t)$ tends to zero uniformly on $[0,T]$. Thus the family $\bar{m}_\epsilon(t)$ is uniformly bounded on $[0,T]$ and tends to zero in the norm almost uniformly on $(0,T]$.

Since the families of operators $A(t)$ and $B(t)$ are assumed to satisfy the requirements of Theorem 2, both quasi-semigroups $U(t,s)$ and $V(t,s)$ are uniformly bounded and

$$\|U(t,s)\| \le \exp[\alpha(t-s)], \quad \|V(t,s)\| \le \exp[\beta(t-s)]$$

with $\alpha < 0$ and $\beta < 0$.

From these inequalities one gets

$$\|\epsilon^{-1}W_\epsilon(t,s)\| \le \epsilon^{-1}\int_s^t ds'\, \exp[(\alpha/\epsilon)(t-s') + \beta(s'-s)] \\ \|P(s')\|\cdot\|Q(s')\| \\ \le [C_1^s/(\epsilon\beta - \alpha)]\,\{\exp[\beta(t-s)] - \exp[(\alpha/\epsilon)(t-s)]\},$$

where C_1 is a constant independent of ϵ. Thus the kernel $\epsilon^{-1}W_\epsilon(t,s)$ is uniformly bounded on $[0,T]$.

With the above properties of $\bar{m}_\epsilon(t)$ and $W_\epsilon(t,s)$ one can get from the Volterra equation the following inequality for the function $\bar{g}_\epsilon(t) = \|\bar{v}_\epsilon(t)\|$

$$\bar{g}_\epsilon(t) \le \bar{\varphi}_\epsilon(t) + C\int_0^t ds\, \bar{g}_\epsilon(s),$$

where

$$C = \sup_{0\le\epsilon,\, 0\le s\le t\le T} \|\epsilon^{-1}W_\epsilon(t,s)\|$$

and

$$\bar{\varphi}_\epsilon(t) = \|\bar{m}_\epsilon(t)\|.$$

The family $\bar{\varphi}_\epsilon(t)$ is uniformly bounded on $[0,T]$ and tends to zero almost uniformly on $(0,T]$.

From the above inequality it follows immediately that the family of functions $\bar{g}_\epsilon(t)$ is uniformly bounded in ϵ and t and

$$\bar{g}_\epsilon(t) \le M\, e^{CT}, \quad 0 \le t \le T, \quad 0 \le \epsilon,$$

where

$$M = \sup_{0\le\epsilon,\, 0\le t\le T} \bar{\varphi}_\epsilon(t).$$

Since $\bar{\varphi}_\epsilon(t)$ is almost uniformly tending to zero, for each $\delta > 0$ there exists $t_1(\delta) > 0$ and $\epsilon_1^{(1)}(\delta) > 0$ such that for $\epsilon < \epsilon_1^{(1)}(\delta)$ and $\epsilon t_1(\delta) \le t \le T$

$$\bar{\varphi}_\epsilon(t) < \tfrac{1}{2}\delta\exp(-CT).$$

Let $\epsilon_1^{(1)}(\delta) > 0$ be such that for $\epsilon < \epsilon_1^{(2)}(\delta)$

$$C\int_0^{\epsilon t_1(\delta)} ds\, \bar{g}_\epsilon(s) < \tfrac{1}{2}\delta\exp(-CT).$$

This is always possible since the family $\bar{g}_\epsilon(t)$ is uniformly bounded.

Thus for $\epsilon < \epsilon_1(\delta) = \min(\epsilon_1^{(1)}(\delta), \epsilon_1^{(2)}(\delta))$ and $\epsilon t_1(\delta) \le t \le T$

$$\bar{g}_\epsilon(t) < \delta e^{-CT} + C\int_{\epsilon t_1(\delta)}^t ds\, \bar{g}_\epsilon(s)$$

from which it follows that

$$\bar{g}_\epsilon(t) < \delta e^{-CT}\exp\{C[T - \epsilon t_1(\delta)]\} \le \delta,$$

for $\epsilon < \epsilon_1(\delta)$ and $\epsilon t_1(\delta) \le t \le T$. This shows that the family $\bar{g}_\epsilon(t) = \|\bar{v}_\epsilon(t)\|$ tends to zero almost uniformly on $(0,T]$.

Taking the norm of $\bar{w}_\epsilon(t)$ expressed in terms of $\bar{v}_\epsilon(t)$ one has

$$\|\bar{w}_\epsilon(t)\| \le C_2\int_0^t ds\, \bar{g}_\epsilon(s),$$

where C_2 is a constant independent of ϵ.

Since $\bar{g}_\epsilon(t)$ is uniformly bounded on $[0,T]$ and tends almost uniformly to zero on $(0,T]$, for each $\delta > 0$ there exist $t_1'(\delta) > 0$ and $\epsilon_1'(\delta) > 0$ such that for $\epsilon < \epsilon_1'(\delta)$ and $\epsilon t_1'(\delta) \le t \le T$

$$g_\epsilon(t) < \delta/2C_2[T - \epsilon t_1'(\delta)]$$

and

$$\int_0^{\epsilon t_1'(\delta)} ds\, \bar{g}_\epsilon(s) < \delta/2C_2.$$

This shows that $\bar{w}_\epsilon(t)$ tends to zero almost uniformly on $(0,T]$. This concludes the proof of the lemma.

Lemma 13: Under the assumptions of Lemma 6 the equation of evolution

$$\epsilon\frac{d\tilde{x}_\epsilon(t)}{dt} = A(0)\tilde{x}_\epsilon(t) + P(0)y_0 + q(0), \quad \tilde{x}_\epsilon(0) = x_0 \in D(A)$$

has the unique strongly differentiable solution $\tilde{x}_\epsilon(t)$. The family $\{\tilde{x}_\epsilon(t), \tilde{y}_\epsilon(t)\}$, where $\tilde{y}_\epsilon(t) = y_0$ is called the inner asymptotic solution.

The family $\tilde{x}_\epsilon(t)$ tends almost uniformly on $(0,T]$ to

$$\tilde{\bar{x}}(t) = -A^{-1}(0)[P(0)y_0 + q(0)]$$

(the intermediate asymptotic solution).

For any $t_2 > 0$ and $\delta > 0$ there exists $\epsilon_2(t_2,\delta) > 0$ such that if $\epsilon < \epsilon_2(t_2,\delta)$ then for $0 \le t \le \epsilon t_2$

$$\|x_\epsilon(t) - \tilde{x}_\epsilon(t)\| < \delta, \quad \|y_\epsilon(t) - \tilde{y}_\epsilon(t)\| < \delta.$$

Proof: The existence of $\tilde{x}_\epsilon(t)$ follows from Lemma 4 and its asymptotic behavior for $\epsilon \to 0$ from Lemma 9.

To prove the last statement of the lemma, consider the functions

$$\tilde{v}_\epsilon(t) = x_\epsilon(t) - \tilde{x}_\epsilon(t), \quad \bar{w}_\epsilon(t) = y_\epsilon(t) - \tilde{y}_\epsilon(t)$$

which satisfy the following system of equations:

$$\epsilon\frac{d\tilde{v}_\epsilon(t)}{dt} = A(t)\tilde{v}_\epsilon(t) + P(t)\tilde{w}_\epsilon(t) + \tilde{q}_\epsilon(t).$$

$$\frac{d\tilde{w}_\epsilon(t)}{dt} = Q(t)\tilde{v}_\epsilon(t) + B(t)\tilde{w}_\epsilon(t) + \tilde{r}_\epsilon(t),$$

with the initial conditions

$$\tilde{v}_\epsilon(0) = \tilde{w}_\epsilon(0) = 0.$$

The functions $\tilde{q}_\epsilon(t)$ and $\tilde{r}_\epsilon(t)$ are defined as

$$\tilde{q}_\epsilon(t) = [A(t) - A(0)]\tilde{x}_\epsilon(t) + [P(t) - P(0)]y_0 + q(t) - q(0),$$

$$\tilde{r}_\epsilon(t) = Q(t)\tilde{x}_\epsilon(t) + B(t)y_0 + r(t).$$

By Lemma 6 the above system has a unique solution $\{\tilde{v}_\epsilon(t); \tilde{w}_\epsilon(t)\}$ and by Lemma 8 it is equivalent to the Volterra type equation

$$\tilde{v}_\epsilon(t) = \tilde{m}_\epsilon(t) + \epsilon^{-1}\int_0^t ds\, W_\epsilon(t,s)\,\tilde{v}_\epsilon(s),$$

where

$$\tilde{m}_\epsilon(t) = \epsilon^{-1}\int_0^t ds\, U_\epsilon(t,s)\,\tilde{q}_\epsilon(s) + \epsilon^{-1}\int_0^t ds\, Z_\epsilon(t,s)\tilde{r}_\epsilon(s),$$

and $Z_\epsilon(t,s)$ and $W_\epsilon(t,s)$ were defined in the proof of Lemma 12.

The family $\tilde{x}_\epsilon(t)$ is uniformly bounded in ϵ and t. The same is also true with $A(0)\tilde{x}_\epsilon(t)$. This follows from the equation for $\tilde{x}_\epsilon(t)$ and the fact that $\epsilon[d\tilde{x}_\epsilon(t)/dt]$ and $P(0)\tilde{y}_\epsilon(t)$ are uniformly bounded on $[0,T]$.

Now the function $t \to A(t)x$ is strongly continuous for any $x \in D(A)$ and the family of operators $A(t)A^{-1}(0)$ is uniformly bounded. Thus the function $t \to A(t)A^{-1}(0)$ is uniformly continuous and the quantity

$$\|[A(t) - A(0)]\tilde{x}_\epsilon(t)\| = \|[A(t) - A(0)]A^{-1}(0)A(0)\tilde{x}_\epsilon(t)\|$$

tends to zero with $t \to 0$ uniformly in ϵ.

By assumptions of Lemma 6 the function $t \to P(t)$ is uniformly continuously differentiable and the function $t \to q(t)$ is strongly continuously differentiable. Thus the remaining terms in $\tilde{q}_\epsilon(t)$ and the family $\tilde{q}_\epsilon(t)$ itself tend to zero in the norm with $t \to 0$ uniformly in ϵ.

The operator family $Q(t)$ is uniformly bounded and the family $\tilde{x}_\epsilon(t)$ is uniformly bounded in t and ϵ. The terms

$$B(t)y_0 = B(t)B^{-1}(0)B(0)y_0$$

and $r(t)$ are also uniformly bounded. Thus it is seen that $\tilde{r}_\epsilon(t)$ is uniformly bounded in t and ϵ.

Introducing notation

$$\tilde{\varphi}_\epsilon(t) = \|\tilde{m}_\epsilon(t)\|$$

and using the inequalities satisfied by the norms of $U(t,s)$ and $V(t,s)$ introduced in the proof of Lemma 12 one gets from the definition of $\tilde{m}_\epsilon(t)$ the inequality

$$\tilde{\varphi}_\epsilon(t) \le \epsilon^{-1}\int_0^t ds\, \exp[(\alpha/\epsilon)(t-s)]\|\tilde{q}_\epsilon(s)\|$$
$$+\,\epsilon^{-1}\int_0^t ds \int_s^t ds'\, \exp[(\alpha/\epsilon)(t-s') + \beta(s'-s)]$$
$$\|P(s')\|\cdot\|\tilde{r}_\epsilon(s)\|.$$

Substituting $t = \epsilon t_2$, changing the dummy variable in the integrals, and denoting

$$C_3 = \sup_{0 \le s, s' \le T} \|P(s')\|\cdot\|\tilde{r}_\epsilon(s)\|,$$

one has

$$\tilde{\varphi}_\epsilon(\epsilon t_2) \le \int_0^{t_2} ds'\, \|\tilde{q}_\epsilon(\epsilon s')\|$$
$$+\, C_3\{[\epsilon/\alpha(\beta\epsilon - \alpha)](1 - e^{\alpha t_2}) - [\beta(\beta\epsilon-\alpha)]^{-1}(1 - e^{\epsilon\beta t_2})\}.$$

Thus for any $t_2 > 0$ and $\delta > 0$ there exists $\epsilon_2(t_2, \delta)$ such that if $\epsilon < \epsilon_2(t_2,\delta)$ and $0 \le t \le \epsilon t_2$ then

$$\varphi_\epsilon(t) < \delta.$$

Introducing now the notation

$$\tilde{g}_\epsilon(t) = \|\tilde{v}_\epsilon(t)\|$$

one has from the Volterra equation satisfied by $\tilde{v}_\epsilon(t)$ the inequality

$$\tilde{g}_\epsilon(t) \le \tilde{\varphi}_\epsilon(t) + C\int_0^t ds\, \tilde{g}_\epsilon(s).$$

The constant C was introduced in the proof of Lemma 12.

From the last inequality it follows that

$$\tilde{g}_\epsilon(t) \le \varphi_\epsilon(t)\, e^{Ct} \le \tilde{\varphi}_\epsilon(t)\, e^{CT}$$

which shows that for any $t_2 > 0$ and $\delta > 0$ there exists $\epsilon_2^{(1)}(t_2,\delta)$ such that if $\epsilon < \epsilon_2^{(1)}(t_2,\delta)$ and $0 \le t \le \epsilon t_2$ then

$$\tilde{g}_\epsilon(t) = \|x_\epsilon(t) - \tilde{x}_\epsilon(t)\| < \delta.$$

It follows from Lemma 8 that

$$\tilde{w}_\epsilon(t) = \int_0^t ds\, V(t,s)\,\tilde{r}_\epsilon(s) + \int_0^t ds\, V(t,s)\,Q(s)\,\tilde{v}_\epsilon(s).$$

Since $\tilde{r}_\epsilon(s)$ and $Q(s)\tilde{v}_\epsilon(s)$ are uniformly bounded on $[0,T]$ one gets from the last equation and from the inequality satisfied by $V(t,s)$

$$\|\tilde{w}_\epsilon(t)\| \le (C_4/\beta)(e^{\beta t} - 1)$$

which again shows that for any $t_2 > 0$ and $\delta > 0$ there exists $\epsilon_2^{(2)}(t_2,\delta)$ such that if $\epsilon < \epsilon_2^{(2)}(t_2,\delta)$ and $0 \le t \le \epsilon t_2$ then

$$\|\tilde{w}_\epsilon(t)\| = \|y_\epsilon(t) - y_\epsilon(t)\| < \delta.$$

Choosing $\epsilon_2 = \min(\epsilon_2^{(1)}, \epsilon_2^{(2)})$ one obtains the proof of the lemma.

Theorm 4: Under the assumptions of Lemma 11 the solution $\{x_\epsilon(t), y_\epsilon(t)\}$ tends with $\epsilon \to 0$ to the asymptotic solution $\{x_\epsilon^{(as)}(t), y_\epsilon^{(as)}(t)\}$, where

$$x_\epsilon^{(as)}(t) = \bar{x}(t) + \tilde{x}_\epsilon(t) - \tilde{\bar{x}}(t),$$

$$y_\epsilon^{(as)}(t) = \bar{y}(t) + \tilde{y}_\epsilon(t) - \tilde{\bar{y}}(t)$$

uniformly on $[0,T]$. The intermediate asymptotic solution $\tilde{\bar{y}}(t) = y_0$.

Proof: It follows from Lemmas 12 and 13 that $\{x_\epsilon(t), y_\epsilon(t)\}$ tends to $\{\bar{x}(t), \bar{y}(t)\}$ and $\{\tilde{x}_\epsilon(t), \tilde{y}_\epsilon(t)\}$ to $\{\tilde{\bar{x}}(t), \tilde{\bar{y}}(t)\}$ almost uniformly on $(0,T]$. Thus $\{V_\epsilon(t), W_\epsilon(t)\}$, where

$$v_\epsilon(t) = x_\epsilon(t) - x_\epsilon^{(as)}(t),$$

$$w_\epsilon(t) = y_\epsilon(t) - y_\epsilon^{(as)}(t)$$

tends to zero almost uniformly on $(0,T]$. This means that for any $\delta > 0$ there exists $t_0(\delta) > 0$ and $\epsilon_1(\delta) > 0$ such that for $\epsilon < \epsilon_1(\delta)$ and $\epsilon t_0(\delta) \le t \le T$

$$\|v_\epsilon(t)\| < \delta, \quad \|w_\epsilon(t)\| < \delta.$$

Next from Lemma 13 it follows that for $\epsilon < \epsilon_2(t_0(\delta), \tfrac{1}{2}\delta)$ and $0 \le t \le \epsilon t_0(\delta)$

$$\|x_\epsilon(t) - \tilde{x}_\epsilon(t)\| < \tfrac{1}{2}\delta, \quad \|y_\epsilon(t) - \tilde{y}_\epsilon(t)\| < \tfrac{1}{2}\delta.$$

Finally, applying the similar technique as was used in the proofs of Lemmas 12 and 13 one can show that there exists such $\epsilon_3(\delta) > 0$ that for $\epsilon < \epsilon_3(\delta)$ and $0 \le t \le \epsilon t_0(\delta)$

$$\|\bar{x}(t)-\tilde{\bar{x}}(t)\|<\tfrac{1}{2}\delta,\quad \|\bar{y}(t)-\tilde{\bar{y}}(t)\|<\tfrac{1}{2}\delta.$$

Now, taking $\epsilon\leqslant\min(\epsilon_1,\epsilon_2,\epsilon_3)$ and $0\leqslant t\leqslant T$ one has

$$\|v_\epsilon(t)\|<\delta,\quad \|w_\epsilon(t)\|<\delta$$

which shows that $\{V_\epsilon(t), W_\epsilon(t)\}$ tends to zero with $\epsilon\to 0$ uniformly on $[0,T]$.

APPLICATION TO THE NEUTRON TRANSPORT THEORY

Let R be a three-dimensional Euclidean space and let each point in R be represented by the vector ρ. Define an open bounded convex set $\Gamma\in R$ and assume that its boundary $\partial\Gamma$ is sufficiently smooth so that for almost all $\rho\in\partial\Gamma$ there exists a normal $\mathbf{n}(\rho)$ pointed outward.

Let R' be another three-dimensional Euclidean space and let each point in R' be represented by the vector $\mathbf{v}$. Denote by Ω the sphere in R' of the radius $|\mathbf{v}|=v_m$.

As the Banach space E used in previous sections take the Hilbert space of complex-valued square integrable functions defined over $\Gamma\times\Omega$.

Let $\nu(\rho,\mathbf{v};t)$ be the function defined over $\Gamma\times\Omega\times[0,T]$. It will be assumed that $\nu(\rho,\mathbf{v};t)$ is for each $(\rho,\mathbf{v})\in\Gamma\times\Omega$ continuously differentiable on $[0,T]$ and for each $t\in[0,T]$ the functions $\nu(\rho,\mathbf{v};t)$ and $\partial\nu(\rho,\mathbf{v};t)/\partial t$ are square summable over $\Gamma\times\Omega$.

Let the functions $K(\rho;\mathbf{v},\mathbf{v}';t)$ and $Q(\rho;\mathbf{v},\mathbf{v}';t)$ are defined over $\Gamma\times\Omega\times\Omega\times[0,T]$ and have the properties analogous to $\nu(\rho,\mathbf{v};t)$ with Ω replaced by $\Omega\times\Omega$.

Let the operator $A(t)$ be defined as

$$(A(t)f)(\rho,\mathbf{v})=-\mathbf{v}\cdot\mathrm{grad}_\rho f(\rho,\mathbf{v})-\nu(\rho,\mathbf{v};t)f(\rho,\mathbf{v})$$
$$+\int_\Omega d\mathbf{v}'\,K(\rho;\mathbf{v},\mathbf{v}';t)f(\rho,\mathbf{v}'),\quad \rho\in\Gamma,\ \mathbf{v}\in\Omega,\ t\in[0,T]$$

with the domain:

$D(A)=\{f\in E;$ (i) $f(\rho-s\mathbf{v},\mathbf{v})$ is absolutely continuous in s for almost all $\rho\in\Gamma$, $\mathbf{v}\in\Omega$ and s such that $\rho-s\mathbf{v}\in\Gamma$, (ii) $Af\in E$, (iii) $f(\rho,\mathbf{v})=0$, $\rho\in\partial\Gamma$, $\mathbf{v}\cdot\mathbf{n}(\rho)<0\}$.

Let finally $Q(t)$ be the bounded operator defined as

$$(Q(t)f)(\rho,\mathbf{v})=\int_\Omega d\mathbf{v}'\,Q(\rho;\mathbf{v},\mathbf{v}';t)f(\rho,\mathbf{v}'),\quad \rho\in\Gamma,\quad \mathbf{v}\in\Omega,$$
$$t\in[0,T]$$

and the operators $P(t)$ and $B(t)$ denote the multiplication by λ and $-\lambda$, respectively, where λ is a positive constant.

With the above definitions the system of equations considered in Lemma 6 describes the behavior of neutrons in a reactor system with one group of delayed neutrons if $\nu(\rho,\mathbf{v};t)$ is the neutron collision frequency, $K(\rho;\mathbf{v},\mathbf{v}';t)$ is the kernel describing the scattering and fission, $Q(\rho;\mathbf{v},\mathbf{v}';t)$ is the kernel describing the production of delayed neutron precursors, λ is the precursor decay constant, $x(\rho,\mathbf{v},t)$ denotes the neutron distribution, $y(\rho,\mathbf{v},t)$ denotes the precursor distribution and finally, $q(\rho,\mathbf{v},t)$ and $r(\rho,\mathbf{v},t)$ are the external sources of neutrons and precursors, respectively.

It has been shown by Marti[12] that the operator $A(\tau)$ as defined above for any fixed $\tau\in[0,T]$ generates a strongly continuous semigroup $G_\tau(t)$. If the system is subcritical then

$$\|G_\tau(t)\|\leqslant\exp(\alpha_\tau t),$$

with $\alpha_\tau<0$ for $\tau\in[0,T]$.

The source functions are rather smooth in practical applications and may be assumed to obey the requirements of Lemma 6.

The small dimensionless parameter $\in>0$ appears at the time derivative of the neutron distribution $dx(t)/dt$ if the system of equations describing the reactor with one group of delayed neutrons is put into the dimensionless form in which the coefficients are independent of the system of units (see Ref. 2).

Now it is seen that all the requirements of previous sections are satisfied for the equations describing the evolution of a reactor system without or with one group of delayed neutrons and the results obtained valid in this case.

The considerations in this paper could be easily extended to the systems containing more than two equations of evolution, thus accounting for more than one group of delayed neutrons.

ACKNOWLEDGMENTS

The author is indebted to Dr. J. Kisyński and Mr. T. Duracz for the critical remarks and helpful suggestions in the course of preparation of the manuscript.

[1]W. L. Hendry and G. I. Bell, Nucl. Sci. Eng. **35**, 240 (1969).
[2]W. L. Hendry, J. Math. Phys. **11**, 1743 (1970).
[3]W. L. Hendry, Nucl. Sci. Eng. **45**, 1 (1971).
[4]M. Van Dyke, *Perturbation Methods in Fluid Mechanics* (Academic, New York, 1964), p. 77.
[5]J. D. Cole, *Perturbation Methods in Applied Mathematics* (Blaisdell, Waltham, Mass., 1968), p. 4.
[6]S. Kaplun, *Fluid Mechanics and Perturbations*, edited by P. A. Lagerstrom *et al.* (Academic, New York, 1967).
[7]N. Dunford and J. T. Schwartz, *Linear Operators* (Interscience, New York, 1958), Part I.
[8]T. Kato, J. Math. Soc. Jap. **5**, 208 (1953).
[9]J. Kisynski, Stud. Math. **23**, 285 (1964).
[10]S. G. Krein, *Linear Differential Equations in Banach Space* (Nauka, Moscow, 1967) (in Russian).
[11]The term quasi-semigroup is not used in the literature but it seems to the author to be quite adequate.
[12]J. T. Marti, Nukleonik **8**, 159 (1966).

On mass zero indecomposable representations of the Poincaré group

P. M. Mathews*

Center for Particle Theory, The University of Texas at Austin, Austin, Texas 78712

M. Seetharaman and M. T. Simon†

Department of Theoretical Physics, University of Madras, Madras-25, India
(Received 4 May 1973)

Considering the space of massless-particle wavefunctions transforming locally according to the irreducible representation $D^{(m,n)}$ of the homogeneous Lorentz group, we determine explicitly, by a direct and elementary method, the matrices constituting the (reducible but indecomposable) representation of Lorentz transformations with respect to a helicity basis.

I. INTRODUCTION

It is a well-known fact that in the group theoretical classification of elementary particles according to Wigner[1] and Bargmann and Wigner,[2] any massless particle is associated with an irreducible representation (IR) of the Poincaré group (PG) labelled by a vanishing value of the Casimir operator $P^\mu P_\mu$ and a definite (integral or half-integral) value of the helicity[3] λ defined through the relation $W^\mu = \lambda P^\mu$. Here the P^μ are the translation generators and $\{W^\mu\} = (\mathbf{J}\cdot\mathbf{P},\ P^0\mathbf{J} - \mathbf{P}\times\mathbf{K})$, where $\mathbf{J}$, $\mathbf{K}$ are the generators of rotations and boosts, respectively. Though an elementary massless particle has thus only a single value of the helicity, it is not customary to describe such a particle by a single-component wavefunction. Rather, one employs in general a multicomponent wavefunction which transforms locally according to some reducible or irreducible representation of the homogeneous Lorentz group (HLG). Such a wavefunction involves as many helicity values (some of which may be coincident) as there are components, unless the components are constrained in some suitable fashion.

The problem of constructing fields characterized by a *unique* helicity has been studied by Weinberg[4] who has shown that the only finite component covariant fields that can be constructed from entities transforming according to the Wigner IR $[0,\lambda]$ are those having the transformation property $D^{(m,m+\lambda)}$ with respect to the HLG, m being arbitrary. Conversely, it has been shown by a number of authors[5-7] that in a massless field which transforms locally according to $D^{(m,n)}$, there is just one helicity eigenstate, with $\lambda = n - m$, which remains invariant under Lorentz transformations. Every state with $\lambda \neq n - m$ develops admixtures with other helicity eigenstates. Thus the representation of the PG provided by wavefunctions transforming according to $D^{(m,n)}$ is not a direct sum of Wigner IR's, though it is reducible. In other words, the representation is indecomposable. The indecomposability has been interpreted in terms of gauge transformations by Shaw[5] and later by McKerrell.[6] Moses[8] has considered this aspect in some detail. However the explicit determination of the representation matrices in the helicity basis does not seem to have been carried out so far. Our objective in this paper is to make such a determination.

It is known that any Lorentz transformation Λ can be decomposed into a boost L and a rotation R. The representation of rotations in a helicity basis is diagonal, with diagonal elements which are known phase factors.[9,10] What needs to be done therefore is to determine the matrices representing boosts. Further, since any momentum vector can be brought into a specified direction by a rotation (whose representation is known, as just mentioned) it is sufficient to confine attention to the transformation under boosts L of helicity eigenstates defined with respect to a specified standard momentum, and we shall do so in this paper. Specifically, considering the space of $(2m+1)(2n+1)$-component wavefunctions transforming as $\psi(p) \to \psi'(\Lambda p) = D^{(m,n)}(\Lambda)\,\psi(p)$, we determine the matrix representing the effect of boosts on the helicity eigenfunctions for a given momentum. In a subsequent paper we shall explore some of the implications of the indecomposability of the representation in the context of the quantum theory of the electromagnetic fields described by the vector field A^μ.

II. THE REPRESENTATION OF BOOSTS IN HELICITY BASIS

Let us recall first that $D^{(m,n)}$ is defined by $\mathbf{M}^2 \to m(m+1)$, $\mathbf{N}^2 \to n(n+1)$, where

$$\mathbf{M} \equiv \tfrac{1}{2}(\mathbf{J} + i\mathbf{K}) \quad \text{and} \quad \mathbf{N} \equiv \tfrac{1}{2}(\mathbf{J} - i\mathbf{K}) \tag{1}$$

are two angular-momentum-like vectors obeying the algebra

$$[M_i, M_j] = i\epsilon_{ijk}M_k, \quad [N_i, N_j] = i\epsilon_{ijk}N_k, \quad [M_i, N_j] = 0. \tag{2}$$

Since $\mathbf{M}$ and $\mathbf{N}$ commute, any finite Lorentz transformation $\exp(i\boldsymbol{\theta}\cdot\mathbf{J} + i\boldsymbol{\alpha}\cdot\mathbf{K})$ can be expressed as $\exp(i\boldsymbol{\zeta}\cdot\mathbf{M}) \times \exp(i\boldsymbol{\zeta}^*\cdot\mathbf{N})$, where $\boldsymbol{\theta}$, $\boldsymbol{\alpha}$ are real vectors, and $\boldsymbol{\zeta} = \boldsymbol{\theta} - i\boldsymbol{\alpha}$. The space of spinors ψ transforming according to $D^{(m,n)}$ may thus be thought of as a direct product of a $(2m+1)$-dimensional "M space" on which the matrix operator $\exp(i\boldsymbol{\zeta}\cdot\mathbf{M})$ acts, and a $(2n+1)$-dimensional "N space" acted on by $\exp(i\boldsymbol{\zeta}\cdot\mathbf{N})$. In the case of a pure Lorentz transformation (boost),

$$\boldsymbol{\zeta} = -\boldsymbol{\zeta}^* = -i\boldsymbol{\alpha} = -i\hat{\mathbf{n}}\alpha, \tag{3}$$

where $\hat{\mathbf{n}}$ is the unit vector in the direction of the boost. Hence the transformation of a wavefunction $\psi(p)$ under boosts is

$$\psi(p) \to \psi'(p') = D(L)\,\psi(p), \tag{4a}$$

$$D(L) = \exp(i\boldsymbol{\zeta}\cdot\mathbf{M})\exp(-i\boldsymbol{\zeta}\cdot\mathbf{N}), \tag{4b}$$

where the two factors in $D(L)$ act on two separate indices labelling the components of ψ.

We now introduce, for any lightlike 4-momentum p, sets of "helicity" eigenfunctions in the M and N spaces, defined by

$$(\mathbf{M}\cdot\mathbf{p})u^{(M)}(p;\mu)=\mu|\mathbf{p}|u^{(M)}(p;\mu),$$

$$\mu=m, m-1, \ldots, -m, \tag{5a}$$

$$(\mathbf{N}\cdot\mathbf{p})u^{(N)}(p;\nu)=\nu|\mathbf{p}|u^{(N)}(p;\nu), \quad \nu=n, n-1, \ldots, -n. \tag{5b}$$

Since $\mathbf{J}=\mathbf{M}+\mathbf{N}$, the direct products

$$u(p;\mu,\nu)\equiv u^{(M)}(p;\mu)u^{(N)}(p;\nu) \tag{6a}$$

of these functions are eigenfunctions of the helicity operator $(\mathbf{J}\cdot\mathbf{p})/|\mathbf{p}|$, and they form a basis for the space of functions $\psi(p)$:

$$(\mathbf{J}\cdot\mathbf{p})u(p;\mu,\nu)=\lambda u(p;\mu,\nu), \quad \lambda=\mu+\nu. \tag{6b}$$

Now, let

$$\exp(i\boldsymbol{\zeta}\cdot\mathbf{M})u^{(M)}(p;\mu)=\sum_\rho u^{(M)}(p';\rho)d^{(M)}_{\rho\mu}, \tag{7a}$$

$$\exp(-i\boldsymbol{\zeta}\cdot\mathbf{N})u^{(N)}(p;\nu)=\sum u^{(N)}(p';\sigma)d^{(N)}_{\sigma\nu}, \tag{7b}$$

where the helicity states appearing on the right-hand side are defined with respect to the transformed momentum $p'=Lp$. Then the matrix $\mathcal{D}(L;p)$ with elements $\mathcal{D}_{\rho\sigma,\mu\nu}(L;p)$ defined by

$$D^{(m,n)}(L)u(p;\mu,\nu) =\sum_{\rho\sigma} d^{(M)}_{\rho\mu}d^{(N)}_{\sigma\nu}u(p';\rho,\sigma)\equiv\sum\mathcal{D}_{\rho\sigma,\mu\nu}u(p';\rho,\sigma), \tag{8}$$

evidently gives a helicity representation of the boost L. Our task now is to determine this matrix $\mathcal{D}$.

Consider the relation

$$\exp(i\boldsymbol{\zeta}\cdot\mathbf{M})\,[\exp(-i\boldsymbol{\zeta}\cdot\mathbf{M})(\mathbf{M}\cdot\mathbf{p}')\exp(i\boldsymbol{\zeta}\cdot\mathbf{M})]u^{(M)}(p;\mu) =\sum_\rho(\mathbf{M}\cdot\mathbf{p}')u^{(M)}(p';\rho)d^{(M)}_{\rho\mu}=\sum_\rho u^{(M)}(p';\rho)\rho|\mathbf{p}'|d^{(M)}_{\rho\mu}, \tag{9}$$

which follows immediately from (7a). As we shall now see, the first member of this equation can be easily expressed in terms of the $u^{(M)}(p';\rho)$ and hence a recurrence relation for the $d^{(M)}_{\rho\mu}$ obtained. For this purpose we need to use only the following elementary facts:

$$\exp(-i\zeta\cdot\mathbf{M})\mathbf{M}\exp(i\zeta\cdot\mathbf{M})=\mathbf{M}\cos\zeta-(\hat{\mathbf{n}}\times\mathbf{M})\sin\zeta -(\mathbf{M}\cdot\hat{\mathbf{n}})\hat{\mathbf{n}}(\cos\zeta-1), \tag{10}$$

$$\mathbf{p}'=\mathbf{p}+(\hat{\mathbf{n}}\cdot\mathbf{p})\hat{\mathbf{n}}(\cos\zeta-1)-i\hat{\mathbf{n}}|\mathbf{p}|\sin\zeta, \tag{11}$$

where n is defined as in (3), and $\zeta\equiv-i\alpha$ is pure imaginary. Equation (10) simply states the transformation property of a vector operator under rotations. (Since this relation depends only on the algebra of the components of $\mathbf{M}$, it holds irrespective of whether ζ is real or complex.) Equation (11) is just the 3-vector part of the boosted lightlike momentum vector. Combining Eqs. (10) and (11), we obtain

$$\exp(-i\zeta\cdot\mathbf{M})(\mathbf{M}\cdot\mathbf{p}')\exp(i\boldsymbol{\zeta}\cdot\mathbf{M})=(\mathbf{M}\cdot\mathbf{p})\cos\zeta-i(\hat{\mathbf{n}}\times\mathbf{M})\cdot\mathbf{p} -i(\hat{\mathbf{n}}\cdot\mathbf{M})|\mathbf{p}|\sin\zeta. \tag{12}$$

Let us specialize now to the standard vector $p=\mathring{p}$,

$$\mathring{p}=(\kappa,0,0,\kappa). \tag{13}$$

In this case $|\mathbf{p}'|=\kappa(\cos\zeta-i\hat{n}_3\sin\zeta)$, and we can write (12) then as

$$\exp(-i\boldsymbol{\zeta}\cdot\mathbf{M})(\mathbf{M}\cdot\mathbf{p}')\exp(i\boldsymbol{\zeta}\cdot\mathbf{M})=M_3|\mathbf{p}'|-(i\kappa\hat{n}_+\sin\zeta)M_-, \tag{14}$$

where

$$M_\pm=M_1\pm iM_2, \quad \hat{n}_\pm=\hat{n}_1\pm i\hat{n}_2. \tag{15}$$

On using (14), the first member of Eq. (9) becomes

$$\mu|\mathbf{p}'|\exp(i\boldsymbol{\zeta}\cdot\mathbf{M})u^{(M)}(\mathring{p};\mu)-(i\kappa\hat{n}_+\sin\zeta) \times[(m+\mu)(m-\mu+1)]^{1/2}\exp(i\boldsymbol{\zeta}\cdot\mathbf{M})u^{(M)}(\mathring{p};\mu-1). \tag{16}$$

Here we have used the fact that with p as in (13), Eq. (5a) reduces to

$$M_3u^{(M)}(\mathring{p};\mu)=\mu u^{(M)}(\mathring{p};\mu) \tag{17a}$$

and hence, from angular momentum theory,[11]

$$M_-u^{(M)}(\mathring{p};\mu)=[(m+\mu)(m-\mu+1)]^{1/2}u^{(M)}(\mathring{p};\mu-1). \tag{17b}$$

Finally, on substituting for the first member of (9) the expression (16), and using (7a), we obtain

$$\sum_\rho[\mu|\mathbf{p}'|d^{(M)}_{\rho\mu}-(i\kappa\hat{n}_+\sin\zeta) \times[(m+\mu)(m-\mu+1)]^{1/2}d^{(M)}_{\rho,\mu-1}]u^{(M)}(p';\rho) =\sum_\rho\rho|\mathbf{p}'|d^{(M)}_{\rho\mu}u^{(M)}(p';\rho). \tag{18}$$

On equating coefficients of the linearly independent functions $u^{(M)}(p';\rho)$, we get the recurrence relation

$$(\mu-\rho)d^{(M)}_{\rho\mu}=c[(m+\mu)(m-\mu+1)]^{1/2}d^{(M)}_{\rho,\mu-1}, \tag{19}$$

$$c=i(\kappa/|\mathbf{p}'|)\hat{n}_+\sin\zeta=(i\hat{n}_+\sin\zeta)/(\cos\zeta-i\hat{n}_3\sin\zeta) =(\hat{n}_+\sinh\alpha)/(\cosh\alpha-\hat{n}_3\sinh\alpha). \tag{20}$$

Equation (19) implies that $d^{(M)}_{\rho\mu}=0$ for all $\rho>\mu$. For any $\rho\leq\mu$ it gives

$$d^{(M)}_{\rho\mu}=\frac{c^{\mu-\rho}}{(\mu-\rho)!}\left(\frac{(m+\mu)!(m-\rho)!}{(m-\mu)!(m+\rho)!}\right)^{1/2}d^{(M)}_{\rho\rho}. \tag{21}$$

Thus any $u^{(M)}(\mathring{p};\mu)$ for given μ gets transformed into a linear combination of all the $u^{(M)}(p';\rho)$ with $\rho\leq\mu$. The only case where there is no admixture with other helicities is when μ itself has the minimum possible value, $\mu=-m$.

The above derivation does not give the values of the diagonal elements $d^{(M)}_{\rho\rho}$. It can be seen however that they must be of the form $\exp(i\rho\Theta)$. This is a consequence of the well-known fact that if a group representation is reduced to a form wherein all nonvanishing elements outside of diagonal blocks (square blocks along the diagonal) are confined to one side of the diagonal, then these diagonal blocks themselves must form representations of the group (even if the given representation is not fully reducible). In the present case the diagonal "blocks" are just the diagonal elements $d^{(M)}_{\rho\rho}$. Thus $d^{(M)}_{\rho\rho}$ for each ρ must give a representation of the group. This is of course just the helicity-ρ Wigner representation, which gives $d^{(M)}_{\rho\rho}=\exp(i\rho\Theta)$, where Θ is characteristic of the particular boost, and depends on p too.[4,9,12]

The coefficients $d^{(N)}_{\sigma\nu}$ in Eq. (7b) can be obtained in exactly the same way. The counterpart of Eq. (10) in this case is obtained by the replacements $\mathbf{M}\to\mathbf{N}$, $\boldsymbol{\zeta}\to-\boldsymbol{\zeta}$ (or $\hat{\mathbf{n}}\to-\hat{\mathbf{n}}$), while Eq. (11) remains unchanged. Then Eq. (14) goes over into one with N_3, N_+, and $\hat{n}_-$ instead of M_3, M_-, and $-\hat{n}_+$. Consequently the $d^{(N)}_{\sigma\nu}$ are nonzero only for σ *greater* than ν. It may be verified that they are given by

$$d^{(N)}_{\sigma\nu} = \frac{(-c^*)^{\sigma-\nu}}{(\sigma-\nu)!}\left(\frac{(n+\sigma)!(n-\nu)!}{(n-\sigma)!(n+\nu)!}\right)^{1/2} d^{(N)}_{\sigma\sigma}, \tag{22}$$

with $d^{(N)}_{\sigma\sigma} = \mathrm{rcp}(i\sigma\Theta)$. Thus we have finally

$$D_{\rho\sigma,\mu\nu} = \frac{c^{\mu-\rho}(-c^*)^{\sigma-\nu}}{(\mu-\rho)!(\sigma-\nu)!} \times \left(\frac{(m+\mu)!(m-\rho)!(n+\sigma)!(n-\nu)!}{(m-\mu)!(m+\rho)!(n-\sigma)!(n+\nu)!}\right)^{1/2} e^{i(\rho+\sigma)\Theta}, \tag{23}$$

where c is given by (20). The value of Θ remains undetermined until a specific prescription is adopted for the definition of the phases of helicity eigenfunctions $u(p;\lambda)$ for arbitrary momentum **p**. The convention used in Ref. 9 is to define $u(p,\lambda)$ as $RZ_0u(\mathring{p};\lambda)$, where Z_0 is the boost in the direction of the standard vector $\mathring{p}$ which changes $|\mathring{\mathbf{p}}|$ into $|\mathbf{p}|$, and the rotation R takes the direction of $\mathring{\mathbf{p}}$ into that of **p**. (The helicity remains unaltered in both these steps.) With the phases fixed in this manner, it follows from the work of Ref. 9 that $\Theta = 1$ for all boosts in the direction of $\mathring{p}$. Any other boost can be decomposed into a boost of this type together with rotations preceding and following it. The corresponding value of Θ is obtained then from the knowledge of the rotation properties of helicity functions.[10]

III. DISCUSSION

Though we have made use of the standard vector (13) in the derivation of Eq. (23), extension of this result to other vectors can be carried out easily. One way is to accomplish the transition from a reference frame S to the boosted one, S', in three steps: first a rotation in the plane containing the z axis and the vector under consideration, to bring the z axis along this vector; then the boost; and finally a rotation to bring the axes in coincidence with S'. For the second step, Eq. (23) applies, but with $\hat{\mathbf{n}}$ replaced by $\hat{\mathbf{n}}'$, the direction of the boost with respect to the axes obtained after the first rotation. The two rotations themselves are of course represented by known diagonal matrices in the helicity representation. An alternative procedure is based on the observation that Eq. (14), on which the remaining derivation rests, holds good for any vector p, provided that $\hat{\mathbf{n}}$ is interpreted as the direction of the boost with respect to an orthonormal triad of vectors of which the third one is along the direction of **p**. The direction cosines $\hat{\mathbf{n}}$ can of course be expressed in terms of those with reference to a fixed coordinate system (i.e., one which is independent of **p**), and with the use of such expressions for the components of $\hat{\mathbf{n}}$, Eq. (23) becomes valid for any arbitrary vector.

Finally, as we have noted in the Introduction, any general Lorentz transformation Λ can be decomposed into $\Lambda = RL$, where R is a rotation and L is a boost, and hence the matrix $D(\Lambda;p)$ representing Λ may be determined. The set of matrices $D(\Lambda;p)$ for all Λ and p provides a helicity representation of the Lorentz group in the sense that $D(\Lambda_2;\Lambda_1 p)D(\Lambda_1;p) = \pm D(\Lambda_2\Lambda_1;p)$. It is well known[1] that ambiguous ($\pm$) signs arise in the case of the double-valued representations corresponding to half-odd-integral values of $(m+n)$. As for translations, they are represented by the familiar phase factors $\exp(ip\cdot a)$, and these taken together with the $D(\Lambda;p)$ corresponding to homogeneous Lorentz transformations, complete the representation of the Poincaré group.

It may be easily verified that the matrix D defined by Eq. (23) for given p is not unitary. This is related to the fact that the little group of the Poincaré group associated with any lightlike four-vector is the noncompact group $E(2)$. It is also a necessary concomitant of indecomposability: if a representation were reducible and unitary, it would be completely reducible. In a sequel to this paper, we shall show that the gauge problem and the appearance of the indefinite metric in the quantum theory of the electromagnetic potentials A_μ have their origin in the indecomposability (and nonunitarity) of the representation of the little group associated with any lightlike vector p over functions transforming according to $D^{(1/2,1/2)}$ under the HLG.

ACKNOWLEDGMENTS

It is a pleasure to thank Professor N. Mukunda and Professor E.C.G. Sudarshan for useful comments on the manuscript. One of us (M.T.S.) gratefully acknowledges the grant of study leave by the authorities of St. Thomas College, and financial support from the University Grants Commission in the form of a Fellowship.

*Permanent address: Department of Theoretical Physics University of Madras, Madras, India.
†Present address: Department of Physics, St. Thomas College, Kozhencherry, Kerala, India.

[1] E.P. Wigner, Ann. Math. 40, 149 (1939).
[2] V. Bargmann and E.P. Wigner, Proc. Natl. Acad. Sci. USA 34, 211 (1948).
[3] The possibility of constructing massless particle theories associated with the "continuous spin" representations has been considered recently by G.J. Iverson and G. Mack, Ann. Phys. (N.Y.) 64, 211 (1971).
[4] S. Weinberg, Phys. Rev. 134, B882 (1964).
[5] R. Shaw, Nuovo Cimento 37, 1086 (1965).
[6] A. McKerrell, Ann. Phys. (N.Y.) 40, 237 (1966).
[7] M.T. Simon, M. Seetharaman, and P.M. Mathews, Int. J. Theor. Phys. 6, 399 (1972). In this paper, it is shown that no invariant helicity states exist in the case of unitary (infinite dimensional) IR's of the HLG, and those associated with infinite dimensional nonunitary IR's are identified.
[8] H.E. Moses, J. Math. Phys. 9, 16 (1968).
[9] A. McKerrell, Proc. R. Soc. A 285, 287 (1965).
[10] M. Jacob and G.C. Wick, Ann. Phys. (N.Y.) 7, 404 (1959).
[11] Equation (17b) is true to within a phase factor which may depend on p. This factor has been chosen unity in accordance with the Condon and Shortley convention, in the case of the standard momentum vector p of (13).
[12] Iu.M. Shirokov, Zh. Eksp. Teor. Fiz. 33, 1208 (1957) [Sov. Phys.—JETP 6, 929 (1958)].

Rigged Hilbert space formalism as an extended mathematical formalism for quantum systems. I. General theory

Olaf Melsheimer

Institut für Theoretische Physik, Universität Marburg, Marburg, Germany
(Received 6 September 1972)

Roberts' proposal of a rigged Hilbert space $\Phi \subset \mathcal{G} \subset \Phi^{\times}$ for a certain class of quantum systems is reinvestigated and developed in order to exhibit various properties of this kind of rigged Hilbert spaces which might be of interest for the application of this formalism to special quantum systems. It is shown that on the basis of this proposal one also obtains a satisfactory solution for a rigged Hilbert space for composite systems. Another part is concerned with topological properties of the so-called eigenoperators $\gamma(t)$ belonging to an A-eigen-integral decomposition of Φ with respect to a self-adjoint operator A on $\mathcal{G}$. We derive a representation of $\gamma(t)$ in terms of the generalized eigenvectors of A and in the same context give a rough topological characterization of these eigenvectors.

I. INTRODUCTION

The present paper is mainly intended to give a systematic description of a mathematical apparatus which, in the opinion of the author, is suited to cope with some problems inherent in the Hilbert space formalism for quantum systems. In the usual terminology problems of this type are named "choice of the representation," "transformation of one representation into another," and "representation of operators in a given representation." As it is well-known all these problems are connected with the presence of continuous parts in the spectra of most of the self-adjoint operators in a separable Hilbert space $\mathcal{G}$ which represent observables (physical quantities) in the mathematical scheme of quantum mechanics. On the other hand, in scattering theory it is sometimes very useful to work with plane wave "states" which, in a certain idealization, may still be regarded as physical states. Since these states no longer belong to the Hilbert space in question one runs into mathematical troubles when applying Hilbert space operators onto these states.

In order to be able to treat all these problems in a unified scheme we present in this paper a mathematical apparatus based on a formalism which for the first time was systematically developed by Gel'fand[1] and is therefore known as the Gel'fand-triplet-formalism; very often it is also called "rigged Hilbert space formalism." In the scope of Gel'fand's work this formalism was intended to provide an extended space $\Phi^{\times}$ of $\mathcal{G}$ which should contain "generalized eigenvectors," corresponding to points of the continuous spectrum of a given self-adjoint operator in $\mathcal{G}$.

The general idea is to replace the Hilbert space structure by a space triplet $\Phi \subset \mathcal{G} \subset \Phi^{\times}$, where Φ is a dense subspace of $\mathcal{G}$ equipped with a finer topology and $\Phi^{\times}$ is the dual space of Φ. The above inclusions then are defined by means of the continuous canonical injection I: $\Phi \to \mathcal{G}$ and its adjoint $I^{\times}: \mathcal{G} \to \Phi^{\times}$. Applications of the rigged Hilbert space formalism to "physics" have already been investigated by Böhm,[2] Roberts,[3] and Antoine.[4] In this context we should also mention the work of Grossman[5] and Prugovečki.[6] These authors have tackled the problem of modifying the usual Hilbert space structure by methods closely related to those employed in the rigged Hilbert space formalism. For an early attempt to treat similar problems we mention also the work of Mayer.[7] However, it is still an unsolved question whether all the new technical features that appear in this formalism admit of a clear-cut physical interpretation, if any at all.

As for the physical relevance of the rigged Hilbert space structure we understand the problem as follows: As long as one deals with ordinary quantum systems the basic mathematical structure which carries all the *physical* interpretation is after all the Hilbert space structure. In special cases some elements of the space $\Phi^{\times}$ may be regarded as idealized physical states, for instances, in scattering theory. However, in doing so one has to be very cautious and always to keep in mind the original physical states from which these idealized states have been abstracted. Thus, at the present stage we would say that there is no general rule of assigning a clear-cut "physical" meaning to all elements of $\Phi^{\times}$. In particular, we cannot agree with the proposal of Antoine[4] to interpret all elements of $\Phi^{\times}$ as potential experiments which can be performed on the physical system in question. We have to refuse this interpretation for the simple reason that the definition of the space $\Phi^{\times}$ itself depends very strongly on the topology of Φ; the choice of this topology is to a certain extent arbitrary and cannot be properly justified by physical arguments. The same argument applies to the space $\mathcal{L}(\Phi, \Phi^{\times})$. Furthermore—again for physical reasons—we cannot accept without further criteria the proposal of Antoine that the set of all "preparable" states of the physical system can be identified with the set Φ since this assumption would necessarily require an additional postulate for the characterization of the "preparable" states of a physical system. We would like to argue as follows: From the purely physical point of view it is always sufficient to work with a dense linear subspace Φ of $\mathcal{G}$ since up to an arbitraryly small error of measurement we can approximate any state by an element of Φ. Then we try to choose Φ in such a way that all elements of Φ have nice properties with respect to certain mathematical requirements. For instance, we can choose Φ as the subset of all infinitely differentiable vectors for a strongly continuous unitary representation of the central extension of the Galilei group in $\mathcal{G}$ which form a dense set in $\mathcal{G}$. This resembles very much the situation in ordinary mechanics where we usually work with differentiable functions. Physics tells us at most that the path of a particle should be represented by a continuous function. However, in order to have a handy formalism we confine ourselves to a subspace of the space of all continuous functions, namely, the space of differentiable or—more restrictively—infinitely differentiable functions. Summarizing the argument we may say that it is largely a matter of mathematical convenience to use the rigged Hilbert space as a formalism for the description of quantum systems instead of the usual Hilbert space structure. In our opinion, the rigged Hilbert space formalism does not bring about any enrichment of the structure of the physical theory proper; however, it may provide powerful means for the solution of certain mathematical problems of quantum mechanics thus possibly leading to new results (e.g., in scattering theory). Also it makes possible a rigorous formulation of this *extended* ("idealized") quantum mechanical scheme.

Now, in order to make the rigged Hilbert space formalism efficient for its application to quantum systems, one has to construct a triplet $\Phi \subset \mathcal{G} \subset \Phi^\times$ satisfying a special requirement: $\Phi^\times$ should contain the generalized eigenvectors for all observables that belong to the system in question. Even such a triplet is in no way unique! For the case where the set of observables is supposed to be an algebra $\mathfrak{A}$ (with Φ as the common invariant domain of definition), a construction of such a triplet was given by Roberts.[3] Since in this case the topology of Φ is essentially the initial topology with respect to $\mathfrak{A}$ which is mathematically well-known, we took this solution as a basis of our investigations.

The essential result within this context is a general representation of the elements of the dual space $\Phi^\times$. We shall need this result in a forthcoming paper in which we will treat the transformation theory of nonrelativistic quantum mechanics.

However, we emphasize that in principle the solution of Roberts[3] could exclude some of the observables which might be of interest to physics. This complex of questions is intimately connected with the problem which self-adjoint operators on Hilbert space correspond to *physical* observables; these problems will form the subject of further research.

In this paper we have adopted the scheme of spectral decompositions originally proposed by Foias[8] and later modified by Roberts[3] by using the Gårding–Maurin theorem.[9] This scheme is essentially based on the notion of the so-called eigenoperator $\gamma(t)$ of an operator A. $\gamma(t)$ is in a way a generalization of the Hilbert space projector on the eigenspace of a Hilbert space operator A. Foias[10] had already given a representation of the operators $\gamma(t)$ in terms of the generalized eigenvectors of A. However, for technical reasons, he had to confine himself to the case where Φ and $\Phi^\times$ are normed spaces. It is, of course, always possible, even in the general case of locally convex spaces Φ and $\Phi^\times$ to reduce this representation problem to the case of normed spaces Φ' and $\Phi'^\times$ (by choosing suitable subspaces Φ' and $\Phi'^\times$ of Φ and $\Phi^\times$, respectively). For physical applications, however, it is desirable to work exclusively in the general structure with topologies corresponding to the duality of the spaces Φ and $\Phi^\times$.

Working in this general framework we have obtained a representation of the operators $\gamma(t)$ with results very similar to those of Foias.[10] We have assumed here that $\Phi^\times$ is equipped with the strong topology $\beta(\Phi^\times, \Phi)$.

Last but not least we emphasize that we regard the work as presented in this paper as an extension and a completion of the work of Roberts[3] and Foias.[10,8]

However, we think that the mathematical apparatus as presented in this paper has now reached a stage of a satisfactory generality and thus can provide a sound basis for the treatment of some open mathematical problems within the Hilbert space formalism of quantum mechanics. As an application of this formalism the transformation theory in nonrelativistic quantum mechanics will be treated in a forthcoming paper.

II. THE NEED FOR A NEW FORMALISM

In almost all physical theories that are based on the Hilbert space formalism one has to deal with unbounded self-adjoint operators. It is well-known that such operators have a unique spectral decomposition

$$A = \int_{-\infty}^{\infty} \lambda dE_\lambda.$$

From the mathematical point of view this decomposition is very satisfactory. But as for practical purposes one is faced with the shortcoming that the "eigenfunctions" belonging to the continuous spectrum do not belong to the Hilbert space in question. This difficulty can partly be overcome through the use of the so-called eigenpackets. However, even then the discrete and the continuous spectrum, respectively, cannot be treated on the same footing. Just for this reason the choice of the representation and the transformation theory still represent unsolved problems in quantum theory.

On the other hand, the formal eigenfunctions belonging to the continuous spectrum are very often well-behaved functions (in particular, for differential operators) although they are not elements of the Hilbert space. The problem we are concerned with may be tentatively formulated as follows: Does there exist a space X that contains the given Hilbert space $\mathcal{G}$ as a subspace and such that the eigenelements of a given unbounded self-adjoint operator A in $\mathcal{G}$, both for the continuous and the discrete spectrum, belong to X and satisfy a completeness relation, i.e., every element of $\mathcal{G}$ can be decomposed with respect to this set of eigenelements in a well-defined fashion. Now the general ideal underlying the construction of such a space X is essentially the following: Let A be a cyclic bounded or unbounded self-adjoint operator A with the domain $D(A)[\Phi \subseteq D(A) \subset \mathcal{G}]$. Let

$$\mathcal{G} \to \hat{\mathcal{G}} = \int \hat{\mathcal{G}}(\lambda) d\mu(\lambda),$$

$$f \to \{\hat{f}(\lambda)\}, \quad f \in \mathcal{G}, \ \hat{f}(\lambda) \in \hat{\mathcal{G}}(\lambda),$$

$$Af \to \{\lambda \hat{f}(\lambda)\}, \quad f \in \Phi,$$

$$(f,g)_{\mathcal{G}} = \int \langle \hat{f}(\lambda), \ \hat{g}(\lambda) \rangle_\lambda d\mu(\lambda),$$

be the direct integral decomposition of $\mathcal{G}$ that originates from the spectral decomposition of A. Then one can search for a locally convex (l.c.) topology τ in Φ which is finer than the topology induced in Φ by the norm-topology of $\mathcal{G}$ such that the linear form $\phi \to \hat{\phi}(\lambda)$ with $\phi \in \Phi$ is a continuous linear form on Φ for every $\lambda \in (-\infty, \infty)$. Then one can write

$$\hat{\phi}(\lambda) = \langle \xi_\lambda, \phi \rangle \quad \text{for } \phi \in \Phi \quad \text{and} \quad \xi_\lambda \in \Phi',$$

where Φ' denotes the space of all continuous linear forms over Φ. Since Φ is dense in $\mathcal{G}$ and the canonical injection from Φ into $\mathcal{G}$ is continuous $\mathcal{G}$ may be densely embedded into Φ'. Thus we are led to the triplet structure

$$\Phi \subset \mathcal{G} \subset \Phi'.$$

Furthermore one can look for such a l.c. topology in which A becomes a continuous map from Φ into itself. Then there exists the transpose A' of the operator A that maps Φ'_β continuously into itself where the subscript β means that Φ' is endowed with the strong topology $\beta(\Phi', \Phi)$. A' is defined by the relation

$$\langle A'\xi, \phi \rangle = \langle \xi, A\phi \rangle \quad \text{for every } \phi \in \Phi \text{ and } \xi \in \Phi'.$$

The relation between A' and A^*, the adjoint of A ($A^x = A$ in $\mathcal{G}$), is not apparent at first sight, and will be cleared up later on. The linear forms ξ_λ introduced above are eigenforms of A'

$$\langle A'\xi_\lambda, \phi \rangle = \langle \xi_\lambda, A\phi \rangle = \lambda \langle \xi_\lambda, \phi \rangle \quad \text{for all } \phi \in \Phi.$$

This relation can be written in a shorter form

$$A'\xi_\lambda = \lambda \xi_\lambda.$$

After these introductory heuristic remarks we may now give a correct mathematical formulation of the problem to be solved.

(I) One has to look for a dense linear manifold Φ in $\mathcal{G}$ which belongs to the domain of definition for A and is invariant under A. Furthermore one has to endow Φ with a l.c. topology τ which is finer than the topology induced by $\mathcal{G}$ on Φ such that the operator A maps Φ continuously into itself.

(II) The dual space Φ' of Φ has to be sufficiently large so as to contain a complete system of eigenforms of A'. Completeness means that each element of Φ can be uniquely decomposed with respect to this complete system of eigenforms of A'. This system fulfills a completeness relation of the following form:

$$\|\phi\|^2 = \int |\langle \xi_\lambda, \phi \rangle|^2 d\mu(\lambda)$$

holds for every $\phi \in \Phi$.

For a given s.a. operator A on $\mathcal{G}$ this problem has already been solved by Foias.[8] We summarize his results in the following.

Proposition 1: Let A be a selfadjoint operator on $\mathcal{G}$ with a domain of definition D with $AD \subset D$ and D dense in $\mathcal{G}$. Then there exists a dense subspace Φ of $\mathcal{G}$ with $\Phi \subset D$ such that Φ can be equipped with a l.c. topology τ such that Φ is a Fréchet space. Furthermore with respect to τ, A becomes a continuous map of Φ into itself.

The situation one frequently encounters in physical applications is somewhat different from the situation described so far. Instead of one operator A one is given an algebra $\mathfrak{A}$ of unbounded operators (observables) which has the following properties:

(a) All elements of $\mathfrak{A}$ have a common, dense domain of definition Φ in $\mathcal{G}$ which is invariant under every element of $\mathfrak{A}$.

(b) Every element $A \in \mathfrak{A}$ admits a closure $\bar{A}$. We shall call Φ the basic domain for $\mathfrak{A}$.

Examples:

(1) In the case of nonrelativistic quantum mechanics of a single particle the algebra $\mathfrak{A}_{QM}$ is given by the envoping algebra of the generators of a unitary irreducible representation of a central extension of the Galilei group for $m \neq 0$. As it will be shown later Φ can be chosen to be the linear hull of the set of all infinitely differentiable vectors for the representation in question.

(2) In an axiomatic approach to quantum field theory one usually postulates that the algebra of field operators has a common dense invariant domain of definition in $\mathcal{G}$.

Now to cope with this situation first one has to settle the following problem: Φ has to be equipped with a l.c. topology τ finer than $\tau_{\mathcal{G}}$ (the topology on Φ induced by $\mathcal{G}$ on Φ) such that every $A \in \mathfrak{A}$ becomes a continuous map from Φ into itself. In the second step one has to find out which conditions should be imposed on $\mathfrak{A}$ in order that Φ becomes a nuclear locally convex space. Then the injection $I: \Phi \to \mathcal{G}$ (which is continuous) is also a nuclear map [Ref. 11, p. 100].

In the sequal we shall be concerned with the situation characterized by conditions (a) and (b). For the solution of (a) and (b) we shall adopt here essentially the method proposed by Roberts.[3] Without loss of generality we can adjoin to $\mathfrak{A}$ the 1-operator in $\mathcal{G}$. Then let Φ be equipped with the initial topology τ_{in} with respect to all elements of $\mathfrak{A}$, i.e., with the coarsest topology for which all the maps

$$A: \Phi \to \mathcal{G}, \quad A \in \mathfrak{A}$$

are continuous. We have at once that τ_{in} is locally convex and that τ_{in} is finer than $\tau_{\mathcal{G}}$ since 1 belongs to $\mathfrak{A}$.

Furthermore, from the general properties of the initial topology we deduce the following properties of Φ[12]:

(1) The space (Φ, τ_{in}) has a o-neighborhood base that consists of all sets of the form

$$\bigcap_{i=1}^{u} A_{\alpha_1}^{-1}(\epsilon_i K)$$

with u arbitrary but finite, and K is the unit ball in $\mathcal{G}$.

(2) (Φ, τ_{in}) is a Hausdorff locally convex space. This follows by the fact that 1 belongs to $\mathfrak{A}$ since for each $\phi \in \Phi$ there is an A, namely 1, such that $A\phi \neq 0$. Then Φ can be identified with a subspace of the topological product space $\prod_{\alpha \in A} \mathcal{G}_\alpha$, where A is an index set which has the cardinality of $\mathfrak{A}$.

(3) The topology τ_{in} on Φ is generated by the following

family of seminorms

$$\phi \to \|A\phi\| \text{ with } A \in \mathfrak{A} \text{ and } \phi \in \Phi.$$

(4) Each map $A: \Phi \to \mathcal{G}$ with $A \in \mathfrak{A}$ is also continuous, considered as a map from Φ into Φ. Due to the initial property[12] a map $A_\alpha \in \mathfrak{A}$

$$A_\alpha: (\Phi, \tau_{\rm in}) \to (\Phi, \tau_{\rm in})$$

is continuous if and only if the maps

$$A_\beta A_\alpha: (\Phi, \tau_{\rm in}) \to (\Phi, \tau_{\rm in}) \to \mathcal{G}$$

are continuous for all $\beta \in A$. Since $\mathfrak{A}$ is assumed to be an algebra, the latter condition is obviously fulfilled.

Next we shall deal with completeness properties of $(\Phi, \tau_{\rm in})$. We assumed at the beginning that every $A \in \mathfrak{A}$ admits a closure $\bar{A}$ in $\mathcal{G}$. In some cases the algebra $\mathfrak{A}$ is generated by a finite set $\mathcal{O}$ of symmetric or even essentially self-adjoint operators on a Hilbert space $\mathcal{G}$ such that all elements of $\mathcal{O}$ have a common invariant dense domain D. We consider the free algebra generated by $\mathcal{O}$ but without taking into account the multiplication by scalars. For the topological properties of Φ only the elements of $\mathfrak{A}$ really matter. Finally one can take into account the multiplication by scalars as well without altering the topology on Φ. Since the closure $\bar{A}$ of an operator A in $\mathcal{G}$ is given by A^{**} we see that every element of $\mathfrak{A}$ admits a closure $\bar{A}$. The maximal invariant domain of definition is given by

$$\Phi = \bigcap_{A \in \mathfrak{A}} D(\bar{A}) \text{ with } D(\bar{A}) \text{ as the domain of } \bar{A}.$$

Then the algebra generated by $\mathcal{O}$ is given by $\mathfrak{A}^\circ = \mathfrak{A}/\Phi$.[3] Since $\Phi \supset D$, Φ is also dense in $\mathcal{G}$. The invariance of Φ under all elements of $\mathfrak{A}^\circ$ follows from the semigroup property of $\mathfrak{A}^\circ$.

Proposition 2: $\Phi = \cap_{A \in \mathfrak{A}} D(\bar{A})$ equipped with the initial topology with respect to the elements of $\mathfrak{A}^\circ$ is a complete locally convex space.

Proof: Let $\mathcal{F}$ be a Cauchy filter in Φ. We know that $\mathcal{F}$ is a Cauchy filter in Φ if and only if $A_\alpha \mathcal{F}$ is a Cauchy filter in $\mathcal{G}$ for every $A_\alpha \in \mathfrak{A}^\circ$. Since $1 \in \mathfrak{A}^\circ$, $\mathcal{F}$ is also a Cauchy filter in $\mathcal{G}$.

By the completeness of $\mathcal{G}$, $\mathcal{F}$ converges to an element $x \in \mathcal{G}$. Therefore every Cauchy filter $A_\alpha \mathcal{F}$ converges to the element $A_\alpha x$ since every $A_\alpha \in \mathfrak{A}^\circ$ is by assumption closable. Hence we have $x \in \Phi$. Now we want to show that $\mathcal{F}$ also converges to x in the topology $\tau_{\rm in}$ on Φ. Let

$$\mathcal{U} = \bigcap_{i=1}^{u} A_{\alpha_i}^{-1}(\epsilon_i K)$$

be a o-neighborhood in Φ. For each $i = 1, 2, \ldots, u$ there exists a $F_i \in \mathcal{F}$ such that

$$A_{\alpha_i} F_i \subset A_{\alpha_i} x + \epsilon_i K.$$

Then we have

$$F = \bigcap_{i=1}^{u} F_i \in \mathcal{F}$$

and $F \subset x + \mathcal{U}$, i.e., the Cauchy filter $\mathcal{F}$ converges to x in the topology in Φ. QED

We denote now by $\Phi^\times$ the space of all continuous antilinear forms over Φ. (In this case the embedding of $\mathcal{G}$ into $\Phi^\times$ is a linear map.) The next proposition deals with the general form of elements of $\Phi^\times$. First of all we see that every expression of the form $g \cdot A = g_A$ with $g \in \mathcal{G}$ and $A \in \mathfrak{A}$ which operates on Φ according to the equation $\langle g_A, \varphi \rangle = \overline{(A\varphi, g)}$ is a continuous antilinear form on Φ. The linearity of this form follows from the linearity of A. For the proof of the continuity of g_A let $\epsilon A^{-1} K$ with K as the unit ball in $\mathcal{G}$ be a o-neighborhood in Φ. Then we have

$$|\langle g_A, \varphi \rangle| = |(A\varphi, g)| \le \|g\| \, \|A\varphi\| \langle \epsilon \|g\|$$

for all $\varphi \in \epsilon A^{-1} K$.

The same assertion is true for all finite linear combinations of antilinear forms g_A. We want to show that every element of $\Phi^\times$ can be represented in the form of a linear combination of continuous antilinear forms g_A defined above.

Proposition 3: Let $(\Phi, \tau_{\rm in})$ be the Hausdorff locally convex space endowed with the initial topology with respect to $\mathfrak{A}$. Then each element $\varphi^\times \in \Phi^\times$ can be represented in the form

$$\varphi^\times = \sum_{k=1}^{u(\varphi^\times)} g_{\alpha_K} \cdot A_{\alpha_K}$$

which operates on Φ in the following way:

$$\langle \varphi^\times, \varphi \rangle = \sum_{k=1}^{u(\varphi^\times)} \overline{(A_{\alpha_K} \varphi, g_{\alpha_K})}$$

for every $\varphi \in \Phi$.

Proof: Since the topology $\tau_{\rm in}$ on Φ is separating the space Φ is topologically isomorphic to a subspace of the topological product space

$$\prod_{\alpha \in A} \mathcal{G}_\alpha.^{12}$$

Again, we denote by A an index set which has the cardinality of the algebra $\mathfrak{A}$. The injective map indicated above

$$i: \Phi \to \prod_{\alpha \in A} \mathcal{G}_\alpha$$

is defined by

$$i(\varphi) = \{A_\alpha \varphi\}_{A_\alpha \in \mathfrak{A}}.$$

Now the dual space of the topological product $\prod_{\alpha \in A} \mathcal{G}_\alpha$ can be identified with the locally convex direct sum of the duals of the spaces $\mathcal{G}_\alpha$, i.e., we have

$$\left(\prod_{\alpha \in A} \mathcal{G}_\alpha\right)' = \sum_{\alpha \in A} \oplus \mathcal{G}'_\alpha.$$

Since there exists a conjugate norm-isomorphism from each $\mathcal{G}'_\alpha$ to $\mathcal{G}_\alpha$ there exists a conjugate canonical identification of

$$\left(\prod_{\alpha \in A} \mathcal{G}_\alpha\right)' \text{ with } \sum_{\alpha \in A} \oplus \mathcal{G}_\alpha.$$

Therefore one can define a linear map

$$j: \sum_{\alpha \in A} \oplus \mathcal{G}_\alpha \to \Phi^\times$$

by the equation

$$j(\{g_\alpha\}) = \sum_{\alpha \in A} A_\alpha^\times (g_\alpha),$$

where $A_\alpha^\times$ denotes the transpose of the operator

$$A_\alpha : \Phi \to \mathcal{G}_\alpha, \text{ i.e., } A_\alpha^\times : \mathcal{G}_\alpha \to \Phi^\times.$$

By the preceding remark we have $j = i^\times$. Indeed, for $\varphi \in \Phi$ and

$$\{g_\alpha\} \in \sum_{\alpha \in A} \oplus \mathcal{G}_\alpha$$

it follows

$$\langle i(\varphi), \{g_\alpha\}\rangle = \sum_{\alpha\in A} (A_\alpha \varphi, g_\alpha)_{\mathcal{G}_\alpha} = \sum_{\alpha\in A} \langle \varphi, A_\alpha^\times g\rangle$$
$$= \langle \varphi, j(\{g_\alpha\})\rangle.$$

Now i is an injective map from Φ into $\Pi_{\alpha\in A}\mathcal{G}_\alpha$ which is continuous in both directions. By means of general properties of continuous mappings it follows that i has the same properties for the topologies $\sigma(\Phi, \Phi^\times)$ and

$$\sigma\left(\prod_{\alpha\in A}\mathcal{G}_\alpha, \sum_{\alpha\in A}\oplus\mathcal{G}_\alpha\right)$$

and finally we arrive at the conclusion that the map j is surjective. Hence for each $\varphi^\times \in \Phi^\times$ there exists a sequence $\{g_\alpha\} \in \{\mathcal{G}_\alpha\}$ such that

$$\varphi^\times = j(\{g_\alpha\}) = \sum_{\alpha\in A} A_\alpha^\times (g_\alpha\}$$

holds, where $g_\alpha = 0$ except for a finite set $\{\alpha_i\}$, $i = 1, 2, \ldots, u$, i.e., for each $\varphi^\times \in \Phi^\times$ there exist finitely many vectors $g_{\alpha_i} \in \mathcal{G}_{\alpha_i}$, $i = 1, \ldots, u(\varphi^\times)$ and finitely many $A_{\alpha_i} \in \mathfrak{A}$, $i = 1, 2, \ldots, u(\varphi^\times)$ such that

$$\varphi^\times = \sum_{i=1}^{u(\varphi^\times)} A_{\alpha_i}^\times (g_{\alpha_i})$$

holds. The functional $\varphi^\times$ on Φ is defined by the equation

$$\langle \varphi^\times, \varphi\rangle = \sum_{i=1}^{u(\varphi^\times)} \langle \varphi, A_{\alpha_i}^\times \quad {}_{\alpha_i}\rangle \text{ for every } \varphi \in \Phi. \qquad \text{QED}$$

In what follows we shall denote by I the canonical injection of Φ into $\mathcal{G}$. Then the transpose $I^\times : \mathcal{G} \to \Phi^\times$ is determined by the relation $(f, I(\varphi))_{\mathcal{G}} = \langle I^\times(f), \varphi\rangle$ for every $\varphi \in \Phi$ and $f \in \mathcal{G}$. We shall always assume that $\Phi^\times$ is equipped with the strong topology with respect to the dual system $\langle \Phi^\times, \Phi\rangle$

Proposition 4: $I^\times$ is an injective map. $I^\times(\mathcal{G})$ is dense in $\Phi^\times$ with respect to any topology compatible with the dual system $\langle \Phi^\times, \Phi\rangle$.

Proof: I is an injective map which maps Φ onto a dense subspace of $\mathcal{G}$. Then $I^\times$ is an injective map. By the same reasoning it follows that $I^\times(\mathcal{G})$ is dense in $\Phi^\times$ with respect to $\sigma(\Phi^\times, \Phi)$ and hence in any topology compatible with the dual system $\langle \Phi^\times, \Phi\rangle$. QED

Proposition 5: Let $\mathfrak{A}$ be an algebra of operators which is generated by a finite set of essentially self-adjoint operators in $\mathcal{G}$ with a common dense invariant domain. The Φ is a complete metrizable, reflexive l.c. space, i.e., a reflexive Fréchet space.

Proof: Under the conditions assumed in the proposition, Φ is complete by Proposition 2. Since $\mathfrak{A}$ is countable Φ has a countable o-neighborhood base whence Φ is metrizable. Furthermore since $\mathcal{G}$ is a reflexive l.c. space it follows that $\Pi_{\alpha\in A}\mathcal{G}_\alpha$ is a reflexive space. But a closed subspace of a reflexive space is also a reflexive space whence Φ identified with a closed subspace of $\Pi_{\alpha\in A}\mathcal{G}_\alpha$ is a reflexive space. QED

Corollary: The antidual space $\Phi^\times$ endowed with the strong topology is a barrelled and complete l.c. space and the strong topology $\beta(\Phi^\times, \Phi)$ is compatible with the duality of $\langle \Phi^\times, \Phi\rangle$. All these assertions are immediate consequences of the reflexivity, metrizability and completness of the space Φ.

Proposition 6: The map $\gamma = I^\times \cdot I : \Phi \to \Phi^\times$ is a continuous injection. $\gamma\Phi$ is dense in $\Phi^\times$ for any topology compatible with the dual system $\langle \Phi^\times, \Phi\rangle$.

Proof: The first assertion is a trivial consequence of Proposition 4.

The second assertion follows from the relation

$$\overline{I^\times(I(\Phi))} \supset I^\times(\overline{I(\Phi)}).$$

Indeed $I^{\times\,-1}\overline{(I^\times(I(\Phi)))}$ is closed and contains $I(\Phi)$ since $I^\times$ is injective. Therefore we have

$$I^{\times\,-1}\overline{(I^\times(I(\Phi)))} \supset \overline{I(\Phi)} \text{ or } \overline{I^\times(I(\Phi))} \supset I^\times\overline{(I(\Phi))}.$$

Since $\overline{I(\Phi)} = \mathcal{G}$ and $I^\times(\mathcal{G})$ is dense in $\Phi^\times$ for any topology compatible with the dual system $\langle \Phi^\times, \Phi\rangle$ we get $\overline{I^\times(I(\Phi))} = \Phi^\times$ whence $I^\times(I(\Phi))$ is dense in $\mathcal{G}$. QED

III. RIGGED HILBERT SPACES

In this section we shall specify which conditions may be imposed on $\mathfrak{A}$ in order that Φ becomes a nuclear l.c. space in the topology τ_{in} defined in II. or that $I: \Phi \to \mathcal{G}$ becomes a nuclear map with respect to τ_{in}. Let now Φ, $\mathcal{G}$, and $\Phi^\times$ be introduced as in II.

Definition 1: Let $\Phi \subset \mathcal{G} \subset \Phi^\times$ be a space triplet where $\Phi \subset \mathcal{G}$ and $\mathcal{G} \subset \Phi^\times$ are defined via the continuous injection $I: \Phi \to \mathcal{G}$ and the conjugate transpose of I, respectively. This space triplet is called a rigged Hilbert space or sometimes a Gel'fand triplet provided the injection I is a nuclear map.

In the first step we shall deal with the case of "elementary systems." In the second step we shall settle the problem of "composite systems." Throughout this section we shall assume situation described in II. First we shall state a result which is due to Roberts.[3]

Proposition 7: Let Φ be equipped with the initial topology with respect to $\mathfrak{A}$. Then the following condition is necessary and sufficient for the nuclearity of the l.c. space Φ: There exists a self-adjoint operator B in $\mathfrak{A}$ whose inverse $B^{-1} = H$ exists and is a nuclear operator in $\mathcal{G}$, i.e., H is an element of the trace-class.

Proposition 8: The canonical injection $I: \Phi \to \mathcal{G}$ (which is continuous) is a nuclear map if and only if $\mathfrak{A}$ contains a self-adjoint operator B whose inverse $B^{-1} = H$ exists and is a nuclear operator in $\mathcal{G}$.

Proof: The sufficiency follows immediately from proposition 7 and the general properties of nuclear spaces:[11]

The condition is necessary:

Let $I: \Phi \to \mathcal{G}$ be nuclear. By a general theorem[11] a linear map from the l.c. space Φ into the l.c. space $\mathcal{G}$ is nuclear if there exists a circled convex o-neighborhood $\mathcal{U}$ in Φ such that $I(\mathcal{U}) \subset B$ where B is a bounded subset of $\mathcal{G}$ for which $\mathcal{G}_B$ is complete and the induced map $\bar{I}_0$: $\tilde{\Phi}_u \to \mathcal{G}_B$ (for the definition of the spaces $\tilde{\Phi}_u$, $\mathcal{G}_B$ we refer to Ref. 11, p. 97) is nuclear. Without loss of generality we can assume $B \subset K$, where K is the closed unit ball in $\mathcal{G}$. It is easy to see that the map $\mathcal{G}_B \to \mathcal{G}_K = \mathcal{G}$ is continuous. Thus the map $T: \tilde{\Phi}_u \to \mathcal{G}$ is nuclear. But this case can be disposed of by repeating the second part of the proof of the preceding proposition given by Roberts.[3] QED

At first sight one could conjecture that the requirement of the nuclearity of the l.c. space Φ is too strong. However, Proposition 8 shows that the condition imposed on $\mathfrak{A}$, namely, that $\mathfrak{A}$ contains a self-adjoint operator whose inverse $B^{-1} = H$ is a nuclear operator in $\mathcal{G}$ cannot be weakened if the canonical injection $I: \Phi \to \mathcal{G}$ is to be nuclear.

In the next step we shall tackle the problem how to deal with composite systems in physical applications. Obviously one might apply the same method as for elementary systems. What we are mainly interested in is the question whether there is a similar procedure for constructing the rigged Hilbert space for a composite system as it is usually applied to the Hilbert space for a composite system, i.e., by taking the tensor product of the Hilbert spaces of the elementary systems.

In what follows we shall always assume that the algebras $\mathfrak{A}_i$ of the elementary systems have the properties indicated in Sec. II. We shall also assume that each $\mathfrak{A}_i$ contains a self-adjoint operator B_i whose inverse B_i^{-1} exists and is a nuclear operator in $\mathcal{G}_i$.

First we sketch the procedure for setting up the Hilbert space formalism for a composite system. For the sake of simplicity we restrict our discussion to the case of two elementary systems. Let $\mathcal{G}_1$, $\mathcal{G}_2$ and $\mathfrak{A}_1 \mathfrak{A}_2$ be the Hilbert spaces and the algebras of observables of the two systems, respectively. Then the Hilbert space of the composed system is built up as follows: First one takes the tensor product $\mathcal{G}_1 \otimes \mathcal{G}_2$ consisting of all elements of the form

$$f = \sum_{i=1}^{u} f_i^{(1)} \otimes f_i^{(2)} \quad \text{with } f_i^{(1)} \in \mathcal{G}_1 \text{ and } f_i^{(2)} \in \mathcal{G}_2.$$

The scalar product in $\mathcal{G}_1 \otimes \mathcal{G}_2$ is defined as follows:

$$(f, g) = \sum (f_i^{(1)}, g_j^{(1)})_1 (f_i^{(2)}, g_j^{(2)}),$$

with

$$f = \sum_{i=1}^{u} f_i^{(1)} \otimes f_i^{(2)} \quad \text{and} \quad g = \sum_{j=1}^{u} g_j^{(1)} \otimes g_j^{(2)}.$$

If the spaces $\mathcal{G}_i$, $i = 1, 2$, are infinite dimensional, $\mathcal{G}_1 \otimes \mathcal{G}_2$ is not necessarily complete. Usually one takes $\mathcal{G}_1 \tilde{\otimes} \mathcal{G}_2$, the completion of $\mathcal{G}_1 \otimes \mathcal{G}_2$ with respect to the norm which is given by the scalar product defined above.

Now the algebra of "observables" operating on $\mathcal{G}_1 \tilde{\otimes} \mathcal{G}_2$ is given by the algebra $\mathfrak{A}$ which is generated by the set $\mathfrak{A}_1 \otimes \mathfrak{A}_2$. An operator $A_1 \otimes A_2 \in \mathfrak{A}_1 \otimes \mathfrak{A}_2$ is defined on $\mathcal{G}_1 \tilde{\otimes} \mathcal{G}_2$ as follows:

$$(A_1 \otimes A_2)(f^{(1)} \otimes f^{(2)}) = A_1 f^{(1)} \otimes A_2 f^{(2)}.$$

The problem we are going to investigate now can be stated as follows:

(I) Does there exist a tensor product $\Phi_1 \otimes \Phi_2$ with Φ_1 and Φ_2 as domain of definition for $\mathfrak{A}_1$ and $\mathfrak{A}_2$, respectively, such that $\Phi_1 \otimes \Phi_2$ is dense in $\mathcal{G}_1 \tilde{\otimes} \mathcal{G}_2$ and the embedding of $\Phi_1 \otimes \Phi_2$ into $\mathcal{G}_1 \tilde{\otimes} \mathcal{G}_2$ is continuous?

(II) Is $\Phi_1 \otimes \Phi_2$ a nuclear l.c. space?

(III) Are the operators in the algebra $\mathfrak{A}$, generated by the set $\mathfrak{A}_1 \otimes \mathfrak{A}_2$ continuous mappings from $\Phi_1 \otimes \Phi_2$ into itself?

Now let $\Phi = \Phi_1 \otimes_\pi \Phi_2$ be the tensor product of the l.c. spaces Φ_1 and Φ_2 endowed with the Π-topology (c.f. Ref. 13). The Π-topology on $\Phi = \Phi_1 \otimes \Phi_2$ is generated by the following family of seminorms:

$$\Pi(A_\alpha^{(1)}, A_\beta^{(2)})^{(\phi)} = \inf \sum_i \| A_\alpha^{(1)} \phi_i^{(1)} \|_1 \| A_\beta^{(2)} \phi_i^{(2)} \|_2 \quad ,$$

where $A_\alpha^{(1)} \in \mathfrak{A}_1$, $A_\beta^{(2)} \in \mathfrak{A}_2$ and the infimum is taken over all possible representations of the element ϕ in the form

$$\phi = \sum_{i=1}^{u} \phi_i^{(1)} \otimes \phi_i^{(2)}.$$

One can ask whether this Π-topology on $\Phi_1 \otimes \Phi_2$ is comparable with τ_{in}, the initial topology on $\Phi_1 \otimes \Phi_2$ with respect to the algebra of operators $\mathfrak{A}$ introduced above. τ_{in} is generated by the following family of seminorms:

$$\phi \to \| (A_\alpha^{(1)} \otimes A_\beta^{(2)}) \phi \| = \| \sum_{i=1}^{u} A_\alpha^{(1)} \phi_i^{(1)} \otimes A_\beta^{(2)} \phi_i^{(2)} \|,$$

with

$$A_\alpha^{(1)} \in \mathfrak{A}_1, \quad A_\beta^{(2)} \in \mathfrak{A}_2, \quad \phi = \sum_{i=1}^{u} \phi_i^{(1)} \otimes \phi_i^{(2)}.$$

Obviously one has

$$\| (A_\alpha^{(1)} \otimes A_\beta^{(2)}) \phi \| \le \Pi(A_\alpha^{(1)}, A_\beta^{(2)})^{(\phi)} \quad \text{for every } \phi \in \Phi_1 \otimes \Phi_2,$$

whence the Π-topology is finer than the topology τ_{in}. Now we shall give an answer to the three problems (I), (II), (III) stated above:

Proposition 9: $\Phi_1 \otimes \Phi_2$ is dense in $\mathcal{G}_1 \otimes \mathcal{G}_2$. The canonical injection of $\Phi \otimes_\Pi \Phi_2$ into $\mathcal{G}_1 \otimes \mathcal{G}_2$ is continuous; therefore the image of $\Phi_1 \tilde{\otimes}_\Pi \Phi_2$ under the extended canonical injection $\tilde{I}: \Phi_1 \tilde{\otimes}_\Pi \Phi_2 \to \mathcal{G}_1 \tilde{\otimes} \mathcal{G}_2$ is dense in $\mathcal{G}_1 \tilde{\otimes} \mathcal{G}_2$.

Proof: In order to prove the first assertion let

$$h = \sum_{i=1}^{u} h_i^{(1)} \otimes h_i^{(2)}$$

be an arbitrary elements of $\mathcal{G}_1 \otimes \mathcal{G}_2$ and let

$$\phi = \sum_{i=1}^{u} \phi_i^{(1)} \otimes \phi_i^{(2)}$$

by any element of $\Phi_1 \otimes \Phi_2$. Then we have

$$\| h - \phi \| = \| \sum_{i=1}^{u} (h_i^{(1)} \otimes h_i^{(2)} - \phi_i^{(1)} \otimes \phi_i^{(2)}) \|$$

$$\leq \sum_{i=1}^{u} \|h_i^{(1)} \otimes h_i^{(2)} - h_i^{(1)} \otimes \phi_i^{(2)}\| + \sum_{i=1}^{u} \|h_i^{(1)} \otimes \phi_i^{(2)} - \phi_i^{(1)} \otimes \phi_i^{(2)}\|$$
$$= \sum_{i=1}^{u} \{\|h_i^{(1)}\| \, \|h_i^{(2)} - \phi_i^{(2)}\| + \|\phi_i^{(2)}\|_2 \|h_i^{(1)} - \phi_i^{(1)}\|\}.$$

Since Φ_1 and Φ_2 are dense in $\mathcal{G}_1$ and $\mathcal{G}_2$, respectively, for a given $\epsilon > 0$ and $h_i^{(K)} \in \mathcal{G}_K$, $K=1, 2$ there exist $\phi_i^{(K)} \in \Phi_K$, $K=1, 2$ such that

$$\|h_i^{(K)} - \phi_i^{(K)}\| < \epsilon.$$

Hence we get

$$\|\phi - h\| \leq \sum_{i=1}^{u} \{\|h_i^{(1)}\| \cdot \epsilon + (\|h_i^{(2)}\| + \epsilon)\epsilon\} = u\epsilon^2 + M \cdot \epsilon = \epsilon',$$

with

$$M = \sum_{i=1}^{u} \{\|h_i^{(1)}\| + \|h_i^{(2)}\|\}.$$

Thus for every $h \in \mathcal{G}_1 \otimes \mathcal{G}_2$ and $\epsilon' > 0$ there exists a $\phi \in \Phi_1 \otimes \Phi_2$ such that $\|\phi - h\| < \epsilon'$ holds [We choose $\epsilon = (M^2 + 4u\epsilon')^{1/2}/2u$.] For the proof of the second assertion it suffices to show that the π-topology on $\Phi_1 \otimes \Phi_2$ is finer than the topology induced by $\mathcal{G}_1 \otimes \mathcal{G}_2$ on $\Phi_1 \otimes \Phi_2$. Now for each

$$\phi = \sum_{i=1}^{u} \phi_i^{(1)} \otimes \phi_i^{(2)} \in \Phi_1 \otimes \Phi_2$$

we have

$$\|\phi\|^2 = \sum_{i,j=1} (\phi_i^{(1)}, \phi_j^{(1)})(\phi_i^{(2)}, \phi_j^{(2)}) \leq \sum \|\phi_i^{(1)}\| \|\phi_j^{(1)}\| \, \|\phi_i^{(2)}\| \, \|\phi_j^{(2)}\|$$
$$\leq \left\{ \sum_{i=1}^{u} \|\phi_i^{(1)}\| \, \|\phi_i^{(2)}\| \right\}^2.$$

Since the canonical injection $I_K: \Phi_K \to \mathcal{G}_K$, $K=1, 2$ are continuous there exist positive numbers C_K and operators $A^{(K)} \in \mathfrak{A}_K$, $K=1, 2$ such that

$$\|\phi^{(K)}\| \leq C_K \|A^{(K)} \phi^{(K)}\| \quad \text{for every } \phi^{(K)} \in \Phi_K \text{ holds.}$$

Thus we have

$$\|\phi\| \leq C_i \cdot C_2 \sum_i \|A^{(1)} \phi_i^{(1)}\| \, \|A^{(2)} \phi_i^{(2)}\|.$$

Since this relation holds for every representation of the element $\phi \in \Phi_1 \otimes \Phi_2$, we obtain

$$\|\phi\| \leq C_i C_2 \Pi_{(A^{(1)}, A^{(2)})}(\phi) \quad \text{for every } \phi \in \Phi_1 \otimes \Phi_2.$$

Hence the injection $I: \Phi_1 \otimes_\Pi \Phi_2 \to \mathcal{G}_1 \otimes \mathcal{G}_2$ is continuous. This mapping can be extended continuously to the map

$$\tilde{I}: \Phi_1 \tilde{\otimes}_\Pi \Phi_2 \to \mathcal{G}_1 \tilde{\otimes} \mathcal{G}_2.$$

It follows that $\tilde{I}(\Phi_1 \tilde{\otimes}_\Pi \Phi_2)$ is dense in $\mathcal{G}_1 \tilde{\otimes} \mathcal{G}_2$. QED

Proposition 10: The elements of the algebra $\mathfrak{A}$ generated by $\mathfrak{A}_1 \otimes \mathfrak{A}_2$ are continuous mappings from $\Phi_1 \tilde{\otimes}_\Pi \Phi_2$ into itself.

Proof: It is immediate that every element of $\mathfrak{A}$ maps $\Phi_1 \otimes \Phi_2$ into itself.

Furthermore $A^{(1)} \otimes A^{(2)}$ maps $\Phi_1 \otimes \Phi_2$ continuously into itself if for each continuous seminorm $\Pi_{(B^{(1)}, B^{(2)})}(\cdot)$ on $\Phi_1 \otimes_\Pi \Phi_2$ there exist another continuous seminorm $\Pi_{(D^{(1)}, D^{(2)})}(\cdot)$ and a positive number M such that the inequality

$$\Pi_{(B^{(1)}, B^{(2)})}((A^{(1)} \otimes A^{(2)})(\phi)) \leq M \Pi_{(D^{(1)}, D^{(2)})}(\phi)$$

holds for every $\phi \in \Phi_1 \otimes \Phi_2$.

Let

$$\phi = \sum_{i=1}^{u} \phi_i^{(1)} \otimes \phi_i^{(2)}$$

be an arbitrary element of $\Phi_1 \otimes \Phi_2$. For if $A^{(1)} \otimes A^{(2)} \in \mathfrak{A}_1 \times \mathfrak{A}_2$ we have

$$\Pi_{(B^{(1)}, B^{(2)})}((A^{(1)} \otimes A^{(2)})(\phi))$$
$$= \inf \left\{ \sum_{i=1}^{u} \|B^{(1)} A^{(1)} \phi_i^{(1)}\| \, \|B^{(2)} A^{(2)} \phi_i^{(2)}\| \right.$$
$$= \Pi_{(B^{(1)} A^{(1)}, B^{(2)} A^{(2)})}(\phi) \quad \text{for every } \phi \in \Phi_1 \otimes \Phi_2.$$

By the saturation property of the seminorms on $\Phi_1 \otimes_\Pi \Phi_2$, we see that also every element of $\mathfrak{A}$ is a continuous map from $\Phi_1 \otimes_\Pi \Phi_2$ into itself. Finally by taking the continuous extension of every element in $\mathfrak{A}$ we see that every element of $\mathfrak{A}$ is also a continuous map from $\Phi_1 \tilde{\otimes}_\Pi \Phi_2$ into itself. QED

Thus we have succeeded in showing that the problems (I) and (III) can actually be settled. For if Φ_1, Φ_2 are nuclear l.c. spaces it follows (c.f. Refs. 11, 13) that also $\Phi_1 \tilde{\otimes} \Phi_2$ is a nuclear l.c. space.

For nuclear Fréchet spaces Φ_1 and Φ_2 it follows (c.f. Ref. 11, p. 175) that the strong antidual space, i.e. $(\Phi_1 \tilde{\otimes}_\Pi \Phi_2)^\times_\beta$ can be identified with the space

$$\Phi_1^\times \tilde{\otimes}_\Pi \Phi_2^\times.$$

One might ask at this point whether the Π-topology and the topology τ_{1n} coincide in case both Φ_1 and Φ_2 are nuclear l.c. spaces. It is known that in this case the ϵ-topology and the Π-topology on $\Phi_1 \otimes \Phi_2$ coincide. Unfortunately, we were not able to show that the ϵ topology (c.f. Ref. 13) is coarser than the topology τ_{1n} which would imply the equivalence of both topologies. It is perhaps worthwhile to note that the method described in Propositions 9, 10 and the preceding remarks applies to a far more general situation. Let Φ_i, $i=1, 2$ be nuclear l.c. spaces which are domains of definition for sets of operators $\mathfrak{A}_i$, $i=1,2$, respectively, such that all elements of $\mathfrak{A}_i$, $i=1, 2$ are continuous mappings from Φ_i, $i=1, 2$, into itself.

Φ_1 and Φ_2 are assumed to be continuously embedded into Hilbert spaces $\mathcal{G}_1$ and $\mathcal{G}_2$, respectively. Then one can follow the same lines as in Propositions 9 and 10 to show that problems (I), (II), and (III) can be settled for this case as well.

IV. EIGENOPERATORS OF AN OPERATOR IN Φ

Let $\mathcal{L}(\Phi)$ be the set of all linear continuous operators from Φ into itself. Then $A^c \in \mathcal{L}(\Phi)$ and $(A^c\phi, \psi) = (\phi, A\psi)$ for all pairs $\phi, \psi \in \Phi$. A is called real if $A^c = A$. The set of all those $A \in \mathcal{L}(\Phi)$ that have a conjugate A^c will be denoted by $\mathcal{L}^c(\Phi)$.

Definition 2: Let $\mathcal{L}(\Phi, \Phi^\times)$ denote the set of all linear continuous mappings from Φ into $\Phi^\times$ where $\Phi^\times$ is assumed to be equipped with the strong topology $\beta(\Phi^\times, \Phi)$. Then an element $\gamma \in \mathcal{L}(\Phi, \Phi^\times)$ is called self-adjoint if $\langle \phi, \gamma\psi\rangle = \overline{\langle \psi, \gamma\phi\rangle}$ for all $\phi,\ \psi \in \Phi$. γ is called positive if $\langle \phi, \gamma\phi\rangle \geq 0$ for every $\phi \in \Phi$. We shall summarize some results mostly obtained by Foias[8] and Roberts[3]:

(1) $\gamma \in \mathcal{L}(\Phi, \Phi^\times)$ is self-adjoint if and only if $\langle \phi, \gamma\phi\rangle$ is real for every $\phi \in \Phi$.[3]

(2) If $A \in \mathcal{L}^c(\Phi)$ then IAI^{-1} (I denotes the continuous injection $\Phi \to \mathcal{G}$) defines a linear operator in $\mathcal{G}$. Obviously I^{-1} is applied to the subset $I(\Phi) \subset \mathcal{G}$ only. The operator IAI^{-1} is densely defined in $\mathcal{G}$ and has an adjoint operator $(IAI^{-1})^* \supset A^c I^{-1}$ which is also densely defined in $\mathcal{G}$. Therefore IAI^{-1} admits a closure in $\mathcal{G}$ which will be denoted by $\overline{A}$ (recall an operator A in $\mathcal{G}$ admits a closure A in $\mathcal{G}$ if and only if A^* is densely defined in $\mathcal{G}$). Let now $A \in \mathcal{L}^c(\Phi)$ be an operator that originates from a linear operator on $\mathcal{G}$ with Φ as domain. At the beginning of the last chapter elements $\phi^\times \in \Phi^\times$ were called eigenforms to the eigenvalue λ of an operator A if the following relation $\langle A\phi, \psi^\times\rangle = \langle \phi, A^\times\psi^\times\rangle = \lambda\langle \phi, \psi^\times\rangle$ for every $\phi \in \Phi$ holds. Within the structure of a rigged Hilbert space one can identify the transpose $A^\times$ of the operator A regarded as a continuous map from $\Phi^\times$ into itself with the continuous extension of A^c on Φ. Indeed, let $\gamma = I^\times \cdot I$. Then as it has been shown in II γ is an injective map from Φ onto a dense subspace of $\Phi^\times$. For all $\phi,\ \psi \in \Phi$ we have

$$
\begin{aligned}
(IA\phi, I\psi) &= \langle \phi, A^\times\gamma\psi\rangle = (A\phi, \psi) = (\phi, A^c\psi) \\
&= (I\phi, IA^c\psi) = \langle \phi, \gamma A^c\psi\rangle .
\end{aligned}
$$

Thus the equation $A^\times\gamma\psi = \gamma A^c\psi$ is valid for every $\psi \in \Phi$ and therefore $A^\times\gamma = \gamma A^c$ or $A^{c\times}\gamma = \gamma A$. By the last relation we can identify $A^{c\times}$ with the continuous extension of A on Φ to an operator on $\Phi^\times$.

Definition 3: A positive element γ of $\mathcal{L}(\Phi, \Phi^*)$ is called an eigenoperator of A belonging to the eigenvalue λ provided γ satisfies the relation

$$A^{c\times}\gamma = \gamma A = \lambda\gamma.$$

(3) Then it is immediate that we also have

$$A^\times\gamma = \gamma A^c = \overline{\lambda}\gamma.$$

If γ is an eigenoperator of A to the eigenvalue λ and if $\gamma\phi \neq 0$ then $\gamma\phi$ is an eigenform of A and A^c both to the eigenvalues λ and $\overline{\lambda}$.

Definition 4: An integral decomposition of Φ is a triplet $\{\gamma(z), Z, \mu\}$ which has the following properties:

(I) $\gamma(z) \in \mathcal{L}(\Phi, \Phi^\times)$ and $\gamma(z)$ is positive $(z \in Z)$,

(II) μ is a positive regular measure on the Borel sets of a locally compact Hausdorff space Z,

(III) $z \to \langle \phi, \gamma(z)\psi\rangle$ is μ-integrable for all $\phi,\ \psi \in \Phi$,

(IV) $(\phi, \psi) = \int_Z \langle \phi, \gamma(z)\psi\rangle\, d\mu(z)$ for all $\phi,\ \psi \in \Phi$.

The following proposition is due to Foias,[8] and Roberts.[3]

Proposition 11: Let $\{\gamma(z),\ z, \mu\}$ be an integral decomposition of Φ. Then there exists a unique operator-valued measure $\hat{m}$ on Z with values in $\mathcal{L}(\Phi, \Phi^\times)$ which is σ-additive with respect to $\sigma(\mathcal{L}(\Phi, \Phi^\times), \Phi \times \Phi$; $\hat{m}$ satisfies $\langle \phi, \hat{m}(\delta)\psi\rangle = \int_\delta \langle \phi, \gamma(z)\psi\rangle\, d\mu(z)$ for all Borel sets δ of Z and all $\phi,\ \psi \in \Phi$.

Definition 5: An operator-valued measure B with values in $\mathcal{L}(\mathcal{G})$ ($\mathcal{G}$ being a Hilbert space) and defined on the Borel sets $\mathcal{B}$ of a locally compact Hausdorff space Z is called a semispectral measure if $B(\sigma)$ is a symmetric operator for every Borel set σ, $0 \leq B(\sigma) \leq 1$ and $(h, B(\sigma)g)$ is a measure on $\mathcal{B}$ for all $h,\ g \in \mathcal{G}$. Furthermore $B(\phi) = 0$ and $B(Z) = 1$. By a theorem of Naĭmark[14] any semispectral measure $\delta \to B(\delta)$ in $\mathcal{G}$ may be extended to a spectral measure $\delta \to E(\delta)$ in an extended Hilbert space $\tilde{\mathcal{G}} \supset \mathcal{G}$. Let P be the orthogonal projection of $\tilde{\mathcal{G}}$ onto $\mathcal{G}$; then $E(\delta)$ is the extension of $B(\delta)$ in the following sense: If the extension of $\mathcal{G}$ to $\tilde{\mathcal{G}}$ is minimal, i.e., if the set $\{E(\delta)h; h \in \mathcal{G}, \delta \in \mathcal{B}\}$ is dense in $\tilde{\mathcal{G}}$ then this extension is unique up to isomorphisms. With the aid of this definition one can derive a sharpening of the preceding proposition which is also due to Roberts[3] and Foias.[8]

Proposition 12: Let Φ be a l.c. space continuously embedded in a Hilbert space $\mathcal{G}$ and let $\Phi^\times$ be the space of all continuous antilinear forms on Φ. Let $\{\gamma(z), Z, \mu\}$ be an integral decomposition of Φ. Then there exists a unique semispectral measure $\delta \to B(\delta)$ defined on $\mathcal{G}$ satisfying the relation

$$(I\phi, B(\delta)I\psi) = \langle \phi, \hat{m}(\delta)\psi\rangle = \int_\delta \langle \phi, \gamma(z)\psi\rangle\, d\mu(z)$$

for all $\phi,\ \psi \in \Phi$ and $\delta \in \mathcal{B}$.

Definition 6: Let A be an element of $\mathcal{L}^c(\Phi)$. Then Φ is said to have an integral A-eigendecomposition if Φ has an integral decomposition with $\gamma(z)$ either equal to zero or $\gamma(z)$ an eigenoperator of A $(z \in Z)$. Φ is said to have a real integral A-eigendecomposition if it has an integral decomposition $\{\gamma(\lambda), R, \mu\}$ such that

$$A^{c\times}\gamma(\lambda) = \gamma(\lambda)A = \lambda\gamma(\lambda), \quad \lambda \in R.$$

This relation shows that the $\gamma(\lambda)$'s, $\lambda \in R$, are in a way a generalization of the Hilbert space projectors onto the eigenspaces belonging to discrete eigenvalues of $\overline{A}$ in $\mathcal{G}$.

Recall that a closed operator A on $\mathcal{G}$ is called formally normal if $D(A) \subset D(A^*)$ and $\|Ax\| = \|A^*x\|$ for every $x \in D(A)$ hold. An operator A is called subnormal if there exists an extension $\tilde{A}$ of A in an extended Hilbert space $\tilde{\mathcal{G}} \supset \mathcal{G}$ such that $\tilde{A}$ is a normal operator.

Proposition 13: Let $A \in \mathcal{L}^c(\Phi)$. If Φ has an integral A-eigendecomposition then the closure $\overline{A}$ of A is subnormal and formally normal.[3] In a rigged Hilbert space as defined in Definition 1 the converse of Proposition 13 is also true. The proof of this assertion is based on the following proposition which is due to Gårding and Maurin.[9]

Proposition 14: Let $\Phi \subset \mathcal{G}$ be a nuclear l.c. space which is continuously embedded into $\mathcal{G}$ and let

$$\mathcal{G} \to \hat{\mathcal{G}} = \int_Z \hat{\mathcal{G}}(z)\, d\mu(z)$$

be a direct integral decomposition of $\mathcal{G}$ with Z as a locally compact Hausdorff space. Then for μ-almost all $z \in Z$ there exist nuclear mappings $I(z)$ satisfying the following conditions

$$I(z): \Phi \to \hat{\mathcal{G}}(z), \quad z \in Z$$

and

$$(I\phi, h) = \int_z (I(z)\phi, \hat{h}(z))_z d\mu(z) \text{ for every } \phi \in \Phi,\ h \in \mathcal{G}.$$

Proof: We present a proof of this proposition since we will need some of the details in the following section. Since the map $I: \Phi \to I\Phi \to \hat{\Phi} \subset \hat{\mathcal{G}}$ is a nuclear map, the map I has a representation of the following form (c.f. Ref. 11, p. 99):

$$I\phi = \hat{\phi} = \sum_{k=1} \lambda_k \langle \phi, \varphi_k^\times \rangle \hat{h}_k,$$

with

$$\sum_{k=1}^{\infty} |\lambda_k| < \infty, \{\varphi_k^\times\}$$

an equicontinuous sequence in $\Phi^\times$, i.e. $\{\varphi_k^\times\} \subset V^0$, where V is a o-neighborhood in Φ and $\{\hat{h}_k\}$ a bounded sequence of vectors in $\hat{\mathcal{G}}$. We denote by V^0 the polar of V (c.f. Ref. 11, p. 125).

Then for $\phi \in V$ we have

$$\|\sum_{k=1}^{\infty} \lambda_k \langle \phi, \varphi_k^\times \rangle \hat{h}_k\| \leq \sum_{k=1}^{\infty} |\lambda_k| \|\hat{h}_k\| < \infty.$$

Without loss of generality we may assume each vector $\hat{h}_k$ to be of norm 1.

Then we derive from the last relation

$$\sum_{k=1}^{\infty} |\lambda_k| \|\hat{h}_k\|^2 < \infty.$$

Let us now consider the function $\|\hat{h}_k(z)\|_z^2$ depending on the descrete variable k and the continuous variable z.

We conclude that the iterated integral

$$\sum_{k=1}^{\infty} |\lambda| \int_z \|\hat{h}_k(z)\|_z^2 d\mu(z) < \infty$$

exists where the summation over $|\lambda_k|$ has been interpreted as an integration with respect to a discrete measure. Since $\|\hat{h}_k(z)\|_z^2 \geq 0$ for all $z \in Z$, it follows by Fubini's theorem (note that $\|\hat{h}_k(z)\|_z^2$ need not be measurable! Ref. 15, p. 204) that $\|\hat{h}_k(z)\|_z^2$ is also integrable with respect to the product measure and the iterated integral with the reversed order of integration also exists, i.e., we have

$$\int_z \sum_{k=1}^{\infty} |\lambda_k| \|\hat{h}_k(z)\|_z^2 d\mu(z) < \infty.$$

From this relation we deduce

$$\sum_{k=1}^{\infty} |\lambda_k| \|\hat{h}_k(z)\|_z^2 < \infty \text{ for all } z \text{ but } z \in N,$$

where N is a subset of Z with μ-measure zero.

Now we want to show that all the mappings

$$I(z)\phi = \hat{\phi}(z) = \sum_{k=1}^{\infty} \lambda_k \langle \phi, \varphi_k^\times \rangle \hat{h}_k(z)$$

are continuous mappings from Φ into $\hat{\mathcal{G}}(z)$ for all $z \in Z$.

These mappings are even nuclear. Indeed, if $\phi \in \Phi$ we have

$$\|I(z)\phi\|_z^2 = \|\sum_{k=1}^{\infty} \lambda_k \langle \phi, \varphi_k^\times \rangle \hat{h}_k(z)\|_z^2 \leq (P_V(\phi) \sum_{k=1}^{\infty} |\lambda_k| \|\hat{h}_k(z)\|_z)^2$$

$$= (P_V(\phi))^2 \Big(\sum_{k=1}^{\infty} |\lambda_k|^{1/2} \|\hat{h}_k(z)\|_z |\lambda_k|^{1/2}\Big)^2,$$

$$\leq (P_V(\phi))^2 \sum_{k=1}^{\infty} |\lambda_k| \sum_{k=1}^{\infty} |\lambda_k| \|\hat{h}_k(z)\|_z^2 = C(z)(P_V(\phi))^2,$$

where $C(z)$ is finite for every $z \in Z$, $z \notin N$. This proves the continuity of the mappings $I(z)$ for all $z \in Z$ but $z \in N$.

In order to prove that the mappings $I(z)$ are nuclear we shall show that

$$\sum_{k=1}^{\infty} \lambda_k \varphi_k^\times \otimes \hat{h}_k(z)$$

is an element of $\tilde{\Phi}_V^\times \otimes_\pi \hat{\mathcal{G}}(z)$ (c.f. Ref. 11, p. 99). But this follows from the relation

$$\Big(\sum_{k=1}^{\infty} |\lambda_k| \|\varphi_k^\times\| \|\hat{h}_k(z)\|_z\Big)^2 \leq \Big(\sum_{k=1}^{\infty} |\lambda_k|^{1/2} |\lambda_k|^{1/2} \|h_k(z)\|_z\Big)^2 \quad \text{(IV.1)}$$

$$\leq \sum_{k=1}^{\infty} |\lambda_k| \sum_{k=1}^{\infty} |\lambda_k| \|\hat{h}_k(z)\|_z^2 < \infty \text{ for } z \in Z,\ z \notin N.$$

In this formula $\|\varphi_k^\times\|$ denotes the norm in the space $\tilde{\Phi}_V^\times \approx \Phi_{V^0}^\times$ which is defined as follows

$$\|\varphi_k^\times\| = \sup_{\phi \in V} \{|\langle \phi, \varphi_k^\times \rangle|\}.$$

For the definition of the spaces Φ_V and Φ_{V^0} we refer to Ref. 11, p. 97. Since $\varphi_k^\times \in V^0$ we have $\|\varphi_k^\times\| \leq 1$.

Thus we conclude from relation (IV.1) that

$$\sum_{k=1}^{\infty} |\lambda_k| \|\varphi_k^\times\| \|\hat{h}_k(z)\|_z$$

is convergent so that

$$\sum_{k=1}^{\infty} \lambda_k \varphi_k^\times \otimes \hat{h}_k(z)$$

defines an element of $\Phi_{V^0}^\times \tilde{\otimes}_\pi \hat{\mathcal{G}}(z)$. Hence we have proved the nuclearity of the mappings $I(z)$ for all $z \in Z$, $z \in N$.

QED

With the aid of the last proposition one can derive the next proposition which is due to Roberts.[3]

Proposition 15: Let Φ be a nuclear, separable l.c. space which is continuously embedded into $\mathcal{G}$. Let $A \in \mathcal{L}^c(\Phi)$ and let $\bar{A}$ be the closure of IAI^{-1} in $\mathcal{G}$ which is subnormal and formally normal. Then Φ has an integral A-eigendecomposition. The eigenoperators $\gamma(t)$ are defined as $\gamma(t) = I^\times(t) \cdot I(t)$.

Corollary: An operator $A \in \mathcal{L}^c(\Phi)$ has a real integral A-eigendecomposition if and only if A is real, i.e., $A = A^c$. The next problem arising in the theory is the following: Under what conditions is the integral A-eigendecomposition unique.

Definition 7: Two integral decompositions $\{\gamma, (z), \mathbf{C}, \mu_1\}$ and $\{\gamma_2(z), \mathbf{C}, \mu_2\}$ of Φ are called equivalent if there exist a positive regular Borel measure μ and measurable functions f_1 and f_2 such that $\mu_1 = f_i \mu$, $\mu_2 = f_i \mu$ and $f_1\gamma_1(z) = f_2\gamma_2(z)$ hold with the exception of a set $N \subset \mathbf{C}$ with μ-measure zero.

Proposition 16: A real operator $A \in \mathcal{L}^c(\Phi)$ has a unique integral A-eigendecomposition up to the equivalence defined in Definition 7 if and only if $\bar{A}$ is a maximal Hermitian operator in $\mathcal{G}$.

Proof: By Proposition 12 an integral decomposition of Φ determines a unique semispectral measure. Obviously two integral A-eigendecompositions are equivalent if and only if they determine the same semispectral measure on $\mathbb{C}$. Therefore Φ has a unique integral A-eigendecomposition up to equivalence if and only if there is a unique generalized spectral measure which is attached to $\bar{A}$.

But this is exactly the case when $\bar{A}$ is maximal Hermitian. For if $\bar{A}$ is already a self-adjoint operator in $\mathcal{G}$ then the conditions of the preceding proposition are fulfilled. It follows at once that an essentially self-adjoint operator on Φ induces a unique integral A-eigendecomposition of Φ up to equivalence. QED

For an operator $A \in \mathcal{L}^c(\Phi)$ which is essentially self-adjoint on Φ we may summarize the results obtained so far: There exists an integral A-eigendecomposition of Φ, i.e.,

$$(\phi, \psi) = \int_R \langle \phi, \gamma(t)\psi\rangle d\mu(t) \quad \text{for all } \phi,\ \psi \in \Phi. \qquad \text{(IV.2)}$$

The semispectral measure $B(\delta)$, which according to Proposition 12 is defined by this integral A-eigendecomposition, is already up to isometry a (uniquely) determined spectral measure in $\mathcal{G}$ and hence unitarily equivalent to $E(\delta)$, i.e., the spectral measure which is given by the spectral representation of $\bar{A}$ in $\mathcal{G}$, i.e.,

$$(\phi, E(\sigma)\psi) = \int_\sigma \langle \phi, \gamma(t)\psi\rangle d\mu(t) = \int_\sigma (I(t)\phi, I(t)\psi)_t\, d\mu(t), \qquad \text{(IV.3)}$$

where $(\cdot,\cdot)_t$ denotes the scalar product in $\hat{\mathcal{G}}(t)$ (cf. Proposition 14).

Recall that a symmetric operator A with the domain $D(A)$ is essentially self-adjoint if the closure $\bar{A}$ in $\mathcal{G}$ is a self-adjoint operator. The most advantageous property of a symmetric operator which is essentially self-adjoint is that it admits a unique self-adjoint extension in $\mathcal{G}$, namely the closure $\bar{A}$. In particular for physical applications it is very essential to have, for an observable, i.e., an (unbounded) symmetric operator, a domain on which this operator is essentially self-adjoint (note that in many physical applications we obtain from physical considerations a priori only a symmetric operator for the mathematical representation of an observable). Otherwise one needs additional criteria in order to choose the correct, i.e., the physically important, self-adjoint extension of the symmetric operator provided self-adjoint extension of the operator exist. We shall come back to the question later when dealing with the case of nonrelativistic quantum mechanics.

We conclude this section by extending the formalism described so far to operators in $\mathcal{G}$ that have Φ as their domain of definition and are e.s.a. on Φ but do not map Φ continuously into itself.

In this case we shall assume that such an operator is at least a continuous map of Φ into a space $\tilde{\Phi}_V$ for such a o-neighborhood V in Φ that the canonical embedding

$$\tilde{\Phi}_V \to \mathcal{G} \text{ is nuclear.}$$

This assumption seems to require some further explanation. We have started with the general assumption that Φ is a nuclear l.c. space.

This entails that the canonical injection

$$I: \Phi \to \mathcal{G}$$

is a nuclear map. I is nuclear if there exists a circled convex o-neighborhoods in Φ such that $I(V) \subset F$, where F is a bounded set in $\mathcal{G}$ for which $\mathcal{G}_F$ (for the definition of the space $\mathcal{G}_F$ we refer to Ref. 11, p. 97) is complete. Without loss of generality we can assume $F = K$, where K is the closed unit ball in $\mathcal{G}$. Therefore, $\mathcal{G}_F = \mathcal{G}$. In addition the map

$$\bar{I}_0 : \tilde{\Phi}_V \to \mathcal{G} \text{ must be nuclear.}$$

Then for all circled convex o-neighborhood U with $U \subset V$ the map

$$\phi_{U,V} \circ \bar{I}_0 : \tilde{\Phi}_U \to \tilde{\Phi}_V \to \mathcal{G} \text{ is nuclear.}$$

These comments are intended to show that o-neighborhoods with the required property always exist.

The transpose $A^\times$ of $A: \Phi \to \tilde{\Phi}_V$ is continuous map from $(\tilde{\Phi}_V) \approx \Phi^\times_{V^0}$ into $\Phi^\times$. The generalized spectral decomposition has to be performed in the triple of spaces.

$$\tilde{\Phi}_V \to \mathcal{G} \to \Phi^\times_{V^0}.$$

V. THE REPRESENTATION OF THE OPERATORS $\gamma(t)$

The aim of this section is, on one hand, to derive a representation of the operators $\gamma(t)$ in terms of eigenfunctionals of the operator A and on the other hand to give a rough characterization of these eigenfunctionals by the fact that for fixed $t \in \mathrm{sp}(A)$ all eigenfunctionals occurring in the representation of $\gamma(t)$ are contained in a certain subspace of $\Phi^\times$. The results in this section have mostly already been obtained by Foiaş.[10] However, Foiaş works with a direct decomposition of the mapping γ into eigenoperators $\gamma(t)$ based on the use of vector-valued measures of bounded variation.

Using the A-eigenintegral decomposition of Roberts which is based on the Gårding–Maurin theorem we are able to rederive some results of Foiaş in a different manner and last but not least to work directly with the strong topology $\beta(\Phi^\times, \Phi)$ on $\Phi^\times$. However, we emphasize that in some parts we can follow the line of the proofs given by Foiaş.[10]

Let $A \in \mathcal{L}^c(\Phi)$ be e.s.a. on Φ and let $E(\sigma)$, $\sigma \in \mathcal{B}$, denote the spectral measure originating from the spectral decomposition of the closed operator $\bar{A}$. $\mathcal{B}$ denotes the lattice of Borel sets on the real line R. Throughout this section we shall assume Φ to be a separable nuclear l.c. space. As we know the fact that Φ is a metrizable nuclear space already entails the separability of Φ.[13]

Now we want to derive a representation of the eigenoperators $\gamma(t)$ of A that fulfill the relation

$$\langle\gamma\phi|\psi\rangle = \int_R \langle\gamma(t)\phi|\psi\rangle d\mu(t) \tag{V.1}$$

in terms of the so-called eigenfunctionals of $A^\times$.

Let $h\in\mathcal{G}$ be a fixed vector and let $\mathcal{G}_h$ be the closed subspace of $\mathcal{G}$ generated by all elements of the form $E(\sigma_i)h$ with $\sigma_i\in\mathcal{B}$. Let P_h denote the orthogonal projection of $\mathcal{G}$ onto $\mathcal{G}_h$. It follows at once that P_h commutes with every $E(\sigma)$, $\sigma\in\mathcal{B}$.

Proposition 17: There exists a function $\phi_h^\times(t)\in\Phi^\times$ which is uniquely defined on R, μ almost everywhere, such that

$$(P_hE(\sigma)I\phi, I\psi) = \int_\sigma \overline{\langle\phi_h^\times(t)|\phi\rangle}\langle\phi_h^\times(t)|\psi\rangle d\mu(t) \tag{V.2}$$

for any pair $\phi,\psi\in\Phi$ and any $\sigma\in\mathcal{B}$ holds.

Proof: First we observe that by Proposition 14

$$(P_hE(\sigma)I\phi, I\psi) = (E(\sigma)P_hI\phi, I\psi) = \int_\sigma\langle(\widehat{P_hI\phi})(t)|I(t)\psi\rangle_t d\mu(t)$$

$$= \int_\sigma\langle I^\times(t)(\widehat{P_hI\phi})(t)|\psi\rangle d\mu(t) \text{ is valid.}$$

In the usual way there is an isometric mapping from $\mathcal{G}_h$ onto L^2_μ. Hence we get

$$\int_\sigma\langle I^\times(t)(\widehat{P_hI\phi})(t)|\psi\rangle d\mu(t) = \int_\sigma u_{P_h\phi}(t)\overline{u_{P_h\psi}(t)}d\mu(t),$$

for all $\phi,\psi\in\Phi$ and $\sigma\in\mathcal{B}$.

It follows that up to a μ-null set $N(\phi,\psi)$

$$\langle I^\times(t)(\widehat{P_hI\phi})(t)|\psi\rangle = u_{P_h\phi}(t)\overline{u_{P_h\psi}(t)}, \tag{V.3}$$

Let Φ_0 be a countable dense set in Φ. Then

$$N = \bigcup_{\phi,\psi\in\Phi_0} N(\phi,\psi)$$

is again a μ-null set and (V3) is valid for all $\phi,\psi\in\Phi_0$ and all $t\in R$ except $t\in N$.

In particular (V. 3) yields

$$|u_{P_h\phi}(t)|^2 = \langle I^\times(t)(\widehat{P_hI\phi})(t)|\phi\rangle.$$

We want to show that $u_{P_h\phi}(t)$ is a continuous linear form over Φ. Now with the aid of the theorem of Beppo Levi we may derive the following inequality (cf. the proof Proposition 14).

$$\langle I^\times(t)(\widehat{P_hI\phi})(t)|\phi\rangle \le \{P_V(\phi)\}^2\Big(\sum_{i=1}^\infty|\lambda_i|\,\|(P_hh_i)(t)\|_t\Big)^2$$

$$\le \{P_V(\phi)\}^2\sum_{i=1}^\infty|\lambda_i|\sum_{i=1}^\infty|\lambda_i|\,\|\hat h_i(t)\|_t^2 = C(t)\{p_V(\phi)\}^2, \tag{V.4}$$

for all $t\in R$ but $t\in N(\phi)$. Let Φ_0 be the countable dense set in Φ defined above and set $N=\cup_{\phi\in\Phi_0}N(\phi)$. Now we know from the proof of Proposition 14 that

$$\sum_{i=1}^\infty\|\hat h_i(t)\|_t^2| < \infty \quad \text{for all } t\in R \text{ but } t\in N'.$$

Hence for all $t\in R$, $t\notin\hat N\cup N'$ we have

$$\langle I^\times(t)(\widehat{P_hI\phi})(t)|\phi\rangle \le c(t)\{p_v(\phi)\}^2,$$

for every $\phi\in\Phi_0$ which means that $\phi\to u_{P_u\phi}(t)$ is continuous linear form over Φ_0. This form can continuously be extended to all $\phi\in\Phi$ whence there exists an $\phi_h^\times(t)\in\Phi^\times$ for all $t\in R$, $t\notin\hat N\cup N'$. If we set now $\phi_h^\times(t)=0$ for all $t\in N\cup\hat N\cup N'$ we get

$$\langle\phi_h^\times(t)|\phi\rangle = \overline{u_{P_h\phi}(t)}, \quad \text{for all } \phi\in\Phi.$$

Moreover, we have $\langle\phi_h^\times(t)|\phi\rangle\in L^2_\mu$ for every $\phi\in\Phi$. The unicity of $\phi_h^\times(t)$ is obtained by checking (V.2) with all $\sigma\in\mathcal{B}$. QED

Proposition 18: $I^\times(t)(\widehat{P_hI\phi})(t)$ defines μ-almost everywhere a continuous positive mapping $\gamma_h(t)\in\mathcal{L}(\Phi,\Phi^\times)$.

Proof: $\gamma_h(t)$ is a mapping since, up to a μ-null set N, the function $I^\times(t)(\widehat{P_hI\phi})(t)$ is uniquely defined. In order to prove the continuity of $\gamma_h(t)$ we observe that the strong topology $\beta(\Phi^\times,\Phi)$ on $\Phi^\times$ is determined by the seminorms

$$P_B(\phi^\times) = \sup\{|\langle\phi^\times|\phi\rangle|,\ \phi\in B\}, \quad \text{where } B\in\hat{\mathcal{B}}$$

($\hat{\mathcal{B}}$ is the class of all bounded sets in Φ). Now we have, with

$$c_h(t) = \sum_{i=1}^\infty|\lambda_i|\,\|(\widehat{P_hh_i})(t)\|_t^2 \text{ and } u=\sum_{i=1}^\infty|\lambda_i|,$$

$$|\langle\gamma_h(t)\phi|\psi\rangle| \le u\cdot c_h(t)\cdot p_V(\phi)p_V(\psi) \quad \text{for all } \phi,\ \psi\in\Phi \tag{V.5}$$

and $t\in R$, $t\notin N$. Since B as a bounded set can be absorbed into the o-neighborhood V it follows at once

$$P_B(\gamma_h(t)\phi) \le \tilde C(t)p_V(\phi), \text{ for every } \phi\in\Phi.$$

The positiveness of $\gamma_h(t)$ is obvious. QED

We remark in passing that $\gamma_h(t)$ is μ-almost everywhere an eigenoperator of the operator A in question. Let now $\{h_n\}$ be a sequence of elements of $\mathcal{G}$ such that $\mathcal{G}=\oplus_{n=1}^\infty\mathcal{G}_n$ holds with each $\mathcal{G}_n=\mathcal{G}_{h_n}$ as the closed subspace of $\mathcal{G}$ generated by all elements of the form $E(\sigma_i)h_n$, $\sigma\in\mathcal{B}$. We shall denote by P_n the orthogonal projection of $\mathcal{G}$ onto $\mathcal{G}_n$. Each P_n commutes with every $E(\sigma)$, $\sigma\in\mathcal{B}$. For the next proposition we shall need the following:

Definition 8: (Ref. 11). Let E be a l.c. space. We say that a sequence $\{x_n\}\subset E$ converges unconditionally to $x\in E$ it for each o-neighborhood $\mathcal{U}$ in E there exists a finite index set $N_U\subset N$ such that for every finite index set $\tilde N$ with $N_U\subset\tilde N\subset N$ $\sum_{n\in\tilde N}x_n\in x+\mathcal{U}$ holds. We say that the sequence $\{x_n\}\subset E$ converges absolutely to $x\in E$ if it converges unconditionally to x and if for each continuous semi-norm p on E the series $\sum_{n\in N}p(x_n)$ is convergent.

Proposition 19: The sequence $\{\gamma_n(t)\}$ is absolutely convergent $\gamma(t)$ for μ-almost all $t\in R$, in the strong topology of $\mathcal{L}(\Phi;\Phi^\times)$. Moreover, $\gamma(t)\phi$ has the representation

$$\gamma(t)\phi = \sum_{n=1}^\infty\overline{\langle\phi_n^\times(t)|\phi\rangle}\,\phi_n^\times(t). \tag{V.6}$$

Proof: We remark that the so-called strong topology in $\mathcal{L}(\Phi,\Phi^\times)$ is given by the following family of seminorms:

$$P_{B,B'}(\gamma) = \sup\{|\langle\gamma\phi|\psi\rangle|;\ \phi\in B,\ \psi\in B'\},$$

where B, B' run through all bounded sets in Φ. Now by (V5) we have at once

$$\sum_n p_{B,B'}(\gamma_n(t)) \le A \cdot \sum_{i,n=1}^{\infty} |\lambda_i| \, \|\widehat{(P_n h_i)}(t)\|_t^2 .$$

With the aid of the theorem of Beppo Levi we get from

$$\sum_{i,n=1}^{\infty} |\lambda_i| \int_R \|\widehat{(P_n h_i)}(t)\|_t^2 d\mu(t) \le \sum_{i=1}^{\infty} |\lambda_i| < \infty$$

that

$$\sum_{i,n=1}^{\infty} |\lambda_i| \, \|\widehat{(P_n h_i)}(t)\|_t^2 < \infty \quad \text{for all } t \in R, \ t \notin N'.$$

Hence

$$\sum_{n=1}^{\infty} p_{B,B'}(\gamma_n(t)) < \infty, \quad \mu \text{ almost everywhere}.$$

Furthermore, we have

$$\sum_{n=1}^{\infty} \int_\sigma |\langle \gamma_n(t)\phi | \psi\rangle| \, d\mu(t) \le P_V(\phi) P_V(\psi) \left(\sum_{i=1}^{\infty} |\lambda_i|\right)^2 < \infty,$$

whence by the theorem of Beppo Levi

$$\sum_{n=1}^{\infty} \int_\sigma \langle \gamma_n(t)\phi | \psi\rangle d\mu(t) = \int_\sigma \sum_{n=1}^{\infty} \langle \gamma_n(t)\phi | \psi\rangle d\mu(t),$$

for all ϕ, $\psi \in \Phi$.

Thus

$$\begin{aligned}(E(\sigma)\phi, \psi) &= \int_\sigma \langle \gamma(t)\phi | \psi\rangle d\mu(t) \\ &= \sum_{n=1}^{\infty} (P_n E(\sigma)\phi, \psi) \\ &= \sum_{n=1}^{\infty} \int_\sigma \langle \gamma_n(t)\phi | \psi\rangle d\mu(t) \\ &= \int_\sigma \sum_{n=1}^{\infty} \langle \gamma_N^c(t)\phi | \psi\rangle d\mu(t).\end{aligned}$$

Now, by standard arguments we can show that

$$\sum_{n=1}^{\infty} \langle \gamma_n(t)\phi | \psi\rangle$$

is a sesquilinear continuous form on $\Phi \times \Phi$. Therefore we get for all $\phi, \psi \in \Phi$ and $t \in R$, $t \notin N$ $[\mu(N) = 0]$

$$\sum_{n=1}^{\infty} \langle \gamma_n(t)\phi | \psi\rangle = \langle \gamma(t)\phi | \psi\rangle .$$

Let $N_U \subset N$ be any finite index set. As before we can derive the inequality

$$\begin{aligned}&\left|\left\langle \left(\sum_{n \in N_U} \gamma_n(t) - \gamma(t)\right)\phi \,\middle|\, \psi \right\rangle\right| \\ &\quad \le P_U(\phi) P_U(\psi) \cdot u \cdot \sum_{i=1}^{\infty} |\lambda_i| \sum_{n \in N_U} \|\widehat{(P_n h_i)}(t)\|_t^2 \Big).\end{aligned}$$

Let now

$$\mathcal{U} = \{\gamma \in \mathcal{L}(\Phi, \Phi^\times);\ p_{B,B'}(\gamma) < 1\}$$

be a o-neighborhood in $\mathcal{L}_b(\Phi, \Phi^\times)$.

Then we can choose an $M > 0$ that

$$\sum_{i=M}^{\infty} |\lambda_i| \sum_{n \in N_U} \|\widehat{(P_n h_i)}(t)\|_t^2 < \frac{1}{2 \cdot U} \frac{1}{\lambda_B \cdot \lambda_{B'}}$$

holds, with $U = \sum_{i=1}^{\infty} |\lambda_i|$. For the remaining part

$$\sum_{i=1}^{M} |\lambda_i| \sum_{n \notin N_U} \|\widehat{(P_n h_i)}(t)\|_t^2$$

we can find such a N_U that for every finite index set $\hat{N} \supset N_U$

$$\sum_{i=1}^{M} |\lambda_i| \sum_{n \notin \hat{N} \supset N_n} \|\widehat{(P_n h_i)}(t)\|_t^2 < \frac{1}{2 \cdot U} \cdot \frac{1}{\lambda_B \cdot \lambda_{B'}}$$

holds, where $\lambda_B, \lambda_{B'}$ fulfill the conditions $B \subset \lambda_B V$ and $B' \subset \lambda_{B'} V$, respectively. Thus

$$\left|\left\langle \left(\sum_{n \in \hat{N} \supset N_U} \gamma_n(t) - \gamma(t)\right)\phi \,\middle|\, \psi \right\rangle\right| \le P_V(\phi) P_V(\psi) \cdot \frac{1}{\lambda_B \cdot \lambda_{B'}},$$

for all $\phi, \psi \in \Phi_0$ and $t \in R$, $t \notin N$ with $\mu(N) = 0$. However, by continuity this relation holds for all $\phi, \psi \in \Phi$. Finally we get

$$P_{B,B'}\left(\sum_{n \in \hat{N} \supset N_U} \gamma_n(t) - \gamma(t)\right) < 1.$$

Obviously $\gamma(t)\phi$ has the representation

$$\gamma(t)\phi = \sum_{n=1}^{\infty} \overline{\langle \phi_n^\times(t) | \phi\rangle} \phi_n^\times(t),$$

for all $t \in R$, $t \notin N$ with $\mu(N) = 0$. QED

Next we want to show that the elements $\phi_n^\times(t)$ (for fixed t) occurring in the representation of $\gamma(t)$ are contained in a certain subspace of $\Phi^\times$.

We shall need the following:

Lemma: Let $\{k_n\} \subset \mathcal{G}$ be a sequence converging to zero in $\mathcal{G}$. Then there exists a subsequence $\{k_{n_p}\}$ such that $I^\times(t)\hat{k}_{n_p}(t)$ converges to zero with respect to the strong topology in $\Phi^\times$ for μ-almost all $t \in R$.

Proof: We have

$$\int_\sigma |\langle I^\times(t)\hat{h}(t) | \phi\rangle| \, d\mu(t) \le [(E(\sigma)h, h)]^{1/2} \left(\int_\sigma \langle \gamma(t)\phi | \phi\rangle d\mu(t)\right)^{1/2},$$

with $h \in \mathcal{G}$, $\phi \in \Phi$, and $\sigma \in \mathcal{B}$. Now since $(E(\sigma)k_n, k_n)$ is absolutely continuous with respect to $\mu(\sigma)$ by the Radon–Nikodym theorem there exists a function $k_n(t)$ such that

$$(E(\sigma)k_n, k_n) = \int_\sigma k_n^2(t) d\mu(t)$$

holds with $\sigma \in \mathcal{B}$. This relation yields

$$\begin{aligned}&\int_\sigma |\langle I^\times(t)\hat{k}_n(t) | \phi\rangle| \, d\mu(t) \\ &\quad \le \left(\int_\sigma k_n^2(t) d\mu(t)\right)^{1/2} \left(\int_\sigma \langle \gamma(t)\phi | \phi\rangle d\mu(t)\right)^{1/2}.\end{aligned}$$

Let now[10]

$$\begin{aligned}\sigma_n(\alpha, \beta, \gamma) = \{t : &|\langle I^\times(t)\hat{k}_n(t) | \phi\rangle| \ge \alpha, \\ &|k_n(t)| \le \beta,\ [\langle \gamma(t)\phi | \phi\rangle]^{1/2} \le \gamma\},\end{aligned}$$

where α, β, γ are assumed to be rational numbers that fulfill the conditions β, $\gamma \ge 0$, $\beta\gamma < \alpha$. Then we have

$$\begin{aligned}\alpha \cdot \mu[\sigma_n(\alpha, \beta, \gamma)] &\le \int_{\sigma_n(\alpha,\beta,\gamma)} |\langle I^\times(t)\hat{k}_n(t) | \phi\rangle| \, d\mu(t) \\ &\le \left(\int_{\sigma_n(\alpha,\beta,\gamma)} k_n^2(t) d\mu(t)\right)^{1/2} \cdot \left(\int_{\sigma_n(\alpha,\beta,\gamma)} \langle \gamma(t)\phi | \phi\rangle d\mu(t)\right)^{1/2}\end{aligned}$$

$$\leq \{\beta^2 \cdot \mu[\sigma_n(\alpha,\beta,\gamma)]\}^{1/2} \cdot \{\gamma^2 \cdot \mu[\sigma_n(\alpha,\beta,\gamma)]\}^{1/2}$$

$$= \beta \cdot \gamma \cdot \mu[\sigma_n(\alpha,\beta,\gamma)].$$

However, this relation is fulfilled if and only if $\mu[\sigma_n(\alpha,\beta,\gamma)]=0$. Since the set $N_n(\phi)$ of all $t\in R$ for which the inequality

$$|\langle I^\times(t)\hat{k}_n(t)|\phi\rangle| \leq |k_n(t)|[\langle\gamma(t)\phi|\phi\rangle]^{1/2} \tag{V.7}$$

is not valid is the countable union of the sets $\sigma_n(\alpha,\beta,\gamma)$, it follows that $\mu(N_n(\phi))=0$. Let Φ_0 be a countable dense set in Φ. Then $N_n = \cup_{\phi\in\Phi_0} N_n(\phi)$ is again a set of μ-measure zero and (V.7) can be extended to hold for all $\phi\in\Phi$. Then for any bounded set B in Φ we get

$$P_B(I^\times(t)\hat{k}_n(t)) \leq |k_n(t)|[P_{B,B}(\gamma(t))]^{1/2}, \tag{V.8}$$

for all $t\in R$ but $t\notin N_n$, where $P_{B,B}(\cdot)$ is a continuous seminorm in the space $\mathcal{L}_b(\Phi,\Phi^\times)$. We have to show that

$$P_{B,B}(\gamma(t))<\infty \text{ for all } t\in R \text{ but } t\in N \text{ with } \mu(N)=0,$$

where N is independent of B. In order to do so we observe $P_{B,B}(\gamma(t))=\sup\{|\langle\gamma(t)\phi|\psi\rangle|;\ \phi\in B,\ \psi\in B\}$. Now we have by Proposition 14

$$\langle\gamma(t)\phi|\psi\rangle$$
$$=\langle I(t)\phi|I(t)\psi\rangle_t$$
$$\sum_{i,k=1}^{\infty}\lambda_k\bar{\lambda}_i\langle\phi_k^\times|\phi\rangle\overline{\langle\phi_j^\times|\psi\rangle}\langle\hat{h}_k(t)|\hat{h}_i(t)\rangle_t,$$

whence

$$|\langle\gamma(t)\phi|\psi\rangle| \leq \sum_{k,i=1}^{\infty}|\lambda_k|\,|\lambda_i|\,|\langle\phi_k^\times|\phi\rangle|\,|\langle\phi_i^\times|\psi\rangle|\,\|\hat{h}_k(t)\|_t\|\hat{h}_i(t)\|_t$$

$$\leq P_V(\phi)P_V(\psi)\sum_{k=1}^{\infty}|\lambda_k|\sum_{k=1}^{\infty}|\lambda_k|\,\|\hat{h}_k(t)\|_t^2.$$

Since B is bounded we can absorb it into the o-neighborhood V determined by the seminorm $P_V(\cdot)$. Referring to the proof of Proposition 14, we obtain

$$P_{B,B}(\gamma(t)) \leq \lambda_B^2\cdot u\cdot C(t)<\infty \text{ for all } t\in R \text{ but } t\in N,$$

where N is the μ-null set indicated in the proof of Proposition 14 and

$$u=\sum_{k=1}^{\infty}|\lambda_k|<\infty.$$

N is in fact independent of the set $B\in\hat{\mathcal{B}}$ in question. Now if we put

$$\hat{N}=\bigcup_{n=1}^{\infty}N_n$$

then we have $\mu(\hat{N})=0$ and the inequality (V.8) is valid for all $t\in R$ but $t\in\hat{N}$. Let $\{k_{n_p}\}$ be a subsequence of $\{k_n\}$ such that

$$\sum_{p=1}^{\infty}\|k_{n_p}\|^2<\infty.$$

From

$$\sum_{p=1}^{\infty}\int_R|k_{n_p}(t)|^2d\mu(t)=\sum_{p=1}^{\infty}\|k_{n_p}\|^2<\infty$$

we conclude by the theorem of Beppo Levi that

$$\sum_{p=1}^{\infty}|k_{n_p}(t)|^2<\infty \text{ for } \mu\text{-almost all } t\in R.$$

Therefore, $|k_{n_p}(t)|\to 0$ if $p\to\infty$ for $t\notin N'$ with $\mu(N')=0$. Up to the μ-null set $N\cup\hat{N}\cup N'$ it follows

$$P_B(I^\times(t)\hat{k}_{n_p}(t))\to 0 \text{ if } p\to\infty$$

which concludes the proof. QED

Henceforth, we shall denote by $\Phi^\times(t)$ the closure of $\gamma(t)\Phi$ with respect to the strong topology $\beta(\Phi^\times,\Phi)$ in $\Phi^\times$.

Proposition 20: For $h\in\mathcal{G}$ it follows $I^\times(t)\hat{h}(t)\in\Phi^\times(t)$ for μ, almost all $t\in R$.

Proof 10: Let

$$h=\sum_i\lambda_iE(\sigma_i)I\phi_i \text{ with } \sigma_i\in\mathcal{B},\ \phi_i\in\Phi.$$

We set $\tilde{h}(t)=\sum_i\lambda_i\chi_{\sigma_i}(t)\phi_i\in\Phi$. Then it follows

$$\sum_i\langle I^\times E(\sigma)\lambda_iE(\sigma_i)I\phi_i|\phi\rangle=\sum_i\lambda_i\langle I^\times E(\sigma\cap\sigma_i)I\phi_i|\phi\rangle$$

$$=\sum_i\lambda_i(E(\sigma\cap\sigma_i)I\phi_i|I\phi)=\sum_i\lambda_i\int_{\sigma\cap\sigma_i}\langle\gamma(t)\phi_i|\phi\rangle d\mu(t)$$

$$=\int_\sigma\langle\gamma(t)\hat{h}(t)|\phi\rangle d\mu(t).$$

From this relation we get $I^\times(t)\hat{h}(t)=\gamma(t)\tilde{h}(t)$ for μ, almost all $t\in R$, which means that $I^\times(t)\hat{h}(t)\in\gamma(t)\Phi$ for μ, almost all $t\in R$.

Let now $\{k_n\}$ be a sequence of elements indicated above which converges to an element $h\in\mathcal{G}$. Up to a μ-null set N if follows

$$I^\times(t)(\widehat{h-k_n})(t)=I^\times(t)\hat{h}(t)-I^\times(t)\hat{k}_n(t)$$

and

$$I^\times(t)\hat{k}_n(t)\in\gamma(t)\Phi \text{ for all } n=1,2,\cdots.$$

Now if we apply the preceding lemma to the sequence $\{h-k_n\}$ we get a subsequence $\{k_{n_p}\}$ such that $I^\times(t)(\widehat{h-k_{n_p}})(t)\to 0$ in the strong topology in $\Phi^\times$ for all $t\in R$ but $t\in N'$ with $\mu(N')=0$. Therefore we obtain $I^\times(t)\hat{h}(t)\in\Phi^\times(t)$ for all $t\in R$ but $t\in N\cup N'$ with $\mu(N\cup N')=0$ QED

Proposition 21: Let $\phi_h^\times(t)$ be the function indicated in Proposition 17. Then for μ, almost all $t\in R$, $\phi_h^\times(t)\in\Phi^\times(t)$.

Proof[10]: We denote by $\tilde{h}(t)$ the image of h under the isomorphic mapping $\mathcal{G}_h\to L_\mu^2$. We set $Z_h=\{t\in R;\ \tilde{h}(t)=0\}$.

Then for every $f\in\mathcal{G}_h$ it follows $u_f(t)=0$ for $t\in Z_h\backslash Z_h(f)$, where $Z_h(f)$ is a μ-null set. Let Φ_0 denote a countable dense set in Φ and set $Z_h'=\cup_{\phi\in\Phi_0}Z_h(P_h\phi)$. It follows $\mu(Z_h')=0$ and

$$\langle\phi_h^\times(t)|\phi\rangle=u_{P_h\phi}(t)=0$$

for every $\phi\in\Phi$ and $t\in Z_h/Z_h'$ whence $\phi_h^\times(t)=0$ for $t\in Z_h/$

Z'_h. On the other hand, we have

$$\langle I^\times E(\sigma)h \,|\, \psi\rangle = \int_\sigma \langle I^\times(t)\hat{h}(t) \,|\, \psi\rangle d\mu(t) = \int_\sigma \tilde{h}(t)\overline{u_{p_n\psi}(t)}d\mu(t),$$

from which we derive $I^\times(t)\hat{h}(t) = \tilde{h}(t)\phi_h^\times(t)$, for all $t\in R$, $t\notin N'$ with $\mu(N')=0$.

Now if $\phi_h^\times(t)=0$ we have already arrived at the conclusion. Let us consider those $t\in R$ with $t\notin N'\cup Z'_h$ and $t\notin Z_h$. Then we have $\tilde{h}(t)\neq 0$ and by Proposition 20 it follows

$$\phi_h^\times(t) = I^\times(t)\hat{h}(t)/\tilde{h}(t) \in \Phi^\times(t). \qquad \text{QED}$$

At the beginning of Sec. II we had introduced the so-called eigenforms of a self-adjoint operator A in $\mathcal{G}$ corresponding to value t in the spectrum of A to be those continuous antilinear forms $\phi_t^\times \in \Phi^\times$ that fulfil the condition

$$\langle \phi_t^\times \,|\, A\phi\rangle = t\langle \phi_t^\times \,|\, \phi\rangle \quad \text{for every } \phi\in\Phi.$$

Now let for a fixed $t\in \mathrm{sp}(A)$ $\Phi_t^\times$ denote the space of all eigenforms of A corresponding to t. $\Phi_t^\times$ is a closed linear subspace of $\Phi^\times$. Indeed, we observe that $\langle \phi_t^\times | (A - t1)\phi\rangle = 0$ for every $\phi\in\Phi^\times$. Since $(A-t1)\Phi = M$ is a subset of Φ we obtain $\Phi_t^\times = M^\perp$ is orthogonal to M. $M^\perp$ is a $\sigma(\Phi^\times,\Phi)$-closed subspace of Φ. Since the strong topology in $\Phi^\times$ is stronger than the weak topology $\sigma(\Phi^\times,\Phi)$ it follows that $\Phi_t^\times$ is also a closed linear subspace of $\Phi^\times$ with respect to the strong topology in $\Phi^\times$.

Proposition 22: For μ, almost all $t\in R$, it follows $\Phi^\times(t)\subset\Phi_t^\times$.

Proof: Recall that $\Phi^\times(t)$ was defined to be the closure of $\gamma(t)\Phi$ in $\Phi^\times$ with respect to the strong topology in $\Phi^\times$. According to the remark following Definition 3 we have $\gamma(t)\Phi\subset\Phi_t^\times$. Since $\Phi_t^\times$ is closed we see immediately that $\Phi^\times(t)\subset\Phi_t^\times$. QED

Corollary: The continuous antilinear forms $\phi_k^\times(t)$, $k=1, 2, \cdots$ originating from the representation of the mappings $\gamma(t)$ are μ-almost everywhere elements of $\Phi_t^\times$. Moreover for every $h\in\mathcal{G}$ we have $I^\times(t)\hat{h}(t)\in\Phi_t^\times$ for μ, almost all $t\in R$. Indeed, by Proposition 21, we have, μ almost everywhere, $\phi_k^\times(t)\in\Phi^\times(t)\subset\Phi_t^\times$. Furthermore, by Proposition 20 it follows $I^\times(t)\hat{h}(t)\in\Phi^\times(t)\subset\Phi_t^\times$, μ almost everywhere.

Proposition 23: $\Phi_t^\times\cap I^\times\mathcal{G}\neq\{0\}$ if and only if t is an eigenvalue of A in $\mathcal{G}$. In this case $\Phi_t^\times\cap I^\times\mathcal{G}$ is the subspace of $I^\times\mathcal{G}$ which is generated by the $I^\times$-image of all eigenvectors of A in $\mathcal{G}$.

Proof: Let $h\in\mathcal{G}$ be an eigenvector of A in $\mathcal{G}$ belonging to the eigenvalue t. Then for every $\phi\in\Phi$ we get

$$\langle I^\times h, A\phi\rangle = (h, IA\phi) = (Ah,\phi) = t(h,\phi) = t\langle I^\times h,\phi\rangle.$$

Thus $I^\times h\in\Phi_t^\times$ and we have shown that $\Phi_t^\times$ contains the image of all eigenvectors of A in $\mathcal{G}$ for an eigenvalue t. On the other hand, if $\Phi_t^\times\cap I^\times\mathcal{G}\neq\{0\}$ it follows that t is an eigenvalue of A and every element of $\Phi_t^\times\cap I^\times\mathcal{G}$ is an eigenvector of A. Indeed, for $0\neq h^\times\in\Phi_t^\times\cap I^\times\mathcal{G}$ we have $\langle h^\times, A\phi\rangle = t\langle h^\times,\phi\rangle$ for every $\phi\in\Phi$. Since A is self-adjoint and Φ is dense in $\mathcal{G}$ we have $A^\times h^\times = th^\times$. QED

Corollary: $\gamma(t)\phi\cap I^\times\mathcal{G}\neq\{0\}$ if and only if t is an eigenvalue of A in $\mathcal{G}$. Indeed, for if $\gamma(t)\phi\cap I^\times\mathcal{G}\neq\{0\}$ it follows $\Phi_t^\times\cap I^\times\mathcal{G}\neq\{0\}$ and hence by the preceding proposition t is an eigenvalue of A in $\mathcal{G}$. Conversely, if $E(\{t\})\neq 0$, we get $\mu(\{t\})\neq 0$. Thus for every $\psi\in\Phi$ we obtain

$$\langle I^\times E(\{t\})I\phi,\psi\rangle = \int_{\{t\}}\langle\gamma(t)\phi,\psi\rangle d\mu(t) = \langle\gamma(t)\phi,\psi\rangle\,\mu(\{t\})$$

and therefore

$$\gamma(t)\Phi\cap I^\times\mathcal{G} = I^\times E(\{t\})I\Phi\cap I^\times\mathcal{G}\neq\{0\}.$$

In addition we emphasize that it is unknown whether $\Phi^\times(t)=\Phi_t^\times$, μ almost everywhere, holds. This fact makes it difficult to set up a simple calculus for the generalized eigenvectors of a self-adjoint operator A in $\mathcal{G}$. By solving the eigenvalue equation $A^\times\phi_t^\times = t\phi_t^\times$ in $\Phi^\times$ one might find solutions which have nothing in common with the generalized spectral decomposition within the Rigged Hilbert space $\Phi\subset\mathcal{G}\subset\Phi^\times$. Further conditions are needed in order to decide whether a solution $\phi_t^\times$ belongs to the space $\Phi^\times(t)$ for μ, almost all $t\in R$. We emphasize that we always restrict our attention to the spectrum of the operator A in $\mathcal{G}$ only. The spectrum of the operator $A^\times$ in $\Phi^\times$ can be larger than the Hilbert space spectrum of A.

At the end of this section we shall investigate somewhat more closely the relationship between the form $I^\times E(\sigma)I\phi$ and $\gamma(t)\phi$ for $\phi\in\Phi$. Before entering into details, a few preparatory remarks are needed.

First we observe that the canonical embedding I of Φ into $\mathcal{G}$ and its conjugate $I^\times:\mathcal{G}\to\Phi^\times$ can be decomposed in the following way:

$$\Phi \xrightarrow[I_V]{} \tilde{\Phi}_V \xrightarrow[\hat{I}]{} \mathcal{G} \xrightarrow[\hat{I}^\times]{} \Phi_V^\times \xrightarrow[I_V^\times]{} \Phi^\times.$$

Note that $\tilde{\Phi}_V^\times$ and $(\Phi^\times)_{V^\circ}$ (cf. Ref. 11, p. 97, and Ref. 12, p. 277) are normisomorphic.

From the proof of Proposition 14 it is obvious that the mapping I: $\Phi\to\hat{\mathcal{G}}(t)$ is given by $\hat{I}(t)\cdot I_V$, with $I(t)$: $\tilde{\Phi}_V\to\mathcal{G}$. Since the norm-topology and the strong topology on $(\Phi^\times)_{V^\circ}$ coincide, it follows that $I_V^\times$ is a continuous map with respect to the norm-topology in $(\Phi^\times)_{V^\circ}$ and $\beta(\Phi^\times,\Phi)$ on $\Phi^\times$. We set

$$\hat{\gamma}(t) = \hat{I}(t)^\times\cdot\hat{I}(t).$$

Let then $\Phi_V^\times(t)$ and $\Phi^\times(t)$ denote the closure of $\hat{\gamma}(t)\tilde{\Phi}_V$ (with respect to the norm in $\Phi_V^\times$) and the closure of $\gamma(t)\Phi$ [with respect to $\beta(\Phi^\times,\Phi)$], respectively.

Proposition 24: Let $t_0\in\mathrm{sp}(A)$. Then there exist sequences $\{t'_n\}$ and $\{t''_n\}$, with $t'_n\uparrow t_0$ and $t''_n\downarrow t_0$ such that for each $\phi\in\Phi$

$$\frac{I^\times E([t'_n,t''_n])I\phi_V}{\mu([t'_n,t''_n])}$$

converges to $\gamma(t_0)\phi$ for $n\to\infty$ with respect to $\beta(\Phi^\times,\Phi)$ in $\Phi^\times$.

Proof: Foias[10] has shown that there exist sequences $\{t''_n\}$ such that

$$\frac{\hat{I}E([t'_n,t''_n])\hat{I}\phi_V}{\mu([t'_n,t''_n])}$$

converges to $\hat{\gamma}(t_0)\phi_V$ with respect to the norm in $(\Phi^\times)_{V^\circ}$. Now, since $I_V^\times$: $(\Phi^\times)_{V^\circ} \to \Phi^\times$ is a continuous map we arrive at the conclusion.

ACKNOWLEDGMENTS

The author would like to thank Professor G. Ludwig for his generous encouragement and Dr. A. Harkämper, Dr. H. Neumann, and Dr. F. Jenč for stimulating discussions and useful comments.

[1]I.M. Gel'fand, G.E. Shilov, and N.I. Vilenkin, *Generalized Functions* (Academic, New York, 1964), Vols. I—IV.
[2]A. Böhm, Lect. Theor. Phys. 9A, 255 (1967).
[3]J.E. Roberts, J. Math. Phys. 7, 1097 (1966); Commun. Math. Phys. 3, 98 (1966)
[4]J.-P. Antoine, J. Math. Phys. 10, 53, 2276 (1969).
[5]A. Grossman, J. Math. Phys. 5, 1025 (1964); 6, 54 (1965); Commun. Math Phys. 2, 1 (1966); 4, 190 (1967).
[6]E. Prugovečki, Can J. Math. 21, 158 (1969).
[7]M.E. Mayer, Proc. of Seminar on Unified Theories of Elementary Particles, Feldafing, July 1965, pp. 515—27.
[8]C. Foiaş, Rev. Math. Pure Appl. 7, 241, 571 (1962).
[9]K. Maurin, Bull. Acad. Pol. Sci. Sér. Sci. Math. 7, 470 (1959).
[10]C. Foiaş, Acta Sci. Math. 20, 117 (1959).
[11]H. Schäfer, *Topological Vector Spaces* (MacMillan, New York, 1967).
[12]J. Horváth, *Topological Vector Spaces and Distribution* (Addison-Wesley, Reading, Mass., 1966), Vol. 1.
[13]A. Pietsch, *Nukleare lokalkonvexe Räume* (Akademie-Verlag, Berlin, 1965).
[14]B. Sz.-Nagy, Appendix to F. Riesz and B. Sz.-Nagy, *Functional Analysis* (Ungar, New York, 1960).
[15]S.K. Berberian, *Measure and Integration* (MacMillan, New York, 1965).

Rigged Hilbert space formalism as an extended mathematical formalism for quantum systems. II. Transformation theory in nonrelativistic quantum mechanics

O. Melsheimer

Institut für Theoretische Physik, Universität Marburg, Marburg, Germany
(Received 6 September 1972)

Results of a previous paper are used to obtain a rigorous mathematical formulation of the transformation theory of nonrelativistic quantum mechanics within the framework of rigged Hilbert spaces.

I. INTRODUCTION

In his original form the rigged Hilbert space formalisms was invented to provide a framework for establishing a generalized eigenfunction decomposition for self-adjoint operators A in a complex Hilbert space $\mathcal{G}$. This essentially offers the possibility to treat both the discrete and the continuous spectrum of a s.a. operator in $\mathcal{G}$ on equal footing. The general idea of this formalism is to have the generalized eigenfunctions as elements of an enlarged space $\Phi^{\times}\supset\mathcal{G}$. To give $\Phi^{\times}$ a mathematical meaning one establishes $\Phi^{\times}$ as the antidual space of a locally convex space $\Phi\subset\mathcal{G}$.

Constructively one can set up a generalized spectral decomposition of a s.a. operator A in $\mathcal{G}$ within the triplet structure $\Phi\subset\mathcal{G}\subset\Phi^{\times}$ provided the operator A is a continuous mapping of Φ into itself and the embedding of Φ into $\mathcal{G}$ is nuclear.

For application of the rigged Hilbert space formalism to the eigenfunction decomposition of a given s.a. operator one has to find such a triplet with the properties indicated above. However, in order to be able to cope with the interrelationships of various observables in a certain quantum mechanical theory Roberts,[1] Böhm,[2] and Antoine[3] investigated the problem whether there is a common triplet for all operators which belong to the class of observables of the quantum system in question. Roberts made the assumption that this class of observables is an algebra. Then under a special assumption on this algebra $\mathfrak{A}$ of unbounded operators in $\mathcal{G}$, namely, that $\mathfrak{A}$ contains an element whose inverse exists and is an element of the trace class he found a very elegant solution for such a triplet. In a previous paper[4] which henceforth will be referred to as I, we have reinvestigated the solution of Roberts. Together with other results of paper I concerning the topological aspects of the generalized spectral decomposition we shall give in the present paper a rigorous treatment of the so-called transformation theory in nonrelativistic quantum mechanics. Although this problem seems at present to be of no direct physical relevance it still seems of interest being an as yet unsolved problem within the Hilbert space formalism of quantum mechanics. For our purpose we identify the Roberts algebra $\mathfrak{A}$ with the enveloping algebra $\Pi(\mathcal{A})$ of a unitary representation of the central extension of the Galilei group which, however, certainly does not represent the full set of observables of nonrelativistic quantum mechanics.

On the basis of well-known properties of $\Pi(\mathcal{A})$ we derive a very handy representation of the elements of $\Phi^{\times}$ which then in fact allows us to set up the desired transformation formulas. It turns out for example, that the mapping of the functions $\langle\phi^{\times}(t)|\phi\rangle$ (A-representation) to the numbers $\langle\psi^{\times}(u)|\phi\rangle$ (B-representation) is in general a distribution.

II. THE BASIC DOMAIN Φ OF NONRELATIVISTIC QUANTUM MECHANICS

We shall assume that the algebra of observables of nonrelativistic quantum mechanics of one particle can be identified with the enveloping algebra of the infinitesimal generators of an irreducible unitary representation of the central extension of the Galilei group for mass m and spin s.

A vector $h\in\mathcal{G}$ is called an infinitely differentiable vector for the strongly continuous unitary representation $\mathfrak{A}(G)$ of the Lie group G if the map $g\to\mathcal{U}(g)h$ belongs to the class $C^{\infty}(G)$. The set of all such vectors will be denoted by $\mathcal{E}$. $\mathcal{E}$ is called the domain of all infinitely differentiable vectors of the representation $\mathfrak{A}(G)$.

Let $\mathfrak{G}$ denote the Lie algebra of the Lie group G. If $A\in\mathfrak{G}$ and if $a(s)$ is the one-parameter subgroup of G that is generated by A then $A_{\mathcal{U}}$ is defined as follows:

$$A_{\mathcal{U}}h=\lim_{s\to 0}S^{-1}\{U[\exp(sA)-1]\}h \quad \text{for every } h\in\mathcal{E}.$$

$\mathcal{E}$ is invariant under all $A_{\mathcal{U}}$ with $A\in\mathcal{G}$. Now the question arises whether $\mathcal{E}$ is dense in $\mathcal{G}$. This question will be answered positively by the following proposition which is due to Gårding.[5]

Proposition 1: Let r be a natural number (finite or infinite) and let $\mathcal{G}_r$ be the set of all elements of $\mathcal{G}$ of the form

$$h(\varphi)=\int_G\varphi(g)\,U(g)h\,d\mu(g) \quad \text{with } \varphi\in C_0^r(G),$$

$h\in\mathcal{G}$, $d\mu(g)$ as the left invariant measure on G, and C_0^r the class of all functions of C^r with compact supports. Then $\mathcal{G}_r$ is dense in $\mathcal{G}$ for all r, $\mathcal{G}_{r+1}\subset D(A_{\mathcal{U}})$ and $A_{\mathcal{U}}\mathcal{G}_{r+1}\subset\mathcal{G}_r$ for every one-parameter subgroup $\{a(s)\}\subset G$. Since $\mathcal{G}_{\infty}\subset\mathcal{E}$ we have at once that $\mathcal{E}$ is dense in $\mathcal{G}$. The set $\mathcal{G}_{\infty}$ is called the Gårding domain of the representation $\mathfrak{A}(G)$. It is well known that the Lie algebra $\mathfrak{G}$ of a Lie group G can be realized by a set of differential $\mathfrak{G}$ operators on $C^{\infty}(G)$. This yields a representation $\hat{\mathfrak{G}}$ of $\mathfrak{G}$. For each one-parameter subgroup $x(t)$ of G one has

$$(Xf)(g)=\lim_{t\to 0}t^{-1}[f(x(t)g)-f(g)] \quad \text{with } f\in C^{\infty}(G).$$

The operator X so defined is right-invariant, i.e.,

$(R_y X)(f) = (XR_y)(f)$ with $R_y f(g) = f(gy)$, for $g \in G$.

Then any pair (X, Y) of such operators can be associated with a commutator $[X, Y]$ defined as follows:

$$[X, Y]f(g) = X[Yf] - Y[Xf].$$

It is immediate that $[X, Y]$ is also a differential operator, i.e.,

$$[X, Y](f\cdot\varphi) = \varphi([X, Y]f) + f([X, Y]\varphi) \quad \text{for } f, \varphi \in C^\infty(G).$$

For a basis $X_1, \ldots, X_n$ of $\mathfrak{G}$ one has

$$[X_1, X_k] = \sum_{\nu=1}^{n} C^{\nu}_{ik} X_\nu.$$

Now one can define the right-invariant enveloping algebra of the Lie group G. This algebra is the associative algebra $\mathcal{A}$ which is generated by all operators of $X_i \in \mathfrak{G}$. The elements of $\mathcal{A}$ are polynomials in the operators $X_i \in \mathfrak{G}$. Furthermore, we have

$$X_i X_k = X_k X_i + \sum_{\nu=1}^{n} C^{\nu}_{ik} X_\nu.$$

By definition, a representation π of the Lie algebra is any homomorphism Π of $\mathfrak{G}$ into the set of Hermitian operators in a Hilbert space $\mathcal{G}$ with D as a dense set in $\mathcal{G}$ that has the following properties:

$$\Pi([X, Y])h = [\Pi(X)\Pi(Y) - \Pi(Y)\Pi(X)]h \quad \text{for every } h \in D,$$

$$\Pi(X) D \subset D \quad \text{for every } X \in \mathfrak{G}.$$

We summarize the results obtained so far in the following:

Proposition 2: Every strongly continuous unitary representation of a Lie group G on a Hilbert space $\mathcal{G}$ induces a representation of the Lie algebra $\mathfrak{G}$ on $\mathcal{G}_\infty$. This representation is given by

$$\Pi(X) = d\mathcal{U}(X) \text{ for } X \in \mathfrak{G},$$

where $d\mathcal{U}(X)$ is defined by the relation

$$d\mathcal{U}(X)\, h(\varphi) = \mathcal{U}(X\varphi)h = -\lim_{t\to 0} t^{-1}[\mathcal{U}(X(t)) - 1]h(\varphi)$$

with

$\mathcal{U}(\varphi)h = \int_G \varphi(g)\mathcal{U}(g)h\, d\mu(g) = h(\varphi)$, $\varphi \in C_0^\infty(G)$ and $h \in \mathcal{G}$. Each $d\mathcal{U}(X)$ is skew-symmetric.

Thus $\mathcal{U}(G)$ induces the so-called differential representation $X \to d\mathcal{U}(X)$, $X \in \mathfrak{G}$ on the Gårding domain $\mathcal{G}_\infty$. This representation can be extended to a representation of the right-invariant enveloping algebra $\mathcal{A}$ on $\mathcal{G}_\infty$.

Now $\Pi(\mathcal{A})$ is an algebra that has all the properties indicated in I, Sec. II.[6]

Proposition 3: The initial topology τ_{in} on $\mathcal{G}_\infty = \Phi$ with respect to all elements of $\Pi(\mathcal{A})$ is given by the following family of seminorms:

$$\phi \to \|\Delta^n \phi\|, \quad n = 0, 1, 2, \ldots, \quad \phi \in \Phi,$$

with

$$\Delta = \textstyle\sum_{k=1}^{n} \Pi(X_k)^2 + 1.$$

Proof: Since every Δ^ν $(\nu = 0, 1, \ldots)$ belongs to $\Pi(\mathcal{A})$ we see that the topology τ_Δ defined by the seminorms $\phi \to \|\Delta^n\phi\|$ is coarser than the topology τ_{in}. In order to prove that τ_Δ is also finer than τ_{in} we have to show that for any $A \in \Pi(\mathcal{A})$ there exists a Δ^{n_A} such that

$$\|A\phi\| \le C_A \|\Delta^{n_A}\phi\| \quad \text{for every } \phi \in \Phi \text{ holds.}$$

However, this relation has explicitly been proved by Nelson.[6] QED

Quite often Δ is called the Nelson operator for $\Pi(\mathcal{A})$.

Corollary 1: The initial topology τ_{in} on Φ with respect to $\Pi(\mathcal{A})$ is equivalent to the initial topology with respect to the algebra $\mathcal{A}_\Delta$ which is generated by the elements 1 and Δ.

The proof is immediate by repeating all the arguments that have been employed for the case of the original algebra $\mathfrak{A}$.

Corollary 2: The space $\Phi^\times$ of all continuous antilinear forms over Φ consists of all elements of the form

$$\Phi^\times = \sum_{u=1}^{u(\phi^\times)} \Delta^{\alpha_u} 0 \mathcal{g}_{\alpha_u}$$

operating on $\phi \in \Phi$ as $\langle \Phi^\times | \phi\rangle = \Sigma_{u=1}^{u(\phi^\times)} (\mathcal{g}_{\alpha_u}, \Delta^{\alpha_u}\phi)$,

where $u(\phi^\times)$ is finite and $\mathcal{g}_{\alpha_u} \in \mathcal{G}$.

Again, the proof of this assertion is achieved by following the same lines as in the proof of I, Proposition 3.

Now we come to the point mentioned at the beginning of this section. We shall refrain here from explaining all the details of the so called central extension of the Galilei group. In order to fix the notation let us define the Galilei group G to be the group of all transformations of space—time (R^3, T) of the form $g = (D, \eta, \mathbf{v}, \mathbf{u})$ defined on (R^3, T) as $\mathbf{x}' = D\mathbf{x} + \mathbf{v}t + \mathbf{u}$, $t' = t + \eta$ and obeying the multiplication law

$$g \cdot g' = (D \cdot D', \eta + \eta', \mathbf{u} + D\mathbf{u}', \mathbf{v} + D\mathbf{v}').$$

Here D denotes an element of the rotation group $SO(3)$, $\mathbf{u}$ is a space translation, $\mathbf{v}$ is a pure Galilei transformation, and η denotes the time translation. Then the central extension $\hat{G}$ of the Galilei group is the set of all elements $\hat{g} = (\Theta, D, \eta, \mathbf{v}, \mathbf{u}) = (\Theta, g)$ satisfying the multiplication law

$$\hat{g} \cdot \hat{g}' = (\Theta + \Theta' + \mathbf{v} D\mathbf{u}' + \eta'(v^2/2), g \cdot g').$$

Let us adopt the following notations:

J_i $(i = 1, 2, 3,)$ are the generators of the rotation group $SO(3)$,
Q_i $(i = 1, 2, 3)$ are the generators of the pure Galilei transformation,
P_i $(i = 1, 2, 3)$ are the generators of the space translation,
H is the generator of the time translation,
M is the generator of the central extension.

Thereby we get the following Lie algebra of the group $\hat{G}$:

$$[J_i, J_j] = e_{ijk} J_k; [J_i, Q_j] = e_{ijk} Q_k; [J_i, P_j] = e_{ijk} P_k; \; (i, j, k = 1, 2, 3);$$

$$[J_i, H] = 0; \; [J_i, M] = 0; \; [Q_i, Q_j] = 0; \; [Q_i, P_j] = \delta_{ij} M;$$

$$[Q_i, H] = P_i; \; [Q_i, M] = 0; \; [P_i, P_j] = 0; \; [P_i, H] = 0;$$

$$[P_i, M] = 0; \; [H, M] = 0; \; (i, j = 1, 2, 3).$$

The unitary irreducible representations of $\hat{G}$ can be classified according to the eigenvalues of the following three Casimir operators:

$$M,\ MH-P^2/2,\ (MJ-\mathbf{Q}\times\mathbf{P})^2=S^2.$$

Vector representations for different values of $MH-P^2/2$ become equivalent when considered as ray representations of the Galilei group G. Since we are interested in ray representations of the Galilei group only we have to consider only the representations of $\hat{G}$ with values of $M(-\infty<M<\infty)$ and $S^2=2s+1$ $(S=0,\frac{1}{2},1,\ldots)$. For a detailed explanation of the representation theory of $\hat{G}$ we refer to Refs. 7 and 8. In the so-called spin representation of $\hat{G}$ the infinitesimal generators indicated above have the following form: Up to a factor the operators Q_i can be identified with the three components of the position operator, P_i with three components of the momentum operator, and $\mathbf{J}$ with the total angular momentum $\mathbf{S}+\mathbf{L}$, where $\mathbf{S}$ and $\mathbf{L}$ denote the spin and the angular momentum, respectively. H and M are the energy and the mass of the free particle, respectively.

The Hilbert space $\mathcal{G}^{[m,s]}$ for a unitary irreducible representation of $\hat{G}$ characterized by the spin s and the mass m can be written as the following tensor product

$$\mathcal{G}^{[m,s]}=\mathcal{G}\otimes R^{2s+1},$$

where $\mathcal{G}$ is the Hilbert space that belongs to the irreducible representation of the operators P_i, Q_i, $i=1,2,3$ and R^{2s+1} is the representation space for $SO(3)$ for the spin s.

In what follows we restrict ourselves to the case $s=0$. Then, as is well-known, the representation in $\mathcal{G}^{[m,0]}$ is unitarily equivalent to $L^2(-\infty,\infty)$ where the operators P_i, Q_i, $\mathbf{L}$, and H are given by $-i\partial/\partial x_i$, x_i, $ix_i\mathbf{x}\partial/\partial x$, and $\partial^2/\partial x^2$, respectively.

Now in this case the enveloping algebra of the infinitesimal generators in the representation $[m,0]$ is equal to the enveloping algebra of the operators $-i\partial/\partial x_i$, x_i $(i=1,2,3)$.

The operator

$$H=\sum_{i=1}^{3}P_i^2+Q_i^2+1$$

has a purely discrete spectrum with finite multiplicity. The spectrum of H is equal to the set N. Therefore the operator $(H^2)^{-1}$ is a nuclear operator since

$$\sum_{n=1}^{\infty}\frac{1}{n^2}<\infty.$$

The operator H is strictly positive, which leads to the conclusion: The topology in Φ is generated by the positive-definite Hermitian forms

$$\{\phi,\phi\}\to(H^k\phi,H^k\phi)=(\phi\,|\,\phi)_k.$$

Φ becomes a pre-Hilbert space with respect to each such form. The completion $\mathcal{G}_k$ of such a space is achieved by taking the closure of the operator H^k in $\mathcal{G}$.

For this case we are able to sharpen the result indicated in Corollary 2 of Proposition 3. Since each element $\phi^\times$ of $\Phi^\times$ fulfills the relation

$$|\langle\phi^\times|\,\phi\rangle|\leq\lambda\|H^k\phi\|,\quad \phi\in\Phi$$

it follows that $\phi^\times$ is a continuous form over the Hilbert space $\mathcal{G}_k$. Therefore $\phi^\times$ can be represented by an element h which is in the domain of definition of the operator H^k. We have

$$\langle\phi^\times|\,\phi\rangle=(H^kh,\ H^k\phi)\ \text{ for all }\phi\in\Phi.$$

Hence the element $\phi^\times$ can be represented by

$$\phi^\times H^{2k}\circ h\quad\text{with } h\in D(\overline{H^k}).$$

Furthermore, we emphasize that the spectrum of

$$H=\sum_{i=1}^{3}P_i^2+Q_i^2+1$$

can be determined completely without referring to any special representation of the operators P_i, Q_i, $i=1,2,3$.[9] The assumptions which have to be made are the commutation relations of the operators P_i, Q_i and the essential self-adjointness of the operator H on a domain $\Phi\subset\mathcal{G}$ which is a common invariant domain of definition for the operators P_i, Q_i.

According to what has been explained in this section the last condition can always be fulfilled. The essential self-adjointness of H on the Gårding domain has been shown in Ref. 10. We use again the Nelson method[6] quoted above in showing that every eigenvector ϕ_n of H belongs to the common invariant domain of all operators of $\mathcal{A}(P_i,Q_i)$. Nelson has shown that, for every $B\in\mathcal{A}(P_i,Q_i)$ there exists a power of H such that for all $h\in\mathcal{G}$ we have

$$\|Bh\|\leq k\|H^mh\|.$$

For if $h=\phi_n$ we get for any $B\in\mathcal{A}(P_i,Q_i)$

$$\|B\phi_n\|\leq k\|H^m\phi_n\|=k_n^m\|\phi_n\|,$$

for all $n\in N$. This relation shows that all eigenvectors of H belong to the maximal invariant domain of definition for all $B\in\mathcal{A}$, i.e.,

$$\Phi=\bigcap_{B\in\mathcal{A}}D(\overline{B}).$$

Now the maximal invariant domain $D\subset L^2(-\infty,\infty)$ for the operators x_i, $-i\partial/\partial x_i$, $i=1,2,3$, consists of all functions from C^∞ with the property

$$x^k\phi^{(m)}(x)\in L^2(-\infty,\infty),$$

where we have used the abbreviations $k=(k_1,k_2,k_3)$,

$$m=(m_1,m_2,m_3),\quad x^k=x_1^{k_1}\cdot x_2^{k_2}\cdot x_3^{k_3},$$

$$\phi^{(m)}=(\partial/\partial x_1)^{m_1}\cdot(\partial/\partial x_2)^{m_2}\cdot(\partial/\partial x_3)^{m_3}\,\phi.$$

The initial topology τ_{in} on D is given by the following family of seminorms:

$$\phi\to\|(\pm i)x^{k_1}\phi^{(m_1)}+\cdots+(\pm i)x^{k_n}\phi^{(m_n)}\|.$$

Let us endow now the space D with a topology that is determined by the following family of seminorms:

$$P_{k,m}(\phi)=\left(\int x^{2k}\,\overline{\phi^{(m)}}(x)\,\phi^{(m)}(x)\,dx\right)^{1/2}.$$

Then D equipped with this topology is homeomorphic to the space S.[1]

III. TRANSFORMATION THEORY

In this section we shall attempt to use the formalism of a rigged Hilbert space to provide a tool for tackling the following problems pertaining to the usual Hilbert space formalism of quantum mechanics: Choice of a representation, transformation of one representation into another, representation of operators in a given representation etc. First, we shall sketch the problems. Let A be an observable (s.a. operator) in $\mathcal{G}$ with a pure discrete spectrum. Then the eigenvectors h_ν of A form a complete orthonormal system in $\mathcal{G}$.

Using this fact, by the application $f \longleftrightarrow \{(h_1, f)\}$ one has an isometric mapping from $\mathcal{G}$ onto $l^2 \cong l^2_A$. l^2_A is called the A-representation of $\mathcal{G}$.

A is diagonal in this representation

$$\hat{A}(h_\nu, f) = (h_\nu, Af) \quad \text{for } f \in D(A).$$

Obviously this procedure fails in the presence of a continuous part in the spectrum of A. It is essentially this problem which can be very easily settled within the framework of a rigged Hilbert space. For the case of nonrelativistic quantum mechanics we shall use the triplet $\Phi \subset \mathcal{G} \subset \Phi^\times$ described in Sec. II. Let $A \in \mathcal{L}^c(\Phi)$ (cf. I, Sec. IV) be an operator which is essentially self-adjoint on Φ.

For the sake of simplicity we shall assume A to possess a simple spectrum. Then according to I, Proposition 18 the existence of the eigenforms $\phi^\times(t) \in \Phi^\times(t)$ implies the existence of an isometric mapping

$$\phi \longleftrightarrow \langle \phi^\times(t) \mid \phi \rangle \quad \text{for all } \phi \in \Phi,$$

with $\langle \phi^\times(t) \mid \phi \rangle \in L^2_\mu(sp(A))$. $\phi^\times(t)$ being an eigenfunctional of A, μ almost everywhere, entails

$$\hat{A}\langle \phi^\times(t) \mid \phi \rangle = \langle \phi^\times(t) \mid A\phi \rangle = t \langle \phi^\times(t) \mid \phi \rangle.$$

Note that this method is confined to the submanifold Φ! For an operator A with spectrum of multiplicity n we have

$$\phi \longleftrightarrow \{\langle \phi_1^\times(t) \mid \phi \rangle, \ldots, \langle \phi_n^\times(t) \mid \phi \rangle\} \in \bigoplus_1^n L^2_\mu.$$

Now we come to the problem of transforming one such representation (according to A) into a representation corresponding to another operator $B \in \mathcal{L}^c(\Phi)$. In the same context we shall be concerned with the representation of an operator $D \in \mathcal{L}^c(\Phi)$ within the given A-representation. Both problems are settled provided we are able to establish relations of the form

$$\langle \psi^\times(u) \mid \phi \rangle = \int_{sp(A)} k(u,t) \langle \phi^\times(t) \mid \phi \rangle \, d\mu(t)$$

and

$$\langle D^\times \phi^\times(t') \mid \phi \rangle = \int_{sp(A)} D(t',t) \langle \phi^\times(t) \mid \phi \rangle \, d\mu(t),$$

where we again assume A to possess a simple spectrum.

By these relations we are led to inquire about the possibility of establishing a general integral decomposition of certain elements of $\Phi^\times$ in the form

$$\psi^\times(u) = \int_{sp(A)} K(u,t)\, \phi^\times(t)\, d\mu(t)$$

and

$$D^\times \phi^\times(t') = \int_{sp(A)} D(t',t)\, \phi^\times(t)\, d\mu(t).$$

In what follows we shall investigate the possibility of decomposing the whole of $\Phi^\times$ with respect to the measure μ related to the direct Hilbert space decomposition with respect to a given operator $A \in \mathcal{L}^c(\Phi)$. This problem has already been investigated in a different context by Foiaş.[11] We recall that in this case Φ is a nuclear Fréchet space. Now we shall give the formal decomposition

$$\phi^\times = \int_{sp(A)} \phi^\times(t)\, d\mu(t) \quad \text{for } \phi^\times \in \Phi^\times$$

a precise mathematical meaning. Within this context we shall follow Bourbaki.[12]

Let μ be a positive Borel measure on R. For every function $\psi(t)$ with values in a l.c. space F we denote by $\langle z', \psi \rangle$ the numerical function $z' \circ \psi$ over R; $z' \in F'$. We shall say that ψ has numerically the property P if $\langle z', \psi \rangle$ has this property for every $z' \in F$. According to this general definition we shall say that a function $\psi(t) \in F$ is numerically μ-integrable if $\langle z', \psi \rangle$ is μ-integrable for every $z' \in F'$. We emphasize that according to this definition the integrability depends essentially on the dual pair $\langle F, F' \rangle$. Any topology compatible with the duality of $\langle F, F' \rangle$ leads to the integrability of ψ if ψ is integrable for the given topology in F. If ψ is numerically μ-integrable then the map

$$z' \to \int \langle z', \psi(t) \rangle \, d\mu(t)$$

is a linear form over F', i.e., an element of F'^*.

Definition 1: $\int \psi(t) d\mu(t)$ is called the integral of ψ with respect to μ where $\int \psi(t) d\mu(t)$ being an element of the algebraic dual space F'^* of F' is defined by

$$\langle z', \int \psi(t)\, d\mu(t) \rangle = \int \langle z', \psi(t) \rangle \, d\mu(t)$$

for every $z' \in F'$.

Now we shall investigate functions $\psi^\times(t)$ over R with values in $\Phi^\times$. We are mainly interested in the problem under what conditions the integral of a function $\psi^\times(t)$ is again an element of $\Phi^\times$. For that purpose let us consider those l.c. spaces which posses the so-called property GDF (*"du graphe denombrablemente fermé"*).

GDF: Let u be a linear map of the l.c. space F into a Banach space B. Then if the limit of any convergent sequence in the graph $\Gamma \subset F \times B$ also belongs to Γ it follows that u is continuous.

We note that every Frechet space has the property GDF.[13] For l.c. spaces that have the property GDF the following theorem is valid.

Gel'fand–Dunford: Let F be an l.c. Hausdorff space which has the property GDF and let F' be the dual space F endowed with the weak topology $\sigma(F', F)$. Then for each each function $\psi(t)$ over R with values in F' which is numerically integrable the integral $\int \psi(t) d\mu(t)$ belongs to F'.[12]

Let now A be an element of $\mathcal{L}^c(\Phi)$ which is an e.s.a. operator on Φ, i.e., induces a unique A-eigenintegral decomposition of Φ.

As in I, Sec. V, we define $\Phi^\times(t)$ to be the $\beta(\Phi^\times, \Phi)$-closure of $\gamma(t)\Phi$, where $\gamma(t)$ are the eigenoperators belonging to the integral A-eigendecomposition of Φ (cf. I, Sec. IV).

Definition 2: $\Phi^\times$ is said to have a unique μ-integral decomposition if every element $\phi^\times \in \Phi^\times$ can be represented in the form

$$\phi^\times = \int \psi^\times(t)\, d\mu(t),$$

with $\psi^\times(t)\in\Phi^\times(t)$ μ-almost everywhere and $\psi^\times(t)$ is unique up to a μ-null set. In what follows we shall investigate whether this definition is useful in the case of the nonrelativistic quantum mechanical rigged Hilbert space.

Recall that in this case all elements of $\Phi^\times$ are of the form

$$\phi^\times=\sum_{k=1}^{n(\phi^\times)} H^{\alpha_k}\circ\mathcal{G}_{\alpha_k}$$

according to Corollary 2 of Proposition 3 or even of the form $\phi^\times=H^n\circ\tilde{h}$ (cf. Sec. I).

Proposition 4: For every operator $A\in\mathcal{L}^c(\Phi)$ being e.s.a. on Φ and strongly commuting with H there is a unique μ-integral decomposition of $\Phi^\times$.

Proof: We define A and H to commute strongly if all spectral operators of H and A commute with each other.

Now let A be such an operator and let

$$\mathcal{G}\longrightarrow\hat{\mathcal{G}}=\int\hat{\mathcal{G}}(t)\,d\mu(t)$$

be a direct integral decomposition of $\mathcal{G}$ which originates from the spectral decomposition of A. Then for any $\phi^\times\in\Phi^\times$ we have

$$\langle\phi^\times|\phi\rangle=(\mathcal{G},H^n\phi)=\int\langle\hat{\mathcal{G}}(t)\,|\,I(t)H^n\phi\rangle_t\,d\mu(t)$$

$$=\int\langle I^\times(t)\hat{\mathcal{G}}(t)\,|\,H^n\phi\rangle\,d\mu(t)=\int\langle H^{n\times}I^\times(t)\hat{\mathcal{G}}(t)\,|\,\phi\rangle\,d\mu(t).$$

By I, Proposition 20 if follows $I^\times(t)\hat{\mathcal{G}}(t)\in\Phi^\times(t)$. We have to show that under the condition of strong commutativity of A and H, $H^{n\times}$ map $\Phi^\times(t)$ into itself. For that purpose we shall show that H^n and $\gamma(t)$ satisfying the following relation $H^{n\times}\gamma(t)\phi=\gamma(t)H^n\phi$ for all $\phi\in\Phi$ and μ-almost all $t\in R$. This relation means that $H^{n\times}$ maps $\gamma(t)\Phi$ into itself. But then $\overline{(\gamma(t)\Phi)}_\beta=\Phi^\times(t)$ being a closed subspace of the complete space $\Phi^\times_\beta$ is mapped continuously into itself by $H^{n\times}$.

Now the strong commutativity entails in particular $[H^n,E(\sigma)]=0$ for all $\sigma\in\mathcal{B}$ and $n\in N$, where $E(\sigma)$ are the spectral operators of A. Therefore we obtain

$$(H^nE(\sigma)I\phi,I\psi)=(E(\sigma)H^nI\phi,I\psi)=(E(\sigma)IH^n\phi,I\psi)$$

$$=\int_\sigma\langle\gamma(t)H^n\phi,\psi\rangle\,d\mu(t)=(E(\sigma)I\phi,H^nI\psi)=(E(\sigma)I\phi,IH^n\psi)$$

$$=\int_\sigma\langle\gamma(t)\phi\,|\,H^n\psi\rangle\,d\mu(t)=\int_\sigma\langle H^{n\times}\gamma(t)\phi\,|\,\psi\rangle\,d\mu(t).$$

Since these equations hold for every $\sigma\in\mathcal{B}$ (class of Borel sets on R), it follows that

$$\langle\gamma(t)H^n\phi\,|\,\psi\rangle=\langle H^{n\times}\gamma(t)\phi\,|\,\psi\rangle$$

for all $\phi,\ \psi\in\Phi$ and $t\in R$ but $t\in N(\phi,\psi)$ with $\mu(N(\phi,\psi))=0$.

Using the separability of Φ, we conclude by the usual reasoning that $\gamma(t)H^n=H^{n\times}\gamma(t)$ holds, μ almost everywhere. Hence we have obtained a representation of any $\phi^\times\in\Phi^\times$ in the form

$$\phi^\times=\int\psi^\times(t)\,d\mu(t)\quad\text{with }\psi^\times(t)\in\Phi^\times(t),$$

μ almost everywhere. According to Definition 1 this integral is well defined. Now we shall prove that this integral decomposition of $\phi^\times$ is unique. To this end, let $\varphi^\times(t)\in\Phi^\times$ be another function for which we have

$$\phi^\times=\int\varphi^\times(t)\,d\mu(t)$$

and $\varphi^\times(t)$ is an eigenfunctional of $A^\times$, μ almost everywhere. Then $h^\times(t)=\psi^\times(t)-\varphi^\times(t)$ is, μ almost everywhere, an eigenform of $A^\times$ and fulfills the relation $\int\langle h^\times(t)\,|\,\phi\rangle\,d\mu(t)=0$ for all $\phi\in\Phi$.

Because of $A\Phi\subset\Phi$ we have

$$0=\int\langle h^\times(t)\,|A\phi\rangle\,d\mu(t)=\int t\langle h^\times(t)\,|\,\phi\rangle\,d\mu(t),\quad\phi\in\Phi.$$

Now let us consider the following expressions[11]:

$$\int[(t+i)/(t-i)]^k h^\times(t)\,d\mu(t)=h_k^\times.$$

Since $\mathcal{U}^k(t)=[(t+i)/(t-i)]^k$ is a bounded function, we conclude that

$$\int\mathcal{U}^k(t)\langle h^\times(t)\,|\,\phi\rangle\,d\mu(t)$$

does exist for all $\phi\in\Phi$. Therefore, by the theorem of Gel'fand—Dunford it follows that $h_k^\times\in\Phi^\times$. We shall show that all $h_k^\times$ are equal to zero. We already have $h_0^\times=0$. Let then $k>0$. If we set $\psi_k=(A+i1)^k\phi$ and $\chi_k=(A-i1)^k\phi$, it follows (note that $\psi_k,\ \chi_k\in\Phi$) that

$$\langle h^k(t)\,|\,\psi_k\rangle=(t-i)^k\langle h^\times(t)\,|\,\phi\rangle=[(t-i)/(t+i)]^k\langle h^\times(t)\,|\,\psi_k\rangle$$

and therefore

$$\langle h_k^\times|\,\psi_k\rangle=\int[(t+i)/(t-i)]^k\langle h^\times(t)\,|\,\psi_k\rangle\,d\mu(t)$$

$$=\int\langle h^\times(t)\,|\,\psi_k\rangle\,d\mu(t)=0.$$

From this relation we derive

$$\langle h_k^\times|\,(A+i1)^k\phi\rangle=0\quad\text{for all }\phi\in\Phi.$$

For the moment we take it for granted that $A^\times$ does not have the eigenvalues $\pm i$ in $\Phi^\times$ whence we conclude $h_k^\times=0$ for all $k=1,\ 2,\ \ldots$. The same reasoning applies to $k=-1,\ -2,\ \ldots$. Now we choose a fixed element $\phi\in\Phi$. Then let us consider the measure

$$\nu(\sigma)=\int_{u(\sigma)}\langle h^\times(t)\,|\,\phi\rangle\,d\mu(t)\quad\text{on }\{\lambda\,|\,|\lambda|=1\},$$

with $u(t)=(t+i)/(t-i)$. We have

$$0=\langle h_k^\times|\,\phi\rangle=\int_{\{\lambda\,|\,|\lambda|=1\}}\lambda^{-k}\,d\nu(\lambda),\quad k=0,\ \pm1,\ \pm2,\ldots.$$

Since the trigonometric polynomials $\Sigma C_k\exp(ik\theta)$ are dense in $C([0,2\pi])$ we arrive at the conclusion $\nu=0$. This means that $\langle h^\times(t)\,|\,\phi\rangle=0$ up to a μ-null set $N(\phi)$. Let now again Φ_0 be a countable dense set in Φ. Then for all $\phi\in\Phi_0$ and all $t\notin N=\bigcup_{\phi\in\Phi_0}N(\phi)$, it follows that $\langle h^\times(t)\,|\,\phi\rangle=0$ whence by continuity arguments we obtain $h^\times(t)=0$, μ almost everywhere. To conclude the proof, we shall show that $A^\times$ cannot have the eigenvalues $\pm i$ in $\Phi^\times$. As it has been shown at the end of Sec. II all elements of $\Phi^\times$ are of the form $\phi^\times=H^{2m}\circ h$ with $h\in D(\overline{H^m})$. Now assume that a relation of the form (note that $[H^m,A]=0$ on Φ)

$$\langle A^\times I^\times H^m h\,|\,H^m\phi\rangle=(H^m h,AH^m\phi)=(h,AH^{2m}\phi)=i(h,H^{2m}\phi)$$

would hold for all $\phi\in\Phi$. Since all eigenvectors ϕ_k of H are contained in Φ, we would have $k^{2m}(h,A\phi_k)=ik^{2m}(h,\phi k)$ or $a_k(h,\phi_k)=(h,A\phi_k)=i(h,\phi_k)$ for all $k=1,\ 2,\ \ldots$, where a_k are real eigenvalues of A. Since $\{\phi_k,\ k=1,2,\ldots\}$ is dense in $\mathcal{G}$, we conclude $h=0$ which completes the proof. QED

In the following proposition we shall essentially show that $\Phi^\times$ is separable with respect to any topology compatible with the duality of $\langle\Phi^\times,\Phi\rangle$. In a quite formal sense this proposition will also provide an answer to the question whether there exists an integral decomposition of $\Phi^\times$ with respect to any operator A being an element of $\mathcal{L}^c(\Phi)$ and e.s.a. on Φ.

Proposition 5 (c.f. Ref. 11): Let

$$H=\sum_{i=1}^{3} P_i^2+Q_i^2+1$$

and let ϕ_n denote the eigenvectors of H. Then $\{\gamma\phi_n, n\in N\}$ is dense in $\Phi^\times$ with respect to any topology compatible with the duality of $\langle\Phi^\times,\Phi\rangle$. We have $\Phi^\times=\Sigma_n\langle\phi^\times|\phi_n\rangle\gamma\phi_n$ for every $\phi^\times\in\Phi^\times$.

Proof: We know from the last proposition that H induces a unique integral decomposition of $\Phi^\times$. Hence every $\phi^\times\in\Phi^\times$, in particular, every eigenform $\phi^\times_{t'}$ with t' real (we know already that $H^\times$ cannot have any complex eigenvalue in $\phi^\times$) of H can be represented in the form

$$\phi^\times_{t'}=\int\phi^\times_{t'}(t)\,d\mu(t)\quad\text{with }\phi^\times_{t'}(t)\in\Phi^\times(t).$$

Hence it follows that

$$t\langle\phi^\times_{t'}|\phi\rangle=\langle H^\times\phi^\times_{t'}|\phi\rangle=\langle\phi^\times_{t'}|H\phi\rangle$$

$$=\int\langle H^\times\phi^\times_{t'}(t)|\phi\rangle\,d\mu(t)=\int t\langle\phi^\times_{t'}(t)|\phi\rangle\,d\mu(t)\quad\text{for all }\phi\in\Phi.$$

This relation yields $\int(t'-t)\,\phi^\times_{t'}(t)\,d\mu(t)=0$. Because of the uniqueness of $\phi^\times_{t'}(t)$ it follows $(t'-t)\,\phi^\times_{t'}(t)=0$, μ almost everywhere. We obtain $\phi^\times_{t'}(t)=0$ except for $t=t'$, whence

$$\phi^\times_{t'}=\phi^\times_{t'}(t)\cdot\mu(\{t'\}).$$

Since μ is a discrete measure, we conclude that this t is a discrete eigenvalue of H. This fact entails that $\Phi^\times_{t_n}=\Phi^\times(t_n)$ (cf. I, Sec. V) for all discrete eigenvalues of H. Then by I, Proposition 23 we obtain that $I^\times\mathcal{G}\cap\Phi^\times(t_n)$ is equal to $I^\times P_{t_n}\mathcal{G}$, where $P_{t_n}\mathcal{G}$ is the whole eigenspace for the eigenvalue t_n. Furthermore, by the discreteness of μ we conclude that each element $\phi^\times\in\Phi^\times$ is fo the form $\phi^\times=\Sigma_n\phi^\times(t_n)\,\mu(\{t_n\})$, where this series converges in the weak topology $\sigma(\Phi^\times,\Phi)$.

Now since $I^\times$ is an injective map there exists for each n a unique element $h_n^{\phi^\times}\in P_{t_n}\mathcal{G}$ such that $\phi^\times(t_n)=I^\times h_n^{\phi^\times}$ holds. However, $P_{t_n}\mathcal{G}$ possess at most finitely many linear independent elements, say h_{nj}; $j=1, 2, \ldots, m$ whence

$$h_n^{\phi^\times}=\sum_j(h_n^{\phi^\times},h_{nj})\,h_{nj}.$$

Therefore, we have

$$\phi^\times=\sum_{n,j}(h_n^{\phi^\times},h_{nj})\,I^\times h_{nj}\,\mu(\{t_n\}).$$

Furthermore, we know that all eigenvectors of H are contained in Φ whence we have $h_{nj}=I\phi_{nj}$.

Thus $\phi^\times$ is of the form

$$\phi^\times=\sum_{n,j}(h_n^{\phi^\times},h_{nj})I^\times\circ I\phi_{nj}\,\mu(\{t_n\})=\sum_{nj}\langle\phi^\times,\phi_{nj}\rangle\gamma\,\phi_{nj}.$$

This leads to the conclusion that the set $\{\gamma\phi_{nj}, n, j\in N\}$ is dense in $\Phi^\times$ in the weak topology $\sigma(\Phi^\times,\Phi)$ and hence in any topology compatible with the duality of $\langle\Phi^\times,\Phi\rangle$. QED

Corollary 1: $\Phi^\times_\beta$ being a complete nuclear space, is separable (Note that $\Phi^\times_\beta$ is not metrizable!). For the proof of this assertion we remark that the strong topology is compatible with the duality of $\langle\Phi^\times,\Phi\rangle$.

Then the assertion follows by I, Proposition 6. By definition[14] a sequence $\{x_i\}$ in a topological vector space E is called a basis in E if for every $x\in E$ there exists a unique sequence $\{\alpha_i\}$ of complex numbers such that

$$x=\lim_{n\to\infty}\sum_{i\le n}\alpha_i x_i$$

holds with respect to the topology of E. Then $x\to\alpha_i$ defines linear functionals over E. A sequence $\{x_i\}$ in a topological vector space E is called a Schauder basis if $\{x_i\}$ is a basis and if the coefficient functionals f_i, $f_i(x)=\alpha_i$ are continuous linear functionals on E. We shall make use of these definitions by assuming the coefficient functionals f_i either being linear or antilinear continuous functionals on the space E in question. These definitions enable us to state the following:

Corollary 2: Let $\{\phi_n\}$ denote the set of eigenvectors of the operator H. Then we have (I) $\{\phi_n\}$ is a Schauder basis in Φ; (II) $\{\gamma\phi_n\}$ is a Schauder basis in $\Phi^\times_\beta$; (III) For every $\phi^\times\in\Phi^\times$, $\{\langle\phi^\times|\phi_n\rangle\gamma\phi_n\}$ is absolutely convergent to $\phi^\times$ with respect to $\beta(\Phi^\times,\Phi)$.

Proof: (I) By the completeness of the system $\{\phi_n\}$ in the Hilbert space $\mathcal{G}$, for each $\phi\in\Phi$ we have the unique decomposition

$$\phi=\sum_{n=1}^{\infty}(\phi_n,\phi)\,\phi_n.$$

In order to prove that

$$\sum_{n=1}^{m}(\phi_n,\phi)\,\phi_n$$

converges to ϕ with respect to the topology of Φ we have only to show (Ref. 15, p. 120) (note that Φ is complete!) that the series

$$\sum_{n=1}^{\infty}|(\phi_n,\phi)|\,\|H^k\phi_n\|$$

are convergent for all $k=0, 1, 2, \cdots$. This is an immediate consequence of the equation

$$\sum_{n=1}^{\infty}|(\phi_n,\phi)|\,\|H^k\phi_n\|=\sum_{n=1}^{\infty}\frac{|(H^{k+2}\phi,\phi_n)|}{n^{k+2}}\,n^k\|\phi_n\|.$$

$\phi=\sum_{n=1}^{\infty}(\phi_n,\phi)\phi_n$ can be rewritten as

$$\phi=\sum_{n=1}^{\infty}\langle\gamma\phi_n,\phi\rangle\phi_n.$$

$\gamma\phi_n$ are continuous antilinear functionals on Φ whence $\{\phi_n\}$ is a Schauder basis for Φ.

(II) By Proposition 5 the decomposition

$$\phi^\times=\sum_{n=1}^{\infty}\langle\phi^\times|\phi_n\rangle\gamma\phi_n$$

is unique. Moreover, the ϕ_n, $n=1, 2, \ldots$, are continuous linear functionals on $\Phi^\times_\beta$ whence $\{\gamma\phi_n\}$ is a Schauder basis in $\Phi^\times_\beta$.

(III) We have to show that for each continuous seminorm $p_B(\cdot)$ on $\Phi^\times_\beta$ (B any bounded set in Φ) it follows that

$$\sum_{n=1}|\langle\phi^\times|\phi_n\rangle|\,P_B(\gamma\phi_n)<\infty.$$

We have

$$\sum_{n=1}^{\infty} |\langle\phi^\times|\phi_n\rangle| P_B(\gamma\phi_n) = \sum_{n=1}^{\infty} |\langle\phi^\times|\phi_n\rangle| \sup_{\phi\in B} |\langle\gamma\phi_n, \phi\rangle|.$$

Now each element of $\Phi^\times$ is of the form $\phi^\times = H^{2k} \circ h$. Therefore we get

$$\sum_{n=1}^{\infty} |(h, H^{2k}\phi_n)| \sup_{\phi\in B} |(\phi_n, \phi) = \sum_{n=1}^{\infty} \frac{|(h, \phi_n)}{n^2}$$

$$\times \sup_{\phi\in B} |(\phi_n, H^{2k+2}\phi)| \le \sum_{n=1}^{\infty} (1/n^2)\|h\| \sup_{\phi\in B}\| H^{2k+2}\phi\|.$$

Since $\|H^{2k+2}\|$ is a continuous seminorm on Φ and B is bounded in Φ, there exists a $\lambda > 0$ such that

$$\sup_{\phi\in B} \| H^{2k+2}\phi\| < \lambda. \qquad \text{QED}$$

At the end of this section we shall discuss the results obtained so far in order to see if they permit the construction of a rigorous formalism for the transformation theory in nonrelativistic quantum mechanics. To this end let $A \in \mathcal{L}^c(\Phi)$ be any operator which is e.s.a. on Φ. Then, Φ has a unique A-eigenintegral decomposition (cf. I, Sec. IV) so that we can write

$$\gamma\phi = \int \gamma(t)\phi\, d\mu(t) \quad \text{for all } \phi \in \Phi.$$

By the theorem of Gelfand—Dunford the right-hand side of the last equation is well-defined.

Then by Proposition 5 we can write

$$\phi^\times = \sum_{n=1}^{\infty} \langle\phi^\times|\phi_n\rangle\gamma\phi_n = \sum_{n=1}^{\infty} \langle\phi^\times|\phi_n\rangle \int \gamma(t)\,\phi_n\, d\mu(t),$$

for every $\phi^\times \in \Phi^\times$, where ϕ_n are the eigenvectors of H. Now if $B \in \mathcal{L}^c(\Phi)$ is any other operator e.s.a. on Φ and $\psi^\times(u) \in \Phi^\times(u)$ is an eigenform originating from the B-eigenintegral decomposition (cf. I, Proposition 17 or 21) then we have

$$\langle\psi^\times(u)|\phi\rangle = \sum_n \langle\psi^\times(u)|\phi_n\rangle \int \langle\gamma(t)\,\phi_n|\phi\rangle\, d\mu(t) \tag{1}$$

for all $\phi \in \Phi$ and the right-hand side of (1) is absolutely convergent for every $\phi \in \Phi$. In the same way for any operator $D \in \mathcal{L}^c(\Phi)$ we have

$$\langle D^\times\phi^\times(t')|\phi\rangle = \sum_n \langle D^\times\phi^\times(t')|\phi_n\rangle \int \langle\gamma(t)\,\phi_n|\phi\rangle\, d\mu(t), \tag{2}$$

for every $\phi \in \Phi$ and again the right-hand side is absolutely convergent for every $\phi \in \Phi$.

We remark in passing that formulas (1) and (2) allow us to deal with any multiplicity of the spectra of the operators A, B, and D in question.

Now, for the sake of simplicity, let us assume that the operators A and B have a simple spectrum. Then according to I, Proposition 17 we can rewrite (1) and (2) in the form

$$\langle\psi^\times(u)|\phi\rangle = \sum_n \langle\psi^\times(u)|\phi_n\rangle \int \overline{\langle\phi^\times(t)|\phi_n\rangle}\langle\phi^\times(t)|\phi\rangle\, d\mu(t) \tag{3}$$

and

$$\langle D^\times\phi^\times(t')|\phi\rangle = \sum_n \langle D^\times\phi^\times(t')|\phi_n\rangle \int \overline{\langle\phi^\times(t)|\phi_n\rangle}\langle\phi^\times(t)|\phi\rangle\, d\mu(t), \tag{4}$$

for all $\phi \in \Phi$. These formulas give us the most general possibility to deal with the transformation theory in nonrelativistic quantum mechanics. These formulas would be essentially simplified if we could rewrite them in the form

$$\langle\psi^\times(u)|\phi\rangle = \int K(u,t)\langle\phi^\times(t)|\phi\rangle\, du(t)$$

and

$$\langle D^\times\phi^\times(t')|\phi\rangle = \int D(t',t)\,\langle\phi^\times(t)|\phi\rangle\, d\mu(t)$$

with mathematically well-defined kernels

$$K(u,t) = \sum_n \langle\psi^\times(u)|\phi_n\rangle\overline{\langle\phi^\times(t)|\phi_n\rangle}$$

and

$$D(t',t) = \sum_n \langle D^\times\phi^\times(t')|\phi_n\rangle\overline{\langle\phi^\times(t)|\phi_n\rangle},$$

respectively. In this formulation we would in fact obtain an integral representation of the operator D. For the general case these equations can not be deduced in our scheme. It would be interesting to see what conditions must be fulfilled by the operators A, B, and D (especially, what spectral properties A, B, and D must posses) in order that the transformation equations (3) and (4) can be written in an integral form.

Finally let us consider the character of the mapping

$$\mathcal{J}_u(\langle\phi^\times(t)|\phi\rangle) = :\sum_n\langle\psi^\times(u)|\phi_n\rangle \int \overline{\langle\phi^\times(t)|\phi_n\rangle}\langle\phi^\times(t)|\phi\rangle\, d\mu(t).$$

We denote by Φ_A the image of Φ under the isometric mapping $U: \Phi \longrightarrow \langle\phi^\times(t)|\phi\rangle$. Φ_A is a dense linear subspace of the space $\mathcal{L}^2_\mu$.

Φ_A is a l.c. space and the l.c. topology is given by the the following family of seminorms:

$$\{\int |\langle\phi^\times(t)|H^n\phi\rangle|^2 d\mu(t)\}^{1/2}, \quad n=0, 1, 2, \ldots .$$

Now, for all $u \in R$, $\mathcal{J}_u$ is a distribution over Φ_A as can easily be checked. The same assertion is true for the mapping

$$\mathcal{D}_{t'}(\langle\phi^\times(t)|\phi\rangle) = :\sum_n \langle D^\times\phi^\times(t')|\phi_n\rangle \int \overline{\langle\phi^\times(t)|\phi_n\rangle}\langle\phi^\times(t)|\phi\rangle\, d\mu(t).$$

IV. FINAL REMARKS

The usefulness of the rigged Hilbert-space $\Phi \subset \mathcal{G} \subset \Phi^\times$ which we have treated in Secs. II and III depends substantially on whether all s.a. operators which are of physical importance are already e.s.a. on the basic domain Φ. At present no solution of this problem is known. Up to now one has not been able to prove that all elements of the enveloping algebra $\mathcal{A}(P_i, Q_i)$, $i = 1, 2, 3$, are e.s.a. operators on Φ. In Ref. 16 it has been proved that all polynomials of second degree in the operators P_i, Q_i, $i = 1, 2, 3$, are e.s.a. on Φ. However, since the property of essential self-adjointness of an s.a. operator is preserved under unitary transformations we can pass to the Q-representation in which we essentially have to deal with differential operators. In Ref. 17 it has been shown that all differential operators that occur in physical applications are e.s.a. on the domain $\Phi \cong S \subset L^2(-\infty, \infty)$. As already explained at the end of I, Sec. IV, it is not necessary that all observables map Φ continuously into itself. The method described at the end of I, Sec. IV enables one to weaken the requirements on the order of differentiability both of the coefficients of a differential operator and the potential function.

Another question which arises in connection with the use of the formalism of a rigged Hilbert space concerns the following fact: In this new formalism only the vectors of Φ have to be used in practical calculations. Then for the practicability of this concept it is very desirable that all eigenvectors and eigenpackets[17] are already contained in Φ. It seems that this requirement cannot be fulfilled in general. Thus in order to make the rigged Hilbert space formalism more handy for practical applications one has to enlarge the basic domain Φ. Obviously this enlarged domain Φ no longer has the property of being invariant under all elements of $\mathcal{A}(P_i, Q_i)$. We shall refrain here from explaining the details of a formalism which is based on an enlarged domain Φ since we have not yet obtained definite results in this direction.

To ensure the full applicability of the rigged Hilbert space formalism some additional problems would have to be solved. First of all for e.s.a. operator A in $\mathcal{G}$ it would be desirable to be able to solve the eigenvalue equation of the map $A^\times$ in $\Phi^\times$ and in this way to obtain the spectrum of the operator A in $\mathcal{G}$. Following Babbitt[18] we may define $\mathrm{sp}(A^\times)$ to be the closure of the set of all eigenvalues of the mapping $A^\times$ in $\Phi^\times$ (cf. also Ref. 19). Then the question arises whether $\mathrm{sp}(A)=\mathrm{sp}(A^\times)$. A necessary and sufficient condition for this equality is that $(A+\lambda 1)\Phi$ is dense in Φ for all $\lambda \notin \mathrm{sp}(A)$.[18] It would be very desirable to show that any operator $A \in \mathcal{L}^c(\Phi)$ which is e.s.a. on Φ already fulfils this condition.

Even if this problem could be settled there is still another problem which, in our opinion, is much more relevant. In I, Sec. V, we have shown that all eigenfunctionals (for an eigenvalue t) of an operator A that belong to the generalized eigenfunction decomposition of A have to be sought in the space $\Phi^\times(t)$. However, as has been pointed out at the end of I, Sec. V, it is by no means clear whether $\Phi^\times(t)=\Phi_t^\times$ (cf. I, Sec. V) for μ-almost all $t \in \mathrm{sp}(A)$. In the same way we do not know of any condition which would lead to this equality.

Despite all these open problems and even of some drawbacks inherent in the rigged Hilbert space formalism we should like to advocate this formalism as a natural enrichment of the usual Hilbert space formalism for quantum systems since it makes the whole formalism more transparent and allows one to deal with mathematical problems which cannot be adquately treated within the pure Hilbert space formalism.

Finally we may ask whether in other realizations of rigged Hilbert spaces $\Phi \subset \mathcal{G} \subset \Phi^\times$ for nonrelativistic quantum mechanics the whole space $\Phi^\times$ is really needed for the formulation of the transformation theory. It is obvious that the argument in III is based on very special structural properties of the spaces Φ and $\Phi^\times$ which belong to the Robert's triplet. In a more general situation it would already be of great help if one could confine oneself to a subspace of $\Phi^\times$ for which a similar decomposition as in III should be derived.

A subspace possessing some of the necessary fundamental properties (let us denote it by $\Psi^\times$) could, e.g., be constructed as follows: suppose Φ is a metrizable separable l.c. space with a nuclear embedding I of Φ into the Hilbert space $\mathcal{G}$. If we denote by $\Psi^\times$ the $\beta(\Phi^\times, \Phi)$-closure of $\gamma\Phi = I^\times \cdot I\,\Phi$ we have

Proposition 6: $\Psi^\times$ has the following properties:

(I) For every A-eigenintegral decomposition $\{\gamma(t), R, \mu\}$ of Φ with $A \in \mathcal{L}^c(\Phi)$ and A e.s.a. on Φ it follows $\Phi^\times(t) \subset \Psi^\times$ μ almost everywhere.

(II) For every $B \in \mathcal{L}^c(\Phi)$, $B^\times$ is a $\beta(\Phi^\times, \Phi)$-continuous mapping of $\Psi^\times$ into itself.

Proof: First, we shall show $\gamma(t)\Phi \subset \Psi^\times$ μ almost everywhere for any operator $A \in \mathcal{L}^c(\Phi)$ which is e.s.a. on Φ. To this end let us consider the norm $\|\cdot\|'$ on the subspace $I^\times \mathcal{G}$ of $\Phi^\times$ defined by

$$\|I^\times h\|' = \|h\| \quad \text{for every } h \in \mathcal{G}.$$

Then $\beta(\Phi^\times, \Phi)$ restricted to $I^\times \mathcal{G}$ is coarser than $\|\cdot\|'$. Indeed, for every continuous seminorm $P_B(\cdot)$ of $\beta(\Phi^\times, \Phi)$ with B as a bounded set in Φ we have

$$P_B(I^\times h) = \sup\{|\langle I^\times h \mid \phi\rangle|;\ \phi \in B\} = \sup\{|(h, I\phi)|;\ \phi \in B\}$$

$$\leq \sup\{\|h\| \cdot \|\phi\|;\ \phi \in B\} \leq \lambda_B \|h\| = \lambda_B \|h\| = \lambda_B \|I^\times h\|'$$

since B, being a bounded set, can be absorbed into the o-neighborhood $\|\phi\| \leq 1$.

Since $\gamma\Phi$ is dense in $I^\times \mathcal{G}$ with respect to $\|\cdot\|'$ it follows that $\Psi^\times$ contains $I^\times \mathcal{G}$. Therefore $\Psi^\times$ contains all elements of the form $I^\times E(\sigma) I\phi$ with $\phi \in \Phi$ and $E(\sigma)(\sigma \in \mathcal{B})$ being a spectral operator of A. This entails (cf. I, Sec. V) $\gamma(t)\phi \in \Psi^\times$, μ almost everywhere. We conclude $\gamma(t)\Phi \subset \Psi^\times$ and therefore $\overline{(\gamma(t)\Phi)}_\beta = \Phi^\times(t) \subset \Psi^\times$. For the proof of (II) we note that $\Phi_\beta^\times$ is complete. Furthermore, for every $B \in \mathcal{L}^c(\Phi)$ we have

$$B^{c\times}\gamma = \gamma B \quad \text{(cf. I, Sec. IV)}.$$

This relation shows that $B^{c\times}$ maps $\gamma\Phi$ into itself. But $\Psi^\times$ being a closed subspace of a complete space is itself complete. Therefore $B^{c\times}$ can be continuously extended to a mapping of $\Psi^\times$ into itself. QED

The existence of a suitable topology of Φ permitting a decomposition of the elements of $\Psi^\times$ should, of course, still be investigated.

To comment upon the last proposition we remark that the subspace $\Psi^\times \subset \Phi^\times$ is large enough to contain all eigenforms of any operator $A \in \mathcal{L}^c(\Phi)$ which is e.s.a. on Φ and to enable the transformation of one representation into another and the representation of any operator $D \in \mathcal{L}^c(\Phi)$ in a given representation (cf. I, Sec. V).

For a nuclear Frechét space the assertion of the last proposition is trivial. Φ being a nuclear Frechét space is a reflexive space. In this case the strong topology $\beta(\Phi^\times, \Phi)$ is compatible with the duality of $\Phi^\times$ and Φ. Thus $\Psi^\times = \Phi^\times$.

ACKNOWLEDGMENTS

The author would like to thank Professor G. Ludwig for his generous encouragement and Dr. A. Hartkämper, Dr. H. Neumann, and Dr. F. Jenč for stimulating discussions and useful comments.

[1]J.E. Roberts, Commun. Math. Phys. 3, 98 (1966).
[2]A. Böhm, Lect. Theor. Phys. 9A, 255 (1967);
[3]J.-P. Antoine, J. Math. Phys. 10, 53 (1969).
[4]O. Melsheimer, J. Math. Phys. 15, 902 (1974).
[5]L. Gårding, Proc. Natl. Acad. Sci. USA 33, 331 (1947).
[6]E. Nelson, Ann. Math. 70, 572 (1959).
[7]J. Voisin, J. Math. Phys. 10, 6 (1965).
[8]J.R.H. Brennich, Ann. Inst. H. Poincaré 13 (2), 137 (1970).
[9]J. Dixmier, Comput. Math. 13, 263 (1958).
[10]E. Nelson and F. Stinespring, Am. J. Math. 81, 547 (1959).
[11]C. Foiaş, Rev. Math. Pure Appl. 7, 241 571 (1962).
[12]N. Bourbaki, *Intégration vectorielle* (Hermann, Paris, 1959), fasc. XXV, livre IV, Chap. 6.
[13]N. Bourbaki, *Espaces vectoriels topologiques* (Hermann, Paris, 1964), fasc. XVIII, livre V, Chaps. 3, 4, and 5.
[14]J.T. Marti, *Introduction to the Theory of Bases*, Springer Tracts in Natural Philosophy, Vol. 18 (Springer-Verlag, Berlin, 1969).
[15]H. Schäfer, *Topological Vector Spaces* (MacMillan, New York, 1967).
[16]H.D. Doebner and O. Melsheimer, J. Math. Phys. 9, 1638 (1968).
[17]G. Hellwig, *Differentialoperatoren der mathematischen Physik* (Springer-Verlag, Berlin, 1964).
[18]D. Babbit, Rept. Math. Phys. 3, 1 (1972).
[19]W.G. Faris, Helv. Phys. Acta 44 (7), 930 (1971).

On the inverse problem in radiative transfer

N. J. McCormick* and I. Kuščer

Department of Physics, University of Ljubljana, Ljubljana, Yugoslavia
(Received 2 February 1974)

The inverse problem in monochromatic radiative transfer is considered for an infinite medium with anisotropic scattering. It is shown that each Legendre coefficient of the scattering function can be related independently of the others to an appropriate integral over space and angles of the intensity due to a monodirectional plane source. This result offers some advantage over the analogous one for an isotropic plane source if the medium is weakly absorbing.

I. INTRODUCTION

In a recent paper, Case has shown how moments of the infinite-medium, azimuth-independent Green's function for monochromatic radiative transfer can be used to extract the expansion coefficients of the scattering function.[1] To obtain the Nth coefficient, it is required to know all lower-order ones plus the $2N$th spatial moment of the angle-integrated intensity due to an isotropic plane source. A systematic technique to determine more of these coefficients is discussed.

The primary purpose of this note is to show that a simplified method enables each coefficient to be related independently of the others to the spatial integral of the corresponding azimuthal Fourier component of the radiation field due to a monodirectional plane source. The importance of azimuthal components for determining the scattering function from measurements of the intensity has also been noted by Pahor in a related context.[2]

II. MONODIRECTIONAL PLANE SOURCE IN AN INFINITE MEDIUM

If the intensity depends only upon one coordinate (τ), on the cosine of the polar angle with respect to the positive τ axis (μ), and on the azimuth (ϕ), the equation of transfer in the absence of a source is[3]

$$\left(\mu\frac{\partial}{\partial\tau}+1\right) I(\tau,\mu,\phi) = \frac{1}{4\pi}\int_{-1}^{1} d\mu' \int_0^{2\pi} d\phi'\, p(\cos\delta)\, I(\tau,\mu',\phi'). \tag{1}$$

Anisotropic scattering of arbitrary but finite order will be admitted, which shall mean that

$$p(\cos\delta)=\sum_{l=0}^{L}\varpi_l P_l(\cos\delta). \tag{2}$$

We assume that some absorption is present, hence $0<\varpi_0<1$, in order that the infinite-medium Green's function be uniquely defined. For a source of unit magnitude, the definition is given by the conditions that $I(\tau,\mu,\phi)$ stays bounded at $\tau\to\pm\infty$ and that

$$I(0^+,\mu,\phi)-I(0^-,\mu,\phi)=\mu_0^{-1}\delta(\mu-\mu_0)\,\delta(\phi), \quad -1\le\mu\le 1. \tag{3}$$

Multiplying both sides of (1) by $(1-\mu^2)^{m/2}\cos(m\phi)\,d\mu\,d\phi$ and integrating, we derive the following identity:

$$\frac{d}{d\tau}\int_{-1}^{1} d\mu \int_0^{2\pi} d\phi\, I(\tau,\mu,\phi)\,\mu(1-\mu^2)^{m/2}\cos(m\phi) + \frac{h_m}{2m+1}J_m(\tau)=0, \tag{4}$$

where $h_m=2m+1-\varpi_m$, and

$$J_m(\tau)=\int_{-1}^{1} d\mu \int_0^{2\pi} d\phi\, I(\tau,\mu,\phi)(1-\mu^2)^{m/2}\cos(m\phi). \tag{5}$$

We now integrate both sides of Eq. (4) over τ, taking account of the discontinuity of I at $\tau=0$. The result is

$$\frac{h_m}{2m+1}\int_{-\infty}^{\infty} J_m(\tau)\,d\tau=(1-\mu_0^{\,2})^{m/2}. \tag{6}$$

In this way the spatial integral of the mth azimuthal Fourier component of the radiation field is directly related to the mth Legendre coefficient of the scattering function.

Another way to derive the above result is to express the solution in terms of singular eigenfunctions,[4–6] and to invoke the closure relation.[5]

III. ISOTROPIC PLANE SOURCE IN AN INFINITE MEDIUM

For the isotropic source of unit magnitude, the function which must be considered is

$$K_{l,n}=2\pi\int_{-\infty}^{\infty} d\tau\,\tau^n \int_{-1}^{1} d\mu\, I(\tau,\mu)\,P_l(\mu), \tag{7}$$

for which we derive the identity

$$2\pi(2l+1)\int_{-\infty}^{\infty} d\tau\,\tau^n\frac{d}{d\tau}\int_{-1}^{1} d\mu\, I(\tau,\mu)\,\mu P_l(\mu)+h_l K_{l,n}=0. \tag{8}$$

From symmetry considerations[7,8] it follows that $K_{l,n}=0$ for $(n+l)$ odd and for $n<l$, while use in Eq. (8) of the recursion relation for Legendre polynomials, followed by an integration by parts, gives[7]

$$(l+1)K_{l+1,n-1}+K_{l-1,n-1}-(h_l/n)K_{l,n}=0, \quad n\ge 1. \tag{9}$$

The starting equation which accompanies this set of equations,

$$K_{0,0}=1/h_0, \tag{10}$$

is found from Eq. (8) and the discontinuity at $\tau=0$.

Equation (9) gives a closed set of $2N+N(N-1)/2$ equations[8] for determining $K_{0,2N}$ for $N\ge 1$. The results of Case (with a correction for $K_{0,2}$) are

$$K_{0,2}=2/(h_0^2 h_1), \tag{11}$$

$$K_{0,4}=(24/h_0^3 h_1^2)(1+4h_0/h_2), \tag{12}$$

and we illustrate the increasing complexity of the moments with the result

$$K_{0,6}=\frac{720}{h_0^4 h_1^3}\left(1+\frac{4h_0}{h_2}\right)^2+\frac{25,920}{h_0^2 h_1^2 h_2^2 h_3}. \tag{13}$$

For weak absorption, when $h_0 \ll 1$, the leading contribution to $K_{0,2N}$ comes from the dominant mode of the radiation field. In the notation of Refs. 4 and 6, this mode is

$$\begin{aligned}2\pi\int_{-1}^{1} I(\tau,\mu)\,d\mu &\approx [2N(\nu_1)]^{-1}\exp(-|\tau|/\nu_1)\\ &=[\nu_1^2 g(\nu_1,\nu_1)\Lambda'(\nu_1)]^{-1}\exp(-|\tau|/\nu_1)\\ &\approx \tfrac{1}{2}h_1\nu_1\exp(-|\tau|/\nu_1),\end{aligned} \tag{14}$$

with $\nu_1^2 \approx 1/h_0 h_1$. Hence

$$K_{0,2N}\approx(2N)!\,h_1\nu_1^{2N+2}\approx(2N)!/(h_0^{N+1}h_1^N). \tag{15}$$

Thus we see that in an evaluation for a weakly absorbing medium the moments $K_{0,2N}$ could hardly reveal much more than the first two coefficients ϖ_0 and ϖ_1 that approximately determine the dominant mode. At larger distances that are relevant for the higher moments, everything else is drowned out by this one mode. Relation (6) offers some advantage in this respect, because the zeroth spatial moment of the higher azimuthal Fourier components might be easier to disentangle.

IV. COMMENTS

While it is encouraging to find simple relations between moments of the radiation field and the scattering coefficients, it would be premature to claim practical applicability. Neither Eq. (6) nor (10) to (13) can be related directly to experiments, since extended plane sources are not available. However, the isotropic plane source considered by Case can be substituted by an isotropic point source, if the point-to-plane transformation[7] is applied. Likewise, the monodirectional plane source and discrete directional detector implied in Sec. 2 may be replaced by a directional point source and a plane distribution of detectors. With proper normalization the responses in both cases are equal, according to the reciprocity principle.[9]

Even with such a modification experiments would not appear worth the effort. One would have to place a directional source (a searchlight) and a strategically distributed set of directional detectors (sensors with little telescopes) inside the medium, say in terrestrial fog, many mean free paths away from boundaries. Such a clumsy method could not be expected to compete with single-scattering experiments carried out with a laser beam.

Clearly spatial integrals of the type derived here are of little use in remote sensing of the terrestrial and planetary atmospheres, where one would need relations between the scattering coefficients and the reflected radiation field. There seems to be little hope for finding exact relations of this kind, so that indirect methods are called for. These are either based upon models with a few adjustable parameters,[10] or upon iterative procedures.[2] Perhaps such procedures can be improved with correction techniques based upon Eq. (6) or (10)–(13).

ACKNOWLEDGMENTS

One of the authors (NJM) expresses his gratitude for the hospitality of the Slovene Academy of Sciences and Arts and the University of Ljubljana under the exchange program of the U.S. National Academy of Sciences and the Council of the Academies of Yugoslavia.

*Present address: Department of Nuclear Engineering, University of Washington, Seattle, Washington 98195.

[1]K. M. Case, Phys. Fluids **16**, 1607 (1973); K. M. Case and S. C. Chiu, J. Math. Phys. **14**, 1643 (1973).
[2]S. Pahor, Phys. Rev. **175**, 218 (1968).
[3]S. Chandrasekhar, *Radiative Transfer* (Oxford U. P., London and New York, 1950; Dover, New York, 1960).
[4]N. J. McCormick and I. Kuščer, J. Math. Phys. **7**, 2036 (1966).
[5]N. J. McCormick and I. Kuščer, in *Advances in Nuclear Science and Technology*, edited by E. J. Henley and J. Lewins (Academic, New York, 1973), Vol. 7, p. 181.
[6]I. Kuščer and N. J. McCormick, in *Radiation and Remote Probing of the Atmosphere* (to be published).
[7]K. M. Case and P. F. Zweifel, *Linear Transport Theory* (Addison-Wesley, Reading, Mass., 1967), pp. 101–04.
[8]E. Canfield, Nucl. Sci. Eng. 53, 137 (1974).
[9]K. M. Case, Rev. Mod. Phys. 29, 651 (1957).
[10]J. G. Kuriyan, in *Radiation and Remote Probing of the Atmosphere* (to be published).

On the geometrization of neutrino fields. I*

Garry Ludwig

Department of Mathematics, University of Alberta, Edmonton, Alberta, Canada
(Received 23 August 1973)

Necessary and sufficient algebraic and differential conditions are obtained for a geometry to have as its source a neutrino field whose energy density relative to any observer is positive or negative definite.

1. INTRODUCTION

Using what turned out to be a refinement of the Plebanski classification[1] of the Ricci tensor we have previously considered the geometrization of nonnull as well as null electromagnetic fields[2] and of massless[3,4] and massive[5] real and complex scalar fields. In this paper we shall employ similar techniques and geometrize those neutrino fields whose energy density is either strictly positive or strictly negative relative to any observer. Partial results to this problem have recently appeared in the literature.[6,7]

In the Appendix we show the relation between our classification[3] of the Ricci tensor $R_{\alpha\beta}$ and the Plebanski one.[1] We also exhibit canonical forms for $R_{\alpha\beta}$ for the various classes. In Sec. 2 we review the neutrino field conditions. We discover that if the Ricci tensor belongs to class A_2, B_5, C_1, D_{1a} or D_{1b} the geometry cannot have a neutrino field as source, and if it belongs to class $B_2, B_3, B_6, D_{1c}, D_{1d}, D_2, D_{3a}, D_{3c}$ or D_4 the energy density of any field for which the trace of the energy-momentum tensor vanishes is positive for some observers and negative for others. The remaining classes $(A_1, A_3, B_1, B_4, C_2, C_3, D_{3b})$ permit neutrino fields with positive (or negative) energy density, which we proceed to study in some detail. Finally, in Sec. 3, we derive the geometrization conditions for such neutrino fields.

The notation used is the same as in previous papers[2,3]; however, the null tetrad corresponding to the spinor dyad $\{k_A, m_A\}$ is now called $\{k_\alpha, n_\alpha, m_\alpha, \overline{m}_\alpha\}$. The Newman–Penrose[8,9] and the spinor[9] formalisms are assumed known. The transformation laws for the Newman–Penrose scalars under a change of spinor dyad are used extensively in the remainder of the paper. For the reader's convenience, they are listed in the Appendix.

2. THE NEUTRINO FIELD

A two-component neutrino field k^A satisfies the equation[10]

$$\nabla_{A\dot{X}} k^A = 0 \tag{2.1}$$

and has an energy–momentum tensor

$$T_{\mu\nu} = i[\sigma_{\mu A\dot{X}}(k^A \bar{k}^{\dot{X}}{}_{;\nu} - \bar{k}^{\dot{X}} k^A{}_{;\nu}) + \sigma_{\nu A\dot{X}}(k^A \bar{k}^{\dot{X}}{}_{;\mu} - \bar{k}^{\dot{X}} k^A{}_{;\mu}) \tag{2.2}$$

whose trace vanishes because of Eq. (2.1). According to the Einstein equations the Ricci scalar must also vanish and

$$R_{\alpha\beta} = - T_{\alpha\beta}. \tag{2.3}$$

If we choose a spinor m^A satisfying $k_A m^A = 1$ (but otherwise arbitrary for the moment) the dyad $\{k_A, m_A\}$ is defined by the neutrino field only up to the null rotation

$$k_A \to k_A, \qquad m_A \to c k_A + m_A. \tag{2.4}$$

From Eqs. (2.2) and (2.3) we find that in such a dyad $R_{\alpha\beta}$ has components

$$\phi_{00} = 0, \quad \phi_{01} = -\frac{i\kappa}{2}, \quad \phi_{02} = - i\sigma,$$
$$\phi_{11} = (i/2)(\bar{\epsilon} - \epsilon), \quad \phi_{12} = (i/2)(\bar{\alpha} - \tau - \beta), \tag{2.5}$$
$$\phi_{22} = i(\bar{\gamma} - \gamma),$$

and Eq. (2.1) reduces to

$$\rho = \epsilon, \qquad \tau = \beta. \tag{2.6}$$

It is not surprising to find that the field equations, Eqs. (2.5) and (2.6), are invariant with respect to the transformation given by Eq. (2.4).

Since in Eq. (2.5) ϕ_{00} vanishes it follows that a geometry which does not possess a null vector field k^α such that $R_{\alpha\beta}k^\alpha k^\beta$ vanishes cannot admit a neutrino field as its source. An easy calculation using the canonical forms given in Table II of the Appendix shows that this is the case for geometries whose Ricci tensor is in class $A_{2\pm}, B_{5a,5b}, C_{1\pm}$, or $D_{1a,1b}$.

The energy density of a field at any point P with respect to an observer whose world line contains P is defined by $E(u) = T_{\alpha\beta}u^\alpha u^\beta$, where $T_{\alpha\beta}$ is the energy-momentum tensor of the field and u^α is the velocity of the observer at P. The flow of energy in the field with respect to this observer is defined by $Q_\alpha(u) = T_{\alpha\beta}u^\beta$. According to Wainwright[11] a field satisfies the strong energy condition if $E(u) > 0$ and $Q_\alpha(u)Q^\alpha(u) \geq 0$, the weak energy condition E_2 if $Q_\alpha(u)Q^\alpha(u) \geq 0$, and the weak energy condition E_1 if $E(u) \neq 0$ (in each case for all time-like vectors u and at each point P for which $T_{\alpha\beta} \neq 0$).

Again using the canonical forms given in the Appendix we easily derive that any field for which T, i.e., $T^\alpha{}_\alpha$, vanishes and for which $R_{\alpha\beta}(= - T_{\alpha\beta})$ belongs to class $A_{2-}, A_{3-}, B_{1b}, B_{4b}, B_{5b}, C_{1a-}, C_{1b-}$ or D_{1a-} satisfies the strong energy condition (i.e., $Q_\alpha Q^\alpha \geq 0$, $E > 0$). If the field is such that $R_{\alpha\beta}$ is in class $A_{2+}, A_{3+}, B_{1a}, B_{4a}, B_{5a}, C_{1a+}, C_{1b+}$ or D_{1a+} it satisfies only the weak energy condition E_2 $(Q_\alpha Q^\alpha > 0, \; E < 0)$. The field satisfies only the weakest energy condition E_1 with $E > 0$ if $R_{\alpha\beta}$ belongs to class $A_{1-}, C_{1c-}, C_{2-}, C_{3-}, D_{1b-}$ or D_{3b-} and with $E < 0$ if $R_{\alpha\beta}$ is in $A_{1+}, C_{1c+}C_{2+}, C_{3+}$, D_{1b} or D_{3b+}. The remaining classes do not allow even the weakest energy condition to be fulfilled. In all cases where $Q_\alpha Q^\alpha$ is strictly nonnegative it is strictly positive except for classes $A_{3\pm}$ when $Q_\alpha(u)Q^\alpha(u) \equiv 0$ for all time-like vectors u. Therefore, if the field is a pure radiation field the Ricci tensor must belong to $A_{3\pm}$.

In the remainder of the paper we shall consider only neutrino fields satisfying at least the weak energy condition E_1. The Ricci tensor is then in class $A_{1\pm}, A_{3\pm}$, $B_{1a,b}, B_{4a,b}, C_{2\pm}, C_{3\pm}$, or $D_{3b\pm}$ and may be written

$$R_{\alpha\beta} = C k_\alpha k_\beta + 2\lambda k_{(\alpha} n_{\beta)} - \lambda_1 V_\alpha V_\beta - \lambda_2 T_\alpha T_\beta - \tfrac{1}{4}(\lambda_1 + \lambda_2 + 2\lambda) g_{\alpha\beta}. \tag{2.7}$$

The signs of the coefficients C, λ, λ_1, and λ_2 are given in Table I below for each class. The vectors T_α (if $\lambda_2 \neq 0$) and V_α (if $\lambda_1 \neq 0$) are normalized space-like eigenvectors of $R_{\alpha\beta}$ corresponding to nonrepeated eigenvalues, n_α (if $\lambda \neq 0$) is a future-pointing null vector orthogonal to the space-like eigenplane of $R_{\alpha\beta}$, and k_α ($C \neq 0$ or $\lambda \neq 0$) is the neutrino flux vector.

The neutrino flux vector k_α, by definition the future-pointing null vector corresponding to the neutrino spinor k_A, is an eigenvector of $R_{\alpha\beta}$ in all cases under consideration and is, of course, orthogonal to T_α (when $\lambda_2 \neq 0$) and V_α (when $\lambda_1 \neq 0$). For example, if $R_{\alpha\beta}$ belongs to class B then, apart from a multiple of $g_{\alpha\beta}$, we can write[3] $R_{\alpha\beta} = R_{(\alpha}S_{\beta)}$. Since $R_{\alpha\beta}k^\alpha k^\beta = 0$ the neutrino flux vector k_α is orthogonal to either R_α or S_α. For class $B_{1a,b}$, R_α and S_α are null eigenvectors and therefore k_α must coincide with one of these. We proceed in a similar fashion in all other cases.

We now tie the dyad $\{k_A, m_A\}$ more closely to the geometry by choosing m_A in such a way that the corresponding null vector n_α is the one appearing in Eqn. (2.7) (if $\lambda \neq 0$) or is orthogonal to T_α (if $\lambda_2 \neq 0$) and V_α (if $\lambda_1 \neq 0$). The dyad is now uniquely defined except when $\lambda = \lambda_1 = 0$. For $\lambda_2 \neq 0$ let us define an angle θ in terms of the complex null vector m_α corresponding to the dyad by $T_\alpha m^\alpha = -2^{-1/2}e^{i\theta}$. For classes $A_{3\pm}, A_{1\pm}$, and $C_{2\pm}$ the dyad may be transformed according to Eq. (2.4) provided that for $A_{1\pm}$ and $C_{2\pm}$ the transformation parameter c is restricted by $\bar{c} = -ce^{2i\theta}$.

The components of $R_{\alpha\beta}$ with respect to the dyad just specified are given by

$$\Lambda = \phi_{00} = \phi_{01} = \phi_{12} = 0,$$
$$\phi_{11} = \tfrac{1}{8}(\lambda_1 + \lambda_2 - 2\lambda), \qquad \phi_{02} = \tfrac{1}{4}(\lambda_2 - \lambda_1)e^{2i\theta}, \tag{2.8}$$
$$\phi_{22} = -C/2,$$

and the field equations, Eqs. (2.5) and (2.6), specialize to

$$\begin{aligned} &\rho = \epsilon, \quad \tau = \beta, \quad \kappa = 0, \quad \bar{\alpha} = 2\beta, \\ &\phi_{02} = -i\sigma, \quad \phi_{11} = \omega, \quad \phi_{22} = i(\bar{\gamma} - \gamma), \end{aligned} \tag{2.9}$$

where ω, the imaginary part of ρ, is the twist of the geodesic null congruence associated with k_α.

Eqs. (2.9) are invariant with respect to the freedom of dyad (if any) described above as well as with respect to rotations through a constant angle ϕ_0 in the plane orthogonal to k_α and n_α, i.e., $k_A \to e^{i\phi_0}k_A, m_A \to e^{-i\phi_0}m_A$, provided, of course, we let $\theta \to \theta + 2\phi_0$ (when $\lambda_2 \neq 0$).

We exhibit the signs of various relevant quantities in Table I. A blank in the last column indicates that $Q_\alpha Q^\alpha$ is positive for some observers and negative for others.

TABLE I.

Class	C	λ	λ_1	λ_2	$\lvert\sigma\rvert$	ω	ϕ_{22}	E	$Q_\alpha Q^\alpha$
A_{1+}	0	0	0	−	+	−	0	−	
A_{1-}	0	0	0	+	+	+	0	+	
A_{3+}	+	0	0	0	0	0	−	−	0
A_{3-}	−	0	0	0	0	0	+	+	0
B_{1a}	0	+	0	0	0	−	0	−	+
B_{1b}	0	−	0	0	0	+	0	+	+
B_{4a}	+	+	0	0	0	−	−	−	+
B_{4b}	−	−	0	0	0	+	+	+	+
C_{2+}	+	0	0	−	+	−	−	−	
C_{2-}	−	0	0	+	+	+	+	+	
C_{3+}	0	0	−	−	+	−	0	−	
C_{3-}	0	0	+	+	+	+	0	+	
D_{3b+}	+	0	−	−	+	−	−	−	
D_{3b-}	−	0	+	+	+	+	+	+	

Results obtained previously[11,12] can now readily be deduced from Table I and Eqs. (2.8) and (2.9): The null congruence defined by the neutrino flux vector is geodesic. It is shearfree if and only if the energy condition E_2 is satisfied. If it is twistfree it is also shearfree and the class is $A_{3\pm}$, i.e., the field is a pure radiation field. In all other cases sign ω = sign E. In all cases $\sigma\bar{\sigma} \leq 4\omega^2$ and $\phi_{22}\omega \geq 0$, but if the strong energy condition is obeyed, then $\phi_{22} \geq 0$, $\omega \geq 0$ (but not both zero).

3. GEOMETRIZATION

A geometry whose source is a null electromagnetic field has a Ricci tensor belonging to class A_{3-}. If the electromagnetic field is nonnull the class is B_{1b}. If the source is a real scalar field, $R_{\alpha\beta}$ is in class A_-, if it is a complex scalar field, $R_{\alpha\beta}$ is in class C_- or B_{1b}.

As we have seen, neutrino fields may belong to a large number of classes. Collinson and Shaw[6] have geometrized neutrino fields in $B_{1b}, B_{4b}, C_{2-}, C_{3-}$, and D_{3b-}, whereas Griffiths and Newing[7] dealt with those neutrino fields in $A_3, B_{2b,c}, B_{3d,e}, B_{4a,b}, C_2$, and D_3 whose flux vector corresponds to a geodesic congruence. Here we shall geometrize all neutrino fields satisfying (at least) the weakest energy condition E_1, namely those in $A_{1\pm}, A_{3\pm}, B_{1a,b}, B_{4a,b}, C_{2\pm}, C_{3\pm}$, and $D_{3b\pm}$.

Let us consider a geometry with Ricci tensor belonging to any of these classes except $A_{3\pm}$. (The latter case will be dealt with separately). If we assume that

$$R = 0, \tag{3.1}$$

then $R_{\alpha\beta}$ may be written as in Eq. (2.7), where the signs of C, λ, λ_1, and λ_2 are as in Table I, T_α (if $\lambda_2 \neq 0$) and V_α (if $\lambda_1 \neq 0$) are normalized space-like eigenvectors of $R_{\alpha\beta}$ corresponding to nonrepeated eigenvalues, k_α (if $C \neq 0$ or $\lambda \neq 0$) is a future-pointing null eigenvector of $R_{\alpha\beta}$, and n_α (if $\lambda \neq 0$) is a future-pointing null vector orthogonal to the space-like eigenplane of $R_{\alpha\beta}$, chosen so that $k_\alpha n^\alpha = 1$.

We define two future-pointing null vectors k_α and n_α (such that $k_\alpha n^\alpha = 1$) for the remaining cases as follows. For class $C_{3\pm}$ we take k_α orthogonal to T_α and V_α. For class $A_{1\pm}$ there are at most four directions orthogonal to T_α for which the corresponding null congruence may be geodesic. This can be seen as follows. Adopting a dyad satisfying $k_\alpha T^\alpha = n_\alpha T^\alpha = 0$ and $m_\alpha - \bar{m}_\alpha = 2^{1/2}iT_\alpha$ we have, as seen from the Appendix, the freedom to make null rotations about m_A with real parameters b. Again using the Appendix we find that under such a dyad change the imaginary part of κ, Im κ, transforms according to

$$\begin{aligned} \operatorname{Im}\kappa' = \operatorname{Im}\kappa + b\operatorname{Im}(\sigma + \rho + 2\epsilon) + b^2\operatorname{Im}(\tau + \pi + 2\alpha + 2\beta) \\ + b^3\operatorname{Im}(\mu + \lambda + 2\gamma) + b^4\operatorname{Im}\nu. \end{aligned}$$

Therefore, there are at most four real values b for which $\operatorname{Im}\kappa' = 0$. We assume there is at least one and choose k_α in the corresponding direction. Finally, we choose n_α orthogonal to T_α (if $\lambda_2 \neq 0$) and V_α (if $\lambda_1 \neq 0$). Note that in classes $B_{1a,b}$ and $C_{3\pm}$ there is a choice of two directions for k_α. In all cases where there is a choice we assume that the conditions below are satisfied for at least one such choice.

Let k_A and m_A be spinors corresponding to k_α and n_α, respectively. The dyad $\{k_A, m_A\}$ corresponds to a null tetrad $\{k_\alpha, n_\alpha, m_\alpha, \bar{m}_\alpha\}$. The components of $R_{\alpha\beta}$ with respect to this tetrad are given by Eq. (2.8), where the angle θ is defined (for $\lambda_2 \neq 0$) by $T_\alpha m^\alpha = -2^{-1/2}e^{i\theta}$.

We now assume that the null congruence associated with k_α is geodesic and that its twist has the same sign as $\lambda_1 + \lambda_2 - 2\lambda$, i.e., that

$$\kappa = 0, \quad \omega/\phi_{11} > 0, \tag{3.2}$$

and use the freedom

$$k_A \to a k_A, \quad m_A \to a^{-1} m_A \quad (a > 0) \tag{3.3}$$

with $a = (\phi_{11}/\omega)^{1/2}$ to make $\phi_{11} = \omega$. Our dyad is still not uniquely determined by the geometry but permits the freedom

$$k_A \to e^{i\varphi} k_A, \quad m_A \to e^{-i\varphi} m_A \tag{3.4}$$

[accompanied by $\theta \to \theta + 2\phi$ when $\lambda = 0$], and, when $\lambda = \lambda_1 = 0$, a null rotation [Eq. (2.4)] with c restricted by $\bar{c} = -ce^{2i\theta}$.

Using the Appendix it is straightforward to verify that the conditions

$$\phi_{02} = -i\sigma \tag{3.5}$$

and

$$\rho + \bar{\rho} - \epsilon - \bar{\epsilon} = 0 = \bar{\alpha} + \beta - 3\tau \tag{3.6}$$

hold with respect to any such dyad if they hold with respect to one. They are, therefore, conditions on the geometry, which we assume satisfied.

Consider next the equations

$$\begin{aligned} D\phi &= i(\epsilon - \rho), \\ \delta\phi &= i(\beta - \tau), \\ \bar{\delta}\phi &= i(\bar{\tau} - \bar{\beta}), \\ \Delta\phi &= \frac{i}{2}(\gamma - \bar{\gamma}) + \frac{\phi_{22}}{2}, \end{aligned} \tag{3.7}$$

i.e., $\phi_{,\alpha} = S_\alpha$, where

$$S_\alpha = i(\epsilon - \rho) n_\alpha - i(\beta - \tau)\bar{m}_\alpha - i(\bar{\tau} - \bar{\beta}) m_\alpha + \tfrac{1}{2}[i(\gamma - \bar{\gamma}) + \phi_{22}] k_\alpha .$$

One can easily verify using the Appendix that two such vectors S_α, obtained in two different permissible tetrads, differ at most by the gradient of a scalar function. Therefore, $S_{[\alpha;\beta]}$ is a geometric quantity. The assumption

$$S_{[\alpha;\beta]} = 0 \tag{3.8}$$

permits us to solve Eq. (3.7) for ϕ, the solution being unique up to a constant. If we perform a phase change using this function ϕ in Eq. (3.4) we arrive at a spinor k_A which satisfies the field equations (2.9).

When the direction k_α is not uniquely defined by the geometry the above assumptions must be valid for at least one of the (at most four) possible directions. Apart from this difficulty the neutrino spinor is then determined uniquely up to a constant phase. That the conditions we have found above are also necessary is easily seen.

In summary, necessary and sufficient conditions that a geometry of class $A_{1\pm}, B_{1a,b}, B_{4a,b}, C_{2\pm}, C_{3\pm}$ or $D_{3b\pm}$ have as its source a neutrino field are given by Eqs. (3.1), (3.2), (3.5), (3.6), and (3.8).

The geometrization procedure for neutrino fields in class $A_{3\pm}$ is slightly different. Given a geometry of class $A_{3\pm}$ with vanishing Ricci scalar we can write $R_{\alpha\beta} = Ck_\alpha k_\beta$ with $C > 0$ for A_{3+} and $C < 0$ for A_{3-}. The geometry defines only the direction of k_α which we take to be future-pointing. Let k_A be any spinor corresponding to k_α and let m_A be such that $k_A m^A = 1$. In such a dyad the only nonzero component of $R_{\alpha\beta}$ is $\phi_{22} = -C/2$. Let us assume that the null congruence determined by k_α is geodesic, shearfree and twistfree, i.e.,

$$\kappa = \sigma = \omega = 0. \tag{3.9}$$

Note that this implies $\psi_0 = 0$. If we further impose the geometric conditions

$$\psi_1 = 0 = \psi_2 = \bar{\psi}_2, \tag{3.10}$$

the compatibility conditions for the equations

$$\begin{aligned} D\psi &= \rho + \bar{\rho} - \epsilon - \bar{\epsilon}, \\ \delta\psi &= 3\tau - \bar{\alpha} - \beta, \\ \bar{\delta}\psi &= 3\bar{\tau} - \alpha - \bar{\beta} \end{aligned}$$

are identically satisfied and a real solution ψ of these equations exists. Transforming to a new dyad using Eq. (3.3) with $a = e^{\psi/2}$ yields

$$\rho + \bar{\rho} - \epsilon - \bar{\epsilon} = 0 = 3\tau - \bar{\alpha} - \beta.$$

The dyad may still be subjected to the transformations given by Eqs. (2.4) and (3.4) as well as Eq. (3.3) for parameters a satisfying

$$Da = \delta a = 0. \tag{3.11}$$

With S_α as defined above one finds that, in general, Eq. (3.8) is not invariant under allowed scale changes [cf. Eqs. (3.3) and (3.11)].

However, if we define H_α by

$$H_\alpha = S_\alpha - \tfrac{1}{2}\phi_{22} k_\alpha$$

we find that $H_{[\alpha;\beta]}$ is a geometric quantity.

We now distinguish two cases. In the ordinary case,

$$(\phi_{22} k_{[\alpha})_{;\beta]} \neq 0. \tag{3.12}$$

It is straightforward to verify that $\frac{1}{2}L$, given by

$$L \equiv 2(\phi_{22} k_\alpha)_{;\beta} n^{[\alpha} m^{\beta]} = \delta\phi_{22} + 2\tau\phi_{22},$$

is the only nonzero component of $(\phi_{22} k_{[\alpha})_{;\beta]}$. If we let

$$M \equiv 4H_{\alpha;\beta} n^{[\alpha} m^{\beta]}$$

and assume the geometric conditions

$$M/L < 0, \quad D(ML^{-1}) = \delta(ML^{-1}) = 0 \tag{3.13}$$

hold, we can, by means of an allowed scale change, satisfy the equation $M + L = 0$, i.e.,

$$S_{[\alpha;\beta]} n^\alpha m^\beta = 0.$$

If we further assume the geometric conditions

$$S_{[\alpha;\beta]} k^\alpha n^\beta = S_{[\alpha;\beta]} k^\alpha m^\beta = S_{[\alpha;\beta]} m^\alpha \bar{m}^\beta, \tag{3.14}$$

then Eqs. (3.7) are compatible and have a solution ϕ that is unique up to an additive constant. Using Eq. (3.4) with this ϕ we arrive at a spinor k_A which satisfies the field equations (2.9). Moreover, this spinor is unique up to a constant phase change. The sufficient conditions we have found are also necessary, as is easily verified.

Therefore, necessary and sufficient conditions that a geometry of class $A_{3\pm}$ have as its source a neutrino field satisfying Eq. (3.12) are given by Eqs. (3.1), (3.9), (3.10), (3.13), and (3.14).

In the exceptional case that

$$(\phi_{22}k_{[\alpha})_{;\beta]} = 0 \qquad (3.15)$$

the tensors $S_{[\alpha;\beta]}$ and $H_{[\alpha;\beta]}$ coincide. Equation (3.8) is once more a geometric condition which we can require the given geometry to satisfy. We can use the solution ϕ of Eq. (3.7) in Eq. (3.4) and obtain a spinor k_A which satisfies the field equations (2.9). Again, the sufficient conditions we have found are also necessary.

Therefore, necessary and sufficient conditions that a geometry of class $A_{3\pm}$ has as its source a neutrino field satisfying Eq. (3.15) are given by Eqs. (3.1), (3.8), (3.9), and (3.10).

In the exceptional case the neutrino spinor is far from unique. For each solution "a" of Eqs. (3.11) we can find a function ϕ (unique up to an additive constant) so that $ak_Ae^{i\phi}$ is also a neutrino spinor if k_A is. This situation, analogous to the exceptional case for the null electromagnetic field,[2,13,14] has been discussed previously by Griffiths and Newing.[15]

In conclusion, it should be remarked that it has not been possible to express the geometrization conditions directly in terms of $R_{\alpha\beta}$ as one can do for a nonnull electromagnetic field (but not, to the author's knowledge, for a null electromagnetic field[2,13]). The conditions given here involve $R_{\alpha\beta}$ only implicitly. Nevertheless, they are geometric in the sense that they prescribe an algorithm for determining whether or not a given geometry has as its source a neutrino field with positive (or negative) definite energy. It seems clear that even if these conditions could be written explicitly in terms of the Ricci tensor they would be so cumbersome as to be of little or no practical value.

APPENDIX

The relationship between our classification[3] of the trace-free Ricci tensor and that of Plebanski[1] is given in Table II. Classes C_1, D_1, and D_3 have been subclassified further for purposes of this paper.

It should be recalled that Z represents a complex

TABLE II.

Ref. 3	Plebanski	Canonical form for $R_{\alpha\beta}$		
	$[4T]_1$		0	
$A_{1\pm}$	$[3T-S]_2$	$S \lessgtr T$	$-\lambda_y y_\alpha y_\beta$	$\lambda_y \lessgtr 0$
$A_{2\pm}$	$[T-3S]_2$	$T \gtrless S$	$\lambda_t t_\alpha t_\beta$	$\lambda_t \gtrless 0$
$A_{3\pm}$	$[4N]_2$	$C \gtrless 0$	$Ck_\alpha k_\beta$	$C \gtrless 0$
$B_{1a,b}$	$[2T-2S]_2$	$T \gtrless S$	$2\lambda_t k_{(\alpha}n_{\beta)}$	$\lambda_t \gtrless 0$
B_{2a}	$[4N]_3$		$k_{(\alpha}y_{\beta)}$	
$B_{2b,c}$	$[2N-2S]_{(2-1)}$	$N \gtrless S,\ C \lessgtr 0$	$(2\lambda_n/A)k_{(\alpha}z_{\beta)}$	$\lambda_n \gtrless 0$
			$[=2\lambda_n k_{(\alpha}n_{\beta)} - (\lambda_n/A^2)k_\alpha k_\beta]$	
$B_{3a,b,c}$	$[2T-S_1-S_2]_3$	$S_1 < T < S_2$	$-\lambda_x x_\alpha x_\beta - \lambda_y y_\alpha y_\beta$	$\lambda_x < 0 < \lambda_y$
		$S_1+S_2 \gtreqqless 2T$		$\lambda_x+\lambda_y \gtreqqless 0$
$B_{3d,e}$	$[3N-S]_3$	$S \gtrless N,\ C \gtrless 0$	$Ck_\alpha k_\beta - \lambda_y y_\alpha y_\beta$	$\lambda_y \gtrless 0,\ C \gtrless 0$
$B_{3f,g}$	$[T-2S_1-S_2]_3$	$S_2 \gtrless T \gtrless S_1$	$-\lambda_z z_\alpha z_\beta + \lambda_t t_\alpha t_\beta$	$\lambda_z \gtrless \lambda_t \gtrless 0$
$B_{4a,b}$	$[2N-2S]_{(2-1)}$	$N \gtrless S,\ C \gtrless 0$	$(2\lambda_n/A)k_{(\alpha}t_{\beta)}$	$\lambda_n \gtrless 0$
			$[=2\lambda_n k_{(\alpha}n_{\beta)} + (\lambda_n/A^2)k_\alpha k_\beta]$	
$B_{5a,b}$	$[T-2S_1-S_2]_3$	$T \gtrless S_2 \gtrless S_1$	$-\lambda_z z_\alpha z_\beta + \lambda_t t_\alpha t_\beta$	$\lambda_t \gtrless \lambda_z \gtrless 0$
$B_{6a,b,c}$	$[Z-\bar{Z}-2S]_3$	$D \gtreqqless 0$	$Bz_{(\alpha}t_{\beta)} + Dt_\alpha t_\beta$	$B^2-D^2>0,\ D \gtreqqless 0$
$C_{1a\pm}$	$[T-2S_1-S_2]_3$	$S_2 \lessgtr S_1 \lessgtr T$	$-\lambda_z z_\alpha z_\beta + \lambda_t t_\alpha t_\beta$	$\lambda_z \lessgtr 0 \lessgtr \lambda_t$
		$S_2+T-2S_1 \gtrless 0$		$\lambda_z+\lambda_t \gtrless 0$
$C_{1b\pm}$	$[T-2S_1-S_2]_3$	$S_2 \lessgtr S_1 \lessgtr T$	$-\lambda_z z_\alpha z_\beta + \lambda_t t_\alpha t_\beta$	$\lambda_z \lessgtr 0 \lessgtr \lambda_t$
		$S_2+T-2S_1 = 0$		$\lambda_z = -\lambda_t$
$C_{1c\pm}$	$[T-2S_1-S_2]_3$	$S_2 \lessgtr S_1 \lessgtr T$	$-\lambda_z z_\alpha z_\beta + \lambda_t t_\alpha t_\beta$	$\lambda_z \lessgtr 0 \lessgtr \lambda_t$
		$S_2+T-2S_1 \lessgtr 0$		$\lambda_z+\lambda_t \lessgtr 0$
$C_{2\pm}$	$[3N-S]_3$	$S \lessgtr N,\ C \gtrless 0$	$Ck_\alpha k_\beta - \lambda_y y_\alpha y_\beta$	$\lambda_y \lessgtr 0,\ C \gtrless 0$
$C_{3\pm}$	$[2T-S_1-S_2]_3$	$S_1, S_2 \lessgtr T$	$-\lambda_x x_\alpha x_\beta - \lambda_y y_\alpha y_\beta$	$\lambda_x, \lambda_y \lessgtr 0$
$D_{1a\pm}$	$[T-S_1-S_2-S_3]_4$	$S_3 \lessgtr S_2 \lessgtr S_1 \lessgtr T$	$-\lambda_x x_\alpha x_\beta - \lambda_y y_\alpha y_\beta - \lambda_z z_\alpha z_\beta$	$\lambda_z \lessgtr \lambda_y \lessgtr \lambda_x \lessgtr 0$
		$S_1+S_2-S_3 \lesseqgtr T$		$\lambda_x+\lambda_y-\lambda_z \lesseqgtr 0$
$D_{1b\pm}$	$[T-S_1-S_2-S_3]_4$	$S_3 \lessgtr S_2 \lessgtr S_1 \lessgtr T$	$-\lambda_x x_\alpha x_\beta - \lambda_y y_\alpha y_\beta - \lambda_z z_\alpha z_\beta$	$\lambda_z \lessgtr \lambda_y \lessgtr \lambda_x \lessgtr 0$
		$S_1+S_2-S_3 \gtrless T$		$\lambda_x+\lambda_y-\lambda_z \gtrless 0$
$D_{1c,d}$	$[T-S_1-S_2-S_3]_4$	$S_1 \lessgtr T \lessgtr S_2 \lessgtr S_3$	$-\lambda_x x_\alpha x_\beta - \lambda_y y_\alpha y_\beta - \lambda_z z_\alpha z_\beta$	$\lambda_x \lessgtr 0 \lessgtr \lambda_y \lessgtr \lambda_z$
D_2	$[Z-\bar{Z}-S_1-S_2]_4$		$-\lambda_x x_\alpha x_\beta + Dt_\alpha t_\beta + Bz_{(\alpha}t_{\beta)}$	$B^2-D^2>0$
$D_{3a\pm}$	$[2N-S_1-S_2]_4$	$S_1, S_2 \gtrless N,\ C \gtrless 0$	$-\lambda_x x_\alpha x_\beta - \lambda_y y_\alpha y_\beta + Ck_\alpha k_\beta$	$\lambda_x, \lambda_y \gtrless 0,\ C \gtrless 0$
$D_{3b\pm}$	$[2N-S_1-S_2]_4$	$S_1, S_2 \lessgtr N,\ C \gtrless 0$	$-\lambda_x x_\alpha x_\beta - \lambda_y y_\alpha y_\beta + Ck_\alpha k_\beta$	$\lambda_x, \lambda_y \lessgtr 0,\ C \gtrless 0$
$D_{3c\pm}$	$[2N-S_1-S_2]_4$	$S_1 < N < S_2,\ C \gtrless 0$	$-\lambda_x x_\alpha x_\beta - \lambda_y y_\alpha y_\beta + Ck_\alpha k_\beta$	$\lambda_x < 0 < \lambda_y,\ C \gtrless 0$
D_4	$[3N-S]_4$		$-\lambda_x x_\alpha x_\beta + By_{(\alpha}k_{\beta)}$	$B \neq 0$

eigenvalue whereas T, N or S represent a real eigenvalue with which is associated, respectively, a time-like eigenvector, no time-like eigenvector but a null eigenvector, only space-like eigenvectors. The number of times an eigenvalue occurs (if more than once) is placed before the symbol T, N or S.

Adding a multiple of $g_{\alpha\beta}$ to $R_{\alpha\beta}$ leaves the eigenvectors unchanged but increases all eigenvalues by that multiple. We choose, in each case, a convenient multiple and exhibit a canonical form for $R_{\alpha\beta}$.

$\{t_\alpha, x_\alpha, y_\alpha, z_\alpha\}$ stands for an orthonormal tetrad adapted to the eigenvectors, and, where appropriate, $k_\alpha = A(t_\alpha - z_\alpha)$, $n_\alpha = (1/2A)(t_\alpha + z_\alpha)$. $\lambda_t, \lambda_x, \lambda_y, \lambda_z$, and λ_n denote corresponding eigenvalues. The quantities C and D are defined to be $R_{\alpha\beta}n^\alpha n^\beta$ and $R_{\alpha\beta}t^\alpha t^\beta$, respectively. The symbol $\lesseqgtr$ means that the inequalities $\leq$ and $\geq$ apply for class D_{1a+} and D_{1a-}, respectively.

Finally, we exhibit the transformation laws for the null tetrad, the spin-coefficients and the components of the Weyl and Ricci tensors under changes of dyad.

For a phase and scale change

$$k'_A = ae^{i\phi}k_A, \quad m'_A = a^{-1}e^{-i\phi}m_A \quad (a > 0),$$

the tetrad and the Newman-Penrose scalars transform as

$$k'_\alpha = a^2 k_\alpha, \quad n'_\alpha = a^{-2}n_\alpha, \quad m'_\alpha = e^{2i\phi}m_\alpha,$$

$$\kappa' = a^4 e^{2i\phi}\kappa, \quad \sigma' = a^2 e^{4i\phi}\sigma, \quad \rho' = a^2\rho, \quad \tau' = e^{2i\phi}\tau,$$

$$\nu' = a^{-4}e^{-2i\phi}\nu, \quad \lambda' = a^{-2}e^{-4i\phi}\lambda, \quad \mu' = a^{-2}\mu, \quad \pi' = e^{-2i\phi}\pi,$$

$$\epsilon' = a^2(\epsilon + D\ln a + iD\phi), \quad \beta' = e^{2i\phi}(\beta + \delta\ln a + i\delta\phi),$$

$$\gamma' = a^{-2}(\gamma + \Delta\ln a + i\Delta\phi), \quad \alpha' = e^{-2i\phi}(\alpha + \bar\delta\ln a + i\bar\delta\phi),$$

$$\psi'_0 = a^4 e^{4i\phi}\psi_0, \quad \psi'_1 = a^2 e^{2i\phi}\psi_1, \quad \psi'_2 = \psi_2,$$
$$\psi'_3 = a^{-2}e^{-2i\phi}\psi_3, \quad \psi'_4 = a^{-4}e^{-4i\phi}\psi_4,$$

$$\phi'_{00} = a^4\phi_{00}, \quad \phi'_{01} = a^2 e^{2i\phi}\phi_{01}, \quad \phi'_{02} = e^{4i\phi}\phi_{02},$$

$$\phi'_{11} = \phi_{11}, \quad \phi'_{12} = a^{-2}e^{2i\phi}\phi_{12}, \quad \phi'_{22} = a^{-4}\phi_{22},$$

whereas for a null rotation about k_A, given by Eq. (2.4), they transform as

$$k'_\alpha = k_\alpha, \quad n'_\alpha = c\bar c k_\alpha + cm_\alpha + \overline{cm}_\alpha + n_\alpha,$$
$$m'_\alpha = \bar c k_\alpha + m_\alpha,$$

$$\kappa' = \kappa, \quad \sigma' = \bar c\kappa + \sigma, \quad \rho' = c\kappa + \rho,$$
$$\tau' = c\bar c\kappa + c\sigma + \bar c\rho + \tau,$$

$$\epsilon' = c\kappa + \epsilon, \quad \alpha' = c^2\kappa + c(\epsilon + \rho) + \alpha,$$
$$\beta' = c\bar c\kappa + c\sigma + \bar c\epsilon + \beta,$$

$$\gamma' = c^2\bar c\kappa + c^2\sigma + c\bar c(\rho + \epsilon) + c(\tau + \beta) + \bar c\alpha + \gamma,$$

$$\pi' = c^2\kappa + 2c\epsilon + \pi + Dc,$$
$$\mu' = c^2\bar c\kappa + c^2\sigma + 2c\bar c\epsilon + 2c\beta + \bar c\pi + \mu + \bar c Dc + \delta c,$$

$$\lambda' = c^3\kappa + c^2(\rho + 2\epsilon) + 2c\alpha + c\pi + \lambda + cDc + \bar\delta c,$$

$$\nu' = c^3\bar c\kappa + c^3\sigma + c^2\bar c(\rho + 2\epsilon) + c^2(\tau + 2\beta) + c\bar c(2\alpha + \pi)$$
$$+ c(2\gamma + \mu) + \bar c\lambda + \nu + c\bar c Dc + c\delta c + \bar c\bar\delta c + \Delta c,$$

$$\psi'_0 = \psi_0, \quad \psi'_1 = c\psi_0 + \psi_1, \quad \psi'_2 = c^2\psi_0 + 2c\psi_1 + \psi_2,$$
$$\psi'_3 = c^3\psi_0 + 3c^2\psi_1 + 3c\psi_2\psi_3,$$
$$\psi'_4 = c^4\psi_0 + 4c^3\psi_1 + 6c^2\psi_2 + 4c\psi_3 + \psi_4,$$

$$\phi'_{00} = \phi_{00}, \quad \phi'_{01} = \bar c\phi_{00} + \phi_{01},$$
$$\phi'_{02} = \bar c^2\phi_{00} + 2\bar c\phi_{01} + \phi_{02},$$

$$\phi'_{11} = c\bar c\phi_{00} + c\phi_{01} + \bar c\phi_{10} + \phi_{11},$$
$$\phi'_{12} = \bar c^2 c\phi_{00} + 2c\bar c\phi_{01} + \bar c^2\phi_{10} + c\phi_{02} + 2\bar c\phi_{11} + \phi_{12},$$
$$\phi'_{22} = c^2\bar c^2\phi_{00} + 2c^2\bar c\phi_{01} + 2\bar c^2 c\phi_{10} + c^2\phi_{02} + \bar c^2\phi_{20}$$
$$+ 4c\bar c\phi_{11} + 2c\phi_{12} + 2\bar c\phi_{21} + \phi_{22}.$$

The transformation laws for a null rotation about m_A can now be derived quite easily from the above if we use the fact that for the change

$$k'_A = im_A, \quad m'_A = ik_A$$

we get the interchanges[16,17]

$$k_\alpha \leftrightarrow n_\alpha, \quad m_\alpha \leftrightarrow \overline{m}_\alpha, \quad D \leftrightarrow \Delta, \quad \delta \leftrightarrow \bar\delta,$$

$$\psi_0 \leftrightarrow \psi_4, \quad \psi_1 \leftrightarrow \psi_3, \quad \psi_2 \leftrightarrow \psi_2, \quad \kappa \leftrightarrow -\nu,$$
$$\sigma \leftrightarrow -\lambda,$$

$$\rho \leftrightarrow -\mu, \quad \tau \leftrightarrow -\pi, \quad \epsilon \leftrightarrow -\gamma, \quad \beta \leftrightarrow -\alpha,$$
$$\phi_{00} \leftrightarrow \phi_{22},$$

$$\phi_{01} \leftrightarrow \phi_{21}, \quad \phi_{02} \leftrightarrow \phi_{20}, \quad \phi_{11} \leftrightarrow \phi_{11}.$$

Thus, for example, under the null rotation

$$k'_A = k_A + bm_A, \quad m'_A = m_A,$$

the spin-coefficient κ transforms as

$$\kappa' = b^3\bar b\nu + b^3\lambda + b^2\bar b(\mu + 2\gamma) + b^2(\pi + 2\alpha)$$
$$+ b\bar b(\tau + 2\beta) + b(\rho + 2\epsilon) + \bar b\sigma + \kappa - b\bar b\Delta b$$
$$- Db - \bar b\delta b - b\bar\delta b.$$

*Supported in part by the National Research Council of Canada.

[1] J. Plebański, Acta Phys. Polon. 26, 963 (1964).
[2] G. Ludwig, Commun. Math. Phys. 17, 98 (1970).
[3] G. Ludwig and G. Scanlan, Commun. Math. Phys. 20, 291 (1971).
[4] G. Ludwig, J. Math. Phys. 12, 1218 (1971).
[5] G. Ludwig, J. Math. Phys. 12, 1220 (1971).
[6] C. D. Collinson and R. Shaw, Int. J. Theor. Phys. 6, 347 (1972).
[7] J. B. Griffiths and R. A. Newing, Int. J. Theor. Phys. 5, 421 (1972).
[8] E. Newman and R. Penrose, J. Math. Phys. 3, 566 (1962).
[9] F. A. E. Pirani, *Lectures on general relativity, Brandeis Summer Institute, 1964* (Prentice-Hall, Englewood Cliffs, N. J., 1965).
[10] J. B. Griffiths, Commun. Math. Phys. 28, 295 (1972).
[11] J. Wainwright, J. Math. Phys. 12, 828 (1971).
[12] J. B. Griffiths and R. A. Newing, J. Phys. A 4, 208 (1971).
[13] A. Peres, Ann. Phys., N. Y., 14, 419 (1961).
[14] R. P. Geroch, Ann. Phys., N. Y., 36, 147 (1966).
[15] J. B. Griffiths and R. A. Newing, J. Phys. A 4, 306 (1971).
[16] G. Ludwig, Indiana Univ. Math. J. 20, 185 (1970).
[17] R. Geroch, A. Held and R. Penrose, J. Math. Phys. 14, 874 (1973).

On the geometrization of neutrino fields. II*

Garry Ludwig

Department of Mathematics, University of Alberta, Edmonton, Alberta, Canada T6G 2G1
(Received 18 September 1973)

In a previous paper neutrino fields with positive (or negative) energy density were geometrized in the sense of Rainich, Misner, and Wheeler. The present paper deals with the geometrization of neutrino fields which do not satisfy such an energy condition.

1. INTRODUCTION

Recently[1] necessary and sufficient conditions were derived in order that a Riemannian geometry admit as source a neutrino field whose energy density is positive (or negative) relative to any observer. Extensive use was made of a certain classification[2] of the Ricci tensor $R_{\alpha\beta}$. The various classes may, for convenience, be grouped as follows:

I: $A_2, B_5, C_1, D_{1a,b}$,

II: $A_1, A_3, B_1, B_4, C_2, C_3, D_{3b}$, 0,

III: $B_2, B_{3a,b,c,d,e}, D_{3a}, D_{3c}, D_4$,

IV: $B_{3f,g}, B_6, D_{1c,d}, D_2$.

A geometry does not admit a neutrino field as source if $R_{\alpha\beta}$ belongs to a class of group I. For geometries in group II the neutrino field has an energy density which is of the same sign for all observers. Neutrino fields in group III or IV do not obey such an energy condition. We have seen that for neutrino fields in group II the neutrino flux is an eigenvector of $R_{\alpha\beta}$ and generates a geodesic null congruence. Neutrino fields in group III may but in general do not; neutrino fields in group IV definitely do not possess this property.

In the previous paper[1] we geometrized all neutrino fields in group II with the exception of those in vacuum (class 0). In Sec. 2 we shall geometrize all remaining neutrino fields for which the flux is an eigenvector of $R_{\alpha\beta}$, and in Sec. 3 those for which it is not.

The notation is the same as in previous papers.[1,2] κ, σ, ρ, etc. are the spin coefficients,[3] ϕ_{00}, ϕ_{01}, etc. the components of $R_{\alpha\beta}$, and Ψ_0, Ψ_1, etc. the components of the Weyl tensor relative to some dyad $\{k_A, m_A\}$. The null tetrad corresponding to such a dyad (which we always normalize by $k_A m^A = 1$) is denoted by $\{k_\alpha, n_\alpha, m_\alpha, \overline{m}_\alpha\}$. T_α and V_α will be unit spacelike vectors which are determined by the given geometry up to a sign and which are mutually orthogonal when they both occur in $R_{\alpha\beta}$. The angle θ will be defined by

$$T_\alpha m^\alpha = -2^{-1/2} e^{i\theta} \tag{1.1}$$

and the vector S_α by

$$S_\alpha = i(\epsilon - \rho) n_\alpha - i(\beta - \tau)\overline{m}_\alpha - i(\overline{\tau} - \overline{\beta}) m_\alpha + \tfrac{1}{2}[i(\gamma - \overline{\gamma}) + \phi_{22}] k_\alpha \tag{1.2}$$

relative to a given dyad. ω stands for Imρ, the imaginary part of the spin coefficient ρ, and Eq. (I3.2), for example, refers to Eq. (3.2) of Ref. 1.

We recall that in any dyad $\{k_A, m_A\}$ in which k_A is the neutrino spinor the neutrino field equations are

$$\rho = \epsilon, \quad \tau = \beta, \tag{1.3}$$

and the gravitational field equations become

$$R = 0 \tag{1.4}$$

and

$$\begin{aligned} &\phi_{00} = 0, \quad 2\phi_{01} = -i\kappa, \quad \phi_{02} = -i\sigma, \quad 2\phi_{11} = i(\overline{\epsilon} - \epsilon), \\ &2\phi_{12} = i(\overline{\alpha} - \tau - \beta), \quad \phi_{22} = i(\overline{\gamma} - \gamma). \end{aligned} \tag{1.5}$$

Our geometrization procedure consists of gradually adapting a dyad to the given geometry until Eqs. (1.3)–(1.5) are satisfied. In the process we make various necessary assumptions that are independent of the remaining freedom of choice in dyad and are, therefore, conditions on the geometry. In proving such independence and in other straightforward calculations (such as deriving compatibility conditions on certain differential equations) of which we give no details, we have made free use of the Ricci and Bianchi identities.[4]

2. NEUTRINO FLUX AN EIGENVECTOR OF $R_{\alpha\beta}$

In this section we shall derive necessary and sufficient conditions that a geometry belonging to class 0 or group III admit as its source a neutrino field whose flux is an eigenvector of the Ricci tensor. The null congruence associated with this flux will turn out to be geodesic. The procedure for all classes but B_{2a} and 0 will follow closely that given in Ref. 1 for fields in group II other than A_3 and 0 and will be carried through in subsection (i). Subsections (ii) and (iii) will deal with neutrino fields in B_{2a} and 0, respetively.

(i) A trace-free Ricci tensor belonging to class D_4 may be written[1]

$$R_{\alpha\beta} = Bk_{(\alpha}T_{\beta)} - \lambda_\alpha V_\beta - \tfrac{1}{4}\lambda g_{\alpha\beta} \quad (B\lambda \neq 0).$$

The null eigenvector k_α is orthogonal to T_α and V_α. The second null vector orthogonal to T_α and V_α will be called n_α.

A trace-free Ricci tensor that belongs to one of the other classes under consideration here can be written as in Eq. (I2.7). The vectors $k_\alpha, n_\alpha, T_\alpha$, and V_α are as described in the paragraph following Eq. (I3.1) but with the signs of C, λ, λ_1, and λ_2 as given in Table I (which also gives the signs of ϕ_{11} and $|\phi_{02}|$ relative to any of the dyads defined below). For classes $B_{3a,b,c}$ we define k_α to be a null vector orthogonal to T_α and V_α and assume the conditions derived below to be valid for at least one of the two choices. Where n_α does not occur in $R_{\alpha\beta}$ we define it to be a null vector orthogonal to T_α (for $\lambda_2 \neq 0$) and V_α (for $\lambda_1 \neq 0$).

TABLE I.

	C	λ	λ_1	λ_2	$\lambda_1+\lambda_2$	$\lvert\phi_{02}\rvert$	ϕ_{11}
B_{2b}	−	+	0	0	0	0	−
B_{2c}	+	−	0	0	0	0	+
B_{3a}	0	0	−	+	0	+	0
B_{3b}	0	0	−	+	+	+	+
B_{3c}	0	0	−	+	−	+	−
B_{3d}	+	0	0	+	+	+	+
B_{3e}	−	0	0	−	−	+	−
D_{3a+}	+	0	+	+	+	+	+
D_{3a-}	−	0	−	−	−	+	−
D_{3c+}	+	0	−	+		+	
D_{3c-}	−	0	−	+		+	

Let k_A and m_A be spinors corresponding, respectively, to the null vectors k_α and n_α, which have been defined for all geometries considered here. The nonzero components of $R_{\alpha\beta}$ relative to such a dyad $\{k_A, m_A\}$ are given by

$$\phi_{02} = -(\lambda/4)e^{2i\theta}, \quad \phi_{12} = (B/4\sqrt{2})e^{i\theta}$$

[with θ defined by Eq. (1.1)] for class D_4 and by Eq. (I 2.8) for the others. The dyad is so far fixed only up to scale changes

$$k_A \to ak_A, \quad m_A \to a^{-1}m_A \quad (a>0), \tag{2.1}$$

phase changes

$$k_A \to e^{i\phi}k_A, \quad m_A \to e^{-i\phi}m_A \tag{2.2}$$

(accompanied, where applicable, by $\theta \to \theta + 2\phi$), and, for class $B_{3d,e}$, a restricted null rotation about k_A, i.e.,

$$k_A \to k_A, \quad m_A \to ck_A + m_A \tag{2.3}$$

subject to

$$\bar{c} = -ce^{2i\theta}. \tag{2.4}$$

Let us now suppose that the null congruence generated by k_α is geodesic, i.e.,

$$\kappa = 0. \tag{2.5}$$

In addition we assume

$$\sigma = 0, \quad \omega\phi_{11}^{-1} > 0 \tag{2.6a}$$

when ϕ_{02} vanishes (i.e., for $B_{2b,c}$),

$$\omega = 0, \quad \mathrm{Im}(i\sigma\phi_{02}^{-1}) = 0, \quad \mathrm{Re}(-i\sigma\phi_{02}^{-1}) > 0 \tag{2.6b}$$

when ϕ_{11} vanishes, or

$$\mathrm{Im}(i\sigma\phi_{02}^{-1}) = 0, \quad \mathrm{Re}(-i\sigma\phi_{02}^{-1}) = \omega\phi_{11}^{-1} > 0 \tag{2.6c}$$

when neither ϕ_{11} nor ϕ_{02} vanish. We can now use the freedom in the choice of dyad given by Eq. (2.1) to satisfy the necessary conditions

$$\phi_{02} = -i\sigma, \quad \phi_{11} = \omega. \tag{2.7}$$

Next we assume that the (geometric)conditions

$$\rho + \bar{\rho} - \epsilon - \bar{\epsilon} = 0, \quad \bar{\alpha} + \beta - 3\tau + 2i\phi_{12} = 0 \tag{2.8}$$

and

$$S_{[\alpha;\beta]} = 0 \tag{2.9}$$

hold, where S_α is defined by Eq. (1.2). Due to Eq. (2.9) the equations

$$\phi_{,\alpha} = S_\alpha \tag{2.10}$$

have a solution ϕ which is unique up to an additive constant. If we make a phase change [Eq. (2.2)] with this ϕ we arrive at a spinor k_A satisfying the field equations (1.3)–(1.5). Apart from the ambiguity in the direction of k_α for class $B_{3a,b,c}$, this spinor is unique up to a constant phase.

It is easily verified that the geometric assumptions given by Eqs. (1.4), (2.5), (2.6), (2.8), and (2.9) are not only sufficient but also necessary ones for a geometry of class $B_{2b,c}$, $B_{3a,b,c,d,e}$, D_{3a}, D_{3c}, or D_4 to have as its source a neutrino field whose flux is an eigenvector of the Ricci tensor. We see from Eq. (2.5) that the null congruence associated with this flux is geodesic.

(ii) We consider next neutrino fields in class B_{2a} whose flux is an eigenvector of $R_{\alpha\beta}$. An example of such a field has been given by Griffiths.[5]

If k_A is the neutrino spinor and k_α the flux, then the Ricci tensor may be written[1]

$$R_{\alpha\beta} = Bk_{(\alpha}T_{\beta)}, \quad B \neq 0, \tag{2.11}$$

where k_α and T_α are orthogonal. If m_A is any spinor such that the associated null direction is orthogonal to T_α, then in the dyad $\{k_A, m_A\}$

$$\phi_{12} = (B/4\sqrt{2})e^{i\theta} \tag{2.12}$$

is the only nonzero component of $R_{\alpha\beta}$. Note that the dyad is determined up to restricted null rotations about k_A [cf. Eqs. (2.3) and (2.4)] and constant phase changes. From the field equations [Eqs. (1.3)–(1.5)] and the Ricci identities[3] we see that the null congruence generated by the flux k_α is geodesic, shear-free, and twist-free and that k_α is a repeated principal null vector of the Weyl tensor, i.e.,

$$\kappa = \sigma = \omega = 0 \tag{2.13}$$

and

$$\Psi_0 = \Psi_1 = 0. \tag{2.14}$$

Let us now investigate the uniqueness of the field. If we look for another neutrino spinor of the form $ae^{i\phi}k_A$, we have to try and solve the equations

$$D\ln a = 0,$$

$$\delta\ln a = i\phi_{12}(1 - 1/a^2) \tag{2.15}$$

and

$$D\phi = \Delta\phi = 0,$$

$$\delta\phi = -\phi_{12}(1 - 1/a^2). \tag{2.16}$$

The compatibility conditions are

$$(1 - 1/a^2)[(\pi + \bar{\tau})\phi_{12} + (\bar{\pi} + \tau)\phi_{21}] = 0, \tag{2.17}$$

$$(1 - 1/a^2)[\delta\phi_{21} - \bar{\delta}\phi_{12} + (\alpha - \bar{\beta})\phi_{12} + (\beta - \bar{\alpha})\phi_{21} + 4ia^{-2}|\phi_{12}|^2] = 0, \tag{2.18}$$

$$(1 - 1/a^2)[\Delta\phi_{12} + (\mu - \gamma + \bar{\gamma})\phi_{12} + \bar{\lambda}\phi_{21}] - 2a^{-2}\phi_{12}\Delta\ln a = 0. \tag{2.19}$$

If we let

$$Q_\alpha = \phi_{12}\bar{m}_\alpha + \phi_{21}m_\alpha \tag{2.20}$$

and

$$N = iQ_{[\alpha;\beta]}m^\alpha\bar{m}^\beta = \tfrac{1}{2}i[\delta\phi_{21} - \bar{\delta}\phi_{12} + (\alpha - \bar{\beta})\phi_{12} + (\beta - \bar{\alpha})\phi_{21}], \tag{2.21}$$

Eqs. (2 17)–(2.19) become, respectively,

$$(1 - 1/a^2)Q_{[\alpha;\beta]}k^\alpha n^\beta = 0,$$

$$-i(1 - 1/a^2)(N - 2a^{-2}|\phi_{12}|^2) = 0,$$

$$(1 - 1/a^2)Q_{[\alpha;\beta]}n^\alpha\bar{m}^\beta - a^{-2}\phi_{12}\Delta\ln a = 0.$$

Also, if we define

$$S = \phi_{12} + \tfrac{1}{2}i(\tau - \bar{\alpha} + \beta), \tag{2.22}$$

evaluate $\bar{\delta}S + \delta\bar{S}$ and set S equal to zero due to Eq. (1.5), we find that necessarily

$$\mathrm{Im}\Psi_2 + (\pi + \bar{\tau})\phi_{12} + (\bar{\pi} + \tau)\phi_{21} = 0. \tag{2.23}$$

It follows that if $\mathrm{Im}\Psi_2 \neq 0$ or if $\mathrm{Im}\Psi_2 = 0$ and $N \leq 0$, then Eqs. (2.15) and (2.16) are satisfied only by $a = 1$, $\phi = \mathrm{const}$, and the neutrino spinor is unique up to a constant phase. This is, in general, also the case if $\mathrm{Im}\Psi_2 = 0$ and $N > 0$ unless $a = (2|\phi_{12}|^2N^{-1})^{1/2}$ satisfies Eqs. (2.15) and (2 19), in which case there are two neutrino spinors.

Conversely, a trace-free Ricci tensor belonging to class B_{2a} may be written as in Eq. (2.11), where k_α is the null eigenvector of $R_{\alpha\beta}$ and T_α is a unit spacelike vector orthogonal to k_α. Adopting a dyad for which k_A corresponds to k_α and m_A to another null direction orthogonal to T_α, we find that ϕ_{12} as given by Eq. (2.12) is the only nonzero component of $R_{\alpha\beta}$. The dyad may be subjected to scale changes, phase changes, and restricted null rotations about k_A. We assume that Eqs. (2.13) and (2.14) are valid and distinguish between the two cases $\mathrm{Im}\Psi_2 \neq 0$ and $\mathrm{Im}\Psi_2 = 0$.

In the first case we define

$$A = -(\mathrm{Im}\Psi_2)^{-1}[\phi_{12}(\pi + \bar{\tau}) + \phi_{21}(\bar{\pi} + \tau)],$$

assume

$$A > 0, \tag{2.24}$$

and make a scale change with $a = A^{1/2}$. In the new dyad, which may still be subjected to phase changes and restricted null rotations about k_A, the necessary condition (2.23) is satisfied. With the further assumptions (2.8) and (2.9) we can, as in subsection (i), determine a spinor k_A which satisfies the field equations and which is unique apart from a constant phase.

The geometric conditions given by Eqs. (1.4), (2.13), (2.14), (2.24), (2.8), and (2.9) are obviously necessary as well. Note, however, that Eq. (2.9) may be relaxed to

$$S_{[\alpha;\beta]}m^\alpha\bar{m}^\beta = S_{[\alpha;\beta]}n^\alpha m^\beta = 0$$

since the other independent components of $S_{[\alpha;\beta]}$ already vanish due to earlier assumptions.

If $\mathrm{Im}\Psi_2$ vanishes Eq. (2.23) becomes the geometric condition

$$(\pi + \bar{\tau})\phi_{12} + (\bar{\pi} + \tau)\phi_{21} = 0 \tag{2.25}$$

which we assume to hold. The equations

$$\begin{aligned} &2D\ln a = \rho + \bar{\rho} - \epsilon - \bar{\epsilon}, \\ &2\delta\ln a = 3\tau - \beta - \bar{\alpha} - 2ia^{-2}\phi_{12} \end{aligned} \tag{2.26}$$

are then compatible. Using a solution of Eq. (2.26) in Eq. (2.1), we can transform to a dyad in which Eqs. (2.8) are valid. Any further scale change must be restricted by Eq. (2.15). We use Eq. (2.2) to make ϵ real; phase changes must now be restricted by $D\phi = 0$.

We now define quantities N and M, respectively, by Eq. (2.21) and

$$-2iM = 2U_{[\alpha;\beta]}m^\alpha\bar{m}^\beta = \delta\bar{S} - \bar{\delta}S + (\alpha - \bar{\beta})S + (\beta - \bar{\alpha})\bar{S} + 2iN,$$

where Q_α, S, and U_α are given by Eqs. (2.20), (2.22) and

$$U_\alpha = (S - \phi_{12})\bar{m}_\alpha + (\bar{S} - \phi_{21})m_\alpha,$$

respectively. In order to satisfy the necessary condition

$$M + N = 0 \tag{2.27}$$

by means of an appropriate scale change we must solve the equation

$$a^4M + a^2(N + 2|\phi_{12}|^2) - 2|\phi_{12}|^2 = 0. \tag{2.28}$$

The solution A (for a^2) is

$$A = 2|\phi_{12}|^2[N + 2|\phi_{12}|^2]^{-1}$$

for $M = 0$ and

$$A = (-2M)^{-1}\{N + 2|\phi_{12}|^2 \pm [(N + 2|\phi_{12}|^2)^2 + 8M|\phi_{12}|^2]^{1/2}\}$$

for $M \neq 0$. We note that M and the sign of $N + 2|\phi_{12}|^2$ are geometric quantities, i.e., they are independent of the freedom in the choice of dyad that we still have. For positive M there are two solutions, only one of which is positive. For negative M there are two real and positive solutions provided the discriminant of the quadratic occurring in Eq. (2.28) is positive. (In this case we must assume that the conditions described below hold for at least one of the two solutions).

Hence for $M \leq 0$ we assume that

$$N + 2|\phi_{12}|^2 > (-8M)^{1/2}|\phi_{12}|. \tag{2.29}$$

Assuming further that $A^{1/2}$ satisfies Eq. (2.15), i.e., that

$$\begin{aligned} &D\ln A = 0, \\ &\delta\ln A = 2i\phi_{12}(1 - 1/A), \end{aligned} \tag{2.30}$$

we can make a scale transformation with $a = A^{1/2}$ and satisfy Eq. (2.27). Equation (2.9), with S_α as in Eq. (1.2) but with $\epsilon = \rho$, is now a geometric condition which we assume satisfied. Making a phase change with the solution ϕ of Eq. (2.10), we arrive at a dyad which is fixed apart from constant phase changes and restricted null rotations about k_A and in which k_A obeys the field Eqs. (1.3)–(1.5).

Apart from constant phase changes and the ambiguity in the solution of Eq. (2.28) for negative M the neutrino spinor was determined uniquely from the geometry.

The necessary and sufficient conditions were Eqs. (1.4), (2.13), (2.14), (2.25), (2.29), (2.30), and (2.9). Condition (2.9) may be replaced by the weaker one

$$S_{[\alpha;\beta]}n^\alpha m^\beta = 0$$

since the other independent components of $S_{[\alpha;\beta]}$ all vanish due to earlier assumptions.

(iii) It has been shown by Griffiths[6] that there exist neutrino fields with zero energy—momentum tensor. We shall now find necessary and sufficient conditions for a geometry with vanishing Ricci tensor to admit such a field.

Since all the components of $R_{\alpha\beta}$ vanish, it is evident from Eq. (1.5) that the null congruence associated with the neutrino flux is geodesic, shear-free, and twist-free. From the Ricci identities (2.14) and the equation

$$\mathrm{Im}\Psi_2 = 0 \qquad (2.31)$$

follow. As far as uniqueness is concerned, it is easy to check that if k_A is a neutrino spinor, so is $ae^{i\phi_0}k_A$ provided ϕ_0 is a constant and the function a satisfies Eq. (2.15) (with $\phi_{12}=0$, of course).

Conversely, if the Ricci tensor vanishes, we assume that the Weyl tensor is algebraically special and let $\{k_A, m_A\}$ be any (normalized) dyad for which k_A corresponds to a repeated principal null vector k_α (of which there are at most two) of the Weyl tensor. According to the Goldberg–Sachs theorem[3,7] the associated null congruence is geodesic and shear-free and we assume it to be twist-free as well. If we also assume the validity of Eq. (2.31), we can solve Eq. (2.26) and use the solution to transform, by means of Eq. (2.1), to a dyad in which Eqs. (2.8) hold. The additional assumption (2.9) allows us to find a spinor k_A satisfying the field equation as we did in subsection (i).

The fact that the final dyad is defined only up to null rotations about k_A, constant phase changes and scale changes restricted by Eq. (2.15) shows once again that if k_A is a neutrino spinor, so is $ae^{i\phi_0}k_A$ provided ϕ_0 is a constant and a satisfies Eq. (2.15). In summary, the necessary and sufficient conditions for such a spinor to exist are given by Eqs. (2.13), (2.31), and (2.9).

3. NEUTRINO FLUX NOT AN EIGENVECTOR OF $R_{\alpha\beta}$

The geometrization of neutrino fields for which the flux is not an eigenvector of the Ricci tensor is more difficult since it is less obvious how to pick a possible direction for the flux in the given geometry. We overcome this problem by choosing the null vector n_α conveniently and satisfying the necessary conditions

$$\phi_{00} = 0 \qquad (3.1)$$

and

$$\mathrm{Im}(i\kappa\phi_{01}^{-1}) = 0 \qquad (3.2)$$

by means of a null rotation about n_α. (κ itself cannot vanish, i.e., the null congruence associated with the neutrino flux is not geodesic.) Once this is accomplished we proceed in the usual manner.

The geometries under consideration in this section belong to group III or IV. In subsection (i) we deal with classes B_{2a} and $B_{3a,b,c,d,e}$, and in subsection (ii) we deal with the rest.

(i) A trace-free Ricci tensor belonging to class B_{2a} or $B_{3a,b,c,d,e}$ may be written[1,2]

$$R_{\alpha\beta} = T_{(\alpha}[AT_{\beta)} + Bn_{\beta)} + CV_{\beta)}] + \tfrac{1}{4}Ag_{\alpha\beta},$$

where, for $B\neq 0$, n_α is a null eigenvector orthogonal to T_α. For $B=0$ we define n_α to be one of the two null eigenvectors orthogonal to T_α and V_α. For B_{2a}, $A=C=0$, and $B\neq 0$; for $B_{3a,b,c}$, $A \lesseqgtr 0$, $B=0$, and $C\neq 0$; for $B_{3d,e}$, $A \lessgtr 0$, $C=0$, and $B\neq 0$. Without loss of generality we may assume that the neutrino flux is to be orthogonal to T_α but not to V_α (when $C\neq 0$). Clearly it cannot be in the direction of n_α since n_α is an eigenvector of $R_{\alpha\beta}$.

Therefore, we choose a dyad such that m_A corresponds to n_α and k_A to a null vector orthogonal to T_α (but not to V_α when $C\neq 0$). Relative to such a dyad the nonzero components of $R_{\alpha\beta}$ are

$$\phi_{01} = (e^{i\theta}/4\sqrt{2})[B + C(k_\alpha V^\alpha)],$$
$$\phi_{02} = \tfrac{1}{4}e^{2i\theta}[-A + iC(V_\alpha W^\alpha)],$$
$$\phi_{11} = -\tfrac{1}{8}A,$$
$$\phi_{12} = [C(n_\alpha V^\alpha)/4\sqrt{2}]e^{i\theta},$$

where W_α is given by $2^{1/2}m_\alpha = (T_\alpha + iW_\alpha)e^{i\theta}$. The dyad is defined only up to scale and phase changes, restricted null rotations about m_A, i.e.,

$$k_A \to k_A + bm_A, \quad m_A \to m_A \qquad (3.3)$$

with the parameter b restricted by

$$\bar{b} = -be^{-2i\theta}, \qquad (3.4)$$

and, when $C\neq 0$, the ambiguity in the choice of n_α and the requirement $k_\alpha V^\alpha \neq 0$.

We use the transformation given by Eqs. (3.3) and (3.4) to try and determine a null direction k_α which is not orthogonal to V_α (when $C\neq 0$) and for which Eqs. (3.1) and (3.2) are valid. The existence of such a k direction is a necessary condition in order that the source of the given geometry be a neutrino field. We shall show that this condition together with the necessary conditions given by Eqs. (1.4), (3.5), (2.7), (2.8), and (2.9) are also sufficient.

However, we must restate the condition at hand in a way which allows us to actually determine whether or not in a given geometry such a null direction exists. If for convenience we take $\theta=0$, then b is pure imaginary, i.e., $b=ib_1$, and, with the aid of the Appendix of Ref. 1, Eq. (3.2) reduces to a quartic equation in b_1 with real coefficients. If this equation has no real solutions, then the given geometry cannot have a neutrino field as source. Hence we must assume as one of our geometrization conditions that at least one of the at most four solutions be real and that for at least one of the at most four k directions which correspond to these real solutions Eq. (3.1) as well as the assumptions below are satisfied.

If we assume that

$$-\tfrac{1}{2}i\kappa\phi_{01}^{-1} > 0, \qquad (3.5)$$

we can make a change of scale with $a = [2i\phi_{01}\kappa^{-1}]^{1/2}$.

With the further assumptions (2.7), (2.8), and (2.9) we can, in the usual fashion, determine a dyad for which Eqs. (1.3)–(1.5) are satisfied. The neutrino spinor k_A is uniquely determined up to a constant phase and up to the possible ambiguity in the direction of the flux.

(ii) The procedure for the remaining geometries ($B_{2b,c}$, $B_{3f,g}$, $B_{6a,b,c}$, $D_{1c,d}$, D_2, D_{3a*}, D_{3c*}, and D_4) is only slightly different. We shall be content with illustrating the method on classes $B_{2b,c}$.

A trace-free Ricci tensor in $B_{2b,c}$ may be written[1]

$$R_{\alpha\beta}=2\lambda k_{(\alpha}n_{\beta)}+Cn_\alpha n_\beta-\tfrac{1}{2}\lambda g_{\alpha\beta},\quad \lambda \gtrless 0,\quad C\lessgtr 0,$$

where n_α is the null eigenvector of $R_{\alpha\beta}$ and the direction of the null vector k_α is also determined by the given geometry. If k_A and m_A are spinors corresponding, respectively, to k_α and n_α, the nonzero components of $R_{\alpha\beta}$ relative to the dyad $\{k_A, m_A\}$ are $\phi_{00}=-\frac{1}{2}C$ and $\phi_{11}=-\frac{1}{4}\lambda$. More generally, relative to a dyad obtained from the present one by a null rotation about m_A [cf. Eq. (3.3)] they are

$$\phi_{00}=-\tfrac{1}{2}C-\lambda(b_0^2+b_1^2),\quad \phi_{01}=-\tfrac{1}{2}\lambda b,\quad \phi_{11}=-\tfrac{1}{4}\lambda,$$

where $b=b_0+ib_1$. The necessary condition (3.2) reduces after a considerable amount of algebra to a fifth degree polynomial equation in b_0 and b_1 which can be combined with the necessary condition given by Eq. (3.1), i.e.,

$$C+2\lambda(b_0^2+b_1^2)=0,$$

to yield a quartic polynomial in b.

Therefore, there are at most four candidates for the direction of k_α and we can verify whether or not for at least one of them Eqs. (3.5), (2.7), (2.8), and (2.9) are satisfied. If so, we can in the now familiar manner determine a dyad in which Eqs. (1.3)–(1.5) hold. The spinor k_A of this dyad is then a neutrino spinor which has been determined uniquely from the given geometry except for a constant phase and the ambiguity in the choice of the direction k_α. Again, the sufficient conditions we have just found are obviously necessary as well.

*Supported in part by the National Research Council of Canada.

[1]G. Ludwig, J. Math. Phys. 15, 928 (1974).
[2]G. Ludwig and G. Scanlan, Commun. Math. Phys. 20, 291 (1971).
[3]E. Newman and R. Penrose, J. Math. Phys. 3, 566 (1962).
[4]F. A. E. Pirani, *Lectures on General Relativity, Brandeis Summer Institute*, 1964 (Prentice-Hall, Englewood Cliffs, N.J., 1965), p. 350.
[5]J. B. Griffiths, Int. J. Theor. Phys. 5, 141 (1972).
[6]J. B. Griffiths, Commun. Math. Phys. 28, 295 (1972).
[7]J. N. Goldberg and R. Sachs, Acta Phys. Polon. 22, Suppl. 13 (1962).

Applications of infinite order perturbation theory in linear systems. I

S. Y. Wu, C. C. Tung, and M. Schwartz

Department of Physics, University of Louisville, Louisville, Kentucky 40208
(Received 23 July 1973)

The applicability of infinite order perturbation theory to linear systems is exhibited. The technique involves a generalization of the method developed by Wu and Taylor and can be used to study systems described by the equations of the following form $V_{nn}u_n + V_{n,n+1}u_{n+1} + V_{n,n-1} \times u_{n-1} = E u_n$, where the coupling coefficients $V_{n'n}$'s depend on n. The wide range of application of the generalized method is demonstrated by using it to study systems as different as the plane rotator in an external field on the one hand and the dynamics of a disordered chain on the other.

I. INTRODUCTION

When perturbation theory is used to study problems in physics, in many cases it has been sufficient to carry the expansions to only the first few orders. There are other problems, however, where no meaningful results can be obtained in any finite order of perturbation, but which yield useful results when certain infinite sets of terms are summed. The method commonly used to evaluate the contributions from the infinite sets of terms is Dyson's equation.[1] It has been applied to problems which arise in various branches of physics.[2] For example, the methods of infinite order perturbation theory based on the multiple scattering theory of Lax and others have been applied to study the excitations in an alloy containing a small concentration c of substitutional defects.[3-7] In these theories, attempts were made to develop a systematic approach which could lead to a sequence of improving approximations, e.g., proper accounting of single-site approximations, pair effects, etc. On the whole, the existing perturbational calculations appear to reproduce fairly accurately spectra which are known to be reasonably smooth, but fail to obtain any detailed structure which is known to exist.[8] This is especially the case for one-dimensional alloys.[6,8]

In this work, we shall be concerned with the application of the infinite order perturbation theory to study systems which can be described by equations of the following form:

$$V_{n,n}u_n + V_{n,n+1}u_{n+1} + V_{n,n-1}u_{n-1} = Eu_n, \tag{1}$$

where the u_n's are the amplitudes for the eigenstate with eigenvalue E and the coupling coefficients $V_{n,n'}$'s depend on n. Our approach is based on the method used by Wu and Taylor.[9] Their method started with setting up a series expansion for the Green's function. They then found that the perturbation series can be summed exactly to infinite order in the case of periodic chains. Here we shall generalize their technique so that it can be applied to Eq. (1). There are many interesting physical systems which are described by this equation. We shall first illustrate the generalized method with application to a familiar physical model, the plane rigid rotator in an external field. Then we shall discuss an application of the method to study the dynamics of a linear isotopically disordered chain. Our present interest is twofold: (i) to demonstrate the wide range of applicability of the method and (ii) to demonstrate that the property of periodicity is not an essential requirement for the use of these techniques.

As it will be seen later (See Sec. IV), the method, when applied to a disordered chain of finite length, can lead to the exact summation of the infinite series of the Green's function. The calculation of the frequency spectrum of the finite chain will depend on the particular configuration of the chain. However, the formulation of the method [Eq. (11) through Eq. (13) in Sec. II] is particularly suitable for the calculation of the ensemble average of the Green's function. Thus the method can be extended to study the dynamics of a very long disordered chain, specifically to recover the fine structure of the localized modes. This latter work will be discussed in a future publication.

II. INFINITE ORDER PERTURBATION THEORY

For a system described by Eq. (1), it is more convenient to rewrite it in the form

$$\sum_{n'} D^{-1}_{nn'}u_{n'} = 0, \tag{2}$$

where

$$D^{-1}_{nn'} = (E - V_{nn})\delta_{nn'} - V_{n,n+1}\delta_{n+1,n'} - V_{n,n-1}\delta_{n-1,n'}. \tag{3}$$

The Green's function $D_{nn'}$ can then be introduced by the equation

$$\sum_{n'} D^{-1}_{nn'}D_{n'n''} = \delta_{nn''}. \tag{4}$$

From Eq. (4), we obtain

$$D_{nn'} = D^\circ(n)\delta_{nn'} + \sum_{n''} D^\circ(n)T_{nn''}D_{n''n'}, \tag{5}$$

where

$$D^\circ(n) = 1/(E - V_{nn}) \tag{6}$$

and

$$T_{nn'} = V_{n,n+1}\delta_{n+1,n'} + V_{n,n-1}\delta_{n-1,n'}. \tag{7}$$

By iteration [using Eq. (5)], we obtain a series expansion for $D_{nn'}$, namely

$$D_{nn'} = D^\circ(n)\delta_{nn'} + D^\circ(n)T_{nn'}D^\circ(n') + D^\circ(n)\sum_{n''} T_{nn''}D^\circ(n'')T_{n''n'}D^\circ(n') + \cdots. \tag{8}$$

If now we limit ourselves to diagonal elements, Eq. (8) reduces to

$$D_{nn} = D^\circ(n) + D^\circ(n)\sum_{n'} T_{nn'}D^\circ(n')T_{n'n}D^\circ(n) + \cdots, \tag{9}$$

where we have made use of the fact that $T_{nn} = 0$.

In order to sum the series of Eq. (9), we use a diagramatic method. For this purpose, let us represent $D^\circ(n)$ by a cross, $T_{n,n+1}$ by a directed line joining the nth and $(n+1)$th sites, and $T_{n,n-1}$ by a directed line joining the nth and $(n-1)$th sites. Then a term like

$$D^\circ(n)T_{n,n+1}D^\circ(n+1)T_{n+1,n+2}D^\circ(n+2)T_{n+2,n+1} \times D^\circ(n+1)T_{n+1,n}D^\circ(n)\, T_{n,n-1}D^\circ(n-1)T_{n,n-1}D^\circ(n)$$

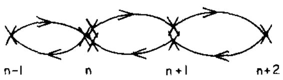


FIG. 1. Typical diagram in the series expansion (9)

may be represented by Fig. 1. Any term on the right of Eq. (9) may be interpreted as a journey starting from the "home" site, going via the transits $T_{n,n\pm 1}$ to all possible sites and finally returning to the home site. The series is thus a sum of all possible journeys. We can then rearrange the series in the following way:

$D_{nn} = D^\circ(n)+$(the sum of all those journeys, starting from n and ending at n, which never returns to the home site on the way) + (the sum of all those journeys, starting from n and ending at n, which returns once to the home site on the way) $+ \cdots$.

If $D^\circ(n)Z_n$ is the sum of all the journeys that never pass the home site n on the way, then

$$D_{nn} = D^\circ(n) + D^\circ(n)(Z_n + Z_n^2 + Z_n^3 + \cdots) = D^\circ(n)/(1 - Z_n). \tag{10}$$

We now define $D^\circ(n)Z_{n,n+1}$ as the sum of all those journeys, starting from n and ending at n, but going in the direction defined by an increase of n, which never return to the home site on the way. Similarly, $D^\circ(n)Z_{n,n-1}$ defines the sum for journeys going in a direction given by a decrease in n. Since $Z_n = Z_{n,n+1} + Z_{n,n-1}$, Eq. (10) becomes

$$D_{nn} = D^\circ(n)/(1 - Z_{n,n+1} - Z_{n,n-1}). \tag{11}$$

In a similar manner, we sum the series representing $Z_{n,n+1}$. We obtain

$$Z_{n,n+1} = V_{n,n+1}D^\circ(n+1)[1 + Z_{n+1,n+2} + \cdots]V_{n+1,n}D^\circ(n) = V_{n,n+1}V_{n+1,n}D^\circ(n)D^\circ(n+1)/(1 - Z_{n+1,n+2}). \tag{12}$$

Likewise

$$Z_{n,n-1} = V_{n,n-1}V_{n-1,n}D^\circ(n)D^\circ(n-1)/(1 - Z_{n-1,n-2}). \tag{13}$$

We shall have use, in addition, for the off-diagonal element $D_{nn'}$. We consider first the simple case $D_{n,n+1}$. Following the same argument as before, we can interpret $D_{n,n+1}$ as the sum of all those journeys starting from the home site n, but ending at site $n+1$. As shown in Fig. 2, the dot at n represents D_{nn}; the directed line from the nth site to the $(n+1)$th site, $V_{n,n+1}$; and $\Delta^{\pm}_{n+1}$ at the $(n+1)$th site, the sum of all those journeys, starting from $n+1$ and ending at $n+1$, but going only in the direction of increasing (or decreasing) n. In the present case,

$$D_{n,n+1} = D_{nn}V_{n,n+1}\Delta^+_{n+1}. \tag{14}$$

Equation (14) implies the following. First we make all possible journeys starting from the home site n and ending at n. Then we go to the site $n+1$ and there make those journeys going away only in one direction. Since

$$\Delta^+_{n+1} = D^\circ(n+1) + D^\circ(n+1)(Z_{n+1,n+2} + Z^2_{n+1,n+2} + \cdots) = D^\circ(n+1)(1 - Z_{n+1,n+2}), \tag{15}$$

it follows that

$$D_{n,n+1} = V_{n,n+1}D_{nn}D^\circ(n+1)/(1 - Z_{n+1,n+2}). \tag{16}$$

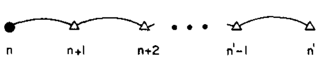


FIG. 2. Diagrams representing $D_{n,n'}$.

For the general case $D_{nn'}$, as shown in Fig. 2, we must consider $\Delta_{n+1}, \Delta_{n+2}, \cdots, \Delta_{n'}$, that is,

$$D_{nn'} = D_{nn}V_{n,n+1}\Delta^+_{n+1}V_{n+1,n+2}\Delta^+_{n+2}\cdots V_{n'-1,n'}\Delta^+_{n'} \quad \text{for } n' > n. \tag{17}$$

Similarly

$$D_{nn'} = D_{nn}V_{n,n-1}\Delta^-_{n-1}V_{n-1,n-2}\Delta^-_{n-2}\cdots V_{n'+1,n'}\Delta^-_{n'} \quad \text{for } n' < n. \tag{18}$$

III. EXAMPLE: A PLANE RIGID ROTATOR IN AN EXTERNAL FIELD

The Schrödinger equation for a plane rotator in an external field can be written as follows:

$$\frac{d^2\psi}{d\phi^2} + \frac{8\pi^2 I}{h^2}(W + \mu F\cos\phi)\psi = 0, \tag{19}$$

where μ is the electric moment, I the moment of inertia, and F the electric field intensity. If we let

$$E \equiv (8\pi^2 I/h^2)W, \qquad G \equiv (4\pi^2 I/h^2)\mu F,$$

Eq. (19) becomes

$$\frac{d^2\psi}{d\phi^2} + (E - 2G\cos\phi)\psi = 0. \tag{20}$$

Equation (20) is a special case of the Mathieu's equation.[10] However, it is our purpose here to illustrate the applicability of infinite perturbation theory to determine the eigenenergies. For this purpose, we note that, in the unperturbed state, the eigenfunctions are

$$\psi^\circ_n = \frac{1}{(2\pi)^{1/2}}e^{in\phi}, \qquad n = 0, \pm 1, \cdots. \tag{21}$$

When ψ°_n is operated on by $d^2/d\phi^2 + (E - 2G\cos\phi)$, we obtain

$$(E - n^2)\psi^\circ_n - G\psi^\circ_{n+1} - G\psi^\circ_{n-1} = 0. \tag{22}$$

We can now express Eq. (22) in the form

$$\sum_{n'} D^{-1}_{nn'}\psi^\circ_{n'} = 0, \tag{23}$$

where

$$D^{-1}_{nn'} = (E - n^2)\delta_{nn'} - G\delta_{n+1,n'} - G\delta_{n-1,n'}. \tag{24}$$

By comparing Eq. (24) with Eq. (13), we see that

$$D^\circ(n) = 1/(E - n^2) \tag{25}$$

and

$$V_{n,n+1} = V_{n,n-1} = G. \tag{26}$$

We now show that

$$D_{nn} = D_{-n,-n}. \tag{27}$$

From Eqs. (11), (12), and (13), it is seen that

$$\begin{aligned} D_{-n,-n} &= D^\circ(-n)/(1 - Z_{-n,-n+1} - Z_{-n,-n-1}) \\ &= D^\circ(-n)/(1 - Z_{-n,-(n-1)} - Z_{-n,-(n+1)}) \\ &= D^\circ(n)/(1 - Z_{n,n-1} - Z_{n,n+1}) = D_{n,n} \end{aligned}$$

since the quantities depend on n through $D^\circ(n)$ which depends on n^2.

By substituting Eq. (26) into Eqs. (17) and (18), we obtain

$$D_{nn'} = G^{(n'-n)} D_{nn} \frac{D^\circ(n+1)}{1-Z_{n+1,n+2}} \frac{D^\circ(n+2)}{1-Z_{n+2,n+3}} \cdots \times \frac{D^\circ(n')}{1-Z_{n',n'+1}} \quad (n'>n)$$

$$= G^{(n-n')} D_{nn} \frac{D^\circ(n-1)}{1-Z_{n-1,n-2}} \frac{D^\circ(n-2)}{1-Z_{n-2,n-3}} \cdots \times \frac{D^\circ(n')}{1-Z_{n',n'-1}} \quad (n'<n). \tag{28}$$

By arguments similar to those given to prove Eq. (27), it follows that

$$D_{-n,n} = D_{n,-n}. \tag{29}$$

In order to compute the energy eigenvalues, we determine the poles of the diagonal elements of the Green's function (see, for example, Kumar[11]). Since $D_{nn} = D_{-n,-n}$, therefore, to determine the perturbed energies for the rigid rotator, we need to diagonalize the submatrix

$$\begin{pmatrix} D_{nn} & D_{n,-n} \\ D_{-n,n} & D_{-n,-n} \end{pmatrix}. \tag{30}$$

From Eqs. (27) and (29), it follows that the matrix is symmetric and its diagonalized form is

$$\begin{pmatrix} D_{nn} + D_{n,-n} & 0 \\ 0 & D_{nn} - D_{n,-n} \end{pmatrix}. \tag{31}$$

The perturbed energies are then the poles of the diagonal elements or the zeros of $(D_{nn} \pm D_{n,-n})^{-1}$.

Equation (28) yields for $D_{nn} \pm D_{n,-n}$

$$D_{nn} \pm D_{n,-n} = D_{nn}\left[1 \pm G^{2n} \frac{D^\circ(n-1)}{1-Z_{n-1,n-2}} \frac{D^\circ(n-2)}{1-Z_{n-2,n-3}} \cdots \times \frac{D^\circ(0)}{1-Z_{0,1}} \frac{D^\circ(1)}{1-Z_{1,2}} \cdots \frac{D^\circ(n)}{1-Z_{n,n+1}}\right]. \tag{32}$$

$D_{nn} \pm D_{n,-n} \simeq D_{nn}[1 \pm G^{2n} D^\circ(n) D^\circ(n-1)^2 D^\circ(n-2)^2 \cdots D^\circ(1)^2 D^\circ(0)]$ to the $(2n)$th order in G. The zeros of $(D_{nn} \pm D_{n,-n})^{-1}$ are

$$D_{nn}^{-1}[1 \mp G^{2n} D^\circ(n) D^\circ(n-1)^2 D^\circ(n-2)^2 \cdots D^\circ(1)^2 D^\circ(0)] = 0, \tag{33}$$

where the approximation $(1-r)^{-1} \simeq 1+r$ for $r \ll 1$ is used here and also later. From Eq. (11), we obtain

$$D_{nn}^{-1} = D^\circ(n)^{-1}[1 - Z_{n,n-1} - Z_{n,n+1}]. \tag{34}$$

Application of Eqs. (12) and (13) to Eq. (34) yields

$$D_{nn}^{-1} \simeq D^\circ(n)^{-1}\left\{1 - \cfrac{G^2 D^\circ(n) D^\circ(n-1)}{1 - \cfrac{G^2 D^\circ(n-1) D^\circ(n-2)}{\ddots \; \overline{1 - G^2 D^\circ(1) D^\circ(0)}}} - \cfrac{G^2 D^\circ(n) D^\circ(n+1)}{1 - \cfrac{G^2 D^\circ(n+1) D^\circ(n+2)}{\ddots \; \overline{1 - G^2 D^\circ(2n-1) D^\circ(2n)}}}\right\}. \tag{35}$$

From Eqs. (33) and (35), we obtain

$$D^\circ(n)^{-1} = \left\{\cfrac{G^2 D^\circ(n-1)}{1 - \cfrac{G^2 D^\circ(n-1) D^\circ(n-2)}{\ddots \; \overline{1 - G^2 D^\circ(1) D^\circ(0)}}} + \cfrac{G^2 D^\circ(n+1)}{1 - \cfrac{G^2 D^\circ(n+1) D^\circ(n+2)}{\ddots \; \overline{1 - G^2 D^\circ(2n-1) D^\circ(2n)}}} \pm \frac{G^{2n} D^\circ(n-1)^2 D^\circ(n-2)^2 \cdots}{D^\circ(1)^2 D^\circ(0)}\right\}. \tag{36}$$

Since $D^\circ(n)^{-1} = E - n^2$, it is evident that the removal of the degeneracy for E_n° depends on the last term containing G^{2n}.

We shall now consider special solutions of Eq. (36) for the cases $n = 1$ and 2. For $n = 1$, we have

$$E - 1 \simeq G^2[D^\circ(0) + D^\circ(2) \pm D^\circ(0)].$$

Since $D^\circ(0) = E^{-1}$ and $D^\circ(2) = (E-4)^{-1}$, we put $E = 1$ in these expressions on the right and obtain

$$E^+ = 1 + \frac{5}{3} G^2 \text{ and } E^- = 1 - G^2/3$$

in accordance with the values determined by Schwartz and Martin.[12]

For $n = 2$, Eq. (36) reduces to

$$E - 4 = \frac{G^2 D^\circ(1)}{1 - G^2 D^\circ(1) D^\circ(0)} + \frac{G^2 D^\circ(3)}{1 - G^2 D^\circ(3) D^\circ(4)} \pm G^4 D^\circ(1)^2 D^\circ(0).$$

Using the approximation $(1-r)^{-1} \simeq 1 + r$, we obtain in this case

$$E - 4 = G^2 D^\circ(1)[1 + G^2 D^\circ(1) D^\circ(0)] + G^2 D^\circ(3)[1 + G^2 D^\circ(3) D^\circ(4)] \pm G^4 D^\circ(1)^2 D^\circ(0).$$

This expression can be written so

$$E^\pm = 4 + G^2[D^\circ(1) + D^\circ(3)] + G^4 \begin{Bmatrix} 2D^\circ(1)^2 D^\circ(0) + D^\circ(3)^2 D^\circ(4) \\ D^\circ(3)^2 D^\circ(4) \end{Bmatrix}.$$

We consider first E^+. We have

$$E^+ = 4 + G^2\left(\frac{1}{E-1} + \frac{1}{E-9}\right) + G^4\left[2 \frac{1}{(E-1)^2} \frac{1}{E} + \frac{1}{(E-9)^2} \frac{1}{(E-16)}\right]$$

Up to G^2, putting $E = 4$, we obtain $E \simeq 4 + \frac{2}{15}G^2$ which must be inserted into the first bracket on the right for E to be correct to G^4. Hence

$$E^+ = 4 + G^2\left(\frac{1}{4 + \frac{2}{15}G^2 - 1} + \frac{1}{4 + \frac{2}{15}G^2 - 9}\right) + G^4\left(\frac{2}{9 \times 4} - \frac{1}{25 \times 12}\right) = 4 + \tfrac{2}{15}G^2 + \tfrac{433}{13500}G^4.$$

Likewise for E^- we obtain

$$E^- = 4 + \tfrac{2}{15}G^2 - \tfrac{317}{13500}G^4.$$

These examples indicate the pattern of reduction of the fractions to the required order in G and the method of evaluation of the respective energy levels. While the procedures remain the same, the reductions for higher orders become more tedious. Nevertheless, by carrying the analysis to the computational level, we have demonstrated the applicability of infinite order perturbation theory to the solution of the problem of the rigid rotator in an electric field. Furthermore, we have shown that the removal of degeneracy can be immediately detected by examination of Eq. (36). Specifically, from Eq. (36), it is obvious that, for E_n°, the lifting of the degeneracy can only be accomplished if the perturbation is carried to the $(2n)$th power in G, a conclusion which is not readily apparent from the conventional method (See Appendix A).

IV. DYNAMICS OF A LINEAR ISOTOPICALLY DISORDERED CHAIN

Lattice dynamics of disordered systems has been a field of research interest since about two decades ago.[13] The problem involves two distinct physical situations, the glasslike disordered systems and alloys. The main difficulty of the problem is the absence of periodicity either in the geometric structure or in the atomic composition, so that methods derived for the crystalline solids are no longer appropriate. In recent years, much thought has been given in searching for a method that could generally be applied to the disordered systems.[4-7,14,15] Because of the complexity of the mathematics, although a variety of new methods have been devised and used, extremely effective methods have not yet been found.

Even though it is still desirable to continue the research for a general theory for the more realistic models, it has become apparent that an exact or near exact treatment of simple models can provide a qualitative understanding of the effects of disorder. A study of a one-dimensional disordered system is just such a case. The one-dimensional problem can be treated by both numerical and analytical methods. In recent years, there were several different approaches which had yielded good results.[16] One of the most successful methods in terms of actual results is the direct numerical calculation devised by Dean[17] for the vibrational spectra of disordered systems. Dean's results have provided us with an insight into the vibrational properties of disordered systems. For example, the frequency spectrum has a fine structure with many distinct peaks and valleys, and some normal modes of vibrations are strongly localized.

On the analytical side, progress has been relatively slow. The approach usually started with setting up a series for the Green's function. Then various schemes were designed to obtain the ensemble average of the "Greenian," from which the frequency spectrum was calculated.[14] Because of the approximation involved, most applications were confined to special systems. The results, in most cases, were not entirely satisfactory.[16]

In this section we wish to discuss the application of the method developed in Sec. II to study the dynamics of a linear isotopically disordered system. The technique can lead to the exact summation of the infinite series of the Green's functions for a chain of finite length, which in turn leads to the exact calculation of the frequency spectrum. Our main interest in this case is to learn, through the study of the finite chain, how the limit of the frequency spectrum corresponding to an infinite disordered chain is reached when the number of atoms in the chain is increased.

The time independent equation of motion for an isotopically disordered chain can be written as

$$-m_l\omega^2 u_l = -2\gamma u_l + \gamma u_{l+1} + \gamma u_{l-1}, \tag{37}$$

where u_l is the displacement of the lth atom, m_l the mass of the lth atom, ω^2 the angular frequency, and γ the nearest neighbor force constant.

Equation (37) can also be rewritten as

$$\sum_{l'} D_{ll'}^{-1} u_{l'} = 0,$$

where

$$D_l^{\circ -1} = 2\gamma/m_l - \omega^2 \tag{38}$$

and

$$T_{ll'} = (\gamma/m_l)\delta_{l+1,l'} + (\gamma/m_l)\delta_{l-1,l'}. \tag{39}$$

Thus from Eq. (12) we obtain

$$Z_{l,l+1} = \left(\frac{\gamma}{m_l}\right)\left(\frac{\gamma}{m_{l+1}}\right)\frac{D_l^\circ D_{l+1}^\circ}{1 - Z_{l+1,l+2}}, \tag{40a}$$

$$Z_{l+1,l+2} = \left(\frac{\gamma}{m_{l+1}}\right)\left(\frac{\gamma}{m_{l+2}}\right)\frac{D_{l+1}^\circ D_{l+2}^\circ}{1 - Z_{l+2,l+3}}, \tag{40b}$$

$$\cdots,$$

from Eq. (13) we have

$$Z_{l,l-1} = \left(\frac{\gamma}{m_l}\right)\left(\frac{\gamma}{m_{l-1}}\right)\frac{D_l^\circ D_{l-1}^\circ}{1 - Z_{l-1,l-2}}, \tag{41a}$$

$$Z_{l-1,l-2} = \left(\frac{\gamma}{m_{l-1}}\right)\left(\frac{\gamma}{m_{l-2}}\right)\frac{D_{l-1}^\circ D_{l-2}^\circ}{1 - Z_{l-1,l-2}}, \tag{41b}$$

$$\cdots,$$

and from Eq. (11) we have

$$D_{ll} = \frac{D_l^\circ}{1 - Z_l} = \frac{D_l^\circ}{1 - Z_{l,l+1} - Z_{l,l-1}}. \tag{42}$$

We shall discuss two different boundary conditions for the system.

A. Fixed-ends boundary condition

For a chain with N atoms, we connect the first and the last atoms on each side to rigid walls. The boundary conditions are then $u_0 = 0$ and $u_{N+1} = 0$. The equations of motion for the two side atoms are

$$-m_1\omega^2 u_1 = -2\gamma u_1 + \gamma u_2 \tag{43}$$

and

$$-m_N\omega^2 u_N = -2\gamma u_N + \gamma u_{N-1}. \tag{44}$$

With these boundary conditions, $T_{ll'}$ becomes

$$T_{ll'} = (\gamma/m_l)(1-\delta_{l,N})\delta_{l+1,l'} + \frac{\gamma}{m_l}(1-\delta_{l,1})\delta_{l-1,l'} \tag{45}$$

so that we obtain

$$T_{N,N+1} = 0 \quad \text{and} \quad T_{1,0} = 0. \tag{46}$$

By applying Eq. (46), Eqs. (40) and (41) become

$$Z_{1,2} = \left(\frac{\gamma}{m_1}\right)\left(\frac{\gamma}{m_2}\right)\frac{D_1^\circ D_2^\circ}{1-Z_{2,3}}, \tag{47a}$$

$$Z_{2,3} = \left(\frac{\gamma}{m_2}\right)\left(\frac{\gamma}{m_3}\right)\frac{D_2^\circ D_3^\circ}{1-Z_{3,4}}, \tag{47b}$$

$$\cdots,$$

$$Z_{N-2,N-1} = \left(\frac{\gamma}{m_{N-2}}\right)\left(\frac{\gamma}{m_{N-1}}\right)\frac{D_{N-2}^\circ D_{N-1}^\circ}{1-Z_{N-1,N}}, \tag{47c}$$

$$Z_{N-1,N} = \left(\frac{\gamma}{m_{N-1}}\right)\left(\frac{\gamma}{m_N}\right) D_{N-1}^\circ D_N^\circ, \tag{47d}$$

$$Z_{N,N+1} = 0 \tag{47e}$$

and

$$Z_{N,N-1} = \left(\frac{\gamma}{m_N}\right)\left(\frac{\gamma}{m_{N-1}}\right)\frac{D_N^\circ D_{N-1}^\circ}{1-Z_{N-1,N-2}}, \tag{48a}$$

$$\cdots,$$

$$Z_{3,2} = \left(\frac{\gamma}{m_3}\right)\left(\frac{\gamma}{m_2}\right)\frac{D_3^\circ D_2^\circ}{1-Z_{2,1}}, \tag{48b}$$

$$Z_{2,1} = \left(\frac{\gamma}{m_2}\right)\left(\frac{\gamma}{m_1}\right) D_2^\circ D_1^\circ, \tag{48c}$$

$$Z_{1,0} = 0. \tag{48d}$$

From Eqs. (47) and (48), we obtain the terms $Z_{l,l+1}$ and $Z_{l,l-1}$ ($l = 1, \cdots, N$) simply by substituting the equations in reverse order. The diagonal elements of the Green's function D_{ll} can then be obtained from Eq. (42).

The eigenfrequencies of the system can be calculated by determining the poles of the diagonal elements of the Green's function. From Eq. (42), the problem then reduces to finding the zeros of the following equation:

$$D_l^{\circ -1} - D_l^{\circ -1}(Z_{l,l+1} + Z_{l,l-1}) = 0. \tag{49}$$

From Eqs. (38), (47) and (48), it is seen that Eq. (49) can be converted into a polynomial of ω^2 of Nth degree. The N solutions of Eq. (49) are the N eigenfrequencies of the system. Thus for a chain with N atoms, there are altogether N normal modes of vibration. From these eigenfrequencies, the histogram of the frequency spectrum can also be obtained.

B. Cyclic boundary condition

For an isotopically disordered chain of N atoms, the cyclic boundary condition is expressed as

$$m_{l+N} = m_l. \tag{50}$$

Equations (40) and (41) then become

$$Z_{1,2} = \left(\frac{\gamma}{m_1}\right)\left(\frac{\gamma}{m_2}\right)\frac{D_1^\circ D_2^\circ}{1-Z_{2,3}}, \tag{51a}$$

$$\cdots,$$

$$Z_{N-1,N} = \left(\frac{\gamma}{m_{N-1}}\right)\left(\frac{\gamma}{m_N}\right)\frac{D_{N-1}^\circ D_N^\circ}{1-Z_{N,N+1}}, \tag{51b}$$

$$Z_{N,N+1} = Z_{N,1} = \left(\frac{\gamma}{m_N}\right)\left(\frac{\gamma}{m_1}\right)\frac{D_N^\circ D_1^\circ}{1-Z_{1,2}} \tag{51c}$$

and

$$Z_{N,N-1} = \left(\frac{\gamma}{m_N}\right)\left(\frac{\gamma}{m_{N-1}}\right)\frac{D_N^\circ D_{N-1}^\circ}{1-Z_{N-1,N-2}}, \tag{52a}$$

$$\cdots,$$

$$Z_{2,1} = \left(\frac{\gamma}{m_2}\right)\left(\frac{\gamma}{m_1}\right)\frac{D_2^\circ D_1^\circ}{1-Z_{1,N}}, \tag{52b}$$

$$Z_{1,N} = \left(\frac{\gamma}{m_1}\right)\left(\frac{\gamma}{m_N}\right)\frac{D_1^\circ D_N^\circ}{1-Z_{N,N-1}}. \tag{52c}$$

By setting

$$A_l = \left(\frac{\gamma}{m_l} D_l^\circ\right)^{-1}, \tag{53}$$

$$x_l = \left(\frac{\gamma}{m_l} D_l^\circ\right)^{-1} Z_{l,l+1}, \tag{54}$$

and

$$y_l = \left(\frac{\gamma}{m_l} D_l^\circ\right)^{-1} Z_{l,l-1}, \tag{55}$$

we obtain the following two equations

$$x_N = \cfrac{1}{A_1 - \cfrac{1}{A_2 - \cfrac{1}{A_3 - \cdots \cfrac{\cdots}{A_{N-1} - \cfrac{1}{A_N - x_N}}}}}, \tag{56}$$

$$y_1 = \cfrac{1}{A_N - \cfrac{1}{A_{N-1} - \cfrac{1}{A_{N-2} - \cdots \cfrac{\cdots}{A_2 - \cfrac{1}{A_1 - y_1}}}}}. \tag{57}$$

By rearranging terms, Eq. (56) or Eq. (57) can be transformed into a simple quadratic equation in x_N (or y_1), i.e.,

$$Px_N^2 + Qx_N + R = 0, \tag{58}$$

where P, Q, and R are functions of ω^2 (see Appendix B). The solutions to these equations then lead to the determination of all the Z's and D_{ll}'s. The frequency spectrum of the system can be obtained from the trace of the matrix D by

$$G(\omega^2) = -\frac{1}{\pi} \lim_{\substack{L\to\infty \\ \epsilon\to 0}} \frac{1}{L} \operatorname{Im} \operatorname{Tr} D(\omega^2 + i\epsilon). \tag{59}$$

Since, in arriving at Eq. (58), no approximation has been used, the frequency spectrum computed by using Eqs. (51), (52), (56), (57), and (59) is then an exact calculation.

V. CALCULATIONS AND DISCUSSION

A. The dynamics of a chain of 22 light atoms and 28 heavy atoms

When the fixed-ends boundary condition is applied to a finite chain of N atoms, the system can be con-

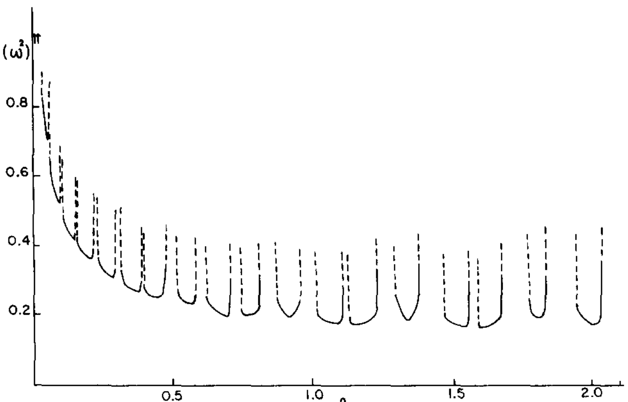


FIG. 3. Part of the frequency spectrum of a disordered chain of 22 heavy and 28 light atoms of mass ratio 3 : 1.

sidered as an isolated giant N-atomic "molecule," and there exist N discrete eigenfrequencies [see Eq. (49)]. By introducing the cyclic boundary condition, however, the system can now be construed as a one-dimensional periodic chain composed of identical giant N-atomic "molecules." In the presence of the other "molecules," the eigenfrequencies of the isolated N-atomic "molecule" are expected to be perturbed. In order to explore the relationship between these two cases, we studied a diatomic linear chain of 22 heavy (H) and 28 light (L) atoms with mass ratio 3 : 1. This model chain was first investigated by Dean.[18] Using his direct numerical method, Dean determined the histogram of the vibrational frequency spectrum of the system. He had also computed the configurations of the normal modes of vibration. The most interesting conclusions drawn from Dean's calculation were:

(a) At low frequencies the modes of vibration are wavelike in character and similar to the waves in the periodic systems and (b) at high frequencies the modes are strongly localized; the transition from extended modes to localized modes occurs at the 26th mode.

The configuration of the chain was given as follows:

HLLLHLLHHHHHLLLLLLLHLHHLHHLHLHHLLLLL
HHLHLLLHHHLHL.

In our calculation, the fixed-ends boundary condition was first used and the 50 squared eigenfrequencies were determined by using Eq. (49) (Table I). Next the cyclic boundary condition was applied. The exact frequency spectrum for the system was computed using Eqs. (51), (52), (56), (57), and (59). The result (Fig. 3) showed that the frequency spectrum was composed of 50 bands, each identifiable with one of the discrete eigenfrequencies obtained by the fixed-ends boundary condition. As pointed out before, the system can be considered as a chain of identical N-atomic "molecules" with the introduction of the cyclic boundary condition. The interactions between the "molecules" then lead to the spreading of the N eigenfrequencies of the isolated "molecule" (corresponding to the fixed-end boundary condition) into N frequency bands.[19] Since the area under each band corresponds to the number of "molecules" in the system, the area covered by each band should be the same. Therefore, the "height" of each band should be roughly inversely proportional to the width of the band. However, the band width, which arises as a result of perturbations of eigenfrequencies of a molecule due to the presence of the other molecules, depends on the degree of overlapping of the normal modes of vibration between neighboring molecules. For extended (wavelike) modes, these overlappings are large, so that the band widths should be wide. Consequently, the bands will be "low." On the other hand, overlappings are small in the case of localized modes and the bands become narrow and high. This is indeed substantiated by the results of our calculations. The frequency bands from the 1st up to the 25th are comparatively wide and low, indicating that the vibration modes for these frequency bands are extended modes, while the bands from the 26th on are narrow and high corresponding to localized modes. The distinction between the 25th and the 26th modes (see Fig. 4) of the chain is indeed very outstanding, signifying the transition from the extended modes to the localized modes. It should be noted here that this picture is in agreement with Dean's numerical calculation.

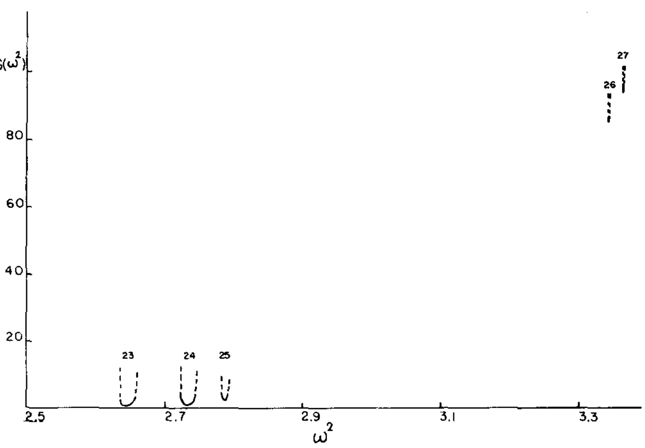


FIG. 4. Part of the frequency spectrum of a disordered chain of 22 heavy and 28 light atoms of mass ratio 3 : 1.

TABLE I. Squared eigenfrequencies for a chain of 22 heavy and 28 light atoms of mass ratio 3 : 1.

Mode	Squared eigenfrequency ω^2	Mode	Squared eigenfrequency ω^2
1	0.006 101 949	26	3.346 300 031
2	0.024 889 029	27	3.372 251 421
3	0.048 228 798	28	3.699 945 229
4	0.099 726 283	29	3.848 558 904
5	0.159 512 345	30	4.360 856 203
6	0.206 974 269	31	5.289 087 201
7	0.292 219 039	32	6.278 387 662
8	0.418 084 996	33	6.410 838 541
9	0.502 648 743	34	6.483 558 854
10	0.601 347 093	35	6.646 408 020
11	0.734 620 530	36	6.670 861 626
12	0.823 075 912	37	6.966 198 211
13	0.973 679 948	38	7.209 902 844
14	1.062 765 390	39	7.340 066 012
15	1.235 319 026	40	7.454 625 914
16	1.387 323 322	41	7.480 533 131
17	1.453 236 970	42	7.627 088 953
18	1.834 712 859	43	9.161 587 033
19	2.009 506 014	44	9.220 072 076
20	2.195 096 297	45	9.404 731 858
21	2.449 245 280	46	10.434 452 225
22	2.510 155 887	47	10.441 714 382
23	2.653 679 749	48	10.667 368 679
24	2.765 475 376	49	11.253 562 390
25	2.827 320 596	50	11.656 096 880

B. The frequency spectrum of a chain of 40 atoms with a 5% concentration of light atoms

The dynamics of all the different configurations of a simple diatomic chain of 40 atoms with a 5% concentration of light atoms was studied using the cyclic boundary conditions. The models consist of two light atoms randomly distributed among 38 heavy atoms. Under cyclic boundary conditions, there are altogether 20 possible distinct configurations. For convenience, we shall use a pair of numbers in parentheses to indicate the positions

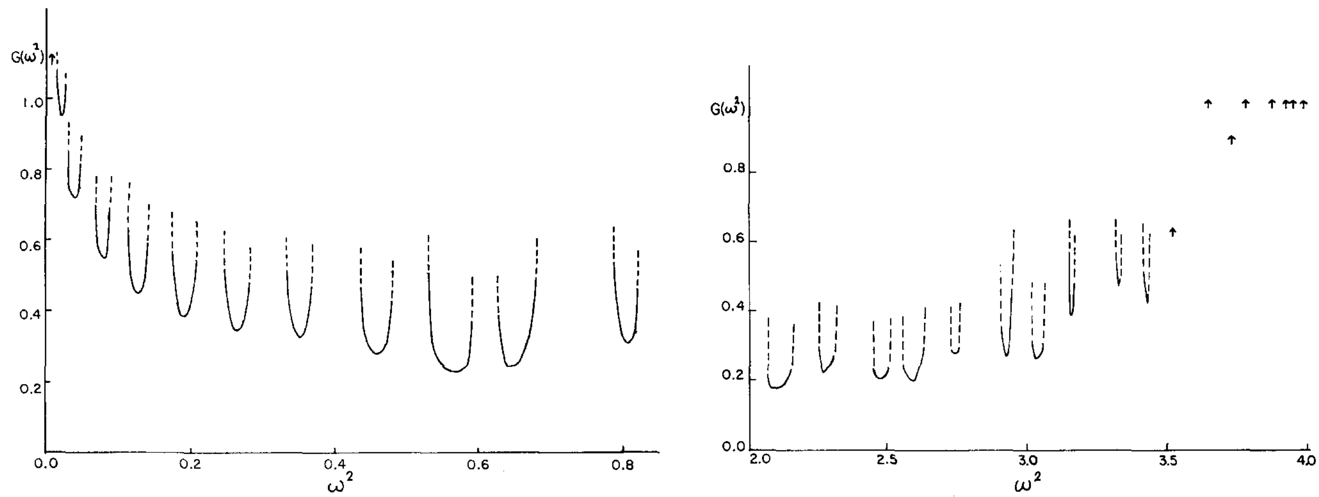


FIG. 5. (a) & (b) Frequency spectrum for the configuration (1, 14) of a chain of 38 heavy and two light atoms with mass ratio 2 : 1 (arrows indicating the locations of the narrow bands).

of the two light atoms along the chain. Thus, the twenty different configurations are (1, 2), (1, 3), $\cdots$, (1, 21).

It is because of the limited number of distinct configurations that this model was chosen for study. In order to investigate the dependence of the vibrational properties of the chain on the configurations, the exact frequency spectra of all the twenty different configurations were computed, using Eqs. (51), (52), (56), (57), and (59). Figure 5 is a typical result of the frequency spectrum for the configuration (1, 14). In this case, the frequency spectrum is formed by 40 distinct bands with forbidden gaps between them (two isolated narrow bands at the high frequency end are not shown in Fig. 5).

Since as the number of atoms in the "molecule" tends to infinity, the envelope obtained by connecting the minima of the bands will approach a limiting curve corresponding to the real frequency spectrum of an infinite disordered chain with the same concentration of light atoms. Thus, to understand the realistic situations of very long chains, it is therefore helpful to study the envelopes of the minima of the bands for all the configurations of our model. It turns out that for most parts these envelopes are smooth curves although for some configurations there are fine structures at the high frequency end. Figures 6 and 7 are the examples of the envelopes of the exact frequency spectra for configurations (1, 2), (1, 4), (1, 6), and (1, 17). From these figures, it is seen that, at the very low frequency end, the envelopes of the spectra are essentially the same. However, they begin to show fluctuations at the intermediate frequencies and the fluctuations are dependent upon configurations. Apparently, these fluctuations are caused by the statistics of the finite number of atoms in the chain. Realizing that a finite chain of 40 atoms with two light atoms can only provide very limited samples to the real situation of a one-dimensional random binary alloy with 5% light impurities, we therefore averaged the envelopes of all the 20 different configurations. The result (Fig. 8) is indeed similar in form to the frequency spectrum of Dean's numerical calculation.[20]

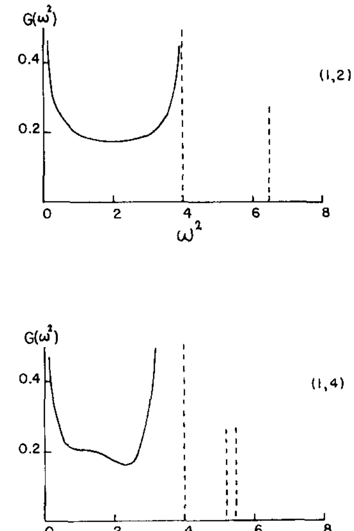


FIG. 6. Envelopes of the frequency spectra for (a) configuration (1, 2) and (b) configuration (1, 4) (dotted lines indicating the localized high frequency modes).

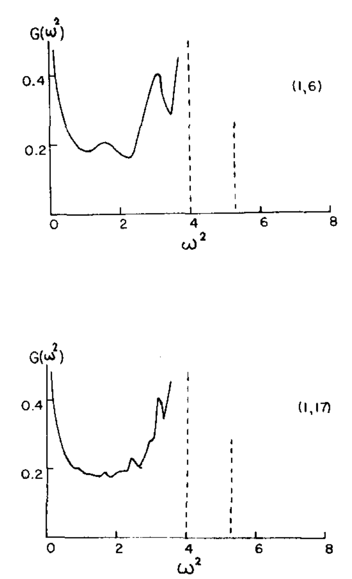


FIG. 7. Envelopes of the frequency spectra for (a) configuration (1, 17) (dotted lines indicating the localized high frequency modes).

VI. CONCLUSIONS

In this work we have generalized the method of Wu and Taylor to study systems which can be described by Eq. (1). We have also demonstrated the wide range of applicability of these techniques by the successful application of the generalized method to study systems as

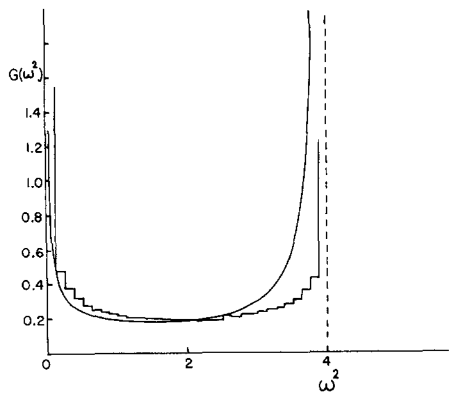


FIG. 8. Average of the envelopes of the frequency spectra (from $\omega^2 = 0.0$ to $\omega^2 = 4.0$) for all possible configurations of a chain of 38 heavy and two light atoms with mass ratio 2 : 1. (The histogram is the frequency spectrum from Dean's mass ratio 1 : 2 impurity concentration.

different as the plane rotator in an external field on the one hand and the dynamics of a disordered chain on the other. In the former case, the application of this method showed that the lifting of the degeneracy for the state E_n° can only be accomplished if the perturbation is carried to the $(2n)$th power in G, a conclusion not readily apparent from the conventional method. In the latter case, the method provides a means to calculate exactly the frequency spectrum of a polyatomic chain regardless of its complexity (See Sec. IVB). By analyzing the results obtained for a disordered diatomic chain, it is learned that the frequency spectrum is fairly stable at low frequencies, but depends on the configuration at the high frequency end. Our selection of 40 atoms is of course too "short" to provide enough statistics for the limiting curve of the spectrum of the infinite disordered chain. However, the conclusions drawn from the study of short chains indicate that as the number of atoms in the "molecule" of the chain increases, the envelope to the frequency spectrum will approach the frequency spectrum of an infinite disordered chain with the same concentration of impurities. Mathematically, there is no difficulty in extending the calculation to treat long chains. Work is currently in progress.

When the method is applied to a disordered chain, the calculation will necessarily depend on the particular configuration of the chain. However, the formulation of the method, in particular Eqs. (11)-(13), provides a framework through which the ensemble average of the Green's function can be calculated in a systematic way. Work is also in progress along this line.

ACKNOWLEDGMENTS

We express our appreciation to our colleagues, Dr. J. S. Chalmers and Dr. R. E. Mills, for stimulating discussions.

APPENDIX A

The problem of the plane rigid rotator in an external electric field can also be approached by using solutions of the form $\cos n\phi$ and $\sin n\phi$ as described below. Since the operator $d^2/d\phi^2 + (E - 2G\cos\phi)$ is an even function of ϕ, it follows that if $\psi(\phi)$ is a solution of Eq. (20), then so is $\psi(-\phi)$. This means $\psi(\phi) \pm \psi(-\phi)$ are solutions of Eq. (20). We may conclude that Eq. (20) has both even and odd solutions.

To obtain the eigenvalues for the even solutions, we assume

$$\psi = \sum_{n=1}^{\infty} A \cos n\phi. \tag{A1}$$

Substitution of Eq. (A1) into Eq. (20) leads to the eigenvalue solutions in continued fraction form

$$E = \frac{2G^2}{E - 1\,\dfrac{G^2}{E-4\cdots}}. \tag{A2}$$

Similarly, to obtain the corresponding equation for the odd solutions, we assume

$$\psi = \sum_{n=1}^{\infty} B_n \sin n\phi \tag{A3}$$

which leads to the equation

$$E = 1 + \frac{G^2}{E - 4 - \dfrac{G^2}{E-9\cdots}}. \tag{A4}$$

The eigenvalues for the perturbed states associate with the original degenerate state E_n° are to be determined from Eqs. (A2) and (A4).

APPENDIX B: FORMULA FOR x_N AND y_1 OF A CONTINUED FRACTIONAL FUNCTION

Define a continued fractional function $x_i'(A_1, A_2, \ldots, A_i)$ by

$$x_i'(A_1, A_2, \ldots, A_i) = \cfrac{1}{A_1 - \cfrac{1}{A_2 - \cfrac{1}{A_3 - \cdots \atop \cdots A_{i-1} - \cfrac{1}{A_i}}}} = \frac{p_i}{q_i}, \quad i > 2,$$

and

$$x_1'(A_1) = 1/A_1.$$

The following relations can be found:

$$x_1'(A_1) = \frac{1}{A_1}, \quad p_1 = 1,\ q_1 = A_1,$$

$$x_2'(A_1, A_2) = \frac{1}{A_1 - \dfrac{1}{A_2}} = \frac{A_2}{A_1A_2 - 1},$$

$$p_2 = A_2,\ q_2 = A_1A_2 - 1,$$

$$x_3'(A_1, A_2, A_3) = \frac{1}{A_1 - \dfrac{1}{A_2 - \dfrac{1}{A_3}}} = \frac{1}{A_1 - x_2'(A_1A_2, A_2A_3)}$$

$$= \frac{1}{A_1 - \dfrac{A_3}{A_2A_3 - 1}} = \frac{A_2A_3 - 1}{A_3(A_1A_2 - 1) - A_1},$$

$$p_3 = A_2A_3 - 1,$$

$$q_3 = A_3(A_1A_2 - 1) - A_1 = A_3q_2 - q_1,$$

...

...

$$x'_N(A_1, A_2, \ldots, A_N) = \cfrac{1}{A_1 - \cfrac{1}{A_2 - \cfrac{1}{A_3 - \cdots \cfrac{}{\cdots A_{N-1} - \cfrac{1}{A_N}}}}} = \frac{p_N}{q_N}.$$

It can be shown that

$$p_N = A_N p_{N-1} - p_{N-2}, \qquad q_N = A_N q_{N-1} - q_{N-2}.$$

By writing Eq. (56) as

$$x_N = \cfrac{1}{A_1 - \cfrac{1}{A_2 - \cfrac{}{A_3 - \cdots \cfrac{}{\cdots A_{N-1} - \cfrac{1}{A'_N}}}}} = \frac{p'_N}{q'_N}$$

where $A'_N = A_N - x_N$,
we obtain

$$p'_N = (A_N - x_N)p_{N-1} - p_{N-2}, \qquad q'_N = (A_N - x_N)q_{N-1} - q_{N-2}.$$

so that

$$x_N = (A_N - x_N)p_{N-1} - p_{N-2}/(A_N - x_N)q_{N-1} - q_{N-2},$$

and the equation can be written as a quadratic form as

$$q_{N-1}x_N^2 - (p_{N-1} + q_N)x_N + p_N = 0.$$

By appropriate programming, p_N, q_N, p_{N-1}, and q_{N-1} can be calculated and the solution x_N determined.

By the same procedure, y_1 in Eq. (57) can be obtained.

[1]R. D. Mattuck, (McGraw-Hill, New York, 1967).
[2]D. A. Kirzhnits, *Field Theoretical Methods in Many-Body Systems* (Pergamon Press, New York, 1967).
[3]M. Lax, Rev. Mod. Phys. **23**, 287 (1951).
[4]P. Soven, Phys. Rev. **156**, 809 (1967).
[5]F. Yonezawa, Progr. Theoret. Phys. (Kyoto) **40**, 734 (1968).
[6]D. W. Taylor, Phys. Rev. **156**, 1017 (1967).
[7]R. N. Aiyer, R. J. Elliott, J. A. Krumhansl, and P. L. Leath, Phys. Rev. **181**, 1006 (1969).
[8]J. Hori, *Spectral Properties of Disordered Chains and Lattices* (Pergamon, Oxford, 1968).
[9]S. Y. Wu and P. L. Taylor, Phys. Rev. **181**, 1136 (1969).
[10]J. Meixner, *Handbuck der Physik* (Springer-Verlag, Berlin, 1956), p. 208.
[11]K. Kumar, *Perturbation Theory and the Nuclear Many Body Problem* (North-Holland, Amsterdam, 1962).
[12]M. Schwartz and M. Martin, Am. J. Phys. **26**, 639 (1958).
[13]E. W. Montroll and R. B. Potts, Phys. Rev. **100**, 525 (1955).
[14]B. G. Nickel and J. A. Krumhansl, Phys. Rev. B **4**, 4354 (1971).
[15]P. L. Taylor and S. Y. Wu, Phys. Rev. B **2**, 1752 (1970).
[16]P. Dean, Rev. Mod. Phys. **44**, 127 (1972).
[17]P. Dean and J. L. Martin, Proc. Roy. Soc. **259**, 409 (1960).
[18]P. Dean and M. D. Bacon, Proc. Phys. Soc. **81**, 642 (1963).
[19]A. A. Maradudin and G. H. Weiss, J. Chem. Phys. **29**, 631 (1958).
[20]P. Dean, Proc. Roy. Soc. **260**, 263 (1961).

Application of infinite order perturbation theory in linear systems. II. The frequency spectrum of disordered chains

Shi-yu Wu

Department of Physics, University of Louisville, Louisville, Kentucky 40208
(Received 23 January 1974)

A formulation is presented such that the ensemble average of the diagonal elements of the Green's function for a one-dimensional disordered system can be calculated to arbitrary accuracy. The method is illustrated with an application to an isotopically disordered chain. An approximate solution to the frequency spectrum of a disordered binary chain is also discussed.

The study of dynamical properties of disordered systems has been a field of interest since about two decades ago.[1] In recent years, there were several different approaches which had yielded good results.[2] Among them are the direct numerical calculation devised by Dean and Green's function approach first applied by Langer.[3] The numerical results have provided us with an insight into the vibrational properties of the disordered systems. For example, the frequency spectrum has a fine structure with many distinct peaks and some normal modes of vibration are strongly localized. The Green's function approach usually starts with expanding the elements of the inverse dynamical matrix, the Green's function or "Greenian." Then various schemes are designed to obtain the ensemble average of the "Greenian," from which the frequency spectrum is computed.[4–7] On the whole, the existing Green's function calculations appear to reproduce fairly accurately spectra which are known to be reasonably smooth but fail in obtaining any detailed structure which is known to exist.[8] This is especially the case for the one-dimensional disordered systems.[5,8] Thus the study of disordered chains provides a testing ground for the analytical theory.

Recently, Wu, Tung, and Schwartz[9] (referred to as I) developed an infinite order perturbation theory for linear systems. Their method, when applied to a disordered chain of finite length, can lead to the exact summation of the infinite series of the Green's function. However, their calculation still depends on the particular configuration of the disordered chain. For the purpose of understanding the properties of a very long disordered chain, in this work we present a method which will lead to an exact calculation of the ensemble average of the Green's function, where the average is taken over an ensemble of chains of all possible configurations. We shall also discuss an approximation scheme for this method.

I. THE FORMULATION

To illustrate the method, we consider an isotopically disordered chain. The equation of motion can be written as

$$-m_l\omega^2 u_l = -2\gamma u_l + \gamma u_{l+1} + \gamma u_{l-1}, \tag{1}$$

where m_l is the mass of the atom at the lth site, u_l is the displacement of the lth atom from its equilibrium position, and γ is the force constant. Note that in writing down Eq. (1) only the nearest-neighbor interaction is considered. Following Wu *et al.*,[9] the Green's function $D(\omega^2)$ is defined as the inverse of the dynamical matrix $D^{-1}(\omega^2)$ such that

$$\sum_{l''} D^{-1}_{ll''} D_{l''l'} = \delta_{ll'}, \tag{2}$$

where

$$D^{-1}_{ll'} = \left(\frac{2\gamma}{m_l} - \omega^2\right)\delta_{ll'} - \frac{\gamma}{m_l}\delta_{l+1,l'} - \frac{\gamma}{m_l}\delta_{l-1,l'}.$$

The formal solution to Eq. (2) has been obtained by Wu *et al.*[9] and can be written as

$$D_{ll} = B_l/(A_l - Z_l) \tag{3}$$

where

$$A_l = 2 - \frac{m_l}{\gamma}\omega^2,\quad B_l = \frac{m_l}{\gamma},\quad \text{and}\quad Z_l = Z_l^+ + Z_l^-$$

with

$$Z_l^\pm = 1/(A_{l\pm1} - Z^\pm_{l\pm1}). \tag{4}$$

The ensemble average of the Green's function can then be obtained as[10]

$$D = \langle D_{ll}\rangle = \sum_i c_i \left\langle \frac{B_i}{A_i - Z}\right\rangle, \tag{5}$$

where c_i is the concentration of the ith specie of the atoms in the chain, $A_i = 2 - (m_i/\gamma)\omega^2$ and $B_i = m_i/\gamma$.

Since the $Z_l^\pm$ satisfy the recurrence relation given by Eq. (4), they must be independent. One can then set up a functional equation for the distribution function $P(Z)$ such that $P(Z)\,dZ$ gives the probability that $Z_l^\pm$ will be in the interval between Z and $Z+dZ$. Following Schmidt,[11] the functional equation can be written as

$$P(Z) = \sum_i c_i \frac{1}{(A_i - Z)^2} P\left(\frac{1}{A_i - Z}\right). \tag{6}$$

The average $\langle B_i/(A_i - Z)\rangle$ is then written as

$$\left\langle \frac{B_i}{A_i - Z}\right\rangle = B_i \int \frac{P(Z^+)P(Z^-)\,dZ^+dZ^-}{A_i - Z^+ - Z^-}. \tag{7}$$

The frequency spectrum of the infinite chain with species concentration c_i's can now be computed using

$$G(\omega^2) = -(1/\pi)\,\mathrm{Im}D,$$

where D is to be calculated from Eqs. (5) and (7). It should be noted that, in deriving Eqs. (5) and (7), no approximation is involved.

The distribution function $P(Z)$ can be determined by solving a system of simultaneous equations either directly[12] or in a self-consistent manner.[13] The analysis of Gubernatis and Taylor[13] indicated that the numerical accuracy can be achieved to an absolute error as low as one part in 10^{10}. The calculation of the frequency spectrum depends entirely on $P(Z)$, and hence the accuracy of the calculation can be expected to arbitrary accuracy. However, the computation is usually tedious and time-consuming. Thus we shall next discuss an approximation scheme which, on the one hand, is easy for computation and, on the other hand, leads to the correct features of the frequency spectrum.

II. THE APPROXIMATION

In Dean's numerical works[2] on the frequency spectra of binary disordered chains, the impurity bands were identified as due to islands of one, two, or three light impurities in the sea of heavy host atoms. The structures of these bands persist even for large impurity concentrations. These features could not be obtained from the calculations of the existing perturbation methods. However, these highly localized impurity modes have been investigated through the study of the dynamics of short chains. Rosenstock and McGill[14] computed the frequency spectrum (using a convenient histogram interval) of binary chains of short lengths, say 10 or 12 atoms, and then averaged over all possible structural configurations. They found that the high frequency region of the resulting average spectrum reproduces the correct features for a very long chain. At low frequencies, the method was not fruitful. Wu *et al.*,[9] using the method of Green's function, computed the average spectrum for chains of 40 atoms with 5% light impurity. At low frequencies, they obtained good agreement with the frequency spectrum of a long chain. They also obtained the correct isolated impurity modes, but the intensities of those modes compared to the continuous part appeared to be too large. The above consideration then suggests that an analytic theory will be successful if it can account for both the properties of the average lattice and the dynamics of the local configurations. The formulation of our method [Eq. (5)] is suitable for these purposes.

To illustrate this point, let us consider the case of a binary chain with concentration c_1 and c_2. Equation (5) can then be written as

$$D = c_1 \left\langle \frac{B_1}{A_1 - Z_l} \right\rangle + c_2 \left\langle \frac{B_2}{A_2 - Z_l} \right\rangle . \tag{8}$$

The ensemble average $\langle B_1/(A_1 - Z_l)\rangle$ can be calculated step by step, using Eq. (4). Thus

$$\begin{aligned}
\left\langle \frac{A_1}{B_1 - Z_l} \right\rangle &= \left\langle B_1 \Big/ \left(A_1 - \frac{1}{A_{l+1} - Z_{l+1}^{+}} - \frac{1}{A_{l-1} - Z_{l-1}^{-}} \right) \right\rangle \\
&= c_1^2 \left\langle B_1 \Big/ \left(A_1 - \frac{1}{A_1 - Z_{l+1}^{+}} - \frac{1}{A_1 - Z_{l-1}^{-}} \right) \right\rangle \\
&\quad + c_1 c_2 \left\{ \left\langle B_1 \Big/ \left(A_1 - \frac{1}{A_1 - Z_{l+1}^{+}} - \frac{1}{A_2 - Z_{l-1}} \right) \right\rangle \right. \\
&\quad \left. + \left\langle B_1 \Big/ \left(A_1 - \frac{1}{A_2 - Z_{l+1}^{+}} - \frac{1}{A_1 - Z_{l-1}^{-}} \right) \right\rangle \right\} \\
&\quad + c_2^2 \left\langle B_1 \Big/ \left(A_1 - \frac{1}{A_2 - Z_{l+1}^{+}} - \frac{1}{A_2 - Z_{l-1}^{-}} \right) \right\rangle .
\end{aligned}$$

This sequential averaging process can be carried out continously so that in principle D can be calculated in this tedious manner. A possible approximate scheme is to truncate the sequential averaging at, say, the rth step by replacing $Z_{l\pm r}^{\pm}$ by an "effective" parameter Σ which reflects the average properties of the system. This parameter Σ can be determined self-consistently by again making use of Eq. (4). In this scheme, the ensemble-averaged Green's function D is given by

$$\begin{aligned}
D \simeq\ & c_1^{2r+1} \frac{B_1}{A_1 - 2\underbrace{Z_{1\cdots 1}}_{r}} + c_1^{2r} c_2 \left(\frac{B_2}{A_2 - 2\underbrace{Z_{1\cdots 1}}_{r}} \right. \\
& + \frac{B_1}{A_1 - \underbrace{Z_{1\cdots 1}}_{r} - \underbrace{Z_{21\cdots 1}}_{r}} \\
& + \cdots + \frac{B_1}{A_1 - \underbrace{Z_{1\cdots 1}}_{r} - \underbrace{Z_{1\cdots 12}}_{r}} + \frac{B_1}{A_1 - \underbrace{Z_{21\cdots 1}}_{r} - \underbrace{Z_{1\cdots 1}}_{r}} \\
& + \cdots + \frac{B_1}{A_1 - \underbrace{Z_{1\cdots 12}}_{r} - \underbrace{Z_{1\cdots 1}}_{r}} + \cdots + c_2^{2r+1} \frac{B_2}{A_2 - 2\underbrace{Z_{2\cdots 2}}_{r}} ,
\end{aligned} \tag{9}$$

where a typical term such as $Z_{121\cdots 12}$ is defined as

$$Z_{121\cdots 12} = \cfrac{1}{A_1 - \cfrac{1}{A_2 - \cfrac{1}{A_1 - \ddots \cfrac{1}{A_1 - \cfrac{1}{A_2 - \Sigma}}}}} . \tag{10}$$

The parameter Σ is to be determined self-consistently by

$$\begin{aligned}
\Sigma =\ & c_1^{r+1} \frac{1}{A_1 - \underbrace{Z_{1\cdots 1}}_{r}} + c_1^{r} c_2 \left(\frac{1}{A_2 - \underbrace{Z_{1\cdots 1}}_{r}} + \frac{1}{A_1 - \underbrace{Z_{21\cdots 1}}_{r}} + \cdots \right. \\
& \left. + \frac{1}{A_1 - \underbrace{Z_{1\cdots 12}}_{r}} \right) + \cdots + c_2^{r+1} \frac{1}{A_2 - \underbrace{Z_{2\cdots 2}}_{r}} .
\end{aligned} \tag{11}$$

We note that Eqs. (9)—(11) can be reduced to the proper equations for a monatomic chain if either c_1 or c_2 is zero. This can be seen as follows. For $c_1 = 1$ and $c_2 = 0$, Eq. (11) becomes

$$\Sigma = 1/(A_1 - \Sigma), \tag{11'}$$

and from Eqs. (9) and (10), we obtain

$$D = B_1 (A_1 - 2\Sigma). \tag{9'}$$

As an example of applying this method to disordered chains, we have computed the frequency spectrum for a random binary chain with 5% light impurity ($c_1 = 95\%$, $c_2 = 5\%$, and $m_2/m_1 = 0.5$). The truncation[15] was made at $r = 4$. In the calculation, a reduced unit was chosen such that $\gamma/m_1 = 1$. The calculated frequency spectrum from $\omega^2 = 4.0$ to $\omega^2 = 8.0$ reproduced the impurity bands at $\omega^2 = 4.83$, 5.18, 5.33, 5.45, 5.65, and 6.46 (Fig. 1). Figure 1 shows that our frequency spectrum exhibits more fine structure in comparison to that obtained by Dean. We believe that this arises from the fact that Dean's histogram was constructed with an interval of $\Delta\omega^2 = 0.125$. To confirm this point, we have computed the frequency spectrum using Dean's method of negative

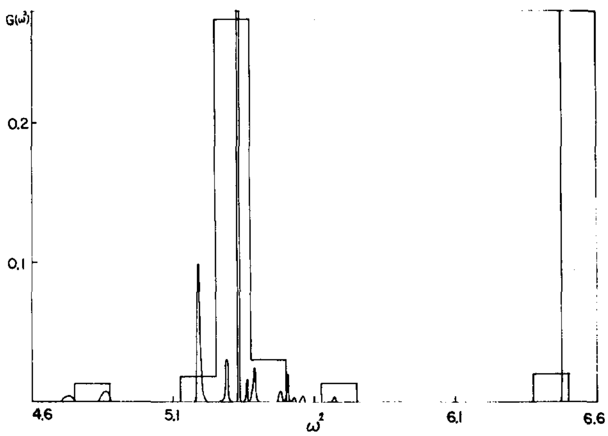


FIG. 1. The frequency spectrum (from $\omega^2=4.6$ to $\omega^2=6.6$) of a binary chain of mass ratio $m_2/m_1=0.5$ and 5% impurity (m_2) concentration. The histogram is Dean's numerical calculation.

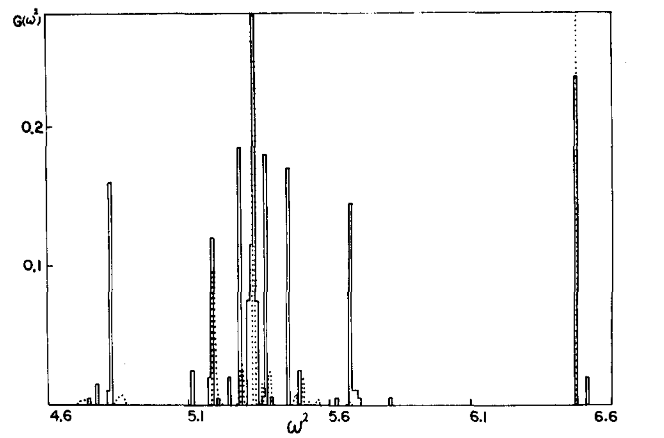


FIG. 2. The computed impurity bands (dotted line) as compared to those obtained by negative mode counting method with $\Delta\omega^2=0.01$.

mode counting at intervals of $\Delta\omega^2=0.01$. The result is shown in Fig. 2. The agreement between the positions of the impurity bands based on our analytic method and those obtained by the new numerical calculation using Dean's method is certainly heartening.

In this calculation, we are mainly concerned with the demonstration of the feasibility of the approximation scheme. Hence the truncation was made as a result of an educated guess based on the work of Rosenstock and McGill.[15] However, a justifiable truncation, say at the rth step, can be established if the results based on it and those obtained by making the truncation at the $(r+1)$th step converge. Work is now in progress along this line. It is expected that, with proper truncation established, even better agreement can be achieved between the results based on our method and those obtained by numerical calculation, especially for the magnitude of the frequency spectrum in the region of impurity bands.

Thus our method, on the one hand, leads to a formulation for an exact calculation of the ensemble average of the Green's function [Eqs. (5)–(7)] which as a rule would be time-consuming but, on the other hand, provides a readily accessible scheme for approximate calculation which gives all the correct features of the frequency spectrum. We also note that, in our approximation scheme, the impurity concentration has not been used as a perturbation parameter.

APPENDIX

Equation (5) can be derived as follows:

Since

$$D=\langle D_{ll}\rangle=\left\langle\frac{B_l}{A_l}\right\rangle+\left\langle\frac{B_l}{A_l}Z_l\frac{1}{A_l}\right\rangle+\left\langle\frac{B_l}{A_l}Z_l\frac{1}{A_l}Z_l\frac{1}{A_l}\right\rangle+\cdots$$

and

$$Z_l^{\pm}=\frac{1}{A_{l\pm1}-Z_{l\pm1}^{\pm}}$$

any term of the structure $\langle(B_l/A_l^{n+1})Z_l^n\rangle$ can be factored into

$$\left\langle\frac{B_l}{A_l^{n+1}}Z_l^n\right\rangle=\left\langle\frac{B_l}{A_l^{n+1}}\right\rangle\left\langle Z_l^n\right\rangle=\sum_i c_i\frac{B_i}{A_i^{n+1}}\left\langle Z_l^n\right\rangle .$$

Therefore D can be expressed as

$$D=D_{ll}=\sum_i c_i\left\langle\frac{B_i}{A_i-Z_l}\right\rangle .$$

[1] E.W. Montroll and R.B. Potts, Phys. Rev. 100, 525 (1955).
[2] P. Dean, Rev. Mod. Phys. 44, 127 (1972).
[3] P. Dean, Proc. Roy. Soc. 254, 507 (1960); J.S. Langer, J. Math. Phys. 2, 584 (1961).
[4] F. Yonezawa and T. Matsubara, Prog. Theoret. Phys. (Kyoto) 35, 759 (1966).
[5] D.W. Taylor, Phys. Rev. 156, 1017 (1967).
[6] R.N. Aiyer, R.J. Elliott, J.A. Krumhansl, and P.L. Leath, Phys. Rev. 181, 1006 (1969).
[7] B.G. Nickel and J.A. Krumhansl, Phys. Rev. B 4, 4354 (1971).
[8] J. Hori, *Spectral Properties of Disordered Chains and Lattices* (Pergamon, Oxford, 1968).
[9] S.Y. Wu, C.C. Tung, and M. Schwartz, "Application of infinite order perturbation theory in linear systems. I," J. Math. Phys. 15, 938 (1974).
[10] See Appendix.
[11] H. Schmidt, Phys. Rev. 105, 425 (1957).
[12] R.L. Agacy and R.E. Borland, Proc. Phys. Soc. 84, 1017 (1964).
[13] J.E. Gubernatis and P.L. Taylor, J. Phys. C: Solid State 4, L94 (1971).
[14] H.B. Rosenstock and R.E. McGill, Phys. Rev. 176, 1004 (1968).
[15] The analysis of Rosenstock and McGill[14] shows that the high frequency modes can be investigated through the study of chains of short length, say 10 or 12 atoms. We therefore chose the truncation at $r=4$ because it corresponds to chains of 9 atoms $(2r+1=2\times4+1=9)$.

Scattering of plane longitudinal elastic wave by a large convex rigid object with a statistically corrugated surface. II. Far field solution

W. Fan and Y. M. Chen

Department of Applied Mathematics and Statistics, State University of New York at Stony Brook, New York 11790

(Received 1 March 1972)

The far field solution for the scattering of plane longitudinal elastic wave by a large convex rigid object with a statistically corrugated surface is obtained and the effective reflection and diffraction coefficients are deduced.

INTRODUCTION

In Paper I[1] the geometrical theory of diffraction,[2,3,4] has been extended to include the scattering of time harmonic plane longitudinal elastic wave by a large convex rigid object with statistical surface irregularities. There the asymptotic expansion of the first-order perturbation solution is expressed in terms of surface integrals of the equivalent source function $\mathbf{S}_1$ andhence it is not explicit enough for the purpose of physical interpretation. In this paper, the aforementioned integrals are first evaluated asymptotically for far field and then the effective reflection and diffraction coefficients are deduced. Finally, the expression of the mean value of $\mathbf{U}_{s1}$ for far field is given.

I. WAVES IN FAR FIELD

A. Geometric optics wave in far field

Since the zeroth-order diffracted wave is asymptotically small in comparison with the geometric optics wave, for simplicity, we ignore the zeroth-order diffracted wave in calculating the asymptotic expansion of the equivalent source function $\mathbf{S}_1$.

From (I. 5. 30), we have that for $\theta = \theta'$

$$\begin{aligned}\left(\frac{\partial \mathbf{U}_i}{\partial r} + \frac{\partial \mathbf{U}^G_{s0}}{\partial r}\right)_{r=a} &\approx ik_1 \cos\theta'(-\cos\theta' \mathbf{1}_{r'} + \sin\theta' \mathbf{1}_{\theta'}) \exp(ik_1 a \cos\theta') \\ &- ik_1 \cos\theta' \frac{\cos(\theta' + \lambda')}{\cos(\theta' - \lambda')} (\cos\theta' \mathbf{1}_{r'} + \sin\theta' \mathbf{1}_{\theta'}) \\ &\qquad \times \exp(ik_1 a \cos\theta') \\ &- ik_2 \cos\lambda' \frac{2\cos\theta' \sin\theta'}{\cos(\theta' - \lambda')} (\sin\lambda' \mathbf{1}_{r'} - \cos\lambda' \mathbf{1}_{\theta'}) \\ &\qquad \times \exp(ik_1 a \cos\theta') \quad \text{for } \theta' < \tfrac{1}{2}\pi \\ &\approx O \quad \text{for } \theta' > \tfrac{1}{2}\pi, \end{aligned} \tag{1.1}$$

where

$$\lambda' = \sin^{-1}[(1/N) \sin\theta']. \tag{1.2}$$

Hence from (I. 4. 18), we obtain the asymptotic expansion of $\mathbf{S}_1$ as

$$\begin{aligned} S_{1r'} &\approx ik_1 a \frac{2\cos\theta' \cos\lambda'}{\cos(\theta' - \lambda')} f(\theta', \varphi', q) \exp(ik_1 a \cos\theta') \quad \text{for } \theta' < \tfrac{1}{2}\pi \\ &\approx O \quad \text{for } \theta' > \tfrac{1}{2}\pi \end{aligned} \tag{1.3}$$

and

$$\begin{aligned} S_{1\theta'} &\approx -ik_2 a \frac{\sin 2\theta'}{\cos(\theta' - \lambda')} f(\theta', \varphi', q) \exp(ik_1 a \cos\theta') \quad \text{for } \theta' < \tfrac{1}{2}\pi \\ &\approx O \quad \text{for } \theta' > \tfrac{1}{2}\pi. \end{aligned} \tag{1.4}$$

We will then evaluate the integral (I. 6. 66) for far field, $r \gg a$. Upon utilizing the saddle point method for double integrals,[5] we find that the saddle point equations are

$$\begin{cases} \sin(\varphi - \varphi') = 0, & (1.5) \\ \sin\theta' + \sin\gamma \dfrac{\partial \gamma}{\partial \theta'} = 0, & (1.6) \end{cases}$$

for the first integral of (I. 6. 66) and

$$\begin{cases} \sin(\varphi - \varphi') = 0, & (1.7) \\ \sin\theta' + N \sin\gamma \dfrac{\partial \gamma}{\partial \theta'} = 0, & (1.8) \end{cases}$$

for the second and third integrals of (I. 6. 66).

It can be shown that equations (1.5) and (1.6) have the real solution

$$\theta' = \theta/2, \quad \varphi' = \varphi, \quad \text{for } \theta' < \pi/2 \tag{1.9}$$

and equations (1.7) and (1.8) have the real solution

$$\theta' = \tan^{-1}[N \sin\theta/(1 + N\cos\theta)], \quad \varphi' = \varphi, \quad \text{for } \theta' < \pi/2. \tag{1.10}$$

Finally, the saddle-point contribution of the first integral of (I. 6. 66) gives the first-order reflected P wave,

$$\begin{aligned} \mathbf{U}^G_{s1p}(\mathbf{r}, q) &\approx i\, \frac{k_1 a^2}{r} f\left(\frac{\theta}{2}, \varphi, q\right) \cos\frac{\theta}{2} \left(N^2 \cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2}\right) \\ &\times \left[\cos\frac{\theta}{2} \left(N^2 - \sin^2\frac{\theta}{2}\right)^{1/2} + \sin\frac{\theta}{2}\right]^{-2} \\ &\times \exp\left\{-ik_1\left(r - 2a\cos\frac{\theta}{2}\right)\right\} \mathbf{1}_r \end{aligned} \tag{1.11}$$

and the saddle-point contribution of the second and third integrals of (I. 6. 66) gives the first-order reflected S wave,

$$\begin{aligned} \mathbf{U}^G_{s1s}(\mathbf{r}, q) &\approx -i\, \frac{2k_1 a^2}{r \sin\theta} f(\hat{\theta}, \varphi, q) \sin\hat{\theta} \cos\hat{\theta} \cos(\theta - \hat{\theta}) \\ &\times \cos^{-3}(2\hat{\theta} - \theta) [(1 - N^2) \sin\hat{\theta} \cos(\theta - \hat{\theta}) \cos 2\hat{\theta} \\ &- N \sin\theta \cos\hat{\theta} \cos(\theta - \hat{\theta}) - N^2 \sin\theta \sin^2\hat{\theta}] \\ &\times \exp\{-ik_2[r - a\cos(\theta - \hat{\theta}) - aN^{-1} \cos\hat{\theta}]\} \mathbf{1}_\theta, \end{aligned} \tag{1.12}$$

where

$$\hat{\theta} = \tan^{-1}[N \sin\theta/(1 + N\cos\theta)]. \tag{1.13}$$

From (1.11) we observe that the main contribution of the far field of the first-order reflected P wave comes from the equivalent source in the neighborhood of Q_p the point of reflection of the zeroth-order P wave. Similarly, it is true for the first-order reflected S wave (Fig. 1).

Upon combining (I. 5. 28) and (1.11), we obtain the far field expression of the scattered geometric optics P wave,

$$\begin{aligned} \mathbf{U}^G_{sp} &\approx \mathbf{1}_r (a/2r) R_{p\text{-}eff} \exp\{-ik_1(r - 2a\cos\tfrac{1}{2}\theta)\} + O(\epsilon^2) \quad \text{in the lit region,} \\ &\approx O(\epsilon^2) \quad \text{in the shadow region,} \end{aligned} \tag{1.14}$$

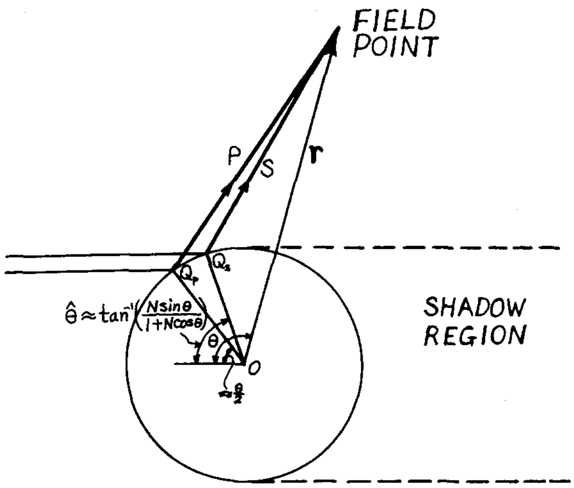


FIG. 1. The first order reflected geometric rays are shown.

where the effective reflection coefficient of P wave

$$R_{p\text{-}eff} = \frac{(N^2 - \sin^2\frac{1}{2}\theta)^{1/2}\cos\frac{1}{2}\theta - \sin^2\frac{1}{2}\theta}{(N^2 - \sin^2\frac{1}{2}\theta)^{1/2}\cos\frac{1}{2}\theta + \sin^2\frac{1}{2}\theta} + \epsilon i 2k_1 a f(\tfrac{1}{2}\theta, \varphi, q) \times \frac{(N^2\cos^2\frac{1}{2}\theta - \sin^2\frac{1}{2}\theta)\cos\frac{1}{2}\theta}{[(N^2 - \sin^2\frac{1}{2}\theta)^{1/2}\cos\frac{1}{2}\theta + \sin\frac{1}{2}\theta]^2}. \tag{1.15}$$

Similarly, from (I. 5.28) and (1.12), we have the far field expression of the scattered geometric optics S wave,

$$\begin{aligned}\mathbf{U}^G_{ss} &\approx \mathbf{1}_\theta \frac{a}{r}\left\{\frac{N^2\cos(\theta-\hat{\theta})\sin\hat{\theta}}{[1+\cos\hat{\theta}/N\cos(\theta-\hat{\theta})]\sin\theta}\right\}^{1/2} \\ &\quad\times R_{s\text{-}eff}\exp\{-ik_2[r - a\cos(\theta-\hat{\theta}) - aN^{-1}\cos\hat{\theta}]\} + O(\epsilon^2) \\ &\qquad \text{for } \theta < (\pi/2) + \sin^{-1}(1/N) \text{ or } \hat{\theta} < (\pi/2) \\ &\approx O(\epsilon^2) \qquad \text{for } \theta > (\pi/2) + \sin^{-1}(1/N) \text{ or } \hat{\theta} > (\pi/2),\end{aligned} \tag{1.16}$$

where the effective reflection coefficient of S wave,

$$\begin{aligned}R_{s\text{-}eff} &= \frac{-2\cos\hat{\theta}\sin\hat{\theta}}{N\cos(\theta-\hat{\theta})\cos\hat{\theta} + \sin^2\hat{\theta}} - \epsilon i 2k_1 a f(\hat{\theta},\varphi,q) \\ &\quad\times \cos\hat{\theta}\left(\frac{N\sin\hat{\theta}}{\sin\theta}\right)^{1/2}[\cos\hat{\theta} + N\cos(\theta-\hat{\theta})]^{1/2} \\ &\quad\times \cos^{-3}(2\hat{\theta}-\theta)[(1-N^2)\sin\hat{\theta}\cos 2\hat{\theta} N\cos(\theta-\hat{\theta}) \\ &\quad - N^2\sin\theta\cos\hat{\theta}\cos(\theta-\hat{\theta}) - N^3\sin\theta\sin^2\hat{\theta}].\end{aligned} \tag{1.17}$$

From (1.14) and (1.16) we observe that in the far field region the reflected waves, both P and S, are essentially spherical waves. Furthermore, from (1.15) and (1.17) we find that up to and including terms of $O(\epsilon)$ the reflected waves depend on $k_1 a\epsilon f$. Hence, in this case we believe that our boundary-perturbation technique is reasonably good for $k_1 a\epsilon|f| < 1$ ($\epsilon|f|$ small in comparison with wavelength). This restriction is stronger than the restriction, $\epsilon|f| < 1$, posed in Sec. 3 of Paper I. However, there is no restriction here on the derivatives of f as those are posed in Sec. 3 of Paper I.

B. Diffracted wave in far field

Upon following the same procedure as that in evaluating the surface integrals of the geometric optics wave, we find that, contrary to the geometric optics wave, $\mathbf{U}_i$ and $\mathbf{U}^G_{s0}$ components of $\mathbf{S}_1$ give no contribution to the far field first-order diffracted wave. From (I. 4.18) and (I. 5.43), the asymptotic expansions of $S_{1r'}$ and $S_{1\theta'}$ due to $\mathbf{U}^D_{s0}$ component are

$$\begin{aligned}S_{1r'} &\approx ik_1 a f(\theta',\varphi',q)\mathfrak{D}^2_{0s}(1/a\sin\theta')^{1/2} \\ &\quad\times(1-1/N^2)^{1/4}(1+e^{-i2\pi\nu_\lambda})^{-1}\{\exp[-i\nu_\lambda(\tfrac{3}{2}\pi+\theta')] \\ &\quad+\exp[-i\nu_\lambda(\tfrac{3}{2}\pi-\theta')-i\tfrac{1}{2}\pi]\} \quad \text{for } \theta' < \tfrac{1}{2}\pi, \\ &\approx ik_1 a f(\theta',\varphi,q)\mathfrak{D}^2_{0s}(1/a\sin\theta')^{1/2}(1-N^2)^{1/4} \\ &\quad\times(1+e^{-i2\pi\nu_\lambda})^{-1}\{\exp[-i\nu_\lambda(\theta'-\tfrac{1}{2}\pi)+i\pi] \\ &\quad+\exp[-i\nu_\lambda(\tfrac{3}{2}\pi-\theta')+i\tfrac{3}{2}\pi]\} \quad \text{for } \theta' > \tfrac{1}{2}\pi,\end{aligned} \tag{1.18}$$

and

$$\begin{aligned}S_{1\theta'} &\approx ik_2 a f(\theta',\varphi,q)\mathfrak{D}^2_{0s}(1/a\sin\theta')^{1/2} \\ &\quad\times(1-1/N^2)^{3/4}(1+e^{-i2\pi\nu_\lambda})^{-1}\{\exp[-i\nu_\lambda(\tfrac{3}{2}\pi+\theta')] \\ &\quad-\exp[-i\nu_\lambda(\tfrac{3}{2}\pi-\theta')-i\tfrac{1}{2}\pi]\} \quad \text{for } \theta' < \tfrac{1}{2}\pi, \\ &\approx ik_2 a f(\theta',\varphi',q)\mathfrak{D}^2_{0s}(1/a\sin\theta')^{1/2} \\ &\quad\times(1-1/N^2)^{3/4}(1+e^{-i2\pi\nu_\lambda})^{-1}\{\exp[-i\nu_\lambda(\theta'-\tfrac{1}{2}\pi)+i\pi] \\ &\quad-\exp[-i\nu_\lambda(\tfrac{3}{2}\pi-\theta')+i\tfrac{3}{2}\pi]\} \quad \text{for } \theta' > \tfrac{1}{2}\pi.\end{aligned} \tag{1.19}$$

In evaluating the double integral (I. 6.89) by the saddle-point method, it is found that the saddle point is independent of θ', hence we can evaluate the integrals with respect to φ' only.

The saddle point equation is

$$\sin(\varphi'-\varphi) = 0, \tag{1.20}$$

which has two solutions

$$\varphi' = \varphi \tag{1.21}$$

and

$$\varphi' = \hat{\varphi} = \begin{cases}\varphi+\pi & \text{for } \varphi < \pi \\ \varphi-\pi & \text{for } \varphi > \pi\end{cases}. \tag{1.22}$$

Finally, the saddle point evaluation leads to the asymptotic expressions for the far field first-order diffracted wave,

$$\mathbf{U}^D_{s1}(\mathbf{r},q) = \mathbf{U}^D_{s1p}(\mathbf{r},q) + \mathbf{U}^D_{s1s}(\mathbf{r},q), \tag{1.23}$$

where the pressure wave

$$\begin{aligned}\mathbf{U}^D_{s1p}(\mathbf{r},q) &\approx \sum_{\nu_\lambda} C_{s1p}\Big[\Big(\exp[-i\nu_\lambda(2\pi+\theta)+i\tfrac{1}{2}\pi]\int_0^{\pi/2} f(\theta',\varphi,q)d\theta' \\ &\quad+\exp[-i\nu_\lambda\theta - i\tfrac{1}{2}\pi]\int_{\pi/2}^{\pi} f(\theta',\varphi,q)d\theta' \\ &\quad+\exp[-i\nu_\lambda(2\pi+\theta)]\int_0^{(\pi/2)-\theta} f(\theta',\hat{\varphi},q)d\theta' \\ &\quad+\exp[-i\nu_\lambda\theta - i\pi]\int_{(\pi/2)-\theta}^{\pi} f(\theta',\hat{\varphi},q)d\theta'\Big) \\ &\quad+\Big(\exp[-i\nu_\lambda(2\pi-\theta)]\int_0^{(\pi/2)+\theta} f(\theta',\varphi,q)d\theta' \\ &\quad+\exp[i\nu_\lambda\theta - i\pi]\int_{(\pi/2)+\theta}^{\pi} f(\theta',\varphi,q)d\theta' \\ &\quad+\exp[-i\nu_\lambda(2\pi-\theta)+i\tfrac{1}{2}\pi]\int_0^{\pi/2} f(\theta',\hat{\varphi},q)d\theta' \\ &\quad+\exp[i\nu_\lambda\theta - i\tfrac{1}{2}\pi]\int_{\pi/2}^{\pi} f(\theta',\hat{\varphi},q)d\theta'\Big)\Big]\mathbf{1}_r \\ &\qquad \text{for } \theta < \tfrac{1}{2}\pi,\end{aligned} \tag{1.24}$$

$$\approx \sum_{\nu_\lambda} C_{s1p}\Big[\Big(\exp[-i\nu_\lambda\theta - i\tfrac{1}{2}\pi]\int_0^{\theta-(\pi/2)} f(\theta',\varphi,q)d\theta'$$
$$+\exp[-i\nu_\lambda(2\pi+\theta)+i\tfrac{1}{2}\pi]\int_{\theta-(\pi/2)}^{\pi/2} f(\theta',\varphi,q)d\theta'$$
$$+\exp[-i\nu_\lambda\theta - i\tfrac{1}{2}\pi]\int_{\pi/2}^{\pi} f(\theta',\varphi,q)d\theta'$$
$$+\exp[-i\nu_\lambda\theta - i\pi]\int_0^{\pi} f(\theta',\hat\varphi,q)d\theta'\Big)$$
$$+\Big(\exp[-i\nu_\lambda(2\pi-\theta)]\int_0^{\pi} f(\theta',\varphi,q)d\theta'$$
$$+\exp[-i\nu_\lambda(2\pi-\theta)+i\tfrac{1}{2}\pi]\int_0^{\pi/2} f(\theta',\hat\varphi,q)d\theta'$$
$$+\exp[i\nu_\lambda\theta - i\tfrac{1}{2}\pi]\int_{\pi/2}^{(3/2)\pi-\theta} f(\theta',\hat\varphi,q)d\theta'$$
$$+\exp[-i\nu_\lambda(2\pi-\theta)+i\tfrac{1}{2}\pi]\int_{(3/2)\pi-\theta}^{\pi} f(\theta',\hat\varphi,q)d\theta'\Big)\Big]\mathbf{1}_r$$
$$\text{for } \theta > \tfrac{1}{2}\pi$$

with

$$C_{s1p} = (a/r)2\pi(N^2-1)^2(2\pi/a\sin\theta)^{1/2}(k_1a/6)^{2/3} \times[A(t_\lambda)]^{-2}(1+e^{-i2\pi\nu_\lambda})^{-2}\exp[-ik_1r - i\pi\nu_\lambda - i\tfrac{5}{12}\pi] \tag{1.25}$$

and the shear wave

$$\mathbf{U}^D_{s1s}(\mathbf{r},q)\approx\sum_{\nu_\lambda} C_{s1s}\Big[\Big(\exp[-i\nu_\lambda(2\pi+\theta)+i\tfrac{1}{2}\pi]\int_0^{\pi/2} f(\theta',\varphi,q)d\theta'$$
$$+\exp[-i\nu_\lambda\theta - i\tfrac{1}{2}\pi]\int_{\pi/2}^{\pi} f(\theta',\varphi,q)d\theta'$$
$$+\exp[-i\nu_\lambda(2\pi+\theta)+i\pi]\int_0^{(\pi/2)-\theta} f(\theta',\hat\varphi,q)d\theta'$$
$$+\exp[-i\nu_\lambda\theta]\int_{(\pi/2)-\theta}^{\pi} f(\theta',\hat\varphi,q)d\theta'\Big)$$
$$+\Big(\exp[-i\nu_\lambda(2\pi-\theta)+i\pi]\int_0^{(\pi/2)+\theta} f(\theta',\varphi,q)d\theta'$$
$$+\exp[i\nu_\lambda\theta]\int_{(\pi/2)+\theta}^{\pi} f(\theta',\varphi,q)d\theta'$$
$$+\exp[-i\nu_\lambda(2\pi-\theta)+i\tfrac{1}{2}\pi]\int_0^{\pi/2} f(\theta',\hat\varphi,q)d\theta'$$
$$+\exp[i\nu_\lambda\theta - i\tfrac{1}{2}\pi]\int_{\pi/2}^{\pi} f(\theta',\hat\varphi,q)d\theta'\Big)\Big]\mathbf{1}_\theta$$
$$\text{for } \theta < \tfrac{1}{2}\pi, \tag{1.26}$$

$$\approx \sum_{\nu_\lambda} C_{s1s}\Big[\Big(\exp[-i\nu_\lambda\theta - i\tfrac{1}{2}\pi]\int^{\theta-(\pi/2)} f(\theta',\varphi,q)d\theta'$$
$$+\exp[-i\nu_\lambda(2\pi+\theta)+i\tfrac{1}{2}\pi]\int_{\theta-(\pi/2)}^{\pi} f(\theta',\varphi,q)d\theta'$$
$$+\exp[-i\nu_\lambda\theta - i\tfrac{1}{2}\pi]\int_{\pi/2}^{\pi} f(\theta',\varphi,q)d\theta'$$
$$+\exp[-i\nu_\lambda\theta]\int_0^{\pi} f(\theta',\hat\varphi,q)d\theta'\Big)$$
$$+\Big(\exp[-i\nu_\lambda(2\pi-\theta)+i\pi]\int_0^{\pi} f(\theta',\varphi,q)d\theta'$$
$$+\exp[-i\nu_\lambda(2\pi-\theta)+i\tfrac{1}{2}\pi]\int_0^{\pi/2} f(\theta',\hat\varphi,q)d\theta'$$
$$+\exp[i\nu_\lambda\theta - i\tfrac{1}{2}\pi]\int_{\pi/2}^{(3/2)\pi-\theta} f(\theta',\hat\varphi,q)d\theta'$$
$$+\exp[-i\nu_\lambda(2\pi-\theta)+i\tfrac{1}{2}\pi]\int_{(3/2)\pi-\theta}^{\pi} f(\theta',\hat\varphi,q)d\theta'\Big)\Big]\mathbf{1}_\theta$$
$$\text{for } \theta > \tfrac{1}{2}\pi$$

with

$$C_{s1s} = \frac{a}{r}\,2\pi N^4\Big(1-\frac{1}{N^2}\Big)^{7/4}\Big(1+\frac{1}{\sqrt{N^2-1}}\Big)\Big(\frac{\pi k_1 a}{a\sin\theta}\Big)^{1/2} \times\Big(\frac{k_1a}{6}\Big)^{1/3}[A(t_\lambda)]^{-1}(1+e^{-i2\pi\nu_\lambda})^{-2} \times\exp\Big[-ik_2r + ik_2a\Big(1-\frac{1}{N^2}\Big)^{1/2} - i\pi\nu_\lambda - i\nu_\lambda\cos^{-1}\frac{1}{N} - i\frac{\pi}{3}\Big]. \tag{1.27}$$

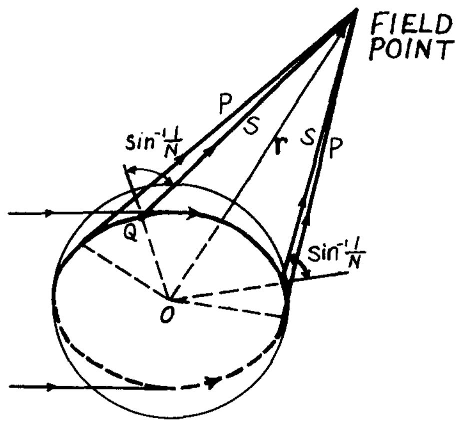


FIG. 2. The first order diffracted rays are shown.

Upon examining (1.24) and (1.26), it is found that the contribution of $\mathbf{S}_1$ to $\mathbf{U}^D_{s1}$ comes from every point on the circle formed by the intersection of the plane, defined by $\mathbf{r}, Q$ and 0, and the sphere of radius a (Fig. 2). Hence the total contribution is obtained by integrating the source distribution along the entire circumference of the circle. The first four terms of (1.24) and (1.26) represent the total contribution of $\mathbf{S}_1$ on the circule to the first-order diffracted rays traveling clockwise on the surface of the sphere and the last four terms of (1.24) and (1.26) represent the first-order diffracted rays traveling counterclockwise. Note that there are only phase factors outside of integrals which means that the phase factors being independent of θ' in various ranges of integration. This is because every point inside the particular range of integration has exactly the same phase distance from the point on the circle, where the incident wave hits the sphere tangentially, to the point, where the first-order diffracted wave leaves the surface tangentially towards the field point.

Upon combining (I. 5.43) and (1.24), we obtain the far field expression of the diffracted P-wave,

$$\mathbf{U}^D_{sp} \approx \mathbf{1}_r\frac{1}{r}\Big(\frac{a}{\sin\theta}\Big)^{1/2}e^{-ik_1r}\sum_{\nu_\lambda}(1+e^{-i2\pi\nu_\lambda})^{-1}$$
$$\times\Big[\Big\{\mathfrak{D}^2_{0p}+\epsilon\Lambda_1\Big[\Big(ie^{-i2\pi\nu_\lambda}\int_0^{\pi/2} f(\theta',\varphi,q)d\theta'$$
$$-i\int_{\pi/2}^{\pi} f(\theta',\varphi,q)d\theta' + e^{-i2\pi\nu_\lambda}\int_0^{(\pi/2)-\theta} f(\theta',\hat\varphi,q)d\theta'$$
$$-\int_{(\pi/2)-\theta}^{\pi} f(\theta',\hat\varphi,q)d\theta'\Big)H\Big(\frac{\pi}{2}-\theta\Big)$$
$$+\Big(-i\int_0^{\theta-(\pi/2)} f(\theta',\varphi,q)d\theta' + ie^{-i2\pi\nu_\lambda}$$
$$\times\int_{\theta-(\pi/2)}^{\pi/2} f(\theta',\varphi,q)d\theta' - i\int_{\pi/2}^{\pi} f(\theta',\varphi,q)d\theta'$$
$$-\int_0^{\pi} f(\theta',\hat\varphi,q)d\theta'\Big)H\Big(\theta-\frac{\pi}{2}\Big)\Big]\Big\}e^{-i\nu_\lambda(\pi+\theta)}$$
$$+\Big\{\mathfrak{D}^2_{0p}+\epsilon\Lambda_1\Big[\Big(e^{-i2\pi\nu_\lambda}\int_0^{(\pi/2)+\theta} f(\theta',\varphi,q)d\theta'$$
$$-\int_{(\pi/2)+\theta}^{\pi} f(\theta',\varphi,q)d\theta' + ie^{-i2\pi\nu_\lambda}\int_0^{\pi/2} f(\theta',\hat\varphi,q)d\theta'$$
$$-i\int_{\pi/2}^{\pi} f(\theta',\hat\varphi,q)d\theta'\Big)H\Big(\frac{\pi}{2}-\theta\Big)$$
$$+\Big(e^{-i2\pi\nu_\lambda}\int_0^{\pi} f(\theta',\varphi,q)d\theta' + ie^{-i2\pi\nu_\lambda}\int_0^{\pi/2} f(\theta',\hat\varphi,q)d\theta'$$
$$-i\int_{\pi/2}^{(3/2)\pi-\theta} f(\theta',\hat\varphi,q)d\theta'$$
$$+ie^{-i2\pi\nu_\lambda}\int_{(3/2)\pi-\theta}^{\pi} f(\theta',\hat\varphi,q)d\theta'\Big)H$$
$$\times\Big(\theta-\frac{\pi}{2}\Big)\Big]\Big\}e^{-i\nu_\lambda(\pi-\theta)}\Big] + O(\epsilon^2)$$
$$\text{in the lit region,} \tag{1.28}$$

$$\approx \mathbf{1}_r \frac{1}{r}\left(\frac{a}{\sin\theta}\right)^{1/2} e^{-ik_1 r}\sum_{\nu_\lambda}(1+e^{-i2\pi\nu_\lambda})^{-1}$$
$$\times\Big\{\Big[\mathfrak{D}_{0p}^2+\epsilon\Lambda_1\Big(-i\int_0^{\theta-(\pi/2)} f(\theta',\varphi,q)d\theta'$$
$$+ie^{-i2\pi\nu_\lambda}\int_{\theta-(\pi/2)}^{\pi/2} f(\theta',\varphi,q)d\theta' - i\int_{\pi/2}^{\pi} f(\theta',\varphi,q)d\theta'$$
$$-\int_0^{\pi} f(\theta',\hat\varphi,q)d\theta'\Big)\Big] e^{-i\nu_\lambda(\pi+\theta)}$$
$$+\Big[\mathfrak{D}_{0p}^2+\epsilon\Lambda_1\Big(e^{-i2\pi\nu_\lambda}\int_0^{\pi} f(\theta',\varphi,q)d\theta'$$
$$+ie^{-i2\pi\nu_\lambda}\int_0^{\pi/2} f(\theta',\hat\varphi,q)d\theta' - i\int_{\pi/2}^{(3/2)\pi-\theta} f(\theta',\hat\varphi,q)d\theta'$$
$$+ie^{-i2\pi\nu_\lambda}\int_{(3/2)\pi-\theta}^{\pi} f(\theta',\hat\varphi,q)d\theta'\Big)\Big] e^{-i\nu_\lambda(\pi-\theta)}\Big\} + O(\epsilon^2)$$

in the shadow region,

where

$$\Lambda_1 = (2\pi)^{3/2}(N^2-1)^2\left(\frac{k_1 a}{6}\right)^{2/3} \times [A(t_\lambda)]^{-2}(1+e^{-i2\pi\nu_\lambda})^{-1}e^{-i5/12\pi}. \quad (1.29)$$

Similarly from (I.5.43) and (1.26), we obtain the far field expression of the diffracted S wave,

$$\mathbf{U}_{ss}^D \approx \mathbf{1}_\theta \frac{1}{r}\left(\frac{a}{\sin\theta}\right)^{1/2}\exp\{-ik_2 r + ik_2 a[1-(1/N^2)]^{1/2}\}$$
$$\times\sum_{\nu_\lambda}(1+e^{-i2\pi\nu_\lambda})^{-1}\Big[\Big\{\mathfrak{D}_{0s}^2+\epsilon\Lambda_2$$
$$\times\Big[\Big(ie^{-i2\pi\nu_\lambda}\int_0^{\pi/2} f(\theta',\varphi,q)d\theta' - i\int_{\pi/2}^{\pi} f(\theta',\varphi,q)d\theta'$$
$$-e^{-i2\pi\nu_\lambda}\int_0^{(\pi/2)-\theta} f(\theta',\hat\varphi,q)d\theta'$$
$$+\int_{(\pi/2)-\theta}^{\pi} f(\theta',\hat\varphi,q)d\theta'\Big)H\Big(\frac{\pi}{2}-\theta\Big)$$
$$+\Big(-i\int_0^{\theta-(\pi/2)} f(\theta',\varphi,q)d\theta'$$
$$+ie^{-i2\pi\nu_\lambda}\int_{\theta-(\pi/2)}^{\pi/2} f(\theta',\varphi,q)d\theta'$$
$$-i\int_{\pi/2}^{\pi} f(\theta',\varphi,q)d\theta' + \int_0^{\pi} f(\theta',\hat\varphi,q)d\theta'\Big)H\Big(\theta-\frac{\pi}{2}\Big)\Big]$$
$$\times\exp\{-i\nu_\lambda[\pi+\theta+\cos^{-1}(1/N)]\}$$
$$+\Big\{\mathfrak{D}_{0s}^2+\epsilon\Lambda_2\Big[\Big(-e^{-i2\pi\nu_\lambda}\int_0^{(\pi/2)+\theta} f(\theta',\varphi,q)d\theta'$$
$$+\int_{(\pi/2)+\theta}^{\pi} f(\theta',\varphi,q)d\theta' + ie^{-i2\pi\nu_\lambda}\int_0^{\pi/2} f(\theta',\hat\varphi,q)d\theta'$$
$$-i\int_{\pi/2}^{\pi} f(\theta',\hat\varphi,q)d\theta'\Big)H\Big(\frac{\pi}{2}-\theta\Big)$$
$$+\Big(-e^{-i2\pi\nu_\lambda}\int_0^{\pi} f(\theta',\varphi,q)d\theta' + ie^{-i2\pi\nu_\lambda}$$
$$\times\int_0^{\pi/2} f(\theta',\hat\varphi,q)d\theta' - i\int_{\pi/2}^{(3/2)\pi-\theta} f(\theta',\hat\varphi,q)d\theta'$$
$$+ie^{-i2\pi\nu_\lambda}\int_{(3/2)\pi-\theta}^{\pi} f(\theta',\hat\varphi,q)d\theta'\Big)H\Big(\theta-\frac{\pi}{2}\Big)\Big]\Big\}$$
$$\times\exp\{i\nu_\lambda[\pi-\theta+\cos^{-1}(1/N)]\}\Big] + O(\epsilon^2)$$

in the lit region, (1.30)

$$\approx \mathbf{1}_\theta \frac{1}{r}\left(\frac{a}{\sin\theta}\right)^{1/2}\exp\{(-ik_2 r + ik_2 a)[1-(1/N^2)]^{1/2}\}$$
$$\times\sum_{\nu_\lambda}(1+e^{-i2\pi\nu_\lambda})^{-1}$$
$$\times\Big\{\Big[\mathfrak{D}_{0s}^2+\epsilon\Lambda_2\Big(-i\int_0^{\theta-(\pi/2)} f(\theta',\varphi,q)d\theta' + ie^{-i2\pi\nu_\lambda}$$
$$\times\int_{\theta-(\pi/2)}^{\pi/2} f(\theta',\varphi,q)d\theta' - i\int_{\pi/2}^{\pi} f(\theta',\varphi,q)d\theta'$$
$$+\int_0^{\pi} f(\theta',\hat\varphi,q)d\theta'\Big)\Big]\exp\{-i\nu_\lambda[\pi+\theta+\cos^{-1}(1/N)]\}$$
$$+\Big[\mathfrak{D}_{0s}^2+\epsilon\Lambda_2\Big(-e^{-i2\pi\nu_\lambda}\int_0^{\pi} f(\theta',\varphi,q)d\theta'$$
$$+ie^{-i2\pi\nu_\lambda}\int_0^{\pi/2} f(\theta',\hat\varphi,q)d\theta'$$
$$-i\int_{(\pi/2)}^{(3/2)\pi-\theta} f(\theta',\hat\varphi,q)d\theta' + ie^{-i2\pi\nu_\lambda}$$
$$\times\int_{(3/2)\pi-\theta}^{\pi} f(\theta',\hat\varphi,q)d\theta'\Big)\Big]$$
$$\times\exp[-i\nu_\lambda[\pi-\theta+\cos^{-1}(1/N)]\Big\} + O(\epsilon^2)$$

in the shadow region,

where

$$\Lambda_2 = 2\pi N^4(1-1/N^2)^{7/4}[1+(N^2-1)^{-1/2}] \times(\pi k_1 a)^{1/2}(k_1 a/6)^{1/3}[A(t_\lambda)]^{-1}(1+e^{-i2\pi\nu_\lambda})^{-1}e^{-i\pi/3}. \quad (1.31)$$

From (1.28) and (1.30) we observe that in the far field region the diffracted waves are essentially spherical waves and their expressions break down at $\theta=0$ and π. Again it seems that our boundary-perturbation technique is reasonably good for $(k_1 a)^{5/6}\epsilon|f|<1$. Due to the appearance of the integrals of f, the effective diffraction coefficients are complicated and different for $\mathbf{r}$ in different regions.

Since (1.14), (1.16), (1.28) and (1.30) contain f linearly, the mean values of $\mathbf{U}_s^G$ and $\mathbf{U}_s^D$ can be simply obtained by replacing f by $\langle f\rangle$. The generalization to large convex rigid objects can be made by following the recipe in Sec. 7 of Paper I.

[1] W. Fan and Y. M. Chen, J. Math. Phys. 14, AIP p. no for Paper I (1973).
[2] J. B. Keller, "A. Geometrical Theory of Diffraction", Symp. Microwave Optics, Eaton Electronics Res. Lab, McGill Univ., Montreal, Canada, 1953.
[3] J. B. Keller, J. Opt. Soc. Am. 52, 116 (1962).
[4] F. Gilbert, J. Acoust. Soc. Am. 32, 841 (1960).
[5] D. S. Jones and M. Kline, Research Report No. EM-100, Division of Electromagnetic Research, Institute of Mathmatical Sciences, New York University (1956).

Thermodynamics of a mixture of fermions and bosons in one dimension with a repulsive δ-function potential

C. K. Lai

Department of Physics, University of Utah, Salt Lake City, Utah 84112
(Received 19 October 1973)

The thermodynamics of a mixture of fermions and bosons is derived on the basis of two ansätze about the roots of a set of algebraic equations.

I. INTRODUCTION

In a previous paper[1] (to be called I), the ground state energy of a mixture of fermions and bosons in one dimension with repulsive δ-function potential was obtained. In this paper, we would like to show that the thermodynamics could alternatively be derived on the basis of two ansätze about the roots of the set of algebraic equations.

The Hamiltonian of the system to be considered is

$$H = -\sum_{1}^{N} \frac{\delta^2}{\delta x_i^2} + 2c \sum_{i<j} \delta(x_i - x_j), \qquad c > 0 \tag{1}$$

for M_1 fermions of species 1, M_2 fermions of species 2, and M_b bosons. The energy levels of the system are determined by the algebraic equations (I. 11a, I. 11b, I. 11g)

$$e^{ipL} = \prod_{\Lambda'}\left(\frac{ip - i\Lambda' - c'}{ip - i\Lambda' + c'}\right), \qquad \text{number of } p = N; \tag{2}$$

$$\prod_{p'}\left(\frac{i\Lambda - ip' + c'}{i\Lambda - ip' - c'}\right) = -\prod_{\Lambda'}\left(\frac{i\Lambda - i\Lambda' + c}{i\Lambda - i\Lambda' - c}\right)\prod_{A'}\left(\frac{i\Lambda - iA' - c'}{i\Lambda - iA' + c}\right),$$
$$\text{number of } \Lambda = M; \tag{3}$$

$$\prod_{\Lambda'}\left(\frac{iA - i\Lambda' + c'}{iA - i\Lambda' - c'}\right) = 1, \qquad \text{number of } A = M_b; \tag{4}$$

where $N = M_1 + M_2 + M_b$, $M = M_2 + M_b$, and $c' = \frac{1}{2}c$. Concerning the solutions of (2)–(4), we propose the following ansätze:

Ansatz 1: When L is very large, the Λ's in the complex plane are located in strings, which are fermion-like: That is, a string $C(\xi, m)$ is of the form

$$C(\xi, m) \qquad \Lambda = \xi_m + \tfrac{1}{2}\, i\mu c^2 + O(e^{-kL}). \tag{5}$$

Ansatz 2: The A's in Eqs. (3) and (4) are real numbers.

Ansatz 1 is precisely the one we used to obtain the thermodynamics of the fermion problem[2] (we will refer to this work as II). As for ansatz 2, it is not obvious from the structure of Eqs. (3) and (4) that the A's should be real. But if one considers that these variables might serve as a kind of pseudomomenta of the bosons, then ansatz 2 is a plausible one.

In the following section, we will obtain the integral equations on the basis of the above ansätze. In Sec. III, the integral equations are solved exactly in special cases [$c \to 0$ and $c \to \infty$]. In Sec. IV, the second virial coefficient in the fugacity expansion is computed. Both the special cases and the second virial coefficient give correct results, and thus help to confirm the ansätze proposed.

II. THERMODYNAMICS

We substitute Eq. (5) into Eqs. (2)–(4) to obtain

$$e^{ipL} = \prod_{C(\xi',n)}\left(\frac{-p + \xi' - in\eta}{-p + \xi' + in\eta}\right), \tag{6}$$

$$\prod_{p'}\left(\frac{-p' + \xi - im\eta}{-p' + \xi + im\eta}\right) = (-1)^m \prod_{A'}\left(\frac{\xi - A' + im\eta}{\xi - A' - im\eta}\right) \times \prod_{nl} \exp\left[i a_{mnl}\, \theta\left(\frac{\xi - \xi'}{n - l}\right)\right], \tag{7}$$

$$\prod_{C(\xi',n)}\left(\frac{A - \xi' - in\eta}{A - \xi' + in\eta}\right) = 1, \tag{8}$$

where

$$a_{mnl} = \begin{cases} 1 & \text{for } l = \pm m, l \neq n, \\ 2 & \text{for } l = -(m-2), -(m-4), \cdots, (m-2), \\ 0 & \text{otherwise;} \end{cases} \tag{9}$$

and

$$\theta(x) = 2\tan^{-1}(x/\eta), \qquad \eta = c/2.$$

Taking the logarithm, one has

$$pL = 2\pi I_p + \sum_{C(\xi',n)} \theta\left(\frac{\xi' - p}{n}\right), \tag{10}$$

$$\sum_{p'} \theta\left(\frac{\xi - p'}{m}\right) = 2\pi J_\xi + \sum_{C(\xi',m)} \sum_{l} \times a_{mnl}\, \theta\left(\frac{\xi - \xi'}{n - l}\right) - \sum_{A'} \theta\left(\frac{\xi - A'}{m}\right), \tag{11}$$

$$\sum_{C(\xi',n)} \theta\left(\frac{A - \xi'}{n}\right) = 2\pi K_A, \tag{12}$$

where I_p, J_ξ, K_A are integers or half-integers coming from the multiples of 2π in logarithm. As $L \to \infty$, the above become integral equations:

$$1 = 2\pi(\rho + \rho_h) - \sum_n \int_{-\infty}^{\infty} \theta'\left(\frac{p - \xi'}{n}\right) \sigma_n(\xi')\, d\xi', \tag{13}$$

$$\int_{-\infty}^{\infty} \theta'\left(\frac{\xi - p}{m}\right) \rho(p)\, dp = 2\pi(\sigma_m + \sigma_{m,h}) + \sum_n a_{mnl} \times \int_{-\infty}^{\infty} \theta'\left(\frac{\xi - \xi'}{n - l}\right)\sigma_n(\xi')\, d\xi' - \int_{-\infty}^{\infty} \theta'\left(\frac{\xi - A}{m}\right) \tau(A)\, dA, \tag{14}$$

$$2\pi(\tau + \tau_h) = \sum_n \int_{-\infty}^{\infty} \theta'\left(\frac{A - \xi}{n}\right) \sigma_n(\xi)\, d\xi, \tag{15}$$

where σ_n, σ_{nh}, etc. are the "particle" density and the "hole" density of the strings $C(\xi, m)$, etc. Taking the Fourier transform of Eqs. (13)–(15), one obtains

$$\tilde{\rho}(\omega)e^{-\eta m|\omega|} = \tilde{\sigma}_{m,h} + \sum_{\nu} \coth|\eta\omega| \, (e^{-\eta|(m-\nu)\omega|} - e^{-\eta|(m+\nu)\omega|})\,\tilde{\sigma}_{\nu} - e^{-\eta m|\omega|}\tilde{\tau}(\omega). \quad (16)$$

$$\sigma(\omega)/2\pi = \tilde{\rho} + \tilde{\rho}_h - \sum_{\nu} e^{-\nu\eta|\omega|}\tilde{\sigma}_{\nu}, \quad (17)$$

$$\tilde{\tau} + \tilde{\tau}_h = \sum_{\nu} e^{-\nu\eta|\omega|}\tilde{\sigma}_{\nu}. \quad (18)$$

Equation (16) can be converted into

$$(\tilde{\sigma}_m + \tilde{\sigma}_{mh})\, 2\cosh\eta\omega = \tilde{\sigma}_{m+1,h} + \tilde{\sigma}_{m-1,h}, \quad (19)$$

where $\tilde{\sigma}_{0h} \equiv \tilde{\rho} + \tilde{\tau}$. Then (17) becomes

$$\delta(\omega)/2\pi = \tilde{\rho} + \tilde{\rho}_h - \tfrac{1}{2}\cosh\eta\omega\,(e^{-\eta|\omega|}\tilde{\sigma}_{0h} - \tilde{\sigma}_{1h}). \quad (20)$$

Now it is straightforward to obtain the thermodynamics following the method of Yang and Yang.[3] By the use of Eqs. (18)–(20) one writes

$$\rho_h/\rho = \exp[\epsilon(p)/T], \qquad \sigma_{nh}/\sigma_h = \exp[\phi_n(p)/T],$$
$$\tau_h/\tau = \exp[\chi(p)T];$$

and minimizes the free energy $F = E - TS$ with the constraints:

$$\frac{N}{L} = \text{const}, \quad (M_1 - M_2)/L = \text{const}, \quad M_b/L = \text{const}.$$

One then obtains integral equations for the $\epsilon(p)$ etc., as in (II. 23):

$$\alpha = p^2 - \epsilon - \tfrac{1}{2}T\int_{-\infty}^{\infty} G_1 \ln(1 + e^{-\epsilon/T})\,dk - \tfrac{1}{2}T\int_{-\infty}^{\infty} G_1 \ln(1 + e^{-\psi/T})\,dk - \tfrac{1}{2}T\int_{-\infty}^{\infty} G_0 \ln\left[1 + \exp\left(\frac{\phi_1}{T}\right)\right] dk, \quad (21a)$$

$$C = -\psi + \alpha - p^2 + \epsilon, \quad (21b)$$

$$\phi_1 = \tfrac{1}{2}T\int_{-\infty}^{\infty} G_0\left\{\ln\left[1 + \exp\left(\frac{\phi_2}{T}\right)\right] - \ln(1 + e^{-\epsilon/T}) - \ln(1 + e^{-\psi/T})\right\} dk, \quad (21c)$$

$$\phi_n = \tfrac{1}{2}T\int_{-\infty}^{\infty} G_0\left\{\ln\left[1 + \exp\left(\frac{\phi_{n+1}}{T}\right)\right] + \ln\left[1 + \exp\left(\frac{\phi_{n-1}}{T}\right)\right]\right\} dk, \quad n \geq 2, \quad (21d)$$

with the asymptotic condition

$$\lim_{n\to\infty} T^{-1}\phi_n/n = \lambda > 0. \quad (22)$$

The G's are kernels defined as in (II. 11). Once the ϵ and ϕ's are obtained, it can easily be shown that from Eqs. (13)–(15), ρ, etc. are given by the following as in (II. 24):

$$\rho = -(2\pi)^{-1}(1 + e^{\epsilon/T})^{-1}\frac{\partial\epsilon}{\partial\alpha}, \quad \tau = -(2\pi)^{-1}(1 + e^{\psi/T})^{-1}\frac{\partial\psi}{\partial\alpha},$$
$$\sigma_{n,h} = -(2\pi)^{-1}[1 + \exp(\phi_n/T)]^{-1}\frac{\partial\phi_n}{\partial\alpha}; \quad (23)$$

and

$$\frac{E}{L} = \int_{-\infty}^{\infty} p^2\rho(p)dp, \quad \frac{N}{L} = \frac{(M_1 + M_2 + M_b)}{L} = \int_{-\infty}^{\infty}\rho dp,$$
$$\frac{M_b}{L} = \int_{-\infty}^{\infty}\tau dp, \quad \frac{(M_1 - M_2)}{L} = \lim_{n\to\infty}\int_{-\infty}^{\infty}\sigma_{n,h}dp.$$

Finally, the free energy is given by

$$\frac{F}{L} = \alpha\frac{N}{L} + C\frac{M_b}{L} - \frac{T}{2\pi}\int_{-\infty}^{\infty}\ln(1 + e^{-\epsilon/T})dk - B\left(\frac{M_1 - M_2}{L}\right). \quad (24)$$

III. SPECIAL CASES

Equation (21) can be solved exactly in the cases $c \to 0$ and $c \to \infty$.

A. $c \to 0$

The solution in this case should give the results of a mixture of free bosons and free fermions. As $c \to 0$, the kernels become delta functions and Eq. (21d) yields

$$\exp(2\phi_\nu) = [1 + \exp(\phi_{\nu+1})][1 + \exp(\phi_{\nu-1})]. \quad (25)$$

Equation (25) has solutions

$$1 + \exp(\phi_\nu) = \sinh^2(n\lambda + \mu)/\sinh^2\lambda, \quad (26)$$

with $\lambda T \equiv B$, and μ is to be determined from Eqs. (21a)–(21c). This leads to

$$[1 + \exp(-\epsilon)][1 + \exp(-\psi)] = \sinh^2\lambda/\sinh^2\mu,$$
$$\exp(-\epsilon + p^2 - \alpha) = \exp(-\psi - C) = \sinh(\lambda + \mu)/\sinh. \quad (27)$$

After some algebraic manipulation, one readily obtains

$$\coth\mu = \frac{\cosh\lambda + \exp C}{\sinh\lambda} + \frac{2(\cosh\lambda + \cosh C)}{\exp(p^2 - \alpha - C) - 1} \quad (28)$$

and

$$2\pi\tau = [\exp(p^2 - \alpha - C) - 1]^{-1}, \quad (29)$$

$$2\pi\rho = 2\pi\tau + [\exp(p^2 - \alpha - B) + 1]^{-1} + [\exp(p^2 - \alpha + B) + 1]^{-1}, \quad (30)$$

$$\frac{N - 2M}{L} = \frac{1}{2\pi}\{[\exp(p^2 - \alpha - B) + 1]^{-1} - [\exp(p^2 - \alpha - B) + 1]^{-1}\}. \quad (31)$$

Equation (29) is just the distribution function for free bosons and Eqs. (30)–(31) are the distribution function for a mixture of free bosons and fermions in a magnetic field B.

B. $c \to \infty$

In this case, one would expect that all particles behave like identical free fermions. This is because as the interaction strength $c \to \infty$, the exchange force due to the symmetry of the wavefunction becomes unimportant. As $c \to \infty$, the integrals $\int G_1 \ln(1 + e^{-\epsilon/T})$ do not contribute, allowing the ψ and the ϕ's to be constants:

$$1 + \exp\phi_n = \sinh^2(n\lambda + \mu)/\sinh^2\lambda,$$
$$1 + \exp(-\psi) = \sinh^2\lambda/\sinh^2\mu. \quad (32)$$

μ is to be determined from Eqs. (21a)–(21b):

$$\exp C = \sinh(\lambda - \mu)/\sinh\mu. \quad (33)$$

Finally Eq. (21a) gives

$$p^2 - \epsilon - T\ln(\exp C + 2\cosh\lambda) = \alpha, \quad (34)$$

$$2\pi\rho = (1 + e^{\epsilon/T})^{-1}. \quad (35)$$

Equation (35) with Eq. (34) gives simply the distribution function for free fermions with each energy level being occupied by only one particle.

IV. THE FUGACITY EXPANSION

One can obtain the fugacity expansion as follows: Let

$$\exp(-\epsilon/T) = \sum_{n=1}^{\infty} a_n(k, T)_z^n,$$
$$\exp(-\psi/T) = \sum \omega_n(k, T)_z^n, \tag{36}$$
$$\exp(\phi_\nu/T) = b_\nu(k, T) + c_\nu(k, T)_z + d_\nu(k, T)_z{}^2 + \cdots .$$

Substituting the above into Eq. (21), to zeroth order in z, one obtains

$$p^2 + \ln a_1 = \tfrac{1}{2}[G_1 \ln\omega_1 + G_0 \ln b_1], \tag{37a}$$

$$C = \ln(\omega_1 - 1) - \tfrac{1}{2}[G_1 \ln\omega_1 + G_0 \ln b_1], \tag{37b}$$

$$\ln(b_\nu - 1) = \tfrac{1}{2}[G_0 \ln b_{\nu+1} + G_0 \ln b_{\nu-1}]; \quad b_0 \equiv 1/\omega_1, \quad \nu \geq 1. \tag{37c}$$

To the first order of z, one obtains

$$\frac{a_2}{a_1} = \frac{1}{2}\left[G_1 a_1 + G_1\left(\frac{\omega_2}{\omega_1}\right) + G_0\left(\frac{c_1}{b_1}\right)\right], \tag{38a}$$

$$\frac{\omega_2}{(\omega_1 - 1)} = \frac{a_2}{a_1}, \tag{38b}$$

$$\frac{c_\nu}{(b_\nu - 1)} = \frac{1}{2} G_0\left[\frac{c_{\nu+1}}{b_{\nu+1}} + \frac{c_{\nu-1}}{b_{\nu-1}}\right] \quad \frac{c_0}{b_0} \equiv -a_1 - \frac{\omega_2}{\omega_1}, \quad \nu \geq 1. \tag{38c}$$

Equation (38c) has solutions

$$b_n = f_n{}^2 = \sinh^2(n\lambda + \mu)/\sinh^2\lambda, \tag{39}$$

where μ can be expressed in terms of C from Eqs. (38b) and (38c). Equation (38c) has solutions

$$c_\nu = 1/2\pi \int \tilde{c}_\nu(\omega) \exp(i\omega k)\, d\omega,$$
$$\tilde{c}_\nu(\omega) = A(\omega)\{f_\nu f_{\nu-1} \exp[-|(\nu + 2)\eta\omega|] - f_\nu f_{\nu+1} \exp(-|\nu\omega\eta|)\}, \tag{40}$$

where $A(\omega)$ can be expressed in terms of the a's and ω's from the initial condition of c_0/b_0 in Eq. (38c). After some algebraic manipulation, one can readily obtain

$$\exp C = \sinh(\lambda - \mu)/\sinh\mu,$$
$$a_1 = (\exp C + 2\cosh\lambda)\exp(-p^2), \tag{41}$$
$$a_2 = [(\exp C + \cosh\lambda)^2 - \sinh^2\lambda]\exp(-p^2/T) \int K_2 \cdot \exp(-p^2/T)\, dp.$$

For simplicity, let $B = 0$ ($\lambda = 0$), then the pressure is given by

$$p = \frac{T}{2\pi} \int dk[a_1 z + (a_2 - a_1{}^2/2)z^2 + \ldots]$$
$$= \frac{(\pi T)^{1/2}}{2\pi}\left\{(2 + \exp C)z + \left[\frac{(1 + \exp C)^2}{\sqrt{2}} \int K_1 \exp\left(\frac{-2p^2}{T}\right) dp - \frac{(2 + \exp C)^2}{2^{2/3}}\right] z^2 + \cdots\right\}.$$

This agrees with the result obtained by standard methods. For higher orders in z, the procedure is more tedious and will not be presented here.

V. CONCLUSION

We have obtained the thermodynamics of a mixture of fermions and bosons on the basis of two ansätze. Ansatz 1 has been used successfully in the pure fermion problem. Ansatz 2 is new. These ansätze are likely to be the correct ones as demonstrated by the correct solutions given in the special cases of the integral equations. The $\epsilon(p)$, $\phi_n(p)$'s, and $\psi(p)$ in the equations can be interpreted as excitation energies in the excitation spectrum at finite temperature. The excitation spectrum has been computed in the pure fermion problem,[4] and thus will not be repeated here.

[1]C. K. Lai and C. N. Yang, Phys. Rev. A 3, 393 (1971).
[2]C. K. Lai, Phys. Rev. Letters 26, 1472 (1971).
[3]C. N. Yang and C. P. Yang, J. Math. Phys. 10, 1115 (1969).
[4]C. K. Lai (to be published).

Elastic scattering in the Kerr metric

S. K. Bose and M. Y. Wang

Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

The differential cross section for scattering from a source whose gravitational field is described by the Kerr metric is evaluated.

I. INTRODUCTION

Recently the problem of elastic scattering in general relativity has been treated by Collins, Delbourgo, and Williams.[1] These authors have studied the differential scattering cross section from a Schwarzschild source and obtained the relativistic correction term to the classical Rutherford formula. In this paper the result of Ref. 1 is generalized to take account of the rotation of the source. The relevant geometry now is that described by the Kerr metric.[2] In view of the fact that black holes are believed to be described by the Kerr metric,[3] our exercise may not be entirely devoid of physical interest.

Several features of the present problem should be mentioned. Two of these, viz., (1) the existence of a critical angular momentum below which capture occurs and (2) multispiral scattering, which contributes to a given final scattering angle and divides the impact parameters into various zones, are general relativistic effects which are also present in the corresponding Schwarzschild problem, as has been already noted in Ref. 1. Additionally, it is known that in the Kerr metric the only planar motion possible is that in the equitorial plane. Throughout this paper we will confine our attention to this plane.

II. GEODESICS AND SCATTERING IN THE KERR METRIC

In the Boyer–Lindquist[4] coordinate the Kerr metric is given by

$$ds^2 = dr^2 + 2a \sin^2\theta dr d\phi + (r^2 + a^2)\sin^2\theta d\phi^2 + (r^2 + a^2\cos^2\theta)d\theta^2 - dt^2 + \frac{2mr}{r^2 + a^2\cos^2\theta}(dr + a\sin^2\theta d\phi + dt)^2, \tag{1}$$

where m is the mass of the scattering center and a its angular momentum per unit mass. If the motion of the test particle is confined to the equitorial plane ($\theta = \pi/2$), then the timelike geodesic equation, according to Boyer and Lindquist, is

$$\frac{d\phi}{du} = \frac{a}{D(u)} \pm \frac{A(u)}{D(u)[B(u)]^{1/2}}, \tag{2}$$

where $u = 1/r$ and $A(u)$, $B(u)$, and $D(u)$ are given by

$$A(u) = \frac{1}{\alpha} - \frac{2m}{\alpha}(1 + \gamma a\alpha)u, \tag{3a}$$

$$D(u) = 1 - 2mu + a^2u^2, \tag{3b}$$

$$B(u) = \gamma^2 - 1 + 2mu + \frac{1}{\alpha^2}(a^2\alpha^2(\gamma^2 - 1) - 1)u^2 + \frac{2m}{\alpha^2}(1 + \gamma a\alpha)^2u^3. \tag{3c}$$

The constants of motion α and γ occurring above are related to the mass μ of the scattered particle, its energy E and its angular momentum l at asymptotic radial distances as

$$\alpha = \mu/l, \quad \gamma = E/\mu. \tag{4}$$

The total deflection $\Delta\phi$ is obtained from Eq. (2) by integrating the right-hand side over a contour from $u = 0$ to the branch point $u = 1/d$ at pericenter and back to $u = 0$ again:

$$\Delta\phi = 2\int_0^{1/d} \frac{A(u)}{D(u)} \frac{du}{[B(u)]^{1/2}}. \tag{5}$$

To bring the above to a standard form write

$$D(u) = a^2(u_+ - u)(u_- - u), \quad u_\pm = \frac{m \pm (m^2 - a^2)^{1/2}}{a^2} \tag{6}$$

and split the rational portion of the integrand into partial fraction:

$$\frac{A(u)}{D(u)} = \frac{A_+}{u_+ - u} + \frac{A_-}{u_- - u}, \quad A_\pm = \frac{\pm 1}{2\alpha(m^2 - a^2)^{1/2}}[-1 + 2m(1 + \gamma a\alpha)u_\pm]. \tag{7}$$

Now write the cubic $B(u)$ as

$$B(u) = \frac{2m}{\alpha^2}(1 + \gamma a\alpha)^2(u - u_1)(u - u_2)(u - u_3), \tag{8}$$

where by definition $u_2 = 1/d$ and introduces further the notation

$$k^2 = \frac{u_2 - u_3}{u_1 - u_3}, \quad \alpha_\pm = \frac{u_2 - u_3}{u_\pm - u_3}, \quad \sin^2\Psi_0 = \frac{-u_3}{u_2 - u_3}, \quad y^2 = \frac{u - u_3}{u_2 - u_3}. \tag{9}$$

Equation (5) now reduces to

$$\Delta\phi = -\frac{4\alpha}{(1 + \gamma a\alpha)(2m)^{1/2}} \frac{1}{(u_1 - u_3)^{1/2}} \left(\frac{A_+}{u_3 - u_+}\int_{\sin\Psi_0}^1 \frac{dy}{(1 - \alpha_+ y^2)[(1 - y^2)(1 - k^2y^2)]^{1/2}} + \frac{A_-}{u_3 - u_-}\int_{\sin\Psi_0}^1 \frac{dy}{(1 - \alpha_- y^2)[(1 - y^2)(1 - k^2y^2)]^{1/2}}\right). \tag{10}$$

The integrals occurring above can be expressed as

$$I_\pm = \int_{\sin\Psi_0}^1 \frac{dy}{(1 - \alpha_\pm y^2)[(1 - y^2)(1 - k^2y^2)]^{1/2}} = \Pi(-\alpha_\pm, k) - \Pi(\Psi_0, -\alpha_\pm, k) \tag{11}$$

when the two terms on the right-hand side of the above are, respectively, the complete and incomplete elliptic functions of the third type. With appropriate (and minor) modifications the above is the result given by Boyer and Lindquist[4] in their treatment of the deflection of light in the Kerr metric.

The expression for scattering cross section is to be calculated from Eq. (10). First, the total angle of deviation χ of the test particle is

$$\chi = \Delta\phi - \pi \tag{12}$$

and the measured scattering angle θ is

$$\theta = |\chi - 2n\pi| \tag{13}$$

when n is the number of loops around the scattering center that the test particle traverses before escaping to infinity. The impact parameter b is $l|p = \alpha^{-1}(\gamma^2 - 1)^{-1/2}$ and as explained in Ref. 1, the differential scattering cross section is given by the formula

$$\frac{d\sigma}{d\Omega} = \frac{1}{\gamma^2 - 1} \sum_n \frac{1}{\alpha_n^3} \left| \frac{d\alpha_n}{d\cos\theta} \right|, \tag{14}$$

where α_n connotes the range of α values which result in an n-spiral scattering having $\theta = \pm(\chi - 2n\pi)$.

If we consider a situation in which the test particle is initially at an infinite distance from the scattering center then two possibilities arise: (i) scattering of the particle, occurring when, $u_3 \le 0 < u_2 < u_1$; (ii) capture of the particle taking place when $u_3 \le 0 < u_2 = u_1$. The critical value of α for which particle capture takes place is denoted by α_c. In the next section we shall derive asymptotic expressions for the cross section in the limit of large and critical angular momentum.

III. ASYMPTOTIC RESULTS

We shall obtain asymptotic expansions about two limits, $\alpha = 0$ and $\alpha = \alpha_c$. First consider the limit around $\alpha = 0$. We expand the roots of the cubic $B(u)$ and obtain:

$$u_1 = \frac{1}{2m} - \frac{\gamma a}{m}\alpha + \left(\frac{2\gamma^2 + 1}{2m} a^2 - 2m\gamma^2\right)\alpha^2 + O(\alpha^3), \tag{15a}$$

$$u_2 = (\gamma^2 - 1)^{1/2}\alpha + \gamma^2 m\alpha^2 + (\gamma^2 - 1)^{1/2}[2\gamma am(\gamma^2 - 1)^{1/2} + X]\alpha^3, \tag{15b}$$

$$u_3 = -(\gamma^2 - 1)^{1/2}\alpha + \gamma^2 m\alpha^2 + (\gamma^2 - 1)^{1/2}[2\gamma am(\gamma^2 - 1)^{1/2} - X]\alpha^3, \tag{15c}$$

with

$$X = 16\gamma^2 m^2 - 4m^2 - 8\gamma^2 a^2 - \frac{Y}{216(\gamma^2 - 1)m^2} \tag{15d}$$

and

$$Y = 324a^2m^2 - 1836a^2m^2\gamma^4 + 3a^4\gamma^4 - 9a^4\gamma^2 + 6a^4 + 2916m^4\gamma^4 + 864m^4 + 1518\gamma^2a^2m^2 - 3888\gamma^2m^4. \tag{15e}$$

The reason for keeping α^3 terms in u_2 and u_3 is that we need these terms to calculate (see below) $\sin^2\Psi_0$ to within α^2 order. From Eqs. (15a)-(15c) and (9) we get:

$$k^2 = 4m\alpha(\gamma^2 - 1)^{1/2}\{1 - 2[m(\gamma^2 - 1)^{1/2} - \gamma a]\alpha\}, \tag{16}$$

$$\alpha_\pm = 2(\gamma^2 - 1)^{1/2}\alpha\left(\frac{1}{u_\pm} - \frac{(\gamma^2 - 1)^{1/2}}{u_\pm^2}\alpha\right), \tag{17}$$

$$\sin^2\Psi_0 = \tfrac{1}{2}\left(1 - \frac{\gamma^2 m}{(\gamma^2 - 1)^{1/2}}\alpha - 2\gamma am(\gamma^2 - 1)^{1/2}\alpha^2\right), \tag{18}$$

$$\Psi_0 = \frac{\pi}{4} - \frac{\gamma^2 m}{2(\gamma^2 - 1)^{1/2}}\alpha - \gamma am(\gamma^2 - 1)^{1/2}\alpha^2. \tag{18$'$}$$

Using the above expressions (16)-(18′), we can carry out an expansion of the integrand in (11) and then evaluate the integral after performing a series of elementary operations. Alternatively, we can calculate the first three terms of the appropriate expansion of the elliptic functions of the third kind[5]

$$\Pi(\Psi_0, -\alpha_\pm, k) = \sum_{\mu=0}^{\infty} (\alpha_\pm)^\mu J_{2_\mu}(\Psi_0) \sum_{\nu=0}^{\mu} \binom{-1/2}{\nu}\left(\frac{k^2}{-\alpha_\pm}\right)^\nu,$$

$$\Pi(-\alpha_\pm, k) = \Pi\left(\Psi_0 = \frac{\pi}{2}, -\alpha_\pm, k\right) \tag{19}$$

with

$$J_{2_\mu}(\Psi_0) = \frac{2\mu - 1}{2\mu} J_{2_{\mu-2}}(\Psi_0) - \frac{1}{2\mu}\sin^{2\mu-1}\Psi_0\cos\Psi_0,$$
$$J_0(\Psi_0) = \Psi_0. \tag{20}$$

Both the above mentioned procedures, of course, yield the same result; which is

$$I_\pm = \frac{\pi}{4} + \left(\frac{\gamma^2 m}{2(\gamma^2 - 1)^{1/2}} + \frac{(\gamma^2 - 1)^{1/2}}{2}(1 + \pi/2)\frac{1}{u_\pm} + \frac{(\gamma^2 - 1)^{1/2}}{2}(1 + \pi/2)m\right)\alpha + \left(2\gamma am(\gamma^2 - 1)^{1/2}(1 + \pi/4) + \frac{\gamma^2 m}{2u_\pm} + \frac{\gamma^2 m^2}{2} + \frac{\gamma^2 - 1}{2u_\pm^2}(1 + \pi/4) + \frac{m^2}{2}(\gamma^2 - 1)(1 + \pi/8) + \frac{m(\gamma^2 - 1)}{u_\pm}(1 + 3\pi/8)\right)\alpha^2. \tag{21}$$

We also note the following expansion derived from Eqs. (6), (7), and (15c):

$$\frac{A_\pm}{u_3 - u_\pm} = \pm\frac{l}{2u_\pm(m^2 - a^2)^{1/2}}\left[1 - 2mu_\pm + \left(2m(\gamma^2 - 1)^{1/2} - \frac{(\gamma^2 - 1)^{1/2}}{u_\pm} - 2m\gamma au_\pm\right)\alpha + \left(2m\gamma a(\gamma^2 - 1)^{1/2} + \frac{\gamma^2 - 1}{u_\pm^2} + \frac{(2 - \gamma^2)m}{u_\pm} - 2m^2\gamma^2\right)\alpha^2\right]. \tag{22}$$

Inserting Eqs. (21) and (22) into Eq. (10) we obtain, after carrying out appropriate expansions of the remaining factors,

$$\Delta\phi = \pi + 2\frac{(2\gamma^2 - 1)m}{(\gamma^2 - 1)^{1/2}}\alpha + \left(3\pi\frac{5\gamma^2 - 1}{4}m^2 + 4m\gamma a(\gamma^2 - 1)^{1/2}\right)\alpha^2. \tag{23}$$

Thus in this case there is very little deviation of the trajectory from a straight line and there is no spiraling effect. The connection between the scattering angle θ and inverse angular momentum α is given by (23):

$$\theta = 2\frac{(2\gamma^2-1)m}{(\gamma^2-1)^{1/2}}\alpha + \left(3\pi\frac{5\gamma^2-1}{4}m^2 + 4m\gamma a(\gamma^2-1)^{1/2}\right)\alpha^2. \tag{24}$$

Substituting Eq. (24) into Eq. (14) we obtain the *small-angle* differential cross section,

$$\frac{d\sigma}{d\Omega} = \left(\frac{2m(2\gamma^2-1)}{\theta^2(\gamma^2-1)}\right)^2 + \frac{3\pi m^2(5\gamma^2-1)}{4\theta^3(\gamma^2-1)} + \frac{4m\gamma a}{\theta^3(\gamma^2-1)^{1/2}} + \cdots. \tag{25}$$

The first term in Eq. (25) corresponds to the Rutherford formula and the second term the general relativistic correction to it, as has been already noted in Ref. 1. The last term in Eq. (25) displays the general relativistic effect due to the rotation of the source.

Now consider the other limit $\alpha = \alpha_c$. Evaluation of α_c will be taken up separately in the next section. Here we will express the cross section in terms of α_c. When $\alpha = \alpha_c$, the corresponding critical values of the roots u_i of the cubic $B(u)$ are the following:

$$u_{1c} = u_{2c} = \frac{1}{6m(1+\gamma a\alpha_c)^2}\{1 - \Gamma a^2\alpha^2{}_c + \left[(1-\Gamma a^2\alpha^2{}_c)^2 - 12m^2\alpha^2{}_c(1+\gamma a\alpha_c)^2\right]^{1/2}\}, \tag{26}$$

$$u_{3c} = \frac{1}{6m(1+\gamma a\alpha_c)^2}\{1 - \Gamma a^2\alpha^2{}_c - 2[(1-\Gamma a^2\alpha^2{}_c)^2 - 12m^2\alpha^2{}_c(1+\gamma a\alpha_c)^2]^{1/2}\},$$

$$\Gamma = \gamma^2 - 1 \tag{26'}$$

and $k^2 = 1$. Thus suitable expansion parameter is $1 - k^2$. We first note that the integrals $I_\pm$ now contain logarithmically growing parts which diverge at $k^2 = 1$. So this (leading) behavior has to be first isolated. The technique for doing this for the elliptic functions of the third type, or the final result, does not seem to be given in the standard mathematical literature, so that it may be worthwhile for us to report some details. Consider first the integral

$$\Pi(-\alpha_\pm, k) = \int_0^1 \frac{dy}{(1-\alpha_\pm y^2)[(1-y^2)(1-k^2y^2)]^{1/2}}$$

which is seen, by inspection, to diverge logarithmically at $k^2 = 1$. A change of variables

$$t = \frac{1}{(1-k^2y^2)^{1/2}}, \quad k' = (1-k^2)^{1/2} \tag{27}$$

casts the integral into the form:

$$\Pi(-\alpha_\pm, k) = \int_1^{1/k'} \frac{k^2t^2}{[\alpha_\pm + (k^2-\alpha_\pm)t^2][(t^2-1)(1-k'^2t^2)]^{1/2}}$$
$$= \sum_{n=0}^{\infty} k^2k'^{2n}\frac{(2n-1)!!}{2n!!}\int_1^{1/k'}\frac{t^{2n+2}dt}{[\alpha_\pm + (k^2-\alpha_\pm)t^2](t^2-1)^{1/2}}. \tag{28}$$

Let us evaluate the above with an accuracy to within $(1-k^2)$. The first two terms of the series have logarithmic behavior, so that it is useful to isolate these terms. Evaluating the desired integrals by elementary means we rewrite (28) as follows:

$$\Pi(-\alpha_\pm, k) = k^2\left(\frac{1}{2(k^2-\alpha_\pm)}\log\frac{1+k}{1-k} - \frac{(\alpha_\pm)^{1/2}}{2k}\log\frac{1+(\alpha_\pm)^{1/2}}{1-(\alpha_\pm)^{1/2}}\right)$$
$$+ \tfrac{1}{2}k^2k'^2\left(\frac{1}{4(k^2-\alpha_\pm)}\log\frac{1+k}{1-k} + \frac{k}{2k'^2(k^2-\alpha_\pm)} - \frac{\alpha_\pm}{(k^2-(\alpha_\pm)^{1/2})^2}\log\frac{1+k}{1-k} + \frac{(\alpha_\pm)^{3/2}}{2k(k^2-\alpha_\pm}\log\frac{1+(\alpha_\pm)^{1/2}}{1-(\alpha_\pm)^{1/2}}\right)$$
$$+ k^2\sum_{n=2}^{\infty}k'^{2n}\frac{(2n-1)!!}{2n!!} \times \int_1^{1/k'}\frac{t^{2n+2}dt}{[\alpha_\pm + (k^2-\alpha_\pm)t^2](t^2-1)^{1/2}}. \tag{29}$$

Consider now a typical integral occurring in the summation in (29),

$$I_{2n+2} = \int_1^{1/k'}\frac{t^{2n+2}dt}{(\alpha_\pm + \gamma_\pm t^2)(t^2-1)^{1/2}},$$
$$\gamma_\pm = k^2 - \alpha_\pm, \quad n \geq 2. \tag{30}$$

Split the rational factor in the integrand above as

$$\frac{t^{2n+2}}{\alpha_\pm + \gamma_\pm t^2} = \frac{1}{\gamma_\pm}\left(t^{2n} - \frac{\alpha_\pm t^{2n}}{\alpha_\pm + \gamma_\pm t^2}\right)$$
$$= \frac{t^{2n}}{\gamma_\pm} - \frac{\alpha_\pm}{\gamma_\pm^2}\left(t^{2n-2} - \frac{t^{2n-2}}{\alpha_\pm + \gamma_\pm t^2}\right) = \text{and so on}, \tag{31}$$

thus generating an infinite series for I_{2n+2}. Each term of this latter series has essentially the same structure and can be written as

$$\int_1^{1/k'}\frac{t^{2n}}{(t^2-1)^{1/2}}dt = \int_1^{1/k'}t^{2n-2}(t^2-1)^{1/2}dt + \int_1^{1/k'}t^{2n-4}(t^2-1)^{1/2} + \cdots. \tag{31'}$$

The individual terms on the right-hand side of Eq. (31′) are now subjected to (repeated) partial integration thus generating a series in powers of $k'^2(=1-k^2)$. Thus to the desired degree of accuracy

$$\int_1^{1/k'}t^{2n-2}(t^2-1)^{1/2}dt = \frac{1}{2n}\frac{k^3}{k'^{2n}} + \frac{(2n-3)}{2n(2n-2)}\frac{k^3}{k'^{2n-2}} + O\left(\frac{1}{k'^{2n-4}}\right) \tag{32}$$

and hence

$$\int_1^{1/k'}\frac{t^{2n}}{(t^2-1)^{1/2}}dt = \frac{1}{2n}\frac{k^3}{k'^{2n}} + \frac{2n-3}{2n(2n-2)}\frac{k^3}{k'^{2n-2}} + \frac{1}{2(n-1)}\frac{k^3}{k'^{2n-2}} + O\left(\frac{1}{k'^{2n-4}}\right). \tag{32'}$$

Substituting (32′) and (31) into (30), we get finally

$$I_{2n+2} = \frac{1}{\gamma_\pm}\left(\frac{k^3}{(2n)k'^{2n}} + \frac{(2n-3)k^3}{2n(2n-2)k'^{2n-2}} + \frac{k^3}{2(n-1)k'^{2n-2}}\right) - \frac{\alpha_\pm}{\gamma_\pm^2}\,\frac{k^3}{(2n-2)k'^{2n-2}} + O\left(\frac{1}{k'^{2n-4}}\right). \tag{33}$$

Inserting (33) into (28) we get $\Pi(-\alpha_\pm, k)$ to the desired accuracy. Retaining, henceforth, terms only up to order $1-k^2$, we get, after some manipulations:

$$\Pi(-\alpha_\pm, k) = \frac{1}{1-\alpha_\pm}\log 2 - \frac{1}{2(1-\alpha_\pm)}\log k'^2 - \frac{1}{2}(\alpha_\pm)^{1/2}\log\frac{1+(\alpha_\pm)^{1/2}}{1-(\alpha_\pm)^{1/2}} + \frac{1}{4(1-\alpha_\pm)} + \frac{1}{(1-\alpha_\pm)}\sum_{n=2}^{\infty}\frac{(2n-1)!!}{2n!!}\frac{1}{2n} + O(1-k^2)$$

$$= \frac{2}{1-\alpha_\pm}\log 2 - \frac{1}{2(1-\alpha_\pm)}\log k'^2 - \frac{(\alpha_\pm)^{1/2}}{2}\log\frac{1+(\alpha_\pm)^{1/2}}{1-(\alpha_\pm)^{1/2}} + O(1-k^2). \tag{34}$$

Using (9) we rewrite (34) as

$$\Pi(-\alpha_\pm, k) = 2\left(\frac{u_\pm - u_{3c}}{u_\pm - u_{2c}}\right)\log 2 - \frac{1}{2}\left(\frac{u_\pm - u_{3c}}{u_\pm - u_{2c}}\right)\log(1-k^2) - \frac{1}{2}\left(\frac{u_{2c}-u_{3c}}{u_\pm - u_{3c}}\right)^{1/2} \times \log\frac{(u_\pm - u_{3c})^{1/2} + (u_{2c}-u_{3c})^{1/2}}{(u_\pm - u_{3c})^{1/2} - (u_{2c}-u_{3c})^{1/2}} + O(1-k^2) \tag{35}$$

which is our final result for $\Pi(-\alpha_\pm, k)$.

The other integral in (11) causes no trouble as this is well-behaved at $k^2 = 1$. Evaluating by straightforward means we have

$$\Pi(\Psi_0, -\alpha_\pm, k) = \int_0^{\sin\Psi_0}\frac{dy}{(1-\alpha_\pm y^2)[(1-y^2)(1-k^2y^2)]^{1/2}}$$

$$= \frac{u_\pm - u_{3c}}{u_\pm - u_{2c}}\tanh^{-1}\left(\frac{u_{3c}}{u_{3c}-u_{2c}}\right)^{1/2} - \frac{[(u_\pm - u_{3c})(u_{2c}-u_{3c})]^{1/2}}{u_\pm - u_{2c}}\tanh^{-1}\left(\frac{u_{3c}}{u_{3c}-u_\pm}\right)^{1/2} + \cdots O(1-k^2). \tag{36}$$

From (36), (35), and (10)-(12) we get, after some algebra,

$$\chi = -\pi - \frac{2}{[2m(u_{2c}-u_3{}^c)]^{1/2}(m^2-a^2)^{1/2}(1+\gamma a\alpha_c)} \times [\xi_+ - \xi_-] + O(1-k^2) \tag{37}$$

with

$$\xi_\pm = \frac{1-2m(1+\gamma a\alpha_c)u_\pm}{u_\pm - u_{3c}}\left[2\left(\frac{u_\pm - u_{3c}}{u_\pm - u_{2c}}\right)\log 2 - \frac{1}{2}\left(\frac{u_\pm - u_{3c}}{u_\pm - u_{2c}}\right)\log(1-k^2) - \frac{1}{2}\left(\frac{u_{2c}-u_{3c}}{u_\pm - u_{3c}}\right)^{1/2}\log\frac{(u_\pm - u_{3c})^{1/2} + (u_{2c}-u_{3c})^{1/2}}{(u_\pm - u_{3c})^{1/2} - (u_{2c}-u_{3c})^{1/2}} - \frac{u_\pm - u_{3c}}{u_\pm - u_{2c}}\tanh^{-1}\left(\frac{u_{3c}}{u_{3c}-u_{2c}}\right)^{1/2} + \frac{[(u_\pm - u_{3c})(u_{2c}-u_{3c})]^{1/2}}{u_\pm - u_{2c}}\tanh^{-1}\left(\frac{u_{3c}}{u_{3c}-u_\pm}\right)^{1/2}\right]. \tag{37'}$$

To get the scattering angle θ we must subtract off from χ the appropriate number of $2\pi n$. Thus as $k^2 \to 1$ $(\alpha \to \alpha_c)$ the term $\log(1-k^2)$ $[\simeq \log(\alpha - \alpha_c)^{1/2}]$ dominates and controls the number of spirals. It therefore follows from (37) and (37′) that

$$\frac{d\chi}{d\alpha_n} = \frac{1-2m(1+\gamma a\alpha_c)u_{2c}}{[2m(u_{2c}-u_{3c})]^{1/2}(1+\gamma a\alpha_c)(1-2mu_{2c}+a^2u^2{}_{2c})(\alpha_c-\alpha_n)} + \cdots$$

$$= \frac{1-2m(1+\gamma a\alpha_c)u_{2c}}{[2m(u_{2c}-u_{3c})]^{1/2}(1+\gamma a\alpha_c)(1-2mu_{2c}+a^2u^2{}_{2c})} \times \exp\left([2m(u_{2c}-u_{3c})]^{1/2}\,\frac{(1+\gamma a\alpha_c)(1-2mu_{2c}+a^2u^2{}_{2c})}{1-2m(1+\gamma a\alpha_c)u_{2c}}(2n\pi \pm \theta)\right). \tag{38}$$

From (38) and (14) we finally write down the differential cross section:

$$\frac{d\sigma}{d\Omega} = \left(\frac{d\sigma}{d\Omega}\right)_{n=0} + \frac{z[2m(u_{2c}-u_{3c})]^{1/2}}{\sin\theta(\gamma^2-1)}\sum_{n=1}^{\infty}\frac{\exp\{-[2m(u_{2c}-u_{3c}])^{1/2}z(2n\pi+\theta)\} + \exp\{-[2m(u_{2c}-u_{3c})]^{1/2}z(2n\pi-\theta)\}}{\alpha^3{}_n}$$

$$= \left(\frac{d\sigma}{d\Omega}\right)_{n=0} + \frac{2z[2m(u_{2c}-u_{3c})]^{1/2}}{(\gamma^2-1)\alpha^3{}_c\sin\theta}\exp\{-2\pi[2m(u_{2c}-u_{3c})]^{1/2}z\}\cosh\{[2m(u_{2c}-u_{3c})]^{1/2}z\theta\} + \cdots. \tag{39}$$

In the above, we have used the abreviation

$$z = \frac{(1 + \gamma a \alpha_c)(1 - 2mu_{2c} + a^2 u^2{}_{2c}}{1 - 2m(1 + \gamma a \alpha_c) u_{2c}} \tag{40}$$

and the expression for the zero spiral cross section $(d\sigma/d\Omega)_{n=0}$ is quite evident from (37). The above result Eq. (39) reduces to that of Ref. 1 when $a \to 0$ $(z \to 1)$. The factor z in (39), as also the explicit expressions for u_{2c} and u_{3c} [Eq. (25)], contain the effect of rotation of the source.

In the above we have expressed the cross section in terms of α_c, which should now be determined in order to complete the discussion of this section. Unlike the case with Schwarzschild metric, it is not possible, in the present instance to closely evaluate α_c in terms of the parameters γ, m, and a of the scattering problem and one is led to take recourse to numerical calculations.

This is done in the next section.

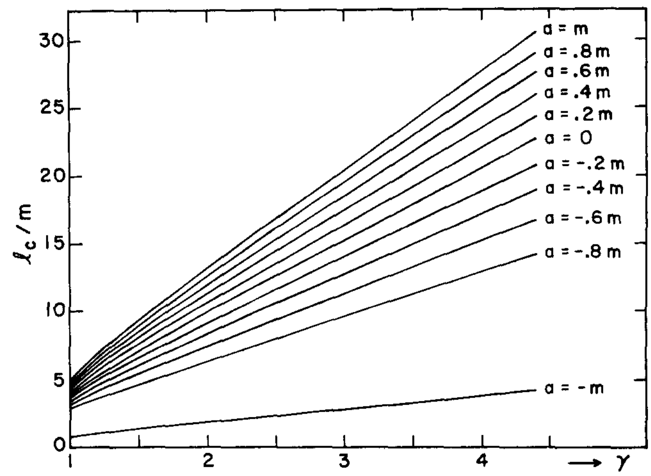


FIG. 1. Plot of l_c/m versus γ, for selected values of a. γ is the energy per unit mass of the test particle at asymptotic radial distances, and a the angular momentum per unit mass of the source. l_c is the critical angular momentum of the best particle for which gravitational capture takes place.

IV. EVALUATION OF α_c

We recall that α_c is the inverse of the critical angular momentum per unit mass of the test particle for which gravitational capture takes place. Evaluation of this quantity in the Kerr metric is of intrinsic interest. From the standard condition for the equality of two roots of a cubic equation we derive the equation for α_c:

$$\frac{\Gamma^2}{16m^2}(1 + \gamma a \alpha_c)^4 \alpha^2_c + \tfrac{1}{27}\alpha^4_c (1 + \gamma a \alpha_c)^2$$

$$- \frac{\Gamma}{24m^2}(\Gamma a^2 \alpha^2_c - 1)(1 + \gamma a \alpha_c)^2 \alpha^2_c$$

$$+ \frac{\Gamma}{432m^4}(\Gamma a^2 \alpha^2_c - 1)^3 - \frac{1}{432m^2}(\Gamma a^2 \alpha^2_c - 1)^2 \alpha^2_c = 0,$$

$$\Gamma = \gamma^2 - 1. \tag{41}$$

In the limiting case $a = 0$, the above collapses essentially to a quadratic equation, whose solution gives the known Schwarzschild metric value of α_c,

$$\alpha_c(a = 0) = \left(\frac{\gamma(9\gamma^2 - 8)^{3/2} - 27\gamma^4 + 36\gamma^2 - 8}{32m^2}\right)^{1/2}. \tag{42}$$

In the general case $a = 0$ it is obviously not possible in general, to obtain analytic solution of the sixth order Eq. (41).

However, in the special case $\gamma = 1$, which corresponds to the test particle having nonrelativistic velocity initially (asymptotic radial distance from the source), Eq. (41) can again be solved and one finds

$$\alpha_c(\gamma = 1) = [-1 + (1 \pm a/m)^{1/2}]/2a. \tag{43}$$

The above generalizes the corresponding result in Schwarzschild metric, $\alpha_c(\gamma = 1, a = 0) = \pm 1/4m$. In the general case the Eq. (41) is solved numerically and the conclusions summarized in Fig. 1.

The dimensionless quantity l_c/m $(l_c = \alpha_c^{-1})$ is plotted against γ. Graphs with positive values of a correspond to the case when the angular momentum of the source is parallel to that of the test particle and those with negative a the case when it is antiparallel. One feature of the graph for $a = -m$ needs mention: When γ lies in the range $3.2 \leq \gamma \leq 4.0$, the corresponding impact parameter falls within the event horizon, i.e., $l_c(\gamma^2 - 1)^{-1/2} < m$. Hence nontrivial gravitational capture does not take place at these energies.

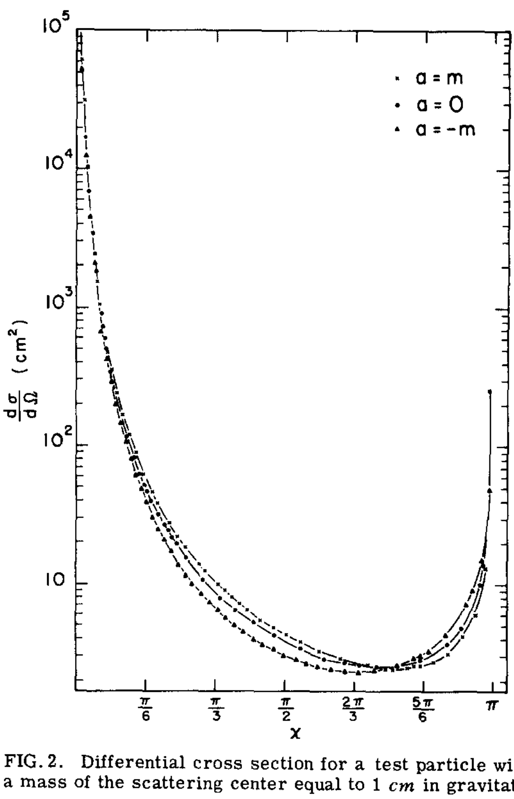


FIG. 2. Differential cross section for a test particle with $\gamma^2 = 5$ and a mass of the scattering center equal to 1 cm in gravitational units.

V. INTERMEDIATE RESULTS

In the above we have obtained asymptotic expressions for the cross section in the limit of large and the critical angular momentum. For general values of the angular momentum lying between these two limiting cases, the differential cross section can be evaluated numerically. This has been done on a computer for the case when the mass of the scattering center is one centimeter in gravitational units and $\gamma^2 = 5$. The result is shown in Fig. 2. The following features of Fig. 2 should be noted:

(1) In the range of angles considered the angle of deviation χ equals the scattering angle θ and there is no spiraling effect. The latter begins to show up as one approaches the backward direction $\theta = \pi$.

(2) The differential cross section diverges in the backward direction. The divergence, however, is of kinematic origin (and hence integrable) coming from the presence of the $\sin\theta$ term in the denominator of the expression for the cross section in Eq. (14). For the Schwarzschild metric the appearance of the divergent differential cross section in the backward direction was found already in Ref. 1.

(3) Fig. 2 shows that the rotation of the source has a very small effect on the differential cross section. Indeed, even for the maximum allowed value of angular momentum of the source, $a = m$, the effect is quite insignificant.

ACKNOWLEDGMENTS

We thank Professor W. R. Johnson and Professor W. D. McGlinn for many conversations from which we have benefited.

[1]P. A. Collins, R. Delbourgo, and R. M. Williams, J. Phys. A **6**, 161 (1973).
[2]R. P. Kerr, Phys. Rev. Lett. **11**, 237 (1963).
[3]B. Carter, Phys. Rev. Lett. **26**, 331 (1971); W. Israel, Phys. Rev. **164**, 1776 (1967).
[4]R. H. Boyer and R. W. Lindquist, J. Math. Phys. **8**, 265 (1967).
[5]W. Gröbner and N. Hofreiter, *Integraltafel*, Erster Teil (Springer-Verlag, Berlin, 1957), p. 60.

Theory of self-reproducing kernel and dispersion inequalities*

Susumu Okubo

Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627
(Received 4 September 1973)

The theory of the self-reproducing kernel by Aronszajn has been investigated for a Hilbert space with norm

$$\|f\|^2 = (1/2\pi)\int_0^{2\pi} d\theta |f(e^{i\theta})|^2 + (1/\pi)\int_{-1}^{1} dx\, \lambda(x)|f(x)|^2,$$

where $f(z)$ is a H^2 function and $\lambda(x)$ is a nonnegative summable function. The self-reproducing kernel of this space satisfies an integral equation. The dispersion inequalities for various problems in the high energy physics can be treated in unified and generalized manner by this theory.

1. SUMMARY OF THE PROBLEM

As we shall show in Sec. 4, there are many problems in high energy physics which can be reduced to the following mathematical question. Let $f(z)$ be a holomorphic function of a complex variable z inside the unit disk $|z| < 1$. Assume that $f(z)$ belongs to the class H^2 with the norm[1]

$$\|f\|_2^2 = \lim_{r\to 1-0} \frac{1}{2\pi}\int_0^{2\pi} d\theta\, |f(re^{i\theta})|^2. \tag{1.1}$$

Let $\lambda(x)$ be a non negative measurable function defined on the real interval $-1 \le x \le 1$, and define a new norm $\|f\|$ by

$$\|f\|^2 = \frac{1}{2\pi}\int_0^{2\pi} d\theta\, |f(e^{i\theta})|^2 + \frac{1}{\pi}\int_{-1}^{1} dx\lambda(x)|f(x)|^2. \tag{1.2}$$

We then ask the following question. Suppose that an upper bound A of $\|f\|^2$ is given, i.e.,

$$\|f\|^2 \le A. \tag{1.3}$$

Moreover, a value or values of $f(z)$ at n interior point $z = z_j$, $|z_j| < 1$ $(j = 1, 2, —, n)$ are known to be

$$f(z_j) = a_j, \quad (j = 1, 2, \dots, n). \tag{1.4}$$

Then, is there any H^2 function $f(z)$ satisfying these conditions? If the answer is yes, then can we give an optimal bound for absolute values of $f(z)$ and its derivatives inside the unit disk $|z| < 1$ in terms of these constants A and a_j? If we have a solution to the problem, then values of a_j must satisfy a constraint inequality. Especially, the case $\lambda(x) \equiv 0$ essentially reduces the problem to the so-called minimum interpolating problem[2] of the H^2-space, where the explicit solution can be readily found. However, the general case with $\lambda(x) \neq 0$ is more complicated and we shall solve it by means of the theory of self-reproducing kernel.[3]

Since $\lambda(x)$ is assumed to be nonnegative we find

$$\|f\| \ge \|f\|_2. \tag{1.5}$$

We shall denote by $H^2(\lambda)$ a set consisting of all H^2 functions $f(z)$ with finite new norm $\|f\|$. Obviously, $H^2(\lambda)$ is a subspace of H^2. As we shall prove in Sec. 3, it is a Hilbert space with self-reproducing kernel $K(z, \xi)$. Following the notation of Aronszajn, [3] we shall often write the inner product in $H^2(\lambda)$ by

$$(f, g) = (f(z), g(z))_z \tag{1.6}$$

so as to indicate that z is the integration variable. Then, we have

$$f(\xi) = (f(z), K(z, \xi))_z \tag{1.7}$$

for all $H^2(\lambda)$-functions $f(z)$ and for all complex values ξ with $|\xi| < 1$. As we shall show in Sec. 3, $K(z, \xi)$ must satisfy an integral equation

$$\frac{1}{1 - z\bar{\xi}} = K(z, \xi) + \frac{1}{\pi}\int_{-1}^{1} dx \frac{\lambda(x)}{1 - xz} K(x, \xi) \tag{1.8}$$

if $\lambda(x)$ is a function belonging to $L^1(1, -1)$.

In order to solve the original problem, let us set

$$g_j(z) = K(z, z_j) \quad (j = 1, 2, \dots, n). \tag{1.9}$$

Then, the condition (1.4) and the equation (1.7) are rewritten as

$$(f, g_j) = a_j \quad (j = 1, 2, \dots, n). \tag{1.10}$$

As we see from (1.8), the n-functions $g_j(z)$ $(j = 1, 2, \dots, n)$ defined by (1.9) are linearly independent, if n-points $z_1, z_2, \dots, z_n$ are all distinct. Then, by Schmidt's orthonormalization procedure, we can construct n orthornormal functions $h_j(z)$ $(j = 1, 2, \dots, n)$ as linear combinations of $g_j(z)$:

$$h_j(z) = \sum_{\mu=1}^{n} C_{j\mu} g_\mu(z), \tag{1.11}$$

$$(h_i, h_j) = \delta_{ij} \quad (i, j = 1, 2, \dots, n). \tag{1.12}$$

The numerical coefficients $C_{j\mu}$ in (1.11) are determined from (1.12) by

$$\sum_{\mu=1}^{n}\sum_{\nu=1}^{n} C_{j\mu}\bar{C}_{k\nu} K(z_\nu, z_\mu) = \delta_{jk}, \tag{1.13}$$

where we have used (1.7) for $f(z) = g_\mu(z) = K(z, z_\mu)$. Therefore, $C_{\mu\nu}$ depend only upon $z_1, z_2, \dots, z_n$. Now, the Bessel inequality

$$A \ge \|f\|^2 \ge \sum_{\mu=1}^{n} |(f, h_\mu)|^2$$

gives the desired bound[4]

$$A \ge \sum_{\mu=1}^{n} \Big|\sum_{\mu=1}^{n} \bar{C}_{\mu\nu} a_\nu\Big|^2 \tag{1.14}$$

for a_μ $(\mu = 1, 2, \dots, n)$. An inequality involving $f(\xi)$ can be obtained by letting $n \to n+1$ in (1.14) with $g_{n+1}(z) = K(z, \xi)$, since we have

$$(f, g_{n+1}) = f(\xi).$$

Similarly, a bound involving the derivative $f'(\xi)$ can be also derived by choosing

$$g_{n+1}(z) = \frac{\partial}{\partial\bar{\xi}} K(z, \xi)$$

Since then we have (see Sec. 2)

$$\frac{d}{d\xi} f(\xi) = (f(z), \frac{\partial}{\partial \xi} K(z, \xi))_z \tag{1.15}$$

These bounds are optimal in a sense that we cannot improve them[4] without additional informations.

For our applications to the high energy physics (see the Sec. 4), $\lambda(x)$ is in general bounded in the interval $-1 \leq x \leq 1$,

$$0 \leq \lambda(x) \leq M. \tag{1.16}$$

In that case, the space $H^2(\lambda)$ is the same set as H^2. To illustrate this, we have only to note the Hilbert–Fejér–Riesz inequality[1,5]

$$\frac{1}{\pi} \int_{-1}^{1} dx |f(x)|^2 \leq \frac{1}{2\pi} \int_0^{2\pi} d\theta |f(e^{i\theta})|^2 = \|f\|_2^2 \tag{1.17}$$

for $f \in H^2$. Hence, we have

$$\|f\|_2^2 \leq \|f\|^2 \leq (1 + M)\|f\|_2^2. \tag{1.18}$$

We shall also discuss in the Sec. 3 conditions under which iterative solution of the integral equation (1.8) converges to and gives the correct answer. One sufficient condition is that we have $M < 1$.

Applications to high energy physics will be discussed in Sec. 4.

2. SELF-REPRODUCING KERNEL

Let $\mathcal{H}$ be a Hilbert space consisting of complex-valued functions $f(x)$ defined in a set D. Following Aronszajn,[3] we shall denote the inner product in $\mathcal{H}$ as

$$(f, g) = (f(x), g(x))_x \tag{2.1}$$

when we wish to emphasize the fact that the integration variable is x. A sum of two functions and a multiplication of a function by a complex number in $\mathcal{H}$ are defined in a natural fashion. Then a function $K(x, y)$ defined in $D \times D$ will be called[3] a self-reproducing kernel of $\mathcal{H}$, if it satisfies the following two conditions:

(i) For any fixed value $y \in D, K(x, y)$ regarded as a function of the variable x is an element of $\mathcal{H}$.
(ii) For all $f \in \mathcal{H}$, we have

$$(f(x), K(x, y))_x = f(y) \tag{2.2}$$

for all $y \in D$.

We shall prove the following theorems[3] for later purposes.

Theorem I: A necessary and sufficient condition that $\mathcal{H}$ has a self-reproducing kernel is that a linear functional defined by

$$x(f) = f(x), \quad f \in \mathcal{H},$$

is continuous for any fixed value $x \in D$.

The necessary part follows from the definition (2.2) since by the Schwarz inequality we have

$$|f(y)| \leq \|f\| \, \|K(x, y)\|_x = \|f\| [K(y, y)]^{1/2}.$$

The sufficiency of the condition results from the theorem of F. Riesz.

Theorem II:

$$(K(x, z), K(x, y))_x = K(y, z), \tag{2.3}$$

$$\overline{K(x, y)} = K(y, x). \tag{2.4}$$

Also for arbitrary n points $x_1, x_2, \ldots, x_n \in D$ and for n complex numbers $C_1, C_2, \ldots, C_n$ we have

$$\sum_{\mu=1}^{n} \sum_{\nu=1}^{n} C_\mu K(x_\mu, x_\nu) \overline{C}_\nu \geq 0. \tag{2.5}$$

Equation (2.3) follows immediately from (2.2) by setting $f(x) = K(x, z)$. Then, (2.4) is an immediate consequence of (2.3). Also, if we set

$$f(x) = \sum_{\nu=1}^{n} \overline{C}_\nu K(x, x_\nu)$$

and if we use (2.3), we compute

$$0 \leq \|f\|^2 = \sum_{\mu=1}^{n} \sum_{\nu=1}^{n} C_\mu K(x_\mu, x_\nu) \overline{C}_\nu.$$

This proves (2.5). Hereafter, we shall designate the inequality (2.5) simply as

$$K \geq 0. \tag{2.6}$$

The special cases $n = 1$ and $n = 2$ in (2.5) give us

$$K(x, x) \geq 0, \quad K(x, x)K(y, y) \geq |K(x, y)|^2. \tag{2.7}$$

The self-reproducing kernel is unique if it exists.

Suppose that we have another self-reproducing kernel $L(x, y)$.

Then from (2.2) we must have

$$(K(x, y), L(x, z))_x = K(z, y),$$

$$(L(x, z), K(x, y))_x = L(y, z).$$

However, the hermiticity condition (2.4) demands then $L(y, z) = K(y, z)$ which proves the uniqueness.

Theorem III: If a sequence $\{f_1, f_2, \ldots\}$ weakly converges to f, it then converges point-wise in D, i.e.,

$$\lim_{n \to \infty} f_n(x) = f(x), \quad x \in D.$$

This follows from (2.2) and the definition of the weak limit.

Theorem IV: Suppose that $\mathcal{H}$ is a separable Hilbert space with the self-reproducing kernel $K(x, y)$. We can then expand

$$K(x, y) = \sum_n \phi_n(x) \overline{\phi_n(y)}$$

for any complete orthonormal set $\phi_n(x) (n = 1, 2, \ldots)$.

To prove this, we expand $K(x, y) \in \mathcal{H}$ for a fixed y into a Fourier series

$$K(x, y) = \sum_n a_n \phi_n(x).$$

Then, the coefficient a_n is computed by

$$a_n = (K(x, y), \phi_n(x))_x = \overline{\phi_n(y)}.$$

Combining this fact with the Theorem III, we prove the desired result. This theorem is sometimes useful to compute $K(x, y)$.

For example, the H^2 space has a complete orthonormal set $\phi_n(z) = z^n$ ($n = 0, 1, 2, \ldots$). Therefore, its self-reproducing kernel $K_0(z, \xi)$ is given by

$$K_0(z, \xi) = 1/(1 - z\bar{\xi}), \tag{2.8}$$

since the space H^2 is known to have self-reproducing kernel. This fact has been utilized[4] to derive dispersion inequalities for the case $\lambda(x) = 0$.

For our physical applications, $f(x)$ is a holomorphic function $f(z)$ in the unit disk $|z| < 1$. We can then define the derivative $d/dz\, f(z)$. More generally, suppose that the derivative

$$\frac{\partial}{\partial y} K(x, y) = \lim_{\Delta y \to 0} \frac{1}{\Delta y}[K(x, y + \Delta y) - K(x, y)]$$

is somehow defined as a weak limit in $\mathcal{H}$. Then the following theorem is an immediate consequence of (2.2).

Theorem V: If $\partial/\partial y\, K(x, y)$ exists as a weak limit in $\mathcal{H}$ for some values of y, then the derivative $d/dy\, f(y)$ also exists and we have

$$\frac{d}{dy} f(y) = \left(f(x), \frac{\partial}{\partial y} K(x, y)\right)_x.$$

Note that $d/dy\, f(y)$ may not be an element of $\mathcal{H}$, nor need it be defined on all points of D.

Next, in order to establish a fact that the space $H^2(\lambda)$ defined in the previous section has the self-reproducing kernel, we shall prove the following theorem.

Theorem VI: Suppose that $\mathcal{H}_1$ is a subspace of a Hilbert space $\mathcal{H}_2$ whose norm is $\|f\|_2$. Moreover, we assume that $\mathcal{H}_1$ is also a Hilbert space with a new norm $\|f\|_1$ satisfying

$$\|f\|_1 \geq \|f\|_2 \tag{2.9}$$

for all $f \in \mathcal{H}_1$. If the larger space $\mathcal{H}_2$ has the self-reproducing kernel $K_2(x, y)$, then the Hilbert space $\mathcal{H}_1$ has also self reproducing kernel $K_1(x, y)$ with property

$$K_2 \geq K_1. \tag{2.10}$$

To establish this theorem, let us define a functional $x(f)$ in $\mathcal{H}_1$ by setting

$$x(f) = f(x).$$

Since $\mathcal{H}_1$ as a set is a subspace of $\mathcal{H}_2$, we compute

$$|x(f)| = |f(x)| \leq \|f\|_2[K_2(x, x)]^{1/2} \leq \|f\|_1[K_2(x, x)]^{1/2}.$$

This shows that the functional $x(f)$ is bounded in $\mathcal{H}_1$ and hence the Theorem I tells us that $\mathcal{H}_1$ must have self-reproducing kernel $K_1(x, y)$. To prove the second half of the theorem, let us set

$$f(x) = \sum_{\mu=1}^{n} C_\mu[K_1(x, x_\mu) - K_2(x, x_\mu)],$$

$$g(x) = \sum_{\mu=1}^{n} C_\mu K_1(x, x_\mu).$$

Regarding $f(x)$ and $g(x)$ as functions in $\mathcal{H}_2$, we compute

$$0 \leq \|f\|_2^2 = \|g\|_2^2 + \sum_{\mu=1}^{n} \sum_{\nu=1}^{n} C_\mu[K_2(x_\nu, x_\mu) - 2K_1(x_\nu, x_\mu)]\bar{C}_\nu.$$

However, noting

$$\|g\|_2^2 \leq \|g\|_1^2 = \sum_{\mu=1}^{n} \sum_{\nu=1}^{n} C_\mu K_1(x_\nu, x_\mu)\bar{C}_\nu,$$

this leads to

$$\sum_{\mu=1}^{n} \sum_{\nu=1}^{n} C_\mu\{K_2(x_\nu, x_\mu) - K_1(x_\nu, x_\mu)\}\bar{C}_\nu \geq 0$$

which is (2.10). In this connection, we should remark that if $\mathcal{H}_1$ and $\mathcal{H}_2$ are the same space as a set, and if we have a stronger condition

$$\|f\|_2 \leq \|f\|_1 \leq C\|f\|_2 \tag{2.11}$$

as in (1.18), then we will have a stronger inequality

$$K_1 \leq K_2 \leq C^2 K_1. \tag{2.12}$$

It is easy to verify[4] that the space H^2 is a Hilbert space with self-reproducing kernel (2.8). Then, identifying now

$$\mathcal{H}_2 = H^2, \qquad \mathcal{H}_1 = H^2(\lambda),$$

and noting (1.5), the Theorem VI guarantees that the space $H^2(\lambda)$ has self-reproducing kernel.

In order to calculate explicit form for the kernel, the following theorem is useful.

Theorem VII: Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two Hilbert spaces as is defined in the previous theorem. Then, there is a linear operator Q which maps $\mathcal{H}_2$ into $\mathcal{H}_1$ with properties

$$(f, g)_2 = (f, Qg)_1, \quad f \in \mathcal{H}_1, g \in \mathcal{H}_2,$$

$$\|Qg\|_1 \leq \|g\|_2. \tag{2.13}$$

The self-reproducing kernel $K_1(x, y)$ of $\mathcal{H}_1$ is given by the formula

$$K_1(x, y) = (Qg_y)(x), \qquad g_y(x) = K_2(x, y). \tag{2.14}$$

To prove this statement, we note

$$|(f, g)_2| \leq \|f\|_2\|g\|_2 \leq \|f\|_1\|g\|_2$$

for any $f \in \mathcal{H}_1$ and $g \in \mathcal{H}_2$. Hence, a linear functional defined by

$$G(f) = (f, g)_2$$

is bounded in $\mathcal{H}_1$. Therefore, F. Riesz's theorem assures us that we can write

$$(f, g)_2 = (f, \tilde{g})_1. \tag{2.15}$$

Setting now

$$\tilde{g} = Qg, \tag{2.16}$$

Q is a linear mapping of $\mathcal{H}_2$ into $\mathcal{H}_1$. Especially, choosing

$$g(x) = g_y(x) = K_2(x, y),$$

then (2.15) gives us

$$f(y) = (f, Qg_y)_1.$$

This proves that Qg_y is indeed the self-reproducing kernel of $\mathcal{K}_1$. The inequality in (2.13) is derived by setting $f = \tilde{g}$ in (2.15) and noting $\|\tilde{g}\|_1^2 = |(\tilde{g}, g)_2| \leq \|\tilde{g}\|_2 \|g\|_2 \leq \|\tilde{g}\|_1 \|g\|_2$.

3. SPACE $H^2(\lambda)$

We must first prove that our space $H^2(\lambda)$ is a Hilbert space. In accordance with the notation used in the previous section, let us define inner products of H^2 and $H^2(\lambda)$ by

$$(f,g)_2 = \lim_{r \to 1-0} \frac{1}{2\pi} \int_0^{2\pi} d\theta\, f(re^{i\theta}) \overline{g(re^{i\theta})} = \frac{1}{2\pi} \int_0^{2\pi} d\theta f(e^{i\theta}) \overline{g(e^{i\theta})} \tag{3.1}$$

and

$$(f,g)_1 = (f,g)_2 + \frac{1}{\pi} \int_{-1}^{1} dx \lambda(x) f(x) \overline{g(x)} \tag{3.2}$$

for two H^2 functions $f(z)$ and $g(z)$. Obviously, $H^2(\lambda)$ is a pre-Hilbert space, and therefore it is only necessary to prove its completeness.

Suppose that a sequence $f_1, f_2, \ldots$ is a Cauchy sequence in $H^2(\lambda)$, i.e.,

$$\|f_n - f_m\|_1 \to 0, \quad (n, m \to \infty). \tag{3.3}$$

Then, since $\lambda(x)$ is nonnegative by assumption, this implies

$$\|f_n - f_m\|_2 \to 0, \tag{3.4}$$

$$\int_{-1}^{1} dx \lambda(x) |f_n(x) - f_m(x)|^2 \to 0 \tag{3.5}$$

for $n, m \to \infty$. Since the space H^2 is complete, the first condition (3.4) implies that we have a holomorphic function $f(z) \in H^2$ satisfying[1]

$$\lim_{n\to\infty} \|f - f_n\|_2 = 0, \tag{3.6}$$

$$\lim_{n\to\infty} f_n(z) = f(z), \quad |z| < 1. \tag{3.7}$$

The Fatou's lemma[6] on nonnegative measurable functions implies

$$\liminf_{m\to\infty} \int_{-1}^{1} dx \lambda(x) |f_n(x) - f_m(x)|^2 \geq \int_{-1}^{1} dx \lambda(x) \liminf_{m\to\infty} |f_n(x) - f_m(x)|^2.$$

Therefore, (3.5) and (3.7) give us

$$\lim_{n\to\infty} \int_{-1}^{1} dx \lambda(x) |f_n(x) - f(x)|^2 = 0.$$

Combining this with (3.6) leads to the desired result

$$\lim_{n\to\infty} \|f_n - f\|_1 = 0$$

which proves the completeness of the space $H^2(\lambda)$. Thus, $H^2(\lambda)$ is a Hilbert space and also we have

$$\|f\|_1 \geq \|f\|_2, \quad f \in H^2(\lambda). \tag{3.8}$$

Therefore, the Theorem VI guarantees that $H^2(\lambda)$ must have the self-reproducing kernel which we rewrite here $K(z, \xi)$ instead of $K_1(x, y)$. Next, we shall prove the following result.

Theorem VIII: Suppose that $\lambda(x)$ is a summable function of x in the interval $-1 \leq x \leq 1$. Then, the self-reproducing kernel $K(z, \xi)$ of the space $H^2(\lambda)$ must satisfy the integral equation

$$K_0(z, \xi) = K(z, \xi) + \frac{1}{\pi} \int_{-1}^{1} dx K_0(z, x) \lambda(x) K(x, \xi), \tag{3.9}$$

$$K_0(z, \xi) = 1/(1 - z\bar{\xi}). \tag{3.10}$$

Conversely, solution of the integral equation (3.9) such that $\lambda(x)K(x, \xi)$ belongs to $L^2(1, -1)$ is unique and we have for all $f \in H^2$

$$f(\xi) = (f(z), K(z, \xi))_1. \tag{3.11}$$

The proof of the first part of the theorem is easy. In order to avoid confusion, let us set

$$h_1(z) = K(z, \xi),$$

$$h_2(z) = K_0(z, \eta)$$

for fixed values of ξ and η with $|\xi| < 1$, and $|\eta| < 1$. If $\lambda(x)$ belongs to $L^1(1, -1)$, then we obviously have

$$\int_{-1}^{1} dx \lambda(x) |K_0(x, \eta)|^2 \leq \frac{1}{(1 - |\eta|)^2} \int_{-1}^{1} dx \lambda(x) < \infty$$

so that $h_2(z)$ is an element of $H^2(\lambda)$. From definition, we then have

$$(h_2, h_1)_1 = (h_2, h_1)_2 + \frac{1}{\pi} \int_{-1}^{1} dx \lambda(x) h_2(x) \overline{h_1(x)}. \tag{3.12}$$

Since h_2 and h_1 are self-reproducing kernels of spaces H^2 and $H^2(\lambda)$, respectively, we have

$$(f, h_1)_1 = f(\xi), \quad f \in H^2(\lambda),$$

$$(g, h_2)_2 = g(\eta), \quad g \in H^2.$$

Therefore, with the help of the hermiticity condition (2.4), we find

$$(h_2, h_1)_1 = h_2(\xi) = K_0(\xi, \eta),$$

$$(h_2, h_1)_2 = \overline{(h_1, h_2)_2} = \overline{h_1(\eta)} = K(\xi, \eta)$$

so that (3.12) is rewritten as

$$K_0(\xi, \eta) = K(\xi, \eta) + \frac{1}{\pi} \int_{-1}^{1} dx \lambda(x) K_0(x, \eta) K(\xi, x). \tag{3.13}$$

Taking the complex conjugate of this equation and noting (2.4), this gives (3.9) if we change the variable η by z.

To prove the second part of the theorem, we need the following lemma.

Lemma:

Let $R(x)$ be a function belonging to $L^2(1, -1)$, i.e.,

$$\int_{-1}^{1} dx |R(x)|^2 < \infty. \tag{3.14}$$

Then, a function $p(z)$ defined by

$$p(z) = \frac{1}{\pi} \int_{-1}^{1} dx \frac{1}{1 - xz} R(x) \tag{3.15}$$

belongs to H^2 with norm

$$\|p\|_2^2 = \frac{1}{\pi^2} \int_{-1}^{1} dx \int_{-1}^{1} dy \frac{R(x)\overline{R(y)}}{1 - xy} \leq \frac{1}{\pi} \int_{-1}^{1} dx |R(x)|^2. \tag{3.16}$$

Moreover, we find

$$\int_{-1}^{1} dx\,|p(x)|^2 \le \int_{-1}^{1} dx\,|R(x)|^2. \tag{3.17}$$

Finally, for any H^2 function $f(z)$, we have

$$(f,p)_2 = \frac{1}{\pi}\int_{-1}^{1} dx f(x)\overline{R(x)}. \tag{3.18}$$

Before we prove this lemma, we shall demonstrate the validity of the second half part of the Theorem VIII. Suppose that the integral equation (3.9) admits a solution such that $\lambda(x)K(x,\xi)$ belongs to $L^2(1,-1)$. Then, setting

$$p(z) = \frac{1}{\pi}\int_{-1}^{1} dx \frac{1}{1-xz}\lambda(x)K(x,\xi),$$

it defines a function belonging to H^2 by our lemma. Moreover, (3.18) gives

$$(f,p)_2 = \frac{1}{\pi}\int_{-1}^{1} dx\lambda(x)f(x)\overline{K(x,\xi)}$$

for any $f \in H^2$. If $K(z,\xi)$ satisfies the equation (3.9), then we have $p(z) = K_0(z,\xi) - K(z,\xi)$ and this implies

$$(f(z),K_0(z,\xi))_2 = (f(z),K(z,\xi))_2 + \frac{1}{\pi}\int_{-1}^{1} dx\lambda(x)f(x)\overline{K(x,\xi)} = (f(z),K(z,\xi))_1,$$

where the integration variable is understood to be z inside any bracket. However, the left-hand side of the above equation is equal to $f(\xi)$ since K_0 is the self-reproducing kernel of H^2, and we discover (3.11), i.e.,

$$f(\xi) = (f(z),K(z,\xi))_1$$

for all H^2 functions $f(z)$. Especially, restricting ourselves to $f \in H^2(\lambda)$, this relation shows that $K(z,\xi)$ must be a self-reproducing kernel of the space $H^2(\lambda)$. Since the self-reproducing kernel is unique, this proves the desired result. Note that (3.11) is valid now not only for $H^2(\lambda)$ functions but also for any $f(z) \in H^2$.

Now, we have to prove our lemma. First, we shall show that $p(z)$ belongs to H^2. Since $R(x)$ is of a class $L^2(1,-1)$, it is also summable by Schwarz inequality. Then, $p(z)$ must be a holomorphic function of z in a cut z plane with cuts on real axis at $\infty > z \ge 1$ and at $-1 \ge z > -\infty$. Especially, it is regular in the unit open disk $|z| < 1$, and we can expand it into a power series.

$$p(z) = \sum_{n=0}^{\infty} a_n z^n, \qquad |z| < 1,$$

$$a_n = \frac{1}{\pi}\int_{-1}^{1} dx x^n R(x).$$

Noting

$$|a_n| \le b_n = \int_0^1 dx x^n S(x),$$

$$S(x) = \frac{1}{\pi}(|R(x)| + |R(-x)|)$$

and using an inequality[7] of Hardy–Littlewood–Polya, we estimate

$$\sum_{n=0}^{\infty} |a_n|^2 \le \sum_{n=0}^{\infty} (b_n)^2 \le \pi\int_0^1 dx (S(x))^2 \le \frac{2}{\pi}\int_{-1}^{1} dx\,|R(x)|^2.$$

This proves that $p(z)$ belongs to H^2. However, it gives an upper bound for $\|p\|_2^2$, which is twice larger than (3.16). We shall improve this bound shortly. To show (3.18), we compute

$$(f,p)_2 = \lim_{r\to 1-0}\frac{1}{2\pi}\int_0^{2\pi} d\theta f(re^{i\theta})\overline{p(re^{i\theta})} = \lim_{r\to 1-0}\frac{1}{2\pi^2}\int_0^{2\pi} d\theta\int_{-1}^{1} dx\frac{1}{1-rxe^{-i\theta}} f(re^{i\theta})\overline{R(x)}.$$

For $r < 1$, the integrand is a summable function of x and θ by our assumptions for $f \in H^2$ and $R(x) \in L^2(1,-1)$. Hence, by the Fubini theorem,[6] we can interchange the order of the integrals to obtain

$$(f,p)_2 = \lim_{r\to 1-0}\frac{1}{\pi}\int_{-1}^{1} dx f_r(x)\overline{R(x)}, \tag{3.19}$$

where we have for simplicity set

$$f_r(z) = f(r^2 z).$$

Using the Schwarz inequality and the Hilbert–Fejér–Riesz inequality[1,5]

$$\frac{1}{\pi}\int_{-1}^{1} dx\,|g(x)|^2 \le \|g\|_2^2, \qquad g \in H^2, \tag{3.20}$$

we now estimate

$$\left|\int_{-1}^{1} dx (f(x) - f_r(x))\overline{R(x)}\right|^2 \le \pi\,\|f - f_r\|_2^2\int_{-1}^{1} dx\,|R(x)|^2.$$

Since $f(z)$ belongs to H^2, we know[1]

$$\lim_{r\to 1-0}\|f - f_r\|_2 = 0$$

and this together with (3.19) gives (3.18). Especially, setting $f(z) = p(z)$, we estimate

$$\|p\|_2^2 = \frac{1}{\pi}\int_{-1}^{1} dx p(x)\overline{R(x)} \le \left(\frac{1}{\pi}\int_{-1}^{1} dx\,|p(x)|^2\right)^{1/2}\left(\frac{1}{\pi}\int_{-1}^{1} dx\,|R(x)|^2\right)^{1/2} \le \|p\|_2\left(\frac{1}{\pi}\int_{-1}^{1} dx\,|R(x)|^2\right)^{1/2}$$

which leads to the desired upper bound in (3.16). The equality in (3.16) also follows from (3.18), if we set again $f(z) = p(z)$ and if we use the Tonelli–Fubini theorem[6] for interchange of order of double integrals. Finally, the inequality (3.17) can be obtained by

$$\frac{1}{\pi}\int_{-1}^{1} dx\,|p(x)|^2 \le \|p\|_2^2 \le \frac{1}{\pi}\int_{-1}^{1} dx\,|R(x)|^2 \tag{3.17'}$$

because of (3.20) and (3.16).

Now let us return to the discussion of (3.9). If we restrict ourselves to real values of z in the interval $-1 \le z \le 1$, then (3.9) represents an integral equation of Fredholm type. However, its kernel is singular and the usual method is not applicable to solve it. If we iterate it once, then it will, however, become a less singular equation

$$K(z,\xi) = K_0(z,\xi) - K_1(z,\xi) + \frac{1}{\pi}\int_{-1}^{1} dx K_1(z,x)\lambda(x)K(x,\xi), \tag{3.21}$$

$$K_1(z,\xi) = \frac{1}{\pi}\int_{-1}^{1} dx\,[\lambda(x)/(1-xz)(1-x\bar{\xi})] \tag{3.22}$$

which can be hopefully solved numerically for practical applications

In what follows, we shall only investigate an iterative solution of (3.9). To this end, let us define the nth iterative kernel $K_n(z,\xi)$ recursively by

$$K_{n+1}(z,\xi) = \frac{1}{\pi}\int_{-1}^{1} dx\frac{\lambda(x)}{1-xz}K_n(x,\xi),$$
$$K_0(z,\xi) = 1/(1-z\bar{\xi}). \tag{3.23}$$

Then, $K_n (n \ge 1)$ will be expressed as a $n$ple integral.

$$K_n(z,\xi)=\frac{1}{\pi^n}\int_{-1}^{1}dx_1\int_{-1}^{1}dx_2\cdots\int_{-1}^{1}dx_n \times\frac{\lambda(x_1)\lambda(x_2)\cdots\lambda(x_{n-1})\lambda(x_n)}{(1-zx_1)(1-x_1x_2)\cdots(1-x_{n-1}x_n)(1-x_n\bar{\xi})}. \tag{3.24}$$

Now we want to interchange orders of multiple integrals in (3.24) as follows. First, if z and ξ are real with $|z|<1$ and $|\xi|<1$, then the integrand of (3.24) is nonnegative.

Therefore, the Tonelli theorem[6] justifies the interchange. For complex values of z and ξ, we notice an inequality

$$|1/(1-xz)|\le 1/[1-x\mathrm{Re}(z)],\qquad |\mathrm{Re}\,z|\le 1, \tag{3.25}$$

so that (3.24) gives us

$$|K_n(z,\xi)|\le K_n(\mathrm{Re}\,z,\mathrm{Re}\,\xi),\qquad n\ge 1. \tag{3.26}$$

Hence, we can justify again interchanges of order of integrals by Fubini–Tonelli's theorem, if the integral is finite.

Immediate consequences of this fact are

$$\overline{K_n(z,\xi)}=K_n(\xi,z) \tag{3.27}$$

and

$$K_{n+m+1}(z,\xi)=\frac{1}{\pi}\int_{-1}^{1}dx K_n(z,x)\lambda(x)K_m(x,\xi). \tag{3.28}$$

If iterative solution of (3.9) exists, then it must be given by

$$K(z,\xi)=\sum_{n=0}^{\infty}(-1)^nK_n(z,\xi) \tag{3.29}$$

In what follows, we shall investigate conditions under which (3.29) represents the correct solution of (3.9). To this end, we assume that $\lambda(x)$ is bounded by

$$0\le\lambda(x)\le M\qquad(-1\le x\le 1). \tag{3.30}$$

Then, first we shall show that $K_n(z,\xi)$ for a fixed ξ belongs to H^2 with a bound

$$\frac{1}{\pi}\int_{-1}^{1}dx\,|K_n(x,\xi)|^2\le\|K_n(z,\xi)\|_2^2\le M^{2n}K_0(\xi,\xi),$$
$$\|K_{n+1}(z,\xi)\|_2\le M\|K_n(z,\xi)\|_2 \tag{3.31}$$

We shall prove this by induction. For $n=0$, this follows from (3.16) and (3.20). Suppose that $K_n(z,\xi)$ is a H^2 function with the bound (3.31). Then since $\lambda(x)$ is bounded by our assumption, $\lambda(x)K_n(x,\xi)$ belongs to $L^2(1,-1)$ and hence

$$K_{n+1}(z,\xi)=\frac{1}{\pi}\int_{-1}^{1}dx\frac{1}{1-xz}\lambda(x)K_n(x,\xi)$$

belongs to H^2 by our lemma. Moreover, (3.16) and (3.20) again give the bound (3.31) for $n\to n+1$. Therefore, by mathematical induction, we have proved that $K_n(z,\xi)$ belongs to H^2 with the bound (3.31).

Applying (3.18) for the case $f(z)=K_n(z,\xi)$ and $p(z)=K_m(z,\xi)$, and noting (3.28), we find then

$$K_{n+m}(\eta,\xi)=(K_n(z,\xi),K_m(z,\eta))_2 \tag{3.32}$$

where the integration variable is z. From this equation and (3.31), we estimate now

$$|K_n(\eta,\xi)|\le\|K_m(z,\xi)\|_2\|K_{n-m}(z,\eta)\|_2 \le M^n[K_0(\xi,\xi)K_0(\eta,\eta)]^{1/2}. \tag{3.33}$$

Therefore, if we have

$$M<1, \tag{3.34}$$

then the series (3.29) converges absolutely and represents a holomorphic function of z in $|z|<1$. Similarly, it is easy to prove that it satisfies the integral equation (3.9), because by the Lebesque's dominant convergent theorem[6] we can interchange the order of the summation and integration. To prove that our $K(z,\xi)$ belongs to H^2, we apply (3.32) again to find

$$\begin{aligned}(K(z,\xi),K(z,\eta))_2&=\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}(-1)^{n+m}(K_n(z,\xi),K_m(z,\eta))_2\\&=\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}(-1)^{n+m}K_{n+m}(\eta,\xi)\\&=\sum_{n=0}^{\infty}(-1)^n(n+1)K_n(\eta,\xi)\end{aligned} \tag{3.35}$$

which absolutely converges again because of (3.33) and (3.34). Thus, $K(z,\xi)$ belongs to H^2. Since $\lambda(x)$ is bounded, the space $H^2(\lambda)$ is the same set as H^2. Then, these facts and theorem VIII of this section guarantee that the iterative solution is indeed the correct solution under conditions (3.30) and (3.34).

We shall simply mention in this connection that if $\lambda(x)$ satisfies

$$B=\frac{1}{\pi}\int_{-1}^{1}dx\frac{1}{1-x^2}\lambda(x)<1$$

instead of (3.34), the iterative solution is still the correct one since we can easily derive a bound

$$|K_n(z,\xi)|\le B^n[K_0(z,z)K_0(\xi,\xi)]^{1/2}$$

for this case. However, this condition presupposes $\lambda(x)=0$ at $x=\pm1$, which is not satisfied in general for our applications to high energy physics.

Actually, we can prove a more general statement that the iterative solution must be the right one even without assuming the boundedness of $\lambda(x)$, as long as $K(x,y)$ in the real interval $-1\le x,y\le 1$ remains nonnegative and we have $\lim_{n\to\infty}K_n(z,z)=0$. To show this, we rewrite (3.13) as

$$K_0(z,\xi)=K(z,\xi)+\frac{1}{\pi}\int_{-1}^{1}dx\,K(z,x)\lambda(x)K_0(x,\xi) \tag{3.36}$$

replacing η by ξ, and ξ by z. Iterating (3.9) once by this equation, we find

$$K(z,\xi)=K_0(z,\xi)-K_1(z,\xi) +\frac{1}{\pi^2}\int_{-1}^{1}dx\int_{-1}^{1}dy\frac{\lambda(x)\lambda(y)}{(1-xz)(1-y\bar{\xi})}K(x,y). \tag{3.37}$$

Further, we iterate (3.37) $2n$-times by repeated uses of (3.9) and (3.36) to obtain

$$K(z,\xi)=\sum_{m=0}^{2n+1}(-1)^m K_m(z,\xi)$$
$$+\frac{1}{\pi^2}\int_{-1}^{1}dx\int_{-1}^{1}dy\,K_n(z,x)\lambda(x)K(x,y)\lambda(y)K_n(y,\xi) \tag{3.38}$$

if we can interchange the order of multiple integrals we encounter. However, this can be justified as we have done for $K_n(z,\xi)$. For real values of z and ξ, the integrand of the right-hand side of (3.38) is nonnegative if $K(x,y)$ is nonnegative. Then, the Tonelli's theorem justifies the interchange. For complex values of z and ξ, we use the inequality (3.25) and repeat the same argument. Obviously, this reasoning is also applicable even if $K(x,y)$ becomes negative but with a finite lower bound, provided that a function $R_n(x)=\lambda(x)K_n(x,\xi)$ is summable in $-1\le x\le 1$.

Theorems II and VI demand that the correct solution must satisfy

$$K_0\ge K\ge 0. \tag{3.39}$$

Then, the structure of Eq. (3.38) implies that this leads to

$$\sum_{m=0}^{2n+2}(-1)^m K_m\ge K\ge \sum_{m=0}^{2n+1}(-1)^m K_m. \tag{3.40}$$

We note now that a relation $A\ge B$ gives

$$A(z,z)\ge B(z,z),$$

$$[A(z,z)-B(z,z)][A(\xi,\xi)-B(\xi,\xi)]\ge |A(z,\xi)-B(z,\xi)|^2$$

as in (2.7). Therefore, letting $n\to\infty$ in (3.40), we discover that the iterative solution

$$\sum_{n=0}^{\infty}(-1)^n K_n(z,\xi)$$

will automatically converge to the correct self-reproducing kernel $K(z,\xi)$, if we have $\lim_{n\to\infty}K_n(z,z)=0$. We notice that if $\lambda(x)$ is bounded, then the function $R_n(x)=\lambda(x)\,K_n(x,\xi)$ is summable as we have proved in (3.31). Hence without assuming $M<1$, the iterative solution always represents the correct solution if we have $\lim_{n\to\infty}K_n(z,z)=0$ and if $K(x,y)$ is bounded below in the real interval $-1\le x,y\le 1$.

Finally, we shall prove the following inequality

$$MK_n\ge K_{n+1}\ge 0. \tag{3.41}$$

Let N be a fixed positive integer. For arbitrary complex numbers $c_1, c_2, \ldots, c_N$, we set

$$\varphi_n(z)=\sum_{\nu=1}^{N}\overline{C}_\nu K_n(z,z_\nu),$$

where $z_1, z_2, \ldots, z_N$ are N arbitrary points in the unit disk. Then, from (3.27), (3.28), and (3.32), we find (by setting $n=m$)

$$\sum_{\mu=1}^{N}\sum_{\nu=1}^{N}C_\mu\,K_{2n+1}(z_\mu,z_\nu)\overline{C}_\nu=\frac{1}{\pi}\int_{-1}^{1}dx\lambda(x)\,|\varphi_n(x)|^2,$$
$$\sum_{\mu=1}^{N}\sum_{\nu=1}^{N}C_\mu K_{2n}(z_\mu,z_\nu)\overline{C}_\nu=\|\varphi_n\|_2^2, \tag{3.42}$$

which immediately gives the lower bound

$$K_n\ge 0$$

in (3.41). To prove the upper bound in (3.41), we have to discuss separately two cases, depending upon whether n is even or odd. Here, for simplicity, we prove it only for the case $N=1$ as follows:

$$\begin{aligned}
K_{2n+1}(\xi,\xi)&=\frac{1}{\pi}\int_{-1}^{1}dx\lambda(x)\,|K_n(x,\xi)|^2\\
&\le\frac{M}{\pi}\int_{-1}^{1}dx\,|K_n(x,\xi)|^2\\
&\le M\|K_n(z,\xi)\|_2^2=MK_{2n}(\xi,\xi),\\
K_{2n}(\xi,\xi)&=\|K_n(z,\xi)\|_2^2\\
&\le\frac{1}{\pi}\int_{-1}^{1}dx\,|\lambda(x)K_{n-1}(x,\xi)|^2\\
&\le\frac{M}{\pi}\int_{-1}^{1}dx\lambda(x)\,|K_{n-1}(x,\xi)|^2=MK_{2n-1}(\xi,\xi)
\end{aligned} \tag{3.43}$$

where we repeatedly used (3.16), (3.17), and (3.20). From (3.41), we can again prove that the iterative solution (3.29) converges absolutely if we have $M<1$.

We remark that the inequality (2.12) gives us

$$K_0\ge K\ge\frac{1}{1+M}K_0 \tag{3.44}$$

which must be valid independent of validity of the iterative solution.

4. APPLICATIONS TO DISPERSION INEQUALITIES

In high energy physics, we often encounter the following problems. Let $F(t)$ be a holomorphic function of a complex variable t in a cut plane with cuts on the right-hand real axis at $\infty\ge t\ge t_0$, and on the left-hand axis at $-t_1\ge t\ge-\infty$, (see Fig. 1),

where t_0 and t_1 are real constants satisfying

$$t_0>0>-t_1. \tag{4.1}$$

For some problems, we have no left-hand cut. In that case, we understand the situation by letting $t_1\to\infty$.

We assume that boundary function $F(t\pm i0)$ on the cut always exists, since it represents physically measurable quantities. Generally, the only other informations available about $F(t)$ are the following:

(i) $F(t)$ is real in the sense that it satisfies

$$F(t)=\overline{F}(\bar{t}). \tag{4.2}$$

Then, we have

$$F(t-i0)=\overline{F}(t+i0). \tag{4.3}$$

This reality condition is usually a consequence of the time-reversal invariance of the theory.

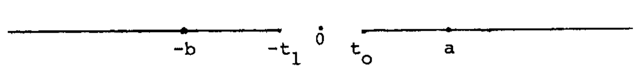


FIG. 1 Cut structures of holomorphic functions $F(t)$ and $F_1(t)$.

(ii) Let $k(t)$ be a given positive (almost everywhere on the cut) measurable function defined on the cut. Then, an upper bound A for the following integral on the cut is given:

$$A \geq \frac{1}{\pi}\left(\int_{t_0}^{\infty} + \int_{-\infty}^{-t_1}\right) dt k(t) |F(t + i0)|^2. \tag{4.4}$$

(iii) We may know a value or values of $F(t)$ at a few points $t = b_j$ $(j = 1, 2, \ldots, n)$ in the real interval $-t_1 < b_j < t_0$ to be

$$F(b_j) = C_j \quad (j = 1, 2, \ldots, n). \tag{4.5}$$

If this information is not available, then we set $n = 0$.

(iv) For some problems, the phase $\delta(t)$ of $F(t + i0)$ is known in a finite interval near the cut at $a \geq t \geq t_0$, and $-b \leq t \leq -t_1$ (see Fig. 1). Hereafter, we define the phase $\delta(t)$ so that we have

$$F(t + i0) = \pm |F(t + i0)| \exp(i\delta(t)), \tag{4.6}$$

where choice of the $\pm$ sign in (4.6) is determined by requiring

$$\delta(t_0) = \delta(-t_1) = 0, \tag{4.7}$$

because of the following physical reason. Generally, $\delta(t)$ represents scattering phase-shifts for reactions such as $\pi - \pi$, $\pi - K$ or $\pi - N$ scatterings so that the condition (4.7) is automatically satisfied because of threshold theorem for scattering phase-shifts.

If we do not want to use this information on the phase, or if the information is not available, then we simply set $a = t_0$ and $b = t_1$.

(v) Some kind of polynomial boundedness condition for $F(t)$ at infinity. A more precise condition will be specified in the end of this section.

Suppose that these are only informations available on $F(t)$. We want to know whether it is possible to give some bounds for $F(t)$ and its derivatives out of these constraints. As we shall see shortly, we can reduce our problem to a form stated in Sec. 1, and we can solve the problem accordingly. Before going into details, we shall mention various problems of this kind. First, let us enumerate those without taking into account the phase condition (iv), since this case corresponds to $\lambda(x) \equiv 0$ and it is easily solvable. We find this kind of problem for

(1) Geshkenbein–Ioffe type problem[8,9] of finding an upper bound of pion-nucleon coupling constant under some technical conditions which we shall not go into detail. Historically, this is the first problem of the dispersion inequality, which has been solved by method due to Meiman.[10]

(2) Finding of an upper bound for the renormalization constant Z_2 of the nucleon[11,12] and of the deuteron.[13]

(3) Determination of an upper bound for hadronic contributions[14] to the anomalous magnetic moment of the muon and to the Lamb–shift.[11]

(4) Existence of a lower bound[11,15] for Schwinger term in the algebra of current.

(5) Upper and lower bounds for scalar K_{l3} decay parameters.[16,17,18]

(6) Upper and lower bound[19] for scaling-dimension of chiral–scale–breaking Hamiltonian theory.

(7) Lower bound[20,21] for low-energy $\pi^0 - \pi^0$ scattering amplitudes.

(8) Existence[22] of absolute upper bound for coupling constant of cubic scalar interaction.

(9) Analysis[23] of chew–Low integral equation

There may be other applications the present author is unaware of. However, this list may be enough to illustrate the usefulness of our method. As we remarked earlier most of these problems have been analyzed without assuming the phase condition (iv). However, for some problems, the inclusion of the additional knowledge (iv) greatly emphasize the usefulness of the bound. Such is indeed the case with analysis of the K_{l3} problem[16] and also with the renormalization constant Z_2 of the nucleon.[24] Also, the same should be applicable to the problem[25] of the types (3) and (4). However, in usual treatment of such problems, the presence of the $\lambda(x)$ term in (1.2) and (1.3) is either neglected or suitably approximated (possibly excepting the analysis of the Ref. 25). As the result, the bounds so obtained are not the best to be achieved. Of course, as we shall see shortly, the expression for $\lambda(x)$ is quite complicated and besides we have to solve the singular integral equation (1.8) to this end. Hopefully, these complications will be overcome and better bounds will be calculated in the future.

Now, returning to the original problem, let us set

$$G(t) = \exp\left[\frac{1}{\pi}(a - t)^{1/2}(b + t)^{1/2}\left(\int_{t_0}^{a} + \int_{-b}^{-t_1}\right) dt' \frac{\rho_0(t')}{t' - t}\right],$$
$$\rho_0(t') = |a - t'|^{-1/2}|b + t'|^{-1/2}\delta(t'). \tag{4.8}$$

If we do not want to use the information (iv), we have only to set $a = t_0$ and $t_1 = b$ with the consequence $G(t) = 1$.

We shall assume on the physical ground that $\delta(t)$ is continuous with Lipschitz condition

$$|\delta(x) - \delta(y)| \leq C|x - y|^n \tag{4.9}$$

for some positive constants c and n in the interval under consideration. We choose the cuts of $(a - t)^{1/2}$ and $(b + t)^{1/2}$ to lie on real axis at $\infty > t \geq a$ and $-b \geq t > -\infty$, respectively. We also choose their branches so that both functions are real and positive in the cut–free interval $a > t > -b$. Then, the function $G(t)$ is real and holomorphic in the same cut plane as of $F(t)$. Because of the conditions (4.9), $G(t \pm i0)$ is continuous on the interval $a \geq t \geq t_0$ and $-b \leq t \leq -t_1$. Further, we have

$$\operatorname{Arg} G(t \pm i0) = \pm\delta(t) \tag{4.10}$$

in the same interval. Moreover, we have

$$|G(t \pm i0)| = 1 \quad \text{for } t > a \text{ or } t < -b. \tag{4.11}$$

If we set

$$F_1(t) = F(t)/G(t), \tag{4.12}$$

then we easily find

$$F_1(t - i0) = F_1(t + i0)$$

in the intervals $a \geq t \geq t_0$ and $-b \leq t \leq -t_1$. Therefore, we conclude that $F_1(t)$ has not cut at all in that interval. Since $G(t)$ has no zero point in the entire cut plane as well as on the interval including the end point $t = t_0$ and $t = -t_1$ because of the conditions (4.7) and (4.9), we find that $F_1(t)$ satisfies the following modified conditions:

(0′) $F_1(t)$ is a holomorphic function of t in a new cut plane with cuts only at $\infty > t \geq a$ and $-b \geq t \geq -\infty$.

(i′) $F_1(t)$ is real, i.e.,

$$F_1(t) = \overline{F}_1(\bar{t}). \tag{4.13}$$

(ii′)

$$A \geq I_1 + I_2 \tag{4.14}$$

where I_1 and I_2 are given by

$$\begin{aligned} I_1 &= \frac{1}{\pi}\left(\int_a^\infty + \int_{-\infty}^{-b}\right) dt k(t) |F_1(t + i0)|^2, \\ I_2 &= \frac{1}{\pi}\left(\int_{t_0}^{a} + \int_{-b}^{-t_1}\right) dt k_1(t) |F_1(t + i0)|^2, \end{aligned} \tag{4.15}$$

$$k_1(t) = k(t) |G(t + i0)|^2. \tag{4.16}$$

Note that we have used (4.11) for the integral I_1.

(iii′) We have

$$F_1(b_j) = \tilde{C}_j = C_j/G(b_j) \quad (j = 1, 2, \ldots, n). \tag{4.17}$$

(v′) Same kind of polynomial boundedness condition for $F_1(t)$, which will be discussed shortly.

In this modified form, we need not consider an analog of the condition (iv) any longer.

Next, we perform the conformal mapping[9]

$$[(a - t)/(b + t)]^{1/2} = (a/b)^{1/2}[(1 - z)/(1 + z)] \tag{4.18}$$

which maps the cut plane of $F_1(t)$ into the unit disk $|z| < 1$. This maps also both upper left and right cuts of the t plane onto the upper-semi circle $0 \leq \arg z \leq \pi$, while both lower cuts are transformed into the lower semi circle $\pi \leq \arg z \leq 2\pi$. Moreover, three points $t = a$, $t = 0$, and $t = -b$ are changed into $z = 1$, $z = 0$, and $z = -1$, respectively. Setting now

$$F_1(t) = f_1(z) \tag{4.19}$$

then $f_1(z)$ is holomorphic in $|z| < 1$. Further, we can rewrite I_1 as[9]

$$\begin{aligned} I_1 &= \frac{1}{2\pi}\int_0^{2\pi} d\theta W_1(\theta) |f_1(e^{i\theta})|^2, \\ W_1(\theta) &= \frac{2ab(a + b)|\sin\theta|}{[(b - a) + (b + a)\cos\theta]^2} k[t(\theta)], \\ t(\theta) &= \frac{2ab}{(b - a) + (b + a)\cos\theta}, \end{aligned} \tag{4.20}$$

while I_2 is transformed into

$$I_2 = \frac{1}{\pi}\left(\int_\alpha^1 + \int_{-1}^{-\beta}\right) dx \gamma_1(x) |f_1(x)|^2. \tag{4.21}$$

Here, α, β, and $\gamma_1(x)$ are defined by

$$\begin{aligned} \alpha &= \frac{(ab + at_0)^{1/2} - (ab - bt_0)^{1/2}}{(ab + at_0)^{1/2} + (ab - bt_0)^{1/2}}, \\ \beta &= \frac{(ab + bt_1)^{1/2} - (ab - at_1)^{1/2}}{(ab + bt_1)^{1/2} + (ab - at_1)^{1/2}}, \\ \gamma_1(x) &= \frac{4ab(a + b)(1 - x^2)}{[(a + b)(1 + x^2) + 2(b - a)x]^2} k_1[t(x)], \\ t(x) &= \frac{4abx}{(a + b)(1 + x^2) + 2(b - a)x}, \end{aligned} \tag{4.22}$$

Note that α and β satisfy the condition

$$1 \geq \alpha > 0 > -\beta \geq -1. \tag{4.23}$$

In order to rewrite I_1 in the form used in the Sec. 1, we further introduce

$$\phi_1(z) = \exp\left(\frac{1}{4\pi}\int_0^{2\pi} d\theta \frac{e^{i\theta} + z}{e^{i\theta} - z} \log W_1(\theta)\right). \tag{4.24}$$

Assuming that $\log w_1(\theta)$ is summable in $(2\pi, 0)$, then $\phi_1(z)$ is a holomorphic function[1,26] of z in $|z| < 1$. Also, from the Poisson's integral formula we have[1,26]

$$|\phi_1(e^{i\theta})|^2 = W_1(\theta) \tag{4.25}$$

almost everywhere on the unit circle. Therefore, finally setting

$$f(z) = \phi_1(z) f_1(z), \tag{4.26}$$

we can rewrite

$$\begin{aligned} I_1 &= \frac{1}{2\pi}\int_0^{2\pi} d\theta |f(e^{i\theta})|^2, \\ I_2 &= \frac{1}{\pi}\int_{-1}^{1} dx \lambda(x) |f(x)|^2, \\ \lambda(x) &= \gamma_1(x) |\phi_1(x)|^{-2}, \end{aligned} \tag{4.27}$$

where we have set

$$\lambda(x) = \gamma_1(x) = 0, \quad \alpha > x > -\beta. \tag{4.28}$$

In this way, we have finally reduced the problem to a form stated in the Sec. 1. However, the mere existence of the integral I_1 does not necessarily guarantee that $f(z)$ belongs automatically to H^2. To insure it, condition (v) is needed. To see it clearly, we rewrite $\phi_1(z)$ in terms of the old variable t as

$$\begin{aligned} \phi_1(z) &= (a + b)^{-1/2}\left(\frac{(a - t)(b + t)}{ab}\right)^{1/4} [(ab + at)^{1/2} + (ab - bt)^{1/2}]\phi(t), \\ \phi(t) &= \exp\left[\frac{1}{2\pi}(a - t)^{1/2}(b + t)^{1/2}\left(\int_a^\infty - \int_{-\infty}^{-b}\right) dt' \frac{\rho(t')}{t' - t}\right], \\ \rho(t') &= |t' - a|^{-1/2} |t' + b|^{-1/2} \log k(t'). \end{aligned} \tag{4.29}$$

Alternatively, we can write $\phi(t)$ also as[9]

$$\begin{aligned} \phi(t) &= \exp\left(\frac{1}{4\pi}\int_0^{2\pi} d\theta \frac{e^{i\theta} + z}{e^{i\theta} - z} \log W(\theta)\right), \\ W(\theta) &= k[t(\theta)], \\ t(\theta) &= \{2ab/[(b - a) + (b + a)\cos\theta]\}. \end{aligned} \tag{4.30}$$

Note that $\lambda(x)$ is finally rewritten as

$$\lambda(x) = |[G(t + i0)/\phi(t)]|^2 k(t),$$
$$t = t(x) = \{4abx/[(a + b)(1 + x^2) + 2(b - a)x]\}. \tag{4.31}$$

In general, $\lambda(x)$ is bounded in the entire interval $-1 \le x \le 1$.

Now, let us investigate asymptotic behaviors of $G(t)$ and $\phi(t)$. First, $G(t)$ becomes a constant at infinity so that it causes no problem at all. Second, in general, $k(t)$ is polynomially bounded at infinity by

$$|k(t)| \le C|t|^n, \qquad t \to \infty,$$

for some constants c and n along the real axis. Then, $\phi(t)$ as a complex function of t is also polynomially bounded at infinity now in all directions in the complex t plane.

Noting that on the cuts at $\infty > t > a$ and $-b \ge t > -\infty$, we have

$$k(t) = |\phi(t \pm i0)|^2, \tag{4.32}$$

so that we can rewrite the integral I_1 of (4.15) as

$$I_1 = \frac{1}{\pi}\left(\int_a^\infty + \int_{-\infty}^{-b}\right)dt|\phi(t + i0)F_1(t + i0)|^2$$
$$= \frac{1}{2\pi}\oint_{\text{cut}} dt|\phi(t)F_1(t)|^2, \tag{4.33}$$

where we used the reality conditions for $\phi(t)$ and $F_1(t)$. If $F(t)$ [and hence $F_1(t)$] is polynomially bounded at infinity, and if I_1 is finite, then the Phragmen–Lindelöf theorem assures that the product $\phi(t)F_1(t)$ will vanish at infinity in the complex t plane. In that case, we can add a large circular term at infinity to I_1 so as to make it a closed contour integral in the complex t plane. Then, we change the variable t into z. In this way, we make sure that $f(z)$ will belong to H^2, after some more lengthy arguments. Regardless, the precise condition for the ansatz (v) should be exactly that $f(z)$ be an element of the space H^2.

Let us briefly remark that if we do not use the information (iv), then we have $\lambda(x) = 0$ by setting $t_0 = a$ and $t_1 = b$. In that case, we can solve the problem explicitly. Defining an inner product of two real holomorphic functions $f(t)$ and $g(t)$ with the given cut structure directly by

$$(f, g) = \frac{1}{\pi}\left(\int_a^\infty + \int_{-\infty}^{-b}\right) dtk(t)f(t + i0)\,\overline{g(t + i0)}, \tag{4.34}$$

then the kernel function $K(t, \eta)$ of this space is calculated to be

$$\frac{1}{K(t, \eta)} = \left(\frac{2}{a + b}\right)(t + b)(\bar{\eta} + b)[\Delta(t) + \Delta(\bar{\eta})][\Delta(t)\Delta(\bar{\eta})]^{1/2} \times \phi(t)\phi(\bar{\eta}), \tag{4.35}$$

where we have set for simplicity

$$\Delta(t) = [(a - t)/(b + t)]^{1/2} \tag{4.36}$$

and $\phi(t)$ is defined by (4.29). For many problems of the high energy physics, $k(t)$ has a simple form

$$k(t) = C\prod_{j=1}^n |t - \gamma_j|, \quad -b < \gamma_j < a. \tag{4.37}$$

In that case, we compute[9]

$$\phi(t) = C^{1/2}(a + b)^{-n/2}\prod_{j=1}^n [(b + \gamma_j)^{1/2}(a - t)^{1/2} + (a - \gamma_j)^{1/2}(b + t)^{1/2}]. \tag{4.38}$$

When the left-hand cut does not exist, then we simply let $b \to \infty$ in Eqs. (4.35) and (4.29) to find

$$\frac{1}{K(t, \eta)} = 2[(a - t)^{1/2} + (a - \bar{\eta})^{1/2}][(a - t)(a - \bar{\eta})]^{1/4}\phi(t)\phi(\bar{\eta}), \tag{4.39}$$

$$\phi(t) = \exp\left(\frac{1}{2\pi}(a - t)^{1/2}\int_a^\infty dt'\frac{1}{(t' - t)|t' - a|^{1/2}}\log k(t')\right).$$

We can easily reproduce various dispersion inequalities[9,23] from these formulas of the self-reproducing kernel.

We may mention that there are other Hilbert spaces[27] with self-reproducing kernels which are of some interest for applications to physical problems. Let $f(z)$ be a holomorphic function of z in a simply connected domain D with a rectifiable curve Γ. Then, we may introduce two norms[27] by

$$\|f\|_B^2 = \frac{1}{\pi}\int_D d^2z|f(z)|^2, \qquad d^2z = d\,(\text{Re}z)d\,(\text{Im}z),$$

$$\|f\|_b^2 = \frac{1}{2\pi}\int_\Gamma |dz||f(z)|^2, \qquad |dz| = [dzd\bar{z}]^{1/2}.$$

The space H^2 corresponds to the norm $\|f\|_b^2$ for a special case with $D = \{z \,|\, |z| < 1\}$. The self-reproducing kernels $K_B(z, \xi)$ and $K_b(z, \xi)$ of two spaces can be shown to satisfy an interesting relation

$$K_B(z, \xi) = [K_b(z, \xi)]^2$$

although we will not go into its proof. Bergman and Schiffer[28] also give some applications of self-reproducing kernels of harmonic functions for problems involving classical physics.

Finally, we briefly remark that there is another class of dispersion inequality in high energy physics. It is essentially reducible to a study of the minimum interpolating problem[2] in the Banach space H^∞. Since it is not a Hilbert space, we have to use a different approach to solve it. Here, we simply quote some relevant references[9,21,29] for its applications, although its relation to the space H^∞ may not be apparent in these papers.

*Work supported in part by the U. S. Atomic Energy Commission.

[1]H. P. Duren, *Theory of Hp Spaces* (Academic Press, New York, 1970). The present author would like to express his gratitude to Professor H. Nussenzveig for calling his attention to this book.

[2]See Ref. 1 Chap. 8.

[3]N. Aronszajn, Trans Amer. Math. Soc. **68**, 337 (1950).

[4]For the special case $\lambda(x) = 0$, this method has been already used by I-Fu. Shih and S. Okubo, Phys. Rev. D **4**, 3519 (1971).

[5]G. H. Hardy, J. E. Littlewood, and G. Polya, *Inequalities* (Cambridge U. P., Cambridge 1967), p. 235.

[6]E. J. McShane, *Integrations*, Princeton U. Pr., Princeton 1947).

[7]Ref. 5, p. 237, theorem 324.

[8]B. V. Geshkenbein and B. L. Ioffe, Phys. Rev. Lett. **11**, 55 (1963). M. Ida, Phys. Rev. **136**, B1767 (1964).

[9]S. Okubo, Prog. Theor. Phys. (Kyoto) **48**, 1986 (1972).

[10]N. N. Meiman, Zh. Eksp. Teor. Fiz. **44**, 1228 (1963) [Soviet Phys. JETP **17**, 830 (1963)].

[11]S. D. Drell, A. C. Finn, and A. C. Hearn, Phys. Rev. **B136**, 1439 (1964).
[12]Y. S. Jin and S. W. MacDowell, Phys. Rev. **B137**, 688 (1965).
[13]T. Celik, to be published.
[14]D. R. Palmer, Phys. Rev. D **4**, 1558 (1971). G. Nenciu and I. Raszillier, Nuovo Cimento **11A**, 319 (1972).
[15]S. Okubo, Phys. Rev. **D3**, 2807 (1971).
[16]C. Bourrely, Nucl. Phys. **B43**, 434 (1972) and **B53**, 289 (1973).
[17]S. Okubo and I-Fu. Shih, Phys. Rev. D **4**, 3519 (1971). M. Micu, Nucl. Phys. **B44**, 531 (1972).
[18]G. J. Aubrecht, D. M. Scott, K. Tanaka, and R. Torgeson, Phys. Rev. D **4**, 1423 (1971). K. Tanaka and R. Torgeson: Phys. Rev. D 5, 1164 (1972).
[19]D. N. Levin, S. Okubo, and D. R. Palmer, Phys. Rev. D **4**, 1847 (1971). D. N. Levin and D. R. Palmer: Phys. Rev. D **4**, 3716 (1971).
[20]A. Martin in *High Energy Physics and Elementary Particles* (International Atomic Energy Agency, Vienna, 1965), p. 155. L. Lukaszuk and A. Martin, Nuovo Cimento **52A**, 122 (1967) especially, Appendix D. P. Dită, Nuovo Cimento **10A**, 753 (1972).
[21]S. Okubo in *Fundamental Interactions in Physics and Astrophysics*, edited by G. Iverson et al. (Plenum, New York, 1973), p. 206.
[22]V. F. Müller, Nuovo Cimento **42**, 185 (1966).
[23]S. Okubo, Nuovo Cimento **16A**, 241 (1973).
[24]V. Baluni and D. J. Broadhurst, Phys., Rev. D **7**, 3738 (1973).
[25]G. Auberson and L.-F. Li, Phys. Rev. D **5**, 2269 (1972). These authors used a different method with different sets of ansätze for the type (3) problem so that it is rather difficult to compare their result to ours.
[26]H. Helson, *Lectures on Invariant Subspaces* (Academic, New York, 1964).
[27]S. Bergman: *The Kernel Function and Conformal Mapping* Mathematical Series V. (American Mathematical Society, Providence, 1950).
[28]S. Bergman and M. Schiffer, *Kernel Functions and Elliptic Differential Equations in Mathematical Physics* (Academic, New York, 1953).
[29]I. Raszillier, Commun. Math. Phys. **26**, 121 (1971). D. N. Levin and S. Okubo, Phys. Rev. **D6**, 3145 (1972).

Scattering theory, orthogonal polynomials, and the transport equation*

K. M. Case

The Rockefeller University, New York, New York 10021
(Received 28 February 1974)

It is shown that under rather weak restrictions the discrete eigenvalues occurring in one-velocity transport theory are real or purely imaginary, simple, ≥ 1 in magnitude, finite in number, and occur in $\pm$ pairs. Proofs are obtained using methods of scattering theory applied to orthogonal polynomials.

1. INTRODUCTION

It has been noted[1] that there is an intimate relation between the solutions of the one-velocity linear transport equation and the theory of orthogonal polynomials. On the other hand, we recently showed that scattering theory throws considerable light on the properties of orthogonal polynomials.[2] The interplay of these results were then used[3] to investigate some simple inverse transport problems. Here we further develop this interplay.

The problem we consider is the following: In the theory of neutron transport with anisotropic scattering the number and position of the eigenvalues of a certain operator play a crucial role.[4] While statements about these eigenvalues appear in the literature[5] few proofs seem available and what statements that have been made apparently apply only to nonmultiplying media with a scattering kernel which is a finite polynomial.

Here using the connection with properties of orthogonal polynomials we wish to give proofs that these eigenvalues

(i) are real or purely imaginary,
(ii) are simple,
(iii) occur in $\pm$ pairs,
(iv) are ≥ 1 in absolute magnitude,
(v) are finite in number.

These properties are proved with rather weak restrictions on the scattering kernel. For nonmultiplying media ($c \leq 1$), it is believed that our restrictions are about as weak as possible for the above properties to hold. If the medium is multiplying, an additional condition (physically not very restrictive) has been applied. It is not known whether this can be relaxed.

The approach throws a certain amount of light on results previously obtained. In particular, those results which are special for the finite polynomial scattering kernel can be isolated.

Our plan is as follows: In Sec. 2 the relation of the eigenvalue problem of transport theory to a corresponding problem involving a three term difference equation eigenvalue problem is sketched. Such equations are typical in the theory of orthogonal polynomials. Accordingly, in Sec. 3 we then derive some properties of such polynomials. Using these results, the following Secs. give detailed proofs of the eigenvalue properties enumerated above for $c<1$, $c=1$, and $c>1$, respectively. Precise sufficient conditions in the various cases are given in Appendix D which sketches the detailed proof of analyticity needed. Appendixes A and B show the connection with previous formulations. In the last, Appendix D, some new and useful formulas are obtained as a by-product.

2. THE CONNECTION WITH ORTHOGONAL POLYNOMIALS

A typical problem[6] of one-velocity neutron transport with anisotropic scattering is to find ψ which satisfies

$$\frac{\partial}{\partial t}\psi(\mathbf{r},\Omega,t)+\Omega\cdot\nabla\psi+\psi=q+c\int d\Omega' f(\Omega\cdot\Omega')\psi(\mathbf{r},\Omega',t) \qquad (2.1)$$

in a volume V subject to appropriate boundary conditions on the bounding surface S. (Here we have assumed a homogeneous medium and the functional dependence of f indicated is such that we have rotational invariance.)

Let us consider the simplified form of Eq. (2.1) when there is

(i) no dependence on t;
(ii) dependence on only one spatial variable $-x$;
(iii) azimuthal symmetry, i.e., dependence on Ω is only thru $\Omega_x=\mu$;
(iv) no inhomogeneous term.

If we expand the scattering function in Legendre polynomials (P_l) so that

$$f(\Omega\cdot\Omega')=\sum_{l=0}^{\infty}\frac{2l+1}{4\pi}f_l P_l(\Omega\cdot\Omega'), \qquad (2.2)$$

then Eq. (2.1) reduces to

$$\mu\frac{\partial\psi}{\partial x}+\psi=c\sum_{l=0}^{\infty}\frac{2l+1}{2}f_l P_l(\mu)\int_{-1}^{1}P_l(\mu')\psi(\mu')\,d\mu'. \qquad (2.3)$$

Looking for infinite medium solutions of this equation in the form $\phi_\nu(\mu)e^{-x/\nu}$, one finds

$$(\nu-\mu)\phi_\nu(\mu)=\frac{c\nu}{2}M(\mu,\nu), \qquad (2.4)$$

where

$$M(\mu,\nu)=\sum_{l=0}^{\infty}(2l+1)f_l P_l(\mu)h_l(\nu) \qquad (2.5)$$

with

$$h_l(\nu)=\int_{-1}^{1}\phi_\nu(\mu)P_l(\mu)\,d\mu. \qquad (2.6)$$

We remark that with no loss of generality we can take $h_0(\nu)=1$ and define $h_{-1}(\nu)$ to be zero.

Now it can be shown[6] that under some conditions a complete set of functions is

$$\phi_\nu(\mu) = \frac{c\nu}{2} P \frac{M(\mu,\nu)}{\nu-\mu} + \lambda(\nu)\delta(\mu-\nu), \quad -1 \le \nu \le 1,$$

and

$$\phi_{\nu_i}(\mu) = \frac{c\nu_i}{2}\frac{M(\mu,\nu_i)}{\nu_i-\mu},$$

where

$$\lambda(\nu) = \frac{\Lambda^+(\nu)+\Lambda^-(\nu)}{2}, \quad \Lambda(\nu) = 1 - \frac{c\nu}{2}\int_{-1}^{1}\frac{M(\mu,\nu)d}{\nu-\mu}. \tag{2.7}$$

Here $\Lambda^\pm$ are the boundary values of Λ as the cut is approached from above and below, respectively. The discrete eigenvalues ν_i are determined by the equation

$$\Lambda(\nu_i) = 0. \tag{2.8}$$

Equations for the $h_l(\nu)$ are obtained[1] by multiplying Eq. (2.4) by $P_l(\mu)$ and using the orthogonality and recursion properties of the Legendre polynomials. The result is the three-term recursion relation

$$(l+1)h_{l+1}(\nu) + l\,h_{l-1}(\nu) = (2l+1)(1-cf_l)\nu h_l, \quad l \ge 0. \tag{2.9}$$

This and the above values of h_0, h_{-1} clearly uniquely determine all $h_l(\nu)$.

An equivalent form of the eigenvalue problem defined by Eq. (2.4) is to find those ν for which Eq. (2.9) has bounded solutions. The solutions which are merely bounded correspond to the continuum. Those ν_i for which there are square-summable solutions are the discrete eigenvalues.

It is useful to consider the more symmetrical polynomials defined by

$$\psi(\nu,n) = [(2n+1)/2]^{1/2}h_n(\nu). \tag{2.10}$$

Then our equations are

$$a^0(n+1)\psi(\nu,n+1) + a^0(n)\psi(\nu,n-1)$$

$$= \nu g(n)\psi(\nu,n), \quad n \ge 0, \tag{2.11}$$

where

$$\psi(\nu,-1)=0, \quad \psi(\nu,0)=\sqrt{\tfrac{1}{2}} \tag{2.12}$$

and

$$a^0(n) = n/[(2n+1)(2n-1)]^{1/2}, \quad g(n) = 1 - cf_n. \tag{2.13}$$

Our main point here is to show that under the conditions given below the discrete eigenvalues have the properties enumerated in the introduction.

3. PRELIMINARIES

A. Properties of the f_l

(i) We assume $\sum_{l=0}^{\infty} l^2|f_l| < \infty$ (This is needed for the analysis below to hold.)

(ii) $f(\mu)$ is continuous and $f(\mu) \ge 0$, $-1 \le \mu \le 1$. (Physical considerations dictate this.) A consequence is that

$$|f_l| < 1, \quad l \ge 1. \tag{3.1}$$

Indeed,

$$f_l = \int_{-1}^{1} f(\mu)P_l(\mu)\,d\mu$$

and therefore

$$|f_l| \le \int_{-1}^{1} f(\mu)\,|P_l(\mu)|\,d\mu < \int_{-1}^{1} f(\mu)\,d\mu = 1 \tag{3.2}$$

$$\text{(iii)}\ g(2m+1) \ge 0, \quad m = 0,1,2,\ldots \tag{3.3}$$

This may appear to be a rather strange restriction. However, for $c \le 1$, Eq. (3.1) implies $g(n) \ge 0$ (all n) and hence this is no additional restriction at all. For $c > 1$ this is a *sufficient* condition for the results listed above to hold. Whether it is also necessary is not known. In any event it is, physically, a rather mild requirement.

B. Some consequences of three-term recursion relations

Let us consider[7] functions $\Phi(\lambda,n)$ which satisfy

$$A(n+1)\Phi(\lambda,n+1) + B(n)\Phi(\lambda,n) + A(n)\Phi(\lambda,n-1)$$

$$= \lambda\Phi(\lambda,n), \quad n \ge n_0, \tag{3.4}$$

where

(i) $A(n)$, $B(n)$ are real and finite for all $n \ge n_0$;
(ii) $\lim_{n\to\infty} A(n) = A(\infty)$ and $\lim_{n\to\infty} B(n) = B(\infty)$ exist.

For our purposes it is sufficient that convergence be such that the limits $\lim_{n\to\infty} n^2[A(n)-A(\infty)]$ and $n^2[B(n) - B(\infty)]$ exist.

Then for large n the solutions of Eq. (3.4) tend to solutions of

$$A(\infty)\{\bar\Phi(\lambda,n+1) + \bar\Phi(\lambda,n-1)\} + B(\infty)\bar\Phi(\lambda,n) = \lambda\bar\Phi(\lambda,n) \tag{3.5}$$

Two linearly independent solutions of this are

$$\bar\Phi(\lambda,n) = z^{\pm n}$$

where

$$\lambda = B(\infty) + A(\infty)[z + z^{-1}]. \tag{3.6}$$

Thus for $|z| = 1$, i.e., $z = e^{i\theta}$ (θ real), all solutions of Eq. (3.5) are bounded as $n \to \infty$. It follows that all solutions of Eq. (3.4) for

$$\lambda = B(\infty) + 2A(\infty)\cos\theta, \quad -1 \le \cos\theta \le 1, \tag{3.7}$$

are bounded.[8] In particular, solutions $\Phi(\lambda,n)$ of Eq. (3.4) subject to $\Phi(\lambda,n_0-1)=0$, $\Phi(\lambda,n_0)=C$ (C a const) are bounded. Thus with these initial conditions the spectrum corresponding to Eq. (3.4) has a continuous part for

$$B(\infty) - 2A(\infty) \le \lambda \le B(\infty) + 2A(\infty). \tag{3.8}$$

Let us define[9] solutions of Eq. (3.4), $\Phi^{\pm}(\lambda, n)$, by the conditions

$$\lim_{n\to\infty} |\Phi^{\pm}(\lambda, n) - z^{\pm n}| = 0 \quad \text{for} \quad |z| \leq 1 \\ |z| \geq 1 \tag{3.9}$$

where λ is related to z by Eq. (3.6).

Further, let us use Eq. (3.4) to define the analog of the Jost function[9] as

$$f^{+}(z) = A(n_0)\Phi^{+}(\lambda, n_0 - 1). \tag{3.10}$$

[We analogously define $f^{-}(z)$].

Then square-summable solutions of Eq. (3.4) subject to the initial conditions exist for those z_i with $|z_i| \leq 1$ such that

$$f^{+}(z_i) = 0. \tag{3.11}$$

The corresponding λ_i are discrete eigenvalues with eigenfunctions

$$\Phi(\lambda_i, n) = C\,\frac{\Phi^{+}(\lambda_i, n)}{\Phi^{+}(\lambda_i, n_0)}. \tag{3.12}$$

We investigate the properties of the eigenvalues and eigenfunctions of Eq. (3.4) subject to the initial conditions in more detail. Consider the following Green type identity.

Let $\Phi^{(1)}(\lambda, n)$, $\Phi^{(2)}(\lambda', n)$ be two solutions of Eq. (3.4) corresponding to λ, λ', respectively. Then it follows in familiar fashion that

$$\begin{aligned}(\lambda - \lambda')\Phi^{(2)}(\lambda', n)\Phi^{(1)}(\lambda, n) &= A(n+1)\{\Phi^{(2)}(\lambda', n)\Phi^{(1)}(\lambda, n+1) - \Phi^{(1)}(\lambda, n)\Phi^{(2)}(\lambda', n+1)\} \\ &\quad + A(n)\{\Phi^{(2)}(\lambda', n)\Phi^{(1)}(\lambda, n-1) - \Phi^{(1)}(\lambda, n)\Phi^{(2)}(\lambda', n-1)\}, \\ &\quad n \geq n_0 \end{aligned} \tag{3.13}$$

Some applications

(i) Let $\lambda = \lambda'$. We conclude that

$$W[\Phi^{(1)}, \Phi^{(2)}] = A(n)\{\Phi^{(1)}(\lambda, n)\Phi^{(2)}(\lambda, n-1) - \Phi^{(1)}(\lambda, n-1)\Phi^{(2)}(\lambda, n)\} \tag{3.14}$$

is independent of n (the Wronskian theorem). Thus for λ in the continuum ($|z| = 1$) we can write

$$\phi(\lambda, n) = C_{+}\Phi^{+}(\lambda, n) + C_{-}\Phi^{-}(\lambda, n), \tag{3.15}$$

where the constants $C_{\pm}$ are respectively

$$C_{\pm} = \pm\frac{W[\Phi, \Phi^{\mp}]}{W[\Phi^{+}, \Phi^{-}]}. \tag{3.16}$$

From the asymptotic behavior of $\Phi^{\pm}$ we find

$$W[\Phi^{+}, \Phi^{-}] = A(\infty)[z - z^{-1}], \tag{3.17}$$

while from the initial conditions on Φ and the definitions of $f^{\pm}$

$$W[\Phi, \Phi^{\pm}] = Cf^{\pm}(z). \tag{3.18}$$

Notice that for

$$z = e^{i\theta}, \quad \Phi^{+} = \Phi^{-*}, \quad \text{and} \quad f^{-} = f^{+*} \tag{3.19}$$

and thus

$$\Phi(\lambda, n) = \frac{C}{A(\infty)[z - z^{-1}]}\{f^{-}(z)\Phi^{+}(\lambda, n) - f^{+}(z)\Phi^{-}(\lambda, n)\}, \qquad \lambda \in \text{continuum} \tag{3.20}$$

In particular, when $z \to z^{-1}$, i.e., $e^{i\theta} \to e^{-i\theta}$, Φ is unchanged. Hence the continuum is characterized by

$$\lambda = B(\infty) + 2A(\infty)\cos\theta, \quad 0 \leq \theta \leq \pi. \tag{3.21}$$

We also note that a discrete eigenvalue cannot occur in the continuum since

$$f^{+}(z_i) = 0 \to f^{-}(z_i) = 0 \to \Phi(\lambda_i, n) \equiv 0. \tag{3.22}$$

The closest a discrete eigenvalue can be is at the edge of the continuum ($z_i = \pm 1$), since then the denominator in Eq. (3.20) also vanishes.

(ii) Let us sum Eq. (3.13) from n_0 to N. With some relabelling we obtain

$$\begin{aligned}(\lambda - \lambda')\sum_{n=n_0}^{N}\Phi^{(2)}(\lambda', n)\Phi^{(1)}(\lambda, n) &= A(N+1)\{\Phi^{(1)}(\lambda, N+1)\Phi^{(2)}(\lambda', N) - \Phi^{(2)}(\lambda', N=1)\Phi^{(1)}(\lambda, N)\} \\ &\quad - A(n_0)\{\Phi^{(2)}(\lambda', n_0 - 1)\Phi^{(2)}(\lambda, n_0) \\ &\quad - \Phi^{(1)}(\lambda, n_0 - 1)\Phi^{(2)}(\lambda', n_0)\}.\end{aligned} \tag{3.23}$$

Suppose $\lambda = \lambda_i$ (a discrete eigenvalue) and $\lambda' = \lambda_i^*$. Letting $N \to \infty$ the terms $\sim A(N+1) \to 0$. Those multiplying $A(n_0)$ are zero by virtue of the initial conditions

$$\therefore 2\,\mathrm{Im}\,\lambda_i \sum_{n=n_0}^{\infty} |\phi(\lambda_i, n)|^2 = 0. \tag{3.24}$$

Thus, *the discrete eigenvalues are real.* Notice that we can (and for simplicity will) choose the $\phi(\lambda_i, n)$ to be real.

(iii) In Eq. (3.23) let $\lambda = \lambda_i$, $\Phi^{(1)}(\lambda, n) = \phi(\lambda_i, n)$, and $\Phi^{(2)}(\lambda', n) = \phi^{+}(\lambda', n)$ where $\lambda' \approx \lambda_i$. Again letting $N \to \infty$, we obtain

$$\sum_{n=n_0}^{\infty}\phi^{+}(\lambda', n)\phi(\lambda_i, n) = \frac{1}{\lambda' - \lambda_i}f^{+}(\lambda'). \tag{3.25}$$

Passing to the limit $\lambda' \to \lambda_i$ yields

$$\sum_{n=n_0}^{\infty}\phi^{+}(\lambda_i, n)\phi(\lambda_i, n) = C\frac{d}{d\lambda'}f^{+}(\lambda')\Big|_{\lambda'=\lambda_i}.$$

Remembering Eq. (3.12) we can rewrite this as

$$\sum_{n=n_0}^{\infty} |\phi(\lambda_i, n)|^2 = \frac{C^2}{\phi^+(\lambda_i, n_0)} \frac{d}{d\lambda'} f^+(\lambda')\Big|_{\lambda'=\lambda_i}. \tag{3.26}$$

Besides giving an explicit expression for the normalization sum of $\phi(\lambda_i, n)$ this also shows that *the discrete eigenvalues are simple*. Indeed, a multiple eigenvalue corresponds to a multiple root of f^+. If this occurred the derivative above would be zero and then $\phi(\lambda_i, n) \equiv 0$.[10]

C. Relevance

It is probably clear that the properties listed in the introduction are to be proved using properties of the related discrete eigenvalues of Eq. (3.4). Thus the fact that the ν_i^2 are real will be related to the reality of the λ_i. The fact that $|\nu_i| \geq 1$ are related to the eigenvalues of Eq. (3.4) being such that $|z_i| \leq 1$ with the only possibilities for $|z_i| = 1$ being $z_i = \pm 1$. The proof, subject to suitable restrictions, that there are only a finite number of the ν_i requires somewhat more analysis. This will be shown by demonstrating that there are only a finite number of the related λ_i. The proof of this, in turn, reduces to showing that $f^+(z)$ is analytic (except perhaps for a simple pole at $z=0$) for $|z| \leq 1$. The method for demonstrating this is illustrated in Appendix (C) where the special case $c<1$ is considered.

4. THE CASE $c < 1$

If in Eq. (2.11) we make the replacement

$$\psi(\nu, n) = \phi(\nu, n)/\sqrt{g(n)}, \tag{4.1}$$

we obtain

$$a(n+1)\phi(\nu, n+1) + a(n)\phi(\nu, n-1) = \nu\phi(\nu, n), \quad n \geq 0, \tag{4.2}$$

where

$$a(n) = a^0(n)/\sqrt{g(n)g(n-1)}. \tag{4.3}$$

We note that this is precisely Eq. (3.4) with the identification

$$A(n) = a(n), \quad B(n) = 0, \quad n_0 = 0. \tag{4.4}$$

[Here the essential use of the condition $c<1$ is that this guarantees $g(n) > 0$ (all n) and hence the coefficients $a(n)$ are real.]

Thus from the previous section we immediately conclude that the discrete eigenvalues ν_i are

(i) real,
(ii) simple, and
(iii) $|\nu_i| \geq 1$ [since $A(\infty) = \frac{1}{2}$, $B(\infty) = 0$].

That the ν_i occur in $\pm$ pairs is readily seen. Thus, if ν_i is a discrete eigenvalue corresponding to the eigenfunction $\phi(\nu_i, n)$, then $\phi'(\nu_i, n) = (-1)^n \Phi(\nu_i, n)$ satisfies

$$a(n+1)\phi'(\nu_i, n+1) + a(n)\phi'(\nu_i, n-1) = -\nu_i\phi'(\nu_i, n)$$

$$\therefore \phi'(\nu_i, n) = \phi(-\nu_i, n). \tag{4.5}$$

In Appendix (C) we prove (subject to the condition $\sum_l l^2 |f_l| < \infty$) that $f^+(z)$ is analytic within the unit circle except that it has a simple pole at $z=0$. Further, subject to $\sum_l l^2 |f_l| < \infty$, $f^+(z)$ is continuous on the unit circle. Hence we conclude it has at most a finite number of zeros in the unit circle and hence there are a finite number of discrete ν_i.

5. THE CASE $c = 1$

The basic equation again is (2.11), but now

$$g(0) = 1 - c = 0, \quad g(n) = 1 - f_n > 0, \quad n > 0. \tag{5.1}$$

From Eq. (2.11) with $n=0$ we conclude that

$$\psi(\nu, 1) = 0. \tag{5.2}$$

With $n=1$ we obtain

$$\psi(\nu, 2) = -\sqrt{\tfrac{5}{8}} \tag{5.3}$$

Thus our problem is to find the bounded solutions of

$$a^0(n+1)\psi(\nu, n+1) + a^0(n)\psi(\nu, n-1) = \nu g(n)\psi(\nu, n), \quad n \geq 2, \tag{5.4}$$

subject to the initial condition of (5.2) and (5.3).

As in Sec. 4 we now define $\phi(\nu, n)$ by

$$\psi(\nu, n) = \phi(\nu, n)/\sqrt{g(n)}, \quad n \geq 2,$$

and obtain

$$a(n+1)\phi(\nu, n+1) + a(n)\phi(\nu, n-1) = \nu\phi(\nu, n), \quad n \geq 2, \tag{5.5}$$

$$\phi(\nu, 2) = \text{given const}, \quad \phi(\nu, 1) = 0.$$

[Here by virtue of Eq. (5.1) the $a(n)$ are again real.] The considerations of Sec. 3 apply and we draw the same conclusions about the discrete eigenvalues as in Sec. 4. Thus, for example, we define a function $f_+'(z)$ as follows: Let $\phi^+(\nu, n)$ be the solution of Eq. (5.5) subject to $\lim_{n\to\infty} |\phi^+(z, n) - z^n| = 0$, $|z| \leq 1$. Then $f^{+\prime}(z) = a(1)\phi(\nu, 1)$. The zeros of $f^{+\prime}$ within the unit circle are the discrete eigenvalues. Also $f^{+\prime}(z)$ is (subject to the same conditions as before) analytic within the unit circle. (One difference from the previous case is that there is no pole at the origin.)

It should, however, be fairly obvious that something is missing! Thus for a given ν we have two fewer coefficients $\psi(\nu, n)$. Alternatively, if we look at $\Lambda(\nu)$ (whose zeros give us the discrete eigenvalues in the case $c<1$, cf. Appendix (A)), we note that in the limit $c \to 1$ this has a double zero at $\nu = \infty$. This suggests that we go back to Eq. (2.3) and, in addition to solutions with x dependence of the form $e^{-x/\nu}$, we also look for solutions of the form

$$\psi(x, \nu) = \alpha(\mu) + \beta(\mu)x. \tag{5.6}$$

Substituting and equating terms proportional to x and independent of it separately, we obtain the two equations

$$\beta(\mu) = \sum_{l=0}^{\infty} \frac{2l+1}{2} f_l P_l(\mu) \int_{-1}^{1} P_l(\mu')\beta(\mu')d\mu' \tag{5.7}$$

and

$$\alpha(\mu)+\mu\beta(\mu)=\sum_{l=0}^{\infty}\frac{2l+1}{2}f_l P_l(\mu)\int_{-1}^{1}P_l(\mu')\beta(\mu')d\mu'. \quad (5.8)$$

If we multiply Eq. (5.7) by $P_m(\mu)$ and integrate, we obtain

$$(1-f_m)\int_{-1}^{1}P_m(\mu')\beta(\mu')d\mu'=0. \quad (5.9)$$

Since $(1-f_m)>0$, $m\neq 0$, we conclude that only the first Legendre coefficient of β is nonvanishing. Therefore, β is a constant independent of μ. Using this and a similar argument in Eq. (5.8), we further conclude that

$$\alpha(\mu)=\alpha_0-[\beta\mu/(1-f_1)] \quad (5.10)$$

where α_0 is some other constant.

Thus in addition to the $e^{-x/\nu}$ type solutions, we have two more which can be taken to be

$$\psi_1(x,\nu)=1 \quad \text{(independent of } x)$$

and

$$\psi_2(x,\nu)=x-[\mu/(1-f_1)].$$

Some remarks

(i) We note that of all our solutions for $c=1$, only ψ_2 has any dependence on f_1. (This has been noted in special cases previously.[11])

(ii) It is readily proved that the functions 1, μ and

$$\phi_\nu(\mu)=\sum_{n=2}^{\infty}\frac{2m+1}{2}^{1/2}\frac{\phi(\nu,n)P_n(\mu)}{(1-f_n)^{1/2}} \quad (5.11)$$

(where ν runs over the discrete and continuous spectrum) yield a complete set of functions of μ.

6. THE CASE $c>1$

Again the basic equation is (2.11). The same substitution $\psi(\nu,n)=\phi(\nu,n)/\sqrt{g(n)}$ yields

$$a(n+1)\phi(\nu,n+1)+a(n)\phi(\nu,n-1)=\nu\phi(\nu,n), \quad n\geq 0, \quad (6.1)$$

where

$$a(n)=a^0(n)/\sqrt{g(n)g(n-1)}, \quad g(-1)=\text{arbitrary}=1,$$

with

$$\phi(\nu,-1)=0, \quad \phi(\nu,0)=\sqrt{g(0)}\,\psi(\nu,0)=(1-c)/2. \quad (6.2)$$

Formal difficulties now arise since some $g(n)$ [certainly $g(0)$] are negative and the $a(n)$ need no longer be real—thus negating some of the results of Sec. 3.

Also we note that the "initial val $\phi(\nu,0)$ is now purely imaginary. However, this i tself plays no role. [We can determine the eigenvalue h the initial condition $\phi(\nu,0)=1$ and then find the cc ect $\phi(\nu,n)$ merely by multiplying by $\sqrt{(1-c)/2}$.]

To circumvent the difficulty of complex $a(n)$ let us iterate Eq. (6.1). One obtains

$$\begin{aligned}\nu^2\phi(\nu,n)&=a(n+2)a(n+1)\phi(\nu,n+2)\\&\quad+[a^2(n+1)+a^2(n)]\phi(\nu,n)\\&\quad+a(n)a(n-1)\phi(\nu,n-2). \quad (6.3)\end{aligned}$$

We note that these are recursion relations which only couple odd (or even) n. (In Appendix (D) these are used to derive what may be called "double" Christoffel–Darboux formulas which are interesting and have certain applications.)

For the present we restrict our attention to the case where $n=2m+1$. Define

$$\lambda=\nu^2, \quad \phi(\nu,2m+1)=\Phi(\lambda,m),$$

$$A(m)=a(2m)a(2m+1), \quad B(m)=a^2(2m+1)+a^2(2m+2).$$

Then Eq. (6.3) is

$$\begin{aligned}\lambda\Phi(\lambda,m)&=A(m+1)\Phi(\lambda,m+1)+B(m)\Phi(\lambda,m)\\&\quad+A(m)\Phi(\lambda,m-1), \quad m\geq 0. \quad (6.4)\end{aligned}$$

We note that $B(m)$ is real by construction, while

$$A(m)=\frac{a^0(2m)a^0(2m+1)}{g(2m)\sqrt{g(2m+1)g(2m-1)}}. \quad (6.5)$$

Hence, *if the only negative $g(n)$ are for even n* the $A(m)$ are all real. We shall assume this is so.[12] [This is not as unreasonable an assumption as one might think. A typical case of weak anisotropy would have $g(0)=1-c<0$ while all other $g(n)>0$.] Whether our results are true with some weaker assumption is not known.

The Eq. (6.4) is now in the form of Eq. (3.4). We still need initial conditions. Since $\phi(\nu,-1)=\Phi(\lambda,-1)$ we have

$$\Phi(\lambda,-1)=0. \quad (6.6)$$

From Eq. (6.1) with $n=0$ we obtain

$$\Phi(\lambda,0)=\phi(\nu,1)=\nu\sqrt{\tfrac{3}{2}}\,g(0)\sqrt{g(1)}. \quad (6.7)$$

(Notice that this is slightly different from before in that the initial condition here *depends* on the eigenvalue. However, it does so in a particularly simple fashion—multiplicatively.) Hence, to determine the λ_i giving square-summable solutions we can just as well replace the initial condition by

$$\Phi(\lambda,0)=1. \quad (6.8)$$

The considerations of Sec. 3 apply. The discrete eigenvalue λ_i of Eq. (6.4) are thus real, simple and $|\lambda_i|\geq 1$. Considerations (and conditions) similar to those for $c<1$ lead us to conclude that the corresponding $f^+(z)$ is analytic within the unit circle (except for the simple pole at zero) and hence that the λ_i are finite in number.

For each simple discrete eigenvalue λ_i we obtain two eigenvalues of Eq. (6.1). These are just

$$\nu_i = \sqrt{\lambda_i} \quad \text{and} \quad \nu_i = -\sqrt{\lambda_i}. \tag{6.9}$$

Thus, these are ± pairs. These eigenvalues are readily seen to be simple. Indeed the eigenfunctions are readily constructed. Thus for n odd

$$\phi(\nu_i, 2m+1) = \phi(-\nu_i, 2m+1) = \Phi(\lambda_i, m), \tag{6.10}$$

while for n even we have from Eq. (6.1)

$$\phi(\pm\nu_i, 2m) = \frac{a(2m+1)\Phi(\lambda_i, m) + a(2m)\Phi(\lambda_i, m-1)}{\pm\nu_i}$$

7. CONCLUSIONS

It is hoped that it has been demonstrated that the methods of scattering theory applied to orthogonal polynomials is a useful tool for investigating solutions of the transport equation. This is not to say that other formulations such as exemplified by Eqs. (2.7) do not also have advantages. However, for obtaining the properties of the discrete eigenvalues the present approach seems particularly simple. Indeed, all but the finiteness of the number of eigenvalues are found by completely elementary means. That this one property requires more sophisticated methods should be obvious. One is trying to find weak conditions on the f_l. Under strong conditions—such as $f_l = 0$, $l \geq N$—the result is immediate.[13]

I would like to thank the Institute for Advanced Study for hospitality during part of the time during which this work was done.

APPENDIX A: THE RELATION BETWEEN $f^+(z)$ AND $\Lambda(\nu)$

In Sec. 2 we described the usual treatment of the transport equation. The discrete eigenvalues there are the zeros of $\Lambda)\nu)$. In the remainder of our discussion these eigenvalues are described as the zeros of $f^+(z)$ within the unit circle. Clearly f^+ and Λ are related. Indeed here we show that

$$f^+(z) = f_0^+(z)\Lambda(\nu)/\sqrt{1-c} \tag{A1}$$

where

$$f_0^+(z) = [(z^{-2}-1)/2\pi]^{1/2}. \tag{A2}$$

(The notation is justified by the fact—to be shown—that f_0^+ is just the Jost function corresponding to the three-term recursion relation characteristic for the Legendre polynomials. In Appendix (B) it is demonstrated that it is natural to split off an explicit factor of f_0^+ for a general Jost function.)

Combining Eqs. (2.5) and (2.7) gives

$$\Lambda(\nu) = 1 - c\nu \sum_{l=0}^{\infty} (2l+1) f_l h_l(\nu) \tfrac{1}{2} \int_{-1}^{1} \frac{P_l(\mu)\, d\mu}{\nu - \mu}. \tag{A3}$$

From the definitions of the associated Legendre functions and Eq. (2.10) this can be rewritten as

$$\Lambda(\nu) = 1 - c\nu\sqrt{2} \sum_{l=0}^{\infty} f_l \sqrt{2l+1}\, \psi(\nu, l) Q_l(\nu). \tag{A4}$$

Now for $n \geq 1$, the function $\psi_0'(\nu, l) = \sqrt{2l+1}\, Q_l(\nu)$ satisfies the relation

$$a^0(l+1)\psi_0'(\nu, l+1) + a^0(l)\psi_0'(\nu, l-1) = \nu\psi_0'(\nu, l), \tag{A5}$$

with $a^0(l)$ given by Eq. (2.13). The asymptotic formula[14] for $\psi_0'(\nu, l)$ then shows that

$$\lim_{n\to\infty} \left| f_0^+(z)\sqrt{2l+1}\, Q_l(\nu) - z^n \right| \to 0, \quad |z| \leq 1. \tag{A6}$$

It is therefore appropriate to write

$$\psi_0^+(z, l) = f_0^+(z)\sqrt{2l+1}\, Q_l(\nu). \tag{A7}$$

The Jost function corresponding to Eq. (A5), $a^0(0)\psi_0^+(z, -1)$, is then given by

$$a^0(0)\psi_0^+(z, -1) = \nu\psi_0^+(z, 0) - a^0(1)\psi_0^+(z, 1). \tag{A8}$$

Using the explicit forms for the associated Legendre functions here, we readily obtain

$$a^0(0)\psi_0^+(z, -1) = f_0^+(z), \tag{A9}$$

thereby justifying the notation.

Now we can write Eq. (A4) as

$$\Lambda(\nu) = 1 - \frac{c\nu\sqrt{2}}{f_0^+} \sum_{l=0}^{\infty} f_l \psi(\nu, l)\psi_0^+(z, l), \tag{A10}$$

where

$$a^0(l+1)\psi_0^+(\nu, l+1) + a^0(l)\psi_0^+(\nu, l-1) = \nu\psi_0^+(\nu, l), \tag{A11}$$

$$a^0(l+1)\psi(\nu, l+1) + a^0(l)\psi(\nu, l-1) = g(\nu)\psi(\nu, l). \tag{A12}$$

Multiplying Eq. (A11) by $\psi(\nu, l)$, Eq. (A12) by $\psi_0^+(\nu, l)$, subtracting and summing yields

$$\begin{aligned} c\nu \sum_{l=0}^{L} f_l \psi_0^+(\nu, l)\psi(\nu, l) &= a^0(L+1)[\psi(\nu, L)\psi_0^+(\nu, L+1) - \psi_0^+(\nu, L)\psi(\nu, L+1)] \\ &\quad + a^0(0)[\psi_0^+(\nu, L-1)\psi(\nu, 0) - \psi(\nu, -1)\psi_0^+(\nu, 0)]. \end{aligned}$$

But the last bracket on the right is just $f_0^+/\sqrt{2}$,

$$\begin{aligned} \therefore\ \Lambda(\nu) = \frac{\sqrt{2}}{f_0^+} \lim_{L\to\infty} a^0(L+1)[&\psi_0^+(\nu, L)\psi(\nu, L+1) \\ &- \psi(\nu, L)\psi_0^+(\nu, L+1)]. \end{aligned}$$

However, we have

(i) $\lim_{L\to\infty} a^0(L+1) = \lim_{L\to\infty} a(L+1)$,

(ii) $\lim_{L\to\infty} \psi_0^+(\nu, L) = \lim_{L\to\infty} \phi^+(\nu, L)$,

(iii) $\lim_{L\to\infty} \psi(\nu, L) = \lim_{L\to\infty} \phi(\nu, L)$;

$$\begin{aligned} \therefore\ \Lambda(\nu) = \frac{\sqrt{2}}{f_0^+} \lim_{L\to\infty} a(L+1)[&\phi^+(\nu, L)\phi(\nu, L+1) \\ &- \phi(\nu, L)\phi^+(\nu, L+1)]. \end{aligned} \tag{A13}$$

The quantity whose limit we are to take is just the Wronskian—which as we have seen is independent of L. In particular, then

$$a(L+1)[\phi^+(\nu,L)\phi(\nu,L+1)-\phi(0,L)\phi^+(\nu,L+1)]$$
$$=a(0)[\phi^+(\nu,-1)\phi(\nu,o)-\phi(\nu,-1)\phi^+(\nu,0)]$$
$$=f^+\phi(\nu,0)=\sqrt{g(0)/2}f^+$$
$$\therefore \Lambda(\nu)=\frac{f^+}{f_0^+}\sqrt{1-c}. \tag{A14}$$

Remark: By simple iteration of the equations for $\phi^+(z,l)$ and $\phi_0^+(z,l)$, we readily find

$$\lim_{z\to 0}\frac{f^+(z)}{f_0^+(z)}=\sqrt{g(0)}\prod_{i=1}^{\infty}g(i) \tag{A15}$$

and then from Eq. (A1)

$$\lim_{\nu\to\infty}\Lambda(\nu)=\prod_{i=0}^{\infty}g(i). \tag{A16}$$

APPENDIX B: THE RELATIONSHIP OF f^+ TO THE SPECTRAL FUNCTION

Here we restrict ourself, for simplicity, to the case $c<1$. We wish to do three things:

(i) Give an explicit, direct calculation of f_0^+.
(ii) Show that it is natural to factor any f^+ in the form $f^+=f_0^+\Gamma$.
(iii) Demonstrate some general relations between f^+ and the spectral function. In particular, we find that f^+ determines the continuous part of the spectral function and the position of the discrete eigenvalues. Conversely, the continuous part of the spectral function and the position of the discrete eigenvalues explicitly determine f^+.

Following closely the methods in the Appendix of Ref. 2, we find that (a) if ν, ν' are in the continuum

$$\sum_{n=0}^{\infty}\phi(\nu',n)\phi(\nu,n)=\frac{\delta(\nu-\nu')}{\rho'(\nu)}, \tag{B1}$$

where

$$\rho'(\nu)=\frac{A(\infty)\sin\theta}{\pi C^2|f^+|^2}, \tag{B2}$$

and (b) if ν_i,ν_j are discrete eigenvalues

$$\sum_{n=0}^{\infty}\phi(\nu_i,n)\phi(\nu_j,n)=\frac{1}{\rho_i}\delta(\nu_i,\nu_j), \tag{B3}$$

where

$$\rho_i=\frac{\phi^+(\nu_i,0)}{C^2\left.\dfrac{d}{d\nu'}f^+(\nu')\right|_{\nu'=\nu_i}}. \tag{B4}$$

Conversely,

$$\int_{-1}^{1}\phi(\nu,n)\phi(\nu,m)\rho'(\nu)d\nu+\sum_i\rho_i\phi(\nu_i.n)\phi(\nu_i,m)=\delta(n,m). \tag{B5}$$

Thus given $f^+(z)$ we see we have $\rho'(\nu)$ and the position of the discrete eigenvalues—since these are the zeros of f^+ in the unit circle. We want to turn this around. Namely, we will construct f^+ given ρ' and the ν_i. [As an aside we note that this implies that given ρ' and the ν_i we can construct the asymptotic form of $\phi(\nu,n)$ when ν is in the continuum.]

The main tool is a version of the Poisson—Jensen formula. Thus let $h(z)$ be

(i) analytic within the unit circle;
(ii) real, i.e., $h^*(z)=h(z^*)$;
(iii) and $h(0)$ is real.

Then we readily find the representation

$$h(r,\theta)=\frac{1}{2\pi}\int_{-\pi}^{\pi}h_r(1,\theta')\frac{[e^{i\theta'}+z]}{[e^{i\theta'}-z]}d\theta',\quad \text{where } z=re^{i\theta}. \tag{B6}$$

Consider

$$g(z)=\frac{zf^+(z)}{\prod_{i=1}^{N}(z_i^2-z^2)},$$

where the real, simple roots of f^+ are at $\pm z_i$.

On the basis of the results in the main text, Appendix (A), and Appendix (C), we conclude that $\ln g(z)$ is a suitable $h(z)$,

$$\therefore \ln g(z)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\{\ln|g(\theta')|\}\frac{[e^{i\theta'}+z]}{[e^{i\theta'}-z]}d\theta'.$$

But on the unit circle

$$|g(z)|=\frac{|f^+|}{\prod_i|z_i^2-z^2|}, \tag{B7}$$

$$\therefore \ln g(z)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln|f^+|\frac{[e^{i\theta'}+z]}{[e^{i\theta'}-z]}d\theta' - \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{[e^{i\theta'}+z]}{[e^{i\theta'}-z]}\sum_i\ln|z_i^2-z'^2|d\theta'. \tag{B8}$$

We note that $\sum_i\ln(1-z_i^2z^2)$ is also an h,

$$\therefore \sum_i\ln(1-z_i^2z^2)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{[e^{i\theta'}+z]}{[e^{i\theta'}-z]}\sum_i\ln|1-z_i^2z'^2|d\theta'. \tag{B9}$$

However, on the unit circle $|1-z_i^2z'^2|=|z_i^2-z'^2|$, therefore Eq. (B9) becomes

$$\ln\prod_i(1-z_i^2z^2)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{[e^{i\theta'}+z]}{[e^{i\theta'}-z]}\sum_i\ln|z_i^2-z'^2|d\theta. \tag{B10}$$

Adding Eqs. (B8) and (B10) yields

$$\ln g(z)\prod_i(1-z_i^2z^2)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln|f_+|\frac{[e^{i\theta'}+z]}{[e^{i\theta'}-z]}d\theta'. \tag{B11}$$

Then from the definition of $g(z)$ and using the result

$f^+(-\theta)=f^{+*}(\theta')$ we obtain the representation

$$f^+(z)=\frac{\Pi_i(z_i^2-z^2)}{\Pi_i(1-z_i^2z^2)}\frac{1}{z}e^I, \tag{B12}$$

where from Eq. (B2)

$$f^+(\theta')=\left(\frac{A(\infty)\sin\theta'}{\pi C^2\rho'}\right)^{1/2}$$

and thus

$$I=\frac{z-z^{-1}}{4\pi}\int_0^\pi\frac{\ln(\sin\theta'/\pi)\,d\theta'}{\cos\theta'-\nu} -\frac{(z-z^{-1})}{4\pi}\int_0^\pi\frac{\ln[C^2\rho'(\nu)/A(\infty)]}{\cos\theta'-\nu}d\theta'. \tag{B13}$$

It is then seen that quite generally $f^+(z)$ has a factor

$$z^{-1}\exp\frac{z-z^{-1}}{4\pi}\int_0^\pi\frac{\ln(\sin\theta'/\pi)\,d\theta'}{\cos\theta'-\nu}. \tag{B14}$$

This always occurs independently of ρ' and the z_i. In particular, for the Legendre case there are no discrete eigenvalues, $\rho'(\nu)=1$, $C^2=A(\infty)=\frac{1}{2}$, and then $f_L^+\equiv f_0^+$.

The integral in Eq. (B14) can be done explicitly. Thus,

$$\int_0^\pi\frac{\ln(\sin\theta'/\pi)\,d\theta'}{\cos\theta'-\nu}=\frac{\Pi}{\sqrt{\nu^2-1}}\left[\ln\nu-\ln\left(\frac{\sqrt{\nu^2-1}}{\nu+\sqrt{\nu^2-1}}\right)\right]$$

and we obtain

$$f_0^+=[(z^{-2}-1)/2\pi]^{1/2}. \tag{B15}$$

The remaining integral in Eq. (B13) is conveniently transformed to one over the spectrum ($\cos\theta'=\nu'$) and thus we finally obtain the desired result:

$$f^+(z)=f_0^+\frac{\Pi_i(z_i^2-z^2)}{\Pi_i(1-z_i^2z^2)}\exp -\frac{(z-z^{-1})}{4\pi}\int_{-1}^1\frac{\ln[C^2\rho'(\nu')/A(\infty)]}{\sqrt{1-\nu'^2}(\nu'-\nu)}d\nu'. \tag{B16}$$

(*Note*: The i here runs only over the positive z_i's.) We see how explicitly $\rho'(\nu)$ and the z_i determine $f^+(z)$.

APPENDIX C: ANALYTIC PROPERTIES

Here we wish to sketch the proof of the various properties mentioned in the main text. To be specific we treat only the case $c<1$.

(1) $\phi(\nu,n)$ is analytic on and within the unit circle in the z plane except for a pole at $z=0$. Indeed, $\phi(\nu,n)$ is a polynomial of order n in $\nu=(z+z^{-1})/2$.

(2) $f^+(z)$ is analytic within the unit circle (except for a simple pole at $z=0$) and is continuous on the circle provided

$$\sum_{m=0}^\infty m^2|f_m|<\infty. \tag{C1}$$

To prove this we first demonstrate that $\phi^+(z,n)$ is analytic within and continuous on the unit circle. From the definition

$$f^+(z)=\frac{z+z^{-1}}{2}\phi^+(z,0)-a(1)\phi^+(z,1), \tag{C2}$$

it then follows that f^+ has at most a simple pole (and it is at zero).

The pole does indeed exist since from Eq. (A15)

$$\lim_{z\to0}zf^+(z)=\left(\frac{1-c}{2}\right)^{1/2}\prod_{l=1}^\infty(1-cf_l)\neq0. \tag{C3}$$

As in discussing the solutions of differential equations, it is useful to transform to an "integral" equation incorporating the boundary conditions. A convenient starting point is Eq. (2.11), i.e.,

$$\nu\psi^+(\nu,n)-a^0(n+1)\psi^+(\nu,n+1)-a^0(n)\psi^+(\nu,n-1)=\nu cf_n\psi^+(\nu,n), \tag{C4}$$

with the condition $\lim_{n\to\infty}|\psi^+(\nu,n)-z^n|=0$, $|z|\leq1$.

We introduce a Green's function defined by

$$G(\nu,n;m)=0,\quad n\geq m$$
$$=\frac{\sqrt{2}}{f_0^+}\{\psi_0^+(m)\psi_0(n)-\psi_0(m)\psi_0^+(n)\},\ n\leq m. \tag{C5}$$

(Here ψ_0, ψ_0^+ are the regular and (+) solution of Eq. (C4) with $c=0$.) Explicitly,

$$\psi_0(n)=[(2n+1)/2]^{1/2}P_n(\nu), \tag{C6}$$

$$\psi_0^+(n)=(2n+1)^{1/2}f_0^+Q_n(\nu).$$

Then

$$\psi^+(\nu,n)=\psi_0^+(\nu,n)+\sum_{m=n+1}^\infty G(\nu,n;m)\nu cf_m\psi^+(\nu,m). \tag{C7}$$

We solve this by iteration. Thus,

$$\psi^+(\nu,n)=\sum_{i=0}^\infty\psi^{+(i)}(\nu,n), \tag{C8}$$

where

$$\psi^{+(0)}(\nu,n)=\psi_0^+(\nu,n)$$

and

$$\psi^{+(i)}(\nu,n)=\sum_{m=n+1}^\infty G(\nu,n;m)\nu cf_m\psi^{+(i-1)}(\nu,m). \tag{C9}$$

Now for $|z|<1$ we have the estimates

$$|\psi_0^+(z,n)|<C_0|z|^n,$$
$$|G(\nu,n;m)\nu|\leq C_1|z|^{n-m}\sqrt{m}. \tag{C10}$$

Thus,

$$|\psi^{+(i)}(\nu,n)|\leq C_1\sum_{m=n+1}^\infty|z|^{n-m}\sqrt{m}\,|cf_m|\,|\psi^{+(i-1)}(\nu,m)|. \tag{C11}$$

In particular,

$$|\psi^{+(1)}(\nu,n)|\leq C_0C_1|z|^n\sum_{m=n+1}^\infty\sqrt{m}\,|cf_m|,$$

$$|\psi^{+(2)}(\nu,n)| \le C_0 C_1^2 |z|^n \sum_{m=n+1}^{\infty} \sqrt{m}\,|cf_m| \sum_{m'=m+1}^{\infty} \sqrt{m'}\,|cf_{m'}|$$

$$= C_0 C_1^2 \frac{|z|^n}{2!} \left(\sum_{m=n+1}^{\infty} \sqrt{m}\,|cf_m|\right)^2,$$

$$|\psi^{+(i)}(\nu,n)| \le C_0 \frac{C_1^i |z|^n}{i!} \left(\sum_{m=n+1}^{\infty} \sqrt{m}\,|cf_m|\right)^i,$$

$$\therefore\ |\psi^{+}(\nu,n)| \le \sum_{i=0}^{\infty} C_0 \frac{C_1^{(i)} |z|^n}{i!} \left(\sum_{m=n+1}^{\infty} \sqrt{m}\,|cf_m|\right)^i$$

$$= C_0 |z|^n \exp C_1 \sum_{m=n+1}^{\infty} \sqrt{m}\,|cf_m|. \tag{C12}$$

Therefore, the series converges uniformly provided

$$\sum_{m=0}^{\infty} \sqrt{m}\,|f_m| < \infty.$$

For analyticity we need to show that the derivative also converges. This proof is as above—with the differentiation introducing another factor of m. Thus we have analyticity within the unit circle provided

$$\sum_{m=0}^{\infty} m^{3/2} |f_m| < \infty. \tag{C13}$$

To discuss the behavior on the unit circle we need a more refined estimate for G. Introducing the form of Eq. (C6), we have

$$G(\nu,n;m) = \sqrt{(2n+1)(2m+1)}\,\{Q_m(\nu)P_n(\nu) - P_m(\nu)Q_n(\nu)\} \tag{C14}$$

But $Q_m(\nu) = P_m(\nu)Q_0(\nu) + W_{m-1}(\nu)$,

$$\therefore\ G(\nu,n;m) = \sqrt{(2n+1)(2m+1)}\,\{W_{m-1}(\nu)P_n(\nu) - W_{n-1}(\nu)P_m(\nu)\}. \tag{C15}$$

Now on the unit circle $|P_n(\nu)| \le 1$, $|W_n(\nu)| < n$.[15] Since always $n \le m$ we then obtain the bound

$$|G(\nu,n,m)| < C_2 m^2, \quad |z| = 1. \tag{C16}$$

This then gives as our convergence criterion

$$\sum_{m=0}^{\infty} m^2 |f_m| < \infty. \tag{C17}$$

Since this is the strongest of our conditions (and is in practice rather weak) we adopt this throughout the work.

APPENDIX D

"Doubled" Christoffel-Darboux formulas

It is well-known[17] that for functions satisfying three-term recursion relations of the form of Eq. (3.4) there exist summation formulas expressing sums like $\sum_{n=0}^{N}\Phi^2(\lambda,N)$ in terms of $\Phi(\lambda,N)$ and $\Phi(\lambda,N+1)$. Indeed, if in Eq. (3.23) we choose

$$\Phi^{(1)}(\lambda,n) = \Phi(\lambda,n),$$

$$\Phi^{(2)}(\lambda',n) = \Phi(\lambda',n)$$

where $\Phi(\lambda,n_0-1)=0$ and pass to the limit $\lambda' \to \lambda$, we find

$$\sum_{n=n_0}^{\infty} \Phi^2(\lambda,n) = A(N+1)\left[\Phi(\lambda,N)\frac{d}{d\lambda}\Phi(\lambda,N+1) - \Phi(\lambda,N+1)\frac{d}{d\lambda}\Phi(\lambda,N)\right]. \tag{D1}$$

Therefore for functions which satisfy

$$a(n+1)\phi(\nu,n+1) + a(n)\phi(\lambda,n-1) = \nu\phi(\lambda,n), \quad n \ge n_0, \tag{D2}$$

we have what we will call the simple C–D formula

$$\sum_{n=n_0}^{N} \phi^2(\nu,n) = a(N+1)\{\phi(\nu,N)\frac{d}{d\nu}\phi(\lambda,N+1) - \phi(\nu,N+1)\frac{d}{d\nu}\Phi(\lambda,N)\}. \tag{D3}$$

However, in Sec. 6 we found on iterating Eq. (D2) that

$$\nu^2\phi(\nu,n) = a(n+2)a(n+1)\phi(\nu,n+2) + [a^2(n+1)a^2(n)]\phi(\nu,n) + a(n)a(n-1)\phi(\nu,n-2). \tag{D4}$$

(i) For odd $n = 2_m + 1$ we define

$$\phi(\nu,2_m+1) = \Phi(\nu,m), \quad A(m) = a(2m)a(2m+1),$$

$$B(m) = [a^2(2m+2) + a^2(2m+1)], \quad \lambda = \nu^2.$$

Then Eq. (D4) is precisely of the form of Eq. (3.4) and Eq. (D1) yields

$$\sum_{m=m_0}^{M} \phi^2(\nu,2m+1) = \frac{a(2M+2)a(2M+3)}{2\nu}\left\{\phi(\nu,2M+1) \times \frac{d}{d\nu}\phi(\nu,2M+3) - \phi(\nu,2M+3) \times \frac{d}{d\nu}\phi(\nu,2M+1)\right\} \tag{D5}$$

(ii) For even $m = 2m$ we obtain analogously

$$\sum_{m=m_0}^{M} \phi^2(\nu,2m) = \frac{a(2M+2)a(2M+1)}{2\nu}\left\{\phi(\nu,2M)\frac{d}{d\nu}\phi(\nu,2M+2) - \phi(\nu,2M+2)\frac{d}{d\nu}\phi(\nu,2M)\right\}. \tag{D6}$$

We note that Eqs. (D5) and (D6) are formally the same, i.e., they say

$$\sum_{n=n_0}^{N}{}' \phi^2(\nu,n) = \frac{a(N+2)a(N+1)}{2\nu}\left\{\phi(\nu,N)\frac{d}{d\nu}\phi(\nu,N+2) - \phi(\nu,N+2)\frac{d}{d\nu}\phi(\nu,N)\right\} \tag{D7}$$

where $\sum'$ means a sum over even n if N is even or a sum over odd n if N is odd.

Let us call Eq. (D7) a "double" Christoffel–Darboux formula. While the "single" formulas follow only from the three-term recursion relation it is seen that the double ones are a consequence of the additional restriction that the "diagonal" terms $B(n)$ are identically zero.

An application

In Ref. 2 a simple example was given. The decisive quantity $K(n,n)$ was obtained in the form

$$\frac{1}{K^2(n,n)}=\alpha\left(1+\frac{(1+\alpha)\phi^2(\bar{\lambda},n)}{\alpha+(1-\alpha)\sum'\phi^2(\bar{\lambda},m)}\right) \tag{D8}$$

where for n even

$$\sum' = \sum_{\substack{1\leq m\leq n-2\\ m\ \text{even}}}$$

and for n odd

$$\sum' = \sum_{\substack{1\leq m\leq n-2\\ m\ \text{odd}}}.$$

Here the $\phi(\bar{\lambda},n)$ are Tchybecheff polynomials satisfying

$$\tfrac{1}{2}\{\phi(\bar{\lambda},n+1)+\phi(\bar{\lambda},n-1)\}=\bar{\lambda}\phi(\bar{\lambda},n) \tag{D9}$$

and

$$\phi(\lambda,0)=0,\quad \phi(\lambda,1)=1.$$

Applying our "double" Christoffel–Darboux formula, we then have

$$\sum{}'\phi^2(\bar{\lambda},m)=\frac{1}{8\bar{\lambda}}\{\phi(\bar{\lambda},n-2)\frac{d}{d\bar{\lambda}}\phi(\bar{\lambda},n) - \bar{\phi}(\bar{\lambda},n)\frac{d}{d\bar{\lambda}}\phi(\bar{\lambda},n-2). \tag{D10}$$

Explicit results are obtained by noting that if $\bar{\lambda}=\cosh\theta$, then

$$\phi(\bar{\lambda},m)=\frac{\sinh m\theta}{\sinh\theta} \tag{D11}$$

Inserting in Eq. (D10) then yields

$$\sum{}'\phi^2(\bar{\lambda},m)=\frac{1}{4\sinh^2\theta}\left((1-n)+\frac{\sinh(2n-2)\theta}{\sinh 2\theta}\right). \tag{D12}$$

Of particular interest is the behavior of $K(n,n)$ for large n. Using Eq. (D12) we obtain

$$1/K^2(n,n)\to 2e^{4\theta}-1. \tag{D13}$$

That $K(n,n)$ becomes independent of n is to be expected from the considerations in Ref. 2.

*This work was supported in part by the Air Force Office of Scientific Research, Grant 722187.

[1]E. Inönü, J. Math. Phys. 11, 568 (1970).

[2]K.M. Case and M. Kac, J. Math. Phys. 14, 594 (1973).

[3]K.M. Case, Phys. Fluids 16, 1607 (1973).

[4]There are mathematically similar questions which arise in other problems, e.g., the so-called X-Y model. See K.M. Case and C.W. Lau, J. Math. Phys. 14, 720 (1973).

[5]I. Kuscer and I. Vidav, J. Math. Appl. 25, 80 (1969) and K.O. Thielman and K. Claussen, Nucl. Sci. Eng. 29, 134 (1970).

[6]See, for example, K.M. Case and P.F. Zweifel, *Linear Transport Theory* (Addison-Wesley, Reading, Mass., 1967).

[7]This may appear to be a peculiar variant of Eq. (2.11) but it will be shown that the cases $c<1$, $c=1$, $c>1$ all reduce to this.

[8]This should be proved. The analysis is quite similar to that used in Appendix C.

[9]The similarity of our treatment to the theory of scattering should be noted [cf., for example, R.G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966)]. Indeed, our whole further discussion is just the discrete version of that theory (cf. Ref. 2).

[10]One might object to this argument by saying that $\phi^+(\lambda_i,n_0)$ could be zero. However, Eq. (3.4) with $f^+(z_i=0)$, $\phi^+(\lambda_i,n_0)=0$ implies $\phi^+(\lambda_i,n)\equiv 0$, in clear contradiction to the definition of $\phi^+(\lambda_i,n)$.

[11]See Ref. 6, p. 168.

[12]We also assume that no $g(n)=0$. If *one* is zero the treatment parallels that for $c=1$.

[13]See Ref. 6, p. 91.

[14]E.W. Hobson, *The Theory of Spherical and Ellipsoidal Harmonics* (Cambridge U.P., Cambridge, 1931), p. 61.

[15]Ref. 14, p. 63.

[16]E.C. Titchmarsh, *The Theory of Functions* (Oxford U.P., Oxford, 1939), p. 124.

[17]G. Szegö, *Orthogonal Polynomials* (American Mathematical Society, New York, 1939), p. 41.

Ising models derived from binary lattice gases*

L. K. Runnels

Department of Chemistry, Louisiana State University, Baton Rouge, Louisiana 70803
(Received 5 March 1974)

Transformations of two-component lattice gases are studied. The original binary lattice gas has the Widom–Rowlinson type of interaction: infinite repulsion between nearby unlike particles and no interaction between like particles (except that multiple site occupancy is forbidden). The first transformation is to an equivalent one-component lattice gas, with many-body interactions—some attractive, some repulsive. The second transformation is to the isomorphic Ising spin system, having many-spin interactions—some ferromagnetic and some antiferromagnetic. All interaction functions are obtained as explicit geometrical derivates of the mutual exclusion "sphere" of the original binary system. The results are discussed in light of recent theories of phase transitions of Ising models with symmetry-breaking many-spin interactions.

INTRODUCTION

There have been several publications recently[1–4] concerned with proofs of the existence of phase transition in Ising systems with many-spin interactions—including interactions of odd order that destroy the up–down symmetry in zero field. The purpose of this communication is to point out a special class of model containing many-spin interactions for which the ferromagnetic transition can be proved in other ways.

The "$A-B$" model of Widom and Rowlinson has proved to be quite valuable.[5,6] The model postulates a two-component system (the A and B particles), with no interaction between like particles and repulsive interactions between unlike particles. The original version had continuous translational coordinates and the repulsion between unlike particles was of the hard sphere type. It has been shown[7] rigorously that this model undergoes a demixing transition at a high activity (same activity for both components). Similar proofs have been given for extensions of the original model to lattice versions[8] and to models with soft repulsions between unlike particles.[9] The lattice version with infinite repulsion between nearby unlike particles is certainly the simplest member of this class of model.

In the initial description of the (continuum) model[5] it was pointed out that the two-component ($A-B$) system with hard core $A-B$ repulsion is isomorphic to a one-component system of "penetrable spheres." The particles of this one-component system may be thought of as the A particles of the $A-B$ system after the B particles have been rendered invisible. Thereby we must attach to each configuration of A particles a weight determined by the totality of B-particle configurations consistent with the A-particle configuration. The effective A-particle interactions induced in this way will be of short range but many-body in character.

Among these effective interactions, attractions will predominate, since clustering of A particles precludes a certain amount of repulsion with the (invisible) B particles. Widom has stated this principle more precisely[6] with the observation that the effective potential energy of an A-particle configuration is just the measure of the region excluded to the B particles, minus the total "volume" of the A particles each regarded as a sphere. So defined, the effective potential energy is always negative, but can be written as:

– (region common to A pairs)

+ (region common to A triples)

– (region common to four A particles)

.

.

.

Although it does not appear to have been done, there is no reason why the two-component-to-one-component transcription cannot also be carried out for lattice versions of the $A-B$ model. An additional incentive in the lattice case for the transcription is the possibility of then effecting a further transformation of the resulting one-component lattice gas to the corresponding Ising spin system. It turns out that both transformations are simpler than might be supposed and provide a collection of Ising systems, with many-spin interactions, rigorously known to undergo phase transitions.

LATTICE GAS

We briefly construct the partition functions for the two-component $A-B$ system and for the corresponding one-component system to show the relationship. For the former, each site of the lattice Λ may be occupied either by an A particle (activity z_A) or a B particle (activity z_B). Hence we have for the grand partition function

$$\Xi_{AB}(z_A, z_B)=\sum_{R\subset\Lambda} z_A{}^{|R|} \sum_{S\subset\Lambda\setminus R^+} z_B{}^{|S|}$$

where R^+ denotes the sites excluded to B particles by virtue of the A particles residing at the sites of $R\subset R^+$. The number of sites in set R is denoted $|R|$. Thus we have

$$\Xi_{AB}(z_A, z_B)=\sum_{R\subset\Lambda} z_A{}^{|R|}\,(1+z_B)^{|\Lambda|-|R^+|}. \tag{1}$$

The region R^+ will be the union of (possibly overlapping) exclusions "spheres" E_x, where $x\in R$ is a site occupied by an A particle. For the simplest, nearest-neighbor exclusion case each E_x consists of the $|E|$ sites consisting of x plus the $2d$ nearest-neighbor sites (d equals dimensionality, two or greater). The demixing transition, however, has been proven for any E_x that is simply connected, convex, and possessing a center of symmetry.[8]

Consider now a one-component lattice gas system on the same lattice Λ. For particles at sites $x \in R$ the energy H is given by:

$$-H(R) = \mu |R| + U(R)\epsilon$$

where μ is the one-body energy (chemical potential), ϵ is a parameter with units of energy, and $-U$ contains all interactions, two-body and higher. The one-component partition function is then

$$\Xi(z, \beta) = \sum_{R \subset \Lambda} z^{|R|} \exp[\beta\epsilon U(R)] \tag{2}$$

where $z = e^{\beta\mu}$ and β is the inverse temperature.

We now notice that the choice

$$U(R) = |R|\,|E| - |R^*| \tag{3}$$

and the identifications

$$\beta\epsilon = \ln(1 + z_B),$$
$$\beta[\mu + |E|\epsilon] = \ln z_A \tag{4}$$

render the partition functions (1) and (2) identical, apart from the analytic factor $\exp(\beta\epsilon|\Lambda|)$.

These identifications mean, of course, that the analyticity or nonanalyticity of the two models go hand-in-hand. Large positive z_A and z_B for the two-component model correspond to low temperature (large β) and not-too-negative μ for the one-component model. It is necessary that ϵ be positive (attraction between two particles) and that μ be greater than $-|E|\epsilon$.

TWO PROPERTIES OF THE PENETRABLE SPHERE LATTICE MODEL

We now show that the exclusion "sphere" E_x contains all the information needed to develop the configurational energy of either the one-component lattice gas or the corresponding Ising spin system. The many-body potentials are simple derivates of the geometry of E_x.

Lattice gas

The interaction term $U(R)$ can be expanded in many-body potentials as

$$U(R) = \sum_{P \subset R} \varphi(P)$$
$$\equiv \sum_{P \subset R} (-)^P \nu_P \tag{5}$$

where ν_P is the number of sites from which B particles are excluded by *each* A particle at the sites of P. The summations are over sets P of two or more sites. (Throughout, whenever a set occurs as an exponent, such as $(-)^P$ and 2^P, the meaning is the corresponding number of sites of the set, or $(-)^{|P|}$ and $2^{|P|}$.) The first property is stated as a lemma.

Lemma 1: If there is some point x such that $P \subset E_x$, where $|P| \geqslant 2$, then ν_P is equal to the number of translates of P also contained in E_x; otherwise $\nu_P = 0$. A translate of P is the set $\{T(x) | x \in P\}$, T being any translation of the lattice.

Proof: Suppose there is an x such that $P \subset E_x$. For any $y \in P$, $\dot{y}$ is in the exclusion sphere of an A particle at x. Consequently, x is in the exclusion sphere of any $y \in P$. Also if $T(P) \subset E_x$, then by the same reasoning $T^{-1}(x)$ is also in the exclusion sphere of all sites of P. If on the other hand there is no x such that $P \subset E_x$, then there is no site in the exclusion sphere of each point of P, and $\nu_P = 0$. This completes the proof.

Ising spin system

We introduce spin variables σ_x for each site x and $\sigma_R = \Pi_{x \in R}\sigma_x$, where $\sigma_x = \pm 1$. Now the (negative) configuration energy of the lattice Λ is expressed as

$$U = \sum_{R \subset \Lambda} J(R)\sigma_R + \text{const} \tag{6}$$

where the functions $J(R)$ are related[10] to the functions $\varphi(P)$ by

$$J(R) = \sum_{P \supset R} 2^{-P} \varphi(P). \tag{7}$$

Our second property is that these two set functions are in fact identical on sets of two or more sites, apart from a constant factor. This of course is not in general true for the lattice-gas-to-spin-system transcription.

Lemma 2: For $|R| \geqslant 2$,

$$J(R) = 2^{-E} \varphi(R). \tag{8}$$

Proof: In view of Eq. (7), the equation to establish is

$$\sum_{P \supset R} 2^{-P}(-)^P \nu_P = 2^{-E}(-)^R \nu_R.$$

If $\nu_R = 0$ and $|R| \geqslant 2$ there is nothing to prove, since ν_P must also vanish for any $P \supset R$, by Lemma 1. Suppose then that $\nu_R \neq 0$ and $R \subset E_{x_i}$ for $1 \leqslant i \leqslant \nu_R$, where the E_{x_i}'s are translates of each other. Consider first

$$\sum_{\substack{P \supset R \\ (E_{x_i} \supset P)}} 2^{-P}(-)^P = \sum_{n=0}^{|E|-|R|} \binom{|E| - |R|}{n} \left(-\frac{1}{2}\right)^{R+n}$$
$$= (1 - 1/2)^{E-R}(-1/2)^R$$
$$= 2^{-E}(-)^R. \tag{9}$$

Now we can write

$$\sum_{P \supset R} 2^{-P}(-)^P \nu_P = \sum_{i=1}^{\nu_R} \left\{ \sum_{\substack{Q \supset R \\ (E_{x_i} \supset Q)}} 2^{-Q}(-)^Q \right\}, \tag{10}$$

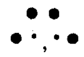
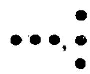
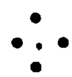
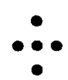
since (a) $\nu_P = 0$ unless P is contained in some translate E_{x_i} and (b) each set P occurs exactly ν_P times in the right-hand side of Eq. (10). But the inner summation is independent of i, so the right-hand side is equal to $2^{-E}(-)^R \nu_R$, by Eq. (9), and the proof is complete.

For convenience we can define the one body terms to be $\varphi(\{x\}) = \mu$ in the lattice gas case and $J(\{x\}) = h$ (magnetic field) in the spin case. Neither lemma applies for these independent field variables. The general relationship between the coefficients, Eq. (7), is still valid, however, and may also be used (in conjunction with the boundary conditions) to determine the constant in Eq. (6).

APPLICATIONS AND DISCUSSION

For the simplest application we consider the two-dimensional square lattice with nearest-neighbor exclusion between an A particle and a B particle. Each exclusion sphere E_x consists of five points.

TABLE I. Square lattice, penetrable sphere model.

Set P	$\varphi(P)$ Lattice gas	$J(P)\times 2^5$ Spin system
$\emptyset$	0	$16\mu+49\epsilon$
•	μ	$16\mu+75\epsilon$
	2ϵ	2ϵ
	2ϵ	2ϵ
	ϵ	ϵ
	$-\epsilon$	$-\epsilon$
, etc.	$-\epsilon$	$-\epsilon$
, etc.	$-\epsilon$	$-\epsilon$
, etc.	ϵ	ϵ
	ϵ	ϵ
	$-\epsilon$	$-\epsilon$

From Lemma 1 we find the φ coefficients in the energy expansion

$$-H(R)=\mu|R|+\sum_{\substack{P\subset R\\|P|\geq 2}}\varphi(P)$$

as shown in Table I. To convert to spin language and the $J(R)$ coefficients of Eq. (6) we can use Lemma 2 for $|R|\geq 2$. The results are also given in Table I. The one-body term (magnetic field) must be obtained by explicit summation of the inversion formula, Eq. (7), for $R=\{x\}$. The chemical potential $\mu=\varphi(\{x\})$ enters here and could be inhomogeneous rather than constant, as indicated in the result shown in Table I. In evaluating Eq. (7) for $J(\{x\})$ there is one term for each set P in the table—as well as sets obtained by rotation—each multiplied by $|P|$ to account for all possible locations of the site x in the set P.

The constant in Eq. (6) may also be computed using the inversion formula, Eq. (7), with the understanding that the constant is the contribution of $J(\emptyset)$, $\emptyset$ being the void set. If we invoke periodic boundary conditions $J(\emptyset)$ is proportional to $|\Lambda|$, and $J(\emptyset)/|\Lambda|$ is a summation over the sets P with nonzero $J(P)$, including rotations but not translations. This is how the entry $J(\emptyset)$ in Table I is obtained.

It is easily learned from the interaction coefficients that the total energy per site $-H/|\Lambda|$ for the (+) configuration (all $\sigma_x=+1$) is $\mu+4\epsilon$, which is also the total energy per site in the lattice gas language for the completely filled lattice. Similarly for the (−) configuration (all $\sigma_x=-1$) or the empty lattice gas, the total energy is zero.

We know from previous work on the $A-B$ lattice model[8] and the isomorphism developed above that this Ising spin system has a phase transition—with ferromagnetic pair and fourth-order interactions and antiferromagnetic third- and fifth-order interactions. It is interesting to interpret this system in light of the recent work of Pirogov and Sinai.[1] In their study of Ising spin systems it is shown that a phase transition occurs if the following conditions are met:

(a) pair interactions are ferromagnetic;

(b) many-body interactions are sufficiently weak;

(c) temperature is sufficiently low.

The actual meaning of (b) is the following. We first define a special value of the magnetic field $h^*=J(\{x\})$ by the condition that in the field h^* the energy of the (+) configuration is identical to that of the (−) configuration. For the special case of only pair interactions, $h^*=0$. We then define the Peierls contours in the usual manner[11]—as if the many-body interactions vanished. With nonvanishing many-body terms, however, the energy that can be attributed to the contour, or "fault line," is no longer the same in the two complementary situations: a "sea" of (−) spins with (+) boundary conditions on the outside of the system, and a "sea" of (+) spins with (−) boundary conditions. The meaning of condition (b) is that the many-body terms be sufficiently weak so that in the field h^* for *either* case, the sign of the contour energy is determined by the pair interaction, i.e., the contour energy is always repulsive.

For the present example, the special field h^* is obtained by setting $\mu=-4\epsilon$ [so that the energy of the (+) configuration and the (−) configuration is zero]. According to the entry for $J(\{x\})$ in Table I this means $h^*=(11/32)\epsilon$. It may now be shown directly that any contour has repulsive energy. For any configuration of spins the total energy may be written as:

$$U_\Lambda=\sum_{x\in\Lambda}W_x+J(\emptyset),$$

where W_x is a modified (negative) energy for the set of five spins: x plus its four nearest neighbors. The modification is that W_x is computed according to Table I except that (a) the nearest- and next-nearest-neighbor pair energy terms are halved—since each is included in two W_x's—and (b) the one-body term applies only to the central site. No other terms are overcounted. We must also remember to interpret $J(\{x\})$ as $h^*=(11/32)\epsilon$.

Thus we find, for example, that with all five spins +1 (or −1), W_x takes the value $+(15/32)\epsilon$. If $\sigma_x=-1$ and the other four are +1, then $W_x=-(9/32)\epsilon$. There are twelve distinct cases, apart from rotationally equivalent configurations, and all except the "pure" configurations +1 or −1 have $W_x<(15/32)\epsilon$.

The Pirogov—Sinai definition of the energy of a single contour in the field h^* is equivalent to

$$\sum_{x\in\Lambda}[W_x-(15/32)\epsilon],$$

which will always be negative for any contour. This means that condition (b) is met and that the conclusions of the Pirogov—Sinai development apply: a phase transition occurs at sufficiently low temperature (spin system), or at sufficiently high activity ($A-B$ system). This amounts to an alternate proof of the phase transition first proved in Ref. 8.

ACKNOWLEDGMENT

Most of this work was done during a visit at Belfer Graduate School, Yeshiva University. Many helpful discussions with J. L. Lebowitz are gratefully acknowledged.

*Supported in part by National Science Foundation Grant No. GP 33535X.

[1]S.A. Pirogov and J.G. Sinai, "Phase Transitions of First Order for Small Perturbation of the Ising Model," preprint (to be published in Russian).
[2]D. Merlini, A. Hintermann, and C. Gruber, "Phase Transitions for an Ising Model with 3-body Interactions," preprint; D. Merlini, "Symmetry Breaking and Peierls Argument for an Ising Model with Three Body Interactions," preprint.
[3]R.J. Baxter and F.Y. Wu, Phys. Rev. Lett. 31, 1294 (1973); R.J. Baxter and F.Y. Wu, "Ising Model on a Triangular Lattice with Three-Spin Interactions. I. The Eigenvalue Equation," preprint; R.J. Baxter, "Ising Model on a Triangular Lattice with Three-Spin Interactions. II. Free Energy and Correlation Length," preprint.
[4]D. Merlini and C. Gruber, J. Math. Phys. 13, 1814 (1972); A. Hintermann and D. Merlini, Phys. Lett. 41A, 208 (1972).
[5]B. Widom and J.S. Rowlinson, J. Chem. Phys. 52, 1670 (1970).
[6]B. Widom, J. Chem. Phys. 54, 3950 (1971).
[7]D. Ruelle, Phys. Rev. Lett. 27, 1040 (1971).
[8]J.L. Lebowitz and G. Gallavotti, J. Math. Phys. 12, 1129 (1971).
[9]J.L. Lebowitz and E.H. Lieb, Phys. Lett. 39A, 98 (1972); L.K. Runnels and B.C. Freasier, Phys. Rev. A8, 2126 (1973).
[10]C.M. Fortuin, P.W. Kasteleyn, and J. Ginibre, Commun. Math. Phys. 22, 89 (1971).
[11]R.B. Griffiths in *Phase Transitions and Critical Phenomena* edited by C. Domb and M.S. Green (Academic, New York, 1972), Vol. I, Chap. 2; G. Gallavotti, Riv. Nuovo Cimento 2, 133 (1972).

A scattering problem for a straight line segment

K. B. Ranger

Department of Mathematics, University of Toronto, Toronto, Canada M5S 1A1
(Received 9 July 1973; revised manuscript received 22 January 1974)

An explicit analytical solution is found for the scattering potential produced by a straight line segment. Integral transformations are introduced which map the mixed boundary value problem into one which can be solved by standard Green's function methods.

1. INTRODUCTION

There have been several attempts to calculate the scattering potential by a straight line segment or needle, see Jones.[1] These attempts are in general unsatisfactory because the needle has been considered as the limiting case of a prolate ellipsoid of revolution which produces a discontinuous solution as the boundary reduces from three-dimensional to one-dimension. For example Bowman, Senior and Uslenghi[2] have described the solution of the Dirichlet problem for a prolate spheroid. This is found as an eigenfunction expansion in spheroidal wave functions and it can be shown that when the spheroid degenerates into a needle the total potential exterior to the straight line segment is the incident wave whilst on the needle the potential vanishes in accordance with the sound soft boundary condition. This result is typical of similarity solutions, that is solutions obtained by "separation of variables" appropriate to the coordinate curves describing the boundary. A similar result was obtained for the electrostatic potential problem in Ref. 3.

In this paper the scattered potential for an incident plane wave advancing parallel to the axis is formulated as a three part mixed boundary value problem in which the potential and its derivative normal to the axis are prescribed at different parts of the axis. At infinity the Sommerfeld radiation condition is to be satisfied. The problem as stated is well posed because using an integral transformation described in Sec. 3, the axially symmetric problem can be mapped in a one—one manner into a two part mixed boundary value problem for the two-dimensional reduced wave equation in which the potential is prescribed on the positive axis and the normal derivative on the negative axis. At infinity the two-dimensional radiation condition is satisfied. This problem can be solved in terms of standard Green's function methods and it can be shown that the scattered potential is finite at the tips of the needle, but its derivatives are infinite. The radiation condition at infinity is satisfied provided the needle is *finite* in length.

It is pointed out the Sec. 2 provides a rationale for the integral transformations used in this paper. These are clearly related to the Mehler–Dirichlet integral representations for the Legendre polynomial. Section 3 generalizes the result of Sec. 2 for the situation in which there are distributions of singularities on the axis.

Finally the continuous potential which is finite at the tips appears to be unique, but if the finiteness condition is relaxed there are infinitely many solutions for the scattered potential which are infinite at the tips. It is conceivable that the solution described here is in fact the limiting case of a solid cylinder finite in length.

2. INTEGRAL TRANSFORMATIONS FOR THE REDUCED WAVE EQUATION

The integral representations used in this paper can be motivated by some elementary considerations. Let $x = r\cos\theta$, $\rho = r\sin\theta$ define spherical polar coordinates; then separable solutions of the axially symmetric reduced wave equation

$$(L_1 + k^2)W = 0\,, \tag{1}$$

where L_1 is the axially symmetric Laplacian defined by

$$L_1 \equiv \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)\,, \tag{2}$$

are given by

$$W^{(n)}(r,\theta) = r^{-1/2}[A_n H^{(1)}_{n+1/2}(kr) + B_n H^{(2)}_{n+1/2}(kr)]P_n(\cos\theta)\,. \tag{3}$$

$H^{(1)}_{n+1/2}(kr)$, $H^{(2)}_{n+1/2}(kr)$ are Hankel functions of fractional order and $P_n(\cos\theta)$ is the Legendre polynomial. Now the Mehler–Dirichlet formulas for the Legendre polynomial are

$$P_n(\cos\theta) = \frac{2^{1/2}}{\pi}\int_0^\theta \frac{\cos(n+\frac{1}{2})\lambda\,d\lambda}{(\cos\lambda - \cos\theta)^{1/2}} = \frac{2^{1/2}}{\pi}\int_\theta^\pi \frac{\sin(n+\frac{1}{2})\lambda\,d\lambda}{(\cos\theta-\cos\lambda)^{1/2}}\,, \tag{4}$$

and it follows that

$$W^{(n)}(r,\theta) = r^{-1/2}\int_0^\theta \frac{u_n(r,\lambda)d\lambda}{(\cos\lambda-\cos\theta)^{1/2}} = r^{-1/2}\int_\theta^\pi \frac{v_n(r,\lambda)d\lambda}{(\cos\theta-\cos\lambda)^{1/2}}\,, \tag{5}$$

where

$$u_n + iv_n = \frac{2^{1/2}}{\pi}[A_n H^{(1)}_{n+1/2}(kr) + B_n H^{(2)}_{n+1/2}(kr)]e^{i(n+1/2)\lambda}\,. \tag{6}$$

Moreover, it is natural to generalize Eqs. (5) by writing

$$W = \sum_n W^{(n)}, \quad (u+iv) = \sum_n (u_n + iv_n)\,; \tag{7}$$

then

$$W(r,\theta) = r^{-1/2}\int_0^\theta \frac{u(r,\lambda)d\lambda}{(\cos\lambda-\cos\theta)^{1/2}} = r^{-1/2}\int_\theta^\pi \frac{v(r,\lambda)d\lambda}{(\cos\theta-\cos\lambda)^{1/2}}\,, \tag{8}$$

where

$$u + iv = \frac{2^{1/2}}{\pi}\sum_n [A_n H^{(1)}_{n+1/2}(kr) + B_n H^{(2)}_{n+1/2}(kr)]e^{i(n+1/2)\lambda} \tag{9}$$

satisfies the equation

$$(L_0+k^2)(u+iv)=0 \tag{10}$$

and the operator

$$L_0 \equiv \frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial^2}{\partial\lambda^2}.$$

Thus the integral representations (or transformations) expressed by Eq. (8) map solutions of the two-dimensional Helmholtz equation into solutions of the axially symmetric form of the same equation. Since the kernels in (8) are of Abel type, the inverse transformations are also of a simple form but these are not required for the present paper. The integral transformations (8) possess the following properties on the axis:

$$\begin{aligned}
&\text{(i)}\quad v(r,0)=\frac{\partial}{\partial\lambda}v(r,\lambda)\big|_{\lambda=\pi}=0,\\
&\text{(ii)}\quad u(r,\pi)=\frac{\partial}{\partial\lambda}u(r,\lambda)\big|_{\lambda=0}=0,\\
&\text{(iii)}\quad W(r,\pi)=(2r)^{-1/2}\pi v(r,\pi),\\
&\text{(iv)}\quad W(r,0)=(2r)^{-1/2}\pi u(r,0),\\
&\text{(v)}\quad \frac{\partial W}{\partial\theta}\bigg|_{\theta=\pi}=\left(\frac{2}{r}\right)^{1/2}\frac{\partial v}{\partial\lambda}\bigg|_{\lambda=\pi},\\
&\text{(vi)}\quad \frac{\partial W}{\partial\theta}\bigg|_{\theta=0}=\left(\frac{2}{r}\right)^{1/2}\frac{\partial u}{\partial\lambda}\bigg|_{\lambda=0}.
\end{aligned} \tag{11}$$

These relations are certainly satisfied for $u+iv=u_n+iv_n$ or a finite combination of the u_n+iv_n. In fact, when there are no sources on the axis $\theta=0$, or π, the expressions in Eq. (11) [(v) and (vi)] are all zero. If there are sources on the axis, or part of the axis, as will be the case when the axis forms part of the boundary it is necessary to proceed in the manner described in the next section.

3. VALIDITY OF THE INTEGRAL TRANSFORMATIONS WHEN THERE ARE SOURCES ON THE AXIS

From Eq. (8) it is natural to define

$$W(r,\theta)=r^{-1/2}\int_\theta^\pi \frac{v(r,\lambda)d\lambda}{(\cos\theta-\cos\lambda)^{1/2}}, \tag{12}$$

where $v(r,\lambda)$ is continuous with its partial derivatives in $r>0$, $0\le\lambda<\pi$ and $v(r,0)=0$. This latter condition ensures convergence of the integral (12) as $\theta\to 0$. It is convenient to introduce new variables defined by $\theta'=\pi-\theta$, $\lambda'=\pi-\lambda$, $u(r,\lambda')=v(r,\pi-\lambda)$, and $W_1(r,\theta')=W(r,\pi-\theta')$; then Eq. (12) can be written as

$$W_1(r,\theta')=r^{-1/2}\int_0^{\theta'}\frac{u(r,\lambda')d\lambda'}{(\cos\lambda'-\cos\theta')^{1/2}}, \tag{13}$$

where it is assumed that $u(r,\lambda')$ is even in λ' [see [9]]. Again it may be verified by replacing the lower limit in (13) by ϵ and considering the limit $\epsilon\to 0\pm$, that if $(L_1+k^2)W_1=0$, then

$$\int_{0\pm}^{\theta'}\frac{(L_0+k^2)u\,d\lambda'}{(\cos\lambda'-\cos\theta')^{1/2}}=\pm\frac{1}{r^2(1-\cos\theta')^{1/2}}\frac{\partial u}{\partial\lambda'}\bigg|_{\lambda'=0+}, \tag{14}$$

since

$$\frac{\partial u}{\partial\lambda'}\bigg|_{\lambda'=0\pm}=\pm\frac{\partial u}{\partial\lambda'}\bigg|_{\lambda'=0+}.$$

Equation (14) is equivalent to

$$(L_0+k^2)u=\pm\frac{2}{r^2}\frac{\partial u}{\partial\lambda'}\bigg|_{\lambda'=0\pm}\delta(\lambda'), \tag{15}$$

where $\delta(\lambda')$ is the Dirac delta function. To determine which sign is appropriate for consistency set $\bar U=e^{-ikx'}u$ and consider the divergence theorem applied to the vector $\mathbf{F}=\operatorname{grad}\bar U+2ik\bar U\hat\imath$ in the (ρ,ϕ) plane over the region

$$r\le\rho\le r+\delta r,\quad -\delta\lambda'\le\phi\le\delta\lambda'.$$

This yields

$$\begin{aligned}
\delta r\int_{-\delta\lambda'}^{\delta\lambda'}&\left(L_0+2ik\frac{\partial}{\partial x'}\right)\bar U r\,d\phi\\
&=\int_{-\delta\lambda'}^{\delta\lambda'}\left(\frac{\partial\bar U}{\partial\rho}+2ik\bar U\cos\phi\right)\bigg|_{\rho=r+\delta r}(r+\delta r)d\phi\\
&\quad-\int_{-\delta\lambda'}^{\delta\lambda'}\left(\frac{\partial\bar U}{\partial\rho}+2ik\bar U\cos\phi\right)\bigg|_{\rho=r}r\,d\phi\\
&\quad+\delta r\left\{\left[\frac{1}{r}\frac{\partial\bar U}{\partial\phi}-2ik\bar U\sin\phi\right]_{\phi=\delta\lambda'}\right.\\
&\quad\left.-\left[\frac{1}{r}\frac{\partial\bar U}{\partial\phi}-2ik\bar U\sin\phi\right]_{\phi=-\delta\lambda'}\right\}.
\end{aligned} \tag{16}$$

Hence as $\delta r,\ \delta\lambda'\to 0$

$$\lim_{\delta\lambda'\to 0}\int_{-\delta\lambda'}^{\delta\lambda'}\left(L_0+2ik\frac{\partial}{\partial x'}\right)\bar U r\,d\phi=\frac{2}{r}\frac{\partial\bar U}{\partial\lambda'}\bigg|_{\lambda'=0+}, \tag{17}$$

or equivalently

$$L_0+2ik\frac{\partial}{\partial x'}\quad\bar U=\frac{2}{r}\frac{\partial u}{\partial\lambda'}\bigg|_{\lambda'=0+}\delta(\lambda'). \tag{18}$$

In terms of u, (18) yields

$$(L_0+k^2)u=\frac{2}{r}\frac{\partial u}{\partial\lambda'}\bigg|_{\lambda'=0+}\delta(\lambda'). \tag{19}$$

Thus when $\partial u/\partial\lambda'|_{\lambda'=0+}\ne 0$, the positive sign in Eq. (15) should be chosen. (See note added in proof.) In terms of $v(r,\lambda)$ it follows that a solution of

$$(L_0+k^2)v=-\frac{2}{r}\frac{\partial v}{\partial\lambda}\bigg|_{\lambda=\pi}\delta(\pi-\lambda), \tag{20}$$

with $v(r,0)=0$, maps into a solution of $(L_1+k^2)W=0$, by the integral transformation (12). Also from (12) it follows that

$$W(r,\pi)=(2r)^{-1/2}\pi v(r,\pi),\quad \frac{\partial W}{\partial\theta}\bigg|_{\theta=\pi}=\left(\frac{2}{r}\right)^{1/2}\frac{\partial v}{\partial\lambda}\bigg|. \tag{21}$$

4. SCATTERING BY A FINITE LINE SEGMENT

Consider an incident plane wave represented by $W_i=e^{-ikx}$ advancing on the needle $\rho=0$, $-1\le x\le 0$, or $\theta=\pi$, $0\le r\le 1$. The total potential $W=W_i+W_s$, where W_s is the scattered potential. The boundary condition (sound soft) requires $W=0$, on the needle, or

$$W_s = -e^{+ikr}, \quad \theta=\pi, \quad 0\le r\le 1. \tag{22}$$

In view of symmetry

$$\begin{aligned} \frac{\partial W_s}{\partial\theta} &= 0, \quad \theta=0, \ r>0, \\ &= 0, \quad \theta=\pi, \ r>1, \end{aligned} \tag{23}$$

and finally the scattered wave should satisfy the Sommerfeld radiation condition at infinity. Now W_s is a solution of the reduced wave equation (1) and a suitable representation for the solution is expressed by equation (12), that is

$$W_s(r,\theta) = r^{-1/2}\int_\theta^\pi \frac{v(r,\lambda)d\lambda}{(\cos\theta-\cos\lambda)^{1/2}}, \tag{24}$$

where $v(r,0)=0$, and $v(r,\lambda)$ is a solution of the two-dimensional reduced wave equation in $r>0$, $0\le\lambda\le\pi$. The boundary conditions for v are

$$\begin{aligned} v(-r,0) &= -\frac{(2r)^{1/2}}{\pi}e^{+ikr}, \quad \lambda=\pi, \quad 0\le r\le 1, \\ \frac{\partial v}{\partial\lambda} &= 0, \quad \lambda=\pi, \quad r>1. \end{aligned} \tag{25}$$

The radiation condition at infinity is satisfied for W_s if the two-dimensional form of the condition is satisfied by v. It is not clear that the condition $\partial W/\partial\theta=0$, on $\theta=0$ is satisfied. To show this write $t=\cos\lambda$, $\beta=\cos\theta$ and consider

$$\frac{\partial W}{\partial\theta} = -\sin\theta\frac{\partial}{\partial\beta}\int_{-1}^{\beta}\frac{v\,dt}{(1-t^2)^{1/2}(\beta-t)^{1/2}}, \tag{26}$$

$$\begin{aligned} &= -\sin\theta\frac{\partial}{\partial\beta}\int_{\alpha}^{\beta}\frac{v\,dt}{(1-t^2)^{1/2}(\beta-t)^{1/2}} \\ &\quad -\sin\theta\frac{\partial}{\partial\beta}\int_{-1}^{\alpha}\frac{v\,dt}{(1-t^2)^{1/2}(\beta-t)^{1/2}}, \end{aligned} \tag{27}$$

where $1\ge\beta>\alpha>-1$. Since

$$\begin{aligned} &\frac{\partial}{\partial\beta}\int_\alpha^\beta\frac{v\,dt}{(1-t^2)^{1/2}(\beta-t)^{1/2}} \\ &= \int_\alpha^\beta\frac{\partial}{\partial t}\left\{\frac{v}{(1-t^2)^{1/2}}\right\}\frac{dt}{(\beta-t)^{1/2}} + \frac{(v)_{t=\alpha}}{(1-\alpha^2)^{1/2}(\beta-\alpha)^{1/2}}; \end{aligned} \tag{28}$$

and

$$\frac{\partial}{\partial\beta}\int_{-1}^{\alpha}\frac{v\,dt}{(1-t^2)^{1/2}(\beta-t)^{1/2}} = -\frac{1}{2}\int_{-1}^{\alpha}\frac{v\,dt}{(1-t^2)^{1/2}(\beta-t)^{3/2}}, \tag{29}$$

and

$$\frac{\partial}{\partial\beta}\int_{-1}^{\alpha}\frac{v\,dt}{(1-t^2)^{1/2}(\beta-t)^{1/2}} = \frac{1}{2}\int_{-1}^{\alpha}\frac{v\,dt}{(1-t^2)^{1/2}(\beta-t)^{3/2}}, \tag{30}$$

both of which remain finite as $\beta\to1$ or $\theta\to0$, it follows that $\partial W/\partial\theta\to0$, as $\theta\to0$, provided

$$\lim_{\lambda\to0}\frac{\partial}{\partial\lambda}\left(\frac{v}{\sin\lambda}\right) \tag{31}$$

exists. This will be fulfilled by the condition $v(r,0)=0$.

It is convenient to pose the problem for v in translated coordinates (r_1,λ_1) defined by

$$r\sin\lambda = r_1\sin\lambda_1, \quad r\cos\lambda+1 = r_1\cos\lambda_1. \tag{32}$$

Regarding v as a function of (r_1,λ_1), the boundary value problem can be stated as follows:

$$(L_0+k^2)v=0, \quad L_0\equiv\frac{\partial^2}{\partial r_1^2}+\frac{1}{r_1}\frac{\partial}{\partial r_1}+\frac{1}{r_1^2}\frac{\partial^2}{\partial\lambda_1^2} \tag{33}$$

except on $\lambda_1=0$, $0\le r_1\le1$, and

$$v=g(r_1), \quad \lambda_1=0, \tag{34}$$

where

$$\begin{aligned} g(r_1) &= -\frac{2^{1/2}}{\pi}(1-r_1)^{1/2}e^{+ik(1-r_1)}, \quad 0\le r_1\le1 \\ &= 0, \quad r_1>1. \end{aligned} \tag{35}$$

Also $\partial v/\partial\lambda_1=0$, $\lambda_1=\pi$, and v satisfies the Sommerfeld condition at infinity. The mixed problem for v can be solved in terms of the Green's function formula

$$v = -\frac{1}{2\pi}\int_0^1\left(\frac{v}{\rho_1}\frac{\partial G}{\partial\phi_1}\right)_{\phi_1=0}d\rho_1, \tag{36}$$

where the Green's function $G(\rho_1,\phi_1|r_1,\lambda_1)$ is defined by

$$(L_0+k^2)\{G(\rho_1,\phi_1|r_1,\lambda_1)\} = \frac{2\pi}{\rho_1}\delta(\rho_1-r_1)\delta(\phi_1-\lambda_1) \tag{37}$$

subject to the mixed conditions

$$G=0, \quad \phi_1=0, \quad \frac{\partial G}{\partial\phi_1}=0, \quad \phi_1=\pi, \tag{38}$$

and G satisfies the two-dimensional Sommerfeld condition as $\rho_1\to\infty$; G is found in the appendix and is represented by the eigenfunction expansions

$$G = 2\pi i\sum_{n=0}^{\infty}H^{(1)}_{n+1/2}(kr_1)[J_{n+1/2}(k\rho_1)]\sin(n+\tfrac12)\phi_1\sin(n+\tfrac12)\lambda_1, \tag{39}$$

for $\rho_1<r_1$, and

$$G = 2\pi i\sum_{n=0}^{\infty}H^{(1)}_{n+1/2}(k\rho_1)[J_{n+1/2}(kr)]\sin(n+\tfrac12)\phi_1\sin(n+\tfrac12)\lambda_1, \tag{40}$$

for $\rho_1>r_1$. Thus v may be expressed in the form

$$\begin{aligned} v &= \frac{2i}{\pi\sqrt2}\sum_{n=0}^{\infty}(n+\tfrac12)H^{(1)}_{n+1/2}(kr_1)\sin(n+\tfrac12)\lambda_1 \\ &\quad\times\int_0^1\frac{(1-\rho_1)^{1/2}}{\rho_1}e^{+ik(1-\rho_1)}[J_{n+1/2}(k\rho_1)]d\rho_1, \end{aligned} \tag{41}$$

for $r_1>1$; and for $0<r_1<1$

$$\begin{aligned} v &= \frac{2i}{\pi\sqrt2}\sum_{n=0}^{\infty}(n+\tfrac12)H^{(1)}_{n+1/2}(kr_1)\sin(n+\tfrac12)\lambda_1\int_0^{r_1}\frac{(1-\rho_1)^{1/2}}{\rho_1} \\ &\quad\times e^{+ik(1-\rho_1)}[J_{n+1/2}(k\rho_1)]d\rho_1 \\ &\quad+\frac{2i}{\pi\sqrt2}\sum_{n=0}^{\infty}(n+\tfrac12)[J_{n+1/2}(kr_1)]\sin(n+\tfrac12)\lambda_1 \end{aligned}$$

$$\times \int_{r_1}^{1} \frac{(1-\rho_1)^{1/2}}{\rho_1} e^{+ik(1-\rho_1)} H^{(1)}_{n+1/2}(k\rho_1)\, d\rho_1 .$$

For large r_1, $r_1 \sim r$, and

$$v \sim \frac{2i e^{ikr}}{(\pi^3 kr)^{1/2}} \sum_{n=0}^{\infty} e^{-i\pi(n+1)/2}(n+\tfrac{1}{2})\sin(n+\tfrac{1}{2})\lambda_1$$

$$\times \int_0^1 \frac{(1-\rho_1)^{1/2}}{\rho_1} e^{+ik(1-\rho_1)} [J_{n+1/2}(k\rho_1)]\, d\rho_1 \qquad (43)$$

so that

$$W_s \sim \frac{2ie^{ikr}}{(\pi^3 k)^{\frac{1}{2}} r} \sum_{n=0}^{\infty} e^{-i\pi(n+1)/2}(n+\tfrac{1}{2}) \int_\theta^\pi \frac{\sin(n+\frac{1}{2})\lambda_1\, d\lambda}{(\cos\theta-\cos\lambda)^{1/2}}$$

$$\times \int_0^1 \frac{(1-\rho_1)^{1/2}}{\rho_1} e^{+ik(1-\rho_1)} [J_{n+1/2}(k\rho_1)]\, d\rho_1 . \qquad (44)$$

Thus the scattered wave does in fact satisfy the radiation condition at infinity. From the manner in which $v(r,\lambda)$ and hence $W(r,\theta)$ has been constructed it follows that $W(r,\theta)$ is continuous at the tips of the needle $r=0$, and $r=1$, $\theta=\pi$. It is difficult from the eigenfunction expansion for the Green's function to determine the nature of the singularity in the derivative of W at the tips and it is useful to this end to write the Green's function as an integral representation. This can be achieved by a slight modification of the method given by Clemmow.[4] In this case G is given by

$$G = e^{ikR'}\left(\int_{-m'}^{\infty} \frac{e^{i\mu^2}d\mu}{(\mu^2+2kR')^{1/2}} - \int_{m'}^{\infty} \frac{e^{i\mu^2}d\mu}{(\mu^2+2kR')^{1/2}}\right) - e^{ikR}\left(\int_{-m}^{\infty} \frac{e^{i\mu^2}d\mu}{(u^2+2kR)^{1/2}} - \int_{m}^{\infty} \frac{e^{i\mu^2}d\mu}{(\mu^2+2kR)^{1/2}}\right), \qquad (45)$$

where

$$m = 2\left(\frac{k\rho_1 r_1}{R_1+R}\right)^{1/2}\cos\left(\frac{\phi_1-\lambda_1}{2}\right), \qquad m' = 2\left(\frac{k\rho_1 r_1}{R_1+R^1}\right)^{1/2}\cos\left(\frac{\phi_1+\lambda_1}{2}\right) \qquad (46)$$

and

$$R = [\rho_1^2 + r_1^2 - 2\rho_1 r_1 \cos(\phi_1-\lambda_1)]^{1/2}, \quad R' = [\rho_1^2 + r_1^2 - 2\rho_1 r_1 \cos(\phi_1+\lambda_1)]^{1/2}, \quad R_1 = \rho_1 + r_1 . \qquad (47)$$

This solution for the Green's function is obtained by adding to Clemmow's solution the Green's function at the image source point so that the derivative cancels out on the negative axis $\phi_1=\pi$. From (44) it is readily shown that as $\rho_1 \to 0$, G is continuous but the derivatives contain singularities like $\rho_1^{-1/2}$. From this it can be shown in a straightforward way that W_s is in fact finite at the tips of the needle but the derivatives have square root singularities. The solution constructed is thus the only one possible if W_s is to be finite at the tips of the needle, but there are clearly infinitely many solutions to the problem if this condition is relaxed at the tips. For example, the functions

$$G_n = H^{(1)}_{n+1/2}(k\rho_1 | H^{(1)}_{n+1/2}(kr_1)\sin(n+\tfrac{1}{2})\phi_1 \sin(n+\tfrac{1}{2})\lambda_1), \qquad (48)$$

$n=0,1,2,\ldots$, can be added to the eigenfunction expansions for the Green's function. G_n clearly satisfy the inner and outer boundary conditions but have singular behavior near $\rho_1=0$.

Note added in proof: It turns out that (14) implies (15) only when the positive sign is chosen, and to determine the scattered potential W_s accurately first write $W_s = W_s^0 + W_s^1$; W_s^0 is the scattered potential found in the paper and W_s^1 is defined by

$$W_s^1 = r_2^{-1/2} \int_0^{\theta_2} \frac{u(r_2,\lambda)\, d\lambda}{(\cos\lambda - \cos\theta_2)^{1/2}},$$

where $r_2 \sin\theta_2 = r\sin\theta$, $1 + r\cos\theta = r_2\cos\theta_2$. On the axis $u(r_2,\pi)=0$, $r_2>0$, $u(r_2,0)=0$, $0\le r_2\le 1$, $\lambda=0$, $\partial u/\partial\lambda = 0$, $r_2>1$, $\lambda=0$, and at infinity u satisfies the radiation condition. In order that W_s be a solution of the reduced wave equation it is found that

$$\int_0^{\theta_2} \frac{(L_0+k^2)u\, d\lambda}{(\cos\lambda-\cos\theta_2)^{1/2}} + \frac{(u_\lambda)_0}{r_2^2(1-\cos\theta_2)^{1/2}} = \left(\frac{r_2}{r}\right)^{1/2} \frac{(v_\lambda)\pi}{r^2(1+\cos\theta)^{1/2}} .$$

This Abel equation is readily inverted and an inhomogeneous Helmholtz equation is obtained for u which can be solved in a straightforward manner subject to the above boundary conditions.

[1] D. S. Jones, *The Theory of Electromagnetism* (Pergamon, New York, 1964), p. 614.
[2] J. J. Bowman, T. B. A. Senior, and P. L. E. Uslenghi, *Electromagnetic and Acoustic Scattering by Simple Shapes* (North-Holland, Amsterdam, 1969), Chap. 11, p. 422.
[3] K. B. Ranger, "The solution of mixed boundary value problems involving a straight line segment," M.R.C. Technical Summary Report, 1324, University of Wisconsin, January 1973.
[4] P. C. Clemmow, Quart. J. Mech. Appl. Maths. 3, 377 (1950).

Expansion-free electromagnetic solutions of the Kerr–Schild class

George Debney

Department of Mathematics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061
(Received 15 February 1974)

Starting with the general Kerr–Schild form of the metric tensor, $ds^2=\eta+l\otimes l$ (where l is null and η is flat space–time), a study is made for those solutions of the Einstein–Maxwell equations in which l is geodesic, shear-free, and expansion-free. It is shown that all resulting solutions must be of Petrov type [4] or type [–] and the Maxwell field must be null. Because of the expansion-free assumption there exist flat and conformally flat gauge conditions on all metrics in this class; i.e., there exist metrics of this Kerr–Schild form which are flat (or conformally flat) but are not Lorentz-related. A method is given for obtaining meaningful solutions to the field equations with the latter gauge equivalence class removed. A simple example of a radiative field of type [4] along a line singularity exhibits how solutions in this class may be generated.

1. INTRODUCTION

This work concerns itself primarily with radiation solutions of the Einstein–Maxwell equations for a metric in Kerr–Schild form. The assumption that the special null congruence is expansion-free bridges the gap left between those electromagnetic solutions of Debney, Kerr, and Schild[1] (hereafter called DKS) and the general expansion-free cases studied by Kundt.[2]

The original Kerr–Schild paper[3] concerned itself with vacuum space–time metrics which have the form (where η is the metric for flat space–time and l is tangent to a null congruence)

$$ds^2=\eta+l\otimes l. \tag{1.1}$$

It assumed the Einstein vacuum field equations plus the condition that l have nonvanishing expansion ($\Leftrightarrow\rho\neq 0$, in Newman–Penrose notation,[4] hereafter called N–P). The general properties possessed by these vacuum space–times include: (a) They are all algebraically special; (b) l is a degenerate principal null direction for the Riemannian curvature tensor (and is both geodesic and shear-free); and (c) the Schwarzschild and Kerr[5] classes of solutions fall into this category.

Later, metrics of the same form satisfying the Einstein–Maxwell source-free equations were studied in DKS, again assuming the condition that l have nonvanishing expansion but also assuming l to be geodesic. The properties implied in general about these space–times turned out to be: (a) They are all algebraically special; (b) l is a principal null direction for the Weyl conformal tensor and l is shear-free; and (c) they contain the Reissner–Nordström and Kerr–Newman[6] classes of solutions.

Before the Kerr-Schild studies appeared, Kundt[2] considered all vacuum, and certain nonvacuum, space–times which possessed a geodesic and shear-free null congruence l with vanishing complex expansion. These fell into two general categories determined by whether the rotation (τ in N–P) of l vanishes or not.[7] It was concluded that such space–times fell into all algebraically special categories and the cases with vanishing rotation were the (type [4]) "pp waves." The general name of "expansion-free radiation fields" characterizes the whole expansion-free class.

It is the purpose of the present paper to examine in more detail the expansion-free constraint on the special null vector in the Kerr–Schild metrics and its implications in the context of Einstein–Maxwell theory. Such studies in vacuum cases have been treated by H. Urbantke[9] and by the author.[10]

Section 2 contains the algebraic preliminaries. Here, also, one assumes the expansion-free condition for l and its alignment with the electromagnetci field. Appendix A supplements this section with the computations to derive the field equations, proving along the way that the Petrov type must be [4] or [–].

Section 3 provides a better coordinate system (at least when the rotation $X\neq 0$) in which to solve the field equations. The work in Appendix B exhibits the flat and conformally flat "gauge" conditions on the metric, providing a way of obtaining type [4] solutions modulo these additional gauge terms. The problem here is that "the" flat background is not unique in these expansion-free cases: some of "$l\otimes l$" can go into "η" to form another flat background, not related to η by a Lorentz transformation. A method given for removing these solutions from the picture allows meaningful examples to be chosen. Section 4 exhibits, as an example of vanishing rotation, an electromagnetic field which falls off radially in cylindrical coordinates and propagates along the z axis; it possesses true singularities on this axis.

2. PRELIMINARIES

The Kerr–Schild[3,1] form for the metric on a four-dimensional Lorentz manifold (C^∞) of signature (+++–) is stated simply as $ds^2=\eta+l\otimes l$, where η is the metric for a flat (Minkowski) background and l^μ is the tangent to a congruence of null curves. Notice that writing

$$g_{\mu\nu}=\eta_{\mu\nu}+l_\mu l_\nu \tag{2.1}$$

tells us that l^μ is null with respect to both $g_{\mu\nu}$ and $\eta_{\mu\nu}$. The field equations for a Einstein–Maxwell space–time

$$R_{\mu\nu}=-8\pi T_{\mu\nu} \tag{2.2}$$

plus the source-free Maxwell equations for the electromagnetic field $F_{\mu\nu}$

$$F^{\mu\nu}{}_{;\nu}=0=F_{[\mu\nu;\sigma]} \tag{2.3}$$

must also be satisfied. [$R_{\mu\nu}\equiv R^\alpha{}_{\mu\nu\alpha}$ is the Ricci tensor,

$T_{\mu\nu} \equiv (\frac{1}{4}\pi)(F_{\mu\alpha}F_{\nu}{}^{\alpha} - \frac{1}{4}g_{\mu\nu}F_{\alpha\beta}F^{\alpha\beta})$ is the electromagnetic stress—energy tensor, and the Riemannian curvature tensor $R_{\mu\nu\rho\alpha}$ satisfies $V_{\sigma;\mu\nu} - V_{\sigma;\nu\mu} = R^{\alpha}{}_{\sigma\mu\nu}V_{\alpha}$ for any vector field V.]

The approach used to find solutions to (2.1)—(2.3) makes use of a complex null tetrad. [This set of four independent vector fields forms a basis (or "frame") in which all geometric objects and equations may be written. Components with respect to such a basis are indicated by Latin indices, whereas components with respect to a coordinate basis are indicated by Greek indices.] The contravariant and covariant components of the tetrad are expressed through

$$e_a \equiv e_a{}^{\mu}\frac{\partial}{\partial x^{\mu}}, \quad \epsilon^b \equiv \epsilon^b{}_{\mu}dx^{\mu}, \tag{2.4}$$

respectively. Since the two systems of vector fields in (2.4) are vector space duals they satisfy, by definition,

$$e_a{}^{\mu}\epsilon^a{}_{\nu} = \delta^{\mu}_{\nu}, \quad e_a{}^{\mu}\epsilon^b{}_{\mu} = \delta^b_a$$

$$g_{\mu\nu} = \epsilon^a{}_{\mu}e_{a\nu}, \quad g_{ab} = e_{a\mu}e_b{}^{\mu}.$$

The "complex null" part comes from the additional relations resulting from a formal complexification, where "bar" denotes complex conjugation:

$$\mathbf{e}_2 = \bar{\mathbf{e}}_1, \quad \mathbf{e}_3 = \bar{\mathbf{e}}_3, \quad \mathbf{e}_4 = \bar{\mathbf{e}}_4.$$

In such a system the metric takes the form

$$ds^2 = g_{\mu\nu}dx^{\mu}dx^{\nu} = 2\epsilon^1\epsilon^2 + 2\epsilon^3\epsilon^4 = g_{ab}\epsilon^a\epsilon^b. \tag{2.5}$$

If $\{x, y, z, t\} = \{x^{\mu}\}$ are Cartesian coordinates in the background Minkowski space, define complex null coordinates $\{\zeta, \bar{\zeta}, u, v\}$ by

$$\sqrt{2}\,\zeta \equiv x + iy, \quad \sqrt{2}\,\bar{\zeta} \equiv x - iy,$$

$$\sqrt{2}\,u \equiv z + t, \quad \sqrt{2}\,v \equiv z - t.$$

Then $\eta_{\mu\nu}dx^{\mu}dx^{\nu} = 2d\zeta d\bar{\zeta} + 2dudv$. By letting h be an unknown scalar and $l^{\mu} \equiv (2h)^{1/2}k^{\mu}$ the metric (2.1) becomes

$$ds^2 = g_{\mu\nu}dx^{\mu}dx^{\nu} = 2d\zeta d\bar{\zeta} + 2dudv + 2h(k_{\mu}dx^{\mu})^2. \tag{2.6}$$

A choice of tetrad is made in terms of these coordinates:

$$\epsilon^1 = d\zeta - Ydv, \quad \epsilon^2 = \bar{\epsilon}^1 = d\bar{\zeta} - \bar{Y}dv,$$

$$\epsilon^3 \equiv k_{\mu}dx^{\mu} = du + \bar{Y}d\zeta + Yd\bar{\zeta} - Y\bar{Y}dv, \tag{2.7}$$

$$\epsilon^4 = dv + h\epsilon^3.$$

As in DKS, the unknown complex function $Y(x^{\mu})$ may be introduced to express any null vector field $k_{\mu}dx^{\mu}$ in Minkowski space. The contravariant tetrad $\{\mathbf{e}_a\}$ is then computed to be

$$\mathbf{e}_1 = \partial_{\zeta} - \bar{Y}\partial_u, \quad \mathbf{e}_2 = \bar{\mathbf{e}}_1 = \partial_{\bar{\zeta}} - Y\partial_u,$$

$$\mathbf{e}_3 = \partial_u - h\mathbf{e}_4, \tag{2.8}$$

$$\mathbf{e}_4 = \partial_v + Y\partial_{\zeta} + \bar{Y}\partial_{\bar{\zeta}} - Y\bar{Y}\partial_u \doteq k^{\mu}\partial_{\mu}.$$

Denoting the operation of $\mathbf{e}_a$ on a scalar ϕ by $\mathbf{e}_a(\phi) \equiv \phi_{,a}$ it is clear from a study of Ricci rotation coefficients for k_{μ} that $Y_{,4} = 0 \Leftrightarrow k_{\mu}$ is geodesic and $Y_{,2} = 0 \Leftrightarrow k_{\mu}$ is shear-free. Furthermore $Y_{,1} = z$, the *complex expansion* of k_{μ} ("ρ" in Newman—Penrose[4]). (See Appendix A.)

The first assumption made for this system is that k_{μ} *is a principal null direction for* $F_{\mu\nu}$; i.e., $F_{\mu\nu}k^{\mu} = \alpha k_{\nu}$ for some scalar α. This is essentially the same as that made in DKS because k^{μ} is geodesic and shear-free if and only if it is a principal null direction for $F_{\mu\nu}$ (see Appendix A). Consequently, the scalar h in (2.6) may be chosen so that k^{μ} is tangent to an affinely parametrized congruence of null geodesics. The congruence is also shear-free.

The next assumption, $z = 0$, restricts the study to those Kerr—Schild electromagnetic solutions for which the vector l^{μ} in the metric is *expansion-free*. As shown in Appendix A, this produces the general theorem: *All source-free vacuum or Einstein—Maxwell fields of the Kerr—Schild class* $ds^2 = \eta + l \otimes l$, *where* l^{μ} *is an expansion-free principal null direction for the Maxwell field* $F_{\mu\nu}$ *are of Petrov type* [4]. Furthermore, the electromagnetic field $F_{\mu\nu}$ is *null*; i.e., $F_{\mu\nu}F^{\mu\nu} = 0 = F^{*}_{\mu\nu}F^{\mu\nu}$

3. NEW COORDINATES: SOLUTIONS OF THE FIELD EQUATIONS

It is shown in Appendix A that the vector fields $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_4\}$ satisfy $[\mathbf{e}_1, \mathbf{e}_2] = [\mathbf{e}_1, \mathbf{e}_4] = [\mathbf{e}_2, \mathbf{e}_4] = 0$. Frobenius' theorem suggests that one may find new coordinates, say $\{\alpha, \bar{\alpha}, \rho, w\} = \{x^{\mu'}\}$, for which

$$\mathbf{e}_1 = \partial_{\alpha}, \quad \mathbf{e}_2 = \partial_{\bar{\alpha}}, \quad \mathbf{e}_4 = \partial_v. \tag{3.1}$$

Indeed this may be accomplished by setting

$$\alpha \equiv \zeta - Yv, \quad \bar{\alpha} \equiv \bar{\zeta} - \bar{Y}v, \quad w \equiv v,$$
$$\rho \equiv u + \bar{Y}\zeta + Y\bar{\zeta} - Y\bar{Y}v. \tag{3.2}$$

Notice that $\alpha\bar{\alpha} + \rho v = \zeta\bar{\zeta} + uv$ but that $\{x^{\mu}\} \to \{x^{\mu'}\}$ is not Lorentz except when Y is a constant. However $\rho = \eta_{\mu\nu}x^{\mu}k^{\nu} = g_{\mu\nu}x^{\mu}k^{\nu}$ so that $\rho = \mathbf{k}\cdot\mathbf{P}$, where $\mathbf{P}$ is a position vector in the original "background" Minkowski space.

The tetrads expressed in the new coordinates become[10]

$$\epsilon^1 = d\alpha + vX\epsilon^3, \quad \epsilon^2 = d\bar{\alpha} + v\bar{X}\epsilon^3,$$
$$\epsilon^3 = r^{-1}d\rho, \qquad \epsilon^4 = dv + h\epsilon^3 \tag{3.3}$$

(with rotation $X \equiv Y_{,3}$ and $r \equiv 1 + \alpha\bar{X} + \bar{\alpha}X$) and

$$\mathbf{e}_1 = \partial_{\alpha}, \quad \mathbf{e}_2 = \partial_{\bar{\alpha}}, \quad \mathbf{e}_4 = \partial_v,$$
$$\mathbf{e}_3 = -v(X\partial_{\alpha} + \bar{X}\partial_{\bar{\alpha}}) + r\partial_{\rho} - h\partial_v. \tag{3.4}$$

The field equations [Eqs. (A14a)—(A14d)] may be written as

$$h_{vv} = 0, \quad F_v = 0, \tag{3.5a}$$

$$h_vX - h_{\bar{\alpha}v} = 0, \tag{3.5b}$$

$$h_{\alpha\bar{\alpha}} - h_{\alpha}X - h_{\bar{\alpha}}\bar{X} = -2F\bar{F}, \tag{3.5c}$$

$$F_{\bar{\alpha}} - XF = 0. \tag{3.5d}$$

The geodesic, shear-free, and expansion-free conditions on k^{μ} clearly imply through (A4) and (A10) that $Y = Y(\rho)$.

Hence, $Y_{,3}=X=r(dY/d\rho)=rY'$ so that

$$r^{-1}=1-\bar{Y}'\alpha-Y'\bar{\alpha}. \tag{3.6}$$

Equation (3.5a) implies that $F=F(\alpha,\bar{\alpha},\rho)$ and that

$$h=a(\alpha,\bar{\alpha},\rho)v+g(\alpha,\bar{\alpha},\rho), \tag{3.7a}$$

where a and g are real-valued functions of their arguments. Equation (3.5b) implies

$$a=rA(\rho), \tag{3.7b}$$

where A is arbitrary. The function a does not enter into (3.5c) so that (3.5c) and (3.5d) reduce to, respectively,

$$(gr^{-1})_{\alpha\bar{\alpha}}=-2r^{-1}F\bar{F}, \tag{3.7c}$$

$$(Fr^{-1})_{\bar{\alpha}}=0. \tag{3.7d}$$

Although it is not obvious, an investigation similar to that of Appendix B gives us that a conformally flat solution must necessarily have $X\equiv 0$, $F=F(\rho)$, and $g=D(\rho)\alpha\bar{\alpha}$. However, these only make up a proper subclass of solutions where $F=F(\rho)$. The latter solutions, even though $X=0$ is implied here too, are not all conformally flat since they admit a more general function $g(\alpha,\bar{\alpha},\rho)$ in the metric.

It is shown in Appendix B that any function (3.7a) having the form

$$\hat{h}=r[A(\rho)v+K(\rho)\alpha+\bar{K}(\rho)\bar{\alpha}+L(\rho)], \tag{3.8}$$

where A, K, L are arbitrary functions of ρ, results in the metric

$$ds^2=\eta+2\hat{h}(k\otimes k) \tag{3.9}$$

being a representation of flat space–time with no field F; i.e., any such $\hat{h}$ term in general makes no contribution to the curvature. Consequently, there exists a coordinate system in which (3.9) may be written manifestly as flat space. Instead of looking for this coordinate transformation we alternatively take the approach that solutions to (3.7c) and (3.7d) *not* containing $\hat{h}=av+\hat{g}$, where $\hat{g}=r[K\alpha+\bar{K}\bar{\alpha}+L]$ for $X\neq 0$ or $\hat{g}=D\alpha\bar{\alpha}+K\alpha+\bar{K}\bar{\alpha}+L$ for $X=0$, are to be regarded as meaningful for the present purposes.

4. A SPECIAL SOLUTION WITH $X=0$ (pp WAVE)

Since $X=0$ implies the system $\{\alpha,\bar{\alpha},\rho,v\}=\{\zeta,\bar{\zeta},u,v\}$, let

$$F_{31}=F=-\gamma(u)/\zeta, \tag{4.1}$$

where $\gamma(u)$ is arbitrary and real. Then by the discussion in Section 3 the metric is not conformally flat. Furthermore,

$$F\bar{F}=\gamma^2/\zeta\bar{\zeta}.$$

By solving (3.7c) the nonflat part of the metric becomes

$$g(\zeta,\bar{\zeta},u)=-2\gamma^2(u)\,|\ln(\zeta)\,|^2. \tag{4.2}$$

Hence, the metric is

$$ds^2=2d\zeta d\bar{\zeta}+2dudv-4\gamma^2|\ln(\zeta)|^2du^2, \tag{4.3}$$

with a curvature singularity at $\zeta=0$ (i.e., $x=y=0$) and possibly elsewhere if $\gamma(u)\to\infty$.

Let $\sqrt{2}\,\zeta=Re^{i\theta}$. Then $R=x^2+y^2$ and cylindrical coordinates $\{R,\theta,z,t\}$ are established. From the relation $F_{\mu\nu}dx^\mu\wedge dx^\nu=F_{ab}\epsilon^a\wedge\epsilon^b$ one obtains the electromagnetic field tensor

$$F_{\mu\nu}dx^\mu\wedge dx^\nu=\sqrt{2}\,(2\gamma/R)[dR\wedge dz+dR\wedge dt]. \tag{4.4}$$

Classically, the electric field $\mathbf{E}$ is along R and the magnetic field $\mathbf{H}$ is along θ; propagation takes place along the z axis and the intensity falls off as R^{-1} in the radial direction. The amplitude $\gamma(z+t)$ determines the longitudinal behavior of the wave.

Notice that many choices for the starting point (4.1) exist and it is possible to study many more type [4] cases when $X=0$ as long as $g_{\zeta\zeta}\neq 0$ and the conformally flat $\hat{g}=D\zeta\bar{\zeta}$ is avoided. Curvature singularities will most likely be determined by the singularities inherent in the choice of F.

ACKNOWLEDGMENTS

The author is particularly grateful to R.J. Adler for many stimulating and clarifying discussions, particularly with regard to similar vacuum cases and comparison with the linearized theory. In addition, H. Urbantke first made the author aware of the nonuniqueness of the flat background, which led to the above discussion of these gauge conditions.

APPENDIX A: THE STRUCTURE EQUATIONS FOR THE KERR-SCHILD CLASS

The following discussion first makes use only of the Kerr–Schild form of the metric. Further assumptions come in later as: (a) l^μ is a principal null direction for the electromagnetic field $F_{\mu\nu}$ and (b) the complex expansion of l^μ vanishes (i.e., $z\equiv 0$), respectively.

The first structure equations are stated concisely as $d\epsilon^a=\Gamma^a{}_{bc}\epsilon^b\wedge\epsilon^c$, which are true for any torsion-free connection $\Gamma^a{}_{bc}$ and any tetrad (or "frame") $\{\epsilon^a\}$ locally throughout the space–time. The condition that the connection be the "metric connection" of Levi-Civita is the equivalent to stating $\Gamma_{bac}=-\Gamma_{abc}$, where $\Gamma_{abc}\equiv g_{am}\Gamma^m{}_{bc}$. $\equiv g_{am}\Gamma^m{}_{bc}$.

For the particular metric (2.1) and tetrad (2.7) the $d\epsilon^3$ equations are written as

$$\begin{aligned} d\epsilon^3&=\tfrac{1}{2}(\Gamma_{4ab}-\Gamma_{4ba})\epsilon^a\wedge\epsilon^b\\ &=(Y_{,1}-\bar{Y}_{,2})\epsilon^1\wedge\epsilon^2+\bar{Y}_{,3}\epsilon^3\wedge\epsilon^1+Y_{,3}\epsilon^3\wedge\epsilon^2\\ &\quad+\bar{Y}_{,4}\epsilon^4\wedge\epsilon^1+Y_{,4}\epsilon^4\wedge\epsilon^2. \end{aligned}$$

Equating coefficients and defining $z\equiv\Gamma_{241}$, $X\equiv Y_{,3}$, we obtain

$$\begin{aligned} &\Gamma_{414}=-Y_{,4}, \qquad \Gamma_{424}=-Y_{,4}, \qquad \Gamma_{434}=0,\\ &z-\bar{z}=Y_{,1}-\bar{Y}_{,2}, \quad \Gamma_{423}=-X-\Gamma_{342}, \quad \Gamma_{413}=-\bar{X}-\Gamma_{341}. \end{aligned} \tag{A1}$$

The $d\epsilon^4$ equations become

$$d\epsilon^4=\tfrac{1}{2}(\Gamma_{3ab}-\Gamma_{3ba})\epsilon^a\wedge\epsilon^b=dh\wedge\epsilon^3+hd\epsilon^3$$
$$=h(z-\bar{z})\epsilon^1\wedge\epsilon^2+(-h_{,4})\epsilon^3\wedge\epsilon^4+(h\bar{X}-h_{,1})\epsilon^3\wedge\epsilon^1$$
$$+(hX-h_{,2})\epsilon^3\wedge\epsilon^2+h\bar{Y}_{,4}\epsilon^4\wedge\epsilon^1+hY_{,4}\epsilon^4\wedge\epsilon^2.$$

Equating coefficients allows one to write

$$\Gamma_{342}-\Gamma_{324}=hY_{,4},\qquad \Gamma_{341}-\Gamma_{314}=h\bar{Y}_{,4},\quad \Gamma_{343}=h_{,4},$$
$$\Gamma_{312}-\Gamma_{321}=h(z-\bar{z}),\quad \Gamma_{313}=h_{,1}-h\bar{X},\quad \Gamma_{323}=h_{,2}-hX. \tag{A2}$$

Finally the $d\epsilon^2$ equations are written as

$$d\epsilon^2=(1/2)(\Gamma_{1ab}-\Gamma_{1ba})\epsilon^a\wedge\epsilon^b$$
$$=-(\bar{X}+h\bar{Y}_{,4})\epsilon^3\wedge\epsilon^4+(-h\bar{Y}_{,1})\epsilon^3\wedge\epsilon^1+(-hY_{,2})\epsilon^3\wedge\epsilon^2$$
$$+\bar{Y}_{,1}\epsilon^4\wedge\epsilon^1+Y_{,2}\epsilon^4\wedge\epsilon^2.$$

Upon equating coefficients we obtain (through the use of complex conjugation)

$$\bar{Y}_{,1}=-\Gamma_{411},\quad Y_{,2}=-\Gamma_{422},\quad h\bar{Y}_{,1}=\Gamma_{311},\quad hY_{,2}=\Gamma_{322},$$
$$\bar{Y}_{,2}=-\Gamma_{124}+\bar{z},\quad Y_{,1}=\Gamma_{124}+z,$$
$$h\bar{Y}_{,2}=\Gamma_{123}+\Gamma_{312},\quad hY_{,1}=-\Gamma_{123}+\Gamma_{321},$$
$$\Gamma_{121}=\Gamma_{122}=0,\qquad \Gamma_{413}=\Gamma_{314}-\bar{X}-h\bar{Y}_{,4},$$
$$\Gamma_{423}=\Gamma_{324}-X-hY_{,4}. \tag{A3}$$

Putting together the information in (A1)—(A3) results in

$$z=Y_{,1},\qquad \bar{z}=\bar{Y}_{,2},\qquad \Gamma_{321}=h\bar{z},\qquad \Gamma_{312}=hz,$$
$$\Gamma_{422}(\text{shear of } k^\mu)=-Y_{,2},\quad \Gamma_{322}=hY_{,2},$$
$$\Gamma_{124}=\Gamma_{121}=\Gamma_{122}=0,\qquad \Gamma_{123}=-h(z-\bar{z}),$$
$$\Gamma_{424}=-Y_{,4},\qquad \Gamma_{342}=hY_{,4},\qquad \Gamma_{324}=0=\Gamma_{344},$$
$$-\Gamma_{423}=hY_{,4}+X,\qquad \Gamma_{323}=h_{,2}-h\bar{X}. \tag{A4}$$

The second structure equations $d\Gamma^a{}_b+\Gamma^a{}_m\wedge\Gamma^m{}_b=\frac{1}{2}R^a{}_{bcd}\epsilon^c\wedge\epsilon^d$ contain implicitly the field equations $R_{ab}=-8\pi T_{ab}$ (where $\Gamma^a{}_b\equiv\Gamma^a{}_{bc}\epsilon^c$ are the connection 1-forms). However, one obtains in particular the relationship

$$-\tfrac{1}{2}R_{\mu\nu}k^\mu k^\nu=-\tfrac{1}{2}R_{44}=R_{4241}=-2h|Y_{,4}|^2. \tag{A5}$$

Hence, making the first assumption that k^μ is a principal null direction for $F_{\mu\nu}$(and therefore $T_{\mu\nu}$) implies $Y_{,4}\equiv 0$. But $Y_{,4}=0$ implies $\Gamma_{424}=\Gamma_{414}=0$, which is equivalent to stating that k^μ is a geodesic. Hence, k^μ is a principal null direction for $F_{\mu\nu}$ if and only if k^μ is a geodesic. Our assumption here (as in the earlier DKS paper) then produces the simplifications

$$\Gamma_{414}=\Gamma_{424}=\Gamma_{341}=\Gamma_{342}=0,\quad \Gamma_{423}=-X,\quad \Gamma_{413}=-\bar{X}. \tag{A6}$$

Furthermore, this forces R_{ab} partly into a canonical form

$$R_{42}=R_{41}=R_{44}=0=R_{11}=R_{22} \tag{A7}$$

since $F_{14}=F_{24}=0$ now, as well. The connection 1-forms of interest reduce to

$$\Gamma_{42}=-dY,$$
$$\Gamma_{12}+\Gamma_{34}=[h_{,4}-h(z-\bar{z})]\epsilon^3, \tag{A8}$$
$$\Gamma_{31}=h\bar{\sigma}\epsilon^1+hz\epsilon^2+(h_{,1}-h\bar{X})\epsilon^3,$$

where $\sigma\equiv-\Gamma_{422}$.

The $d\Gamma_{42}+\Gamma_{4m}\wedge\Gamma^m{}_2=\frac{1}{2}R_{42ab}\epsilon^a\wedge\epsilon^b$ equation results in

$$R_{4242}=0\Rightarrow C^{(5)}=0\quad(\psi_0 \text{ in Newman—Penrose}),$$
$$R_{4212}=R_{4234}=0\Rightarrow C^{(4)}=0\quad(\psi_1 \text{ in N—P}),$$

implying that the space—time is algebraically special and that k^μ is a principal null direction for the Weyl conformal curvature tensor. Also $2R_{4231}=C^{(3)}$ (ψ_2 in N—P) so that

$$-C^{(3)}=2z[h_{,4}-(z-\bar{z})h]. \tag{A9}$$

The field equation $-\frac{1}{2}R_{22}=R_{4232}=0$ is not identically satisfied; it becomes

$$\sigma[h_{,4}-(z-\bar{z})h]=0.$$

As in DKS, $\sigma\neq 0$ gives rise to a contradiction (i.e., vanishing electromagnetic field, algebraically special space—time, $\sigma\neq 0$ are incompatible). Therefore $\sigma=0$ must result. Note that we have derived the relation that k^μ must be geodesic ($\Gamma_{424}=0$) and shear-free ($\Gamma_{422}=0$).

At this stage the work of DKS and that discussed here differ in that the assumption $z=0$ (k^μ is expansion-free) is imposed. From (A9) it is clear that $C^{(3)}=0$ so that Petrov types [3, 1], [4], or [–] are the only possibilities. We choose to exclude the conformally flat cases (type [–]) since these have been solved completely in the Einstein—Maxwell context (see, for example, Cahen and Leroy[11]).

The special relations imposed above imply the following relations from (A4) (omitting complex conjugates):

$$\Gamma_{122}=\Gamma_{121}=\Gamma_{123}=\Gamma_{124}=0=\Gamma_{421}=\Gamma_{422}=\Gamma_{424}=\Gamma_{414},$$
$$\Gamma_{311}=\Gamma_{312}=\Gamma_{314}=0=\Gamma_{341}=\Gamma_{344}, \tag{A10}$$
$$\Gamma_{343}=h_{,4},\quad \Gamma_{313}=h_{,1}-h\bar{X},\quad \Gamma_{423}=-X.$$

In tetrad form the Maxwell equations are written with $z=0$:

$$(F_{12}+F_{34})_{,1}-2F_{31,4}=0,$$
$$(F_{12}+F_{34})_{,2}-2(F_{12}+F_{34})X=0,$$
$$(F_{12}+F_{34})_{,3}+2F_{31,2}-2F_{31}X=0, \tag{A11}$$
$$(F_{12}+F_{34})_{,4}=0.$$

Since F_{12} is pure imaginary and F_{34} is real, the field equations $R_{12}=-8\pi T_{12}$ and $R_{34}=-8\pi T_{34}$ become

$$|F_{12}|^2+|F_{34}|^2=0,$$

$$h_{,44}=-|F_{12}|^2-|F_{34}|^2$$

so that $h_{,44}=0=F_{12}=F_{34}$. Therefore the only nonzero components of F_{ab} are F_{31} and F_{32}. Study of canonical forms for F_{ab} reveals that F_{ab} is a *null* electromagnetic field. Furthermore, the only nonzero component of T_{ab} is $T_{33}=(4\pi)^{-1}F_{31}F_{32}=(4\pi)^{-1}|F_{31}|^2$. Hence the equations left are (letting $F_{31}\equiv F$)

$$h_{,44}=0=F_{,4}, \quad F_{,2}-XF=0. \tag{A12}$$

The equations $R_{31}=R_{32}=0$ come from the $d(\Gamma_{12}+\Gamma_{34})$ structure equations to give

$$h_{,4}\bar{X}-h_{,41}=C^{(2)} \quad (\psi_3 \text{ in N—P}),$$

$$h_{,4}X-h_{,42}=0.$$

The reality of h implies therefore that $C^{(2)}=0$, the Petrov type is [4], and radiation solutions are to be expected.

The $d\Gamma_{31}$ equation completes the set of field equations. These are given by

$$(hX-h_{,2})_{,4}=0,$$

$$h_{,1}X-hX\bar{X}+(h\bar{X})_{,2}-h_{,12}=\tfrac{1}{2}R_{33}=-2|F|^2,$$

$$h_{,1}\bar{X}-h\bar{X}^2+(h\bar{X})_{,1}-h_{,11}=\tfrac{1}{2}C^{(1)} \quad (\psi_4 \text{ in N—P}).$$

The general relationships $\phi_{,ab}-\phi_{,ba}=\phi_{,m}(\Gamma^m{}_{ab}-\Gamma^m{}_{ba})$ are true in any tetrad system for any scalar ϕ. These correspond to the Lie brackets $[e_a,e_b]=C^m{}_{ab}e_m$ on the basis fields $\{e_a\}$. Applying this to e_1, e_2, e_4 and using (A10), we find that

$$[e_1,e_2]=[e_1,e_4]=[e_2,e_4]=0; \tag{A13}$$

i.e., $\phi_{,12}=\phi_{,21}$, $\phi_{,14}=\phi_{,41}$, and $\phi_{,24}=\phi_{,42}$. In further applying this to the function Y in the metric one obtains algorithms for derivatives of X:

$$X_{,1}=X\bar{X}, \quad X_{,2}=X^2, \quad X_{,4}=0.$$

The simplifications above reduce the set of field equations for the expansion-free Kerr—Schild case with electromagnetism to

$$h_{,44}=0, \tag{A14a}$$

$$h_{,4}X-h_{,24}=0=h_{,4}\bar{X}-h_{,14}, \tag{A14b}$$

$$h_{,12}-h_{,1}X-h_{,2}X=-2F\bar{F}, \tag{A14c}$$

$$F_{,2}-XF=0=F_{,4}, \tag{A14d}$$

with the [–] case excluded through the constraint $C^{(1)}\neq 0$; i.e.,

$$h_{,11}-2h_{,1}\bar{X}\neq 0. \tag{A15}$$

The only relationships to come from the Bianchi identities turn out to be

$$C^{(1)}{}_{,2}-C^{(1)}X+R_{33,1}-R_{33}\bar{X}=0,$$

$$C^{(1)}{}_{,4}=0=R_{33,4},$$

which are all identically satisfied.

APPENDIX B: THE GAUGE CONDITIONS FOR FLAT SPACE

In the $\{\alpha,\bar{\alpha},\rho,v\}$ coordinates the general solution for the function h in the metric is given by the real-valued functions

$$h=av+g, \quad a=a(\alpha,\bar{\alpha},\rho), \quad g=g(\alpha,\bar{\alpha},\rho). \tag{B1}$$

Furthermore, $a=rA(\rho)$, where $A(\rho)$ is arbitrary and $r^{-1}=1-\bar{Y}'\alpha-Y'\bar{\alpha}$. Thus far the relations (A14a) and (A14b) are the ones satisfied.

Consider the cases with *no* electromagnetism present ($F=0$). Then the equation (A14c) is the condition that

$$(gr^{-1})_{\alpha\bar{\alpha}}=0. \tag{B2}$$

Hence, $gr^{-1}=S(\alpha,\rho)+\bar{S}(\bar{\alpha},\rho)$, where S is arbitrary. However, there will be a large amount of repetition of solutions because of a gauge condition implicit in (A15); i.e., the set of all functions $\hat{g}$ for which

$$\hat{g}_{\alpha\alpha}-2\hat{g}_{\alpha}\bar{X}=0 \quad (\Leftrightarrow C^{(1)}=0) \tag{B3}$$

give a solution $h=av+\hat{g}$ which is necessarily flat space, completely independent of "a". [When inserting $h=av+g$ into (A15) one finds that the "av" terms cancel each other identically in all cases, leaving only a constraint on g.] It would make things much simpler if somehow one might "divide out" these flat-space solutions.

The condition(B3) is the gauge condition for flat space. The general solution for all such functions $\hat{g}$ satisfying (B2) and (B3) is

$$\hat{g}=r[K(\rho)\alpha+\bar{K}(\rho)\bar{\alpha}+L(\rho)], \tag{B4}$$

where K, L are arbitrary and L is real. This works in a vacuum ($F=0$), but it also holds true in the present electromagnetic case. This is most easily seen by observing that (A14c) is the differential equation

$$(gr^{-1})_{\alpha\bar{\alpha}}=-2F\bar{F}. \tag{B5}$$

Hence, any function $\hat{g}$ added to g will contribute nothing to the field equation (B5) since $(\hat{g}r^{-1})_{\alpha\bar{\alpha}}=0$.

In summary we have shown that the metric

$$\begin{aligned} ds^2 &= 2\epsilon^1\epsilon^2+2\epsilon^3\epsilon^4 \\ &= 2(d\alpha+vXr^{-1}d\rho)(d\bar{\alpha}+v\bar{X}r^{-1}d\rho)+2r^{-1}d\rho dv \\ &\quad +2r^{-1}(Av+K\alpha+\bar{K}\bar{\alpha}+L)d\rho^2 \end{aligned}$$

is a representation of *flat space* (where A,K,L are arbitrary functions of the coordinate ρ). It illustrates a peculiarity of these expansion-free wave solutions in that "the" flat background is by no means unique. In fact, a comparison of results in a vacuum with the linearized theory, even in the $X=0$ case, is doomed to frustration until one writes his particular metric in the correct coordinate system. (Such a coordinate system is usually *not* manifestly $\eta+l\otimes l$. See, for example, Misner, Thorne, and Wheeler[12].)

[1]G. Debney, R. P. Kerr, and A. Schild, J. Math. Phys. **10**, 1842 (1969).
[2]W. Kundt, Z. Phys. **163**, 77 (1961).
[3]R. P. Kerr and A. Schild, in *Atti del Convegno sulla Relatività Generale: Problemi dell'Energia e Onde Gravitazionali*, edited by G. Barbera (Firenze, 1965).
[4]E. T. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).
[5]R. P. Kerr, Phys. Rev. Lett. **11**, 237 (1963).
[6]E. T. Newman, E. Couch, R. Chinnapared, A. Exton, A. Prakash, and R. Torrence, J. Math. Phys. **6**, 918 (1965).
[7]This is a meaningless distinction for algebraically special cases of $\rho \neq 0$ since the rotation can be transformed to zero in such event. This property was evident in the proof of the Goldberg—Sachs theorem,[8] and possibly earlier.
[8]J. N. Goldberg and R. K. Sachs, Acta Phys. Polon. **22**, 13-23 (1962).
[9]H. K. Urbantke, Acta Phys. Austr. **35**, 396 (1972).
[10]G. Debney, Lett. Nuovo Cimento 8, 337 (1973). [After that article went to press, the author discovered an unfortunate algebraic error in his last tetrad, resulting in some incorrect field equations. The correction is most easily made by putting $F \equiv 0$ in the present work. The case $X = 0$ certainly does not give flat space—time, in that event, but gives the pp waves.]
[11]M. Cahen and J. LeRoy, J. Math. Mech. **16**, 501 (1966).
[12]C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973), Sec. 35.9.

A simple proof of certain FKG inequalities*

James L. Monroe

Department of Physics, Northwestern University, Evanston, Illinois 60201
(Received 29 January 1974)

We present a proof of a set of FKG inequalities, namely, $\langle \Pi_{j=1}^N(\rho_i)^{\nu_i}\ \Pi_{j=1}^N(\rho_j)^{\delta_j}\rangle - \langle \Pi_{i=1}^N(\rho_i)^{\nu_i}\rangle\ \langle \Pi_{j=1}^N(\rho_j)^{\delta_j}\rangle \geq 0$. The proof is obtained using Gaussian random variables and holds for systems whose Hamiltonian contains a positive quadratic form and one-body interactions on which no restrictions are placed.

In a recent paper[1] we presented a new proof of the first and second Griffiths–Kelly–Sherman (hereafter GKS) inequalities[2] for a class of general spin Ising systems. The proof was based on the rewriting of the correlation function averages in terms of averages of Gaussian random variables. Subsequent to the initial proof of the GKS inequalities by Griffiths, a different set of inequalities, the FKG inequalities, were proven by Fortuin, Kasteleyn, and Ginibre.[3] Proofs of these inequalities being rather abstract, we offer here a simple derivation of a subclass of the FKG inequalities, the method of proof being similar to the approach used in Ref. 1 to prove the GKS inequalities.

Let Ω be a finite lattice whose sites will be designated by $i=1,2,\ldots,N$. On each site there is a spin variable $s_i = p, p-2, \ldots, -p+2, -p$ or equivalently an occupation number variable $\rho_i = 0, 1, \ldots, p$, where

$$\rho_i = (s_i + p)/2. \tag{1}$$

The spin Hamiltonian of the system is

$$H(\{s\}) = -\tfrac{1}{2}\sum_{i\neq j} J(i,j)s_i s_j - \sum_{i=1}^N \sum_{\substack{m\\ \text{odd}}} h_m(i)s_i{}^m - \sum_{i=1}^N \sum_{\substack{n\\ \text{even}}} \mu_n(i)s_i{}^n, \tag{2}$$

where m and n are respectively, odd and even integers, $J(i,j) \geq 0$, and $J(i,j) = J(j,i)$. The spin Hamiltonian $H(\{s\})$ can be transformed into an equivalent lattice gas Hamiltonian $H(\{\rho\})$ by use of Eq. (1). Also, any other function $f(\{s_i\})$ of the spin variables can be written as a function, $f(\{\rho_i\})$, of the occupation number variables. The thermal averages of such functions are

$$\begin{aligned}\langle f(\{s_i\})\rangle &= Z_N^{-1} \sum_{\{s\}} f(\{s_i\}) \exp[-H(\{s\})]\\ &= Z_N^{-1} \sum_{\{\rho\}} f(\{\rho_i\}) \exp[-H(\{\rho\})] = \langle f(\{\rho_i\})\rangle\end{aligned} \tag{3}$$

with

$$Z_N = \sum_{\{s\}} \exp[-H(\{s\})] = \sum_{\{\rho\}} \exp[-H(\{\rho\})], \tag{4}$$

where $\{s\}$ is a configuration of the set of N spins and we have for simplicity set $\beta = 1$. We shall be interested in thermal averages of products of the ρ_i's given by $\Pi_{i=1}^N(\rho_i)^{\nu_i}$, where ν_i is a multiplicity function assigning a nonnegative integer to each site i.

The method of Gaussian random variables is based on the identity[4]

$$\begin{aligned}\exp[\tfrac{1}{2}\sum_{k,l} \xi_k \alpha_{kl}\xi_l] &= (2\pi)^{-N/2}(\det\alpha)^{-1/2}\\ &\times \int\cdots\int \exp[-\tfrac{1}{2}\sum_{k,l} x_k(\alpha^{-1})_{kl}x_l + \sum_j \xi_j x_j]\prod_{i=1}^N dx_i,\end{aligned} \tag{5}$$

valid for any symmetric, real, and positive definite matrix α and for any N complex variables ξ_k. The sign of $(\det\alpha)^{-1/2}$ is to be chosen positive. The right-hand side of Eq. (5) can be considered as the average $\langle \exp\sum_{j=1}^N x_j\xi_j\rangle_{\text{Av}\alpha}$, where $\langle\ \rangle_{\text{Av}\alpha}$ denotes the average with respect to the probability density

$$W_N(\mathbf{x}) = (2\pi)^{-N/2}(\det\alpha)^{-1/2}\exp[-\tfrac{1}{2}\sum_{kl} x_k(\alpha^{-1})_{kl}x_l], \tag{6}$$

where $\mathbf{x}$ is the vector with components x_j. The identity (5) can be used to rewrite the Boltzmann factor, $\exp(-H)$, by identifying the variable ξ_j with the spin variables s_j and forming a matrix $J=\alpha$ with off-diagonal elements $J(i,j)$ and all diagonal elements equal to a number $J_0 = J(i,i)$ large enough to guarantee that J is positive definite. The Boltzmann factor with the Hamiltonian (2) is then

$$\begin{aligned}\exp(-H) = \Big\langle \exp\sum_{k=1}^N [x_k s_k &+ \sum_{\substack{m\\ \text{odd}}} h_m(k)s_k{}^m + \sum_{\substack{n\\ \text{even}}} \mu_n(k)s_k{}^n\\ &- \tfrac{1}{2}J_0 s_k{}^2]\Big\rangle_{\text{Av}J}.\end{aligned} \tag{7}$$

We then prove the following.

Theorem: Given the system described above, one has the inequalities

$$\left\langle \prod_{i=1}^N(\rho_i)^{\nu_i}\prod_{j=1}^N(\rho_j)^{\delta_j}\right\rangle - \left\langle\prod_{i=1}^N(\rho_i)^{\nu_i}\right\rangle\left\langle\prod_{j=1}^N(\rho_j)^{\delta_j}\right\rangle \geq 0 \tag{8}$$

for any multiplicity functions ν_i and δ_j.

Proof: Using Eqs. (5) and (6) and the relationship between ρ_i and s_i, we have

$$\begin{aligned}&Z_N^2\left[\left\langle\prod_{i=1}^N(\rho_i)^{\nu_i}\prod_{j=1}^N(\rho_j)^{\delta_j}\right\rangle - \left\langle\prod_{j=i}^N(\rho_i)^{\nu_i}\right\rangle\left\langle\prod_{j=1}^N(\rho_j)^{\delta_j}\right\rangle\right]\\ &= e^{-NJ_0}\sum_{\{s\}}\sum_{\{s'\}}\int\cdots\int d^N x\, d^N y\, W_N(x)W_N(y)\\ &\times\prod_{i,j=1}^N\left(\frac{1}{2}\right)^{\nu_i+\delta_j}\left(\frac{\partial}{\partial x_i}+p\right)^{\nu_i}\left[\left(\frac{\partial}{\partial x_j}+p\right)^{\delta_j} - \left(\frac{\partial}{\partial y_j}+p\right)^{\delta_j}\right]\\ &\times\exp(s_i x_i + s_i' y_i)\\ &\times\exp\left\{\sum_{\substack{m\\ \text{odd}}} h_m(i)[s_i^m + (s_i')^m] + \sum_{\substack{n\\ \text{even}}}\mu_n(i)[s_i^n + (s_i')^n]\right\}.\end{aligned} \tag{9}$$

Defining new variables

$$\eta_k = (1/\sqrt{2})(x_k + y_k), \quad \xi_k = (1/\sqrt{2})(x_k - y_k), \tag{10}$$

one can rewrite (9) as

$$Z_n^2\left[\left\langle\prod_{i=1}^N(\rho_i)^{\nu_i}\prod_{j=1}^N(\rho_j)^{\delta_j}\right\rangle-\left\langle\prod_{i=1}^N(\rho_i)^{\nu_i}\right\rangle\left\langle\prod_{j=1}^N(\rho_j)^{\delta_j}\right\rangle\right]$$

$$=e^{-NJ_0}\sum_{\{s\}}\sum_{\{s'\}}\int\cdots\int d^N\eta d^N\xi W_N(\eta)W_N(\xi)$$

$$\times\prod_{i,j=1}^N\left\{\left(\frac{1}{2}\right)^{\nu_i+\delta_j}\left[\frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial\eta_i}+\frac{\partial}{\partial\xi_i}\right)+p\right]^{\nu_i}\right.$$

$$\left.\times\left(\left[\frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial\eta_j}+\frac{\partial}{\partial\xi_j}\right)+p\right]^{\delta_j}-\left[\frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial\eta_j}-\frac{\partial}{\partial\xi_j}\right)+p\right]^{\delta_j}\right)\right\}$$

$$\times\prod_{k=1}^N\exp\left\{s_kx_k+s'_ky_k+\sum_m h_m(k)\,[s_k^m+(s'_k)^m]\right.$$

$$\left.+\sum_n\mu_n(k)[(s_k)^n+(s'_k)^n]\right\}. \tag{11}$$

By looking only at the terms arising from expressing $(\rho_i)^{\nu_i}$ and $(\rho_j)^{\delta_j}$ as partial derivatives,

$$\prod_{i,j=1}^N\left\{\left(\frac{1}{2}\right)^{\nu_i+\delta_j}\left[\frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial\eta_i}+\frac{\partial}{\partial\xi_i}\right)+p\right]^{\nu_i}\right.$$

$$\times\left(\left[\frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial\eta_j}+\frac{\partial}{\partial\xi_j}\right)+p\right]^{\delta_j}\right.$$

$$\left.\left.-\left[\frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial\eta_j}-\frac{\partial}{\partial\xi_j}\right)+p\right]^{\delta_j}\right)\right\}, \tag{12}$$

and defining

$$\gamma_i=\frac{1}{\sqrt{2}}\frac{\partial}{\partial\eta_i}+p,\quad \epsilon=\frac{1}{\sqrt{2}}\frac{\partial}{\partial\xi_i}, \tag{13}$$

the preceding expression (12) becomes

$$\prod_{i,j=1}^N(\tfrac{1}{2})^{\nu_i+\delta_j}(\gamma_i+\epsilon_i)^{\nu_i}\sum_{r\ \mathrm{odd}}\frac{\delta_j!}{(\delta_j-r)!\,r!}(\gamma_j)^{\delta_j-r}(\epsilon_j)^r\ . \tag{14}$$

The above expression can be written as

$$\sum_{l_1\cdots K_N}C(l_1,k_1,\ldots,l_N,k_N)\prod_{i=1}^N\gamma_i^{l_i}\epsilon_i^{k_i}$$

where l_i and k_i are nonnegative integers and where $C(l_1,k_1,\ldots,l_N,k_N)$ is a positive constant. We can therefore consider two separate averages, one of the γ_i's which operate on the variables η_i and the other of the ϵ_i's which operate on the ξ_i's. The γ_i's and the ϵ_i's act on the terms involving s_i and s'_i being summed from $-p$ to $+p$. This expression can be written in terms of η_i and ξ_i as

$$\sum_{s_i,s'_i}\cosh\left(\frac{s_i+s'_i}{\sqrt{2}}\eta_i+\sum_{m\ \mathrm{odd}}h_m(i)[s_i^m+(s'_i)^m]\right)$$

$$\times\cosh\left(\frac{|s_i-s'_i|}{\sqrt{2}}\right)\xi_i$$

$$\times\exp\left(\sum_{n\ \mathrm{even}}\mu_n(i)[s_i^n+(s'_i)^n]\right). \tag{15}$$

Any $\gamma_i^{l_i}$ or $\epsilon_i^{k_i}$ acting on this quantity gives respectively

$$\sum_{s_i,s'_i}\left\{\frac{1}{2}\left(\frac{s_i+s'_i}{2}+p\right)^{l_i}\right.$$

$$\times\exp\left(\frac{s_i+s'_i}{\sqrt{2}}\eta_i+\sum_m h_m(i)[s_i^m+(s'_i)^m]\right)$$

$$+\frac{1}{2}\left(-\frac{s_i+s'_i}{2}+p\right)^{l_i}$$

$$\left.\times\exp\left(-\frac{s_i+s'_i}{\sqrt{2}}\eta_i-\sum_m h_m(i)\,[s_i^m+(s'_i)^m]\right)\right\}$$

$$\times\cosh\left(\frac{|s_i-s'_i|}{\sqrt{2}}\xi_i\right)\exp\left(\sum_n\mu_n(i)[s_i^n+(s'_i)^n]\right) \tag{16}$$

$$\sum_{s_i,s'_i}\cosh\left(\frac{s_i+s'_i}{\sqrt{2}}\eta_i+\sum_{m\ \mathrm{odd}}h_m(i)[s_i^m+(s'_i)^m]\right)\left(\frac{|s_i-s'_i|}{2}\right)^{k_i}$$

$$\times\Theta\left(\frac{|s_i-s'_i|}{\sqrt{2}}\xi_i\right)\exp\left(\sum_{n\ \mathrm{even}}\mu_n(i)[s_1^n+(s'_i)^n]\right), \tag{17}$$

where $\Theta(x)=\sinh(x)$ if k_i is odd and $\Theta(x)=\cosh(x)$ if k_i is even. One now sees that for the terms involving the η_i variables in (16) and (17) we have only nonnegative functions of the η_i's irrespective of the values of s_i, s'_i, and the $h_m(i)$'s. This is true in (16) since $(s_i+s'_i)/2\leqslant p$ for all allowed values of s_i and s'_i, and it is true in (17) since $\cosh(x)\geqslant 0$ for all x. Hence their average is nonnegative when taken with respect to $W_N(\eta)$. For the remaining random variables, ξ_i's, we can expand the cosh and sinh terms as a power series in the ξ_i's. We therefore have a sum of averages of the form $C(\alpha_1,\ldots,\alpha_N)\xi_1^{\alpha_1}\xi_2^{\alpha_2}\cdots\xi_N^{\alpha_N}$, where the constant $C(\alpha_1,\alpha_2,\ldots,\alpha_N)$ is nonnegative. The average of any product of the ξ_i's being nonnegative whenever $J(i,j)\geqslant 0$ the average of the cosh and sinh terms in the ξ_i variables is a sum of nonnegative terms. Clearly the terms involving the $\mu_n(i)$'s are nonnegative and, hence, the right-hand side of Eq. (9) is nonnegative and the inequality is proven.

Note that for these inequalities there is no longer any restrictions on the one-body interactions, a feature common to FKG inequalities. Such freedom has been exploited by Lebowitz[5] to enable him to prove correlation inequalities for a specific class of antiferromagnets. Using the above approach and a transformation of Lebowitz these inequalities can be proven directly by the above method.

ACKNOWLEDGMENT

The author would like to thank Professor A. J. F. Siegert for a critical reading of the manuscript.

*Supported by NSF Grant GP-36564 X1.

[1] J. L. Monroe and A. J. F. Siegert, "*GKS* Inequalities for Arbitrary Spin Ising Ferromagnets," J. Stat. Phys., to be published.
[2] R. B. Griffiths, J. Math. Phys. 8, 478, 484 (1967). D. G. Kelly and S. Sherman, J. Math. Phys. 9, 466 (1968).
[3] C. M. Fortuin, P. W. Kasteleyn, and J. Ginibre, Comm. Math. Phys. 22, 89 (1971).
[4] H. Cramer, *Mathematical Methods in Statistics* (Princeton U. P., Princeton, N. J., 1951), p. 118.
[5] J. L. Lebowitz, Phys. Lett. A 36, 99 (1971).

Unification of the external conformal symmetry group and the internal conformal dynamical group

A. O. Barut* and G. L. Bornzin*

International Centre for Theoretical Physics, Trieste, Italy
Institute for Theoretical Physics, University of Colorado, Boulder, Colorado 80302
(Received 5 November 1973)

The two common applications of $O(4,2)$ as a conformal group on external space–time coordinates and as a dynamical group on internal relative coordinates are combined into a unified algebraic structure for composite systems. A method is given for obtaining from this structure infinite-component wave equations and a discrete, linearly increasing mass spectrum.

1. INTRODUCTION

The four-plus-two-dimensional orthogonal group $O(4,2)$ has been applied in physics both as a symmetry group of transformations of global or external space—time coordinates and as a dynamical group of transformations of rest frame states of a system having internal degrees of freedom. The first application arose historically from the demonstration by Bateman and Cunningham[1] in 1910 that Maxwell's equations are invariant under the inversion

$$X'_\mu = k^2 X_\mu/(X_\nu X^\nu). \tag{1.1}$$

When this observation is combined with the well-known Poincaré and dilatation invariance, Maxwell's equations are found to be invariant under the 15-parameter group of transformations called the conformal group—a *particular* nonlinear realization of $O(4,2)$.[2] A conformally invariant formulation of classical electrodynamics, both for massless and massive particles, has been given recently.[3]

The second application of $O(4,2)$—as a dynamical group—has been continuing over the past decade, growing out of studies of such composite systems as the hydrogen atom and, more recently, the relativistic models for composite hadrons.[4] The quantum numbers of the system which characterize certain internal degrees of freedom can be made to correspond with eigenvalues of certain generators, or combinations of generators of the group. Other generators then serve as transition operators between eigenstates. It seems that precisely $O(4,2)$ is both necessary and sufficient to characterize the system of two spinless particles.[5] In the case of the hydrogen atom, a specific realization of the generators in terms of internal coordinate variables is known.[6]

The question arises whether there may be underlying physical significance or a physical "reason" for the occurrence of the same group in these two different but complementary aspects of particle dynamics. This is the primary question investigated in this paper. Although a conclusive and final answer is yet to be found, we shall demonstrate the relationship between these two roles of $O(4,2)$ and indicate some surprising and interesting interconnections.. Moreover, we shall show that the combined external and internal groups can be used to generate a discrete mass spectrum in a manner which avoids the restrictions of O'Raifeartaigh's theorem.[7]

2. THE CONFORMAL GROUP—EXTERNAL

The Lorentz group $O(3,1)$ of rotations and boosts has six generators $L_{\mu\nu}$ which close under commutation to form a Lie algebra. If we add to these the four generators of translation P_μ, we obtain the Poincaré group or inhomogeneous Lorentz group $IO(3,1)$. The additional possibility of inversion, a single discrete operation, requires the addition of five new generators in order finally to achieve a closed group $O(4,2)$, the conformal group. The new generators are: first, a 4-vector K_μ which generates the so-called special conformal transformations

$$X'_\mu = \frac{X_\mu - C_\mu X^2}{1 - 2C_\nu X^\nu + C^2 X^2}, \tag{2.1}$$

obtained as the product of inversion, translation, and reinversion; and second, a scalar D which generates dilatations

$$X'_\mu = \rho X_\mu \tag{2.2}$$

and occurs in the commutator of P_μ and K_μ.

If X_μ and Π_μ are conjugate variables,

$$[\Pi_\mu, X_\nu] = i g_{\mu\nu}, \quad g = (+,-,-,-), \tag{2.3}$$

we may write the conformal generators as follows:

$$\begin{aligned} L_{\mu\nu} &= X_\mu \Pi_\nu - X_\nu \Pi_\mu, \\ P_\mu &= L_{\mu 6} + L_{\mu 4} = \Pi_\mu, \\ K_\mu &= L_{\mu 6} - L_{\mu 4} = 2X_\mu (X_\nu \Pi^\nu + iH) - (X_\nu X^\nu)\Pi_\mu, \\ D &= L_{64} = X_\nu \Pi^\nu + iH. \end{aligned} \tag{2.4}$$

Here H is a number called the homogeneity (for reasons given below) and is related to the Casimir invariant of the algebra by

$$Q = \tfrac{1}{2} L_{ab} L^{ab} = H^2 - 4H. \tag{2.5}$$

Acting as differential operators on a function space of X, Π_μ would be represented by $i\partial/(\partial X^\mu)$. Direct verification shows that the generators L_{ab} obey the $O(4,2)$ commutation relation

$$[L_{ab}, L_{cd}] = -i(g_{ac} L_{bd} + g_{bd} L_{ac} - g_{ad} L_{bc} - g_{bc} L_{ad}), \tag{2.6}$$

where g is the diagonal metric

$$g = \begin{pmatrix} + & - & - & - & - & + \\ 0=5 & 1 & 2 & 3 & 4 & 6 \end{pmatrix}. \tag{2.7}$$

The specific commutators are

$$[L_{\mu\nu}, L_{\lambda\rho}] = -i(g_{\mu\lambda}L_{\nu\rho} + g_{\nu\rho}L_{\mu\lambda} - g_{\mu\rho}L_{\nu\lambda} - g_{\nu\lambda}L_{\mu\rho}),$$
$$[P_\mu, K_\nu] = 2i(g_{\mu\nu}D - L_{\mu\nu}),$$
$$[L_{\mu\nu}, P_\lambda] = -i(g_{\mu\lambda}P_\nu - g_{\nu\lambda}P_\mu), \quad [D, P_\lambda] = -iP_\lambda,$$
$$[L_{\mu\nu}, K_\lambda] = -i(g_{\mu\lambda}K_\nu - g_{\nu\lambda}K_\mu), \quad [D, K_\lambda] = iK_\lambda. \tag{2.8}$$

All other commutators vanish.

One can readily show for the representation (2.4) that $P_\mu P^\mu$, while not an invariant of this algebra, nevertheless gives some quantity times itself when commuted with each of the generators (for the case $H=1$). Hence, *in the special case* that $P_\mu P^\mu$ equals zero when acting on the carrier space, $P_\mu P^\mu$ is invariant and the conformal symmetry is exact. This observation has led to the customary view that the presence of mass terms breaks conformal symmetry, and thus limits the use of the conformal group to the high-energy domain where rest masses effect only minor symmetry breaking. It has been shown, however, that by considering the usual rest mass to be a scale factor times a new, conformally-invariant mass, and by interpreting conformal transformations as a space- and time-dependent change of scale, conformally invariant equations of motion can be written for *massive* particles as well.[3] (Recently, Dirac[8] also considered a transformation of mass with dilatations proportional to the age of the universe.) Physical processes are considered to take place in a six-dimensional space in which the usual Minkowski sapce is a special projected subspace wherein the scale or unit of measure remains constant from point to point. Although the following analysis is valid regardless of whether or not one chooses to apply this latter interpretation, such an interpretation significantly extends the applicability of all conclusions regarding conformal symmetry. We therefore consider the case where $P_\mu P^\mu \neq 0$, i.e., $P_\mu P^\mu$ has then a continuous spectrum.

The special conformal transformation (2.1) may be linearized by introducing K as a scale parameter along with the following new coordinates

$$Y^\mu \equiv KX^\mu,$$
$$Y^4 - Y^6 \equiv K, \tag{2.9}$$
$$Y^4 + Y^6 \equiv \Lambda \equiv KX_\mu X^\mu,$$

having the property

$$Y \cdot Y \equiv Y_a Y^a = Y_\mu Y^\mu - (Y^4 - Y^6)(Y^4 + Y^6) = Y_\mu Y^\mu - K\Lambda = 0. \tag{2.10}$$

The generators (2.4) may then be written in the simple form

$$L_{ab} = Y_a Q_b - Y_b Q_a, \tag{2.11}$$

where Q is the variable conjugate to Y. With the interpretation indicated above, physics takes place on the five-dimensional hypercone Eq. (2.10) in the six-dimensional Y-space.

The procedure for passing from six-dimensions down to four[3,9] involves, in addition to the change of variables, the imposing of two constraints. First, $S = Y \cdot Y$ is an invariant of the algebra and so may take a particular fixed value—zero, in this case. In fact, formal change of variables from Y_a to Y_μ, K, and S in Eq. (2.11) yields generators independent of Q_S or $\partial/(\partial S)$. Hence the value of S cannot affect the commutation relations; one variable is effectively eliminated. Secondly, the carrier space of functions are taken to be homogeneous in Y. With the change of variables (2.9), the homogeneity property allows one to factor out the K dependence as K^{-H}, where H is the homogeneity. The K-dependence in the generators occurs only in terms of the form $KQ_K \to iK\partial/\partial K$, which may be replaced everywhere by $-iH$, its effect upon such homogeneous functions. Thus, one is left with generators in four variables—the conformal group in X-space, Eq. (2.4). We shall call this procedure "reduction with respect to Y." [The same net result is more readily, but less rigorously, achieved by making the change of variables Eq. (2.9) and then simply replacing Λ by $KX_\mu K^\mu$ directly, and setting $Q_\lambda = 0$, thus avoiding the variable S.]

3. THE DYNAMICAL GROUP—INTERNAL

A dynamical group G is a group whose subgroup structure and multiplicities in a representation T_g can be made to correspond with all the rest frame states (bound and continuum) of a quantum system having internal degrees of freedom, and which contains current operators representing interactions and effecting transitions between states. For example, $O(4)$ is an exact symmetry group of the nonrelativistic H atom and contains the quantum numbers, n, l, and m describing the rest states. But each irreducible representation of $O(4)$ corresponds to a different n and different energy. Enlarging the group to $O(4,1)$ is necessary to allow for transitions between states of different energy. Finally, enlarging to $O(4,2)$ is necessary in order to include continuum states and current operators.

The dynamical group approach and associated techniques have been applied successfully to a number of problems of composite systems. From the quantum mechanical point of view, a knowledge of the states (i.e., wavefunctions) and of the interaction current operators is all there is to know about a system. Extension of the method from known systems such as the nonrelativistic H atom and the Dirac atom to strongly "bound" highly relativistic systems such as pion and proton has enabled calculations of mass spectra of excited states, form factors, structure functions, and cross sections of various electromagnetic processes, all in agreement with experiment—all using $O(4,2)$.[10]

The representations generally used as dynamical groups are characterized by having generators which obey the representation relation[11]

$$\{L_{ab}, L^a{}_c\} = -2\alpha g_{bc}, \quad \alpha \text{ const.} \tag{3.1}$$

The representations with minimum spin-zero and spin-half in this class have found most frequent applications. For strongly bound and highly relativistic systems in which concepts such as localizability and constituent particles become poorly understood and perhaps useless, the particular *realization* of the group representation becomes a matter of convenience as the postulated algebraic structure takes precedence over the detailed physical interpretation of the generators. For a given algebraic structure, the introduction of internal coordinates sometimes becomes ambiguous.[12] But in nonrelativistic problems, from which we gain intuition on how to approach the strongly bound and relativistic problems, the realization is determined by the known constituents and coordinates and by the interaction.

For the H atom, one can build the $O(4,2)$ realization somewhat as outlined at the beginning of this section, starting with $O(4)$. Or one can combine the appropriate radial $O(2,1)$ algebra with the rotation group $O(3)$ using knowledge of either the Runge—Lenz vector or the internal Galilei booster. We shall indicate an alternate approach which reveals a curious connection with the conformal group of Sec. 2.

Conformal group in momentum space

Consider the generators of the conformal group of transformations of an internal, four-dimensional *momentum* vector π_μ conjugate to the difference four-vector of two particles $x_\mu = x_{1\mu} - x_{2\mu}$. These can be written directly from the generators (2.4) by letting $\Pi_\mu \to x_\mu$ and $X_\mu \to -\pi_\mu$ (the minus sign occurs to preserve the commutator in the form $[\pi_\mu, x_\nu] = ig_{\mu\nu}$) and reordering with π's to the right by convention to obtain

$$\begin{aligned} l_{\mu\nu} &= x_\mu \pi_\nu - x_\nu \pi_\mu, \\ l_{\mu 6} + l_{\mu 4} &= x_\mu, \\ l_{\mu 6} - l_{\mu 4} &= 2(x_\nu \pi^\nu + ih)\pi_\mu - x_\mu \pi_\nu \pi^\nu, \\ l_{64} &= -(x_\nu \pi^\nu + ih). \end{aligned} \tag{3.2}$$

Just as the special case $P_\mu P^\mu = 0$ was an invariant of the conformal group in X-space, likewise here $x_\mu x^\mu = 0$ is an invariant of the conformal group in π-space (for $h=1$). We may incorporate this special case into the algebra by changing to the three-dimensional variable $\mathbf{r} = x^i$ and $s = x_\mu x^\mu$ and afterwards set $s = 0$ and $r = |\mathbf{r}|$ to obtain[13]

$$\begin{aligned} l_{ij} &= r_i \pi_j - r_j \pi_i = \epsilon_{ijk}(\mathbf{r} \times \pi)^k, \\ l_{i0} &= r\pi, \\ l_{i6} + l_{i4} &= -\mathbf{r}, \\ l_{06} + l_{04} &= r, \\ l_{i6} - l_{i4} &= 2(\mathbf{r} \cdot \pi - i)\pi - \mathbf{r}\pi^2, \\ l_{06} - l_{04} &= r\pi^2, \\ l_{64} &= \mathbf{r} \cdot \pi - i. \end{aligned} \tag{3.3}$$

Up to sign conventions and labeling of indices, this is the usual H atom $O(4,2)$, useful in both the nonrelativistic and relativistic cases.[14] This realization is also useful in establishing the correspondence between the traditional formulations of Schrödinger, Klein—Gordon, and Dirac with the method of infinite-component wave equations, giving strong motivation to the choice of coefficients required in the latter approach.[15]

The generators (3.3) no longer generate conformal transformations of an internal 4-momentum; one $O(4,1)$ subalgebra, however, still generates the conformal group of transformations of the three-dimensional internal π vector. This group includes the well-known $O(4)$ rotational symmetry of the Fock-transformed π vector.[16]

Direct calculation shows that the general conformal algebra with which we began [Eq. (3.2)] does not satisfy the representation relation (3.1). However, after limiting our concern to the special case $x_\mu x^\mu = 0$, we find the generators (3.3) do indeed satisfy the relation (3.1) and generate the spin-zero representation of this class. Likewise, although the conformal algebra in X-space [Eq. (2.4)] does not satisfy the relation (3.1), when restricted to act upon solutions of $(P_\mu P^\mu)\psi = 0$, and with $H=1$, the relation (3.1) does hold on this function space, and the spin-zero representation is again realized.

The special case $x_\mu x^\mu = 0$ is a very natural and desirable one, as this condition is also obtained as that necessary for electromagnetic interaction to take place between the particles according to the Green's function $\delta(x^2)$ in the action integral or in the vector potential, or simply by the assumption that the interacting field propagates at the speed of light.

The question of the "physical cause" of the so-called accidental degeneracy or $O(4)$ symmetry of the nonrelativistic H atom is thus equivalent to seeking an explanation for the occurrence of the conformal group in internal four-dimensional momentum space (reduced to three dimensions with the condition $x_\mu x^\mu = 0$) as the dynamical group of hydrogenic systems.

4. COMBINED EXTERNAL AND INTERNAL GROUPS

A. The six dimensional approach

We shall now give two approaches towards unifying the concepts of Secs. 2 and 3 by investigating group structures which combine external and internal properties of a two-particle composite system.

First we shall apply the interpretation of Ref. 3 and assume that for each of the two particles physical processes actually take place in a six-dimensional space. Our task is to understand how these processes appear projected into the observed four-dimensional Minkowski space. If in six dimensions the equations of motion for each particle are "rotationally invariant" (i.e., conformally invariant in Minkowski space), then we may write the total invariance algebra for the system as

$$\mathcal{L}_{ab} = (L_{ab})_1 + (L_{ab})_2. \tag{4.1}$$

From Eq. (2.11)

$$(L_{ab})_i = (y_a q_b - y_b q_a)_i, \tag{4.2}$$

where $(y_a)_i$ and $(q_a)_i$ are the six pairs of conjugate variables ($a = 1, \ldots, 6$) for particle i ($i = 1, 2$). We further require [see Eq. (2.10)] that

$$y_1 \cdot y_1 = y_2 \cdot y_2 = 0 \tag{4.3}$$

in order that each particle algebra reduces correctly to Minkowski space. We now introduce usual external and internal variables

$$\begin{aligned} Y &= wy_1 + \bar{w}y_2, \quad & Q &= q_1 + q_2, \\ y &= y_1 - y_2, \quad & q &= \bar{w}q_1 - wq_2, \end{aligned} \tag{4.4}$$

where w and $\overline{w}$ are arbitrary weight functions with the property

$$w+\overline{w}=1. \tag{4.5}$$

(We do not call Y a center-of-mass variable as we wish to leave unspecified the possible relation of w and $\overline{w}$ to the constituent masses.) It is readily shown that

$$[Q_a, Y_b]=[q_a, y_b]=ig_{ab}, \tag{4.6}$$

while all other commutators vanish.

With these definitions the algebra (4.1) splits into commuting external and internal algebras; i.e.,

$$\begin{aligned} L_{ab} &= (L_{ab})_1+(L_{ab})_2 \\ &= (y_aq_b-y_bq_a)_1+(y_aq_b-y_bq_a)_2 \\ &= (Y_aQ_b-Y_bQ_a)+(y_aq_b-y_bq_a) \\ &= L_{ab}+l_{ab}. \end{aligned} \tag{4.7}$$

We would like the external algebra to generate the conformal group in X-space when reduced to four dimensions, so that the composite particle would behave like an "elementary" particle as far as external coordinates are concerned. Recall that reduction to Minkowski space requires the constraint $Y\cdot Y=0$. The compatibility of this constraint with $y_1\cdot y_1=0$ and $y_2\cdot y_2=0$ implies some interesting consequences:

$$\begin{aligned} 0 = Y\cdot Y &= (wy_1+\overline{w}y_2)\cdot(wy_1+\overline{w}y_2) \\ &= w^2y_1\cdot y_1+\overline{w}^2y_2\cdot y_2+2w\overline{w}y_1\cdot y_2 \\ &= 2w\overline{w}y_1\cdot y_2. \end{aligned} \tag{4.8}$$

Either w or $\overline{w}$ equals zero, or $y_1\cdot y_2=0$. For w or $\overline{w}$ equal to zero, Y equals y_2 or y_1, respectively; Y satisfies the constraint trvially, and the reduction to the conformal group in X-space is equivalent to the single particle reduction. But the interesting case $y_1\cdot y_2=0$ implies [using Eqs. (2.9) and (2.10)]

$$\begin{aligned} 0 = y_1\cdot y_2 &= y_{1\mu}y_2{}^{\mu}-\tfrac{1}{2}\kappa_1\lambda_2-\tfrac{1}{2}\kappa_2\lambda_1 \\ &= \tfrac{1}{2}\kappa_1\kappa_2(2x_{1\mu}x_2{}^{\mu}-x_1{}^2-x_2{}^2) \\ &= -\tfrac{1}{2}\kappa_1\kappa_2(x_1-x_2)^2. \end{aligned} \tag{4.9}$$

Thus we must have κ_1 or κ_2 identically equal to zero (which contradicts their interpretation as scale parameters), or we must have $(x_1-x_2)^2=0$. If we define the internal coordinate to be

$$x_\mu \equiv x_{1\mu}-x_{2\mu}, \tag{4.10}$$

we have precisely the desired restriction $x_\mu x^\mu=0$ discussed in Sec. 3 for the internal algebra. Thus we derive a rather unexpected and non-trivial result from applying the six-dimensional interpretation to conformal symmetry of a two-particle system.

We may now reduce the external $O(4,2)$ with respect to Y, using the condition $Y\cdot Y=0$, and arrive at the conformal group in X-space according to the procedure concluding Sec. 2. We then focus our attention on the internal $O(4,2)$. If we attempt a similar reduction internally, we find ourselves forced away from the usual and desirable definition of internal coordinates (4.10) to the definition [according to Eqs. (2.9) and (4.4)]

$$x_\mu=\frac{y_\mu}{\kappa}=\frac{y_{1\mu}-y_{2\mu}}{\kappa_1-\kappa_2}=\frac{\kappa_1x_{1\mu}-\kappa_2x_{2\mu}}{\kappa_1-\kappa_2}. \tag{4.11}$$

Not only is this a strange and undesirable internal coordinate, but it becomes totally useless in the case $\kappa_1=\kappa_2$—a case we would like to interpret as both particles being in the same Minkowski space, using the same unit of measure. In order to avoid this definition, therefore, and to recover instead the definition (4.10), it seems necessary to reduce the internal $O(4,2)$ *with respect to* q, assuming the condition $q\cdot q=0$. That is, we consider $(2q_\lambda)$ to be the scale parameter analogous to κ, define[17]

$$\begin{aligned} q_\mu &= 2q_\lambda\lambda_\mu, \\ q_4+q_6 &= 2q_\lambda, \\ q_4-q_6 &= 2q_\kappa, \end{aligned} \tag{4.12}$$

and proceed in analogy with the reduction procedure in Sec. 2, treating the q's as coordinates and the y's as differential operators. The result of this reduction is, of course, the conformal algebra in π-space, Eq. (3.2). The rigorous reduction procedure yields the same result as replacing q_κ by $q_\lambda\pi_\mu\pi^\mu$ and setting $\kappa=0$ (i.e., $\kappa_1=\kappa_2$, as desired) in analogy with the prescription concluding Sec. 2.

In order to have this conformal albegra in internal π-space identical to the dynamical algebra discussed in Sec. 3, we must make the natural requirement that all scale factors be the same. Physically, this is simply the requirement that all measurements—x_1, x_2, X and x—be made with the same "units." Mathematically, this is

$$\begin{aligned} \kappa_1 &= \kappa_2=\kappa_0, \\ K &\equiv w\kappa_1+\overline{w}\kappa_2=(w+\overline{w})\,\kappa_0=\kappa_0, \\ \kappa &= \kappa_1-\kappa_2=0, \\ (2q_\lambda)^{-1} &= \kappa_0. \end{aligned} \tag{4.13}$$

With these "constant units" the internal coordinate variable x_μ is

$$x_\mu=y_\mu(2q_\lambda)=(\kappa_1x_{1\mu}-\kappa_2x_{2\mu})(2q_\lambda)=x_{1\mu}-x_{2\mu}. \tag{4.14}$$

So the restriction $(x_{1\mu}-x_{2\mu})^2=0$ obtained from reducing the external algebra may be applied to the internal algebra to give us finally the dynamical algebra of Eq. (3.3).

In summary, we began this section with the hypothesis that the position coordinate of each of two particles transforms according to the conformal group and that the motion of each particle may be considered to take place in a six-dimensional space, constrained to a five-dimensional hypercone. Upon introducing external and internal variables, the total system algebra splits into commuting external and internal algebras. The external algebra may be projected to Minkowski space to describe a composite particle which transforms conformally as did the original constituents, but the possibility of such a projection implies $(x_1-x_2)^2=0$ (or the trivial case w or $\overline{w}=0$). The internal algebra cannot be reduced in the same manner, due to the meaningless coordinate definition which results and due to the possibility that $\kappa=0$. It can, however, be reduced with respect to the q-space conjugate to y in a totally analogous manner, so that the internal algebra becomes the conformal algebra in a four-dimensional internal momentum space. When all quantities are measured by the same scale, the internal position coordinate is precisely $x_1 - x_2$; so the condition $(x_1 - x_2)^2 = 0$, which happens to be an invariant of the conformal group in momentum space, may be used to reduce the algebra to the usual r-space realization of the two-particle dynamical group. The total system algebra, therefore, is the direct sum of the conformal algebra in the space of the external position coordinate [Eq. (2.4)] and the usual dynamical algebra in the space of the internal position difference coordinate [Eq. (3.3)].

B. The Casimir operator and mass spectrum

The invariant Casimir operator of this total $O(4,2)$ algebra which combines external and internal algebras would be expected to indicate something of the nature of the system or to indicate nothing, being trivial. In fact, it is quite complex and can be diagonalized only after considerable algebraic manipulation which we outline here.

The external group is generated by $L_{\mu\nu}$, P_μ, K_μ, and D [Eq. (2.4)]. For convenience, we shall employ the same letters in lower case—$l_{\mu\nu}$, p_μ, k_μ, and d—to represent the generators of the internal group [Eq. (3.3)], and in script—$\mathcal{L}_{\mu\nu}$, $\mathcal{P}_\mu$, $\mathcal{K}_\mu$, and $\mathcal{D}$—to represent the sum of the two algebras. Since all three are $O(4,2)$ algebras, all obey the commutation relations (2.8).

The total Casimir operator is given by

$$\begin{aligned}\mathcal{Q} &= \tfrac{1}{2}\mathcal{L}_{ab}\mathcal{L}^{ab}\\ &= \tfrac{1}{2}\mathcal{L}_{\mu\nu}\mathcal{L}^{\mu\nu} + \tfrac{1}{2}\mathcal{P}_\mu\mathcal{K}^\mu + \tfrac{1}{2}\mathcal{K}_\mu\mathcal{P}^\mu - \mathcal{D}^2\\ &= \tfrac{1}{2}\mathcal{L}_{\mu\nu}\mathcal{L}^{\mu\nu} + \mathcal{K}_\mu\mathcal{P}^\mu - \mathcal{D}^2 + 4i\mathcal{D}\\ &\doteq Q + q + L_{\mu\nu}l^{\mu\nu} + K_\mu p^\mu + k_\mu P^\mu - 2Dd,\end{aligned} \tag{4.15}$$

where Q and q are the external and internal Casimir invariants. Slight simplification is obtained by transforming each generator by

$$T_1 = \exp(iX_\mu p^\nu) \tag{4.16}$$

using the explicit form of the generators for the external algebra but the commutation relations only for the internal algebra. The transformation T_1 leaves $\mathcal{L}_{\mu\nu}$ and $\mathcal{D}$ unchanged, but translates P_μ by an amount $(-p_\mu)$ so that

$$\mathcal{P}'_\mu = (P_\mu - p_\mu) + p_\mu = P_\mu. \tag{4.17}$$

$\mathcal{K}_\mu$ becomes

$$\mathcal{K}'_\mu = K_\mu + k_\mu + 2(X_\mu d + X^\nu l_{\mu\nu}). \tag{4.18}$$

The effect on the form of the Casimir operator is the seemingly miraculous cancellation of the terms $L_{\mu\nu}l^{\mu\nu}$, $2Dd$, and $K_\mu p^\mu$, leaving

$$\mathcal{Q}' = Q + q - k_\mu p^\mu + k_\mu P^\mu - 2iHd. \tag{4.19}$$

We diagonalize Q and P_μ, i.e., we let $\mathcal{Q}'$ act on eigenstates of external momentum P_μ (which after the transformation T_1 is also the new total momentum $\mathcal{P}'_\mu$). We may transform $\mathcal{Q}'$ to the rest frame using the "deboost" operator

$$T_2 = \exp(i\xi^i l_{i0}), \tag{4.20}$$

where ξ is the rapidity given by

$$P^\mu = M(\cosh\xi, \hat{\xi}\sinh\xi). \tag{4.21}$$

$\mathcal{Q}'$ then takes on a form entirely in terms of the internal algebra:

$$\mathcal{Q}'' = Q + q - k_\mu p^\mu + Mk_0 - 2iHd. \tag{4.22}$$

We next take advantage of the fact that the internal algebra obeys the representation relation (3.1) with $\alpha = 1$. From the special case ($b = c = \mu = 0, 1, 2, 3$) we obtain the identity

$$k_\mu p^\mu = -l_{ij}l^{ij} + 2k_0 p_0 - 2id - 2, \tag{4.23}$$

which we substitute into $\mathcal{Q}''$, replacing $l_{ij}l^{ij} = 2\mathbf{l}^2$ with its eigenvalue $2l(l+1)$, and q with its eigenvalue -3:

$$\mathcal{Q}'' = Q - 1 + 2l(l+1) - 2k_0p_0 + 2i(1-H)d + Mk_0. \tag{4.24}$$

Finally, to eliminate the d term, we transform with the operator

$$T_3 = k_0^{(1-H)/2}, \tag{4.25}$$

using the relations

$$\begin{aligned}k_0^{-(1-H)/2}p_0k_0^{(1-H)/2} &= p_0 + ik_0^{-1}[(1-H)d - \tfrac{1}{4}i(1-H)(1+H)],\\ k_0^{-(1-H)/2}dk_0^{(1-H)/2} &= d + \tfrac{1}{2}i(1-H),\end{aligned} \tag{4.26}$$

to obtain

$$\begin{aligned}\mathcal{Q}''' &= Q - 1 + 2l(l+1) - 2k_0p_0 - \tfrac{1}{2}(1-H)(3-H) + Mk_0\\ &= s - 2k_0p_0 + Mk_0,\end{aligned} \tag{4.27}$$

where in the last step we have gathered all the constants into one term s.

If we go on to diagonalize $\mathcal{Q}'''$ itself, we find a continuous spectrum for M. However, if we diagonalize $k_0^{-2}\mathcal{Q}'''$ (or, equivalently, if we postulate an additional term k_0^2 considered to represent an interaction to produce bound states), we may postulate a wave equation such as

$$\mathcal{Q}'''\tilde{\psi} = (2\beta^2)k_0^2\tilde{\psi}, \quad \beta = \text{const}, \tag{4.28}$$

$$k_0[(p_0 - \tfrac{1}{2}sk_0^{-1}) + \beta^2k_0 - \tfrac{1}{2}M]\tilde{\psi} = 0, \tag{4.29}$$

To solve this equation, we define $p'_0 = p_0 - \frac{1}{2}sk_0^{-1}$ and note that $\frac{1}{2}(k_0 + p'_0)$, $\frac{1}{2}(k_0 - p'_0)$, and d form an $O(2,1)$ algebra with compact generator $\frac{1}{2}(k_0 + p'_0)$, as one may show by direct calculation. So this equation is diagonalized by transforming with

$$T_4 = \exp\left[id\tanh^{-1}\left(\frac{\beta^2-1}{\beta^2+1}\right)\right], \tag{4.30}$$

to give

$$\beta[\tfrac{1}{2}(k_0 + p'_0)]\psi = (\tfrac{1}{4}M)\psi. \tag{4.31}$$

The spectrum of $\frac{1}{2}(k_0 + p'_0)$ is of the form $n + \varphi$, where n is a nonnegative integer and φ is a constant depending on the $O(2,1)$ Casimir invariant. Letting $M_0 = 4\beta\varphi$ and $M_1 = 4\beta$, we find a linear rising mass spectrum of the form

$$M = M_0 + nM_1, \quad n = 0, 1, 2, \cdots. \tag{4.32}$$

The interesting aspect of this result is not the exact form of the spectrum or the parameters involved, nor the form of the wave equation or the tedious algebra required to solve it. The result is remarkable rather because it indicates the capability of this approach to produce a *discrete* mass spectrum despite O'Raifeartaigh's theorem from an algebraic framework containing both Poincaré transformations and an internal dynamical group.

There is no contradiction, however, as $P_\mu P^\mu$ still commutes with the internal algebra; it is the additional postulated wave equation which selects out particular discrete masses. Hopefully, with further investigation and refinement, the approach will yield quantitative results as well.

C. Relation to infinite-component wave equations

Equations (4.15), or (4.19), or (4.22), or (4.27) show that, for the fixed value of the total Casimir invariant Q(or Q' or Q'' or Q'''), we have an infinite-component wave equation for the composite system. As is well known, the infinite-component wave equations of interest have the form

$$(J_\mu P^\mu + K)\,\psi(p) = 0,$$

where K is an operator on the space of rest frame internal states of the system [e.g., an internal $O(4,2)$ operator]. The current operator is

$$J_\mu = \alpha_1 \Gamma_\mu + \alpha_2 P_\mu + \alpha_3 P_\mu \Gamma_4,$$

where Γ_μ and Γ_4 are also operators like K. In the rest frame we have

$$(J_0 M + K)\,\psi(0) = 0.$$

In our case, Eq. (4.27) is precisely of this form with $J_0 = k_0$ (i.e., $\alpha_1 = 1$, $\alpha_2 = 0$, $\alpha_3 = -1/M$) and $K = s - 2k_0 p_0 - Q'''$, because k_0 and p_0 are operators on the rest frame states, as we know. Note that no such interlocking of external and internal operators is obtained if we use the Poincaré group alone.

The current operator corresponding to $O(4,2)$, Eq. (4.27), $J_\mu = \Gamma_\mu - (1/M) P_\mu \Gamma_4$, is typical of a zero-energy two-body system. Indeed, so far the only "dynamics" that we have put in has been the condition $x_\mu x^\mu = 0$, which gave us a three-dimensional internal space and the dynamical group $SO(4,2)$ representing the space of rest-frame states. The operator $(\Gamma_0 - \Gamma_4)$ cannot be put into the form $c\Gamma_0$ or $c'\Gamma_4$ by tilting. It has continuous spectrum. The second part of the dynamics comes in the assumption of the operator form of the total transformed Casimir operator Q''' in the enveloping algebra of the internal dynamical group $O(4,2)$, as was done, for example, in Eq. (4.28). The exact form of Eq. (4.28) will depend on the further properties of the two-body system, such as masses, strength, and the type of coupling.

The purely covariant steps leading from the two-body problem to the infinite-component wave equations developed in this work complement the previously given approach, where one has generalized the noncovariant Schrödinger, or Klein—Gordon or Dirac equations into a covariant infinite-component wave equation.

D. The four-dimensional approach

As a totally different approach towards combining external and internal groups, we may confine our attention to Minkowski space from the beginning and avoid the need for passing from six coordinate variables to four. We begin with the direct sum of two conformal algebras,

$$\begin{aligned}
\mathcal{L}_{\mu\nu} &= (x_\mu \pi_\nu - x_\nu \pi_\mu)_1 + (x_\mu \pi_\nu - x_\nu \pi_\mu)_2,\\
\mathcal{P}_\mu &= (\pi_\mu)_1 + (\pi_\mu)_2,\\
\mathcal{D} &= (x_\nu \pi^\nu + ih)_1 + (x_\nu \pi^\nu + ih)_2,\\
\mathcal{K}_\mu &= [2x_\mu (x_\nu \pi^\nu + ih) - x^2 \pi_\mu]_1 + [2x_\mu (x_\nu \pi^\nu + ih) - x^2 \pi_\mu]_2,
\end{aligned} \tag{4.33}$$

and introduce external and internal variables directly in 4-space:

$$\begin{aligned}
X &= w x_1 + \bar{w} x_2, \quad \Pi = \pi_1 + \pi_2,\\
x &= x_1 - x_2, \qquad \pi = \bar{w}\pi_1 - w\pi_2.
\end{aligned} \tag{4.34}$$

We find that all generators separate very nicely except for $\mathcal{K}_\mu$:

$$\begin{aligned}
\mathcal{L}_{\mu\nu} &= (X_\mu \Pi_\nu - X_\nu \Pi_\mu) + (x_\mu \pi_\nu - x_\nu \pi_\mu) = L_{\mu\nu} + l_{\mu\nu},\\
\mathcal{P}_\mu &= \Pi_\mu = P_\mu,\\
\mathcal{D} &= (X_\nu \Pi^\nu + iH) + (x_\nu \pi^\nu + ih) = D + d,\\
\mathcal{K}_\mu &= [2X_\mu (X_\nu \Pi^\nu + iH) - X^2 \Pi_\mu] + (\bar{w} - w)[2x_\mu (x_\nu \pi^\nu + ih) - x^2 \pi_\mu]\\
&\quad + 2(X_\mu d + X^\nu l_{\mu\nu}) + 2w\bar{w}(2x_\mu x_\nu \Pi^\nu - x^2 \Pi_\mu).
\end{aligned} \tag{4.35}$$

For the particular case $\bar{w} = 1$ $(w = 0)$, $\mathcal{K}_\mu$ simplifies to

$$\mathcal{K}_\mu = K_\mu + k_\mu + 2(X_\mu d + X^\nu l_{\mu\nu}). \tag{4.36}$$

Noting $\mathcal{P}_\mu$ and $\mathcal{K}_\mu$, we see this is precisely the form of the generators found in Sec. 4 B, Eqs. (4.17) and (4.18), following transformation by $\exp(iX_\nu p^\nu)$. There, however, lower case letters referred to generators of the conformal group in π-space (or of the dynamical group after taking $x_\mu x^\mu = 0$), whereas here they refer to generators of the conformal group in x-space. Nevertheless, we see that the inverse transformation $T_1^{-1} = \exp(-iX_\nu p^\nu)$ upon this algebra makes it exactly separable into external and internal conformal groups (in X-space and x-space, respectively). This neat separation occurs because the condition $\bar{w} = 1$ implies $X = x_2$, and so naturally transforms conformally. The transformation $\exp(-iX_\nu p^\nu)$ translates x by an amount $X = x_2$; thus $x' = x + X = (x_1 - x_2) + x_2 = x_1$, so it, too, transforms conformally.

There is an especially interesting feature of this algebra. Recall that while $x_\mu x^\mu = 0$ is an invariant of the conformal group in π space, it is not an invariant of the conformal group in x-space. Indeed, the conformal group in x-space includes translations which certainly change $x_\mu x^\mu = 0$. Since the above algebra after transformation by $\exp(iX_\nu p^\nu)$ contains the full conformal algebra in x-space, it does not conserve $x_\mu x^\mu = 0$. Nevertheless, $x_\mu x^\mu = 0$ is an invariant of the algebra *before* transformation. This feature enables us once again to reduce the effective internal dimensionality to three.

This algebra, like that of Secs. 4 A and 4 B, can also be made, after various manipulations, to yield a discrete mass spectrum. Although it does not contain the usual internal dynamical algebra, it nevertheless does manifest certain desirable characteristics which tempt one to speculate on the likelihood of its eventual usefulness.

ACKNOWLEDGMENTS

We wish to thank Dr. Rolf Haugen for helpful discussions. We are grateful to Professor Abdus Salam, the International Atomic Energy Agency, and UNESCO for hospitality at the International Centre for Theoretical Physics, Trieste.

*Supported in part by the Air Force Office of Scientific Research AFOSR-72-2289.

[1]H. Bateman, Proc. London Math. Soc. 8, 228, 469 (1910); E. Cunningham, *ibid.* 8, 77 (1910).

[2]P.A.M. Dirac, Ann. Math. 37, 657 (1936).

[3]A.O. Barut and R.B. Haugen, Ann. Phys. (N.Y.) **71**, 519 (1972), Lett. Nuovo Cimento **7**, 625 (1973), Nuovo Cimento A **18**, 495, 511 (1973).

[4]For a discussion of the development and applications of dynamical groups and further references, see A.O. Barut, *Dynamical Groups* (University of Canterbury Press, Christchurch, New Zealand, 1972). The most recent reference is A.O. Barut and W. Rasmussen, J. Phys. B **6**, 1695, 1713 (1973).

[5]See, for example, C.E. Wulfman, in *Dynamical Groups in Atomic and Molecular Physics*, Vol. 2, *Group Theory and Its Applications*, edited by E. Loebl (Academic, New York, 1971), p. 145.

[6]The algebra of the Dirac matrices is isomorphic to the Lie algebra of $SO(4,2)$; in fact, it is the four-dimensional representation of $SO(4,2)$. This occurrence of $SO(4,2)$ is also in the sense of dynamical groups. See, for example, A.O. Barut, P. Cordero, and G.C. Ghirardi, Phys. Rev. D **9**, 536 (1970).

[7]L. O'Raifeartaigh, Phys. Rev. Lett. **14**, 575 (1965); Phys. Rev. **139B**, 1052 (1965).

[8]P.A.M. Dirac, Proc. Roy. Soc. (London) A **333**, 403 (1973).

[9]G. Mack and Abdus Salam, Ann. Phys. (N.Y.) **53**, 174 (1969).

[10]See, for example, A.O. Barut, D. Corrigan, and H. Kleinert, Phys. Rev. Lett. **20**, 167 (1968); A.O. Barut, Nucl. Phys. B **4**, 455 (1968); A.O. Barut and A. Baiquni, Phys. Lett. B **46**, 423 (1973) and Phys. Rev. D **9**, 1084 (1974).

[11]A.O. Barut and A. Böhm, J. Math. Phys. **11**, 2938 (1970).

[12]A.O. Barut and I.H. Duru, Proc. Roy. Soc. (London) A **333**, 217 (1973). The ambiguity can be removed by considering electromagnetic interactions. A.O. Barut and I.H. Duru (to be published).

[13]Y. Nambu, Phys. Rev. **160**, 1171 (1969).

[14]See, for example, A.O. Barut and G.L. Bornzin, J. Math. Phys. **12**, 842 (1971).

[15]A.O. Barut and W. Rasmussen, J. Phys. B **6**, 1695, 1713 (1973).

[16]See, for example, M. Bander and C. Itzykson, Rev. Mod. Phys. **38**, 330, 346 (1966).

[17]The annoying factor of 2 which seems to spoil the otherwise neat symmetry between reducing with respect to y or with respect to q results from the convention of defining $\kappa = y^4 - y^6$ and $\lambda = y^4 + y^6$. The factor could be shared equally between y-space and q-space by defining $\kappa = (1/\sqrt{2})(y^4 - y^6)$, $\lambda = (1/\sqrt{2}) \times (y^4 + y^6)$, $q_\kappa = (1/\sqrt{2})(q_4 - q_6)$, and $q_\lambda = (1/\sqrt{2})(q_4 + q_6)$, thereby making the transformation from y^4 and y^6 to κ and λ a proper rotation of $\pi/4$ in the 4—6-plane.

Quantum field theory on a seven-dimensional homogeneous space of the Poincaré group*

C. P. Boyer† and G. N. Fleming

Department of Physics, The Pennsylvania State University, University Park, Pennsylvania 16802
(Received 17 July 1973)

Field theories over the homogeneous space of the Poincaré group consisting of the direct product of ordinary Minkowski space $\mathcal{M}$ and a three-dimensional timelike hyperboloid H^3 are investigated in an effort to provide a description of a relativistic extended body. The locality properties of such fields are investigated, and it is found that such fields are in general nonlocal with the nonlocal behavior governed by the type of representations used. It is also found that an internal $O(4)$ symmetry arises quite naturally and is used to construct a model describing an infinite number of particles with an almost linear mass spectrum. This model provides no unphysical solutions, decreasing electromagnetic vertex functions, positive definite field energy, and weak nonlocality. All the vertex functions are calculated and discussed.

1. INTRODUCTION

The idea that hadrons are some sort of composite objects is widely accepted even though the nature of their constituents is currently being debated. Indeed, the very meaningfulness of the term "constituent" is in doubt when in order to "see" such an object one has to perform experiments whose energies exceed the production thresholds of strongly interacting particles. One method which has been designed to circumvent such a problem is the bootstrap hypothesis.[1] However, there are other approaches[2–5] which make use of the valuable aid of field theory. Such an approach describes the composite objects themselves in terms of a field theory, more specifically by fields built over homogeneous spaces of the Poincaré group which are larger than and indeed contain the ordinary Minkowski space. The added underlying variables are then the relativistic analog of the Euler angles of an extended solid, for example, and provide for a relativistic description of an extended body,[6] without specifying the details of the constituents.

Moreover, such an approach allows the spin of a particle to play a role more analogous to that of its mass,[7] the additional degrees of freedom acting as the carrier space for the spin. In this way one can obtain fields associated with families of particles, the members of which have a definite mass spin relationship. Although this is not a necessary consequence of such field theories, this approach offers the advantage of explicitly incorporating Regge trajectories as well as providing for the large proliferation of strongly interacting particles now seen. However, whether or not this type of infinite multiplet approach is used, fields on a large enough homogeneous space of the Poincaré group have enough structure in first order perturbation theory to provide at least qualitatively for the decreasing electromagnetic form factors of hadrons.

In the present work, we investigate fields built over the homogeneous space of the Poincaré group consisting of the direct product of ordinary Minkowski space $\mathcal{M}$ and the three-dimensional two-sheeted hyperboloid H^3. Such fields are in general nonlocal and become local only when the finite dimensional representations of the auxilliary Lorentz group $O(3,1)$ are used. In the local case the usual spin statistics result follows. In the nonlocal case it is found that the highly reducible unitary representations of $O(3,1)$ obtained by making full use of the function space available on H^3 provide a much more acceptable nonlocality.

On the internal space H^3, one can introduce an additional $O(4)$ symmetry in a natural way. In this way we are led to an infinite multiplet theory describing an almost linear spectrum of particles with hydrogenlike degeneracies by a wave equation similar to an infinite component wave equation of Casalbuoni, Gatto, and Longhi.[8] However, in our case the basic fields are quite different enabling one to use the canonical decomposition where the particle properties are manifest. All the first order vertex functions are calculated and discussed. They exhibit the noncrossing symmetric behavior typical of such theories.[3,9] The calculation of scattering amplitudes and the completeness problem are briefly discussed.

2. PRELIMINARIES

A homogeneous space[10] E of a group G is a topological space on which the group action is transitive, i.e., given $x_1, x_2 \in E$, there exists $g \in G$ such that $x_2 = gx_1$. The homogeneous spaces of G are homeomorphic to the coset spaces of G. To demonstrate this, the concept of stability or stationary subgroup is introduced. For a fixed point $x_0 \in E$, consider all the elements $h \in G$ which leave x_0 invariant, i.e., $hx_0 = x_0$. All such elements h form a subgroup H of G, called the stability subgroup at the point x_0. Now if $x_1 = gx_0$, then all other transformations g_1 carrying x_0 to x_1 are of the form $g_1 = gh$, $h \in H$. Since G acts transitively on E, there is a one-to-one correspondence between E and the left cosets of G/H. (Of course, right cosets could have been used instead of left cosets.) Notice for any other point $x' \in E$ related to x_0 by $x' = gx_0$, the stability subgroup is just given by the conjugate subgroup gHg^{-1}.

In the following, we are concerned in general with the homogeneous space obtained from the complete Poincaré group $\mathcal{P}$ and the subgroup $O(3)$, i.e., $IO(3,1)/O(3)$. As a topological space it is homeomorphic to the direct product of Minkowski space $\mathcal{M}$ and the two-sheeted hyperboloid $H^3 = \{\eta^\mu : \eta_\mu \eta^\mu = 1\}$. We also consider the homogeneous space obtained from the restricted Poincaré group and its subgroup $SO(3)$; i.e., $ISO_0(3,1)/SO(3)$, which is homeomorphic to the direct product of $\mathcal{M}$ and the single sheet of H^3, H^3_+. Any further discussion of the discrete symmetries is postponed until Sec. 4.

The basic objects of our theory then are the wavefunctions (to be quantized later) over $\mathcal{M}\otimes H^3$ described by the scalar-valued functions $\phi(x,\eta)$. The action of the Poincaré group on such functions is given by the linear representation

$$U(\Lambda^{-1},a)\phi(x,\eta)=\phi(\Lambda x+a,\Lambda\eta),\quad \Lambda,a\in IO(3,1),\quad (2.1)$$

where for now $\phi(x,\eta)$ is assumed to be a sufficiently smooth function of x and η.

The generators of the Lie algebra of $\mathcal{P}$ are obtained in the usual way via the one parameter subgroups, yielding

$$M_{\mu\nu}=L_{\mu\nu}+S_{\mu\nu},\quad L_{\mu\nu}=i\left(x_\mu\frac{\partial}{\partial x^\nu}-x_\nu\frac{\partial}{\partial x^\mu}\right),\quad (2.2)$$

$$S_{\mu\nu}=i\left(\eta_\mu\frac{\partial}{\partial\eta^\nu}-\eta_\nu\frac{\partial}{\partial\eta^\mu}\right),\quad P_\mu=i\frac{\partial}{\partial x^\mu},$$

which, of course, satisfy

$$\begin{aligned}&[M_{\mu\nu},M_{\sigma\lambda}]=i(g_{\nu\sigma}M_{\mu\lambda}-g_{\mu\sigma}M_{\nu\lambda}-g_{\nu\lambda}M_{\mu\sigma}+g_{\mu\lambda}M_{\nu\sigma}),\\&[M_{\mu\nu},P_\lambda]=i(g_{\nu\lambda}P_\mu-g_{\mu\lambda}P_\nu),\\&[P_\mu,P_\nu]=0,\end{aligned}\quad (2.3)$$

with the Casimir invariants given by

$$\begin{aligned}&p^2=P_\mu P^\mu=-\frac{\partial^2}{\partial x^2},\quad W^2=W_\mu W^\mu,\\&W_\mu=\tfrac{1}{2}\epsilon_{\mu\nu\lambda\rho}M^{\nu\lambda}P^\rho=\tfrac{1}{2}\epsilon_{\mu\nu\lambda\rho}S^{\nu\lambda}P^\rho.\end{aligned}\quad (2.4)$$

The canonical basis is then given in the usual way by diagonalizing the operators P^2, W^2, $\vec{P}$, W_3. It is clear, however, that the eigenvalues of these six operators cannot specify completely the functions ϕ over the seven-dimensional space $\mathcal{M}\otimes H^3$. Leaving aside the additional degree of freedom[11] for the moment the above-mentioned decomposition can be implemented for time-like momenta by introducing the spherical coordinates on H^3:

$$\begin{aligned}&\eta^0=\pm\cosh a, &&\\&\eta^1=\sinh a\,\sin\theta\sin\phi, && 0\le a<\infty,\\&\eta^2=\sinh a\sin\theta\cos\phi, && 0\le\theta<\pi,\\&\eta^3=\sinh a\,\cos\theta, && 0\le\phi<2\pi,\end{aligned}\quad (2.5)$$

where the $\pm$ sign in η^0 corresponds to $H^3_\pm$. It is easily seen that the intersection of H^3, in this coordinate system, with a hyperplane defined by $\eta^0=$ const as shown in Fig. 1a defines the two-dimensional sphere S^2. The operator W^2 reduces in the rest frame to the Laplace–Beltrami operator $\Delta(S^2)$ on this sphere, i.e.,

$$W^2=P^2\Delta(S^2),\qquad\text{rest frame},\quad (2.6)$$

where

$$\Delta(S^2)=\frac{\partial^2}{\partial\theta^2}+\cot\theta\frac{\partial}{\partial\theta}-\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\quad (2.7)$$

with eigenvalues $p^2s(s+1)$. The regular solutions to this eigenvalue problem are the well known spherical harmonics, $Y_{sm}(\theta,\phi)$. The canonical basis states are then given in the rest frame by

$$|\vec{p}=0,M\,sm\rangle\sim\exp(-iMx_0)\rho(\eta_0)Y_{sm}(\theta,\phi).\quad (2.8)$$

In a frame with arbitrary momentum, one obtains by applying a Wigner boost $L(p)$

$$|\vec{p},M\,sm\rangle\sim\exp(-ip\cdot x)\rho(\eta\cdot p)Y_{sm}(\theta_p,\phi_p),\quad (2.9)$$

where θ_p, ϕ_p indicate that the variables θ, ϕ have been boosted by $\vec{p}$, for example for $\vec{p}$ in the z direction

$$\tan\theta_p=\frac{\sinh a\sin\theta}{p_3\cosh a+p_0\sinh a\cos\theta},\quad \phi_p=\phi.\quad (2.10)$$

The functions $\rho(\eta\cdot p)$ are yet to be determined and their specification determines the representation of the Wigner boosts. It should be emphasized, however, that the transformation properties of the wavefunctions, Eq. (2.9), under the Poincaré group do not depend on the choice of the representation of the boosts, but only on the little group $O(3)$, viz.,

$$U(\Lambda^{-1})\phi_{sm}(p,\eta)=\sum_{m'}D^s_{m'm}(R)\phi_{sm'}(\Lambda_p,\Lambda_\eta)\quad (2.11)$$

where R is the Wigner rotation,

$$R=L^{-1}(\Lambda_p)\Lambda L(p),$$

and

$$\phi_{sm}(p,\eta)=\rho(\eta\cdot p)Y_{sm}(\theta_p,\phi_p).$$

The first approach which comes to mind for specifying the functions $\rho(\eta,p)$ is to just pick an irreducible representation of the auxilliary Lorentz group $O(3,1)$. Indeed this approach treats the two spaces $\mathcal{M}$ and H^3 on quite an equal footing, for one then diagonalizes the two invariants, the Laplace–Beltrami operators on each of the two spaces $\mathcal{M}$ and H^3. The relevant group is then the group of metric automorphisms on $\mathcal{M}\otimes H^3_+$, $\mathcal{P}_0\otimes O_I(3,1)\approx\mathcal{P}\circledS O_I(3,1)$, where $\circledS$ denotes the semidirect product and $\mathcal{P}_0$ is the "orbital" Poincaré group generated by $L_{\mu\nu}$, P_μ, and $O_I(3,1)$ is the auxiliary Lorentz group[12–14] generated by $S_{\mu\nu}$. However, it will be seen in Sec. 4 that these fields yield undesirable non-

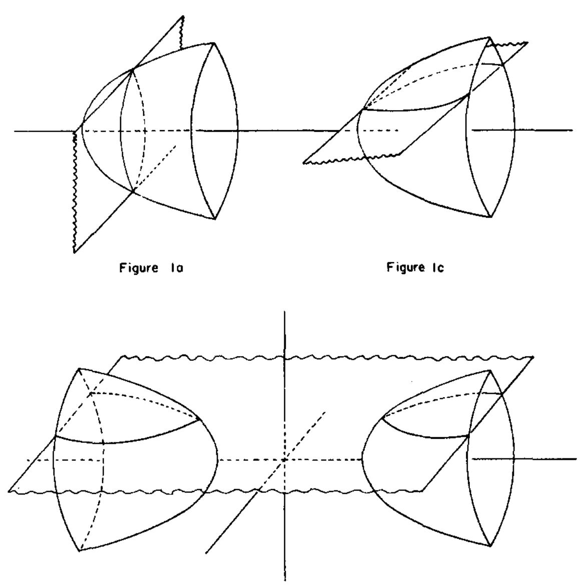


FIG. 1. The intersection of the hyperboloid H^3 by various hyperplanes corresponding to different subgroup reductions as described in the text.

local properties. Therefore, it becomes advantageous to consider highly reducible representations of $O_I(3,1)$. Perhaps the easiest way to deal with such reducible representations is to replace $O_I(3,1)$ by a larger group $G\supset O_I(3,1)$. This is done for several reasons. First, many calculations are simplified by using powerful group theoretical techniques. Second and perhaps more important is the belief that the additional symmetries introduced by a larger group represent hidden dynamical symmetries which are valid to at least first order, such as the degeneracies in the hydrogen atom spectrum.

The group $\mathcal{P}_0\otimes O_I(3,1)$ on $\mathcal{M}\otimes H^3$ is labeled by the two invariants

$$P^2=-\frac{\partial^2}{\partial x^2} \tag{2.12}$$

and

$$-\Delta(H^3)=\frac{\partial^2}{\partial a^2}+2\coth a\,\frac{\partial}{\partial a}-\frac{1}{\sinh^2 a}\,\Delta(S^2). \tag{2.13}$$

To obtain the extension[15,16] to a larger group $\mathcal{P}_0\otimes G$, we introduce the operators

$$\Gamma_\mu=-\frac{1}{2i}\,[\Delta,\eta_\mu]-\rho\eta_\mu=i\left(\frac{\partial}{\partial\eta^\mu}-\eta_\mu\eta\cdot\frac{\partial}{\partial\eta}+\sigma\eta_\mu\right) \tag{2.14}$$

$$\equiv i(\delta_\mu+\sigma\eta_\mu),\quad \sigma=-\tfrac{3}{2}+i\rho.$$

It is not difficult to show that the operator Γ_μ along with the $S_{\mu\nu}$ formally span an irreducible representation of the Lie algebra of $O(4,1)$. These representations are Hermitian with respect to $\mathcal{L}^2(H^3)$ for ρ real, and give rise to the degenerate continuous principal series of $O(4,1)$. Of course, here $G=O(4,1)$ and the conditions necessary to integrate the above representation of the Lie algebra to a representation of the group have been discussed elsewhere.[7] Instead of the operator (2.14) we could consider the operator

$$T_{\mu\nu}=-(1/2i)\,[\Delta,\eta_\mu\eta_\nu]-\rho\eta_\mu\eta_\nu$$
$$=i(\eta_\mu\delta_\nu+\eta_\nu\delta_\mu+g_{\mu\nu}+\sigma\eta_\mu\eta_\nu),\quad \sigma=-4+i\rho, \tag{2.15}$$

and obtain a twofold reducible representation of the Lie algebra of $SL(4,R)$, i.e., $G=SL(4,R)$. In both of the above cases an additional $O(4)$ symmetry has been introduced.

The basic requirement that probabilities be independent of the Lorentz frame in which they are observed demands that the representation of the Poincaré group be unitary. To construct unitary representations of $\mathcal{P}$, we consider a general inner product defined over the Fourier transformed momentum space[18] which is local in p,

$$(\phi_1,\phi_2)=\int_+\frac{d^3p}{p_0}\int\frac{d^3\eta}{\eta_0}\int\frac{d^3\eta'}{\eta_0'}$$
$$\times\phi_1^*(p,\eta')K(p\cdot\eta,p\cdot\eta';\eta\cdot\eta')\phi_1(p,\eta), \tag{2.16}$$

where K is a Poincaré invariant kernel which must, of course, allow (ϕ_1,ϕ_2) to satisfy the required properties of an inner product. We shall confine our attention to three cases of special interest.

(1)

$$(\phi_1,\phi_2)=\int_+\frac{d^3p}{p_0}\int_+\frac{d^3\eta}{\eta_0}\;\frac{\phi_1^*(p,\eta)\phi_2(p,\eta)}{(\eta\cdot p)^\lambda}, \tag{2.17}$$

where λ is real and is chosen to guarantee the proper convergence properties. With this inner product $L_{\mu\nu}$ and $S_{\mu\nu}$ in Eq. (2.2) are not separately Hermitian although their sum $M_{\mu\nu}$ is. Hence, the representation of $O_I(3,1)$ is nonunitary; an example would be the finite-dimensional representations denoted by the Gel'fand[19] notation $(0,n+1)$.

(2)

$$(\phi_1,\phi_2)=\int_+\frac{d^3p}{p_0}\int\frac{d^3\eta}{\eta_0}\;\phi_1^*(p,\eta)\phi_2(p,\eta). \tag{2.18}$$

This inner product is $\mathcal{L}^2(H_+^3\times H^3)$. The representation of both $L_{\mu\nu}$ and $S_{\mu\nu}$ are Hermitian separately. The representation of the internal $O(3,1)$ generated by $S_{\mu\nu}$ consists of two copies of the highly reducible quasiregular representation whose decomposition into irreducible parts is given by the Gel'fand–Graev transform.[20] Moreover, a unitary representation of $O(4,1)$ can be constructed on $\mathcal{L}^2(H^3)$ [14,16,17,21] whose infinitesimal generators are, along with $S_{\mu\nu}$, the Γ_μ given by Eq. (2.13), viz.,

$$U^\sigma(g^{-1})f(\eta)=\left|g^4{}_\nu\eta^\nu+g^4{}_4\right|^\sigma f\left(\frac{g^\mu{}_\nu\eta^\nu+g^\mu{}_4}{g^4{}_\nu\eta^\nu+g^4{}_4}\right), \tag{2.19}$$

where $f\in\mathcal{L}^2(H^3)$, $g\in SO_0(4,1)$.

(3)

$$(\phi_1,\phi_2)=\int_+\frac{d^3p}{p_0}\int\frac{d^3\eta}{\eta_0}\,\frac{d^3\eta'}{\eta_0'}\;\frac{\phi_1^*(p,\eta)\phi_2(p,\eta)}{(1-\eta\cdot\eta')^{\sigma+3}}\,. \tag{2.20}$$

This inner product with $\sigma=-2$ is relevant for the relativistic Coulomb problem. Again the representations of $L_{\mu\nu}$ and $S_{\mu\nu}$ are separately Hermitian. The representation of Γ_μ is Hermitian with respect to this inner product for $-3<\sigma<0$. The decomposition with respect to the $O(3,1)$ subgroup is the same as case 2 for the range $-2\leqslant\sigma\leqslant-1$, and contains an additional point for the remaining ranges.[17]

3. WAVE EQUATIONS

In the preceding section, representations of the Poincaré group were built over the space $\mathcal{M}\otimes H^3$. From such a procedure, two types of theories can be developed. First, a theory in which the wavefunctions describe a particle with a definite mass and spin. Second, a theory in which the wavefunctions contain an infinite ladder of spin states. While most of the present article is concerned with the infinite dimensional theories, a brief presentation of the finite dimensional theories is also given. This presentation for integer n seems somewhat similar in spirit to the work of Nilsson and Beskow[13] for more general representations. However, we do not get to the point of the introduction of interactions, where all the difficulties of such theories lie. Thus our discussion of these theories is very incomplete. Be that as it may, the investigation of the finite-dimensional models from our vantage point is, nevertheless, instructive.

Finite-dimensional models

We describe here only those models in which the fields describe particles of one mass M and one spin s reserving any multiplet structure for the infinite-dimensional theories. In this context, any group larger than the Lorentz group is superfluous; hence only $O(3,1)$ models are discussed.

In the standard theories on Minkowski space, one describes the spin by choosing a representation of the internal Lorentz group $O(3,1)$ in terms of spinor or tensor valued functions over $\mathcal{M}$. The mass $\mathcal{M}$ is then selected by the choice of a wave equation. However, in our theories there is a carrier space for both M and s. Thus in order to select out wavefunctions describing particles of mass M and spin s, two equations are needed. These wave equations indeed play a very symmetrical role, each being the eigenvalue problem for the Laplace—Beltrami operator on the respective spaces, $\mathcal{M}$ and H^3_+. Thus at this level the mass M and spin s are on quite an equal footing. The wave equations are then

$$(\Box + M^2)\,\phi(x,\eta) = 0, \tag{3.1}$$

$$[\Delta + n(n+2)]\,\phi(x,\eta) = 0, \tag{3.2}$$

where $\Box = (\partial/\partial x^\mu)(\partial/\partial x_\mu)$ and $\Delta = \Delta(H^3_+) = \delta^2$ has been given previously.

The rest frame solutions of Eq. (3.2) which are regular at the origin $\eta_\mu = (1,0,0,0)$ are given by

$$\phi_{nsm}(\eta) = \rho_{ns}(\cosh a)\,Y_{sm}(\theta,\phi) \tag{3.3}$$

with

$$\rho_{ns}(\cosh a) = N'(\sinh a)^{-1/2}\,P^{-(s+1/2)}_{n+1/2}(\cosh a)\,Y_{sm}(\theta,\phi)$$
$$= N \sinh a^s\, C^{s+1}_{n-s}(\cosh a)\,Y_{sm}(\theta,\phi),$$

where N and N' are the appropriate normalization factors and the functions P^μ_ν and C^μ_ν are Legendre and Gegenbauer functions[22] respectively. These solutions will be well defined with respect to the norm defined by Eq. (2.17) provided the real parameter λ satisfies $\lambda > 2\,\mathrm{Re}(n) + 2$. This is obtained by using the asymptotic expansion for Legendre functions (22). It should also be mentioned that the solution to Eqs. (3.1) and (3.2) does not describe a system with one spin s. For such a system we need a covariant spin projection operator.[23]

We now summarize the relevant material needed to construct the finite-dimensional wavefunctions over $\mathcal{M}\otimes H^3_+$ for particles of mass M and spin s, i.e., a Poincaré irreducible system: The two wave equations (3.1) and (3.2) whose solutions are given by

$$\phi(x,\eta) = \sum_s \sum_m \int_+ \frac{d^3p}{p_0}$$
$$\times [\exp(-ip\cdot x)\,\phi_{nsm}(p,\eta)\,a_{sm}(\vec{p}) + \exp(ip\cdot x)$$
$$\times \phi^*_{nsm}(p,\eta)\,b^*_{sm}(\vec{p})] \tag{3.4}$$

with

$$\phi_{nsm}(p,\eta) = N'\{[\eta\cdot(p/M)]^2 - 1\}^{-1/2}\;P^{-(S+1/2)}_{n+1/2}[(\eta\cdot p)/M]$$
$$\times Y_{sm}(\theta_p,\phi);$$

the covariant spin projection operator

$$\Theta^n_s = \Pi'_{s'}\left(\frac{W^2 - P^2 s'(s'+1)}{P^2 s(s+1) - P^2 s'(s'+1)}\right), \tag{3.5}$$

where $s' = 0,\cdots, s-1, s+1, \cdots, n$ if n is a positive integer and $s' = 0, \cdots, s-1, s+1, \cdots$ if n is otherwise; and the relevant inner product (2.17) with $\lambda > \mathrm{Re}(2n+2)$. Notice that in the case $n=0$ the ϕ functions have no η dependence and the internal space becomes innocuous; the theory reduces to ordinary Minkowski theory. Notice, too, the projection operator is an infinite product when $n \neq$ positive integer, but it is well defined on the space of solutions of Eq. (3.2).

Infinite-dimensional models

At this point it is convenient to clarify exactly what is meant by a wave equation over $\mathcal{M}\otimes H^3$. In the finite-dimensional case, we were interested in wavefunctions which described particles of one given mass M, and we obtained such by writing two wave equations, each involving the respective Laplace—Beltrami operators on $\mathcal{M}$ on H^3_+. This is a case where the wavefunctions transformed as representations of $\mathcal{P}_0 \otimes O_I(3,1) \approx \mathcal{P} \circledS O_I(3,1)$. In the infinite-dimensional case, with internal symmetry group G, there is an analogous case—the mass degenerate case. However, instead of using the Laplace—Beltrami operator on H^3 which has a purely continous spectrum with respect to the inner product [Eq. (2.17)], the H^3 wave equation analogous to Eq. (3.2) is obtained by using a covariant form of the Casimir operator of the maximal compact subgroup of G. The mass degenerate equations are then

$$(\Box + M^2)\,\phi(x,\eta) = 0, \tag{3.1}$$

$$(C_{MG} + \kappa)\,\phi(x,\eta) = 0, \tag{3.6}$$

where C_{MG} in the rest frame is the product of p^2 with the Casimir, operator of the maximal compact subgroup of G, and the fields $\phi(x,\eta)$ transform as unitary representations of $\mathcal{P} \circledS G$. For example, if $G = O(3,1)$, then $P^2 C_{O(3)} = W^2$ and $\kappa = p^2 s(s+1)$; if $G = O(4,1)$, $P^2 C_{O(4)} = \widetilde{W}^2 \equiv (P\cdot\Gamma)^2 - P^2\Gamma^2 + W^2$ and $\kappa = p^2\, n(n+2)$, n positive integer. Although the mass degenerate case is not very realistic, it does have some nice properties which makes it worthy of further discussion.

Indeed, the most interesting case physically and the most difficult mathematically is provided by one wave equation over $\mathcal{M}\otimes H^3$ which yields a desired mass spectrum. Such a wave equation breaks the symmetry under $\mathcal{P} \circledS G$ since constraining oneself to the solutions of such an equation introduces a more complicated algebraic structure.[24] Indeed a complicated algebraic structure (not forming a Lie algebra) is necessary in order to circumvent the no-go theorems[25] prohibiting the unification of internal symmetries and mass splittings. It should be mentioned that most of the wave equations that follow can be derived from a suitable Lagrangian, although this is not the procedure followed in the text.[14]

When writing down wave equations, it is much more convenient to work in the Fourier transformed momentum space. The simplest nontrivial wave equation then is one linear both in the momentum p_μ and the operator Γ_μ such as

$$(p\cdot\Gamma-\kappa)\,\phi(p,\eta)=0. \tag{3.7}$$

Such an equation, however, is unacceptable since its rest frame solutions are eigenfunctions of the noncompact operator Γ_0 given in Eq. (2.14). Thus Eq. (3.7) has a purely continuous spectrum with respect to the inner product, Eq. (2.34). Another wave equation with a purely continuous spectrum would be

$$(p^2+c_0\Delta)\,\phi(p,\eta)=0, \tag{3.8}$$

since as mentioned previously the Laplace–Beltrami operator $\Delta(H^3)$ has a purely continuous spectrum. Both Eqs. (3.7) and (3.8) are unacceptable whether the underlying space is H^3 or H^3_+.

The next simplest equation would be one which is a combination of Eqs. (3.7) and (3.8)

$$[\delta^2+c_2(p\cdot\delta)+c_1(p\cdot\eta)+c_0]\,\phi(p,\eta)=0, \tag{3.9}$$

where c_0, c_1, c_2 could depend on p^2 if desired. Such an equation is, however, somewhat unwieldy. In the rest frame putting $x=\eta_0=\cosh a$ and solving for the angular part in terms of the $Y_{sm}(\theta,\phi)$ yields for the ρ functions the equation

$$(1-x^2)\frac{d^2\rho}{dx^2}+[(1-x^2)c_2-(2s+3)x]\frac{d\rho}{dx}$$
$$+[c_0-s(s+2)+(c_1-c_2s)x]\rho=0.$$

This equation has regular singular points at $x=\pm 1$ and an irregular singular point[26] at $x=\infty$. Due to the difficulty of equations with irregular singular points, we will not obtain an explicit solution. However, it should be mentioned that an equation of the form (3.9) with $c_2=0$ has been used in the quasipotential approach to the harmonic oscillator by Donkov, Kadyshevsky, Mateev, and Mir-Kasimov.[27] In this paper, the authors obtained a solution in terms of an expansion in Whittaker functions and an approximate mass spectrum. In our case, however, rewriting Eq. (3.9) in terms of the $O(4,1)$ generator Γ_μ,

$$(\Gamma^2+ic_2p\cdot\Gamma+\sigma(\sigma+3)-c_0]\,\phi(p,\eta)=0, \tag{3.10}$$

and choosing $\sigma=c_1/c_2\neq$ positive integer suggests expanding the solutions in a series of the basis functions for an irreducible representation of $SO(4,1)$,

$$\phi_{Nsm}(p,\eta)=\sum_{n=0}^{\infty} b_n\rho_n(p\cdot\eta)Y_{sm}(\theta',\phi), \tag{3.11}$$

where the $\rho_n(p\cdot\eta)$ will be given explicitly for the $SO(4,1)$ case later. Demanding the functions (3.11) to be solutions of Eq. (3.10) will then yield a set of recursion relations involving five b terms at most, i.e., b_{n+2}, b_{n+1}, b_n, b_{n-1}, b_{n-2}. These equations could in principle be solved and along with the convergence property demanded by Eq. (2.17) possibly yield a discrete mass spectrum. The unwieldy mathematics, however, seems too high a price to pay unless there is some good physical reason for choosing Eq. (3.10) over other wave equations with simpler solutions.

Before proceeding on to the analysis of the $O(4,1)$ equations, mention is made of a model obtained by Fleming.[28] This model, obtained from a Lagrangian approach, yielded a mass spectra which is linear in the spin s. The wave equation is defined only over H^3_+ as

$$[p^2\delta^2-(p\cdot\delta)^2+3(\eta\cdot p)(p\cdot\delta)-c_1p^2-c_2p^4]\,\phi(p,\eta)=0. \tag{3.12}$$

The timelike solutions[4] normalized in $\mathcal{L}^2(H^3_+)$ can be written as[28]

$$\phi_{Nsm}=N(\eta\cdot p)^{-s+4}\;P_N^{(2,s+1/2)}\,[1-2p^2/(\eta\cdot p)^2]\;Y_{sm}(\theta',\phi), \tag{3.13}$$

where $P_N^{(2,s+1/2)}$ are Jacobi polynomials and provides the mass spectrum

$$p^2=(1/c_2)\,[(2N+3)(2s+2N+4)-c_1].$$

This model seems to have nothing at all to do with a higher group other than being Poincaré invariant, i.e., it is not written strictly in terms of the generators of any group or for that matter any Lie algebra of finite dimensionality. This arises from the fact that the commutator of δ_μ with η_ν generates higher polynomials in η, thus forming an infinite-dimensional algebra. It is, however, interesting to note that the solutions (3.13) in the rest frame form a complete orthogonal basis in $\mathcal{L}^2(H^3_+)$ which provides a discrete decomposition of the quasiregular representation of the connected component of the homogeneous Lorentz group $SO_0(3,1)$.

Needless to say many other models could be constructed which yield a discrete basis. For example, on the single sheet H^3_+, models can be built which involve the Lie algebra of say $SO(4,1)$ but which cannot be integrated to representations of the group. Such models when restricted to the compact subalgebra $so(4)$ contain only half the usual number of representations, i.e., they only involve functions even under $\eta_0\to-\eta_0$. In what follows for reasons explained previously, we shall confine our attention to models which have a group theoretical interpretation, i.e., the wave equation involves differential operators which are the generators of a certain Lie group [usually $SO(4,1)$] and the representation of the Lie algebra provided by such differential operators can be integrated to a representation of the corresponding Lie group.

O (4,1) models

A Poincaré invariant wave equation, second order in the generators of $O(4,1)$ and polynomial in p, which yields a discrete timelike spectrum is given by

$$[C(p^2)+p^\mu p^\nu S_{\mu\lambda}S^\lambda{}_\nu+p^\mu p^\nu\Gamma_\mu\Gamma_\nu+(\sigma^2+3\sigma+1)p^2]\,\phi(p,\eta)$$
$$=0, \tag{3.14}$$

where $C(p^2)$ is for now an arbitrary polynomial in p^2 from which the last term is kept separate for future convenience. It will be seen shortly that such an equation yields a mass spectrum which depends on only one quantum number n [in the timelike case, the quantum number which lables the irreducible representations of the compact group $O(4)$]. A slight modification of Eq. (3.14) could yield an equation giving a mass spectrum dependent on both n and s. Such a modification is straightforward; for example, add a term proportional to W^2 to the wave equation. We will not consider any such modifications further.

We now proceed with an analysis[8] of the spectrum of Eq. (3.14) for unitary representations of $O(4,1)$.

(i) *Timelike case*: $p^2>0$: The standard frame is $p_\mu^t=(p_0,0,0,0)$ for which the little group is $O(3)$ generated by $S_k=\epsilon_{ijk}S_{ij}/2$. In the rest frame the second Poincaré invariant W^2 is given by $p^2\vec{S}^2$, the wave equation (3.14) reduces to

$$[C(p^2)+p^2(\vec{S}^2+\vec{\Gamma}^2+1)]\,\phi(p^t,\eta)=0. \tag{3.15}$$

The operator $\vec{S}^2+\vec{\Gamma}^2+1$ is the Casimir operator of the compact subgroup $O(4)$ generated by S_i, Γ_i. Hence, in the frame $(M,0,0,0)$ the basis is given by the canonical decomposition $O(4,1)\supset O(4)\supset O(3)\supset O(2)$. The spectrum of the $O(4)$ Casimir operator $\vec{S}^2+\vec{\Gamma}^2+1$ is given by $(n+1)^2$ with $n=$positive integer. Thus Eq. (3.15) becomes

$$[C(p^2)+p^2(n+1)^2]\,\phi(p^t,\eta)=0. \tag{3.16}$$

Hence, in order to have a timelike spectrum, we must demand[29]

$$C(p^2)<0,\quad p^2\geqslant M_0^2>0. \tag{3.17}$$

(ii) *Spacelike case*: $p^2<0$: The standard frame is $p_\mu^s=(0,0,0,k)$ for which the little group is $O(2,1)$ generated by S_3, N_1, N_2. In this frame the second Poincaré invariant W^2 is given by $W^2=k^2\,(S_3^2-N_1^2-N_2^2)$ with eigenvalues $-k^2(1/4+\lambda^2)$, and the wave equation (3.14) reduces to

$$[C(p^2)+p^2(\Gamma_1^2+\Gamma_2^2+S_3^2-\Gamma_0^2-N_1^2-N_2^2+1)\,\phi(p^s,\eta)=0. \tag{3.18}$$

The operator in parenthesis is just the Casimir operator of the $O(3,1)$ subgroup generated by Γ_1, Γ_2, S_3; Γ_0, N_1, N_2. Thus the basis is provided by the subgroup decomposition $O(4,1)\supset O(3,1)\supset O(2,1)\supset O(2)$. The spectrum of the $O(3,1)$ Casimir operator is continuous[17] with $n=-1+i\nu$, ν real, except for a possible additional point for certain values of σ as mentioned previously. Thus Eq. (3.18) becomes

$$[C(p^2)+k^2\nu^2]\,\phi(p^s,\eta)=0. \tag{3.19}$$

If we wish to *avoid* spacelike solutions, we must choose

$$C(p^2)>0,\quad p^2<0. \tag{3.20}$$

(iii) *Lightlike case*: $p^2=0$, $p_\mu\neq 0$: The standard frame is $p_\mu^l=(k,0,0,-k)$ for which the little group is $E(2)$ generated by S_2+N_1, N_2-S_1, S_3. The Poincaré invariant becomes

$$W^2=k^2\,[(N_1+S_2)^2+(N_2-S_1)^2]$$

with eigenvalues $k^2\kappa^2$, $\kappa>0$ and the wave equation reduces to

$$[C_0+k^2\,((\Gamma_0+\Gamma_3)^2+(N_1+S_2)^2+(N_2-S_1)^2)]\ \phi(p^l,\eta)=0 \tag{3.21}$$

with $C_0=C(p^2=0)$. The operator in parenthesis in Eq. (3.21) is the Casimir operator of the $E(3)$ subgroup generated by $\Gamma_0-\Gamma_3$, S_2+N_1, S_1-N_2; Γ_1, Γ_2, S_3, and the basis states are given by the subgroup decomposition $O(4,1)\supset E(3)\supset E(2)\supset O(2)$. The spectrum of the $E(3)$ Casimir operator is continuous and labeled by a nonvanishing real number ϵ. Hence Eq. (3.21) reduces to

$$[C_0+k^2\epsilon^2]\,\phi(p^l,\eta)=0. \tag{3.22}$$

In order to *avoid* lightlike solutions, we must take $C_0\geqslant 0$.

(iv) *Vacuumlike case*: $p_\mu\equiv 0$: The little group is the entire homogeneous Lorentz group $O(3,1)$. Equation (3.14) is satisfied by any $\phi(\eta)$ in the representation space. In order to *avoid* vacuumlike solutions we must have $C_0\neq 0$.

It is interesting to note that the $C(p^2)$ which yields a linearly rising trajectory in the discrete quantum number n is given by the CGL equation obtained by choosing

$$C(p^2)=-\beta_1^{-2}p^2(p^2-\beta)^2\quad\text{with } \beta_1>0, \tag{3.23}$$

and thus avoids spacelike and lightlike solutions but has vacuumlike solutions. We can get rid of the vacuumlike solutions with a slight deviation from linearty of the trajectories by choosing

$$C(p^2)=C_0-\beta_1^{-2}p^2(p^2-\beta)^2 \tag{3.24}$$

with C_0 small and positive. With such a choice, the mass spectrum has discrete timelike solutions only and satisfies the cubic equation in p^2,

$$p^6-2\beta p^4+[\beta^2-\beta_1^2(n+1)^2]\ p^2-\beta_1^2C_0=0. \tag{3.25}$$

Now in the limit $C_0\to 0$, Eq. (3.25) yields the CGL mass spectrum

$$p^2=\pm\beta_1(n+1)+\beta. \tag{3.26}$$

To avoid the negative branch completely[30] we must take $\beta_1>\beta$. This will provide the zeroeth order solution. Let p_1^2, p_2^2, p_3^2 be the roots of Eq. (3.25). They must satisfy the following equations:

$$\begin{aligned} &p_1^2p_2^2p_3^2=\beta_1^2C_0,\\ &p_1^2+p_2^2+p_3^2=2\beta,\\ &p_1^2p_2^2+p_1^2p_3^2+p_2^2p_3^2=\beta^2-\beta_1^2(n+1)^2. \end{aligned} \tag{3.27}$$

Satisfying these equations to first order in C_0 yields the roots

$$p_3^2=-\frac{C_0}{(n+1)^2-(\beta/\beta_1)^2},$$

$$p_1^2=\beta_1(n+1)+\beta+\frac{C_0}{2(n+1)(n+1+\beta/\beta_1)}, \tag{3.28}$$

$$p_2^2=-\beta_1(n+1)+\beta+\frac{C_0}{2(n+1)(n+1-\beta/\beta_1)}. \tag{3.29}$$

p_2^2 and p_3^2 will be negative for all solutions if we demand

$$(\beta_1-\beta)^2>C_0\beta_1/2\quad\text{and}\quad\beta_1>\beta. \tag{3.30}$$

Thus it is seen by choosing $C(p^2)$ as in Eq. (3.24) with the parameters constrained by Eq. (3.30), one obtains an infinitely rising mass spectrum with discrete timelike states only given by

$$p^2=M_n^2=\beta_1(n+1)+\beta+\frac{C_0}{2(n+1)(n+1+\beta/\beta_1)}. \tag{3.31}$$

In several ways this is a very desirable spectrum: It has discrete timelike solutions only, thus avoiding the vacuumlike solutions of the CGL equation; it is approximately linear.

However, our preceding analysis is in many ways independent of the specific form of the polynomial $C(p^2)$; therefore, it is desirable not to specify $C(p^2)$ except when referring to a particular case. Indeed, one ultimately wishes to understand the relationship between the completeness problem, which has a direct bearing on whether the theory is local or not, and the existence of spacelike solutions. For this as well as pedagogical reasons, it is desirable to keep $C(p^2)$ quite general. With this in mind we next write down the single particle states for the cases previously discussed. The relevant mathematics is performed in Ref. 14. The timelike single particle states are up to a normalization factor given in the rest frame by

$$|nsm;\sigma\rangle_t \sim \exp(-ip_0x_0)\cosh^\sigma a \tanh^s a\, C^{s+1}_{n-s}(\pm 1/\cosh a) \times Y_{sm}(\theta,\phi). \tag{3.32}$$

The states with arbitrary momentum are obtained from the rest frame states by applying the Wigner boost, $\exp(-i\vec{\alpha}\cdot\vec{N})$, taken here for simplicity to be in the third direction, so that the parameter α is defined by

$$\cosh\alpha = p_0/(p^2)^{1/2}, \qquad \sinh\alpha = q/(p^2)^{1/2} \tag{3.33}$$

with $p_\mu = (p_0, 0, 0, q)$. With this prescription we find for a state with momentum $(p_0, 0, 0, q)$

$$|q;nsm;\sigma\rangle_t \sim \exp(-ip\cdot x)\left(\frac{\eta\cdot p}{(p^2)^{1/2}}\right)^\sigma \left(\frac{(\eta\cdot p)^2 - p^2}{(\eta\cdot p)^2}\right)^{s/2} C^{s+1}_{n-s} \times\left(\frac{(p^2)^{1/2}}{\eta\cdot p}\right) Y_{sm}(\theta',\phi) \tag{3.34}$$

with

$$\tan\theta' = \frac{\sinh a \sin\theta}{\pm\sinh\alpha\cosh a + \cosh\alpha\cosh a\cos\theta}.$$

It must be remembered that p_0 and p^2 and hence α are functions of n through the mass spectrum imposed by the wave equation.

However, for the spacelike and lightlike cases, the spherical coordinate system, Eq. (2.5), does not separate the relevant Casimir operators. Hence in the spacelike case we introduce the hyperbolic coordinate system defined by

$$\begin{aligned} \eta_0 &= \pm\cosh a\cosh b, & -\infty < a < \infty,\\ \eta^1 &= \cosh a \sinh b \sin\theta, & \\ \eta^2 &= \cosh a\sinh b\cos\theta, & 0 \le b < \infty,\\ \eta^3 &= \sinh a, & 0\le\theta<2\pi. \end{aligned} \tag{3.35}$$

Notice that this coordinate system corresponds to intersecting the hyperboloid with a hyperplane $\eta^3 = \pm$ const shown in Fig. 1b. The single particle states for the spacelike case are given in the standard frame by

$$|k;\nu\lambda m;\sigma\rangle_s \sim \exp(-ikx^3)\sinh^\sigma a\,(\coth a)^{-1/2+i\lambda} \times[A\, C^{i\lambda+1/2}_{-1/2+i\nu-i\lambda}(i/\sinh a) + B\, D^{i\lambda+1/2}_{-1/2+i\nu-i\lambda}(i/\sinh a)] \times P^m_{-1/2+i\lambda}(\cosh b)\exp(-im\phi). \tag{3.36}$$

The functions $D^{i\lambda+1/2}_{-1/2+i\nu-i\lambda}$ are Gegenbauer functions of the second kind, and A and B are constants. Both the Gegenbauer and Legendre functions appearing in this expression are the appropriate analytic continuation of the same functions appearing in Eq. (3.32) in both the discrete indices n and s as well as the relevant arguments of the functions themselves. The explicit details of this continuation are discussed in Ref. 14. The states with momentum $(p_0, 0, 0, q)$ are again obtained from the standard frame states by applying the Wigner boost. In this case the boost parameter α is defined by

$$\cosh\alpha = q/(-p^2)^{1/2}, \qquad \sinh\alpha = p_0/(-p^2)^{1/2}. \tag{3.37}$$

Then the states at momentum $(p_0, 0, 0, q)$ become

$$|p;\nu\lambda m;\sigma\rangle_s \sim \exp(-ip\cdot x)\left(\frac{\eta\cdot p}{(-p^2)^{1/2}}\right)^\sigma\left(\frac{(\eta\cdot p)^2+p^2}{(\eta\cdot p)^2}\right)^{-1/2+i\lambda} \times\left[A\,C^{i\lambda+1/2}_{-1/2+i\nu-i\lambda}\left(\frac{(p^2)^{1/2}}{\eta\cdot p}\right) + B\,D^{i\lambda+1/2}_{-1/2+i\nu-i\lambda}\left(\frac{(p^2)^{1/2}}{\eta\cdot p}\right)\right] \times P^m_{-1/2+i\lambda}(\cosh b')\exp(-im\phi) \tag{3.38}$$

with

$$\tanh b' = \frac{\cosh a\sinh b}{\pm\cosh\alpha\cosh a\cosh b + \sinh\alpha\sinh a}.$$

Again it must be remembered that p^2, p_0, and α are functions of ν.

For a discussion of the lightlike case, the parabolic coordinates of H^3 are appropriate:

$$\begin{aligned} \eta_0 &= \pm(\cosh a + \tfrac{1}{2}r^2e^a), & 0\le a<\infty,\\ \eta^1 &= r\,e^a\sin\phi, & \\ \eta^2 &= r\,e^a\cos\phi, & 0\le r<\infty,\\ \eta^3 &= \mp(\sinh a - \tfrac{1}{2}r^2e^a), & 0\le\phi<2\pi. \end{aligned} \tag{3.39}$$

This system corresponds to the intersection of the hyperboloid with a hyperplane parallel to the light-cone defined by $\eta^0 - \eta^3 =$ const as shown in Fig. 1c. The single particle states for the lightlike case are given in the standard frame by

$$|k;\epsilon\kappa m;\sigma\rangle_l \sim \exp[-ik(x_0 - x^3)]\,\exp[a(\sigma-1/2)]J_{1/2}((\epsilon^2-\kappa^2)^{1/2}e^{-a}) \times J_m(\kappa r)\exp(-im\phi), \tag{3.40}$$

where the J's are Bessel functions.[22] Again by applying the Wigner boost parameterized now by

$$\alpha = \ln(p_0/k), \tag{3.41}$$

one obtains

$$|p;\epsilon\kappa M;\sigma\rangle_l \sim \exp(-ip\cdot x)\left(\frac{p\cdot\eta}{k}\right)^{\sigma-1/2} J_{1/2}\left(\frac{(\epsilon^2-\kappa^2)^{1/2}k}{\eta\cdot p}\right)J_m(e^{-\alpha}\kappa r) \times\exp(-im\phi). \tag{3.42}$$

Before ending this section, we briefly mention various $O(4,2)$ models obtained by others.[31] For the representations $\sigma=-2,\ -1$ of the group $O(4,1)$, the $O(4,1)$ algebra can be extended to $O(4,2)$ by considering $L_0 \equiv(\vec{S}^2+\vec{\Gamma}^2+1)^{1/2}$, $L_i\equiv i[S_{0i},L_0]$, and $L_4\equiv -i[\Gamma_0,L_0]$. These operators form an $O(4,1)$ vector and close weakly with the generators of $O(4,1)$ to form an $O(4,2)$ algebra only for the above-mentioned representations. As an operator on $\mathcal{L}^2(H^3)$, $1/L_4$ is an integral operator whose kernel is $1/(\eta-\eta')^2$ and $L_\mu = L_4\eta_\mu$. Hence, many wave equations employing L_μ, L_4 can be written down.[31] Such wave equations differ from those like Eq. (3.14) since they have a continuous contribution to the timelike portion of the spectrum. In the quasipotential approach,[31] the above-mentioned kernel is related in an obvious way to the Fourier transform of a scalar Coulomb potential.

4. QUANTUM FIELD THEORY

In this section the passage from the unquantized to the second quantized theory[32] is made via the standard relativistic Fock space. That is, the Fourier coefficients occurring in the various decompositions in the previous sections are considered as the creation and annihilation operators for the corresponding single particle states of the theory. No rigorous justification in terms of smeared fields is provided here, however, and only the quantization of timelike states is treated. The quantization proceeds as

$$[a_{nsm}(\vec{p}),\ a^\dagger_{nsm}(\vec{p}')]=[b_{nsm}(\vec{p}),\ b^\dagger_{nsm}(\vec{p}')] = p_0\delta_{nn'}\delta_{ss'}\delta_{mm'}\delta^3(\vec{p}-\vec{p}') \tag{4.1}$$

with all other commutators identically zero. The single particle states are then defined as

$$a^\dagger_{nsm}(\vec{p})|0\rangle = |\vec{p};nsm\rangle \tag{4.2}$$

and

$$b^\dagger_{nsm}(\vec{p})|0\rangle = |\vec{p};nsm\rangle \tag{4.3}$$

for antiparticles.

Notice the quantization proceeded using commutation relations, as opposed to anticommutation relations, which is the usual case for bosons. The question arises as to whether or not there exists a spin-statistics theorem which demands the use of the commutation relations. In the usual theory,[32,33] this is a consequence of the demand for locality of the fields. The theorem states that the fields are local if and only if integer spin fields are quantized by commutation relations and half-odd integer spin fields are quantized by anticommutation relations, and only applies when *finite*-dimensional representations of the auxiliary Lorentz group are used.

However, before the spin statistics problem is discussed, we investigate the discrete transformations.

Discrete transformations

Up until now, we have not specified the transformation properties of the fields under the discrete transformations, charge conjugation **C**, parity **P**, and time reversal **T**, except to say that for theories defined over both sheets of H^3, **P**, and **T** are members of the complete Poincaré group.[34] Indeed, on the relativistic Fock space of physical particle states, we know up to a phase factor what these transformations should do. Furthermore, the transformation properties of the Minkowski space are well known; however, there is some ambiguity with regard to the internal space H^3, especially regarding time reversal. If we demand that time reversal is a member of the complete homogeneous Lorentz group with the representation specified by Eq. (2.1), then both sheets of H^3 are needed and under T

$$x_0\to -x_0 \quad \vec{x}\to\vec{x},\quad \eta\to-\eta_0,\quad \vec{\eta}\to\vec{\eta}.$$

However, if only a single sheet is used, one can define **T** such that

$$\eta_0\to\eta_0 \text{ and } \vec{\eta}\to-\vec{\eta}.$$

In both cases discrete transformations with the correct properties regarding the physical single particle states can be defined which also yield a CPT invariant field theory. The single particle annihilation operators are now defined to have the following discrete transformation properties:

parity:

$$\begin{aligned}\mathbf{P}\,a_{nsm}(\vec{p})\mathbf{P}^{-1}&=(-1)^s\zeta_p\,a_{nsm}(-\vec{p}),\\ \mathbf{P}\,b_{nsm}(\vec{p})\mathbf{P}^{-1}&=(-1)^s\zeta_p^*\,b_{nsm}(-\vec{p});\end{aligned} \tag{4.4}$$

charge conjugation:

$$\begin{aligned}\mathbf{C}\,a_{nsm}(\vec{p})\mathbf{C}^{-1}&=\zeta_c\ b_{nsm}(\vec{p}),\\ \mathbf{C}\,b_{nsm}(\vec{p})\mathbf{C}^{-1}&=\zeta_c^*\ a_{nsm}(\vec{p});\end{aligned} \tag{4.5}$$

time reversal (both sheets):

$$\begin{aligned}\mathbf{T}\,a_{nsm}(\vec{p})\mathbf{T}^{-1}&=(-1)^{n-s+m}\ \zeta_T\ a_{ns-m}(-\vec{p}),\\ \mathbf{T}\,b_{nsm}(\vec{p})\mathbf{T}^{-1}&=(-1)^{n-s+m}\zeta_T^*\ b_{ns-m}(-\vec{p});\end{aligned} \tag{4.6}$$

time reversal (single sheet):

$$\begin{aligned}\mathbf{T}\,a_{nsm}(\vec{p})\mathbf{T}^{-1}&=(-1)^{s+m}\zeta_T\ a_{ns-m}(-\vec{p}),\\ \mathbf{T}\,b_{nsm}(\vec{p})\mathbf{T}^{-1}&=(-1)^{s+m}\zeta_T^*\ b_{ns-m}(-\vec{p}).\end{aligned} \tag{4.7}$$

The transformation properties for the creation operators can be obtained from the above. Of course, the operators **P** and **C** are unitary whereas **T** is antiunitary. It should also be mentioned that the above construction of the charge conjugate operator will have the required properties when interactions with the electromagnetic field are introduced, i.e., the em current and charge change sign under charge conjugation. In addition to the above properties, we find explicitly from the wavefunctions of Eqs. (3.3) and (3.34)

$$\phi_{nsm}(-\vec{p};\ \eta_0,\ -\vec{\eta})=(-1)^s\phi_{nsm}(p,\eta) \tag{4.8}$$

and for $O(4,1)$ theories (double sheet)

$$\begin{aligned}\phi^\sigma_{ns-m}(-\vec{p};\ -\eta_0,\ \vec{\eta})&=(-1)^{n-s+m}(\phi^{\sigma*}_{nsm}(p,\eta))^*\\ &\equiv(-1)^{n-s+m}\ \tilde{\phi}^*_{nsm}(p,\eta),\end{aligned} \tag{4.9}$$

whereas for $O(3,1)$ theories (single sheet)

$$\begin{aligned}\phi_{nsm}(-\vec{p};\eta_0,-\eta)&=(-1)^{s+m}(\phi^*_{n^*sm}(p,\eta))^*\\ &\equiv(-1)^{s+m}\ \tilde{\phi}^*_{nsm}(p,\eta).\end{aligned} \tag{4.10}$$

Equation (4.9) necessitates the doubling of the representation space to incorporate time reversal in $O(4,1)$ theories transforming as a member of the principal series, $\sigma=-\frac{3}{2}+i\rho, \rho\neq 0$, and real and only if σ is real is doubling unnecessary. Equation (4.10) implies that if n is real, no doubling is necessary, but also if $n=-1+i\nu$ it is seen that again no doubling is necessary due to the relation[22]

$$P_\nu^\mu(z)=P_{-\nu-1}^\mu(z)$$

for Legendre functions. Thus for $O(3,1)$ theories doubling of the representation space to include time reversal is only necessary when n is complex and $\neq -1\pm i\nu$, i.e., only for certain nonunitary infinite dimensional representations. Also, Eq. (4.8) implies that the parity of the particles alternate with spin.

The decomposition for the second quantized Poincaré irreducible theories then follows:

$$\begin{aligned}\phi_s^{(n)}(x,\eta)&=\Theta_s^{(n)}\phi^{(n)}(x,\eta)\\&=\sum_{m=-s}^{s}\int\frac{d^3p}{p_0}\,[\exp(-ip\cdot x)\,\phi_{nsm}(p,\eta)\,a_{sm}(\vec{p})\\&\quad+\exp(ip\cdot x)\,\tilde{\phi}^*_{nsm}(p,\eta)\,b^\dagger_{sm}(p)],\\\phi_s^{(n)\dagger}(x,\eta)&=\sum_{m=-s}^{s}\int\frac{d^3p}{p_0}\times[\exp(ip\cdot x)\,\phi^*_{nsm}(p,\eta)a^\dagger_{sm}(\vec{p})\\&\quad+\exp(-ip\cdot x)\,\tilde{\phi}_{nsm}(p,\eta)\,b_{sm}(\vec{p})].\end{aligned}\tag{4.11}$$

Under the **C**, **P**, and **T** transformations, one then finds, combining Eq. (4.11) with Eqs. (4.4), (4.5), (4.7), (4.8), and (4.10),

$$\begin{aligned}\mathbf{P}\,\phi(x,\eta)\mathbf{P}^{-1}&=\zeta_p\phi(x_0,-x;\eta_0,-\eta),\\\mathbf{C}\,\phi(x,\eta)\mathbf{C}^{-1}&=\zeta_c\phi^\dagger(x,\eta),\\\mathbf{T}\,\phi(x,\eta)\mathbf{T}^{-1}&=\zeta_T\,\phi(-x_0,x;\eta_0,-\eta),\end{aligned}\tag{4.12}$$

and similarly for the Hermitian conjugate field with the phase factors replaced by their complex conjugates.

For the $O(4,1)$ infinite multiplet fields, one has

$$\begin{aligned}\phi(x,\eta)&=\sum_{nsm}\int\frac{d^3p}{p_0}\,[\exp(-ip\cdot x)\,\phi_{nsm}(p,\eta)\,a_{nsm}(\vec{p})\\&\quad+\exp(ip\cdot x)\,\tilde{\phi}^*_{nsm}(p,\eta)\,b^\dagger_{nsm}(\vec{p})],\\\phi^\dagger(x,\eta)&=\sum_{nsm}\int\frac{d^3p}{p_0}\,[\exp(+ip\cdot x)\,\phi^*_{nsm}(p,\eta)\,a^\dagger_{nsm}(\vec{p})\\&\quad+\exp(-ip\cdot x)\,\tilde{\phi}_{nsm}(p,\eta)b_{nsm}(\vec{p})]\end{aligned}\tag{4.13}$$

and similarly one obtains for the **C**, **P**, and **T** transformed fields the first two of Eqs. (4.12) with the **T** transformation replaced by

$$\mathbf{T}\,\phi(x,\eta)\mathbf{T}^{-1}=\zeta_T\;\phi(-x_0,x;\eta_0,\eta).\tag{4.14}$$

Equations (4.14) and (4.12) express the CPT invariance of the corresponding theories. However, it is anticipated that only for the finite-dimensional representations of the auxiliary Lorentz group will a TCP theorem be valid.[33,35] We will see in the next section that the $O(4,1)$ matrix elements are singular when analytically continued to the point which would correspond to a CPT reflection, indicating the invalidity of the TCP theorem[35] in this case.

Spin, statistics, and locality

Let us now compute what is usually referred to as the causal commutator,

$$i\Delta_\pm(x-x';\eta,\eta')\equiv[\phi(x,\eta),\,\phi^\dagger(x',\eta')]_\pm.\tag{4.15}$$

The c-number property of this commutator along with the covariance of the fields implies that as indicated it is a function of x and x' through $x-x'$ only as well as the condition

$$\Delta_\pm(x-x';\eta,\eta')=\Delta_\pm(\Lambda^{-1}(x-x');\Lambda^{-1}\eta,\Lambda^{-1}\eta')\tag{4.16}$$

for any homogeneous Lorentz transformation Λ. Combining Eqs. (4.1) and (4.11) or (4.13) with Eq. (4.15), one finds

$$i\Delta_\pm(x;\eta,\eta')=\int\frac{d^3p}{p_0}\,(\exp(-ip\cdot x)\pm\exp(ip\cdot x))\,G(p;\eta,\eta'),\tag{4.17}$$

where

$$G(p;\eta,\eta')=\sum_{m=-s}^{s}\phi_{nsm}(p,\eta)\phi^*_{nsm}(p,\eta')$$

or

$$\begin{aligned}i\Delta_\pm(x;\eta,\eta')&=\sum_{n=0}^{\infty}\int\frac{d^3p}{p_0}\,[\exp(-ip\cdot x)\pm\exp(ip\cdot x)]\\&\quad\times G^{(n)}(p;\eta,\eta'),\end{aligned}\tag{4.18}$$

$$G^{(n)}(p,\eta,\eta')=\sum_{s=0}^{\infty}\sum_{m=-s}^{s}\phi_{nsm}(p,\eta)\phi^*_{nsm}(p,\eta'),$$

depending on which type theory is used. Also, Eq. (4.16) implies for the G functions

$$G(p,\eta,\eta')=G(\Lambda^{-1}p,\Lambda^{-1}\eta,\Lambda^{-1}\eta).\tag{4.19}$$

Consider first the finite-dimensional Poincaré irreducible theories. Now the G-function given in Eq. (4.17) can be evaluated using the explicit expression for the ϕ_{nsm} given by Eq. (3,3) along with the well-known addition theorem for spherical harmonics,[31] yielding

$$\begin{aligned}G(p;\eta,\eta')&=\frac{(2s+1)}{4\pi}\,|N_{n,s}|^2\left[\left(\frac{\eta\cdot p}{M}\right)^2-1\right]^{s/2}\\&\quad\times\left[\left(\frac{\eta'\cdot p}{M}^2\right)-1\right]^{s/2}\\&\quad\times C_{n-s}^{s+1}\left(\frac{\eta\cdot p}{M}\right)C_{n-s}^{s+1}\left(\frac{\eta'\cdot p}{M}\right)P_s(\cos\beta_p)\end{aligned}\tag{4.20}$$

where

$$\cos\beta_p=\cos\theta_p\cos\theta'_p+\sin\theta_p\sin\theta'_p\cos(\phi_p-\phi'_p)$$

and the subscripts p on the angular variables indicate that the rest frame variables for η and η' have been boosted to momentum p. In general, the function G is not a polynomial in p, and $\Delta_\pm$ cannot be expressed as certain derivatives of a causal function; hence the theory is nonlocal. However, if the parameter n is a positive integer, the theory is local corresponding to the usual spinor type theories.[33] This can be seen by rewriting Eq. (4.20) after some algebraic manipulations as

$$G(p;\eta,\eta')=\frac{(2s+1)}{\pi 2^{s+2}}\,|N_{n,s}|^2C_{n-s}^{s+1}\left(\frac{\eta\cdot p}{M}\right)C_{n-s}^{s+1}\left(\frac{\eta'\cdot p}{M}\right)$$

$$\times \sum_{k=0}^{[s/2]} \binom{s}{k}\binom{2s-2k}{s} [(\eta\cdot p)(\eta'\cdot p)-M^2\eta\cdot\eta']^{s-2k}$$
$$\times[(\eta\cdot p)^2-M^2]^k\,[(\eta'\cdot p)^2-M^2]^k, \qquad (4.21)$$

which can be readily seen as an even polynomial in p_μ when n is a positive integer. Thus, this case reproduces the usual results with the spin statistics theorem remaining intact. However, when n is not a positive integer the Gegenbauer functions in Eq. (4.20) are not polynomials in $\eta\cdot p$ and the theory is nonlocal. To study the nonlocal behavior, we extract the polynomial factors from the integral for Δ in the usual manner yielding

$$i\Delta(x;\eta,\eta')=P\left(\eta\cdot\frac{\partial}{\partial x},\eta'\cdot\frac{\partial}{\partial x}\right)\int\frac{d^3p}{p_0}\,[\exp(ip\cdot x)$$
$$-\exp(-ip\cdot x)]\times C^{s+1}_{n-s}\left(\frac{\eta\cdot p}{M}\right)C^{s+1}_{n-s}\left(\frac{\eta'\cdot p}{M}\right), \qquad (4.22)$$

where the function P indicates a polynomial operator in $\eta\cdot\partial/\partial x$ and $\eta'\cdot\partial/\partial x$ each of degree s and is not to be confused with the Legendre polynomials occurring previously.

In order to illustrate explicitly the nonlocal behavior, we put $n=-1$ and $s=0$, in which case the relevant integral becomes

$$\int\frac{d^3p}{p_0}\,[\exp(ip\cdot x)-\exp(-ip\cdot x)]$$
$$\left[\left(\frac{\eta\cdot p}{M}\right)^2-1\right]^{-1/2}\left[\left(\frac{\eta'\cdot p}{M}^2\right)-1\right]^{-1/2}. \qquad (4.23)$$

If in addition we put $\eta=\eta'=(1,0,0,0)$, the integral becomes much more manageable, reducing to

$$\int\frac{d^3p}{p_0}\left(\frac{\exp(ip\cdot x)-\exp(-ip\cdot x)}{\vec{p}^2}\right),$$

which can be written, using the usual manipulations, [31] as

$$\int_0^\infty\frac{dk}{(k^2+M^2)^{1/2}}\,\frac{\sin kr}{kr}\,\sin(k^2+M^2)^{1/2}\,x_0,$$

where

$$r=(x_1^2+x_2^2+x_3^2)^{1/2}.$$

The difference between this expression and the usual causal function is the appearance of the k in the denominator of the integrand which can be extracted via the equality

$$\int_0^\infty\frac{dk}{(k^2+M^2)^{1/2}}\,\frac{\sin kr}{kr}\,\sin(k^2+M^2)^{1/2}x_0$$
$$=\frac{1}{r}\int_0^r dr'\int_0^\infty\frac{dk}{(k^2+M^2)^{1/2}}\cos kr'\,\sin(k^2+M^2)^{1/2}x_0$$
$$\propto\frac{1}{r}\int_0^r dr'\,J_0(M(x_0^2-r'^2)^{1/2})\begin{cases}1, & x_0>r'\\ 0, & -r'<x_0<r'\\ -1, & x_0<-r'\end{cases} \qquad (4.24)$$

with J_0 the zeroth-order Bessel function. Clearly, the above expression does not vanish for arbitrary spacelike x, i.e., $-r<x_0<r$. It is not difficult to perform the integral for spacelike x, yielding[36]

$$\Delta_-\propto\frac{1}{r}\int_0^{x_0}J_0(M(x_0^2-r^2)^{1/2})=\frac{\sin Mx_0}{r},\quad r>x_0. \qquad (4.25)$$

Hence, we find causal behavior only for $x_0=0$, $\pm\pi/M$, $\pm 2\pi/M$, $\cdots$, that is, the function $\Delta_-(x;\eta,\eta')$ vanishes for $x^2<0$, $x_0=0$, $\pm\pi/M$, $\cdots$, $\eta=\eta'=(1,0,0,0)$, but not for arbitrary spacelike x, where the period of oscillation is of the order of the strong interaction time, 10^{-23} sec, a curious effect indeed. Although the above acausal behavior removes the light cone divergence present in ordinary theories, it exhibits undersirable long range features, i.e., it vanishes only as $1/r$ for large spacelike separation. The case for $n\neq$ integer or -1 is somewhat more tedious, but the net results are the same. For example, for $s=0$, we find

$$i\Delta(x;\eta,\eta')\propto\int\frac{d^3p}{p_0}\,[\exp(ip\cdot x)-\exp(-ip\cdot x)]$$
$$\times\sinh\left((n+1)\frac{\eta\cdot p}{M}\right)\sinh\left((n+1)\frac{\eta'\cdot p}{M}\right)$$
$$\times\left[\left(\frac{\eta\cdot p}{M}\right)^2-1\right]^{-1/2}\left[\left(\frac{\eta'\cdot p}{M}^2\right)-1\right]^{-1/2}. \qquad (4.26)$$

The trick, as in the polynomial case, is to treat the divergent factors in the numerator of the integrand as derivative with respect to x, although now this procedure is very dubious due to the high degree of divergence except for $n=-1+i\nu$. However, in this way Eq. (4.26) can be written formally as

$$i\Delta_-(x,1,1)\sim\cosh\left(2(n+1)i\frac{d}{dMx_0}\right)\frac{\sin Mx_0}{r}$$
$$=\cosh[2(n+1)]\,\frac{\sin Mx_0}{r}. \qquad (4.27)$$

The crucial point is that the r dependence remains the same. It is not expected that the higher spin case will change this result much. For the above and aforementioned reasons, we reject such theories transforming irreducibly under $\mathcal{P}$ or $\mathcal{P}_0\otimes O_I(3,1)$.

It will now be seen that nonlocal behavior in $\mathcal{P}_0\otimes O(4,1)$ type theories is much more palatable. The quantization as in the previous case proceeds via Eq. (4.1) with the single particle states given by Eqs. (4.2) and (4.3). Of course, the corresponding wavefunctions are different according to one's choice; thus the quantum number n can have a different meaning, accordingly. The decomposition of the fields valid for the $O(4,1)$ theories without spacelike or lightlike solutions satisfying the wave, Eq. (3.14), is given by Eq. (4.13) with $\phi_{nsm}(p,\eta)$ given in Eq. (3.34). Of course, if there are spacelike or lightlike solutions, the sums over n and s become integrals and the corresponding wavefunctions are given by the appropriate analytic continuation as discussed previously. Also in the case of spacelike solutions, the momentum integral is over the single sheeted hyperboloid and the standard interpretation of the antiparticle states is destroyed.[37]

We first discuss the locality of the exact symmetry theory, i.e., the fields transform as irreducible unitary representations of $\mathcal{P}_0 \otimes O(4,1)$ and hence are mass degenerate. It is known[38] that such theories may or may not be local and may or may not enjoy the usual spin statistics result. Upon calculation of the commutator of the fields, we find

$$i\Delta_{\mp}(x-x';\eta,\eta')$$
$$=2^{-1/2}(2\pi)^{-3/2}\int \frac{d^3p}{p_0}\,[\exp(-ip\cdot x)\mp\exp(ip\cdot x)]\,\delta^{\mathrm{hyp}}(\eta-\eta')$$
$$=\delta^{\mathrm{hyp}}(\eta-\eta')\,i\Delta_{\mp}(x-x'), \tag{4.28}$$

where

$$\delta^{\mathrm{hyp}}(\eta-\eta')=\delta(\cosh a-\cosh a')\,\delta(\cos\theta-\cos\theta')\,\delta(\phi-\phi').$$

Again the theory is local upon quantization with commutation relations only, and if we combine this result with the corresponding result from a spinor theory[28] over $\mathcal{M}\otimes H^3$, we obtain the usual spin statistics theorem.

Notice the above results made used of the completeness of the functions $\phi_{nsm}(p,\eta)$ on the hyperboloid H^3. Such a result is easy to come by for the mass degenerate case, since it just reduces to the completeness of the rest frame states; however, when the masses are split the completeness problem is much more complicated.[39] Consider now the above commutator for the mass nondegenerate case. Now $G^{(n)}(p;\eta,\eta')$ given by Eq. (4.18) can be written with the aid of the addition theorems for Gegenbauer polynomials,[22] as

$$G^{(n)}(p,\eta,\eta')$$
$$=\frac{n+1}{(2\pi)^5}\left(\frac{\eta\cdot p_n}{M_n}\right)^{\sigma}\left(\frac{\eta'\cdot p_n}{M_n}\right)^{\sigma^*}C_n^1\left\{\frac{M_n}{(\eta\cdot p_n)}\,\frac{M_n}{(\eta'\cdot p_n)}\right.$$
$$\left.+\left[1-\left(\frac{M_n}{\eta\cdot p_n}\right)^2\right]^{1/2}\left[1-\left(\frac{M_n}{\eta'\cdot p_n}\right)^2\right]^{1/2}\cos\beta_b\right\}. \tag{4.29}$$

Again specializing to $\eta=\eta'$, Eq. (4.18) becomes

$$i\Delta_{\mp}(x;\eta,\eta)=(2\pi)^{-5}\sum_{n=0}^{\infty}(n+1)^2$$
$$\int\frac{d^3p}{p_0}[\exp(-ip\cdot x)\mp\exp(ip\cdot x)]\left(\frac{\eta\cdot p_n}{M_n}\right)^{\sigma+\sigma^*}. \tag{4.30}$$

Furthermore, taking $\sigma=-\frac{3}{2}+i\rho$ [principal series of $O(4,1)$] and $\eta=(1,0,0,0)$ and using the standard techniques, Eq. (4.30) becomes, for spacelike x,

$$\Delta_-(x;1,1)=-\sum_{n=0}^{\infty}(2\pi)^{-4}\frac{(n+1)^2}{r}M_n^3$$
$$\times\int_0^{\infty}\frac{p\,dp\,\sin(x_0\sqrt{p^2+M_n^2})}{(p^2+M_n^2)^2}\sin pr$$
$$=2^{-5}\pi^{-3}\sum_{n=0}^{\infty}(n+1)^2M_n^2x_0\,\frac{e^{-M_nr}}{r},\qquad r>x_0 \tag{4.31}$$

with the aid of integral tables.[36] It can be noted that taking $\sigma=-1$ yields the same result. The series (4.31) converges uniformly for $r>\epsilon>0$ for any increasing mass spectrum and the dominating term is $r^{-1}x_0\times\exp(-M_0r)$; hence, we see that, for spacelike distance, the commutator of the fields falls off faster than any polynomial—the nonlocality is short range. Again the theory is causal for $x_0=0$, $\vec{x}^2<0$, $\eta=\eta'=(1,0,0,0)$.

However, it must be emphasized that, in the mass nondegenerate case, the above discussion does not exclude the possibility of having a causal theory. The reason for this is that the demand for causality falls upon not the fields themselves necessarily but certain bilinear functionals of the fields—currents. In the next section, it is shown that the fields are orthogonal with respect to a current norm, but not necessarily with respect to the $O(4,1)$ group representation space norm. Hence, if the timelike solutions to any of the wave equations are complete, it is expected that they would be complete with respect to the current norm. As a result, the commutator of currents could be proportional to $\delta^{\mathrm{hyp}}(\eta-\eta')$ and thus vanish for $x-x'$ spacelike, whereas the commutator of the fields would not vanish for spacelike $x-x'$. However, this is perhaps too much to ask for wave equations without spacelike solutions. Indeed, we should expect some sort of nonlocal behavior for observables corresponding to a composite system. Furthermore, it has been demonstrated by Fronsdal,[39] in a similar infinite-component model using the group $U(3,1)$, that a wave equation with discrete timelike solutions only is not complete.

5. CURRENTS AND VERTEX FUNCTIONS

Orthogonality and the current norm

Although it has been mentioned that the wave equation, Eq. (3.14), can be derived by a Lagrangian approach, perhaps the easiest method for constructing currents is via Takahashi's generalized Ward identity.[40] This approach is very easy to apply in our case and leads directly to a conserved current as well as orthogonality conditions on the wavefunctions.

We begin by defining the wave operator corresponding to the wave equation [Eq. (3.14)]

$$L(p)\equiv C(p^2)+p^\mu p^\nu S_{\mu\lambda}S^\lambda{}_\nu+p^\mu p^\nu\Gamma_\mu\Gamma_\nu+(\sigma^2+3\sigma+1)p^2. \tag{5.1}$$

By calculating the difference $L(p')-L(p)$, one can easily arrive at the relation

$$L(p')-L(p)=(p'-p)^\mu I_\mu(p',p) \tag{5.2}$$

with

$$I_\mu(p',p)$$
$$=(p'+p)^\nu\big(g_{\mu\nu}[C_1+C_2(p'^2+p^2)+C_3(p'^4+p^2p'^2+p^4)+\cdots$$
$$+(n-1)\text{ term}+(\sigma^2+3\sigma+1)]+\tfrac{1}{2}\{S_{\mu\lambda},S^\lambda_\nu\}+\tfrac{1}{2}\{\Gamma_\mu,\Gamma_\nu\}\big), \tag{5.3}$$

where the expansion

$$C(p^2)=\sum_{i=0}^{n}C_i(p^2)^i$$

for the polynomial function $C(p^2)$ has been used. Equation (5.2) is of the form of a generalized Ward identity previously derived by Takahashi and Kazes[40] for ordinary fields. Of course, now the operators $L(p)$ and $I_\mu(p', p)$ are differential operators acting in $\mathcal{L}^2(H^3)$. And as with the usual case, the current operator $I_\mu(p'p)$ is not unique since Eq. (5.2) is left unchanged by the gauge transformation

$$I_\mu(p', p) \to I_\mu(p', p) + \Lambda_\mu(p', p)$$

if $(p' - p)^\mu \Lambda_\mu(p', p) = 0$.

The real significance of Eq. (5.2) is that one can use it to derive current conservation and orthogonality of the wavefunctions. By sandwiching Eq. (5.2) between wavefunctions of momentum p' and p, one finds

$$(p' - p)^\mu \langle p' \; n' s' m' \,|I_\mu(p', p)|\, p\, nsm\rangle = 0 \tag{5.4}$$

which is a statement of current conservation. By putting $p_\mu = (p_n^0, \vec{p})$ and $p'_\mu = (p_{n'}^0, \vec{p})$ with $p_n^0 = (M_n^2 + \vec{p}^2)^{1/2}$ and $p_{n'}^0 = (M_{n'}^2 + \vec{p}^2)^{1/2}$, one finds the orthogonality relation

$$(p_{n'}^0 - p_n^0)\, \langle \vec{p}\; n' s' m' |\, I_0(\vec{p}, \vec{p})\, | \vec{p}\; nsm\rangle = 0,$$

which implies

$$\langle \vec{p}\;\; n' s' m' \,|I_0(\vec{p}, \vec{p})|\, \vec{p}\; nsm\rangle \propto \delta_{n' n}\, \delta_{s' s}\, \delta_{m' m}. \tag{5.5}$$

The Kronecker deltas on the quantum numbers s and m follows from the fact that the terms which are off diagonal in s and m are also off diagonal in n. This can be seen in the following calculation, which also demonstrates positive definiteness of the current. Without loss of generality we take $\vec{p}$ in the third direction and extract the boost from the expression for the current in Eq. (5.5) obtaining

$$\langle \vec{p}\; ns' m' \,|I_0(\vec{p}, \vec{p})|\, \vec{p}\; nsm\rangle$$
$$= \langle n\; s' m' \,|\exp(i\alpha N_3) I_0 \exp(-i\alpha N_3)\,|nsm\rangle$$

with α parametrized as per Eq. (3.33), and the explicit expression for I_0 is

$$I_0(p, p) = 2p_0\left(\frac{\partial C(p^2)}{\partial p^2} + (\vec{S}^2 + \Gamma^2 + 1)\right)$$
$$- q[\{\Gamma_0, \Gamma_3\} + \{N_2, S_2\} - \{N_2, S_1\}], \tag{5.6}$$

where

$$\frac{\partial C}{\partial p^2}(p^2) = [C_1 + C_2(p'^2 + p^2) + C_3(p'^4 + p^2 p'^2 + p^4) + \cdots]_{p'^2 = p^2}$$

and

$$p_\mu = (p_0, 0, 0,).$$

Using some automorphisms of the algebra,[14] we find

$$\exp(i\alpha N_3) I_0(q, q) \exp(-i\alpha N_3)$$
$$= 2p_0 \frac{\partial C}{\partial p^2}(p^2) + 2p(\vec{S}^2 + \vec{\Gamma}^2 + 1) + q[\{\Gamma_0, \Gamma_3\} + \{N_1, S_2\} - \{N_2, S_1\}]. \tag{5.7}$$

As can be seen from the action of the generators on the rest frame states,[14] the last term vanishes between states with equal n, hence,

$$\langle \vec{p}\; n' s' m' \,|I_0|\, \vec{p}\; nsm\rangle$$
$$= \delta_{nn'}\, \delta_{ss'}\, \delta_{mm'} \cdot 2p_0\left(\frac{\partial C(p^2)}{\partial p^2} - \frac{C(p^2)}{p^2}\right). \tag{5.8}$$

In order that the operator $-I_0(p, p)$ be positive definite, we must assure that

$$\frac{\partial C(p^2)}{\partial p^2} - \frac{C(p^2)}{p^2} < 0 \tag{5.9}$$

for all timelike solutions. Indeed, this is the case for the models described by Eqs. (3.24) and (3.30). Using Eq. (5.8) with the constraint, Eq. (5.9), we can define a new Hilbert space with the inner product given by

$$(\phi_1, \; -I_0\phi_2) = \int \frac{d^3p}{p_0} \int \frac{d^3\eta}{\eta_0}\, \phi_1^*(p, \eta) I_0(p, p)\, \phi_2(p, \eta). \tag{5.10}$$

It is with respect to this Hilbert space that the timelike solutions of the wave equation, Eq. (3.14) are orthogonal. Notice that the operator I_0 defining the norm given in Eq. (5.10) is proportional to p_0 when sandwiched between single particle states—the usual result for boson field theories.[31] Explicit calculation also yields the result

$$\langle -\vec{p}\; n' s' m' \,|I_0|\, \vec{p}\; nsm\rangle \propto \delta_{nn'}\, \delta_{ss'}\, \delta_{mm'}. \tag{5.11}$$

Similar type calculations can be used to demonstrate for theories containing spacelike and lightlike solutions (treated as the appropriate limit of the spacelike solutions) that such solutions are orthogonal with respect to I_0 norm and have a delta function norm of opposite sign to that of the timelike solutions. For example, transforming the current $I_0(p, p)$ of Eq. (5.6) with a spacelike parametrized boost one obtains in lieu of Eq. (5.7),

$$\exp(i\alpha N_3) I_0(p, p) \exp(-i\alpha N_3)$$
$$= 2p_0 \frac{\partial C(p^2)}{\partial p^2} + 2p_0(\Gamma_1^2 + \Gamma_2^2 + S_3^2 - N_1^2 - N_2^2 - \Gamma_0^2)$$
$$- q[\{\Gamma_0, \Gamma_3\} + \{N_1, S_2\} - \{N_2, S_1\}], \tag{5.12}$$

which by Eq. (3.18) again yields Eq. (5.8). Thus demanding Eq. (5.9) for spacelike solutions yields a norm opposite in sign to the timelike solutions. Also it follows easily from Eq. (5.2) that all spacelike, lightlike, and timelike solutions are orthogonal for fixed $\vec{p}$.

When constructing a field theory, invariant under a certain transformation group, a crucial question to ask is whether one can construct the Poincaré generators as bilinear functionals of the fields themselves. In this way one obtains currents associated with Poincaré invariance as well as the charge current associated with gauge invariance. Indeed, an important problem associated with this construction is to ascertain whether the energy corresponding to the field is positive definite. In order to write down the currents in the x representation, we make the usual identification

$$p_\mu \to i\vec{\partial}_\mu, \quad p'_\mu \to -i\overleftarrow{\partial}_\mu, \tag{5.13}$$

and the current operator given by Eq. (5.3) is written as $I_\mu(-i\overleftarrow{\partial}, \; i\vec{\partial})$. Thus one obtains the conserved current

$$J_\mu(x, \eta) = :\phi^\dagger(x, \eta) I_\mu(-i\overleftarrow{\partial}, i\vec{\partial})\, \phi(x, \eta): \tag{5.14}$$

as a bilinear functional of the fields. The conservation equation, Eq. (5.4), then becomes

$$\partial^\mu J_\mu(x, \eta) = 0, \tag{5.15}$$

yielding the conserved electric charge

$$Q=\int \frac{d^3\eta}{\eta_0}\int d^3x J_0(x,\eta)$$
$$=\int \frac{d^3\eta}{\eta_0}\int d^3x\, \phi^\dagger(x,\eta) I_0(-i\overleftrightarrow{\partial}, i\overleftrightarrow{\partial})\,\phi(x,\eta). \tag{5.16}$$

One can now make use of the orthogonality relations, (5.8) and (5.11) and upon normal ordering obtain Q in the Fock space representation as

$$:Q:=\sum_{nsm}\int_+ \frac{d^3p}{p_0}\,[a^\dagger_{nsm}(\vec{p})\,a_{nsm}(\vec{p})-b^\dagger_{nsm}(\vec{p})\,b_{nsm}(\vec{p})]. \tag{5.17}$$

One constructs the Poincaré generators from the fields in the usual way yielding

$$P_\mu=i\int\frac{d^3\eta}{\eta_0}\int d^3x\,\phi^\dagger(x,\eta)\overleftrightarrow{\partial}_\mu I_0(-i\overleftrightarrow{\partial}, i\overleftrightarrow{\partial})\,\phi(x,\eta)$$

$$M_{\mu\nu}=i\int\frac{d^3\eta}{\eta_0}\int d^3x\Big[\phi^\dagger(x,\eta)I_0(-i\overleftrightarrow{\partial}, i\overleftrightarrow{\partial})\Big(x_\mu\partial_\nu-x_\nu\partial_\mu+\eta_\mu\frac{\partial}{\partial\eta^\nu}-\eta_\nu\frac{\partial}{\partial\eta^\mu}\Big)\,\phi(x,\eta)\;-\;\phi^\dagger(x,\eta)\times\Big(x_\mu\overleftarrow{\partial}_\nu-x_\nu\overleftarrow{\partial}_\mu+\eta_\mu\frac{\overleftarrow{\partial}}{\partial\eta^\nu}-\eta_\nu\frac{\overleftarrow{\partial}}{\partial\eta^\mu}\Big)I_0(-i\overleftrightarrow{\partial}, i\overleftrightarrow{\partial})\,\phi(x,\eta)\Big]. \tag{5.18}$$

Again using the orthogonality conditions, one obtains these generators in the Fock space representation. Exhibiting this explicitly for P_μ yields the expression

$$:P_\mu:=\sum_{nsm}\int_+\frac{d^3p}{p_0}\,p_\mu\,[a^\dagger_{nsm}(\vec{p})\,a_{nsm}(\vec{p})+b^\dagger_{nsm}(\vec{p})\,b_{nsm}(\vec{p})], \tag{5.19}$$

whose zeroth component is positive definite only upon quantization with boson commutation rules.

It should be mentioned that using the above machinery all the standard derivations go through, that is, $M_{\mu\nu}$ and P_μ obey the Lie algebra of the Poincaré group and express the invariance of the theory under arbitrary Poincaré transformations. Moreover, in the exact symmetry mass degenerate case, one could exhibit all the generators of $\mathcal{P}_0\otimes O(3,1)$ as conserved quantities, thus obtaining the maximal number of metric automorphisms of each of the underlying spaces, $\mathcal{M}$ and H^3.

Vertex functions

In the previous section, it was mentioned that the integrated charge Q arriving from the conserved current J_μ was to be associated with the electric charge; hence, J_μ is the electromagnetic current. Indeed, this is the usual interpretation in the analogous infinite component field theories.[30,41] However, in our case we make the further interpretation that any hadronic interaction necessarily couples the η dependence whereas the electromagnetic interaction is independent of the internal η space. This reflects the fact that the electromagnetic current, Eq. (5.14), is obtained from a variational principle in which only the Minkowski endpoints are allowed to vary.[14] Consequently, we obtain the invariant nonminimal electromagnetic coupling

$$J_\mu(x,\eta)A^\mu(x). \tag{5.20}$$

This coupling provides us with the momentum space form factors or electromagnetic vertex functions given by

$$\langle p'\; n's'm'\,|\,-I_0(p',p)\,|\,p\;nsm\rangle.$$

To calculate these functions, we notice that without loss of generality we can take $p'_\mu=(M',0,0,0)$ and $p_\mu=(p_0,0,0,q)$ and normalize the single-particle states according to the Poincaré invariant norm, Eq. (5.9). Then, upon extracting the boost, one finds

$$\langle n's'm'\,|I_0(p',p)|\,p\,nsm\rangle=\sum_{n''s''}\langle n's'm'\,|I_0|\,n''s''m\rangle\times\langle n''s''m\,|\exp(-i\alpha N_3)\,|nsm\rangle. \tag{5.21}$$

The first term in the sum can be evaluated using Eq. (5.3) and the action of the $O(4,1)$ generators on the basis states,[14] yielding

$$\langle n's'm\,|-I_0(p',p')\,|n''s''m\rangle\equiv I_{n''s''}(\alpha)=\delta_{n'n''}\,\delta_{s's''}(M_{n'}+P^0_n)\times\frac{[C_1+C_2(M^2_{n'}+M^2_n)+C_3(M^4_{n'}+M^2_{n'}M^2_n+M^4_n)+\cdots+(n'+1)^2]}{N(n',n)}$$
$$-\frac{q}{2}\Big((\sigma-n'')\delta_{n',n''+1}\frac{(\delta_{s',s''+1}A_{n'',s''}-\delta_{s's''-1}A_{n'',-s''-1})}{N(n',n)}-(\sigma+n''+2)\,\delta_{n',n''-1}\frac{(\delta_{s',s''+1}A_{n-1,-s''-2}-\delta_{s's''-1}A_{n''-1,s''-1})}{N(n',n)}\Big), \tag{5.22}$$

where

$$A_{n,s}=(n+\tfrac{3}{2})\Big(\frac{(s+m+1)(s-m+1)(n+s+3)(n+s+2)}{(2s+3)(2s+1)(n+2)(n+1)}\Big)^{1/2}$$

and

$$N(n',n)=\Big[4p^0_nM_{n'}\Big(\frac{\partial C}{\partial p^2_n}+(n+1)\Big)\Big(\frac{\partial C}{\partial p^2_{n'}}+(n'+1)^2\Big)\Big]^{1/2}.$$

The matrix elements

$$D^\sigma_{n's'm,nsm}(\alpha)\equiv\langle n's'm\,|\exp(-i\alpha N_3)\,|nsm\rangle$$

in Eq. (5.21) can be obtained by group theoretical techniques. The above matrix elements can be related to the double coset matrix elements,

$$T^\sigma_{n'ns}(\alpha)\equiv\langle n'sm\,|\exp(-i\alpha\Gamma_0)\,|nsm\rangle, \tag{5.23}$$

obtained previously as a finite series of hypergeometric functions[42] and exhibited explicitly in the Appendix via the identity

$$\exp(i\pi\Gamma_3/2)\exp(-i\alpha\Gamma_0)\exp(-i\pi\Gamma_3/2)=\exp(-i\alpha N_3), \tag{5.24}$$

yielding

$$D^\sigma_{n's'm;nsm}(\alpha)=\sum_{s''=0}^{\min(n'n)}d^{n'}_{s's''m}(\theta=\tfrac{1}{2}\pi)\,T^\sigma_{n'ns''}(\alpha)\,d^n_{s''sm}(\theta=-\tfrac{1}{2}\pi), \tag{5.25}$$

where the d functions are the $O(4)$ representation functions[43] and can be related to the Clebsh–Gordon coefficients[43] for $SU(2)$ by using the local isomorphism between $O(4)$ and $SU(2)\otimes SU(2)$, yielding finally

$$D^{\sigma}_{n's'm;nsm}(\alpha)=\sum_{s''=0}^{\min(n',n)}\sum_{\mu}\sum_{\mu'}C(\tfrac{1}{2}n',\tfrac{1}{2}n',\ s';\mu,m-\mu)$$
$$\times C(\tfrac{1}{2}n',\tfrac{1}{2}n',s'';\mu,m-\mu)$$
$$\times C(\tfrac{1}{2}n,\tfrac{1}{2}n,s'';\mu',m-\mu')$$
$$\times C(\tfrac{1}{2}n,\tfrac{1}{2}n,s;\mu',m-\mu')\exp[i(\mu-\mu')\pi\, T^{\sigma}_{n'ns''}(\alpha). \tag{5.26}$$

Combining this result with Eqs. (5.21) and (5.22) yields the complete expression for the electromagnetic vertex functions,

$$\langle n's'm\,|I_0(p',p)|\vec{p}\ nsm\rangle$$
$$=\sum_{\substack{n''=n'\pm1\\ s''=s'\pm1}} I_{n''s''}(\alpha)D_{n''s''m;nsm}(\alpha). \tag{5.27}$$

It is easy to relate the boost parameter α to the invariant momentum transfer $t=(p'-p)^2$, obtaining

$$\cosh\alpha=(M^2_{n'}+M^2_n-t)/2M_nM_{n'}, \tag{5.28}$$

where the physical region is $-\infty<t\leq(M'-M)^2$. Equation (5.27) thus provides all of the electromagnetic form factors as well as transition amplitudes for the particles of our theory. We find, using Eq. (A1) of the Appendix and applying one of Kummer's identities[22] for the hypergeometric functions, the asymptotic behavior

$$T_{n'ns}(\alpha)\ \underset{\alpha\to\infty}{\simeq}\ \lambda_1\exp[-\alpha(-\sigma+s)]+\lambda_2\exp[-\alpha(\sigma+3+s)] \tag{5.29}$$

with the exception of $\sigma=-\frac{3}{2},\ -\frac{1}{2}$ where an extra multiplicative factor of α occurs in one of the terms arising from the degeneracy in the hypergeometric functions. We thus obtain the asymptotic behavior of the vertex functions, e.g., for spin-0 particles

$$\langle n'00\,|I_0(p'p)|\vec{p}\ n00\rangle$$
$$\underset{t\to\infty}{\simeq}(-t)^{\sigma+1/2}\begin{cases}\ln(-t), & \sigma=-\frac{3}{2},\ -\frac{1}{2},\\ 1, & \text{otherwise.}\end{cases} \tag{5.30}$$

For the $\sigma=-1$ representation of the supplementary series, the asymptotic behavior is $(-t)^{-1/2}$ and, for the principal series with $\rho\neq0$, the asymptotic behavior is $(-t)^{-1}$ multiplying an oscillating term. While experimental data[44] for boson form factors is scarce, there are indications that the pion form factor, for example, falls off faster than the simple pole behavior above, indeed even faster than the proton form factors. Thus this model should be modified to describe the pion. In this respect the $O(4,2)$ models[41] are somewhat better. Also notice from Eq. (5.30) that, for the supplementary series in the range $-\frac{1}{2}\leq\sigma<0$, the asymptotic behavior of the vertex functions is totally unacceptable.

One common feature which seems to be shared by all infinite multiplet theories[9,41] is that the analytic continuation of the vertex functions into the region $(M'-M)^2<t<\infty$ has nothing at all to do with the pair annihilation process. This is indeed the case in our theory also, as can be seen by noticing that again this simply involves the overlap

$$\int\frac{d^3\eta}{\eta_0}\ \phi^*(p'\eta)I_0(p',p)\phi(p,\eta),$$

where now $s=(p'+p)^2$ is related to the boost parameter α by

$$\cosh\alpha=(M^2_{n'}+M^2_n-s)/2M_nM_{n'},$$

and we obtain the same function in this channel exhibiting no singularities in the physical region for the annihilation process. However, the vertex functions have a cut in the complex t plane from $(M'-M)^2$ to ∞; hence, the form factors and annihilation process cannot be the analytic continuation of one another. This seems to be indicative of a composite structure where the branch cut in the vertex functions refers to an entirely different process in terms of the constituents, perhaps; the branch point at $(M'-M)^2$ could indicate an anomalous threshold.[9,41]

Another common feature of infinite multiplet theories is the inapplicability of the TCP theorem.[35] The proof of this theorem[32] requires the analytic continuation of the boosts to the point $\alpha=i\pi$. However, it is not too difficult to see that the vertex functions have a pole at this point. In the Appendix the vertex functions are written as finite series of hypergeometric functions, $F(a,b,c;z)$ of the argument $z=1-\exp(-2\alpha)$, with a cut from $z=1$ to $z=+\infty$. Now to implement the above-mentioned continuation, let $\alpha=i\delta/2$. Then $z=1-\cos\delta+i\sin\delta$ which describes a circle of unit radius with center $z=1$. Hence, as δ varies from 0 to 2π, z moves on this circle starting and ending at $z=0$; however, it must pass through the branch cut of the hypergeometric function at the point $z=2$, and the second sheet function exhibits a pole at $z=0$. (See the Appendix for details.)

6. CONCLUSION AND DISCUSSION

In the preceding pages a relativistic quantum field theory mostly at the free field level was formulated over the manifold $\mathcal{M}\otimes H^3$ in an attempt to describe hadrons as a relativistic composite structure in a field theoretic context. This formulation at the outset is quite general enabling one to consider both Poincaré irreducible fields as well as the highly reducible infinite multiplet theories. The latter type theories allow one to make full use of the available function space and by allowing both sheets of the hyperboloid one is led quite naturally to the inclusion of an internal $O(4)$ symmetry via the noncompact groups $O(4,1)$ and $SL(4,R)$. Realistic mass splittings are attained by postulating a wave equation which links in a very complicated way the $\mathcal{M}$ and H^3 structure of the fields.

A viable criterion for selecting between the infinite multiplet and Poincaré irreducible theories was found in the locality properties. While the Poincaré irreducible theories that are nonequivalent to the usual spinor theories displayed an irritating long range nonlocality, the $O(4,1)$ theories with mass splittings enjoyed a weak nonlocality which become local when the masses became degenerate. In addition the theories equivalent to the usual spinor theories were shown to be local as indeed they must; furthermore, these theories afford the possibility of providing various nonstandard couplings when interactions are introduced.

The infinite multiplet theories here are very closely related to infinite component theories; however, the fundamental objects, fields, are quite different. We consider this a distinct advantage, for we deal exclusively in the canonical basis where particle properties are manifest whereas the infinite component theories employ the spinor basis. However, the same general features found in infinite component theories are present. Indeed, the vertex functions describing the interaction of the multiplets with the electromagnetic field were calculated and shown to exhibit the characteristic noncrossing symmetric behavior, an apparent indication of a composite structure.[9] In this context one of our models has many desirable features: It has discrete timelike states only, a positive-definite metric, a positive-definite energy, obeys a spin statistics theorem in the mass degenerate limit, CPT invariance, an almost linearly rising mass spectrum, decreasing electromagnetic form factors though probably not fast enough, an internal space which can be interpreted as the underlying manifold for the hadronic interactions. The price we must pay—weak nonlocality at the field level and probably at the current level.

It is at this stage that a crucial period in the development of our theory is reached. The central question here is the question of the completeness of the physical states in the physical I_0 norm. If there exists a complete set of states to a given wave equation, then the theory will be formally local at the current level and a theory for interactions could be developed in the usual covariant manner. For example, then the time-ordered products of fields at different points should be covariant aside from possible contact terms. However, if the wave equation does not provide a complete set of states, the current commutator will not vanish for all spacelike separated points and the theory will be nonlocal as at the field level. It is probably not too hazardous to say that such a theory would be weakly nonlocal at the current level as well as the field level, thus enabling one perhaps to develop a reasonable physical theory. However, the problem with covariance is manifest; the

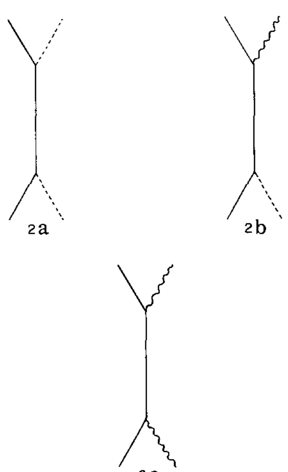


FIG. 2. Compton scattering in the s channel. Solid lines indicate members of an infinite multiplet; dotted lines indicate a simple scalar field; wavy lines indicate the electromagnetic field.

time-ordered product of currents will be noncovariant. The problem is somewhat more confused when spacelike solutions to a wave equation are allowed to enter. If both timelike and spacelike (possibly lightlike, too) form a complete set, it is still not clear that one can develop a causal theory even if the current commutator vanishes for spacelike separation and the time-ordered product is fully covariant.

This problem of completeness has been formulated in two related works. Mukunda[39] has shown that the solutions to the Majorana equation form a complete set in the current norm. Some of these solutions, however, are spacelike and lightlike. This approach is not very well suited to our wave equation due to the fact that our currents contain bilinear products of the $O(4,1)$ generators. A more natural approach for us is that of Fronsdal.[39] His technique differs from that of Mukunda's only in the evaluation of the resolvent Green's function operator. Moreover, his method can handle more complicated currents.

To begin with, consider the simplest nontrivial (infinite multiplet propagator) scattering amplitude, that of s-channel Compton scattering of an ordinary scalar field from the ground state of the multiplet as shown in Fig. 2a. This can be written down using the resolvent operator $L^{-1}(p)$ as the propagator as suggested by the generalized Ward identity, Eq. (5.2),

$$T(p',p;q)=\langle \vec{p}'\ 000|[1/L(q^2)]|\vec{p}\ 000\rangle. \tag{6.1}$$

Transforming $L(q^2)$ to the rest frame by

$$C(p^2)+p^2(\vec{S}^2+\vec{\Gamma}^2+1)=\exp(-i\vec{\alpha}_q\cdot\vec{N})L(q^2)\exp(i\vec{\alpha}_q\cdot\vec{N}) \tag{6.2}$$

and inserting a complete set of rest frame states one finds

$$\begin{aligned}T(p',p;q)&\\ =\sum_{nsm}\{&\langle 000|\exp(i\vec{\alpha}^0_{p'}\cdot\vec{N})\exp(-i\vec{\alpha}_q\cdot\vec{N})|nsm\rangle\\ &\times\langle nsm|\exp(i\vec{\alpha}_q\cdot\vec{N})\exp(-i\vec{\alpha}^0_{p'}\cdot\vec{N})|000\rangle\\ &\times[C(q^2)+q^2(n+1)^2]^{-1}\}.\end{aligned} \tag{6.3}$$

The superscript 0 on α indicates the state $n=0$. But the bra–ket functions are just the scalar vertex functions for the transition of a state nsm to the ground state; we indicate these functions as $V^{(0,n)}(p\cdot q,q^2)$. Therefore,

$$T(p',p;q)=\sum_{nsm}\frac{V^{(0,n)}(p'\cdot q,q^2)\,V^{(n,0)}(p\cdot q,q^2)}{C(q^2)+q^2(n+1)^2}. \tag{6.4}$$

The relevant amplitude for completeness in the I_0 norm is

$$T_0(p',p;q)=\sum_{nsm}\frac{V^{(0,n)}(p'\cdot q,q^2)\,V_0^{(n,0)}(p,q)}{C(q^2)+q^2(n+1)^2}, \tag{6.5}$$

where

$$V_\mu^{(n,0)}(p,q)=\langle nsm|\exp(i\vec{\alpha}_q\cdot\vec{N})I_\mu(q,p)\exp(-i\vec{\alpha}^0_p\cdot\vec{N}|000\rangle.$$

Similarly for Compton scattering from the electromagnetic field, the full amplitude is

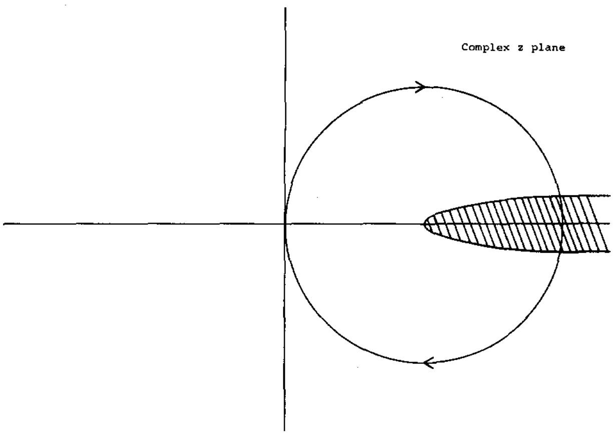


FIG. 3. The complex z plane of the vertex functions. The cut runs from $z=1$ to ∞. The circle indicates the path of analytic continuation to the point $\alpha=i\pi$ as described in the Appendix.

$$T_{\mu\nu}(p',p;)=\sum_{nsm}\frac{V_{\mu}^{(0,n)}(p',q)\,V_{\nu}^{(n,0)}(p,q)}{C(q^2)+q^2(n+1)^2}. \tag{6.6}$$

Notice that if $C(q^2)$ is of the form (3.23), i.e., the CGL wave equation, all of the above amplitudes exhibit a pole in the complex q_0 plane at $q_0=0$. We do not try to evaluate these amplitudes or solve the completeness problem here, but merely indicate what has to be done. The vertex functions $V^{(0,n)}$ in Eq. (6.4) are essentially Legendre functions and should be manageable. The crucial points in any further investigation are (i) to decide whether Fronsdal's conclusion[39] for $U(3,1)$ that a discrete set of timelike solutions is not complete is applicable in our case or not, (ii) to understand the nature of the relationship between completeness and spacelike solutions.

Before closing a few words are in order concerning the relationship between our work and that of CGL.[8] Their approach is toward the saturation of current algebra. Indeed they use the canonical quantization following from the Schwinger variational principal to obtain a current algebra with Schwinger terms, but unless a completeness relation holds, the Fock space decomposition will be invalid. On the other hand we have a Fock space decomposition, but unless a completeness relation holds, we will never get canonical commutation relations or current algebra: Fock space quantization and canonical quantization are equivalent only when a completeness relation is valid.

Note added in proof: Recently the question of completeness of the states in the current norm has been resolved for an $O(3,1)$ version of the CGL infinite component wave equation by R. Casalbuoni and G. Longhi, Nuovo Cimento A 15, 591 (1973). They have shown that the timelike states are *not* complete, and one has contributions from states with complex momentum. This result is intimately related to the existence of the pole in the vertex functions at the point $\alpha=i\pi$ as discussed at the end of Sec. 5. Another recent work related to ours is R. Y. Cusson and L. P. Staunton, Nuovo Cimento A 17, 303 (1973).

ACKNOWLEDGMENTS

The authors would like to thank Dr. F. Ardalan, Dr. H. Grotch, and Dr. K. H. Wang and Professor E. Kazes for many valuable discussions.

APPENDIX: VERTEX FUNCTIONS

In Sec. 5 we discussed the vertex functions of our theory. They were expressed in terms of the $SO_0(4,1)$ matrix elements in the double coset $SO(4)\ SO_0(4,1)/SO(4)$. Such matrix elements were calculated for the groups $SO_0(p,1)$ in Ref. 42, and so we merely restate the results for $p=4$ here:

$$\begin{aligned}T^{\sigma}_{nn's}(\alpha) &= N\sum_{k=0}^{n-s}\sum_{k'=0}^{n'-s}\frac{(-1)^{k+k'}\exp[\alpha(\sigma-n-2k')]\,\Gamma(s+\frac{3}{2}+k+k')}{k'!\,k!\,(n+\frac{3}{2}-k)}\\ &\times\frac{\Gamma(n'+n-s+\frac{3}{2}-k-k')}{\Gamma(n'+\frac{3}{2}-k)\,\Gamma(s+\frac{3}{2}+k)}\\ &\times\frac{{}_2F_1(n'-\sigma;s+\frac{3}{2}+k+k';n'+n+3;1-\exp(-2\alpha))}{\Gamma(s+\frac{3}{2}+k')\,\Gamma(n-s+1-k)\,\Gamma(n-s+1-k')}\end{aligned} \tag{A1}$$

with

$$N=(-1)^{n+s}\frac{[(n-s)!\,(n'-s)!\,(n+s+1)!\,(n'+s+1)!\,(n+1)}{(n+n'+2)!}$$
$$\times(n'+1)]^{1/2}\,\Gamma(n+\tfrac{3}{2})\,\Gamma(n'+\tfrac{3}{2})$$

and, when $n=s=0$, they reduce to Legendre functions,

$$T^{\sigma}_{0n}(\alpha)=2\,\frac{(n+\sigma+3)}{\Gamma(\sigma+3)}\,(n+1)\,\frac{P^{-1-n}_{\sigma+1}(\cosh\alpha)}{\sinh\alpha}. \tag{A2}$$

From Eq. (A1) it is seen that the functions $T^{\sigma}_{nn's}(\alpha)$ have a branch cut taken along the real axis of the variable $z=1-\exp(-2\alpha)$ from 1 to ∞. In terms of this variable, the momentum transfer t is

$$t=M'^2+M^2-M'M[(1-z)^{1/2}+(1-z)^{-1/2}]. \tag{A3}$$

Choosing the cut for the square root function again from $z=1$ to ∞, we see the physical region

$$-\infty<t\leqslant(M'-M)^2$$

corresponds to the range

$$1>z\geqslant 0$$

and the cut in the t plane runs from $(M'-M)^2$ to infinity. The cut of the hypergeometric function corresponds to imaginary values of t.

Now consider the analytic continuation necessary for the CPT theorem as discussed in the text. We take $\alpha=i\delta/2$; then $z=1-\cos\delta+i\sin\delta$. As we continue α from o to $i\pi$, z describes the circle shown in Fig. 3. The hypergeometric function in Eq. (A1) can be continued[22] into the region about the singular point $z=1$. If $\sigma\neq$ integer, half-interger, this is straightforward. When σ = integer or half-interger, the hypergeometric functions are degenerate and care must be taken, but the results will be the same. We then continue these previously continued functions back to the region about $z=0$. Upon doing this our original function

$$U_{1(0)} = F(n' - \sigma;\ s + \tfrac{3}{2} + k + k';\ n' + n + 3;\ z)$$

becomes[22]

$$U_{1(0)} \to A\,U_{1(0)} + B\,U_{2(0)},$$

where

$$\begin{aligned} U_{2(0)} &= \ln z\, U_{1(0)} \\ &\quad + \sum_{i=1}^{\infty} \frac{(n' - \sigma)_i (s + \frac{3}{2} + k + k')_i}{(n' + n + 3)_i\, i\,!}\, z^i \,[\psi(n' - \sigma + i) \\ &\quad - \psi(n' - \sigma) + \psi(s + \tfrac{3}{2} + k + k' + i) - \psi(s + \tfrac{3}{2} + k + k') \\ &\quad - \psi(n' + n + 3 + 2) + \psi(n' + n + 3) - \psi(i + 1) - \psi(1)] \\ &\quad - \sum_{i=1}^{n'+n+2} \frac{(i-1)!\,(-n' - n - 3)_i\, z^{-i}}{(1 + \sigma - n')_i (-s - k - k' - \frac{1}{2})_i}, \end{aligned} \qquad \text{(A4)}$$

where

$$(a)_i = \Gamma(a + i)/\Gamma(a)$$

is Pockhammer's symbol and

$$\psi(x) = \frac{d \ln \Gamma(x)}{dx}.$$

Consequently, $T^{\sigma}_{nn's}$ has a pole of order $z^{-(n'+n+2)}$ when $\alpha = i\pi$.

*Submitted in part for Ph.D. dissertation, The Pennsylvania State University, March 1972, by C.P. Boyer.

†Supported in part by an NDEA Fellowship. Present address: C.I.M.A.S., Universidad Nacional Autónoma de México, México 20, D.F.

[1]G. Chew, *The Analytic S-Matrix* (Benjamin, New York, 1966).

[2]H. Yukawa, Phys. Rev. **77**, 219 (1950); **91**, 415 (1953).

[3]A. Kihlberg, Ann. Inst. Henri Poincaré 13, 57 (1970); Goteborg preprints.

[4]H. Bacry and A. Kihlberg, J. Math. Phys. **10**, 2132 (1969); H. Bacry and J. Nuyts, Phys. Rev. **157**, 1471 (1967).

[5]G. Fuchs, Ecole Polytechnique preprints.

[6]D. Finkelstein, Phys. Rev. **100**, 924 (1955).

[7]F. Lurçat, Physics 7, 95 (1964).

[8]R. Casalbuoni, R. Gatto, and G. Longhi, Phys. Rev. D 3, 1499 (1971); Nuovo Cimento Lett. **2**, 159, 166 (1969). Referred to hereafter as CGL.

[9]C. Fronsdal and R. White, Phys. Rev. **163**, 1835 (1967); C. Fronsdal, Phys. Rev. **168**, 1845 (1969); **171**, 1811 (1968).

[10]L.S. Pontryagin, *Topological Groups* (Gordon and Breach, New York (1966).

[11]E. Wigner, Ann. Math. **40**, 149 (1939). The specification of this freedom is tantamount in the Wigner theory to the choice of tensors used. It is also noted that our carrier space does not carry half-integer spin short of the use spinor valued wavefunctions.

[12]H. Joos, Fortschr. Phys. **10**, 65 (1962).

[13]J. Nilsson and Beskow, Ark. Fys. **34**, 307 (1967).

[14]C.P. Boyer, Ph.D. Thesis, The Pennsylvania State University (1972).

[15]A. Sankaranarayanan, Nuovo Cimento 38, 1441 (1965); J. Math. Phys. **9**, 611 (1968); A. Böhm, Phys. Rev. **145**, 1212 (1966); J. Rosen and P. Roman, J. Math. Phys. **7**, 2072 (1966); A. Chakrabarti, J. Math. Phys. **9**, 2087 (1968).

[16]M. Bander and C. Itzykson, Rev. Mod. Phys. **38**, 330, 346 (1966).

[17]C.P. Boyer and F. Ardalan, J. Math. Phys. **12**, 2070 (1971); C.P. Boyer, J. Math. Phys. **14**, 609 (1973).

[18]A. Kihlberg, Nuovo Cimento **53**, 592 (1968).

[19]I.M. Gel'fand, R.A. Minlos, and Z. Ja. Shapiro, *Representations of the Rotation and Lorentz Group* (Pergamon, New York, 1963).

[20]I.M. Gel'fand, M.I. Graev, and N. Ja. Vilenkin, *Generalized Functions* (Academic, New York, 1966), Vol. 5; N. Ja. Vilenkin and Ja. A. Smorodinski, Zh. Eksp. Teor. Fiz. **46**, 1793 (1963) [Sov. Phys. JETP 19, 1209 (1964)].

[21]N. Ja Vilenkin, *Special Functions and the Theory of Group Representations* (Am. Math. Soc., Providence, R.I., 1968).

[22]A. Erdélyi, *et al.*, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953).

[23]A. Aurilia and H. Umezawa, Nuovo Cimento A 51, 14 (1967); Phys. Rev. **182**, 1682 (1969); C. Fronsdal, Nuovo Cimento Suppl. 9, 416 (1958).

[24]A. Böhm, *Lectures in Theoretical Physics*, edited by A.O. Barut and W. Brittin (Gordon and Breach, New York, 1968), Vol. 10B.

[25]L. O'Raifeartaigh, Phys. Rev. Lett. **14**, 575 (1965); Phys. Rev. **161**, 1571 (1967); **164**, 2000 (1967); R. Jost, Helv. Phys. Acta. **39**, 369 (1966); I. Segal, J. Funct. Anal. **7**, 1 (1967).

[26]E.L. Ince, *Ordinary Differential Equations* (Dover, New York, 1956).

[27]A.D. Donkov, V.G. Kadyshevsky, M.D. Mateev, and R.M. Mir-Kasimov, Dubna Preprint E2-5339 (1970).

[28]G.N. Fleming, Phys. Rev. D 1, 542 (1970); G.N. Fleming and F. Ardalan, The Pennsylvania State University preprint; F. Ardalan, Ph.D. Thesis, The Pennsylvania State University (1970). There is also a continuum of spacelike solutions as well as vacuumlike solutions for Eq. (3.12).

[29]The case $C(p^2) = C_0 + C_1 p^2$ must be handled separately. In order to have a meaningful timelike spectrum, $C_0 < 0$ with C_1 arbitrary.

[30]If, for example, we took $\beta_1 < \beta < 2\beta_1$, we would have another single timelike state which one could perhaps fit nicely into a scheme describing meson masses; however, it can be seen from the analysis in Sec. 5 that such a state will have negative norm.

[31]C. Itzykson, V.G. Kadyshevsky, and I.T. Todorov, Phys. Rev. D 1, 2823 (1970); I.T. Todorov, *Group Representations in Mathematics and Physics*, edited by V. Bargmann (Springer-Verlag, Berlin, 1970); C. Itzykson and I.T. Todorov, *Coral Gables Conference on Fundamental Interactions at High Energy*, edited by T. Gudehus, *et al.* (Gordon and Breach, New York, 1969); C. Fronsdal and R.W. Huff, Phys. Rev. D 3, 933 (1970); C. Fronsdal and L.E. Lundberg, Phys. Rev. D 1, 3247 (1970); Y. Nambu, Phys. Rev. **160**, 1171 (1967); Prog. Theor. Phys. Suppls. **37** and **38**, 268 (1966); A.O. Barut, D. Corrigan, and H. Kleinert, Phys. Rev. **167**, 1527 (1968); for a review and complete set of references, see A.O. Barut, XV-th International Conference on High Energy Physics, Kiev, 1970.

[32]S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Harper, New York, 1961); R.F. Streater, A.S. Wightman, *PCT, Spin and Statistics and All That* (Benjamin, New York, 1964); R. Jost, *Theory of Quantized Fields* (Amer. Math. Soc., Providence, R.I., 1965).

[33]S. Weinberg, Phys. Rev. **133**, B1318 (1964).

[34]Y.M.P. Lam, Ann. Phys. **50**, 323 (1968).

[35]D. Tz. Stoyanov and I.T. Todorov, J. Math. Phys. **9**, 2146 (1968); A.I. Oksak and I.T. Todorov, Commun. Math. Phys. **11**, 125 (1968).

[36]I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals Series and Products* (Academic, New York, 1965).

[37]M. E. Arons and E. C. G. Sudarshan, Phys. Rev. **173**, 1662 (1968); J. Dhar and E. C. G. Sudarshan, Phys. Rev. **174**, 1808 (1968); E. C. G. Sudarshan and N. Mukunda, Phys. Rev. D **1**, 571 (1970).
[38]C. Chi, D. vong Duc, and N. van Hieu, Ann. Phys. **49**, 173 (1968); C. Fronsdal, Phys. Rev. **156**, 1653 (1967); G. Feldman and P. T. Matthews, Phys. Rev. **151**, 1176 (1968); **154**, 1241 (1967).
[39]I. T. Todorov and R. P. Zaikov, J. Math. Phys. **10**, 2014 (1969); A. I. Oksak and I. T. Todorov, Commun. Math. Phys. **14**, 271 (1969); H. D. I. Abarbanel and Y. Frishman, Phys. Rev. **171**, 1442 (1968); I. T. Grodsky and R. F. Streator, Phys. Rev. Lett. **20**, 695 (1968); A. I. Oksak and I. T. Todorov, Phys. Rev. D **1**, 3511 (1970); C. Fronsdal, Phys. Rev. **182**, 1564 (1969); **185**, 1768 (1969); N. Mukunda, Phys. Rev. **183**, 1486 (1969).
[40]Y. Takahashi, *An Introduction to Field Quantization* (Pergamon, New York, 1969); Nuovo Cimento **6**, 370 (1957); E. Kazes, Nuovo Cimento **13**, 1226 (1959).
[41]A. O. Barut and H. Kleinert, Phys. Rev. **161**, 1464 (1967); A. O. Barut, D. Corrigan, and H. Kleinert, Phys. Rev. **167**, 1527 (1968); Phys. Rev. Lett. **20**, 167 (1968).
[42]C. P. Boyer, J. Math. Phys. **12** 1599 (1971).
[43]D. Z. Freedman and J. M. Wang, Phys. Rev. **160**, 1560 (1967); M. E. Rose, *Elementary Theory of Angular Momentum* (Wiley, New York, 1967).
[44]C. W. Akerlof, *et al.*, Phys. Rev. **163**, 1482 (1967).

Lie theory and separation of variables. 3. The equation $f_{tt}-f_{ss} = \gamma^2 f$

E. G. Kalnins and W. Miller Jr.*

Centre de Recherches Mathématiques, Université de Montréal, Montréal 101, P.Q., Canada
(Received 6 November 1973)

Kalnins has related the 11 coordinate systems in which variables separate in the equation $f_{tt}-f_{ss}=\gamma^2 f$ to 11 symmetric quadratic operators L in the enveloping algebra of the Lie algebra of the pseudo-Euclidean group in the plane $E(1,1)$. There are, up to equivalence, only 12 such operators and one of them, L_E, is not associated with a separation of variables. Corresponding to each faithful unitary irreducible representation of $E(1,1)$ we compute the spectral resolution and matrix elements in an L basis for seven cases of interest and also give overlap functions between different bases: Of the remaining five operators three are related to Mathieu functions and two are related to exponential solutions corresponding to Cartesian type coordinates. We then use these results to derive addition and expansion theorems for special solutions of $f_{tt}-f_{ss}=\gamma^2 f$ obtained via separation of variables, e.g., products of Bessel, Macdonald and Bessel, Airy and parabolic cylinder functions. The exceptional operator L_E is also treated in detail.

INTRODUCTION

In Refs. 1 and 2, Winternitz and coworkers introduced a group theoretical method for the description of separation of variables in the principal partial differential equations of mathematical physics. We apply their idea in this paper to study several coordinate systems in which separation of variables is possible in the equation

$$(*)\quad (\partial_s^2-\partial_t^2)f(s,t)=-\gamma^2 f(s,t),\quad \gamma>0.$$

The symmetry group of $(*)$ is $E(1,1)$ the pseudo-Euclidean group in the plane. Its Lie algebra $e(1,1)$ is three-dimensional with basis P_1, P_2, M and commutation relations

$$[M,P_1]=P_2,\ [M,P_2]=P_1,\ [P_1,P_2]=0.$$

A two-variable model of $e(1,1)$ is

$$(**)\quad P_1=\partial_s,\quad P_2=\partial_t,\quad M=-s\partial_t-t\partial_s$$

in which case $(*)$ becomes

$$(P_1^2-P_2^2)f=-\gamma^2 f.$$

According to the prescription in Refs. 1 and 2 one should characterize solutions f of $(*)$ by requiring in addition that f is an eigenfunction of an operator L, $Lf=\lambda f$, where L belongs to the factor space $T=S/S\cap C$. Here, C is the center of the universal enveloping algebra U of $e(1,1)$ and S is the space of all symmetric second order elements in U. In our case, $S\cap C=\{\alpha(P_1^2-P_2^2)\}$, α any constant. $E(1,1)$ acts on T via the adjoint representation and we do not distinguish between operators L on the same orbit.

From the examples presented in Refs. 1 and 2 one might expect that each system of equations

$$(P_1^2-P_2^2)f=-\gamma^2 f,\quad Lf=\lambda f$$

where P_1, P_2, M are given by $(**)$, is related to a coordinate system in which $(*)$ separates, that all separable coordinate systems can be so obtained, and that there is a one-to-one relationship between orbits and separable coordinate systems. However, in Ref. 3 Kalnins has shown that this is not quite true. In fact, there are 12 orbits and 11 coordinate systems in which $(*)$ separates. One orbit (with representative L_E in this paper) does not correspond to a separable coordinate system. Of the separable coordinate systems two, Cartesian and spherical polar, have well-known group theoretical interpretations (Ref. 4, Chap. V), three lead to various types of Mathieu equations, and two correspond to other Cartesian-type coordinates. The remaining four systems are related to parabolic cylinder, Bessel, Macdonald, and Airy functions, respectively, and correspond to operators L_D, L_B, L_K, L_A in T.

A study of parabolic coordinates with respect to the spectral resolution of L_D was carried out in Ref. 5. Here we undertake an analogous study of L_B, L_K, L_A and L_E. In Secs. 1 and 2 we compute the spectral resolutions of the self-adjoint operators L_G, $G=B,K,A,E$, corresponding to each of the irreducible faithful unitary representations of $E(1,1)$. In particular, we compute the matrix elements of the unitary group representation operators in an L_G-basis and we calculate the overlap functions relating two different bases.

In Sec. 3 we show how to construct models of the irreducible representations of $E(1,1)$ in which the Lie algebra operators take the form $(**)$ and the Hilbert space vectors f satisfy $(*)$. These models allow us to apply the results of Sec. 1 to obtain properties of those special solutions of $(*)$ which can be obtained through separation of variables. (Of special interest here is L_E which does not lead to separation of variables.)

Finally, in Sec. 4 we study the spectral resolution of L_K corresponding to nonunitary representations of the complex Euclidean group $CE(2)$ and obtain a series of identities for products of modified Bessel and Macdonald functions.

1. THE REPRESENTATIONS OF $E(1,1)$

The pseudo-Euclidean group $E(1,1)$ is the group of all real matrices

$$A(\theta,a,b,)=\begin{pmatrix}\cosh\theta & \sinh\theta & a\\ \sinh\theta & \cosh\theta & b\\ 0 & 0 & 1\end{pmatrix},\quad -\infty<\theta,a,b<\infty.$$

It acts on the pseudo-Euclidean plane via the transformation $\mathbf{z}\to A\mathbf{z}$ where

$$\mathbf{z}=\begin{pmatrix}t\\ s\\ 1\end{pmatrix}$$

and preserves the form $(t_1-t_2)^2-(s_1-s_2)^2$.

The irreducible faithful unitary representations of $E(1,1)$ are well-known to be indexed by a parameter $\gamma>0$. Each such representation can be defined by operators $\mathbf{T}(\theta,a,b)$,

$$\mathbf{T}(\theta,a,b)f(x)=\exp[i\gamma(a\cosh x+b\sinh x)]f(x+\theta) \tag{1.2}$$

acting on the Hilbert space $L_2(R)$ of Lebesgue square integrable functions $f(x)$ on the real line. The inner product is

$$\langle f,g\rangle=\int_{-\infty}^{\infty}f(x)\overline{g(x)}\,dx,\quad f,g\in L_2(R).$$

The Lie algebra $e(1,1)$ of $E(1,1)$ contains a basis $\{P_1,P_2,M\}$ with commutation relations

$$[M,P_1]=P_2,\quad [M,P_2]=P_1,\quad [P_1,P_2]=0 \tag{1.3}$$

and related to the group via the exponential mapping

$$A(\theta,a,b)=\exp(aP_1+bP_2)\exp(\theta M).$$

The corresponding operators in $L_2(R)$ induced by the group action (1.2) are easily shown to be

$$P_1=i\gamma\cosh x,\quad P_2=i\gamma\sinh x,\quad M=\partial_x. \tag{1.4}$$

Vilenkin (Ref. 4, Chap. V) has studied the unitary representation of $E(1,1)$ in terms of the spectral resolution for the operator

$$L_M=M^2$$

(or M) on $L_2(R)$. In particular, he has determined the matrix elements of the group operators (1.2) with respect to this resolution. In Ref. 5 the representations of $E(1,1)$ were examined with respect to the spectral resolution of the operator

$$L_D=MP_1+P_1M. \tag{1.5}$$

It was shown that L_D has a one parameter family of self-adjoint extensions $L_{D,\alpha}$, $0\le\alpha<2$. Each $L_{D,\alpha}$ has discrete spectrum $-2\gamma(\alpha+2n)$, $n=0,\pm1,\pm2,\cdots$ and normalized eigenfunctions

$$f_n^{D,\alpha}(x)=\sqrt{2\pi}\,\exp(x/2)(1+ie^x)^{\alpha+2n-1/2}(1-ie^x)^{-\alpha-2n-1/2}. \tag{1.6}$$

(In every example treated in this paper the L-operator is initially defined on the subspace of $L_2(R)$ consisting of C^∞-functions with compact support. One then searches for all self-adjoint extensions of this symmetric operator.)

This case was in sharp contrast to that of L_M where there was a single self-adjoint extension with continuous spectrum covering the negative real axis with generalized eigenfunctions

$$f_\lambda^M(x)=\frac{\exp(i\lambda x)}{\sqrt{2\pi i}},\quad -\infty<\lambda<\infty, \tag{1.7}$$

$$Mf_\lambda^M=i\lambda f_\lambda^M,\quad \langle f_\lambda^M,f_\mu^M\rangle=\delta(\lambda-\mu).$$

The spectral resolution was obtained via the Fourier transform. The relationship between these two bases was computed in Ref. 5.

In this paper we study the spectral resolutions in $L_2(R)$ of self-adjoint extensions of the symmetric operators found in Ref. 3:

$$\begin{aligned}L_B&=M^2-(P_1+P_2)^2,\\ L_K&=M^2+(P_1+P_2)^2,\\ L_E&=M(P_1-P_2)+(P_1-P_2)M,\\ L_A&=M(P_1-P_2)+(P_1-P_2)M+(P_1+P_2)^2.\end{aligned} \tag{1.8}$$

For each resolution we compute the matrix elements of the unitary operators (1.2). In addition we determine the unitary transformations which allow us to pass from one spectral resolution to another.

A. The Bessel function or B basis

It follows from (1.4) that

$$\begin{aligned}L_B&=D_x^2+\gamma^2\exp(2x)=v^2D_v^2+vD_v+\gamma^2v^2,\\ v&=\exp(x),\quad D_x=\frac{d}{dx},\quad D_v=\frac{d}{dv}.\end{aligned} \tag{1.9}$$

This operator is symmetric on $L_2(R)$ with deficiency indices $(1,1)$. Thus there is a one-parameter family $L_{B,\alpha'}$, $0\le\alpha'<2\pi$, of self-adjoint extensions of L_B. The domain of each extension is

$$\mathcal{D}_\alpha=\{f\in\mathcal{D}_{L^*_B}:\lim_{v\to+\infty}v[h_{\alpha'}(v)D_vf(v)-f(v)D_vh_{\alpha'}(v)]=0\}$$

where $\mathcal{D}_{L^*_B}$ is the domain of the adjoint of L_B in $L_2(R)$ and

$$h_{\alpha'}(v)=J_\beta(\gamma v)+\exp(i\alpha')J_{\bar\beta}(\gamma v),\quad \beta=\exp(i\pi/4),$$

where $J_\nu(z)$ is a Bessel function. (All special functions in this paper are defined as in Ref. 6.)

Each $L_{B,\alpha'}$ has discrete spectrum and an orthonormal basis of eigenfunctions

$$\begin{aligned}f_n^{B,\alpha}(v)&=\sqrt{2(\alpha+2n)}\,J_{\alpha+2n}(\gamma v),\\ v&=\exp(x),\quad n=0,1,2,\cdots,\end{aligned} \tag{1.10}$$

where $0<\alpha\le2$ and the fixed parameters α,α' are related by

$$\tan\left(\frac{\pi\alpha}{2}-\frac{\pi}{2\sqrt2}\right)=\left(\frac{1+\exp(\pi/\sqrt2)}{1-\exp(\pi/\sqrt2)}\right)\tan\frac{\alpha'}{2}.$$

(Our computations of spectral resolutions for first and second order ordinary differential operators, while certainly nontrivial, are straightforward,[7] so we omit the details.)

The relationship between different bases is easily computed:

$$f_m^{B,\alpha_1}(v)=\sum_{n=0}^{\infty}\langle f_m^{B,\alpha_1},f_n^{B,\alpha_2}\rangle f_n^{B,\alpha_2}(v),$$

$$\begin{aligned}\langle f_m^{B,\alpha_1},f_n^{B,\alpha_2}\rangle&=2\sqrt{(\alpha_1+2m)(\alpha_2+2n)}\int_0^\infty J_{\alpha_1+2m}(v)J_{\alpha_2+2n}(v)\frac{dv}{v}\\ &=\frac{\sqrt{(\alpha_1+2m)(\alpha_2+2n)}\sin\pi[(\alpha_1-\alpha_2)/2+m-n]}{\pi[(\alpha_1-\alpha_2)/2+m-n][(\alpha_1+\alpha_2)/2+m+n]}.\end{aligned} \tag{1.11}$$

The matrix elements of the unitary operator $\mathbf{T}(0,a,a)$, $a>0$ are

$$T_{mn}^{B,\alpha}(0,a,a)=\langle\exp a(P_1+P_2)f_n^{B,\alpha},f_m^{B,\alpha}\rangle$$

$$= 2\sqrt{(\alpha+2n)(\alpha+2m)}\int_0^\infty e^{iav}J_{\alpha+2n}(v)J_{\alpha+2m}(v)\frac{dv}{v}$$

$$= \frac{2\sqrt{(\alpha+2n)(\alpha+2m)}}{\Gamma(\alpha+2n+1)\Gamma(\alpha+2m)}(4a^2)^{\alpha+m+n}e^{-i\pi(\alpha+m+n)}$$

$$\times {}_4F_3\left(\begin{matrix}\alpha+n+m, \alpha+n+m+\frac{1}{2}, \alpha+n+m+\frac{1}{2}, \\ \alpha+n+m+1 \\ \alpha+2m+1, \alpha+2n+1, 2\alpha+2m+2n+1\end{matrix}\middle| \frac{1}{4a^2}\right) \tag{1.12}$$

where $\Gamma(z)$ is the gamma function and ${}_pF_q$ is a generalized hypergeometric function.

Further,

$$T^{B,\alpha}_{mn}(0,-a,-a) = \overline{T^{B,\alpha}_{nm}}(0,a,a).$$

The integral in (1.12) is evaluated with the help of Lebesgue's dominated convergence theorem and the device of expanding $J_{\alpha+2n}(v)J_{\alpha+2m}(v)$ into a power series in v and integrating term by term. There is a similar unenlightening expression for the matrix elements $T^{B,\alpha}_{mn}(0,a,-a)$ which we omit.

The matrix elements of the operator $\mathbf{T}(\theta,0,0,)$ are

$$T^{B,\alpha}_{mn}(\theta,0,0) = \langle \exp(\theta M) f^{B,\alpha}_n, f^{B,\alpha}_m\rangle$$

$$= 2\sqrt{(\alpha+2n)(\alpha+2m)}\int_0^\infty J_{\alpha+2n}(e^\theta v)J_{\alpha+2m}(v)\frac{dv}{v}$$

$$= e^{-(\alpha+2m)\theta}\frac{\sqrt{(\alpha+2n)(\alpha+2m)}\,\Gamma(\alpha+n+m)}{\Gamma(\alpha+2m+1)\Gamma(1+n-m)}$$

$$\times {}_2F_1\left(\begin{matrix}\alpha+n+m, m-n \\ \alpha+2m+1\end{matrix}\middle| e^{-2\theta}\right)$$

for $\theta \geq 0$. (1.13)

[This is a Weber-Schafheithin integral (Ref. 6, Vol. II.)] Furthermore,

$$T^{B,\alpha}_{mn}(-\theta,0,0) = \overline{T^{B,\alpha}_{nm}}(\theta,0,0).$$

Note that the matrix elements (1.13) vanish if $m \geq n+1$.

B. The Macdonald function or K basis

From (1.4) it follows that

$$L_K = D_x^2 - \gamma^2 e^{2x} = v^2 D_v^2 + vD_v - \gamma^2 v^2. \tag{1.14}$$

This operator is symmetric on $L_2(R)$ and has deficiency indices (0,0). Thus L_K has a unique self-adjoint extension (which we also call L_K) and a complete set of orthonormal eigenfunctions of L_K, f^K_x, which form a basis for the representation space. The spectral resolution of L_K can be obtained from the known form of the Lebedev integral transform (Ref. 6, Vol. II). The spectrum of L_K is continuous and an orthonormal basis of eigenfunctions is ($\mathrm{sh}z = \sinh z$, $\mathrm{ch}z = \cosh z$)

$$f^K_z(v) = \frac{1}{\pi}\sqrt{2z\,\mathrm{sh}\pi z}\,K_{iz}(\gamma v), \quad 0 < z < \infty. \tag{1.15}$$

These basis functions satisfy the delta function normalization

$$\langle f^K_x, f^K_y\rangle = \delta(x-y).$$

The matrix elements of the unitary operator $\mathbf{T}(0,a,a)$, $a > 0$, are in this basis

$$T^K_{xy}(0,a,a) = \langle \exp a(P_1+P_2) f^K_y, f^K_x\rangle$$

$$= \frac{\pi^2}{2}\sqrt{xy\,\mathrm{sh}\pi x\,\mathrm{sh}\pi y}\int_0^\infty \exp(iav)K_{ix}(\gamma v)K_{iy}(\gamma v)\frac{dv}{v}$$

$$= \delta(x-y)$$

$$+ \frac{1}{4\pi^2}\sqrt{xy\,\mathrm{sh}\pi x\,\mathrm{sh}\pi y}\Big[\tfrac{1}{4}a(\mathrm{ch}\pi x - \mathrm{ch}\pi y)$$

$$\times {}_4F_3\Big(\frac{1+ix+iy}{2}, \frac{1+ix-iy}{2}, \frac{1+iy-ix}{2},$$

$$\frac{1-ix-iy}{2}; \tfrac{1}{2}, 1, \tfrac{3}{2}; -\tfrac{1}{4}a^2\Big)$$

$$+ a^2\frac{(\mathrm{ch}\pi x + \mathrm{ch}\pi y)}{y^2-x^2}\,{}_4F_3\Big(1+\frac{i(x+y)}{2},$$

$$1+\frac{i(x-y)}{2}, 1+\frac{i(x-y)}{2},$$

$$1-\frac{i(x+y)}{2}; \tfrac{3}{2}, \tfrac{3}{2}, 2; -\tfrac{1}{4}a^2\Big)\Big]. \tag{1.16}$$

This integral is evaluated by expanding the exponential in a power series in v and integrating term by term. We omit the evaluation of the matrix elements of the operator $\mathbf{T}(0,a,-a)$.

The matrix elements of the operator $\mathbf{T}(\theta,0,0)$ are

$$T^K_{xy}(\theta,0,0) = \langle \exp(-\theta M) f^K_y, f^K_x\rangle$$

$$= \frac{2}{\pi^2}\sqrt{xy\,\mathrm{sh}\pi x\,\mathrm{sh}\pi y}\int_0^\infty K_{ix}(v)K_{iy}(e^\theta v)\frac{dv}{v}$$

$$= \cos(\theta y)\delta(x-y)$$

$$+ \frac{1}{4\pi^2}\sqrt{xy\,\mathrm{sh}\pi x\,\mathrm{sh}\pi y}$$

$$\times\Big[e^{iy\theta}\frac{\Gamma(1-iy)}{\Gamma[(ix-iy-1)/2]\Gamma[(-ix-iy-1)/2]}$$

$$\times {}_2F_1\Big(\frac{ix-iy-1}{2}, \frac{-ix-iy-1}{2}; -iy; \exp(-2\theta)\Big)$$

$$+ \exp(-iy\theta)\frac{\Gamma(1+iy)}{\Gamma[(ix+iy-1)/2]\Gamma[(-ix+iy-1)/2]}$$

$$\times {}_2F_1\Big(\frac{ix+iy-1}{2}, \frac{-ix+iy-1}{2}; iy; \exp(-2\theta)\Big)\Big] \tag{1.17}$$

for $\theta > 0$.

For $\theta < 0$ the matrix elements can be obtained from the relation

$$T^K_{xy}(-\theta,0,0)=\overline{T^K_{yx}}(\theta,0,0).$$

C. The exponential or E basis

The basis defining operator in the realization (1.4) has the form

$$L_E=i\gamma(2e^{-x}D_x-e^{-x})=i\gamma(2D_v-v^{-1}).$$

Solutions of the eigenfunction equation $L_EF(v)=\lambda F(v)$ then are

$$F(v)=Cv^{1/2}\exp\left(-\frac{i\lambda}{2\gamma}v\right). \tag{1.18}$$

These eigenfunctions do not form a complete set on the Hilbert space $\mathcal{H}^+$ on which the representation (1.2) is defined, i.e., the space of functions $f(v)$ for $0\le v<\infty$, $v=e^x$. The correct group[3] in which to realize this basis is $E'=E(1,1)\oplus\{R,I\}$ where R is the reflection operator in the pseudo-Euclidean plane and I is the identity operator. R acts on the generators of $E(1,1)$ according to

$$R:M\to M,\quad R:P_i\to -P_i\quad (i=1,2) \tag{1.19}$$

The Hilbert space H on which the irreducible representation labelled by γ is realized is now the direct sum of two Hilbert spaces $H=\mathcal{H}^+\oplus\mathcal{H}^-$ with $\mathcal{H}^-$ the space of functions $f(v)$ for $-\infty<v\le 0$ which are square integrable with respect to the measure dv/v and transform under the group E' according to (1.2) with $v=e^x$ (remember $R:e^x\to -e^x$). In fact, we can write symbolically $\mathcal{H}^-=R\mathcal{H}^+$. Accordingly, each $f(v)\in H$ $(-\infty<v<\infty)$ satisfies the integrability condition

$$\int_{-\infty}^{\infty}|f(v)|^2\frac{dv}{v}<\infty \tag{1.20}$$

with the group action given quite generally by (1.2) with $e^x=v$. The operator L_E is then essentially self adjoint and the eigenfunctions correspond to a form of the exponential solutions of the momentum operator. The spectrum of L_E is the real axis and a complete set of orthonomal eigenfunctions is

$$f^E_\lambda(v)=\frac{1}{2}\left(\frac{-v}{2\pi}\right)^{1/2}\exp\left(-\frac{i\lambda}{2\gamma}v\right) \tag{1.21}$$

where

$$\langle f^E_\lambda,f^E_{\lambda'}\rangle=\int_{-\infty}^{\infty}f^E_\lambda(v)\overline{f^E_{\lambda'}(v)}\frac{dv}{v}$$

$$=\delta(\lambda-\lambda').$$

In (1.21) we make the consistent convention that the square root $(-v)^{1/2}$ for v positive be taken as $+|v|^{1/2}$. The matrix elements in the E basis can be easily calculated.

The matrix element for the unitary operator $\mathbf{T}(0,a,a)$ is

$$T^E_{\lambda\lambda'}(0,a,a)=\frac{-1}{4\pi\gamma}\int_{-\infty}^{\infty}\exp\left(\frac{i}{2\gamma}(\lambda'-\lambda)+i\gamma a\right)v\,dv$$

$$=\delta(\lambda-\lambda'-2\gamma^2a). \tag{1.22}$$

For the operator $T(0,a,-a)$ we have the result

$$T^E_{\lambda\lambda'}(0,a,-a)=\frac{-1}{4\pi\gamma}\int_{-\infty}^{\infty}\exp\left(\frac{i}{2\gamma}(\lambda'-\lambda)v+\frac{i\gamma a}{v}\right)dv$$

$$=\delta\ =\delta(\lambda-\lambda')+\gamma\sqrt{2}a/(\lambda'-\lambda)\,\mathrm{sign}(\lambda'-\lambda)$$

$$\times J_1(2a\sqrt{\lambda'-\lambda}). \tag{1.23}$$

This matrix element can be evaluated by expanding $\exp(i\gamma a/v)$ in a power series then integrating term by term in the sense of generalized functions.[8] Alternatively, contour integration of the regular part of the matrix element will give the same result.

The matrix element of the operator $\mathbf{T}(\theta,0,0)$ is

$$T^E_{\lambda\lambda'}(\theta,0,0)=\frac{-1}{4\pi\gamma}\int_{-\infty}^{\infty}\exp\left(\frac{i}{2\gamma}(\lambda'-e^\theta\lambda)v\right)dv$$

$$=\delta(e^\theta\lambda-\lambda'). \tag{1.24}$$

D. The Airy function or A basis

In the realization (1.4) L_A has the form

$$L_A=i\gamma(2e^{-x}D_x-e^{-x})-\gamma^2e^{2x} \tag{1.25}$$

$$=i\gamma(2D_v-v^{-1})-\gamma^2v^2.$$

The solutions of the eigenfunction equation $L_AF(v)=\lambda F(v)$ are

$$F(v)=v^{1/2}\exp\left(\frac{i}{6}\gamma v^3-\frac{i}{2}\frac{\lambda}{\gamma}v\right). \tag{1.26}$$

As with the E basis these eigenfunctions do not form a complete set on the space $\mathcal{H}^+$ of functions $f(v)$ with $0\le v<\infty$. This space is extended in exactly the same way as for the E basis. A complete set of orthonormal eigenfunctions on $H=\mathcal{H}^+\oplus\mathcal{H}^-$ is then

$$f^A_\lambda=\frac{1}{2}\left(\frac{-v}{\pi\gamma}\right)^{1/2}\exp\left(\frac{i}{6}\gamma v^3-\frac{i}{2}\frac{\lambda}{\gamma}v\right) \tag{1.27}$$

with

$$\langle f^A_\lambda,f^A_{\lambda'}\rangle=\delta(\lambda-\lambda').$$

The matrix elements in this basis can be easily calculated. For the translations $\mathbf{T}(0,a,\pm a)$ the results are the same as for the E basis, viz., (1.22) and (1.23). For the matrix element of the operator $\mathbf{T}(\theta,0,0)$ we have a new result:

$$T^A_{\lambda\lambda'}(\theta,0,0)=\frac{-1}{4\pi\gamma}\int_{-\infty}^{\infty}\exp\left(\frac{i}{6}\gamma(e^{3\theta}-1)v^3+\frac{i}{2\gamma}(\lambda'-\lambda)v\right)dv$$

$$=\frac{-1}{4\gamma}\left(\frac{2}{\gamma}\right)^{1/3}(e^{3\theta}-1)^{-1/3}\,\mathrm{Ai}\left(\frac{\lambda'-\lambda}{[2\gamma^2(e^{3\theta}-1)]^{1/3}}\right) \tag{1.28}$$

for $\theta>0$ and where $\mathrm{Ai}(z)$ is an Airy function. The matrix element for $\theta<0$ can be obtained by using the result

$$T^A_{\lambda\lambda'}(-\theta,0,0)=\overline{T^A_{\lambda'\lambda}}(\theta,0,0).$$

2. OVERLAP FUNCTIONS

In this section we compute functions of the form

$$U^{G,H}_{n,m}=\langle f^G_n,f^H_m\rangle=\int_{-\infty}^{\infty}f^G_n(x)\overline{f^H_m}(x)dx \tag{2.1}$$

which allow us to pass from the $\{f^G\}$ basis to the $\{f^H\}$ basis via the expression

$$f_n^G(x)=\sum_m U_{n,m}^{G,H} f_m^H(x). \tag{2.2}$$

(For $H=M,K,A.E$ the sum should be replaced by an integral.) Note that

$$U_{n,m}^{G,H}=\overline{U}_{m,n}^{H,G}, \tag{2.3a}$$

$$U_{n,m}^{G,G}=\delta_{n,m}, \quad n,m \text{ discrete,}$$
$$=\delta(n-m), \quad n,m \text{ continuous,} \tag{2.3b}$$

$$U_{n,m}^{G,H}=\sum_p U_{n,p}^{G,J} U_{p,m}^{J,H}. \tag{2.3c}$$

In the following we compute the various $U_{n,m}^{G,H}$ by substituting the explicit expressions for $f_n^G(x)$, $f_m^H(x)$ from Sec. 1 into (2.1) and evaluating the integral. In case both L_G and L_H have continuous spectrum then expressions (2.1), (2.2) must be interpreted in the sense of generalized functions.

First we relate all bases to the standard M basis:

$$U_{n,\lambda}^{G,M}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f_n^G(x)\exp(-i\lambda x)dx.$$

The results are

$$U_{n,\lambda}^{B,M}=\langle f_n^{B,\alpha}, f_\lambda^M\rangle=\frac{\sqrt{\alpha+2n}}{\pi}\int_0^\infty J_{\alpha+2n}(\gamma v)v^{-1-i\lambda}dv \tag{2.4}$$

$$=\left(\frac{\alpha+2n}{\pi}\right)^{1/2}\frac{(\gamma/2)^{i\lambda}}{2}\frac{\Gamma(\alpha+n-i\lambda/2)}{\Gamma[1+n+(\alpha+i\lambda)/2]},$$

$$U_{x,\lambda}^{K,M}=\langle f_x^K, f_\lambda^M\rangle$$

$$=\frac{1}{\pi^{3/2}}\sqrt{x\operatorname{sh}\pi x}\int_0^\infty v^{-i\lambda-1}K_{ix}(\gamma v)\,dv$$

$$=\frac{i}{2}\left(\frac{x}{\pi\operatorname{sh}\pi x}\right)^{1/2}\left(\frac{(2\gamma)^{+ix}}{\Gamma(1+ix)}\delta(x-\lambda)-\frac{(2\gamma)^{-ix}}{\Gamma(1-ix)}\delta(x+\lambda)\right)$$

$$+\frac{1}{4\pi}\left(\frac{x\operatorname{sh}\pi x}{\pi}\right)\left(\frac{\gamma}{2}\right)^{i\lambda}\Gamma\left(\frac{ix-i\lambda}{2}\right)\Gamma\left(\frac{ix+i\lambda}{2}\right), \tag{2.5}$$

$$U_{K,\lambda}^{E,M}=\frac{i}{2\pi(2\gamma)^{1/2}}\int_0^\infty v^{-i\lambda-1/2}\exp\left(-\frac{iK}{2\gamma}v\right)dv$$

$$=-\frac{\Gamma(\frac{1}{2}-i\lambda)}{2\pi\sqrt{2\gamma}}e^{\epsilon(i\pi/4+\pi\lambda/2)}\left|\frac{K}{2\gamma}\right|^{-1/2+i\lambda} \tag{2.6}$$

where $\epsilon=+1$ if $K<0$ and -1 if $K>0$. We have

$$U_{K,\lambda}^{A,M}=\frac{i}{2\pi(2\gamma)^{1/2}}\int_0^\infty v^{-i\lambda-1/2}\exp\left(\frac{i\gamma}{6}v^3-\frac{iK}{2\gamma}v\right)dv$$

$$=\frac{i}{2\pi(2\lambda)^{1/2}}\left(\frac{\gamma}{6i}\right)^{(2i\lambda-1)/6}\sum_{n=0}^{\infty}\frac{\Gamma[(n-i\lambda)/3+\frac{1}{6}]}{n!}$$

$$\times\left(\frac{e^{-i\pi/6}K}{2\gamma}\right)^n\left(\frac{6}{\gamma}\right)^{n/3}. \tag{2.7}$$

This expression can also be written as a sum of three ${}_1F_2(a;b,c;z)$ hypergeometric functions.

It should be mentioned here that the overlap coefficients which we have given relating the E and A bases to the M basis are not the complete coefficients with respect to the group E' and hence are not unitary. The coefficients we have calculated only relate these bases on the Hilbert space $\mathcal{H}^+$. A similar calculation for the space $\mathcal{H}^-$ can be made but we do not do this here. For the unitary irreducible representations of E' in a B, K, or M basis we have basis functions labelled by an additional discrete label corresponding to the eigenvalues ± 1 of the reflection operator. This is because R commutes with the operators L_B, L_K, and L_M. For the A and E bases however R does not commute with L_A or L_E. Hence no such labels are required. For the purposes of this paper we have not introduced this discrete label, it being understood whenever we give an overlap function.

We now give a number of further overlap functions of interest:

$$U_{n,x}^{B,K}=\langle f_n^{B,\alpha}, f_x^K\rangle$$

$$=\frac{1}{2\pi}\sqrt{(\alpha+2n)x\operatorname{sh}\pi x}$$

$$\times\frac{\Gamma[n+(\alpha+ix)/2]\Gamma[n+(\alpha-ix)/2]}{\Gamma(1+\alpha+2n)}$$

$$\times{}_2F_1\left(n+\frac{\alpha+ix}{2},\ n+\frac{\alpha-ix}{2};\ 1+\alpha+2n;-1\right), \tag{2.8}$$

$$U_{\lambda,\lambda'}^{E,A}=\langle f_\lambda^E, f_{\lambda'}^A\rangle$$

$$=-(2\gamma^2)^{-1/3}\mathrm{Ai}((\lambda-\lambda')(2\gamma^2)^{-1/3}),$$

$$U_{x,\lambda}^{K,E}=-\frac{i\sqrt{x\operatorname{sh}\pi x}}{\pi}\frac{(2\gamma)^{2ix-1/2}}{(2\gamma^2+i\lambda)^{ix+1/2}}\Gamma(\tfrac{1}{2}+ix)\Gamma(\tfrac{1}{2}-ix)$$

$$\times{}_2F_1\left(\tfrac{1}{2}+ix,\tfrac{1}{2}+ix;1;\frac{2\gamma^2+i\lambda}{2\gamma^2-i\lambda}\right). \tag{2.9}$$

3. A TWO-VARIABLE MODEL FOR $E(1,1)$

As mentioned in Sec. 1, $E(1,1)$ acts as a transformation group in the pseudo-Euclidean plane. We choose this action in the s-t plane so that the Lie derivatives corresponding to the Lie algebra basis $\{P_1,P_2,M\}$ are

$$P_1=\partial_s,\quad P_2=\partial_t,\quad M=-t\partial_s-s\partial_t. \tag{3.1}$$

We now construct models of the irreducible representations of $E(1,1)$ where the Lie algebra acts via the operators (3.1) rather than (1.4). In particular, we construct the two-variable analogs of the basis functions $\{f_n^G\}$.

In the one-variable model we have $(P_2^2-P_1^2)f_n^G=\gamma^2 f_n^G$ for each basis function f, so we would expect the same equation to hold in the two-variable model, i.e.,

$$(\partial_t^2-\partial_s^2)F_n^G(s,t)=\gamma^2 F_n^G(s,t),$$

where $F_n^G(s,t)$ is the two-variable function corresponding to $f_n^G(x)$. In the following we will define a mapping $f(x)\to F(s,t)$ from $L_2(R)$ to functions on the pseudo-Euclidean plane such that $(\partial_t^2-\partial_s^2)F=\gamma^2F$ and such the eigenfunctions $f_n^G(x)$ of L_G map to eigenfunctions $F_n^G(s,t)$ of the corresponding operator L_G constructed from (3.1). Because of the close relationship between separation of variables and operators L_G we can find simple expressions for the $\{F_n^G(s,t)\}$ in terms of the coordinates associated with L_G. (The single exception to this statement is the case $G=E$ where there is no associated coordinate system in which the variables separate.)

To make our construction precise we introduce the functions

$$h_{s,t}(x)=\exp[i\gamma(s\cosh x+t\sinh x)],\quad s,\ t\in\mathbf{C}. \tag{3.2}$$

which belong to $L_2(R)$ for $\operatorname{Im}\gamma(s\pm t)>0$. Given $f(x)\in L_2(R)$, we define a function $F(s,t)$ by

$$F(s,t)=\langle f,\overline{h_{s,t}}\rangle=\int_{-\infty}^{\infty}f(x)h_{s,t}(x)dx. \tag{3.3}$$

In particular, corresponding to a basis $\{f_n^G\}$ for $L_2(R)$ we obtain functions

$$F_n^G(s,t)=\langle f_n^G,\overline{h_{s,t}}\rangle. \tag{3.4}$$

The action (1.2) of $E(1,1)$ on $L_2(R)$ induces an action on the $F(s,t)$ which satisfies the homomorphism property:

$$\begin{aligned}[\mathrm{T}(\theta,a,b)F](s,t)&\equiv\langle \mathrm{T}(\theta,a,b)f,\overline{h_{s,t}}\rangle\\ &=\langle f,\mathrm{T}(\theta,a,b)^{-1}\overline{h_{s,t}}\rangle\\ &=F((s+a)\cosh\theta-(t+b)\sinh\theta,\\ &\qquad(t+b)\cosh\theta-(s+a)\sinh\theta).\end{aligned} \tag{3.5}$$

It is easy to check that the Lie derivatives corresponding to the group action (3.5) coincide with (3.1). Thus the operators (1.4) acting on f correspond to the operators (3.1) acting on F.

On the other hand, for f a basis vector f_n^G we have

$$\begin{aligned}[\mathrm{T}(\theta,a,b)F_n^G](s,t)&=\langle \mathrm{T}(\theta,a,b)f_n^G,\overline{h}_{s,t}\rangle\\ &=\sum_m T_{mn}^G(\theta,a,b)F_m^G(s,t)\end{aligned} \tag{3.6}$$

where the T_{mn}^G are the G-basis matrix elements. It follows from (3.5) and (3.6) that the $\{F_n^G\}$ transform under $E(1,1)$ exactly as the basis vectors $\{f_n^G\}$. In particular,

$$(P_2^2-P_1^2)F_n^G=\gamma^2F_n^G,\quad L_GF_n^G=\lambda_nF_n^G \tag{3.7}$$

[where P_1,P_2,L_G are expressed in terms of the operators (3.1)], provided $L_Gf_n^G=\lambda_nf_n^G$. Relations (3.6) also hold even for $\operatorname{Im}\gamma(s\pm t)=0$ if the $\{f_n^G\}$ belong to $L_1(R)$.

If $h_{s,t}\in L_2(R)$ it follows immediately that

$$h_{s,t}(x)=\sum_n\overline{f}_n^G(x)F_n^G(s,t) \tag{3.8}$$

where the right-hand side converges in $L_2(R)$ and also pointwise. (As usual, if L_G has continuous spectrum we replace the sum by an integral.) We can consider (3.8) as the expansion of a plane wave in a $\{F_n^G\}$ basis of solutions of the Helmholtz equation.

It follows directly from the definition of $h_{s,t}(x)$ that

$$\langle h_{s_1,t_1},h_{\overline{s}_2,\overline{t}_2}\rangle=2K_0(i\gamma[(s_1-s_2)^2-(t_1-t_2)^2]^{1/2}). \tag{3.9}$$

On the other hand, use of (3.8) yields

$$\langle h_{s_1,t_1},h_{\overline{s}_2,\overline{t}_2}\rangle=\sum_nF_n^G(s_1,t_1)\overline{F}_n^G(\overline{s}_2,\overline{t}_2). \tag{3.10}$$

Comparison of (3.9) and (3.10) yields another generating function for the $\{F_n^G\}$.

The overlap functions computed in Sec. 2 carry over immediately to the two variable model. Indeed, expression (2.2) relating the bases $\{f_n^G\}$ and $\{f_m^H\}$ yields

$$F_n^G(s,t)=\sum_mU_{n,m}^{G,H}F_m^H(s,t) \tag{3.11}$$

with the same overlap functions $U_{n,m}^{G,H}$.

It follows from these remarks that the functions $\{F_n^G\}$ will necessarily satisfy the identities (3.5)—(3.11) where the matrix elements $T_{mn}^G(\theta,a,b)$ and overlap functions $U_{n,m}^{G,H}$ have already been computed from the one-variable model. Moreover, due to the relationship between the operators L_G and separation of variables for the Helmholtz equation we can find simple expressions for the function $\{F_n^G\}$ in terms of the coordinate system related to L_G. Indeed, evaluating the integral (3.4) in each case, we find

$$F_\lambda^M[\rho,\theta]=\sqrt{2}/\pi e^{-i\lambda\theta}K_{i\lambda}(i\gamma\rho),$$

$$s=\rho\cosh\theta,\quad t=\rho\sinh\theta, \tag{3.12}$$

$$\begin{aligned}F_n^{D,\alpha}[\xi,\eta]&=2\exp[3i\pi(\alpha+2n-1)/2]D_{-\alpha-2n-1/2}(\sqrt{-2\gamma}\,\xi)\\ &\quad\times D_{\alpha+2n-1/2}(\sqrt{-2\gamma}\,\eta),\end{aligned} \tag{3.13}$$

$$s=i\xi\eta,\quad t=(\eta^2-\xi^2)/2,$$

$$F_n^{B,\alpha}[u,v]=2\sqrt{2(\alpha+2n)}\,J_{\alpha+2n}(\gamma u)K_{\alpha+2n}(-i\gamma v), \tag{3.14}$$

$$s=\frac{u^2+u^2v^2+v^2}{2uv},\quad t=\frac{u^2-u^2v^2+v^2}{2uv},\quad \left|\frac{u}{v}\right|<1,$$

$$F_x^K[u,v]=\frac{2}{\pi}\sqrt{x\operatorname{sh}\pi x}\,K_{ix}(\gamma u)K_{ix}(-i\gamma v), \tag{3.15}$$

$$s=\frac{u^2-u^2v^2-v^2}{2uv},\quad t=\frac{u^2+u^2v^2-v^2}{2uv},\quad \left|\frac{u}{v}\right|>1,$$

$$\begin{aligned}F_\lambda^E[s,t]&=\int_{-\infty}^{\infty}\exp\left(\frac{i\gamma(s+t)}{2}v+\frac{i\gamma(s-t)}{2v}\right)f_\lambda^E(v)\frac{dv}{v}\\ &=\frac{i}{2\sqrt{2\gamma^2(s+t)-2\lambda}}\exp\sqrt{\gamma(t^2-s^2)-(\lambda/\gamma)(t-s)},\\ &\qquad s+t>\lambda/\gamma^2,\quad t>s,\\ &=\frac{i}{2\sqrt{2\gamma^2(s+t)-2\lambda}}\cos\sqrt{\gamma(s^2-t^2)-(\lambda/\gamma)(s-t)},\\ &\qquad s+t>\lambda/\gamma^2,\quad s>t.\end{aligned} \tag{3.16}$$

Similar expressions can be given for the other ranges of s and t:

$$F_\lambda^A[x_1,x_2]=\int_0^{\infty}\exp\left(\frac{i\gamma(s+t)}{2}v+\frac{i\gamma(s-t)}{2v}\right)f_\lambda^A(v)\frac{dv}{v}$$

$$=A\phi_1(y_1)\phi_1(y_2)+B(\phi_1(y_1)\phi_2(y_2)+\phi_1(y_2)\phi_2(y_1))+C\phi_2(y_1)\phi_2(y_2), \tag{3.17}$$

where

$$\phi_1(y)={}_0F_1(\tfrac{2}{3},\tfrac{1}{9}y_1^3),\quad \phi_2(y)={}_0F_1(\tfrac{4}{3},\tfrac{1}{9}y_1^3).$$

The coefficients are given by

$$A=\frac{\Gamma(\frac{1}{6})}{6\sqrt{\pi\gamma}}\left(\frac{6}{\gamma}\right)^{1/6},\quad B=-\frac{2i\gamma^{5/2}}{3}\left(\frac{6}{\gamma}\right)^{2/3},$$

$$C=\frac{16i\gamma^{11/2}}{3}\left(\frac{6}{\gamma}\right)^{5/6}\Gamma\left(\frac{5}{6}\right),$$

s and t are given by the relations

$$2s=-\tfrac{1}{2}(x_1-x_2)^2+(x_1+x_2),$$

$$2t=-\tfrac{1}{2}(x_1-x_2)^2-(x_1+x_2),$$

and

$$y_i=\frac{1}{4\gamma^2}\left(x_i\frac{\lambda}{4\gamma^2}\right)\quad (i=1,2).$$

The expression we have given for the A basis functions in the two parameter model can also be written as a sum of products of Airy functions. One comment should be made here concerning the $F_\lambda^E[s,t]$ functions. These functions indicate that for the two variable model the E basis functions do not afford a separation of variables. This is in agreement with an earlier result.[3]

4. REPRESENTATIONS OF $CE(2)$

For the purpose of relating Lie group theory to special functions it is imperative to study group representations which have no Hilbert space structure, in particular representations defined on spaces of analytic functions. Some example of these were given in Refs. 5 and 9. For such representations one can always assume that the group is complex and we shall do so here by allowing the parameters θ, a, b in (1.1) to take arbitrary complex values. Thus, we shall consider representations of the complex Euclidean group $CE(2)$.

The Lie algebra $ce(2)$ of $CE(2)$ consists of all complex linear combinations of the generators M, P_1, P_2 with commutation relations

$$[M,P_1]=P_2,\quad [M,P_2]=P_1,\quad [P_1,P_2]=0. \tag{4.1}$$

We consider a model of this algebra in which the generators are given by

$$M=z\frac{d}{dz},\quad P^+=\rho z,\quad P^-=\rho z^{-1},\qquad P^\pm=P_1\pm P_2 \tag{4.2}$$

acting on the space $\mathcal{J}^\nu$ of functions $f(z)$ analytic in the domain $|z|>0$ with periodicity $f(e^{2\pi i}z)=e^{2\pi i\nu}f(z)$. Here $\nu\in\mathbf{C}$ is not an integer. The eigenfunctions of the operator $L_K=M^2+(P_1+P_2)^2=M^2+(P^+)^2$ on this space are easily seen to be

$$f_n^K(z)=J_{\nu+n}(\rho z),\quad n=0,\pm1,\cdots,$$

$$L_Kf_n^K=(\nu+n)^2f_n^K \tag{4.3}$$

to within a constant factor. Moreover, as shown in Ref. 10, p. 204, every $f\in\mathcal{J}^\nu$ can be expanded as an infinite series in the eigenfunctions f_n^K

$$f(z)=\sum_{n=-\infty}^{\infty}c_nJ_{\nu+n}(\rho z), \tag{4.4}$$

where the coefficients c_n are given by

$$c_n=\frac{1}{2\pi i}\frac{\pi(\nu+n)}{\sin\pi(\nu+n)}\int f(z)J_{-\nu-n}(\rho z)\frac{dz}{z} \tag{4.5}$$

and the pointwise convergence in (4.4) is uniform on compact subsets of the annulus. The path of integration in (4.5) can be chosen as a circle centered at the origin with radius $r>0$.

It follows from (4.2) that the action of $CE(2)$ on $\mathcal{J}^\nu$ is given by operators $\mathbf{T}(\theta,a,b)$,

$$[\mathbf{T}(\theta,a,b)f](z)=\exp\frac{\rho}{2}[(a+b)z+(a-b)/z]f(e^\theta z), \tag{4.6}$$

and that $\mathcal{J}^\nu$ is invariant under this action. Thus, we can use expressions (4.4) and (4.5) to compute the matrix elements of the operators $M,P^\pm$ and $\mathbf{T}(\theta,a,b)$ in the $\{f_n^K\}$ basis. It is straightforward to show

$$P^+f_n^K=\sum_{m=0}^{\infty}(-1)^m(\nu+n+2m+1)f_{n+2m+1}^K,$$

$$P^-f_n^K=\frac{\rho^2}{2(\nu+n)}(f_{n-1}^K+f_{n+1}^K), \tag{4.7}$$

$$Mf_n^K=\frac{(\nu+n)}{2}f_n^K-\sum_{m=0}^{\infty}(-1)^m(\nu+n+2m+2)f_{n+2m+2}^K,$$

$$T(0,a,a)_{m,n}=\begin{cases}\dfrac{(2a)^l\Gamma(\nu+m+1)}{\Gamma(\nu+n+1)l!}\,{}_4F_3\left(\begin{matrix}-\frac{l}{2},\ \frac{1-l}{2},\ 1-\frac{l}{2},\frac{1}{2}-\frac{l}{2}\\ 1-l,\ -\nu-m+1,\ \nu+n+1\end{matrix}\middle|\ \frac{-4}{a^2}\right), & \text{if } m-n=l\geq 0,\\ 0 & \text{if } m-n<0,\end{cases} \tag{4.8}$$

$$T(0,a,-a)_{m,n}=\frac{(\rho^2a/2)^{n-m}\Gamma(\nu+m+1)}{\Gamma(\nu+n+1)\Gamma(n-m+1)}\times{}_0F_3\left(-\nu-m+1,\ \nu+n+1,\ n-m+1;\ -\frac{\rho^4a^2}{4}\right),$$

$$T(\theta,0,0)_{m,n}=\begin{cases}\dfrac{(-1)^le^{(\nu+n)\theta}\Gamma(\nu+m+1)}{l!\,\Gamma(\nu+n+1)}\,{}_2F_1\left(\begin{matrix}-l,\nu+m-l\\ \nu+n+1\end{matrix}\middle|\ e^{2\theta}\right), & \text{if } m-n=2l, l=0,1,2,\ldots,\\ 0 & \text{otherwise.}\end{cases}$$

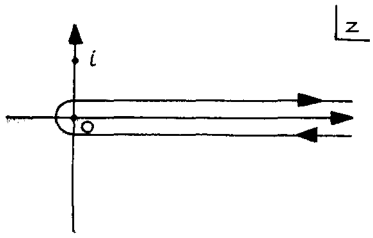


FIG. 1. Contour of Integration.

Note that in our model $P_1^2 - P_2^2 = P^+P^- = \rho^2$.

Next we construct a model of this representation in terms of functions $F_n^K(s,t)$ in the complex $s-t$ plane. Here,

$$P_1 = \partial_s, \quad P_2 = \partial_t, \quad M = -t\partial_s - s\partial_t, \tag{4.9}$$

so the basis functions $F_n^K(s,t)$ analogous to $f_n^K(z)$ must satisfy the equations

$$(\partial_t^2 - \partial_s^2)F_n^K(s,t) = -\rho^2 F_n^K(s,t), \quad L_K F_n^K(s,t) = (\nu+n)F_n^K(s,t). \tag{4.10}$$

In analogy with a similar construction in Ref. 5 and (3.3), we set

$$F_n^K(s,t) = \int_C \exp\frac{\rho}{2}[z(s+t) + z^{-1}(s-t)] f_n^K(z)\frac{dz}{z}, \quad \mathrm{Re}(s+t) < 0, \tag{4.11}$$

where C is the contour in the complex z plane (see Fig. 1). By differentiating under the integral sign in (4.11) and integration by parts it is easy to show that the generators (4.2) acting on $f_n^K(z)$ correspond to the generators (4.9) acting on $F_n^K(s,t)$. Thus Eqs. (4.10) must hold. This suggests that the $F_n^K(s,t)$ are simply expressable in terms of the u-v coordinates,

$$s = \frac{u^2 + u^2v^2 + v^2}{2uv}, \quad t = \frac{u^2 - u^2v^2 + v^2}{2uv}.$$

Indeed, a direct computation yields

$$F_n^K[u,v] = 4i \exp i(\pi/2)(n-\nu) \sin\pi\nu\sqrt{2(\nu+n)}\, I_{\nu+n}(-\rho u)K_{\nu+n}(-\rho v), \quad |u/v| < 1. \tag{4.12}$$

From (4.9) it follows that the action of $CE(2)$ on the basis $F_n^K(s,t)$ is given by

$$[\mathrm{T}(\theta,a,b)F_n^K](s,t) = F_n^K((s+a)\cosh\theta - (t+b)\sinh\theta, (t+b)\cosh\theta - (s+a)\sinh\theta). \tag{4.13}$$

On the other hand, by construction we have

$$[\mathrm{T}(\theta,a,b)F_n^K](s,t) = \sum_{m=-\infty}^{\infty} T(\theta,a,b)_{mn}F_m^K(s,t), \quad \mathrm{Re}[(s+t+a+b)e^{-\theta}] < 0, \tag{4.14}$$

where the matrix elements $T(\theta,a,b)_{mn}$ are given by (4.8). Comparison of expressions (4.12)–(4.14) yields addition theorems for the basis (4.12) whose direct derivation is not at all obvious. Other choices of the contour C in (4.11) will yield different bases satisfying (4.13) and (4.14).

*Permanent address: School of Mathematics, University of Minnesota, Minneapolis, Minn. 55455, USA.

[1]P. Winternitz and I. Fris, Sov. J. Nucl. Phys. 1, 636 (1965).
[2]P. Winternitz, I. Lukac, and Y. Smorodinskii, Sov. J. Nucl. Phys. 7, 139 (1968).
[3]E.G. Kalnins, "On the Separation of Variables for the Laplace Equation $\Delta\psi + \chi^2\psi = 0$ in Two and Three Dimensional Minkowski Space," Université de Montréal preprint CRM 319.
[4]N. Vilenkin, *Special Functions and the Theory of Group Representations* (American Mathematical Society, Providence, Rhode Island, 1968).
[5]W. Miller, Jr., "Lie Theory and Separation of Variables. 1. Parabolic Cylinder Coordinates," SIAM J. Math. Anal. (to be published).
[6]A. Erdelyt *et al.*, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vols. I and II.
[7]N. Dunford and J. Schwartz, *Linear Operators*, Part II (Wiley-Interscience, New York, 1963).
[8]I.M. Gel'fand and G.E. Shilov, *Generalized Functions* (Academic, New York, 1964), Vol. I.
[9]W. Miller, Jr., "Lie Theory and Separation of Variables. 2. Parabolic Coordinates," SIAM J. Math. Anal. (to be published).
[10]F.W. Schäfke, *Einfukrung in die Theorie der Speziallen Functions der Mathematischen Physik* (Springer-Verlag, Berlin, 1963).

Hamiltonian description of relativistically interacting two-particle systems*

H. P. Künzle

Department of Mathematics, University of Alberta, Edmonton, Canada T6G 2G1
(Received 26 October 1973)

The state space of an isolated system of two massive point particles (without spin) in instantaneous action-at-a-distance theory is provided with a Poincaré invariant symplectic (Hamiltonian, canonical) structure. This is possible in such a way that the position coordinates commute with respect to the Poisson brackets, and for a nontrivial interaction, if the state space is chosen as the set of initial data not for spacelike, but for lightlike separation of the two particles. The conserved total 4-momentum and polarization vector of the system can be explicitly calculated and used to define a center of momentum frame in which the relative motion is most conveniently described. Under some mild additional restrictions it is shown that the possible invariant Hamiltonian interactions are all obtained by choosing arbitrarily one function of three Lorentz invariant scalars on the state space. Some evidence is presented for the conjecture that for given equations of motion the symplectic structure on the state space is determined up to arbitrary values for the rest masses of the two particles. It also follows that the relative motion of the particles can be described in a plane orthogonal to the 4-momentum and polarization vector. A simple example is given that seems physically interesting, since to first order in the coupling constant the obtained equations of motion agree with those derived from the advanced–retarded Liénard–Wiechert potentials.

1. INTRODUCTION

In spite of the no-interaction theorems[1–4] proved some ten years ago interest in the instantaneous action-at-a-distance theory of the classical dynamics of Poincaré invariant isolated particle systems has not subsided. While Wheeler–Feynman type action-at-a-distance theories[5] may be fairly acceptable as an alternative to the field theoretical description, the Newtonian or Laplacean causality concept inherent in instantaneous action-at-a-distance theories (or predictive relativistic mechanics[6]) is not easily reconciled with the relativistic causality axiom that any event can only influence other events to its own future. Nevertheless, since the description of the particle motion as a dynamical system on a finite-dimensional manifold (the state or phase space) is so much simpler technically than by means of field theories or hereditary action-at-a-distance theories (e.g., based on the Fokker action principle) it seems still worthwhile to continue the study of this approach. In particular, if the state space can be equipped with a natural canonical structure a statistical mechanics and quantum mechanics for relativistic particle systems can be developed in complete analogy to the respective Galilei invariant theories.

Even if not all the premises of this approach should be directly acceptable physically the study of such instantaneously interacting systems could be relevant from a different point of view. It has recently been noted that there is a great analogy between quantum mechanical and classical elementary systems (i.e., irreducible projective representations and transitive symplectic actions of an invariance group[7]) such that at least for these free particle systems the quantum mechanical wave equations can be reconstructed from the "mechanical" particle models according to the methods of Souriau[8] and Kostant.[9] Elementary classical systems have been classified for different invariance groups by Souriau,[10,8] Renouard,[11] Arens,[12] Rawnsley.[13] Such a purely group theoretical approach is probably not (yet) feasible for interacting multiparticle systems. The problem is still to find a sufficiently simple but physically not too unreasonable interacting system that can be described in terms of a finite-dimensional state manifold on which the Poincaré group acts by symplectic transformations. For a sufficiently simple model a larger symplectic symmetry group might then be found and studied along the lines of the recent discussions of the nonrelativistic Kepler problem.[14,15]

The conditions for a system of n point particles that is supposed to be fully described by $3n$ second order ordinary differential equations (and $6n$ initial data) to be invariant under the Poincaré group have been studied by many authors in various formalisms.[16–24] For $n=2$ they have been explicitly solved and resulted in expressions for the acceleration components in terms of several arbitrary functions of three invariant scalars.[23,24]

Casting this dynamical system into a canonical form such that the canonical structure is also Poincaré invariant is considerably less straightforward. The first attempts to introduce a Poisson bracket on the set of dynamical variables that was invariant under the Poincaré group in a seemingly natural way lead to the so-called no-interaction theorems.[1–4] It was, however, soon realized[25–28] that some of the conditions needed to derive these theorems may not be entirely justified. In fact, the crucial condition, namely that the "position coordinates" commute with respect to the Poisson bracket, depends on the way these position coordinates are defined, which as it turns out, is not completely unambiguous. For example, if the most straightforward definition is adopted that is possible in the completely covariant multitime formalism and a vanishing commutator is required for all these covariant position coordinates then a no-interaction theorem results independently of any group invariance requirements, in particular also for Galilei invariant systems.[29,24] On the other hand, it has been argued[30,31] that without this commutation rule the principle of relativistic invariance becomes vacuous.

The situation is not quite so bad, however. At least in the covariant multitime formalism the action of the Poincaré group on the evolution space of the system can be naturally and unambiguously defined and then induces an action on the state space $\mathcal{M}$. It is thus clear what relativistic invariance of the symplectic form $\tilde{\omega}$ on $\mathcal{M}$ means. The remaining problem is simply that there are too many invariant canonical formalisms (or $\tilde{\omega}$'s) that are compatible with a given set of invariant equations of motion. More relaxed conditions have therefore been proposed to make $\tilde{\omega}$ less arbitrary. Hill and Kerner[26] imposed an asymptotic condition, Bel[32] suggested that only the commutators of the differences of position coordinates should be required to vanish.

The aim of this paper is to show that even the original requirement that all position coordinates of the particles commute is compatible with nontrivial (in fact, quite realistic) Poincaré invariant interactions of a two-particles system. But the position coordinates are not defined on a surface of equal world time for the two particles, but are supposed to be measured when the two particles are in lightlike separation. Thus rather than relaxing the commutator condition—which after all works perfectly for Galilei invariant systems (being implied in the one-time Lagrangian formalism[24]) where the surface of constant world time has an intrinsic meaning—we impose it on a surface that seems more natural in a relativistic space–time. This approach has the added advantage that on a thus defined state space the expressions for known equations of motion like the one derived from the Liénard–Wiechert potentials become particularly simple. Moreover, there are several indications that to a given set of equations of motion there exists at most a two-parameter set of compatible symplectic structures satisfying the commutator condition in this sense (the two parameters being the rest masses of the particles). On the other hand our approach introduces a certain asymmetry between the two particles which may, however, be only apparent.

Instead of investigating systematically the possible Poincaré invariant force laws and canonical structures on the state space of instantaneously interacting systems it is possible to arrive at similar models starting from a Fokker-type variational principle[33] which also leads to the conservation laws associated to the generators of the Poincaré group action. In order to obtain instantaneous interaction (and thus a finite-dimensional state space) the Lagrangean must be suitably restricted. No attempt seems to have been made to deduce the full symplectic structure on the state space from this type of variational principle.

In Sec. 2 we review the multitime formalism of Ref. 24 in which relativistic invariance is most conveniently defined. The relation of this description with the equivalent one as a dynamical system on the chosen state manifold Σ is explicitly exhibited. In Sec. 3 an invariant symplectic form is assumed to be given on Σ and the integrals of motion that can be obtained by means of Noether's theorem are calculated. We see in which sense one can conclude that the relative motion of the two particles takes place in a plane under some fairly general assumptions. The effect of the commutator condition is studied in the fourth section. It is shown with some simplifying assumptions that invariant symplectic forms satisfying the commutation rule are in one-to-one correspondence with functions of three invariant scalars. A particularly simple example is given that agrees to first order in the coupling constant with the electromagnetic interaction. This example is probably the best possible Poincaré invariant analog of the classical Kepler problem (if no infinite mass ratio is assumed). An explicit study of the possible orbits will be made elsewhere.

2. A SUITABLE STATE SPACE FOR THE POINCARÉ INVARIANT TWO-PARTICLE SYSTEM

In this section we review the basic assumptions of instantaneous action-at-a-distance theory in the form developed in Ref. 24 without assuming the existence of a canonical structure. A motion of the two-particle system is supposed to be fully described by a pair of worldlines in space–time V that are determined as solutions of a system of second order ordinary differential equations with the initial data consisting of twelve real numbers giving positions and velocities at one time for each particle. Thus the configuration space–time is $\tilde{V}=V\times V$ and the evolution space its tangent bundle $E=T\tilde{V}\approx TV\times TV$, a 16-dimensional manifold. The possible motions of the system are now obtained as the leaves (=maximal connected integral manifolds) of a second order system $\mathcal{E}$ that is spanned by vector fields of the form[34]

$$X=\sum_{k=1}^{2}[a_k(v_k^\alpha\partial_{\alpha_k}+\xi_k^\alpha\partial_{\dot{\alpha}_k})+b_kv_k^\alpha\partial_{\dot{\alpha}_k}] \tag{2.1}$$

for arbitrary functions a_k and b_k on E where the "accelerations" ξ_k^α are given functions on E, subject to the conditions

$$v_k^\rho\partial_{\dot{\rho}_k}\xi_k^\alpha=2\xi_k^\alpha+\alpha_kv_k^\alpha, \tag{2.2}$$

$$v_l^\rho\partial_{\dot{\rho}_l}\xi_k^\alpha=\beta_{kl}v_k^\alpha,\qquad l\neq k, \tag{2.3}$$

and

$$(v_l^\rho\partial_{\rho_l}+\xi_l^\rho\partial_{\dot{\rho}_l})\xi_k^\alpha=\gamma_{kl}v_k^\alpha,\quad l\neq k, \tag{2.4}$$

for some functions α_k, β_{kl}, and γ_{kl} on E.[24,35]

The advantage of this homogeneous multitime formalism is not only that the description of the motion is four-dimensionally covariant, but also that Poincaré (or Galilei) invariance of the equations can be easily and unambiguously formulated. There is a unique action of the Poincaré group on V which leaves its Lorentz structure invariant (we assume now that V is Minkowski space) which induces the product action on $\tilde{V}=V\times V$ and then an action on the tangent bundle $T\tilde{V}=E$. Explicitly, the infinitesimal generators corresponding to the ten basis vectors of the Lie algebra $\mathfrak{g}$ of the Poincaré group G are

$$\mathcal{T}_\alpha=\sum_k\partial_{\alpha_k}\ \text{(translations)} \tag{2.5}$$

and

$$\Omega_{\alpha\beta}=-2\sum_k(x_k^\gamma\eta_{\gamma[\alpha}\partial_{\beta_k]}+v_k^\gamma\eta_{\gamma[\alpha}\partial_{\dot{\beta}_k]})$$

(homogeneous Lorentz transformations).[36]

$$\tag{2.6}$$

Now, if ψ: $G\times E\to E$: $(a,p)\to\psi_a p$ denotes the action of the Poincaré group on E, the second order system $\mathcal{E}$ is invariant under this action if the (tangent map of) ψ_a maps the subspace $\mathcal{E}_p$ of T_pE onto the subspace $\mathcal{E}_{\psi_a p}$ of $T_{\psi_a p}E$, for all $a\in G$, or if $\mathcal{L}_{\underset{\sim}{A}}X\in\mathcal{E}$ for any vector field $X\in\mathcal{E}$ (i.e., $X_p\in\mathcal{E}_p\,\forall p\in\mathcal{E}$) and for all infinitesimal generators $\underset{\sim}{A}$ of the group action.

The general form of a Poincaré invariant second order system on E has been calculated in Ref. 24 (and in a different formalism by Arens[23]). We state here only the results. Instead of using the fiber coordinates (x_k^α, v_k^α) in E, let

$$z^\alpha:=\tfrac{1}{2}(x_1^\alpha+x_2^\alpha), \tag{2.7}$$

$$u_k^\alpha:=v_k^\alpha,\quad u_3^\alpha:=r^\alpha:=x_2^\alpha-x_1^\alpha, \tag{2.8}$$

$$\upsilon_k^2:=u_k\cdot u_k=u_k^A\delta_{AB}u_k^B,\ \upsilon_3=r, \tag{2.9}$$

$$\tau_k:=(u_k^{02}-\upsilon_k^2)^{1/2}, \tag{2.10}$$

$$\alpha_{kl}:=\tau_k^{-2}\tau_l^{-2}(u_k^0u_l^0-u_k\cdot u_l)-1, \tag{2.11}$$

$$\beta_{kl}:=1-(u_k^0u_l^0)^{-1}u_k\cdot u_l, \tag{2.12}$$

where $k\neq l=1,2,3$. We will also write α_1, α_2, α_3 for α_{23}, α_{31}, α_{12}, respectively, and similarly for β_{kl}. Then z^α, v_1^2, v_1^3, α_k, β_k, and τ_k ($k=1,2,3$) from a new coordinate system on an open and dense set of E, where α_k and τ_k are invariant scalars that label the group orbits in E.

Poincaré invariance of the second order system now implies that the quantities ξ_k^α in (2.1) must be of the form

$$\xi_k^\alpha=u_\Sigma^\alpha\tilde{\xi}_k^\Sigma, \tag{2.13}$$

where the $\tilde{\xi}_k^\Sigma$ ($\Sigma=1,2,3,4$) are arbitrary functions of α_m and τ_m ($m=1,2,3$) and u_4^α is defined by

$$u_4^\alpha\equiv w^\alpha:=-\epsilon^\alpha{}_{\lambda\mu\nu}u_1^\lambda u_2^\mu u_3^\nu. \tag{2.14}$$

Equations (2.2) and (2.3) imply that

$$\tilde{\xi}_k^l=\tau_k^2\tau_l^{-1}\xi_k^l(\tau,\alpha_m),\quad \tilde{\xi}_k^3=\tau_k^2\xi_k^3(\tau,\alpha_m),$$
$$\tilde{\xi}_k^4=\tau_k\tau_l^{-1}\xi_k^4(\tau,\alpha_m) \tag{2.15}$$

where from now on $\tau\equiv\tau_3$. The functions $\tilde{\xi}_k^k$ are completely arbitrary and not required since the ξ_k^α need only be known up to a term parallel to v_k^α. The conditions (2.4) cannot be solved but only restated in terms of the ξ_k^Λ.[24] If the latter's dependence on α_m is known, (2.4) yields a highly coupled nonlinear system of six ordinary differential equations in τ for the six quantities ξ_k^Λ ($\Lambda\neq k$). It is easy to see that this system determines the τ dependence uniquely[35] (except in pathological cases), but it seems to be a most inconvenient starting point for trying to find physically interesting nontrivial equations of motion.[23]

Instead, it is preferable for many purposes to switch to a description of the motion as a dynamical system on a suitable state space as has always been customary in nonrelativistic mechanics. However, for relativistic systems this second type of formalism is somewhat ambiguous, since there is no really distinguished choice of such a state space nor is it obvious what the time parameter should be for a multiparticle system. But since we have clearly established the physical meaning of the coordinates and the relativistic invariance in a space—time framework we will now at least be able to identify the specific assumptions that are necessary for a description as dynamical system.

The *state space* can be defined abstractly as the set of all motions according to Souriau,[8] that is the set of all leaves of the foliation defined by $\mathcal{E}$ on the evolution space E which has a natural differentiable structure as a quotient manifold $\mathcal{M}=E/\mathcal{E}$ if some global condition is satisfied.[37] Equivalently $\mathcal{M}$ can be considered as the set of all initial data necessary to determine a motion uniquely, once an equation of motion is given, or, more precisely (since these initial data do not necessarily have to be taken at the same "time") there exists a (possibly only local) diffeomorphism of $\mathcal{M}$ onto a 12-dimensional submanifold Σ of E that intersects each leaf of $\mathcal{E}$ in exactly one point. Such a surface Σ, which can be chosen highly arbitrarily, will be called a *Cauchy surface*. While less fundamental than $\mathcal{M}$ for purposes of studying and classifying all motions of a particular system a Cauchy surface has more structure than the state space. For example, it makes sense to distinguish between position and velocity coordinates on Σ since Σ is still fibered by the restriction of the projection map of the tangent bundle $E=T\tilde{V}$ onto $\tilde{V}$.[38] Moreover, for most practical calculations one must introduce specific coordinates on $\mathcal{M}$ and therefore might as well work on a suitable Σ.

Considered as set of motions or initial data neither $\mathcal{M}$ nor Σ describes the dynamics of the system. But any invariance group G of $\mathcal{E}$ on E induces an action $\tilde{\psi}:G\times\mathcal{M}\to\mathcal{M}$ on $\mathcal{M}$ or similarly on any Σ. In particular, any one-parameter group of time translations on E induces a time flow on $\mathcal{M}$ whose infinitesimal generator $\mathcal{X}$ makes $\mathcal{M}$ into a dynamical system.[39] There is a slight ambiguity in the choice of the time translation subgroup—instead of the generator $\mathcal{T}_0$ one could choose $\tilde{\mathcal{T}}_0=ad_a\mathcal{T}_0=a_0^\alpha\mathcal{T}_\alpha$ ($a\in G$, $a_0^\alpha\eta_{\alpha\beta}a_0^\beta=-1$)—but this would lead to an equivalent dynamical system on $\mathcal{M}$ related to the previous one by the diffeomorphism $\tilde{\psi}_a$ of $\mathcal{M}$ onto itself.

If, however, we want to construct this vector field $\mathcal{X}$ explicitly on a particular Σ where we want to distinguish position and velocity coordinates it does make a difference which Σ we choose. While for a Galilei invariant system the surface

$$\Sigma_0=\{x_k^0=0,\ v_k^0=1\} \tag{2.16}$$

is distinguished in terms of the space—time structure[40] this is certainly not so for a Poincaré invariant system. If the space-time has a Minkowski structure it would seem at least as natural to choose the surface

$$\Sigma=\{z^0=0=\tau,\ \tau_1=\tau_2=1\} \tag{2.17}$$

[in terms of the coordinates introduced in (2.7) to (2.12)], where for simplicity we will also assume that r^0, $v_k^0>0$, although that destroys somewhat the symmetry between the two particles, the second one being required to be on the future null cone of the first.[41] We will from now on work on this particular Cauchy surface and show later that on it nontrivial dynamical systems can be brought into a canonical form in such a way that the position coordinates (as measured on Σ) commute. The so-called no-interaction theorems[1-4] state essentially that this is impossible if the position coordinates are defined on the surface Σ_0.

In the rest of this section we calculate explicitly the form of the generators of the action $\tilde{\psi}$ of G on the Cauchy surface Σ given by (2.17), thus, in particular, the form a Poincaré invariant two-particle interaction takes as a dynamical system on this advanced—retarded state space.

Let the surface Σ be parametrized as follows:

$$\iota_\Sigma : \Sigma \to E : (\bar{x}^A_k ; \bar{v}^A_k) \to (x^0_k = \bar{x}^0_k := (-1)^k \tfrac{1}{2}\bar{r},\ x^A_k = \bar{x}^A_k ;$$

$$v^0_k = \bar{v}^0_k := (1+\bar{v}^2_k)^{1/2},\ v^A_k = \bar{v}^A_k) \tag{2.18}$$

where

$$\bar{r}^A := \bar{x}^A_2 - \bar{x}^A_1, \quad \bar{r}^2 = \delta_{AB}\bar{r}^A\bar{r}^B, \quad \bar{v}^2_k = \delta_{AB}\bar{v}^A_k\bar{v}^B_k. \tag{2.19}$$

We consider Σ as a surface of initial data. Thus there will be one leaf C of $\mathcal{E}$ going through each point $(\bar{x},\bar{v})$ of Σ, namely the 4-manifold generated by all curves $t \to (x^\alpha_k(t), v^\alpha_k(t))$ in E that satisfy

$$\frac{dx^\alpha_k}{dt} = a_k v^\alpha_k, \quad \frac{dv^\alpha_k}{dt} = a_k \xi^\alpha_k + b_k v^\alpha_k,$$
$$x^\alpha_k(0) = \bar{x}^\alpha_k, \quad v^\alpha_k(0) = \bar{v}^\alpha_k, \tag{2.20}$$

where a_k and b_k are again arbitrary functions on E. We wish to calculate the tangent $\pi_{\Sigma *}$ of the projection map $\pi_\Sigma : E \to \Sigma$ that takes a point $p \in E$ into the point of intersection of the leaf C containing p with Σ. Since only the restriction of $\pi_{\Sigma *}$ to points on $\Sigma \subset E$ will be needed, p can be chosen arbitrarily near to Σ so that it will be enough to solve Eq. (2.20) to first order in t which gives

$$x^\alpha_k = \bar{x}^\alpha_k + \bar{a}_k \bar{v}^\alpha_k t + O(t^2), \quad v^\alpha_k = \bar{v}^\alpha_k + (\bar{a}_k \bar{\xi}^\alpha_k + \bar{b}_k \bar{v}^\alpha_k)t + O(t^2), \tag{2.21}$$

where all barred quantities are to be evaluated on Σ. We need the differentials of $\bar{x}^A_k$ and $\bar{v}^A_k$ expressed in terms of dx^α_k and dv^α_k.

Using (2.18) and (2.21) we find *on* Σ

$$dx^0_k = \tfrac{1}{2}(-1)^k d\bar{r} + \bar{v}^0_k d(\bar{a}_k t), \quad dx^A_k = d\bar{x}^A_k + \bar{v}^A_k d(\bar{a}_k t), \tag{2.22}$$

$$dv^0_k = d\bar{v}^0_k + \bar{\xi}^0_k d(\bar{a}_k t) + \bar{v}^0_k d(\bar{b}_k t), \quad dv^A_k = d\bar{v}^A_k + \bar{\xi}^A_k d(\bar{a}_k t) + \bar{v}^A_k d(\bar{b}_k t). \tag{2.23}$$

Elimination of $d(\bar{a}_k t)$ and $d(\bar{b}_k t)$ from the first of (2.22) and (2.23) and substitution into the second gives equations involving only the differentials of x^α_k, v^α_k, $\bar{x}^A_k$, and $\bar{v}^A_k$ that can be solved for the latter two, yielding, again *on* Σ,

$$d\bar{x}^A_k = dx^A_k - (v^0_k)^{-1} v^A_k dx^0_k + \tfrac{1}{2}(-1)^k (v^0_k)^{-1} v^A_k d\bar{r} \tag{2.24}$$

and

$$d\bar{v}^A_k = dv^A_k - (v^0_k)^{-1} v^A_k dv^0_k - (v^0_k)^{-1}\hat{\xi}^A_k dx^0_k + (v^0_k)^{-1} v^A_k d\bar{v}^0_k + \tfrac{1}{2}(-1)^k (v^0_k)^{-1}\hat{\xi}^A_k d\bar{r}, \tag{2.25}$$

where

$$\hat{\xi}^A_k := \xi^A_k - (v^0_k)^{-1} v^A_k \xi^0_k, \tag{2.26}$$

$$d\bar{r} = D^{-1}(r^{-1} r_A dr^A + (v^0_1)^{-1} v^A_1 r_A dx^0_1 - (v^0_2)^{-1} v^A_2 r_A dx^0_2), \tag{2.27}$$

and

$$d\bar{v}^0_k = v^0_k v_A dv^A_k - v^2_k dv^0_k - v_A \hat{\xi}^A_k (dx^0_k - \tfrac{1}{2}(-1)^k d\bar{r}) \tag{2.28}$$

with

$$D := 1 - \tfrac{1}{2}\bar{w}_1 - \tfrac{1}{2}\bar{w}_2, \quad \bar{w}_k := (\bar{r})^{-1}(\bar{v}^0_l)^{-1}\bar{v}^A_l \bar{r}_A. \tag{2.29}$$

The expressions on Σ for

$$\pi_* \partial_{\alpha_k} = \sum_l \left(\frac{\partial \bar{x}^B_l}{\partial x^\alpha_k}\bar{\partial}_{B_l} + \frac{\partial \bar{v}^B_l}{\partial x^\alpha_k}\bar{\partial}_{\dot{B}_l}\right)$$

and

$$\pi_* \partial_{\dot{\alpha}_k} = \sum_l \left(\frac{\partial \bar{x}^B_l}{\partial v^\alpha_k}\bar{\partial}_{B_l} + \frac{\partial \bar{v}^B_l}{\partial v^\alpha_k}\bar{\partial}_{\dot{B}_l}\right)$$

can now be read off (2.24) and (2.25). Together with (2.5) and (2.6) one finds the following expressions for the generators $\bar{T}_\alpha := \pi_{\Sigma *} T_\alpha$ and $\bar{\Omega}_{\alpha\beta} := \pi_{\Sigma *}\Omega_{\alpha\beta}$ of the Poincaré group action on Σ:

$$X := -\bar{T}_0 = D^{-1}\sum_k (\bar{v}^0_k)^{-1}(1-\bar{w}_k)(\bar{v}^A_k \bar{\partial}_{A_k} + \bar{\xi}^A_k \bar{\partial}_{\dot{A}_k}), \tag{2.30}$$

$$\bar{T}_A = \sum_k \bar{\partial}_{A_k}, \tag{2.31}$$

$$\bar{\Omega}_{0A} = \sum_k (\bar{x}^0_k \bar{\partial}_{A_k} + \bar{v}^0_k \bar{\partial}_{\dot{A}_k}) - \tfrac{1}{2}(x^A_1 + x^A_2)\bar{T}_0, \tag{2.32}$$

$$\bar{\Omega}_{AB} = -2\sum_k (\bar{x}_{k[A}\bar{\partial}_{B_k]} + \bar{v}_{k[A}\bar{\partial}_{\dot{B}_k]}), \tag{2.33}$$

where now $\bar{\xi}^A_k = \hat{\xi}^A_k + v^A_k v_B \hat{\xi}^B_k = \xi^A_k - v_{k\alpha}\xi^\alpha_k v^A_k = \xi^A_k - \tilde{\xi}^k v^A_k$ [see (2.13)]. In particular, the "time" flow in terms of the coordinates $\bar{x}^A_k$ and $\bar{v}^A_k$ on Σ follows from the equations

$$\dot{\bar{x}}^A_k \equiv \frac{d\bar{x}^A_k}{ds} = X(\bar{x}^A_k) = D^{-1}(1-\bar{w}_k)(\bar{v}^0_k)^{-1}\bar{v}^A_k,$$
$$\dot{\bar{v}}^A_k \equiv \frac{d\bar{v}^A_k}{ds} = X(\bar{x}^A_k) = D^{-1}(1-\bar{w}_k)(\bar{v}^0_k)^{-1}\bar{\xi}^A_k. \tag{2.34}$$

Sometimes it will be more convenient to use on Σ the coordinates $\bar{r}^A$ and $\bar{z}^A := \tfrac{1}{2}(\bar{x}^A_1 + \bar{x}^A_2)$ instead of the $\bar{x}^A_k$. Also, the bars will be dropped when there is no danger of confusion.

Finally, we recall the form of the invariant coordinates on Σ as introduced in Ref. 24. Note that $\tau_1 = \tau_2 = 1$ and $\tau_3 \equiv \tau = 0 = z^0$ on Σ while the coordinates α_k in E tend to infinity as the point approaches Σ. But we can use $\bar{z}^A$, $\bar{v}^2_1$, $\bar{v}^3_1$, $\bar{v}^3_2$, β_1, β_2, β_3, and

$$\lambda := (1+\alpha_3)^{1/2} (\geq 1) \quad \text{and} \quad \rho_k := \tau(1+\alpha_k)^{1/2} \quad (\geq 0) \tag{2.35}$$

instead of α_1, α_2, α_3. More explicitly, it follows from (2.35) that on Σ

$$\lambda = \bar{v}_1^0 \bar{v}_2^0 - v_1 \cdot v_2, \quad \rho_k = \bar{r} \bar{v}_l^0 - r \cdot v_l. \tag{2.36}$$

(ρ_k is just the luminosity distance between the two particles, as measured by the kth particle.) From (2.15) we know that the ξ_k^A $(\Lambda \neq k)$ are given functions of λ, ρ_1, ρ_2 only and from (2.13) we find explicitly[24]

$$\bar{\xi}_k^A = \xi_k^l(\bar{v}_l^A - \lambda \bar{v}_k^A) + \xi_k^3(\bar{r}^A - \rho_l \bar{v}_k^A) + \xi_k^4 \bar{w}^A. \tag{2.37}$$

Sometimes it will be more convenient to replace the coordinates β_1, β_2, β_3 by $\bar{r}$ and $\bar{v}_k^0$. Then we find from (2.29)

$$\bar{w}_k = 1 - (\bar{r}\bar{v}_l^0)^{-1}\rho_k,$$

$$D = \tfrac{1}{2}(\bar{r}\bar{v}_1^0\bar{v}_2^0)^{-1}(\bar{v}_1^0\rho_1 + \bar{v}_2^0\rho_2) =: \tfrac{1}{2}(\bar{r}\bar{v}_1^0\bar{v}_2^0)^{-1} N$$

and for the "time" derivatives $[\dot{f} \equiv df/ds \equiv X(f)]$ of the various coordinates[42]

$$\dot{x}_k^A = 2N^{-1}\rho_k v_k^A, \tag{2.38}$$

$$\dot{v}_k^A = 2N^{-1}\rho_k \bar{\xi}_k^A, \tag{2.39}$$

$$\dot{r}^A = 2N^{-1}(\rho_2 v_2^A - \rho_1 v_1^A), \tag{2.40}$$

$$\dot{r} = 2N^{-1}(\rho_2 v_2^0 - \rho_1 v_1^0), \tag{2.41}$$

$$\dot{v}_k^0 = 2N^{-1}\rho_k[(v_l^0 - \lambda v_k^0)\xi_k^l + (r - v_k^0\rho_l)\xi_k^3 + w^0\xi_k^4], \tag{2.42}$$

$$\dot{\lambda} = -2N^{-1}[\rho_1(\alpha\,\xi_1^2 + \hat{\rho}_2\,\xi_1^3) + \rho_2(\alpha\,\xi_2^1 + \hat{\rho}_1\,\xi_2^3)], \tag{2.43}$$

$$\dot{\rho}_k = 2N^{-1}[(-1)^k\hat{\rho}_k - \hat{\rho}_k\rho_l\,\xi_l^k - \rho_k^2\rho_l\,\xi_l^3], \tag{2.44}$$

where

$$\hat{\rho}_k := \lambda\rho_k - \rho_l, \quad \alpha := \lambda^2 - 1. \tag{2.45}$$

If the functions ξ_k^A $(\lambda, \rho_1, \rho_2)$ are explicitly given, this system of ordinary differential equations can be integrated in principle at least, but in all but the simplest cases it is technically not very easy.[43] The symplectic structure on Σ to be introduced in the next section will be helpful as it enables us to find first integrals via Noether's theorem.[44]

For the free particle system, where all $\xi_k^A = 0$, Eqs. (2.39), (2.42), and (2.43) imply that v_k^α and λ are constants of the motion and it is then not very hard to find explicitly

$$x_k^A(s) = x_k^A(0) + (2\lambda + v_1^0(v_2^0)^{-1} + (v_1^0)^{-1}v_2^0)^{-1}[2(\lambda(v_k^0)^{-1} + (v_l^0)^{-1})s + (-1)^k(v_k^0)^{-1}(N - N(0))]v_k^A \tag{2.46}$$

where

$$N = (4\lambda^2 s^2 + 2bs + c)^{-1/2},$$

$$b = (N\dot{N})(0) = 2(v_2^0\hat{\rho}_2 - v_1^0\hat{\rho}_1)(0),$$

$$c = N^2(0) = (v_1^0\rho_1 + v_2^0\rho_2)(0).$$

Thus the orbits in the "configuration space" $\mathbb{R}^3 \times \mathbb{R}^3$, parametrized by (x_1^A, x_2^A), are straight lines according to (46), as expected, but the dependence on the "time" s is complicated except for very special initial conditions.

3. CANONICAL STRUCTURE AND INTEGRALS OF MOTION

The general idea and technique of equipping E with a presymplectic structure ω such that

$$\mathcal{E} = \ker\omega \tag{3.1}$$

($\ker_p\omega := \{X \in T_pE \mid X \lrcorner \omega = 0\}$) has been discussed at length in Ref. 24. It is equivalent to giving a symplectic form $\tilde{\omega}$ on the state space $\mathcal{M}$ that is related to ω by

$$\omega = \pi^*\tilde{\omega} \tag{3.2}$$

where $\pi: E \to \mathcal{M}$ is the canonical projection map. Moreover, if $\iota_\Sigma: \Sigma \to E$ is a Cauchy surface then $\pi \circ \iota_\Sigma$ is a (possibly local) diffeomorphism and the induced 2-form $\bar{\omega}$ on Σ satisfies $\bar{\omega} := \iota_\Sigma^*\omega = \iota_\Sigma^*\pi^*\tilde{\omega} = (\pi \circ \iota_\Sigma)^*\tilde{\omega}$, thus $(\mathcal{M}, \tilde{\omega})$ and $(\Sigma, \bar{\omega})$ are (locally) symplectomorphic. A presymplectic ω satisfying (3.1) always exists locally[24] and need not be very interesting or useful except if some additional requirements are met. Here we are only interested in canonical structures that are also invariant under the Poincaré group, i.e., satisfy $\psi_{a*}\omega = \omega \; \forall a \in G$ or $\mathcal{L}_{\underset{\sim}{A}}\omega = 0 \; \forall A \in \mathfrak{g}$. Then also $\mathcal{L}_{\pi_*\underset{\sim}{A}}\tilde{\omega} = 0$ and $\mathcal{L}_{\pi_{\Sigma*}\underset{\sim}{A}}\bar{\omega} = 0$. Conversely, if $\mathcal{M}$ or Σ carry a symplectic form $\tilde{\omega}$ or $\bar{\omega}$, respectively, invariant under the induced group action, then ω on E, defined by (3.2) is invariant.

For the construction of a Poincaré invariant 2-form ω on E one can proceed by a technique similar to the one used for the second order system.[24] But the solutions of $d\omega = 0$ are rather cumbersome to determine, in general. Fortunately, for Poincaré invariant systems physically interesting second order systems (among them the one for the noninteracting case) exist for which $\omega = -d\theta$ with a 1-form θ that is itself invariant under the group action.[45] We will only investigate this case here. It has the added advantage that Noether's theorem then takes a particularly simple form, namely the functions

$$\mu_A := \underset{\sim}{A} \lrcorner \theta, \quad A \in \mathfrak{g}, \tag{3.3}$$

on E are integrals of motion, i.e., constant on each leaf and hence of the form $\mu_A = \pi^*\tilde{\mu}_A$ with $\tilde{\mu}_A: \mathcal{M} \to \mathbb{R}$. [The map $\mu: E \to \mathfrak{g}^*$, the dual vector space of the Lie algebra $\mathfrak{g}$, such that $\mu(x)(A) = \mu_A(x)$ is called a *moment* according to Souriau.[8]]

The most general invariant 1-form on E can be given in the form[24]

$$\theta = \sum_k (\underset{k}{P}_\alpha dx_k^\alpha + \underset{k}{Q}_\alpha dv_k^\alpha) \tag{3.4}$$

with

$$\underset{k}{P}_\alpha = \underset{k}{\tilde{P}}_\Lambda U_\alpha^\Lambda, \quad \underset{k}{Q}_\alpha = \underset{k}{\tilde{Q}}_\Lambda U_\alpha^\Lambda, \tag{3.5}$$

where (U_α^Λ) is the inverse matrix to (u_Λ^α) and $\underset{k}{\tilde{P}}_\Lambda$ and $\underset{k}{\tilde{Q}}_\Lambda$ are arbitrary functions of τ_m, α_m $(m = 1, 2, 3)$. Since we are only interested in the restriction of θ to Σ we can regard $\underset{k}{\tilde{P}}_\Lambda$ and $\underset{k}{\tilde{Q}}_\Lambda$ as functions of λ, ρ_1 and ρ_2 on Σ.

We calculate the integrals of motion according to (3.3) in the covariant space—time formalism (working, however, at a point on Σ where $\tau_1=\tau_2=1$, $\tau=0$). The 4-vector of total *momentum* becomes

$$P_\alpha := \mathcal{T}_\alpha \lrcorner\, \theta = \sum_k \underset{k}{P}_\alpha = \Big(\sum_k \underset{k}{\tilde{P}}_\Lambda\Big) U^\Lambda_\alpha =: \tilde{P}_\Lambda U^\Lambda_\alpha. \tag{3.6}$$

The matrix U^Λ_α is rather awkward to calculate directly. Instead observe that if

$$K_\alpha = \tilde{K}_\Lambda U^\Lambda_\alpha \tag{3.7}$$

and

$$K^\alpha := \eta^{\alpha\beta} K_\beta = \tilde{K}^\Lambda u^\alpha_\Lambda \tag{3.8}$$

then $\tilde{K}^\Lambda = U^\Lambda_\alpha K^\alpha = U^\Lambda_\alpha \eta^{\alpha\beta} K_\beta = U^\Lambda_\alpha \eta^{\alpha\beta} U^\Sigma_\beta \tilde{K}_\Sigma$. The symmetric matrix $\eta^{\Lambda\Sigma} = U^\Lambda_\alpha \eta^{\alpha\beta} U^\Sigma_\beta$ is easily seen to be the inverse of

$$\eta_{\Lambda\Sigma} := u^\alpha_\Lambda \eta_{\alpha\beta} u^\beta_\Sigma = \begin{pmatrix} -1 & -\lambda & -\rho_2 & 0 \\ -\lambda & -1 & -\rho_1 & 0 \\ -\rho_2 & -\rho_1 & 0 & 0 \\ 0 & 0 & 0 & M \end{pmatrix} \tag{3.9}$$

on Σ [in view of (2.14) and (2.36)] where

$$M := w^\alpha w_\alpha = -\rho_1^2 - \rho_2^2 + 2\lambda\rho_1\rho_2. \tag{3.10}$$

Thus

$$\eta^{\Lambda\Sigma} = M^{-1} \begin{pmatrix} \rho_1^2 & -\rho_1\rho_2 & -\hat{\rho}_1 & 0 \\ -\rho_1\rho_2 & \rho_2^2 & -\hat{\rho}_2 & 0 \\ -\hat{\rho}_1 & -\hat{\rho}_2 & \alpha & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.11}$$

The contravariant components of the 4-momentum $\tilde{P}^\Lambda = \eta^{\Lambda\Sigma}\tilde{P}_\Sigma$ turn out to be often simpler as is seen in the case of no interaction where

$$\theta = \sum_k \theta_k \quad \text{and} \quad \theta_k = m_k \tau_k^{-1} \eta_{\alpha\beta} v^\beta_k dx^\alpha_k \tag{3.12}$$

($m_k=$const) is the only invariant 1-form that defines a second order equation on $E_k = TV_k$ for the one-particle system[24]. Then, simply $P^\alpha = \sum_k m_k v^\alpha_k$ or $\tilde{P}^\Lambda = \sum_k m_k \delta^\Lambda_k$ on Σ.

The conserved quantities corresponding to the invariance under the homogeneous Lorentz group become

$$M_{\alpha\beta} := \Omega_{\alpha\beta} \lrcorner\, \theta = -2\sum_k (x^\gamma_k \eta_{\gamma[\alpha} \underset{k}{P}_{\beta]} + v^\gamma_k \eta_{\gamma[\alpha} \underset{k}{Q}_{\beta]}). \tag{3.13}$$

They depend on the origin of the coordinate system x^α_k. We will only evaluate the more useful *polarization 4-vector*

$$W^\alpha := M^{*\alpha\beta} P_\beta = \tfrac{1}{2}\epsilon^{\alpha\beta\lambda\mu} P_\beta M_{\lambda\mu} \tag{3.14}$$

(which is manifestly orthogonal to P^α). If again

$$W^\alpha = u^\alpha_\Lambda \tilde{W}^\Lambda, \tag{3.15}$$

then (3.13) and (3.14) give

$$\tilde{W}^k = M^{-1}\{\rho_k(\underset{k}{\tilde{P}}_3 \underset{l}{\tilde{P}}_4 - \underset{l}{\tilde{P}}_3 \underset{k}{\tilde{P}}_4) + (-1)^k[\tilde{P}_l(\rho_l \underset{k}{\tilde{Q}}_4 + \rho_k \underset{l}{\tilde{Q}}_4)$$
$$- \tilde{P}_4(\rho_l \underset{k}{\tilde{Q}}_l + \rho_k \underset{l}{\tilde{Q}}_l) - \tilde{P}_3(\lambda \underset{k}{\tilde{Q}}_4 + \underset{l}{\tilde{Q}}_4) + \tilde{P}_4(\lambda \underset{k}{\tilde{Q}}_3 + \underset{l}{\tilde{Q}}_3)]\}, \tag{3.16}$$

$$\tilde{W}^3 = M^{-1}\{-\rho_1(\underset{1}{\tilde{P}}_1 \underset{2}{\tilde{P}}_4 - \underset{2}{\tilde{P}}_1 \underset{1}{\tilde{P}}_4) + \rho_2(\underset{1}{\tilde{P}}_2 \underset{2}{\tilde{P}}_4 - \underset{2}{\tilde{P}}_2 \underset{1}{\tilde{P}}_4) + \tilde{P}_2 \underset{1}{\tilde{Q}}_4 - \tilde{P}_4 \underset{1}{\tilde{Q}}_2$$
$$- \tilde{P}_1 \underset{2}{\tilde{Q}}_4 + \tilde{P}_4 \underset{2}{\tilde{Q}}_1 - \lambda(\tilde{P}_1 \underset{1}{\tilde{Q}}_4 - \tilde{P}_4 \underset{1}{\tilde{Q}}_1 - \tilde{P}_2 \underset{2}{\tilde{Q}}_4 + \tilde{P}_4 \underset{2}{\tilde{Q}}_2)\}, \tag{3.17}$$

and

$$\tilde{W}^4 = M^{-1}\{\rho_1(\underset{1}{\tilde{P}}_1 \underset{2}{\tilde{P}}_3 - \underset{2}{\tilde{P}}_1 \underset{1}{\tilde{P}}_3) - \rho_2(\underset{1}{\tilde{P}}_2 \underset{2}{\tilde{P}}_3 - \underset{2}{\tilde{P}}_2 \underset{1}{\tilde{P}}_3) - \tilde{P}_2 \underset{1}{\tilde{Q}}_3 + \tilde{P}_3 \underset{1}{\tilde{Q}}_2$$
$$+ \tilde{P}_1 \underset{2}{\tilde{Q}}_3 - \tilde{P}_3 \underset{2}{\tilde{Q}}_1 + \lambda(\tilde{P}_1 \underset{1}{\tilde{Q}}_3 - \tilde{P}_3 \underset{1}{\tilde{Q}}_1 - \tilde{P}_2 \underset{2}{\tilde{Q}}_3 + \tilde{P}_3 \underset{2}{\tilde{Q}}_2)$$
$$- \rho_2(\tilde{P}_2 \underset{1}{\tilde{Q}}_2 - \tilde{P}_2 \underset{1}{\tilde{Q}}_1) - \rho_1(\tilde{P}_1 \underset{2}{\tilde{Q}}_2 - \tilde{P}_2 \underset{2}{\tilde{Q}}_1)\}. \tag{3.18}$$

Note that if $\underset{k}{\tilde{P}}_4 = 0 = \underset{k}{\tilde{Q}}_4$ then only $\tilde{W}^4$ does not vanish. In particular, for the noninteracting system

$$W^\alpha = -m_1 m_2 w^\alpha. \tag{3.19}$$

The two 4-vectors P^α and W^α are indeed integrals of motion (corresponding to the conservation of energy, momentum and angular momentum) in the strict sense of commuting with the Hamiltonian in the description on Σ.[46] For, if the Poisson bracket on Σ is defined by[47]

$$\{\bar{f}, \bar{g}\} := \mathcal{L}_{\bar{X}_f} \bar{g} \quad \text{with} \quad \bar{X}_f \lrcorner\, \bar{\omega} := d\bar{f},$$

then $\bar{X}_{\mu_A} = \pi_{\Sigma *} \underset{\sim}{A} =: \bar{\underset{\sim}{A}}$ and $\{\bar{\mu}_A, \bar{\mu}_B\} = \bar{\mu}_{[A,B]}$ for all $A, B \in \mathfrak{g}$. But, in particular,

$$dP^0 = -dP_0 = -\bar{\mathcal{T}}_0 \lrcorner\, \bar{\omega} = X \lrcorner\, \omega. \tag{3.20}$$

Thus, we identify P^0 with the Hamiltonian H. Then $\dot{P}^A = \dot{P}_A = \mathcal{L}_X P_A = \{H, P_A\} = -\{P_0, P_A\} = -\mu_{[\mathcal{T}_0, \mathcal{T}_A]} = 0$, since all translations commute and

$$\dot{W}^\alpha = \tfrac{1}{2}\epsilon^{\alpha\beta\lambda\mu} P_\beta \dot{M}_{\lambda\mu} = -\tfrac{1}{2}\epsilon^{\alpha\beta\lambda\mu} P_\beta \{P_0, M_{\lambda\mu}\}$$
$$= \tfrac{1}{2}\epsilon^{\alpha\beta\lambda\mu} P_\beta \mu_{[\mathcal{T}_0, \Omega_{\lambda\mu}]} = -\epsilon^{\alpha\beta\lambda\mu} P_\beta \eta_{0\lambda} P_\mu = 0.$$

For the description of the orbits in Σ it is therefore convenient to introduce "center of momentum" coordinates. Since it is the relative motion of the two particles that is of main interest we can choose coordinates with respect to a (fixed) Lorentz frame $\{\underset{p}{e}^\alpha\}$ such that

$$P^\alpha = m \underset{0}{e}^\alpha \quad (m > 0) \tag{3.21}$$

and

$$W^\alpha = m s e^\alpha_3 \quad (s \geq 0) \tag{3.22}$$

where m can be regarded as the *total mass-energy* of the system and s as the magnitude of the *total spin* (the case $W^\alpha=0$ must be treated separately). Then, if

$$\underset{k}{\tilde P}_4=0=\underset{k}{\tilde Q}_A, \tag{3.23}$$

$W^\alpha=\tilde W^4 w^\alpha = ms\underset{3}{e}^\alpha$ and since r^α and v_k^α are orthogonal to w^α the *relative position vector* r^A *is confined to the plane* spanned by $\underset{1}{e}$ and $\underset{2}{e}$, *orthogonal to the spin vector* se_3. It is not obvious that only plane motions are possible for Poincaré invariant interactions that do not satisfy (3.23). But we will see in Sec. 4 that (3.23) is implied by the proposed commutator rule.

For the Hamiltonian description of the *relative motion* one would expect to need a six-dimensional state space Σ_r, parametrized by the coordinates r^A and, for example,

$$v^A := 2N^{-1}(\rho_2 v_2^A - \rho_1 v_1^A) \quad (N=\rho_1 v_1^0+\rho_2 v_2^0) \tag{3.24}$$

[then, by (2.40), $dr^A/ds = v^A$]. Such a space can be constructed by solving equations (3.24) and

$$\tilde P^1 v_1^A + \tilde P^2 v_2^A + \tilde P^3 r^A + \tilde P^4 w^A = P^A = \text{const} \tag{3.25}$$

for v_k^A, obtaining $v_k^A = f_k^A(r^B, v^B)$, say. Then let

$$\iota_r : \Sigma_r \to \Sigma : (r^A, v^A) \to (x_k^A = \tfrac12(-1)^k r^A, v_k^A = f_k^A) \tag{3.26}$$

define the submanifold $\Sigma_r \subset \Sigma$. There is an induced vector field X_r on Σ_r, namely the one giving the derivatives $\dot r^A$ and $\dot v^A$, as well as a symplectic structure $\omega_r = \iota_r^* \bar\omega$ and an induced action of the rotation group (not of the whole homogeneous Lorentz group, however). This construction is just an explicit form (using submanifolds as local representatives of quotient manifolds) of the general reduction of dynamical groups introduced by Marsden and Weinstein.[48] (Dividing out the spacelike translations reduces the 12-dimensional symplectic manifold Σ to the six-dimensional "reduced phase space" Σ_r.) Their theorem 1 shows that the construction is independent of some special choices made, like putting $z^A=0$ in (3.26), though the symplectic structure ω_r may depend on the total momentum P^A.

Geometric studies of (Σ_r, ω_r)—with time translations and $SO(3)$ still acting on it—can be made and compared with those of the Kepler manifold.[15] First, however, we need to find a simple and still physically not too unreasonable Poincaré invariant Hamiltonian interaction.

4. THE CONDITION OF COMMUTING POSITION COORDINATES

In Sec. 2 the general form of a Poincaré invariant vectorfield X giving the time flow on Σ was constructed in the form [see Eqs. (2.30), (2.38), and (2.39)]

$$X = 2N^{-1}\sum_k \rho_k (v_k^A \bar\partial_{A_k} + \bar\xi_k^A \bar\partial_{\dot A_k})$$

with $\bar\xi_k^A$ given by (2.37) and involving the six arbitrary functions $\xi_k^\Lambda(\lambda,\rho_1,\rho_2)(\Lambda\neq k)$. It is not quite obvious that a θ of the form (3.4), (3.5) can be found such that (3.20),

$$dH \equiv dP^0 = X \lrcorner \bar\omega, \tag{4.1}$$

holds with $\bar\omega = -d\bar\theta = -d\iota_\Sigma^*\theta$ for arbitrary ξ_k^Λ. If further conditions are imposed on $\bar\omega$ this is even less so. However, Eq. (4.1) for given $\bar\omega$ and H or given $\underset{k}{\tilde P}^\Lambda$ and $\underset{k}{\tilde Q}^\Lambda$ determines X. We could thus choose these quantities on Σ and calculate $\bar\omega$ and X from (4.1). But in practice this procedure is extremely tedious and it seems that the space–time approach developed in Ref. 24 is a little more efficient, mainly because the differentiation with respect to the Poincaré invariant scalars is somewhat easier in the four-dimensional formalism.

The aim is to find nontrivial Hamiltonian systems on Σ, i.e., quantities $\bar\omega$, X, and H that are Poincaré invariant and satisfy the condition that the Poisson brackets of the position coordinates commute, namely that

$$\{\bar x_k^A, \bar x_l^B\} = 0$$

(for all k, $l=1,2$). This is equivalent to saying that $\bar\omega$ contains no term in $d\bar v_k^A \wedge d\bar v_l^B$ or that it satisfies $\bar\omega \wedge d\bar x_1^1 \wedge d\bar x_1^2 \wedge d\bar x_1^3 \wedge d\bar x_2^1 \wedge d\bar x_2^2 \wedge d\bar x_2^3 = 0$.[49] The simplest way to achieve this for $\bar\omega = -d\bar\theta$ would be to assume that $\bar\theta = \sum_k \underset{k}{P}_A d\bar x_k^A$. We will eventually consider only this case. However, in order to calculate $\omega = -d\theta$ in E rather than in Σ only, it cannot be assumed that the $\underset{k}{Q}_\alpha$ in (3.4) vanish identically because that would imply that $\{x_k^A, x_l^B\}=0$ not only on Σ but also for any other Cauchy surface, which is only possible for noninteracting systems.[24,29]

Invariant ω in E

Adopt therefore again for θ the general form (3.4)–(3.5). Then $\omega = -d\theta$ is of the form

$$\omega = \sum_{k=1}^{2} \{\tfrac12 \underset{k}{\omega}_{\alpha\beta} dx_k^\alpha \wedge dx_k^\beta + \underset{k}{\sigma}_{\alpha\beta} dx_k^\alpha \wedge dv_k^\beta + \tfrac12 \underset{k}{\pi}_{\alpha\beta} dv_k^\alpha \wedge dv_k^\beta + \underset{k}{\lambda}_{\alpha\beta} dx_l^\alpha \wedge dv_k^\beta\} + \kappa_{\alpha\beta} dx_1^\alpha \wedge dx_2^\beta + \rho_{\alpha\beta} dv_1^\alpha \wedge dv_2^\beta \tag{4.2}$$

with

$$\underset{k}{\omega}_{\alpha\beta} = -2\partial_{[\alpha_k} \underset{k}{P}_{\beta]}, \tag{4.3}$$

$$\underset{k}{\sigma}_{\alpha\beta} = \partial_{\dot\beta_k} \underset{k}{P}_\alpha - \partial_{\alpha_k} \underset{k}{Q}_\beta, \tag{4.4}$$

$$\underset{k}{\pi}_{\alpha\beta} = -2\partial_{[\dot\alpha_k} \underset{k}{Q}_{\beta]}, \tag{4.5}$$

$$\underset{k}{\lambda}_{\alpha\beta} = \partial_{\dot\beta_k} \underset{l}{P}_\alpha - \partial_{\alpha_l} \underset{k}{Q}_\beta, \tag{4.6}$$

$$\kappa_{\alpha\beta} = -\partial_{\alpha_1} \underset{2}{P}_\beta + \partial_{\beta_2} \underset{1}{P}_\alpha, \tag{4.7}$$

$$\rho_{\alpha\beta} = -\partial_{\dot\alpha_1} \underset{2}{Q}_\beta + \partial_{\dot\beta_2} \underset{1}{Q}_\alpha. \tag{4.8}$$

Contracting Eqs. (4.3) to (4.8) with $u_\Lambda^\alpha u_\Pi^\beta$ gives the invariant components of ω directly in terms of the invariant components $\underset{k}{\tilde P}_\Lambda$ and $\underset{k}{\tilde Q}_\Lambda$:

$$\underset{k}{\tilde\omega}_{\Lambda\Pi} = -2(-1)^k[\tilde\partial^3{}_{[\Lambda} \underset{k}{\tilde P}_{\Pi]} + \partial^3{}_{[\Lambda}(\underset{k}{\tilde P}_{\Pi]} - \delta^4_{\rho]} \underset{k}{\tilde P}_4)], \tag{4.9}$$

$$\underset{k}{\tilde\sigma}_{\Lambda\Pi} = \tilde\partial^k_\Pi \underset{k}{\tilde P}_\Lambda - \delta^k_\Lambda \underset{k}{\tilde P}_\Pi - \delta^4_\Lambda \delta^k_\Pi \underset{k}{P}_4 - \delta^4_\Lambda \delta^4_\Pi M^{k\Sigma} \underset{k}{\tilde P}_\Sigma$$
$$+ (-1)^k[-\tilde\partial^3_\Lambda \underset{k}{\tilde Q}_\Pi + \underset{k}{\tilde Q}_\Lambda \delta^3_\Pi + \delta^3_\Lambda \delta^4_\Pi \underset{k}{\tilde Q}_4 + \delta^4_\Lambda \delta^4_\Pi M^{3\Sigma} \underset{k}{\tilde Q}_\Sigma], \tag{4.10}$$

$$\underset{k}{\tilde\Pi}_\Lambda = -2\tilde\partial^k_{[\Lambda} \underset{k}{\tilde Q}_{\Pi]} - 2\partial^k_{[\Lambda}(\underset{k}{\tilde Q}_{\Pi]} - \delta^4_{\Pi]} \underset{k}{Q}_4), \tag{4.11}$$

$$\underset{k}{\tilde\lambda}_{\Lambda\Pi} = \tilde\partial^k_\Pi \underset{l}{\tilde P}_\Lambda - \delta^k_\Lambda \underset{l}{\tilde P}_\Pi - \delta^4_\Lambda \delta^k_\Pi \underset{l}{\tilde P}_4 - \delta^4_\Lambda \delta^4_\Pi M^{k\Sigma} \underset{l}{\tilde P}_\Sigma$$
$$+ (-1)^k[\tilde\partial^3_\Lambda \underset{k}{\tilde Q}_\Pi - \underset{k}{\tilde Q}_\Lambda \delta^3_\Pi - \delta^3_\Lambda \delta^4_\Pi \underset{k}{\tilde Q}_4 - \delta^4_\Lambda \delta^4_\Pi M^{3\Sigma} \underset{k}{\tilde Q}_\Sigma], \tag{4.12}$$

$$\tilde\kappa_{\Lambda\Pi} = \tilde\partial^3_\Lambda \underset{2}{\tilde P}_\Pi + \tilde\partial^3_\Pi \underset{1}{\tilde P}_\Lambda - \delta^3_\Lambda(\underset{1}{\tilde P}_\Pi + \delta^4_\Pi \underset{2}{\tilde P}_4) - \delta^3_\Pi(\delta^4_\Lambda \underset{1}{\tilde P}_4 + \underset{2}{\tilde P}_\Lambda)$$
$$- \delta^4_\Lambda \delta^4_\Pi M^{3\Sigma}(\underset{1}{\tilde P}_\Sigma + \underset{2}{\tilde P}_\Sigma), \tag{4.13}$$

$$\tilde\rho_{\Lambda\Pi} = -\tilde\partial^1_\Lambda \underset{2}{\tilde Q}_\Pi + \tilde\partial^2_\Pi \underset{1}{\tilde Q}_\Lambda - \delta^2_\Lambda \underset{1}{\tilde Q}_\Pi + \underset{2}{\tilde Q}_\Lambda \delta^1_\Pi + \delta^1_\Lambda \delta^4_\Pi \underset{2}{\tilde Q}_4 - \delta^4_\Lambda \delta^2_\Pi \underset{1}{\tilde Q}_4$$
$$+ \delta^4_\Lambda \delta^4_\Pi (M^{1\Sigma} \underset{2}{\tilde Q}_\Sigma - M^{2\Sigma} \underset{1}{\tilde Q}_\Sigma) \tag{4.14}$$

where[24]

$$\tilde\partial^k_\Lambda := u^\rho_\Lambda \partial_{u^\rho_k}, \quad M^{11} = -\alpha_1 \tau_2^2 \tau_3^2, \cdots, M^{12} = M^{21}$$
$$= \tau_1 \tau_2 \tau_3^2 \Lambda_3, \cdots, M^{4m} = M^{m4} = 0$$

with $\lambda_m = (1+\alpha_m)^{1/2}$, $\Lambda_1 := \lambda_2\lambda_3 - \lambda_1, \cdots$ $(m = 1, 2, 3)$.

Algebraic conditions on ω

What the condition (3.1) means for an invariant ω has also been derived in Ref. 24. First, from $d\omega = 0$ it can be deduced that

$$\underset{k}{\tilde\omega}_{\Gamma\Delta} = \tau_\Gamma \tau_\Delta \underset{k}{\omega}_{\Gamma\Delta}, \quad \tilde\kappa_{\Gamma\Delta} = \tau_\Gamma \tau_\Delta \kappa_{\Gamma\Delta},$$

$$\underset{k}{\tilde\sigma}_{\Gamma\Delta} = \tau_k^{-1} \tau_\Gamma \tau_\Delta \underset{k}{\sigma}_{\Gamma\Delta}, \quad \underset{k}{\tilde\lambda}_{\Gamma\Delta} = \tau_k^{-1} \tau_\Gamma \tau_\Delta \underset{k}{\lambda}_{\Gamma\Delta},$$

$$\underset{k}{\tilde\pi}_{\Gamma\Delta} = \tau_k^{-2} \tau_\Gamma \tau_\Delta \underset{k}{\pi}_{\Gamma\Delta}, \quad \tilde\rho_{\Gamma\Delta} = \tau_1^{-1}\tau_2^{-1} \tau_\Gamma \tau_\Delta \rho_{\Gamma\Delta},$$

where Γ, $\Delta = 1, 2, 3, 4$ and $\tau_\Gamma = \delta^1_\Gamma \tau_1 + \delta^2_\Gamma \tau_2 + \delta^4_\Gamma \tau_1 \tau_2$ and all the $K_{\Gamma\Delta}$ are functions of $\tau \equiv \tau_3$ and α_m $(m = 1, 2, 3)$ only. Then (3.1) gives

$$\underset{k}{\sigma}_{\Lambda k} = \underset{k}{\lambda}_{\Lambda k} = \underset{k}{\pi}_{\Lambda k} = \rho_{1\Lambda} = \rho_{\Lambda 2} = 0 \tag{4.15}$$

and

$$\underset{k}{\sigma}_{\Lambda\Sigma} \xi^\Sigma_k = \underset{k}{\omega}_{k\Lambda}, \tag{4.16}$$

$$\underset{1}{\lambda}_{\Lambda\Sigma} \xi^\Sigma_1 = \kappa_{1\Lambda}, \quad \underset{2}{\lambda}_{\Lambda\Sigma} \xi^\Sigma_2 = -\kappa_{\Lambda 2}, \tag{4.17}$$

$$\underset{k}{\pi}_{\Lambda\Sigma} \xi^\Sigma_k = \underset{k}{\sigma}_{k\Lambda}, \tag{4.18}$$

$$\rho_{\Lambda\Sigma} \xi^\Sigma_2 = \underset{1}{\lambda}_{2\Lambda}, \quad \xi^\Sigma_1 \rho_{\Sigma\Lambda} = -\underset{2}{\lambda}_{1\Lambda}. \tag{4.19}$$

Substituting the expressions (4.9) to (4.14) into (4.15) and noting that applied to invariant scalars[24]

$$\tilde\partial^m_m = \tau_m \partial_{\tau_m}, \quad \tilde\partial^m_4 = 0, \tag{4.20}$$

$$\tilde\partial^m_n = \lambda_{mn} \tau_n \partial_{\tau_m} - 2\tau_n \tau_m^{-1} \partial^m_n \tag{4.21}$$

where

$$\partial^m_n = \Lambda_m \lambda_n \partial_{\alpha_n} + \alpha_{mn} \lambda_{mn} \partial_{\alpha_{mn}} \tag{4.22}$$

and $m \neq n = 1, 2, 3$, we find that

$$\underset{k}{\tilde Q}_\Lambda = \tilde\partial^k_\Lambda \tilde Q + \tau_k^{-1} \tau_\Lambda \underset{k}{Q}_\Lambda, \quad \underset{k}{Q}_k = 0, \tag{4.23}$$

and

$$\underset{k}{\tilde P}_\Lambda = (-1)^k \tilde\partial^3_\Lambda \tilde Q + \tau_\Lambda \underset{k}{P}_\Lambda \tag{4.24}$$

where $\tilde Q$ is an arbitrary function of τ_m and α_m and $\underset{k}{P}_\Lambda$ and $\underset{k}{Q}_\Lambda$ do not depend on τ_1 and τ_2.

Restriction to Σ

Next we wish to calculate the components of ω explicitly on Σ in order to impose there the conditions that

$$\underset{k}{\bar\pi}_{AB} = 0 = \bar\rho_{AB}. \tag{4.25}$$

Note that if near Σ again the coordinates λ, ρ_k are used instead of α_m [cf. (2.35)] then any sufficiently regular invariant scalar ϕ can be expanded in the form[24] (assume $\tau_1 = \tau_2 = 1$)

$$\phi = \bar\phi + \tau\overset{1}{\phi} + \tfrac{1}{2}\tau^2\overset{2}{\phi} + O(\tau^3) \tag{4.26}$$

with $\overset{m}{\phi} = \overset{m}{\phi}(\lambda, \rho_1, \rho_2)$. Moreover, since $\partial_{\alpha_k} = \frac{1}{2}\tau^2\rho_k^{-1}\partial_{\rho_k}$, $\partial_{\alpha_3} = \frac{1}{2}\lambda^{-1}\partial_\lambda$, and $\tilde\partial^3_3 = \tau_3\partial_{\tau_3} = \tau\partial_\tau + \rho_1\partial_{\rho_1} + \rho_2\partial_{\rho_2}$ we find from (4.21) and (4.22) that

$$\partial^k_l = \tfrac{1}{2}(\alpha\,\partial_\lambda + \hat\rho_l \partial_{\rho_l}),$$

$$\tau\partial^k_3 = \tfrac{1}{2}[\hat\rho_l \partial_\lambda + (\rho_l^2 - \tau^2)\partial_{\rho_l}],$$

$$\tau\tilde\partial^3_k = \tau_k[\rho_l \partial_\tau + \tau(\lambda\partial_{\rho_k} + \partial_{\rho_l})].$$

If these operators are applied to the expansion (4.26) of an invariant scalar, they give (again restricted to $\tau_1 = \tau_2 = 1$)

$$\tilde\partial^3_3\phi = \partial_0\bar\phi + \tau(\overset{1}{\phi} + \partial_0\overset{1}{\phi}) + O(\tau^2), \tag{4.27}$$

$$\partial^k_l\phi = \tfrac{1}{2}\partial_l\bar\phi + \tfrac{1}{2}\tau\partial_l\overset{1}{\phi} + O(\tau^2), \tag{4.28}$$

$$\tau\partial_3^k = \tfrac{1}{2}\eth_l\bar{\phi} + \tfrac{1}{2}\tau\eth_l\overset{1}{\phi} + O(\tau^2) \tag{4.29}$$

and

$$\tau\tilde{\partial}_k^3\phi = \rho_l\overset{1}{\phi} + \tau(\partial_{kl}\bar{\phi} + \rho_l\overset{2}{\phi}) + O(\tau^2) \tag{4.30}$$

with

$$\partial_0 := \rho_1\partial_{\rho_1} + \rho_2\partial_{\rho_2}, \tag{4.31}$$

$$\partial_k := \alpha\,\partial_\lambda + \hat{\rho}_k\partial_{\rho_k}, \tag{4.32}$$

$$\eth_k := \hat{\rho}_k\partial_\lambda + \rho_k^2\partial_{\rho_k}, \tag{4.33}$$

$$\partial_{kl} := \lambda\partial_{\rho_k} + \partial_{\rho_l}. \tag{4.34}$$

Substituting Eqs. (4.23) and (4.24) into (4.9) to (4.14) one can express the components of ω restricted to Σ in terms of the $\underset{k}{\overset{m}{P}}{}_\Lambda$ and $\underset{k}{\overset{m}{Q}}{}_\Lambda$ given on Σ as functions of λ, ρ_1, and ρ_2. Actually they involve only the restricted functions $\underset{k}{\bar{P}}{}_\Lambda$ and $\underset{k}{\bar{Q}}{}_\Lambda$ and the combinations

$$\underset{k}{\overset{\times}{P}}{}_\Lambda := \tau^{-1}\underset{k}{\overset{1}{P}}{}_\Lambda + \underset{k}{\overset{2}{P}}{}_\Lambda \quad\text{and}\quad \underset{k}{\overset{\times}{Q}}{}_\Lambda := \tau^{-1}\underset{k}{\overset{1}{Q}}{}_\Lambda + \underset{k}{\overset{2}{Q}}{}_\Lambda \tag{4.35}$$

of τ-derivatives which must be assumed to be regular on Σ. The explicit expressions are listed in an appendix.

Effect of the commutator condition

Note that in view of (4.15) the conditions (4.25) simply amount to

$$\underset{k}{\pi}{}_{\Gamma\Delta} = 0 = \rho_{\Gamma\Delta} \quad \text{on } \Sigma. \tag{4.36}$$

A simple way of assuring this is by requiring that

$$\underset{k}{\bar{Q}}{}_\Lambda = 0, \tag{4.37}$$

which will be assumed from now on. We also let $\tilde{Q}$ in (4.23) and (4.24) be zero since this quantity does not affect ω.

Then, substitution of the expressions for the invariant components of ω in the Appendix into Eqs. (4.16) to (4.19) enables us to eliminate the $\underset{k}{\overset{\times}{P}}{}_\Lambda$ and $\underset{k}{\overset{\times}{Q}}{}_\Lambda$ and to find the following relations between the $\underset{k}{\bar{P}}{}_\Lambda$ and the ξ_k^Λ:

$$\underset{k}{\bar{P}}{}_4 = 0, \quad \xi_k^4 = 0, \tag{4.38}$$

$$\underset{k}{\bar{P}}{}_3 = \rho_k^{-1}(-\eth_l Z + \rho_l Z), \tag{4.39}$$

$$\bar{P}_k = \rho_l^{-1}(-\partial_k Z + \lambda Z), \tag{4.40}$$

and

$$D\bar{P}_k = (\lambda\rho_k\xi_k^l + \rho_k\rho_l\xi_k^3)\bar{P}_k - \rho_k\xi_k^l\bar{P}_l - \rho_k\xi_k^3\bar{P}_3, \tag{4.41}$$

$$D\bar{P}_3 = \rho_1\bar{P}_1 - \rho_2\bar{P}_2, \tag{4.42}$$

$$D\bar{\mathcal{P}}_3 + 2\partial_0 Z = \rho_1\bar{P}_1 + \rho_2\bar{P}_2, \tag{4.43}$$

where

$$P_\Lambda := \underset{1}{P}{}_\Lambda + \underset{2}{P}{}_\Lambda, \quad \mathcal{P}_\Lambda := \underset{2}{P}{}_\Lambda - \underset{1}{P}{}_\Lambda, \tag{4.44}$$

$$Z := \rho_1\underset{1}{\bar{P}}{}_1 + \rho_2\underset{2}{\bar{P}}{}_2 = \tfrac{1}{2}(\rho_1\bar{P}_1 + \rho_2\bar{P}_2 - \rho_1\bar{\mathcal{P}}_1 + \rho_2\bar{\mathcal{P}}_2), \tag{4.45}$$

and

$$D := \partial_1 - \partial_2 + \rho_1\xi_1^2\partial_2 + \rho_1\xi_1^3\eth_2 + \rho_2\xi_2^1\partial_1 + \rho_2\xi_2^3\eth_1. \tag{4.46}$$

The $\underset{k}{\overset{\times}{P}}{}_\Lambda$ and $\underset{k}{\overset{\times}{Q}}{}_\Lambda$ (which are needed to compute $\bar{\omega}$) then become

$$\overset{\times}{P}_k = -\rho_l^{-1}\partial_{kl}\bar{P}_k + (-1)^k\rho_l^{-1}[(\partial_l\bar{P}_k - \lambda\bar{P}_k + \bar{P}_l)\xi_k^l + (\eth_l\bar{P}_k - \rho_l\bar{P}_k + \bar{P}_3)\xi_k^3], \tag{4.47}$$

$$\overset{\times}{P}_3 = \rho_2^{-1}[-\partial_{12}\bar{P}_3 + \bar{P}_1 - \xi_1^2\partial_2\bar{P}_3 - \xi_1^3\eth_2\bar{P}_3], \tag{4.48}$$

$$\underset{k}{\overset{\times}{P}}{}_4 = 0, \tag{4.49}$$

$$\rho_2\overset{\times}{\mathcal{P}}_2 - \rho_1\overset{\times}{\mathcal{P}}_1 = -\partial_{21}\bar{P}_1 - \partial_{12}\bar{P}_2 + \partial_{21}\bar{\mathcal{P}}_1 - \partial_{12}\bar{\mathcal{P}}_2 - (\rho_1\overset{\times}{P}_1 + \rho_2\overset{\times}{P}_2), \tag{4.50}$$

$$\overset{\times}{\mathcal{P}}_3 = \rho_2^{-1}[-\partial_0\bar{P}_1 + \partial_0\bar{\mathcal{P}}_1 + \bar{\mathcal{P}}_1 - \partial_{12}\bar{\mathcal{P}}_3 - (\xi_1^2\partial_2 + \xi_1^3\eth_2)\bar{\mathcal{P}}_3], \tag{4.51}$$

$$\underset{k}{\overset{\times}{Q}}{}_l = \tfrac{1}{2}\rho_l^{-1}[(-1)^k(-\partial_l\bar{P}_k + \lambda\bar{P}_k - \bar{P}_l) - \partial_l\bar{\mathcal{P}}_k + \lambda\bar{\mathcal{P}}_k - \bar{\mathcal{P}}_l], \tag{4.52}$$

$$\underset{k}{\overset{\times}{Q}}{}_3 = \tfrac{1}{2}\rho_l^{-1}[(-1)^k(-\eth_l\bar{P}_k + \rho_l\bar{P}_k - \bar{P}_3) - \eth_l\bar{\mathcal{P}}_k + \rho_l\bar{\mathcal{P}}_k - \bar{\mathcal{P}}_3], \tag{4.53}$$

$$\underset{k}{\overset{\times}{Q}}{}_4 = 0. \tag{4.54}$$

We can interprete these equations as follows: The function $Z = Z(\lambda, \rho_1, \rho_2)$ can be arbitrarily chosen. It then determines $\underset{k}{\bar{P}}{}_3$ and $\bar{P}_k$ by (4.39) and (4.40), respectively, and then also $\rho_1\bar{\mathcal{P}}_1 - \rho_2\bar{\mathcal{P}}_2$ by (4.45). Next, Eqs. (4.41) to (4.43) determine (at a generic point) the four unknown components ξ_k^l and ξ_k^3 uniquely. The combination $\rho_1\bar{\mathcal{P}}_1 + \rho_2\bar{\mathcal{P}}_2$ remains undetermined, if it is chosen arbitrarily, all $\underset{k}{\overset{\times}{P}}{}_\Lambda$ and $\underset{k}{\overset{\times}{Q}}{}_\Lambda$ are obtained from (4.47) to (4.54), except $\rho_1\overset{\times}{\mathcal{P}}_1 + \rho_2\overset{\times}{\mathcal{P}}_2$. However, it turns out that these quantities do not enter into the expressions for ω, nor into those for the interesting integrals of motion, namely P^α and W^α. Thus, under the simplifying assumptions $\omega = -d\theta$ in E with θ invariant and $\underset{k}{Q}{}^\alpha = 0$ on Σ just *one invariant function on Σ determines the equations of motion as well as the canonical structure of the state space* uniquely. The function Z is not the Hamiltonian, but it seems likely that prescribing the Hamiltonian would have a similar effect, except that it is somewhat less convenient since $H \equiv P^0 = u^0_\Lambda \eta^{\Lambda\Sigma} \bar{P}_\Sigma$ [cf. (3.8) and (3.11)] does not depend on the invariant scalars λ, ρ_1, ρ_2 only.

That one function determines equations of motion and the symplectic structure of the state space comes as little surprise to anyone familiar with Lagrangian mechanics. Here, however, while it is not immediately clear whether the whole formalism can be brought into Lagrangian form,[50] we can also solve the converse problem,[51] namely *for given equations of motion there seems to exist at most* a two-parameter set of such functions Z and hence *a two-parameter set of invariant compatible symplectic structures on the state space.* We will not attempt to prove this here in general, but simply look at the case of no interaction, where $\xi^\Lambda_k = 0$. Then the operator D given by (4.46) has a very simple form and a straightforward though tedious analysis of the integrability conditions of the system (4.39) to (4.43) of first order partial differential equations shows that necessarily

$$Z = -(m_1\rho_1 + m_2\rho_2) \tag{4.55}$$

for two arbitrary constants m_1 and m_2—which is precisely the expression one gets for the noninteracting system starting with a θ of the form (3.12). It thus follows that in this case at least the conditions we have imposed on ω determine it for given equations of motion just as much as desired, namely up to arbitrary values for the masses of the particles. It seems likely that for nonzero ξ^Λ_k the symplectic form is similarly determined, a conjecture that is confirmed by an analysis of Eqs. (4.39) to (4.43) for weak interactions (small ξ^Λ_k).

It has turned out that the invariant 1-form θ on Σ is not fully determined by the equations of motion even if condition (4.37) is imposed, although ω and the integrals of motion P^α and W^α are. It would seem not unlikely that the more natural conditions (4.36) alone determine ω as much as desired while θ then contains even more arbitrary terms. The calculations for this case unfortunately become considerably more involved; however, it can be seen easily [directly from (4.36) and the expressions in the Appendix] that $\underset{k}{\bar{Q}}_4 = 0 = \underset{k}{\bar{P}}_4$ and hence also $\xi^4_k = 0$. According to Sec. 3 *the relative motion therefore takes place in a plane orthogonal to the spin vector.*

A simple example of a Hamiltonian interacting system

It is already clear that there exist interacting systems that are Hamiltonian according to our definition because Eqs. (4.39) to (4.43) will lead to $\xi^\Lambda_k \neq 0$ for any choice of $Z = Z(\lambda, \rho_1, \rho_2)$ other than (4.55). The problem is to find a simple interacting system that seems physically not too unreasonable and can be studied more explicitly. In Ref. 24 the electromagnetic interaction obtained from the Liénard–Wiechert fields by neglecting the acceleration terms[52] was cast into the present formalism. It corresponds to

$$m_k \xi^l_k = (-1)^k g \rho_k^{-3} \rho_l, \quad m_k \xi^3_k = -(-1)^k g \lambda \rho_k^{-3}, \quad \xi^4_k = 0 \tag{4.56}$$

where $g = e_1 e_2$ is the coupling constant (if e_k is the charge of the kth particle). We can now try to solve (4.39) to (4.43) for this choice of the ξ^Λ_k's, but this does not seem to lead to very simple expressions for Z and the $\underset{k}{\bar{P}}_\Lambda$. To first order in g, however, Z takes the simple form

$$Z = -(m_1\rho_1 + m_2\rho_2 - g\lambda). \tag{4.57}$$

Adopting this as a new starting point one can now calculate ξ^Λ_k and ω, exactly obtaining

$$m_k \xi^l_k = (-1)^k g \rho_k^{-3} \rho_l \Delta^{-1}[1 + g m_l^{-1} \rho_k \rho_l^{-2}], \tag{4.58}$$

$$m_k \xi^3_k = -(-1)^k g \rho_k^{-3} \Delta^{-1}[\lambda + g m_l^{-1} \rho_k \rho_l^{-3} \hat{\rho}_l - g^2 m_k^{-1} m_l^{-1} \rho_l^{-2}], \tag{4.59}$$

$$\xi^4_k = 0, \tag{4.60}$$

where $\Delta = 1 - g^2 m_1^{-1} m_2^{-1} \rho_1^{-1} \rho_2^{-1}$. Thus the equations of motion agree to first order with the Liénard–Wiechert equations only and are somewhat complicated.[53] On the other hand—what is more important—the symplectic structure and the integrals of motion are very simple, for example,

$$P^\alpha = (m_1 - g\rho_2^{-1}) v_1^\alpha + (m_2 - g\rho_1^{-1}) v_2^\alpha + g\lambda \rho_1^{-1}\rho_2^{-1} r^\alpha \tag{4.61}$$

and

$$W^\alpha = -(m_1 m_2 - g^2 \rho_1^{-1} \rho_2^{-1}) w^\alpha. \tag{4.62}$$

Recalling that ρ_1 and ρ_2 stand for the luminosity distance between the two particles, one notes a certain similarity of the Hamiltonian $H \equiv P^0$ with the one for the nonrelativistic Kepler problem. A further investigation of this system may therefore be quite interesting.

5. CONCLUSION

It was shown that a consistent Hamiltonian description of a great variety of nontrivial relativistic instantaneous two-particle interactions is possible. The dynamical system on the state space Σ can be discussed with the full apparatus of analytical mechanics, in particular, all the integrals of motion arising from the Poincaré invariance can be explicitly given in terms of no more than three variables and can be used to find the orbits.

If the commutator condition is imposed on Σ there seems to correspond only a two-parameter set of symplectic structures to a given equation of motion, namely the two particle rest masses emerge at this point as the parameters that characterize the possible symplectic structures to a given equation of motion—just as mass and spin do for an elementary (free) one-particle system. Moreover, the result that the relative motion can be described as taking place in a fixed plane (whose orientation is determined by total momentum and spin vector) depends on the commutator condition.

A drawback of the proposed formulation seems to be that it does not treat the two particles symmetrically, a difficulty that becomes more serious for systems with more than two particles. However, the imbedding of the asymmetrically defined state space into the fully symmetric evolution manifold is known explicitly. It is therefore possible to check in the space–time formalism whether the introduced asymmetry is in fact physical or only apparent. In view of the remarkable symmetry of all the equations obtained the latter seems just as likely.

Apart from a clarification of this point the method used to find the general invariant Hamiltonian systems is clearly capable of improvement. A somewhat more efficient formalism should also help to determine the status of the conjectures arrived at in Sec. 4.

APPENDIX

The nonvanishing invariant components of ω on Σ in terms of $\underset{k}{P}_\Lambda$ and $\underset{k}{Q}_\Lambda$ (see Sec. 4) become if

$$\underset{k}{\overset{\times}{P}}_\Lambda := \tau^{-1}\underset{k}{\overset{1}{P}}_\Lambda + \underset{k}{\overset{2}{P}}_\Lambda \quad\text{and}\quad \underset{k}{\overset{\times}{Q}}_\Lambda := \tau^{-1}\underset{k}{\overset{1}{Q}}_\Lambda + \underset{k}{\overset{2}{Q}}_\Lambda$$

are assumed regular on Σ (for $k \neq l = 1, 2$)

$$\underset{k}{\omega}_{kl} = (-1)^k[-\partial_{kl}\underset{k}{\bar{P}}_l + \partial_{lk}\underset{k}{\bar{P}}_k - \rho_l\underset{k}{\overset{\times}{P}}_l + \rho_k\underset{k}{\overset{\times}{P}}_k],$$

$$\underset{k}{\omega}_{k3} = (-1)^k[+\partial_0\underset{k}{\bar{P}}_k + \underset{k}{\bar{P}}_k - \partial_{kl}\underset{k}{\bar{P}}_3 - \rho_l\underset{k}{\overset{\times}{P}}_3],$$

$$\underset{k}{\omega}_{l3} = (-1)^k[+\partial_0\underset{k}{\bar{P}}_l + \underset{k}{\bar{P}}_l - \partial_{lk}\underset{k}{\bar{P}}_3 - \rho_k\underset{k}{P}_3],$$

$$\underset{k}{\omega}_{k4} = -(-1)^k[\partial_{kl}\underset{k}{\bar{P}}_4 + \rho_l\underset{k}{\overset{\times}{P}}_4],$$

$$\underset{k}{\omega}_{l4} = -(-1)^k[\partial_{lk}\underset{k}{\bar{P}}_4 + \rho_k\underset{k}{\overset{\times}{P}}_4],$$

$$\underset{k}{\omega}_{34} = -(-1)^k\partial_0\underset{k}{\bar{P}}_4,$$

$$\kappa_{11} = \partial_{12}(\underset{1}{\bar{P}}_1 + \underset{2}{\bar{P}}_1) + \rho_2(\underset{1}{\overset{\times}{P}}_1 + \underset{2}{\overset{\times}{P}}_1),$$

$$\kappa_{12} = \partial_{21}\underset{1}{\bar{P}}_1 + \partial_{12}\underset{2}{\bar{P}}_2 + \rho_1\underset{1}{\overset{\times}{P}}_1 + \rho_2\underset{2}{\overset{\times}{P}}_2,$$

$$\kappa_{13} = \partial_0\underset{1}{\bar{P}}_1 + \partial_{12}\underset{2}{\bar{P}}_3 - \underset{2}{\bar{P}}_1 + \rho_2\underset{2}{\overset{\times}{P}}_3,$$

$$\kappa_{14} = \kappa_{41} = \partial_{12}\underset{2}{\bar{P}}_4 + \rho_2\underset{2}{\overset{\times}{P}}_4,$$

$$\kappa_{21} = \partial_{12}\underset{1}{\bar{P}}_2 + \partial_{21}\underset{2}{\bar{P}}_1 + \rho_2\underset{1}{\overset{\times}{P}}_2 + \rho_1\underset{2}{\overset{\times}{P}}_1,$$

$$\kappa_{22} = \partial_{21}(\underset{1}{\bar{P}}_2 + \underset{2}{\bar{P}}_2) + \rho_1(\underset{1}{\overset{\times}{P}}_2 + \underset{2}{\overset{\times}{P}}_2),$$

$$\kappa_{23} = \partial_0\underset{1}{\bar{P}}_2 - \underset{2}{\bar{P}}_2 + \partial_{21}\underset{2}{\bar{P}}_3 + \rho_1\underset{2}{\overset{\times}{P}}_3,$$

$$\kappa_{24} = \kappa_{42} = \partial_{21}\underset{1}{\bar{P}}_4 + \rho_1\underset{1}{\overset{\times}{P}}_4,$$

$$\kappa_{31} = \partial_0\underset{2}{P}_1 - \underset{1}{P}_1 + \partial_{12}\underset{1}{P}_3 + \rho_2\underset{1}{P}_3,$$

$$\kappa_{32} = \partial_0\underset{2}{\bar{P}}_2 + \partial_{21}\underset{1}{\bar{P}}_3 - \underset{1}{\bar{P}}_2 + \rho_1\underset{1}{\overset{\times}{P}}_3,$$

$$\kappa_{33} = \partial_0(\underset{1}{\bar{P}}_3 + \underset{2}{\bar{P}}_3) - (\underset{1}{\bar{P}}_3 + \underset{2}{\bar{P}}_3),$$

$$\kappa_{34} = \partial_0\underset{2}{\bar{P}}_4 - (\underset{1}{\bar{P}}_4 + \underset{2}{\bar{P}}_4),$$

$$\kappa_{43} = \partial_0\underset{1}{\bar{P}}_4 - (\underset{1}{\bar{P}}_4 + \underset{2}{\bar{P}}_4),$$

$$\kappa_{44} = -\hat{\rho}_1(\underset{1}{\bar{P}}_1 + \underset{2}{\bar{P}}_1) - \hat{\rho}_2(\underset{1}{\bar{P}}_2 + \underset{2}{\bar{P}}_2) + \alpha(\underset{1}{\bar{P}}_3 + \underset{2}{\bar{P}}_3),$$

$$\underset{k}{\sigma}_{kl} = -\partial_l\underset{k}{\bar{P}}_k + \lambda\underset{k}{\bar{P}}_k - \underset{k}{\bar{P}}_l - (-1)^k[\partial_{kl}\underset{k}{\bar{Q}}_l + \rho_l\underset{k}{\overset{\times}{Q}}_l],$$

$$\underset{k}{\sigma}_{k3} = -\tilde{\partial}_l\underset{k}{\bar{P}}_k + \rho_l\underset{k}{\bar{P}}_k - \underset{k}{\bar{P}}_3 - (-1)^k[\partial_{kl}\underset{k}{\bar{Q}}_3 + \rho_l\underset{k}{\overset{\times}{Q}}_3],$$

$$\underset{k}{\sigma}_{k4} = -\underset{k}{\bar{P}}_4 - (-1)^k[\partial_{kl}\underset{k}{\bar{Q}}_4 + \rho_l\underset{k}{\overset{\times}{Q}}_4],$$

$$\underset{k}{\sigma}_{ll} = -\partial_l\underset{k}{\bar{P}}_l - (-1)^k[\partial_{lk}\underset{k}{\bar{Q}}_l + \rho_k\underset{k}{\overset{\times}{Q}}_l],$$

$$\underset{k}{\sigma}_{l3} = -\partial_l\underset{k}{\bar{P}}_l - (-1)^k[\partial_{lk}\underset{k}{\bar{Q}}_3 - \underset{k}{\bar{Q}}_l + \rho_k\underset{k}{\overset{\times}{Q}}_3],$$

$$\underset{k}{\sigma}_{l4} = -(-1)^k[\partial_{lk}\underset{k}{\bar{Q}}_4 + \rho_k\underset{k}{\overset{\times}{Q}}_4],$$

$$\underset{k}{\sigma}_{3l} = -\partial_l\underset{k}{\bar{P}}_3 - (-1)^k\partial_0\underset{k}{\bar{Q}}_l,$$

$$\underset{k}{\sigma}_{33} = -\tilde{\partial}_l\underset{k}{\bar{P}}_3 - (-1)^k[\partial_0\underset{k}{\bar{Q}}_3 - \underset{k}{\bar{Q}}_3],$$

$$\underset{k}{\sigma}_{34} = (-1)^k[-\partial_0\underset{k}{\bar{Q}}_4 + \underset{k}{\bar{Q}}_4],$$

$$\underset{k}{\sigma}_{4l} = -\partial_l\underset{k}{\bar{P}}_4 + \lambda\underset{k}{\bar{P}}_4,$$

$$\underset{k}{\sigma}_{43} = -\tilde{\partial}_l\underset{k}{\bar{P}}_4 + \rho_l\underset{k}{\bar{P}}_4 + (-1)^k\underset{k}{\bar{Q}}_4,$$

$$\underset{k}{\sigma}_{44} = +\rho_k^2\underset{k}{\bar{P}}_k - \rho_k\rho_l\underset{k}{\bar{P}}_l - \hat{\rho}_k\underset{k}{\bar{P}}_3 + (-1)^k[-\alpha\underset{k}{\bar{Q}}_3 + \hat{\rho}_l\underset{k}{\bar{Q}}_l],$$

$$\underset{k}{\lambda}_{kl} = -\partial_l\underset{l}{\bar{P}}_k + \lambda\underset{l}{\bar{P}}_k - \underset{l}{\bar{P}}_l + (-1)^k[\partial_{kl}\underset{k}{\bar{Q}}_l + \rho_l\underset{k}{\overset{\times}{Q}}_l],$$

$$\underset{k}{\lambda}_{k3} = -\tilde{\partial}_l\underset{l}{\bar{P}}_k + \rho_l\underset{l}{\bar{P}}_k - \bar{P}_3 + (-1)^k[\partial_{kl}\underset{k}{\bar{Q}}_3 + \rho_l\underset{k}{\overset{\times}{Q}}_3],$$

$$\underset{k}{\lambda}_{k4} = -\underset{l}{\bar{P}}_4 + (-1)^k[\partial_{kl}\underset{k}{\bar{Q}}_4 + \rho_l\underset{k}{\overset{\times}{Q}}_4],$$

$$\underset{k}{\lambda}_{ll} = -\partial_l\underset{l}{\bar{P}}_l + (-1)^k[\partial_{lk}\underset{k}{\bar{Q}}_l + \rho_k\underset{k}{\overset{\times}{Q}}_l],$$

$$\underset{k}{\lambda}_{l3} = -\tilde{\partial}_l\underset{l}{\bar{P}}_l + (-1)^k[\partial_{lk}\underset{k}{\bar{Q}}_3 - \underset{k}{\bar{Q}}_l + \rho_k\underset{k}{\overset{\times}{Q}}_3],$$

$$\underset{k}{\lambda}_{33} = -\tilde{\partial}_l\underset{l}{\bar{P}}_3 + (-1)^k[\partial_0\underset{k}{\bar{Q}}_3 - \underset{k}{\bar{Q}}_3],$$

$$\underset{k}{\lambda}_{34} = +(-1)^k[\partial_0\underset{k}{\bar{Q}}_4 - \underset{k}{\bar{Q}}_4],$$

$$\underset{k}{\lambda}_{4l} = -\partial_l\underset{l}{\bar{P}}_4 + \lambda\underset{l}{\bar{P}}_4,$$

$$\underset{k}{\lambda}_{43} = -\partial_l\underset{l}{\bar{P}}_4 + \rho_l\underset{l}{\bar{P}}_4 - (-1)^k\underset{k}{\bar{Q}}_4,$$

$$\underset{k}{\lambda}_{44} = +\rho_k^2\underset{l}{\bar{P}}_k - \rho_k\rho_l\underset{l}{\bar{P}}_l - \hat{\rho}_k\bar{P}_3 + (-1)^k[\alpha\underset{k}{\bar{Q}}_3 - \hat{\rho}_l\underset{k}{\bar{Q}}_l],$$

$$\underset{k}{\pi}_{l3} = -\tilde{\partial}_l\underset{k}{\bar{Q}}_l + \partial_l\underset{k}{\bar{Q}}_3 - \rho_l\underset{k}{\bar{Q}}_l + \lambda\underset{k}{\bar{Q}}_3,$$

$$\underset{k}{\pi}_{l4} = \partial_l\underset{k}{\bar{Q}}_4,$$

$$\underset{k}{\pi}_{34} = \tilde{\partial}_l\underset{k}{\bar{Q}}_4,$$

$$\rho_{21} = -\partial_1 \underset{1}{\bar{Q}_2} + \partial_2 \underset{2}{\bar{Q}_1} - \lambda(\underset{2}{\bar{Q}_1} - \underset{1}{\bar{Q}_2}),$$

$$\rho_{23} = -\not{\partial}_1 \underset{1}{\bar{Q}_2} + \partial_2 \underset{2}{\bar{Q}_3} - \underset{1}{\bar{Q}_3} + \rho_1 \underset{1}{\bar{Q}_2},$$

$$\rho_{24} = \partial_2 \underset{2}{\bar{Q}_4} - \lambda \underset{2}{\bar{Q}_4} - \underset{1}{\bar{Q}_4},$$

$$\rho_{31} = \not{\partial}_2 \underset{2}{\bar{Q}_1} - \partial_1 \underset{1}{\bar{Q}_3} + \underset{2}{\bar{Q}_3} - \rho_2 \underset{2}{\bar{Q}_1},$$

$$\rho_{33} = -\not{\partial}_1 \underset{1}{\bar{Q}_3} + \not{\partial}_2 \underset{2}{\bar{Q}_3},$$

$$\rho_{34} = \partial_2 \underset{2}{\bar{Q}_4} - \rho_2 \underset{2}{\bar{Q}_4},$$

$$\rho_{41} = -\partial_1 \underset{1}{\bar{Q}_4} + \lambda \underset{1}{\bar{Q}_4} + \underset{2}{\bar{Q}_4},$$

$$\rho_{43} = -\partial_1 \underset{1}{\bar{Q}_4} + \rho_1 \underset{1}{\bar{Q}_4},$$

$$\rho_{44} = \rho_2^2 \underset{1}{\bar{Q}_2} - \rho_1^2 \underset{2}{\bar{Q}_1} - \hat{\rho}_2 \underset{1}{\bar{Q}_3} + \hat{\rho}_1 \underset{2}{\bar{Q}_3}.$$

*Supported in part by the National Research Council of Canada.

[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]J.A. Wheeler and R.P. Feynman, Rev. Mod. Phys. 17, 157 (1945).
[6]L. Bel, Ann. Inst. Henri Poincaré 14 A, 189 (1971).
[7]For an excellent summary see D.J. Simms, Z. Naturforsch. 28a, 538 (1973).
[8]J.M. Souriau, *Structure des systèmes dynamiques* (Dunod, Paris, 1970).
[9]B. Kostant, in *Lectures in Modern Analysis and Applications III*, edited by C.T. Taam (Springer, Berlin, 1970).
[10]J.M. Souriau, C.R. Acad. Sci. Paris 263 B, 1191 (1966) and C.R. Acad. Sci. Paris 265 B, 165 (1967).
[11]P. Renouard, "Variétés symplectiques et quantification," Thèse, Orsay (1969).
[12]R. Arens, Commun. Math. Phys. 21, 139 (1971) and J. Math. Phys. 12, 2415 (1971).
[13]J. Rawnsley, "Some applications of quantisation," Thesis, Oxford (1972).
[14]R. Hermann, *Lie Groups for Physicists* (Benjamin, New York, 1966).
[15]J.M. Souriau, "Sur la variété de Kepler," J. Elhadad, "Sur l'interpretation en géometrie symplectique des états quantiques de l'atome d'hydrogène," and D.J. Simms, "Geometric Quantization of Energy Levels in the Kepler Problem," in *Symposia Mathematica* (Academic, New York) (to be published); F.A.E. Pirani, "Once more the Kepler problem," preprint.
[16]P. Havas and J. Plebanski, Bull. Amer. Phys. Soc. 5, 433 (1961).
[17]D.G. Currie, Phys. Rev. 142, 817 (1966).
[18]R.N. Hill, J. Math. Phys. 8, 201 (1967).
[19]D.G. Currie and T.F. Jordan, Phys. Rev. 167, 1178 (1968).
[20]R. Arens and D.G. Babbitt, Pacific J. Math. 28, 243 (1969).
[21]P. Droz-Vincent, Physica Scripta 2, 129 (1970).
[22]L. Bel, Ann. Inst. Henri Poincaré 12 A, 307 (1970).
[23]R. Arens, Arch. Rat. Mech. 47, 255 (1972).
[24]H.P. Künzle, "Galilei and Lorentz invariance of classical particle interactions," in *Symposia Mathematica* (Academic, New York) (to be published).
[25]E.H. Kerner, J. Math. Phys. 6, 1218 (1965).
[26]R.N. Hill and E.H. Kerner, Phys. Rev. Lett. 17, 1156 (1966).
[27]R.N. Hill, J. Math. Phys. 8, 1756 (1967).
[28]A.N. Beard and R. Fong, Phys. Rev. 182, 1397 (1969).
[29]P. Droz-Vincent, Nuovo Cimento B 12, 1 (1972).
[30]A. Peres, Phys. Rev. Lett. 27, 1666 (1971) and 28, 392 (1972).
[31]H.W. Woodcock and P. Havas, Phys. Rev. D 6, 3422 (1972).
[32]L. Bel, C.R. Acad. Sci. Paris 273 A, 405 (1971) and Ann. Inst. Henri Poincaré 18 A, 57 (1973).
[33]For a survey see, for example, P. Havas in *Problems in the Foundations of Physics*, edited by M. Bunge (Springer, New York, 1971).
[34](x_k^α, v_k^α) is a fiber coordinate system of $E = T\tilde{V}$, $\partial_{\alpha k} := \partial/\partial x_k^\alpha$, $\partial_{\dot{\alpha} k} := \partial/\partial v_k^\alpha$ ($\alpha, \beta, \ldots$ range from 0 to 3, $A, B, \ldots$ from 1 to 3; $k, l = 1, 2$ with $k \neq l$ unless otherwise specified).
[35]P. Droz-Vincent, C.R. Acad. Sci. Paris 275 A, 1263 (1972) and "Local existence for finitely predictive two-body interactions," preprint (1973).
[36]$\eta_{\alpha\beta} := \mathrm{diag}(-1, 1, 1, 1)$, will be used to lower and raise indices.
[37]R.S. Palais, *A Global Formulation of the Lie Theory of Transformation Groups* (American Mathematical Society, Providence, 1957).
[38]However, this distinction is less clear than for E. For example, the imbedding map $\iota_\Sigma : \Sigma \to E$ does not appear to induce an almost tangent structure on Σ except in the most special cases like, e.g., $\Sigma = \{p \in E \mid x_1^0 = x_2^0 = 0,\ v_1^0 = v_2^0 = 1\}$.
[39]If the second order system $\mathcal{E}$ is not invariant under a time translation group, i.e., if the force law is explicitly time dependent, this construction obviously fails. However it can be argued that every isolated system must be at least invariant under time translations and that systems with time dependent forces can always be considered as part of a larger system that would have some invariance group to make this construction possible.
[40]H.P. Künzle, Ann. Inst. Henri Poincaré 17 A, 337 (1972).
[41]Alternatively, one might instead of $\tau = 0$ let $\tau^2 = -r^\alpha r_\alpha = \mu^2 > 0$ (cf. Droz-Vincent, Ref. 35).
[42]Dropping all the bars for quantities on Σ from now on.
[43]Explicit solutions of equivalent systems (not formulated on Σ) have been obtained for one-dimensional motion by R.N. Hill, J. Math. Phys. 11, 1918 (1970); R.A. Rudd and R.N. Hill, J. Math. Phys. 11, 2740 (1970); C.S. Shukre and T.F. Jordan, J. Math. Phys. 13, 868 (1972). L. Bel, A. Salas, and J.M. Sanchez have solved the system formally by expanding ξ_k^α in a power series in the coupling constant.
[44]A. Staruszkiewicz, Ann. Inst. Henri Poincaré 14, 69 (1971) and very recently B. Bruhns, Phys. Rev. D 8, 2370 (1973) have introduced a Fokker-type Lagrangian on Σ for a special interaction (essentially the same as our example in Sec. 4) and obtained some and all, respectively, of these first integrals.
[45]For a Galilei invariant one-particle-system there is no invariant θ such that $\omega = -d\theta$ defines a second order system (cf., for example, Ref. 24). Hence there is no such θ for the noninteracting two-particle-system and therefore none for at least a weakly interacting system (cf. Refs. 8 and 24).
[46]This cannot be said of all the quantities $M_{\alpha\beta}$ some of which correspond to an (origin dependent) angular momentum vector, the others to the "center of mass."
[47]R. Abraham and J.E. Marsden, *Foundations of Mechanics* (Benjamin, New York, 1967).
[48]J. Marsden and A. Weinstein, "Reduction of symplectic manifolds with symmetry," preprint (1972).
[49]See Ref. 32 or 24. The condition can be stated somewhat more elegantly in the terminology of almost tangent structures [cf. J. Klein and A. Voutier, Ann. Inst. Fourier 18, 241 (1968)].
[50]In the restricted sense in which we use this term here this would consist in showing that there exists a function $L : \Sigma \to R$ and coordinates $\tilde{v}_k^A$ complementary to the x_k^A such that $\dot{x}_k^A = \tilde{v}_k^A$ [cf. P2.38)] and $P_{kA} = \partial L/\partial \tilde{v}_k^A$.
[51]See, for example, P. Havas, Suppl. Nuovo Cimento Ser. 10, 5, 364 (1957).
[52]As was done already by J.L. Synge, Proc. Roy. Soc. A 177, 118 (1940) (who, however, used only retarded interactions). The present example in advanced—retarded form was apparently first considered by A.D. Fokker, Physica 9, 33 (1929).
[53]Also Staruskiewicz and Bruhns (Ref. 44) have chosen a simple Lagrangian, rather than the simple equations of motion.

Exponentially small scattering amplitude in high energy potential scattering. II*

A. Paliov (Lev)

Department of Physics, Technion-Israel Institute of Technology, Haifa, Israel

S. Rosendorff

Institute of Physics, University of Aarhus, Aarhus, Denmark

Department of Physics, Technion-Israel Institute of Technology, Haifa, Israel'

(Received 14 December 1972)

The high energy, small angle Glauber–Molière scattering amplitude of spherical symmetric potentials, which are expandable in ascending even powers of r and singular in coordinate space, is calculated. The singularities are poles of order $n(=1,2,\cdots)$ off the real axis. As in the nonsingular case the amplitude decreases like $\exp(-q\,b)$ as a function of the momentum transfer q. However, the dependenc of b on q is quite different. Unlike in the nonsingular case it reaches a finite value at infinite energy. Also, in contradistinction to the nonsingular case, the first Born approximation under certain conditions holds in the whole range of validity of the scattering angles.

1. INTRODUCTION

In a previous paper[1] the high energy scattering amplitude of spherical symmetry potentials, expandable in ascending even power of r and nonsingular in coordinate space, has been calculated.

In the present paper we shall discuss potentials which are also even in r but are singular in coordinate space. The singularities considered are poles of order n $(=1,2,\cdots)$ off the real axis. For simplicity we have chosen them to be on the imaginary axis. We calculate the corresponding amplitude by making use of the well-known Glauber–Molière impact parameter representation.[2,3] Our result is thus valid for small scattering angles only. Large angle scattering amplitudes due to even power singular potentials will be examined in another paper,[4] where use will be made of the Landau–Lifshitz approach discussed extensively in I.

The scattering amplitudes calculated in the present paper turn out to have an exponential decrease as function of the momentum transfer q, similar to the behavior of the scattering amplitudes due to potentials which are nonsingular in the finite coordinate plane discussed in I. There are however some well-defined differences to which attention should be given. In both cases the dependence of the amplitude on q is of the form $\exp(-qb)$, where b may be considered to be the effective range of the interaction. In both cases b depends on q. However, in the nonsingular case, b increases slowly (as the square root of the logarithmic function) but indefinitely when $q\to\infty$, whereas in the singular case b approaches a definite limit r_0. Here r_0 is the distance of the singularity of the interaction from the origin. Another important difference between the nonsingular and the singular interaction is that in the former case the first Born approximation never holds, except for very small scattering angles. In the latter case, however, under certain conditions, the first Born approximation holds essentially in the whole range of validity of the scattering angles allowed by the Glauber–Molière representation.

We assume throughout the paper that the coupling constant has a power law dependence on the momentum, i.e., $g(k)=g_0k^m$. The $m=1$ case is equivalent to an energy independent potential in which the relativistic dependence of the mass has been taken into account. It is important to point out that the Glauber–Molière representation is a valid small-angle, high energy approximation of the scattering amplitude if $|g(k)|/k^2 \ll 1$.[1-3] We shall therefore always assume that $m<2$. The Born approximation, on the other hand, is valid only when $m<1$.

In Sec. 2 the conditions for the validity or nonvalidity of the Born approximation are discussed. In Sec. 3 the scattering amplitude is evaluated by the saddle point method when the Born approximation breaks down. The main points of the paper are summarized in Sec. 4.

Throughout this paper, we use $\hbar=c=2m=1$.

2. THE BORN APPROXIMATION

As pointed out in the Introduction, we wish to calculate the high energy, Glauber–Molière scattering amplitude due to an even-power potential function $V(r)$ which has poles of order n in the r plane excluding the real axis. There are many functions which satisfy this condition. Probably the simplest function is of the form

$$\left(\frac{1}{r+ir_0}-\frac{1}{r-ir_0}\right)^n.$$

It is, however, rather difficult to deal with, as far as calculation of the amplitude is concerned.

A better choice turns out to be the function

$$V(r)=g(r^2+r_0^2)^{-n},\qquad g=|g|\exp(-i\gamma),\quad 0\le\gamma\le\pi,\quad n=1,2,3,\cdots. \tag{1}$$

We assume, for simplicity, that r_0 is real. It is well known that the Glauber–Molière amplitude is represented by

$$A=ik\int_0^\infty \rho\,d\rho a(\rho)J_0(q\rho)$$
$$a(\rho)=1-\exp[2i\delta(\rho)] \tag{2}$$
$$\delta(\rho)=-1/k\int_0^\infty V(\sqrt{\rho^2+z^2})\,dz,$$

where $q=2k\sin\vartheta/2$ is the momentum transfer, and the variable of integration ρ is the impact parameter. The evaluation of the phase shift is straightforward:

$$\delta(\rho) = -\frac{g}{k}\int_0^\infty \frac{dz}{(z^2+\rho^2+r_0^2)^n}$$
$$= -\frac{1}{k}G_n(\rho^2+r_0^2)^{-n+1/2}, \qquad (3)$$

where[5]

$$G_n = g[\pi(2n-3)!!/2^n(n-1)!]. \qquad (3')$$

The singularity of the interaction in coordinate plane thus gives rise to a singularity of the S matrix in impact parameter plane at the same point ir_0. Substituting for $\delta(\rho)$ and expanding the exponential function in powers of G_n yields for the amplitude

$$A = ik\int_0^\infty \sum_{\mu=1}^\infty \rho\, d\rho J_0(q\rho)\left(-2i\frac{G_n}{k}\right)^\mu \frac{(\rho^2+r_0^2)^{-(n-1/2)\mu}}{\mu!}. \qquad (4)$$

The integral over the impact parameter ρ is known,

$$\int_0^\infty \frac{\rho d\rho J_0(q\rho)}{(\rho^2+r_0^2)^\nu} = \left(\frac{q}{2r_0}\right)^{\nu-1}\frac{1}{\Gamma(\nu)}K_{-(\nu-1)}(r_0q), \qquad (5)$$

where $K_\mu(x)$ is the modified Bessel function of the third kind. Thus we obtain for the amplitude, using the relation $K_{-\nu}(x)=K_\nu(x)$

$$A = -ik\sum_{\mu=1}^\infty A^{(\mu)}, \qquad (6)$$

where the μ's order Born amplitude is given by[6]

$$A^{(\mu)} = \left(-2i\frac{G_n}{k}\right)^\mu \left(\frac{q}{2r_0}\right)^{(n-1/2)\mu-1}\frac{K_{(n-1/2)\mu-1}(r_0q)}{\Gamma(\mu+1)\Gamma[(n-1/2)\mu]}. \qquad (6')$$

The ratio $|A^{(2)}/A^{(1)}|$ represents a measure for the condition of the range of validity of the Born approximation. We have

$$\left|\frac{A^{(2)}}{A^{(1)}}\right| = \frac{|G_n|}{k}\left(\frac{q}{2r_0}\right)^{n-1/2}\frac{\Gamma(n-1/2)}{\Gamma(2n-1)}\frac{K_{2n-2}(r_0q)}{K_{n-3/2}(r_0q)}. \qquad (7)$$

$K_\nu(x)$ is independent of ν when $x \gg \nu^2$. We therefore get, in the region of large momentum transfer, $r_0q \gg n^2$, the condition for the validity of the Born approximation

$$\left|\frac{A^{(2)}}{A^{(1)}}\right| = \frac{\Gamma(n-1/2)}{\Gamma(2n-1)}\frac{|G_n|}{k}\left(\frac{q}{2r_0}\right)^{n-1/2} \ll 1. \qquad (8)$$

For fixed momentum transfer q, the above condition is always satisfied, provided the dependence of the coupling constant on k is such that

$$|g(k)|/k \ll 1,$$

i.e., under this condition $A^{(1)}$ is a good approximation to the amplitude. In the framework of the Glauber–Molière theory this statement is true, in general, provided the amplitudes $A^{(1)}$ and $A^{(2)}$ are definable. If the coupling constant has a power law momentum dependence, i.e., if $g=g_0k^m$ then the above condition holds for $m<1$. For $m>1$, Eq. (8) cannot be satisfied (for fixed q).

For large but fixed k, Eq. (8) yields an upper limit on the scattering angle ϑ. We find

$$\vartheta \ll u(k), \qquad (9)$$

where $u(k)$ expresses the explicit dependence of the upper limit on k

$$u(k) \propto (kr_0)^{-(n-3/2+m)/(n-1/2)}. \qquad (9')$$

We must, however bear in mind that in the Glauber–Molière theory ϑ is confined to small values, $\vartheta \ll (kr_0)^{-1/2}$. Hence the upper limit on ϑ is given by

$$\vartheta \ll \mathrm{Min}[u(k), (kr_0)^{-1/2}].$$

Call

$$p = m + \tfrac{1}{2}n - \tfrac{5}{4}, \qquad (10)$$

then for $p>0$, we have $u(k)<(kr_0)^{-1/2}$, and for $p<0$ we have $u(k)>(kr_0)^{-1/2}$. Therefore, when $p<0$ the first Born approximation is valid for all angles which satisfy

$$n^2/kr_0 \ll \vartheta \ll (kr_0)^{-1/2}. \qquad (11)$$

The lower limit is derived from the requirement, $r_0q \gg n^2$, as already mentioned. On the other hand, when $p>0$, we have to distinguish between two cases:

(i) $m<1$. In this case the first Born approximation is valid for angles which satisfy

$$n^2/kr_0 \ll \vartheta \ll u(k), \qquad (12)$$

and is not valid for angles which satisfy

$$u(k) \ll \vartheta \ll (kr_0)^{-1/2}. \qquad (12')$$

It is easy to show that in the extreme forward direction, $\vartheta \ll 1/kr_0$, the Born approximation always holds as long as $m<1$. This follows immediately from the small argument behavior of the modified Bessel function, $K_\nu(x) \sim \frac{1}{2}\Gamma(\nu)(2/x)^\nu$.

(ii) $m \geq 1$. In this case the first Born approximation breaks down in the whole angular range $\vartheta \ll (kr_0)^{-1/2}$.

As to the amplitude itself, when condition (11) or (12) holds, it is well approximated by $A^{(1)}$, Eq. (6′),

$$A \approx -ikA^{(1)} = -[\sqrt{\pi}G_n/\Gamma(n-1/2)r_0](q/2r_0)^{n-2}\exp(-r_0q). \qquad (13)$$

Here we have made use of the asymptotic expansion of $K_\nu(x)$,

$$K_\nu(x) = \sqrt{\pi/2x}\,\exp(-x)[1+(4\nu^2-1/8x)+\cdots]. \qquad (14)$$

When the Born approximation is not valid [case (i), Eq. (12′), and case (ii)], the amplitude can be evaluated by the saddle point method. This will be done in the next section.

In the extreme forward direction, $\vartheta \ll 1/kr_0$, the amplitude is given by

$$A \approx -ikA^{(1)} = -G_n/(n-3/2)r_0^{2n-3}, \qquad (15)$$

provided $m<1$. The method of the present paper does not provide a satisfactory solution for the amplitude in the extreme forward direction when $m \geq 1$.

3. THE SADDLE POINT METHOD

In this section the scattering amplitude will be calculated by the saddle point method. Our starting point is again the Glauber–Molière impact parameter representation, Eqs. (2), (3), (3′). For $\vartheta \neq 0$ we have

$$A=-\frac{ik}{2}\int_0^{\infty}\rho\,d\rho[H_0^{(1)}(q\rho)+H_0^{(2)}(q\rho)]\times\exp\left(-\frac{2iG_n}{k}(\rho^2+r_0^2)^{-n+1/2}\right). \quad (16)$$

r_0 is taken to be real. The integrand has two branch points on the imaginary axis at $\pm ir_0$, and therefore two cuts, one from $+ir_0$ to $+i\infty$, and the other from $-ir_0$ to $-i\infty$. Now, in order to evaluate the amplitude the original paths of integration along the real axis will be opened in such a way that the contours of the $H_0^{(1)}$ and $H_0^{(2)}$ integrals will be along the positive and negative imaginary axes, repectively. Both contours are on the right-hand side of the cuts. It is very easy to see that there are no contributions from the paths which connect the real axis at infinity with the imaginary axis at infinity. As the Hankel functions satisfy the relation $H_0^{(1)}(z)=-H_0^{(2)}(-z)$, and $\delta(\rho)$ is an even function of its argument, the two integrals along the imaginary axis from the origin to ir_0 and $-ir_0$, respectively, cancel each other. Therefore the contour of the $H_0^{(1)}$ integral consists of a contour parallel with and to the right-hand side of the cut from ir_0 to $i\infty$, and a small semicircle which connects this contour with the imaginary axis in the immediate vicinity of the branch point. The contour of the $H_0^{(2)}$ integral is the mirror image of the $H_0^{(1)}$ contour through the real axis. Now, because of the above relation between the two Hankel functions, it is convenient to make the transformation $\rho\to-\rho$ in the $H_0^{(2)}$ integral and thus convert it into a $H_0^{(1)}$ integral along the left upper cut proceeding from $i\infty$ to ir_0. In conclusion the scattering amplitude will be represented by

$$A=-(ik/2)\int_C\rho\,d\rho H_0^{(1)}(q\rho)\exp[-(2iG_n/k)(\rho^2+r_0^2)^{-n+1/2}], \quad (17)$$

where C is a path of integration which starts from $i\infty$, proceeds along the imaginary axis, describes a small circle counterclockwise around ir_0 and returns to $i\infty$ along the imaginary axis. Given the fact that the square root $(\rho^2+r_0^2)^{1/2}$ is positive on the real axis, its phase on the right side of the cut will be $+i$, and on the left side $-i$.

So far everything is exact in the framework of the Glauber–Molière theory. However, when we are in the large momentum transfer region, $r_0q\gg 1$, it is advantageous to make use of the asymptotic expansion of the Hankel function[7]

$$H_0^{(1)}(\rho q)=\sqrt{2/\pi\rho q}\,\exp[i(\rho q-\pi/4)]. \quad (18)$$

Thus the amplitude becomes

$$A=-ik\sqrt{1/2\pi q}\,\exp[-i(\pi/4)]\int_C\sqrt{\rho}\,d\rho\times\exp\{i[\rho q-(2G_n/k)(\rho^2+r_0^2)^{-n+1/2}]\}. \quad (19)$$

Next, let us make the transformation $\rho=i(r_0+s/q)$, then the amplitude assumes the form

$$A=(r_0k^2/2\pi q^3)^{1/2}\exp(-r_0q)\int_C(1+s/r_0q)^{1/2}\,ds\times\exp\left(-s-\frac{2iG_n/k}{(-2r_0s/q)^{n-1/2}(1+s/2r_0q)^{n-1/2}}\right). \quad (20)$$

The cut is now along the positive real axis. The path of integration starts from infinity, proceeds along the real axis, describes a small circle counterclockwise around the origin and returns to infinity along the real axis. The left and right side of the cut in the ρ-plane transform to the upper and lower side of the real axis in the s plane, respectively. Therefore the phase of $\sqrt{-s}$ on the positive side is $(-i)$, and on the negative side is $(+i)$. It follows that $-s=|s|\exp[i(\alpha-\pi)]$.

Let us make the scale transformation

$$s=\mu_0t. \quad (21)$$

We shall determine the parameter μ_0 in such a way that the location of the saddle points in the t plane becomes practically independent of q, k, G_n, and r_0. With this in mind we find

$$\mu_0=\left[\frac{2iG_n}{k}\left(\frac{q}{2r_0}\right)^{n-1/2}\right]^{(n+1/2)^{-1}}, \quad (22)$$

and the amplitude will be of the form

$$A=\sqrt{r_0k^2/2\pi q^3}\,\mu_0\exp(-r_0q)\int_C\sqrt{1+(\mu_0/r_0q)t}\,dt\times\exp[\mu_0f_n(t)] \quad (23)$$

with

$$f_n(t)=-t-(-t)^{-n+1/2}(1+\mu_0t/2r_0q)^{-n+1/2}. \quad (23')$$

If we now make the assumption

$$|\mu_0|/2r_0q\ll 1, \quad (24)$$

then the saddle points are fixed points in the t plane. Before we proceed with the actual calculation of the amplitude, let us find out what the physical implication of Eq. (24) is. The impact parameter ρ_0 at the saddle point is given by

$$\rho_0=ir_0[1+2(\mu_0/2r_0q)t_0]. \quad (25)$$

We show below that $|t_0|\approx 1$. Hence Eq. (24) implies that the saddle points are in the vicinity of the pole (ir_0) of the potential function. In other words, the relevant values of the impact parameter are complex, even when the potential function is real $(\gamma=0,\pi)$. This means that the scattering phenomena are of nonclassical nature. Thus it turns out that the condition of Eq. (24) brings about the fact that the region of angles discussed in the present paper will be inaccessible for a classical particle. The reason why Eq. (24) gives rise to nonclassical scattering can be seen directly from the relationship between impact parameter ρ and scattering angle ϑ. It is determined by

$$\frac{d}{d\rho}[\pm q\rho+2\delta(\rho)]=0. \quad (26)$$

Thus by Eq. (3),

$$\vartheta=G_{n0}(2n-1)[\rho/(\rho^2+r_0^2)^{n+1/2}]k^{m-2}. \quad (27)$$

With condition Eq. (24) in mind, this gives rise to complex values of ρ.

In order to apply the saddle point method, the expansion parameter Eq. (22), has to be a large number. It is easy to verify that it is large when

$$\vartheta\gg\vartheta_l, \quad (28)$$

where

$$\vartheta_l = u(k)/a_n^{(n+1/2)/(n-1/2)} \tag{28'}$$

and

$$a_n = (|G_{n0}| 2^{3/2-n}/r_0^{2n+m-2})^{(n+1/2)^{-1}}. \tag{28''}$$

$u(k)$ is given by Eq. (9'). Thus the saddle point method is complementary to the Born approximation as far as the range of the scattering angle is concerned. This follows from the fact that the Born approximation is valid when $\vartheta \ll u(k)$, as explained in Sec. 2. Also Eq. (24) gives rise to a lower limit on the angle,

$$\vartheta^{(n+1/2)^{-1}} \gg \vartheta_l'^{(n+1/2)^{-1}}, \tag{29}$$

where

$$\vartheta_l' = (a_n/2)^{n+1/2}(r_0 k)^{m-2}. \tag{29'}$$

When $m \ll 1$, the lower limit is determined by Eq. (28). When $m \lesssim 1$, then the relevant lower limit on ϑ is determined by Eq. (29), although $\vartheta_l > \vartheta_l'$. The reason is that in Eq. (29) both sides of the inequality are raised to the power $1/(n+\frac{1}{2})$. When $m > 1$ then $\vartheta_l < \vartheta_l'$, and the lower limit on ϑ is of course given by Eq. (29). The angle ϑ is always smaller than the Glauber limit, $\vartheta \ll (kr_0)^{-1/2}$, therefore there is an upper limit on m, namely $m < 3/2$.

We return now to the expression of the amplitude, Eq. (23). The saddle points are defined by

$$\left.\frac{df_n(t)}{dt}\right|_{t_0} = 0. \tag{30}$$

Taking Eq. (24) into account, they are the solution of

$$\begin{aligned} &\frac{d}{dt}[-t-(-t)^{-n+1/2}] = 0, \\ &-t = |t| \exp[i(\alpha - \pi)]. \end{aligned} \tag{31}$$

We find

$$t_0 = (n-1/2)^{(n+1/2)^{-1}}(-1)^{(n+1/2)^{-1}} \exp(i\pi). \tag{32}$$

Note that $|t_0|$ is of order one. Thus, the omission of the factor $(\mu_0/2r_0 q)t$ in the equation of t_0 is a consistent approximation.

It is obvious that more than one saddle point exists on the same Riemann sheet. Let us put

$$-1 = \exp[i\pi(1-2j)], \quad |j| = 0, 1, 2, \cdots$$

then the phases of the respective saddle points are, according to Eq. (32), given by

$$\alpha_j = [(2n+3-4j)/(2n+1)]\pi. \tag{33}$$

Our Riemann sheet is defined by $0 < \alpha < 2\pi$, which gives rise to upper and lower limits on j

$$-2n-3 < -4j < 2n-1.$$

It follows that the number of saddle points is equal to $(n+1)$ for odd values of n, and equal to n for even values of n. One-half of these points ($j > 0$) are in the upper half-plane. The other half ($j \leqslant 0$) in the lower half-plane are the corresponding complex conjugate points. Their respective phases are α_j and α_{1-j}.

We shall now discuss in detail the contour of integration through the saddle points. Let us first consider the special case of a purely absorptive potential, $\gamma = \pi/2$. Then by Eq. (22) the expansion parameter μ_0 is real, and the contour are determined by $\mathrm{Im} f_n(t) = \mathrm{Im} f_n(t_0)$, or explicitly, taking Eq. (24) into account, by

$$\begin{aligned} &|t| \sin(\alpha) + |t|^{-n+1/2} \sin[(n-1/2)(\pi-\alpha)] \\ &\quad = (n+1/2)(n-1/2)^{-(n-1/2)/(n+1/2)} \sin\alpha_j, \end{aligned} \tag{34}$$

together with the requirement that

$$\mathrm{Re} f(t) = -|t| \cos\alpha - |t|^{-n+1/2} \cos[(n-1/2)(\pi-\alpha)] \tag{34'}$$

decreases monotonically from its peak value at t_0 to $(-\infty)$ on both sides.

Equations (34), (34') permit solutions at infinity for $\alpha \to 0$, and at the origin for $\alpha \to \pi$. There are no solutions for finite $|t|$ and $\alpha = \pi$. Therefore the contour should pass through the positive imaginary axis in order to proceed to infinity. However, this is not always possible. The contours through the saddle points with $j = 1$ cannot pass the positive imaginary axis if

$$\{\sin[(n-1/2)\pi/2]\}^{(n+1/2)^{-1}} > \sin\{[(n-1/2)/(n+1/2)]\pi\},$$

and

$$\sin[(n-1/2)\pi/2] > 0. \tag{35}$$

This follows from Eq. (34). For example, when $n=5$ the contour through the saddle point, $\alpha_1 = 9\pi/11$, satisfies the above condition. The conclusion therefore is that it returns to the origin, or, in other words, it forms a closed loop. On the other hand, for all values of n for which $\sin(n-\frac{1}{2})\pi/2 < 0$ a solution can be found on the positive imaginary axis, which means that the contours will either proceed to infinity or return to the origin in the first quadrant.

In Appendix A we discuss the contours of integration through the saddle points in detail. The contours in the upper and lower half-plane are mirror images of each other through the real axis. In addition we prove that the condition for the contour defined by saddle point α_j [8] to be a closed loop is

$$(\sin\beta_j)^{n-1/2} > (\sin\alpha_j)^{n+1/2}, \tag{A7}$$

where β_j and β_{j-1} are the exit and entrance angles of the loop at the origin, respectively. The angles β_j are defined by

$$\beta_j = [(2n-1-4j)/(2n-1)]\pi, \quad n \geqslant 2|j|+1. \tag{A1}$$

Application of the above result shows that there is no loop for $n=3$. For $n=4,5,6,7$ there is one loop, for $n=8,9,10,11$ there are two loops, etc. According to Eq. (A7) all saddle points which appear in the second quadrant give rise to loops, provided the corresponding angles β_j are larger than $\pi/2$. Furthermore, saddle points which appear in the first quadrant do not give rise to loops but to open contours. Finally, in those cases in which the saddle points are in the second quadrant, but $\beta_j < \pi/2$, there will be either loops or open contours according to whether Eq. (A7) holds or does not hold.

Suppose now that for given n there are altogether N loops through the saddle points $\alpha_1, \ldots, \alpha_N$. Then according to the above the original contour in the upper half-plane can be split up into N loops plus one open contour. The N loops "cover" the range $\pi \geqslant \alpha \geqslant \beta_N$, and the open contour covers the range $\beta_N \geq \alpha \geq 0$. The remaining saddle points do not contribute to the scattering amplitude. They can, therefore, simply be ignored. For example, for $n=8$, the loop through $\alpha_1 = 15\pi/17$ covers the range $\pi \geq \alpha \geq 11\pi/15$, the second loops through $\alpha_2 = 11\pi/17$ "covers" the range $11\pi/15 \geq \alpha \geq 7\pi/15$, and the remaining part $7\pi/15 \geq \alpha \geq 0$ is "covered" by the open contour which goes through $\alpha_3 = 7\pi/17$. The fourth saddle point $\alpha_4 = 3\pi/17$ covers only the range $3\pi/15 \geq \alpha \geq 0$. Its contribution to the amplitude is already included in the contribution from α_3.

Until now we have discussed saddle points and corresponding contours for a purely absorptive potential, $\gamma = \pi/2$. In the more general case the phase χ of the expansion parameter μ_0, Eq. (22), is given by

$$\chi = (\pi - 2\gamma)/(2n+1) \tag{36}$$

$$-\pi/(2n+1) \leq \chi \leq \pi/(2n+1).$$

The location of the saddle points is of course independent of χ. However, the contours depend on it. The relevant function in the exponent of the integrand of the amplitude, Eq. (23), is now $\exp(i\chi)f_n(t)$. Thus the contours are determined by

$$|t|\sin(\alpha+\chi) + |t|^{-n+1/2}\sin[(n-1/2)(\pi-\alpha)+\chi] = [(n+1/2)/(n-1/2)](n-1/2)^{(n+1/2)^{-1}}\sin(\alpha_j+\chi), \tag{37}$$

together with the requirement that

$$\mathrm{Re}[\exp(i\chi)f_n(t)] = -|t|\cos(\alpha+\chi) - |t|^{-n+1/2} \times\cos[(n-1/2)(\pi-\alpha)+\chi] \tag{37'}$$

decreases monotonically from its peak value at t_{0j} to $(-\infty)$ on both sides. The above expressions are valid for any χ given by Eq. (36), whereas the corresponding expression Eqs. (34), (34′) are valid only when $\chi=0$. From the expressions of α_j and χ, Eqs. (33) and (36), respectively, it follows that the right-hand side of Eq. (37) is positive for all saddle points in the upper half plane ($j>0$), and is negative for all saddle points in the lower half-plane ($j \leq 0$). Consequently, the cut and the contour of integration C are moved into the direction $(-\chi)$. Thus the Riemann sheet is defined by $-\chi \leq \alpha \leq (2\pi-\chi)$. The contours determined by the upper saddle points will approach the cut at infinity asymptotically from above, and the contours determined by the lower saddle points will approach the cut at infinity asymptotically from below. We have pointed out before that for the special case, $\chi=0$, the contours in the upper and lower half-plane are mirror images of each other. A glance at Eq. (37) reveals that, in general, when $\chi \neq 0$, this statement is no longer true. However, Eq. (37) is invariant under the transformation $\alpha \to 2\pi - \alpha$, $\chi \to -\chi$. This follows simply from the fact that the saddle points in the upper and lower half-plane are complex conjugates of each other. This invariance is no more than another way of expressing the relationship[1]

$$A_1(\gamma) = -A_2^*(\pi-\gamma), \tag{38}$$

where the two amplitudes are defined by

$$A_{1,2}(\gamma) = -\frac{ik}{2}\int_0^\infty \rho\, d\rho \exp[2i\delta(\rho)]H_0^{(1,2)}(q\rho). \tag{38'}$$

In the Glauber–Molière representation, Eq. (38) is an exact statement for all interactions for which the phase is independent of the coordinate. The contours of integration for the general case ($\chi \neq 0$) are also discussed in Appendix A. We find that for saddle points in the upper half-plane ($j>0$), the condition for closed loops is

$$(\sin(\beta_j+\chi))^{n+1/2} > (\sin(\alpha_j+\chi))^{n-1/2} \tag{A14}$$

and for saddle points in the lower half-plane ($j<0$), the condition is

$$|\sin(\beta_j+\chi)|^{n+1/2} > |\sin(\alpha_{j+1}+\chi)|^{n-1/2}. \tag{A14'}$$

The exit and entrance angles, β_j and β_{j-1}, are determined by

$$(n-1/2)(\pi-\beta_j)+\chi = 2\pi j, \quad n \geq 2|j|+1. \tag{A13}$$

Our conclusion is, therefore, that the scattering amplitude, Eq. (23), breaks up into a sum of $(L+2)$ integrals, where L is the number of closed loops. Each integral is estimated by the standard method of steepest descent. Hence the amplitude will be of the form

$$A = (r_0k^2/q^3)^{1/2}\mu_0\exp(-r_0q)\sum_j[|\mu_0 f_n''(t_{0j})|]^{-1/2} \times\exp[\mu_0 f_n(t_{0j}) + i\varphi_j], \tag{39}$$

where φ_j is the inclination of the contour at saddle points α_j. Call $f_n'' = |f_n''|\exp(i\vartheta_j)$, then φ_j is determined by

$$\cos(\vartheta_j + 2\varphi_j + \chi) = -1.$$

The functions $f_n(t_{0j})$ at saddle points α_j are according to Eqs. (23′) and (32) given by

$$f_n(t_{0j}) = (-t_{0j})\{1 + [1/(n-1/2)][1+(\mu_0/2r_0q)t_{0j}]^{-n+1/2}\}. \tag{40}$$

Note that these functions appear in the exponent multiplied by $\mu_0 \gg 1$. Therefore the small term $(\mu_0 t_{0j}/2r_0q)$ cannot be neglected. In the expression of the second derivatives, on the other hand, it is negligible. We thus find

$$f_n''(t_{0j}) = (n+1/2)/(-t_{0j}). \tag{41}$$

In order to check the reliability of the saddle point method the next term has been calculated. We find that in Eq. (39) each term has to be multiplied by $(1+C_{nj}/2\mu_0)$, where the parameters C_{nj} are defined by the functions $f_n(t_{0j})$ according to Ref. 1, Eq. (A7′). For the problem under discussion it is equal to

$$C_{nj} = (1/6t_{0j})[(n^2+9n/2+3)/(n+1/2)]\exp(-i\chi). \tag{42}$$

Equation (39) is thus a reliable representation of the scattering amplitude if

$$|C_{nj}/2\mu_0| \ll 1. \tag{43}$$

This inequality measures the accuracy of the saddle point method. It gives rise to a lower limit on the scattering angle. We find

$$\vartheta \gg \vartheta_l'',$$

$$\vartheta_l'' = (C_n/2a_n)^{(n+1/2)/(n-1/2)}u(k), \tag{44}$$

where $C_n = |C_{nj}|$, and $u(k)$ is given by Eq. (9′). It replaces Eq. (28) which was derived earlier by requiring simply $|\mu_0| \gg 1$. Note that as a function of n, a_n does not depend much upon it, whereas C_n increases with it linearly. This means that there is an upper limit on n keeping all the other quantities at a fixed value.

Let us show now that the main contribution to the amplitude, Eq. (39), comes from the two saddle points α_1 and α_0. The rest of the saddle points contribute very little under conditions which are similar to those outlined above. Obviously, the contribution of saddle point α_2 is negligible if

$$|\exp[\mu_0 f_n(t_{01})]| \gg |\exp[\mu_0 f_n(t_{02})|, \tag{45}$$

and we have a corresponding condition for the omission of α_{-1} in comparison with α_0. The contribution from the other saddle points are even smaller than from α_2 and α_{-1}. From Eq. (45) we get

$$|\text{Re}[\mu_0(f_n(t_{01}) - f_n(t_{02}))]| \gg 1,$$

which by Eqs. (40) and (36) becomes

$$|\mu_0| \gg \tfrac{1}{2}|C_n'|, \tag{46}$$

$$|C_n'| = \left(\frac{n-1/2}{n+1/2}\right)\frac{2}{|t_0|[\cos(\alpha_2+\chi) - \cos(\alpha_1+\chi)]}. \tag{46'}$$

This condition is very similar to that of Eq. (43). It is easily verified that $|C_n'|$ increases with n, when $n \gg 1$, like n^2, and that for $n \gtrsim 5$

$$|C_n'| > C_n.$$

Thus the only change to be made is that for $n \gtrsim 5$, C_n is replaced by $|C_n'|$ in the expression of ϑ_l'', Eq. (44).

Consequently, the amplitude Eq. (39) assumes the form

$$A = -(r_0k^2/q^3)^{1/2}\mu_0 \exp(-r_0q)[|\mu_0 f_n''(t_0)|]^{-1/2} \times \{\exp[\mu_0 f_n(t_{01}) + i\varphi_1] + \exp[\mu_0 f_n(t_{00}) + i\varphi_0]\},$$

which by Eqs. (40), (41), (33) becomes

$$A = -\{(r_0k^2/q^3)[(n-1/2)/(n+1/2)]^{1/(n+1/2)}|\mu_0|\}^{1/2} \times \exp(i\chi/2)\exp(-r_0q)\{\exp[|\mu_0|F(\chi;\alpha_1) + i\alpha_1/2] - \exp[|\mu_0|F(\chi;-\alpha_1) - i\alpha_1/2]\}. \tag{47}$$

Here α_1, and χ are given by Eqs. (33) and (36), respectively, $F(\chi;\alpha)$ is defined by

$$F(\chi;\alpha) = -|t_0|\exp[i(\chi+\alpha)](1 + [1/(n-1/2)] \times \{1 + (|\mu_0||t_0|/2r_0q)\exp[i(\chi+\alpha)]\}^{-n+1/2}), \tag{48}$$

and $|\mu_0|$ is the absolute value of μ_0, Eq. (22).

In the immediate neighborhood of $\chi=0$ (i.e., $\gamma=\pi/2$) both saddle points contribute equally, as far as the magnitudes are concerned. When $\chi>0$ and not very close to zero, the main contribution comes from saddle point α_1. When $\chi<0$ (and not very close to zero) the main contribution comes from saddle point α_0. Clearly the condition that only one saddle point contributes significantly is given by

$$|\mu_0| \cdot |\text{Re}(F(\chi;\alpha_1) - F(\chi;-\alpha_1))| \gg 1,$$

or equivalently

$$2|\mu_0||t_0||\sin\chi|\sin\alpha_1 \gg 1. \tag{49}$$

For $\chi \neq 0$, this condition is essentially the same as that of Eq. (46) for the omission of the rest of the saddle points, $\alpha_2, \alpha_3, \cdots$. Therefore when Eq. (49) is satisfied, the amplitude Eq. (47) assumes the simplified form

$$A = \mp\{(r_0k^2/q^3)[(n-1/2)/(n+1/2)]^{(n+1/2)^{-1}}|\mu_0|\}^{1/2} \times \exp[-r_0q + |\mu_0|F(\chi;\pm\alpha_1) + i/2(\chi\pm\alpha_1)], \tag{50}$$

where the upper sign is to be taken when $\chi>0$, and the lower sign when $\chi<0$.

4. SUMMARY

We have calculated the high-energy Glauber–Molière scattering amplitude of potentials $V(r)$ which are even powered and singular in coordinate space. The singularities of V are poles of order n $(=1,2,\cdots)$ off the real axis. For simplicity we have chosen them to be on the imaginary axis. The coupling constant g is assumed to have a power law energy dependence $g \propto k^m$. It is well-known that the Glauber–Molière representation of the amplitude is a small angle scattering theory. The scattering angles ϑ have an upper limit $\vartheta_G = (\sqrt{kr_0})^{-1}$. Only scattering through angles for which $\vartheta \gg \vartheta_{\min} \approx 1/kr_0$ has been considered.

In Sec. 2 the conditions for the validity of the Born approximation as function of n, m, and ϑ have been found. The result is, loosely speaking, the well-known fact that the Born approximation is valid when the effective interaction is weak. In the problem under discussion this means, small values of n, m, and ϑ (small angles correspond to big impact parameters, thus weak interaction). The exact result established is that if the parameter $p=n+m-5/4<0$, then the Born approximation is valid in the angular region $\vartheta_{\min} \ll \vartheta \ll \vartheta_G$. However, for $p>0$ one has to distinguish between the case $m<1$ when the Born approximation is valid for $\vartheta_{\min} \ll \vartheta \ll u(k)$ and the case $m \geq 1$ when the Born approximation is not valid in the whole angular range, $\vartheta \gg \vartheta_{\min}$. $u(k)$ is given by Eq. (9′). The condition for the validity of the Born approximation thus depends critically on m. The dependence on n is much weaker and smoother. When n becomes very large, $u \to 1/kr_0$ and the angular region of the Born approximation shrinks to zero.

The amplitude itself, when the Born approximation is not valid, has been calculated in Sec. III. It has been evalued by the saddle point method. Although the total number of saddle points is equal to n for n even, and $(n+1)$ for n odd, the number of saddle points which contribute dominantly is very much smaller. For an absorptive potential $(\gamma \approx \pi/2)$ there are only two saddle points of importance and for $\gamma \neq \pi/2$ only one saddle point contributes dominantly. The corresponding amplitudes are given by Eqs. (47) and (50), respectively. Note that

$$\begin{aligned} &\cos(|\chi|+\alpha_1)<0, \quad \text{for } n \geq 2, \quad \text{and} \quad 0 \leq \gamma \leq \pi, \\ &\cos(|\chi|+\alpha_1)<0, \quad \text{for } n=1 \quad \text{and} \quad \begin{cases} 0 \leq \gamma < \pi/4 \\ 3\pi/4 < \gamma \leq \pi. \end{cases} \end{aligned} \tag{51}$$

It follows that $\text{Re}F(\chi;\pm\alpha_1)>0$, and therefore the differential cross section, decreases much more slowly than $\exp(-2r_0q)$.[9] The factor $\exp(|\mu_0|F)$, which measures

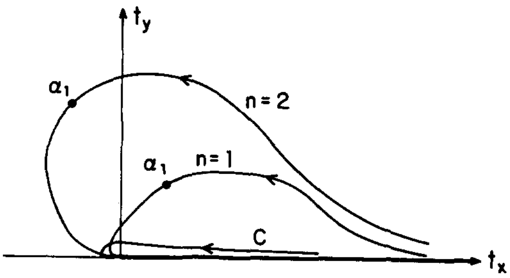


FIG. 1. The contours of steepest descent in the upper half plane through the saddle points α_1 for the absorptive ($\gamma=\pi/2$) potentials $n=1$ and $n=2$, respectively. The arrows show the direction of integration.

the deviation from the exponential law, contains most of the information concerning the interaction, namely, the coupling constant $g(k)$, including its energy dependence and phase γ, and the order of the singularity n. As

$$|\mu_0| \propto [(|g|/k)q^{n-1/2}]^{(n+1/2)^{-1}}$$

we conclude that the stronger the singularity of the interaction and the greater the strength of the interaction itself, the larger the elastic cross section. Furthermore, for given n and given $|g|$, the elastic cross section increases with decreasing absorption ($\gamma\to 0$, or $\gamma\to\pi$), i.e., the smaller the number of open inelastic channels, the greater the elastic cross section. However, when $n=1$, and $\pi/4<\gamma<3\pi/4$, we have $\cos(|\chi|+\alpha_1)>0$. In this case the elastic cross section as function of ϑ decreases faster than $\exp(-2r_0q)$. It also decreases with increasing $|g|$.

The calculation of large angle scattering amplitudes due to potentials of the same kind as those discussed in the present paper, will be dealt with in a later communication.

ACKNOWLEDGMENT

One of us (S.R.) is indebted to Professor Jens Lindhard for the hospitality of the Institute of Physics at the University of Aarhus.

APPENDIX

In this Appendix we investigate the question of loops systematically. First we discuss the contour in the upper half plane. The contours in the lower half plane are their mirror images through the real axis. We first consider the contours for $n=1$ and $n=2$. In both cases there is only one saddle point; for $n=1$ it is at $\alpha_1=\pi/3$, and for $n=2$ it is at $\alpha_1=3\pi/5$. Both contours start at the origin with $\alpha=\pi$ and end at infinity with $\alpha=0$. They are sketched in Fig. 1.

For $n=3$ we have two saddle points at $\alpha_1=5\pi/7$, and at $\alpha_2=\pi/7$. The contour through α_1 is similar to the contours of $n=1$ and $n=2$. The contours through α_2, however, is different. It emerges with the origin at an angle $\beta_1=\pi/5$. The two contours are given in Fig. 2. It is obvious that the contour C_1 through α_1 is the correct one, because the original contour can be deformed into it. In other words, the contour through α_1 "covers" the original contour C. On the other hand, the original contour cannot be deformed into the contour C_2 through α_2 simply because the segment of the circle in the neighborhood of the origin, $\pi>\alpha>\pi/5$ is not "covered" by it.

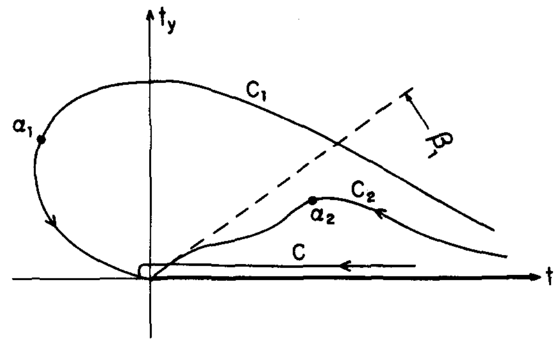


FIG. 2. The contours of steepest descent C_1 and C_2 in the upper half plane through the saddle points α_1 and α_2, respectively, for the absorptive potential $n=3$. C_1 'covers' the original contour C, wheras C_2 does not.

Next let us consider the case of $n=4$. There are again two saddle points, at $\alpha_1=7\pi/9$ and $\alpha_2=\pi/3$. Exact analysis reveals that the contour through α_1 is a closed loop, which begins and ends at the origin in directions π and $3\pi/7$, respectively. The other contour emerges from the origin in the direction $3\pi/7$ and proceeds through the saddle point, α_2, to infinity. The two contours are given in Fig. 3. It follows that the original contour C is covered by the sum of the two contours C_1 and C_2.

We shall now consider the question of loops in general. The first loop emerges from the origin in direction π, proceeds through the saddle point α_1 and returns to the origin at an angle β_1. This angle is determined by Eq. (34) and the condition $\mathrm{Re} f(t)\to-\infty$. Thus by Eq. (34′) we have $(n-\frac{1}{2})(\pi-\beta_1)=2\pi$, or

$$\beta_1=[(2n-5)/(2n-1)]\pi, \quad n\geq 3,$$

from which it immediately follows that no loops exist for $n=1$ and $n=2$. By the same argument, the next loop which goes through the saddle point α_2 emerges from the origin in the direction β_1 and returns to it at an angle β_2, determined by $(n-\frac{1}{2})(\pi-\beta_2)=4\pi$, i.e.,

$$\beta_2=[(2n-9)/(2n-1)]\pi, \quad n\geq 5.$$

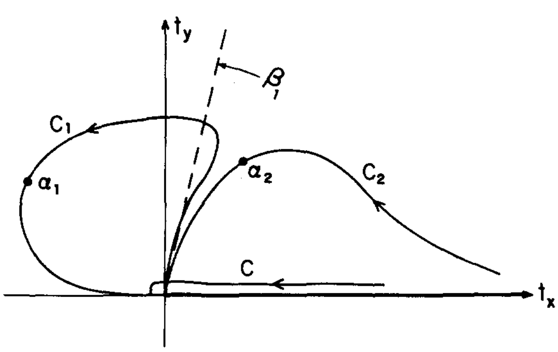


FIG. 3. The same as in Fig. 2. for the absorptive potential $n=4$. C_1 is a closed loop. The original contour C is "covered" by the sum (C_1+C_2).

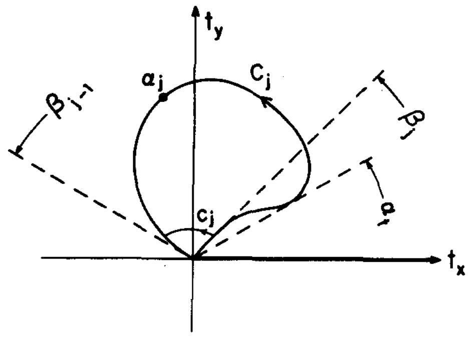


FIG. 4. This figure shows schematically a closed loop C_j through the saddle point α_j with exit and entrance angles β_j and β_{j-1}, respectively. α_t is the angle of the tangent to the loop on the right-hand side.

In general, the jth loop emerges from the origin at angle β_j and returns to it at angle β_{j-1}, where

$$\beta_j=[(2n-1-4j)/(2n-1)]\pi, \quad n\geq 2|j|+1 \tag{A1}$$

However, not all contours are closed loops. Let us find the condition of closed loops. Suppose we are in the vicinity of direction β_j. We then define a small positive angle δ by

$$(n-1/2)(\pi-\alpha)=2\pi j+\delta. \tag{A2}$$

Therefore the contour in this region is according to Eq. (34) determined by

$$|t|\sin\beta_j+|t|^{-n+1/2}\sin\delta=[(n+1/2)/(n-1/2)] \times(n-1/2)^{(n+1/2)^{-1}}\sin\alpha_j. \tag{A3}$$

Keeping the angles constant, this expression becomes an equation for $|t|$. Call the left-hand side $g(|t|)$. It is always positive. Its minimum occurs at

$$\rho_0=[(n-1/2)\sin\delta/\sin\beta_j]^{(n+1/2)^{-1}},$$

and we have

$$g(\rho_0)=[(n+1/2)/(n-1/2)](n-1/2)^{(n+1/2)^{-1}} \times(\sin\delta)^{(n+1/2)^{-1}}(\sin\beta_j)^{(n-1/2)/(n+1/2)}. \tag{A4}$$

It follows that if a range of angles δ can be defined such that

$$(\sin\delta)^{(n+1/2)^{-1}}(\sin\beta_j)^{(n-1/2)/(n+1/2)}>\sin\alpha_j, \tag{A5}$$

there will be *no* solution to the above equation. In other words we have a loop through the saddle point α_j.

Let the tangent to the loop on the right hand side be in the direction α_t (see Fig. 4). Clearly, the corresponding angle δ_t is determined by

$$\sin\delta_t=(\sin\alpha_j)^{n+1/2}/(\sin\beta_j)^{n-1/2}, \tag{A6}$$

from which, by Eq. (A2), α_t can be calculated. We therefore conclude that if the condition

$$(\sin\beta_j)^{n-1/2}>(\sin\alpha_j)^{n+1/2} \tag{A7}$$

is fulfilled, there will exist a loop through saddle point α_j with exit angle β_j. To complete the picture we have to determine the angle at which the contour returns to the origin. It is not difficult to show that this angle is equal to β_{j-1}. The proof is as follows: First, we notice that it cannot return to the origin at an angle smaller than β_{j-1}. This follows from Eq. (34′). Next, we show that the contour cannot pass the direction β_{j-1}. At this angle the value of $|t|$ is according to Eq. (34)

$$|t_1|=b_n(\sin\alpha_j/\sin\beta_{j-1}), \tag{A8}$$

where

$$b_n=[(n+1/2)/(n-1/2)](n-1/2)^{(n+1/2)^{-1}}. \tag{A8′}$$

The corresponding value of $\mathrm{Re}f(t)$ at this point is according to Eq. (34′)

$$-b_n\sin\alpha_j ctg\beta_{j-1}-(b_n\sin\alpha_j/\sin\beta_{j-1})^{-n+1/2}. \tag{A9}$$

This should be compared to $\mathrm{Re}f(t)$ at the saddle point α_j. We have, according to Eqs. (32), (33),

$$\mathrm{Re}f(t_{0j})=-b_n\cos\alpha_j. \tag{A10}$$

Clearly, the contour cannot cross the direction β_{j-1} if

$$\mathrm{Re}f(t_1)>\mathrm{Re}f(t_{0j}), \tag{A11}$$

which by Eqs. (A9) and (A10) can be put in the form

$$\frac{\sin(\beta_{j-1}-\alpha_j)}{\sin\alpha_j}>\left(\frac{\sin\beta_{j-1}}{b_n\sin\alpha_j}\right)^{n+1/2}. \tag{A12}$$

This inequality always holds. For n, not small, the left-hand side of Eq. (A12) becomes essentially independent of n, whereas the right-hand side decreases like $1/n$. We conclude that if Eq. (A7) is satisfied, there will be a loop, emerging from the origin at angle β_j and returning to it at angle β_{j-1}. On the other hand, if Eq. (A7) is not satisfied the corresponding contour will start from infinity, proceed through the saddle point α_j, and end at the origin at angle β_{j-1}.

We now proceed to determine the condition for the existence of loops when $\chi\neq 0$. Let us define the angles β_j by

$$(n-1/2)(\pi-\beta_j)+\chi=2\pi j, \quad n\geq 2|j|+1, \tag{A13}$$

which is a generalization of Eq. (A1). The exit and entrance angles of the loop through α_j are again β_j and β_{j-1}, respectively. The same reasoning which led to Eq. (A7) will give rise to the condition for the existence of loops in the general case. For saddle points in the up-

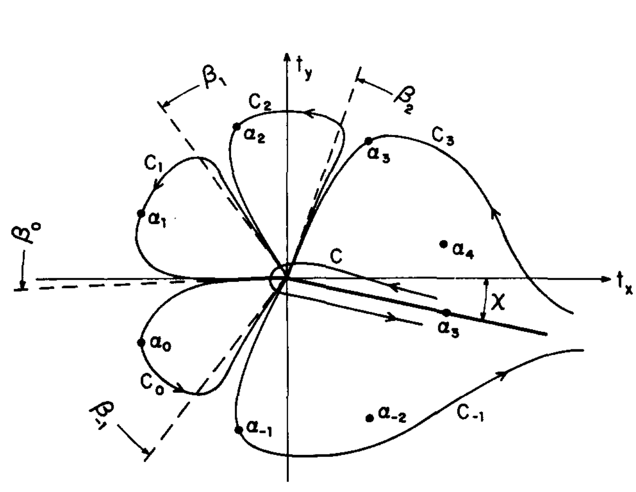


FIG. 5. The contour of steepest descent for the repulsive ($\gamma=0$) potential $n=7$. There are three loops and two open contours. The original contour C is "covered" by the sum $(C_3+C_2+C_1+C_0+C_{-1})$. The cut is in the direction $-\chi=-\pi/15$.

per half-plane ($j>0$), we obtain

$$(\sin(\beta_j+\chi))^{n+1/2} > (\sin(\alpha_j+\chi))^{n-1/2}, \tag{A14}$$

and for saddle points in the lower half plane ($j<0$), we obtain

$$(|\sin(\beta_j+\chi)|)^{n+1/2} > (|\sin(\alpha_{j+1}+\chi)|)^{n-1/2}. \tag{A14$'$}$$

Application of these conditions shows that for $\chi=\chi_{max}=\pi/(2n+1)$, the first loop appears in the upper half-plane when $n=3$ and the second loop appears when $n=7$. For the same value of χ the first loop in the lower half-plane appears when $n=5$, and the second loop appears when $n=9$.

From the equations of the contours, Eqs. (37), (37′), it follows that the contour (loop or open contour, as the case may be) through α_1 enters the origin at angle β_0 which by Eq. (A13) is given by

$$\beta_0=\pi+2\chi/(2n-1). \tag{A15}$$

The contour through α_0 in the lower half-plane leaves the origin at the same angle. As to all the other contours, it can be shown that in the upper half-plane the contour through α_j, ($j>0$), enters the origin at β_{j-1}, and in the lower half-plane the contour through α_j, ($j<0$) leaves the origin at β_j. These conclusions follow from arguments which are similar to those which preceded Eq. (A12). The generalization of Eq. (A12) for $\chi\neq 0$ is straightforward. For saddle points in the upper half-plane it is

$$\sin(\beta_{j-1}-\alpha_j)/\sin(\alpha_j+\chi) > [\sin(\beta_{j-1}+\chi)/b_n \sin(\alpha_j+\chi)]^{n+1/2}, \quad j>1. \tag{A16}$$

and in the lower half-plane it is

$$\sin(\beta_j-\alpha_j)/\sin(\alpha_j+\chi) > [\sin(\beta_j+\chi)/b_n \sin(\alpha_j+\chi)]^{n+1/2}, \quad j\leq 1. \tag{A16$'$}$$

These inequalities hold for all saddle points. Our above statement is therefore correct. In Fig. 5 the contours for $n=7$ and $\chi=\chi_{max}=\pi/15$ are shown. We have two loops, C_1 and C_2 in the upper half-plane and one loop C_4 in the lower half-plane. It is obvious, from the figure, that the three loops and the two additional open contours C_3 and C_5 "cover" the original contour C completely.

*This article is based in part on a chapter of the thesis submitted to the Technion by A. P. in partial fulfillment of the D.Sc. degree.

†Permanent address.

[1]A. Paliov (Lev) and S. Rosendorff, Ann. Phys. (N.Y.) **66**, (1971), 189. It will be referred to as I hereafter. Most of the relevant references are listed in this paper.

[2]R.J. Glauber, in *Lectures in Theoretical Physics*, edited by W.E. Brittin *et al.* (Interscience, New York, 1958), Vol. I.

[3]G. Molière, Z. Naturforsch. **A2**, (1947), 133.

[4]S. Rosendorff, "Exponentially small scattering amplitude in high energy potential scattering. III" (to be published).

[5]With $(-1)!!=1$ the case $n=1$ is included.

[6]The Born series, Eq. (6) converges absolutely for every finite value of r_0q and $|G_n|/k$. This follows from the behavior of $K_\nu(x)\sim\frac{1}{2}\Gamma(\nu)(2/x)^\nu$ for $\nu\gg x$. Therefore

$$|A^{(\mu)}|\sim\frac{(r_0^2/2n-1)(2|G_n|/kr_0^{2n-1})^\mu}{\mu\mu!}.$$

In other words the Born series converges faster than the exponential law.

[7]Further on we show that the main contribution to the amplitude comes from the vicinity of saddle points for which $|\rho|\simeq r_0$. Therefore the evaluation of the amplitude by this method is consistent with the assumption $r_0q\gg 1$.

[8]From now on most of the time the saddle points will be denoted by their phases α_j.

[9]This factor measures the location of the singularity of the interaction; in other words it gives the effective range at infinite energy.

Poincaré analytic vectors in axiomatic quantum field theory

Håkan Snellman*

Physique Mathématique, Collège de France, Paris 5^e, France
(Received 23 October 1972)

It is demonstrated that in a Wightman quantum field theory with a denumerable set of field operators over the test function space $\mathcal{S}(\mathbb{R}^4)$, transforming covariantly among themselves under a unitary representation $U(\Lambda, a)$ of the covering group of the Poincaré group, the domain D_0 of vectors generated by the polynomial ring over the field operators applied on the vacuum state contains a dense invariant set of analytic vectors for $U(\Lambda, a)$.

1. INTRODUCTION

Regular vectors for Lie-group representations have been studied extensively in the literature.[1,2] For a unitary representation $U(\mathcal{G})$ of a real Lie group $\mathcal{G}$ with Lie algebra g in a Hilbert space $\mathcal{H}$ one knows, e.g., that both the C^∞ domain $D^\infty(U(\mathcal{G}))$, the Gårding domain $D^G(U(\mathcal{G}))$ and the domain $D^\omega(U(\mathcal{G}))$ of analytic vectors for $U(\mathcal{G})$ are dense in $\mathcal{H}$.

On all these domains the differential $dU(\mathrm{g})$ of $U(\mathcal{G})$ defines a representation of g by operators essentially skew-adjoint. The domain $D^\omega(U(\mathcal{G}))$ also admits free passage between the Lie-algebra representation $dU(\mathrm{g})$ and the Lie-group representation $U(\mathcal{G})$ by exponentiation.

These results are extremely useful in quantum physics, where the invariance groups of the systems under consideration are implemented by unitary representations in the Hilbert space of states, but where the corresponding Lie-algebra representation has a more direct physical interpretation in terms of observables.

In the Wightman formulation of quantum field theory one assumes the existence in the Hilbert space of states of a unitary representation of the universal covering group $\tilde{\mathcal{P}}_+^\uparrow$ of the restricted Poincaré group $\mathcal{P}_+^\uparrow$. In the present paper we are going to discuss analytic vectors for this representation and for definiteness we formulate below the axioms.[3,4]

A 0: The space of states is a Hilbert space $\mathcal{H}$ over the complex field $\mathbb{C}$.

A 1: The test function-space $\mathcal{S}(\mathbb{R}^4)$ [or $\mathcal{D}(\mathbb{R}^4)$] and the set of fields $\mathcal{A}=\{\phi_i(x);\ i\in I\}$ are mapped into linear operators $\phi_i(\varphi)$; $\varphi\in\mathcal{S}(R^4)$ over $\mathcal{H}$. The operators are defined on a common invariant dense domain D. For $\Phi,\ \Psi\in D,\ (\Phi,\phi_i(\cdot)\Psi)\in\mathcal{S}'(\mathbb{R}^4)$ [or $\mathcal{D}'(\mathbb{R}^4)$].

A 2: There exists a continuous unitary representation $U(\Lambda, a)$ of the universal covering group $\tilde{\mathcal{P}}_+^\uparrow = T^4\otimes SL(2,C)$ of the restricted Poincaré group $\mathcal{P}_+^\uparrow$ on $\mathcal{H}$ such that

$$U(\Lambda,a)D\subset D,\quad (\Lambda,a)\in\tilde{\mathcal{P}}_+^\uparrow. \tag{1.1}$$

A 3: There is one unique state $\Omega\in\mathcal{H}$ (up to normalization and a phase factor) which satisfies

$$U(\Lambda, a)\Omega=\Omega. \tag{1.2}$$

$\Omega\in D$ and the generators of translations, P^μ, have their spectral support in the set $\mathrm{Sp}\{P^\mu\}=\{0\}\cup\{p;p^2=p_0^2-\mathbf{p}^2\geqslant u_0^2>0\}$.

A 4: The fields transform covariantly under $U(\Lambda, a)$, i.e.,

$$U(\Lambda, a)\phi_i(\varphi)U^+(\Lambda, a)=\sum_{i'}S^{-1}_{ii'}(\Lambda)\phi_{i'}(\varphi_{(\Lambda,a)}), \tag{1.3}$$

where $\varphi_{(\Lambda,a)}(x)=\varphi[\Lambda^{-1}(x-a)]$ and S is a representation of $SL(2,\mathbb{C})$ which is at most a direct sum of finite-dimensional irreducible representations.[5]

A 5: Let I be a denumerable infinite index set. Then the polynomial ring $\mathfrak{B}_0$ over the smeared fields applied on Ω is dense in $\mathcal{H}$. The space $\mathfrak{B}_0\Omega$ is denoted D_0.

A 6: Let φ_1 and $\varphi_2\in\mathcal{S}(\mathbb{R}^4)$, and let supp φ_1 be spacelike relative to supp φ_2. Then one of the two relations

$$\phi_i(\varphi_1)\phi_j(\varphi_2)\pm\phi_j(\varphi_2)\phi_i(\varphi_1)=0$$

holds. This axiom as well as the spectral condition in *A*3 will play no further role in this paper.

Now, let $\varphi=\varphi_1(x_1)\otimes\cdots\varphi_k(x_k)\in\otimes_k\mathcal{S}(\mathbb{R}^4)$ [or $\otimes_k\mathcal{D}(\mathbb{R}^4)$]. Then by *A* 1 the mapping

$$\varphi\rightarrow\phi_{i_1}(\varphi_1)\cdots\phi_{i_k}(\varphi_k)\Omega\in\mathcal{H}$$

can be shown to be a strongly continuous mapping of $\otimes_k\mathcal{S}(\mathbb{R}^4)$ into $\mathcal{H}$.[3,4,6]

By using Schwartz nuclear theorem[7] and the fact that $\otimes_k\mathcal{S}(\mathbb{R}^4)$ is dense in $\mathcal{S}(\mathbb{R}^{4k})$ (the same is true for the $\mathcal{D}$ spaces), the domain D_0 can be extended to a domain D_1, generated by the $*$-algebra $\mathfrak{B}$ of all quasilocal (or strictly local) operators of the form

$$\sum_{m=0}^{M}\int d^4x_1\cdots d^4x_m\varphi(x_1,\ldots,x_m)\phi_{i_1}(x_1)\cdots\phi_{i_m}(x_m),$$

$M\in\mathbb{N},\ i_k\in I$, with $\varphi(x_1,\ldots,x_m)\in\mathcal{S}(\mathbb{R}^{4m})$ [or $\mathcal{D}(\mathbb{R}^{4m})$] applied to Ω. The mapping

$$\varphi(x_1,\ldots,x_m)\rightarrow\int d^4x_1\cdots d^4x_m\varphi(x_1,\ldots,x_m)\phi_{i_1}(x_1)\cdots\phi_{i_m}(x_m)\Omega$$

is again strongly continuous and the domain $D_1=\mathfrak{B}\Omega$ is dense in $\mathcal{H}$ since by construction $D_0\subset D_1\subset\mathcal{H}$ and $\overline{D_0}=\mathcal{H}$.

Both D_0 and D_1 are C^∞ domains for the field operators and can be shown to be invariant C^∞ domains for $U(\Lambda, a)$ in $\mathcal{H}$. In this paper we furthermore show that the quasilocal domain D_0, generated by $\mathfrak{B}_0$ with test functions from $\mathcal{S}(\mathbb{R}^4)$, contains a dense invariant domain of analytic vectors for $U(\Lambda, a)$, i.e., $D^\omega(U(\Lambda,a))\cap D_0$ is dense in $\mathcal{H}$. The same result does not however seem to apply for a theory defined on $\mathcal{D}$.[8]

The proof that $D^\omega(U(\Lambda,a))\cap D_0$ is dense in $\mathcal{H}$ does not involve *A* 6 and utilizes only the infinitesimal form of Eq. (1.3) on D_0. The result therefore immediately raises the question of how to formulate a nontrivial generalization of the Wightman axioms $A0$–$A5$ to a Poincaré algebra covariant field theory. Some remarks on this problem are contained in Sec. 3.

2. ANALYTIC VECTORS FOR THE POINCARE GROUP

A. Preliminary discussion

Let us consider for simplicity a Wightman theory with a denumerable set of spinless field operators satisfying axioms $A0$–$A5$. The transformation properties of the field operators under $U(\Lambda, a)$ are then given by

$$U(\Lambda,a)\phi_i(f)U^+(\Lambda,a)=\phi_i(f_{(\Lambda,a)}),\quad f\in \mathcal{S}(\mathbb{R}^4), \tag{2.1}$$

where $f_{(\Lambda,a)}(x)=f(\Lambda^{-1}(x-a))$.

The action of $U(\Lambda, a)$ on a vector $\phi_{i_1}(f_1)\cdots\phi_{i_k}(f_k)\Omega \in D_0$ is thus given by

$$U(\Lambda,a)\phi_{i_1}(f_1)\cdots\phi_{i_k}(f_k)\Omega=\phi_{i_1}(f_{1(\Lambda,a)})\ldots\phi_{i_k}(f_{k(\Lambda,a)})\Omega \tag{2.2}$$

and since $f_{(\Lambda,a)}$ again belongs to $\mathcal{S}(\mathbb{R}^4)$ [or $\mathcal{D}(\mathbb{R}^4)$] this shows that D_0 is invariant under $U(\Lambda, a)$. Equation (2.2) also shows that $U(\Lambda, a)$ is C^∞ on D_0 (the same is true for D_1) since the right-hand side is C^∞ in any local coordinate system around the identity in $T^4\otimes SL(2,C)$. For the translations we have, e.g.,

$$P^\mu\phi_{i_1}(f_1)\cdots\phi_{i_k}(f_k)\Omega =\sum_{m=1}^{k}\phi_{i_1}(f_1)\cdots\phi_{i_m}(i\partial^\mu f_m)\cdots\phi_{i_k}(f_k)\Omega.$$

The generators P^μ and $M^{\mu\nu}$ of $U(\Lambda, a)$ can thus be applied freely any number of times on D_0.

By differentiating relation (2.1) applied to any vector $\Psi\in D_0$, we get

$$[P^\mu,\phi_k(f)]\Psi=\phi_k(i\partial^\mu f)\Psi \tag{2.3}$$

$$[M^{\mu\nu},\phi_k(f)]\Psi=\phi_k(i(x^\mu\partial^\nu-x^\nu\partial^\mu)f)\Psi, \tag{2.4}$$

where $k\in I$.

In the following we shall show that D_0 contains a dense invariant domain of analytic vectors for $U(\Lambda, a)$. We recall that an analytic vector for a continuous unitary one-parameter group $U(t)$; $t\in\mathbb{R}$ in a Hilbert space $\mathcal{H}$ is any vector $\Psi\in\mathcal{H}$ such that $U(t)\Psi$ is analytic in t. By Stone's theorem[9] we have $U(t)=e^{itA}$, with A self-adjoint. Thus the analyticity of $U(t)\Psi$ is equivalent to the convergence of the series

$$\left\|\sum_{m=0}^{\infty}\frac{(it)^m}{m!}A^m\Psi\right\|\leq\sum_{m=0}^{\infty}\frac{|t|^m}{m!}\|A^m\Psi\|, \tag{2.5}$$

for some $|t|>0$.

The vectors in D_0 have the form

$$\Phi=\sum_{m=0}^{M}\phi_{i_1}(\varphi_{i_1})\ldots\phi_{i_m}(\varphi_{i_m})\Omega,\quad M\in\mathbb{N},\varphi_{i_k}\in\mathcal{S}(\mathbb{R}^4), \tag{2.6}$$

$i_k\in I$, with $\Phi=c\cdot\Omega$ $(c\in\mathbb{C})$ for $M=0$.

We shall work with a particular basis of these vectors formed by

$$D_0'=\{\Psi\in\mathcal{H};\Psi=\phi_{i_1}(\varphi_{n_{i_1}})\cdots\phi_{i_m}(\varphi_{n_{i_m}})\Omega;\ m\in\mathbb{N},i_k\in I,\ \text{and}\ \varphi_{n_k}\in H(\mathbb{R}^4)\}, \tag{2.7}$$

where $\Psi=c\Omega$ $(c\in\mathbb{C})$ when $m=0$. $H(\mathbb{R}^4)$ is the space of Hermite functions in four dimensions and n_k is the corresponding multi-index. Evidently $D_0'\subset D_0$ and the linear hull of D_0' is dense in D_0 as well as in $\mathcal{H}$.

Now from $A3$ and the fact that $\Omega\in D_0$ it follows that

$$P^\mu\Omega=M^{\mu\nu}\Omega=0. \tag{2.8}$$

If A is anyone of P^μ or $M^{\mu\nu}$ and if Φ is a vector in D_0' of degree k which we want to prove to be an analytic vector for the unitary one-parameter subgroup of $U(\Lambda, a)$ generated by A, then from Eq. (2.5) we see that we will have to estimate expressions of the form

$$\begin{aligned}\|A^m\Phi\|&=\|A^m\phi_{i_1}(\varphi_{n_{i_1}})\ldots\phi_{i_k}(\varphi_{n_{i_k}})\Omega\|\\&=\|[A^m,\phi_{i_1}(\varphi_{n_{i_1}})\ldots\phi_{i_k}(\varphi_{n_{i_k}})]\Omega\|,\end{aligned} \tag{2.9}$$

for all $m\in\mathbb{N}$. Lemmas for estimating such expressions are collected in the next section.

Finally, an analytic vector for a continuous unitary representation $U(\mathcal{G})$ of a real Lie group $\mathcal{G}$ in $\mathcal{H}$ is any vector $\Psi\in\mathcal{H}$ such that the mapping $\mathcal{G}\ni g\to U(g)\Psi$ is analytic.

B. Lemmas for Hermite functions

Let us recall that the Hermite functions $\varphi_n(x)$ are defined by

$$\varphi_n(x)=[\sqrt{\pi}2^n n!]^{-1/2}e^{1/2x^2}\left(\frac{d}{dx}\right)^n e^{-x^2}. \tag{2.10}$$

We define two operators A and B that act as follows on the system $H(\mathbb{R}^1)$ of Hermite functions:

$$A\varphi_n(x)=\sqrt{n+1}\,\varphi_{n+1}(x), \tag{2.11a}$$

$$B\varphi_n(x)=\sqrt{n}\,\varphi_{n-1}(x),\ n\in\mathbb{N}. \tag{2.11b}$$

These operators satisfy $(AB-BA)\varphi_n=\varphi_n$ and are related to the operators x and $\partial=d/dx$ as follows:

$$\partial\varphi_n=[(A+B)/\sqrt{2}]\varphi_n, \tag{2.12a}$$

$$x\varphi_n=[(-A+B)/\sqrt{2}]\varphi_n. \tag{2.12b}$$

By introducing the functions

$$\begin{aligned}\varphi_{\{n\}}(x)&=\varphi_{(n_0,n_1,n_2,n_3)}(x^0,x^1,x^2,x^3)\\&=\varphi_{n_0}(x^0)\varphi_{n_1}(x^1)\varphi_{n_2}(x^2)\varphi_{n_3}(x^3)\in\bigotimes_4 H(\mathbb{R}^1)\equiv H(\mathbb{R}^4)\end{aligned}$$

and the operators A^μ and B^μ defined as in (2.11a) and (2.11b) for $x=x^\mu$; $\mu=0,1,2,3$, we can write

$$\partial^\mu\varphi_{\{n\}}(x)=[(A^\mu+B^\mu)/\sqrt{2}]\varphi_{\{n\}}(x), \tag{2.13a}$$

$$(x^\mu\partial^\nu-x^\nu\partial^\mu)\varphi_{\{n\}}(x)=-(A^\mu B^\nu-A^\nu B^\mu)\varphi_{\{n\}}(x), \tag{2.13b}$$

where $\varphi_{\{n\}}(x)\in H(\mathbb{R}^4)$.

Now, by $A1$ the Hilbert space norm of the vectors $\Phi=\phi_{i_1}(\varphi_1)\ldots\phi_{i_k}(\varphi_k)$ in D_0' defines tempered distributions over $\otimes_{2k}\mathcal{S}(\mathbb{R}^4)$. The basis in $\otimes_{2k}\mathcal{S}(\mathbb{R}^4)$ shall be taken to be $\otimes_{2k}H(\mathbb{R}^4)$ and we need some estimate for these norms. The following lemma due to Simon[10] is useful.

Lemma 2.1: Let $T\in \mathcal{S}'(\mathbb{R}^k)$ and $\varphi_{n_i}(x)\in H(\mathbb{R}^1)$. We then have

$$|T(\varphi_{n_1},\ldots,\varphi_{n_k})|\leq c(1+n_1)^{p_1}\ldots(1+n_k)^{p_k}, \qquad (2.14)$$

for some $c\in\mathbb{R}_+$ and some $p_i: s\in\mathbb{N}$. For the norm of the vectors in D_0' we then have

Lemma 2.2: Let $\phi_{i_1}(\varphi_1)\ldots\phi_{i_k}(\varphi_k)\Omega\in D_0$. Then for $\varphi_{\{n_j\}}\in H(\mathbb{R}^4)$ and $1\leq j\leq k$ we have

$$\|\phi_{i_1}(\varphi_{\{n_j\}})\ldots\phi_{i_k}(\varphi_{\{n_k\}})\Omega\|\leq c(1+\{n_1\})^{\{p_1\}}\ldots(1+\{n_k\})^{\{p_k\}}, \qquad (2.15)$$

where

$$(1+\{n_j\})^{\{p_j\}}=\prod_{\mu=0}^{3}(1+n_{j,\mu})^{p_{j,\mu}} \text{ and } c\in\mathbb{R}_+.$$

Proof: We have

$$\|\phi_{i_1}(\varphi_1)\ldots\phi_{i_k}(\varphi_k)\Omega\| = |(\Omega,\phi^*_{i_k}(\tilde\varphi_k)\ldots\phi^*_{i_1}(\tilde\varphi_1)\phi_{i_1}(\varphi_1)\ldots\phi_{i_k}(\varphi_k)\Omega)|^{1/2}, \qquad (2.16)$$

where $\tilde\varphi_j$ is the complex conjugate of φ_j. From the symmetry properties of the right-hand side of (2.16), the reality of the $\varphi_{\{n_j\}}: s$ and by use of (2.14) we get (2.16).

QED

Next we are interested in test functions of the type $\partial^m_\mu\varphi_n$ and $(x_\mu\partial_\nu-x_\nu\partial_\mu)^m\varphi_n$ expressed in terms of the operators A_μ and B_μ. For tempered distributions over such test functions we have the following lemmas.

Lemma 2.3: Let $T\in\mathcal{S}'(\mathbb{R}^k)$. Then if $\varphi_{n_i}\in H(\mathbb{R}^1)$ we have

$$\begin{aligned}|T(\partial^{m_1}_{\mu_1}\varphi_{n_1},\ldots,\partial^{m_k}_{\mu_k}\varphi_{n_k})| &= (\sqrt2)^{-|m|}|T((A_{\mu_1}+B_{\mu_1})^{m_1}\varphi_{n_1},\ldots,(A_{\mu_k}+B_{\mu_k})^{m_k}\varphi_{n_k})| \\ &\leq(\sqrt2)^{|m|}c\prod_{j=1}^{k}\left(\frac{(n_j+m_j)!}{n_j!}\right)^{1/2}(1+n_j+m_j)^{p_j},\end{aligned} \qquad (2.17)$$

where $c\in\mathrm{R}_+$, $|m|=\Sigma m_j$, and $p_j\in\mathbb{N}$ for $1\leq j\leq k$.

Proof: Since T is separately linear in each variable we see from (2.14) that it is enough to prove (2.17) for $k=1$. The first equality follows from (2.13a). Thus we need to prove the relation

$$|T((A+B)^m\varphi_n)|\leq c\cdot 2^m[(n+m)!/n!]^{1/2}(1+n+m)^p. \qquad (2.18)$$

Now $(A+B)^m\varphi_n$ can be expanded in a noncommutative binomial series, each term of which has the form $C\varphi_n$, where C is a monomial of degree m in A and B. By using the triangle inequality after insertion of the expansion into (2.18) we see that we have to estimate the terms $|T(C\varphi_n)|$. From (2.14) we get $|T(\varphi_n)|\leq|T(\varphi_m)|$ for $n\leq m$. By using this and (2.11a) and (2.11b) we find that $|T(BC\varphi_n)|\leq|T(AC\varphi_n)|$ for any monomial C, since $C\varphi_n=f_n\varphi_n$, with $f_n\geq 0$. Thus if C is a monomial of degree m we find by induction

$$|T(C\varphi_n)|\leq|T(A^m\varphi_n)|\leq c[(n+m)!/n!]^{1/2}(1+n+m)^p. \qquad (2.19)$$

Since there are 2^m terms in the expansion the estimate (2.18) follows. QED

Lemma 2.3 can be combined with Lemma 2.2 to give

Lemma 2.4: The following estimate holds:

$$\begin{aligned}&\|\phi_{i_1}((A_\mu+B_\mu)^{m_1}\varphi_{\{n_1\}})\ldots\phi_{i_k}((A_\mu+B_\mu)^{m_k}\varphi_{\{n_k\}})\Omega\| \\ &\leq c\cdot 2^{|m|}\left(\frac{(n_{1,\mu}+m_1)!\ldots(n_{k,\mu}+m_k)!}{n_{1,\mu}!\ldots n_{k,\mu}!}\right)^{1/2} \\ &\times(1+[m_j]_{\max}+[n_{j,\mu}]_{\max})^{k[p_j]_{\max}},\end{aligned} \qquad (2.20)$$

where $|m|=\Sigma m_j$, $\mu=0,1,2,3$, and $[m_j]_{\max}$ is the maximum of the $m_i: s$ and $[n_{j,\mu}]_{\max}$ is the maximum of the $n_{j,\mu}: s$.

Entirely similar techniques, especially estimates of the type in (2.19) are used to prove

Lemma 2.5: The following estimate holds:

$$\begin{aligned}&\|\phi_{i_1}[(A_\nu B_\mu+A_\mu B_\nu)^{m_1}\varphi_{\{n_1\}}] \\ &\times\ldots\phi_{i_k}[(A_\nu B_\mu+A_\mu B_\nu)^{m_k}\varphi_{\{n_k\}}]\Omega\|\leq c\cdot 2^{|m|} \\ &\times\left(\frac{(n_{1,\mu}+m_1)!\ldots(n_{k,\mu}+m_k)!(n_{1,\nu}+m_1)!\ldots(n_{k,\nu}+m_k)!}{n_{1,\mu}!\ldots n_{k,\mu}!n_{1,\nu}!\ldots n_{k,\nu}!}\right)^{1/2} \\ &\times(1+[n_{j,\mu}]_{\max}+[m_j]_{\max})^{k[p_{j,\mu}]_{\max}} \\ &\times(1+[n_{j,\nu}]_{\max}+[m_j]_{\max})^{k[p_{j,\nu}]_{\max}}.\end{aligned} \qquad (2.21)$$

C. Analytic vectors for one-parameter subgroups of $U(\Lambda,a)$

On the basis of the estimates in Sec. 2B we now show that the set of vectors in D_0' defined in (2.7) are analytic vectors for the unitary one-parameter groups of $U(\Lambda,a)$ corresponding to the generators P^μ and $M^{\mu\nu}$.

Theorem 1: Given a Wightman field theory with a denumerable set of field operators over the test function space $\mathcal{S}(\mathbb{R}^4)$, satisfying axioms $A\,0$—$A\,5$ and transforming as scalar fields under a continuous unitary representation $U(\Lambda,a)$ of the universal covering group of the restricted Poincaré group in the Hilbert space $\mathcal{H}$, then the domain D_0 defined in $A\,5$ contains a dense invariant domain of analytic vectors for the unitary one-parameter groups of $U(\Lambda,a)$ corresponding to P^μ and $M^{\mu\nu}$, where P^μ and $M^{\mu\nu}$ are the usual infinitesimal generators of $U(\Lambda,a)$.

Proof: I. We start by showing that the vectors in D_0' are analytic vectors for the translation subgroup $U(1,a)$. Since the generators P^μ are defined and commute on D_0 it is sufficient to consider $\mu=1$, say. Let us first treat the simplest nontrivial case of a vector $\phi_j(\varphi_{\{n\}})\Omega\in D_0'$. If $a_1\in\mathbb{R}^1$ we have, using (2.3)

$$\begin{aligned}\left\|\sum_{m=0}^{\infty}\frac{1}{m!}(ia_1)^m P_1^m\phi_j(\varphi_{\{n\}})\Omega\right\| &\leq\sum_{m=0}^{\infty}\frac{1}{m!}|a_1|^m\|\phi_j(\partial_1^m\varphi_{\{n\}})\Omega\| \qquad (2.22)\\ &=\sum_{m=0}^{\infty}\frac{1}{m!}|a_1|^m(\sqrt2)^{-m}\|\phi_j[(A_1+B_1)^m\varphi_{\{n\}}]\Omega\| \\ &\leq c\sum_{m=0}^{\infty}\frac{1}{m!}|\sqrt2\,a_1|^m\left(\frac{(n_1+m)!}{n_1!}\right)^{1/2}(1+n_1+m)^{p_1},\end{aligned}$$

where the last inequality is given by Lemma 2.4 for $k=1$. The power series converges by the root test, since for $m>n_1$ the coefficient of the mth power does not grow faster than

$$\frac{1}{m!}\left(\frac{(n_1+m)!}{n_1!}\right)^{1/2}(2m)^{p_1}\leq\frac{1}{\sqrt{n_1!}}(\sqrt2)^{n_1+2p_1}\frac{(\sqrt m)^{n_1+2p_1}}{\sqrt{m!}}\quad(m>n_1).$$

The case of an arbitrary vector $\phi_{i_1}(\varphi_{\{n_1\}})\ldots\phi_{i_k}(\varphi_{\{n_k\}})\Omega \in D_0'$ is only slightly more complicated. Using Lemma 2.4 we have

$$\|\sum_{m=0}\frac{1}{m!}(ia_1)^m P_1^m \phi_{i_1}(\varphi_{\{n_1\}})\cdots\phi_{i_k}(\varphi_{\{n_k\}})\Omega\|$$
$$=\sum_{m=0}\frac{1}{m!}|a_1|^m k^m \max_{\substack{0\le j_1\le\cdots\le j_k\le m\\ j_1+\cdots+j_k=m}}$$
$$\times\{\|\phi_{i_1}(\partial_1^{j_1}\varphi_{\{n_1\}})\ldots\phi_{i_k}(\partial_1^{j_k}\varphi_{\{n_k\}})\Omega\|\} \tag{2.23}$$
$$\le c\sum_{m=0}\frac{1}{m!}|\sqrt{2}\,ka_1|^m\left(\frac{([n_{j,1}]_{\max}+m)!}{[n_{j,1}]_{\max}!}\right)^{1/2}$$
$$\times(1+[n_{j,1}]_{\max}+m)^{k[p_{j,1}]_{\max}}.$$

For $m>[n_{j,1}]_{\max}$ we have for the coefficient of the mth power:

$$\left(\frac{([n_{j,1}]_{\max}+m)!}{[n_{j,1}]_{\max}!}\right)^{1/2}\frac{(1+[n_{j,1}]_{\max}+m)^{k[p_{j,1}]_{\max}}}{m!} \tag{2.24}$$
$$\le([n_{j,1}]_{\max}!)^{-1/2}2^{k[p_{j,1}]_{\max}+[n_{j,1}]_{\max}/2}\frac{m^{k[p_{j,1}]_{\max}+[n_{j,1}]_{\max}/2}}{\sqrt{m!}},$$

which again shows by the root test that the series converges. This concludes the proof for the translations.

II. Next we treat the subgroups corresponding to rotations (with generators $M_{ij}, i,j=1,2,3.$) and Lorentz transformations (with generators $M_{0l}, l=1,2,3.$) simultaneously. From Eqs. (2.4) and (2.13b) we see that the same formalism and the same estimates will hold for both cases. Let us choose for definiteness M_{12}. Again we start with the vector $\phi_j(\varphi_{\{n\}})\Omega\in D_0'$. With $\theta\in\mathbb{R}^1$ we have

$$\|\sum_{m=0}^{\infty}\frac{1}{m!}(i\theta)^m M_{12}^m\phi_j(\varphi_{\{n\}}\Omega\|$$
$$\le\sum_{m=0}^{\infty}\frac{1}{m!}|\theta|^m\|\phi_j((A_1B_2-A_2B_1)^m\varphi_{\{n\}}\Omega\| \tag{2.25}$$
$$\le c\cdot\sum_{m=0}^{\infty}\frac{1}{m!}|2\theta|^m\left(\frac{(n_1+m)!(n_2+m)!}{n_1!n_2!}\right)^{1/2}$$
$$\times(1+n_1+m)^{p_1}(1+n_2+m)^{p_2}.$$

The last line follows by use of Lemma 2.5. For $m>\max(n_1,n_2)$ the coefficient of the mth power behaves as

$$\frac{1}{m!}\left(\frac{(n_1+m)!(n_2+m)!}{n_1!n_2!}\right)^{1/2}(1+n_1+m)^{p_1}(1+n_2+m)^{p_2}$$
$$\le 2^{p_1+p_2+(n_1+n_2)/2}m^{p_1+p_2+(n_1+n_2)/2}(n_1!n_2!)^{-1/2}.$$

By the root test, this shows that the series converges for $|\theta|<\frac{1}{2}$.

Next, consider the series

$$\left\|\sum_{m=0}\frac{1}{m!}(i\theta)^m M_{12}^m\phi_{j_1}(\varphi_{\{n_1\}})\ldots\phi_{j_k}(\varphi_{\{n_k\}})\Omega\right\|$$
$$\le c'\cdot\sum_{m=0}^{\infty}\frac{1}{m!}(2|\theta|k)^m\left(\frac{([n_{i,1}]_{\max}+m)!([n_{i,2}]_{\max}+m)!}{[n_{i,1}]_{\max}![n_{i,2}]_{\max}!}\right)^{1/2}$$
$$\times(1+[n_{i,1}]_{\max}+m)^{k[p_{i,1}]_{\max}}(1+[n_{i,2}]_{\max}+m)^{k[p_{i,2}]_{\max}}$$
$$\lesssim c''\sum_{m=0}^{\infty}|2\theta k|^m m^{k([p_{i,1}]_{\max}+[p_{i,2}]_{\max})+1/2([n_{i,1}]_{\max}+[n_{i,2}]_{\max})}.$$
$$\times 2^{k([p_{i,1}]_{\max}+[p_{i,2}]_{\max})}, \tag{2.26}$$

which converges for $|\theta|<1/(2k)$.

Hence D_0' consists of analytic vectors for the unitary one-parameter subgroups corresponding to P^μ and $M^{\mu\nu}$ and the same holds for the linear hull of D_0' which is contained in D_0. Furthermore the set $D_0^A=$ linear hull of $\{U(\Lambda,a)\psi;\ (\Lambda,a)\in P_+^\uparrow$ and $\psi\in D_0'\}$ is an invariant dense domain of analytic vectors for P^μ and $M^{\mu\nu}$ contained in D_0. QED

D. Analytic vectors for $U(\Lambda,a)$

The following stronger version of Theorem 1 holds.

Theorem 1′: Under the same assumptions as in Theorem 1, D_0 contains a dense invariant set of analytic vectors for $U(\Lambda,a)$.

We shall prove that the vectors in D_0' are analytic for $U(\Lambda,a)$. For this proof we need four lemmas.

Lemma 2.6 (Goodman[11]): Let the Lie group $\mathcal{G}$ be a semidirect product of the two subgroups $\mathcal{G}_1$ and $\mathcal{G}_2$ and let $U(\mathcal{G})$ be a unitary representation of $\mathcal{G}$ in the Hilbert space $\mathcal{H}$. If $U(\mathcal{G}_1)$ and $U(\mathcal{G}_2)$ are the restrictions of $U(\mathcal{G})$ to $\mathcal{G}_1$ and $\mathcal{G}_2$, respectively, then

$$D^\omega(U(\mathcal{G}))=D^\omega(U(\mathcal{G}_1))\cap D^\omega(U(\mathcal{G}_2)).$$

This is a special case of *Lemma* 2.7 (Flato, Simon[12]): Let $\mathcal{G}$ be a Lie group with Lie algebra g and let $\mathcal{G}_1$ and $\mathcal{G}_2$ be two subgroups of $\mathcal{G}$ with Lie algebras g_1 and g_2 such that g is a unification of g_1 and g_2. Let $U(\mathcal{G})$ be a unitary representation of $\mathcal{G}$ in the Hilbert space $\mathcal{H}$. If $U(\mathcal{G}_1)$ and $U(\mathcal{G}_2)$ are the restrictions of $U(\mathcal{G})$ to $\mathcal{G}_1$ and $\mathcal{G}_2$, respectively, and if $D^\omega(U(\mathcal{G}_1))\cap D^\omega(U(\mathcal{G}_2))$ is dense in $\mathcal{H}$, then

$$D^\omega(U(\mathcal{G}))=D^\omega(U(\mathcal{G}_1))\cap D^\omega(U(\mathcal{G}_2)).$$

Lemma 2.8 (Flato, Simon[12]): Let $\mathcal{G}$ be a Lie group with Lie algebra g and let $U(\mathcal{G})$ be a unitary representation of $\mathcal{G}$ in $\mathcal{H}$. Let $\{x_1,\ldots,x_n\}$ be an arbitrary basis for g and let $U(G_i(t))$, $t\in\mathbb{R}$ be the unitary representations of the one-parameter subgroups e^{tx_i}, $i=1,\ldots,n$. Let $D^\omega(U(G_i(t)))$ be a dense invariant domain of analytic vectors for $U(G_i(t))$. Then if $\mathcal{G}$ is compact

$$D^\omega(U(\mathcal{G}))=\bigcap_{i=1}^{n}D^\omega(U(G_i(t))).$$

Lemma 2.9 (Goodman[11]): Let $\mathcal{G}$ be a solvable Lie group with Lie algebra g. Let $\{x_j\}_{j=1}^d$ be a Jordan basis for g, i.e., a basis such that $\mathfrak{L}_i=\mathrm{span}\{x_j\}_{j\le i}$ is a subalgebra of g with $\mathfrak{L}_i$ an ideal in $\mathfrak{L}_{i+1}$. Let $U(\mathcal{G})$ be a unitary representation of $\mathcal{G}$ in the Hilbert space $\mathcal{H}$ and let $U(G_j(t))=U(e^{tx_j})$, $t\in\mathbb{R}$ be the representations of the unitary one-parameter groups e^{tx_j} of $\mathcal{G}$ for $j=1,\ldots d$. Then

$$D^\omega(U(\mathcal{G}))=\bigcap_{j=1}^{d}D^\omega(U(G_j(t))).$$

Proof of Theorem 1′: The representation $U(\Lambda,a)$ of the universal covering group $T^4\otimes SL(2,\mathbb{C})$ of the Poincaré group in $\mathcal{H}$ is composed into

$$U(\Lambda,a)=U(T^4)\otimes U(SL(2,\mathbb{C})).$$

The representation $U(SL(2,\mathbb{C}))$ is further decomposed according to Iwasawa in

$$U(SL(2,\mathbb{C}))=U(K)U(R),$$

where K is a compact Lie group and R is a solvable one. The generators of $U(K)$ are M_{23}, M_{31}, and M_{12} and those of $U(R)$ are M_{03}, $M_{23}+M_{02}$, and $-(M_{13}+M_{01})$, which is easily seen to be a Jordan basis for the Lie algebra of

$U(R)$. Now, the domain D_0^A is a dense invariant set of analytic vectors for P^μ and $M^{\mu\nu}$ according to Theorem 1. Let us assume for a moment that the vectors in D_0' are also analytic for the one-parameter groups corresponding to $M_{23}+M_{02}$ and $-(M_{13}+M_{01})$. Then by Lemma 2.8 and Lemma 2.9 the domain D_0^A is also a common dense invariant domain of analytic vectors for $U(K)$ and $U(R)$. By Lemma 2.7 the domain D_0^A is therefore a dense invariant domain of analytic vectors for $U(SL(2, \mathrm{C}))$, and finally, by Lemma 2.6 and the fact that the translation group is Abelian, it is a dense invariant domain of analytic vectors for $U(\Lambda, a)$.

It remains to show that the vectors in D_0' are analytic for the unitary one-parameter groups corresponding to $M_{23}+M_{02}$ and $-(M_{13}+M_{01})$.

I. We consider first $M_{23}+M_{02}$. The vector $\phi_j(\varphi_{\{n\}})\Omega \in D_0'$ and with $b\in \mathbb{R}$ we have

$$\Big\|\sum_{m=0}^{\infty}\frac{1}{m!}(ib)^m(M_{23}+M_{02})^m\phi_j(\varphi_{\{n\}})\Omega\Big\|$$

$$\leq \sum_{m=0}^{\infty}\frac{1}{m!}|b|^m\|\phi_j((x_2\partial_3-x_3\partial_2+x_0\partial_2-x_2\partial_0)^m\varphi_{\{n\}})\Omega\|$$

$$=\sum_{m=0}^{\infty}\frac{1}{m!}|b|^m\|\phi_j((A_3B_2-A_2B_3+A_2B_0-A_0B_2)^m\varphi_{\{n\}})\Omega\|.$$

Now $\|\phi_j([(A_3-A_0)B_2+A_2(B_0-B_3)]^m\varphi_{\{n\}})\Omega\|$ is expanded into a series of terms containing A_i : s and B_i : s. After application of the triangle inequality every term in this series is majorized by the corresponding term where any B_i is replaced by a corresponding A_i, $i=0,2,3$. Thus

$$\|\phi_j([(A_3-A_0)B_2+A_3(B_0-B_3)]^m\varphi_{\{n\}})\Omega\|$$

$$\leq 2^m\left(\frac{(n_2+m)!}{n_2!}\right)^{1/2}\sum_{k=0}^{m}\binom{m}{k}\left(\frac{(n_0+k)!(n_3+m-k)!}{n_0!n_3!}\right)^{1/2}$$

$$\times\|\phi_j(\varphi_{(n_0+k,n_1,n_2+m,n_3+m-k)})\Omega\|.$$

Now, let $\bar n=\max(n_0,n_3)$. Since $m\geq k$, we have

$$\left(\frac{(n_0+k)!(n_3+m-k)!}{n_0!n_3!}\right)^{1/2}\leq\left(\frac{(\bar n+k)!(\bar n+m-k)!}{\bar n!\bar n!}\right)^{1/2}$$

$$=\left(\frac{(\bar n+m)!}{\bar n!}\right)^{1/2}\left(\frac{(\bar n+k)!(\bar n+m-k)!}{\bar n!(\bar n+m)!}\right)^{1/2}\leq\left(\frac{(\bar n+m)!}{\bar n!}\right)^{1/2}.$$

Thus

$$\|\phi_j([(A_3-A_0)B_2+A_3(B_0-B_3)]^m\varphi_{\{n\}})\Omega\|$$

$$\leq c\cdot 4^m\frac{(n+m)!}{n!}(1+n+m)^{3p},$$

where $n=\max(n_2,\bar n)$ and $p=\max(p_0,p_2,p_3)\in \mathbb{N}$.

From this follows

$$\Big\|\sum_{m=0}^{\infty}\frac{1}{m!}(ib)^m(M_{23}+M_{02})^m\phi_j(\varphi_{\{n\}})\Omega\Big\|$$

$$\leq c\sum_{m=0}^{\infty}\frac{|4b|^m}{m!}\frac{(n+m)!}{n!}(1+n+m)^{3p}<\infty\quad\text{for } b<1/4.$$

Next we consider an arbitrary vector $\phi_{i_1}(\varphi_{\{n_{i_1}\}})\cdots\phi_{i_k}(\varphi_{\{n_{i_k}\}})\Omega\in D_0'$. Let $b\in\mathbb{R}$. We then have

$$\Big\|\sum_{m=0}^{\infty}\frac{(ib)^m}{m!}(M_{23}+M_{02})^m\phi_{i_1}(\varphi_{\{n_{i_1}\}})\ldots\phi_{i_k}(\varphi_{\{n_{i_k}\}})\Omega\Big\|\leq\sum_{m=0}^{\infty}\frac{|b|^m}{m!}k^m$$

$$\times\max_{\substack{0\leq j_1\leq\cdots j_k\leq m\\ j_1+\cdots+j_k=m}}\{\|\phi_{i_1}((A_3B_2-A_2B_3+A_2B_0-A_0B_2)^{j_1}\varphi_{\{n_{i_1}\}})$$

$$\times\cdots\phi_{i_k}((A_3B_2-A_2B_3+A_2B_0-A_0B_2)^{j_k}\varphi_{\{n_{i_k}\}})\Omega\|\}$$

$$\leq c\sum_{m=0}^{\infty}\frac{|4bk|^m}{m!}\max_{\substack{0\leq j_1\leq\cdots j_k\leq m\\ j_1+\cdots+j_k=m}}\left(\frac{(n_{i_1}+j_1)!\ldots(n_{i_k}+j_k)!}{n_{i_1}!\ldots n_{i_k}!}\right.$$

$$\left.\times(1+n_{i_1}+j_1)^{3p_{i_1}}\ldots(1+n_{i_k}+j_k)^{3p_{i_k}}\right)$$

$$\leq c\sum_{m=0}^{\infty}\frac{|4bk|^m}{m!}\frac{(n+m)!}{n!}(1+n+m)^{3kp}<\infty,$$

for $b<1/(4k)$, where $n=\max_j(n_{i_j})$ and $n_{i_j}=\max(n_{i_j,0},n_{i_j,2},n_{i_j,3})$ and $p=\max_j(p_{i_j})$ with $p_{i_j}=\max(p_{i_j,0},p_{i_j,2},p_{i_j,3})$.

This concludes the proof that D_0' consists of analytic vectors for the unitary one-parameter group corresponding to $M_{23}+M_{02}$.

II. From the preceding calculation it is immediately clear that the vectors in D_0' are also analytic for the unitary one-parameter group corresponding to $-(M_{13}+M_{01})$ the only difference in the proof is that we have to exchange the index 2 by 1 everywhere. Thus the proof of Theorem 1′ is complete. QED

In both Theorem 1 and Theorem 1′ we have for simplicity considered spinless field operators. It is evident, however, that spin degrees of freedom can be introduced as in $A4$ without changing the validity of the theorems, since the spin generators act as index transformations on the fields.

Remark

The existence of a dense invariant subset of analytic vectors for the translations $U(\mathbb{1},a)$ in D_0 can also be proved directly by using Eq. (2.2) with $\Lambda=\mathbb{1}$. Since the mapping from $\otimes_k\mathcal{S}(\mathbb{R}^4)$ to $\mathcal{H}$ is strongly continuous, the vector $\phi_{i_1}(f_1)\ldots\phi_{i_k}(f_k)\Omega\in D_0$ is analytic for $U(\mathbb{1},a)$ if the $f_j(x)$:s are nice real-analytic functions. As is well known there is in $\mathcal{S}(\mathbb{R}^4)$ a dense subclass of analytic functions which provides D_0 with a dense set of analytic vectors for $U(1,a)$ in this case. This argument suggests that when the test function space is $\mathcal{D}(\mathrm{R}^4)$, then the domain D_0 does not contain a dense subset of analytic vectors for $U(\mathbb{1},a)$.

3. DISCUSSION

The result that the quasilocal domain D_0 contains an invariant dense set of analytic vectors for the unitary representation $U(\Lambda,a)$ of $T^4\otimes SL(2,\mathbb{C})$ first shows that the Wightman axioms $A0$–$A5$, especially the cyclicity of Ω in $A5$, are rich enough to give a concrete domain on which free passage between the Poincaré algebra representation and the Poincaré group representation is possible. Secondly it might eventually permit a distinction between theories where the test function space is $\mathcal{S}(\mathbb{R}^4)$ and where it is $\mathcal{D}(\mathbb{R}^4)$. Thirdly it indicates possible departures for generalizations of the axioms to more general physical situations. One possible type of generalization is to consider a Poincaré algebra covariant quantum field theory defined by replacing axioms $A2$, $A3$, and $A4$ by

$A\,2'$. There exists a (continuous) representation $\mathcal{T}(\mathrm{p})$ of the Poincaré algebra $\mathrm{p}=t^4 \otimes sl(2,\mathbb{C})$ on D such that

$$\mathcal{T}(\mathrm{p})D\subset D. \tag{3.1}$$

$A\,3'$. There is one unique state Ω (up to normalization and a phase factor) $\in \mathcal{H}$ and satisfying

$$P^\mu\Omega=0, \tag{3.2}$$

$$M^{\mu\nu}\Omega=0, \tag{3.3}$$

where $\{iP^\mu, iM^{\mu\nu}\}$ is the usual basis in $\mathcal{T}(\mathrm{p})$. $\Omega\in D$ and for any vector $\Psi\in D$, $(\Omega,\Psi)=0$ we have $(\Psi,P_0\Psi)\geq \mu_0\|\Psi\|^2$, $\mu_0>0$.

$A\,4'$. The fields ϕ_i transform covariantly among themselves under commutation with operators from $\mathcal{T}(\mathrm{p})$, i.e., for $\Psi\in D$ we have

$$[P^\mu,\phi_k(\varphi)]\Psi=\phi_k(i\partial^\mu\varphi)\Psi, \tag{3.4}$$

$$[M^{\mu\nu},\phi_k(\varphi)]\Psi=\sum_{k'=0}^{\infty} T^{\mu\nu}_{kk'}\phi_{k'}(\varphi)\Psi+\phi_k(i(x^\mu\partial^\nu-x^\nu\partial^\mu)\varphi)\Psi, \tag{3.5}$$

where $k\in I$ and $T^{\mu\nu}_{kk'}$ are matrix elements of a representation of $sl(2,\mathbb{C})$.

If for D we chose D_0, since this is the concretely given domain, then the following remarks apply.

For spin zero fields over $\mathcal{S}(\mathrm{R}^4)$, the calculations in Sec. 2C show that D_0' is a set of analytic vectors for the basis $\{iP^\mu, iM^{\mu\nu}\}$ of $\mathcal{T}(\mathrm{p})$ and since the linear hull of D_0' is dense in $\mathcal{H}$, $\mathcal{T}(\mathrm{p})$ is integrable.[13] This is actually also true when the test function space is $\mathcal{D}(\mathbb{R}^4)$, since we can always define a unitary representation $U(\Lambda,a)$ of $T^4 \otimes SL(2,\mathbb{C})$ transforming the fields as in (2.1) and giving rise to a representation $dU(\mathrm{p})$ of p which coincides with $\mathcal{T}(\mathrm{p})$ on D_0. The generalization is therefore trivial in this case.

For fields with spin, the situation is more complex. The requirement $\mathcal{T}(\mathrm{p})D_0\subset D_0$, however, puts strong restrictions on the representation T of $sl(2,\mathbb{C})$. In general this condition can be always fulfilled only if the representation T of $sl(2,\mathbb{C})$ is decomposable into a direct sum of finite-dimensional irreducible representations, since an infinite sum in (3.5) will take vectors in D_0 out of D_0. If the representation $\mathcal{T}(\mathrm{p})$ is symmetric then it is in this case integrable to a unitary representation of $T^4 \otimes SL(2,\mathbb{C})$, since the definition in (2.1) is exactly the one required.

When $\mathcal{T}(\mathrm{p})$ is not required to be symmetric on D_0, then nontrivial generalizations exist. This is of course also true when D is a more general domain than D_0.

ACKNOWLEDGMENTS

It is a pleasure to thank Professor M. Flato for numerous discussions and suggestions concerning various parts of this work. I also want to thank Dr. J. Simon and Dr. D. Sternheimer for helpful comments and discussions. This work has been done with financial support from the French Centre National de la Recherche Scientifique which is gratefully acknowledged.

*Present address: Department of Theoretical Physics, Royal Institute of Technology, S-10044 Stockholm, Sweden.

[1]L. Gårding, Proc. Natl. Acad. Sci. USA **33**, 331 (1947). F. Bruhat, Bull. Soc. Math. Fr. **84**, 97 (1956).
[2]Harish-Chandra, Trans. Am. Math. Soc. **75**, 185 (1953); P. Cartier and J. Dixmier, Am. J. Math. **80**, 131 (1958); E. Nelson, Ann. Math. **81**, 547 (1959); L. Gårding, Bull. Soc. Math. Fr. **88**, 73 (1960).
[3]R. F. Streater and A. S. Wightman, *PCT, Spin & Statistics, and All That* (Benjamin, New York, 1964).
[4]R. Jost, *The General Theory of Quantized Fields* (AMS, Providence, R. I., 1965).
[5]It is understood that the fields are indexed so that only fields in finite subsets transform among themselves.
[6]L. Schwartz, Ann. Inst. Fourier **7**, 1 (1957).
[7]L. Schwartz, *Théorie des distribution* (Hermann, Paris, 1957 and 1959), Vols. I. and II.
[8]A. Jaffe, Phys. Rev. **158**, 1454 (1967).
[9]F. Riesz and B. Sz.-Nagy, *Vorlesungen über Funktional-analysis* (Deutscher Verlag der Wissenschaften, Berlin, 1965).
[10]B. Simon, J. Math. Phys. **12**, 140 (1971).
[11]R. Goodman, Trans. Am. Math. Soc. **143**, 55 (1969).
[12]M. Flato and J. Simon (to be published).
[13]M. Flato, J. Simon, H. Snellman, and D. Sternheimer, Annales Sc. de l'Ecole Normale Superieure 5, 423 (1972).

Factorizable and infinitely divisible PUA representations of locally compact groups

B. J. Falkowski

Department of Mathematics, The University, Oxford Road, Manchester, England

(Received 9 November 1972)

Factorizable and infinitely divisible projective unitary–antiunitary representations of locally compact groups are analyzed in terms of their "expectation values." A connection between these and first order cocycles is established. As a consequence a very general analog of the Araki–Woods embedding theorem is proved.

INTRODUCTION

In this paper we are concerned with factorizable and infinitely divisible projective unitary—antiunitary (PUA) representations. These are of great interest in quantum field theory. The former ones arise as the simplest representations of "current groups." The latter ones are closely linked to continuous tensor products. Factorizable representations were first discussed by Araki.[1] While the connection of infinitely divisible representations and continuous tensor products was exhibited by Streater.[2] Then Parthasarathy and Schmidt gave an almost complete analysis of both kinds of representations.[3] Here we give an extension of the result in Ref. 3 to the case where antiunitary operators may occur in the representation. This is of interest since with a sufficiently large symmetry group in quantum mechanics (e.g., the extended Poincaré group) antiunitary operators occur in the representation. The methods of proof used in Ref. 3 apply for most of our theorems with more or sometimes less trivial modifications. The main difficulty consisted of finding suitable "positive definite Kernels" (as opposed to positive definite functions) to describe our representations. In Sec. 1 we exhibit a correspondence between "UA-Araki multipliers" and "Araki-s functions," on the one hand, and unitary—antiunitary (UA) representations and their associated first order cocycles in a Hilbert space, on the other hand. In Sec. 2 we make a detailed analysis of "factorizable PUA representations." In Sec. 4 we make use of the results in Secs. 1 and 2 to redefine the "UA current group" and prove that even under these most general circumstances an analogue of the "Araki—Woods embedding theorem" still holds. That is, we give an explicit construction which allows us to embed a "factorizable PUA representation" in a symmetric Fock space over a Hilbert space.

Preliminaries: Given a locally compact group G, we consider a projective unitary—antiunitary representation (PUA representation) in a separable complex Hilbert space H. If $G=G^+\cup G^-$ is the associated UA-decomposition (for details see Ref. 4) where $G^+=\{g\in G: g$ is mapped into a unitary operator under the representation$\}$, $G^-=\{g\in G: g$ is mapped into an antiunitary operator$\}$, then we have the following. Our representation is a map $G\to$ unitary/antiunitary operators in H sending $g\to U_g$ which is measurable w.r.t. suitable Borel structures (see Ref. 4) together with a map $\sigma: G\times G\to S^1$ (the complex unit circle) which is also measurable (again in a suitable sense) such that the following holds:

(i) $U_{g_1}U_{g_2}=\sigma(g_1,g_2)U_{g_1g_2}\ \forall g_1,g_2\in G$,

(ii) $\sigma(g_1,g_2)\sigma(g_1g_2,g_3)=\begin{cases}\sigma(g_1,g_2g_3)\sigma(g_2,g_3), & g_1\in G^+,\\ \sigma(g_1,g_2g_3)\overline{\sigma}(g_2,g_3), & g_1\in G^-,\end{cases}$

(iii) $\sigma(e,g)=\sigma(g,e)=1\quad \forall g\in G$

σ is then called a *UA multiplier* and is in fact a second order cocycle.

An *additive UA multiplier* is then a function $s: G\times G\to\mathbb{R}$ which is measurable (or continuous) and satisfies

(i) $s(g_1,g_2)+s(g_1g_2,g_3)=\begin{cases}s(g_1,g_2g_3)+s(g_2,g_3), & g_1\in G^+,\\ s(g_1,g_2g_3)-s(g_2,g_3), & g_1\in G^-.\end{cases}$

(ii) $s(e,g)=s(g,e)=0\quad\forall g\in G.$

Note: $\exp(is)$ is then an ordinary UA multiplier. We shall also need the notion of positive definite (conditionally positive definite) Kernels:

A function K from $X\times X\to\mathbb{C}$, where X is a topological space (Borel space) is called *positive definite* if

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\overline{\alpha}_jK(x_i,x_j)\geq 0$$

$\forall$ choices of points $x_1,\cdots x_n\in X$ and arbitrary complex numbers $\alpha_1,\ldots,\alpha_n$. It is *conditionally positive definite* if this inequality holds whenever $\sum_{i=1}^n\alpha_i=0$. Usually our Kernels will be continuous in the product topology. This then leads to the concept of a "conditionally s-positive definite function":

Given a continuous function $\varphi: G\to\mathbb{C}$ and an additive UA multiplier s, which is continuous, we say φ is "conditionally s-positive definite" if the following holds:

(i) $\varphi(g^{-1})=\begin{cases}\overline{\varphi}(g), & g\in G^+,\\ \varphi(g), & g\in G^-,\end{cases}\quad \varphi(e)=0$

($e=$ identity in the group),

(ii) $s(g,h)=\begin{cases}-s(h^{-1},g^{-1}) & (g,h)\in G^+\times G^+\cup G^-\times G^-,\\ s(h^{-1},g^{-1}) & (g,h)\in G^+\times G^-\cup G^-\times G^+,\end{cases}$

(iii) the Kernel

$$K(g,h)=\begin{cases}\varphi(h^{-1}g)+is(h^{-1},g), & h\in G^+,\\ \overline{\varphi}(h^{-1}g)-is(h^{-1},g), & h\in G^-,\end{cases}$$

is conditionally positive definite.

A kernel K is called Hermitian if $K(x,y)=\overline{K}(y,x)$, $\forall x,y\in X$. From (iii) and Lemma (2.2) in Ref. 5 it then follows immediately that the following Kernel is positive definite and Hermitian and satisfies $\rho_\varphi(e,e)=0$:

$$\rho_\varphi(g,h)=\begin{cases}\varphi(h^{-1}g)-\varphi(g)-\varphi(h^{-1})+is(h^{-1},g), & h\in G^+,\\ \overline{\varphi}(h^{-1}g)-\varphi(g)-\overline{\varphi}(h^{-1})-is(h^{-1},g), & h\in G^-,\end{cases}$$

for a motivation and further details of these preliminaries see Refs. 3 and 5.

I. MULTIPLIER-VALUED MEASURES, CONDITIONALLY S-POSITIVE FUNCTION-VALUED MEASURES AND COCYCLES

In order to analyze "factorizable PUA representations" we have to study additive-UA-multiplier-valued measures and conditionally s-positive function-valued measures.

We introduce the following definition.

Definition (1.1): Let $(T, \mathcal{S})$ be a standard Borel space. A function $s: \mathcal{S}\times G\times G \to \mathbb{R}$ is called a *UA-Araki multiplier* (cf. Ref. 6) if the following conditions hold:

(i) For every fixed $(g_1, g_2) \in G\times G$, $s(\cdot, g_1, g_2)$ is a totally finite signed measure in $\mathcal{S}$.

(ii) For every fixed $A \in \mathcal{S}$, the function $s(A, \cdot, \cdot)$ is an additive UA multiplier on $G\times G$.

For a given UA-Araki multiplier s, a function $\varphi: \mathcal{S}\times G \to \mathbb{C}$ is called an Araki s function if the following conditions hold:

(i) For every fixed $g \in G$, $\varphi(\cdot, g)$ is a totally finite complex measure on $\mathcal{S}$.

(ii) For every fixed $A \in \mathcal{S}$, $\varphi(A, \cdot)$ is a conditionally $s(A, \cdot, \cdot)$-positive function on G.

[All multipliers are defined w.r.t. the same normal subgroup G^+ (cf. Ref. 4) of G.]

We choose and fix a standard Borel space $(T, \mathcal{S})$ and a pair (s, φ) of a UA-Araki multiplier and an Araki s function. We now define a kernel K_φ on the space $(\mathcal{S}\times G)\times(\mathcal{S}\times G)$ by the following equation:

$$K_\varphi(A, g; B, h) = \begin{cases} \varphi(A\cap B, h^{-1}g) - \varphi(A\cap B, g) - \varphi(A\cap B, h^{-1}) + is(A\cap B, h^{-1}, g), & h\in G^+, \\ \overline{\varphi}(A\cap B, h^{-1}g) - \varphi(A\cap B, g) - \overline{\varphi}(A\cap B, h^{-1}) - is(A\cap B, h^{-1}, g), & h\in G^-. \end{cases}$$

Lemma (1.2): The kernel K_φ defined above is positive definite in the space $(\mathcal{S}\times G)\times(\mathcal{S}\times G)$.

Proof: Let $A_1, A_2, \ldots, A_n \in \mathcal{S}$, $g_1, g_2, \ldots, g_n \in G$ and $a_1, a_2, \ldots, a_n$ be complex numbers. Let $B_1, B_2, \ldots, B_m$ be the atoms of the ring generated by $A_1, A_2, \ldots, A_n$. We set

$$\chi(i,k) = \begin{cases} 1 & \text{if } B_k \subset A_i, \qquad i=1,2,\cdots n, \\ 0 & \text{otherwise}, \qquad k=1,2,\cdots m. \end{cases}$$

Thus we have

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_i\overline{a}_j K_\varphi(A_i, g_i, A_j, g_j)$$

$$= \sum_{k=1}^{m}\left\{\sum_{i=1}^{n}\sum_{j=1}^{n} a_i\chi(i,k)\overline{a}_j\overline{\chi}(j,k)\begin{bmatrix}\varphi(B_k, g_j^{-1}g_i) - \varphi(B_k, g_i) - \varphi(B_k, g_j^{-1}) + is(B_k, g_j^{-1}, g_i) \\ \overline{\varphi}(B_k, g_j^{-1}g_i) - \varphi(B_k, g_i) - \overline{\varphi}(B_k, g_j^{-1}) - is(B_k, g_j^{-1}, g_i)\end{bmatrix}\right\}, \quad \begin{matrix} g_j\in G^+, \\ g_j \in G^-. \end{matrix}$$

Since φ is an Araki s function we have that every term inside the brackets is nonnegative. QED

Lemma (1.3): Let φ be an Araki s function. Then there exists a separable Hilbert space H spanned by vectors $Y(A, g, h)$, $A \in \mathcal{S}$, $g, h \in G$ such that the inner product is given by the following equations:

$\langle Y(A, g_1, h_1), Y(B, g_2, h_2)\rangle$

$$\text{(i)} = \varphi(A\cap B, h_2^{-1}g_2^{-1}g_1h_1) - \varphi(A\cap B, h_2^{-1}g_2^{-1}g_1) - \varphi(A\cap B, g_2^{-1}g_1h_1) + \varphi(A\cap B, g_2^{-1}g_1) + i[s(A\cap B, h_2^{-1}, g_2^{-1}g_1h_1) - s(A\cap B, h_2^{-1}, g_2^{-1}g_1)], \quad \text{if } (g_2, h_2)\in G^+\times G^+$$

$$\text{(ii)} = \overline{\varphi}(A\cap B, h_2^{-1}g_2^{-1}g_1h_1) - \overline{\varphi}(A\cap B, h_2^{-1}g_2^{-1}g_1) - \varphi(A\cap B, g_2^{-1}g_1h_1) + \varphi(A\cap B, g_2^{-1}g_1) + i[s(A\cap B, h_2^{-1}, g_2^{-1}g_1) - s(A\cap B, h_2^{-1}, g_2^{-1}g_1h_1), \quad \text{if } (g_2, h_2)\in G^+\times G^-$$

$$\text{(iii)} = \overline{\varphi}(A\cap B, h_2^{-1}g_2^{-1}g_1h_1) - \overline{\varphi}(A\cap B, h_2^{-1}g_2^{-1}g_1) - \overline{\varphi}(A\cap B, g_2^{-1}g_1h_1) + \overline{\varphi}(A\cap B, g_2^{-1}g_1) + i[s(A\cap B, h_2^{-1}, g_2^{-1}g_1) - s(A\cap B, h_2^{-1}, g_2^{-1}g_1h_1)], \quad \text{if } (g_2, h_2)\in G^-\times G^+$$

$$\text{(iv)} = \varphi(A\cap B, h_2^{-1}g_2^{-1}g_1h_1) - \varphi(A\cap B, h_2^{-1}g_2^{-1}g_1) - \overline{\varphi}(A\cap B, g_2^{-1}g_1h_1) + \overline{\varphi}(A\cap B, g_2^{-1}g_1) + i[s(A\cap B, h_2^{-1}, g_2^{-1}g_1h_1) - s(A\cap B, h_2^{-1}, g_2^{-1}g)] \quad \text{if } (g_2, h_2)\in G^-\times G^-$$

$\forall A, B \in \mathcal{S}$, $g_1, g_2, h_1, h_2 \in G$.

Proof: This follows by considering K_φ as the covariance function of a complex Gaussian stochastic process $X(A,g)$ where $(A,g)\in \mathcal{S}\times G$ cf. Ref. 6 and then setting $Y(A,g,h) = X(A,gh) - X(A,g)$. Let $H(A)$ be the span of the $Y(A,g,h)$ as g,h vary over G. Let $P(A)$ be the projection onto $H(A)$. Then we have the following lemma:

Lemma (1.4): $A \to P(A)$ is a projection-valued measure on $(T, \mathcal{S})$. We also have

$$P(A)Y(B,g,h) = Y(A\cap B, g, h) \quad \forall A, B \in \mathcal{S}.$$

This leads us to the following *main result* in this section which is obtained in the manner similar to Theorem (2.1) in Ref. 6 with some fairly obvious modifications:

Theorem (1.5): Let $G = G^+\cup G^-$ be a locally compact second countable group and $(T, \mathcal{S})$ be a standard Borel space. Let further (s, φ) be a pair consisting of a UA-Araki multiplier s and an Araki s function φ on $\mathcal{S}\times G\times G$ and $\mathcal{S}\times G$, respectively. Then there exists a complex separable Hilbert space H, a projection-valued measure $A \to P(A)$ on $\mathcal{S}$, a continuous UA representation $g \to U_g$ of G in H, and a continuous function $g\to\delta(g)$ in G with values in H such that the following holds:

(i) $U_g\delta(h) = \delta(gh) - \delta(g)$ $\forall g, h \in G$.

(ii) The subspaces $H(A) = P(A)H$ are invariant under all U_g.

(iii) for every $A \in \mathcal{S}$; $g, h_1, h_2 \in G$ we have

$\langle P(A)U_g\delta(h_1), \delta(h_2)\rangle$

$$=\begin{cases}\varphi(A,h_2^{-1}gh_1)-\varphi(A,h_2^{-1}g)-\varphi(A,gh_1)+\varphi(A,g)+i[s(A,h_2^{-1},gh_1)-s(A,h_2^{-1},g)], & h_2\in G^+,\\ \overline{\varphi}(A,h_2^{-1}gh_1)-\overline{\varphi}(A,h_2^{-1}g)-\varphi(A,gh_1)+\varphi(A,g)-i[s(A,h_2^{-1},gh_1)-s(A,h_2^{-1},g)], & h_2\in G^-.\end{cases}$$

Further the Hilbert space H can be written as a direct integral $H=\int_T H_t\,d\mu(t)$ with respect to a totally finite measure μ on $\mathcal{S}$, where the family $\{H_t : t\in T\}$ satisfies: (i)′ For all $t\in T$ there exists a representation $g\to \mathcal{U}(t,g)$ of G in H_t such that $U_g=\int_T \mathcal{U}(g,t)\,d\mu(t)$. (ii)′ The projection valued measure $A\to P(A)$ is given by

$$P(A)=\int_T \chi_A(t)I_t\,d\mu(t)\quad (I_t=\text{identity operator in } H_t).$$

(iii)′ For every t there exists a continuous map $g\to\delta(t,g)$ from G into H_t such that

$$\mathcal{U}(t,g)\delta(t,h)=\delta(t,gh)-\delta(t,g)\quad \forall g,h\in G$$

and

$$\delta(g)=\int_T \delta(t,g)\,d\mu(t)\quad \forall g\in G.$$

The measure μ is determined up to equivalence. The map $t\to(\mathcal{U}(t,\cdot),\delta(t,\cdot))$ is determined up to unitary equivalence, a.e., $[\mu]$.

Conversely: Given a totally finite measure μ and a triplet $(H_t,\mathcal{U}(t,\cdot),\delta(t,\cdot))$ for every t such that (i)′, (ii)′, (iii)′ hold and the direct integrals are well-defined, then we can construct an UA-Araki multiplier s and an Araki-s function φ such that $(U_g, P(A),\delta)$ defined by (i)′, (ii)′, (iii)′ satisfies (i), (ii), and (iii). If (s',φ') is another pair satisfying the same conditions then $(s-s')$ is trivial for every fixed A and $\mathrm{Re}\varphi=\mathrm{Re}\varphi'$.

Note: A continuous multiplier $s(g,h)$ is trivial if there exists a continuous real-valued function $a: G\to\mathbb{R}$ such that

$$s(g,h)=\begin{cases}a(gh)-a(g)-a(h), & g\in G^+,\\ a(gh)-a(g)+a(h), & g\in G^-.\end{cases}$$

That is to say, a trivial multiplier is just a coboundary.

Proof: We only note that for the converse one sets

$$\varphi(A,g)=-\tfrac{1}{2}\langle P(A)\delta(g),\delta(g)\rangle,$$

$$s(A,g,h)=\begin{cases}\mathrm{Im}\langle P(A)\delta(h),\delta(g^{-1})\rangle, & g\in G^+,\\ -\mathrm{Im}\langle P(A)\delta(h),\delta(g^{-1})\rangle, & g\in G^-.\end{cases}$$

Remark: This gives the generalization of Theorem (2.1) in Ref. 6.

2. MULTIPLICATIVE MEASURES AND FACTORIZABLE FAMILIES OF σ-POSITIVE DEFINITE FUNCTIONS

In order to study factorizable representations we have to analyse "nonatomic complex-valued multiplicative measures" and "σ-positive definite functions." We need a definition and two lemmas from Ref. 6.

Definition (2.1): Let $(T,\mathcal{S})$ be a standard Borel space. A function $M:\mathcal{S}\to\mathbb{C}$ is called *nonatomic complex-valued multiplicative measure* if the following holds

(i) $0<|M(A)|\leq 1\quad\forall A\in\mathcal{S}$,

(ii) $M(\phi)=1$,

(iii) $M(\cup_{i=1}^{\infty}A_i)=\Pi_{i=1}^{\infty}M(A_i)$ for any disjoint sequence $\{A_n\}\in\mathcal{S}$,

(iv) for every single point set $\{t\}$, $t\in T$, $M(\{t\})=1$.

Throughout the remainder, by a multiplicative measure we shall always mean a nonatomic complex-valued one. The following lemma shows that every multiplicative measure is the exponential of an additive measure.

Lemma (2.2)[6]: Let M be a multiplicative measure on $(T,\mathcal{S})$. Then there exists a unique nonatomic complex-valued totally finite measure m such that

$$M(A)=\exp[m(A)]\quad\forall A\in\mathcal{S}.$$

We shall also need the following.

Lemma (2.3)[6]: Let X be an arbitrary topological space and $K:\mathcal{S}\times X\times X\to\mathbb{C}$ have the following properties:

(i) $K(\cdot,x,y)$ is a multiplicative measure on $\mathcal{S}$ $\forall x,y\in X$.

(ii) $K(A,\cdot,\cdot)$ is a continuous positive definite Kernel on $X\times X$ $\forall A\in\mathcal{S}$.

Then there exists a unique function $K':\mathcal{S}\times X\times X\to\mathbb{C}$ satisfying the following conditions:

(a) $K'(\cdot,x,y)$ is a totally finite nonatomic complex-valued measure on $\mathcal{S}$ $\forall x,y\in X$;

(b) $K'(A,\cdot,\cdot)$ is a conditionally positive definite Kernel on $X\times X$ whose real part is continuous;

(c) $K(A,x,y)=\exp K'(A,x,y)\quad\forall A\in\mathcal{S},\ x,y\in X$.

Definition (2.4): A function $\varphi:G\to\mathbb{C}$ is called σ *positive* for a UA-multiplier σ if the following holds:

(i)

$$\varphi(e)=1,\quad \varphi(g)=\begin{cases}\overline{\varphi}(g^{-1}), & g\in G^+,\\ \varphi(g^{-1}), & g\in G^-,\end{cases}$$

(ii) the kernel

$$K(g,h)=\begin{cases}\sigma(h^{-1},g)\varphi(h^{-1}g), & h\in G^+,\\ \overline{\sigma}(h^{-1},g)\overline{\varphi}(h^{-1}g), & h\in G^-,\end{cases}$$

is positive definite in $G\times G$.

Definition (2.5): A family $\{\sigma(A,\cdot,\cdot);\varphi(A,\cdot)\}A\in\mathcal{S}$ of UA multipliers on $G\times G$ and $\sigma(A,\cdot,\cdot)$-positive functions $\varphi(A,\cdot)$ on G is said to be *factorizable* if

(i) $\sigma(\cdot,g,h)$ and $\varphi(\cdot,h)$ are multiplicative measures on $\mathcal{S}$ $\forall g,h\in G$,

(ii)

$$\sigma(A,g,h)=\begin{cases}\overline{\sigma}(A,h^{-1},g^{-1}), & (g,h)\in G^+\times G^+\cup G^-\times G^-,\\ \sigma(A,h^{-1},g^{-1}), & (g,h)\in G^+\times G^-\cup G^-\times G^+\end{cases}$$

With the preceding lemmas and definitions it is then fairly easy to prove (see Ref. 6) the following.

Theorem (2.6): Let G be a locally compact group, $(T,\mathcal{S})$ be a standard Borel space. Let $\{\sigma(A,\cdot,\cdot);\varphi(A,\cdot)\}A\in\mathcal{S}$ be a factorizable family of continuous UA multipliers $\sigma(A,\cdot,\cdot)$ defined w.r.t. the same normal subgroup G^+ on $G\times G$ and $\sigma(A,\cdot,\cdot)$-positive continuous functions $\varphi(A,\cdot)$ on G. Then $\exists$ a map $\beta:\mathcal{S}\times G\to S^1$ such that $\beta(A,\cdot)$ is a continuous function on G for every $A\in\mathcal{S}$, $\beta(\cdot,g)$ is a multiplicative measure for every $g\in G$ and

$$\exp[\psi(A,gh)+is(A,g,h)]$$

$$=\begin{cases}\sigma(A,g,h)\varphi(A,gh)\beta(A,g)\beta(A,h), & g\in G^+,\\ \sigma(A,g,h)\varphi(A,gh)\beta(A,g)\bar\beta(A,h), & g\in G^-,\end{cases}$$

where s is a UA-Araki multiplier and ψ is an Araki s function. Further if $\sigma\equiv 1$ we may choose $\beta\equiv 1$, $s\equiv 0$. We need another two definitions.

Definition (2.7): Let $(W^{(i)},\sigma_i,x_i)$ be two cyclic UA-multiplier representations acting in Hilbert spaces H_i, $i=1,2$, respectively, where both representations are defined w.r.t. the same normal subgroup G^+ of G. Their *convolution* denoted by $(W^{(1)},\sigma_1,x_1)*(W^{(2)},\sigma_2,x_2)$ is the UA-multiplier representation $W^{(1)}\otimes W^{(2)}$ restricted to the cyclic subspace generated by $x_1\otimes x_2$ in $H_1\otimes H_2$. The multiplier of the convolution is clearly $\sigma_1\sigma_2$.

Definition (2.8): For any cyclic UA-multiplier representation (W,σ,x) of G the function $f(g)=\langle W_g x,x\rangle$ is called its expectation value. We then have the following:

Lemma (2.9): Let (W,σ,x) be a cyclic UA-multiplier representation where σ satisfies

(i) $\sigma(h^{-1},g^{-1})=\begin{cases}\bar\sigma(g,h), & (g,h)\in G^+\times G^+\cup G^-\times G^-,\\ \sigma(g,h), & (g,h)\in G^+\times G^-\cup G^-\times G^+,\end{cases}$

(ii) $\sigma(g,g^{-1})\equiv\sigma(g^{-1},g)\equiv 1$;

then the expectation $f(g)=\langle W_g x,x\rangle$ is σ positive.

Proof: Straightforward computation.

Definition (2.10): Let $(T,\mathcal{S})$ be a standard Borel space and for every $A\in\mathcal{S}$ let $(W^A,\sigma(A,\cdot,\cdot),x_A)$ be cyclic UA-multiplier representations of G. (All defined w.r.t. the same G^+.) The family $\{W^A,\sigma(A,\cdot,\cdot),x_A\}A\in\mathcal{S}$ is said to be *factorizable* if

(i) for every sequence $\{A_n\}$ of sets in $\mathcal{S}$ descending to a single point set we have

$$\lim_{n\to\infty}\langle W_g^{A_n}x_{A_n},x_{A_n}\rangle=1$$

and therefore

$$\lim_{n\to\infty}\sigma(A_n,g,h)=1$$

uniformly on compact sets of G and $G\times G$, respectively.

(ii) for every $A\in\mathcal{S}$ and any finite measurable partition of A into sets $A_1,A_2,\ldots,A_k$, the cyclic representations $(W^A,\sigma(A,\cdot,\cdot),x_A)$ and $(W^{A_1},\sigma(A_1,\cdot,\cdot),x_{A_1})*\cdots*(W^{A_k},\sigma(A_k,\cdot,\cdot),x_{A_k})$ are unitarily equivalent.

In order to see how Araki s functions and factorizable representations are connected, we need another lemma.

Lemma (2.11): Let $(T,\mathcal{S})$ be a standard Borel space, $G=G^+\cup G^-$, locally compact and second countable as usual, and let $\{W^A,\sigma(A,\cdot,\cdot),x_A\}$ be a family of factorizable UA-multiplier representations of G then $\{g:f(A,g)=\langle W_g^A x_A,x_A\rangle\neq 0\}A\in\mathcal{S}$ forms an open subgroup of G.

Proof: Since $f(T,g)=f(A,g)f(A',g)$ it is sufficient to show $\{g:f(T,g)\neq 0\}$ is an open subgroup of G. We first show that $|f(T,g)|^2$ is positive definite and continuous. For this we need the $\sigma(T,\cdot,\cdot)$ extension of G (denoted by G^σ) which is defined as follows: $G^\sigma=G\times S^1$ as a set with group operation given by

$$(g_1,\lambda_1)\circ(g_2,\lambda_2)=\begin{cases}(g_1g_2,\lambda_1\lambda_2\sigma(T,g_1,g_2)), & g_1\in G^+,\\ (g_1g_2,\lambda_1\bar\lambda_2\sigma(T,g_1,g_2)), & g_1\in G^-,\end{cases}$$

$$\forall g_1,g_2\in G,\quad \lambda_1,\lambda_2\in S^1.$$

It is then easy to see that furnished with the product Borel structure G^σ is a standard group. (This definition was also given by E. Khabie independently of the author in an unpublished paper). Now the proof proceeds along the lines of Lemma (4.1) in Ref. 6. We are now ready to state the *main result* connecting Araki-s functions and factorizable cyclic UA-multiplier representations.

Theorem (2.12): Let $(T,\mathcal{S})$ be a standard Borel space and $\{W^A,\sigma(A,\cdot,\cdot),x_A\}$ be a factorizable family of cyclic UA-multiplier representations of G. Then there exists another factorizable family $\{\hat W^A,\hat\sigma(A,\cdot,\cdot),\hat x_A\}$ such that

(i) for each A, $(W^A,\sigma(A,\cdot,\cdot),x_A)$ and $(\hat W^A,\hat\sigma(A,\cdot,\cdot),\hat x_A)$ are projectively equivalent (i.e., unitarily equivalent up to a scalar factor of modulus 1)

(ii) $\{g:f(A,g)=\langle\hat W_g\hat x_A,\hat x_A\rangle\neq 0\}$ is an open and closed subgroup N of G and on that subgroup $f(A,g)$ is $\sigma(A,\cdot,\cdot)$-positive $\forall A\in\mathcal{S}$.

(iii) The family $\{\hat\sigma(A,\cdot,\cdot),f(A,\cdot)\}$ is factorizable in the sense of Definition (2.5) on N and both $\hat\sigma(A,\cdot,\cdot)$ and $f(A,\cdot)$ are continuous on $G\times G$ and G, respectively.

Conversely: Every continuous factorizable family on G $\{\sigma(A,\cdot,\cdot),f(A,\cdot)\}$ yields a factorizable family $\{W^A,\sigma(A,\cdot,\cdot),x_A\}$ by the equation $f(A,g)=\langle W_g^A x_A,x_A\rangle$. [Here the $f(A,\cdot)$ are $\sigma(A,\cdot,\cdot)$-positive.]

Proof: We set

$$\hat W_g^A=\begin{cases}(|\langle W_g^A x_A,x_A\rangle|/\langle W_g^A x_A,x_A\rangle)W_g^A & \text{if } \langle W_g^A x_A,x_A\rangle\neq 0,\\ W_g^A & \text{otherwise,}\end{cases}$$

changing $\sigma(A,\cdot,\cdot)$ accordingly to $\hat\sigma(A,\cdot,\cdot)$ and setting $x_A=\hat x_A$. Then it is fairly easy to see that (i), (ii), and (iii) hold. For the converse we use again the σ extension G^σ of G. We note that the following Kernel is positive definite on G^σ (for fixed A):

$$K_A[(g,\lambda),(h,\mu)]=\begin{cases}\psi_A[(h,\mu)^{-1}\circ(g,\lambda)], & h\in G^+,\\ \bar\psi_A[(h,\mu)^{-1}\circ(g,\lambda)], & h\in G^-,\end{cases}$$

where $\psi_A(g,\lambda)=\lambda f(A,g)$. Also note that

$$\psi_A[(g,\lambda)^{-1}]=\begin{cases}\bar\psi_A[(g,\lambda)], & g\in G^+,\\ \psi_A[g,\lambda] & g\in G^-.\end{cases}$$

Thus considering K_A as the covariance function of a Gaussian stochastic process $X(g,\lambda)$ we can find a UA representation of G^σ [for details see Theorem (3.7) in Ref. 8] defined by

$$U^A_{(g_1,\lambda_1)}X(g,\lambda)=X((g_1,\lambda_1)\circ(g,\lambda))$$

and extended by linearity (antilinearity). This satisfies

$$\langle U^A_{(g,\lambda)}X_A,X_A\rangle=\psi_A(g,\lambda)=\lambda f(A,g).$$

Now the rhs is a product in λ and g thus $U^A_{(g,\lambda)}=\lambda V_g^A$ for some unitary (or antiunitary) V_g^A and $g\to V_g^A$ has multiplier $\sigma(A,\cdot,\cdot)$ and satisfies

$$\langle V_g^A x_A,x_A\rangle=f(A,g).$$

Now let $A_1,\ldots,A_n$ be any measurable partition of A then

$$f(A,g)=\prod_{i=1}^n f(A_i,g)\text{——(a)},$$
$$\sigma(A,\cdot,\cdot)=\prod_{i=1}^n\sigma(A_i,\cdot,\cdot)\text{——(b)}.$$

So for each we construct $V_g^{A_i}$ but from (a) and (b) and uniqueness in Theorem (3.7) in Ref. 8 it then follows that $\{V_g^A, \sigma(A,\cdot,\cdot), x_A\}$ is a factorizable family. QED

Corollary: Let $\{W^A, \sigma(A,\cdot,\cdot), x_A\} A\in \mathcal{S}$ be a factorizable family of cyclic UA-multiplier representations of G. Then $\exists$ another factorizable family $\{\hat{W}^A, \hat{\sigma}(A,\cdot,\cdot), \hat{x}_A\} A \in \mathcal{S}$ such that (i) $(W^A, \sigma(A,\cdot,\cdot), x_A)$ and $(\hat{W}^A, \hat{\sigma}(A,\cdot,\cdot), \hat{x}_A)$ are projectively equivalent $\forall A\in\mathcal{S}$ and (ii) there exists a UA-Araki multiplier $s(A,\cdot,\cdot)$ and an Araki s function $\varphi(A,\cdot)$ such that

$$\langle \hat{W}_g^A \hat{x}_A, \hat{x}_A\rangle = \exp\varphi(A,g), \quad \text{where } \langle W_g^A x_A, x_A\rangle \neq 0$$

and

$$\hat{\sigma}(A,g,h) = \exp is(A,g,h) \quad \forall A\in\mathcal{S},\ g,h\in G.$$

Proof: Define $\hat{W}^A$ and $\hat{x}_A$ as above. Then an application of Theorem (2.6) where $\beta\equiv 1$ immediately implies the result.

3. INFINITELY DIVISIBLE PUA REPRESENTATIONS

"Infinitely divisible" representations are closely related to "factorizable representations." They are also of interest in quantum mechanics. See, e.g., Ref. 2. So we'll briefly outline their theory in this chapter. It will turn out that they are described by conditionally s-positive functions and these in turn by first order cocycles. Thus we need the Theorem [corresponding to Theorem (1.5)].

Theorem (3.1): To every conditionally s-positive definite function ϕ defined on G, there corresponds a UA representation $g\to U_g$ in a Hilbert space H with inner product $\langle\cdot,\cdot\rangle$ and a $\delta: G\to H$ such that

(i) $U_g\delta(h) = \delta(gh) - \delta(g)$ (i.e., δ is a first order cocyle) $\forall g.h\in G$,

(ii) $\langle\delta(h_1), \delta(h_2)\rangle$

$$=\begin{cases}\phi(h_2^{-1}h_1) - \phi(h_1) - \phi(h_2^{-1}) + is(h_2^{-1},h_1), & h_2\in G^+,\\ \overline{\phi}(h_2^{-1}h_1) - \phi(h_1) - \overline{\phi}(h_2^{-1}) - is(h_2^{-1},h_1), & h_2\in G^-.\end{cases}$$

The vectors $\{\delta(g): g\in G\}$ may be assumed to span H. If U^1, δ^1 is another pair satisfying the above conditions, for the same ϕ and s then U and U^1 are unitarily equivalent and δ and δ^1 correspond under this equivalence.

Conversely: Given (U,δ) as above then there exists a pair (ϕ, s) such that (ii) holds.

Proof: Take the Borel-space T in Theorem (1.5) to be a point.

Definition (3.2): A cyclic (in the obvious sense) PUA representation is called *infinitely divisible* if, for every positive integer n, there exists a cyclic PUA representation $g\to U_g^{1/n}$ with cyclic vector $x^{1/n}$ [denoted by $(U_g^{1/n}, x^{1/n})$] such that U_g and

$$\underbrace{U_g^{1/n}\otimes\cdots\otimes U_g^{1/n}}_{n},$$

restricted to the cyclic subspace generated by $x^{1/n}\otimes\cdots\otimes x^{1/n}$ are unitarily equivalent and the vectors x and $x^{1/n}\otimes\cdots\otimes x^{1/n}$ correspond under this equivalence. The pair $(U_g^{1/n}, x^{1/n})$ is called an nth root of (U_g, x). If σ is the multiplier of (U_g,x) and σ_n the multiplier of $(U_g^{1/n}, x^{1/n})$, then it follows immediately that we may assume without loss of generality $\sigma = \sigma_n^n\ \forall n$. Such a σ is called *infinitely divisible*.

We shall define an extension G_σ^* of G to see how infinitely divisible PUA representations of G may be extended to UA representations of G_σ^*. Let $\{\sigma_k\}_{k=1,2,\cdots}$ be such that $\sigma_k^k = \sigma\ \forall k$ and let $\sigma=\sigma_1$. Define:

$$G_\sigma^* = \{(g, t_1, t_2, \cdots): g\in G, t_k\in S^1, t_k^k = t_1,\ k=1,2,\cdots\}$$

then G_σ^* becomes a group with operation given by

$$(g,t_1,t_2,\cdots)\,(g',t_1',t_2',\cdots)$$
$$=\begin{cases}(gg', t_1t_1'\sigma_1(g,g'), t_2t_2'\sigma_2(g,g'),\cdots), & g\in G^+,\\ (gg', t_1\overline{t}_1'\sigma_1(g,g'), t_2\overline{t}_2'\sigma_2(g,g'),\cdots), & g\in G^-.\end{cases}$$

The set $T_\sigma^* = \{(e,t_1,t_2,\cdots)\in G_\sigma^*\}$ is a closed subgroup of the infinite dimensional torus and therefore compact. The arguments used in Ref. 9, show that G_σ^* is locally compact in the Weil topology, T_σ^* is a normal subgroup and G_σ^*/T_σ^* is isomorphic to G. We then have the following.

Theorem (3.3): Let $g\to U_g$ be an infinitely divisible PUA representation with multiplier σ of a group G. Let $\sigma=\sigma_k^k$, $k=1,2,\cdots$; then there exists an infinitely divisible UA representation of G_σ^* $(g,t_1,\ldots)\to V_{(g,t_1,\cdots)}$ given by $V_{(g_1,t_1,t_2,\cdots)} = t_1U_g$.

Definition (3.4): A continuous positive definite Hermitian Kernel $K(g,h)$ satisfying $K(e,e)=1$ is called *infinitely djvisible* if for each integer n there exists a continuous positive definite Hermitian Kernel K_n such that

$$K_n(e,e)=1;\quad K_n^n = K.$$

Remark: Infinitely divisible kernels of the form

$$K(g,h)=\begin{cases}\phi(h^{-1}g), & h\in G^+,\\ \overline{\phi}(h^{-1}g), & h\in G^-,\end{cases}\quad K_n=\begin{cases}\phi_n(h^{-1}g), & h\in G^+,\\ \overline{\phi}_n(h^{-1}g), & h\in G^-,\end{cases}$$

where ϕ,ϕ_n are continuous function $G\to\mathbb{C}$ and satisfy

(i) $\phi_{(n)}(g^{-1}) = \begin{cases}\overline{\phi}_{(n)}(g), & g\in G^+,\\ \phi_{(n)}(g), & g\in G^-,\end{cases}$ (ii) $\phi_{(n)}(e)=1$,

give a one-to-one correspondence between infinitely divisible positive Kernels and equivalence classes of cyclic infinitely divisible UA representation. This is easily proved by considering the $K(g,h)$ as covariance functions of Gaussian stochastic processes $X(g)$. See Theorem (2.12). This leads to

Theorem (3.5): Let $G=G^+\cup G^-$ be locally compact and second countable and let (s,ϕ) be a pair consisting of an additive UA multiplier s and a conditionally s-positive function ϕ. Then, up to equivalence, there exists a unique infinitely divisible PUA representation (U_g, x) with multiplier $\exp is=\sigma$ such $\langle U_gx, x\rangle = \exp\phi(g)$

Proof: This makes use of the above remark, Lemma (2.2) in Ref. 5 and then proceeds along the lines of Theorem (5.4) in Ref. 6. Making use of G_σ^* and proceeding similarly as in the first part of the proof of Theorem (2.12) we obtain a "canonical form" for U_g namely:

Theorem (3.6): Let $G=G^+\cup G^-$ as usual. Let (V_g,σ,x) be an infinitely divisible cyclic PUA representation.

Let $G_1=\{g:\langle V_g x, x\rangle \neq 0\}$, $G_1^+=G^+\cap G_1$; $G_1^-=G^-\cap G_1$. Then there exists a projectively equivalent infinitely divisible cyclic PUA representation (U_g, σ', x), where σ' is continuous, such that the function $f(g)=\langle U_g x, x\rangle$ is real and positive on G_1. Further there exists a sequence $\{\sigma_n'\}$ of multipliers which are continuous on $G_1\times G_1$ such that $\sigma_n'^n=\sigma'\ \forall n$ and σ_n' and σ satisfy

$$\sigma'_{(n)}(h^{-1},g^{-1})=\begin{cases}\bar{\sigma}'_{(n)}(g,h), & (g,h)\in G_1^+\times G_1^+\cup G_1^-\times G_1^-,\\ \sigma'_{(n)}(g,h), & (g,h)\in G_1^+\times G_1^-\cup G_1^-\times G_1^+.\end{cases}$$

Now it is fairly easy by adapting the arguments in the proof of Theorem (9.10) in Ref. 3 and considering the positive definite Kernel

$$K(g,h)=\begin{cases}\sigma(h^{-1},g)f(h^{-1}g)[f(g)f(h^{-1})]^{-1}, & h\in G^+\\ \bar{\sigma}(h^{-1},g)f(h^{-1}g)[f(g)f(h^{-1})]^{-1}, & h\in G^-,\end{cases}$$

where $f(g)=\langle U_g x, x\rangle$ to arrive at a converse of Theorem (3.5) namely:

Theorem (3.7): Suppose (U_g,σ,x) is a "canonical" [in the sense of theorem (3.6)] infinitely divisible PUA representation of G. Suppose the expectation value $f(g)=\langle U_g x, x\rangle$ does not vanish on the subgroup N of G. Then there exists a real-valued, conditionally s-positive definite function ψ and an additive UA multiplier s defined on N and $N\times N$, respectively such that

$$f(g)=\langle U_g x, x\rangle=\exp\psi(g)\quad \forall g\in N$$

$$\sigma(g,h)=\exp is(g,h)\quad \forall (g,h)\in N\times N.$$

Remark: This is the main result in this section and generalizes the results of Parthasarathy and Schmidt in Ref. 5.

4. FACTORIZABLE REPRESENTATIONS OF CURRENT GROUPS AND THE ARAKI-WOODS EMBEDDING THEOREM

In Ref. 6 the definition of the *weak current group* of G over T [$(T,\mathcal{S})$ is a standard Borel space], is given as follows: $F(T,G)$, the weak current group of G over T is the set of all measurable maps $\gamma: T\to G$ which take only finitely many values. Multiplication is defined pointwise. In order to consider UA representations we redefine the *weak UA-current* group $F(T,G)$ as follows: Let $G=G^+\cup G^-$ then we set $F^+(T,G)=$ set of measurable $\gamma: T\to G^+$ taking only finitely many values; $F^-(T,G)=$ set of measurable $\gamma:T\to G^-$ taking only finitely many values in G^-. $F(T,G)=F^+(T,G)\cup F^-(T,G)$ is then a group under pointwise multiplication. Note that this definition reduces to the one given in Ref. 6 if G^- is empty.

For each $A\in\mathcal{S}$ we define the following functions:

$$\chi_g^A: T\to G,\quad \chi_g^A(t)=\begin{cases}g, & t\in A,\\ e, & t\notin A.\end{cases}$$

Note: If $g\in G^+$ then $\chi_g^A\in F^+(T,g)$.

Now any function $\gamma\in F(T,G)$ may be written as

$$\gamma=\prod_{i=1}^n \chi_{g_i}^{A_i},\quad\text{where}\quad\begin{cases}g_i\in G^+ & \forall i\ \text{ if } \gamma\in F^+,\\ g_i\in G^- & \forall i\ \text{ if } \gamma\in F^-,\end{cases}$$

and $\{A_i\}_{i=1}^n$ is a measurable partition of T. Now suppose that $\{W^A, \sigma(A,\cdot,\cdot), x_A\}$ is a factorizable family which is in the "canonical form" described in Theorem (2.12) and suppose, for simplicity, that the expectation values $\langle W_g^A x_A, x_A\rangle$ are nonvanishing, (otherwise we have to restrict our attention to N).

For any two elements

$$\gamma_1=\prod_{i=1}^n\chi_{g_i}^{A_i},\quad \gamma_2=\prod_{j=1}^m\chi_{h_j}^{B_j}\in F(T,G)$$

we set

$$\tilde{\sigma}(\gamma_1,\gamma_2)=\prod_{i=1}^n\prod_{j=1}^m\sigma(A_i\cap B_j, g_i, h_j);$$

then we obtain easily (by computation):

Lemma (4.1): $\tilde{\sigma}$ is a UA multiplier for $F(T,G)$ such that

(i) $$\tilde{\sigma}(\gamma_2^{-1},\gamma_1^{-1})=\begin{cases}\bar{\tilde{\sigma}}(\gamma_1,\gamma_2), & (\gamma_1,\gamma_2)\in F^+\times F^+\cup F^-\times F^-,\\ \tilde{\sigma}(\gamma_1,\gamma_2), & (\gamma_1,\gamma_2)\in F^+\times F^-\cup F^-\times F^+,\end{cases}$$

(ii) $\tilde{\sigma}(\gamma,\gamma^{-1})\equiv\tilde{\sigma}(\gamma^{-1},\gamma)\equiv 1\quad\forall\gamma,\gamma_1,\gamma_2\in F.$

Now let us define a function $\psi: F(T,G)\to\mathbb{R}$ as follows. Let

$$\gamma=\prod_{i=1}^n\chi_{g_i}^{A_i};$$

then

$$\psi(\gamma)=\prod_{i=1}^n\langle W_{g_i}^{A_i}x_{A_i}, x_{A_i}\rangle,$$

then ψ is $\tilde{\sigma}$ positive on $F(T,G)$, and we also see that $\psi(\gamma^{-1})=\psi(\gamma)$. Thus using the $\tilde{\sigma}$ extension of F we may again [as in Theorem (2.12)] construct a $\tilde{\sigma}$ representation $\tilde{W}_\gamma$ of F such that $\langle\tilde{W}_\gamma\tilde{x},\tilde{x}\rangle=\psi(\gamma)$. It seems natural to define $\tilde{\sigma}$ representation of $F(T,G)$ to be *factorizable* if they can be constructed in the way defined above (and, of course, $\tilde{\sigma}$ representations which are projectively equivalent to these). Combining this definition and the corollary to Theorem (2.12) we obtain:

Theorem (4.2): Let $G=G^+\cup G^-$ as usual, let $(T,\mathcal{S})$ be a standard Borel space, and let $F(T,G)=F^+(T,G)\cup F^-(T,G)$ be the weak UA-current group of G over T. Let $(\tilde{W}',\tilde{\sigma}',\tilde{x}')$ be a factorizable cyclic UA-multiplier representation of $F(T,G)$ in a separable Hilbert space H; then there exists a projectively equivalent factorizable representation $(\tilde{W},\tilde{\sigma},\tilde{x})$ such that if

$$A_i, B_j\in\mathcal{S},\quad \gamma_1=\prod_{i=1}^n\chi_{g_i}^{A_i},\quad \gamma_2=\prod_{j=1}^m\chi_{h_j}^{B_j},\quad g_i,h_j\in G,$$

we have

$$\langle\tilde{W}_{\gamma_1}\tilde{x},\tilde{x}\rangle=\exp\left\{\sum_{i=1}^n\varphi(A_i,g_i)\right\},$$

$$\tilde{\sigma}(\gamma_1,\gamma_2)=\exp\left\{\sum_{i=1}^n\sum_{j=1}^m s(A_i\cap B_j,g_i,h_j)\right\},$$

where $\varphi(A,\cdot)$ is an Araki s function and $s(A,\cdot,\cdot)$ is a UA-Araki multiplier.

Conversely: Given an Araki s function and a UA-Araki multiplier we can construct a factorizable UA representation $(\tilde{W},\tilde{\sigma},\tilde{x})$ satisfying the above equations.

Proof: The first part of the theorem just follows by construction. For the converse we set

$$\tilde{\varphi}(\gamma_1)=\sum_{i=1}^n\varphi(A_i,g_i),\quad \tilde{s}(\gamma_1,\gamma_2)=\sum_{i=1}^n\sum_{j=1}^m s(A_i\cap B_j,g_i,h_j)$$

whenever

$$\gamma_1=\prod_{i=1}^{n}\chi_{g_i}^{A_i},\quad \gamma_2=\sum_{j=1}^{m}\chi_{h_j}^{B_j}.$$

It then follows that $\tilde{\varphi}$ is conditionally $\tilde{s}$-positive definite and we construct the UA representation by a method analogous to the one used in Theorem (2.12). Note that by construction $\varphi(A,\cdot)$ is real valued $\forall A\in\mathcal{S}$ and $\tilde{\varphi}$ is real.

Remark: The above theorem together with Theorem (1.4) gives a complete description of all factorizable UA-multiplier representations of weak UA-current groups with nonvanishing expectation value. Let us now proceed to show how factorizable UA multiplier representations can be embedded in the symmetric Fock space over a separable Hilbert space H. If we consider the UA-Araki multiplier s and the Araki s function φ of the above theorem then, using Theorem (1.4) we obtain a measure μ on $(T,\mathcal{S})$ and UA representations $\mathcal{U}(t,g)$ of G in Hilbert spaces H_t and cocycles $\delta(t,\cdot)$ connected with (s,φ). Let us write for

$$\gamma=\prod_{i=1}^{n}\chi_{g_i}^{A_i},\quad \Delta(\gamma)=\{\delta(t,\gamma(t)):t\in T\}\in\int_T H_t\,d\mu(t)$$

$$\mathcal{U}(A,g)=\int_T\chi_A(t)\mathcal{U}(t,g)\,d\mu(t)$$

$$U_\gamma=\bigoplus_{i=1}^{n}\mathcal{U}(A_i,g_i)\tag{1}$$

thus

$$\Delta(\gamma)=\bigoplus_{i=1}^{n}\delta(A_i,g_i).\tag{2}$$

Then using (1) and (2) we see

$$U_{\gamma_1}\Delta(\gamma_2)=\Delta(\gamma_1\gamma_2)-\Delta(\gamma_1).$$

So Δ is a first order cocycle associated with U_γ.

Now let us define a metric in $F(T,G)$ by

$$\rho(\gamma_1,\gamma_2)=\{\int_T\|\delta(t,\gamma_1(t))-\delta(t,\gamma_2(t))\|^2\,d\mu(t)\}^{1/2}$$

and complete $F(T,G)$ under this metric to obtain the full UA-current group $\Gamma(T,G)$ (cf. Ref. 6). Then it is clear that our representation and the associated cocycle can be extended by continuity to Γ. Now let us set as before

$$\tilde{\varphi}(\gamma_1)=\sum_{i=1}^{n}\varphi(A_i,g_i),\quad \tilde{s}(\gamma_1,\gamma_2)=\sum_{i=1}^{n}\sum_{j=1}^{n}s(A_i\cap B_j,g_i,h_j),$$

where

$$\gamma_1=\prod_{i=1}^{n}\chi_{g_i}^{A_i},\quad \gamma_2=\prod_{j=1}^{m}\chi_{h_j}^{B_j},$$

then we have

$$\langle\Delta(\gamma_1),\Delta(\gamma_2)\rangle=\tilde{\varphi}(\gamma_2^{-1}\gamma_1)-\tilde{\varphi}(\gamma_1)-\tilde{\varphi}(\gamma_2^{-1})\pm i\tilde{s}(\gamma_2^{-1},\gamma_1),$$

where the $(-)$ sign applies if $\gamma_2\in\Gamma^-$.

Note: Here we have used the same notation for the natural extension of $\tilde{s}$ and $\tilde{\varphi}$ to Γ.

Let $H=\int H_t\,d\mu(t)$ be the Hilbert space where U_γ and Δ are defined. We construct the symmetric Fock space $\exp[H]$ over H: We set $X(\gamma)=\exp\tilde{\varphi}(\gamma)\exp[\Delta(\gamma)]$ then we have:

$$\langle X(\gamma_1),X(\gamma_2)\rangle=\exp\{\tilde{\varphi}(\gamma_2^{-1}\gamma_1)\pm i\tilde{s}(\gamma_2^{-1},\gamma_1)\},$$

where $(-)$ sign applies if $\gamma_2\in\Gamma^-$. We define the map $W_{\gamma_1}:X(\gamma)\to X(\gamma_1\gamma)\exp i\tilde{s}(\gamma_1,\gamma)$. This is unitary (antiunitary) if $\gamma_1\in\Gamma^+(\Gamma^-)$. Hence it can be extended to a unitary (antiunitary) operator on the closed linear span of this set. Further we have

$$W_{\gamma_1}W_{\gamma_2}=\exp i\tilde{s}(\gamma_1,\gamma_2)W_{\gamma_1\gamma_2},$$

$$\langle W_\gamma X(e),X(e)\rangle=\exp\tilde{\varphi}(\gamma).$$

So we get a factorizable UA-multiplier representation of Γ associated with the pair (s,φ). This is the analogue of the Araki—Woods embedding theorem for factorizable UA-multiplier representations with nonvanishing expectation (cf. Refs. 2 and 6).

ACKNOWLEDGMENT

The author would like to thank Professor K. R. Parthasarathy and Dr. K. Schmidt for several helpful discussions.

[1]H. Araki, *Factorizable Representations of Current Algebra*. RIMS-41, preprint Kyoto (1969).
[2]R. F. Streater, Sci. Int. Fis. "E. Fermi" **XI**, 247 (1969).
[3]K. R. Parthasarathy and K. Schmidt, "Positive Definite Kernels, Continuous Tensor Products, and Central Limit Theorems of Probability Theory," *Lecture Notes in Mathematics* (Springer, Berlin, 1972), Vol. 272.
[4]K. R. Parthasarathy, Commun. Math. Phys. **15**, 305 (1969).
[5]K. R. Parthasarathy and K. Schmidt, "Infinitely Divisible Projective Representations, Cocycles and Levy—Khinchine—Araki Formula on Locally Compact Groups," Research Report 17, Manchester—Sheffield School of Probability and Statistics, 1970 (unpublished).
[6]K. R. Parthasarathy, and K. Schmidt, Acta Math. 128, 53 (1972). Most of the results in Refs. 5, 6, and 7 have also been published in Ref. 3.
[7]K. Schmidt, Math. Ann. **192**, 107 (1971).
[8]B. J. Falkowski, Ph. D. thesis, University of Manchester, 1971 (unpublished).
[9]K. R. Parthasarathy, *Multipliers on Locally Compact Groups*. *Lecture Notes in Mathematics* 93 (Springer, Berlin, 1969).

A new system of Casimir operators for $U(n)$

R. L. Hudson

Mathematics Department, University of Nottingham, University Park, Nottingham NG7 2RD, Great Britain

(Received 30 August 1972)

Explicit formulas are found for the Casimir operators for $U(n)$ whose eigenvalues in the irreducible representation for which the highest weight has components $l_1 - n + 1$, $l_2 - n + 2$, ..., l_n are the elementary symmetric polynomials in l_1, l_2, ..., l_n.

1. INTRODUCTION

The eigenvalues in an irreducible representation of a complete system of Casimir operators for the unitary group $U(n)$ of order n have been calculated by Perelomov and Popov.[1] The systemof Casimir operators used is that of Gel'fand.[2] An alternative and more explicit expression for these eigenvalues was given by Louck,[3] the proof being supplied by Louck and Biedenharn.[4] Numerous earlier authors obtained partial results for eigenvalues of Gel'fand operators of low orders.[5] Now if the components of the highest weight of the representation are written as $l_1 - n + 1,\ l_2 - n + 2, \ldots, l_n$, it is known[6] that the eigenvalue of an arbitrary Casimir operator must be a symmetric polynomial in the quantities $l_1, l_2, \ldots, l_n$. The purpose of this work is to construct explicitly the system of Casimir operators for which the corresponding polynomials are the elementary symmetric polynomials:

$$\sigma_0^n(l_1, l_2, \ldots, l_n) = 1,$$

$$\sigma_r^n(l_1, l_2, \ldots, l_n) = \sum_{1 \leq j_1 < j_2 < \ldots < j_r \leq n} l_{j_1} l_{j_2} \ldots l_{j_r} \qquad (1.1)$$

$$(r = 1, 2, \ldots, n).$$

Using this system of Casimir operators it is clearly possible to construct Casimir operators whose eigenvalues are arbitrary symmetric polynomials in the quantities $l_1, l_2, \ldots, l_n$; in particular, Casimir operators for which the eigenvalues are sums of powers of $l_1, \ldots, l_n$ are readily obtained. Our method is to follow[7] Louck and Biedenharn[4] in exploiting certain behavioral characteristics of the eigenvalues of Casimir operators under uniform translation of $l_1, \ldots, l_n$, and under the summation process of the Weyl branching formula,[9] which are analogs of properties of the elementary symmetric polynomials.

2. THE CASIMIR OPERATORS G_r^n

The Lie group $U(n)$ is generated by elements A_j^i, $i, j = 1, \ldots, n$, satisfying

$$[A_j^i, A_l^k] = \delta_j^k A_l^i - \delta_l^i A_j^k. \qquad (2.1)$$

A *Casimir operator* is an element of C of the complex associative algebra $\mathscr{A}$ generated by these elements, subject to the commutation rule (2.1), such that for all $i, j = 1, \ldots, n$

$$[C, A_j^i] = CA_j^i - A_j^i C = 0.$$

C is *real* if it is invariant under the natural involution of $\mathscr{A}$ which makes A_j^i into A_i^j and reverses the order of products.

We denote by $\mathscr{S}_n$ the group of permutations of $\{1, \ldots, n\}$ and by ϵ_α the sign of the permutation $\alpha \in \mathscr{S}_n$. We write, when needed, $\binom{j_1 \ldots j_n}{k_1 \ldots k_n}$ for the permutation $j_1 \mapsto k_1, \ldots, j_n \mapsto k_n$, and (j, k) for the permutation which exchanges j and k.

Proposition 1:

$$G_n^n = \sum_{\alpha, \beta \in \mathscr{S}_n} \epsilon_{\alpha\beta} A_{1\beta}^{1\alpha} \cdots A_{n\beta}^{n\alpha} \qquad (2.2)$$

is a real Casimir operator.

Proof:

$$[G_n^n, A_j^i]$$

$$= \sum_{\alpha, \beta \in \mathscr{S}_n} \epsilon_{\alpha\beta} \sum_{r=1}^{n} \{\delta_{r\beta}^i A_{1\beta}^{1\alpha} \ldots A_j^{r\alpha} \ldots A_{n\beta}^{n\alpha} - \delta_j^{r\alpha} A_{1\beta}^{1\alpha} \ldots A_{r\beta}^i \ldots A_{n\beta}^{n\alpha}\}$$

$$= \sum_{r=1}^{n} \sum_{\alpha, \beta \in \mathscr{S}_n, r\beta = i} \epsilon_{\alpha\beta} A_{1\beta}^{1\alpha} \ldots A_j^{r\alpha} \ldots A_{n\beta}^{n\alpha}$$

$$- \sum_{s=1}^{n} \sum_{\alpha', \beta' \in \mathscr{S}_n, s\alpha' = j} \epsilon_{\alpha'\beta'} A_{1\beta'}^{1\alpha'} \ldots A_{s\beta'}^{i} \ldots A_{n\beta'}^{n\alpha'}.$$

If $i = j$, the term corresponding to the choice (r, α, β) in the first sum cancels that corresponding to $(i\alpha^{-1}, \alpha, \beta)$ in the second sum. If $i \neq j$ the terms in the first sum corresponding to the choices (r, α, β), (r, α_1, β), where $\alpha_1 = \alpha(j\beta^{-1}\alpha, r\alpha)$, cancel, and the terms in the second sum corresponding to the choices (s, α', β'), (s, α', β_1'), where $\beta_1' = \beta'(i\alpha^{-1}\beta', s\beta')$, cancel. It follows that G_n^n is a Casimir element. Moreover, the image of G_n^n under the natural involution is

$$\sum_{\alpha, \beta \in \mathscr{S}_n} \epsilon_{\alpha\beta} A_{n\alpha}^{n\beta} \ldots A_{1\alpha}^{1\beta} = \sum_{\alpha', \beta' \in \mathscr{S}_n} \epsilon_{\alpha'\beta'} A_{1\beta'}^{1\alpha'} \cdots A_{n\beta'}^{n\alpha'} = G_n^n,$$

where

$$\alpha' = \begin{pmatrix} 1, \ldots, n \\ n, \ldots, 1 \end{pmatrix} \beta, \quad \beta' = \begin{pmatrix} 1, \ldots, n \\ n, \ldots, 1 \end{pmatrix} \alpha. \qquad \text{QED}$$

For each real number θ, the map

$$A_j^i \mapsto A_j^i + \theta\delta_j^i \qquad (2.3)$$

determines an automorphism t_θ of the algebra $\mathscr{A}$ which preserves the natural involution on $\mathscr{A}$. The image under this automorphism of a real Casimir operator is thus also a real Casimir operator. Thus, for all real θ,

$$G_n^{n(\theta)} = \sum_{\alpha, \beta \in \mathscr{S}_n} \epsilon_{\alpha\beta} (A_{1\beta}^{1\alpha} + \delta_{1\beta}^{1\alpha}\theta) \cdots (A_{n\beta}^{n\alpha} + \delta_{n\beta}^{n\alpha}\theta) \qquad (2.4)$$

is a real Casimir operator. It follows that the coefficients of the powers of θ occurring in $G_n^{n(\theta)}$ are likewise real Casimir operators; these are easily seen to be the elements $(n!/r!)\,G_r^n$, where

$$G_0^n = 1,$$

$$G_r^n = \sum_{J,K\in P_r} \epsilon_{J,K} A_{k_1}^{j_1}\cdots A_{k_r}^{j_r}, \quad r=1,\ldots,n, \tag{2.5}$$

P_r is the class of ordered r-tuples

$$J=(j_1,\ldots,j_r), \quad K=(k_1,\ldots,k_r)$$

of distinct elements of $\{1,\ldots,n\}$, and

$$\epsilon_{J,K} = \begin{cases} 0 & \text{if } \{j_1,\ldots,j_r\}\neq\{k_1,\ldots,k_r\}, \\ \operatorname{sign}\begin{pmatrix} j_1,\ldots,j_r \\ k_1,\ldots,k_r\end{pmatrix} & \text{otherwise.}\end{cases}$$

Denoting by $G_r^{n(\theta)}$ the image of the Casimir operator G_r^n under the automorphism t_θ, we prove

Proposition 2:

$$G_r^{n(\theta)} = \sum_{s=0}^{r} \theta^{r-s} \frac{(n-s)!\,r!}{(n-r)!\,(r-s)!\,s!} G_s^n. \tag{2.6}$$

Proof: The G_r^n are defined by

$$G_n^{n(\theta)} = \sum_{r=0}^{n} \theta^r \frac{n!}{(n-r)!} G_{n-r}^n.$$

Using the fact that $t_{\theta+\phi} = t_\theta t_\phi$, we have

$$\sum_{r=0}^{n} \phi^r \frac{n!}{(n-r)!} G_{n-r}^{n(\theta)} = \sum_{r=0}^{n} (\phi+\theta)^r \frac{n!}{(n-r)!} G_{n-r}^n.$$

Comparing coefficients of θ^r on the two sides yields the result. QED

We record for use below the obvious identity

$$G_n^{n+1} = \sum_{r=1}^{n+1} \sum_{J,K\in P_n^{(r)}} \epsilon_{J,K} A_{k_1}^{j_1}\cdots A_{k_n}^{j_n}, \tag{2.7}$$

where $P_n^{(r)}$ is the subclass of P_n consisting of those ordered n-tuples which exclude r. We observe that for each r the inner sum in (2.7) is the Casimir operator G_n^n for the Lie subgroup of $U(n+1)$ generated by the A_j^i with $i,j\neq r$.

3. TRANSLATION AND SUMMATION PROPERTIES

The general theory of representations[10] of Lie groups, specialized to $U(n)$, yields the following characterization of the irreducible representations of $U(n)$. Every irreducible representation π is characterized by a strictly decreasing n-tuple of integers

$$l_1 > l_2 > \cdots > l_n.$$

Moreover, there exists a nonzero vector, called the vector of highest weight, in the representation space such that for $i,j=1,\ldots,n$,

$$d\pi(A_j^j)\psi = (l_j - n + j)\psi, \tag{3.1}$$

$$d\pi(A_j^i)\psi = 0 \quad (i<j), \tag{3.2}$$

where $d\pi$ is the differential of π.

For every Casimir operator C we denote by $c(l_1,\ldots,l_n)$ the unique eigenvalue of C in the representation characterized by $(l_1,\ldots,l_n)$; in particular, $g_r^n(l_1,\ldots,l_n)$ is the eigenvalue of G_r^n. Comparison of (2.3) and (3.1) shows that

$$g_r^{n(\theta)}(l_1,\ldots,l_n) = g_r^n(l_1+\theta,\ldots,l_n+\theta).$$

Hence from Proposition 2 we obtain

$$g_r^n(l_1+\theta,\ldots,l_n+\theta) = \sum_{s=0}^{r} \theta^{r-s} \frac{(n-s)!\,r!}{(n-r)!\,(r-s)!\,s!} g_s^n(l_1,\ldots,l_n). \tag{3.3}$$

According to the Weyl branching formula[9] the irreducible representation of $U(n+1)$ characterized by the $(n+1)$-tuple $l_1,\ldots,l_{n+1}$ decomposes, when restricted to any Lie subgroup of $U(n+1)$ isomorphic to $U(n)$, into the direct sum of irreducible representations of $U(n)$ characterized by the n-tuples $l_1',\ldots,l_n'$, where

$$l_1 > l_1' \geq l_2 > l_2' \geq l_3 > l_3' \geq l_4 \cdots \geq l_n > l_n' \geq l_{n+1}.$$

Taking the traces of the representatives of the two sides of (2.7) in the representation of $U(n+1)$ characterized by $l_1,\ldots,l_{n+1}$, and recalling the observation following (2.7), it follows that

$$g_n^{n+1}(l_1,\ldots,l_{n+1})\,d^{n+1}(l_1,\ldots,l_{n+1}) = (n+1)\sum_{l_1'=l_2}^{l_1-1}\sum_{l_2'=l_3}^{l_2-1}\cdots\sum_{l_n'=l_{n+1}}^{l_n-1} g_n^n(l_1',\ldots,l_n')\,d^n(l_1',\ldots,l_n'),$$

where $d^n(l_1,\ldots,l_n)$ is the dimension of the representation of $U(n)$ corresponding to $l_1,\ldots,l_n$, which is given[10] by

$$d^n(l_1,\ldots,l_n) = \prod_{1\leq i<j\leq n} (l_i - l_j). \tag{3.4}$$

We write this equation as

$$g_n^{n+1} d^{n+1} = (n+1) \mathop{\Sigma\!\!\!\!\!\!\;|}\, g_n^n d^n, \tag{3.5}$$

where the summation operator $\mathop{\Sigma\!\!\!\!\!\!\;|}$ yields a function of $n+1$ arguments from one of n arguments.

It follows from (3.5) and the fact that the dimension functions d^n, d^{n+1} are invariant under translation of their arguments through θ that

$$g_n^{n+1(\theta)} d^{n+1} = (n+1) \mathop{\Sigma\!\!\!\!\!\!\;|}\, g_n^{n(\theta)} d^n.$$

Using (3.3) and comparing coefficients of powers of θ gives

$$g_r^{n+1} d^{n+1} = [(n+1)/(n+1-r)] \mathop{\Sigma\!\!\!\!\!\!\;|}\, g_r^n d^n. \tag{3.6}$$

The eigenvalue $g_n^n(l_1,\ldots,l_n)$ can, in principle, be computed from the definition (2.2) of G_n^n and the Eqs. (3.1), (3.2) by repeated use of the commutation relation (2.1) to rewrite each monomial in G_n^n in a form in which the last term is of form A_j^i with $i\leq j$. Now it is known[6] that the eigenvalue of a Casimir operator is necessarily a symmetric polynomial in the quantities $l_1,\ldots,l_n$. Since no term of G_n^n contains a repeated upper or lower suffix, and application of the commutation rule (2.1) does not create new suffices, consideration of (3.1) shows that the polynomial $g_n^n(l_1,\ldots,l_n)$ cannot involve any argument l_j to more than the first power. We have proved the first assertion of

Proposition 3: $g_n^n(l_1,\ldots,l_n)$ is a linear combination of the elementary symmetric polynomials (1.1); moreover, the coefficient of $\sigma_n^n(l_1,\ldots,l_n)$ is nonzero.

Proof: To prove the latter assertion, observe that in the computation of $g_n^n(l_1,\ldots,l_n)$, contributions to the term in $\sigma_n^n(l_n,\ldots,l_n)$ will come only from those terms in G_n^n for which $\alpha=\beta$. Since the sign of each such term is positive there can be no cancellation, so that the coefficient of $\sigma_n^n(l_1, \ldots, l_n)$ is indeed nonzero.

Proposition 4: For $r=0,\ldots,n$, $g_r^n(l_1,\ldots,l_n)$ is a linear combination of the elementary symmetric polynomials $\sigma_s^n(l_1,\ldots,l_n)$, with $0 \le s \le r$, moreover, the coefficient of $\sigma_r^n(l_1,\ldots,l_n)$ is nonzero.

Proof: Comparing coefficients of x^{n-r} in the identity

$$\sum_{r=0}^{n} \sigma_r^n(l_1+\theta,\ldots,l_n+\theta)x^{n-r} = \sum_{r=0}^{n} \sigma_s^n(l_1,\ldots,l_n)(\theta+x)^{n-r},$$

gives the formula

$$\sigma_r^n(l_1+\theta,\ldots,l_n+\theta) = \sum_{s=0}^{r} \frac{(n-s)!}{(n-r)!\,(r-s)!}\,\theta^{r-s}\sigma_s^n(l_1,\ldots,l_n). \tag{3.7}$$

In accordance with Proposition 3, we can write

$$g_n^n(l_n,\ldots,l_n) = \sum_{r=0}^{n} b_r^n \sigma_r^n(l_1,\ldots,l_n),$$

with $b_n^n \ne 0$. Replacing each l_j by $l_j+\theta$,

$$g_n^n(l_1+\theta),\ldots,l_n+\theta) = \sum_{r=0}^{n} b_r^n \sigma_r^n(l_1+\theta,\ldots,l_n+\theta).$$

Substituting from (3.3) into the left-hand side, and from (3.7) into the right-hand side of this equation and comparing coefficients of θ^{n-r} yields $g_r^n(l_1,\ldots,l_n)$ as a linear combination of $\sigma_0^n(l_1,\ldots,l_n),\ldots,\sigma_r^n(l_1,\ldots,l_n)$, in which the coefficient of the latter is a nonzero multiple of b_n^n. QED

4. THE CASIMIR OPERATORS H_r^n

From Proposition 4, it follows that for $r=0,\ldots,n$, $\sigma_r^n(l_1,\ldots,l_n)$ can be expressed as a linear combination of $g_s^n(l_1,\ldots,l_n)$, with $0 \le s \le r$. The corresponding linear combination of the Casimir operators G_s^n, which we denote by H_r^n, will then have the desired property that their eigenvalues in each irreducible representation are the elementary symmetric polynomials in the weights of that representation. We proceed to determine the H_r^n explicitly.

We write

$$\sigma_n^n(l_1,\ldots,l_n) = \sum_{r=0}^{n} a_r^n g_r^n(l_1,\ldots,l_n). \tag{4.1}$$

Replacing each l_j by $l_j+\theta$, using (3.3) and (3.7), and comparing coefficients of powers of θ in (4.1) gives

$$\sigma_r^n(l_1,\ldots,l_n) = \sum_{s=0}^{r} a_{n-r+s}^n \frac{(n-s)!\,(n-r+s)!}{(r-s)!\,(n-r)!\,s!}\, g_s^n(l_1,\ldots,l_n). \tag{4.2}$$

Thus knowledge of the coefficients a_r^n suffices to determine all the H_r^n.

Before determining the a_r^n we require the following identity:

Proposition 5:

$$\frac{1}{(n+1)!}\sum_{r=0}^{n}(-)^r \sigma_{n-r}^{n+1} d^{n+1} = \sum \sigma_n^n d^n. \tag{4.3}$$

Proof: The dimension function d^n is given[10] by

$$d^n(l_1,\ldots,l_n) = \prod_{1 \le i<j \le n}(l_i - l_j) = \det\begin{bmatrix} l_1^{n-1} & \ldots & l_n^{n-1} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ 1 & \ldots & 1 \end{bmatrix}.$$

Thus we must prove that

$$\frac{1}{(n+1)!}\sum_{r=0}^{n}(-)^r\sigma_{n-r}^{n+1}(l_1,\ldots,l_{n+1})\det\begin{bmatrix} l_1^{n} & \ldots & l_{n+1}^{n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ 1 & \ldots & 1 \end{bmatrix}$$

$$= \sum_{l_1'=l_2}^{l_1-1}\cdots\sum_{l_n'=l_{n+1}}^{l_n-1} l_1' l_2' \cdots l_n' \det\begin{bmatrix} l_1'^{n-1} & \ldots & l_n'^{n-1} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ 1 & \ldots & 1 \end{bmatrix}. \tag{4.4}$$

The summand on the right-hand side of (4.4) is

$$\det\begin{bmatrix} l_1'^n & \cdots & l_n'^n \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ l_1' & & l_1' \end{bmatrix} = \det\begin{bmatrix} l_1'^n & \cdots & l_n'^n & 0 \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ l_1' & \cdots & l_n' & 0 \\ 1 & \cdots & 1 & 1 \end{bmatrix}$$

$$= \frac{1}{(n+1)!}\det\begin{bmatrix} (l_1'+1)^{n+1} - l_1'^n \ldots & (l_n'+1)^{n+1} - l_n'^{n+1} & l^{n+1} - 0 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ (l_1'+1) - l_1' \ldots & (l_n'+1) - l_n' & l - 0 \end{bmatrix},$$

where the extra terms, other than the appropriate power of l_j', in each row are linear combinations of the row below. The summation can now be carried out to give

$$\frac{1}{(n+1)!}\det\begin{bmatrix} l_1^{n+1} - l_2^{n+1} \cdots & l_n^{n+1} - l_{n+1}^{n+1} & 1 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ l_1 - l_2 \quad \cdots & l_n - l_{n+1} & 1 \end{bmatrix}$$

$$= \frac{1}{(n+1)!}\sum_{r=1}^{n+1}(-)^{n+1-r}\det\begin{bmatrix} l_1^{n+1} \cdots \overset{r}{\wedge} \cdots & l_{n+1}^{n+1} & 1 \\ & \cdot & \cdot \\ & \cdot & \cdot \\ & \cdot & \cdot \\ l_1 \cdots \overset{r}{\wedge} \cdots & l_{n+1} & 1 \end{bmatrix}$$

$$= \frac{1}{(n+1)!}(-)^n\left\{\det\begin{bmatrix} l_1^{n+1} \cdots & l_{n+1}^{n+1} & 1 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ l_1 \cdots & l_{n+1} & 1 \\ 1 \cdots & 1 & 1 \end{bmatrix} - \det\begin{bmatrix} l_1^{n+1} & \cdots & l_{n+1}^{n+1} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ l_1 & \cdots & l_{n+1} \end{bmatrix}\right\},$$

as is seen by expanding the $(n+2)\times(n+2)$ determinant by its last row,

$$= \frac{1}{(n+1)!}(-)^n\Big(d^{n+2}(l_1,l_2,\ldots,l_{n+1},1) - \prod_{j=1}^{n+1} l_j\, d^{n+1}(l_1,\ldots,l_{n+1})$$

$$= \frac{1}{(n+1)!}(-)^n\left(\prod_{j=1}^{n+1}(l_j-1)-\prod_{j=1}^{n+1}l_j\right)d^{n+1}(l_1,\ldots,l_{n+1})$$

$$= \frac{1}{(n+1)!}\sum_{r=0}^{n}(-)^r\sigma_{n-r}^{n+1}(l_1,\ldots,l_{n+1})d^{n+1}(l_1,\ldots,l_{n+1}).\ \text{QED}$$

We now obtain a system of linear equations for the coefficients a_r^n. Using (4.1) and (4.2) we substitute for the functions σ_{n-r}^{n+1}, σ_n^n occuring on the two sides of (4.3). The right-hand side of the resulting equation can be expressed, using (3.5), as a linear combination of the functions g_r^{n+1}. These functions inherit linear independence from the functions σ_r^{n+1}, in view of Proposition 4. Hence we may compare coefficients of g_r^{n+1} on the two sides of the resulting equation. We obtain finally

$$a_r^n = \frac{1}{n+1-r}\sum_{s=0}^{n-r}(-)^{n-s+r}a_{n+1-s}^{n+1}\frac{(n+1-r)!(n+1-s)!}{(n+1-r-s)!r!s!} \quad (r=0,\ldots,n). \tag{4.5}$$

Now in the trivial representation of $U(n)$, in which every element of $U(n)$ is mapped into the identity, every A_j^i is represented by the zero operator. This representation characterized by the n-tuple $n-1$, $n-2,\ldots,0$. Substituting these values into (4.1), we see that

$$a_0^n = \begin{cases} 1 & \text{if } n=0,\\ 0 & \text{if } n\neq 0.\end{cases} \tag{4.6}$$

Now if $a_0^n,\ldots,a_r^n$ and a_0^{n+1} are known, then (4.5) gives a triangular system of inhomogeneous linear equations for $a_1^{n+1},\ldots,a_{n+1}^{n+1}$ and hence determines these quantities uniquely. It follows by induction on n that if $a_0^0, a_0^1,\cdots$ are all specified, then the system (4.5) determines all a_r^n uniquely. Hence we need only exhibit a solution to the system (4.5), consistent with the condition (4.6), to determine the a_r^n. This is done in

Proposition 6: The quantities

$$a_r^n = [(n-r)!/n!]\sigma_{n-r}^n(0,1,\ldots,n-1) \tag{4.7}$$

satisfy (4.5), (4.6).

Proof: That (4.6) is satisfied is clear by inspection. Substituting into the right-hand side of (4.5) the values for a_{n+1-s}^{n+1} given by (4.7), we obtain

$$\frac{(n-r)!}{n!}\cdot\frac{1}{n+1}\sum_{s=0}^{n-r}\frac{(n+1-s)!}{(n+1-s-r)!r!}(-)^{n-s+r}\sigma_s^{n+1}(0,\ldots,n)$$

$$=\frac{(n-r)!}{n!}\cdot\frac{1}{n+1}\left(\sum_{s=0}^{n+1-r}\frac{(n+1-s)!}{(n+1-s-r)!r!}(-)^{n-s+r}\sigma_s^{n+1}(0,\ldots,n)\right.$$
$$\left.+\sigma_{n+1-r}^{n+1}(0,\ldots,n)\right)$$

$$=\frac{(n-r)!}{n!}\frac{1}{n+1}[-\sigma_{n+1-r}^{n+1}(-1,0,1,\ldots,n-1)$$
$$+\sigma_{n+1}^{n+1}(0,\ldots,n)],$$

using the formula (3.7), with θ taken as -1,

$$=\frac{(n-r)!}{n!}\frac{1}{n+1}[-\sigma_{n+1-r}^{n}(0,1,\ldots,n-1)$$
$$+\sigma_{n-r}^{n}(0,1,\ldots,n-1)+n\sigma_{n-r}^{n}(0,1,\ldots,n-1)$$
$$+\sigma_{n+1-r}^{n}(0,1,\ldots,n-1)]$$
$$=\frac{(n-r)!}{n!}\sigma_{n-r}^{n}(0,1,\ldots,n-1),$$
$$=a_r^n,$$

according to (4.7). Hence (4.5) is satisfied. QED

5. CONCLUSION

We conclude from (4.2) and (4.7) that:

The system of Casimir operators

$$H_r^n=\sum_{s=0}^{r}\frac{(n-s)!(n-r+s)!}{n!(n-r)!s!}\sigma_{r-s}^n(0,1,\ldots,n-1)G_s^n, \quad r=0,1,\ldots,n,$$

where

$$G_0^n=1,$$
$$G_s^n=\sum_{J,K\in P_s}\epsilon_{J,K}A_{k_1}^{j_1}\cdots A_{k_s}^{j_s}, \quad s=1,2,\ldots,n,$$

P_s is the class of ordered s-tuples $J=(j_1,j_2,\ldots,j_s)$, $K=(k_1,k_2,\ldots,k_s)$ of distinct elements of $\{1,\ldots,n\}$, and

$$\epsilon_{J,K}=\begin{cases}0 & \text{if } \{j_1,\ldots,j_s\}\neq\{k_1,\ldots,k_s\}\\ \operatorname{sign}\begin{pmatrix}j_1,\ldots,j_s\\ k_1,\ldots,k_s\end{pmatrix} & \text{otherwise}\end{cases}$$

is such that in the irreducible representation characterized by $l_1,\ldots,l_n$ of $U(n)$, H_r^n is represented by the scalar operator

$$\sigma_r^n(l_1,\ldots,l_n)I.$$

[1]A.M. Perelomov and V.S. Popov, Yad. Fiz. 3, 924 (1966) [Sov. J. Nucl. Phys. 3, 676 (1966)].
[2]I.M. Gel'fand, Mat. Sb. 26, 103 (1950).
[3]J.D. Louck, Am. J. Phys. 38, 3 (1970).
[4]J.D. Louck and L.C. Biedengarn, J. Math. Phys. 11, 2368 (1970).
[5]J.D. Louck, J. Math. Phys. 6, 1786 (1965); M. Moshinsky, Nucl. Phys. 31, 384 (1962); M. Umezawa, Nucl. Phys. 48, 111 (1963); 53, 54 (1964); 57, 65 (1964).
[6]G. Racah, *Group Theory and Spectroscopy*, Lecture notes (Institute for Advanced Study, Princeton, N.J., 1951).
[7]The author is grateful to the referee for bringing to his attention the fact that this technique has already been used in the reference cited,[8] and is not original to the author as he had thought.
[8]See Ref. 4, particularly Appendix B.
[9]H. Boerner, *Representations of Groups* (North-Holland, Amsterdam, 1963).
[10]N. Jacobson, *Lie Algebras* (Interscience, New York, 1967).

Remarks on the Klauder phenomenon

B. DeFacio* and C. L. Hammer

Ames Laboratory-USAEC and Department of Physics, Iowa State University, Ames, Iowa 50010
(Received 14 December 1973)

An elementary discussion is given of a phenomenon discovered by Klauder. In Klauder's phenomenon $H=H_0+\lambda V$ does not converge to H_0 as the positive real parameter $\lambda \to 0^+$. We discuss domain questions of the operators, and also operator forms and operator extensions. The implications of this phenomenon for "gauge theories" are given, and new examples, which include the massless limit for particles of arbitrary spin and an integral operator, are given.

1. INTRODUCTION

Recently Klauder[1,2] has shown that sufficiently singular potentials V *cannot* be turned off in the Hamiltonian $H=H_0+\lambda V$ to restore the free Hamiltonian H_0. Thus, one may have that

$$\operatorname*{s-lim}_{\lambda\to 0^+}(H_0+\lambda V)\neq H_0 \tag{1.1}$$

for λ a positive real parameter. One example of the Klauder *phenomenon* can be exhibited with a one-dimensional simple harmonic oscillator as the free system H_0,

$$H_0=P^2+\omega^2Q^2=-\frac{d^2}{dx^2}+\omega^2x^2, \tag{1.2}$$

and the singular interaction

$$V=1/|x|^3, \tag{1.3}$$

so that

$$H=H_0+\lambda V, \tag{1.4}$$

with λ a positive real parameter. In units with $\hbar=1$ configuration space eigenvectors of H_0 are given by

$$\psi_n^{(0)}(x)=\langle n,0|x\rangle=h_n(x)\exp(-x^2/2), \tag{1.5}$$

where h_n are the Hermite polynominals of order n ($n\in \mathbf{Z}^+$, positive integers including zero).

If the energy units further have $\omega=1$, the nth energy eigenvalue $E_n^{(0)}$ of H_0 is given by

$$E_n^{(0)}=2n+1. \tag{1.6}$$

The energy eigenvectors $\Phi_n(x,\lambda)$ of H, with energy eigenvalues $E_n(\lambda)$ are solutions to the operator equation

$$H\Phi_n(x,\lambda)=E_n(\lambda)\Phi_n(x,\lambda). \tag{1.7}$$

Neglecting domain questions, Klauder showed that for $n\in Z^+$ and n odd

$$\lim_{\lambda\to 0^+}[\Phi_n(x,\lambda)]=\psi_n^{(0)}(x), \tag{1.8}$$

$$\lim_{\lambda\to 0^+}[E_n(\lambda)]=E_n^{(0)}, \tag{1.9}$$

whereas for $n\in \mathbf{Z}^+$ and n even

$$\lim_{\lambda\to 0^+}[\Phi_n(x,\lambda)]=\psi_{n+1}^{(0)}(x),$$

$$\lim_{\lambda\to 0^+}[E_n(\lambda)]=E_{n+1}^{(0)}. \tag{1.10}$$

Therefore, one sees that only the odd eigenvectors remain. This occurs because $\Phi_n(x,\lambda)$ must vanish at $x=0$ if it is to be $L^2(R)$. Thus,

$$\operatorname*{s-lim}_{\lambda\to 0^+}[H(\lambda)]\neq H_0, \tag{1.11}$$

a result which we refer to as the *Klauder phenomenon*.

In the present paper, we wish to give a simple discussion of the Klauder phenomenon including domain questions and two possible meanings for the sum "+" in $H_0+\lambda V$ namely operator extensions[3] and operator forms.[4-7] The present note[8] contains some distinct results, e.g., the integral operator in Sec. 4, and elaborates on several points made independently by Simon and us. We are presenting our discussion partly on the assumption that it is more easily accessible to the working physicists. In Sec 3 and 4 several examples are presented and in Sec. 5 our conclusions are given. Useful definitions are given in the Appendix.

2. NOTATION, THEOREM, REMARKS AND EXAMPLES

Notation

$\mathcal{H}$ is a complex Hilbert space, $\mathbf{R}$, $\mathbf{C}$, $\mathbf{R}^n$ are the real, complex, and n-dimensional real number systems respectively.

H_0, H, V, A are (possibly unbounded) operators in $\mathcal{H}$, i.e., $H:\mathcal{H}\to\mathcal{H}$.

A^* is the *adjoint* of the operator A.

$\mathcal{D}(A)$ is the *domain* of A in $\mathcal{H}$, i.e., $\mathcal{D}(A)\subset\mathcal{H}$.

$\mathcal{R}(A)$ is the *range* of A in $\mathcal{H}$.

$\Gamma(A)$ is the *graph* of a in $\mathcal{H}$.

$f_n\uparrow f$ means that the sequence of real valued functions $f_n(x)$ increase to the real function $f(x)$ for all real x.

$A_1\subset A$ means that operator A *extends* the operator A_1.

$A|\mathcal{D}(A_1)=A_1$ is the restriction of A to the domain of A_1.

$\mathbf{Z}$, $\mathbf{Z}^+$ are the integers and positive integers, including zero, respectively.

Theorems

Theorem 1: Let $\mathfrak{q}_n$, $n=1,\cdots$, and $\mathfrak{q}$ be closed densely defined quadratic forms [see definition (15) in the Appendix] in $\mathcal{H}\times\mathcal{H}$ and satisfying the following conditions:

(i) $\mathcal{D}(\mathfrak{q}_n)\subset\mathcal{D}(\mathfrak{q})$, $\forall\, n$.

(ii) $\mathfrak{q}_n' = \mathfrak{q} - \mathfrak{q}_n$ satisfies

$$|\mathrm{Im}(\mathfrak{q}_n'[U])| \leq M\ \mathrm{Re}(\mathfrak{q}_n'[U]),$$

$$U \in \mathcal{D}(\mathfrak{q}_n),\ \forall\ n,\ \ M>0.$$

(iii) There is a core $\mathcal{D}$ of q such that $\mathcal{D} \subset \lim\inf \mathcal{D}(\mathfrak{q}_n)$ and $\lim_{n\to\infty} \mathfrak{q}_n[U] = \mathfrak{q}[U]$ if $U \in \tilde{\mathcal{D}}$.

The densely defined closed sesquilinear forms for self-adjoint T_n, T, given by

$$\mathfrak{q}_n[U,V] = (U, T_n V),\ \forall\ n,$$

$$\mathfrak{q}[U,V] = (U, TV),\ \forall\ n,$$

are identified with $\mathfrak{q}_n$ and $\mathfrak{q}$, respectively. Then T_n and T_n^* converge strongly to T and T^*, i.e.,

$$\mathrm{s\text{-}lim}_{n\to\infty} (T_n) = T,$$

$$\mathrm{s\text{-}lim}_{n\to\infty} (T_n^*) = T^*.$$

Thus, if Δ_s and Δ_s^* are the regions of strong convergence of the operator sequences $\{T_n\}$, $\{T_n^*\}$, then $\Delta_s{}^*$ is the mirror image of Δ_s with respect to the real axis and both contain the half-plane $\mathrm{Re}(z) < \gamma$, where γ is a vertex for $\mathfrak{q}$. The resolvents $R_z(T_n)$, $R_z(T_n^*)$ converge strongly to the resolvents in Δ_s, Δ_s^*.

Proof: Ref. 5, p.454—456.

Theorem 2: Let $\{\mathfrak{q}_n\}$ be a nonincreasing sequence of densely defined closed symmetric forms uniformly bounded from below $\mathfrak{q}_n \geq \lambda$ (λ constant). If T_n is a self-adjoint operator associated with $\mathfrak{q}_n$ [see definition (16)], then

$$\mathrm{s\text{-}lim}_{n\to\infty} (T_n) = T.$$

Proof: Ref. 5, p.458.

*Theorem*3: If the lower part $\lambda < \beta$ of the spectrum of T of Theorem 2 consists of finitely degenerate eigenvalues, these eigenvalues are stable under a form perturbation.

Proof: Ref. 5, p.460, 461.

Remarks

One notices that Theorems 1 and 2 do not hold for the operator $H = H_0 + \lambda V$ unless $\mathcal{D}(H_0) \cap \mathcal{D}(V)$ is dense and is a form core.[3] If these theorems do not hold, the representation in definition (16) cannot be used. If $\mathcal{D}(H_0) \cap \mathcal{D}(V)$ is not a form core, the unbounded self-adjoint operator H may have a strong limit

$$\mathrm{s\text{-}lim}_{\lambda\to 0^+} [H(\lambda)] \equiv H_{PF}. \tag{2.1}$$

It is just that $H_{PF} \neq H_0$, which is the Klauder phenomenon. Mathematically, these conditions have been given by Simon[4,8] and physically correspond to singular potentials, i.e., nonrenormalizable interactions.[9,10,11]

Examples

In view of the popularity of the so called *gauge field theories* we repeat an example due to Klauder[1,2] concerning the generating functional formalism. If the Langrangion

$$\mathcal{L} = \mathcal{L}_0 + \lambda V \tag{2.2}$$

with $V(\Phi) \geq 0$ and free of derivatives, the generating functional $Z(\Phi)$, for the Euclidean scalar field Φ, can be written as

$$Z(h) = \int \exp\{i[(h,\Phi) - \mathcal{L}_0(\Phi) - \lambda V(\Phi)]\}\mathcal{D}\Phi, \tag{2.3}$$

where $\mathcal{D}\Phi$ is a *formal*, translationally invariant measure with

$$(h,\Phi) = \int h(x)\Phi(x) d^4x, \tag{2.4}$$

and

$$\mathcal{L}_0(\Phi) = \tfrac{1}{2}\int\left[\sum_{i=1}^{n}\left(\frac{\partial\Phi}{\partial x_i}\right)^2 + m^2\Phi^2\right] d^4x. \tag{2.5}$$

If $Z(h)$ is positive definite, continuous, and satisfies $Z(0) = 1$, a unique normalized measure on the fields exists such that

$$Z(h) = \int \exp[i(h,\Phi)] d\sigma(\Phi). \tag{2.6}$$

When $Z(h)$ fulfills translation invarience and the cluster property, the support of σ is concentrated on those fields Φ for which

$$\lim_{\Omega\to\infty} |\Omega|^{-1} \int_\Omega dz \exp[i(h_z, \varphi)] = Z(h), \tag{2.7}$$

where $h_z(x) \equiv h(x-z)$, $\Omega \subset R^n$, and $|\Omega| = \int_\Omega dz$. For the free field, for example,

$$\begin{aligned} Z_F(h) &= \int \exp[i(h,\varphi)]\, d\sigma_F(\Phi) \\ &= \exp[-\tfrac{1}{2}\int (k^2+m^2)^{-1} |\hat{h}(k)|^2 dk], \end{aligned} \tag{2.8}$$

where the support of σ_F is as indicated by Eq. (2.7).

Klauder introduces, in a formal sense,

$$d\sigma(\Phi) = \chi(\Phi)\exp[-\mathcal{L}_0(\Phi) - \lambda V(\Phi)]\mathcal{D}\Phi, \tag{2.9}$$

where

$$\begin{aligned} &\chi(\Phi) = 1,\ \ \mathcal{L}_0 < \infty,\ \ V(\Phi) < \infty, \\ &\chi(\Phi) = 0,\ \ \mathcal{L}_0 < \infty,\ \ V(\Phi) = \infty. \end{aligned} \tag{2.10}$$

Equations (2.9) and (2.10) define $\chi(\Phi)$ as a projection operator which suppresses those fields (nonrenormalizable) which cause the interaction to be infinite. This suggests that as $\lambda \to 0^+$ the supports for the free and interacting systems coincide, i.e.,

$$d\sigma(\Phi) = \chi(\Phi) d\sigma(\Phi) \to \chi(\Phi) d\sigma_F(\Phi) \equiv d\sigma_{PF}(\Phi),$$

and that

$$\begin{aligned} Z(h) \to \int \exp[i(h,\Phi)]\chi(\Phi) d\sigma_F(\Phi) &\equiv Z_{PF}(h) \\ &\neq Z_F(h). \end{aligned} \tag{2.11}$$

These equations formerly characterize the psuedofree system which is not equivalent to the free system. One sees that even as $\lambda \to 0^+$ the potential suppresses any contributions to $Z(h)$ from the nonrenormalizable fields.

Klauder has shown[1,12–14] that this analysis applies with full rigor to ultralocal field theories. One therefore expects full field theories to exhibit this same phenomena. If this is the case then $d\sigma(\Phi)$ cannot be treated simply as an infinite "constant" as is sometimes done.[15] The distinction between the psuedofree system and the customary free theory must be carefully investigated.

The Klauder phenomenon can also occur in the $\mathcal{L}_0$ part of the Lagrangian. It is well known that a massless particle, regardless of its spin, has only two helicity states, one aligned with and the other against its direction of motion. This is a typical example of the Klauder phenomenon with the mass parameter m acting as a Klauder projector $\chi(\Phi)$ on the helicity states. To see this, consider the free particle, "rest" frame field equation[16,17]

$$E\beta\varphi(x,m) = i\frac{\partial}{\partial t}\varphi(x,m), \tag{2.12}$$

where $E = (-\nabla^2 + m^2)^{1/2}$ and β is the $2(2s+1)$ generalization of the Dirac β matrix. Weaver, Hammer, and Good[16] have shown that a generalized Foldy—Wouthysen transformation exists which relates the "rest" frame field to the "lab" frame field

$$\psi(x,m) = [m^s E^{-1/2} S]\varphi(x,m), \tag{2.13}$$

where $\psi(x,m)$ satisfies a Schrödinger-like field equation with the Hamiltonian operator defined by

$$H = S(E\beta)S^{-1}. \tag{2.14}$$

Williams, Draayer, and Weber[18] (WDW) have examined the massless limit of these equations in considerable detail. They show that in the massless limit, when operating on $\psi(x,m)$, the Hamiltonian becomes

$$\operatorname*{s-lim}_{m\to 0^+} H = H_0, \tag{2.15}$$

$$H_0 = \alpha\cdot\mathbf{p}, \quad \alpha = -\gamma_5(\mathbf{s}/s), \tag{2.16}$$

where γ_5 is the $2(2s+1)$ generalization of the Dirac γ_5 matrix. This is the expected free particle Hamiltonian for a massless particle of spin s. The Klauder phenomena arises because not all solutions of H_0 survive the massless limit.

It is worth repeating the WDW argument to illustrate this point. They find

$$S = \cosh(\beta\hat{\mathbf{p}}\cdot\mathbf{s}\omega) - \gamma_5\sinh(\beta\hat{\mathbf{p}}\cdot\mathbf{s}\omega) \tag{2.17}$$

where

$$\exp\omega = (E+p)/m.$$

The "rest" frame field can be expanded in helicity states

$$\varphi(x,m) = (2\pi)^{-3/2}\sum_{k,\epsilon}\int d\mathbf{p}\, a_{k,\epsilon}(\epsilon\mathbf{p})u_{k,\epsilon}\exp[i\epsilon(\mathbf{p}\cdot\mathbf{x} - Et)], \tag{2.18}$$

where

$$a_{k,-}(-\mathbf{p}) = b_k^\dagger(\mathbf{p}),$$

$$\beta u_{k,\epsilon} = \epsilon u_{k,\epsilon}, \quad \epsilon = \pm 1,$$

$$\beta\mathbf{s}\cdot\hat{\mathbf{p}}u_{k,\epsilon} = k u_{k,\epsilon}, \quad k = s, s-1, \cdots, -s. \tag{2.19}$$

Therefore, when S acts on $\varphi(x,m)$, under the integral sign, $\beta\hat{\mathbf{p}}\cdot\mathbf{s}$ can be replaced everywhere by k, giving

$$S = \cosh(k\omega) - \gamma_5\sinh(k\omega),$$

$$= \frac{(1-\gamma_5)}{2}\left(\frac{E+p}{m}\right)^k + \frac{(1+\gamma_5)}{2}\left(\frac{E+p}{m}\right)^{-k}. \tag{2.20}$$

It is therefore clear that unless $k = \pm s$, the generalized Foldy—Wouthysen transformation, and therefore $\text{s-lim}_{m\to 0^+}\psi(x,m)$, vanishes in the massless limit,

$$\operatorname*{s-lim}_{m\to 0^+}(m^s E^{-1/2} S u_{k,\epsilon})$$

$$= (2p)^{s-1}[(1-\gamma_5)\delta_{s,k} + (1+\gamma_5)\delta_{s,-k}]u_{k,\epsilon},$$

and

$$\operatorname*{s-lim}_{m\to 0^+}\psi(x,m)$$

$$= (2\pi)^{-3/2}\sum_{\epsilon}\int d\mathbf{p}(2p)^{s-1}[(1-\gamma_5)a_{s,\epsilon}(\epsilon\mathbf{p})u_{s,\epsilon}$$

$$+ (1-\gamma_5)a_{-s,\epsilon}(\epsilon,\mathbf{p})u_{-s,\epsilon}]\exp[i\epsilon(\mathbf{p}\cdot\mathbf{x} - Et)]. \tag{2.21}$$

Note that this is an example of the Klauder phenomenon which takes place *inside* the $\mathcal{L}_0$-term, enforcing the fact that only helicity $k = \pm s$ states survive the massless limit.

3. ADDITIONAL EXAMPLES

Integral example

Consider the integral operator

$$H_0 = \int_{-\infty}^{\infty}\exp[-|x| - |y|](\cdot)dy \tag{3.1}$$

with the unperturbed eigenvalue equation

$$H_0\psi_n^{(0)}(x) = E_n^{(0)}\psi_n^{(0)}(x). \tag{3.2}$$

Let V denote the perturbing operator

$$V(x) = x. \tag{3.3}$$

Let λ denote a positive real parameter and define the full Hamiltonian H as the form sum

$$H = H_0 + \lambda V. \tag{3.4}$$

We impose the usual boundary conditions on both $\psi_n^{(0)}(x)$ and the eigenstates of the full Hamiltonian, $\psi_n(x,\lambda)$,

$$\lim_{\Omega\to\infty}|\Omega|^{-1}\int_\Omega dx|\varphi(x)|^2 < M, \tag{3.5}$$

where M is some positive, finite real number and

$$|\Omega| = \int_\Omega dx.$$

Note this integrability condition allows $\psi_n^{(0)}(x)$ and $\psi_n(x,\lambda)$ to be continuous as well as bound state solutions.

Proposition: The unperturbed Hamiltonian H_0 has two eigenvalues $E_0^{(0)} = 0$ and $E_1^{(0)} = 1$ (there is no continuum). The degeneracy of $E_0^{(0)}$ is infinite and the eigenvalue $E_1^{(0)}$ is nondegenerate.

Proof: The eigenvalue equation is

$$\int_{-\infty}^{\infty}\exp(-|x|-|y|)\,\psi_n^{(0)}(y)dy=E_n^{(0)}\psi_n^{(0)}(x). \tag{3.6}$$

It follows from Eq. (3.5) that a finite number b exists such that

$$b=\int_{-\infty}^{\infty}\exp(-|y|)\psi_n^{(0)}(y)dy. \tag{3.7}$$

Imposing this condition upon Eq. (3.6) then gives

$$E_n^{(0)}\psi_n^{(0)}(x)=be^{-|x|}. \tag{3.8}$$

The eigenvalues $E_n^{(0)}$ are obtained by substitution for $[E_n^{(0)}\psi_n^{(0)}(x)]$ from Eq. (3.8) into Eq. (3.7). This result is

$$\begin{aligned} E_n^{(0)}b&=b\int_{-\infty}^{\infty}dy\exp(-2|y|),\\ &=b. \end{aligned} \tag{3.9}$$

If $b\neq 0$, $E_1^{(0)}=1$ and the corresponding eigenvector is

$$\psi_1^{(0)}(x)=\exp(-|x|). \tag{3.10}$$

If $b=0$, the only nontrival solution ($\psi_n^{(0)}(0)\neq 0$, $\forall x$) is $E_0^{(0)}=0$ with the corresponding eigenvectors being all functions which satisfy Eq. (3.5) and for which

$$\int_{-\infty}^{\infty}dx\exp(-|x|)\psi_0^{(0)}(x)dx=0. \tag{3.11}$$

If one considers $x^n\exp(-|x|)$, one sees that the family of eigenvectors $\psi_0^{(0)}(x)$ is nonempty (n positive integers). Since there are an infinite number of n's, the degeneracy of $E_0^{(0)}$ is of infinite order.

Finally we note that Eq. (3.7) applies for continuous as well as for discrete solutions to Eq. (2.6). Since $E_1^{(0)}$ and $E_0^{(0)}$ are the only two possible eigenvalues, there are no continuous solutions.

Proposition: The full Hamiltonian has only the trivial eigenstate $\psi_n(x,\lambda)=0$ and the indeterminate eigenvalue $E_n(\lambda)$. The limit $\lambda\to 0^+$of $\psi_n(x,\lambda)$ therefore corresponds only to the trivial solution of the unperturbed Hamiltonian and Eq. (3.11). The eigenstate corresponding to $E_1^{(0)}$, as well as all nontrivial eigenstates corresponding to $E_0^{(0)}$have been lost. Thus the Klauder phenomenon occurs and

$$\text{s-lim}[H(\lambda)]=H_{PF}\neq H_0. \tag{3.12}$$

Proof: The full eigenvalue problem is

$$\int_{-\infty}^{\infty}\exp(-|x|-|y|)\psi_n(y,\lambda)dy+\lambda x\psi_n(x,\lambda)=E_n(\lambda)\psi_n(x,\lambda). \tag{3.13}$$

Proceeding as before, we note that if $\psi_n(x,\lambda)$ is to satisfy Eq. (3.5), a finite number d exists such that

$$d=\int_{-\infty}^{\infty}\exp(-|y|)\psi_n(y,\lambda)dy. \tag{3.14}$$

The full eigenvalue equation thus implies

$$\psi_n(x,\lambda)=-d\exp(-|x|/[\lambda x-E_n(\lambda)]. \tag{3.15}$$

The combination of Eqs. (3.14) and (3.15) give

$$d\int_{-\infty}^{\infty}dy\,\frac{\exp(-2|y|)}{E_n(\lambda)-\lambda y}=d. \tag{3.16}$$

If $d\neq 0$,

$$\int_{-\infty}^{\infty}dy\,\frac{\exp(-2|y|)}{E_n(\lambda)-\lambda y}=1. \tag{3.17}$$

However this equation has no solution for $E_n(\lambda)$, $\lambda\in\mathcal{D}(R)$. Thus the integrability condition forces the choice $d=0$, with the corresponding solution

$$\psi_n(x,\lambda)=0\quad(\forall x), \tag{3.18}$$

for any λ, however small.

It is interesting to note that if a perturbation apporach to the solution of Eq. (3.13) is attempted, a *formal series* is obtained,

$$E_1(\lambda)\overset{F}{=}[1+\tfrac{1}{2}\lambda^2+\cdots], \tag{3.19}$$

and

$$\psi_1(x,\lambda)\overset{F}{=}[1+x\lambda+(x^2-\tfrac{3}{4})\lambda^2+\cdots]\exp(-|x|), \tag{3.20}$$

which apparantly corresponds to $E_1^{(0)}$ and $\psi_1^{(0)}(x)$ in the limit $\lambda\to 0^+$. The "F" over the equal sign is defined as a *formal equality*, meaning that if the individual terms exist and if further their sum exists then the object on the left-hand side is equal to the object on the right-hand side. However, these series diverge for all λ, however small, as can be seen from examination of Eqs. (3.15) and (3.17). Thus no hint that the Klauder phenomenon is taking place can be obtained from such a perturbation approach.

This example can be studied further by considering a more general interaction $V=w(x)$ for Eq. (3.4), where $w(x)$ is a piecewise continuous function with compact support. Repeating the arguments from Eq. (3.13) to Eq. (3.16), one has

$$\psi_n(x,\lambda)=-d\exp(-|x|)/[\lambda w(x)-E_n(\lambda)], \tag{3.21}$$

and

$$d\int_{-\infty}^{\infty}dx\,\frac{\exp(-2|x|)}{E_n(\lambda)-\lambda w(x)}=d. \tag{3.22}$$

Again if $d=0$, only the trivial solution $\psi_n(x,\lambda)=0$ ($\forall x$) is obtained. However, if $d\neq 0$ then Eq. (3.22) can be written as

$$\int_{-\infty}^{\infty}\frac{dx\exp(-2|x|)}{[1-\mu w(x)]}=E_n(\lambda), \tag{3.23}$$

where $\mu=[\lambda/E_n(\lambda)]$.

If w_0 is given by

$$w_0=\sup_{x\in R}|w(x)|=\text{l.u.b.}_{w\in R}\,|w(x)|, \tag{3.24}$$

then solutions to Eq. (3.23) exist for all λ, $E_n(\lambda)$ such that $\mu<w_0$.

Note that a perturbation approach to this full Hamiltonian will result in *formal series* similar to Eqs. (3.19) and (3.20) that will converge. This follows directly from Eq. (3.21) which can be expanded in a convergent Taylor series if $\mu<w_0$ and from Eq. (3.23), the integrand of which can be expanded in a convergent series and integrated terms by term.

In summary, whenever the form sum $H_0 + \lambda V$ is *small in norm* no eigenvalues of H_0 are lost by adding the form perturbations. When the form sum is large in norm, some (possibly all!) eigenvalues can be lost even if the perturbation is turned off.

The trouble with forms

Since the boundary conditions are part of the very essence of a self-adjoint operator, the boundary conditions can force $\mathcal{D}(H_0) \cap \mathcal{D}(V)$ to contain only the zero vector of $\mathcal{H}$. For example, consider $\mathcal{H} = L^2(R)$; Ω a smooth open set in R^3 and

$$H_0 = \nabla^2, \tag{3.25}$$

and

$$V = \nabla^2, \tag{3.26}$$

where $H0$ satisfies Neumann boundary conditions and V satisfies Dirichlet boundary conditions.[19] Since

$$H = H_0 + \lambda V = (1+\lambda)V, \tag{3.27}$$

one gets the result that

$$\operatorname*{s-lim}_{\lambda\to 0^+} [H(\lambda)] = V \neq H_0. \tag{3.28}$$

Thus the boundary conditions have caused the domains to have so little in common that the Klauder phenomena occurs.

Operator extensions

Because of these examples one searches for alternate meanings for the "+" in $H_0 + \lambda V$. Operator extensions are one candidate. Let $\mathcal{J}_\lambda$ denote a family of self-adjoint extensions of $H_0 + \lambda V$ on $\mathcal{D}(H_0) \cap \mathcal{D}(V)$ so that

$$\operatorname*{s-lim}_{\lambda\to 0^+} [H(\lambda)] = H_{PF}. \tag{3.29}$$

Then the Klauder phenomenon occurs when $H_{PF} \neq H_0$. Consider $\mathcal{H} = L^2(R)$,

$$H_0 = -\frac{d^2}{dx^2} + x^2, \tag{3.30}$$

and

$$V = 1/|x|^\gamma. \tag{3.31}$$

The example given in the Introduction corresponds to $\gamma = 3$. Now for $-2 < \gamma < 1$, Eq. (3.31) is a form which is small in the norm[1,2] and no Klauder phenomenon can occur. For $\gamma \geq 1$, all vectors $\psi(x) \in \mathcal{D}(H_0)$ satisfy a Holder condition so that

$$\int |x|^{-\gamma} |\psi(x)|^2 d\mu(x) < \infty, \tag{3.32}$$

which requires Dirichlet boundary conditions, i.e.,

$$\psi(0) = 0. \tag{3.33}$$

Therefore, one has

$$q(H_0) \neq H_0. \tag{3.34}$$

However $q(H_0)$ is that subset of $\mathcal{D}(H_0)$ that satisfies Eq. (3.33). Therefore an alternative definition to the meaning of "+" in $H_0 + \lambda V$ used thus far, namely that $\mathcal{D}(H_0) \cap \mathcal{D}(V)$ be well defined, is that of operator extensions, e.g., $\mathcal{D}(H_0 + \lambda V) \subset \mathcal{D}(H_0)$ as in Klauder's example given in Sec. 1.

Note that for $\gamma > 2$ the interaction is nonrenormalizable. A particularly clear discussion of these cases is given by Aly and Taylor.[9]

4. CONCLUSIONS

Since nonrenormalizable interactions are very much a part of present day physical theories, e.g., nongauge weak interactions, charged massive vector mesons, and so on, it seems rather clear that perturbation approaches and cutoff dependent theories do not deserve such widespread usage. Also, the measure questions which are so glibly assumed away in *gauge theories* are complicated and make a significant difference in the physical fields in nonrenormalizable cases. Furthermore, the domains of essential self-adjointness of the unbounded operators are not a mathematical luxury but rather are a physical necessity, even for such venerable problems as the massless limit of a general spin free particle equation.

ACKNOWLEDGMENTS

One of us (BDF) wishes to thank the Ames Laboratory-USAEC and the Iowa State University Physics Department for their hospitality during Summer 1973 when this work was done. Both of us thank Dr. Klauder for pointing out Professor Simon's work and Professor Bing-Lin Young and Professor Derek Pursey for useful discussions.

APPENDIX

This appendix contains the definitions required for our discussion of unbounded operators and the Klauder phenomenon. For a comprehensive introduction to unbounded operators, we refer the interested reader to Simon[20] and Reed and Simon.[21]

(1) The *graph* of the operator A, written as $\Gamma(A)$, is the set of all ordered pairs

$$\Gamma(A) = \{\varphi, A\varphi \mid \varphi \in \mathcal{D}(A)\}$$

and is a subset of $\mathcal{H} \times \mathcal{H}$.

(2) The operator A is called a *closed operator* if $\Gamma(A)$ is closed in $\mathcal{H} \times \mathcal{H}$.

(3) The *adjoint* of an operator A, written as A^*, is defined as $(\Phi, A\psi) = (A^*\Phi, \psi)$; $\forall\, \Phi$, $\psi \in \mathcal{D}(A) \supset \mathcal{D}(A^*)$

(4) If A_1 and A are operators on $\mathcal{H}$ and $\Gamma(A_1) \subset \Gamma(A)$, then A is called the *extension* of A_1, written as $A_1 \subset A$. Also, A_1 is called the restriction of A to $\mathcal{D}(A_1)$, written as $A|\mathcal{D}(A_1)$.

(5) An operator A is called *closable* if it has a closed extension. The smallest closed extension is called the closure of A and is written as $\bar{A}$.

(6) An operator P is called *positive* if

$$(\varphi, P\varphi) \geq 0 \quad \text{for each } \varphi \in \mathcal{D}(P).$$

(7) An operator H is called *Hermitian* (or symmetric) if

$$\mathcal{D}(H) \subset \mathcal{D}(H^*), \quad H\varphi = H^*\varphi$$

for each $\varphi \in \mathcal{D}(H)$.

(8) An operator H is called *self-adjoint* if

$$H = H^*, \quad \mathcal{D}(H) = \mathcal{D}(H^*)$$

Remark: One Hermitian operator may have several self-adjoint extensions; see esp. pp. 204, 205 of Ref. 6 for a discussion of this point.

(9) A Hermitian operator H is called *essentially self-adjoint* if its closure $\bar{H}$ is self-adjoint.

(10) If H is a closed operator, the subset $\mathcal{D}_c \subset \mathcal{D}(H)$ is called a *core* or domain of essential self-adjointness if $H|\mathcal{D}_c = H$.

(11) An operator S is called *densely defined* in $\mathcal{H}$ if $\mathcal{D}(S)$ is dense in $\mathcal{H}$.

(12) The *resolvent* set $\rho(H)$ of H in $\mathcal{H}$ is the set of all complex numbers z such that $(z1 - H)$ has an inverse.

(13) The *resolvent operator* $R(z)$ of H in $\mathcal{H}$ at z is

$$R_z(H) = (z1 - H)^{-1}$$

(14) The *spectrum* $\sigma(z)$ is the complement of the resolvent set $\rho(H)$ in $\mathbf{C}$.

(15) The *quadratic form* q is a map $\mathcal{D}(q) \times \mathcal{D}(q) \to \mathbf{C}$, where $\mathcal{D}$ (q) is called the *form domain* and is a dense linear subset of $\mathcal{H}$. For each φ, $\psi \in \mathcal{D}$ (q) the *sesquilinear form*

$$q(\varphi, \psi) \to C$$

is conjugate linear in ψ and conjugate antilinear in φ. The form q is called *semibounded* if

$$q(\varphi, \varphi) \geq 0$$

for each $\varphi \in \mathcal{D}$ (q).

(16) The quadratic form q_A associated with the self-adjoint operator A on $\mathcal{D}(A) \subset \mathcal{H} = L^2(\mathbf{R}, \mu)$ is

$$q_A(\varphi, \psi) = (\varphi, A\psi)$$

for each $\varphi, \psi \in \mathcal{D}(A)$. The quantity $L^2(\mathbf{R}, \mu)$ is the real Lebesgue square integrable functions with measure μ.

(17) If H_0 and V are positive self-adjoint operators, if $H = H_0 + \lambda V$, and if $\mathcal{D}(H) = \mathcal{D}(H_0) \cap D(V)$ is dense in $\mathcal{H}$, then the form sum $(H_0 + \lambda V)$ is

$$q_H'(\varphi, \psi) = (H_0^{1/2}\varphi,\ H_0^{1/2}\varphi) + \lambda(V^{1/2}\psi,\ V^{1/2}\psi)$$

for each $\varphi \in \mathcal{D}(H_0)$ and for each $\psi \in \mathcal{D}(V)$.

(18) The operator T is called *normal* if $T^*T = TT^*$. Every normal operator can be written uniquely as

$$T = \mathrm{Re}(T) + i\,\mathrm{Im}(T)$$

where $\mathrm{Re}(T)$ and $\mathrm{Im}(T)$ are Hermitian.

(19) The complex number γ is called a *vertex* and the real number θ is a semiangle for a normal operator T if

$$|\arg(z - \gamma)| \leq \theta < \pi/2$$

whenever the complex number z is in the spectrum of T, i.e., $\mathrm{Spec}(T) = \{z \,|\, (z - T)^{-1} \text{ does not exist}\}$.

(20) Given $H_0 + \lambda V$ and that $\mathcal{D}(H) = \mathcal{D}(H_0) \cap \mathcal{D}(V)$ is dense in $\mathcal{H}$ the maximal self-adjoint extension of H on $\mathcal{D}(H)$ is written as

$$H \upharpoonright \mathcal{D}(H),$$

or

$$(H_0 + \lambda V) \upharpoonright \mathcal{D}(H_0) \cap \mathcal{D}(V).$$

(21) The sequence of operators T_n *uniformly* or *norm converge* to the operator T

$$n\text{-}\lim_{n\to\infty} T_n = T$$

if for $\epsilon > 0$

$$\|T_n - T\| < \epsilon$$

whenever $n > N$, where $\|\cdot\|_{\mathcal{O}}$ is the operator norm, e.g., $\mathrm{Tr}(\omega T^*T)$ for ω a real function.

(22) The sequence of operators T_n are *strongly convergent* to the operator T,

$$\text{s-}\lim_{n\to\infty} T_n = T,$$

if for $\epsilon > 0$ and $\psi \in \mathcal{D}(T) \supset \mathcal{D}(T_n)$,

$$\|(T_n - T)\psi\|_{\mathcal{H}} < \epsilon,$$

whenever $n > N$, where $\|\cdot\|_{\mathcal{H}}$ is the Hilbert space norm of $\mathcal{H}$.

(23) The sequence of operators T_n *weakly converge* to the operator T

$$\text{w-}\lim_{n\to\infty} T_n = T$$

if for $\epsilon > 0$, and $\psi, \varphi \in \mathcal{D}(T) \supset \mathcal{D}(T_n)$

$$|(\varphi, T_n\psi) - (\varphi, T\psi)| < \epsilon$$

whenever $n > N$, where (,) is the inner product of $\mathcal{H}$.

Remark: Every norm convergent sequence of operators is strongly convergent and every strongly convergent sequence of operators is weakly convergent. That the converse propositions do not hold can be seen from the next two examples.

*Example A*1: A strongly convergent sequence which is *not* norm convergent. Let $|n\rangle$, $n \in Z^+$, be a complete orthonormal basis for $\mathcal{H}$ and let P_n be the projection operator onto the n-dimensional subspace spanned by $|1\rangle \cdots |n\rangle$, e.g., $P_n|r\rangle = |r\rangle$, $r \leq n$, and $P_n|r = 0$, $r > n$. Then

$$\text{s-}\lim_{n\to\infty}(P_n) = 1$$

but

$$\|P_n - P_{n+1}\| = 1$$

for each n so that the norm limit diverges, or equivalently, no real positive $\epsilon < 1$ bounds $\|P_n - P_{n-1}\|$ for any n.

Example A2: A weakly convergent sequence which is not strongly convergent. Keep $|n>$ as in the last example and let S_r denote the shift operator by r units ($r \in Z^+$). Then

$$S_r|n> = |n+r>, \quad \|S_r - S_{r+1}\|_{\mathcal{H}} = 2$$

whereas

$$\text{w-lim}_{r\to\infty}(S_r) = 0.$$

*Present address: Department of Physics, University of Missouri-Columbia, Columbia, Missouri 65201.

[1]J.R. Klauder, Acta Phys. Austr. Supp. XI, 341 (1973).
[2]J.R. Klauder, Phys. Lett. B 47, 523 (1973).
[3]B. Simon. J. Funct. Anal. 14, 295 (1973).
[4]F. Riesz and B. Sz-Nagy, *Functional Analysis* (Ungar, New York, 1955), p. 296 *et seq.*
[5]T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1966), p. 308 *et seq.*
[6]B. Simon, *Quantum Mechanics for Hamiltonians Defined as Quadratic Forms* (Princeton U.P., Princeton, N.J., 1971).
[7]Ref. 6 is an excellent second or third semester graduate quantum mechanics text or reference. One starts with the Appendix, "Some Mathematical Background," pp. 201—22 and then resumes with page 1. A rigorous treatment is given of one- and many-particle Schrödinger systems including negative and positive energies and including a number of new results.
[8]After we had completed the present work, Dr. J.R. Klauder kindly communicated a Marseille preprint by Professor B. Simon, published since then (see Ref. 3), which overlaps our results somewhat.
[9]H.H. Aly and J.G. Taylor, J. Math. Phys. 9, 781 (1968).
[10]H.H. Aly, H.J.W. Muller, and N. Vahedi-Faradi, Lett. Nuovo Cimento (Sene 1) 2, 485 (1969).
[11]H.H. Aly and F. Odeh, Lett. Nuovo Cimento 2, 193 (1971).
[12]J.R. Klauder, Commun. Math. Phys. 18, 307 (1970).
[13]J.R. Klauder, Acta Physica Austr. Suppl. VIII, 227 (1971).
[14]J.R. Klauder, 1971 Boulder Summer School (to be published).
[15]B.W. Lee, Phys. Rev. D 5, 823 (1972).
[16]D.L. Weaver, C.L. Hammer, and R.H. Good, Jr., Phys. Rev. 135, 241 (1964).
[17]T.J. Nelson and R.H. Good, Jr., Rev. Mod. Phys. 80, 508 (1968).
[18]S.A. Williams, J.P. Draayer, and T.A. Weber, Phys. Rev. 152, 1207 (1966).
[19]Discovered first by Simon and independently by us.
[20]B. Simon, in *Mathematics of Contemporary Physics*, edited by R.F. Streater (Academic, New York, 1972).
[21]M. Reed and B. Simon, *Functional Analysis* (Academic, New York, 1972). See especially Chap. VIII, p. 249, for a discussion of unbounded operators.

Polarization of electrons in screened Coulomb scattering

F. L. Ridener Jr. and R. H. Good Jr.

Ames Laboratory-USAEC and Department of Physics, Iowa State University, Ames, Iowa 50010

Department of Physics, The Pennsylvania State University, University Park, Pennsylvania 16802
(Received 23 January 1974)

In this paper the effects of screened-Coulomb scattering on electron polarization are calculated. The results apply to elastic electron scattering by light atoms at moderate energies. The calculation is carried out to second order in αZ; in this order the change in magnitude of the polarization first occurs. For arbitrary initial polarization, the final polarization vector is expressed explicitly as a function of initial polarization, momentum, and scattering angle. The effects of screening on polarization is discussed. In the low energy region, a few hundred eV, there is a significant effect on the output polarization of an input unpolarized beam.

I. INTRODUCTION

The problem of calculating polarization effects in scattering by the screened Coulomb potential $Ze\, e^{-\lambda r}/4\pi r$ is interesting because it gives an approximation to the elastic atomic scattering problem and because of the insight it gives into the Coulomb scattering problem.

To second order the problem can be solved in terms of elementary functions; the complete results are given below. An arbitrarily polarized input beam is considered and the polarization of the output beam is calculated, as a function of the input polarization, momentum, and scattering angle, to second order in αZ. It is essential to go to this order to begin to understand the change in magnitude of the polarization. In first order there is only a rotation of the polarization 3-vector about the normal to the scattering plane. The second order contribution consists of a correction to that rotation angle, an additional rotation toward the normal, and a change in magnitude of the polarization. For small αZ, the higher order terms are expected to be small with no qualitatively new effects coming in.

This calculation proceeds from a much earlier one by Dalitz.[1] He obtained the cross section for scattering of unpolarized beams in this order.

The limit of small λ, properly taken, gives the Coulomb scattering results. A second order calculation with a pure Coulomb field leads to divergent matrix elements for the scattering. As shown by Dalitz,[1] the divergence is circumvented by calculating the probabilities in the screened Coulomb problem and only then taking the limit of small λ.

The results which we derive below for the Coulomb scattering agree with those obtained earlier by Gürsey[2]; he treated the case of longitudinally polarized electrons incident.

II. TRANSITION OPERATOR

We shall use Bjorken and Drell's notation and follow their presentation of scattering theory.[3]

The initial wavefunction is

$$\psi_i = (m/E_i V)^{1/2} u(p_i, s_i) \exp(-ip_i \cdot y). \tag{1}$$

The state $\psi(x)$ that evolves from this is given by (Ref. 3, p. 96):

$$\psi(x) = \psi_i(x) + e \int d^4y S_F(x-y)\gamma\cdot A(y)\psi(y), \tag{2}$$

where S_F is the Feynman propagator and $A(y)$ is the external 4-potential. To two orders the scattered wave is then

$$\begin{aligned}\psi_f(x) &= e\int d^4y\, S_F(x-y)\gamma\cdot A(y)\psi_i(y)\\ &\quad + e^2\int d^4y\, S_F(x-y)\gamma\cdot A(y)\\ &\quad \int d^4z\, S_F(y-z)\gamma\cdot A(z)\psi_i(z).\end{aligned} \tag{3}$$

Here one sets

$$\mathbf{A}(y)=0, \quad A_0 = -(Ze/4\pi|\mathbf{y}|)\exp(-\lambda|\mathbf{y}|) \tag{4}$$

and uses the formula (Ref. 3, p. 95)

$$S_F(x-y) = \int \frac{d^4p}{(2\pi)^4}\frac{\exp[-ip\cdot(x-y)]}{p^2-m^2+i\epsilon}(\gamma\cdot p+m). \tag{5}$$

After simplifying in the usual way, one finds that

$$\begin{aligned}\psi_f(x) &= \int \frac{d^4p_f}{(2\pi)^3}\frac{\exp(-ip_f\cdot x)}{p_f^2-m^2+i\epsilon}(\gamma\cdot p_f+m)\delta(E_f-E_i)\\ &\quad \times M(p_f,p_i)\left(\frac{m}{E_iV}\right)^{1/2} u(p_i,s_i),\end{aligned} \tag{6}$$

where

$$\begin{aligned}M(p_f,p_i) &= -Ze^2\bigg(\frac{\gamma^0}{\lambda^2+|\mathbf{p}_f-\mathbf{p}_i|^2} - \frac{Ze^2}{(2\pi)^3}\int d^3q\frac{1}{\lambda^2+|\mathbf{p}_f-\mathbf{q}|^2}\\ &\quad \times \frac{E_i\gamma^0+\mathbf{q}\cdot\boldsymbol{\gamma}+m}{E_i^2-|\mathbf{q}|^2-m^2+i\epsilon}\,\frac{1}{\lambda^2+|\mathbf{q}-\mathbf{p}_i|^2}\bigg).\end{aligned} \tag{7}$$

This transition operator M is related to the Bjorken and Drell S-matrix element S_{fi} by

$$S_{fi} = -\frac{i}{V}\left(\frac{m^2}{E_fE_i}\right)^{1/2} 2\pi\delta(E_f-E_i)\bar{u}(p_fs_f)Mu(p_is_i).$$

In the integration over p_f^μ in Eq. (6), there is a contribution only when $E_i = E_f$ and, because of the pole in the denominator, only when $E_f = (|\mathbf{p}_f|^2+m^2)^{1/2}$. Consequently there is really only an integration over the direction of $\mathbf{p}_f$ involved, where the output wavefunction in the direction $\mathbf{p}_f/|\mathbf{p}_f|$ is of the form

$$\begin{aligned}\psi_{f\mathbf{p}}(x) &= C\exp(-ip_f\cdot x)(\gamma\cdot p_f+m)\\ &\quad \times M(p_f,p_i)(m/E_fV)^{1/2}u(p_is_i),\end{aligned} \tag{8}$$

C being a normalization factor.

The integrals that occur in Eq. (7) are

$$I=\int\frac{d^3q}{(\lambda^2+|\mathbf{p}_f-\mathbf{q}|^2)(\lambda^2+|\mathbf{q}-\mathbf{p}_i|^2)(p^2-q^2+i\epsilon)},$$

$$\mathbf{K}=\int\frac{\mathbf{q}d^3q}{(\lambda^2+|\mathbf{p}_f-\mathbf{q}|^2)(\lambda^2+|\mathbf{q}-\mathbf{p}_i|^2)(p^2-q^2+i\epsilon)},$$

where p denotes $(E_i^2-m^2)^{1/2}$ and q is $|\mathbf{q}|$. The lengths $|\mathbf{p}_i|$ and $|\mathbf{p}_f|$ are both p. (In this respect we depart from Bjorken and Drell's notation.) In the Appendix it is shown that $\mathbf{K}$ is of the form $(\mathbf{p}_i+\mathbf{p}_f)\,J$ and the complete evaluation of I and J in terms of elementary functions is given, Eqs. (A17) and (A24). In terms of I and J the transition operator is

$$M(p_f,p_i)=-Ze^2\left(\frac{\gamma^0}{\lambda^2+|\mathbf{p}_f-\mathbf{p}_i|}-\frac{Ze^2}{(2\pi)^3}(E_i\gamma^0+m)I\right.$$
$$\left.-\frac{Ze^2}{2(2\pi)^3}\gamma\cdot(\mathbf{p}_i+\mathbf{p}_f)J\right). \tag{9}$$

The matrix dependence in M can be simplified using the fact that M, as it occurs in Eq. (6), only relates electron wavefunctions. That is, since

$$\gamma\cdot\mathbf{p}_i\,u(p_i,s_i)=(m-E\gamma^0)u(p_i,s_i),$$

$$(\gamma\cdot p_f+m)\gamma\cdot\mathbf{p}_f=(\gamma\cdot p_f+m)(m-E\gamma^0),$$

one can replace $\gamma\cdot(\mathbf{p}_i+\mathbf{p}_f)$ in Eq. (9) by $2(m-E\gamma^0)$. An equivalent expression for M is then

$$M=-Ze^2\left[\frac{\gamma^0}{\lambda^2+4p^2\sin^2\frac{1}{2}\theta}-\frac{Ze^2}{(2\pi)^3}E\gamma^0(I+J)-\frac{Ze^2}{(2\pi)^3}m(I-J)\right]; \tag{10}$$

here θ is the angle between $\mathbf{p}_i$ and $\mathbf{p}_f$.

An alternative way of writing this expression, valid to second order, is

$$M=-\frac{Ze^2}{\lambda^2+4p^2\sin^2\frac{1}{2}\theta}$$
$$\exp\left[-\frac{Ze^2}{(2\pi)^3}2EI(\lambda^2+4p^2\sin^2\tfrac{1}{2}\theta)\right]$$
$$\cdot\left[\gamma^0+\frac{Ze^2}{(2\pi)^3}E\gamma^0(\lambda^2+4p^2\sin^2\tfrac{1}{2}\theta)(I-J)\right. \tag{11}$$
$$\left.-\frac{Ze^2}{(2\pi)^3}m(\lambda^2+4p^2\sin^2\tfrac{1}{2}\theta)(I-J)\right].$$

There is freedom in choosing the exponential that is factored out at this point. This particular choice was arrived at in a pragmatic way by finding an exponent that makes the term of order $(\alpha Z)^1$ in the final equation for the helicity amplitude ratio, Eq. (17) below, small compared to the $(\alpha Z)^0$ term. This choice of exponential factor agrees with Dalitz[1] in the Coulomb limit where the exponent is proportional to $\log\lambda$. Discussions of this factor were given recently by Gasiorowicz[4] and Überall.[5]

Some remarks can be made about the transition operator in the general form

$$M=C(\gamma^0+a) \tag{12}$$

independent of the expansion on αZ. For any wavefunction of the form of Eq. (1) the amplitude is

$$u=\begin{pmatrix}\chi\\ [\sigma\cdot\mathbf{p}/(E+m)]\chi\end{pmatrix}\left(\frac{E+m}{2}\right)^{1/2}$$

and the polarization of the state is given by the spin of the two-component χ. In the present case Eqs. (8) and (12) imply that the initial and final amplitudes are related by

$$u_f=C(\gamma\cdot p_f+m)(\gamma_0+a)u_i$$

and so the spin functions are related by

$$\chi_f=M\chi_i,$$

where

$$M=C\left(1+\frac{1-a}{1+a}\,\frac{E-m}{E+m}\sigma\cdot\hat{\mathbf{p}}_f\sigma\cdot\hat{\mathbf{p}}_i\right). \tag{13}$$

This M is two-by-two and this coefficient C accumulates several factors. Alternatively one may write

$$M=C\{f_{++}\exp[-\tfrac{1}{2}i\theta\mathbf{n}\cdot\sigma]-f_{+-}\exp[-\tfrac{1}{2}i(\theta+\pi)\mathbf{n}\cdot\sigma]\},$$

where the f's are helicity amplitudes in the notation of Martin and Spearman[6] and $\mathbf{n}$ is the unit vector in the direction of $\mathbf{p}_i\times\mathbf{p}_f$. Only the ratio of the terms can affect the polarization; the parameter

$$\eta=f_{+-}/f_{++}$$

is suggested. In terms of η the matrix is

$$M=C\exp(-\tfrac{1}{2}i\theta\mathbf{n}\cdot\sigma)(1+i\eta\mathbf{n}\cdot\sigma) \tag{14}$$

and by comparing Eqs. (13) and (14) one finds that

$$\eta=\frac{aE+m}{E+am}\tan\tfrac{1}{2}\theta. \tag{15}$$

We return now to the specific problem. Equation (11) implies

$$a=\frac{-[Ze^2/(2\pi)^3]m(\lambda^2+4p^2\sin^2\frac{1}{2}\theta)(I-J)}{1+[Ze^2/(2\pi)^3]E(\lambda^2+4p^2\sin^2\frac{1}{2}\theta)(I-J)}, \tag{16}$$

and Eq. (15) gives

$$\eta=\frac{m\tan\frac{1}{2}\theta}{E+(\alpha Z/2\pi^2)(E^2-m^2)(\lambda^2+4p^2\sin^2\frac{1}{2}\theta)(I-J)}, \tag{17}$$

where α is $e^2/4\pi$. The expansion for small αZ gives here

$$\eta=\frac{m}{E}\tan\tfrac{1}{2}\theta\left(1-\frac{\alpha Z}{2\pi^2}\,\frac{(E^2-m^2)}{E}(\lambda^2+4p^2\sin^2\tfrac{1}{2}\theta)(I-J)\right). \tag{18}$$

III. FINAL POLARIZATION

The final and initial density matrices for the polarization are related by

$$\rho_f=M\rho_iM^\dagger.$$

In general such a density matrix is written as $\frac{1}{2}(1+\sigma\cdot\zeta)$, where ζ is the polarization 3-vector for the state. Accordingly, one writes, using Eq. (14) for M,

$$1+\sigma\cdot\zeta_f=C(1+i\eta\mathbf{n}\cdot\sigma)\exp(-\tfrac{1}{2}i\theta\mathbf{n}\cdot\sigma)(1+\sigma\cdot\zeta_i)$$

$$\times\exp(\tfrac{1}{2}i\theta\mathbf{n}\cdot\sigma)(1-i\eta^*\mathbf{n}\cdot\sigma). \tag{19}$$

By simplifying the right side into an expression linear in σ, one finds C and the relation between the initial and final polarizations.

A standard form for organizing polarization results in this type of problem[7] is

$$\begin{aligned}\boldsymbol{\zeta}_f=(1+S\mathbf{n}\cdot\boldsymbol{\zeta}_i)^{-1}\{&(S+\mathbf{n}\cdot\boldsymbol{\zeta}_i)\mathbf{n}\\ &+T[\cos\theta\,\mathbf{n}\times\boldsymbol{\zeta}_i+\sin\theta\,\mathbf{n}\times(\mathbf{n}\times\boldsymbol{\zeta}_i)]\\ &+U[\sin\theta\,\mathbf{n}\times\boldsymbol{\zeta}_i-\cos\theta(\mathbf{n}\times(\mathbf{n}\times\boldsymbol{\zeta}_i)]\}.\end{aligned} \tag{20}$$

Here S, T, U are functions of the energy and scattering angle related by

$$S^2+T^2+U^2=1. \tag{21}$$

The quantities T and $-U$ give the components of $\boldsymbol{\zeta}_f$ in the directions $\cos\theta\,\mathbf{n}\times\boldsymbol{\zeta}_i+\sin\theta\,\mathbf{n}\times(\mathbf{n}\times\boldsymbol{\zeta}_i)$ and $-\sin\theta\,\mathbf{n}\times\boldsymbol{\zeta}_i+\cos\theta\,\mathbf{n}\times(\mathbf{n}\times\boldsymbol{\zeta}_i)$; these are the vectors $\mathbf{n}\times\boldsymbol{\zeta}_i$ and $\mathbf{n}\times(\mathbf{n}\times\boldsymbol{\zeta}_i)$ rotated about $\mathbf{n}$ through the scattering angle θ. Another standard way to express the results[2] relates the components of $\boldsymbol{\zeta}_f$ in a coordinate system built from $\mathbf{n}$ and $\mathbf{p}_f$ to the components of $\boldsymbol{\zeta}_i$ in a coordinate system built from $\mathbf{n}$ and $\mathbf{p}_i$:

$$\begin{aligned}\boldsymbol{\zeta}_f=(1+S\boldsymbol{\zeta}_i\cdot\mathbf{n})^{-1}\{&(S+\mathbf{n}\cdot\boldsymbol{\zeta}_i)\mathbf{n}\\ &+[U\boldsymbol{\zeta}_i\cdot(\mathbf{n}_i\times\mathbf{n})-T\boldsymbol{\zeta}_i\cdot\mathbf{n}_i]\mathbf{n}_f\times\mathbf{n}\\ &+[T\boldsymbol{\zeta}_i\cdot(\mathbf{n}_i\times\mathbf{n})+U\boldsymbol{\zeta}_i\cdot\mathbf{n}_i]\mathbf{n}_f\},\end{aligned} \tag{22}$$

where $\mathbf{n}_i$ and $\mathbf{n}_f$ are unit vectors in the $\mathbf{p}_i$ and $\mathbf{p}_f$ directions.

In the present problem Eq. (19) leads directly to Eq. (20) with

$$S=-2\,\mathrm{Im}\,\eta/(1+|\eta|^2), \tag{23a}$$

$$T=-2\,\mathrm{Re}\,\eta/(1+|\eta|^2), \tag{23b}$$

$$U=(1-|\eta|^2)/(1+|\eta|^2). \tag{23c}$$

The problem is in principle solved. Equation (22) gives the final polarization; Eqs. (23) gives S, T, U in terms of η; Eq. (18) gives η in terms of $I-J$; I and J are given in Eqs. (A17) and (A24).

IV. DISCUSSION

A. General remarks

For comparison we shall quote the results for cross section in second order. For scattering of an unpolarized beam without observation of the final polarization, Dalitz[1] found

$$\begin{aligned}\left(\frac{d\sigma}{d\Omega}\right)_{\mathrm{unp}}=\frac{4\alpha^2Z^2E^2}{(\lambda^2+4p^2\sin^2\frac{1}{2}\theta)^2}\Big\{&(1-\beta^2\sin^2\tfrac{1}{2}\theta)\\ &\times\left[1-\left(\frac{\alpha ZE}{\pi^2}\right)(\lambda^2+4p^2\sin^2\tfrac{1}{2}\theta)\,\mathrm{Re}(I+J)\right]\\ &-\left(\frac{\alpha Zm^2}{\pi^2E}\right)(\lambda^2+4p^2\sin^2\tfrac{1}{2}\theta)\,\mathrm{Re}(I-J)\Big\}.\end{aligned} \tag{24}$$

For scattering of a particle with polarization $\boldsymbol{\zeta}_i$ into any final polarization $\boldsymbol{\zeta}_f'$ the cross section is

$$d\sigma(\boldsymbol{\zeta}_i,\boldsymbol{\zeta}_f',\mathbf{n})/d\Omega=(d\sigma/d\Omega)_{\mathrm{unp}}(1+S\,\mathbf{n}\cdot\boldsymbol{\zeta}_i)\tfrac{1}{2}(1+\boldsymbol{\zeta}_f\cdot\boldsymbol{\zeta}_f'), \tag{25}$$

where $\boldsymbol{\zeta}_f$ is understood to be written in terms of $\boldsymbol{\zeta}_i$, as, for example, in Eq. (20) or (22). From the standard form for the polarization, Eq. (22), the degree of output polarization is found to be

$$\zeta_f=[(1-\zeta_i^2)(S^2-1)/(1+S\,\boldsymbol{\zeta}_i\cdot\mathbf{n})^2+1]^{1/2}. \tag{26}$$

If the input particle is in a pure state so that $\zeta_i=1$, then the output particle also is in a pure state.

In the especially interesting case of an unpolarized incident beam, $\boldsymbol{\zeta}_i=0$, the output polarization is

$$\boldsymbol{\zeta}_f=S\mathbf{n}. \tag{27}$$

B. The Coulomb limit

The screening constant λ appears in the cross section and polarization results always in the ratio λ/p. The limit as λ/p goes to zero is identified as the Coulomb scattering.

For small λ/p Eqs. (A17) and (A24) lead easily to

$$I=\frac{\pi^2 i}{2p^3\sin^2\frac{1}{2}\theta}\ln\frac{\lambda}{2p}-\frac{\pi^2 i}{2p^3\sin^2\frac{1}{2}\theta}\ln(\sin\tfrac{1}{2}\theta), \tag{28}$$

$$\begin{aligned}J=&\frac{\pi^2 i}{2p^3\sin^2\frac{1}{2}\theta}\ln\frac{\lambda}{2p}\\ &-\frac{\pi^2 i}{2p^3\sin^2\frac{1}{2}\theta\cos^2\frac{1}{2}\theta}\ln(\sin\tfrac{1}{2}\theta)\\ &+\frac{\pi^3}{4p^3\cos^2\frac{1}{2}\theta}(1-\csc\tfrac{1}{2}\theta).\end{aligned} \tag{29}$$

Here one sees that I and J and hence M diverge as $\ln(\lambda/p)$. However cross section and polarization results depend only on the real parts of I and J and $Im(I-J)$ and so are finite.

The Coulomb cross section in this order results from using Eqs. (28) and (29) in Eq. (24):

$$\left(\frac{d\sigma}{d\Omega}\right)_{\mathrm{unp}}=\left(\frac{\alpha^2Z^2(1-\beta^2)}{4m^2\beta^4}\right)\times[\csc^4\tfrac{1}{2}\theta-\beta^2\csc^2\tfrac{1}{2}\theta+\pi\alpha Z\beta(\csc^3\tfrac{1}{2}\theta-\csc^2\tfrac{1}{2}\theta)]. \tag{30}$$

To get the polarization effects, one combines Eqs. (18), (28), and (29) to get, in the $\lambda/p\to 0$ limit,

$$\eta=(m/E)\tan\tfrac{1}{2}\theta\{1+\tfrac{1}{2}\alpha Z\beta\tan^2\tfrac{1}{2}\theta[\pi(1-\csc\tfrac{1}{2}\theta)-2i\ln(\sin\tfrac{1}{2}\theta)]\}. \tag{31}$$

Finally Eqs. (23) give, for the Coulomb limit,

$$S=\frac{\alpha Z\beta(1-\beta^2)^{1/2}\tan^3\frac{1}{2}\theta\ln(\sin^2\frac{1}{2}\theta)}{1+(1-\beta^2)\tan^2\frac{1}{2}\theta}, \tag{32a}$$

$$T=-\frac{2(1-\beta^2)^{1/2}\tan\frac{1}{2}\theta}{1+(1-\beta^2)\tan^2\frac{1}{2}\theta}+\frac{\pi\alpha Z\beta(1-\beta^2)^{1/2}\tan^3\frac{1}{2}\theta(\csc\frac{1}{2}\theta-1)}{1+(1-\beta^2)\tan^2\frac{1}{2}\theta} \tag{32b}$$

$$\times\frac{1-(1-\beta^2)\tan^2\frac{1}{2}\theta}{1+(1-\beta^2)\tan^2\frac{1}{2}\theta},$$

$$U=\frac{1-(1-\beta^2)\tan^2\frac{1}{2}\theta}{1+(1-\beta^2)\tan^2\frac{1}{2}\theta}+\frac{\pi\alpha Z\beta(1-\beta^2)^{1/2}\tan^3\frac{1}{2}\theta(\csc\frac{1}{2}\theta-1)}{1+(1-\beta^2)\tan^2\frac{1}{2}\theta}\times\frac{2(1-\beta^2)^{1/2}\tan\frac{1}{2}\theta}{1+(1-\beta^2)\tan^2\frac{1}{2}\theta}. \tag{32c}$$

From this result it is easy to see qualitatively the amount of polarization that you get from Coulomb scattering. The function S, as given by Eq. (32a), is zero at $\beta=0$ and $\beta=1$ and is zero at $\theta=0$ and $\theta=\pi$. For intermediate values it varies smoothly and has a minimum value near αZ. A graphical discussion of S, T, and U is given in the next subsection.

At high energy, $\beta\to 1$, the formulas simplify to

$$\boldsymbol{\zeta}_f=\mathbf{n}\cdot\boldsymbol{\zeta}_i\mathbf{n}+\boldsymbol{\zeta}_i\cdot(\mathbf{n}_i\times\mathbf{n})\,\mathbf{n}_f\times\mathbf{n}+\boldsymbol{\zeta}_i\cdot\mathbf{n}_i\mathbf{n}_f.$$

Thus the final polarization is found by rotating the initial polarization through scattering angle θ about the normal to the scatttering plane $\mathbf{n}$. The helicity is hard.

At low energy, $\beta\to 0$, the formulas simplify to

$$\boldsymbol{\zeta}_f=\boldsymbol{\zeta}_i,$$

and here the polarization is hard.

C. The atomic case

In the case of finite λ/p Eqs. (23) give, to order αZ,

$$S=\frac{\alpha Z\beta(1-\beta^2)^{1/2}\tan\frac{1}{2}\theta}{1+(1-\beta^2)\tan^2\frac{1}{2}\theta}\,\frac{\lambda^2/p^2+4\sin^2\frac{1}{2}\theta}{2\cos^2\frac{1}{2}\theta}\left[\frac{\lambda^2/p^2+2\sin^2\frac{1}{2}\theta}{2\sin\frac{1}{2}\theta[\lambda^4/p^4+4(\lambda^2/p^2+\sin^2\frac{1}{2}\theta)]^{1/2}}\times\ln\left(\frac{[\lambda^4/p^4+4(\lambda^2/p^2+\sin^2\frac{1}{2}\theta)]^{1/2}+2\sin\frac{1}{2}\theta}{[\lambda^4/p^4+4(\lambda^2/p^2+\sin^2\frac{1}{2}\theta)]^{1/2}-2\sin\frac{1}{2}\theta}\right)+\ln\left(\frac{\lambda/p}{(4+\lambda^2/p^2)^{1/2}}\right)\right], \tag{33a}$$

$$T=-\frac{2(1-\beta^2)^{1/2}\tan\frac{1}{2}\theta}{1+(1-\beta^2)\tan^2\frac{1}{2}\theta}+\frac{\alpha Z\beta(1-\beta^2)^{1/2}\tan\frac{1}{2}\theta}{1+(1-\beta^2)\tan^2\frac{1}{2}\theta}\times\frac{\lambda^2/p^2+4\sin^2\frac{1}{2}\theta}{2\cos^2\frac{1}{2}\theta}\,\frac{1-(1-\beta^2)\tan^2\frac{1}{2}\theta}{1+(1-\beta^2)\tan^2\frac{1}{2}\theta}\times\left[\frac{\lambda^2/p^2+2\sin^2\frac{1}{2}\theta}{\sin\frac{1}{2}\theta[\lambda^4/p^4+4(\lambda^2/p^2+\sin^2\frac{1}{2}\theta)]^{1/2}}\times\arctan\left(\frac{(\lambda/p)\sin\frac{1}{2}\theta}{[\lambda^4/p^4+4(\lambda^2/p^2+\sin^2\frac{1}{2}\theta)]^{1/2}}\right)+\frac{1}{\sin\frac{1}{2}\theta}\arctan\left(\frac{\sin\frac{1}{2}\theta}{\lambda/p}\right)-\arctan\left(\frac{2}{\lambda/p}\right)\right], \tag{33b}$$

$$U=\frac{1-(1-\beta^2)\tan^2\frac{1}{2}\theta}{1+(1-\beta^2)\tan^2\frac{1}{2}\theta}+\frac{\alpha Z\beta(1-\beta^2)^{1/2}\tan\frac{1}{2}\theta}{1+(1-\beta^2)\tan^2\frac{1}{2}\theta}\times\frac{\lambda^2/p^2+4\sin^2\frac{1}{2}\theta}{2\cos^2\frac{1}{2}\theta}\,\frac{2(1-\beta^2)^{1/2}\tan\frac{1}{2}\theta}{1+(1-\beta^2)\tan\frac{1}{2}\theta}\times\left[\frac{\lambda^2/p^2+2\sin^2\frac{1}{2}\theta}{\sin\frac{1}{2}\theta[\lambda^4/p^4+(4\lambda^2/p^2+\sin^2\frac{1}{2}\theta)]^{1/2}}\times\arctan\left(\frac{(\lambda/p)\sin\frac{1}{2}\theta}{(\lambda^4/p^4+4[\lambda^2/p^2+\sin^2\frac{1}{2}\theta)]^{1/2}}\right)+\frac{1}{\sin\frac{1}{2}\theta}\arctan\left(\frac{\sin\frac{1}{2}\theta}{\lambda/p}\right)-\arctan\left(\frac{2}{\lambda/p}\right)\right]. \tag{33c}$$

These results are expected to apply to the elastic scattering of electrons by low-Z atoms at moderate energies. Only small αZ can be considered since this is a perturbation calculation. Thus, for a 10% calculation, Z should not be much greater than 10. For such low-Z atoms the screened-Coulomb potential is a good approximation to the Hartree potential,[8] with the screening parameter λ given by the Thomas–Fermi value,

$$\lambda=1.13\,Z^{1/3}/a_0,$$

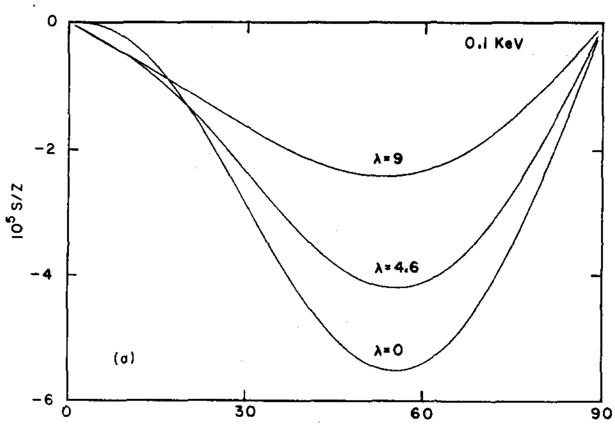
where a_0 is the Bohr radius. The parameter ranges from zero for a Coulomb field to 4.6 Å^{-1} for neon at $Z=10$. A static potential is expected to provide a valid model of the atom at kinetic energies greater than 100 eV; below that energy, electron exchange and shell polarization effects probably become significant.[9] As long as the kinetic energy is greater than 100 eV the wavelength p^{-1} is less than 0.20 Å so λ/p ranges between zero and unity in the applicable region.

With the applicable range of the variables established, the discussion of the exponential factored out in Eq. (11) can be completed. With the factorization as given one obtains Eq. (17) for η and it is sensible to expand the denominator in Eq. (17), to obtain Eq. (18), provided the quantity

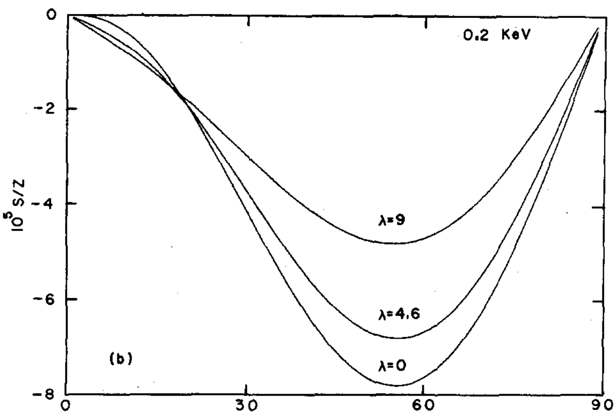
$$\kappa=(\alpha Zp^2/2\pi^2E)(\lambda^2+4p^2\sin^2\tfrac{1}{2}\theta)(I-J)$$

is small. A numerical discussion establishes that the absolute value of this quantity is less than αZ at all energies and angles provided λ is less than p. On the other hand, if an inappropriate exponential is factored out, one arrives at an expression for η where such an expansion is not justified. For example, if no exponential at all is factored out, the same procedure leads to

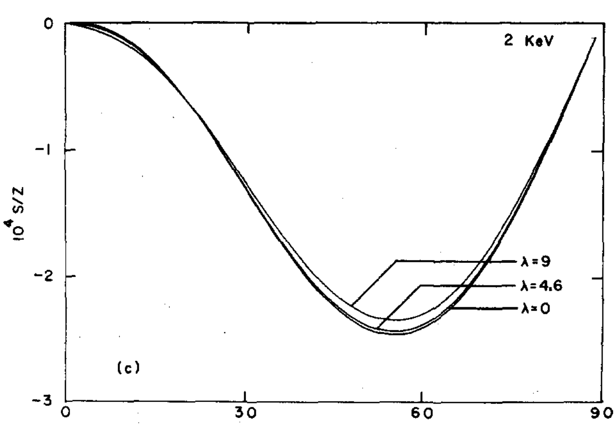
$$\eta=\frac{1-(\alpha Z/\pi^2)E(\lambda^2+4p^2\sin^2\frac{1}{2}\theta)I}{1+(\alpha Z/2\pi^2)(p^2/E)(\lambda^2+4p^2\sin^2\frac{1}{2}\theta)(I-J)-(\alpha Z/\pi^2)E(\lambda^2+4p^2\sin^2\frac{1}{2}\theta)I}\,\frac{m}{E}\tan\tfrac{1}{2}\theta.$$



(a)



(b)



(c)

FIG. 1. Effect of screening on S. Here $S(\theta)/Z$ is plotted at kinetic energies of 0.1, 0.2, and 2.0 KeV for screening parameters $\lambda=0$, 4.6, and 9.

Here the expansion can be made provided the quantity

$$\kappa' = (\alpha Z/\pi^2)E(\lambda^2 + 4p^2 \sin^2\tfrac{1}{2}\theta)I$$

also is small. However, κ' diverges for small λ/p as $\log(\lambda/p)$ and diverges at low energy as $1/\beta$.

Graphs of S, T, and U, showing the dependence on screening parameter, kinetic energy, and scattering angle, are given in Figs. 1, 2, and 3. It is interesting

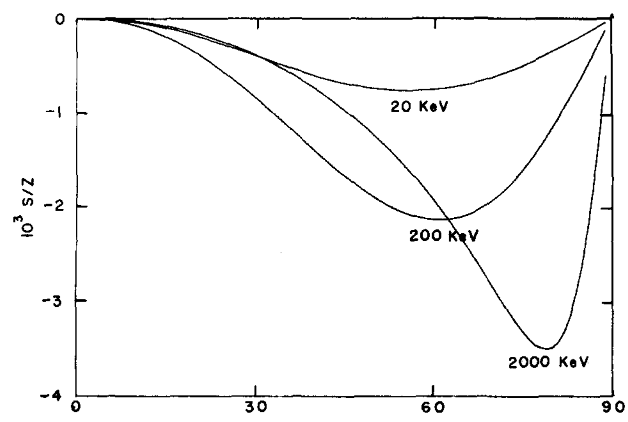
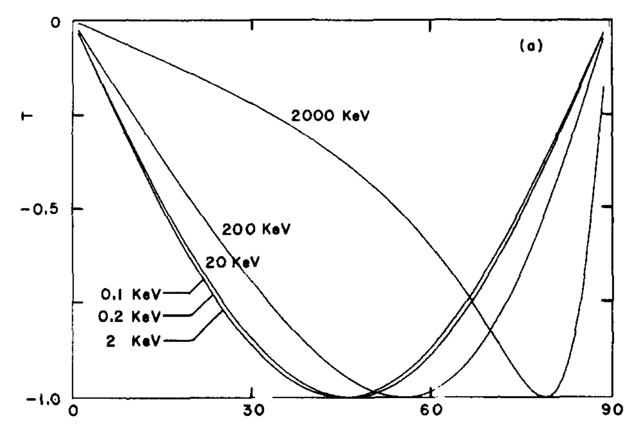
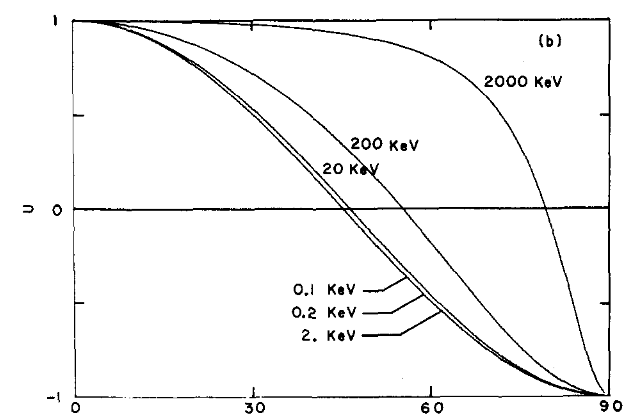


FIG. 2. Polarization at higher energies. Here $S(\theta)/Z$ is plotted at kinetic energies of 20, 200, and 2000 KeV. The dependence on λ is too small to be seen at these energies.

that the only significant dependence on λ is in S at kinetic energies below about 2 KeV. Thus, except in that respect, the screened Coulomb potential gives the same results as the pure Coulomb potential. The Z dependence in T and U is too small to be seen on these graphs; the $(\alpha Z)^1$ terms are too small compared to the $(\alpha Z)^0$ terms in these functions. However, S is proportional to αZ, in the order calculated.



(a)



(b)

FIG. 3. Energy dependence of $T(\theta)$ and $U(\theta)$. The functions are plotted at kinetic energies of 0.1, 0.2, 2.0, 20, 200, and 2000 KeV. The dependence on λ and Z is too small to be seen.

APPENDIX

The integrals that occur in Eq. (7) were evaluated by Dalitz.[1] We give here an alternate and more complete description of how to do them.

The integrals are

$$I=\int\frac{d^3q}{(\lambda^2+|\mathbf{p}_f-\mathbf{q}|^2)(\lambda^2+|\mathbf{q}-\mathbf{p}_i|^2)(p^2-q^2+i\epsilon)}, \tag{A1}$$

$$\mathbf{K}=\int\frac{\mathbf{q}d^3q}{(\lambda^2+|\mathbf{p}_f-\mathbf{q}|^2)(\lambda^2+|\mathbf{q}-\mathbf{p}_i|^2)(p^2-q^2+i\epsilon)}, \tag{A2}$$

where p denotes $(E_i^2-m^2)^{1/2}$ and q is $|\mathbf{q}|$. The lengths $|\mathbf{p}_i|$ and $|\mathbf{p}_f|$ are both p.

The first step is to use the identity,

$$\frac{1}{ab}=\int_{-1}^{1}\frac{dZ}{2}\left(\frac{2}{a(1+Z)+b(1-Z)}\right)^2 \tag{A3}$$

to bring the integrals to the forms

$$I=\int_{-1}^{1}\frac{dZ}{2}\int\frac{d^3q}{(|\mathbf{q}-\mathbf{P}|^2+\Lambda^2)^2(p^2-q^2+i\epsilon)}, \tag{A4}$$

$$\mathbf{K}=\int_{-1}^{1}\frac{dZ}{2}\int\frac{\mathbf{q}d^3q}{(|\mathbf{q}-\mathbf{P}|^2+\Lambda^2)^2(p^2-q^2+i\epsilon)}, \tag{A5}$$

where

$$\mathbf{P}=\tfrac{1}{2}[(1+Z)\mathbf{p}_i+(1-Z)\mathbf{p}_f], \tag{A6}$$

$$\Lambda=[\lambda^2+p^2(1-Z^2)\sin^2\tfrac{1}{2}\theta]^{1/2}, \tag{A7}$$

and θ is the angle between $\mathbf{p}_i$ and $\mathbf{p}_f$, the scattering angle. The length of $\mathbf{P}$ is

$$P=[p^2-p^2(1-Z^2)\sin^2\tfrac{1}{2}\theta]^{1/2}, \tag{A8}$$

and one notices that

$$P^2+\Lambda^2=p^2+\lambda^2. \tag{A9}$$

The second step is to recognize that

$$\int\frac{d^3q}{(|\mathbf{q}-\mathbf{P}|^2+\Lambda^2)^2(p^2-q^2+i\epsilon)}=-\frac{dL}{d\lambda^2}, \tag{A10}$$

$$\int\frac{q_id^3q}{(|\mathbf{q}-\mathbf{P}|^2+\Lambda^2)^2(p^2-q^2+i\epsilon)}=\frac{1}{2}\frac{\partial L}{\partial P_i}, \tag{A11}$$

where L is defined by

$$L=\int\frac{d^3q}{(|\mathbf{q}-\mathbf{P}|^2+\Lambda^2)(p^2-q^2+i\epsilon)}$$

$$=\int\frac{d^3q}{(q^2-2\mathbf{q}\cdot\mathbf{P}+p^2+\lambda^2)(p^2-q^2+i\epsilon)}. \tag{A12}$$

The problem is reduced then to the evaluation of L.

In terms of spherical polar coordinates, z axis chosen in the $\mathbf{P}$ direction, L becomes

$$L=\pi\int_{-1}^{1}d(\cos\theta)\times\int_{-\infty}^{\infty}\frac{q^2dq}{(q^2-2qP\cos\theta+p^2+\lambda^2)(p^2-q^2+i\epsilon)}. \tag{A13}$$

The q integration is done with complex variable technique. For large q the integrand goes like q^{-2} so that the contour may be closed upward at infinity. The denominator has zeros at $\pm(p+i\epsilon)$ and $P\cos\theta\pm i(p^2+\lambda^2-P^2\cos^2\theta)^{1/2}$ so the contour includes poles at $p+i\epsilon$ and $P\cos\theta+i(p^2+\lambda^2-P^2\cos^2\theta)^{1/2}$. The evaluation of the residues leads to

$$L=-2\pi^2i\int_{-1}^{1}dt\times\left\{\frac{2P^2t^2-p^2-\lambda^2+2iPt(p^2+\lambda^2-P^2t^2)^{1/2}}{2i(p^2+\lambda^2-P^2t^2)^{1/2}[2P^2t^2-2p^2-\lambda^2+2iPt(p^2+\lambda^2-p^2t^2)^{1/2}]}+\frac{p}{2(2p^2+\lambda^2-2pPt)}\right\}. \tag{A14}$$

From here on the integration is elementary. The result is

$$L(P)=\frac{\pi^2i}{P}\ln\frac{p-P+i\Lambda}{p+P+i\Lambda}. \tag{A15}$$

Here we are thinking of L as a function of P, $\Lambda(P)$ given by Eq. (A9). The logarithm function is defined to have $-\pi<\arg\log z<\pi$.

One differentiates Eq. (A15) by λ^2, as required by Eq. (A10), to obtain

$$\int\frac{d^3q}{(|\mathbf{q}-\mathbf{P}|^2+\Lambda^2)^2(p^2-q^2+i\epsilon)}=\frac{\pi^2}{\Lambda(-\lambda^2+2pi\Lambda)}, \tag{A16}$$

and then integrates over Z as required by Eq. (A4). The final expression for I is

$$I=\frac{-\pi^2}{p\sin\frac{1}{2}\theta[\lambda^4+4p^2(\lambda^2+p^2\sin^2\frac{1}{2}\theta)]^{1/2}}\times\left(\arctan\frac{\lambda p\sin\frac{1}{2}\theta}{[\lambda^4+4p^2(\lambda^2+p^2\sin^2\frac{1}{2}\theta)]^{1/2}}+\frac{1}{2}i\ln\frac{[\lambda^4+4p^2(\lambda^2+p^2\sin^2\frac{1}{2}\theta)]^{1/2}+2p^2\sin\frac{1}{2}\theta}{[\lambda^4+4p^2(\lambda^2+p^2\sin^2\frac{1}{2}\theta)]^{1/2}-2p^2\sin\frac{1}{2}\theta}\right), \tag{A17}$$

the arctangent defined to be in the first quadrant.

To get $\mathbf{K}$, one first writes

$$K_i=\int_{-1}^{1}\frac{dZ}{4}\frac{\partial L}{\partial P_i}=\int_{-1}^{1}\frac{dZ}{4}\frac{P_i}{P}\frac{dL}{dP}. \tag{A18}$$

However, P and dL/dP are even in Z so only the even part of P_i contributes and

$$\mathbf{K}=\tfrac{1}{2}(\mathbf{p}_i+\mathbf{p}_f)J, \tag{A19}$$

where

$$J=\int_{-1}^{1}\frac{dZ}{4}\frac{1}{P}\frac{dL}{dP}=p^{-2}\sin^{-2}\tfrac{1}{2}\theta\int_{-1}^{1}\frac{dZ}{4Z}\frac{dL}{dZ}. \tag{A20}$$

Two integrations by parts bring this to a form that can be treated by elementary methods. First of all

$$J=\frac{1}{4p^2\sin^2\frac{1}{2}\theta}\frac{L}{Z}\bigg|_{-1}^{1}+\frac{1}{p^2\sin^2\frac{1}{2}\theta}\int_{-1}^{1}\frac{dZ}{4Z^2}L$$

$$=\frac{\pi^2i}{2p^3\sin^2\frac{1}{2}\theta}\ln\frac{i\lambda}{2p+i\lambda}$$

$$+\frac{\pi^2 i}{2p^2\sin^2\frac{1}{2}\theta}\int_{-1}^{1}\frac{dZ}{2Z^2P}\ln\frac{p-P+i\Lambda}{p+P+i\Lambda}. \tag{A21}$$

Second, one recognizes that dZ/PZ^2 is $-p^{-2}\cos^{-2}\frac{1}{2}\theta d(P/Z)$ so that

$$J=\frac{\pi^2 i}{2p^3\sin^2\frac{1}{2}\theta}\ln\frac{i\lambda}{2p+i\lambda}$$

$$-\frac{\pi^2 i}{2p^2\sin^2\frac{1}{2}\theta}\,\frac{1}{2p^2\cos^2\frac{1}{2}\theta}\,\frac{P}{Z}\ln\frac{p-P+i\Lambda}{p+P+i\Lambda}\Bigg|_{-1}^{1}$$

$$+\frac{\pi^2 i}{2p^2\sin^2\frac{1}{2}\theta}\,\frac{1}{2p^2\cos^2\frac{1}{2}\theta}$$

$$\int_{-1}^{1}\frac{P}{Z}\,dZ\,\frac{d}{dZ}\ln\frac{p-P+i\Lambda}{p+P+i\Lambda}. \tag{A22}$$

This simplifies down to

$$J=-\frac{\pi^2 i}{2p^2\cos^2\frac{1}{2}\theta}\left(\frac{1}{p}\ln\frac{i\lambda}{2p+i\lambda}+\int_{-1}^{1}\frac{dZ}{\Lambda}\,\frac{p\Lambda+i(\lambda^2+p^2)}{-\lambda^2+2pi\Lambda}\right). \tag{A23}$$

The final result is

$$J=\frac{\lambda^2+2p^2}{2p^2\cos^2\frac{1}{2}\theta}I-\frac{i\pi^2}{2p^3\cos^2\frac{1}{2}\theta}$$

$$\left(\ln\frac{i\lambda}{2p+i\lambda}-\frac{i}{\sin\frac{1}{2}\theta}\arctan\frac{p\sin\frac{1}{2}\theta}{\lambda}\right). \tag{A24}$$

[1]R.H. Dalitz, Proc. Roy. Soc. A **206**, 509 (1951).
[2]F. Gürsey, Phys. Rev. **107**, 1734 (1957).
[3]J.D. Bjorken and S.D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, New York, 1964).
[4]S. Gasiorowicz, *Elementary Particle Physics* (Wiley, New York, 1966), p. 173.
[5]H. Überall, *Electron Scattering from Complex Nuclei* (Academic, New York, 1971), Part A, pp. 55–56.
[6]A.D. Martin and T.D. Spearman, *Elementary Particle Theory* (Wiley, New York, 1970).
[7]See, for example, Überall, Ref. 5, Eq. (2-133a).
[8]W.J. Byatt, Phys. Rev. **104**, 1298 (1956).
[9]J. Kessler, Rev. Mod. Phys. **41**, 3 (1969).

Representation and differential geometry of the semisimple Lie groups

Hsin-Yang Yeh

Department of Physics and Astronomy, Moorhead State College, Moorhead, Minnesota 56560
(Received 3 January 1974)

A systematic method is presented whereby any compact Lie group of n-real parameters is dealt with from an infinitesimal approach with the representative matrix method based on a group of inner automorphisms suggested in a previous paper. The group manifold, defined in terms of a metric of group parameters, is identified as a Riemannian one in which these parameters play a role of n curvilinear coordinates. Riemannian geometry is thus valid in the group manifold, and geometric quantities are explicitly calculated in terms of the symmetric or (0) connection by a straightforward application of the ordinary procedure of tensor analysis. A new and simpler method of computing the invariant volume element is presented within this framework. Furthermore, we discuss in detail the group of inner automorphisms for the calculation of the matrix element of finite rotations for any irreducible representation (abbreviated MEFRIR). It is found that our method works very well and yields right and left vector fields together with a set of $2n$ equations to be satisfied by the MEFRIR. The global properties of the group may, therefore, be obtained as a solution to these equations. Moreover, it provides not only the generalized Maurer–Cartan equations, the Lie structure formulas, two parameter groups of point transformations and the adjoint group, but also two additional nonsymmetric (+) and (−) connections with zero curvature, which do not possess any preassigned metric but possess two absolute parallelisms. Thus, our results on differential geometry completely agree with Cartan and Schouten's. A link between differential geometry and representations is presented by the right and left vector fields which are explicitly calculable in terms of the n parameters and through which geometric quantities, e.g., the Riemann tensor, the Ricci tensor, and the scalar curvature, of any connection are explicitly displayed. A theorem relating both the vector fields to the metric tensors is also included. Finally, the l (rank of the group) invariant differential equations to be satisfied by the MEFRIR are cast in the covariant (or Lie derivative) forms in any connection. Examples of the invariant equations are given for $SU(2)$, $SO(3)$, and $SU(3)$. The two invariant equations of the latter can be cast in terms of the eigenvalues of isospin and hypercharge upon carrying out charge and hypercharge quantizations; in this connection, a new nonrelativistic wave equation to be satisfied by $SU(3)$ multiplets as a generalization of the Schrödinger equation of the symmetric top is also proposed.

I. INTRODUCTION

In spite of some prominent works by Lie,[1] Cartan and Schouten[2,3] (abbreviated CS), and Eisenhart,[4,5] on differential geometry of the semisimple Lie groups, little progress has been made along this line of research. Attempts have been made to introduce the Hilbert space of complex functions[6–8] to the unitary group, symmetric spaces,[9] special functions,[10] and many others.[11]

An infinitesimal and differential geometric approach has a definite advantage in that the motion of geometric quantities may be explicitly describable in terms of the variables at any point in group space, as compared with the algebraic method, the Hilbert space method, and others, in which there exist no geometry. Recently, there were attempts to introduce a metric to the unitary group $U(m)$[12] without success and to geometrize $U(m)$.[13]

Yet, no systematic method exists by which all semisimple Lie groups may be dealt with from an infinitesimal and differential geometric point of view. An apparent difficulty seems to be that although the transformation of a semisimple Lie group with n-real parameters has been dealt with and the existence of three [(0), (+), and (−)] connections has been shown, in practice it is not known how these connections can be calculated, how they are related to the group manifold and the MEFRIR. On the other hand, recent developments in the representation theory of the semisimple Lie groups indicate that a separate treatment for each of A_l, B_l, C_l, D_l and the five exceptional groups G_2, F_4, E_6, E_7 and E_8 would be necessary—this is accomplished by our method. In this connection, representative matrices of these groups have been parametrized quite arbitrarily.[14,15]

In this paper we present a method whereby the representation theory and differential geometry of any semisimple Lie group are treated in a systematic way from an infinitesimal point of view. This is accomplished by generalizing the representative matrix method[16] based on a group of inner automorphisms which was used in treating the representation theory of $SU(3)$ where it was suggested that such a generalization would indeed be possible. Our method is completely different from Lie's original approach, in which a finite continuous transformation group is defined as a system of transformations

$$x_i' = f_i(x_1 \ldots x_s,\ a_1 \ldots a_n) \quad \text{for } i = 1 \ldots S,$$

in that we need neither the variables x_i (or x_i') nor the functional equations derived thereby. However, we require a representative matrix for a group, which is a continuous and differentiable function of group coordinates (parameters), through which a Riemann space can be introduced. Different representative matrices of a group are regarded as different sets of coordinates in the same Riemannian manifold which provides the same quadratic line element so that the method has universal applicability to any representative matrix of any semisimple Lie group.

While it is important to recognize that CS's treatment applies to the semisimple Lie group in general, it is more important to recognize that our treatment applies particularly to a specific group, e.g., the unitary group $U(m)$ if we start with a representative matrix of $U(m)$. Our method also has an advantage over CS's in that the metric tensor of a group manifold is explicitly calculable as functions of coordinate variables and thereby making possible the explicit calculation of the right and left vector fields as well as three connections and geometric quantities, such as the Riemann tensors, the Ricci tensors, and the scalar curvatures for any connection.

We also present a new method simpler than the ordinary ones by Wigner, Boener, and Murnaghan of computing the invariant volume element of a group. Especially important is the fact that the differential equations to be satisfied by the MEFRIR may be integrated; once this is done, our method establishes a link between the differential approach and the global approach of Schur and Weyl and thus, when solved, provides us with a complete set of orthonormal functions in a Riemannian group manifold. Consequently, any function of this manifold may be expansible in terms of these functions. Our method therefore generalizes the method of Fourier series to one with a larger class of series expansible in terms of a complete set of orthonormal functions defined in a general Riemannian space.

It is apparent that the important global properties of group manifolds cannot be introduced independent of the infinitesimal approach and can only be understood as solutions to differential equations to be satisfied by the MEFRIR; this paper indeed details such a method.

This paper is divided into five sections and the Appendix as follows:

In Sec. II, we introduce, through a representative matrix, a quadratic line element of a semisimple Lie group of n real variables α^μ ($\mu=1\ldots n$). The quadratic line element is invariant under a coordinate transformation $\alpha^\mu \to \alpha'^\mu$, therefore provides a symmetric Riemannian metric, and defines an invariant group space of n-real dimensions. Cyclic coordinates are also defined.

We can then utilize the ordinary method of tensor calculus of Riemannian geometry to obtain the symmetric or (0) connection, the Riemann tensors, the Ricci tensors, and other geometric quantities of the symmetric connection associated with this metric. The invariant volume element of a group is also introduced within the framework of Riemannian geometry.

In Sec. III, a group of inner automorphisms induced by parameter change is considered which is a generalization of the method used in treating $SU(3)$.[16] By applying with the generators from the right and left sides on the representative matrix, the right and left translations (or rotations) are generated, which, in turn, induce two contravariant vector fields $\gamma_I^\mu(\alpha)$ and $l_I^\mu(\alpha)$. We thus obtain a set of $2n$ differential equations to be satisfied by the MEFRIR. Two sets of covariant vector fields $\gamma_\mu^i(\alpha)$ and $l_\mu^I(\alpha)$ are then introduced through orthonormality conditions. Furthermore, another set of $2n$ equations to be satisfied by the MEFRIR are derived which is shown to include the generalized Maurer—Cartan equations and the Lie structure formulas of the first and second parameter groups. The adjoint group is also discussed in this connection.

Section IV, treated in parallel with CS and Schouten, deals with two nonsymmetric (±) connections and their geometric properties generated by the real vector fields $\gamma_i^\mu(\alpha)$, $l_I^\mu(\alpha)$, $\gamma_\mu^i(\alpha)$, and $l_\mu^I(\alpha)$. It is shown that these connections with zero curvature do not possess any preassigned metric and therefore are non-Riemannian but possess two absolute parallelisms of Levi-Civita, which is in complete agreement with CS's results. The relationships between these connections and the curvature tensor of the (0) connection are also given.

Section V deals with invariant differential equations to be satisfied by the MEFRIR. A theorem relating the vector fields to the metric of a group is proved. By making use of this theorem, the second order invariant equation of any group is cast in the covariant forms in any connection. For the higher order invariant equations we take $U(m)$ as an example and cast them in similar covariant (or Lie derivative) forms. Some special examples of the invariant differential equations are explicitly given, including those of $SU(2)$ and $SO(3)$ (the Schrödinger equation of the symmetric top) as well as those of $SU(3)$. The latter example also includes new quantization rules for charge and hypercharge so that the invariant equations can be expressed in terms of their eigenvalues. A new nonrelativistic wave equation to be satisfied by $SU(3)$ multiplets is also displayed.

Section VI contains our conclusions which summarize some new results and compare our results with CS's.

In the Appendix we briefly discuss some properties of the MEFRIR, including its orthonormality, completeness and transformation properties in R_n.

The general theory developed in this paper will be applied to the groups $SU(2)$ and $SO(3)$ in another article.[17]

II. RIEMANNIAN SPACE, GEOMETRIC PROPERTIES, AND SYMMETRIC (0) CONNECTION

In this section we identify the group manifold of a compact Lie group as a Riemannian one with a metric provided by a representative matrix. It is shown that the quadratic line element defined is invariant under any transformation of its representative matrix. Thus we can apply the ordinary method of tensor calculus on Riemannian geometry to obtain its symmetric (0) connection and all other geometric quantities. The invariant volume element is also introduced and its invariance under a coordinate transformation as well as right and left translations is shown.

A. The Riemannian space

Let G_l be a semisimple Lie group of rank l with a representative matrix $M(\alpha)=(m_{tt'}(\alpha))$ of rank $N-m_{tt'}$ are complex numbers and the α collectively denotes the α^μ's—expressed in terms of n (order of the group) linearly independent real variables α^μ ($\mu=1, 2, \ldots, n$) such that the variability domains of the α^μ's are appropriately chosen to cover the entire group manifold once and only once. We assume the continuity and differentiability of the representative matrix with respect to the group parameters α^μ. The $M(\alpha)$ satisfies the relavant group conditions. For example, $M^\dagger M=1$ and $\det M=1$ for SU_{l+1}, the dagger (†) indicating the Hermitian conjugate; the rank N of matrix $M(\alpha)$ is $l+1$ and the order n is l^2+2l.

With the help of $M(\alpha)$ we define a symmetric metric $g_{\mu\nu}(\alpha)$ and a positive-definite quadratic line element between two adjacent points in group space in terms of the real coordinate variables α^μ by

$$ds^2 \equiv dm^*_{tt'}, \quad dm_{tt'} = g_{\mu\nu}(\alpha)\, d\alpha^\mu\, d\alpha^\nu, \tag{2.1}$$

the asterisk indicating complex conjugation. Throughout this paper, indices t and t' run from 1 to N, whereas italic indices [both lower case $(i,j,)$ and capitals $(I,J,)$] as well as Greek indices (μ,ν) run from 1 to n. The summation convention for a repeated index is used unless otherwise stated. By definition, $g_{\mu\nu}(\alpha) \equiv (\partial m^*_{tt'}/\partial\alpha^\mu)(\partial m_{tt'}/\partial\alpha^\nu)$; the metric tensor $g_{\mu\nu}(\alpha)$ is clearly symmetric and therefore Riemannian.

Let $\alpha^\mu \to \alpha'^\mu$ induced by $M(\alpha)\to M(\alpha')$ of the same G_l be a transformation to another set of coordinates of the same point; let a neighboring point have coordinates $\alpha^\mu + d\alpha^\mu$. The $d\alpha^\mu$ then transform according to a contravariant vector

$$d\alpha'^\mu = \frac{\partial\alpha'^\mu}{\partial\alpha^\nu}\, d\alpha^\nu, \tag{2.2}$$

and $g_{\mu\nu}(\alpha)$[18] according to

$$g_{\mu\nu}'(\alpha') = \frac{\partial\alpha^\rho}{\partial\alpha'^\mu}\frac{\partial\alpha^\gamma}{\partial\alpha'^\nu}\, g_{\rho\gamma}(\alpha). \tag{2.3}$$

Accordingly, for the same G_l with different $M(\alpha)$ and $M(\alpha')$ we get the same line element

$$ds^2 = g_{\mu\nu}(\alpha)\, d\alpha^\mu\, d\alpha^\nu = g_{\rho\delta}'(\alpha')\, d\alpha'^\rho\, d\alpha'^\delta, \tag{2.4}$$

i.e., the quadratic line element is invariant under a coordinate transformation $\alpha^\mu \to \alpha'^\mu$. Since there is a symmetric (0) connection associated with the metric $g_{\mu\nu}(\alpha)$, we have therefore defined the group manifold, called R_n, of a Riemann space in the n real variables α^μ.

It is convenient to define a cyclic coordinate α^μ of R_n; any coordinate α^μ which does not appear explicitly in $g_{\mu\nu}(\alpha)$ is said to be cyclic (or ignorable). This concept of cyclic coordinates will play an important role in quantization procedures—in both the ordinary quantum theory and a new quantum theory of quantized charge and hypercharge.

Mathematically, no set of coordinate system is preferred to another; any set of n coordinates which satisfies the relevant group conditions may be acceptable mathematically in defining the group manifold of a n-dimensional Riemannian space. However, physically there is very often a preferred set of coordinates. Physical requirements, such as the requirement of quantizing the z component of angular momentum, often impose restrictions on a choice of coordinate system. Two examples are given in another article[17] on preferred sets of coordinates for $SU(2)$ and $SO(3)$ which are compatible with the quantization of angular momentum in quantum mechanics. Another example has been displayed[16] of a preferred set of coordinates for $SU(3)$ where isospin and hypercharge angular variables have been introduced to make the chosen coordinate system be compatible with the quantization of electric charge and hypercharge.

B. Geometric properties of the (0) connection

Once the metric of G_l is known, one can follow the general procedure[19,20] of tensor calculus to obtain general mixed tensors and geometric quantities in R_n. A mixed tensor of rank $r+s$ with r contravariant indices and s covariant indices may be defined in the ordinary way. The contravariant metric tensor $g^{\mu\nu}(\alpha)$ are defined by

$$g_{\rho\mu}\, g^{\mu\nu} = \delta_\rho{}^\nu \tag{2.5}$$

The covariant (or contravariant) derivatives of $g_{\mu\nu}$ and $g^{\mu\nu}$, denoted by ∇_ρ (or ∇^ρ), then vanish

$$\nabla_\rho g_{\mu\nu} = 0, \quad \nabla_\rho g^{\mu\nu} = 0. \tag{2.6}$$

These equations lead to the expressions of the Christoffel's symbols of the first and second kinds for the symmetric or (0) connection, $\Gamma_{\alpha\beta,\gamma}$ and $\Gamma^\gamma_{\alpha\beta}$:

$$\begin{aligned}\Gamma_{\alpha\beta,\gamma} &\equiv \tfrac{1}{2}(\partial_\beta g_{\alpha\gamma} + \partial_\alpha g\gamma_\beta - \partial_\gamma g_{\alpha\beta}),\\ \Gamma^\gamma_{\alpha\beta} &= g^{\gamma\delta}\,\Gamma_{\alpha\beta,\delta}\end{aligned} \tag{2.7}$$

where ∂_γ denotes $\partial/\partial\alpha^\gamma$. Clearly, $\Gamma_{\alpha\beta,\gamma}$ and $\Gamma^\gamma_{\alpha\beta}$ are symmetric in α and β.

Through the (0) connection a tensor is parallel displaced in group space. For clarity, we write out geometric quantities of the (0) connection as follows.

The Riemann tensor of the (0) connection is defined in the ordinary way as

$$R^\rho{}_{\mu\nu\delta} = \partial_\nu \Gamma^\rho{}_{\mu\delta} - \partial_\delta \Gamma_\mu{}^\rho{}_\nu + \Gamma_\mu{}^\alpha{}_\delta\, \Gamma_\alpha{}^\rho{}_\nu - \Gamma^\gamma_{\alpha\beta}\Gamma_\alpha{}^\rho{}_\delta \tag{2.8}$$

which has the symmetric properties

$$R_{\mu\nu\rho\delta} = R_{\rho\delta\mu\nu} = -R_{\nu\mu\rho\delta} = -R_{\mu\nu\delta\rho} \tag{2.9}$$

and satisfies the two Bianchi identities

$$\begin{aligned}&R_{\mu\nu\rho\delta} + R_{\mu\rho\delta\nu} + R_{\mu\delta\nu\rho} = 0,\\ &\nabla_\epsilon R_{\mu\nu\rho\delta} + \nabla_\rho R_{\mu\nu\delta\epsilon} + \nabla_\delta R_{\mu\nu\epsilon\rho} = 0.\end{aligned} \tag{2.10}$$

The Ricci tensor and the scalar curvature of the (0) connection are obtained by contraction

$$R_{\mu\nu} = R^\alpha{}_{\mu\alpha\nu}, \quad R = R^\mu{}_\mu, \tag{2.11}$$

which do not vanish in general. Geodesics in R_n can be introduced in a similar way.

C. The invariant integral

Wigner[21] has generalized a method of calculating the invariant volume elements of finite groups to those of continuous groups, and Murnaghan has utilized the method of characteristic matrices (generators). An alternative to these methods is presented which makes use of Riemannian geometry.

When the determinant of the metric of G_l defined in (2.1) does not vanish, i.e., $g = |g_{\mu\nu}| \neq 0$,[22] which is actually the case, the invariant volume element of G_l is defined by the one that is invariant under any coordinate transformation $\alpha^\mu \to \alpha'^\mu$, i.e., $\sqrt{g}\prod_{\mu=1}^n d\alpha^\mu$. The volume element is also invariant under left (or right) translation inasmuch as the translation is also regarded as a special case of coordinate transformation and the group space is invariant of the transformation. The invariant volume element defined in this way is indeed equivalent to the one defined by Murnaghan. Note that in our method the translation invariance follows automatically (not required) from the invariance under a coordinate transformation; in contrast, this is a required assumption[23] if one tries to generalize finite groups to continuous groups.

We use the notation $\langle F(\alpha)\rangle$ to denote the average value of a function $F(\alpha)$ taken over the variability domains of the α^μ's which is denoted by R_n,

$$\langle F(\alpha)\rangle = \int_{R_n} F(\alpha)\sqrt{g}\prod_{\mu=1}^{n} d\alpha^\mu / V \tag{2.12}$$

where $V=\int_{R_n}\sqrt{g}\prod_{\mu=1}^{n} d\alpha^\mu$ is the entire volume of R_n.

Our definition is more concrete and definite than the customary ones in that it is indeed possible to carry out the integration over the group space of G_l.

III. THE AUTOMORPHISM GROUP

Let X_i be a set of linearly independent generators which may be obtained from the representation matrix by

$$X_i = \frac{\partial M(\alpha)}{\partial \alpha^i}\Bigg|_{\text{all } \alpha^i=0}.$$

Once the generators are given we can make use of the following method to obtain the first and second parameter groups as well as the adjoint group. When taken over all transformations of inner automorphisms, we have a group of inner automorphisms which yields a set of $2n$ matrix equations to be satisfied by the MEFRIR. Also included are the right and left vector fields obtained thereby and the generalized Maurer–Cartan equations together with the Lie structure formulas for each of the right and left vectors to be satisfied by the MEFRIR.

A. Right and left translation

Any representative matrix $M(\alpha)$ of G_l is a function of a group element g of G_l and by this we write $M(g)$. When we apply a group element g_l on g, $M(g)$ is transformed into $M(gg_l)$ which is a function of another set of parameters ϵ^μ (or α'^μ).

A group of inner automorphisms is a one-to-one mapping of G_l onto itself which preserves multiplication and which also induces parameter change. One may regard this as a group of infinitesimal inner automorphisms as a result of first performing a right translation $M\to MX$ and then a left translation $M\to XM$. Each of these translations may be regarded as the special case of a coordinate transformation $\alpha^\mu\to\alpha'^\mu$ or a point transformation. By these translations, the element of the representation matrix $M(\alpha)$ obeys two systems of first-order partial differential equations, which are obtained by finite right and left translations with the generators X_i applied respectively from the right and left side to $M(\alpha)$. If we denote the two sets of coordinate variables before and after the transformation by $\epsilon^{j(\text{or } J)}$ and α^β respectively, the right and left translations so defined are expressed by

$$M(\alpha)X_j = \sigma\left(\frac{\partial\alpha^\beta}{\partial\epsilon^j}\right)_\gamma \frac{\partial M(\alpha)}{\partial\alpha^\beta}, \tag{3.1}$$

$$X_J M(\alpha) = \sigma\left(\frac{\partial\alpha^\beta}{\partial_\epsilon J}\right)_l \frac{\partial M(\alpha)}{\partial\alpha^\beta}, \tag{3.2}$$

where σ is an indicator which is either 1 or i depending on the type of group under consideration so that $\gamma_j{}^\beta$ and $l_J{}^\beta$ in Eqs. (3.5) and (3.6) become real quantities. For the left translation, we use a capital italic letter J to comply with Schouten's notation.[3] The coefficients of the right and left translations $(\partial\alpha^\beta/\partial\epsilon^j)_\gamma$ and $(\partial\alpha^\beta/\partial\epsilon^J)_l$ are obtained by regarding α^β as functions of $\epsilon^{j(\text{or } J)}$ and setting all $\epsilon^{j(\text{or } J)}=0$. These coefficients naturally reduce to $\delta^\beta{}_{j(\text{or } J)}$ when $X_{j(\text{or } J)}$ is an identity operation.

These coefficients should be expressed as explicit functions of the variables α^β. To attain this, each corresponding element on both sides of each matrix equation has to be equated to each other and the resulting system of equations must be solved for the $2n^2$ unknown coefficients. In practice, solving these equations in most cases is extremely involved.

By solving these matrix equations we map $M(\alpha)$ into the irreducible vector space prescribed by the l-row partition label $[p]\equiv(p_1, p_2, \ldots, p_l)$ in R_n. When the matrix element of any infinitesimal generator X_j of G_l is denoted by $\langle m|X_j|m'\rangle$ where[24] $|m'\rangle$ (or $\langle m|$) is a Gelfand (or its conjugate) state prescribed by the partition label $[p]$, we obtain the following $2n$ equations to be satisfied by the MEFRIR $\mathcal{D}^{[p]}_{\langle m|m'\rangle}(\alpha)$ (Ref. 25):

$$\mathcal{D}^{[p]}_{\langle m|m''\rangle}(\alpha)\langle m''|X_j|m'\rangle = \sigma\gamma_j{}^\beta\partial_\beta\mathcal{D}^{[p]}_{\langle m|m'\rangle}(\alpha), \tag{3.3}$$

$$\langle m|X_J|m''\rangle\mathcal{D}^{[p]}_{\langle m''|m'\rangle}(\alpha) = \sigma l_J{}^\beta\partial_\beta\mathcal{D}^{[p]}_{\langle m|m'\rangle}(\alpha). \tag{3.4}$$

Here the contravariant vectors $\gamma_j{}^\beta$ and $l_J{}^\beta$ are defined by

$$\gamma_j{}^\beta \equiv \left(\frac{\partial\alpha^\beta}{\partial\epsilon^J}\right)_\tau, \tag{3.5}$$

$$l_J{}^\beta \equiv \left(\frac{\partial\alpha^\beta}{\partial\epsilon^J}\right)_l, \tag{3.6}$$

where the right sides of (3.5) and (3.6) are regarded as explicit functions of the α^μ's as a result of solving (3.1) and (3.2).

The matrix elements of the generators $\langle m|X_i|m'\rangle$ for all $U(m)$ and $O(m)$ were first obtained by Gelfand and Zetlin.[26,27] A complete global solution to the MEFRIR of any semisimple Lie group may be obtained by solving either the set of $2n$ equations, (3.3) and (3.4), or the l invariant equations, which are also satisfied by the MEFRIR (refer to Sec. V); our method thus unifies the integral approach of Schur and Weyl with the differential one of CS.

The problem of $SU(3)$ along this line of work in a preferred system of coordinates, which is compatible with charge and hypercharge quantizations, will be treated, and the MEFRIR $\mathcal{D}^{[p]}_{\langle m|m'\rangle}(\alpha)$ will be dealt with elsewhere. The corresponding vectors $\gamma_i{}^\beta$ and $l_I{}^\beta$ have already been displayed.[28]

For each of $SU(2)$ and $SO(3)$, the vectors $\gamma_i{}^\beta$, $\gamma^i{}_\beta$, $l_I{}^\beta$ and $l^I{}_\beta$ are explicitly given[29,30] for a preferred set of coordinates, which is consistent with the quantization of the Z component of angular momentum in quantum theory.

Although σ appears whenever we map the generators X_i into differential operators according to (3.3) and (3.4), hereafter we simply set $\sigma=1$ for the sake of convenience.

B. Right and left vectors $\gamma_i{}^\beta$, $\gamma^j{}_\beta$, $l_I{}^\beta$, and $l^J{}_\beta$

From (3.5) and (3.6) it is clear that under a coordinate transformation $\alpha^\beta \to \alpha'^\beta$ in R_n, $\gamma_j{}^\beta$ and l_J^β transform according to

$$\gamma_j'^\beta = \frac{\partial \alpha'^\beta}{\partial \alpha^\gamma}\gamma_j{}^\gamma, \tag{3.7}$$

$$l_J'^\beta = \frac{\partial \alpha'^\beta}{\partial \alpha^\gamma} l^\gamma{}_J \tag{3.8}$$

and, therefore, are contravariant vectors. Hence, for each j (or J), j (or J) being the vector label, the quantities γ_j^β (or l_J^β) are the components of a contravariant vector, β indicating the component. Thus at every point in R_n, $\gamma_j{}^\alpha$ (or $l_J{}^\alpha$) may be chosen as a set of n linearly independent contravariant vectors.

Because all the right (or left) contravariant vectors $\gamma_j{}^\alpha$ (or $l_J{}^\alpha$) are linearly independent, their determinants $|\gamma_j{}^\alpha|$ (or $|l_J{}^\alpha|$) are different from zero. At every point of R_n we now introduce the covariant vectors $\gamma^j{}_\mu$ (or $l_\mu{}^J$) by dividing the cofactor of $\gamma^\alpha{}_j$ (or $l_J{}^\alpha$) in $|\gamma^\alpha{}_j|$ (or $|l_J{}^\alpha|$) by $|\gamma_j^\alpha|$ (or $|l_J{}^\alpha|$), i.e.,

$$\gamma_i{}^\mu \gamma^j{}_\mu = \delta_i{}^j, \tag{3.9}$$

$$l_I{}^\mu l^J{}_\mu = \delta^J{}_I. \tag{3.10}$$

Clearly, for each value of j (or J) the quantities $\gamma^j{}_\mu$ (or $l_\mu{}^J$) are the components of a covariant vector.

By the same token, we can easily show

$$\gamma^\mu{}_i \gamma^i{}_\nu = \delta^\mu{}_\nu, \tag{3.11}$$

$$l_I{}^\mu l^I{}_\nu = \delta^\mu{}_\nu. \tag{3.12}$$

C. The generalized Maurer–Cartan equations and the Lie structure formulas

Operating with the commutator $[Xi, Xj] = C_i{}^k{}_j X_k$ on $M(\alpha)$ from the right side, we obtain from (3.3)

$$\begin{aligned}&[\gamma_i{}^\mu \partial_\mu(\gamma_j{}^\nu \partial^\mu_\nu) - \gamma_j{}^\mu \partial_\mu(\gamma_i{}^\nu \partial^\mu_\nu)] D^{[p]}_{\langle m|m'\rangle}(\alpha)\\ &\quad = C_i{}^k{}_j\, \gamma_k{}^\nu \partial_\nu D^{[p]}_{\langle m|m'\rangle}(\alpha).\end{aligned} \tag{3.13}$$

Likewise, operating with the same commutator on $M(\alpha)$ from the left side, we obtain from (3.4)

$$\begin{aligned}&[l^\mu{}_I \partial_\mu (l^\nu{}_J \partial_\nu) - l_J{}^\mu \partial_\mu (l_I{}^\nu \partial_\nu)] D^{[p]}_{\langle m|m'\rangle}(\alpha)\\ &\quad = C_I'^K{}_J l_K{}^\nu \partial_\nu D^{[p]}_{\langle m|m'\rangle}(\alpha)\\ &\quad = -C_I{}^K{}_J l_K{}^\nu \partial_\nu D^{[p]}_{\langle m|m'\rangle}(\alpha).\end{aligned} \tag{3.14}$$

Note that in (3.14) the structure constants for the left sets of differential operators $C'_I{}^K{}_J$ are the negative of the original structure constants, i.e., $-C_I{}^K{}_J$. This can be proved as follows: Consider the operation of the commutator $[X_I, X_J] = C_{IJ}{}^K X_K$ from the left side on $M(\alpha)$. Operating with X_J on $X_I M(\alpha)$, i.e., $X_J X_I M(\alpha)$, we obtain $l_J{}^\mu \partial_\mu [l_I{}^\nu \partial_\nu D^{[p]}_{\langle m|m'\rangle}(\alpha)]$. Also, note that the order of the subscripts I and J in $X_J X_I$ are interchanged in the corresponding operation of the group elements $M(gg_I g_J)$. Thus, operating with $g_I g_J - g_J g_I = C_I{}^K{}_J g_K$ or correspondingly with $X_I X_J - X_J X_I = C'_I{}^K{}_J X_K = -C_I{}^K{}_J X_K$ on $M(g)$ results in

$$\begin{aligned}&[l_I{}^\mu \partial_\mu (l^\nu{}_J \partial_\nu) - l_J{}^\mu \partial_\mu (l_I{}^\nu \partial_\nu)] D^{[p]}_{\langle m|m'\rangle}(\alpha)\\ &\quad = C'_{IJ}{}^K l_K{}^\nu \partial_\nu D^{[p]}_{\langle m|m'\rangle}(\alpha)\\ &\quad = -C_I{}^K{}_J l_K{}^\nu \partial_\nu D_{\langle m|m'\rangle}(\alpha)\end{aligned}$$

which is (3.14).

Leaving out $D^{[p]}_{\langle m|m'\rangle}(\alpha)$ from both sides of (3.13) and (3.14), we obtain the Lie structure formulas[31] for the first and second parameter groups:

$$\partial_i \partial_j - \partial_j \partial_i = C_i{}^k{}_j \partial_k, \tag{3.15}$$

$$\partial_I \partial_J - \partial_J \partial_I = -C_I{}^K{}_J \partial_K, \tag{3.16}$$

where we have used $\partial_i = \gamma_i{}^\mu \partial_\mu$ and $\partial_I = l^\nu{}_I \partial_\nu$.

These may be written in the Lie derivative forms (refer to Sec. IV)

$$[\pounds_i, \pounds_j] = C_i{}^k{}_j \pounds_k, \tag{3.17}$$

$$[\pounds_I, \pounds_J] = -C_I{}^K{}_J \pounds_K. \tag{3.18}$$

Corresponding to these parameter groups, there exist two one-parameter groups of point transformations in R_n which are $\alpha^\mu \to \alpha'^\mu + \gamma_i{}^\mu t$ and $\alpha^\mu \to \alpha'^\mu + l_I{}^\mu t$ and transformations of which are in one-to-one correspondence with those of the original group. The left and right vectors serve to form two anholonomic coordinate systems in group space. On the left sides of (3.15) and (3.16) the second-order differential operators cancel out, when developed. Hence, we have the generalized Maurer–Cartan equations[32]

$$[\gamma_i{}^\mu (\partial_\mu \gamma_j{}^\nu) - \gamma_j{}^\mu (\partial_\mu \gamma_i{}^\nu)] \partial_\nu = C_i{}^k{}_j \gamma_k{}^\nu \partial_\nu, \tag{3.19}$$

$$[l_I{}^\mu (\partial_\mu l_J{}^\nu) - l_J{}^\mu (\partial_\mu l_I{}^\nu)] \partial_\nu = -C_I{}^K{}_J l_K{}^\nu \partial_\nu. \tag{3.20}$$

Substitution of the results of differentiating (3.9) with respect to α^δ, i.e., $\partial_\delta \gamma_i{}^\mu = -\gamma_i{}^\nu (\partial_\delta \gamma_\nu{}^j) \gamma_j{}^\mu$, into (3.9) yields

$$\partial_\nu \gamma_\mu{}^i - \partial_\mu \gamma_\nu{}^i = C_j{}^i{}_k \gamma^j{}_\mu \gamma^k{}_\nu. \tag{3.21}$$

Likewise, from (3.10) and (3.20) we obtain

$$\partial_\nu l^I{}_\mu - \partial_\mu l_\nu{}^I = -C_J{}^I{}_K l^J{}_\mu l^K{}_\nu. \tag{3.22}$$

The conditions (3.13), (3.14), (3.21), and (3.22) show that the vectors $\gamma^\alpha{}_i$, $\gamma^i{}_\alpha$, $l_I{}^\alpha$, and $l^I{}_\alpha$ form anholonomic coordinate systems, i.e., there does not exist a set of coordinates α^i such that $\partial_\beta \alpha^i = \gamma_\beta{}^i$ or $\partial_\beta \alpha^I = l^I{}_\beta$.

Equations (3.19), (3.20), (3.21) and (3.22) are actually the generalized Maurer–Cartan equations in specific forms inasmuch as $\gamma^\mu{}_i$, $\gamma^i{}_\mu$, $l_I{}^\mu$, and $l^I{}_\mu$ are the vectors in a more restricted sense, e.g., those of $U(m)$ if one starts with $M(\alpha)$ of $U(m)$. This corresponds to the first fundamental theorem of Lie which provides the structure constants from the vector fields. Conversely, once the structure constants are known, the Lie structure formulas as well as the Maurer–Cartan equations are merely identities which do not provide us any new information as far as the MEFRIR is concerned. In other words, being identities these equations are not useful for solving the problem of G_l despite that they have been widely quoted in the literature.

Examples of (3.19) and (3.20) and also of (3.21) and (3.22) for $SU(2)$ and $SO(3)$ are explicitly given[33] whereas those for $SU(3)$ can be obtained by the explicit expressions of $\gamma_i{}^\mu$, $\gamma^i{}_\mu$, $l_I{}^\mu$, and $l^I{}_\mu$.[28]

Besides the two parameter groups aforementioned, there is another transformation group known as the adjoint group whose group manifold is also given by R_n. This group transforms the right vector $\gamma_i{}^\mu$ (or $\gamma^i{}_\mu$) at a point in group space into the left vector $l_I{}^\mu$ (or $l^I{}_\mu$) at the same point.

IV. THE (+) AND (-) CONNECTIONS AND THE CURVATURE TENSORS

This section, added for the sake of completeness, discusses briefly the salient features of the (+) and (−) connections generated by the vector fields $\gamma_i{}^\mu$, $\gamma_\mu{}^i$, $l_I{}^\mu$, and $l_\mu{}^I$ and their relationships to the curvature tensors. A parallel method to that of CS is developed from the vector fields derived in the previous section. However, it should be kept in mind throughout that our vector fields are generated by the group of inner automorphisms, and our differential operators are always to be understood to operate on the MEFRIR $D^{[p]}_{\langle m|m'\rangle}(\alpha)$ although sometimes $D^{[p]}_{\langle m|m'\rangle}(\alpha)$ is omitted. In the next section, we shall make use of some results obtained in this section.

To any pair of elements X, Y there always belong two elements XY^{-1} and $Y^{-1}X$. Two pairs X, Y and X_1, Y_1 are called (+) equipollent if $XY^{-1}=X_1Y_1^{-1}$ and (−) equipollent if $Y^{-1}X=Y_1^{-1}X_1$. Now consider a right translation from a point in group space represented by X_β to another point $X'_{\alpha+\beta}$ through $X'_{\alpha+\beta}=M_\alpha X_\beta$. Since $X'_{\alpha+\beta}X_\beta^{-1}=M_\alpha$ is independent of any point β, we have $X'_{\alpha+\beta}X_\beta^{-1}=Y'_{\alpha+\gamma}Y_\gamma^{-1}$, γ being another point. That means

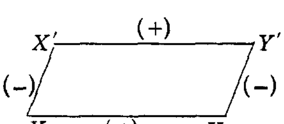
$$X'X^{-1}=Y'Y^{-1},$$

and, therefore,

$$Y^{-1}X=Y'^{-1}X'.$$

Thus X, Y and X', Y' are (−) equipollent. From $XX'^{-1}=YY'^{-1}$ it also follows that X, X' and Y, Y' are (+) equipollent. Hence, there exist (−) geodesics which transform X to X' and Y to Y' and (+) geodesics which transform X to Y and X' to Y'.

A right translation therefore transforms any element X into a (−) equipollent element X'.



Right translation

Likewise, for left translation X, Y and X', Y' are (+) equipollent and hence X, X' and Y, Y' are (−) equipollent. A left translation transforms any element into its (+) equipollent element.

With this preliminary for a proof of the existence of nonsymmetric connections we now pass to the (+) and (−) connections defined by

$$\bar{\Gamma}^{\gamma}_{\beta\alpha}\equiv-\gamma_\beta{}^i\partial_\alpha\gamma_i{}^\gamma=\gamma_i{}^\gamma\partial_\alpha\gamma_\beta{}^i, \tag{4.1}$$

$$\overset{+}{\Gamma}{}^{\gamma}_{\beta\alpha}\equiv-l_\beta{}^I\partial_\alpha l^\gamma{}_I=l_I{}^\gamma\partial_\alpha l_\beta{}^I, \tag{4.2}$$

where the second equality in each expression has been obtained by use of the orthonormality relations (3.9) and (3.10). There are altogether $2n^3$ independent components for $\overset{\pm}{\Gamma}{}^{\gamma}_{\beta\alpha}$ as compared with $n^2(n+1)/2$ independent components for $\Gamma^\gamma{}_{\beta\alpha}$.

The fact that these quantities are indeed linear affine connections can easily be proved by showing that under a coordinate transformation $\alpha^\mu\to\alpha'^\mu$, $\overset{\pm}{\Gamma}{}^\gamma{}_{\alpha\beta}$ respectively transform according to

$$\overset{\pm}{\Gamma}{}'^{\epsilon}_{\alpha\beta}\frac{\partial\alpha^\gamma}{\partial\alpha'^\epsilon}=\frac{\partial^2\alpha^\gamma}{\partial\alpha'^\alpha\partial\alpha'^\beta}+\overset{\pm}{\Gamma}{}^{\gamma}_{\delta\epsilon}\frac{\partial\alpha^\delta}{\partial\alpha'^\alpha}\frac{\partial\alpha^\epsilon}{\partial\alpha'^\beta}, \tag{4.3}$$

which are the conditions that $\overset{\pm}{\Gamma}{}^\gamma{}_{\alpha\beta}$ must satisfy in order that they determine linear connections.

Next, (4.1) and (4.2) together with the orthonormality relations (3.9) and (3.10) yield

$$\bar{\nabla}_\rho\gamma^j{}_\delta\equiv\partial_\rho\gamma^j{}_\delta-\bar{\Gamma}^\beta{}_{\delta\rho}\gamma^j{}_\beta=0, \tag{4.4}$$

$$\bar{\nabla}_\rho\gamma^\delta{}_j\equiv\partial_\rho\gamma_j{}^\delta+\bar{\Gamma}^\delta{}_{\beta\rho}\gamma^\beta{}_j=0, \tag{4.5}$$

$$\overset{+}{\nabla}_\rho l^I{}_\delta\equiv\partial_\rho l^I{}_\delta-\overset{+}{\Gamma}{}^\beta{}_{\delta\rho}l^I{}_\beta=0, \tag{4.6}$$

$$\overset{+}{\nabla}_\rho l^\delta{}_I\equiv\partial_\rho l_I{}^\delta+\overset{+}{\Gamma}{}^\delta{}_{\beta\rho}l^\beta{}_I=0, \tag{4.7}$$

where $\overset{\pm}{\nabla}_\rho$ are respectively the covariant derivatives with respect to the (+) and (−) connections. Because these are the conditions that the first covariant derivatives of $\gamma^i{}_\delta$ and $\gamma_i{}^\delta$ (or $l^I{}_\delta$ and $l^\delta{}_I$) with respect to the (−) [or (+)] connection be zero, each of the right fields $\gamma^i{}_\delta$ and $\gamma_i{}^\delta$ (or the left fields $l^I{}_\delta$ and $l_I{}^\delta$) forms a parallel field with respect to the (−) [or (+)] connection. In other words, vectors at different points can be transformed into each other by a combination of (±) parallel transformations.

In view of the fact that $\overset{\pm}{\Gamma}{}^\rho{}_{\mu\nu}$ are nonsymmetric, it is convenient to have the relationships between $\overset{\pm}{\Gamma}{}^\rho{}_{\mu\nu}$ and $\overset{\pm}{\Gamma}{}^\rho{}_{\nu\mu}$ (μ and ν exchanged) which are obtained as follows: Multiplying (3.21) [or (3.22)] by $\gamma^\rho{}_i$ (or $l^\rho{}_I$) and summing over i (or I), we get

$$\gamma^i{}_\nu\partial_\mu\gamma^\rho{}_i-\gamma^i{}_\mu\partial_\nu\gamma_i{}^\rho=C_j{}^i{}_k\gamma^\rho{}_i\gamma^j{}_\mu\gamma_\nu{}^k, \tag{4.8}$$

$$l_\nu{}^I\partial_\mu l_I{}^\rho-l^I{}_\mu\partial_\nu l^\rho{}_I=-C_J{}^I{}_K l^\rho{}_I l^J{}_\mu l_\nu{}^K, \tag{4.9}$$

which, upon taking (4.4) and (4.5) into consideration, lead to

$$\bar{\Gamma}^\rho{}_{\mu\nu}=\bar{\Gamma}^\rho{}_{\nu\mu}+C_j{}^i{}_k\gamma^\rho{}_i\gamma^j{}_\mu\gamma^k{}_\nu, \tag{4.10}$$

$$\overset{+}{\Gamma}{}^\rho{}_{\mu\nu}=\overset{+}{\Gamma}{}^\rho{}_{\nu\mu}-C_J{}^I{}_K l_I{}^\rho l^J{}_\mu l^K{}_\nu. \tag{4.11}$$

Now we observe that the right contravariant vector field $\gamma_i{}^\alpha$ is absolutely invariant with respect to the left contravariant field $l_I{}^\beta$ and vice versa, i.e., the Lie derivative of $\gamma_i{}^\alpha$ with respect to $l_I{}^\beta$ is zero and vice versa:

$$\mathfrak{L}_J\gamma_i{}^\beta\equiv l_J{}^\gamma\partial_\gamma\gamma_i{}^\beta-\gamma_i{}^\beta\partial_\gamma l_J{}^\beta=0, \tag{4.12}$$

$$\mathfrak{L}_j l_I{}^\beta\equiv\gamma_j{}^\gamma\partial_\gamma l_I{}^\beta-l^\gamma{}_I\partial_\gamma\gamma_j{}^\beta=0. \tag{4.13}$$

Consequently, we can write

$$\pounds_J \gamma_i{}^\beta = l_J{}^\gamma(\partial_\gamma \gamma_i{}^\beta - \overset{+}{\Gamma}{}_\delta{}^\beta{}_\gamma \gamma_i{}^\delta) = 0, \tag{4.14}$$

$$\pounds_j l_I{}^\beta = \gamma^\gamma{}_j(\partial_\gamma l_I{}^\beta + \bar{\Gamma}{}_\delta{}^\beta{}_\gamma l_I{}^\delta) = 0. \tag{4.15}$$

Making use of (4.5), (4.7), (4.14), and (4.15) leads to

$$\overset{+}{\Gamma}{}^\alpha{}_{\beta\gamma} = \bar{\Gamma}_\gamma{}^\alpha{}_\beta. \tag{4.16}$$

This condition reduces the number of independent components of $\overset{\pm}{\Gamma}{}_\alpha{}^\gamma{}_\beta$ from $2n^3$ to n^3.

If $\overset{\pm}{\Gamma}{}_\alpha{}^\gamma{}_\beta$ is decomposed into its symmetric ($\overset{\pm}{S}{}_\alpha{}^\gamma{}_\beta$) and antisymmetric ($\overset{\pm}{A}{}_\alpha{}^\gamma{}_\beta$) parts, then

$$\overset{\pm}{\Gamma}{}_\alpha{}^\gamma{}_\beta = \overset{\pm}{S}{}_\alpha{}^\gamma{}_\beta + \overset{\pm}{A}{}_\alpha{}^\gamma{}_\beta \equiv \tfrac{1}{2}(\overset{\pm}{\Gamma}{}_\alpha{}^\gamma{}_\beta + \overset{\pm}{\Gamma}{}_\beta{}^\gamma{}_\alpha) + \tfrac{1}{2}(\overset{\pm}{\Gamma}{}_\alpha{}^\gamma{}_\beta - \overset{\pm}{\Gamma}{}_\beta{}^\gamma{}_\alpha). \tag{4.17}$$

From (4.14), (4.15), and (4.16)

$$\overset{+}{S}{}_\alpha{}^\gamma{}_\beta = \bar{S}_\alpha{}^\gamma{}_\beta = \Gamma_\alpha{}^\gamma{}_\beta, \tag{4.18}$$

$$\overset{+}{A}{}_\alpha{}^\gamma{}_\beta = \bar{A}_\beta{}^\gamma{}_\alpha \equiv A_\beta{}^\gamma{}_\alpha. \tag{4.19}$$

The $A_\beta{}^\gamma{}_\alpha$ is also known as the torsion tensor.

The last equality of (4.18) means that the symmetric parts of the (±) connections are, in fact, the (0) connection in R_n. This can be shown in a way similar to *CS* or Yano and Bochner[34] and therefore is omitted. It may be shown in the ordinary way that the symmetric parts $\overset{\pm}{S}{}_\alpha{}^\gamma{}_\beta$ are not tensors whereas the antisymmetric parts $\overset{\pm}{A}{}_\alpha{}^\gamma{}_\beta$ are tensors in R_n.

Making use of (3.9), (4.8), and (4.12), we have

$$\partial_\nu \gamma_\mu{}^j - \partial_\mu \gamma_\nu{}^j = C_k{}^j{}_l \gamma_\mu{}^k \gamma^l{}_\nu \tag{4.20}$$

i.e., from (4.1)

$$(\bar{\Gamma}_\mu{}^\beta{}_\nu - \bar{\Gamma}_\nu{}^\beta{}_\mu)\gamma_\beta{}^j = C^j{}_{kl}\gamma_\mu{}^k \gamma^l{}_\nu,$$

or from (4.17)

$$\bar{A}^\beta{}_{\mu\nu} = \tfrac{1}{2} C_k{}^j{}_l \gamma_\mu{}^k \gamma_\nu{}^l \gamma^\beta{}_j. \tag{4.21}$$

Since the right vector is invariant for dragging along over the left field,

$$\pounds_I \bar{A}^\beta{}_{\mu\nu} = \tfrac{1}{2} C_k{}^j{}_l \pounds_I(\gamma_\mu{}^k \gamma_\nu{}^l \gamma^\beta{}_j) = 0 \text{ or } \overset{+}{\nabla}_\gamma \bar{A}^\beta{}_{\mu\nu} = 0. \tag{4.22}$$

Likewise, from (4.19)

$$\overset{+}{A}{}^\beta{}_{\mu\nu} = -\tfrac{1}{2} C^J{}_{KL} l^K{}_\mu l^L{}_\nu l^\beta{}_J \tag{4.23}$$

and

$$\pounds_j \overset{+}{A}{}^\beta{}_{\mu\nu} = 0 \quad\text{or}\quad \bar{\nabla}_\gamma \overset{+}{A}{}^\beta{}_{\mu\nu} = 0. \tag{4.24}$$

From (4.19), (4.22), and (4.24) we get

$$\overset{\pm}{\nabla}_\gamma A_\mu{}^\beta{}_\nu = 0, \quad \nabla_\gamma A_\mu{}^\beta{}_\nu = 0. \tag{4.25}$$

The last equality is obtained by the first two equalities upon taking $\nabla_\gamma = \tfrac{1}{2}(\overset{+}{\nabla}_\gamma + \bar{\nabla}_\gamma)$ into account.

The Riemann tensors of the (±) connections are defined by

$$\overset{\pm}{R}{}^\delta{}_{\alpha\beta\gamma} \equiv \partial_\beta \overset{\pm}{\Gamma}{}^\delta{}_{\alpha\gamma} - \partial_\gamma \overset{\pm}{\Gamma}{}^\delta{}_{\alpha\beta} + \overset{\pm}{\Gamma}{}^\epsilon{}_{\alpha\gamma}\overset{\pm}{\Gamma}{}^\delta{}_{\epsilon\beta} - \overset{\pm}{\Gamma}{}^\epsilon{}_{\alpha\beta}\overset{\pm}{\Gamma}{}^\delta{}_{\epsilon\gamma} \tag{4.26}$$

which vanish identically—the proof can be shown in a way similar to Eisenhart.[35]

The relation between the curvature tensor of the (0) connection and that of the (−) connection, for instance, is obtained in consequence of $\Gamma^\alpha{}_{\beta\gamma} = \overset{+}{\Gamma}{}^\alpha{}_{\beta\gamma} - \overset{+}{A}{}^\alpha{}_{\beta\gamma}$:

$$R^\alpha{}_{\beta\gamma\delta} = \bar{R}^\alpha{}_{\beta\gamma\delta} - \nabla_\gamma A^\alpha{}_{\beta\delta} + \nabla_\delta A^\alpha{}_{\beta\gamma} + A_{\beta\delta}{}^\epsilon A^\alpha{}_{\epsilon\gamma} - A_{\beta\gamma}{}^\epsilon A_\epsilon{}^\alpha{}_\delta + \bar{\Gamma}_{\gamma\delta}{}^\epsilon A^\alpha{}_{\beta\epsilon} - \bar{\Gamma}^\epsilon{}_{\delta\gamma} A^\alpha{}_{\beta\epsilon}. \tag{4.27}$$

Since the first three terms on the right side vanish, it follows (4.17) and Jacobi's identity that

$$R^\alpha{}_{\beta\gamma\delta} = -A^\epsilon{}_{\gamma\delta} A^\alpha{}_{\epsilon\beta}. \tag{4.28}$$

Combining with (4.25), we obtain

$$\overset{\pm}{\nabla}_\rho R^\alpha{}_{\beta\gamma\delta} = 0, \quad \nabla_\rho R^\alpha{}_{\beta\gamma\delta} = 0, \tag{4.29}$$

i.e., the Riemann tensor of the (0) connection is constant in any connection. The last condition of (4.29) is the one required for a symmetric space.

In consequence of the Theorem [Eq. (5.1), Sec. V] we may write (4.21) as

$$C_i{}^k{}_l \gamma_\mu{}^i = -2A^\beta{}_{\mu\nu}\gamma^k{}_\beta \gamma^\nu{}_l,$$

so that

$$g_{\mu\nu} = 4A^\beta{}_{\mu\gamma} A_\nu{}^\gamma{}_\beta. \tag{4.30}$$

On the other hand, from (4.28) the Ricci tensor can be written as

$$R_{\beta\delta} = -A_\beta{}^\alpha{}_\epsilon A_\delta{}^\epsilon{}_\alpha. \tag{4.31}$$

Comparing (4.30) with (4.31), we have

$$R_{\mu\nu} = -4g_{\mu\nu}, \tag{4.32}$$

i.e., $R_{\mu\nu}$ is symmetrical and proportional to the metric tensor $g_{\mu\nu}$. Accordingly, we have the symmetric Ricci tensor for the group manifold of G_l. Using the scalar curvature $R = R^\mu{}_\mu$, we may rewrite (4.30) in the form

$$R_{\mu\nu} = \frac{R}{n} g_{\mu\nu}. \tag{4.33}$$

Therefore, we conclude that the group space of a semisimple Lie group is an Einstein space, i.e., a space in which $R_{\mu\nu}$ differs from $g_{\mu\nu}$ only by a scalar factor which is equal to -4 in this case. Clearly, the group space is homogeneous with respect to the Ricci tensor. In other words, the principal directions for the Ricci tensor become indeterminate.[36]

It might be of interest and useful to classify group spaces algebraically, but this would require further knowledge on elementary divisors and therefore is not within the scope of the present work.

V. THE INVARIANT EQUATIONS

It has been shown that a group of rank l has l group invariants, which are usually expressed in terms of sums of products of generators. In particular, Gelfand[37] and Biedenharn[38] have explicitly given these invariants for $U(m)$ whereas Racah,[39] Gruber and O'Raifearteigh[40] those for $O(m)$. We derive by induction the invariant differential equations to be satisfied by the MEFRIR of G_l.

First, we present the following theorem to which we have already had recourse in the previous section.

Theorem: The metric tensors of R_n are related to the real vectors $\gamma_i{}^\alpha$, $\gamma^i{}_\alpha$, $l_i{}^\alpha$, and $l_\alpha{}^i$ by

$$g^{\alpha\beta} = G^{IJ} l_I{}^\alpha l_J{}^\beta = G^{ij}\gamma_i{}^\alpha \gamma_j{}^\beta$$

and

$$g_{\alpha\beta}=G_{IJ}\,l^I{}_\alpha\,l^J{}_\beta=G_{ij}\,\gamma^i{}_\alpha\,\gamma^j{}_\beta, \qquad (5.1)$$

where

$$G_{ij}=C_i{}^k{}_l\,C_j{}^l{}_k=C_{IJ}$$

and G^{ij} (or G^{IJ}) are given by $G^{ij}\,G_{jl}=\delta^i{}_l$.

Proof: The quantity $g^{\alpha\beta}=G^{ij}\gamma_i{}^\alpha\gamma_j{}^\beta$ (or $G^{IJ}l^\alpha{}_I l^\beta{}_J$) has the determinant of rank n since $|G^{ij}|$ is of the same rank. Also, $\nabla_\gamma g^{\alpha\beta}=0$ from (4.21), (4.29) and (4.30). It can then be shown that $\Gamma_{\alpha\beta,\gamma}$ and $\Gamma_\alpha{}^\gamma{}_\beta$ defined by (2.7) are respectively the Christoffel's sumbols of the first and second kinds in R_n and therefore the symmetric parts of $\overset{+}{\Gamma}_{\alpha\beta,\gamma}$ (or $\bar{\Gamma}_{\alpha\beta,\gamma}$) and $\overset{+}{\Gamma}_\alpha{}^\gamma{}_\beta$). Consequently, $g^{\alpha\beta}$ is none other than the metric tensor of R_n. The proof of $g_{\alpha\beta}$ can be carried out in like manner.

This theorem proved, we now proceed to the derivation of the second-order invariant equation for G_l with the help of the vectors $\gamma_i{}^\alpha$ and $l_I{}^\alpha$. Operation with the invariant product of generators[41] $2G^{ij}\,X_iX_j$ from the right side on $M(\alpha)$ is equivalent to operation with $2G^{ij}\gamma_i{}^\gamma\partial_\gamma\gamma_j{}^\beta\partial_\beta$ from the left side on $\mathcal{D}^{[p]}_{(m|m')}(\alpha)$. Hence we obtain

$$\begin{aligned}2G^{ij}\gamma_i{}^\alpha\,\partial_\alpha(\gamma^\beta{}_j\,\partial_\beta\mathcal{D}^{[p]}_{(m|m')}(\alpha))\\ &=2(g^{\alpha\beta}\partial^2{}_{\alpha\beta}-\bar{\Gamma}_{\beta\gamma}{}^\gamma g^{\alpha\beta}\partial_\gamma)\,\mathcal{D}^{[p]}_{(m|m')}(\alpha)\\ &=2(g^{\alpha\beta}\partial^2{}_{\alpha\beta}-\Gamma_\alpha{}^\gamma{}_\beta g^{\alpha\beta}\partial_\gamma)\,\mathcal{D}^{[p]}_{(m|m')}(\alpha),\end{aligned} \qquad (5.2)$$

where we have used the above theorem together with (4.1), (4.17), and $A_\alpha{}^\gamma{}_\beta\,g^{\alpha\beta}=0$. Upon taking (2.6) and (2.7) into consideration, it follows that the above equation can be cast in the covariant form

$$\begin{aligned}(5.2)&=2\,\frac{1}{\sqrt{g}}\,\frac{\partial}{\partial\alpha^\alpha}\left(\sqrt{g}\,g^{\alpha\beta}\frac{\partial}{\partial\alpha^\beta}\,\mathcal{D}^{[p]}_{(m|m')}(\alpha)\right)\\ &=2\nabla^\gamma\,\nabla_\gamma\mathcal{D}^{[p]}_{(m|m')}(\alpha)=2\bar{\nabla}^\gamma\,\bar{\nabla}_\gamma\mathcal{D}^{[p]}_{(m|m')}(\alpha).\end{aligned} \qquad (5.3)$$

The operation of $2G^{IJ}X_IX_J$ from the left side on $M(\alpha)$ can be treated in the same way except that $\gamma_i{}^\alpha$ and $\bar{\Gamma}_\beta{}^\alpha{}_\gamma$ have to be replaced respectively with $l_I{}^\alpha$ and $\overset{+}{\Gamma}{}^\alpha{}_{\beta\gamma}$, and then we obtain

$$2\;\overset{+}{\nabla}{}^\gamma\;\overset{+}{\nabla}_\gamma\mathcal{D}^{[p]}_{(m|m')}(\alpha). \qquad (5.4)$$

By equating (5.2)—(5.4) to an invariant function of the partition label $[p]$, $-2f_1([p])$, we get the second-order invariant equation of G_l in the covariant form of any $[(+),\ (-),\ \text{or}\ (0)]$ connection

$$\begin{aligned}\nabla^\gamma\nabla_\gamma\mathcal{D}^{[p]}_{(m|m')}(\alpha)&=\overset{+}{\nabla}{}^\gamma\overset{+}{\nabla}_\gamma\mathcal{D}^{[p]}_{(m|m')}(\alpha)=\bar{\nabla}^\gamma\,\bar{\nabla}_\gamma\mathcal{D}^{[p]}_{(m|m')}(\alpha)\\ &=-f_1([p])\mathcal{D}^{[p]}_{(m|m')}(\alpha).\end{aligned} \qquad (5.5)$$

In the following, we use CS's symmetrically coupled coefficients[42] $g_{ij\cdots k}$—Biedenharn's notation $[ij\cdots k]$ is actually preferred—to derive, for example, the third- and higher-order invariant equations of $U(m)$ although those of other groups can be derived in a similar way.

Applying the third-order invariant operator $[ijk]X_iX_jX_k$ from the right side on $M(\alpha)$ leads to

$$[ijk]\gamma_i{}^\alpha\,\partial_\alpha\,\gamma_j{}^\beta\,\partial_\beta\gamma_k{}^\gamma\,\partial_\gamma\mathcal{D}^{[p]}_{(m|m')}(\alpha)=-f_2([p])\mathcal{D}^{[p]}_{(m|m')}(\alpha) \qquad (5.6a)$$

where $f_2([p])$ is another function of the partition label $[p]$. From (4.5)

$$\partial_\beta(\gamma_k{}^\gamma\psi)=(\gamma_k{}^\gamma\,\partial_\beta+\bar{\Gamma}_\delta{}^\gamma{}_\beta\,\gamma_k{}^\delta)\psi=\gamma^\gamma{}_k\;\bar{\nabla}_\beta\psi$$

where ψ is any function of the α^β's. Again, from (4.5) we have

$$\gamma_i{}^\alpha\,\partial_\alpha(\gamma_j{}^\beta\,\gamma_k{}^\gamma\,\bar{\nabla}_\beta\psi)=\gamma_i{}^\alpha\,\gamma_j{}^\beta\,\bar{\nabla}_\alpha(\gamma_k{}^\gamma\,\bar{\nabla}_\beta\psi)=\gamma^\alpha{}_i\gamma_j{}^\beta\gamma_k{}^\gamma\,\bar{\nabla}_\alpha\,\bar{\nabla}_\beta\psi.$$

Similarly, ∂_γ can also be replaced by $\bar{\nabla}_\gamma$ since the invariant operator is completely symmetric in i, j, and k. Hence, the third-order invariant equation can be written in terms of the (−) connection as

$$[ijk]\gamma_i{}^\alpha\,\gamma^\beta{}_j\,\gamma_k{}^\gamma\,\bar{\nabla}_\alpha\;\bar{\nabla}_\beta\,\bar{\nabla}_\gamma\mathcal{D}^{[p]}_{(m|m')}(\alpha)=-f_2([p])\mathcal{D}^{[p]}_{(m|m')}(\alpha). \qquad (5.6b)$$

In the above equation we can replace $\bar{\nabla}_\alpha\;\bar{\nabla}_\beta\;\bar{\nabla}_\gamma$ by $\nabla_\alpha\nabla_\beta\nabla_\gamma$ because of the complete symmetry in i, j, and k, so that the left side in (5.6b) becomes

$$[ijk]\gamma_i{}^\alpha\,\gamma_j{}^\beta\gamma_k{}^\gamma\,\nabla_\alpha\;\nabla_\beta\,\nabla_\gamma\mathcal{D}^{[p]}_{(m|m')}(\alpha). \qquad (5.6c)$$

By the same token, $\bar{\nabla}_\gamma$ (or ∇_γ) can also be replaced by $\overset{+}{\nabla}_\gamma$ (or ∇_γ) when the $\gamma^\alpha{}_i$'s are replaced by the $l^\alpha{}_I$'s. In consequence of the foregoing discussion, it follows that the third-order invariant equation takes the covariant form

$$[ijk]\,t_i{}^\alpha\,t_j{}^\beta t^\gamma{}_k\Lambda_\alpha\,\Lambda_\beta\Lambda_\gamma\mathcal{D}^{[p]}_{(m|m')}(\alpha)=-f_2([p])\mathcal{D}^{[p]}_{(m|m')}(\alpha) \qquad (5.7)$$

where $t_i{}^\alpha$ is either $l_I{}^\alpha$, when $\Lambda_\alpha=\nabla_\alpha$ or $\overset{+}{\nabla}_\alpha$, or $\gamma_i{}^\alpha$, when $\Lambda_\alpha=\nabla_\alpha$ or $\bar{\nabla}_\alpha$.

It is worth noting that, as an example, in the expression $[ijk]\;\gamma_i{}^\alpha\,\gamma_j{}^\beta\,\gamma_k{}^\gamma\;\nabla_\alpha\;\;\nabla_\beta\;\nabla_\gamma$ the symmetry with respect to any pair of indices, say α and β, is not obvious since $\nabla_\alpha\;\nabla_\beta\neq\nabla_\beta\;\nabla_\alpha$. However, by interchanging α and β, we get[43] $\nabla_\alpha\;\nabla_\beta\nabla_\gamma=\nabla_\beta\nabla_\alpha\;\nabla_\gamma+R^\delta{}_{\gamma\alpha\beta}\,\nabla_\delta$ so that

$$\begin{aligned}[ijk]\gamma_i{}^\alpha\,\gamma^\beta{}_j\gamma^\gamma{}_k\,\nabla_\alpha\;\nabla_\beta\,\nabla_\gamma&=[ijk]\gamma^\beta{}_i\,\gamma^\alpha{}_j\,\gamma^\gamma{}_k\,\nabla_\beta\,\nabla_\alpha\;\nabla_\gamma\\ &=[ijk]\gamma^\beta{}_i\gamma^\alpha{}_j\,\gamma^\gamma{}_k(\nabla_\alpha\;\nabla_\beta\,\nabla_\gamma-R^\delta{}_{\gamma\alpha\beta}\,\nabla_\delta).\end{aligned} \qquad (5.8)$$

The last term on the right side vanishes identically because $[ijk]\;\gamma^\beta{}_i\;\gamma^\alpha{}_j\;\gamma^\gamma{}_k$ is symmetric in α and β, but $R^\delta{}_{\gamma\alpha\beta}$ is antisymmetric. The complete symmetry under a permutation of any pair of indices α, β and γ has therefore been proved. This proof, however, is not limited to the third-order invariant equation and holds in general to any order.

When we apply to the qth-order invariant ($q=2,\ 3,\ \ldots,\ l+1$) reasoning similar to that which led to the third-order invariant differential equation (5.7), we have the qth-order invariant differential equation

$$\begin{aligned}[i_1,i_2\cdots i_q]\,t^{\alpha_1}_{i_1}\,t^{\alpha_2}_{i_2}\cdots t^{\alpha_q}_{i_q}\Lambda_{\alpha_1}\,\Lambda_{\alpha_2}\cdots\Lambda_{\alpha_q}\mathcal{D}^{[p]}_{(m|m')}(\alpha)\\ =-f_{q-1}([p])\mathcal{D}^{[p]}_{(m|m')}(\alpha),\end{aligned} \qquad (5.9)$$

where $t^\alpha{}_I$ is either $l^\alpha{}_I$, when $\Lambda^\alpha=\nabla^\alpha$ or $\overset{+}{\nabla}{}^\alpha$, or $\gamma^\alpha{}_i$, when $\Lambda_\alpha=\nabla_\alpha$ or $\bar{\nabla}_\alpha$, and $f_{q-1}([p])$ is a function of the partition label $[p]$, which we shall call the $(q-1)$th partition function of G_l.

Similarly, with the help of the covariant vectors $\gamma_\alpha{}^i$ and $l^I{}_\alpha$, we can write Eq. (5.9) as

$$\begin{aligned}[i_1,\ i_2\cdots i_q]\,t^{i_1}_{\alpha_1}\,t^{i_2}_{\alpha_2}\cdots t^{i_q}_{\alpha_q}\Lambda^{\alpha_1}\Lambda^{\alpha_2}\cdots\Lambda^{\alpha_q}\mathcal{D}^{[p]}_{(m|m')}(\alpha)\\ =-f_{q-1}([p])\mathcal{D}^{[p]}_{(m|m')}(\alpha)\end{aligned} \qquad (5.10)$$

where $t^i{}_\alpha$ is either $l^I{}_\alpha$ when $\Lambda^\alpha=\nabla^\alpha$ or $\overset{+}{\nabla}{}^\alpha$, or $\gamma^i{}_\alpha$, when $\Lambda^\alpha=\nabla^\alpha$ or $\bar{\nabla}^\alpha$.

Finally, by inspection we can easily show that the qth-order invariant equation of any group can be written in the forms of Lie derivatives:

$$[i_1, i_2 \cdots i_q] \mathfrak{L}_{i_1} \mathfrak{L}_{i_2} \cdots \mathfrak{L}_{i_q} D^{[p]}_{\langle m|m'\rangle}(\alpha) = [I_1 I_2 \cdots I_q] \times \mathfrak{L}_{I_1} \mathfrak{L}_{I_2} \cdots \mathfrak{L}_{I_q} D^{[p]}_{\langle m|m'\rangle}(\alpha) = -f_{q-1}([p]) D^{[p]}_{\langle m|m'\rangle}(\alpha). \quad (5.11)$$

If a point transformation belong to a coordinate transformation $\alpha^\mu \to \alpha'^\mu$ we say that α^μ is dragged along by this point transformation. The numerical values of the coordinates of a point remain invariant if any point transformation is applied and if the coordinate system is dragged along by the same transformation. Equation (5.11) thus means that the MEFRIR is invariant by dragging along over either the left or the right field.

Examples will make the meaning of these invariant equation easily understandable, and here we choose the invariant equations of the groups $SU(2)$, $SO(3)$, and $SU(3)$ for this purpose.

Example 1: For each of the groups $SU(2)$ and $SO(3)$,[44] a partition is given either by a positive integer or by a half-odd integer, denoted by j. The $f_1(j)$ is then $j(j+1)$ and the second-order invariant equation becomes

$$\nabla^\mu \nabla_\mu D^j_{\langle m|m'\rangle}(\alpha) = -j(j+1) D^j_{\langle m|m'\rangle}(\alpha)$$

which is the Schrödinger equation of the symmetric top for stationary states. Here we have used the natural units so that no physical constant appears in the equation. A detailed account of this equation and other properties of these groups will be published elsewhere.

Example 2: The group $SU(3)$ has two invariant equations, one of the second-order and the other of the third-order. The partition is specified by the two positive integers p_1 and p_2 where $p_1 \geq p_2 \geq 0$. With the known expressions of $f_1(p_1, p_2)$ and $f_2(p_1, p_2)$,[45] the second- and third-order invariant equations become

$$\nabla^\mu \nabla_\mu D^{(p_1,p_2)}_{\langle I_1 m_1 y_1 | I_2 m_2 y_2\rangle}(\alpha) = -f_1(p_1, p_2) D^{(p_1,p_2)}_{\langle I_1 m_1 y_1 | I_2 m_2 y_2\rangle}(\alpha), \quad (5.12)$$

$$[ijk] t^\mu{}_i t^\nu{}_j t^\rho{}_k \nabla_\mu \nabla_\nu \nabla_\rho D^{(p_1,p_2)}_{\langle I_1 m_1 y_1 | I_2 m_2 y_2\rangle}(\alpha) = -f_2(p_1,p_2) D^{(p_1,p_2)}_{\langle I_1 m_1 y_1 | I_2 m_2 y_2\rangle}(\alpha) \quad (5.13)$$

where $t_j{}^\mu$ is either $l_J{}^\mu = L_J{}^\mu / i$ or $\gamma_j{}^\mu = R_j{}^\mu / i$ of (A.1) or (A.2) in Ref. 16.

If the total isospin operator $\mathbb{I}^2$, its third component I_z, and the hypercharge operator Y are chosen as a complete set of commuting operators with the eigenvalues $I(I+1)$, m, and y, respectively, four of the octet angles can be eliminated and replaced by the two sets of eigenvalues I, m, y and I', m', y'. The simultaneous quantization conditions then are[46]

$$[I_z, \alpha_I] = -i\not{e}, \quad [Y, \alpha_H] = -i\not{g} \quad (5.14)$$

where $\not{e}$ (or $\not{g}$) is the fundamental unit of the electric (or magnetic) charge divided by 2π, α_I (or α_H) being any one of the isospin (or hypercharge) angles.

The following equation may then be regarded as a new nonrelativistic wave equation to be satisfied by a new wave function ψ[47,48]

$$\frac{\hbar^2}{2I} \nabla^\mu \nabla_\mu \psi = E\psi = i\hbar \frac{\partial}{\partial t}\psi \quad (5.15)$$

where t is the time variable and I the moment of inertia of a $SU(3)$ multiplet. An alternative to Eq. (5.15) is to replace E by E^2 on the right side. If one keeps in Eq. (5.15) the constants $\not{e}$ and $\not{g}$ introduced through the conditions (5.14), the structure constants $\not{e}^2/\hbar c$ and $\not{g}^2/\hbar c$ then appear in the invariant equations. In stationary states, Eq. (5.15) reduces to Eq. (5.12), i.e., the energy eigenvalues are specified by the partition function $f_1([p])$ and eigenfunctions become solutions of (5.12), namely, the MEFRIR $D^{[p]}_{\langle m|m'\rangle}(\alpha)$ of $SU(3)$. This new quantum theory is based on a new correspondence principle which asserts that the motion of a system as described by the new quantum theory and by the ordinary quantum theory should agree in the limit when e and g can be neglected. We shall discuss this new theory in detail elsewhere.

The left side of Eq. (5.15), when fully developed, consists of about 90 terms, whereas the third-order invariant equation (5.13) contains over 200 terms.

One characteristic feature of $SU(3)$ which is completely different from $SU(2)$ or $SO(3)$ is that the invariant equations (5.12) and (5.13), when expressed in terms of the eigenvalues I, m, y and I', m', y', become essentially complex.

VI. CONCLUSIONS

In conclusion, we list some new results obtained by our method and some of the differences between our method and CS's.

1. We have started with the concept of a quadratic line element of G_l through the representative matrix method to identify the group manifold R_n as a Riemann space to make Riemann geometry available to G_l, and to construct explicitly the Riemann tensor, the Ricci tensor, and other geometric quantities for the (0) connection. A new and simpler method of calculating the invariant volume element and integral has been introduced within R_n.

2. From the right and left translations, induced by the group of inner automorphisms regarded as an infinitesimal automorphisms, we obtain a set of $2n$ equations (3.3) and (3.4) to be satisfied by the MEFRIR and by the vectors $\gamma_i{}^\mu$, $\gamma_\mu{}^i$, $l_I{}^\mu$, and $l^I{}_\mu$; through the latter a link between representations and differential geometry have been established. Moreover, also derived are Eqs. (3.13) and (3.14) to be satisfied by the MEFRIR which include the generalized Maurer-Cartan equations and the Lie structure formulas.

3. We have introduced the nonsymmetric (+) and (−) connections associated respectively with the right and left vector fields, and geometric properties of these connections have been detailed. The Rieman tensors of these nonsymmetric connections vanish identically. Each (+) or (−) connection, being non-Riemannian, has been divided into the symmetric and antisymmetric parts in Eq. (4.17); the symmetric part agrees with that of the (0) connection [Eq. (4.18)] and the antisymmetric part is related to the Riemann tensor [Eq. (4.28)] of the (0) connection; the group space is therefore an Einstein

and symmetric space. The differential geometry of the group manifold thus agrees with CS's. But our method makes the explicit calculations of these quantities possible in terms of the group parameters.

4. The theorem (5.1) has been proved which relates the right (or left) vectors to the metric tensors of R_n. With the help of this theorem, the second-order invariant equation of any semisimple Lie group has been cast in the covariant or Lie derivative forms of any connection in Eq. (5.5). For the higher-order invariant equations, we have taken $U(m)$ as an example and have displayed them in the covariant (or Lie derivative) forms in Eq. (5.9) [or (5.10)]. Moreover, examples of the invariant equations for $SU(2)$, $SO(3)$, and $SU(3)$ are given. For $SU(2)$, $SO(3)$ they are the Schrödinger equation of the symmetric top of ordinary quantum theory. For $SU(3)$ we have shown that after carrying out the quantizations of charge and hypercharge, the invariant equations become the eigenvalue equations [Eqs. (5.12) and (5.13)] expressed in terms of the eigenvalues of isospin and hypercharge. A new nonrelativistic wave equation based on the second-order invariant equations has also been proposed for $SU(3)$ multiplets.

5. The orthonormality and completeness properties as well as the transformation in R_n of the MEFRIR are briefly discussed in the Appendix.

APPENDIX: SOME PROPERTIES OF $\mathcal{D}^{[p]}_{\langle m/m'\rangle}(\alpha)$

The MEFRIR $\mathcal{D}^{[p]}_{\langle m|m'\rangle}(\alpha)$ which satisfies (3.3) and (3.4) is no longer a matrix but is a mixed tensor in R_n (and also in a subspace of R_n); its transformation property under a coordinate transformation $\alpha^\mu \to \alpha'^\mu$ is determined by the dimension of the partition label $[p]$. For example, in $SU(3)$ the three-dimensional representations $(p_1=1,\ p_2=0)$ and $(p_1=0,\ p_2=1)$ are the fundamental irreducible representations of a $SU(2)$ or $SO(3)$ subgroup; teh corresponding MEFRIR transforms according to a contravariant or a covariant vector of this subgroup. For the regular representation, it transforms like a contravariant vector in R_n with a prescribed state $\langle m|$ and like a covariant vector with another prescribed state $|m'\rangle$, whereas $\mathcal{D}^{*[p]}_{\langle m|m'\rangle}(\alpha)$ transforms like a covariant vector with the prescribed state $\langle m|$ and like a contravariant vector with $|m'\rangle$.

The $\mathcal{D}^{[p]}_{\langle m|m'\rangle}(\alpha)$ satisfies the generalized orthogonality relation

$$\int_{R_n} \mathcal{D}^{*[p]}_{\langle m|m'\rangle}(\alpha)\, \mathcal{D}^{[p_1]}_{\langle m_1|m_1'\rangle}(\alpha)\, \sqrt{g}\prod_{\mu=1}^{n} d\alpha^\mu = \frac{\delta_{[p],[p_1]}\,\delta_{\langle m|,\langle m_1|}}{d}\,\delta_{|m'\rangle,|m_1'\rangle}$$

where d is the dimension of the irreducible representation with the partition $[p]$ or $[p_1]$ and the integration is carried out over R_n. The $\delta_{|m\rangle,|m'\rangle}$ means $\delta_{m_{11},m'_{11}}\,\delta_{m_{12},m'_{12}},\ \delta_{m_{13},m'_{13}}\cdots$ and $\delta_{[p],[p']}=\delta_{p_1,p_1'},\ \delta_{p_2,p_2}\cdots\delta_{p_l,p_l'}$. Being a unitary matrix of dimension d, the $\mathcal{D}^{[p]}_{\langle m|m'\rangle}(\alpha)$ is a complex function of many periodic variables α^μ in R_n. The functional space carrying IR label $[p]$ is itself a vector space over the l invariant differential equations to be satisfied by $\mathcal{D}^{[p]}_{\langle m|m'\rangle}(\alpha)$.

[1]S. Lie, *Theorie der Transformationsgruppen* (Druck und Verlag Von B.G. Treubner, Leibzig, 1888).
[2]E. Cartan and J.A. Schouten, Acad. Van Wetens, Amsterdam, Proc. **29**, 53 (1926).
[3]J.A. Schouten, *Ricci Calculus* (Springer-Verlag, Berlin, 1954), Chap. IV.
[4]L.P. Eisenhart, *Continuous Group of Transformation* (Dover, New York, 1933).
[5]L.P. Eisenhart, *Non-Riemannian Geometry* (Am. Math. Soc. Colloquim Publications 8) (AMS, Providence, R.I., 1927).
[6]V. Bargmann, Rev. Mod. Phys. **34**, 829 (1962).
[7]M. Resnikopf, J. Math. Phys. **8**, 67, 79 (1967).
[8]S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic, New York, 1962).
[9]S. Helgason, "Lie Groups and Symmetric Spaces" in *Battelle Recontres*, edited by C.M. De Witt and J.A. Wheeler (Benjamin, New York, 1968).
[10]W. Miller, *Lie Theory and Special Functions* (Academic, New York, 1968).
[11]Refer to references of L.C. Biedenharn, *Group Theoretical Approach to Nuclear Spectroscopy* (University of Colorado, Boulder, Colorado, 1962).
[12]M. Beg and H. Ruegg, J. Math. Phys. **6**, 677 (1967).
[13]S. Heskia and S.A. Sofroniou, J. Math. Phys. **12**, 2280 (1971). They attempted to geometrize $U(m)$. Their approach, however, is quite different from ours and therefore from Cartan and Schouten's.
[14]F.D. Murnaghan, *The Unitary and Rotation Groups*, Lectures on Applied Mathematics (Spartan, Washington, D.C., 1962), Vol. 3.
[15]P. Pajas and R. Reczka, J. Math. Phys. **8**, 857 (1967).
[16]H.Y. Yeh, Ann. N.Y. Acad. Sci. **172**, 721 (1971).
[17]H.Y. Yeh, "Representation and Differential Geometry of the Groups $SU(2)$ and $SO(3)$" (to be published).
[18]L.K. Hua, *Harmonic Analysis of Functions of Several Variables in the Classical Domains* (AMS, Providence, R.I.), , p. 52. For $U(m)$ Hua has proved that the quadratic line element is independent of choices of representative matrices. Recently, an unsuccessful attempt (Ref. 12) was made to introduce a metric to $U(m)$.
[19]J.L. Synge and A. Schild, *Tensor Calculus* (University of Toronto Press, Toronto, 1949).
[20]L.D. Landau and E.M. Lifschitz, *The Classical Theory of Fields* (Pergamon, New York, 1962).
[21]E.P. Wigner, *Group Theory and Its Applications to the Quantum Mechanics of Atomic Spectra* (Academic, New York, 1959), Chaps. 10 and 14.
[22]The condition $g=0$ contradicts to the fact that all n variables are linearly independent.
[23]H. Boerner, *Representations of Groups* (North-Holland, Amsterdam, 1970), p. 43, Eq. (6.5). Boerner requires translation invariance on the group manifold. Such an assumption is not needed in our method.
[24]G.E. Baird and L.C. Biedenharn, J. Math. Phys. **4**, 1449 (1963).
[25]The $\mathcal{D}^{[p]}_{\langle m|m'\rangle}(\alpha)$ is no longer a matrix; its transformation properties in R_n are discussed in the Appendix. For $SU(3)$, the equations equivalent to (3.3) and (3.4) have been obtained in Ref. 16.
[26]I.M. Gelfand and M.L. Zetlin, Dokl. Nauk. USSR **71**, 825, 1017 (1950). Also, for $U(m)$ see Ref. 24 and for $O(m)$ see Ref. 27.
[27]M.K.F. Wong, J. Math. Phys. **8**, 1899 (1967).
[28]The Appendix of Ref. 16.
[29]Reference 17, Sec. IV.
[30]For $SO(3)$ see U. Fano and G. Racah, *Irreducible Tensor Sets* (Academic, New York, 1959), p. 144, Appendix E.
[31]Reference 3, p. 196, Eqs. (2.25) and (2.26).
[32]L. Maurer, Bayer. Akad. Wis., Berlin **18**, 117 (1888).
[33]Reference 17, Sec. V.
[34]K. Yano and S. Bochner, *Curvature and Betti Numbers* (Krans Reprint, New York, 1953), Sec. 8, Chap. II.
[35]L.P. Eisenhart, Proc. Natl. Acad. Sci. **11**, 246 (1925).
[36]L.P. Eisenhart, *Riemannian Geometry* (Princeton U.P., 1925), p. 114.

[37] I. M. Gelfand, Mat. Sb. **26**, 103 (1950).

[38] L. C. Biedenharn, J. Math. Phys. 4, 436 (1963).

[39] G. Racah, Atti. Acad. Nezl. Lincei, Rend., Classe. Sci. Fis. Mat. Nat. 8, 108 (1950).

[40] B. Gruber and L. O'Raifearteigh, J. Math. Phys. 5, 1796 (1964).

[41] The Casimir invariant corresponds to $2G^{ij}X_iX_j$.

[42] Reference 3, p. 220, Eq. (6.4). The $g_{i_1,i_2\ldots i_m}$ of Cartan and Schoulten differs from Biedenharn's $[i_1i_2\ldots i_m]$ by a constant factor of $1/2^m$.

[43] Here we regard $\nabla_\alpha\nabla_\beta\nabla_\gamma$ as operating on a scalar function ψ so that $\nabla_\gamma\psi$ is a vector. When ψ is a tensor, a proof can be given in a similar way. If one exchanges β and γ, Eq. (4.29) has to be taken into consideration.

[44] Reference 17, Sec. VI.

[45] L. C. Biedenharn, "Group Theory and Classification," CERN Report 65-41 (1965).

[46] These conditions (5.14) lead to a new uncertainty principle which asserts that it is impossible to measure simultaneously and with abritrary accuracy both the eigenvalue of I_z and its conjugate isospin angle and also both the eigenvalue of hypercharge and its conjugate hypercharge angle. The rules for the minimum uncertainties of measurements are

$$\langle\Delta I_z\rangle\langle\Delta\alpha_I\rangle F(\langle\Delta\alpha_I\rangle)\sim\hbar \quad\text{and}\quad \langle\Delta Y\rangle\langle\Delta\alpha_H\rangle F(\langle\Delta\alpha_H\rangle)\sim\hbar$$

where $F(\langle\Delta\alpha_I\rangle)$ and $F(\langle\Delta\alpha_H\rangle)$ are correction factors and $\langle F\rangle$ denotes the expectation value of F in R_8.

[47] H. Y. Yeh, Bull. Am. Phys. Soc. **15**, 50 (1970). For stationary states the equation reduces to Eq. (5.12).

[48] H. Y. Yeh, Bull. Am. Phys. Soc. **14**, 543 (1969). Baryon mass formulas can be obtained from Eq. (5.12).

On the stationary gravitational fields

S. Kloster, M. M. Som*, and A. Das†

Department of Mathematics, Simon Fraser University, Burnaby 2, British Columbia, Canada
(Received 5 September 1973; revised manuscript received 21 February 1974)

The stationary gravitational equations in vacuum are expressed in five different forms. A necessary integral condition on the twist potential ϕ is derived. The Papapetrou–Ehlers class of stationary solutions is rederived in a different way. In the study of the complex potential theory it is proved from the field equations that a body admitting an arbitrary symmetry must satisfy an integral condition analogous to the equilibrium criterion. It is proved that the vanishing of the scalar curvature of the associated space implies the flatness of the space–time metric. A proof is given for the fact that the only analytic functions of the complex potential F which preserve the field equations form a four-parameter Möbius group. It is also shown that any differentiable function of F and $\bar{F}$ which preserves the field equations must either be an analytic function of F or the conjugate of such a function. Next the conformastationary vacuum metrics are classified. In the study of the axially symmetric stationary fields a class of metrics (outside the Papapetrou–Ehlers class) is found depending on Euclidean harmonic functions.

1. INTRODUCTION

In recent years the stationary gravitational fields have drawn attention on account of the collapse theorems.[1] In this paper a detailed study of the stationary field equations is made. We adopt the convenient metric form

$$\Phi \equiv -\exp[-\omega(x)]\, g_{\alpha\beta}(x)\, dx^\alpha dx^\beta + \exp(\omega)[a_\alpha(x)\, dx^\alpha + dt]^2,$$

which has an associated space with metric tensor $g_{\alpha\beta}$.

In the second section the Riemann tensor components are exhibited. Then five different forms of the field equations are given. First the equations which are analogous to the magnetovac form[2] are written. Next the twist potential ϕ is introduced and the equations resembling the electrovac case are given. Then the complex potential[3] $F \equiv \exp(\omega) + i\phi$ is brought in and the corresponding field equations are obtained. Finally by using the triad formalism the invariant form of the field equations is provided. The present form, though mathematically equivalent to that of Perjes,[4] does not contain explicitly the conformal invariants ψ_0, ψ_1, etc.

In the third section the variational derivation of the field equations is given. The determinacy with regard to the number of unknown functions versus the number of independent field equations is sorted out. A necessary integral condition on the twist potential ϕ is obtained. The eigenvalues of the Ricci subtensor $R_{\alpha\beta}$ of the associated space are explored. One of the eigenvalues turns out to be zero and the other two are nonnegative. This result may have some implications on the existence of groups of motions in the space.[5]

In the fourth section we look into the class of stationary metrics such that ω and ϕ are functionally related. Papapetrou[6] (in the axially symmetric case) and Ehlers[7] (in a more general case) have already shown that the field equations boil down to the static form in such a situation. Following the analogy of the electrovac problem and Majumdar's[8] work, a different derivation of the Papapetrou–Ehlers results is given. Furthermore, the integral condition on ϕ in this case implies the vanishing of the total mass of the static sources. One way to get around this difficulty is to mathematically impose semi-infinite branch cuts or physically adjoin semi-infinite, massless, rotating tails to the finite bodies.

In Theorem 5.4 we investigate the class of C_ω transformations of F which preserve the field equations. (This problem was essentially suggested by Matzner and Misner.[9]) It is proved that the only analytic functions which preserve the equations are $f(F) = (AF + iB)/(iCF + D)$, where A, B, C, D are real constants satisfying $AD + BC > 0$. These transformations form a subgroup of the fractional linear group (Möbius). Geroch[10] gave this result without proof. Kinnersley[11] has found some transformations which preserve the stationary electrovac equations and which include the group mentioned above.

In the next section the class of conformastationary universes is investigated. These are the metrics where the associated space is conformally flat. It turns out that only three such metrics exist and all of these belong to the Papapetrou–Ehlers class.

In the study of the axially symmetric stationary case, a new class of solutions is obtained. In this class $a_1 = a_2 = a_2 = 0$, a_3 is functionally related to ω, and the equations boil down to the Newtonian potential equation.

2. NOTATIONS AND PRELIMINARIES

The metric of a stationary V_4 will be written in the following form:

$$\Phi = \gamma_{ij}\, dx^i dx^j = -\exp[-\omega(x)]\, g_{\alpha\beta}(x)\, dx^\alpha dx^\beta + \exp(\omega)[a_\alpha(x)\, dx^\alpha + dx^4]^2, \tag{2.1}$$

where the Latin indices range from 1 to 4, the Greek indices range from 1 to 3, x denotes (x_1, x_2, x_3), $a^\alpha = g^{\alpha\beta} a_\beta$, and the associated space V_3 with metric $g_{\alpha\beta}$ is positive definite.

The form of (2.1) is preserved under the following transformations:

$$x'^\alpha = f^\alpha(x), \quad t' = t + \lambda(x), \quad a'_\alpha = a_\alpha - \lambda_{,\alpha}, \tag{2.2}$$

where λ and the f^α's are arbitrary functions and partial derivatives are denoted by commas. We write $\frac{1}{2}(T^\alpha{}_{\beta\gamma\delta}$

$-T^{\alpha}{}_{\beta\delta\gamma}) = T^{\alpha}{}_{\beta[\gamma\delta]}$, $a_{\alpha,\beta} - a_{\beta,\alpha} = f_{\alpha\beta}$, and $g^{\alpha\beta}\omega_{,\alpha}\omega_{,\beta} = \Delta_1\omega$. The indices on $\omega_{,\alpha}$ and $f_{\alpha\beta}$ are raised and lowered by $g_{\alpha\beta}$. The covariant derivative with respect to $g_{\alpha\beta}$ is denoted by a slash. The Riemann tensors for γ_{ij} and $g_{\alpha\beta}$ are denoted ${}^{(4)}R^{i}{}_{jkl}$ and $R^{\alpha}{}_{\beta\gamma\delta}$, respectively, the Ricci tensors by ${}^{(4)}R_{ij}$ and $R_{\alpha\beta}$, and the curvature invariants by ${}^{(4)}R$ and R.

The components ${}^{(4)}R^{i}{}_{jkl}$ can be calculated from (2.1):

$$
\begin{aligned}
{}^{(4)}R^{\mu}{}_{44\nu} = {} & [\exp(2\omega)/4][-2\omega_{|}{}^{\mu}{}_{\nu} - 3\omega_{,}{}^{\mu}\omega_{,\nu} + \delta^{\mu}_{\nu}\Delta_1\omega] \\
& + [\exp(4\omega)/4][f^{\alpha}{}_{\nu}f^{\mu}{}_{\alpha}],
\end{aligned} \tag{2.3}
$$

$$
\begin{aligned}
{}^{(4)}R^{4}{}_{\beta\gamma 4} = {} & \tfrac{1}{4}[2\omega_{|\beta\gamma} + 3\omega_{,\beta}\omega_{,\gamma} - g_{\beta\gamma}\Delta_1\omega] + [\exp(2\omega)/4][-2a_{\beta}a^{\mu}\omega_{|\mu\gamma} \\
& + 2a^{\mu}f_{\mu\beta|\gamma} - 3a_{\beta}a_{\gamma}a^{\mu}\omega_{,\mu} + a_{\beta}a_{\gamma}\Delta_1\omega + 4a^{\mu}f_{\mu\beta}\omega_{,\gamma} \\
& + 2a^{\mu}f_{\mu\gamma}\omega_{,\beta} + a_{\gamma}f_{\beta\mu}\omega_{,}{}^{\mu} + g_{\beta\gamma}\omega_{,}{}^{\alpha}f_{\alpha\mu}a^{\mu} + 2f_{\gamma\beta}a^{\mu}\omega_{,\mu} \\
& + f_{\mu\gamma}f^{\mu}{}_{\beta}] + [\exp(4\omega)/4][a_{\beta}a^{\alpha}f_{\alpha\mu}f^{\mu}{}_{\gamma}],
\end{aligned} \tag{2.4}
$$

$$
\begin{aligned}
{}^{(4)}R^{\mu}{}_{\alpha 4\nu} = {} & [\exp(2\omega)/4][-2a_{\alpha}\omega_{|}{}^{\mu}{}_{\nu} + 2f^{\mu}{}_{\alpha|\nu} - 3a_{\alpha}\omega_{,\nu}\omega_{,}{}^{\mu} \\
& + \delta^{\mu}_{\nu}a_{\alpha}\Delta_1\omega + 4\omega_{,\nu}f^{\mu}{}_{\alpha} + 2\omega_{,\alpha}f^{\mu}{}_{\nu} + 2\omega_{,}{}^{\mu}f_{\nu\alpha} \\
& + g_{\alpha\nu}\omega_{,\beta}f^{\beta\mu} + \delta^{\mu}_{\nu}\omega_{,}{}^{\beta}f_{\alpha\beta}] + [\exp(4\omega)/4][a_{\alpha}f^{\beta}{}_{\nu}f^{\mu}{}_{\beta}],
\end{aligned} \tag{2.5}
$$

$$
\begin{aligned}
{}^{(4)}R^{\alpha}{}_{\beta\gamma\delta} = {} & R^{\alpha}{}_{\beta\gamma\delta} + \tfrac{1}{2}[2g_{\beta[\delta}\omega_{|\gamma]}{}^{\alpha} + 2\delta^{\alpha}_{[\gamma}\omega_{|\delta]\beta} + g_{\beta[\gamma}\delta^{\alpha}_{\delta]}\Delta_1\omega \\
& + \omega^{\alpha}g_{\beta[\delta}\omega_{,\gamma]} + \omega_{,\beta}\delta^{\alpha}_{[\gamma}\omega_{,\delta]}] + [\exp(2\omega)/2][2a_{\beta}a_{[\delta}\omega_{|\gamma]}{}^{\alpha} \\
& + 2a_{\beta}f^{\alpha}{}_{[\gamma|\delta]} + 2f^{\alpha}{}_{\beta|[\delta}a_{\gamma]} + \delta^{\alpha}_{[\delta}a_{\gamma]}a_{\beta}\Delta_1\omega \\
& + 3a_{\beta}\omega_{,}{}^{\alpha}a_{[\delta}\omega_{,\gamma]} + 4f_{\delta\gamma}a_{\beta}\omega_{,}{}^{\alpha} + 2a_{[\gamma}f_{\delta]\beta}\omega_{,}{}^{\alpha} \\
& + 2f^{\alpha}{}_{[\delta}a_{\gamma]}\omega_{,\beta} + 2a_{\beta}f^{\alpha}{}_{[\gamma}\omega_{,\delta]} + 4f^{\alpha}{}_{\beta}a_{[\gamma}\omega_{,\delta]} \\
& + g_{\beta[\gamma}a_{\delta]}\omega_{,}{}^{\mu}f^{\alpha}{}_{\mu} + a^{\beta}\omega_{,}{}^{\mu}f_{\mu[\delta}\delta^{\alpha}_{\gamma]} + \delta^{\alpha}_{[\gamma}a_{\delta]}\omega_{,}{}^{\mu}f_{\mu\beta} \\
& + f_{\beta[\gamma}f_{\delta]}{}^{\alpha} + 2f_{\gamma\delta}f^{\alpha}{}_{\beta}] \\
& + [\exp(4\omega)/2][a_{\beta}f^{\alpha}{}_{\mu}f^{\mu}{}_{[\delta}a_{\gamma]}].
\end{aligned} \tag{2.6}
$$

The field equations can be written as follows:

$$
\begin{aligned}
\sigma_{\mu\nu} & \equiv [\exp(-\omega)/2]g_{\mu\nu}{}^{(4)}R + {}^{(4)}R_{\mu\nu} - a_{\mu}{}^{(4)}R_{\nu 4} \\
& \quad - a_{\nu}{}^{(4)}R_{\mu 4} + a_{\mu}a_{\nu}{}^{(4)}R_{44} \\
& \equiv G_{\mu\nu} + \tfrac{1}{2}(\omega_{,\mu}\omega_{,\nu} - \tfrac{1}{2}g_{\mu\nu}\Delta_1\omega) \\
& \quad + [\exp(2\omega)/2](-f_{\mu}{}^{\alpha}f_{\nu\alpha} + \tfrac{1}{4}g_{\mu\nu}f^{\alpha\beta}f_{\alpha\beta}), \\
\rho & \equiv -2[\exp(-2\omega)]\,{}^{(4)}R_{44} \equiv \Delta_2\omega + [\exp(2\omega)/2]f^{\alpha\beta}f_{\alpha\beta} = 0, \\
D_{\mu} & \equiv 2a_{\mu}{}^{(4)}R_{44} - 2\,{}^{(4)}R_{\mu 4} \equiv [\exp(2\omega)f_{\mu}{}^{\nu}]_{|\nu} = 0, \\
E_{\mu\nu\lambda} & \equiv f_{\mu\nu|\lambda} + f_{\nu\lambda|\mu} + f_{\lambda\mu|\nu} = 0,
\end{aligned} \tag{F_1}
$$

where $\Delta_2\omega = g^{\alpha\beta}\omega_{|\alpha\beta}$ and $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}$. These equations are almost the static magnetovac form.[2]

Defining $\phi_{\alpha} \equiv \frac{1}{2}\exp(2\omega)\eta_{\alpha\beta\gamma}f^{\beta\gamma}$, where $\eta_{\alpha\beta\gamma} = \sqrt{g}\,\epsilon_{\alpha\beta\gamma}$, it can be shown from the equations $D_{\mu} = 0$ that $\phi_{\alpha} = \phi_{,\alpha}$ for some function ϕ (called the twist potential). Now the field equations become

$$
\begin{aligned}
\sigma_{\mu\nu} & \equiv G_{\mu\nu} + \tfrac{1}{2}(\omega_{,\mu}\omega_{,\nu} - \tfrac{1}{2}g_{\mu\nu}\Delta_1\omega) \\
& \quad + \tfrac{1}{2}\exp(-2\omega)(\phi_{,\mu}\phi_{,\nu} - \tfrac{1}{2}g_{\mu\nu}\Delta_1\phi) = 0, \\
\rho & \equiv \Delta_2\omega + \exp(-2\omega)\Delta_1\phi = 0, \\
\rho' & \equiv \Delta_2\phi - 2\Delta_1(\omega, \phi) = 0,
\end{aligned} \tag{F_2}
$$

where $\Delta_1(\omega, \phi) \equiv g^{\alpha\beta}\omega_{,\alpha}\phi_{,\beta}$. These equations are almost the static electrovac form.[8]

They can be written more compactly by using a complex potential[3] $F \equiv \exp(\omega) + i\phi$:

$$
\begin{aligned}
\tilde{\sigma}_{\mu\nu} & \equiv R_{\mu\nu} + \tfrac{1}{4}(\mathrm{Re}F)^{-2}[F_{,\mu}\bar{F}_{,\nu} + \bar{F}_{,\mu}F_{,\nu}] = 0, \\
\mu & \equiv \Delta_2 F - (\mathrm{Re}F)^{-1}\Delta_1 F = 0.
\end{aligned} \tag{F_3}
$$

Here the bar stands for complex conjugation.

Another form is obtained since the conformal curvature tensor vanishes in three dimensions:

$$
\begin{aligned}
\sigma_{\mu\nu\lambda\rho} \equiv {} & R_{\mu\nu\lambda\rho} - \tfrac{1}{2}(\mathrm{Re}F)^{-2}[g_{\mu[\lambda}(\bar{F}_{,\rho]}F_{,\nu} + F_{,\rho]}\bar{F}_{,\nu}) \\
& + g_{\nu[\rho}(\bar{F}_{,\lambda]}F_{,\mu} + F_{,\lambda]}\bar{F}_{,\mu}) \\
& + \tfrac{1}{2}\Delta_1(F, \bar{F})g_{\mu[\rho}g_{\lambda]\nu}] = 0, \quad \mu = 0.
\end{aligned} \tag{F_4}
$$

The field equations can also be put into an invariant form. Triad labels will be denoted by capital Roman indices which take the values 1, 2, 3, and the summation convention on these indices is also adopted. An orthonormal triad field $e_{A}{}^{\alpha}(x)$ in V_3 satisfies $e_{A}{}^{\alpha}e_{B\alpha} = \delta_{AB}$. Furthermore, with the following definitions,

$$
\begin{aligned}
T_{AB\cdots} & \equiv e_{A\alpha}e_{B\beta\cdots}T^{\alpha\beta\cdots}, \\
\gamma_{ABC} & \equiv e_{A\mu|\nu}e_{B}{}^{\mu}e_{C}{}^{\nu}, \\
T_{\cdots,A} & \equiv e_{A}{}^{\alpha}T_{\cdots,\alpha},
\end{aligned}
$$

the invariant form of (F_4) is the following:

$$
\begin{aligned}
\sigma_{MNLR} \equiv {} & 2[\gamma_{MN[L,R]} + \gamma_{MNA}\gamma_{A[LR]} + \gamma_{AM[R}\gamma_{|AN|L]}] \\
& - \tfrac{1}{2}(\mathrm{Re}F)^{-2}[\delta_{M[L}(\bar{F}_{,R]}F_{,N} + F_{,R]}\bar{F}_{,N}) \\
& + \delta_{N[R}(\bar{F}_{,L]}F_{,M} + F_{,L]}\bar{F}_{,M}) + \tfrac{1}{2}(F_{,A}\bar{F}_{,A})\delta_{M[R}\delta_{L]N}] = 0.
\end{aligned} \tag{F_5}
$$

Perjes[4] has also written invariant forms of the stationary equations. But our formulation, though mathematically equivalent to his, does not contain explicitly the conformal invariants ψ_0, ψ_1, etc.

3. DISCUSSION OF THE FIELD EQUATIONS

A. Lagrangian

The Lagrangian from (2.1) is

$$
\begin{aligned}
& \int_{D\times\{t|t_1<t<t_2\}} {}^{(4)}R\,d_4v \\
& \quad = -(t_2 - t_1)\int_D [R + \tfrac{1}{2}\Delta_1\omega - \tfrac{1}{4}\exp(2\omega)f^{\alpha\beta}f_{\alpha\beta} \\
& \qquad - \Delta_2\omega]\,d_3v.
\end{aligned} \tag{3.1}
$$

Note that $\int_D \Delta_2\omega\,d_3v$ may be converted to a surface interal and hence neglected. Thus the effective Lagrangian is

$$
\int_D [R + \tfrac{1}{2}\Delta_1\omega - \tfrac{1}{2}\exp(2\omega)f^{\alpha\beta}f_{\alpha\beta}]\,d_3v. \tag{3.2}
$$

This is invariant under (2.2) and can be used to derive the system (F_1).

B. Counting equations

The system (F_1) has ten unknown functions: six $g_{\alpha\beta}$'s, three a_{α}'s, and ω. There are 11 equations, but they are related by certain identities. These are

$$
\begin{aligned}
& \sigma^{\mu\nu}{}_{|\nu} - \tfrac{1}{2}\omega^{\mu}\rho - \tfrac{1}{2}\exp(-2\omega)\,\phi^{\mu}\rho' \equiv 0, \\
& \quad D^{\mu}_{|\mu} \equiv 0, \quad \text{and} \quad E_{\mu\nu\lambda} \equiv 0,
\end{aligned} \tag{3.3}
$$

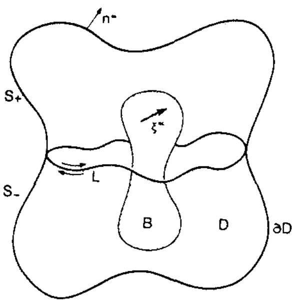


FIG. 1.

a total of five identities. The number of independent equations is reduced to six, so that, to make the system determinate, we are entitled to put three coordinate conditions on the $g_{\alpha\beta}$'s and one gauge condition on the a_α's.

C. Integral condition

We derive a necessary integral condition. We will say that a body is a region of V_3 where ρ, ρ' or one of the $\sigma_{\alpha\beta}$'s is nonzero. Suppose that B is a finite body and D is a region containing B and no other bodies. Suppose furthermore that ∂D is C_2^{ϕ}, orientable, closed, bounded, and simply connected, and L is a curve in ∂D which divides ∂D into two surfaces S_+ and S_- (see Fig. 1). Let n_α be the outward normal of D. Now $0 = [\oint_{L_+} + \oint_{L_-}] a_\alpha dx^\alpha$, so

$$\begin{aligned} 0 &= \int_{S_+\cup S_-} \eta^{\alpha\beta\gamma} f_{\beta\gamma} n_\alpha d_2S = 2\int_{\partial D} \exp(-2\omega)\phi|^\alpha n_\alpha d_2 S \\ &= 2\int_D [\exp(-2\omega)\phi|^\alpha]_{|\alpha} d_3 v = 2\int_D \rho' \exp(-2\omega)\, d_3 v \\ &= 2\int_B \rho' \exp(-2\omega)\, d_3 v. \end{aligned} \tag{3.4}$$

D. Eigenvalues of the Ricci subtensor $R_{\alpha\beta}$

From (F_3) we obtain

$$-\det[R^\alpha{}_\beta - \lambda\delta^\alpha_\beta] \equiv \lambda^3 + \tfrac{1}{2}\lambda^2(\mathrm{Re}F)^{-2}\Delta_1(F,\bar F) + \frac{\lambda}{16}(\mathrm{Re}F)^{-4}[(\Delta_1(F,\bar F))^2 - |\Delta_1 F|^2] = 0. \tag{3.5}$$

The eigenvalues are

$$\begin{aligned} \lambda_1 &= 0, \\ \lambda_2 &= -\tfrac{1}{4}(\mathrm{Re}F)^{-2}[\Delta_1(F,\bar F) + |\Delta_1 F|], \\ \lambda_3 &= -\tfrac{1}{4}(\mathrm{Re}F)^{-2}[\Delta_1(F,\bar F) - |\Delta_1 F|]. \end{aligned} \tag{3.6}$$

Now $\lambda_3 = 0$ iff $\Delta_1(F,\bar F) = |\Delta_1 F|$, and this is the case iff ReF and ImF are functionally related. $\lambda_2 = \lambda_3$ iff $\Delta_1 F = 0$. In this case F becomes a complex harmonic function in V_3. Note that λ_1, λ_2, and λ_3 are all nonpositive. This is clear for λ_1 and λ_2. Applying Schwartz's inequality, we have

$$0 \le \{\Delta_1[\exp(\omega)]\,\Delta_1\phi - (\Delta_1[\exp(\omega), \phi])^2\} \equiv 4(\mathrm{Re}F)^4\lambda_2\lambda_3. \tag{3.7}$$

Therefore λ_3 is nonpositive. This result may have some implications on the existence of groups of motion in V_3[5].

4. PAPAPETROU-EHLERS (P.E.) CLASS OF SOLUTIONS

We investigate (F_2) under the assumption that ω is a function of ϕ and $\omega'(\phi) \ne 0$. We find

$$(\omega')^{-1}\rho - \rho' \equiv (\omega')^{-1}[\omega'' + \exp(-2\omega) + 2(\omega')^2]\,\Delta_1\phi = 0. \tag{4.1}$$

Thus $\omega'' + \exp(-2\omega) + 2(\omega')^2 = 0$ since $\Delta_1\phi \ne 0$ in general. The general solution is $\exp(2\omega) = c + 2b\phi - \phi^2$, where b and c are arbitrary real constants such that $b^2 + c > 0$. In this case

$$\sigma_{\mu\nu} \equiv G_{\mu\nu} + \tfrac{1}{2}(b^2+c)(c + 2b\phi - \phi^2)^{-2}(\phi_{,\mu}\phi_{,\nu} - \tfrac{1}{2}g_{\mu\nu}\Delta_1\phi) = 0. \tag{4.2}$$

Defining $\chi(\phi) = \pm k\int(c + 2b\phi - \phi^2)^{-1}d\phi$, where $k = (b^2+c)^{1/2}$, we obtain

$$\phi = b \pm k\tanh\chi \tag{4.3}$$

and

$$\sigma_{\mu\nu} \equiv G_{\mu\nu} + \tfrac{1}{2}(\chi_{,\mu}\chi_{,\nu} - \tfrac{1}{2}g_{\mu\nu}\Delta_1\chi) = 0. \tag{4.4}$$

Equations (4.4) are just the static field equations for the metric

$$\Phi_0 = -\exp[-\chi(x)]g_{\alpha\beta}(x)\,dx^\alpha dx^\beta + \exp[\chi(x)](dx^4)^2; \tag{4.5}$$

therefore we have the Papapetrou–Ehlers theorem[6,7]: Given a static vacuum metric (4.5), a stationary vacuum metric

$$\Phi = -k^{-1}(\cosh\chi)g_{\alpha\beta}dx^\alpha dx^\beta + k(\mathrm{sech}\chi)(a_\alpha dx^\alpha + dx^4)^2 \tag{4.6}$$

can be generated provided one can solve

$$a_{\alpha,\beta} - a_{\beta,\alpha} = \pm k^{-1}\eta_{\alpha\beta\gamma}\chi|^\gamma. \tag{4.7}$$

Comments: (i) If the source for $\chi(x)$ is a finite body, then the integral condition (3.4) boils down to $\int_B \Delta_2\chi\, d_3 v = 0$. This means that the total mass of the static body must be zero. But any solution $\chi(x)$ due to a finite body can be made single-valued by interpreting it as due to that finite body joined with a semi-infinite, massless, rotating tail.[12] This construction would puncture any enclosing surface $S_+\cup S_-$, so that the above integral condition need not be satisfied.

(ii) The P.E. condition can be written in terms of the complex potential F as $|F - ib| = k$. This shows that both $\exp(\omega)$ and ϕ are bounded everywhere. However, ω itself need not be bounded; in fact, if it were, the space–time metric would be flat.

(iii) As an example of the P.E. class which is asymptotically flat and due to a finite source, we cite the following solution:

$$\begin{aligned} \Phi = &-k^{-1}\cosh(2m\cos\theta/r^2)[\exp(m^2\sin^2\theta\,(\sin^2\theta \\ &- 8\cos^2\theta)/2r^4)(dr^2 + r^2 d\theta^2) + r^2\sin^2\theta\, d\phi^2] \\ &+ k\,\mathrm{sech}(2m\cos\theta/r^2)[-(2m\sin^2\theta/kr)\,d\theta + dt]^2. \end{aligned} \tag{4.8}$$

This is generated by a dipole source.

(iv) The eigenvalue λ_3 of the Ricci subtensor $R^\alpha{}_\beta$ is zero iff the stationary metric is of the P.E. class.

(v) In the static electrovac case Majumdar[8] succeeded in reducing all the field equations to a single Laplace equation in Euclidean space E_3 by choosing the functional relation $g_{44}=(b+A_4)^2$. The analog of that case here would be to choose $k=0$ so that $\exp(2\omega)=-(b-\phi)^2$. But this choice makes $\exp(\omega)$ imaginary, and thus is not allowable. However, it may be mentioned that in a positive-definite V_4 the metric form

$$\Phi = U(x)((dx^1)^2+(dx^2)^2+(dx^2)^2)+U^{-1}(a_\alpha(x)\,dx^\alpha+dt)^2 \tag{4.9}$$

satisfies ${}^{(4)}R_{ij}=0$, where $\mathbf{curl}\,\mathbf{a}=\mathbf{grad}\,U$ and $\nabla^2U=0$ in E_3.

5. POTENTIAL THEORY

The action integral for the field equations in the form (F_3) is the Dirichlet type integral

$$\int_D (R+\tfrac{1}{2}(\mathrm{Re}F)^{-2}\Delta_1(F,\bar F))\,d_3v. \tag{5.1}$$

This does not follow from (3.1) because some of the field equations have been used. The boundary conditions which go with the variational principle are

$$\{[g^{\alpha\beta}\delta\{\alpha^\gamma\gamma\}-g^{\alpha\gamma}\delta\{\alpha^\beta\gamma\}+(\mathrm{Re}F)^{-2}\,\mathrm{Re}(g^{\alpha\beta}\bar F_{,\alpha}\delta F)]\,n_\beta\}_{\partial D}=0. \tag{5.2}$$

For a unique solution of any boundary value problem for the stationary field equations, the above conditions must be fulfilled. An example of such boundary conditions is $\{\alpha^\beta\gamma\}_{\partial D}=0$ and $[F_{,\alpha}n^\alpha]_{\partial D}=0$. But any boundary value problem in general relativity is different from the usual boundary value problems in applied mathematics in one respect. Though the boundary ∂D can be prescribed analytically, it is geometrically unknown until the boundary value problem itself is solved.

From the equation $\rho=0$ in (F_2) it can be concluded that ω is superharmonic. But from $\rho'=0$, no conclusion about ϕ can be drawn. By Hopf's theorem,[13] the regularity of ω throughout the whole V_3 implies that ω is constant. This in turn implies $\Delta_1\phi=0$, so that ϕ is constant and V_4 is flat. This is the content of the Einstein–Pauli–Lichnerowicz theorem.[14,15]

For a finite body generating a stationary field which permits some group of motions, the following result is true.

Theorem 5.1: Let the interior of a regular body B be simply connected and have a piecewise smooth, orientable boundary ∂B. Suppose there exists a Killing vector ξ^α in B. Let n^β be the outward unit normal to ∂B and let

$$\sigma=(\mathrm{Re}F)^{-2}[\Delta_2F-(\mathrm{Re}F)^{-1}\Delta_1F]. \tag{5.3}$$

If $\sigma_{\alpha\beta}n^\beta$ is continuous across ∂B, then

$$\mathrm{Re}\int_B\bar\sigma F_{,\alpha}\xi^\alpha d_3v=0. \tag{5.4}$$

Proof: Note that $\sigma_{\alpha\beta}=0$ in the neighboring exterior points of the body (see Fig. 1). From the assumption of continuity, it follows that $\sigma_{\alpha\beta}n^\beta=0$ on ∂B. Applying the divergence theorem, we have

$$0=\int_{\partial B}\sigma_{\alpha\beta}\xi^\alpha n^\beta d_2S=\int_B(\sigma_\alpha{}^\beta\xi^\alpha)_{|\beta}d_3v$$

$$=\int_B\sigma_\alpha{}^\beta{}_{|\beta}\,\xi^\alpha d_3v=\tfrac{1}{2}\mathrm{Re}\int_B\bar\sigma F_{,\alpha}\xi^\alpha d_3v. \tag{5.5}$$

For the physical meaning of this integral condition, one can mention that in Euclidean space E_3 a Killing vector can be expressed as $\boldsymbol{\xi}=\mathbf{t}+(\boldsymbol{\omega}\times\mathbf{r})$, where $\mathbf{t}$ and $\boldsymbol{\omega}$ are arbitrary constant vectors. In this case the integral condition (5.4) becomes

$$\mathrm{Re}\int_B\bar\sigma\,(\mathbf{grad}\,F)\,d_3v=0 \tag{5.6}$$

and

$$\mathrm{Re}\int_B\bar\sigma\,(\mathbf{r}\times\mathbf{grad}\,F)\,d_3v=0; \tag{5.7}$$

i.e., the total force and the total torque are zero.

Theorem 5.2: Let V_3 be the associated space of a stationary metric and $D\subseteq V_3$. Let $T=\{t:\,-\infty<t<\infty\}$. If $R=0$ throughout D, then $D\times T$ is flat.

Proof: Using (F_3), we find

$$R=-\tfrac{1}{2}[\Delta_1\omega+\exp(-2\omega)\Delta_1\phi]. \tag{5.8}$$

Since $g^{\alpha\beta}$ is positive definite, $\Delta_1\omega=0$ and $\Delta_1\phi=0$. Hence ω and ϕ are constant. Thus all of the $f^{\beta\gamma}$'s vanish and by (2.3)–(2.6) and (F_4), all of the ${}^{(4)}R^i{}_{jkl}$'s vanish.

The converse is not true. This is shown by the following example:

$$\Phi=-(x+x^2+x^3)^{-2}[(x^1+x^2+x^3)^2((dx^1)^2+(dx^2)^2+(dx^3)^2)]$$
$$+(x^1+x^2+x^3)^2[\lambda_{,\alpha}dx^\alpha+dx^4]^2, \tag{5.9}$$

where $\lambda(x^1,x^2,x^3)$ is an arbitrary function. This gives a flat V_4, but $R=-6(x^1+x^2+x^3)^{-4}$, so that it is nonzero.

To see the physical meaning of Theorem 5.2, note that

$$R=-\tfrac{1}{2}(\mathrm{Re}F)^{-2}\Delta_1(F,\bar F). \tag{5.10}$$

Thus R is proportional to the modulus squared of the complex force, so when it is zero, there is no gravity, and V_4 is flat.

Theorem 5.3: With V_3 and D as in Theorem 2, if $[\exp(-\omega)]_{|\alpha}$ is a Killing vector in D, then $D\times T$ is flat.

Proof: If $[\exp(-\omega)]_{|\alpha}$ is a Killing vector, then $[\exp(-\omega)_{|\alpha\beta}=0$ and hence $\Delta_1\omega-\Delta_2\omega=0$. Using (F_2), we have

$$R=-\tfrac{1}{2}(\Delta_1\omega-\Delta_2\omega)=0. \tag{5.11}$$

By Theorem 5.2, $D\times T$ is flat.

A transformation $f(F)$ is said to preserve solutions of (F_3) if $f(F)$ is a solution of (F_3) whenever F is. A nonanalytic transformation $f(F,\bar F)$ is said to preserve solutions of (F_3) if $f(F,\bar F)$ is a solution of (F_3) whenever F is.

Theorem 5.4: Let the stationary field equations (F_3) be valid in a compact, regular domain D in the associated space V_3 and let $g_{\alpha\beta}(x)$, $F(x)$ belong to $C^3(D)$. Let f be a differentiable function which preserves solutions of (F_3).

(i) If, furthermore, f is analytic, then

$$f(F)=(AF+iB)/(iCF+D), \tag{5.12}$$

where A, B, C, and D are real constants such that $AD+BC>0$.

(ii) If f is conjugate analytic, then

$$f(F)=(A\bar{F}+iB)/(iC\bar{F}+D), \tag{5.13}$$

where A, B, C, and D are as in (i).

(iii) There does not exist any nonanalytic f which preserves (F_3).

Proof: (i) For f to preserve (F_3), the following equations are necessary and sufficient:

$$f'\bar{f}'/(f+\bar{f})^2=1/(F+\bar{F})^2, \tag{5.14}$$

$$(\ln f')'+2/(F+\bar{F})=2f'/(f+\bar{f}). \tag{5.15}$$

Integrating (5.14) with respect to $\bar{F}$ gives

$$f'/(f+\bar{f})=1/(F+\bar{F})+[\ln K(F)]', \tag{5.16}$$

where K is an analytic function of F. Integrating (5.16) with respect to F gives

$$f+\bar{f}=(F+\bar{F})K(F)\bar{G}(F), \tag{5.17}$$

where G is an analytic function of F. Neither K nor G can be identically zero, lest f be constant. Since $K\bar{G}$ is real, $G=aK$, for some real constant a.

To find $K(F)$, we differentiate (5.17) with respect to F and $\bar{F}$, obtaining

$$0=(K/K')+(\bar{K}/\bar{K}')+F+\bar{F}. \tag{5.18}$$

Hence $(K/K')=-F+ib$ and $K=q/(F-ib)$, for some real constant b and complex constant q.

Now (5.17) becomes

$$f+\bar{f}=aq\bar{q}(F+\bar{F})/(F-ib)(\bar{F}+ib)$$
$$=aq\bar{q}/(F-ib)+aq\bar{q}/(\bar{F}+ib); \tag{5.19}$$

thus f must be $(aq\bar{q}/(F-ib))+ic$, for some real constant c, and this can be written as (5.12), the condition that $AD+BC>0$ will ensure that Ref remains positive and ω remains real. Substitution shows that (5.12) satisfies (5.15) as well as (5.14). The transformations (5.12) form a subgroup of the Möbius group and are generated by the following transformations: (1) $F\to AF$, A a positive real (magnification), (2) $F\to F+iB$, B real (imaginary boost), and (3) $F\to 1/F$ (inversion).

Notice that in the static case $F=\exp(\omega)$, so that the inversion transformation yields Buchdahl's theorem.[16]

New solutions can be generated by (5.12). For example, let $f(F)=(F+i)/(iF+1)$ and $F=\exp(\chi)$ be a static solution. Then $f(F)=\operatorname{sech}\chi-i\tanh\chi$, which is the stationary solution belonging to the P.E. class.

(ii) The proof is similar to (i).

(iii) Under the transformation[3] $F=(\xi-1)/(\xi+1)$, the first of Eqs. (F_3) becomes

$$R_{\alpha\beta}+(\xi\bar{\xi}-1)^{-2}(\xi_{,\alpha}\bar{\xi}_{,\beta}+\bar{\xi}_{,\alpha}\xi_{,\beta})=0. \tag{F_3'}$$

Suppose $f(\xi,\bar{\xi})$ is a nonanalytic function of ξ and $\bar{\xi}$ which preserves solutions of (F_3'). Then we must have

$$\tau_{\alpha\beta}\equiv(|f_{,\xi}|^2+|f_{,\bar{\xi}}|^2)(\xi_{,\alpha}\bar{\xi}_{,\beta}+\bar{\xi}_{,\alpha}\xi_{,\beta})$$
$$+2f_{,\xi}\bar{f}_{,\xi}\xi_{,\alpha}\xi_{,\beta}+2\bar{f}_{,\bar{\xi}}f_{,\bar{\xi}}\bar{\xi}_{,\alpha}\bar{\xi}_{,\alpha}$$
$$-(f\bar{f}-1)^2(\xi\bar{\xi}-1)^{-2}(\xi_{,\alpha}\bar{\xi}_{,\beta}+\bar{\xi}_{,\alpha}\xi_{,\beta})=0.$$

for all solutions ξ. In particular, these equations hold for the Kerr[17,3] solution $\xi=px^1+iqx^2$. The τ_{12} equation gives $f_{,\xi}\bar{f}_{,\xi}-\bar{f}_{,\bar{\xi}}f_{,\bar{\xi}}=0$, while the difference of the τ_{11} and τ_{22} equations gives $f_{,\xi}\bar{f}_{,\xi}+\bar{f}_{,\bar{\xi}}f_{,\bar{\xi}}=0$. Thus $f_{,\bar{\xi}}=0$ or $f_{,\xi}=0$, which means that f is either an analytic function of ξ or else the conjugate of such a function. These cases are covered by parts (i) and (ii).

6. CONFORMASTATIONARY SOLUTIONS

A *conformastationary* metric is one of the form

$$\Phi=-\exp(-\omega)g_{\alpha\beta}dx^\alpha dx^\beta+\exp(\omega)(a_\alpha dx^\alpha+dt)^2 \tag{6.1}$$

for which the associated metric $g_{\alpha\beta}dx^\alpha dx^\beta$ is conformally flat. (For an electrovac example see Israel and Wilson[18]).

Theorem 6.1: If a nonflat conformastationary metric form which is not static satisfies Einstein's equations ${}^{(4)}R_{ij}=0$, then it must be reducible to one of the following forms:

(i) (NUT solution)

$$\Phi=-(2k)^{-1}[(1-2m/r)+(1-2m/r)^{-1}][dr^2+(1-2m/r)$$
$$\times(d\theta^2+r^2\sin^2\theta\,d\phi^2)]+2k[(1-2m/r)+(1-2m/r)^{-1}]^{-1}$$
$$\times[(2m/k)\cos\theta\,d\phi+dt]^2, \tag{6.2}$$

(ii)
$$\Phi=-(2k)^{-1}[1+(1-mx^1)^4][(dx^1)^2+(dx^2)^2+(dx^3)^2]$$
$$+2k[(1-mx^1)^2+(1-mx^1)^{-2}]^{-1}$$
$$\times[(m/k)(x^3dx^2-x^2dx^3)+dt]^2, \tag{6.3}$$

(iii)
$$\Phi=-(2k)^{-1}[(1+m/2R)^4+(1-m/2R)^4]$$
$$\times[dR^2+R^2(d\psi^2+\sinh^2\psi\,d\phi^2)]+2k(1-m^2/4R^2)^2$$
$$\times[(1+m/2R)^4+(1-m/2R)^4]^{-1}$$
$$\times[(2m/k)\cosh\psi\,d\phi+dt]^2, \tag{6.4}$$

where m and k are constants.

Proof: First consider those solutions F for which ReF and ImF are functionally related (the P.E. class). From the P.E. theorem, it is obvious that such a solution must be generated from a static solution which has a conformally flat associated space. This is called a conformastat metric, and it has been shown[19] that only three such solutions for Einstein's equations exist. From these three static solutions via the P.E. theorem the solutions (6.2)–(6.4) follow.

Now suppose ReF and ImF are not functionally related. Then $x^1=\mathrm{Re}F$ and $x^2=\mathrm{Im}F$ are allowable coordinate conditions. As the third condition we take $g^{11}+g^{22}=2(x^1)^2$. From (F_3) it follows that $R=-1$. Thus Schouten's[20] condition for conformal flatness becomes $R_{\alpha\beta|\gamma}-R_{\alpha\gamma|\beta}=0$. This yields nine equations, six of which are $\{3^1\alpha\}=0$ and $\{3^2\alpha\}=0$, $\alpha=1, 2, 3$. The calculation of $R^1{}_{313}$ and $R^1{}_{323}$ from the definition shows these are both zero. But the conformal flatness condition in V_3 implies $R^1{}_{313}=\frac{1}{2}(g^{11}-(x^1)^2)g_{33}(x^1)^{-2}$ and $R^1{}_{323}=\frac{1}{2}g^{12}g_{33}(x^1)^{-2}$. Hence $g^{11}=(x^1)^2$, $g^{12}=0$, and $g^{22}=(x^1)^2$. Using conformal flatness, $R^1{}_{212}=\frac{1}{2}(x^1)^{-2}$. By using the definition, this is found to be $(x^1)^{-2}$ (after a long calculation). This contradiction means there cannot be any solutions outside of the P.E. class.

7. AXIALLY SYMMETRIC STATIONARY FIELDS

We now consider metrics of the following form:

$$\Phi = -\exp[-\omega(\rho,z)]\{\exp[2\nu(\rho,z)](d\rho^2+dz^2)+\rho^2 dz^2\} + \exp(\omega)[a(\rho,z)\,d\theta+dt]^2. \tag{7.1}$$

This will be called the Weyl—Lewis[21]—Papapetrou[6] (W. L. P.) coordinate system. The field equations reduce to

$$\nu_{,\rho}=\tfrac{1}{4}\rho[(\omega_{,\rho})^2-(\omega_{,z})^2+\exp(-2\omega)((\phi_{,\rho})^2-(\phi_{,z})^2)], \tag{7.2}$$

$$\nu_{,z}=\tfrac{1}{2}\rho[\omega_{,\rho}\omega_{,z}+\exp(-2\omega)\phi_{,\rho}\phi_{,z}], \tag{7.3}$$

$$\nabla^2\omega+\exp(-2\omega)[(\phi_{,\rho})^2+(\phi_{,z})^2]=0, \tag{7.4}$$

$$\nabla^2\phi-2[\phi_{,\rho}\omega_{,\rho}+\phi_{,z}\omega_{,z}]=0, \tag{7.5}$$

where

$$\nabla^2\equiv\frac{\partial^2}{\partial\rho^2}+\frac{1}{\rho}\frac{\partial}{\partial\rho}+\frac{\partial^2}{\partial z^2}. \tag{7.6}$$

The two real equations (7.4) and (7.5) can be combined into the complex equation

$$\nabla^2 F=(\mathrm{Re}F)^{-1}((F_{,\rho})^2+(F_{,z})^2). \tag{7.7}$$

The known solutions of (7.7) are those of Van Stockum,[22] Kerr,[17] Tomimatsu and Sato,[23] and Papapetrou[6] and Ehlers[7]. The following solutions may be mentioned:

(i) In the case $F=V(\rho)+i\phi(z)$, we must have $\phi=ez+d$ and

$$V=b^{-1}\rho\sinh(eb\ln\rho+c) \tag{7.8}$$

or

$$V=b^{-1}\rho\sin(eb\ln\rho+c), \tag{7.9}$$

where b, c, d, and e are real constants (cf. Van Stockum[22]).

(ii) If the eigenvalues λ_2 and λ_3 are equal, then $\Delta_1 F=0$. The space—time turns out to be flat in this case.

(iii) When the equations $f_{\alpha\beta}=\exp(-2\omega)\eta_{\alpha\beta\gamma}\phi^{|\gamma}$ are written in W. L. P. coordinates, we find

$$a_{,\rho}=\rho\exp(-2\omega)\phi_{,z} \tag{7.10}$$

and

$$a_{,z}=-\rho\exp(-2\omega)\phi_{,\rho}. \tag{7.11}$$

These equations are the pseudo-Cauchy-Riemann equations, and they show that a and ϕ are conjugate pseudoharmonic functions.[24] Thus $\Omega\equiv\phi+ia$ is a pseudo-analytic function of the second kind and

$$f\equiv\phi_{,\rho}-i\phi_{,z}=\rho^{-1}\exp(2\omega)(a_{,z}+ia_{,\rho}) \tag{7.12}$$

is a pseudo-analytic function of the first kind. Moreover, if $\zeta\equiv\rho+iz$, then

$$g\equiv f+\frac{1}{\pi}\int\frac{(\omega_{,\bar\zeta'}-1/4\rho)f+(\omega_{,\zeta'}-1/4\rho)\bar f}{\zeta'-\zeta}\,d\rho\,dz \tag{7.13}$$

is an analytic function of ζ.

From (7.10) and (7.11) it follows that a and ϕ are functionally independent. Thus we shall attempt to find a class of solutions where ω and a are functionally related and this class must be outside the P. E. class. To do that, we introduce $U=\omega-\ln\rho$. Then the field equations become

$$\nu_{,\rho}=\tfrac{1}{4}\rho[(U_{,\rho})^2-(U_{,z})^2+(2/\rho)U_{,\rho}+1/\rho^2 + \exp(2U)((a_{,z})^2-(a_{,\rho})^2)], \tag{7.14}$$

$$\nu_{,z}=\tfrac{1}{2}\rho[U_{,\rho}U_{,z}+U_{,z}/\rho+\exp(2U)a_{,\rho}a_{,z}], \tag{7.15}$$

$$\nabla^2U+\exp(2U)[(a_{,\rho})^2+(a_{,z})^2]=0, \tag{7.16}$$

$$\nabla^2 a+2[U_{,\rho}a_{,\rho}+U_{,z}a_{,z}]=0. \tag{7.17}$$

If U and a are functionally related, the only possibility is

$$\exp(-2U)=c_1+2c_2a+a^2, \tag{7.18}$$

where c_1 and c_2 are real constants. Equation (7.16) becomes

$$\nabla^2 a-2(c_2+a)(c_1+2c_2a+a^2)^{-1}[(a_{,\rho})^2+(a_{,z})^2]=0. \tag{7.19}$$

Let $b\equiv(c_1-c_2^2)^{1/2}$ and

$$W(a)\equiv-\int^a(c_1+2c_2a+a^2)^{-1}da=b^{-1}\cot^{-1}[(a+c_2)/b]. \tag{7.20}$$

Now equation (7.19) becomes $\nabla^2W=0$ and (7.1) becomes

$$\Phi=-b\rho^{-1}\sec bW[\exp(2\nu)(d\rho^2+dz^2)+\rho^2d\theta^2] + \rho b^{-1}\cos bW[((b\tan bW)-c_2)\,d\theta+dt]^2, \tag{7.21}$$

$$\nu=\int(\rho/2)[(2b^2(\cot^2bW)+1)\,W_{,\rho}W_{,z}+(1/\rho)(b\cot bW)\,W_{,z}]dz + (\rho/4)[((W_{,z})^2-(W_{,\rho})^2)+(2/\rho)(b\cot bW)\,W_{,\rho} + (1/\rho^2)]\,d\rho.$$

Before concluding we would like to discuss Ernst's potential equation.[3] This can be obtained from (7.7) by putting $F=(\xi-1)/(\xi+1)$ and using general coordinates in E_3, and it is

$$\nabla^2\xi=[2\bar\xi/(|\xi|^2-1)](\mathbf{grad}\,\xi)\cdot(\mathbf{grad}\,\xi). \tag{7.22}$$

In the Cartesian coordinates a set of solutions of (7.22) depending on two variables x,y is generated by arbitrary analytic or conjugate analytic functions of the complex variable $x+iy$. For the three-dimensional analog we can mention that any analytic or conjugate analytic function of the complex variable $x+i(\cos c)y+i(\sin c)z$, (where x,y,z are Cartesian coordinates, c is a real number) solves (7.22). In all these solutions $\mathrm{Re}\xi$, $\mathrm{Im}\xi$ are functionally independent. Moreover, for a pole of a prescribed order at infinity, an infinite subset of these solutions exists. However, none of these solutions apply directly to relativity. For that purpose only the axially symmetric solutions are permitted. Furthermore, the solutions of physical interest are those for which $\mathrm{Re}\xi$ and $\mathrm{Im}\xi$ are functionally independent and which generate asymptotically flat metrics. Even in such cases, in view of the present solutions of (7.22) just mentioned, it is reasonable to conjecture that an infinite set of solutions of Ernst's equation exists.

ACKNOWLEDGMENTS

The authors are indebted to the NRC Grant No. 3993. One of them (A.D.) would like to thank Professor A.H. Taub (Berkeley), Professor W. Israel (Edmonton), and Dr. A.S. Farooqui for the benefit of informal discussions.

*Present address: Department of Physics, City College, Calcutta-9, India.

†On sabbatical leave to Department of Mathematics, University of Washington, Seattle, Washington, 98195.

[1]W. Israel, Gen. Rel. Grav. **2**, 53 (1971).
[2]W. B. Bonnor, Proc. Phys. Soc. A **67**, 225 (1954).
[3]F. J. Ernst, Phys. Rev. **167**, 1175 (1968).
[4]Z. Perjes, J. Math. Phys. **11**, 3383 (1970).
[5]H. Guggenheimer, *Differential Geometry* (McGraw-Hill, New York, 1963), p. 351.
[6]A. Papapetrou, Ann. Physik **12**, 309 (1953).
[7]J. Ehlers, *Les théories relativistes de la gravitation* (Gauthier-Villars, Paris, 1952), p. 275.
[8]S. D. Majumdar, Phys. Rev. **72**, 390 (1947); Proc. Phys. Soc. 79, 657 (1962).
[9]R. A. Matzner and C. W. Misner, Phys. Rev. **154**, 1229 (1967).
[10]R. Geroch, J. Math. Phys. **12**, 918 (1971).
[11]W. Kinnersley, J. Math. Phys. **14**, 651 (1973).
[12]W. B. Bonner, Proc. Camb. Phil. Soc. **66**, 145 (1969).
[13]K. Yano and S. Bochner, *Curvature and Betti Numbers* Annals of Math. Studies No. 32 (Princeton U. P. Princeton, N. J., 1953), p. 26.
[14]A. Einstein and W. Pauli, Ann. Math. **44**, 131 (1943).
[15]A. Lichenerowicz, *Théories relativistes de la gravitation et de l'électromagnétisme* (Masson, Paris, 1955), p. 138.
[16]H. A. Buchdahl, J. Math. Oxford Ser. 5, 116 (1954).
[17]R. P. Kerr, Phys. Rev. Lett. **11**, 237 (1963).
[18]W. Israel and G. A. Wilson, J. Math. Phys. 13, 865 (1972).
[19]A. Das, J. Math. Phys. 12, 1136 (1971).
[20]J. A. Schouten, Math. Z. **11**, 58 (1921).
[21]T. Lewis, Proc. Roy. Soc. (London) A **136**, 176 (1932).
[22]W. J. Van Stockum, Proc. Roy. Soc. (Edinburgh) 57, 135 (1937).
[23]A. Tomimatsu and H. Sato, Phys. Rev. Lett. **29**, 1344 (1972).
[24]R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1966), Vol. II, p. 375.

Complexification of the algebraically special gravitational fields

Robert W. Lind*

Department of Physics, Temple University, Philadelphia, Pennsylvania 19122

Department of Physics, Florida State University, Tallahassee, Florida 32306†

Ezra T. Newman‡

Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania 15213

(Received 28 January 1974)

The fact that the Maxwell equations can be analytically extended into complex Minkowski space is used to show that a class of solutions to the real Maxwell equations exists which can be viewed as arising from a monopole source moving along a *complex* world line in the complex Minkowski space. This class of solutions is the natural analog of the class of regular, algebraically special type II, twisting metrics in Einstein's general theory of relativity, in that the two cases are characterized geometrically by the fact that the Maxwell and Weyl tensors, respectively, both possess a shear-free, diverging, geodesic principal null vector field l, which is twisting. By analytically extending the algebraically special metrics into a complex manifold, we show that the analogy runs even deeper than this. Aside from the constants, charge and mass, the solutions in both cases are completely determined by a single complex function ϕ. In both analytically extended manifolds the surfaces, $\phi=$const, are *complex* null surfaces and the complexified versions of both the Maxwell and Weyl tensors now have a *nontwisting* principal null vector field l^*, equal to the gradient of ϕ. We introduce the natural coordinate and tetrad systems associated with l and l^* and show the relationship between them in both the flat and curved complex manifolds. The class of solutions to the Maxwell equations is solved in both systems. The algebraically special metrics are treated in detail, and the Kerr metric is given as an explicit example.

1. INTRODUCTION

We have recently shown[1] that the Lienard–Wiechert (LW) solutions of the Maxwell equations have a precise analogy in the class of solutions to the vacuum Einstein equations known as the Robinson–Trautman (RT) metrics.[2] It is one of the purposes of this paper to generalize this analogy.

From the fact that the Maxwell equations can be analytically extended into complex Minkowski space,[3] we will show that a class of solutions to the ordinary (that is, real) Maxwell equations exists which can be viewed as arising from a monopole source moving along a complex world line in the complex Minkowski space. These solutions [hereafter to be referred to as complexified Lienard–Wiechert (CLW) solutions] can be geometrically characterized by the fact that they possess a principal null vector field (p.n.v.f.) of the Maxwell tensor with the following properties:

(1) the p.n.v.f. is the tangent field to a congruence of null geodesics;
(2) the p.n.v.f. has nonvanishing divergence,
(3) the shear of the p.n.v.f. vanishes.

If the further condition is satisfied that

(4) the twist (or curl) of the p.n.v.f. vanishes, then the Maxwell field is LW.

The solutions to the vacuum Einstein equations which are analogous to the CLW Maxwell fields [in the sense that a p.n.v.f. of the Weyl tensor satisfies conditions (1)–(3)] are the regular, algebraically special type II, twisting metrics.[4] (As R. Kerr was the first to systematically study these metrics, we will refer to them as Kerr-type metrics; *the Kerr metric* is a special case of the Kerr-type metric.)

Once it is shown that CLW Maxwell fields exist, the analogy [via conditions (1)–(3)] is then displayed. This, however, only superficially touches the full analogy. In both the Maxwell and gravitational cases the solutions are completely determined (aside from the constants, charge and mass) by a single complex function ϕ. The gradient of $\phi(l^*_\mu \equiv \phi_{,\mu})$ is a *complex* null vector field which is a p.n.v.f., not of the Maxwell or Weyl tensors, but of their complexified versions, i.e., of $F_{\alpha\beta} + iF^*_{\alpha\beta}$ and $C_{\alpha\beta\gamma\delta} + iC^*_{\alpha\beta\gamma\delta}$, the asterisk indicating dual.

If the real manifolds are analytically extended, in one case to complex Minkowski space and in the other to a complexified Kerr-type manifold, then l^*_μ not only satisfies conditions (1)–(3) but condition (4) as well; the surfaces $\phi=$ const being complex null surfaces. Complex null coordinates can be introduced by using ϕ to label the (complex) null surfaces, r^* as the (complex) affine parameter along l^* and ζ^*, η^* as two complex angles, constant along each l^* ray. In this new coordinate (and associated tetrad) system the Maxwell and Weyl tensors have respectively poles of the form r^{*-2} and r^{*-3}, so that one can (in some sense) view the solutions as being similar to the LW fields or the RT metrics but with the pole (i.e., $r^*=0$) tracing out a complex, rather than real, world line.

In Sec. 2 we describe and show the relationship between two null vector fields, l^* and l (in complex Minkowski space), which are defined from an arbitrary complex world line and which satisfy conditions (1)–(4) and (1)–(3), respectively. In addition, natural coordinate and tetrad systems associated with l^* and l are introduced. In Sec. 3 we find the solutions to the (real) Maxwell equations that have l as a p.n.v.f. and then show that they can be analytically extended into complex Minkowski space where they can be viewed as CLW fields. Section 4 is devoted to reviewing the Kerr-type metrics and in Sec. 5 we show how these can also be extended into a complex manifold with properties analogous to the CLW fields. In an appendix we show how many of the ideas developed here can be understood from the point of view of spinors and the Penrose theory of twistors.

2. TWISTING, SHEAR-FREE CONGRUENCES IN MINKOWSKI SPACE

We begin with a brief description of some properties of complex Minkowski space. It is a four-dimensional complex manifold (8 real dimensions) that can be covered by a single complex chart, with coordinates $z^\mu = x^\mu + iy^\mu$ (x^μ, y^μ real) and endowed with a complex line element,

$$ds^2 = \eta_{\mu\nu} dz^\mu dz^\nu, \quad \eta_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1). \tag{2.1}$$

The group of isometries is the ten (complex) parameter Poincaré group

$$z'^\mu = a^\mu_\nu z^\nu + b^\mu \tag{2.2}$$

with b^μ being a complex vector and a^μ_ν being a complex matrix satisfying

$$a^\mu_\alpha a^\nu_\beta \eta_{\mu\nu} = \eta_{\alpha\beta}. \tag{2.3}$$

(Although there is no intrinsic method of singling it out, we will always consider the real four-dimensional subspace $y^\mu = 0$ as the physical or real Minkowski space. This effectively reduces the group of isometries to the real Poincaré group.)

A complex world line (which will play a fundamental role here) is defined by the four analytic functions of a single complex variable ϕ

$$z^\mu = \xi^\mu(\phi). \tag{2.4}$$

If ϕ is chosen to be the complex proper length (times $\sqrt{2}/2$) along the world line, then (assuming the path is not null)

$$\eta_{\mu\nu} \xi'^\mu \xi'^\nu = 2, \quad \xi'^\mu = d\xi^\mu/d\phi. \tag{2.5}$$

(It should be realized that the complex world line is not really a line but a two-dimensional surface.)

We now introduce two different, but closely related, types of analytic coordinate systems in the complex space (analytic in the sense of the new coordinates being analytic functions of z^μ), which are associated with an arbitrary complex time-like world line. The first system uses complex null coordinates ϕ, r^*, ζ^*, and η^* based on the complex line (analogous to real null coordinates based on a real line) and introduced by

$$z^\mu = \xi^\mu(\phi) + r^* l^{*\mu}(\phi, \zeta^*, \eta^*), \; l^* = \hat{l}^*/V^* \tag{2.6}$$

with

$$\hat{l}^*_\mu = (\sqrt{2}/4P^*_0)(1 + \zeta^*\eta^*, \zeta^* + \eta^*, ((\zeta^* - \eta^*)/i), -1 + \zeta^*\eta^*),$$
$$P^*_0 = \tfrac{1}{2}(1 + \zeta^*\eta^*), \tag{2.7}$$

and

$$V^*(\phi, \zeta^*, \eta^*) = \hat{l}^*_\mu \xi'^\mu. \tag{2.8}$$

l^* is a complex null vector (i.e., $l^*_\mu l^{*\mu} = 0$) which sweeps out the complex light cone as ζ^* and η^* move over their respective extended complex planes. ϕ = const defines a complex cone, while ζ^* and η^* constant singles out a particular generator of the cone and r^* is the (normalized) complex affine length (along each generator) from the world line $\xi^\mu(\phi)$. We remark that "real" values of these coordinates (i.e., ϕ and r^* real, $\eta^* = \bar{\zeta}^*$) do not correspond to real Minkowski space. In the new coordinates the line element (2.1) becomes

$$ds^2 = 2\left(1 - \frac{V'^*}{V^*} r^*\right) d\phi^2 - 2d\phi dr^* - \frac{r^{*2}}{2V^{*2}} \frac{d\zeta^* d\eta^*}{P_0^{*2}} \tag{2.9}$$

and l^* becomes

$$l^* = \frac{\partial}{\partial r^*} \quad \text{or} \quad l^* = d\phi. \tag{2.10a}$$

It is convenient to introduce a complete null tetrad m^*, $\overline{m}^*$, and n^* in addition to l^* by

$$m^* = -\frac{d\eta^*}{2P^*_0 V^* \rho^*}, \tag{2.10b}$$

$$\overline{m}^* = -\frac{d\zeta^*}{2P^*_0 V^* \rho^*}, \tag{2.10c}$$

$$n^* = dr^* + \left(1 - \frac{V'^* r^*}{V^*}\right) d\phi, \tag{2.10d}$$

where $\rho^* = -1/r^*$.

Considered as a field on complex Minkowski space, l^* satisfies conditions (1)—(4). l^* does not, however, have immediate geometrical or physical significance for us because on the real Minkowski space ($y^\mu = 0$) it is a complex null field. It will nevertheless be possible to find a real null vector field l (on the real Minkowski space) which satisfies conditions (1)—(3), by means of a "projection" on l^*. (See Appendix A for a description of this in terms of spinors.)

Perhaps the easiest way to understand the relationship of l to l^* is to start with a description of the second coordinate system.

We begin by first describing a particular parametrization of an arbitrary null geodesic congruence, C. Take a *time-like* geodesic with tangent vector t^μ and the family of null cones emanating from it. The generators of these cones form a special null geodesic congruence S (not C) and are labeled by the three parameters u, ζ, and $\bar{\zeta}$ where u is ($\sqrt{2}/2$ times) the proper time at the apex of the cone and ζ and $\bar{\zeta}$ are "angular" variables (stereographic coordinates on the sphere) labeling the direction of each generator. To each geodesic, s, of S there is associated a unique null hyperplane in which s lies. The hyperplane can thus be labeled in the same manner as its associated geodesic, i.e., by u, ζ, and $\bar{\zeta}$.

An arbitrary geodesic c from C will lie in one of these hyperplanes. Its position, relative to the geodesic, s, of S can be described by a connecting vector η_μ lying in the hypersurface (i.e., with $\eta^\mu l_\mu = 0$, where l_μ is the tangent vector to s) and made unique by requiring $\eta^\mu t_\mu = 0$.

The parametric form of c can then be written as

$$x^\mu = \sqrt{2} t^\mu u + \eta^\mu + (r - r_0) l^\mu, \quad t^\mu t_\mu = 1,$$

where r is an affine parameter along c and r_0 specifies an arbitrary affine origin. l^μ is normalized by $l^\mu t_\mu = 1$. The connecting vector can be written as

$$\eta^\mu = -L\overline{m}^\mu - \overline{L} m^\mu$$

where m^μ and its complex conjugate $\overline{m}^\mu$ are two spacelike complex vectors defined (up to the transformation $m^\mu \to e^{i\phi} m^\mu$) by

$$m^\mu \overline{m}_\mu + 1 = m^\mu m_\mu = m^\mu l_\mu = m^\mu t_\mu = 0.$$

and $L = m_\mu \eta^\mu$ is a spin weight 1 quantity.

The congruence C can thus be described by taking the connecting vector as a known function of the hyperplane and thereby obtaining for the description of the geodesics of C, the explicit form

$$x^\mu = \sqrt{2} t^\mu u - L(u, \zeta, \overline{\zeta}) \overline{m}^\mu(\zeta, \overline{\zeta}) - \overline{L}(u, \zeta, \overline{\zeta}) m^\mu(\zeta, \overline{\zeta})$$
$$+ [r - r_0(u, \zeta, \overline{\zeta})] l^\mu(\zeta, \overline{\zeta}), \tag{2.11}$$

where

$$l^\mu = \frac{\sqrt{2}}{4P_0}\left(1 + \zeta\overline{\zeta},\ \zeta + \overline{\zeta},\ \frac{\zeta - \overline{\zeta}}{i},\ -1 + \zeta\overline{\zeta}\right), \quad P_0 = 1/2(1 + \zeta\overline{\zeta}), \tag{2.12a}$$

$$m^\mu = \eth_0 l^\mu = \frac{\sqrt{2}}{4P_0}(0, 1 - \overline{\zeta}^2, -i(1 + \overline{\zeta}^2), 2\overline{\zeta}), \tag{2.12b}$$

and r_0 is an arbitrary function of u, ζ, and $\overline{\zeta}$.

For future reference, in order for the divergence of the tangent field to C, i.e., $l^\mu(\zeta, \overline{\zeta})$, to have a conventional asymptotic form, we choose

$$r_0 = -\tfrac{1}{2}(\overline{\eth}_0 L + L\mathring{\overline{L}} + \eth_0 \overline{L} + \overline{L}\mathring{L}). \tag{2.12c}$$

If $L(u, \zeta, \overline{\zeta})$ is analytic separately in the three arguments, we can consider the "freeing" of $\overline{\zeta}$ from the complex conjugate of ζ and thus view L as an analytic function of the three independent complex arguments u, ζ, and $\eta = \overline{\zeta}$. This is turn, permits us to replace in (2.12) $\overline{\zeta}$ by η and consider (2.11) to be a three complex parameter family of null geodesics in complex Minkowski space. The affine parameter r is also permitted to take on complex values.

It is now assumed that when working with (2.11) and (2.12) $\overline{\zeta}$ and x^μ are replaced by η and z^μ.

Equation (2.11) can be considered not only as the parametric form of the congruence C but also as the analytic coordinate transformation from the complex Minkowski coordinates z^μ to the natural coordinates associated with the congruence, namely u, r, ζ, η.

The function L will be chosen by the following procedure. Beginning with the complex world line (2.4), $z^\mu = \xi^\mu(\phi)$, we define

$$u = X(\phi, \zeta, \eta) \equiv \xi^\mu(\phi) l_\mu(\zeta, \eta) \tag{2.13}$$

and thereby (implicitly) define an analytic scalar function of three complex variables, namely

$$\phi = \phi(u, \zeta, \eta). \tag{2.14}$$

We also need a second analytic scalar $\overline{\phi}(u, \zeta, \eta)$ which is obtained from the "complex conjugate" world line $\overline{\xi}^\mu(\overline{\phi})$ by

$$u = \overline{X}(\overline{\phi}, \zeta, \eta) \equiv \overline{\xi}^\mu(\overline{\phi}) l_\mu(\zeta, \eta). \tag{2.15}$$

(The idea of a "complex conjugate" world line depends on our retaining the identity of the real Minkowski space. We emphasize that $\phi(u, \zeta, \eta)$ and $\overline{\phi}(u, \zeta, \eta)$ are *not* in general complex conjugates of each other except when u is real and $\eta = \overline{\zeta}$. This eventually leads to the result that ϕ and $\overline{\phi}$ as fields in complex Minkowski space are complex conjugates only on the real subspace.)

Using these scalar functions ϕ and $\overline{\phi}$ we choose L and $\overline{L}$ to be

$$L = -\frac{\eth_0 \phi}{\mathring{\phi}}, \quad \overline{L} = -\frac{\overline{\eth}_0 \overline{\phi}}{\mathring{\overline{\phi}}}$$

with $\mathring{\phi} \equiv \partial\phi/\partial u$. With this choice, the transformation between the z^μ and the new coordinates can (after some manipulation) be put into the form

$$z^\mu = \xi^\mu(\phi(u, \zeta, \eta)) + (i\Sigma + Q^0 \mathring{L}) l^\mu$$
$$- Q^0 \eth_0 l^\mu + r l^\mu \tag{2.16}$$

where l^μ and $\eth_0 l^\mu$ are given by Eq. (2.12) and Σ, Q^0, and L ($\mathring{L} = \partial L/\partial u$) depend on ϕ and $\overline{\phi}$ by

$$2i\Sigma = \eth_0 \overline{L} + L\mathring{\overline{L}} - \overline{\eth}_0 L - \overline{L}\mathring{L}, \tag{2.17}$$

$$L = -\frac{\eth_0 \phi}{\mathring{\phi}}, \quad \overline{L} = -\frac{\overline{\eth}_0 \overline{\phi}}{\mathring{\overline{\phi}}}, \tag{2.18}$$

$$Q^0 = \frac{\overline{\eth}_0 \phi}{\mathring{\phi}} - \frac{\overline{\eth}_0 \overline{\phi}}{\mathring{\overline{\phi}}}. \tag{2.19}$$

[Alternate expressions for these quantities are easily obtained by implicit differentiation of (2.14) and (2.15), e.g.,

$$L = \eth_0 X, \quad \mathring{\phi} = X'^{-1}.] \tag{2.18'}$$

The line element (2.1) becomes (after a lengthy calculation)

$$ds^2 = 2(ln - m\overline{m}), \tag{2.20}$$

with

$$l = l_\mu dx^\mu = du - (1/2P_0)(L d\zeta + \overline{L} d\eta), \tag{2.21a}$$

$$m = m_\mu dx^\mu = -\frac{1}{2P_0 \rho} d\eta, \tag{2.21b}$$

$$\overline{m} = \overline{m}_\mu dx^\mu = -(1/2P_0 \overline{\rho}) d\eta, \tag{2.21c}$$

$$n = n_\mu dx^\mu = dr - (1/2P_0)[(\omega^0 + L/\overline{\rho}) d\zeta + (\overline{\omega}^0 + \overline{L}/\rho) d\eta]$$
$$+ [1 + \tfrac{1}{2}(\eth_0 \mathring{\overline{L}} + L\ddot{\overline{L}} + \overline{\eth}_0 \mathring{L} + \overline{L}\ddot{L})] l_\mu dx^\mu \tag{2.21d}$$

where

$$\rho = -\frac{1}{(r + i\Sigma)}, \quad \overline{\rho} = \frac{-1}{(r - i\Sigma)}, \tag{2.22}$$

and

$$\omega^0 = -i(\eth_0 \Sigma + L\mathring{\Sigma} + 2\mathring{L}\Sigma). \tag{2.23}$$

ρ is the complex divergence of l and is defined by

$$\rho = l_{\mu;\nu} m^\mu \bar{m}^\nu.$$

In order to obtain (2.20) and (2.21), considerable use had to be made of

$$\eth_0 L + L\mathring{L} = 0, \quad (2.24)$$

which follows directly from (2.13) and (2.18). Equation (2.24) states that the shear of the null vector field l vanishes.[5] (See Appendix B.) Thus l, in general, satisfies conditions (1)—(3). The function Σ is a measure of the twist of l, with $\Sigma = 0$ implying that l is proportional to a gradient. One can show that if the world line $\xi^\mu(\phi)$ were real (for real ϕ) then $\Sigma = 0$, and l would be the null tangent vectors to the ordinary light cones from the (real) world line and thus would, in addition, satisfy condition (4).

We have introduced two null vector fields and their associated coordinate (and tetrad) systems in complex Minkowski space which are associated with a complex world line. The first, the simpler and more natural one, suffers from two drawbacks; real values of the coordinates correspond to complex points and the vector field l^* at real points is complex. The second system is vastly more complicated to derive and work with but possesses the major advantages of having a real field l at real points and having real coordinates (u and r real, $\eta = \bar{\zeta}$) correspond to real points.

The tranformation connecting the two coordinate systems is

$$\phi = \phi(u, \zeta, \eta) \quad \text{or} \quad u = X(\phi, \zeta, \eta), \quad (2.25\text{a})$$

$$r^* = [r + i\Sigma(u, \zeta, \eta)]/\mathring{\phi}, \quad (2.25\text{b})$$

$$\eta^* = \eta, \quad (2.25\text{c})$$

$$\zeta^* = [\zeta + \mathring{\phi} Q^0 \rho A(\phi, \eta)]/[1 - \mathring{\phi} Q^0 \rho B(\phi, \eta)], \quad (2.25\text{d})$$

where A and B are defined from

$$V^*(\phi, \zeta^*, \eta^*) = \frac{A(\phi, \eta^*) + \zeta^* B(\phi, \eta^*)}{1 + \zeta^* \eta^*},$$

which follows from the definition of V^*, (2.8).

The tetrad systems are related by

$$l^* = (l - Qm)\mathring{\phi}, \quad (2.26\text{a})$$

$$m^* = S^{-1} m, \quad (2.26\text{b})$$

$$\bar{m}^* = S(\bar{m} - Qn + Ql - Q^2 m), \quad (2.26\text{c})$$

$$n^* = (n + Qm)\mathring{\phi}^{-1}, \quad (2.26\text{d})$$

where $Q = Q_0 \rho$, and

$$S = \frac{P_0^*(\zeta^* \eta^*) V^*(\phi, \zeta^*, \eta^*)}{P_0(\zeta, \eta)\mathring{\phi}^{-1}(u, \zeta, \eta)}. \quad (2.27)$$

From (2.26) we have

$$l = \mathring{\phi}^{-1} l^* + QSm^* \quad (2.28)$$

which on the real Minkowski space gives the projection of the complex l^* onto the real l.

It is the real l field on the real Minkowski space that is of direct physical interest to us.

3. COMPLEXIFIED LIENARD-WIECHERT FIELDS

In this section we will first find the (regular) solutions of Maxwell equations in real Minkowski space such that the l field of the previous section is a p.n.v.f. of the Maxwell tensor. Then we will study the analytic extension of these solutions.

The easiest way to accomplish this is to use the spin-coefficient formulation[6] of the Maxwell equations in the tetrad and coordinate system associated with l, i.e., with (2.20) and (2.21). We then have as basic variables

$$\phi_0 = F_{\mu\nu} l^\mu m^\nu,$$
$$\phi_1 = \tfrac{1}{2} F_{\mu\nu}(l^\mu n^\nu + m^\mu \bar{m}^\nu), \quad (3.1)$$
$$\phi_2 = F_{\mu\nu} \bar{m}^\mu n^\nu.$$

The condition that l be a p.n.v.f. of the Maxwell tensor is

$$\phi_0 = 0. \quad (3.2)$$

The Maxwell equations [under condition (3.2)] have already been partially integrated.[7] We simply quote the results:

$$\phi_0 = 0,$$
$$\phi_1 = \phi_1^0 \rho^2. \quad (3.3)$$
$$\phi_2 = \phi_2^0 \rho + \phi_2^1 \rho^2 + \phi_2^2 \rho^3,$$

with

$$\rho = -\frac{1}{r + i\Sigma}$$

and

$$\phi_2^1 = \bar{\eth}_0 \phi_1^0 + \bar{L}\mathring{\phi}_1^0 + 2\mathring{\bar{L}}\phi_1^0, \quad (3.4\text{a})$$

$$\phi_2^2 = 2i\phi_1^0(\bar{\eth}_0 \Sigma + \bar{L}\mathring{\Sigma} + \mathring{\bar{L}}\Sigma). \quad (3.4\text{b})$$

ϕ_1^0 and ϕ_2^0 satisfy the equations

$$\eth_0 \phi_1^0 + L\mathring{\phi}_1^0 + 2\mathring{L}\phi_1^0 = 0, \quad (3.5)$$

$$\eth_0 \phi_2^0 + L\mathring{\phi}_2^0 + \mathring{L}\phi_2^0 = \mathring{\phi}_1^0. \quad (3.6)$$

Although these equations appear rather formidable, they can be solved rather easily by changing the independent variables u, ζ, η to ϕ, ζ, η, with $\phi = \phi(u, \zeta, \eta)$ defined by (2.13). [It should be noted that ϕ is complex for real u and $\eta = \bar{\zeta}$. Although now this is just a formal means of solving (3.5) and (3.6), eventually we will be interested in the same field but expressed in the ϕ, r^*, ζ^*, η^* coordinates.]

Equations (3.5) and (3.6) become

$$\eth_0'(\phi_1^0 V^2) = 0, \quad (3.7)$$

$$\eth_0'(\phi_2^0 V) = (\phi_1^0)', \quad (3.8)$$

where, as earlier $'$ denotes differentiation with respect to ϕ, $\eth_0'$ is $\eth_0$ holding ϕ constant and

$$V(\phi, \zeta, \eta) \equiv X'(\phi, \zeta, \eta) = \xi'^\mu(\phi) l_\mu(\zeta, \eta) = \mathring{\phi}^{-1}(u, \zeta, \eta). \quad (3.9)$$

At this point we make the assumption that the complex world line is "time-like". By "time-like" we mean that the function $V(\phi, \zeta, \eta)$ has no zeros for the values ϕ, ζ, and η taken in the real Minkowski space, i.e., ϕ takes on the values [from its defining equation (2.14)] $\phi = \phi(u, \zeta, \eta = \bar{\zeta})$, u real, and $\eta = \bar{\zeta}$. For a real world line this definition of time-like coincides with the usual one.

The regular solution of (3.7) is

$$\phi_1^0 = e(\phi)/V^2, \tag{3.10}$$

for which (3.8) becomes

$$\eth_0'(\phi_2^0 V) = (e/V^2)'. \tag{3.11}$$

When (3.11) is integrated over the surface of a sphere (using the properties of $\eth_0$) we obtain

$$\frac{d}{d\phi}[e\int V^{-2} d\Omega] = 0. \tag{3.12}$$

By explicit integration (or by more general methods) $\int V^{-2} d\Omega = 4\pi$, so that $e' = 0$, or $e = \text{const} = \text{charge}$.

The final equation to be solved is then

$$V^2 \eth_0'(\phi_2^0 V) = -2e(V'/V). \tag{3.13}$$

One can readily check that V from its defining equation (3.9) satisfies

$$V^2 \eth_0' \bar{\eth}_0' \log P_0 V = 1,$$

which, when differentiated, becomes

$$2V'/V = -V^2 \eth_0' \bar{\eth}_0'(V'/V). \tag{3.14}$$

By comparing (3.13) and (3.14), we have

$$\phi_2^0 = e[(1/V)\bar{\eth}_0(V'/V)], \tag{3.15}$$

or in terms of the coordinates u, ζ, η,

$$\phi_2^0 = -e(\bar{\eth}_0 \phi/\dot{\phi})^{\cdot\cdot}. \tag{3.16}$$

This class of solutions of the Maxwell equations (Kerr-type Maxwell fields), i.e., (3.3), (3.4), (3.10) (with $e \neq 0$), and (3.15) constitutes the class of general regular solutions whose p.n.v.f. satisfies conditions (1)–(3) and which has nonvanishing charge. Note that the solution is determined completely by e and the function $\phi(u, \zeta, \eta)$, which is defined by the complex (time-like) world line.

We now show that this solution analytically extended into complex Minkowski space coincides with the solution of the complexified, analytically extended Maxwell equations with an electric monopole source moving along an arbitrary, time-like, complex world line, i.e., that it is a complexified Lienard–Wiechert solution.

The complexified and analytically extended Maxwell field $W^{\mu\nu}(z^\mu)$ is, on the real Minkowski space, given by $W^{\mu\nu}(x^\mu) = \frac{1}{2}(F^{\mu\nu} + iF^{*\mu\nu})$ and satisfies (in the vacuum region)

$$\frac{\partial}{\partial z^\mu} W^{\mu\nu} = 0. \tag{3.17}$$

If one introduces a complex null tetrad (in the complex space) say, l^*, m^*, $\bar{m}^*$, n^* satisfying the conditions $l^* \cdot n^* = -m^* \cdot \bar{m}^* = 1$, all other scalar products being zero, then one can define

$$\begin{aligned} \phi_0^* &= W_{\mu\nu} l^{*\mu} m^{*\nu}, \\ \phi_1^* &= \tfrac{1}{2} W_{\mu\nu}(l^{*\mu} n^{*\nu} + \bar{m}^{*\mu} m^{*\nu}), \\ \phi_2^* &= W_{\mu\nu} \bar{m}^{*\mu} n^{*\nu}. \end{aligned} \tag{3.18}$$

[If the tetrad is real in the real space, then (3.18) coincides with the usual spin coefficient form of the Maxwell field.]

The spin coefficient form of the extended Maxwell equations (3.17) is identical to the usual spin coefficient form except that the spin coefficients are obtained from the extended complex null tetrad.

. . .

On real space the p.n.v.f. for the real Maxwell field just discussed was real, unique, and given by

$$F_{\mu\nu} l^\nu = \lambda l_\mu.$$

In the complexified version there is no meaning to $F_{\mu\nu}$ (except on the real space), but only to $W_{\mu\nu}$. This suggests the eigenvalue equation

$$W_{\mu\nu} L^\nu = \lambda L_\mu.$$

This, however, does not lead to a unique solution for L; there is, in fact, a complex plane of solutions (see Appendix A) at each point for each eigenvalue λ. If we demanded that L be real on the real space we would obtain by analytic continuation that $L = l$. There is, however, another natural choice for L, namely the l^* of the previous section, which has simple geometric meaning.

We shall solve the extended Maxwell equations (in spin coefficient form) using the * coordinate and tetrad system of Sec. 2, assuming the l^* is a p.n.v.f. of $W^{\mu\nu}$ (i.e., assuming that $\phi_0^* = 0$) and then show that this solution coincides with the one just obtained. The equations take the form (after a lengthy calculation)

$$D^* \phi_1^* + 2\phi_1^*/r^* = 0, \tag{3.19a}$$

$$D^* \phi_2^* + \phi_2^*/r^* = -(V^*/r^*)\bar{\eth}_0^* \phi_1^*, \tag{3.19b}$$

$$0 = -(V^*/r^*)\eth_0^* \phi_1^*, \tag{3.19c}$$

$$\phi_1^{*\prime} - \left(1 - \frac{V'^*}{V^*} r^*\right) D^* \phi_1^* - 2\frac{\phi_1^*}{r^*} = -\frac{V^{*2}}{r^*} \eth_0^* \left(\frac{\phi_2^*}{V^*}\right). \tag{3.19d}$$

Integrating the first two of (3.19) yields

$$\begin{aligned} \phi_0^* &= 0, \\ \phi_1^* &= \phi_1^{0*}(\phi, \zeta^*, \eta^*)/r^{*2}, \\ \phi_2^* &= \phi_2^{0*}(\phi, \zeta^*, \eta^*)/r^* + (V^*/r^{*2})\bar{\eth}_0^* \phi_1^{0*}, \end{aligned} \tag{3.20}$$

which when substituted into the second pair of (3.19) yields

$$\eth_0^* \phi_1^{0*} = 0, \tag{3.21}$$

$$(\phi_1^{0*} V^{*-2})' = -\eth_0^*(\phi_2^{0*}/V^*). \tag{3.22}$$

We have used in (3.22) the result from (3.21) that $\phi_1^{0*} = \phi_1^{0*}(\phi)$ which follows from the regularity assumption. The analysis here is almost identical to that following Eq. (3.11) with the result that

$$\begin{aligned} \phi_1^{0*} &= e^* = \text{const}, \\ \phi_2^{0*} &= -e^* V^* \bar{\eth}_0^*(V^{*\prime}/V^*). \end{aligned} \tag{3.23}$$

This solution of the extended Maxwell equations can be naturally thought of as a complexified Lienard—Wiechert field.[1]

By subjecting the ϕ^*'s to the coordinate transformation (2.25) and the tetrad transformation (2.26) one can simply check that the CLW field is the same as the Kerr-type Maxwell field.

We have thus proved our contention that a regular solution of the (real) Maxwell equations whose p.n.v.f. satisfies conditions (1)—(3) when analytically extended can be looked upon as a CLW field.

4. THE KERR-TYPE METRICS

In this section we present a brief review of the spin coefficient formulation of the real Kerr-type metrics.

In a four-dimensional Riemannian manifold with signature $(+,-,-,-)$ a null tetrad $Z_{a\mu}=(l_\mu, n_\mu, m_\mu, \overline{m}_\mu)$ is introduced composed of two real (l,n) and two complex $(m,\overline{m})$ null vectors satisfying

$$l\cdot n = -m\cdot\overline{m}=1, \tag{4.1}$$

all other scalar products vanishing. Equation (4.1) implies the completeness relation

$$g_{\mu\nu}=2(l_{(\mu}n_{\nu)}-m_{(\mu}\overline{m}_{\nu)}). \tag{4.2}$$

From the tetrad one can define the Ricci rotation coefficients

$$\gamma^{abc}=Z^a_{\mu;\nu}Z^{b\mu}Z^{c\nu} \tag{4.3}$$

and the spin coefficients

$$\begin{aligned}
&\kappa=l_{\mu;\nu}m^\mu l^\nu, \qquad \nu=-n_{\mu;\nu}\overline{m}^\mu n^\nu,\\
&\rho=l_{\mu;\nu}m^\mu\overline{m}^\nu, \quad \mu=-n_{\mu;\nu}\overline{m}^\mu m^\nu,\\
&\sigma=l_{\mu;\nu}m^\mu m^\nu, \quad \lambda=-n_{\mu;\nu}\overline{m}^\mu\overline{m}^\nu,\\
&\tau=l_{\mu;\nu}m^\mu n^\nu, \qquad \pi=-n_{\mu;\nu}\overline{m}^\mu l^\nu,\\
&\alpha=1/2(l_{\mu;\nu}n^\mu\overline{m}^\nu-m_{\mu;\nu}\overline{m}^\mu\overline{m}^\nu),\\
&\beta=1/2(l_{\mu;\nu}n^\mu m^\nu-m_{\mu;\nu}\overline{m}^\mu m^\nu),\\
&\gamma=1/2(l_{\mu;\nu}n^\mu n^\nu-m_{\mu;\nu}\overline{m}^\mu n^\nu),\\
&\epsilon=1/2(l_{\mu;\nu}n^\mu l^\nu-m_{\mu;\nu}\overline{m}^\mu l^\nu).
\end{aligned} \tag{4.4}$$

The tetrad components of the Weyl tensor are defined by

$$\begin{aligned}
&\psi_0=-C_{\mu\nu\rho\sigma}l^\mu m^\nu l^\rho m^\sigma,\\
&\psi_1=-C_{\mu\nu\rho\sigma}l^\mu n^\nu l^\rho m^\sigma,\\
&\psi_2=-C_{\mu\nu\rho\sigma}\overline{m}^\mu n^\nu l^\rho m^\sigma,\\
&\psi_3=-C_{\mu\nu\rho\sigma}\overline{m}^\mu n^\nu l^\rho n^\sigma,\\
&\psi_4=-C_{\mu\nu\rho\sigma}\overline{m}^\mu n^\nu\overline{m}^\rho n^\sigma.
\end{aligned} \tag{4.5}$$

Directional derivatives have the form

$$\begin{aligned}
&D\phi\equiv\phi_{;\mu}l^\mu, \quad \Delta\phi=\phi_{;\mu}n^\mu,\\
&\delta\phi=\phi_{;\mu}m^\mu, \ \overline{\delta}\phi=\phi_{;\mu}\overline{m}^\mu.
\end{aligned} \tag{4.6}$$

The spin coefficient formalism then consists of three sets of first order [in the derivatives (4.6)] differential equations (equivalent to the vacuum Einstein equations) for the three sets of variables: the spin coefficients, the Weyl tensor components, and the tetrad or metric components.

The formalism can be readily adapted to yield the Kerr-type metrics by simply choosing the null vector field l to be a p.n.v.f. of the Weyl tensor satisfying conditions (1)—(3). By condition (3) l is shearfree. Therefore, according to the Goldberg—Sachs[8] theorem, l is a degenerate principal null vector field of the Weyl tensor satisfying

$$C_{\mu\nu\rho[\sigma}l_{\tau]}l^\nu l^\rho=0 \tag{4.7}$$

or, equivalently,

$$\psi_0=0=\psi_1. \tag{4.8}$$

We also introduce a null coordinate system, $x^\mu=(u,r,\zeta,\eta), \eta=\overline{\zeta}$, associated with the null tetrad such that r is an affine parameter along the null geodesics, labelled by (u,ζ,η), to which l is tangent.

Integration of the three sets of equations may now proceed under the above assumption on l. In performing the integrations considerable simplification of the results is gained through the use of available coordinate and tetrad freedom. For instance, the tetrad vectors m and n may be chosen to be parallelly propagated along l, which leads to $\epsilon=\pi=0$. Because the details of these integrations appear elsewhere only the results[7] will be given here.

The tetrad components of the Weyl tensor have the form

$$\psi_0=\psi_1=0, \tag{4.9a}$$

$$\psi_2=\psi_2^0\rho^3, \tag{4.9b}$$

$$\psi_3=\psi_3^0\rho^2+\psi_3^1\rho^3+\psi_3^2\rho^4, \tag{4.9c}$$

$$\psi_4=\psi_4^0\rho+\psi_4^1\rho^2+\tfrac{1}{2}\psi_4^2\rho^3+\tfrac{1}{3}\psi_4^3\rho^4+\tfrac{1}{4}\psi_4^4\rho^5, \tag{4.9d}$$

with

$$\psi_3^0=\eth_0 R+L\dot{R}, \tag{4.10a}$$

$$\psi_3^1=\overline{\eth}_0\psi_2^0+\overline{L}\mathring{\psi}_2^0+3\dot{\overline{L}}\psi_2^0, \tag{4.10b}$$

$$\psi_3^2=3i\psi_2^0(\overline{\eth}_0\Sigma+\overline{L}\mathring{\Sigma}+\mathring{\overline{L}}\Sigma), \tag{4.10c}$$

$$\psi_4^0=\mathring{R}, \tag{4.10d}$$

$$\psi_4^1=\overline{\eth}_0\psi_3^0+\overline{L}\dot{\psi}_3^0+4\mathring{\overline{L}}\psi_3^0, \tag{4.10e}$$

$$\psi_4^2=\overline{\eth}_0\psi_3^1+\overline{L}\mathring{\psi}_3^1+5\mathring{\overline{L}}\psi_3^1+4i\psi_3^0(\overline{\eth}_0\Sigma+\overline{L}\mathring{\Sigma}+\mathring{\overline{L}}\Sigma), \tag{4.10f}$$

$$\psi_4^3=\overline{\eth}_0\psi_3^2+\overline{L}\dot{\psi}_3^2+6\mathring{\overline{L}}\psi_3^2+6i\psi_3^1(\overline{\eth}_0\Sigma+\overline{L}\mathring{\Sigma}+\mathring{\overline{L}}\Sigma), \tag{4.10g}$$

$$\psi_4^4=8i\psi_3^2(\overline{\eth}_0\Sigma+\overline{L}\dot{\Sigma}+\mathring{\overline{L}}\Sigma). \tag{4.10h}$$

The spin coefficients become

$$\kappa=\epsilon=\pi=\sigma=\tau=\lambda=0, \tag{4.11a}$$

$$\rho=-1/(r+i\Sigma), \tag{4.11b}$$

$$\alpha=\tfrac{1}{2}(\overline{\eth}_0\log P_0+2\mathring{\overline{L}})\rho, \tag{4.11c}$$

$$\beta=-\tfrac{1}{2}(\eth_0\log P_0)\overline{\rho}, \tag{4.11d}$$

$$\gamma=\tfrac{1}{2}\psi_2^0\rho^2, \tag{4.11e}$$

$$\mu = (\eth_0 N + L\mathring{N})\bar{\rho} + \tfrac{1}{2}\psi_2^0(\rho^2 + \rho\bar{\rho}), \tag{4.11f}$$

$$\nu = \mathring{N} + \psi_3^0\rho + \tfrac{1}{2}\psi_3^1\rho^3 + \tfrac{1}{3}\psi_3^2\rho_1^3 \tag{4.11g}$$

with

$$L = -\eth_0\phi/\mathring{\phi}, \tag{4.12a}$$

$$2i\Sigma = \eth_0\bar{L} + L\mathring{\bar{L}} - \bar{\eth}_0 L - \bar{L}\mathring{L}, \tag{4.12b}$$

$$N = \bar{\eth}_0 \log P_0 + \mathring{\bar{L}}, \tag{4.12c}$$

$$R = \bar{\eth}_0 N + \bar{L}\mathring{N} + N^2 - 2N\bar{\eth}_0 \log P_0. \tag{4.12d}$$

The metric may be written in the form

$$ds^2 = 2(ln - m\bar{m}), \tag{4.13}$$

with

$$l = l_\mu dx^\mu - du - \frac{1}{2}\frac{L d\zeta}{P_0} - \frac{1}{2}\frac{\bar{L} d\eta}{\rho_0}, \tag{4.14a}$$

$$m = m_\mu dx^\mu = -(1/2P_0\rho)d\eta, \tag{4.14b}$$

$$\bar{m} = \bar{m}_\mu dx^\mu = -(1/2P_0\bar{\rho})d\zeta, \tag{4.14c}$$

$$n = n_\mu dx^\mu = dr - \frac{1}{2P_0}\left[\left(\omega^0 + \frac{\mathring{L}}{\bar{\rho}}\right)d\zeta + \left(\bar{\omega}^0 + \frac{\mathring{\bar{L}}}{\rho}\right)d\eta\right] \tag{4.14d}$$

$$+ [1 + \tfrac{1}{2}(\eth_0\mathring{\bar{L}} + L\mathring{\bar{L}} + \bar{\eth}_0\mathring{L} + \bar{L}\mathring{L}) + \tfrac{1}{2}(\psi_2^0\rho + \bar{\psi}_2^0\bar{\rho})]l_\mu dx^\mu$$

where

$$\omega^0 = -i(\eth_0\Sigma + L\mathring{\Sigma} + 2\mathring{L}\Sigma). \tag{4.14e}$$

Finally, ψ_2^0 must satisfy the differential equations

$$\eth_0\psi_2^0 + L\mathring{\psi}_2^0 + 3\mathring{L}\psi_2^0 = 0, \tag{4.15}$$

$$\mathring{\psi}_2^0 = \eth_0\psi_3^0 + L\mathring{\psi}_3^0 + 2\mathring{L}\psi_3^0, \tag{4.16}$$

with

$$\mathrm{Im}\left[\psi_2^0 - \left(\eth_0 + L\frac{\partial}{\partial u}\right)\left(\eth_0 + L\frac{\partial}{\partial u}\right)(\bar{\eth}_0\bar{L} + \bar{L}\mathring{\bar{L}})\right] = 0. \tag{4.17}$$

In summary then the Kerr-type metrics constitute the most general solution to the vacuum Einstein equations for which a p.n.v.f. of the Weyl tensor satisfies conditions (1)—(3).

Equation (4.15) can be easily integrated [compare with Eq. (3.7)] and has the general regular solution

$$\psi_2^0 = m(\phi)\mathring{\phi}^3. \tag{4.18}$$

From its defining equation (4.12a) we see that ϕ can always be replaced by an arbitrary function of ϕ. This freedom can be used, for example, to put m in (4.18) equal to a constant.

Therefore, aside from m, the entire class of regular Kerr-type solutions to the vacuum Einstein equations is completely determined by the single complex function $\phi(u, \zeta, \eta)$, which in turn must satisfy the differential equation (4.16).

Thus, the analogy between the solutions of this section and the Kerr-type Maxwell solutions discussed in the first part of Sec. 3 is established. The coordinate-tetrad system used here, i.e., (4.13) and (4.14), may also be compared with its flat space analog (2.20) and (2.21).

Finally, if ϕ is real, the class of solutions reduces to the RT metrics. That is, l becomes twist free (or proportional to a gradient) and thus satisfies condition (4) in addition to conditions (1)—(3). The twist of l is given by the imaginary part of ρ, (4.11b), so that the function Σ is a measure of the twist. When ϕ is real, Eqs. (4.12a) and (4.12b) immediately lead to $\Sigma = 0$.

5. THE COMPLEXIFIED KERR-TYPE METRICS

The real Kerr-type metrics as given in the previous section can easily be analytically extended into a complex manifold. First, we replace the Weyl tensor by its complexified analytically extended version $W_{\mu\nu\rho\sigma}$ which on the real space is given by

$$W_{\mu\nu\rho\sigma} = \tfrac{1}{2}(C_{\mu\nu\rho\sigma} + iC^*_{\mu\nu\rho\sigma}). \tag{5.1}$$

Then we allow the coordinates to assume the complex values $z^\mu = (u, r, \zeta, \eta)$ so that the metric (4.13) and all of the null tetrad vectors (4.14), although unchanged in form, will now be complex. We are tacitly assuming that all functions considered in Sec. 4 are real analytic.

The tetrad components of the Weyl tensor will be defined in terms of the extended complex tetrad as

$$\begin{aligned} \psi_0 &= -W_{\mu\nu\rho\sigma}l^\mu m^\nu l^\rho m^\sigma, \\ \psi_1 &= -W_{\mu\nu\rho\sigma}l^\mu n^\nu l^\rho n^\sigma, \\ \psi_2 &= -W_{\mu\nu\rho\sigma}\bar{m}^\mu n^\nu l^\rho m^\sigma, \\ \psi_3 &= -W_{\mu\nu\rho\sigma}\bar{m}^\mu n^\nu l^\rho n^\sigma, \\ \psi_4 &= -W_{\mu\nu\rho\sigma}\bar{m}^\mu n^\nu \bar{m}^\rho n^\sigma. \end{aligned} \tag{5.2}$$

The spin coefficients will also be obtained from the extended tetrad. Because m and $\bar{m}$ are no longer complex conjugates, the list of spin coefficients will have to be doubled to separately include the now independent "conjugate" versions of those appearing in (4.4).

On the real subspace of the complex space (characterized by u and r real and $\eta = \bar{\zeta}$) the tetrad components of the Weyl tensor given by (4.5) and (5.2) coincide and everything else reduces to the results given in Sec. 4.

For the general Kerr-type solutions, the field l is a unique p.n.v.f. of the Weyl tensor on the real subspace. In the analytically extended complex space, however, we have

$$W_{\mu\nu\rho[\sigma}l_{\tau]}l^\nu l^\rho = 0, \tag{5.3}$$

or, equivalently,

$$\psi_0 = -W_{\mu\nu\rho\sigma}l^\mu m^\nu l^\rho m^\sigma = 0$$

and

$$\psi_1 = -W_{\mu\nu\rho\sigma}l^\mu n^\nu l^\rho m^\sigma = 0. \tag{5.4}$$

We know that l and the null tetrad associated with it satisfy (5.3) and (5.4). But the question is are there other solutions as well in the complex space? We can find this out by determining the most general tetrad transformation we can make which preserves (5.3) and

(5.4). This transformation has the form

$$l^* = R(l + Am), \qquad (5.5)$$
$$n^* = R^{-1}(n + Bm),$$
$$m^* = S^{-1}m,$$
$$\bar{m}^* = S[\bar{m} + Al + Bn + ABm],$$

where A, B, R, S are arbitrarily complex functions of the coordinates z^μ.

Just as in the corresponding situation in the Maxwell case, l in the analytically extended space is no longer unique. Instead there is a complex plane of solutions at each point and any null vector of the form $R(l + Am)$ will be a p.n.v.f. of the complexified Weyl tensor.

We will now show that among these possibilities, there is one complex p.n.v.f. of $W_{\mu\nu\rho\sigma}$ which is a gradient and thus satisfies conditions (1)–(4). In the process we will obtain the transformations from the complex coordinate-tetrad system associated with l to the one associated with l^*.

In order for l^* to be a p.n.v.f. of $W_{\mu\nu\rho\sigma}$ it must have the form $l^* = R(l + Am)$. After substituting L from (4.12a) into the expression for l, (4.14a), examination of m, (4.14b), leads to the result that if we choose $R = \dot{\phi}$ and $A = -Q$, where $Q = Q_0\rho$ and Q_0 is given by (2.19), then $l^*_\mu = \phi_{,\mu}$.

Writing the new coordinate system as $z^{\mu *} = (u^*, r^*, \zeta^*, \eta^*)$, we may choose r^* to be a (complex) affine parameter along the l^*. This means that r^* must satisfy

$$D^* r^* = \dot{\phi}(D - Q\delta) r^* = 1, \qquad (5.6)$$

where $D^* \equiv l^{*\mu}(\partial/\partial z_\mu)$. The remaining three coordinates $z^{a*} = (u^*, \zeta^*, \eta^*)$ must satisfy

$$D^* z^{a*} = 0 \qquad (5.7)$$

in order to be constant on each ray l^*.

Equation (5.6) has the solution

$$r^* = (r + i\Sigma)/\dot{\phi}, \qquad (5.8)$$

while both $u^* = \phi$ and $\eta^* = \eta$ satisfy Eq. (5.7). Before finding $\zeta^* = F(u, r, \zeta, \eta)$ [which is also to be a solution of Eq. (5.7)], we will first choose convenient forms for the coefficients S and B of Eq. (5.5).

From the function $u = X(\phi, \zeta, \eta)$, which is implicitly defined by $\phi = \phi(u, \zeta, \eta)$ [compare with (2.13) and (2.14) in the flat space case], we can define [as in (3.9)], the quantity

$$V(\phi, \zeta, \eta) \equiv \frac{\partial X}{\partial \phi}(\phi, \zeta, \eta) = \dot{\phi}^{-1}(u, \zeta, \eta).$$

In terms of V we see from (4.14b) that m may be written as

$$m = -\frac{d\eta}{2P_0\rho} = \frac{(r + i\Sigma)d\eta}{(1 + \zeta\eta)} = \frac{r^* d\eta^*}{\dot{\phi}^{-1}(1 + \zeta\eta)} = \frac{r^* d\eta^*}{V(\phi, \zeta, \eta)(1 + \zeta\eta)}.$$

If we define the function $V^*(\phi, \zeta^*, \eta^*)$ by

$$(1 + \zeta^*\eta^*) V^*(\phi, \zeta^*, \eta^*) \equiv (1 + \zeta\eta) V(\phi, \zeta, \eta), \qquad (5.9)$$

we obtain from (5.5) the result that

$$S = \frac{(1 + \zeta^*\eta^*) V^*(\phi, \zeta^*, \eta^*)}{(1 + \zeta\eta)\dot{\phi}^{-1}} = \frac{P^*(\phi, \zeta^*, \eta^*)}{P_0 \dot{\phi}^{-1}}, \qquad (5.10)$$

which may be compared with Eq. (2.27) in Sec. 2.

Because the choice of the function B in (5.5) affects neither l^* nor the metric we will simply choose $B = Q$.

The only remaining problem now is to determine the transformation $\zeta^* = F(u, r, \zeta, \eta)$. This can be accomplished by considering the form $\bar{m}^*$, which, from (5.5), is given by

$$\bar{m}^* = \bar{m}^*_\mu dz^{\mu *} = S(\bar{m}l + Q - Qn - Q^2 m) \qquad (5.11)$$
$$= S(\bar{m}_\mu dz^\mu + Q l_\mu dz^\mu - Q n_\mu dz^\mu - Q^2 m_\mu dz^\mu).$$

Using the facts that $\phi_{,r} = \eta^*_{,r} = 0$ [$\bar{m}^*_1 = 0$ from $(\bar{m}^* \cdot l^*) = 0$, $\bar{m}^*_2 = r^*/2P^*$ from $(\bar{m}^* \cdot m^*) = -1$] and equating the coefficients of the dr term on both sides of (5.11), we obtain

$$\frac{r^*}{2P^*} F_{,r} = -SQ. \qquad (5.12)$$

By using (5.10) and (2.19), Eq. (5.12) may be rewritten in the form

$$F_{,r}/(P^*)^2 = 2(\dot{\phi}^2/P^0) Q_0 \rho^2,$$

where $\dot{\phi}$, Q_0, and P_0 are all independent of r. Integrating we obtain

$$G(\phi, F, \eta) = \frac{2\dot{\phi}^2 Q}{P_0} + G(\phi, \zeta, \eta), \qquad (5.13a)$$

where

$$G(\phi, F, \eta) = \int \frac{dF}{[P^*(\phi, F, \eta^*)]^2} \qquad (5.13b)$$

and

$$G(\phi, \zeta, \eta) = \int \frac{d\zeta}{[P_0/\dot{\phi}]^2}, \qquad (5.13c)$$

so that in the limit $r \to \infty$, $\zeta^* \to \zeta$.

The complete tetrad and coordinate transformations relating the l and l^* systems are given by

$$l^* = \dot{\phi}(l - Qm),$$
$$n^* = \dot{\phi}^{-1}(n + Qm),$$
$$m^* = S^{-1}m, \qquad (5.14)$$
$$\bar{m}^* = S(\bar{m} - Qn + Ql - Q^2 m)$$

with $Q = Q_0\rho$, where Q_0 and S given by (2.19) and (5.10), respectively, and

$$u^* = \phi(u, \zeta, \eta),$$
$$r^* = (r + i\Sigma)/\dot{\phi}, \qquad (5.15)$$
$$\zeta^* = F,$$
$$\eta^* = \eta,$$

with F given implicitly by (5.13).

Under the above transformations

$$\psi^*_2 = \psi_2 = \psi^0_2 \rho^3.$$

But $\overset{\circ}{\phi}\rho = -1/r^*$ and $\psi_2^0/\overset{\circ}{\phi}{}^3 = m(\phi)$ from (4.18). Therefore,

$$\psi_2^* = -m(u^*)/(r^*)^3 \tag{5.16}$$

in the new coordinates. This is the same form (but with complex coordinates) that ψ_2 has in the class of RT solutions.

In this way the analytically extended Kerr-type metrics can be thought of as complexified RT solutions and the analogy with the CLW solutions of Sec. 3 is complete. As an explicit example of how this works, the analytically extended Kerr metric is given in both coordinate systems in Appendix C. The starred coordinate-tetrad system of this section is the analog of the flat space system discussed in Sec. 2.

APPENDIX A

In the text, two null vector fields l and l^* were introduced on complex Minkowski space, both being defined in terms of a complex world line $[z^\mu = \xi^\mu(\phi)]$. The l^* field was the tangent field to the generators of the complex light cones with apexes on the complex world line; l was, in some sense, a projection of l^* such that for l^* on the real Minkowski subspace the projection yields a real l. [See Eq. (2.28).] We wish to point out here (with no proof) that there is a simple way in terms of spinors to see the meaning of this projection. Using spinor notation, the complex world line is written as

$$z^{AA'} = \xi^{AA'}(\phi) \tag{A1}$$

and a point on its complex cones by

$$X^{AA'} = \xi^{AA'}(\phi) + \pi^A\bar{\omega}^{A'} \tag{A2}$$

where π^A and $\bar{\omega}^{A'}$ are two arbitrary spinors and $\pi^A\bar{\omega}^{A'}$ represents an arbitrary complex null displacement. (A2) is the spinor version of Eq. 2.6 where $r^*l^{*\mu} \longleftrightarrow \pi^A\bar{\omega}^{A'}$. For a π^A and $\bar{\omega}^{A'}$ such that $X^{AA'}$ is Hermitian (i.e., represents a real point) our "projection" is simply $\pi^A\omega^{A'} \to \pi^A\bar{\pi}^{A'} \to l^\mu$. Though it is not obvious, a real null field so constructed satisfies conditions (1)—(3).

An alternate way to construct the l field from the complex world line is to use the Penrose theory of twistors.[9] Since each point in complex Minkowski space is a line in twistor space, the complex world line is equivalent to a one complex parameter family of lines in twistor space, or an analytic (and in particular a developable) surface. By the theorem of R. Kerr[9] this generates a shear free congruence in the real Minkowski space, namely the l field.

As a final point we wish to show that although $F^{\mu\nu}$ has (in the nondegenerate case) two unique (real) null eigenvectors, the eigenvectors of $W^{\mu\nu} = \frac{1}{2}(F^{\mu\nu} + iF^{*\mu\nu})$ are degenerate, forming null planes. (See Sec. 3.)

In spinor notation

$$F^{\mu\nu} \longleftrightarrow \epsilon^{A'B'} i\lambda^{(A}\mu^{B)} - \epsilon^{AB} i\bar{\lambda}^{(A'}\bar{\mu}^{B')}. \tag{A3}$$

It is known (and easy to check) that $l^\mu \longleftrightarrow \lambda^A\bar{\lambda}^{A'}$ and $n^\mu \longleftrightarrow \mu^A\bar{\mu}^{A'}$ are the two real null eigenvectors.

The spinor form of $W^{\mu\nu}$ is

$$W^{\mu\nu} \longleftrightarrow \epsilon^{A'B'} i\lambda^{(A}\mu^{B)} \tag{A4}$$

from which it is seen that

$$\epsilon^{A'B'} i\lambda^{(A}\mu^{B)}\lambda_A\bar{\sigma}_{A'} = i\lambda_A\mu^A\lambda^B\bar{\sigma}^{B'}$$

for arbitrary $\bar{\sigma}_{A'}$. This is simply the statement that $\lambda_A\bar{\sigma}_{A'} \longleftrightarrow L$ is an eigenvector of $W^{\mu\nu}$(with eigenvalue $i\lambda_A\mu^A$). From the assumption that $F^{\mu\nu}$ is nondegenerate, the σ_A can be written as a linear combination of λ_A and μ_A, from which it follows that

$$l^* = \alpha l + \beta m$$

which α and β arbitrary and $m \longleftrightarrow \lambda_A\bar{\mu}_{A'}$.

APPENDIX B

One can prove that $\eth L + L\overset{\circ}{L} = 0$ is necessary and sufficient for the congruence C to be shear free by simply computing the shear

$$\sigma = l_{\mu;\nu}m^\mu m^\nu,$$

where l_μ and m^μ are given by (2.12) and

$$l_{\mu;\nu} = l_{\mu,\zeta}\frac{\partial\zeta}{\partial x^\nu} + l_{\mu,\bar{\zeta}}\frac{\partial\bar{\zeta}}{\partial x^\nu}.$$

$\partial\zeta/\partial x^\nu$ and $\partial\bar{\zeta}/\partial x^\nu$ are obtained by differentiating (2.11) with respect to x^ν. The result of this rather lengthy calculation is

$$\sigma = \frac{\sigma^0}{(r + i\Sigma)(r - i\Sigma) - \sigma^0\bar{\sigma}^0}, \quad \sigma^0 = \eth L + L\overset{\circ}{L}$$

from which it follows that $\sigma = 0 \longleftrightarrow \eth L + L\overset{\circ}{L} = 0$.

APPENDIX C

In this appendix we will apply the results of Secs. 4 and 5 to the special case of the analytically extended Kerr metric, which is characterized in the l system by

$$\phi = u - ia[(1 - \zeta\eta)/(1 + \zeta\eta)]. \tag{C1a}$$

$$\bar{\phi} = u + ia[(1 - \zeta\eta)/(1 + \zeta\eta)], \tag{C1b}$$

with the only nonvanishing tetrad component of the Weyl tensor being

$$\psi_2 = -W_{\mu\nu\rho\sigma}\bar{m}^\mu n^\nu l^\rho m^\sigma = -\frac{m}{\{r + 2ia[(1 - \zeta\eta)/(1 + \zeta\eta)]\}^3}, \tag{C2}$$

where a and m are constants.

After substituting (C1) and (C2) into (4.14), using the definitions (4.12), we obtain

$$l = du + \frac{2ia\eta}{(1 + \zeta\eta)^2}d\zeta - \frac{2ia\zeta}{(1 + \zeta\eta)^2}d\eta, \tag{C3a}$$

$$m = \frac{\{r + 2ia[(1 - \zeta\eta)/(1 + \zeta\eta)]\}}{(1 + \zeta\eta)}d\eta, \tag{C3b}$$

$$\bar{m} = \frac{\{r - 2ia[(1 - \zeta\eta)/(1 + \zeta\eta)]\}}{(1 + \zeta\eta)}d\zeta, \tag{C3c}$$

$$n = \left(1 - \frac{mr}{\{r^2 + 4a^2[(1 - \zeta\eta)^2/(1 + \zeta\eta)^2]\}}\right)d\mu + dr \tag{C3d}$$

$$-\left(1+\frac{mr}{\{r^2+4a^2[(1-\zeta\eta)^2/(1+\zeta\eta)^2]\}}\right)\left(\frac{2ia\eta d\zeta}{(1+\zeta\eta)^2}-\frac{2ia\zeta d\eta}{(1+\zeta\eta)^2}\right).$$

From (4.13) the complexified Kerr metric is then given by

$$ds^2=2\left[du^2+dudr+\frac{2iadr}{(1+\zeta\eta)^2}[\eta d\zeta-\zeta d\eta]+\frac{4a^2}{(1+\zeta\eta)^4}[\eta d\zeta-\zeta d\eta]^2-\left(r^2+4a^2\frac{(1-\zeta\eta)^2}{(1+\zeta\eta)^2}\right)\frac{d\zeta d\eta}{(1+\zeta\eta)^2}\right]-\frac{2mr}{\{r^2+4a^2[(1-\zeta\eta)^2/(1+\zeta\eta)]^2\}}\left(du+\frac{2ia\eta d\zeta}{(1+\zeta\eta)^2}-\frac{2ia\zeta d\eta}{(1+\zeta\eta)^2}\right)^2 \qquad \text{(C4)}$$

and the real Kerr solution can immediately be recovered from (C1)—(C4) by simply taking u and r real and $\eta=\bar{\zeta}$.

We will now transform the above solution into the l^* system.

By substituting (C1) into the definitions of the various quantities involved, we find that the tetrad transformation relating the two systems is given by (5.13) with

$$\mathring{\phi}=1,$$

$$Q=\frac{-4ia\zeta}{(1+\zeta\eta)}\frac{1}{\{r+2ia[(1-\zeta\eta)/(1+\zeta\eta)\}}, \qquad \text{(C5)}$$

and

$$V^*=1,$$

so that

$$S=(1+\zeta^*\eta^*)/(1+\zeta\eta).$$

Similarly, after making use of (5.12), the coordinate transformation (5.14) becomes

$$\begin{aligned} u^*&=\phi=u-ia[(1-\zeta\eta)/(1+\zeta\eta)],\\ r^*&=r+2ia[(1-\zeta\eta)/(1+\zeta\eta)],\\ \zeta^*&=\zeta[(r-2ia)/(r+2ia)],\\ \eta^*&=\eta, \end{aligned} \qquad \text{(C6)}$$

or, in the form more useful to our purposes,

$$\begin{aligned} u&=u^*+\tfrac{1}{4}[r^*-R],\\ r&=\tfrac{1}{2}[r^*+R],\\ \zeta&=\zeta^*[r^*+R+4ia]/[r^*+R-4ia]\\ \eta&=\eta^*, \end{aligned} \qquad \text{(C7)}$$

where

$$R\equiv\left(r^{*2}-16a^2-8iar^*\frac{(1-\zeta^*\eta^*)}{(1+\zeta^*\eta^*)}\right)^{1/2}\equiv r-2ia\frac{(1-\zeta\eta)}{(1+\zeta\eta)}.$$

Performing these transformations on the tetrad (C3), we obtain

$$l^*=du^*,$$

$$m^*=r^*d\eta^*/(1+\zeta^*\eta^*),$$

$$\bar{m}^*=\frac{r^*}{(1+\zeta^*\eta^*)}\left(d\zeta^*-\frac{128ma^2\zeta^{*2}r^*(r^*+R)}{(1+\zeta^*\eta^*)^2R[(r^*+R)^2+16a^2]}d\eta^*\right), \qquad \text{(C8)}$$

$$n^*=\left(1-\frac{m(r^*+R)}{2r^*R}\right)du^*+dr^*+\frac{8ima\zeta^*r^*(r^*+R)}{(1+\zeta^*\eta^*)^2R[r^*+R)^2+16a^2]}d\eta^*,$$

so that the complexified Kerr metric in the * system becomes

$$ds^2=2\left(du^{*2}+2du^*dr^*-\frac{r^{*2}d\zeta^*d\eta^*}{(1+\zeta^*\eta^*)^2}\right)-\frac{m(r^*+R)}{r^*R}\left(du^*-\frac{16ia\zeta^*r^{*2}}{[(r^*+R)^2+16a^2]}\frac{d\eta^*}{(1+\zeta^*\eta^*)^2}\right)^2 \qquad \text{(C9)}$$

and the only nonvanishing component of the Weyl tensor is

$$\psi_2^*=-m/r^{*3}. \qquad \text{(C10)}$$

*Research supported in part by the National Science Foundation and the U.S. Atomic Energy Commission.
†Present address.
‡Research supported by the National Science Foundation under Grant No. GP-35773.

[1]E. T. Newman, J. Math. Phys. 15, 44 (1974).
[2]I. Robinson and A. Trautman, Proc. Roy. Soc. (London) 265, 463 (1962).
[3]E. T. Newman, J. Math. Phys. 14, 102 (1973).
[4]G. C. Debney, R. P. Kerr, and A. Schild, J. Math. Phys. 10, 1842 (1969).
[5]B. Aronson, R. Lind, J. Messmer, and E.T. Newman, J. Math. Phys. 12, 2462 (1971).
[6]E. T. Newman and R. Penrose, J. Math. Phys. 3, 566 (1962).
[7]R. Lind Gen. Relativity GravitationJ. 5, 25 (1974).
[8]J. N. Goldberg and R.K. Sachs, Acta Phys. Polon. Suppl. 22, 13 (1962).
[9]R. Penrose, J. Math. Phys. 8, 345 (1967).

A curiosity concerning angular momentum*

Ezra T. Newman and Jeffrey Winicour

Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania 15260
(Received 6 February 1974)

We show that for classical, relativistic particles (or systems of noninteracting particles) one can interpret the intrinsic (or spin) angular momentum as arising from a center of mass world line displaced from the real Minkowski space into complex Minkowski space. This interpretation is then connected with the Penrose theory of twistors. It is further shown that if the complex center of charge coincides with the complex center of mass, the resulting particle has the Dirac value of the gyromagnetic ratio, i.e., (e/mc). For massless particles, there is no unique complex line about which the angular momentum vanishes—instead, there is a complex 2-surface, which can be considered to be a Penrose twistor.

I. INTRODUCTION

It is the purpose of this note to point out the unusual behavior of classical, relativistic angular momentum when Minkowski space is viewed as a (real) subspace of *complex* Minkowski space.[1]

We consider a classical particle (or system of noninteracting particles) which is characterized by a conserved momentum, P^a, and angular momentum M^{ab}. The values of M^{ab} depends on the choice of origin. Under the origin displacement X^a, one obtains

$$P'^a = P^a, \quad M'^{ab} = M^{ab} - 2X^{[a}P^{b]}. \tag{1}$$

It is well known[2] that with a proper choice of X^a the *orbital* part of the angular momentum can be made zero. [The spin part of M^{ab} is uneffected by (1).] In fact if P^a is timelike then a line (center of mass world line) can be found such that the orbital momentum is zero about any point on the line. In the case of a null P^a ($P^aP_a = 0$) a unique center of mass line cannot, in general, be found.

If, however, we now permit *complex* translations (i.e., consider the complexification of Minkowski space and allow the origin to be complex) then not only can we always find a unique complex center of mass line in the massive case, but we also discover that the *total* angular momentum is zero about each point of the line. The spin angular momentum can thus be conceived as arising, like orbital momentum, from a translation away from the center of mass but now an imaginary one. In the massless (null momentum) case, there does not exist a unique complex line about which the total angular momentum vanishes. Instead there is a totally null complex 2-surface for which the angular momentum vanishes about each point on the surface. This surface is equivalent to a Penrose twistor.

In Sec. II the ideas concerning the (real) center of mass world lines are reviewed while Sec. III is devoted to the complex center of mass lines. In Sec. IV we point out the connection between the complex center of mass planes and the Penrose theory of twistors. Finally in Sec. V we show that one can define a complex center of charge for the system. If the centers of charge and mass coincide then one obtains the Dirac value of the gyromagnetic ratio.

II. CENTER OF MASS

The material of this chapter is essentially a restatement of the work of Penrose and MacCallum.[2]

We consider separately the two cases $P^aP_a > 0$ and $P^aP_a = 0$.

A. $P^aP_a > 0$

In this case the angular momentum can be decomposed into orbital and spin parts

$$M^{ab} = L^{ab} + S^{ab}, \tag{2}$$

the orbital part being given by

$$L^{ab} = 2A^{[a}P^{b]}(P^cP_c)^{-1}, \tag{3}$$

with

$$A^a = M^{ab}P_b, \quad P^aA_a = 0; \tag{4}$$

and the spin part by

$$S^{ab} = \eta^{abcd}S_cP_d(P^eP_e)^{-1} \tag{5}$$

with

$$S_a = \overset{*}{M}_{ab}P^b = -\tfrac{1}{2}\eta_{abcd}M^{cd}P^b, \quad P^aS_a = 0, \tag{6}$$

with η_{abcd} the totally skew tensor defined by $\eta_{0123} = (-g)^{1/2}$, $g = \det g_{ab}$. S_a which is referred to as the spin vector is unchanged by (1).

From (1) and (4) one sees immediately that

$$0 = A'^a = M'^{ab}P_b = M^{ab}P_b - X^aP^bP_b + P^aX^bP_b \tag{7}$$

is the equation for the center of mass, i.e., around the new origin one has $L'^{ab} = 2A'^{[a}P^{b]} = 0$. The general solution for the center of mass line is thus

$$X^a = A^a(P^bP_b)^{-1} + \lambda P^a \tag{8}$$

with λ an arbitrary parameter along the world line.

B. $P^aP_a = 0$

This case is slightly more complicated than case A. We assume that

$$M^{ab}P_b = lP^a \tag{9}$$

from which it follows that

$$S^a \equiv \overset{*}{M}{}^{ab}P_b = sP^a. \tag{10}$$

This assumption is necessary in the theory of representations of the Poincaré group in order to avoid the continuous spin representations. From our point of view it is necessary in order to find points about which $M'^{ab}P_b = 0$.

From (1) we then have

$$0 = M'^{ab}P_b = M^{ab}P_b + P^aX^bP_b \tag{11}$$

or

$$X^a P_a = -l \tag{12}$$

as the condition for the displacement to the center of mass. As (12) is the equation for a null hypersurface, one obtains the result for null momentum that the center of mass is a three-dimensional region.

In the special case when $s=0$ (or $S^a=0$) one can show that

$$M^{ab} = 2A^{[a}P^{b]} \tag{13}$$

for some vector A^a. Then from (1)

$$0 = M'^{ab} = 2A^{[a}P^{b]} - 2X^{[a}P^{b]}$$

and one does obtain a unique center of mass line

$$X^a = A^a + \lambda P^a. \tag{14}$$

However when $s \neq 0$ there is no (real) means of localizing the null hypersurface (12) to a center of mass line.

III. COMPLEX CENTER OF MASS

Again we consider the two cases $P^aP_a > 0$ and $P^aP_a = 0$ separately.

A. $P^aP_a > 0$

If we define the self-dual, angular momentum tensor

$$W^{ab} = M^{ab} + i\overset{*}{M}{}^{ab}, \tag{15}$$

then from (1) we have

$$W'^{ab} = W^{ab} - 2X^{[a}P^{b]} - 2iX^{[a}\overset{*}{P}{}^{b]}. \tag{16}$$

Now let $X^a = Z^a$ be complex and demand that

$$W'^{ab}P_b = 0 = W^{ab}P_b - Z^aP^bP_b + P^aZ^bP_b \tag{17}$$

or

$$Z^a = W^{ab}P_b(P^eP_e)^{-1} + \lambda P^a \tag{18}$$

as the equation for the complex center of mass.

Using the notation of (2)—(6), one easily checks that

$$Z^a = (A^a + iS^a)(P^bP_b)^{-1} + \lambda P^a. \tag{19}$$

If (19) is substituted into (16), then

$$W'^{ab} = 0, \tag{20}$$

i.e., around the complex world line the total angular momentum vanishes.

B. $P^aP_a = 0$

There is first the trivial case when P^a is a degenerate (double) eigenvector of M^{ab} to be considered. One then has that

$$M^{ab} = 2A^{[a}P^{b]}, \quad A^aP_a = 0 \tag{21}$$

and

$$Z^a = X^a = A^a + \lambda P^a \tag{22}$$

gives the center of mass line. This is the $s=0$ case of Sec. II.

In the nondegenerate case M^{ab} has two real null eigenvectors P^a and N^a; N^a is normalized by

$$N^aP_a = 1. \tag{23}$$

One can show[3] that M^{ab} has the canonical form

$$M^{ab} = 2lP^{[a}N^{b]} - 2sP^{[a}\overset{*}{N}{}^{b]} \tag{24}$$

and therefore

$$W^{ab} = 2\phi(P^{[a}N^{b]} + iP^{[a}\overset{*}{N}{}^{b]}), \quad \phi = l + is. \tag{25}$$

It immediately follows from

$$W'^{ab} = 0 = W^{ab} - 2Z^{[a}P^{b]} - 2iZ^{[a}\overset{*}{P}{}^{b]} \tag{26}$$

that

$$Z^a = -\phi N^a + \lambda P^a + \nu \bar{M}^a, \tag{27}$$

where ν and λ are arbitrary and $\bar{M}^a$ is a complex null eigenvector of M^{ab} with eigenvalue is. Thus we do not obtain a complex null line but instead a complex plane which is totally null, i.e., $\Delta Z^a \Delta Z_a = 0$ with $\Delta Z^a = P^a\Delta\lambda + \bar{M}^a\Delta\nu$

IV. TWISTORS

If (27) is written in spinor form, i.e.,

$$Z^{AA'} = -\phi N^{AA'} + \lambda\bar{\pi}^A\pi^{A'} + \nu\iota^A\pi^{A'} \quad \text{with } \iota^A\bar{\pi}_A = 1$$

or

$$Z^{AA'} = -\phi N^{AA'} + \xi^A\pi^{A'} \qquad \text{with } \xi^A = \lambda\bar{\pi}^A + \nu\iota^A, \tag{28}$$

then one can immediately define the associated twistor[2]

$$Z^\mu = (\mu^A, \pi_{A'}) \tag{29}$$

with

$$\mu^A = -i\phi N^{AA'}\pi_{A'}. \tag{30}$$

From the normalization $N^aP_a = 1$, one also has

$$\mu^A\bar{\pi}_A = -i\phi \tag{31}$$

and thus the twistor scalar product is

$$Z^\mu\bar{Z}_\mu = \mu^A\bar{\pi}_A + \bar{\mu}^{A'}\pi_{A'} = i(\bar{\phi} - \phi) = 2s, \tag{32}$$

A twistor can thus be thought of as the complex center of mass line "null plane" of a null momentum particle of spin s.

V. ELECTRIC AND MAGNETIC DIPOLE TENSOR

If the classical particle (with $\dot{P}^a = 0$, $P^a = mcV^a$, $V^aV_a = 1$) has a charge $e \neq 0$ and a magnetic moment, the dipole tensor D^{ab} can be defined by

$$D^{ab} = 2D^{[a}V^{b]} + \eta^{abcd}M_cV_d, \tag{33}$$

where D^a and M^a are respectively the electric and magnetic dipole moments with $D^aV_a = M^aV_a = 0$. (Alternatively, one could begin with a skew tensor D^{ab} and define $D^a \equiv D^{ab}V_b$ and $M^a = \overset{*}{D}{}^{ab}V_b$.) Under a shift in origin, X^a, (since $D'^a = D^a - eX^a$) it follows that

$$D'^{ab} = D^{ab} - 2eX^{[a}V^{b]} \tag{34}$$

and an analogy is established between D^{ab} and M^{ab}. By repeating the arguments of Secs. II and III the (real) center of charge line and complex center of charge line can be defined by

$$X^a = e^{-1}D^a + \lambda P^a \tag{35}$$

and

$$Z^a = e^{-1}(D^a + iM^a) + \lambda P^a . \tag{36}$$

A further curiosity arises if D^{ab} is proportional to M^{ab}, i.e., if the complex center of mass line coincides with the complex center of charge line. It then follows from (19) and (36) that

$$e^{-1}(D^a + iM^a) = (A^a + iS^a)(P^b P_b)^{-1} \tag{37}$$

or that the real centers of mass and charge coincide, i.e., $e^{-1}D^a = A^a(P^b P_b)^{-1}$, and moreover

$$M^a = (e/mc)(S^a/mc). \tag{38}$$

In terms of usual dimensions, S^a/mc corresponds to the spin-angular momentum vector. We thus see, from (38), that one obtains the Dirac value of the gyromagnetic ratio (e/mc) when the complex center of charge and mass lines coincide.[4]

*Supported by the National Science Foundation, Grant No. GP-35773.

[1]E.T. Newman and Jeffrey Winicour, J. Math. Phys. 15, 426 (1974).
[2]R. Penrose and M. MacCallum, Phys. Rep. 6C, 243 (1973).
[3]J.L. Synge, *Relativity, The Special Theory* (North-Holland, Amsterdam, 1958).
[4]For a related, but different, way of obtaining this result, see E.T. Newman, J. Math. Phys. 14, 102 (1973).

Relativistic Grad polynomials*

James L. Anderson

Department of Physics, Stevens Institute of Technology, Hoboken, New Jersey 07030
(Received 31 October 1973; revised manuscript received 6 March 1974)

A new method for constructing the relativistic Grad polynomials employed in relativistic kinetic theory is presented. It simplifies considerably the labor involved in their construction compared to previous methods and at the same time exhibits orthogonality properties between the components of the polynomials that were not apparent from the forms obtained by these earlier methods.

INTRODUCTION

In attempting to obtain approximate solutions of the relativistic Boltzmann equation, a number of authors[1] have made use of a family of orthogonal polynomials defined on 4-momentum space analogous to the Hermite—Grad polynomials[2] employed in constructing approximate solutions to the classical Boltzmann equation. In particular, these polynomials, like their classical counterparts, have been used most extensively in conjunction with the Grad moment method to calculate transport coefficients for a relativistic gas. However, although it has been shown[3] that these polynomials reduce to the Hermite—Grad polynomials in the classical limit, they have not been shown to possess, in their relativistic form, a number of properties possessed by the Hermite—Grad polynomials. Furthermore, while there exists a well-defined algorithm for their construction, it is extremely tedious to apply in practice, and hence explicit expressions for only the first few polynomials appear in the literature. In this paper we will give an alternate prescription for constructing these polynomials which is much easier to apply and which, at the same time, leads to closed form expressions for the polynomials that are completely analogous to those for the Hermite—Grad polynomials.

1. COVARIANT SPHERICAL HARMONICS

As a first step in our construction of a family of orthogonal polynomials of the 4-momentum p^μ of a particle of mass m, we introduce a family of polynomials that are the covariant form of the generalized spherical functions of three dimensions. For this purpose we assume that there exists a timelike unit vector field $u^\mu(x)$ in the region of space—time occupied by a gas of particles of mass m, which we take to be the local 4-velocity of the gas. With its help we decompose p^μ at a point according to

$$p^\mu = Eu^\mu + pl^\mu, \tag{1.1}$$

where l^μ is a unit spacelike vector orthogonal to u^μ:

$$l^\mu l_\mu = -1, \quad l^\mu u_\mu = 0. \tag{1.2}$$

(Indices are raised and lowered with the space—time metric $g_{\mu\nu}$, which we take to have a signature $+2$.) Since $u^\mu u_\mu = 1$ and $p^\mu p_\mu = m^2$, it follows that

$$E = u_\mu p^\mu \tag{1.3}$$

and

$$p = (E^2 - m^2)^{1/2}. \tag{1.4}$$

Also

$$l^\mu = (1/p)\, h^\mu{}_\nu p^\nu, \tag{1.5}$$

where

$$h^\mu{}_\nu \equiv u^\mu u_\nu - \delta^\mu{}_\nu \tag{1.6}$$

is the projection operator onto the hyperplane orthogonal to u^μ.

The desired polynomials are now constructed with the help of the generating function $1/l$, where $l = (h_{\mu\nu} l^\mu l^\nu)^{1/2}$. We first define the quantities

$$X^{\mu_1\mu_2\cdots\mu_n} \equiv \frac{1}{n!}\, l^{n+1} \frac{\partial^n}{\partial l_{\mu_1}\partial l_{\mu_2}\cdots\partial l_{\mu_n}} \frac{1}{l}. \tag{1.7}$$

The spherical function $Y^{\mu_1\cdots\mu_n}$ is then the value of $X^{\mu_1\cdots\mu_n}$ on the unit sphere $l=1$:

$$Y^{\mu_1\cdots\mu_n} = X^{\mu_1\cdots\mu_n}\big|_{l=1}. \tag{1.8}$$

From its definition, $Y^{\mu_1\cdots\mu_n}$ is seen to be a polynomial of the components l^μ of order n. More specifically it follows from Eqs. (1.7) and (1.8) that

$$Y^{\mu_1\cdots\mu_{2n}} = \sum_{m=0}^{n} A^e_{mn}\, l^{(\mu_1}\cdots l^{\mu_{2m}} h^{\mu_{2m+1}\mu_{2m+2}}\cdots h^{\mu_{2n-1}\mu_{2n})}, \tag{1.9a}$$

where

$$A^e_{mn} = (-1)^{n-m} \frac{1}{2^{2n}} \frac{(2n+2m)!}{(2m)!\,(n+m)!\,(n-m)!},$$

$$Y^{\mu_1\cdots\mu_{2n+1}} = \sum_{m=0}^{n} A^o_{mn}\, l^{(\mu_1}\cdots l^{\mu_{2m+1}} h^{\mu_{2m+2}\mu_{2m+3}}\cdots h^{\mu_{2n}\mu_{2n+1})}, \tag{1.9b}$$

where

$$A^e_{mn} = (-1)^{n-m} \frac{1}{2^{2n}} \frac{(2n+2m)!}{(2m+1)!\,(n+m)!\,(n-m)!},$$

and where round brackets around a set of indices indicates the symmetrization of the terms enclosed, e.g., $A^{(\mu}B^\nu C^{\rho)} = (1/3!)(A^\mu B^\nu C^\rho + A^\mu B^\rho C^\nu + A^\nu B^\rho C^\mu + A^\nu B^\rho C^\mu + A^\rho B^\mu C^\nu + A^\rho B^\nu C^\mu)$. Thus we have, for example,

$$Y = 1, \quad Y^\mu = l^\mu, \quad Y^{\mu\nu} = \tfrac{1}{2}(3 l^\mu l^\nu - h^{\mu\nu}),$$

$$Y^{\mu\nu\rho} = \tfrac{1}{6}[15\, l^\mu l^\nu l^\rho - 3(h^{\mu\nu} l^\rho + h^{\rho\mu} l^\nu + h^{\nu\rho} l^\mu)], \quad \text{etc.}$$

The above expressions for A^e_{mn} and A^o_{mn} are most easily obtained by noting that, in the frame in which $u^\mu = (1,0,0,0)$, $Y^{\mu_1\cdots\mu_n}|_{\mu_1=\mu_2=\cdots=\mu_n=3} = P_n$, where P_n is the nth order Legendre polynomial, and by using the expansion for $P_n(z)$ in powers of z.

We see from its definition that $Y^{\mu_1\cdots\mu_n}$ is completely symmetric with respect to interchange of its indices. It also follows from Eqs. (1.2) and (1.3) that

$$Y^{\mu_1\mu_2\cdots\mu_n} u_{\mu_1} = 0, \quad Y^{\mu_1\mu_2\cdots\mu_n} g_{\mu_1\mu_2} = 0. \tag{1.10}$$

There are, therefore, $2n+1$ independent polynomials of order n corresponding to the $2n+1$ spherical harmonics of nth order. Finally, from the definition of $X^{\mu_1\cdots\mu_n}$ one can show that

$$l^2\nabla^2 X^{\mu_1\cdots\mu_n}=n(n+1)X^{\mu_1\cdots\mu_n}, \tag{1.11}$$

where $\nabla^2\equiv -h^{\mu\nu}\partial^2/\partial l^\mu\partial l^\nu$. It therefore follows that

$$\int_S\int Y^{\mu_1\cdots\mu_n}Y^{\nu_1\cdots\nu_m}\,d\Omega=0, \quad n\neq m, \tag{1.12}$$

where the integration is over the unit sphere $l^2=1$. (The integral appearing here will be defined more precisely in the next section.) The polynomials of different order are thus seen to be orthogonal to each other on the unit sphere.

The properties of $Y^{\mu_1\cdots\mu_n}$ lead to the conclusion that it is a linear combination of spherical harmonics $Y_n^m(\theta,\phi)=P_n^m(\cos\theta)\exp(im\phi)$, where $P_n^m(x)$ is an associated Legendre polynomial. To determine this linear combination, we introduce an orthogonal tetrad n_i^μ $(i=1,2,3)$ orthogonal to u^μ so that

$$u_\mu n_i^\mu=0, \quad g_{\mu\nu}n_i^\mu n_j^\nu=-\delta_{ij}. \tag{1.13}$$

Then l^μ can be expressed in terms of n_i^μ and the angles θ and ϕ as

$$l^\mu=\sin\theta\cos\phi\, n_1{}^\mu+\sin\theta\sin\phi\, n_2{}^\mu+\cos\theta\, n_3{}^\mu. \tag{1.14}$$

We next define the complex unit vectors n_+^μ and n_-^μ by

$$n_\pm^\mu=2^{-1/2}(n_1{}^\mu\pm i n_2{}^\mu)$$

to obtain the relations

$$n_-{}^{\mu_1}\cdots n_-{}^{\mu_l}n_+{}^{\mu_{l+1}}\cdots n_+{}^{\mu_{2l+m}}n_3{}^{\mu_{2l+m+1}}\cdots n_3{}^{\mu_n}Y_{\mu_1\cdots\mu_n}=\left(\frac{1}{\sqrt2}\right)^{m+2l}\frac{(n-m)!}{n!}\,Y_n{}^m(\theta,\phi). \tag{1.15}$$

These relations are most easily verified by transforming to the local rest frame where $u^\mu=(1,0,0,0)$, taking $n_i^\mu=\delta_i^\mu$ and making use of the known[4] relations between $Y_n{}^m(\theta,\phi)$ and the derivatives of $1/r$. It then follows that

$$Y^{\mu_1\cdots\mu_n}=\sum_{m=-n}^{n}\sum_l n_+{}^{(\mu_1}\cdots n_+{}^{\mu_l}\times n_-{}^{\mu_{l+1}}\cdots n_-{}^{\mu_{2l+m}}n_3{}^{\mu_{2l+m+1}}\cdots n_3{}^{\mu_n)}\times(-1)^k\left(\frac{1}{\sqrt2}\right)^{2l+m}\frac{1}{l!(l+m)!}\,Y_n{}^m(\theta,\phi), \tag{1.16}$$

where

$$k=l,\quad m\geq 0,$$
$$=l+m,\quad m<0,$$

and where the sum on l is over all positive integer values of l that satisfy the conditions

$$l+m\geq 0,\quad 2l+m\leq n.$$

Finally we will need the expansion of products of the l's in terms of the Y's. They are given by

$$l^{\mu_1}\cdots l^{\mu_n}=\sum_{m=0}^{n}B_{mn}^e\,Y^{(\mu_1\cdots\mu_{2m}}h^{\mu_{2m+1}\mu_{2m+2}}\cdots h^{\mu_{2n-1}\mu_{2n})} \tag{1.17a}$$

where

$$B_{mn}^e=\frac{(2m+1)\,2n\,(2n-2)\cdots(2n-2m+2)}{(2n+1)(2n+3)\cdots(2n+2m+1)},$$

$$l^{\mu_1}\cdots l^{\mu_{2n+1}}=\sum_{m=0}^{n}B_{mn}^o\,Y^{(\mu_1\cdots\mu_{2m+1}}h^{\mu_{2m+2}\mu_{2m+3}}\cdots h^{\mu_{2n}\mu_{2n+1})}, \tag{1.17b}$$

where

$$B_{mn}^o=\frac{(2m+1)\,2n\,(2n-2)\cdots(2n-2m+2)}{(2n+3)(2n+5)\cdots(2n+2m+1)}.$$

Again, the above expressions for B_{mn}^e and B_{mn}^o are most easily derived in the frame in which $u^\mu=(1,0,0,0)$ using the expansion of z^n in terms of P_m.

2. REALTIVISTIC GRAD POLYNOMIALS

The relativistic generalization of the Hermite–Grad polynomials as developed by Marle and Anderson and Stewart are a set of polynomials $H^{\mu_1\cdots\mu_n}$, $n=0,1,2,\ldots,\infty$ with the following two defining properties:

(a) $H^{\mu_1\cdots\mu_n}=p^{\mu_1}\cdots p^{\mu_n}+$ terms of lower order in the p's, (2.1)

(b) $\int H^{\mu_1\cdots\mu_n}H^{\nu_1\cdots\nu_m}w(x,p)\pi=0, \quad m\neq n,$ (2.2)

where $w(x,p)$ is a weight function to be specified,

$$\pi=\sqrt{-g}\;\delta_+(p_\mu p^\mu-m^2)\,d^4p \tag{2.3}$$

is the relativistic volume element in momentum space, and δ_+ is the positive-frequency δ function defined by

$$\delta_+(x^2-a^2)=(1/x)\,\delta(x-a). \tag{2.4}$$

In addition to these properties $H^{\mu_1\cdots\mu_n}$ can be shown to be completely symmetric under interchange of its indicies and to satisfy

$$g_{\mu_1\mu_2}H^{\mu_1\cdots\mu_n}=0. \tag{2.5}$$

The weight function w appearing in Eq. (2.2) is required to vanish sufficiently rapidly as $t_\mu p^\mu\to\infty$ for any timelike vector t^μ that all of the integrals used to construct the H's exist but is otherwise arbitrary. In practice w is usually taken to be a local equilibrium distribution for the gas that is isotropic about its local 4-velocity u^μ, that is, a function only of position and $E=u_\mu p^\mu$. In what follows we shall consider only such isotropic weight functions.

Marle has shown that the two properties (a) and (b) are sufficient to determine completely and uniquely the full set of H's. The method used to date to construct the H's amounts to essentially a Schmidt orthogonalization process and although in principle it is possible to proceed in this fashion to obtain the components of $H^{\mu_1\cdots\mu_n}$ for arbitrary large n, the fact that the H's are functions of four variables makes the labor involved prohibitive for all but the first few polynomials. In this section we shall give a new method of constructing these polynomials in terms of their expansions in terms of the spherical functions $Y^{\mu_1\cdots\mu_n}$ developed in the preceeding section that greatly reduces the labor involved in their construction and at the same time exhibits the orthogonality properties of polynomials of the same order which is lacking in the previous construction.

It is clear that if the weight function appearing in Eq. (2.2) is a function only of E and position, $H^{\mu_1\cdots\mu_n}$ will be a linear combination of nth order and lower spherical polynomials with coefficients that are also only functions of E and position. The problem then is to determine these coefficients in such a way that Eqs. (2.1) and (2.2) are satisfied. The uniqueness of the polynomials satisfying these conditions will then assure that these coefficients are unique.

We begin our construction by noting that any product of p's can be written as a unique linear combination of spherical polynomials. The coefficients are determined by substituting for each p^μ its decomposition (1.1) and then expressing the various products l's in terms of Y's by means of Eqs. (1.17). In this way we obtain

$$p^{\mu_1}\cdots p^{\mu_{2n}}=\sum_{m=0}^{n} E^{2n-2m}p^{2m}\frac{(2n)!}{(2m)!(2n-2m)!}\sum_{p=0}^{m} B^e_{pm}Y^{(\mu_1\cdots\mu_{2p}}h^{\mu_{2p+1}\mu_{2p+2}}\cdots h^{\mu_{2m-1}\mu_{2m}}u^{\mu_{2m+1}}\cdots u^{\mu_{2n})}$$
$$+\sum_{m=0}^{n-1} E^{2n-2m-1}p^{2m+1}\frac{(2n)!}{(2m+1)!(2n-2m-1)!}\times\sum_{p=0}^{m} B^o_{pm}Y^{(\mu_1\cdots\mu_{2p+1}}h^{\mu_{2p+2}\mu_{2p+3}}\cdots h^{\mu_{2m}\mu_{2m+1}}u^{\mu_{2m+2}}\cdots u^{\mu_{2n})} \quad (2.6a)$$

and

$$p^{\mu_1}\cdots p^{\mu_{2n+1}}=\sum_{m=0}^{n} E^{2n-2m+1}p^{2m}\frac{(2n+1)!}{(2m)!(2n-2m+1)!}\times\sum_{p=0}^{m} B^e_{pm}Y^{(\mu_1\cdots\mu_{2p}}h^{\mu_{2p+1}\mu_{2p+2}}\cdots h^{\mu_{2m-1}\mu_{2m}}u^{\mu_{2m+1}}\cdots u^{\mu_{2n+1})}$$
$$+\sum_{m=0}^{n} E^{2n-2m}p^{2m+1}\frac{(2n+1)!}{(2m+1)!(2n-2m)!}\times\sum_{p=0}^{m} B^o_{pm}Y^{\mu_1\cdots\mu_{2p+1}}h^{\mu_{2p+2}\mu_{2p+3}}\cdots h^{\mu_{2m}\mu_{2m+1}}u^{\mu_{2m+2}}\cdots u^{\mu_{2n+1})}. \quad (2.6b)$$

We next replace each factor of p multiplying a Y in the above sums in excess of the order of this Y by its value $(E^2-m^2)^{1/2}$. It will be seen that in all cases the number of such factors is even. This fact allows us to rewrite the above equations in the form

$$p^{\mu_1}\cdots p^{\mu_{2n}}=\sum_{q=0}^{n} E^{2n-2q}p^{2q}T^{\mu_1\cdots\mu_{2n}}_{qe}+\sum_{q=0}^{n-1} E^{2n-2q-1}p^{2q+1}T^{\mu_1\cdots\mu_{2n}}_{qo}+\text{terms in } m \quad (2.7a)$$

and

$$p^{\mu_1}\cdots p^{\mu_{2n+1}}=\sum_{q=0}^{n} E^{2n-2q+1}p^{2q}S^{\mu_1\cdots\mu_{2n+1}}_{qe}+\sum_{q=0}^{n} E^{2n-2q}p^{2q+1}S^{\mu_1\cdots\mu_{2n+1}}_{qo}+\text{terms in } m \quad (2.7b)$$

where

$$T^{\mu_1\cdots\mu_{2n}}_{qe}=\sum_{r=q}^{n}\frac{(2n)!}{(2r)!(2n-2r)!}\times B^e_{qr}Y^{(\mu_1\cdots\mu_{2q}}h^{\mu_{2q+1}\mu_{2q+2}}\cdots h^{\mu_{2r-1}\mu_{2r}}u^{\mu_{2r+1}}\cdots u^{\mu_{2n})}.$$

$$T^{\mu_1\cdots\mu_{2n}}_{qo}=\sum_{r=q}^{n-1}\frac{(2n)!}{(2r)!(2n-2r-1)!}\times B^o_{qr}Y^{(\mu_1\cdots\mu_{2q+1}}h^{\mu_{2q+2}\mu_{2q+3}}\cdots h^{\mu_{2r}\mu_{2r+1}}u^{\mu_{2r+2}}\cdots u^{\mu_{2n})}.$$

$$S^{\mu_1\cdots\mu_{2n+1}}_{qe}=\sum_{r=q}^{n}\frac{(2n+1)!}{(2r)!(2n-2r+1)!}\times B^e_{qr}Y^{(\mu_1\cdots\mu_{2q}}h^{\mu_{2q+1}\mu_{2q+2}}\cdots h^{\mu_{2r-1}\mu_{2r}}u^{\mu_{2r+1}}\cdots u^{\mu_{2n+1})},$$

and

$$S^{\mu_1\cdots\mu_{2n+1}}_{qo}=\sum_{r=q}^{n}\frac{(2n+1)!}{(2r+1)!(2n-2r)!}\times B^o_{qr}Y^{(\mu_1\cdots\mu_{2q+1}}h^{\mu_{2q+2}\mu_{2q+3}}\cdots h^{\mu_{2r}\mu_{2r+1}}u^{\mu_{2r+2}}\cdots u^{\mu_{2n+1})}.$$

In these equations "terms in m" indicate terms which contain at least one power of m^2 and hence at least two powers of p less than does the product of the p's on the left side of the equation in which they appear. If now we contract both sides of one of the above equations over any two free indices, we get a factor of m^2 times two powers of p less than before the contraction on the left side. It follows therefore that the contraction of any term in the sums appearing in Eqs. (2.7) over any two indices must vanish since these terms contain no factors of m.

We now assert that the expression for $H^{\mu_1\cdots\mu_n}$ is equal to the sum of the terms in the expression for $p^{\mu_1}\cdots p^{\mu_n}$ which contain no factors of m with each power E^m in the sum replaced by a polynomial in E of the form $E^m+a_{m-1}E^{m-1}+\cdots+a_0$. The resulting expression for $H^{\mu_1\cdots\mu_n}$ clearly satisfies Eqs. (2.1) and (2.5). Furthermore, there are just enough coefficients a_k in this expression to satisfy Eqs. (2.2). There are in fact $n(n+1)/2$ such coefficients in the expression for $H^{\mu_1\cdots\mu_n}$ and, because of the orthogonality conditions (1.12) which the Y's satisfy, the orthogonality conditions (2.2) which must be satisfied by $H^{\mu_1\cdots\mu_n}$ reduce to this same number of linear equations in the a's.

In order to find the equations for the a's, we must reexpress the volume element π given by Eq. (2.3) in terms of the variables E, p, and l^μ or E, p, θ, and ϕ in order to make use of the orthogonality of spherical functions of different orders. It is not difficult to see that

$$\pi = \sqrt{-g}\,\delta_+(E^2 - p^2 - m^2)\,p^2\,dEdpd\Omega, \tag{2.8}$$

where $d\Omega = \sin\theta d\theta d\phi$ when the Y's are expressed as functions of θ and ϕ by means of Eqs. (1.16) or $d\Omega = \delta(u_\mu l^\mu)\delta_+(l^2 - 1)\,d^4l$ when they are expressed as functions of the l^μ by means of Eqs. (1.9). The equations for the a's will then be seen to involve the integrals

$$I_{kn} \equiv \int\int E^n p^k w(E)\,\delta_+(E^2 - p^2 - m^2)\,p^2 dEdp$$
$$= m^{n+k+2}\int_0^\infty \cosh^n x \sinh^{k+2} x\; w(m\cosh x)\,dx. \tag{2.9}$$

If w is a local Maxwellian distribution $\exp(\alpha - \beta u_\mu p^\mu)$, these integrals can all be expressed in terms of modified Bessel functions of the second kind.

Consider now the expression for H^μ. From Eq. (2.7b) with $n = 0$ we obtain, according to our above prescription,

$$H^\mu = (E - a)\,u^\mu + pY^\mu. \tag{2.10}$$

The second term is already orthogonal to $H = 1$ because of the orthogonality of the Y's. The first term will also be orthogonal to H provided that

$$I_{01} - aI_{00} = 0. \tag{2.11}$$

The expression for $H^{\mu\nu}$ is gotten from the sum in Eq. (2.7a) with $n = 1$ and is of the form

$$H^{\mu\nu} = \tfrac{1}{3}(E^2 - a_1E - a_2)(4u^\mu u^\nu + g^{\mu\nu}) + (E - b_1)p(u^\mu Y^\nu + u^\nu Y^\mu) + \tfrac{2}{3}p^2 Y^{\mu\nu} \tag{2.12}$$

and will be orthogonal to H and H^μ if

$$\begin{aligned} &I_{02} - I_{01}a_1 - I_{00}a_2 = 0,\\ &I_{03} - I_{02}a_1 - I_{01}a_2 = 0,\\ &I_{21} - I_{20}b_1 = 0. \end{aligned} \tag{2.13}$$

More generally, if we designate by $\phi_{nk}(E)$ the polynomial in E which is a coefficient of $Y^{\nu_1\cdots\nu_k}$ in the expansion for $H^{\mu_1\cdots\mu_n}$, we have

$$\phi_{nk} = \frac{1}{G_{nk}}\begin{vmatrix} I_{2k,0} & I_{2k,1}\cdots & I_{2k,n-k} \\ I_{2k,1} & I_{2k,2}\cdots & I_{2k,n-k+1} \\ \cdots & \cdots & \cdots \\ I_{2k,n-k-1} & I_{2k,n-k}\cdots & I_{2k,2n-2k-1} \\ 1 & E & E^{n-k} \end{vmatrix}, \tag{2.14}$$

where G_{nk} is the minor of E^{n-k} in the determinant above.

The polynomials so obtained are seen to form families of orthogonal polynomials analogous to the associated Laguerre polynomials and satisfying the orthogonality conditions

$$\int_m^\infty w(E)(E^2 - m^2)^{(2k+1)/2}\,\phi_{nk}(E)\,\phi_{mk}(E)\,dE = 0, \quad n \neq n. \tag{2.15}$$

Finally then we have the desired expressions for the H's:

$$H^{\mu_1\cdots\mu_{2n}} = \sum_{q=0}^{n}\phi_{2n,2q}\,T_{qe}^{\mu_1\cdots\mu_{2n}} + \sum_{q=0}^{n-1}\phi_{2n,2q+1}\,T_{qo}^{\mu_1\cdots\mu_{2n}} \tag{2.16}$$

and

$$H^{\mu_1\cdots\mu_{2n+1}} = \sum_{q=0}^{n}\phi_{2n+1,2q}\,S_{qe}^{\mu_1\cdots\mu_{2n+1}} + \sum_{q=0}^{n}\phi_{2n+1,2q+1}\,S_{qo}^{\mu_1\cdots\mu_{2n+1}}, \tag{2.17}$$

where the T's and the S's are given by the expressions following Eq. (2.7b).

Unfortunately the polynomials $\phi_{nk}(E)$ introduced above are not equivalent to any of the classical orthogonal polynomials[5] and hence do not satisfy self-adjoint second-order differential equations nor do generalized Rodrigues' formulae exist for them. Nevertheless, Eq. (2.15) allows us to construct any one ϕ without a knowledge of the others and together with Eqs. (2.16) and (2.17) yield closed form expressions for the relativistic Hermite–Grad polynomials.

*Supported in part by National Science Foundation Grant No. GP-34621-1.

[1]C. Marle, Ann. Inst. H. Poincaré 10, 127 (1969); J.L. Anderson, "Relativistic Boltzmann Theory and the Grad Method of Moments" in *Relativity*, Proc. of the Midwest Conference on Relativity, edited by Carmeli, Witten, Fickler (Plenum, New York, 1971); J.M. Stewart, *Non-Equilibrium Relativistic Kinetic Theory*, Springer Lecture Notes in Physics No. 10 (Springer, New York, 1971)

[2]H. Grad, Comm. Pure Appl. Math. 2, 331 (1949).

[3]C. Marle, Ref. 1.

[4]P.M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), p. 1281.

[5]A. Erdelyi *et al.*, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. II, Chap. 10.

Spinor connections in general relativity

C. P. Luehr

Department of Mathematics, University of Florida, Gainesville, Florida 32601

M. Rosenbaum*

Centro de Estudios Nucleares, Universidad de México, México 20, D.F., Mexico
(Received 17 April 1973; revised manuscript received 23 October 1973)

An axiomatic foundation is provided for the theory of spinor calculus on the space–time manifold of general relativity. The methods, which deal directly with concepts in a coordinate free manner, allow not only elegant and compact statements of definitions and formulas but also have served as powerful analytical tools for the derivation of some interesting new results and for the unification and clarification of the previous work of other authors. The most general spinor connections are defined and related to the standard spinor connection (the unique spinor connection which is compatible with the spinor inner product and generates the Riemann 4-vector connection). By means of the intrinsic formalism presented here an interpretation is given to the spinor theory of Infeld and van der Waerden. The most general spinor curvature tensors are derived, and two alternate expressions result from two bispinor connections which satisfy the desirable requirement of producing the standard 4-vector connection by two different prescriptions. Another application of the techniques developed here results in an interesting expression for the spinor connection coefficients in terms of Dirac gamma matrices for four component spinors and arbitrary spinor connections which is much simpler and more general than others given in the literature.

I. INTRODUCTION

Modern differential geometry has provided powerful tools for tensor analysis which are based on abstract methods dealing directly with concepts in a coordinate free manner. The advanced level of development of the subject can be seen by consulting, for example, Hicks.[1] In particular, two areas elegantly treated by modern methods are: (1) algebra of tensors including the definition of tensors and algebraic operations on tensors, (2) calculus of tensors including the linear connection and covariant differentiation of tensors.

The old method of defining tensors in terms of the transformation properties of its components is replaced in a modern development by a direct definition as multilinear functionals. Transformation properties of components of a tensor is not a part of the definition of a tensor but is easily proved as a consequence of the definition. This conceptual approach to the definition of tensors as well as algebraic operations on tensors is simple and direct.

The old method of defining a linear connection in terms of components is replaced in a modern development by a direct method which first treats vectors axiomatically as derivation operators and then defines a linear connection as a map that associates with each derivation operator a covariant differentiation operator on vector fields such that a certain set of basic axioms is satisfied. The operation of covariant differentiation of tensor fields is next defined very simply. The discussion is coordinate free, but the expression of the results in terms of coordinates is easily derived.

Although the theory of spinors has been discussed extensively in the literature,[2] it has not reached the high level of modern development as the theory of tensors. As a first step in the development of an intrinsic spinor formalism, the algebra of two and four component spinors from an axiomatic viewpoint has been discussed in a previous paper[3] (hereafter referred to as I). On the other hand, the calculus of spinors has not been given an axiomatic foundation as has been done for tensor calculus in works of modern differential geometry. The present paper is intended to provide this foundation in the case of spinor calculus on the space—time manifold of general relativity. The formalism developed here has greatly aided in the derivation of some useful new results which are included in the paper. Also it unifies and clarifies the previous work of others on spinor calculus.

The presentation is divided into four parts. It begins with a discussion in Sec. II of the various spinor spaces in which the quantities used later in the paper are located. This section also serves the purpose of making the material in the remaining sections more self contained.

In Sec. III we consider connections and curvature in the general relativity manifold. It is intended as a summary of methods and results needed for the rest of the paper.

Section IV deals with spinor connections. First the most general connections are introduced for two- and four-component spinors, and their relation to the standard spinor connection (the unique spinor connection which is compatible with the spinor inner product and generates the standard four-vector connection) is established. By introducing separately the requirements of compatibility with the spinor inner product and of generating the standard four-vector connection, four theorems are proved to permit the discussion of the important cases which comprise the various spinor connections considered in the literature. Also the general spinor coefficients of connection are derived and special cases are considered.

The spinor theory of Infeld and van der Waerden[4] is interpreted from the point of view of our axiomatic approach, and the differences between the two formalisms are discussed. Bispinor connections are also discussed intrinsically, and an expression for the bispinor connection coefficients in terms of Dirac gamma matrices is derived using arbitrary spinor bases. To conclude this section, the Fock—Ivanenko coefficients[5] are obtained by a further particularization of these results.

Finally in Sec. V, we discuss spinor and bispinor curvature tensors, and show how they are related to the Riemann curvature tensor in the special cases of the spinor connections discussed in the previous section. These results are compared then with those given in the literature.[6]

II. SPINOR SPACES AND SPINOR ALGEBRA

In order to make the discussion of the following sections more self-contained, we review here spinors and spinor-tensors from the abstract point of view introduced in I. Although some of the material contained here has already been discussed in I, it is felt that a more complete and systematic presentation is desirable, particularly since the intrinsic notation that we employ, with its natural advantages over the standard component notation, is not commonly utilized.

At the end of this section we give a summary of the spinor and tensor spaces discussed in the following, as well as the special intrinsic notation required for later sections.

A. Spinor spaces $\mathcal{S}_2$ and $\overline{\mathcal{S}}_2$

$\mathcal{S}_2$ is a two-dimensional vector space over the field $\mathcal{C}$ of complex numbers with an antisymmetric inner product and with a conjugate spinor space $\overline{\mathcal{S}}_2$ associated with it.

The inner product in $\mathcal{S}_2$ associates a complex number $u\bullet v$ with every pair of spinors u and v in $\mathcal{S}_2$ in such a way that the following axioms are satisfied:

$$\begin{aligned}
&\text{(a) } u\blacktriangle v=-v\blacktriangle u \quad \text{(antisymmetry)},\\
&\text{(b) } (\alpha u+\beta w)\blacktriangle v=\alpha(u\blacktriangle v)+\beta(w\blacktriangle v)\\
&\qquad u\blacktriangle(\alpha v+\beta w)=\alpha(u\blacktriangle v)+\beta(u\blacktriangle w) \quad \text{(linearity)},\\
&\text{(c) } u\blacktriangle v=0 \text{ for fixed } u \text{ and all } v \text{ implies } u=0\\
&\qquad \text{(nondegeneracy)},
\end{aligned} \tag{1}$$

where $u,v,w\in\mathcal{S}_2$ and $\alpha,\beta\in\mathcal{C}$.

The conjugate spinor space $\overline{\mathcal{S}}_2$ is also a two-dimensional vector space over the field of complex numbers with an inner product satisfying axioms (a), (b), (c). The spaces $\mathcal{S}_2$ and $\overline{\mathcal{S}}_2$ are related by a map (conjugation operation) $u\in\mathcal{S}_2\to\bar{u}\in\overline{\mathcal{S}}_2$ and $w\in\overline{\mathcal{S}}_2\to\bar{w}\in\mathcal{S}_2$ such that

$$\begin{aligned}
&\text{(d) } \overline{\alpha u}=\bar{\alpha}\bar{u},\\
&\text{(e) } \overline{u+v}=\bar{u}+\bar{v},\\
&\text{(f) } \overline{u\blacktriangle v}=\bar{u}\blacktriangle\bar{v},\\
&\text{(g) } \bar{\bar{u}}=u
\end{aligned} \tag{2}$$

for all $u,v\in\mathcal{S}_2$ and $\alpha\in\mathcal{C}$,

where $\bar{\alpha}$ is the ordinary complex conjugate of the complex number α. It follows that the properties (d), (e), (f), (g) are also satisfied for u,v in $\overline{\mathcal{S}}_2$. We shall usually write elements of $\overline{\mathcal{S}}_2$ with a bar, e.g., $\bar{u}$, expressing it as the conjugate of an element u in $\mathcal{S}_2$.

A basis h_1,h_2 of $\mathcal{S}_2$ is formed by any two independent spinors in $\mathcal{S}_2$. The reciprocal basis h^1,h^2 is defined to satisfy

$$h^a\blacktriangle h_b=\delta^a_b.$$

The relations between a spinor u in $\mathcal{S}_2$ to its contravariant components u^1, u^2, and its covariant components u_1, u_2, are given in I [Eqs. (5) and (6)].

The conjugate $\bar{h}_1,\bar{h}_2$ of any basis for $\mathcal{S}_2$ is a basis in $\overline{\mathcal{S}}_2$. The corresponding reciprocal basis is $\bar{h}^1,\bar{h}^2$.

B. Spinor space $\mathcal{S}_4$

The direct sum spinor space $\mathcal{S}_4=\mathcal{S}_2\oplus\overline{\mathcal{S}}_2$ consists of all pairs $u,\bar{v}$ of spinors with $u\in\mathcal{S}_2$, $\bar{v}\in\overline{\mathcal{S}}_2$, which we write formally as $u+\bar{v}$. In the special cases where $u=0$ or $\bar{v}=\bar{o}$ we write $\bar{v}\equiv 0+\bar{v}$ and $u\equiv u+\bar{0}$. Addition and multiplication by complex numbers are defined by the equations

$$\begin{aligned}
&(u+\bar{v})+(y+\bar{z})=(u+y)+(\bar{v}+\bar{z}),\\
&\alpha(u+\bar{v})=\alpha u+\alpha\bar{v}.
\end{aligned} \tag{3}$$

The inner product in $\mathcal{S}_4$ is defined by the equation

$$(u+\bar{v})\blacktriangle(y+\bar{z})=u\blacktriangle y+\bar{v}\blacktriangle\bar{z}. \tag{4}$$

Note that $u\blacktriangle\bar{z}=0$ and $\bar{v}\blacktriangle y=0$. It follows that the properties (a), (b), (c) for spinors in $\mathcal{S}_2$ are satisfied if u, v, w are replaced by arbitrary spinors in $\mathcal{S}_4$. The conjugate of $\psi=u+\bar{v}$ in $\mathcal{S}_4$ is defined as

$$\bar{\psi}=\overline{(u+\bar{v})}=\bar{u}+\bar{\bar{v}}=\bar{u}+v=v+\bar{u} \tag{5}$$

and is also an element of $\mathcal{S}_4$. This operation in $\mathcal{S}_4$ has the properties (d), (e), (f), (g) if u,v are replaced by arbitrary spinors in $\mathcal{S}_4$. We shall regard $\mathcal{S}_2$ and the subspace $\mathcal{S}_2\oplus\{\bar{0}\}$ of $\mathcal{S}_4$ as being identical and shall write $\mathcal{S}_2\equiv\mathcal{S}_2\oplus\{\bar{0}\}$; likewise we write $\overline{\mathcal{S}}_2\equiv\{0\}\oplus\overline{\mathcal{S}}_2$.

Given any basis l_1,l_2,l_3,l_4, in $\mathcal{S}_4$, the reciprocal basis l^1,l^2,l^3,l^4 is defined by the equation

$$l^\alpha\blacktriangle l_\beta=\delta^\alpha_\beta, \tag{6}$$

where $\alpha,\beta=1,\ldots,4$. We have the following relations for a spinor $\psi\in\mathcal{S}_4$, its contravariant and covariant components ψ^α and ψ_α respectively:

$$\begin{aligned}
&\psi=\psi^\alpha l_\alpha=\psi_\alpha l^\alpha,\\
&\psi^\alpha=l^\alpha\blacktriangle\psi,\\
&\psi_\alpha=\psi\blacktriangle l_\alpha,
\end{aligned} \tag{7}$$

and for any two spinors ψ and ξ in $\mathcal{S}_4$:

$$\psi\blacktriangle\xi=\psi_\alpha\xi^\alpha=-\psi^\alpha\xi_\alpha \tag{8}$$

C. Spinor-tensor spaces

A second order spinor-tensor $\mathbf{T}$ is defined to be a complex bilinear functional $T(\xi,\eta)$ for $\xi,\eta\in\mathcal{S}_4$, i.e., it has the properties

$$\begin{aligned}
&T(\alpha\chi+\beta\xi,\eta)=\alpha T(\chi,\eta)+\beta T(\xi,\eta),\\
&T(\xi,\alpha\chi+\beta\eta)=\alpha T(\xi,\chi)+\beta T(\xi,\eta)
\end{aligned} \tag{9}$$

for all $\chi,\xi,\eta\in\mathcal{S}_4$ and $\alpha,\beta\in\mathcal{C}$. We shall use the notation $\mathbf{T}$ for the spinor-tensor, $T(\xi,\eta)\in\mathcal{C}$ for the value it associates with the pair $\xi,\eta\in\mathcal{S}_4$, and $\mathcal{S}_4\otimes\mathcal{S}_4$ for the space of all second-order spinor-tensors. The sum $\mathbf{R}=\mathbf{S}+\mathbf{T}$ for $\mathbf{S},\mathbf{T}\in\mathcal{S}_4\otimes\mathcal{S}_4$ is an element of $\mathcal{S}_4\otimes\mathcal{S}_4$ defined by the equation

$$R(\xi,\eta)=S(\xi,\eta)+T(\xi,\eta) \tag{10}$$

for every $\xi, \eta \in \mathcal{S}_4$. For $\alpha \in \mathcal{C}$ and $\mathbf{T} \in \mathcal{S}_4 \otimes \mathcal{S}_4$, the product $\mathbf{P} \doteq \alpha\mathbf{T}$ is an element of $\mathcal{S}_4 \otimes \mathcal{S}_4$ defined by the equation

$$P(\xi, \eta) = \alpha T(\xi, \eta) \tag{11}$$

for every $\xi, \eta \in \mathcal{S}_4$.

For an arbitrary pair of spinors $\psi, \zeta \in \mathcal{S}_4$, its tensor product $\mathbf{T} = \psi \otimes \zeta \in \mathcal{S}_4 \otimes \mathcal{S}_4$ is defined by the equation

$$T(\xi, \eta) = \psi \blacktriangle \xi \zeta \blacktriangle \eta \tag{12}$$

for all $\xi, \eta \in \mathcal{S}_4$. We shall use the abbreviated notation $\psi\zeta$ for $\psi \otimes \zeta$.

The subspace $\overline{\mathcal{S}}_2 \otimes \mathcal{S}_2$ of $\mathcal{S}_4 \otimes \mathcal{S}_4$ is the set of all elements which can be expressed as a linear combination of tensors of the form $\bar{\boldsymbol{u}}\boldsymbol{v}$ with $\bar{u} \in \overline{\mathcal{S}}_2$ and $v \in \mathcal{S}_2$. The subspaces $\mathcal{S}_2 \otimes \overline{\mathcal{S}}_2$, $\mathcal{S}_2 \otimes \mathcal{S}_2$, and $\overline{\mathcal{S}}_2 \otimes \overline{\mathcal{S}}_2$ can be defined in a similar way.

The space $\mathcal{S}_4^{\otimes r}$ of rth order spinor-tensors is the set of all complex functionals of r arguments in $\mathcal{S}_4$ which is linear in each argument. Addition of rth order tensors, multiplication of rth order tensors by complex numbers, the tensor product of r spinors in $\mathcal{S}_4$ can be defined by a straightforward generalization of the definitions we have already given.

Given $\mathbf{A} \in \mathcal{S}_4 \otimes \mathcal{S}_4$ and $\xi \in \mathcal{S}_4$ we define $\mathbf{A}\blacktriangle\xi$ and $\xi\blacktriangle\mathbf{A}$ in the special case where $\mathbf{A} = \psi\zeta$ as

$$\begin{aligned} (\psi\zeta)\blacktriangle\xi &= (\zeta\blacktriangle\xi)\psi, \\ \xi\blacktriangle(\psi\zeta) &= (\xi\blacktriangle\psi)\zeta \end{aligned} \tag{13}$$

and since each $\mathbf{A}$ is a linear combination of elements of the form $\psi\zeta$, the assumption of linearity in $\mathbf{A}$,

$$\begin{aligned} (\alpha\mathbf{A} + \beta\mathbf{B})\blacktriangle\xi &= \alpha(\mathbf{A}\blacktriangle\xi) + \beta(\mathbf{B}\blacktriangle\xi), \\ \xi\blacktriangle(\alpha\mathbf{A} + \beta\mathbf{B}) &= \alpha(\xi\blacktriangle\mathbf{A}) + \beta(\xi\blacktriangle\mathbf{B}), \end{aligned} \tag{14}$$

is sufficient to extend the definition to include all $\mathbf{A}$.

The operations $\mathbf{A}_s$ (scalar of A), $\mathbf{A}\blacktriangle\mathbf{B}$ (product of A and B), $\mathbf{A} \overset{\blacktriangle}{\blacktriangle} \mathbf{B}$ (double product of A and B), $\tilde{\mathbf{A}} \equiv \mathbf{A}_T$ (transpose of A), and $\overline{\mathbf{A}}$ (complex conjugate of A) for $\mathbf{A}, \mathbf{B} \in \mathcal{S}_4 \otimes \mathcal{S}_4$ are defined in the special cases where $\mathbf{A} = \psi\zeta$ and $\mathbf{B} = \chi\xi$; the assumption of linearity in A and B extends the definitions to arbitrary A and B in all the above operations except $\overline{\mathbf{A}}$ in which case antilinearity is assumed:

$$\begin{aligned} (\psi\zeta)_s &= \psi\blacktriangle\zeta, \\ (\psi\zeta)\blacktriangle(\chi\xi) &= (\zeta\blacktriangle\chi)(\psi\xi), \\ (\psi\zeta)\overset{\blacktriangle}{\blacktriangle}(\chi\xi) &= (\psi\blacktriangle\chi)(\zeta\blacktriangle\xi), \\ (\psi\zeta)_T &= \zeta\psi, \\ (\overline{\psi\zeta}) &= \bar{\psi}\bar{\zeta}. \end{aligned} \tag{15}$$

We also define

$$\begin{aligned} &\mathbf{A}^+ = \tilde{\overline{\mathbf{A}}} && \text{(Hermitian conjugate of } \mathbf{A}), \\ &\mathbf{A} \odot \mathbf{B} = -\mathbf{A} : \mathbf{B} && \text{(inner product of A and B).} \end{aligned} \tag{16}$$

The unit spinor tensor $\mathbf{I} = \boldsymbol{l}_\alpha \boldsymbol{l}^\alpha = -\boldsymbol{l}^\alpha \boldsymbol{l}_\alpha$ in $\mathcal{S}_4 \otimes \mathcal{S}_4$ is defined to satisfy

$$\mathbf{I}\blacktriangle\psi = \psi\blacktriangle\mathbf{I} = \psi \tag{17}$$

for all spinors ψ in $\mathcal{S}_4$. Moreover, we also have

$$\mathbf{I} = \mathbf{I}_2 + \overline{\mathbf{I}}_2 \tag{18}$$

where $\mathbf{I}_2$ and $\overline{\mathbf{I}}_2$ are the unit spinor tensors in $\mathcal{S}_2 \otimes \mathcal{S}_2$ and $\overline{\mathcal{S}}_2 \otimes \overline{\mathcal{S}}_2$, respectively, which we introduced in I [Eqs. (22), (23), (24), (25)].

Bases and components of spinor tensors in $\mathcal{S}_2 \otimes \mathcal{S}_2$ are discussed in I [Eqs. (19), (20)]. The generalization of these equations to $\mathcal{S}_4 \otimes \mathcal{S}_4$ and various subspaces such as $\overline{\mathcal{S}}_2 \otimes \mathcal{S}_2$, etc., is self-evident.

We define the space $\mathcal{M}_4 = \overline{\mathcal{S}}_2 \otimes_H \mathcal{S}_2$ to be the set of all Hermitian tensors in $\overline{\mathcal{S}}_2 \otimes \mathcal{S}_2$. Each element in $\mathcal{M}_4$ is a real linear combination of tensors of the form $\bar{\boldsymbol{u}}\boldsymbol{u}$. The space $\mathcal{M}_4$ is closed under the operations of addition and multiplication by real numbers; it also follows that $\mathcal{M}_4$ is a vector space over the reals, and has dimension four. The inner product $\mathbf{A} \odot \mathbf{B}$ for $\mathbf{A}, \mathbf{B}$ in $\mathcal{M}_4$ is real, symmetric, bilinear, and nondegenerate. Taking the basis $\mathbf{E}_0, \mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3$ given by Eq. (52) in I,[7]

$$\begin{aligned} \mathbf{E}_0 &= (2)^{-1/2}(\bar{\boldsymbol{h}}_1\boldsymbol{h}_1 + \bar{\boldsymbol{h}}_2\boldsymbol{h}_2), \\ \mathbf{E}_1 &= -(2)^{-1/2}(\bar{\boldsymbol{h}}_1\boldsymbol{h}_2 + \bar{\boldsymbol{h}}_2\boldsymbol{h}_1), \\ \mathbf{E}_2 &= -i(2)^{-1/2}(\bar{\boldsymbol{h}}_1\boldsymbol{h}_2 - \bar{\boldsymbol{h}}_2\boldsymbol{h}_1), \\ \mathbf{E}_3 &= -(2)^{-1/2}(\bar{\boldsymbol{h}}_1\boldsymbol{h}_1 - \bar{\boldsymbol{h}}_2\boldsymbol{h}_2), \end{aligned}$$

with $\boldsymbol{h}_1, \boldsymbol{h}_2$ chosen such that $\boldsymbol{h}_1\blacktriangle\boldsymbol{h}_2 = 1$, one can easily show that

$$\begin{aligned} &g_{11} = g_{22} = g_{33} = -g_{00} = 1, \\ &g_{\mu\nu} = 0 \quad \text{for } \mu \neq \nu, \end{aligned}$$

where

$$g_{\mu\nu} = E_\mu \odot E_\nu. \tag{19}$$

Therefore the vector space $\mathcal{M}_4$ with the inner product $\mathbf{A} \odot \mathbf{B}$ is a Minkowski space. We shall disregard isomophisms, and regard this space as identical to the space of four-vectors used in special relativity theory. We shall also sometimes use the notation a in place of A and $\mathbf{a}\cdot\mathbf{b}$ in place of $\mathbf{A} \odot \mathbf{B}$.

With the above basis $\mathbf{E}_\mu$, the reciprocal basis $\mathbf{E}^\mu$ defined by the relation $\mathbf{E}^\mu \odot \mathbf{E}_\nu = \delta^\mu_\nu$ becomes $\mathbf{E}^0 = -\mathbf{E}_0$, $\mathbf{E}^k = \mathbf{E}_k$ for $k = 1, 2, 3$.

Elements of the space $\mathcal{CM}_4 = \overline{\mathcal{S}}_2 \otimes \mathcal{S}_2$ can be expressed in the form $\mathbf{Z} = \mathbf{A} + i\mathbf{B}$ with $\mathbf{A}, \mathbf{B} \in \overline{\mathcal{S}}_2 \otimes_H \mathcal{S}_2$, thus $\overline{\mathcal{S}}_2 \otimes \mathcal{S}_2$ is the space of complex four-vectors. The Hermitian conjugate of Z is

$$\mathbf{Z}^+ = \mathbf{A}^+ - i\mathbf{B}^+ = \mathbf{A} - i\mathbf{B}.$$

Thus Hermitian conjugation is the appropriate operation in this space which corresponds to the usual concept of complex conjugation of complex four-vectors.

Elements A in $\mathcal{M}_4$ and also in $\mathcal{CM}_4$ can be expressed in terms of the basis $\mathbf{E}_\mu$ of $\mathcal{M}_4$ or the reciprocal basis $\mathbf{E}^\mu$ as

$$\mathbf{A} = A^\mu \mathbf{E}_\mu = A_\mu \mathbf{E}^\mu, \tag{20}$$

and it follows that

$$A^\mu = \mathbf{E}^\mu \odot \mathbf{A}, \quad A_\mu = \mathbf{E}_\mu \odot \mathbf{A}. \tag{21}$$

The space $\widetilde{\mathcal{M}}_4 = \mathcal{S}_2 \otimes_H \overline{\mathcal{S}}_2$ is the set of all Hermitian

TABLE I. Spinor and tensor spaces.

Space	Notation and expressions of elements of space	Inner product and properties	Bases and reciprocal bases
$\mathcal{S}_2$	u, v, s, t $\bar{\bar{u}} = u$	$u\blacktriangle v = -v\blacktriangle u$	h_1, h_2 and h^1, h^2 $(h^a\blacktriangle h_b = \delta^a_b)$
$\bar{\mathcal{S}}_2$	$\bar{u}, \bar{v}, \bar{s}, \bar{t}$ $(\overline{\alpha u + \beta v}) = \bar{\alpha}\bar{u} + \bar{\beta}\bar{v}$ for $\alpha, \beta \in \mathcal{C}$	$\bar{u}\blacktriangle\bar{v} = -\bar{v}\blacktriangle\bar{u}$ $\bar{u}\blacktriangle\bar{v} = \overline{u\blacktriangle v}$	$\bar{h}_1, \bar{h}_2$ and $\bar{h}^1, \bar{h}^2$ $(\bar{h}^a\blacktriangle\bar{h}_b = \delta^a_b)$
$\mathcal{S}_4 = \mathcal{S}_2 \oplus \bar{\mathcal{S}}_2$	$\psi = u + \bar{v}$, $\zeta = s + \bar{t}$ $\bar{\psi} = v + \bar{u}$, $\bar{\bar{\psi}} = \psi$	$\psi\blacktriangle\zeta = -\zeta\blacktriangle\psi = u\blacktriangle s + \bar{v}\blacktriangle\bar{t}$ $\bar{\psi}\blacktriangle\bar{\zeta} = \overline{\psi\blacktriangle\zeta}$	l_α and l^α $(\alpha = 1, \cdots, 4)$ $(l^\alpha\blacktriangle l_\beta = \delta^\alpha_\beta)$
$\mathcal{S}_4 \otimes \mathcal{S}_4$	$\mathbf{S} = \psi\zeta + \cdots$, $\mathbf{T} = \xi\eta + \cdots$ $\bar{\mathbf{S}} = \bar{\psi}\bar{\zeta} + \cdots$, $\mathbf{S}^+ = \tilde{\bar{\mathbf{S}}}$ $\mathbf{S}_T \equiv \tilde{\mathbf{S}} = \zeta\psi + \cdots$	$\mathbf{S}\odot\mathbf{T} = -[(\psi\blacktriangle\xi)(\zeta\blacktriangle\eta) + \cdots]$ $= \mathbf{T}\odot\mathbf{S} = \tilde{\mathbf{S}}\odot\tilde{\mathbf{T}}$ $\bar{\mathbf{S}}\odot\bar{\mathbf{T}} = \overline{\mathbf{S}\odot\mathbf{T}}$	$l_\alpha l_\beta$ and $-l^\alpha l^\beta$ $l_\alpha l^\beta$ and $l^\alpha l_\beta$
$\mathcal{S}_2 \otimes \mathcal{S}_2$	$\mathbf{M} = uv + \cdots$, $\mathbf{N}$	$\mathbf{M}\odot\mathbf{N}$	$h_a h_b$ and $-h^a h^b$ $h_a h^b$ and $h^a h_b$
$\bar{\mathcal{S}}_2 \otimes \bar{\mathcal{S}}_2$	$\bar{\mathbf{M}} = \overline{uv} + \cdots$, $\bar{\mathbf{N}}$	$\bar{\mathbf{M}}\odot\bar{\mathbf{N}} = \overline{\mathbf{M}\odot\mathbf{N}}$	$\bar{h}_a\bar{h}_b$ and $-\bar{h}^a\bar{h}^b$ $\bar{h}_a\bar{h}^b$ and $\bar{h}^a\bar{h}_b$
$\bar{\mathcal{S}}_2 \otimes \mathcal{S}_2$	$\mathbf{Y} = \bar{u}v + \cdots$, $\mathbf{Z}$	$\mathbf{Y}\odot\mathbf{Z}$	$\bar{h}_a h_b$ and $-\bar{h}^a h^b$ $\bar{h}_a h^b$ and $\bar{h}^a h_b$
$\mathcal{S}_2 \otimes \bar{\mathcal{S}}_2$	$\tilde{\mathbf{Y}} = v\bar{u} + \cdots$, $\tilde{\mathbf{Z}}$ $\bar{\mathbf{Y}} = u\bar{v} + \cdots$, $\bar{\mathbf{Z}}$	$\tilde{\mathbf{Y}}\odot\tilde{\mathbf{Z}} = \mathbf{Y}\odot\mathbf{Z}$, $\mathbf{Y}\odot\tilde{\mathbf{Z}} = 0$ $\bar{\mathbf{Y}}\odot\bar{\mathbf{Z}} = \overline{\mathbf{Y}\odot\mathbf{Z}}$, $\mathbf{Y}\odot\bar{\mathbf{Z}} = 0$	$h_a\bar{h}_b$ and $-h_a\bar{h}^b$ $h_a\bar{h}^b$ and $h^a\bar{h}_b$
$\mathcal{M}_4 = \bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2$	$\mathbf{a} \equiv \mathbf{A} = \mathbf{A}^+ = \bar{u}u + \cdots$ $= \bar{s}t + \bar{t}s + \cdots$ $\mathbf{b} \equiv \mathbf{B}$, $\mathbf{c}, \mathbf{d}$	$\mathbf{a}\cdot\mathbf{b} \equiv \mathbf{A}\odot\mathbf{B}$	$\mathbf{e}_\mu \equiv \mathbf{E}_\mu$ and $\mathbf{e}^\mu \equiv \mathbf{E}^\mu$ $(\mu = 0, \cdots, 3)$ $(\mathbf{e}^\mu\cdot\mathbf{e}_\nu \equiv \mathbf{E}^\mu\odot\mathbf{E}_\nu = \delta^\mu_\nu)$ $(\mathbf{e}_\mu\cdot\mathbf{e}_\nu \equiv \mathbf{E}_\mu\odot\mathbf{E}_\nu = g_{\mu\nu})$
$\tilde{\mathcal{M}}_4 = \mathcal{S}_2 \otimes_H \bar{\mathcal{S}}_2$	$\tilde{\mathbf{A}} = \bar{\mathbf{A}} = u\bar{u} + \cdots = t\bar{s} + s\bar{t} + \cdots$	$\tilde{\mathbf{A}}\odot\tilde{\mathbf{B}} = \mathbf{A}\odot\mathbf{B}$	$\tilde{\mathbf{E}}_\mu$ and $\tilde{\mathbf{E}}^\mu$, etc.
$\mathcal{M}_S = \mathcal{M}_4 \oplus_S \tilde{\mathcal{M}}_4$	$\mathbf{A}_S = \tilde{\mathbf{A}}_S = \bar{\mathbf{A}}_S = (1/\sqrt{2})(\mathbf{A} + \tilde{\mathbf{A}})$	$\mathbf{A}_S\odot\mathbf{B}_S = \mathbf{A}\odot\mathbf{B}$	$(\mathbf{E}_\mu)_S$ and $(\mathbf{E}^\mu)_S$, etc.
$i\mathcal{M}_S$	$\mathbf{A}_I = \tilde{\mathbf{A}}_I = -\bar{\mathbf{A}}_I = \sqrt{2}i(\mathbf{A} + \tilde{\mathbf{A}})$	$-\frac{1}{4}\mathbf{A}_I\odot\mathbf{B}_I = \mathbf{A}\odot\mathbf{B}$	$\mathbf{\Gamma}_\mu = (\mathbf{E}_\mu)_I$ and $\mathbf{\Gamma}^\mu = (\mathbf{E}^\mu)_I$
$\mathcal{C}\mathcal{M}_4 = \bar{\mathcal{S}}_2 \otimes \mathcal{S}_2$	$\mathbf{y} \equiv \mathbf{Y} = \mathbf{a} + i\mathbf{b} \equiv \mathbf{A} + i\mathbf{B}$ $\mathbf{z} \equiv \mathbf{Z}$, $\mathbf{Y}^+ = \mathbf{A} - i\mathbf{B}$	$\mathbf{y}\cdot\mathbf{z} \equiv \mathbf{Y}\odot\mathbf{Z}$	$\mathbf{e}_\mu \equiv \mathbf{E}_\mu$ and $\mathbf{e}^\mu \equiv \mathbf{E}^\mu$
$\mathcal{C}\tilde{\mathcal{M}}_4 = \mathcal{S}_2 \otimes \bar{\mathcal{S}}_2$	$\tilde{\mathbf{Y}} = \tilde{\mathbf{A}} + i\tilde{\mathbf{B}}$	$\tilde{\mathbf{Y}}\odot\tilde{\mathbf{Z}} = \mathbf{Y}\odot\mathbf{Z}$	$\tilde{\mathbf{E}}_\mu$ and $\tilde{\mathbf{E}}^\mu$
$\mathcal{M}_4 \otimes \mathcal{M}_4$	$\mathbf{W} = \mathbf{ab} + \cdots$, $\mathbf{V} = \mathbf{cd} + \cdots$	$\mathbf{W} : \mathbf{V} = (\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d}) + \cdots$	$\mathbf{e}_\mu\mathbf{e}_\nu$ and $\mathbf{e}^\mu\mathbf{e}^\nu$ $\mathbf{e}_\mu\mathbf{e}^\nu$ and $\mathbf{e}^\mu\mathbf{e}_\nu$

tensors in $\mathcal{S}_2 \otimes \bar{\mathcal{S}}_2$. Each element of $\tilde{\mathcal{M}}_4$ is a real linear combination of tensors of the form $u\bar{u}$. The one-to-one map $\mathbf{A} \to \tilde{\mathbf{A}}$ from $\mathcal{M}_4$ onto $\tilde{\mathcal{M}}_4$ preserves real linear combinations and inner products, i.e., $\mathbf{A}\odot\mathbf{B} = \tilde{\mathbf{A}}\odot\tilde{\mathbf{B}}$; thus $\tilde{\mathcal{M}}_4$ is isomorphic to to $\mathcal{M}_4$. Similarly the space $\mathcal{C}\tilde{\mathcal{M}}_4 = \mathcal{S}_2 \otimes \bar{\mathcal{S}}_2$ is isomorphic to $\mathcal{C}\mathcal{M}_4 = \bar{\mathcal{S}}_2 \otimes \mathcal{S}_2$. Also note that if $\mathbf{A} \in \mathcal{M}_4$, then $\bar{\mathbf{A}} \in \tilde{\mathcal{M}}_4$ because $\bar{\mathbf{A}} = \tilde{\mathbf{A}}$.

The space $\mathcal{M}_S = \mathcal{M}_4 \oplus_S \tilde{\mathcal{M}}_4 = (\bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2) \oplus_S (\mathcal{S}_2 \otimes_H \bar{\mathcal{S}}_2)$ is the set of all symmetric tensors in $\mathcal{M}_4 \oplus \tilde{\mathcal{M}}_4 = (\bar{\mathcal{S}}_2 \otimes_H \mathcal{S}_2) \oplus (\mathcal{S}_2 \otimes_H \bar{\mathcal{S}}_2)$. Each element in $\mathcal{M}_S$ is a real linear combination of tensors of the form $\bar{u}u + u\bar{u}$. The one-to-one map $\mathbf{A} \to \mathbf{A}_S = 2^{-1/2}(\mathbf{A} + \tilde{\mathbf{A}})$ from $\mathcal{M}_4$ onto $\mathcal{M}_S$ preserves real linear combinations and inner products, i.e., $\mathbf{A}\odot\mathbf{B} = \mathbf{A}_S \odot \mathbf{B}_S$; thus $\mathcal{M}_S$ is isomorphic to $\mathcal{M}_4$. Similarly $\mathcal{C}\mathcal{M}_S = \mathcal{C}\mathcal{M}_4 \oplus_S \mathcal{C}\tilde{\mathcal{M}}_4 = (\bar{\mathcal{S}}_2 \otimes \mathcal{S}_2) \oplus_S (\mathcal{S}_2 \otimes \bar{\mathcal{S}}_2)$ is isomorphic to $\mathcal{C}\mathcal{M}_4 = \bar{\mathcal{S}}_2 \otimes \mathcal{S}_2$. Also note that $\tilde{\mathbf{A}}_S = \bar{\mathbf{A}}_S = \mathbf{A}_S$ for $\mathbf{A}_S \in \mathcal{M}_S$.

The space $i\mathcal{M}_S$ is the set of all elements of the form $i\mathbf{A}_S$ with $\mathbf{A}_S$ in $\mathcal{M}_S$. The one-to-one map $\mathbf{A} \to \mathbf{A}_I = 2i\mathbf{A}_S = 2^{1/2}i(\mathbf{A} + \tilde{\mathbf{A}})$ from $\mathcal{M}_4$ onto $i\mathcal{M}_S$ preserves real linear combinations, but inner products are related as $\mathbf{A}_I \odot \mathbf{B}_I = -4\mathbf{A}\odot\mathbf{B}$. We also have the useful relation[8]

$$\mathbf{A}_I\blacktriangle\mathbf{B}_I + \mathbf{B}_I\blacktriangle\mathbf{A}_I = -2\mathbf{A}\odot\mathbf{B}\mathbf{I} \tag{22}$$

The operators

$$\mathbf{\Gamma}^\mu = \mathbf{E}^\mu_I = 2i\mathbf{E}^\mu_S = 2^{1/2}i(\mathbf{E}^\mu + \tilde{\mathbf{E}}^\mu) \tag{23}$$

form a basis for $i\mathcal{M}_S$, and Eq. (22) for these operators is

$$\mathbf{\Gamma}^\mu\blacktriangle\mathbf{\Gamma}^\nu + \mathbf{\Gamma}^\nu\blacktriangle\mathbf{\Gamma}^\mu = -2(\mathbf{E}^\mu \odot \mathbf{E}^\nu)\mathbf{I} = -2g^{\mu\nu}\mathbf{I} \tag{24}$$

Hence the $\mathbf{\Gamma}^\mu$ are the Dirac gamma operators.

Note that

$$\tilde{\mathbf{A}}_I = -\bar{\mathbf{A}}_I = \mathbf{A}_I \quad \text{for all } \mathbf{A}_I \text{ in } i\mathcal{M}_S.$$

Recall now the linear operation

$$(\xi\eta\varphi\chi)^\dagger = -\xi\varphi\eta\chi \tag{25}$$

defined in Eq. (73a) in I. It follows from the linearity of this operation that if A, $\mathbf{B} \in \mathcal{S}_4 \otimes \mathcal{S}_4$, then

$$(\mathbf{AB})^\dagger \odot \psi\zeta = \mathbf{A}\blacktriangle\psi\mathbf{B}\blacktriangle\zeta. \tag{26}$$

The permutation operator (12) on $\xi\psi$ and on $\xi\varphi\eta\chi$ by definition gives

$$\begin{aligned} (12)\ \xi\varphi &= \varphi\,\xi, \\ (12)\ \xi\varphi\eta\chi &= \varphi\,\xi\eta\chi, \end{aligned} \tag{27}$$

and therefore

$$(12)\ (\xi\eta\varphi\chi)^\dagger = -\varphi\,\xi\eta\chi. \tag{28}$$

It follows from the linearity of the (12) operation that

$$[(12)\,(\mathbf{AB})^\dagger]\odot\psi\xi=(12)[\mathbf{A}\blacktriangle\psi\mathbf{B}\blacktriangle\xi]=\mathbf{B}\blacktriangle\xi\mathbf{A}\blacktriangle\psi. \tag{29}$$

The unit tensor $\mathbf{I}_4$ in $\mathcal{M}_4\otimes\mathcal{M}_4$, defined to satisfy

$$\mathbf{I}_4\odot\mathbf{A}=\mathbf{A} \tag{30}$$

for all $\mathbf{A}$ in $\mathcal{M}_4$, also acts as a projection operator onto $C\mathcal{M}_4$. It can be expressed as

$$\mathbf{I}_4=(\bar{\mathrm{I}}_2\mathrm{I}_2)^\dagger=-\bar{h}_a h_b\bar{h}^a h^b. \tag{31}$$

Its action on $\psi\xi$ where $\psi=\mathrm{u}+\bar{\mathrm{v}}$ and $\xi=\mathrm{s}+\bar{\mathrm{t}}$ is as follows:

$$\mathbf{I}_4\odot\psi\xi=(\bar{\mathrm{I}}_2\mathrm{I}_2)^\dagger\odot(u+\bar{v})(s+\bar{t})=\bar{\mathrm{I}}_2\blacktriangle(u+\bar{v})\mathrm{I}_2\blacktriangle(s+\bar{t})=\bar{v}s. \tag{32}$$

The properties of the unit tensor $\bar{\mathbf{I}}_4$ in $\tilde{\mathcal{M}}_4\otimes\tilde{\mathcal{M}}_4$ are analogous to those of $\mathbf{I}_4$ in $\mathcal{M}_4\otimes\mathcal{M}_4$, and we have

$$\bar{\mathbf{I}}_4=(\mathrm{I}_2\bar{\mathrm{I}}_2)^\dagger, \tag{33}$$

$$\bar{\mathbf{I}}_4\odot\psi\xi=u\bar{t}. \tag{34}$$

The unit tensor $\mathbf{I}_{4s}$ in $\mathcal{M}_s\otimes\mathcal{M}_s$, defined to satisfy

$$\mathbf{I}_{4S}\odot\mathbf{A}_S=\mathbf{A}_S \tag{35}$$

for all $\mathbf{A}_S$ in $\mathcal{M}_S$, also acts as a projection operator onto $C\mathcal{M}_S$. Since $\mathbf{I}_4+\bar{\mathbf{I}}_4$ is a projection operator onto $C\mathcal{M}_4\oplus C\tilde{\mathcal{M}}_4$, and the symmetrizer $\frac{1}{2}[1+(12)]$ acting in $C\mathcal{M}_4\oplus C\tilde{\mathcal{M}}_4$ projects onto $C\mathcal{M}_4\oplus_S C\tilde{\mathcal{M}}_4$, it follows that these two operators acting in succession give an expression for $\mathbf{I}_{4S}$; i.e.,

$$\begin{aligned}\mathbf{I}_{4S}&=\tfrac{1}{2}[1+(12)](\mathbf{I}_4+\bar{\mathbf{I}}_4)\\&=\tfrac{1}{2}[1+(12)](\bar{\mathrm{I}}_2\mathrm{I}_2+\mathrm{I}_2\bar{\mathrm{I}}_2)^\dagger.\end{aligned} \tag{36}$$

The action of $\mathbf{I}_{4S}$ on $\psi\xi$ is

$$\begin{aligned}\mathbf{I}_{4S}\odot\psi\xi&=\tfrac{1}{2}[1+(12)](\mathbf{I}_4+\bar{\mathbf{I}}_4)\odot(u+\bar{v})(s+\bar{t})\\&=\tfrac{1}{2}[1+(12)](\bar{v}s+u\bar{t})\\&=\tfrac{1}{2}(\bar{v}s+s\bar{v}+u\bar{t}+\bar{t}u)\end{aligned} \tag{37}$$

D. Summary

Some additional operations not in the table:

$$\mathbf{S}\blacktriangle\xi=(\psi\xi+\cdots)\blacktriangle\xi=(\xi\blacktriangle\xi)\psi+\cdots$$

$$\xi\blacktriangle\mathbf{S}=\xi\blacktriangle(\psi\xi+\cdots)=(\xi\blacktriangle\psi)\xi+\cdots$$

$$\mathbf{S}_s=(\psi\xi+\cdots)_s=\psi\blacktriangle\xi+\cdots$$

$$\mathbf{S}\blacktriangle\mathbf{T}=(\psi\xi+\cdots)\blacktriangle(\xi\eta+\cdots)=(\xi\blacktriangle\xi)\psi\eta+\cdots$$

$$\mathbf{S}\overset{\blacktriangle}{\blacktriangle}\mathbf{T}=(\psi\xi+\cdots)\overset{\blacktriangle}{\blacktriangle}(\xi\eta+\cdots)=(\psi\blacktriangle\xi)(\xi\blacktriangle\eta)+\cdots=-\mathbf{S}\odot\mathbf{T}$$

$$\mathbf{W}\cdot\mathbf{c}=(\mathbf{ab}+\cdots)\cdot\mathbf{c}=(\mathbf{b}\cdot\mathbf{c})\mathbf{a}+\cdots$$

$$\mathbf{c}\cdot\mathbf{W}=\mathbf{c}\cdot(\mathbf{ab}+\cdots)=(\mathbf{c}\cdot\mathbf{a})\mathbf{b}+\cdots$$

Unit tensors:

$\bar{\mathrm{I}}_2=h_a h^a=-h^a h_a$ in $S_2\otimes S_2$, $\mathrm{I}_2\blacktriangle u=u$, $\mathrm{I}_2\blacktriangle\bar{v}=\bar{0}$

$\bar{\mathrm{I}}_2=\bar{h}_a\bar{h}^a=-\bar{h}^a\bar{h}_a$ in $\bar{S}_2\otimes\bar{S}_2$, $\bar{\mathrm{I}}_2\blacktriangle u=0$, $\bar{\mathrm{I}}_2\blacktriangle\bar{v}=\bar{v}$

$\mathrm{I}=\mathrm{I}_2+\bar{\mathrm{I}}_2=l_\alpha l^\alpha=-l^\alpha l_\alpha$ in $S_4\otimes S_4$, $\mathrm{I}\blacktriangle\psi=\psi$

$\mathbf{I}_4=\mathbf{e}_\mu\mathbf{e}^\mu=\mathbf{e}^\mu\mathbf{e}_\mu$ in $\mathcal{M}_4\otimes\mathcal{M}_4$, $\mathbf{I}_4\cdot\mathbf{a}=\mathbf{a}$

and in $(C\mathcal{M}_4)\otimes(C\mathcal{M}_4)$, $\mathbf{I}_4\cdot\mathbf{z}=\mathbf{z}$

$\mathbf{I}_{4S}=(\mathbf{E}_\mu)_S(\mathbf{E}^\mu)_S=(\mathbf{E}^\mu)_S(\mathbf{E}_\mu)_S$ in $\mathcal{M}_S\otimes\mathcal{M}_S$, $\mathbf{I}_{4S}\odot\mathbf{A}_S=\mathbf{A}_S$

Property of elements of $i\mathcal{M}_S$:

$$\mathbf{A}_I\blacktriangle\mathbf{B}_I+\mathbf{B}_I\blacktriangle\mathbf{A}_I=-2\mathbf{A}\odot\mathbf{B}\mathrm{I}$$

$$\mathbf{\Gamma}^\mu\blacktriangle\mathbf{\Gamma}^\nu+\mathbf{\Gamma}^\nu\blacktriangle\mathbf{\Gamma}^\mu=-2g^{\mu\nu}\mathrm{I}$$

III. CONNECTIONS AND CURVATURE IN THE GENERAL RELATIVITY MANIFOLD

As a preliminary to the discussion of spinor connections and spinor curvature tensors we summarize some modern methods of differential geometry in the treatment of the four-vector connection and the Riemann tensor. This section serves the twofold purpose of establishing the formalism and deriving equations which are utilized in the next two sections. First we shall briefly introduce tangent vector spaces and vector fields. We will be using the general relativity manifold $\mathcal{M}$ which is a four-dimensional manifold in which the tangent vector space $\mathcal{M}_q$ at each point q in $\mathcal{M}$ has a Minkowski inner product.

A *tangent vector* X_q at a point q in $\mathcal{M}$ is a map that associates a real number $X_q f$ with each differentiable real function f on $\mathcal{M}$ for which the following properties hold:

$$\begin{aligned}&X_q(f+g)=X_q f+X_q g,\\&X_q(\alpha f)=\alpha X_q f,\\&X_q(fg)=(X_q f)g(q)+f(q)(X_q g)\end{aligned} \tag{38}$$

for each real number α and differentiable functions f and g. The product αX_q of a real number α with X_q and the sum X_q+Y_q of two tangent vectors at q are defined by the equations

$$\begin{aligned}&(\alpha X_q)f=\alpha(X_q f),\\&(X_q+Y_q)f=X_q f+Y_q f\end{aligned} \tag{39}$$

for all f. It can be shown that the space $\mathcal{M}_q$ of all tangent vectors at q is a vector space over the real numbers and that $\mathcal{M}_q$ has the same dimension as the dimension of the manifold $\mathcal{M}$, which in this case is four. The domain of operation of a tangent vector X_q can be extended to complex differentiable functions $f=m+in$, where m and n are real functions, by the definition

$$X_q f=X_q m+iX_q n,$$

and the properties given by Eqs. (38) are easily shown to hold, where α is now any complex number, and f and g are complex functions. Furthermore, $\overline{(X_q f)}=X_q\bar{f}$.

We shall use two notations for tangent vectors, for example X_q and $\mathbf{x}(q)$ will stand for the same tangent vector at q. Whenever we are using it to operate on a function, we shall write it as X_q and call it a derivation operator at q. Otherwise we shall use the notation $\mathbf{x}(q)$ and call it a tangent vector or a four-vector or simply a vector at q. For example, when we are considering its inner product with another tangent vector $y(q)$ at q we shall write $\mathbf{x}(q)\cdot\mathbf{y}(q)$. We shall relate the two notations by the formal relation $X_q\rho=\mathbf{x}(q)$, where X_q is regarded as operating on the identity map $\rho(r)=r$ of $\mathcal{M}$ onto $\mathcal{M}$.

A vector field $\mathbf{x}$ associates a vector $\mathbf{x}(q)$ in $\mathcal{M}_q$ with each point q in $\mathcal{M}$. The other symbol X for the same vector field will be called a derivation operator on $\mathcal{M}$ and it associates a derivation operator X_q at q with each point q in $\mathcal{M}$. The derivation operator X operates on each differentiable function f on $\mathcal{M}$ to produce another function Xf on $\mathcal{M}$. We also have the formal relation $\mathbf{x}=X\rho$.

It can be proved easily that the Lie bracket $[X, Y]$ of two derivation operators X and Y defined by the equation

$$[X, Y]f=X(Yf)-Y(Xf)$$

for all functions f is also a derivation operator. Furthermore, it can be proved that in any coordinate system q^0, q^1, q^2, q^3 any derivation operator X can be expressed in the form

$$X=\lambda^\mu\partial_\mu,$$

where $\lambda^\mu(q)$ are real functions on $\mathcal{M}$ and $\partial_\mu\equiv\partial/\partial q^\mu$ (we are using the summation convention on repeated indices). Also, we can define the *natural basis* $e_\mu(q)$ of $\mathcal{M}_q$ for arbitrary q by the equation

$$\mathbf{e}_\mu=\partial_\mu\rho, \tag{40}$$

and, using this, we get

$$\mathbf{x}=X\rho=\lambda^\mu\partial_\mu\rho=\lambda^\mu\mathbf{e}_\mu.$$

The *reciprocal basis* $\mathbf{e}^\mu(q)$ of $\mathcal{M}_q$ corresponding to $\mathbf{e}_\mu(q)$ is defined uniquely by the equation

$$\mathbf{e}^\mu\cdot\mathbf{e}_\nu=\delta^\mu_\nu. \tag{41}$$

A *four-vector connection* D'_X on $\mathcal{M}$ associates an operator D'_X on vector fields with each derivation operator X such that, for each vector field $\mathbf{v}(q)$, $D'_X\mathbf{v}$ is another vector field having the following properties:

$$\begin{aligned}&D'_X(\mathbf{v}+\mathbf{w})=D'_X\mathbf{v}+D'_X\mathbf{w},\\&D'_X(f\mathbf{v})=(Xf)\mathbf{v}+fD'_X\mathbf{v},\\&D'_{X+Y}\mathbf{v}=D'_X\mathbf{v}+D'_Y\mathbf{v},\\&D'_{gX}\mathbf{v}=gD'_X\mathbf{v}\end{aligned}\tag{42}$$

for arbitrary real vector functions $\mathbf{v}(q)$ and $\mathbf{w}(q)$ and arbitrary real functions $f(q)$ and $g(q)$. The operator D'_X is also called the covariant derivative operator in the direction of $\mathbf{x}$. For the derivation operator $X=\partial_\mu$ in any coordinate system q^μ, we shall use the abbreviated notation $D'_\mu\equiv D'_{\partial_\mu}$, and call this the covariant derivative operator with respect to q^μ.

The operation of D'_X on complex four-vector fields $\mathbf{v}(q)=\mathbf{m}(q)+i\mathbf{n}(q)$, where $\mathbf{m}(q)$ and $\mathbf{n}(q)$ are real four-vector fields, is defined as

$$D'_X\mathbf{v}=D'_X\mathbf{m}+iD'_X\mathbf{n}.$$

It is easily shown that the same properties of Eqs. (42) hold, where $\mathbf{v}(q)$ and $\mathbf{w}(q)$ are now arbitrary complex four-vector fields, $f(q)$ is complex, and $g(q)$ is still real.

The *torsion tensor* $\mathbf{T}(q)$ of a connection D'_X, which has values in $\mathcal{M}_q\otimes\mathcal{M}_q\otimes\mathcal{M}_q$, is defined by the equation

$$\mathbf{xy}:\mathbf{T}=D'_X\mathbf{y}-D'_Y\mathbf{x}-[X, Y]\rho \tag{43}$$

for arbitrary vector fields $\mathbf{x}(q)$ and $\mathbf{y}(q)$. It can be shown that the value of the right-hand side of Eq. (43) at any point q depends on the value of $\mathbf{x}$ and $\mathbf{y}$ at q only. It can also be shown to be linear in $\mathbf{x}$ and in $\mathbf{y}$; therefore the tensor character of this expression is established.

A connection D'_X is said to be *symmetric* iff[9] the torsion tensor $\mathbf{T}(q)$ vanishes for all q, i.e.,

$$D'_X\mathbf{y}-D'_Y\mathbf{x}-[X, Y]\rho=0 \tag{44}$$

everywhere for all vector fields $\mathbf{x}$ and $\mathbf{y}$.

A connection D'_X is said to be *compatible with the inner product* iff[9]

$$X(\mathbf{y}\cdot\mathbf{z})=(D'_X\mathbf{y})\cdot\mathbf{z}+\mathbf{y}\cdot(D'_X\mathbf{z}) \tag{45}$$

for arbitrary vector fields $\mathbf{x}(q)$, $\mathbf{y}(q)$, and $\mathbf{z}(q)$.

It can be shown that there exists a unique connection D_X on $\mathcal{M}$ which has both properties of being symmetric and being compatible with the inner product. We shall call this the *standard four-vector connection*.

We now show how the expression for the covariant derivative of a vector field $\mathbf{v}(q)$ with respect to a set of coordinates q^μ can be converted to component form in the case of the standard connection. First we can write $\mathbf{v}(q)$ in terms of contravariant components $v^\mu(q)$ and covariant components $v_\mu(q)$ respectively as

$$\mathbf{v}=v^\mu\mathbf{e}_\mu=v_\mu\mathbf{e}^\mu. \tag{46}$$

From this equation and Eq. (41), we get

$$v^\mu=\mathbf{e}^\mu\cdot\mathbf{v},\quad v_\mu=\mathbf{e}_\mu\cdot\mathbf{v}. \tag{47}$$

The contravariant components $v^\mu{}_{;\nu}(q)$ of the covariant derivative of $\mathbf{v}$ with respect to q^ν are then

$$\begin{aligned}v^\mu{}_{;\nu}&=(D_\nu\mathbf{v})\cdot\mathbf{e}^\mu=[D_\nu(v^\lambda\mathbf{e}_\lambda)]\cdot\mathbf{e}^\mu\\&=[(\partial_\nu v^\lambda)\mathbf{e}_\lambda+v^\lambda(D_\nu\mathbf{e}_\lambda)]\cdot\mathbf{e}^\mu\\&=\partial_\nu v^\mu+\Gamma^\mu_{\nu\lambda}v^\lambda,\end{aligned}\tag{48}$$

where the Christoffel symbols $\Gamma^\mu_{\nu\lambda}(q)$ are defined by the equation

$$D_\nu\mathbf{e}_\lambda=\Gamma^\mu_{\nu\lambda}\mathbf{e}_\mu, \tag{49}$$

from which we immediately obtain

$$\Gamma^\mu_{\nu\lambda}=(D_\nu\mathbf{e}_\lambda)\cdot\mathbf{e}^\mu. \tag{50}$$

The covariant components $v_{\mu;\nu}(q)$ of the covariant derivative of $\mathbf{v}$ with respect to q^ν are

$$\begin{aligned}v_{\mu;\nu}&=(D_\nu\mathbf{v})\cdot\mathbf{e}_\mu=[D_\nu(v_\lambda\mathbf{e}^\lambda)]\cdot\mathbf{e}_\mu\\&=[(\partial_\nu v_\lambda)\mathbf{e}^\lambda+v_\lambda(D_\nu\mathbf{e}^\lambda)]\cdot\mathbf{e}_\mu,\end{aligned}$$

and, since

$$\begin{aligned}(D_\nu\mathbf{e}^\lambda)\cdot\mathbf{e}_\mu&=\partial_\nu(\mathbf{e}^\lambda\cdot\mathbf{e}_\mu)-\mathbf{e}^\lambda\cdot(D_\nu\mathbf{e}_\mu)\\&=\partial_\nu\delta^\lambda_\mu-\Gamma^\lambda_{\nu\mu}=-\Gamma^\lambda_{\nu\mu},\end{aligned}$$

we get

$$v_{\mu;\nu}=\partial_\nu v_\mu-\Gamma^\lambda_{\nu\mu}v_\lambda. \tag{51}$$

Equations (48) and (51) agree with the usual definitions of covariant derivatives of contravariant and covariant vectors in the component notation.

An arbitrary connection D'_X differs from the standard connection D_X by an operator $\mathbf{C}_X(q)$ which is just a linear transformation on $\mathcal{M}_q$ for each q. To prove this, first note that the equation

$$(D'_X - D_X)f\mathbf{v} = f(D'_X - D_X)\mathbf{v} \tag{52}$$

follows from the axioms for a connection. Taking any coordinate system q^μ and using Eqs. (46) and (47), we obtain

$$\begin{aligned}(D'_X - D_X)\mathbf{v} &= (D'_X - D_X)(v^\mu \mathbf{e}_\mu) \\ &= v^\mu (D'_X - D_X)\mathbf{e}_\mu = [(D'_X - D_X)\mathbf{e}_\mu]\mathbf{e}^\mu \circ \mathbf{v} \\ &= \mathbf{C}_X \circ \mathbf{v},\end{aligned}$$

where

$$\mathbf{C}_X(q) = [(D'_X - D_X)\mathbf{e}_\mu(q)]\mathbf{e}^\mu(q) \tag{53}$$

is in $\mathcal{M}_q \otimes \mathcal{M}_q$ for each q. Therefore,[10]

$$D'_X\mathbf{v} = D_X\mathbf{v} + \mathbf{C}_X \circ \mathbf{v} \tag{54}$$

for an arbitrary vector field $\mathbf{v}(q)$.

The operation of the standard covariant derivative operator D_X on any second order tensor field $\mathbf{M}(q)$ in $\mathcal{M}_q \otimes \mathcal{M}_q$ is defined by the equation

$$(D_X\mathbf{M}) : \mathbf{yz} = X(\mathbf{M} : \mathbf{yz}) - \mathbf{M} : (D_X\mathbf{y})\mathbf{z} - \mathbf{M} : \mathbf{y}(D_X\mathbf{z}) \tag{55}$$

for arbitrary vector fields $\mathbf{x}(q)$, $\mathbf{y}(q)$, $\mathbf{z}(q)$. The value of the expression on the right at each q can be shown to depend on the values of $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ at the point q only. It can also be shown to be linear in $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$; therefore, the tensor character of this expression is established.

The Riemann tensor $\mathfrak{R}(q)$ is in the space $\mathcal{M}_q \otimes \mathcal{M}_q \otimes \mathcal{M}_q \otimes \mathcal{M}_q$ for each q, and in the case of the standard connection is defined by the equation

$$\mathbf{xy} : \mathfrak{R} \circ \mathbf{z} = (D_X D_Y - D_Y D_X - D_{[X,Y]})\mathbf{z} \tag{56}$$

for arbitrary vector fields $\mathbf{x}(q)$, $\mathbf{y}(q)$, $\mathbf{z}(q)$. It can be shown that the value of the expression on the right side at q depends on the values of $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ at the point q only. It can also be shown to be linear in $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$; thus its tensor character is established and the definition is therefore valid.

For the next part of the discussion, we need the property

$$D_X(\mathbf{y} \circ \mathbf{M}) = (D_X\mathbf{y}) \circ \mathbf{M} + \mathbf{y} \circ (D_X\mathbf{M}) \tag{57}$$

for vector fields $\mathbf{y}(q)$, and second order tensor fields $\mathbf{M}(q)$. Equation (57) follows easily from the fact that the standard connection is compatible with the inner product.

The expression for the Riemann tensor $\mathfrak{R}(q)$ can now be put in another convenient form. Using the covariant gradient operator $\mathbf{D}$ defined by the equations

$$\begin{aligned}&\mathbf{x} \circ (\mathbf{D}\mathbf{v}) = D_X\mathbf{v}, \\ &\mathbf{x} \circ (\mathbf{D}\mathbf{M}) = D_X\mathbf{M}\end{aligned} \tag{58}$$

for vector fields $\mathbf{v}(q)$ and tensor fields $\mathbf{M}(q)$ of second order or higher, we can rewrite Eq. (56) with the aid of Eq. (57) and the symmetry property of the standard connection as follows:

$$\begin{aligned}\mathbf{xy} : \mathfrak{R} \circ \mathbf{z} &= D_X(\mathbf{y} \circ \mathbf{Dz}) - D_Y(\mathbf{x} \circ \mathbf{Dz}) - ([X, Y]\rho) \circ (\mathbf{Dz}) \\ &= (D_X\mathbf{y} - D_Y\mathbf{x} - [X, Y]\rho) \circ (\mathbf{Dz}) + \mathbf{y} \circ (D_X\mathbf{Dz}) \\ &\quad - \mathbf{x} \circ (D_Y\mathbf{Dz}) \\ &= \mathbf{y} \cdot (\mathbf{x} \cdot \mathbf{DDz}) - \mathbf{x} \cdot (\mathbf{y} \cdot \mathbf{DDz}) \\ &= \mathbf{xy} : (\mathbf{DDz}) - \mathbf{yx} : (\mathbf{DDz}) \\ &= \mathbf{xy} : (\mathbf{D} \wedge \mathbf{Dz}).\end{aligned}$$

Therefore,

$$\mathfrak{R} \circ \mathbf{z} = \mathbf{D} \wedge \mathbf{Dz} \tag{59}$$

for arbitrary vector fields $\mathbf{z}(q)$. Observe also that Eqs. (56) and (59) also hold true when $\mathbf{z}(q)$ is a complex vector field.

Still another expression for $\mathfrak{R}$ is obtained by noting that $\mathbf{e}_\mu \mathbf{e}^\mu$ is the unit tensor; then by using Eq. (59) with $\mathbf{z} = \mathbf{e}_\mu$, we get

$$\mathfrak{R} = \mathfrak{R} \circ \mathbf{e}_\mu \mathbf{e}^\mu = (\mathbf{D} \wedge \mathbf{De}_\mu)\mathbf{e}^\mu. \tag{60}$$

IV. SPINOR CONNECTIONS

At each point q in the general relativity manifold $\mathcal{M}$ we associate a spinor space $(\mathcal{S}_2)_q$ and its corresponding conjugate spinor space $(\bar{\mathcal{S}}_2)_q$. The space $(\bar{\mathcal{S}}_2)_q \otimes_H (\mathcal{S}_2)_q$ of Hermitian tensors at q is isomorphic to the tangent space $\mathcal{M}_q$, and we shall treat them as being identical.[11] In addition to the notations such as $\mathbf{y}(q)$ and Y_q for a tangent vector at q, we will also use the notation $\mathbf{Y}(q)$ for the same vector whenever it is useful to emphasize the fact that it is also a spin tensor. In addition to the notation $\mathbf{y}(q) \circ \mathbf{z}(q)$ for inner products, we shall also use the notation $\mathbf{Y}(q) \odot \mathbf{Z}(q)$, which was introduced in I for the inner product of spinor tensors.

A *spinor connection* D'_X on $\mathcal{M}$ associates with each derivation operator X an operator D'_X on spinor fields such that if $u(q)$ is any spinor field with values in $(\mathcal{S}_2)_q$ then $D'_X u$ is also a spinor field with values in $(\mathcal{S}_2)_q$ where the following properties are satisfied:

$$\begin{aligned}&\text{(a)}\ D'_X(u+v) = D'_X u + D'_X v, \\ &\text{(b)}\ D'_X(fu) = (Xf)u + fD'_X u, \\ &\text{(c)}\ D'_{X+Y}u = D'_X u + D'_Y u, \\ &\text{(d)}\ D'_{gX}u = gD'_X u,\end{aligned} \tag{61}$$

where X, Y are derivation operators, $f(q)$ is any complex function, $g(q)$ is any real function, and $u(q)$ and $v(q)$ are any spinor fields with values in $(\mathcal{S}_2)_q$. The operator D'_X is also called the covariant derivative operator on spinors in the direction of $\mathbf{x}$. In a coordinate system q^μ we shall use the notation $D'_\mu \equiv D'_{\partial_\mu}$.

The covariant derivative $D'_X \bar{u}$ of spinor field $\bar{u}(q)$ with values in $(\bar{\mathcal{S}}_2)_q$ is defined as

$$D'_X\bar{u} = \overline{(D'_X u)}. \tag{62}$$

It follows that axioms (61) and Eq. (62) are satisfied if u and v are replaced by $\bar{u}$ and $\bar{v}$ respectively.

The covariant derivative $D'_X\psi$ of spinor field $\psi(q) = u(q) + \bar{v}(q)$ with values in $(\mathcal{S}_4)_q = (\mathcal{S}_2)_q \oplus (\bar{\mathcal{S}}_2)_q$ is defined as

$$D'_X\psi = D'_X u + D'_X \bar{v}. \tag{63}$$

It follows that axioms (61) and Eq. (62) are satisfied if u and v are replaced by ψ and ξ which have values in $(\mathcal{S}_4)_q$.

To any connection D'_X can be associated another connection D'^*_X, which we call the *dual connection*, and is defined by the equation

$$(D_X'^* u)\circ v = X(u\circ v) - u\circ(D_X' v) \tag{64}$$

for arbitrary $u(q)$ and $v(q)$. It is easily verified from this definition that $D_X'^*$ satisfies the axioms (61) for a connection. (The relation of the dual connection to the corresponding concept used in the literature is discussed in Appendix A.) Using the defining Equations (62) and (63) we extend the operation $D_X'^*$ to spinor fields $\bar{u}(q)$ and $\psi(q)$, and it follows that

$$(D_X'^*\bar{u})\blacktriangle\bar{v} = X(\bar{u}\blacktriangle\bar{v}) - \bar{u}\blacktriangle(D_X'\bar{v}) \tag{65}$$

and

$$(D_X'^*\psi)\blacktriangle\zeta = X(\psi\blacktriangle\zeta) - \psi\blacktriangle(D_X'\zeta) \tag{66}$$

The covariant derivative $D_X'\mathbf{A}$ for any second-order spinor tensor field $\mathbf{A}(q)$ with values in $(\mathcal{S}_4)_q\otimes(\mathcal{S}_4)_q$ can be defined by the equation

$$(D_X'\mathbf{A})\blacktriangle\!\!\blacktriangle\xi\eta = X(\mathbf{A}\blacktriangle\!\!\blacktriangle\xi\eta) - \mathbf{A}\blacktriangle\!\!\blacktriangle(D_X'^*\xi)\eta - \mathbf{A}\blacktriangle\!\!\blacktriangle\xi(D_X'^*\eta) \tag{67}$$

for arbitrary spinor fields $\xi(q)$ and $\eta(q)$ with values in $(\mathcal{S}_4)_q$. Note that if $\mathbf{A}(q)$ has its values in any of the subspaces of $(\mathcal{S}_4)_q\otimes(\mathcal{S}_4)_q$ such as $(\bar{\mathcal{S}}_2)_q\otimes(\mathcal{S}_2)_q$ or $(\bar{\mathcal{S}}_2)_q\otimes_H(\mathcal{S}_2)_q$ for all q then $D_X\mathbf{A}(q)$ will also have its values in the same subspace for all q.

It easily follows from the above definition and Eq. (66) that

$$\begin{aligned}
&D_X'(\psi\xi) = (D_X'\psi)\xi + \psi(D_X'\xi),\\
&D_X'(\mathbf{A}\blacktriangle\eta) = (D_X'\mathbf{A})\blacktriangle\eta + \mathbf{A}\blacktriangle(D_X'^*\eta),\\
&D_X'(\xi\blacktriangle\mathbf{A}) = (D_X'^*\xi)\blacktriangle\mathbf{A} + \xi\blacktriangle(D_X'\mathbf{A}),\\
&\overline{(D_X'\mathbf{A})} = D_X'\overline{\mathbf{A}}.
\end{aligned} \tag{68}$$

An obvious generalization of Eq. (67) can be used for defining the covariant derivative of higher order spinor tensor fields.

A spinor connection D_X' is said to be *compatible with the inner product* iff

$$X(\psi\blacktriangle\zeta) = (D_X'\psi)\blacktriangle\zeta + \psi\blacktriangle(D'_X\zeta) \tag{69}$$

for all spinor fields $\psi(q)$ and $\zeta(q)$. Note that D_X' is compatible with the inner product iff $D_X'^* = D_X'$.

We now show that a spinor connection D_X' is compatible with the spinor inner product iff $D_X'\mathbf{I}_2 = 0$ for every derivation operator X, where $\mathbf{I}_2(q)$ is the unit tensor in $(\mathcal{S}_2)_q\otimes(\mathcal{S}_2)_q$ for each q. From the definition of covariant differentiation of spinor tensors and the definition of $D_X'^*$ in Eq. (64), we get

$$\begin{aligned}
u\blacktriangle(D_X'\mathbf{I}_2)\blacktriangle v &= X(u\blacktriangle\mathbf{I}_2\blacktriangle v) - (D_X'^*u)\blacktriangle\mathbf{I}_2\cdot v - u\blacktriangle\mathbf{I}_2\blacktriangle(D_X'^*v)\\
&= X(u\blacktriangle v) - (D_X'^*u)\blacktriangle v - u\blacktriangle(D_X'^*v)\\
&= u\blacktriangle(D_X'v - D_X'^*v)
\end{aligned}$$

for arbitrary spinor fields $u(q)$ and $v(q)$. Consequently, $D_X'\mathbf{I}_2 = 0$ for arbitrary X iff $D_X' = D_X'^*$ for arbitrary X, which is true if D_X' is compatible with the spinor inner product. QED

A spinor connection D_X' is said to *generate the standard four-vector connection* iff $D_X'\mathbf{A} = D_X\mathbf{A}$ for all spinor tensor fields $\mathbf{A}(q)$ with values in $(\bar{\mathcal{S}}_2)_q\otimes_H(\mathcal{S}_2)_q$, i.e., for four-vector fields, where D_X is the standard four-vector connection.

We shall later show that there exists a unique spinor connection which has both properties of being compatible with the inner product and of generating the standard four-vector connection. From now on we shall use the symbol D_X, when operating on spinor and spinor tensors, to denote this spinor connection, and we shall call it the *standard spinor connection*.

Now we shall derive an expression for the contravariant components of the covariant derivative $D_\mu'u$ of a spinor field $u(q)$ with respect to a set of coordinates q^μ. Let $h_1(q)$, $h_2(q)$ be spinor fields that form a basis for $(\mathcal{S}_2)_q$ for each q. The spinor fields $h^1(q)$, $h^2(q)$ are defined to be a reciprocal basis for $(\mathcal{S}_2)_q$ for each q, i.e.,

$$h^a\blacktriangle h_b = \delta^a_b. \tag{70}$$

We can write the spinor field $u(q)$ in terms of its contravariant components $u^a(q)$ and covariant components $u_a(q)$ respectively as

$$u = u^a h_a = u_a h^a.$$

From this equation and Eq. (70) we get

$$u^a = h^a\blacktriangle u, \quad u_a = u\blacktriangle h_a.$$

The contravariant components $u'^a{}_{|\mu}$ of $D_\mu' u$, defined by the equation

$$D_\mu' u = u'^a{}_{|\mu} h_a,$$

are then

$$\begin{aligned}
u'^a{}_{|\mu} &= h^a\blacktriangle(D_\mu' u) = h^a\blacktriangle[D_\mu'(u^b h_b)]\\
&= h^a\blacktriangle[(\partial_\mu u^b)h_b + u^b(D_\mu' h_b)]\\
&= \partial_\mu u^a + \Lambda'^a_{\mu b}u^b,
\end{aligned} \tag{71}$$

where the spinor connection coefficients $\Lambda'^a_{\mu b}$ are defined by the equation

$$D_\mu' h_b = \Lambda'^a_{\mu b} h_a, \tag{72}$$

from which we get

$$\Lambda'^a_{\mu b} = h^a\blacktriangle(D_\mu' h_b). \tag{73}$$

For the covariant components $u'_{a|\mu}$ of $D_\mu' u$, defined by the equation

$$D_\mu' u = u'_{a|\mu} h^a,$$

we get

$$\begin{aligned}
u'_{a|\mu} &= (D_\mu' u)\blacktriangle h_a = [D_\mu'(u_b h^b)]\blacktriangle h_a\\
&= [(\partial_\mu u_b)h^b + u_b(D_\mu' h^b)]\blacktriangle h_a.
\end{aligned}$$

Moreover, by making use of

$$\begin{aligned}
(D_\mu' h^b)\blacktriangle h_a &= \partial_\mu(h^b\blacktriangle h_a) - h^b\blacktriangle(D_\mu'^* h_a)\\
&= \partial_\mu\delta^b_a - \Lambda'^{*b}_{\mu a} = -\Lambda'^{*b}_{\mu a},
\end{aligned}$$

where $\Lambda'^{*b}_{\mu a}$ are the spinor connection coefficients for the dual connection $D_\mu'^*$, we find

$$u'_{a|\mu} = \partial_\mu u_a - \Lambda'^{*b}_{\mu a} u_b. \tag{74}$$

In the case where the connection D_μ' is compatible with the spinor inner product, it immediately follows that

$$\Lambda'^{*b}_{\mu a} = \Lambda'^b_{\mu a}.$$

In the case of the standard spinor connection D_μ we shall denote the connection coefficients by $\Lambda^b_{\mu a}$ [an explicit expression for these in terms of the Christoffel symbols is given in Eq. (119) below].

By arguments analogous to those used for arriving at Eq. (54) we shall prove that any spinor connection D_X' differs from D_X by a linear transformation. To this end, take a spinor basis $\boldsymbol{h}_a(q)$ for each $(\mathcal{S}_2)_q$. Then by virtue of the axioms for a spinor connection, we get

$$\begin{aligned}(D_X' - D_X)\boldsymbol{u} &= (D_X' - D_X)(u^a \boldsymbol{h}_a) = u^a (D_X' - D_X)\boldsymbol{h}_a \\ &= [(D_X' - D_X)\boldsymbol{h}_a]\boldsymbol{h}^a \blacktriangle \boldsymbol{u} = \mathbf{K}_X \blacktriangle \boldsymbol{u},\end{aligned}$$

where the operator $\mathbf{K}_X(q)$ defined as

$$\mathbf{K}_X = [(D_X' - D_X)\boldsymbol{h}_a]\boldsymbol{h}^a \tag{75}$$

is an element of $(\mathcal{S}_2)_q \otimes (\mathcal{S}_2)_q$ for each q, and does not depend on the spinor field $\boldsymbol{u}$. Hence $(D_X' - D_X)\boldsymbol{u}$ is linear in u and its value at q depends on the value of u only at q. Since D_X' and D_X are linear in X, so is $\mathbf{K}_X$. We can therefore write

$$D_X'\boldsymbol{u} = D_X \boldsymbol{u} + \mathbf{K}_X \blacktriangle \boldsymbol{u}. \tag{76}$$

The complex conjugate of this equation with $\bar u$ replaced by $\bar v$ is

$$D_X'\bar{v} = D_X \bar{v} + \bar{\mathbf{K}}_X \blacktriangle \bar{v}. \tag{77}$$

Summing these two equations gives

$$D_X'\psi = D_X\psi + \mathbf{M}_X \blacktriangle \psi, \tag{78}$$

where

$$\mathbf{M}_X(q) = \mathbf{K}_X(q) + \bar{\mathbf{K}}_X(q), \tag{79}$$

and $\psi(q) = \boldsymbol{u}(q) + \bar{v}(q)$ is an arbitrary spinor field with values in $(\mathcal{S}_4)_q$.

The above results allow us to express the dual connection $D_X'^*$ in terms of D_X. We merely substitute Eq. (76) into Eq. (64) to get

$$\begin{aligned}(D_X'^*\boldsymbol{u}) \blacktriangle \boldsymbol{v} &= X(\boldsymbol{u} \blacktriangle \boldsymbol{v}) - \boldsymbol{u} \blacktriangle (D_X' v) \\ &= X(\boldsymbol{u} \blacktriangle \boldsymbol{v}) - \boldsymbol{u} \blacktriangle (D_X v + \mathbf{K}_X \blacktriangle \boldsymbol{v}) \\ &= (D_X \boldsymbol{u}) \blacktriangle \boldsymbol{v} - \boldsymbol{u} \blacktriangle \mathbf{K}_X \blacktriangle \boldsymbol{v} \\ &= (D_X u + \tilde{K}_X \blacktriangle u) \blacktriangle v,\end{aligned}$$

for arbitrary spinor fields v. Hence

$$D_X'^*\boldsymbol{u} = D_X \boldsymbol{u} + \tilde{\mathbf{K}}_X \blacktriangle \boldsymbol{u}. \tag{80}$$

From this it follows also that

$$D_X'^*\bar{v} = D_X \bar{v} + \tilde{\bar{K}}_X \blacktriangle \bar{v} \tag{81}$$

and

$$D_X'^*\psi = D_X\psi + \tilde{\mathbf{M}}_X \blacktriangle \psi. \tag{82}$$

Moreover, by virtue of Eq. (76), Eq. (80) can be expressed as

$$\begin{aligned}D_X'^*\boldsymbol{u} &= D_X'\boldsymbol{u} - (\mathbf{K}_X - \tilde{\mathbf{K}}_X) \blacktriangle \boldsymbol{u} \\ &= D_X'\boldsymbol{u} + \tfrac{1}{2}(\mathbf{K}_X - \tilde{\mathbf{K}}_X)_s \mathbf{I}_2 \blacktriangle \boldsymbol{u} \\ &= D_X'\boldsymbol{u} + (\mathbf{K}_X)_s \mathbf{I}_2 \blacktriangle \boldsymbol{u},\end{aligned} \tag{83}$$

where we have made use of Eq. (A9) of I. We also have

$$D_X'^*\bar{v} = D_X'\bar{v} + (\bar{\mathbf{K}}_X)_s \bar{\mathbf{I}}_2 \blacktriangle \bar{v} \tag{84}$$

and

$$D_X'^*\psi = D_X'\psi + [(\mathbf{K}_X)_s \mathbf{I}_2 + (\bar{\mathbf{K}}_X)_s \bar{\mathbf{I}}_2] \blacktriangle \psi \tag{85}$$

We can now rewrite Eq. (64) as

$$\begin{aligned}X(\boldsymbol{u} \blacktriangle \boldsymbol{v}) &= (D_X'^*\boldsymbol{u}) \blacktriangle \boldsymbol{v} + \boldsymbol{u} \blacktriangle (D_X' \boldsymbol{v}) \\ &= (D_X'\boldsymbol{u}) \blacktriangle \boldsymbol{v} + \boldsymbol{u} \blacktriangle (D_X' \boldsymbol{v}) + (\mathbf{K}_X)_s \boldsymbol{u} \blacktriangle \boldsymbol{v}\end{aligned} \tag{86}$$

and Eq. (67) as

$$\begin{aligned}\zeta \blacktriangle (D_X'\mathbf{A}) \blacktriangle \eta &= X(\zeta \blacktriangle \mathbf{A} \blacktriangle \eta) - (D_X'^*\zeta) \blacktriangle \mathbf{A} \blacktriangle \eta - \zeta \blacktriangle \mathbf{A} \blacktriangle (D_X'^*\eta) \\ &= X(\zeta \blacktriangle \mathbf{A} \blacktriangle \eta) - (D_X\zeta) \blacktriangle \mathbf{A} \blacktriangle \eta - (\tilde{\mathbf{M}}_X \blacktriangle \zeta) \blacktriangle \mathbf{A} \blacktriangle \eta \\ &\quad - \zeta \blacktriangle \mathbf{A} \blacktriangle (D_X\eta) - \zeta \blacktriangle \mathbf{A} \blacktriangle \tilde{\mathbf{M}}_X \blacktriangle \eta \\ &= \zeta \blacktriangle (D_X\mathbf{A}) \blacktriangle \eta + \zeta \blacktriangle \mathbf{M}_X \blacktriangle \mathbf{A} \blacktriangle \eta - \zeta \blacktriangle \mathbf{A} \blacktriangle \tilde{\mathbf{M}}_X \blacktriangle \eta.\end{aligned}$$

Therefore

$$D_X'\mathbf{A} = D_X\mathbf{A} + \mathbf{M}_X \blacktriangle \mathbf{A} - \mathbf{A} \blacktriangle \tilde{\mathbf{M}}_X. \tag{87}$$

In particular if $\mathbf{A}(q)$ has values in the subspace $(\bar{\mathcal{S}}_2)_q \otimes (\mathcal{S}_2)_q$, then Eq. (87) becomes

$$D_X'\mathbf{A} = D_X\mathbf{A} + \bar{\mathbf{K}}_X \blacktriangle \mathbf{A} - \mathbf{A} \blacktriangle \tilde{\mathbf{K}}_X. \tag{88}$$

We next prove three theorems based on the above results.

Theorem 1: A spinor connection D_X' is compatible with the spinor inner product iff $\tilde{\mathbf{K}}_X = \mathbf{K}_X$ (i.e., $\mathbf{K}_X$ is symmetric).

Proof: From Eq. (86) we see that the connection D_X' is compatible with the inner product iff $(\mathbf{K}_X)_s = 0$. Moreover, in the appendix of I we have shown that $(\mathbf{K}_X)_s = 0$ iff $\mathbf{K}_X$ is symmetric. QED

Theorem 2: A spinor connection D_X' generates the standard four-vector connection iff $\mathbf{K}_X$ can be expressed as

$$\mathbf{K}_X = i(\mathbf{x} \cdot \varphi)\mathbf{I}_2 \tag{89}$$

where $\varphi(q)$ is a real four-vector field.

Proof: From Eq. (88) we see that the spinor connection D_X' generates the standard four-vector connection iff

$$\bar{\mathbf{K}}_X \blacktriangle \mathbf{A} - \mathbf{A} \blacktriangle \tilde{\mathbf{K}}_X = 0 \tag{90}$$

for arbitrary $\mathbf{A}(q)$ with values in $(\bar{\mathcal{S}}_2)_q \otimes_H (\mathcal{S}_2)_q$. To prove the necessity of Eq. (89), let us suppose that D_X' generates the standard four-vector connection. Choose $\mathbf{A} = \bar{\boldsymbol{u}}\boldsymbol{u}$ where $\boldsymbol{u}(q)$ is an arbitrary spinor field. Then Eq. (90) becomes

$$(\bar{\mathbf{K}}_X \blacktriangle \bar{\boldsymbol{u}})\boldsymbol{u} + \bar{u}(\mathbf{K}_X \blacktriangle \boldsymbol{u}) = 0.$$

It follows from this that

$$\mathbf{K}_X \blacktriangle \boldsymbol{u} = \alpha_X \boldsymbol{u} \tag{91}$$

and

$$\bar{\alpha}_X \bar{\boldsymbol{u}}\boldsymbol{u} + \alpha_X \bar{\boldsymbol{u}}\boldsymbol{u} = 0,$$

where $\alpha_X(q)$ is scalar function. Consequently, $\alpha_X(q)$ must be linear in X and pure imaginary; thus it can be written as $\alpha_X = i\mathbf{x} \cdot \varphi$ with $\varphi(q)$ being a real four-vector field. Equation (89) is an immediate consequence of Eq. (91). To prove the sufficiency of Eq. (89), we need only to substitute it into the left-hand side of Eq. (90) to verify that it is true. QED

Theorem 3: A spinor connection D'_X has both the properties of being compatible with the spinor inner product and of generating the standard four-vector connection iff $\mathbf{K}_X = 0$ everywhere.

Proof: This follows from the two previous theorems and the fact that $\mathbf{K}_X$ can be both symmetric and antisymmetric iff $\mathbf{K}_X = 0$. QED

Theorem 3 establishes the uniqueness of the standard spinor connection.

The consequences of these three theorems on the form of Eqs. (86) and (88) can be summarized in the following four cases:

Case 1: $\widetilde{\mathbf{K}}_X = \mathbf{K}_X \neq 0$. i.e., $\mathbf{K}_X$ is symmetric.

$$X(\boldsymbol{u}\blacktriangle\boldsymbol{v}) = (D'_X\boldsymbol{u})\blacktriangle\boldsymbol{v} + \boldsymbol{u}\blacktriangle(D'_X\boldsymbol{v}), \tag{92}$$

$$D'_X\mathbf{A} = D_X\mathbf{A} + \overline{\widetilde{\mathbf{K}}}_X\blacktriangle\mathbf{A} - \mathbf{A}\blacktriangle\widetilde{\mathbf{K}}_X; \tag{93}$$

thus $D'_X\mathbf{A} \neq D_X\mathbf{A}$ except where $\mathbf{K}_X = 0$ or $\mathbf{A} = 0$.

Case 2: $\mathbf{K}_X = i\mathbf{x}\cdot\varphi\mathbf{I}_2$, where $\varphi(q) \neq 0$ is a real four-vector field:

$$X(\boldsymbol{u}\blacktriangle\boldsymbol{v}) = (D'_X\boldsymbol{u})\blacktriangle\boldsymbol{v} + \boldsymbol{u}\blacktriangle(D'_X\boldsymbol{v}) - 2i\mathbf{x}\cdot\varphi(\boldsymbol{u}\blacktriangle\boldsymbol{v}), \tag{94}$$

$$D'_X\mathbf{A} = D_X\mathbf{A}. \tag{95}$$

Case 3: $\mathbf{K}_X \equiv \mathbf{x}\cdot\varphi\mathbf{I}_2$ where $\varphi(q) \neq 0$ is a real four-vector field:

$$X(\boldsymbol{u}\blacktriangle\boldsymbol{v}) = (D'_X\boldsymbol{u})\blacktriangle\boldsymbol{v} + \boldsymbol{u}\blacktriangle(D'_X\boldsymbol{v}) - 2\mathbf{x}\cdot\varphi\boldsymbol{u}\blacktriangle\boldsymbol{v}, \tag{96}$$

$$D'_X\mathbf{A} = D_X\mathbf{A} + 2\mathbf{x}\cdot\varphi\mathbf{A}. \tag{97}$$

Case 4: $\mathbf{K}_X \equiv 0$. Then $D'_X\boldsymbol{u} = D_X\boldsymbol{u}$ and

$$X(\boldsymbol{u}\blacktriangle\boldsymbol{v}) = (D'_X\boldsymbol{u})\blacktriangle\boldsymbol{v} + \boldsymbol{u}\blacktriangle(D'_X\boldsymbol{v}), \tag{98}$$

$$D'_X\mathbf{A} = D_X\mathbf{A}. \tag{99}$$

The covariant derivatives $D'_X\mathbf{A}$ [defined by Eq. (67)] and $D'^{*}_X\mathbf{A}$ of a spinor tensor field $\mathbf{A}(q)$ with values in $(\mathcal{S}_4)_q \otimes (\mathcal{S}_4)_q$ are not the only ones possible within the general framework of our formalism. We can also define the covariant derivatives $D'^{(\cdot*)}_X\mathbf{A}$ and $D'^{(*\cdot)}_X\mathbf{A}$, by means of the equations

$$\begin{aligned}
(D'^{(\cdot*)}_X\mathbf{A})\stackrel{\blacktriangle}{\blacktriangle}\xi\eta &= X(\mathbf{A}\stackrel{\blacktriangle}{\blacktriangle}\xi\eta) - \mathbf{A}\stackrel{\blacktriangle}{\blacktriangle}(D'^{*}_X\xi)\eta - \mathbf{A}\stackrel{\blacktriangle}{\blacktriangle}\xi(D'_X\eta),\\
(D'^{(*\cdot)}_X\mathbf{A})\stackrel{\blacktriangle}{\blacktriangle}\xi\eta &= X(\mathbf{A}\stackrel{\blacktriangle}{\blacktriangle}\xi\eta) - \mathbf{A}\stackrel{\blacktriangle}{\blacktriangle}(D'_X\xi)\eta - \mathbf{A}\stackrel{\blacktriangle}{\blacktriangle}\xi(D'^{*}_X\eta)
\end{aligned} \tag{100}$$

for arbitrary spinor fields $\xi(q)$ and $\eta(q)$ with values in $(\mathcal{S}_4)_q$. Also, for purposes of comparison, we include the defining equation for $D'^{*}_X\mathbf{A}$,

$$(D'^{*}_X\mathbf{A})\stackrel{\blacktriangle}{\blacktriangle}\xi\eta = X(\mathbf{A}\stackrel{\blacktriangle}{\blacktriangle}\xi\eta) - \mathbf{A}\stackrel{\blacktriangle}{\blacktriangle}(D'_X\xi)\eta - \mathbf{A}\stackrel{\blacktriangle}{\blacktriangle}\xi(D'_X\eta)$$

obtained by interchanging D'_X and D'^{*}_X in Eq. (67).

The following identities, which are easily proved, illustrate how these covariant derivatives operate on the tensor files by consideration of the simple case where $\mathbf{A} = \psi\zeta$:

$$\begin{aligned}
D'_X(\psi\zeta) &= (D'_X\psi)\zeta + \psi(D'_X\zeta),\\
D'^{*}_X(\psi\zeta) &= (D'^{*}_X\psi)\zeta + \psi(D'^{*}_X\zeta),\\
D'^{(\cdot*)}_X(\psi\zeta) &= (D'_X\psi)\zeta + \psi(D'^{*}_X\zeta),\\
D'^{(*\cdot)}_X(\psi\zeta) &= (D'^{*}_X\psi)\zeta + \psi(D'_X\zeta).
\end{aligned} \tag{101}$$

It also follows readily that

$$D'_X(\mathbf{A}\blacktriangle\xi) = (D'^{(\cdot*)}_X\mathbf{A})\blacktriangle\xi + \mathbf{A}\blacktriangle(D'_X\xi). \tag{102}$$

This shows how the operator $D'^{(\cdot*)}_X$ on second order spinor tensors $\mathbf{A}$ arises naturally when that tensor field is being used for expressing a linear transformation $\mathbf{A}\blacktriangle\xi$ of the spinor field ξ.

That $D'^{(\cdot*)}_X$ is the natural operator to use on second order spinor tensors when viewed as linear transformations on spinor fields can also be seen when considering the parallel transportation of spinors along a curve $q = q(t)$ in $\mathcal{M}$. To this end let $\xi(t)$ and $\eta(t)$ be functions with values in $(\mathcal{S}_4)_{q(t)}$ and $\mathbf{A}(t)$ a function with values in $(\mathcal{S}_4)_{q(t)} \otimes (\mathcal{S}_4)_{q(t)}$ for each t. Suppose that

$$\eta(t_0) = \mathbf{A}(t_0)\blacktriangle\xi(t_0).$$

Here $\mathbf{A}(t_0)$ is acting as a linear transformation on $\xi(t_0)$ to produce $\eta(t_0)$. Suppose that $\xi(t)$, for each t, is given as the parallel transport of $\xi(t_0)$ via the connection D'_X from $q(t_0)$ along the curve to $q(t)$, and likewise $\eta(t)$ for each t is given as the parallel transport of $\eta(t_0)$ via the connection D'_X along the curve from $q(t_0)$ to $q(t)$, i.e., $\xi(t)$ and $\eta(t)$ satisfy the differential equations

$$D'_{(d/dt)}\xi(t) = 0, \quad D'_{(d/dt)}\eta(t) = 0.$$

Now, if $\mathbf{A}(t)$ is given as the parallel transport of $\mathbf{A}(t_0)$ via the operator $D'^{(\cdot*)}_X$ along the curve from $q(t_0)$ to $q(t)$, i.e., it satisfies the differential equation

$$D'^{(\cdot*)}_{(d/dt)}\mathbf{A}(t) = 0;$$

then by using Eq. (102) we get

$$\begin{aligned}
D'_{(d/dt)}[\mathbf{A}(t)\blacktriangle\xi(t)] &= [D'^{(\cdot*)}_{(d/dt)}\mathbf{A}(t)]\blacktriangle\xi(t) + \mathbf{A}(t)\blacktriangle[D'_{(d/dt)}\xi(t)]\\
&= 0
\end{aligned}$$

which is the same differential equation for $\mathbf{A}(t)\circ\xi(t)$ as the differential equation for $\eta(t)$ with the same initial values at t_0. Consequently, by a uniqueness theorem for differential equations, we have

$$\eta(t) = \mathbf{A}(t)\blacktriangle\xi(t).$$

In summary, if $\xi(t)$ and $\eta(t)$ are given by parallel transport of their initial values at $q(t_0)$ via the operator D'_X along the curve, and if $\mathbf{A}(t)$ is given by parallel transport of its initial value at $q(t_0)$ via the operator $D'^{(\cdot*)}_X$ along the curve, then the relation $\eta(t_0) = \mathbf{A}(t_0)\blacktriangle\xi(t_0)$ at $q(t_0)$ implies that the corresponding relation $\eta(t) = \mathbf{A}(t)\blacktriangle\xi(t)$ also holds at all points $q(t)$ along the curve.

The covariant derivative operator $D'^{(\cdot*)}_X$ acting on four-vector fields leads to an alternate way of generating the standard four-vector connection. We shall say that a spinor connection D'_X *alternately generates the standard four-vector connection* iff $D'^{(\cdot*)}_X\mathbf{A} = D_X\mathbf{A}$ for all spinor tensor fields $\mathbf{A}(q)$ with values in $(\overline{\mathcal{S}}_2)_q \otimes_H (\mathcal{S}_2)_q$, where D_X is the standard four-vector connection.

In this context, the alternate to Theorem 2 is

Theorem 4: A spinor connection D'_X alternately generates the standard four-vector connection iff $\mathbf{K}_X$ can be expressed as

$$\mathbf{K}_X = \mathbf{x}\cdot\varphi\mathbf{I}_2$$

where $\varphi(q)$ is a real four-vector field.

Proof: By definition of $D'^{(\cdot*)}_X\mathbf{A}$ for $\mathbf{A}(q)$ having values in

$(\mathcal{S}_4)_q \otimes (\mathcal{S}_4)_q$ and using Eqs. (78) and (82) gives

$$\begin{aligned}\zeta\blacktriangle(D_X'^{(\cdot*)}\mathbf{A})\blacktriangle\eta &= X(\zeta\blacktriangle\mathbf{A}\blacktriangle\eta) - (D_X'^{*}\zeta)\blacktriangle\mathbf{A}\blacktriangle\eta - \zeta\blacktriangle\mathbf{A}\blacktriangle(D_X'\eta)\\ &= X(\zeta\blacktriangle\mathbf{A}\blacktriangle\eta) - (D_X\zeta)\blacktriangle\mathbf{A}\blacktriangle\eta - \zeta\blacktriangle\mathbf{A}\blacktriangle(D_X\eta)\\ &\quad - (\widetilde{\mathbf{M}}_X\blacktriangle\zeta)\blacktriangle\mathbf{A}\blacktriangle\eta - \zeta\blacktriangle\mathbf{A}\blacktriangle(\mathbf{M}_X\blacktriangle\eta)\\ &= \zeta\blacktriangle(D_X\mathbf{A})\blacktriangle\eta + \zeta\blacktriangle(\mathbf{M}_X\blacktriangle\mathbf{A} - \mathbf{A}\blacktriangle\mathbf{M}_X)\blacktriangle\eta\end{aligned}$$

for arbitrary spinor fields ζ, η; therefore,

$$D_X'^{(\cdot*)}\mathbf{A} = D_X\mathbf{A} + \mathbf{M}_X\blacktriangle\mathbf{A} - \mathbf{A}\blacktriangle\mathbf{M}_X. \tag{103}$$

In particular if $\mathbf{A}(q)$ has its values in $(\bar{\mathcal{S}}_2)_q \otimes (\mathcal{S}_2)_q$, then by Eq. (79) we get

$$D_X'^{(\cdot*)}\mathbf{A} = D_X\mathbf{A} + \overline{\mathbf{K}}_X\blacktriangle\mathbf{A} - \mathbf{A}\blacktriangle\mathbf{K}_X \tag{104}$$

Thus, in order that $D_X'^{(\cdot*)}\mathbf{A} = D_X\mathbf{A}$, it is necessary and sufficient that

$$\overline{\mathbf{K}}_X\blacktriangle\mathbf{A} - \mathbf{A}\blacktriangle\mathbf{K}_X = 0.$$

By an argument entirely analogous to the one used in Theorem 2 we can show that this is true for all $\mathbf{A}$ iff $\mathbf{K}_X$ has the form

$$\mathbf{K}_X = \mathbf{x}\cdot\varphi\mathbf{I}_2 \qquad \text{QED}$$

We can now give a correct interpretation of the paper of Infeld and van der Waerden, the review paper of Bade and Jehle, as well as other papers in the literature based on theirs.[4] In essence, their approach consists in taking the covariant derivative according to one connection when the spinor is expressed by contravariant components, and taking the covariant derivative according to what we call the dual connection when the spinor is expressed by covariant components (see Appendix A regarding our interpretation of expressions involving both covariant and contravariant spinors). To elaborate this point, let D_X' be an arbitrary spinor connection that generates the standard four-vector connection but is not necessarily compatible with the spinor inner product. It follows from Eq. (80) and Theorem 2 that the dual connection $D_X'^{*}$ also generates the standard four-vector connection. For a spinor field $\boldsymbol{u}(q)$ with values in $(\mathcal{S}_2)_q$, the contravariant components $u'^{a}{}_{|\mu}$ and the covariant components $u'_{a|\mu}$ of $D_\mu' u$ are defined by the equations

$$D_\mu' u = u'^{a}{}_{|\mu}\, h_a = u'_{a|\mu} h^a.$$

Also, the contravariant components $u'^{*a}{}_{|\mu}$ and the covariant components $\mu'^{*}{}_{a|\mu}$ of $D_\mu'^{*}u$ are defined by the equations

$$D_\mu'^{*}u = u'^{*a}{}_{|\mu} h_a = u'^{*}{}_{a|\mu} h^a.$$

Infeld and van der Waerden use only the contravariant components $\boldsymbol{u}'^{a}{}_{|\mu}$ of $D_\mu' \boldsymbol{u}$ and the covariant components $u'^{*}{}_{a|\mu}$ of $D_\mu'^{*}u$ in their theory. Their Eqs. (16) would then be expressions for these two quantities, but without any marks such as the asterisk to make a distinction between the dual covariant derivative and the original covariant derivative. According to this view of their formalism, the covariant derivative operator D_μ' is the indicated one to use whenever contravariant spinor indices occur, and the covariant derivative operation $D_\mu'^{*}$ is indicated whenever covariant spinor indices occur. The product rule then follows easily from the equation

$$\partial_\mu(\boldsymbol{u}\blacktriangle\boldsymbol{v}) = (D_\mu'^{*}\boldsymbol{u})\blacktriangle\boldsymbol{v} + \boldsymbol{u}\blacktriangle(D_\mu'\boldsymbol{v})$$

to give

$$\partial_\mu(u_a v^a) = u'^{*}_{a|\mu} v^a + u_a v'^{a}{}_{|\mu}. \tag{105}$$

The generalization of these results to spinor tensors of second or higher order in the Infeld—van der Waerden formalism when indices are all contravariant or all covariant is evident: D_μ' is used in the former case and $D_\mu'^{*}$ in the latter.

In the case of mixed spinor indices their formalism implies that the operation of covariant differentiation on such a tensor is the joint operation of D_μ' and $D_\mu'^{*}$ on the files with contravariant and covariant indices, respectively. For example, for a spinor tensor field $\mathbf{N}(q)$ with values in $(\mathcal{S}_2)_q \otimes (\mathcal{S}_2)_q$ take the covariant derivative $D_\mu'^{(\cdot*)}\mathbf{N}$ as defined by Eq. (100). In the Infeld—van der Waerden component formalism this would be denoted by making the first spinor index contravariant and the second spinor index covariant, i.e., it would be written as $\mathbf{N}^{a}{}_{b|\mu}$ in their notation.

In summary a desirable feature for a spinor connection is that it produce the standard four-vector connection by some prescription; thus the usual four-vector calculus will result in a natural way from spinor calculus. Both the spinor connections for which $\mathbf{K}_X = i\mathbf{x}\cdot\varphi\mathbf{I}_2$ and $\mathbf{K}_X = \mathbf{x}\cdot\varphi\mathbf{I}_2$ have this desirable feature, but in different ways. The first one "generates" the standard four-vector connection as seen by Theorem 2, and the second one "alternatively generates" the standard four-vector connection as seen by Theorem 4. The first one corresponds to Infeld and van der Waerden's connection as is indicated by the requirement stated in their Eqs. (19) or (20). This is also the connection considered by Schmutzer,[12] who follows closely Infeld and van der Waerden's approach.

Moreover, for the extension of arbitrary spinor connections D_X' on $(\mathcal{S}_2)_q$ valued spinors to give a bispinor connection [i.e., a connection for $(\mathcal{S}_4)_q$ valued spinors], our Eq. (63) leads to

$$D_X'\psi = D_X'(u + \bar{v}) = D_X'u + D_X'\bar{v} = D_X\psi + \mathbf{M}_X\blacktriangle\psi,$$

where

$$\mathbf{M}_X = \mathbf{K}_X + \overline{\mathbf{K}}_X.$$

Using the connection which "alternately generates" the standard four-vector connection, i.e., taking $\mathbf{K}_X = \mathbf{x}\cdot\varphi\mathbf{I}_2$ yields

$$\mathbf{M}_X = \mathbf{x}\cdot\varphi\mathbf{I}. \tag{106}$$

Using the connection which "generates" the standard four-vector connection, i.e., taking

$$\mathbf{K}_X = i\mathbf{x}\cdot\varphi\mathbf{I}_2,$$

yields

$$\mathbf{M}_X = \mathbf{x}\cdot\varphi\mathbf{\Gamma}^5. \tag{107}$$

On the other hand, according to our intrinsic interpretation of Schmutzer's[13] equations (37), (38), (39) for his bispinor connection, which we denote as $D_X'^{(B)}$, one has

$$D_X'^{(B)}\psi = D_X'u + D_X'^{*}\bar{v} = D_X\psi + \mathbf{N}_X\blacktriangle\psi, \tag{108}$$

where $\mathbf{M}_X$ is replaced by $\mathbf{N}_X$ given as

$$\mathbf{N}_X = \mathbf{K}_X + \tilde{\tilde{\mathbf{K}}}_X$$

and $\mathbf{K}_X = i\mathbf{x}\cdot\varphi\mathbf{I}_2$. Hence

$$\mathbf{N}_X = i\mathbf{x}\cdot\varphi\mathbf{I}.$$

As another interesting result of our formalism we show how a simple and elegant derivation of an equation relating the standard spinor connection to the standard four-vector connection may be obtained. For any derivation operator X and spinor fields $t(q)$, $u(q)$, $v(q)$ with values in $(\mathcal{S}_2)_q$, we have

$$\begin{aligned}\bar{t}\blacktriangle\bar{u}(D_Xv) &= \tfrac{1}{2}\{\bar{t}\blacktriangle\bar{u}(D_Xv) - \bar{u}\blacktriangle\bar{t}(D_Xv)\}\\ &= \tfrac{1}{2}\{\bar{t}\blacktriangle[D_X(\bar{u}v) - (D_X\bar{u})v] - \bar{u}\blacktriangle[D_X(\bar{t}v) - (D_X\bar{t})v]\}\\ &= \tfrac{1}{2}\{\bar{t}\blacktriangle[D_X(\bar{u}v)] - \bar{u}\blacktriangle[D_X(\bar{t}v)] - \bar{t}\blacktriangle(D_X\bar{u})v\\ &\quad - (D_X\bar{t})\blacktriangle\bar{u}v\}\\ &= \tfrac{1}{2}\{\bar{t}\blacktriangle[D_X(\bar{u}v)] - \bar{u}\blacktriangle[D_X(\bar{t}v)] - [X(\bar{t}\blacktriangle\bar{u})]v\},\end{aligned}$$

i.e.,

$$\bar{t}\blacktriangle\bar{u}(D_Xv) = \tfrac{1}{2}\{\bar{t}\blacktriangle[D_X(\bar{u}v)] - \bar{u}\blacktriangle[D_X(\bar{t}v)] - [X(\bar{t}\blacktriangle\bar{u})]v\}. \tag{109}$$

On the left side of this equation, D_X is the standard spinor connection operating on the spinor field v. On the right side, D_X is the standard spinor connection operating on the spinor tensor fields $\bar{u}v$ and $\bar{t}v$. But such spinor tensors are also regarded as complex four-vectors; thus, due to the fact that the standard spinor connection generates the standard four-vector connection, the operator D_X on the right side is also interpreted as the standard four-vector connection operating on complex four-vector fields.

In order to derive Eq. (109), the existence of the standard spinor connection was assumed. A brief sketch of this existence proof is as follows: We start by defining D_Xv for spinor fields v by means of Eq. (109), where the operator D_X on the right side of the equation is the standard four-vector connection operating on complex four-vector fields. We must show that the equation is consistent in that the right side is actually a quantity that can be expressed in the form appearing on the left side. It is easily shown that the right side is linear in $\bar{t}$ and in $\bar{u}$ and its value at every point q depends on the value of $\bar{t}$ and $\bar{u}$ at the point q only. Furthermore, the right side is also antisymmetric under the interchange of $\bar{t}$ and $\bar{u}$. It follows from these facts that the right side is expressible in the form $\bar{t}\blacktriangle\bar{u}$ times a spinor; this is consistent with the form of the expression on the left side in that it is written as $\bar{t}\blacktriangle\bar{u}$ times a spinor D_Xv. Thus this equation is consistent and therefore defines the spinor field D_Xv, i.e., the operation of D_X on each spinor field v is defined. One can then proceed to show that this operator D_X on spinor fields satisfies all the properties of a spinor connection, that it is compatible with the spinor inner product, and that it generates the standard four-vector connection. This completes the proof that the operator D_X on spinor fields is the standard spinor connection, and its existence is then established.

For the purpose of re-expressing the above results in another convenient form, let us take a basis $h_a(q)$ for each $(\mathcal{S}_2)_q$ and note that h_dh^d is the unit tensor in $(\mathcal{S}_2)_q \otimes (\mathcal{S}_2)_q$ for each q. Then

$$-2 = h_c\blacktriangle h^c = h_c\blacktriangle(h_dh^d)\blacktriangle h^c = -h_c\blacktriangle h_dh^c\blacktriangle h^d,$$

from which we get

$$\begin{aligned}&\tfrac{1}{2}h^c\blacktriangle h^dh_c\blacktriangle h_d = 1,\\ &\tfrac{1}{2}\bar{h}^c\blacktriangle\bar{h}^d\bar{h}_c\blacktriangle\bar{h}_d = 1.\end{aligned} \tag{110}$$

Using this, we write

$$D_Xv = \tfrac{1}{2}\bar{h}^c\blacktriangle\bar{h}^d\bar{h}_c\blacktriangle\bar{h}_dD_Xv,$$

and from Eq. (109) with $\bar{t}=\bar{h}_c$ and $\bar{u}=\bar{h}_d$ we find

$$\begin{aligned}D_Xv &= \tfrac{1}{4}\bar{h}^c\blacktriangle\bar{h}^d\{\bar{h}_c\blacktriangle[D_X(\bar{h}_dv)] - \bar{h}_d\blacktriangle[D_X(\bar{h}_cv)] - [X(\bar{h}_c\blacktriangle\bar{h}_d)]v\}\\ &= -\tfrac{1}{4}\bar{h}^d\blacktriangle\bar{h}^c\bar{h}_c\blacktriangle[D_X(\bar{h}_dv)] - \tfrac{1}{4}\bar{h}^c\blacktriangle\bar{h}^d\bar{h}_d\blacktriangle[D_X(\bar{h}_cv)]\\ &\quad - \tfrac{1}{4}\bar{h}^c\blacktriangle\bar{h}^d[X(\bar{h}_c\blacktriangle\bar{h}_d)]v\\ &= \tfrac{1}{2}\bar{h}^c\blacktriangle[D_X(\bar{h}_cv)] - \tfrac{1}{4}\bar{h}^c\blacktriangle\bar{h}^d[X(\bar{h}_c\blacktriangle\bar{h}_d)]v.\end{aligned}$$

But since $\bar{h}^c\blacktriangle\bar{h}^d$ and $\bar{h}_c\blacktriangle\bar{h}_d$ are zero when $c=d$, we can write the preceding equation as

$$D_Xv = \tfrac{1}{2}\bar{h}^c\blacktriangle[D_X(\bar{h}_cv)] - \tfrac{1}{2}\bar{h}^1\blacktriangle\bar{h}^2[X(\bar{h}_1\blacktriangle\bar{h}_2)]v. \tag{111}$$

For the same reason, Eq. (110) becomes

$$\bar{h}^1\blacktriangle\bar{h}^2\bar{h}_1\blacktriangle\bar{h}_2 = 1.$$

Therefore,

$$\bar{h}^1\blacktriangle\bar{h}^2 = (\bar{h}_1\blacktriangle\bar{h}_2)^{-1}. \tag{112}$$

Putting this in Eq. (111) yields

$$D_Xv = \tfrac{1}{2}\bar{h}^c\blacktriangle[D_X(\bar{h}_cv)] - \tfrac{1}{2}[X\ln(\bar{h}_1\blacktriangle\bar{h}_2)]v. \tag{113}$$

We shall use this result to derive an expression for the spinor connection coefficients in terms of the Christoffel symbols. From Eqs. (73) and (113) with $X=\partial_\mu$ and $v=h_a$ we have

$$\begin{aligned}\Lambda^b_{\mu a} &= -(D_\mu h_a)\blacktriangle h^b\\ &= -\tfrac{1}{2}\bar{h}^c\blacktriangle[D_\mu(\bar{h}_ch_a)]\blacktriangle h^b + \tfrac{1}{2}[\partial_\mu\ln(\bar{h}_1\blacktriangle\bar{h}_2)]h_a\blacktriangle h^b\\ &= -\tfrac{1}{2}[D_\mu(\bar{h}_ch_a)]\odot(\bar{h}^ch^b) - \tfrac{1}{2}\partial_\mu\ln(\bar{h}_1\blacktriangle\bar{h}_2)\delta^b_a.\end{aligned} \tag{114}$$

From Eq. (48) we get

$$D_\mu\mathbf{z} = z^\lambda{}_{;\mu}\mathbf{e}_\lambda = (\partial_\mu z^\lambda + \Gamma^\lambda_{\mu\nu}z^\nu)\mathbf{e}_\lambda \tag{115}$$

with $\mathbf{z}(q)$ in place of $\mathbf{v}(q)$. Recall[7] now the definition of some of the hybrid components of the unit tensor field $\mathbf{I}_4(q)$ with values in $\mathcal{M}_q\otimes\mathcal{M}_q$ given as

$$\mathbf{I}_4 = I^\lambda{}_{\dot{c}a}\mathbf{e}_\lambda\bar{h}^ch^a = I_\lambda{}^{\dot{c}b}\mathbf{e}^\lambda\bar{h}_ch_b, \tag{116}$$

from which we find

$$\begin{aligned}I^\lambda{}_{\dot{c}a} &= -\mathbf{e}^\lambda\odot(\bar{h}_ch_a),\\ I_\lambda{}^{\dot{c}b} &= -\mathbf{e}_\lambda\odot(\bar{h}^ch^b).\end{aligned} \tag{117}$$

If we let $\mathbf{z}=\bar{h}_ch_a$, then

$$z^\lambda = \mathbf{e}^\lambda\cdot\mathbf{z} = \mathbf{e}^\lambda\circledast(\bar{h}_ch_a) = -I^\lambda{}_{\dot{c}a}.$$

Putting this into Eq. (115) gives

$$D_\mu(\bar{h}_ch_a) = -(\partial_\mu I^\lambda{}_{\dot{c}a} + \Gamma^\lambda_{\mu\nu}I^\nu{}_{\dot{c}a})\mathbf{e}_\lambda.$$

Substituting this result into Eq. (114) and using Eq. (117) results in[14]

$$\Lambda^b_{\mu a} = -\tfrac{1}{2}(\partial_\mu I^\lambda{}_{\dot{c}a} + \Gamma^\lambda_{\mu\nu}I^\nu{}_{\dot{c}a})I_\lambda{}^{\dot{c}b} - \tfrac{1}{2}\partial_\mu\ln(\bar{h}_1\blacktriangle\bar{h}_2)\delta^b_a. \tag{118}$$

If we choose the spinor basis h_a such that $h_a \blacktriangle h_b = \text{const}$ for all q, then we get

$$\Lambda^b_{\mu a} = -\tfrac{1}{2}(\partial_\mu l^\lambda{}_{\dot c a} + \Gamma^\lambda_{\mu\nu} l^\nu{}_{\dot c a}) l_\lambda{}^{\dot c b}, \tag{119}$$

which agrees with the expression given by Ruse[15] except for sign.

To conclude this section, we derive an expression for $\mathcal{S}_4$-spinor connection coefficients in terms of Dirac gamma matrices in an arbitrary spin representation. For $(\mathcal{S}_4)_q$-valued spinors our results will be shown to be much simpler and more general than those obtained by Fletcher.[16] To this end, note that the procedure used in deriving Eq. (109) can be generalized to give

$$\xi \blacktriangle \eta D_X \zeta = \tfrac{1}{2}\{\xi \blacktriangle [D_X(\eta\zeta)] - \eta \blacktriangle [D_X(\xi\zeta)] - [X(\xi \blacktriangle \eta)]\zeta\}. \tag{120}$$

Moreover, since

$$1 = -\tfrac{1}{8}(\mathbf{I}_2)_s = -\tfrac{1}{2}(\mathbf{I} \blacktriangle \mathbf{I}_2)_s = \tfrac{1}{2}(l^\sigma l_\sigma \blacktriangle \mathbf{I}_2)_s = -\tfrac{1}{2} l_\sigma \blacktriangle \mathbf{I}_2 \blacktriangle l^\sigma$$
$$= -\tfrac{1}{2}(l_\sigma \blacktriangle \mathbf{I}_2) \blacktriangle (l^\sigma \blacktriangle \mathbf{I}_2),$$

making use of Eq. (120) with $\xi = l_\sigma \blacktriangle \mathbf{I}_2$, $\eta = l^\sigma \blacktriangle \mathbf{I}_2$, $\zeta = \bar{\mathbf{I}}_2 \blacktriangle \psi$, and $X = \partial_\mu$, we can write

$$D_\mu(\bar{\mathbf{I}}_2 \blacktriangle \psi) = -\tfrac{1}{2}(l_\sigma \blacktriangle \mathbf{I}_2) \blacktriangle (l^\sigma \blacktriangle \mathbf{I}_2) D_\mu(\bar{\mathbf{I}}_2 \blacktriangle \psi)$$
$$= -\tfrac{1}{4}\{(l_\sigma \blacktriangle \mathbf{I}_2) \blacktriangle [D_\mu(l^\sigma \blacktriangle \mathbf{I}_2 \bar{\mathbf{I}}_2 \blacktriangle \psi)] - (l^\sigma \blacktriangle \mathbf{I}_2) \blacktriangle [D_\mu(l_\sigma \blacktriangle \mathbf{I}_2 \bar{\mathbf{I}}_2 \blacktriangle \psi)]\}, \tag{121}$$

where the contribution from the last term in Eq. (120) drops out because

$$\partial_\mu[(l_\sigma \blacktriangle \mathbf{I}_2) \blacktriangle (l^\sigma \blacktriangle \mathbf{I}_2)] = \partial_\mu(-2) = 0.$$

Also, since $\mathbf{I}_2 \blacktriangle \bar{\mathbf{I}}_2 = 0$, we have

$$0 = -\tfrac{1}{4} l_\sigma \blacktriangle \mathbf{I}_2 \blacktriangle \bar{\mathbf{I}}_2 \blacktriangle [D_\mu(l^\sigma \mathbf{I}_2 \blacktriangle \psi)] = -\tfrac{1}{4}(l_\sigma \blacktriangle \mathbf{I}_2) \blacktriangle [D_\mu(\bar{\mathbf{I}}_2 \blacktriangle l^\sigma \mathbf{I}_2 \blacktriangle \psi)]$$
$$= -\tfrac{1}{4}(l_\sigma \blacktriangle \mathbf{I}_2) \blacktriangle [D_\mu(l^\sigma \blacktriangle \bar{\mathbf{I}}_2 \mathbf{I}_2 \blacktriangle \psi)], \tag{122}$$

and

$$0 = \tfrac{1}{4}(l^\sigma \blacktriangle \mathbf{I}_2) \blacktriangle [D_\mu(l_\sigma \blacktriangle \bar{\mathbf{I}}_2 \mathbf{I}_2 \blacktriangle \psi)]. \tag{123}$$

Adding Eqs. (121), (122), and (123) gives

$$D_\mu(\bar{\mathbf{I}}_2 \blacktriangle \psi) = -\tfrac{1}{4}(l_\sigma \blacktriangle \mathbf{I}_2) \blacktriangle \{D_\mu[l^\sigma \blacktriangle (\mathbf{I}_2\bar{\mathbf{I}}_2 + \bar{\mathbf{I}}_2\mathbf{I}_2) \blacktriangle \psi]\}$$
$$+ \tfrac{1}{4}(l^\sigma \blacktriangle \mathbf{I}_2) \blacktriangle \{D_\mu[l_\sigma \blacktriangle (\mathbf{I}_2\bar{\mathbf{I}}_2 + \bar{\mathbf{I}}_2\mathbf{I}_2) \blacktriangle \psi]\}. \tag{124}$$

Now recall that by virtue of Eq. (74) of Paper I and Eqs. (31), (33) of Sec. II we can write

$$l^\sigma \blacktriangle (\mathbf{I}_2\bar{\mathbf{I}}_2 + \bar{\mathbf{I}}_2\mathbf{I}_2) \blacktriangle \psi = (\mathbf{I}_2\bar{\mathbf{I}}_2 + \bar{\mathbf{I}}_2\mathbf{I}_2) \diamond l^\sigma \psi$$
$$= (\mathbf{I}_2\bar{\mathbf{I}}_2 + \bar{\mathbf{I}}_2\mathbf{I}_2)^\dagger \odot l^\sigma \psi = (\mathbf{I}_4 + \bar{\mathbf{I}}_4) \odot l^\sigma \psi$$

and corresponding equations with l^σ replaced by l_σ. Consequently,

$$D_\mu(\bar{\mathbf{I}}_2 \blacktriangle \psi) = -\tfrac{1}{4} l_\sigma \blacktriangle \mathbf{I}_2 \blacktriangle \{D_\mu[(\mathbf{I}_4 + \bar{\mathbf{I}}_4) \odot (l^\sigma \psi)]\}$$
$$+ \tfrac{1}{4} l^\sigma \blacktriangle \mathbf{I}_2 \blacktriangle \{D_\mu[(\mathbf{I}_4 + \bar{\mathbf{I}}_4) \odot (l_\sigma \psi)]\}. \tag{125}$$

Similarly

$$D_\mu(\mathbf{I}_2 \blacktriangle \psi) = -\tfrac{1}{4} l_\sigma \blacktriangle \bar{\mathbf{I}}_2 \blacktriangle \{D_\mu[(\mathbf{I}_4 + \bar{\mathbf{I}}_4) \odot (l^\sigma \psi)]\}$$
$$+ \tfrac{1}{4} l^\sigma \blacktriangle \bar{\mathbf{I}}_2 \blacktriangle \{D_\mu[(\mathbf{I}_4 + \bar{\mathbf{I}}_4) \odot (l_\sigma \psi)]\}. \tag{126}$$

Adding the last two equations and using $\mathbf{I}_2 + \bar{\mathbf{I}}_2 = \mathbf{I}$ gives

$$D_\mu \psi = -\tfrac{1}{4} l_\sigma \blacktriangle \{D_\mu[(\mathbf{I}_4 + \bar{\mathbf{I}}_4) \odot (l^\sigma \psi)]\}$$
$$+ \tfrac{1}{4} l^\sigma \blacktriangle \{D_\mu[(\mathbf{I}_4 + \bar{\mathbf{I}}_4) \odot (l_\sigma \psi)]\}. \tag{127}$$

It then follows immediately that

$$\Lambda^\alpha_{\mu\beta} = l^\alpha \blacktriangle (D_\mu l_\beta) = \tfrac{1}{4}(l_\sigma l^\alpha) \odot \{D_\mu[(\mathbf{I}_4 + \bar{\mathbf{I}}_4) \odot (l^\sigma l_\beta)]\}$$
$$- \tfrac{1}{4}(l^\sigma l^\alpha) \odot \{D_\mu[(\mathbf{I}_4 + \bar{\mathbf{I}}_4) \odot (l_\sigma l_\beta)]\}. \tag{128}$$

Furthermore, with the help of Eq. (48), we can write for any $\mathbf{A}, \mathbf{B} \in (\mathcal{S}_4)_q \otimes (\mathcal{S}_4)_q$

$$\mathbf{A} \odot [D_\mu(\mathbf{I}_4 \odot \mathbf{B})] = \mathbf{A} \odot [D_\mu(\mathbf{E}_\nu \mathbf{E}^\nu \odot \mathbf{B})]$$
$$= \mathbf{A} \odot \mathbf{E}_\nu(\mathbf{E}^\nu \odot \mathbf{B})_{;\mu},$$

where

$$(\mathbf{E}^\nu \odot \mathbf{B})_{;\mu} = \partial_\mu(\mathbf{E}^\nu \odot \mathbf{B}) + \Gamma^\nu_{\mu\lambda}(\mathbf{E}^\lambda \odot \mathbf{B}),$$

and

$$\Gamma^\nu_{\mu\lambda} = \mathbf{E}^\nu \odot (D_\mu E_\lambda) = \tilde{\mathbf{E}}^\nu \odot (D_\mu \tilde{\mathbf{E}}_\lambda).$$

In the same fashion

$$\mathbf{A} \odot [D_\mu(\bar{\mathbf{I}}_4 \odot \mathbf{B})] = \mathbf{A} \odot [D_\mu(\tilde{\mathbf{E}}_\nu \tilde{\mathbf{E}}^\nu \odot \mathbf{B})] = \mathbf{A} \odot \tilde{\mathbf{E}}_\nu(\tilde{\mathbf{E}}^\nu \odot \mathbf{B})_{;\mu}.$$

Hence, Eq. (128) can be expressed in the form

$$\Lambda^\alpha_{\mu\beta} = \tfrac{1}{4}[l_\sigma \blacktriangle \mathbf{E}_\nu \blacktriangle l^\alpha (l^\sigma \blacktriangle \mathbf{E}^\nu \blacktriangle l_\beta)_{;\mu} + l_\sigma \blacktriangle \tilde{\mathbf{E}}_\nu \blacktriangle l^\alpha (l^\sigma \blacktriangle \tilde{\mathbf{E}}^\nu \blacktriangle l_\beta)_{;\mu}$$
$$- l^\sigma \blacktriangle \mathbf{E}_\nu \blacktriangle l^\alpha (l_\sigma \blacktriangle \mathbf{E}^\nu \blacktriangle l_\beta)_{;\mu} - l^\sigma \blacktriangle \tilde{\mathbf{E}}_\nu \blacktriangle l^\alpha (l_\sigma \blacktriangle \tilde{\mathbf{E}}^\nu \blacktriangle l_\beta)_{;\mu}]. \tag{129}$$

In order to put the above result in terms of Dirac gamma matrices, observe that from [cf. Eqs. (100) and (132) of I]

$$\mathbf{\Gamma}^\mu = (2)^{1/2} i(\mathbf{E}^\mu + \tilde{\mathbf{E}}^\mu), \quad \mathbf{\Gamma}^5 = i(\mathbf{I}_2 - \bar{\mathbf{I}}_2)$$

we get

$$\mathbf{E}^\mu = -(2)^{-3/2} i(\mathbf{I} + i\mathbf{\Gamma}^5) \blacktriangle \mathbf{\Gamma}^\mu,$$
$$\tilde{\mathbf{E}}^\mu = -(2)^{-3/2} i(\mathbf{I} - i\mathbf{\Gamma}^5) \blacktriangle \mathbf{\Gamma}^\mu. \tag{130}$$

Substituting these expressions into Eq. (129) yields, after some straightforward operations and rearrangement of terms,

$$\Lambda^\alpha_{\mu\beta} = -\tfrac{1}{16}[(l_\sigma \blacktriangle \mathbf{\Gamma}_\nu \blacktriangle l^\alpha)(l^\sigma \blacktriangle \mathbf{\Gamma}^\nu \blacktriangle l_\beta)_{;\mu} - (l^\sigma \blacktriangle \mathbf{\Gamma}_\nu \blacktriangle l^\alpha)(l_\sigma \blacktriangle \mathbf{\Gamma}^\nu \blacktriangle l_\beta)_{;\mu}]$$
$$+ \tfrac{1}{16}[(l_\sigma \blacktriangle \mathbf{\Gamma}^5 \blacktriangle \mathbf{\Gamma}_\nu \blacktriangle l^\alpha)(l^\sigma \blacktriangle \mathbf{\Gamma}^5 \blacktriangle \mathbf{\Gamma}^\nu \blacktriangle l_\beta)_{;\mu}$$
$$- (l^\sigma \blacktriangle \mathbf{\Gamma}^5 \blacktriangle \mathbf{\Gamma}_\nu \blacktriangle l^\alpha)(l_\sigma \blacktriangle \mathbf{\Gamma}^5 \blacktriangle \mathbf{\Gamma}^\nu \blacktriangle l_\beta)_{;\mu}]. \tag{131}$$

Although this result is only valid for the standard spinor connection, its generalization to arbitrary spinor connections follows most simply from Eq. (78). We thus have

$$D'_\mu l_\beta = D_\mu l_\beta + \mathbf{M}_\mu \blacktriangle l_\beta,$$

or

$$\Lambda'^\alpha_{\mu\beta} \equiv l^\alpha \blacktriangle D'_\mu l_\beta = \Lambda^\alpha_{\mu\beta} + l^\alpha \blacktriangle \mathbf{M}_\mu \blacktriangle l_\beta, \tag{132}$$

where $\Lambda^\alpha_{\mu\beta}$ is given by Eq. (131) above, and $\mathbf{M}_\mu(q) = \mathbf{K}_\mu(q) + \bar{\mathbf{K}}_\mu(q)$. In particular, for the specific cases $\mathbf{K}_\mu = i\varphi_\mu \mathbf{I}_2$ and $\mathbf{K}_\mu = \varphi_\mu \mathbf{I}_2$, Eq. (132) becomes

$$\Lambda'^\alpha_{\mu\beta} = \Lambda^\alpha_{\mu\beta} + \varphi_\mu l^\alpha \blacktriangle \mathbf{\Gamma}^5 \blacktriangle l_\beta \tag{133}$$

and

$$\Lambda'^\alpha_{\mu\beta} = \Lambda^\alpha_{\mu\beta} + \varphi_\mu \delta^\alpha_\beta, \tag{134}$$

respectively.

In order to relate these results to those obtained by

Fletcher, note first that the derivation of his Eqs. (30) and (37) is based on his Eq. (7). This equation, however, is not true for an arbitrary spinor connection. In Appendix B we derive an alternate expression which is always valid, and show that for the Cases 3 and 4 discussed previously in this section it reduces to Fletcher's Eq. (7). Thus, his Eqs. (30) and (37), for $n=4$, correspond to our Eqs. (134) and (131). Equation (133) is not contained in Fletcher's results because for $\mathbf{K}_\mu = i\varphi_\mu \mathbf{I}_2$ his Eq. (7) is not valid. A comparison also shows that our results are considerably simpler without any loss in the arbitrariness of the four-dimensional representation of the Dirac gamma matrices. Note, moreover, that our Eq. (132) has been derived without making any restrictive assumption. It is, therefore, completely general.

Equation (131) can be expressed in a somewhat different form by noting that

$$\begin{aligned}(l_\sigma\blacktriangle\Gamma^5\blacktriangle\Gamma^\nu\blacktriangle l_\beta)_{;\mu} &= (l_\sigma\blacktriangle\Gamma^5\blacktriangle l_\tau l^\tau\blacktriangle\Gamma^\nu\blacktriangle l_\beta)_{;\mu}\\ &= [\partial_\mu(l_\sigma\blacktriangle\Gamma^5\blacktriangle l_\tau)](l^\tau\blacktriangle\Gamma^\nu\blacktriangle l_\beta)\\ &\quad + (l_\sigma\blacktriangle\Gamma^5\blacktriangle l_\tau)(l^\tau\blacktriangle\Gamma^\nu\blacktriangle l_\beta)_{;\mu}.\end{aligned}$$

Hence, recalling that $\Gamma^5\blacktriangle\Gamma^5 = -\mathrm{I}$, and also that $\xi\blacktriangle\Gamma^5\blacktriangle\eta = -\eta\blacktriangle\Gamma^5\blacktriangle\xi$ and $\xi\blacktriangle\Gamma_\mu\blacktriangle\eta = \eta\blacktriangle\Gamma_\mu\blacktriangle\xi$ for arbitrary ξ,η, we get

$$\begin{aligned}&(l^\sigma\blacktriangle\Gamma^5\blacktriangle\Gamma_\nu\blacktriangle l^\alpha)(l_\sigma\blacktriangle\Gamma^5\blacktriangle\Gamma^\nu\blacktriangle l_\beta)_{;\mu}\\ &\quad = [\partial_\mu(l_\sigma\blacktriangle\Gamma^5\blacktriangle l_\tau)](l^\tau\blacktriangle\Gamma^\nu\blacktriangle l_\beta)(l^\sigma\blacktriangle\Gamma^5\blacktriangle\Gamma_\nu\blacktriangle l^\alpha)\\ &\qquad + (l^\alpha\blacktriangle\Gamma_\nu\blacktriangle l_\tau)(l^\tau\blacktriangle\Gamma^\nu\blacktriangle l_\beta)_{;\mu}.\end{aligned} \tag{135}$$

Similarly

$$\begin{aligned}&(l_\sigma\blacktriangle\Gamma^5\blacktriangle\Gamma_\nu\blacktriangle l^\alpha)(l^\sigma\blacktriangle\Gamma^5\blacktriangle\Gamma^\nu\blacktriangle l_\beta)_{;\mu}\\ &\quad = \partial_\mu(l^\sigma\blacktriangle\Gamma^5\blacktriangle l_\tau)(l^\tau\blacktriangle\Gamma^\nu\blacktriangle l_\beta)(l_\sigma\blacktriangle\Gamma^5\blacktriangle\Gamma_\nu\blacktriangle l^\alpha)\\ &\qquad - (l^\alpha\blacktriangle\Gamma_\nu\blacktriangle l_\tau)(l^\tau\blacktriangle\Gamma^\nu\blacktriangle l_\beta)_{;\mu}.\end{aligned} \tag{136}$$

Also note that

$$\begin{aligned}(l^\sigma\blacktriangle\Gamma_\nu\blacktriangle l^\alpha)(l_\sigma\blacktriangle\Gamma^\nu\blacktriangle l_\beta)_{;\mu} &= (l^\sigma\blacktriangle\Gamma_\nu\blacktriangle l^\alpha)(l_\sigma\blacktriangle l_\tau l^\tau\blacktriangle\Gamma^\nu\blacktriangle l_\beta)_{;\mu}\\ &= (l^\sigma\blacktriangle\Gamma_\nu\blacktriangle l^\alpha)(l^\tau\blacktriangle\Gamma^\nu\blacktriangle l_\beta)\partial_\mu(l_\sigma\blacktriangle l_\tau)\\ &\quad - (l^\alpha\blacktriangle\Gamma_\nu\blacktriangle l_\tau)(l^\tau\blacktriangle\Gamma^\nu\blacktriangle l_\beta)_{;\mu}.\end{aligned} \tag{137}$$

Substituting Eqs. (135), (136), (137) into Eq. (131) yields

$$\begin{aligned}\Lambda^\alpha_{\mu\beta} &= -\tfrac{1}{4}(l^\alpha\blacktriangle\Gamma_\nu\blacktriangle l_\sigma)(l^\alpha\blacktriangle\Gamma^\nu\blacktriangle l_\beta)_{;\mu}\\ &\quad + \tfrac{1}{16}(l^\sigma\blacktriangle\Gamma_\nu\blacktriangle l^\alpha)(l^\tau\blacktriangle\Gamma^\nu\blacktriangle l_\beta)\partial_\mu(l_\sigma\blacktriangle l_\tau)\\ &\quad + \tfrac{1}{16}(l^\tau\blacktriangle\Gamma^\nu\blacktriangle l_\beta)(l_\sigma\blacktriangle\Gamma^5\blacktriangle\Gamma_\nu\blacktriangle l^\alpha)\partial_\mu(l^\sigma\blacktriangle\Gamma^5\blacktriangle l_\tau)\\ &\quad - \tfrac{1}{16}(l^\tau\blacktriangle\Gamma^\nu\blacktriangle l_\beta)(l^\sigma\blacktriangle\Gamma^5\blacktriangle\Gamma_\nu\blacktriangle l^\alpha)\partial_\mu(l_\sigma\blacktriangle\Gamma^5\blacktriangle l_\tau).\end{aligned} \tag{138}$$

This expression can be simplified further if the spinor basis is chosen so that

$$\begin{aligned}&l^\sigma\blacktriangle\mathbf{I}_2\blacktriangle l^\tau = \text{const},\\ &l^\sigma\blacktriangle\bar{\mathbf{I}}_2\blacktriangle l^\tau = \text{const}\end{aligned} \tag{139}$$

for all σ,τ. It readily follows from these assumptions that

$$l^\sigma\blacktriangle l^\tau = \text{const},\quad l_\sigma\blacktriangle l_\tau = \text{const},$$

$$l^\sigma\blacktriangle\mathbf{I}_2\blacktriangle l_\tau = \text{const},\quad l^\sigma\blacktriangle\bar{\mathbf{I}}_2\blacktriangle l_\tau = \text{const},$$

$$l^\sigma\blacktriangle\Gamma^5\blacktriangle l_\tau = il^\sigma\blacktriangle(\mathbf{I}_2-\bar{\mathbf{I}}_2)\blacktriangle l_\tau = \text{const}.$$

Consequently, Eq. (138) reduces in this case to

$$\begin{aligned}\Lambda^\alpha_{\mu\beta} &= -\tfrac{1}{4}(l^\alpha\blacktriangle\Gamma_\nu\blacktriangle l_\sigma)(l^\sigma\blacktriangle\Gamma^\nu\blacktriangle l_\beta)_{;\mu}\\ &= -\tfrac{1}{4}(\gamma_\nu\gamma^\nu{}_{;\mu})^\alpha{}_\beta\end{aligned} \tag{140}$$

(using the matrix notation with suppressed spinor indices introduced in Appendix B). The quantities in this last result are the so called Fock–Ivanenko coefficients,[5] and are quoted in the literature in various forms.[17]

V. SPINOR CURVATURE TENSORS

By analogy with the Riemann tensor defined in Sec. III, we introduce spinor curvatures $\mathfrak{P}'$, and $\mathfrak{Q}'$ for a general spinor connection, D'_X, by the following equations:

$$\mathbf{xy}:\mathfrak{P}'\blacktriangle u = (D'_X D'_Y - D'_Y D'_X - D'_{[X,Y]})u \tag{141}$$

and

$$\mathbf{xy}:\mathfrak{Q}'\blacktriangle\bar{v} = (D'_X D'_Y - D'_Y D'_X - D'_{[X,Y]})\bar{v} \tag{142}$$

where $\mathfrak{P}'(q)$ takes values in $\mathcal{M}_q\otimes\mathcal{M}_q\otimes(\mathcal{S}_2)_q\otimes(\mathcal{S}_2)_q$, and $\mathfrak{Q}'(q)$ takes values[18] in $\mathcal{M}_q\otimes\mathcal{M}_q\otimes(\bar{\mathcal{S}}_2)_q\otimes(\bar{\mathcal{S}}_2)_q$. Adding the above defining equations yields

$$\mathbf{xy}:\mathfrak{S}'\blacktriangle\psi = (D'_X D'_Y - D'_Y D'_X - D'_{[X,Y]})\psi, \tag{143}$$

where $\psi = u+\bar{v}$, and $\mathfrak{S}' = \mathfrak{P}'+\mathfrak{Q}'$ has values in $\mathcal{M}_q\otimes\mathcal{M}_q\otimes(\mathcal{S}_4)_q\otimes(\mathcal{S}_4)_q$. The curvature tensors $\mathfrak{P},\mathfrak{Q},\mathfrak{S}$ for the standard spinor connection D_X are defined in exactly the same manner as above but without the primes. We shall now establish the relationship between these standard curvature tensors and the Riemann tensor and later extend these relations to general spinor curvature tensors.

Putting $\mathbf{z}=\bar{u}v$ in Eq. (56) gives after some simple operations

$$\begin{aligned}\mathbf{xy}:\mathfrak{R}\circledcirc(\bar{u}v) &= (D_XD_Y - D_YD_X - D_{[X,Y]})\bar{u}v\\ &= [(D_XD_Y - D_YD_X - D_{[X,Y]})\bar{u}]v\\ &\quad + \bar{u}[(D_XD_Y - D_YD_X - D_{[X,Y]})v]\\ &= (\mathbf{xy}:\mathfrak{Q}\blacktriangle\bar{u})v + \bar{u}(\mathbf{xy}:\mathfrak{P}\blacktriangle v)\\ &= [(\mathbf{xy}:\mathfrak{Q})\mathbf{I}_2 + \bar{\mathbf{I}}_2(\mathbf{xy}:\mathfrak{P})]\diamond(\bar{u}v)\\ &= [(\mathbf{xy}:\mathfrak{Q})\mathbf{I}_2 + \bar{\mathbf{I}}_2(\mathbf{xy}:\mathfrak{P})]^\dagger\circledcirc(\bar{u}v),\end{aligned}$$

where we have used the special operations previously defined by Eqs. (73) and (74) of I. Since $\bar{u}$ and v are arbitrary we can therefore write

$$\mathbf{xy}:\mathfrak{R} = [(\mathbf{xy}:\mathfrak{Q})\mathbf{I}_2 + \bar{\mathbf{I}}_2(\mathbf{xy}:\mathfrak{P})]^\dagger \tag{144}$$

By an inner multiplication with $\mathbf{I}_2$ we get

$$\begin{aligned}\mathbf{xy}:\mathfrak{R}\diamond\mathbf{I}_2 &= [(\mathbf{xy}:\mathfrak{Q})\mathbf{I}_2 + \bar{\mathbf{I}}_2(\mathbf{xy}:\mathfrak{P})]^\dagger\diamond\mathbf{I}_2\\ &= [(\mathbf{xy}:\mathfrak{Q})\mathbf{I}_2 + \bar{\mathbf{I}}_2(\mathbf{xy}:\mathfrak{P})]\circledcirc\mathbf{I}_2\\ &= -2\mathbf{xy}:\mathfrak{Q} + (\mathbf{xy}:\mathfrak{P}\circledcirc\mathbf{I}_2)\bar{\mathbf{I}}_2.\end{aligned}$$

Consequently,

$$\mathfrak{R}\diamond\mathbf{I}_2 = -2\mathfrak{Q} + (\mathfrak{P}\circledcirc\mathbf{I}_2)\bar{\mathbf{I}}_2 = -2\mathfrak{Q} + \mathbf{P}\bar{\mathbf{I}}_2, \tag{145}$$

where $\mathbf{P}\equiv\mathfrak{P}\odot\mathbf{I}_2$. Alternatively, we can also write

$$-C(6,8)\mathfrak{R} = -2\mathfrak{Q} + \mathbf{P}\bar{\mathbf{I}}_2. \tag{146}$$

Here the symbol $C(i,j)$ denotes contraction of the ith and jth files of the tensor $\mathfrak{R}$, regarded as a spinor tensor in the space $(\bar{S}_2\otimes_H S_2)\otimes(\bar{S}_2\otimes_H S_2)\otimes(\bar{S}_2\otimes_H S_2)\otimes(\bar{S}_2\otimes_H S_2)$, where the files are numbered from left to right.[19] Similarly by an inner multiplication with $\bar{\mathbf{I}}_2$ from the left, Eq. (144) yields

$$\begin{aligned}\bar{\mathbf{I}}_2 \diamond (\mathbf{xy}:\mathfrak{R}) &= \bar{\mathbf{I}}_2 \diamond [(\mathbf{xy}:\mathfrak{Q})\mathbf{I}_2 + \bar{\mathbf{I}}_2(\mathbf{xy}:\mathfrak{P})]^{\ddagger}\\ &= \bar{\mathbf{I}}_2 \odot [(\mathbf{xy}:\mathfrak{Q})\mathbf{I}_2 + \bar{\mathbf{I}}_2(\mathbf{xy}:\mathfrak{P})]\\ &= \mathbf{xy}:\mathfrak{Q}\odot\bar{\mathbf{I}}_2\mathbf{I}_2 - 2\mathbf{xy}:\mathfrak{P}\\ &= \mathbf{xy}:\mathbf{Q}\mathbf{I}_2 - 2\mathbf{xy}:\mathfrak{P}\end{aligned}$$

or

$$-C(5,7)\mathfrak{R} = \mathbf{Q}\mathbf{I}_2 - 2\mathfrak{P}, \tag{147}$$

where

$$\mathbf{Q} \equiv \mathfrak{Q}\odot\bar{\mathbf{I}}_2.$$

Next we show that both $\mathbf{P}$ and $\mathbf{Q}$ are zero. To this end, note that:

$$\mathbf{xy}:\mathfrak{P}\overset{\blacktriangle}{\blacktriangle}(uv) = (\mathbf{xy}:\mathfrak{P}\blacktriangle v)\blacktriangle u = [(D_XD_Y - D_YD_X - D_{[X,Y]})v]\blacktriangle u. \tag{148}$$

The first term on the right can be put in the form

$$\begin{aligned}(D_XD_Yv)\blacktriangle u &= X[(D_Yv)\blacktriangle u] - (D_Yv)\blacktriangle(D_Xu)\\ &= X[(D_Yv)\blacktriangle u] - Y[v\blacktriangle(D_Xu)] + v\blacktriangle(D_YD_Xu).\end{aligned}$$

By exchanging X and Y we also have

$$(D_YD_Xv)\blacktriangle u = Y[(D_Xv)\blacktriangle u] - X[v\blacktriangle(D_Yu)] + v\blacktriangle(D_XD_Yu).$$

Finally, rewriting the last term in the right of Eq. (148) as

$$(D_{[X,Y]}v)\blacktriangle u = [X,Y](v\blacktriangle u) - v\blacktriangle(D_{[X,Y]}u),$$

and substituting these results back into Eq. (148) yields

$$\begin{aligned}\mathbf{xy}:\mathfrak{P}:(uv) &= -v\blacktriangle[(D_XD_Y - D_YD_X - D_{[X,Y]})u]\\ &\quad + XY(v\blacktriangle u) - YX(v\blacktriangle u) - [X,Y](v\blacktriangle u)\\ &= [(D_XD_Y - D_YD_X - D_{[X,Y]})u]\blacktriangle v\\ &= \mathbf{xy}:\mathfrak{P}\overset{\blacktriangle}{\blacktriangle}(vu).\end{aligned}$$

Therefore

$$(\mathbf{xy}:\mathfrak{P})_T = (\mathbf{xy}:\mathfrak{P}), \tag{149}$$

i.e., $\mathfrak{P}$ is symmetric in the last two files. It follows immediately from this result that

$$\mathbf{P} = \mathfrak{P}\odot\mathbf{I}_2 = 0 \tag{150}$$

and similarly that

$$\mathbf{Q} = \mathfrak{Q}\odot\bar{\mathbf{I}}_2 = 0. \tag{151}$$

Consequently, Eqs. (146) and (147) become

$$C(6,8)\mathfrak{R} = 2\mathfrak{Q} \tag{152}$$

and

$$C(5,7)\mathfrak{R} = 2\mathfrak{P}. \tag{153}$$

Combining these two results and using $\mathfrak{S} = \mathfrak{P} + \mathfrak{Q}$, we also have

$$[C(6,8) + C(5,7)]\mathfrak{R} = 2\mathfrak{S} \tag{154}$$

In order to modify these results so that they apply to the general spinor connections, we need to find the relationship between $\mathfrak{P}$ and $\mathfrak{P}'$. This follows straightforwardly by substituting Eq. (76) into Eq. (141); we thus get

$$\begin{aligned}\mathbf{xy}:\mathfrak{P}'\blacktriangle v &= D_X(D_Yv + \mathrm{K}_Y\blacktriangle v) + \mathrm{K}_X\blacktriangle(D_Yv + \mathrm{K}_Y\blacktriangle v)\\ &\quad - D_Y(D_Xv + \mathrm{K}_X\blacktriangle v) - \mathrm{K}_Y\blacktriangle(D_Xv + \mathrm{K}_X\blacktriangle v)\\ &\quad - D_{[X,Y]}v - \mathrm{K}_{[X,Y]}\blacktriangle v\\ &= \mathbf{xy}:\mathfrak{P}\blacktriangle v + (D_X\mathrm{K}_Y - D_Y\mathrm{K}_X + \mathrm{K}_X\blacktriangle\mathrm{K}_Y\\ &\quad - \mathrm{K}_Y\blacktriangle\mathrm{K}_X - \mathrm{K}_{[X,Y]})\blacktriangle v,\end{aligned}$$

or

$$\begin{aligned}\mathbf{xy}:\mathfrak{P}' &= \mathbf{xy}:\mathfrak{P} + (D_X\mathrm{K}_Y - D_Y\mathrm{K}_X + \mathrm{K}_X\blacktriangle\mathrm{K}_Y\\ &\quad - \mathrm{K}_Y\blacktriangle\mathrm{K}_X - \mathrm{K}_{[X,Y]}).\end{aligned} \tag{155}$$

Similarly, for $\mathfrak{Q}'$ we obtain

$$\begin{aligned}\mathbf{xy}:\mathfrak{Q}' &= \mathbf{xy}:\mathfrak{Q} + (D_X\bar{\mathrm{K}}_Y - D_Y\bar{\mathrm{K}}_X + \bar{\mathrm{K}}_X\blacktriangle\bar{\mathrm{K}}_Y\\ &\quad - \bar{\mathrm{K}}_Y\blacktriangle\bar{\mathrm{K}}_X - \bar{\mathrm{K}}_{[X,Y]}),\end{aligned} \tag{156}$$

and adding the two results gives

$$\begin{aligned}\mathbf{xy}:\mathfrak{S}' &= \mathbf{xy}:\mathfrak{S} + (D_X\mathrm{M}_Y - D_Y\mathrm{M}_X + \mathrm{M}_X\blacktriangle\mathrm{M}_Y\\ &\quad - \mathrm{M}_Y\blacktriangle\mathrm{M}_X - \mathrm{M}_{[X,Y]}).\end{aligned} \tag{157}$$

Clearly then, the generalization of Eqs. (152), (153), (154) follows merely by substituting into them the expressions for $\mathfrak{Q}$, $\mathfrak{P}$, and $\mathfrak{S}$ in terms of $\mathfrak{Q}'$, $\mathfrak{P}'$ and $\mathfrak{S}$ respectively, as given by the above results.

For the special case $\mathrm{K}_X = i\mathbf{x}\cdot\boldsymbol{\varphi}\mathbf{I}_2$ (Case 2 in the previous section) we find[20]

$$\begin{aligned}\mathbf{xy}:\mathfrak{P}' &= \mathbf{xy}:\mathfrak{P} + [iX(\mathbf{y}\cdot\boldsymbol{\varphi}) - iY(\mathbf{x}\cdot\boldsymbol{\varphi}) - i([X,Y]\rho)\cdot\boldsymbol{\varphi}]\mathbf{I}_2\\ &= \mathbf{xy}:\mathfrak{P} + i[\mathbf{y}\cdot(D_X\boldsymbol{\varphi}) - \mathbf{x}\cdot(D_Y\boldsymbol{\varphi})]\mathbf{I}_2\\ &= \mathbf{xy}:[\mathfrak{P} + i(\mathbf{D}\wedge\boldsymbol{\varphi})\mathbf{I}_2]\end{aligned}$$

or

$$\mathfrak{P}' = \mathfrak{P} + i(\mathbf{D}\wedge\boldsymbol{\varphi})\mathbf{I}_2. \tag{158}$$

Similarly

$$\mathfrak{Q}' = \mathfrak{Q} - i(\mathbf{D}\wedge\boldsymbol{\varphi})\bar{\mathbf{I}}_2, \tag{159}$$

and furthermore

$$\begin{aligned}\mathfrak{S}' &= \mathfrak{S} + i(\mathbf{D}\wedge\boldsymbol{\varphi})(\mathbf{I}_2 - \bar{\mathbf{I}}_2)\\ &= \mathfrak{S} + (\mathbf{D}\wedge\boldsymbol{\varphi})\boldsymbol{\Gamma}^5,\end{aligned} \tag{160}$$

where $\boldsymbol{\Gamma}^5$ is the Dirac operator defined by Eq. (130) in I. For this particular case Eqs. (153), (152), and (154) generalize to

$$C(5,7)\mathfrak{R} = 2\mathfrak{P}' - 2i(\mathbf{D}\wedge\boldsymbol{\varphi})\mathbf{I}_2, \tag{161}$$

$$C(6,8)\mathfrak{R} = 2\mathfrak{Q}' + 2i(\mathbf{D}\wedge\boldsymbol{\varphi})\bar{\mathbf{I}}_2, \tag{162}$$

$$[C(6,8) + C(5,7)]\mathfrak{R} = 2\mathfrak{S}' - 2(\mathbf{D}\wedge\boldsymbol{\varphi})\boldsymbol{\Gamma}^5. \tag{163}$$

It is readily seen that for the case $\mathrm{K}_X = \mathbf{x}\cdot\boldsymbol{\varphi}\mathbf{I}_2$ (Case 3 in the previous section) one obtains

$$C(5,7)\mathfrak{R} = 2\mathfrak{P}' - 2(\mathbf{D}\wedge\boldsymbol{\varphi})\mathbf{I}_2, \tag{164}$$

$$C(6,8)\mathfrak{R} = 2\mathfrak{Q}' - 2(\mathbf{D}\wedge\boldsymbol{\varphi})\bar{\mathbf{I}}_2, \tag{165}$$

$$[C(6,8) + C(5,7)]\mathfrak{R} = 2\mathfrak{S}' - 2(\mathbf{D}\wedge\boldsymbol{\varphi})\mathbf{I}_2, \tag{166}$$

where $\mathbf{I}=\mathbf{I}_2+\bar{\mathbf{I}}_2$. We emphasize at this point the fact that by virtue of Theorem 2 of the previous section, Eqs. (161), (162), and (163) are the ones that apply for a spinor connection D'_X which generates the standard four-vector connection, while by virtue of Theorem 4, Eqs. (164), (165), and (166) are the ones that correspond to a spinor connection which "alternately generates" the standard four-vector connection.

In order to compare our results with others appearing in the literature, note that

$$C(6,8)\mathfrak{R}=C(6,8)(\mathbf{D}\wedge\mathbf{DE}_\mu)\mathbf{E}^\mu=(\mathbf{D}\wedge\mathbf{DE}_\mu)\blacktriangle\tilde{\mathbf{E}}^\mu$$

and

$$C(5,7)\mathfrak{R}=(\mathbf{D}\wedge\mathbf{D}\tilde{\mathbf{E}}_\mu)\blacktriangle\mathbf{E}^\mu$$

Moreover, recalling the definition of the Dirac operators in Eq. (23), we find

$$\begin{aligned}[C(6,8)+C(5,7)]\mathfrak{R}&=(\mathbf{D}\wedge\mathbf{DE}_\mu)\blacktriangle\tilde{\mathbf{E}}^\mu+(\mathbf{D}\wedge\mathbf{D}\tilde{\mathbf{E}}^\mu)\blacktriangle\mathbf{E}^\mu\\&=[\mathbf{D}\wedge\mathbf{D}(\mathbf{E}_\mu+\tilde{\mathbf{E}}_\mu)]\blacktriangle(\mathbf{E}^\mu+\tilde{\mathbf{E}}^\mu)\\&=-\tfrac{1}{2}(\mathbf{D}\wedge\mathbf{D}\boldsymbol{\Gamma}_\mu)\blacktriangle\boldsymbol{\Gamma}^\mu.\end{aligned}$$

Consequently, Eqs. (163) and (166) can be written as

$$-\tfrac{1}{2}(\mathbf{D}\wedge\mathbf{D}\boldsymbol{\Gamma}_\mu)\blacktriangle\boldsymbol{\Gamma}^\mu=2\mathfrak{S}'-2(\mathbf{D}\wedge\varphi)\boldsymbol{\Gamma}^5$$

or

$$\mathfrak{S}'=-\tfrac{1}{4}(\mathbf{D}\wedge\mathbf{D}\boldsymbol{\Gamma}_\mu)\blacktriangle\boldsymbol{\Gamma}^\mu+(\mathbf{D}\wedge\varphi)\boldsymbol{\Gamma}^5 \tag{167}$$

and

$$\mathfrak{S}'=-\tfrac{1}{4}(\mathbf{D}\wedge\mathbf{D}\boldsymbol{\Gamma}_\mu)\blacktriangle\boldsymbol{\Gamma}^\mu+(\mathbf{D}\wedge\varphi)\mathbf{I} \tag{168}$$

respectively. In terms of the usual component notation these last two results become

$$\mathbf{S}'_{\kappa\lambda}=\mathbf{E}_\kappa\mathbf{E}_\lambda:\mathfrak{S}'=-\tfrac{1}{4}R_{\kappa\lambda\sigma\mu}\boldsymbol{\Gamma}^\sigma\blacktriangle\boldsymbol{\Gamma}^\mu+(D_\kappa\varphi_\lambda-D_\lambda\varphi_\kappa)\boldsymbol{\Gamma}^5, \tag{169}$$

$$\mathbf{S}'_{\kappa\lambda}=\mathbf{E}_\kappa\mathbf{E}_\lambda:\mathfrak{S}'=-\tfrac{1}{4}R_{\kappa\lambda\sigma\mu}\boldsymbol{\Gamma}^\sigma\blacktriangle\boldsymbol{\Gamma}^\mu+(D_\kappa\varphi_\lambda-D_\lambda\varphi_\kappa)\mathbf{I}. \tag{170}$$

Note that Eq. (170) differs from the spinor curvature tensor usually given in the literature[6,13] by a factor of i in the last term. Referring to Schmutzer's derivation of the spinor curvature tensor, which agrees with Schrodinger's result, shows that the factor i can be explained by the fact that the bispinor connection he uses is not the same as the one we use in arriving at Eq. (170). Specifically, this factor i is direct consequence of the fact that the $\mathbf{N}_X$ in Eq. (108) differs from $\mathbf{M}_X$ in Eq. (106) by the same factor. It is also interesting to note that although in the derivation of Eq. (169) we start with the same spinor connection for $(\mathcal{S}_2)_q$ valued spinors as Schmutzer, the different prescription we used for obtaining the bispinor connection leads to the factor $\boldsymbol{\Gamma}^5$ in the last term in place of $i\mathbf{I}$.

APPENDIX A

In the intrinsic spinor formalism introduced in I and further elaborated in Sec. II of this paper, we have omitted consideration of the dual space $\mathcal{S}_2^\star$ of $\mathcal{S}_2$; i.e., the space of linear functionals on $\mathcal{S}_2$. As will be seen from the following discussion, this omission is quite intentional and leads to a simpler abstract formalism without the need for unnecessary complications.

In order to relate our point of view with the approach followed in the literature reviewed in this paper, we first remark that, in this literature, elements of $\mathcal{S}_2$ are called contravariant spinors and elements of $\mathcal{S}_2^\star$ are called covariant spinors. For each element $\boldsymbol{w}^\star$ in $\mathcal{S}_2^\star$, its operation as a linear functional on $\mathcal{S}_2$ associates a complex number $\boldsymbol{w}^\star\circ\boldsymbol{u}$ with each $\boldsymbol{u}$ in $\mathcal{S}_2$. The quantity $\boldsymbol{w}^\star\circ\boldsymbol{u}$ is linear in $\boldsymbol{u}$; furthermore, the operations of multiplication by scalars in $\mathcal{S}_2^\star$ and addition in $\mathcal{S}_2^\star$ are defined such that $\boldsymbol{w}^\star\circ\boldsymbol{u}$ is linear also in $\boldsymbol{w}^\star$. The inner product in $\mathcal{S}_2$ gives rise to a natural one-to-one map B from $\mathcal{S}_2$ onto $\mathcal{S}_2^\star$ which takes each spinor $\boldsymbol{v}$ in $\mathcal{S}_2$ into a unique spinor $\boldsymbol{v}^\star=\mathrm{B}\boldsymbol{v}$ in $\mathcal{S}_2^\star$ according to the equation

$$(\mathrm{B}\boldsymbol{v})\circ\boldsymbol{u}=\boldsymbol{v}\blacktriangle\boldsymbol{u} \tag{A1}$$

for all $\boldsymbol{u}$ in $\mathcal{S}_2$. Thus, wherever other writers have an expression involving spinors $\boldsymbol{v}^\star$ in $\mathcal{S}_2^\star$, we reinterpret the expression as an equivalent one in which each $\boldsymbol{v}^\star$ is replaced by $\boldsymbol{v}=\mathrm{B}^{-1}\boldsymbol{v}^\star$ in $\mathcal{S}_2$; for example, the expression $\boldsymbol{v}^\star\circ\boldsymbol{u}$ is reinterpreted as $\boldsymbol{v}\blacktriangle\boldsymbol{u}$. To explain this in terms of components, let us take a basis $\boldsymbol{h}_1,\boldsymbol{h}_2$ for $\mathcal{S}_2$, its corresponding reciprocal basis $\boldsymbol{h}^1,\boldsymbol{h}^2$ for $\mathcal{S}_2$ and its dual basis $\boldsymbol{h}^{\star1}$, $\boldsymbol{h}^{\star2}$ for $\mathcal{S}_2^\star$. By definition, we have

$$\boldsymbol{h}^a\blacktriangle\boldsymbol{h}_b=\delta^a_b,\quad \boldsymbol{h}^{\star a}\circ\boldsymbol{h}_b=\delta^a_b, \tag{A2}$$

and it follows that

$$\boldsymbol{h}^{\star a}=B\boldsymbol{h}^a. \tag{A3}$$

A spinor $\boldsymbol{v}$ in $\mathcal{S}_2$ can be expressed in terms of components as

$$\boldsymbol{v}=v^a\boldsymbol{h}_a=v_a\boldsymbol{h}^a$$

and its image $\boldsymbol{v}^\star$ in $\mathcal{S}_2^\star$ as

$$\boldsymbol{v}^\star=v^\star_a\boldsymbol{h}^{\star a}. \tag{A4}$$

In the literature reviewed here only the components v^a of $\boldsymbol{v}$, and the components $v^\star_a$ of $\boldsymbol{v}^\star$ are used, but without the star on $v^\star_a$. Thus there the symbol v^a denotes the components of a contravariant spinor (in $\mathcal{S}_2$) with respect to the basis $\boldsymbol{h}_a$, and $v^\star_a$ (with the star omitted) denotes the components of a covariant spinor (in $\mathcal{S}_2^\star$) with respect to the basis $\boldsymbol{h}^{\star a}$. In our work we do not consider $\boldsymbol{v}^\star$ but use only $\boldsymbol{v}$, and we use both the components v^a and v_a; i.e., we are using only the space $\mathcal{S}_2$, and the words contravariant and covariant are used in reference to the components only and not to the spinors themselves. We therefore use the symbol $\boldsymbol{v}^a$ to denote the "contravariant components" of the spinor $\boldsymbol{v}$ in $\mathcal{S}_2$ (i.e., with respect to the basis $\boldsymbol{h}_a$), and v_a to denote the "covariant components" of the same spinor $\boldsymbol{v}$ in $\mathcal{S}_2$ (i.e., with respect to the basis $\boldsymbol{h}^a$). It follows from the above that

$$\boldsymbol{v}^\star\circ\boldsymbol{u}=v^\star_a u^a,\quad \boldsymbol{v}\blacktriangle\boldsymbol{u}=v_a u^a,$$

and since $\boldsymbol{v}^\star\circ\boldsymbol{u}=\boldsymbol{v}\blacktriangle\boldsymbol{u}$, then $v^\star_a u^a=v_a u^a$. Hence, expressions such as $v^\star_a u^a$ appearing in the literature (with the star on $v^\star_a$ omitted) with the above meaning for $v^\star_a$ and u^a are reinterpreted in our formalism as $v_a u^a$ with the above meaning for v_a and u^a. We therefore have an equal expression with $v^\star_a$ replaced by v_a. Moreover,

$$v^\star_a=\boldsymbol{v}^\star\circ\boldsymbol{h}_a=\boldsymbol{v}\blacktriangle\boldsymbol{h}_a=v_a, \tag{A5}$$

i.e., $v^\star_a$ and v_a are equal regardless of their different meanings.

In a like manner in our discussion we have also omitted consideration of the dual space $\bar{\mathcal{S}}_2^\star$ of $\bar{\mathcal{S}}_2$. Consequently, expressions in the literature involving $v_{\dot{a}}^{\star}$ (star omitted) and $v^{\dot{a}}$, which are the complex conjugates of $v_a^{\star}$ and v^a respectively, are replaced in our interpretation by an equivalent expression involving the corresponding equal quantities $v_{\dot{a}} \equiv (\bar{v}_a)$ and $v^{\dot{a}} \equiv (\bar{v}^a)$.

The operation of a spinor connection D'_X for $(\mathcal{S}_2)_q$ valued spinors is extended to $(\mathcal{S}_2^{\star})_q$ valued spinors by the equation

$$(D'_X v^{\star}) \circ u = X(v^{\star} \circ u) - v^{\star} \circ (D'_X u), \tag{A6}$$

where $u(q)$ is an arbitrary $(\mathcal{S}_2)_q$ valued spinor field, and $v^{\star}(q)$ is any $(\mathcal{S}_2^{\star})_q$ valued spinor field. Now for each q we have a map $\mathrm{B}(q)$ from $(\mathcal{S}_2)_q$ to $(\mathcal{S}_2^{\star})_q$ which takes each $v(q)$ in $(\mathcal{S}_2)_q$ into $v^{\star}(q) = \mathrm{B}(q)v(q)$ according to the equation

$$(\mathrm{B}v) \circ u = v \blacktriangle u$$

for all spinor fields $u(q)$. Given any expression involving the $(\mathcal{S}_2^{\star})_q$ valued spinor fields $v^{\star}$ and $D'_X v^{\star}$, we reinterpret it as an equivalent one in which $v^{\star}$ is replaced by the $(\mathcal{S}_2)_q$ valued spinor field $v = \mathrm{B}^{-1} v^{\star}$ and likewise $D'_X v^{\star}$ is replaced by $\mathrm{B}^{-1}(D'_X v^{\star})$. It is important however to observe that

$$\mathrm{B}^{-1}(D'_X v^{\star}) = D'^{*}_X v, \tag{A7}$$

where D'^{*}_X is the dual connection introduced in Sec. IV, as we shall now prove.

To this end, note that the equivalent of Eq. (A6) in our reinterpretation is

$$[\mathrm{B}^{-1}(D'_X v^{\star})] \blacktriangle u = X(v \blacktriangle u) - v \blacktriangle (D'_X u) \tag{A8}$$

and the definition of $D'^{*}_X v$ is [cf. Eq. (64)]

$$(D'^{*}_X v) \blacktriangle u = X(v \blacktriangle u) - v \blacktriangle (D'_X u).$$

Consequently,

$$[\mathrm{B}^{-1}(D'_X v^{\star})] \blacktriangle u = (D'^{*}_X v) \blacktriangle u$$

for arbitrary u, i.e.,

$$\mathrm{B}^{-1}(D'_X v^{\star}) = D'^{*}_X v. \tag{A9}$$

For the sake of completeness, and in order to facilitate comparison with the usual formalism given in the literature, we list the definitions of the components of the following various covariantly differentiated quantities:

$$\begin{aligned} D'_\mu v^{\star} &= v'^{\star}_{a|\mu} h^{\star a}, \\ D'_\mu u &= u'^{a}{}_{|\mu} h_a = u'_{a|\mu} h^a, \\ D'^{*}_\mu u &= u'^{*a}{}_{|\mu} h_a = u'^{*}_{a|\mu} h^a. \end{aligned} \tag{A10}$$

We see then that the familiar Leibnitz product rule

$$\partial_\mu (v_a^{\star} u^a) = v'^{\star}_{a|\mu} u^a + v_a^{\star} u'^{a}{}_{|\mu} \tag{A11}$$

(with the star usually omitted) is the component form of the equation

$$\partial_\mu (v^{\star} \circ u) = (D'_\mu v^{\star}) \circ u + v^{\star} \circ (D'_\mu u),$$

which is the defining equation for the operation of D'_X on $v^{\star}$ as already stated in Eq. (A6). The equivalent of this equation in our reinterpretation is

$$\partial_\mu (v \blacktriangle u) = (D'^{*}_\mu v) \blacktriangle u + v \blacktriangle (D'_\mu u),$$

which is just the defining equation we have used [Eq. (64) in the text] for the dual connection D'^{*}_μ. Furthermore, since the map $v^{\star} \to v = \mathrm{B}^{-1} v^{\star}$ preserves the value of the components, i.e., $v_a^{\star} = v_a$, it follows from Eq. (A9) that

$$v'^{\star}_{a|\mu} = v'^{*}_{a|\mu}. \tag{A12}$$

The reinterpretation of spinor tensor equations is illustrated by the following example. A tensor $\mathbf{L}$ in $\mathcal{S}_2^{\star} \otimes \mathcal{S}_2$ is reinterpreted as a tensor $\mathbf{M}$ in $\mathcal{S}_2 \otimes \mathcal{S}_2$, and, by recalling that $\mathbf{L}$ can be defined as a bilinear functional of the form $L(u, v^{\star})$ and $\mathbf{M}$ can be defined as a bilinear functional of the form $M(u, v)$, the correspondence is made such that

$$L(u, v^{\star}) = M(u, -v). \tag{A13}$$

It will follow then that tensors of the form $s^{\star} t$ in $\mathcal{S}_2^{\star} \otimes \mathcal{S}_2$ will correspond to st in $\mathcal{S}_2 \otimes \mathcal{S}_2$. In addition, the components $L_a{}^b$ and $M_a{}^b$ defined by the equations

$$\mathbf{L} = L_a{}^b h^{\star a} h_b$$

$$\mathbf{M} = M_a{}^b h^a h_b$$

of the corresponding tensors $\mathbf{L}$ and $\mathbf{M}$ are equal, i.e.,

$$L_a{}^b = M_a{}^b.$$

Finally, if a spinor tensor field $\mathbf{L}(q)$ is $(\mathcal{S}_2)_q \otimes (\mathcal{S}_2)_q$, $(\mathcal{S}_2^{\star})_q \otimes (\mathcal{S}_2^{\star})_q$, $(\mathcal{S}_2^{\star})_q \otimes (\mathcal{S}_2)_q$, or $(\mathcal{S}_2)_q \otimes (\mathcal{S}_2^{\star})_q$ valued and if it corresponds in our reinterpretation to the $(\mathcal{S}_2)_q \otimes (\mathcal{S}_2)$ valued tensor field $\mathbf{M}(q)$, it will follow that the covariant derivative $D'_\mu \mathbf{L}$ will correspond to $D'_\mu \mathbf{M}$, $D'^{*}_\mu \mathbf{M}$, $D'^{(*\cdot)}_\mu \mathbf{M}$, or $D'^{(\cdot *)}_\mu \mathbf{M}$ respectively.

APPENDIX B

Schrödinger[6] has derived an expression for what he calls the "covariant derivative" of the Dirac gamma matrices which is quoted and used by several authors.[16,21] This expression is not true for arbitrary spinor connections; the purpose of this appendix is to derive an alternate result which is generally valid. To help in the presentation, we list some additional results and definitions that we will be using: First we have

$$D_\mu \mathbf{\Gamma}_\nu = \Gamma^\lambda_{\mu\nu} \mathbf{\Gamma}_\lambda \tag{B1}$$

which was obtained by adding Eq. (49) to its transpose and using Eq. (23). We define the standard spinor connection matrix Λ_μ by

$$\Lambda_\mu \equiv \Lambda^\alpha_{\mu\beta} l_\alpha l^\beta = l^\alpha \blacktriangle (D_\mu l_\beta) l_\alpha l^\beta = (D_\mu l_\beta) l^\beta, \tag{B2}$$

which can also be expressed in the form

$$\Lambda_\mu = D_\mu (l_\beta l^\beta) - l_\beta (D_\mu l^\beta) = - l_\beta (D_\mu l^\beta). \tag{B3}$$

Note that

$$\Lambda_\mu \blacktriangle l_\beta = D_\mu l_\beta, \quad l^\alpha \blacktriangle \Lambda_\mu = - D_\mu l^\alpha. \tag{B4}$$

In analogy to Eq. (B2) we define an arbitrary spinor connection matrix by

$$\Lambda'_\mu \equiv \Lambda'^{\alpha}_{\mu\beta} l_\alpha l^\beta = \Lambda_\mu + \mathrm{M}_\mu, \tag{B5}$$

where we have made use of Eq. (132).

Using now Eq. (67) in the case of the standard spinor connection, we can write

$$\partial_\mu(l^\alpha \blacktriangle \Gamma_\nu \blacktriangle l_\beta) = l^\alpha \blacktriangle (D_\mu \Gamma_\nu) \blacktriangle l_\beta + (D_\mu l^\alpha) \blacktriangle \Gamma_\nu \blacktriangle l_\beta + l^\alpha \blacktriangle \Gamma_\nu \blacktriangle (D_\mu l_\beta)$$
$$= \Gamma^\lambda_{\mu\nu} l^\alpha \blacktriangle \Gamma_\lambda \blacktriangle l_\beta - l^\alpha \blacktriangle \Lambda_\mu \blacktriangle \Gamma_\nu \blacktriangle l_\beta + l^\alpha \blacktriangle \Gamma_\nu \blacktriangle \Lambda_\mu \blacktriangle l_\beta.$$

If we define

$$(l^\alpha \blacktriangle \Gamma_\nu \blacktriangle l_\beta)_{;\,\mu} \equiv \partial_\mu(l^\alpha \blacktriangle \Gamma_\nu \blacktriangle l_\beta) - \Gamma^\lambda_{\mu\nu} l^\alpha \blacktriangle \Gamma_\lambda \blacktriangle l_\beta, \tag{B6}$$

then the above expression can be put in the form

$$(l^\alpha \blacktriangle \Gamma_\nu \blacktriangle l_\beta)_{;\,\mu} = l^\alpha \blacktriangle (\Gamma_\nu \blacktriangle \Lambda_\mu - \Lambda_\mu \blacktriangle \Gamma_\nu) \blacktriangle l_\beta. \tag{B7}$$

In terms of the Λ'_μ Eq. (B7) becomes

$$(l^\alpha \blacktriangle \Gamma_\nu \blacktriangle l_\beta)_{;\,\mu} - l^\alpha \blacktriangle (\Gamma_\nu \blacktriangle \Lambda'_\mu - \Lambda'_\mu \blacktriangle \Gamma_\nu) \blacktriangle l_\beta$$
$$= l^\alpha \blacktriangle (\mathrm{M}_\mu \blacktriangle \Gamma_\nu - \Gamma_\nu \blacktriangle \mathrm{M}_\mu) \blacktriangle l_\beta \tag{B8}$$

after making use of Eq. (B5). By introducing the operator

$$\nabla'_\mu(l^\alpha \blacktriangle \Gamma_\nu \blacktriangle l_\beta) \equiv (l^\alpha \blacktriangle \Gamma_\nu \blacktriangle l_\beta)_{;\,\mu} - l^\alpha \blacktriangle (\Gamma_\nu \blacktriangle \Lambda'_\mu - \Lambda'_\mu \blacktriangle \Gamma_\nu) \blacktriangle l_\beta, \tag{B9}$$

Eq. (B8) reads

$$\nabla'_\mu(l^\alpha \blacktriangle \Gamma_\nu \blacktriangle l_\beta) = l^\alpha \blacktriangle (\mathrm{M}_\mu \blacktriangle \Gamma_\nu - \Gamma_\nu \blacktriangle \mathrm{M}_\mu) \blacktriangle l_\beta. \tag{B10}$$

Observe that the right side of Eq. (B10) is always zero iff $\mathrm{M}_\mu = \varphi_\mu \mathrm{I}$; thus we have

Theorem. The equation

$$\nabla'_\mu(l^\alpha \blacktriangle \Gamma_\nu \blacktriangle l_\beta) = 0 \tag{B11}$$

is always true iff K_μ has the form $\mathrm{K}_\mu = \varphi_\mu \mathrm{I}_2$.

Consequently, the requirement in Eq. (B11) restricts the spinor connection to cases 3 and 4 in Sec. IV of the text. Note, in particular, that for case 2 of Sec. IV where $\mathrm{K}_\mu = i\varphi_\mu \mathrm{I}_2$, we get

$$\nabla'_\mu(l^\alpha \blacktriangle \Gamma_\nu \blacktriangle l_\beta) = \varphi_\mu l^\alpha \blacktriangle (\Gamma^5 \blacktriangle \Gamma_\nu - \Gamma_\nu \blacktriangle \Gamma^5) \blacktriangle l_\beta$$
$$= 2\varphi_\mu\, l^\alpha \blacktriangle \Gamma^5 \blacktriangle \Gamma_\nu \blacktriangle l_\beta. \tag{B12}$$

For the purpose of comparing our results with Schrödinger's matrix notation, let γ_ν, Λ'_μ and M_μ be the matrices whose σ, τ-th elements are

$$(\gamma_\nu)^\sigma{}_\tau = l^\sigma \blacktriangle \Gamma_\nu \blacktriangle l_\tau, \quad (\Lambda'_\mu)^\sigma{}_\tau = l^\sigma \blacktriangle \Lambda'_\mu \blacktriangle l_\tau, \quad (M_\mu)^\sigma{}_\tau = l^\sigma \blacktriangle \mathrm{M}_\mu \blacktriangle l_\tau$$

respectively. Then Eqs. (B6), (B7), (B9), (B10) become

$$\gamma_{\nu;\,\mu} \equiv \partial_\mu \gamma_\nu - \Gamma^\lambda_{\mu\nu} \gamma_\lambda, \tag{B13}$$

$$\gamma_{\nu;\,\mu} = \gamma_\nu \Lambda_\mu - \Lambda_\mu \gamma_\nu, \tag{B14}$$

$$\nabla'_\mu \gamma_\nu \equiv \gamma_{\nu;\,\mu} - (\gamma_\nu \Lambda'_\mu - \Lambda'_\mu \gamma_\nu), \tag{B15}$$

$$\nabla'_\mu \gamma_\nu = M_\mu \gamma_\nu - \gamma_\nu M_\mu. \tag{B16}$$

The quantity $\nabla'_\mu \gamma_\nu$ defined by Eq. (B15) is what Schrödinger and others call the covariant derivative of the Dirac gamma matrices (in their notation they have $-\Gamma_\mu$ for our Λ'_μ). Note, however, that by virtue of Eq. (B16), $\nabla'_\mu \gamma_\nu \neq 0$ in general. Hence Schrödinger's Eq. (8), which corresponds to our Eq. (B11), is not valid for arbitrary spinor connections, in which case it should be replaced by Eq. (B16).

*Also at the Instituto Nacional de Energía Nuclear.

[1]N. J. Hicks, *Notes on Differential Geometry* (Van Nostrand Reinhold, New York, 1965); also see R. L. Bishop and S. I. Goldberg, *Tensor Analysis on Manifolds* (Macmillan, New York, 1968).

[2]E. Cartan, Bull. Soc. Math. France **41**, 53 (1913); E. Cartan, *Theory of Spinors* (MIT Press, Cambridge, Mass. 1966); O. Laporte and G. E. Uhlenbeck, Phys. Rev. **37**, 1380 (1931); L. Witten, Phys. Rev. **113**, 357 (1959) and **115**, 206 (1959); R. Penrose, Ann. Phys. (N.Y.) **10**, 171 (1960); F. A. E. Pirani in *Lectures on General Relatively*, S. Deser and K. W. Ford, eds., *Brandeis Summer Institute in Theoretical Physics*, 1964, *Vol* 1 (Prentice Hall, Englewood Cliffs, New Jersey 1965); A. Lichnerowitz in *Relatvity, Groups, and Topology*, C. DeWitt and B. DeWitt, eds., Les Houches 1963 (Gordon and Breach, New York, 1964), and references therein; C. G. Oliveira, J. Math. Phys. **11**, 2440 (1970); E. Schmutzer, *Relativtische Physik* (B. G. Teubner Verlagsgesellschaft, Leipzig 1968). Also see the references in Refs. 4–6, 12–17, 21.

[3]C. P. Luehr and M. Rosenbaum, J. Math. Phys. **9**, 2225 (1968).

[4]L. Infeld and B. L. van der Waerden, Sitzber. Preuss. Akad. Wiss. Phys.-Math. Kl., 380 (1933); W. L. Bade and H. Jehle, Rev. Mod. Phys. **25**, 714 (1953) and references therein; W. C. Parke and H. Jehle, in *Lectures in Theoretical Physics*, *Vol VII A*, edited by W. E. Brittin and A. O. Barut, University of Colorado Summer Institute for Theoretical Physics, 1964 (Gordon and Breach, New York, 1968); also see Schmutzer, Ref. 12.

[5]V. Fock, Z. Physik **57**, 261 (1929).

[6]E. Schrödinger, Sitzber. Preuss. Akad. Wiss. Phys.-Math. Kl, XI, 105 (1932); V. Bargmann, Sitzber. Preuss. Akad. Wiss. Phys.-Math. Kl. 346 (1932); H. Loos, Ann Phys. (N.Y.) **25**, 91 (1963).

[7]Covariant spinor components $(\mathrm{E}_\mu)_{\dot{a}b}$ of E are defined by $E_\mu = (E_\mu)_{\dot{a}b} \bar{h}^a h^b$, from where it follows that

$$(\mathbf{E}_\mu)_{\dot{a}b} = \mathbf{E}_\mu \blacktriangle \bar{h}_a h_b = \mathbf{E}_\mu \odot \mathbf{I}_4 \blacktriangle \bar{h}_a h_b = (\mathbf{I}_4)_{\mu\dot{a}b}.$$

The quantities $(\mathbf{I}_4)_{\mu\dot{a}b}$ are the hybrid four-vector spinor components of $\mathbf{I}_4$ introduced in I [Eq. (58b)], and explicitly shown there in [Eq. (61)] to be proportional to the identity and the three Pauli matrices. Note that $(\mathbf{I}_4)_{\mu\dot{a}b}$ corresponds to the $\sigma_{\mu\dot{a}b}$ used in the literature.

[8]This relation follows from the observation that the left side, $\mathbf{L}$, is an element of $(\mathcal{S}_2 \otimes \mathcal{S}_2) \oplus (\bar{\mathcal{S}}_2 \otimes \bar{\mathcal{S}}_2)$ having the property $\tilde{\mathbf{L}} = -\mathbf{L}$, and must therefore have the form $\alpha \mathbf{I}_2 + \beta \bar{\mathbf{I}}_2$. The fact that $(\mathbf{I}_2 \blacktriangle \mathbf{L})_s = (\bar{\mathbf{I}}_2 \blacktriangle \mathbf{L})_s = 4\mathbf{A} \odot \mathbf{B}$ is sufficient to evaluate α and β in order to obtain the right side.

[9]We use the convention "iff" to stand for "if and only if."

[10]A simple generalization of Eq. (54) is the well known fact that the difference $D'_X - D''_X$ of two arbitrary connections is a linear transformation at each point (i.e., a tensor field).

[11]Precisely, we assume that for each point q we are given the spaces $(\mathcal{S}_2)_q$ and $(\bar{\mathcal{S}}_2)_q$, and just one isomorphic map from $(\bar{\mathcal{S}}_2)_q \otimes_H (\mathcal{S}_2)_q$ onto $\mathcal{M}_q$. Under this map the image of an element $\mathrm{Y}(q)$ in $(\bar{\mathcal{S}}_2)_q \otimes_H (\mathcal{S}_2)_q$ is an element in $\mathcal{M}_q$ denoted as $\mathrm{y}(q)$ or Y_q.

[12]E. Schmutzer, Z. Naturforsch. **15a**, 355 (1960).

[13]E. Schmutzer, Z. Naturforsch. **17a**, 707 (1962).

[14]Note by virtue of Eqs. (73) and (76) we can write for the general spinor connection coefficients:

$$\Lambda'^{\,b}_{\mu a} = \Lambda^b_{\mu a} + h^b \blacktriangle \mathrm{K}_\mu \blacktriangle h_a,$$

which for $\mathbf{K}_\mu = i\varphi_\mu \mathbf{I}_2$ becomes

$$\Lambda'^{\,b}_{\mu a} = \Lambda^b_{\mu a} + i\varphi_\mu \delta^b_a.$$

Inserting Eq. (118) into this result gives an expression which is equivalent to the one derived by Schmutzer (Eq. 11) in Acta Phys. Hungar. **17**, 57 (1964).

[15]H. S. Ruse, Proc. Roy. Soc. (Edinburgh) **57**, 97 (1937).

[16]J. G. Fletcher, Nuovo Cimento **8**, 451 (1958).

[17]J. L. Anderson, *Principles of Relativity Physics* (Academic, New York, 1967).

[18]Note that $\mathbf{xy} : \mathfrak{P}' = \mathbf{xy} : \mathfrak{Q}'$

[19]E.g., $C(6,8)\, [(\bar{u}_1 u_1) \otimes (\bar{u}_2 u_2) \otimes (\bar{u}_3 u_3) \otimes (\bar{u}_4 u_4)] = (u_3 \blacktriangle u_4) [(\bar{u}_1 u_1) \otimes (\bar{u}_2 u_2) \otimes (\bar{u}_3 \bar{u}_4)]$.

[20]Equations (158) and (159) are the equivalent in intrinsic notation of the spinor curvature tensors derived by Infeld and Van der Waerden.

[21]D. R. Brill and J. A. Wheeler, Rev. Mod. Phys. **29**, 465 (1957).

The Hamiltonian and generating functional for a nonrelativistic local current algebra*†

Ralph Menikoff

Department of Physics, University of Pennsylvania, Philadelphia, Pennsylvania 19174
(Received 11 February 1974)

The nonrelativistic current algebra with conserved current consisting of $\rho(\mathbf{x})$, the particle number density, and $\mathbf{J}(\mathbf{x})$, the flux density of particles, is studied. The Hamiltonian for any time reversal invariant system of spinless particles, interacting via a two-body interaction potential, is expressed as a Hermitian form in the currents. This leads to a functional equation for the generating functional, which is the ground state expectation value of $\exp[i\int d\mathbf{x}\rho(\mathbf{x})f(\mathbf{x})]$. In the N/V limit an expression for the generating functional in terms of correlation functions is given. Representations of the exponentiated current algebra which are translation invariant satisfy the cluster decomposition property, and those which have different Hamiltonians are shown to be unitarily inequivalent.

1. INTRODUCTION

Several physicists[1-5] have investigated the possibility of expressing field theory in terms of local currents instead of the canonical fields. To gain further insight into writing field theory in terms of local currents, we study in this paper the nonrelativistic equal-time current algebra consisting of $\rho(\mathbf{x})$, the particle number density, and $\mathbf{J}(\mathbf{x})$, the flux density of particles. We seek to determine representations of the current algebra suitable for describing physical systems associated with a specific Hamiltonian H. A generating functional is used for this purpose. The representation incorporates certain general physical constraints on the system, such as current conservation, time reversal invariance, and translation invariance. The dynamics, which is not studied here, would be obtained by considering the time dependent local currents, $\rho(\mathbf{x}, t) = \exp(itH)\rho(\mathbf{x})\exp(-itH)$ and $\mathbf{J}(\mathbf{x}, t) = \exp(itH)\mathbf{J}(\mathbf{x})\exp(-itH)$, in the representation determined by the equal-time current algebra and the Hamiltonian.

In this approach we start with nonrelativistic quantum mechanics in second quantized form. Then $\rho(\mathbf{x})$ and $\mathbf{J}(\mathbf{x})$ can be written in terms of the canonical annihilation and creation field operators, and their commutation relations computed. The commutation relations between $\rho(\mathbf{x})$ and $\mathbf{J}(\mathbf{x})$ are taken as our starting point.[1] We will be especially interested in representations corresponding to the "N/V limit," since they describe systems with "an infinite number of degrees of freedom" and have many features similar to those of quantum field theory. In this case the quantum mechanics of N particles in a box of volume V is considered. The limit is taken as $N \to \infty$ and $V \to \infty$ in such a way that $N/V \to \bar{\rho}$, the average density of the system. In statistical mechanics this is known as the thermodynamic limit. It is applicable to systems with a large number of particles when surface effects can be neglected. In this paper we deal only with the case of zero temperature.

In Sec. 2 the ρ, J current algebra is defined as in Ref. 1. For our purposes it is more convenient to deal with the group obtained by exponentiating the currents. This is reviewed along with its unitary representations as given by Goldin.[6] The generating functional $L(f)$, the ground state expectation value of $\exp[i\int d\mathbf{x}\rho(\mathbf{x})f(\mathbf{x})]$, is introduced and its use in defining a representation is discussed.

In Sec. 3 we consider the Hamiltonian for a time reversal invariant system of spinless particles. Dashen and Sharp[1] have given a formal expression for the Hamiltonian in terms of currents as the sum of a kinetic energy term plus a potential energy term. A rigorous definition for the kinetic energy term has been given by Goldin and Sharp[7] for the Hamiltonian of a system of free bosons by considering it as a densely defined Hermitian form. We generalize this form to obtain the Hamiltonian for a system of interacting particles. The resulting expression for the Hamiltonian combines the kinetic energy and potential energy into one factored term. Two points of view may be taken in this section:

(i) Given a representation in which a Hamiltonian exists, the Hamiltonian is expressed in terms of $\rho(\mathbf{x})$ and $\mathbf{J}(\mathbf{x})$ as a densely defined Hermitian form, or

(ii) given a representation, an operator with all the properties of a Hamiltonian is defined from a densely defined Hermitian form.

The form of the Hamiltonian leads in Sec. 4 to a functional equation for the generating functional. Supplemented by the appropriate boundary conditions, this equation determines a representation associated with the Hamiltonian.

In Sec. 5 the generating functional for a representation corresponding to a system of N particles is expressed in terms of correlation functions. This form of the generating functional is extended to the N/V limit representations. Next, we consider the consequences of translation invariance and the cluster decomposition property. The results are analogous to those in field theory[8]; the ground state is unique and is the only momentum eigenfunction. Furthermore, it is shown that representations corresponding to different Hamiltonians are unitarily inequivalent. Finally, the particle nature of the N/V limit representations is studied. The representation restricted to a finite volume is found to be the direct sum of N-particle representations. Thus the N/V limit representation is "locally Fock."

These results are illustrated by examples in the following paper (to be published) where, in the N/V limit, the generating functional along with the Hamiltonian and functional equation are given exactly in the following cases: (i) free Bose gas, (ii) noninteracting bosons in an external potential, (iii) free Fermi gas, (iv) bosons in one dimension with the two-body interacting potential $U(x) = 2/x^2$.

2. REVIEW OF THE NONRELATIVISTIC CURRENT ALGEBRA

This section contains a brief review of the nonrelativistic current algebra and its representations. (For a more extensive review see Refs. 7 and 9.)

In terms of the canonical field operators $\psi(\mathbf{x})$ and $\psi^\dagger(\mathbf{x})$ which satisfy either the commutation (−) or anticommutation (+) relations

$$[\psi(\mathbf{x}),\ \psi(\mathbf{y})]_\pm=[\psi^\dagger(\mathbf{x}),\ \psi^\dagger(\mathbf{y})]_\pm=0,$$
$$[\psi(\mathbf{x}),\ \psi^\dagger(\mathbf{y})]_\pm=\delta(\mathbf{x}-\mathbf{y}), \tag{2.1}$$

the particle density and flux density are given by

$$\rho(\mathbf{x})=\psi^\dagger(\mathbf{x})\,\psi(\mathbf{x}),$$
$$\mathbf{J}(\mathbf{x})=(\hbar/2im)\,[\psi^\dagger(\mathbf{x})\nabla\psi(\mathbf{x})-\nabla\psi^\dagger(\mathbf{x})\,\psi(\mathbf{x})]. \tag{2.2}$$

Henceforth the mass of the particles and $\hbar$ will be set equal to 1. Dashen and Sharp[1] showed that the equal-time commutation relations between $\rho(\mathbf{x})$ and $\mathbf{J}(\mathbf{y})$ are given by

$$[\rho(f_1),\ \rho(f_2)]=0,$$
$$[\rho(f),\ J(\mathbf{g})]=i\rho(\mathbf{g}\cdot\nabla f), \tag{2.3}$$
$$[J(\mathbf{g}_1),\ J(\mathbf{g}_2)]=iJ(\mathbf{g}_2\cdot\nabla\mathbf{g}_1-\mathbf{g}_1\cdot\nabla\mathbf{g}_2)$$

for both bosons and fermions. We have used the smeared currents $\rho(f)=\int d\mathbf{x}\,\rho(\mathbf{x})f(\mathbf{x})$ and $J(\mathbf{g})=\int d\mathbf{x}\,\mathbf{J}(\mathbf{x})\cdot\mathbf{g}(\mathbf{x})$, where $f(\mathbf{x})$ and each component of $\mathbf{g}(\mathbf{x})$ belong to a suitable class of test functions, for example, Schwartz's space $\mathcal{S}$, the set of C^∞ functions of fast decrease at infinity.

The commutation relations (2.3) will be taken as the starting for the work of this paper. We will also assume current conservation, $(d/dt)\,\rho(\mathbf{x},t)+\nabla\cdot\mathbf{J}(\mathbf{x},t)=0$. This is expressed in terms of a Hamiltonian by

$$[H,\rho(f)]=-iJ(\nabla f). \tag{2.4}$$

Since the local currents correspond to physical observables, we require them to be self-adjoint operators, $\rho(f)^\dagger=\rho(f)$ and $J(\mathbf{g})^\dagger=J(\mathbf{g})$. However, they may be unbounded operators. For this reason it is convenient to work with the unitary operators formed by exponentiating the currents,[6]

$$U(f)=\exp[i\rho(f)]\quad\text{and}\quad V(\varphi_t^{\mathbf{g}})=\exp[itJ(\mathbf{g})], \tag{2.5}$$

where $(d/dt)\,\varphi_t^{\mathbf{g}}(\mathbf{x})=\mathbf{g}\circ\varphi_t^{\mathbf{g}}(\mathbf{x})$, $\varphi_0^{\mathbf{g}}(\mathbf{x})=\mathbf{x}$, and "$\circ$" stands for composition, i.e., $\mathbf{g}\circ\varphi(\mathbf{x})=\mathbf{g}(\varphi(\mathbf{x}))$.

Remark: $\varphi_t^{\mathbf{g}}(\mathbf{x})$ is the flow corresponding to the vector field $\mathbf{g}(\mathbf{x})$. This has the following physical interpretation. Imagine a fluid with velocity field $\mathbf{v}=\mathbf{g}(\mathbf{x})$. Then $\varphi_t^{\mathbf{g}}(\mathbf{x})$ is the position of a particle which starts at point $\mathbf{x}$, after a time t.

The exponentiated currents form a group with the following multiplication law:

$$U(f_1)U(f_2)=U(f_1+f_2),$$
$$V(\varphi)\,U(f)=U(f\circ\varphi)\,V(\varphi), \tag{2.6}$$
$$V(\varphi_1)\,V(\varphi_2)=V(\varphi_2\circ\varphi_1).$$

Throughout the rest of this paper we will be concerned with representations of the group of exponentiated currents. Goldin[6] has analyzed these representations using the Gel'fand–Vilenkin formalism for "nuclear Lie groups."[10] The results listed below will be used in our study.

The Hilbert space for every continuous representation of $U(f)$ and $V(\varphi)$ is unitarily equivalent to one with direct sum decomposition,

$$\mathcal{H}=\int^{\oplus}_{F\in\mathcal{S}'}d\mu(F)\mathcal{H}_F,$$

where μ is a cylindrical measure on $\mathcal{S}'$, the continuous dual of $\mathcal{S}$ (i.e., $\mathcal{S}'$ is the set of continuous real linear functionals on $\mathcal{S}$). For physical reasons explained below we will only be concerned with the case when $\dim\mathcal{H}_F=1$. The Hilbert space is then the space of square integrable functions on $\mathcal{S}'$ with respect to the measure μ; i.e., $\mathcal{H}=L^2{}_\mu(\mathcal{S}')$.

$U(f)$ acts as a multiplication operator on elements of $\mathcal{H}$, i.e.,

$$U(f)\Psi(F)=\exp[i(F,f)]\,\Psi(F),\quad\forall\,\Psi(F)\in\mathcal{H}. \tag{2.7}$$

In order to express the action of $V(\varphi)$, we need the mapping φ^* from $\mathcal{S}'$ onto $\mathcal{S}'$ defined by

$$(\varphi^*F,f)=(F,f\circ\varphi),\qquad\forall F\in\mathcal{S}'\text{ and }f\in\mathcal{S}.$$

The action of $V(\varphi)$ is then given by

$$V(\varphi)\Psi(F)=\chi_\varphi(F)\Psi(\varphi^*F)\left(\frac{d\mu(\varphi^*F)}{d\mu(F)}\right)^{1/2},\quad\forall\Psi(F)\in\mathcal{H}, \tag{2.8}$$

where $d\mu(\varphi^*F)/d\mu(F)$ is the Radon–Nikodym derivative of $\mu(\varphi^*F)$ with respect to $\mu(F)$ and $\chi_\varphi(F)$, called the multiplier, is a complex valued function of modulus one. In order for the Radon–Nikodym derivative to exist, the measure μ must be quasi-invariant with respect to the set of flows; i.e., for any measurable set $X\subset\mathcal{S}'$ and any flow φ, $\mu(X)=0$ iff $\mu(\varphi^*X)=0$. The group law requires the multipliers to satisfy the equation

$$\chi_{\varphi_2}(F)\chi_{\varphi_1}(\varphi_2{}^*F)=\chi_{\varphi_1\circ\varphi_2}(F)\ \text{ a.e.} \tag{2.9}$$

A representation of $U(f)$ and $V(\varphi)$ is thus completely determined by a measure μ and a system of multipliers $\chi_\varphi(F)$.

The representation corresponding to the quantum mechanics of N identical particles has a measure concentrated on delta functions[6,11]; i.e., the measure is only nonzero on functionals of the form

$$F(\mathbf{x})=\sum_{j=1}^{N}\delta(\mathbf{x}-\mathbf{x}_j)\quad\text{and}\quad d\mu(F)=d\sigma(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N).$$

By a suitable choice of measure the ground state for a given Hamiltonian may be taken as $\Omega(F)=1$. [In the N-particle representation the measure is given by $d\mu(F)=d\psi^*\psi(\mathbf{x}_1\cdots\mathbf{x}_N)$ where $\psi(\mathbf{x}_1\cdots\mathbf{x}_N)$ is the ground state wave function.]

Remarks: (1) The ground state $|\Omega\rangle$ is cyclic with respect to $U(f)$. In other words, the set of states of the form $\sum_{j=1}^N a_jU(f_j)|\Omega\rangle$ is dense in $\mathcal{H}=L^2{}_\mu(\mathcal{S}')$. The continuity of the representation then implies $\mathcal{H}$ is separable.

(2) Dicke and Goldin[12] have proposed a definition of

statistics for representations of the exponentiated current algebra based on the multipliers. They found that the only "well-behaved" irreducible representations of $U(f)$ and $V(\varphi)$ with $\dim \mathcal{H}_F = 1$ are those corresponding to either bosons or fermions.

(3) $\chi_\varphi(F) = 1$ always satisfies Eq. (2.9). This corresponds to a representation for bosons.[12] Thus a boson representation can be completely defined by giving a measure μ and setting $\chi_\varphi(F) = 1$. There may be other systems of multipliers corresponding to bosons.

(4) The representations with $\dim \mathcal{H}_F > 1$ have the following physical significance:

(i) If $U(f)$ and $V(\varphi)$ are reducible, the representation can correspond to particles with different masses or with internal degrees of freedom (e.g., spin). In the latter case, additional local currents need to be added to obtain a complete set of observables (e.g., spin density). Spin has been treated briefly by Grodnik and Sharp[13] and Goldin.[6]

(ii) If $U(f)$ and $V(\varphi)$ are "well behaved" and irreducible, the representation corresponds to parastatistics.[14]

Thus by restricting ourself to the case $\dim \mathcal{H}_F = 1$, we only will be considering identical spinless particles (either bosons or fermions).

Much information about the representation can be obtained from the ground state expectation value of $U(f)$. This is known as the generating functional and is denoted by $L(f)$. Thus,

$$L(f) = (\Omega, U(f)\Omega) = \int_{\mathcal{S}'} d\mu(F) \exp[i(F, f)]. \tag{2.10}$$

The generating functional for any representation has the following properties:

$$\text{(i)}\quad L(f) = L(-f)^*. \tag{2.11}$$

This follows from the relation $U(f)^\dagger = U(-f)$.

$$\text{(ii)}\quad L(0) = 1. \tag{2.12}$$

Since the ground state is normalized, $(\Omega, \Omega) = 1$.

$$\text{(iii)}\quad |L(f)| \leq 1. \tag{2.13}$$

This follows from the condition that $U(f)$ be a unitary operator. (iv) $L(f)$ is a positive functional. This means

$$\sum_{j,k=1}^{N} a_j^* a_k L(f_k - f_j) \geq 0, \quad \forall a_j \in C,\ f_j \in \mathcal{S}, \text{ and finite } N. \tag{2.14}$$

This property follows from the requirement that the inner product on $\mathcal{H}$ be positive: i.e.,

$$\left(\sum_{j=1}^{N} a_j U(f_j)\Omega,\ \sum_{k=1}^{N} a_k U(f_k)\Omega\right) \geq 0.$$

It can be shown that a continuous functional $L(f)$ satisfying the above four properties determines a measure μ for a representation of $U(f)$.[6] If μ is a quasi-invariant measure and the multipliers are known (e.g., this is the case for bosons), a representation of both $U(f)$ and $V(\varphi)$ is completely determined. Otherwise, it is necessary to know, $L(f, \varphi) = (\Omega, U(f)V(\varphi)\Omega)$, in order to completely determine a representation of the exponentiated currents.

Remark: The exponentiated algebra and generating functional techniques we will be using are similar to those introduced by Araki[8] in studying the CCR's. They have been applied to find representations of the canonical commutation relations describing a nonrelativistic infinite free bose gas by Araki and Woods.[15] A similar approach was used in a study of the CAR's by Araki and Wyss.[16]

3. THE HAMILTONIAN EXPRESSED IN TERMS OF CURRENTS

In this section we will express the Hamiltonian of a physical system in terms of the currents $\rho(\mathbf{x})$ and $\mathbf{J}(\mathbf{x})$. A formal expression for the Hamiltonian abstracted from canonical field theory was given by Dashen and Sharp.[1] In terms of the canonical field operators (satisfying either the CCR's or CAR's) the Hamiltonian for a system of particles with a two-body interaction potential $V(\mathbf{x})$ is given by

$$H = \tfrac{1}{2}\int d\mathbf{x}\, \nabla\psi^\dagger(\mathbf{x}) \cdot \nabla\psi(\mathbf{x}) + \tfrac{1}{2}\int\int d\mathbf{x}d\mathbf{y}\, \psi^\dagger(\mathbf{x})\psi^\dagger(\mathbf{y})V(\mathbf{x}-\mathbf{y})\psi(\mathbf{y})\psi(\mathbf{x}). \tag{3.1}$$

The potential energy term can be written as

$$\text{P.E.} = \tfrac{1}{2}\int\int d\mathbf{x}d\mathbf{y}\, \rho(\mathbf{x})[\rho(\mathbf{y}) - \delta(\mathbf{x}-\mathbf{y})]V(\mathbf{x}-\mathbf{y}). \tag{3.2}$$

To obtain the kinetic energy term, we introduce the quantity $\mathbf{K}(\mathbf{x}) = \nabla\rho(\mathbf{x}) + 2i\mathbf{J}(\mathbf{x})$. In terms of the canonical fields $\mathbf{K}(\mathbf{x}) = 2\psi^\dagger(\mathbf{x})\nabla\psi(\mathbf{x})$. Then formally the kinetic energy is given by

$$\text{K.E.} = \tfrac{1}{8}\int d\mathbf{x}\, \mathbf{K}(\mathbf{x})^\dagger \frac{1}{\rho(\mathbf{x})}\mathbf{K}(\mathbf{x}). \tag{3.3}$$

By combining Eqs. (3.2) and (3.3) the Hamiltonian is given by

$$H = \tfrac{1}{8}\int d\mathbf{x}\, \mathbf{K}(\mathbf{x})^\dagger \frac{1}{\rho(\mathbf{x})}\mathbf{K}(\mathbf{x}) + \tfrac{1}{2}\int\int d\mathbf{x}d\mathbf{y}\, \rho(\mathbf{x})[\rho(\mathbf{y}) - \delta(\mathbf{x}-\mathbf{y})]V(\mathbf{x}-\mathbf{y}). \tag{3.4}$$

In the N/V limit there are two problems with writing H as the sum of the total K.E. plus the total P.E.:

(i) The K.E./particle and the P.E./particle are finite. However, the total K.E. and the total P.E. are infinite. Therefore, it is unclear just how each term in Eq. (3.4) is to be defined.

(ii) From statistical mechanics the ground state energy is proportional to the number of particles; $E_0 \to \epsilon N$ as N becomes large. In the limit, $E_0 = \infty$. Thus, the sum of the two terms in Eq. (3.4) is also ill defined as it stands.

These problems lead us to consider an alternative expression for the Hamiltonian. First, it is necessary to define the quantity "$1/\rho(\mathbf{x})$" which appears in the kinetic energy term. In the representation corresponding to a free Bose gas, a rigorous definition has been given by Goldin and Sharp.[7] By extending their definition we can combine the K.E. and P.E. into one term and obtain a well-defined expression for the Hamiltonian as a densely defined Hermitian form.

We denote the Hamiltonian for a free Bose gas by

$$H_0=\tfrac{1}{8}\int d\mathbf{x}\,\mathbf{K}(\mathbf{x})^\dagger \frac{1}{\rho(\mathbf{x})}\mathbf{K}(\mathbf{x}).$$

It is defined as follows[7]:

Let

$\nu=\mathrm{Span}\{w(\mathbf{x})\rho(\mathbf{x})\Phi;\ \Phi\in\mathcal{H}$ and $w(\mathbf{x})\in C^\infty$ functions of polynomial growth at infinity$\}$.

ν is a set of vector valued distributions "proportional" to $\rho(\mathbf{x})$. "$1/\rho(\mathbf{x})$" is defined as a map from $\nu\times\nu\to\mathcal{S}'$ in the following way: Let $v_1=w_1(\mathbf{x})\rho(\mathbf{x})\Phi_1$ and $v_2=w_2(\mathbf{x})\rho(\mathbf{x})\Phi_2$. Then

$$\left(v_1,\ \frac{1}{\rho(\mathbf{x})}v_2\right)=(\Phi_1,\ w_1(\mathbf{x})w_2(\mathbf{x})\rho(\mathbf{x})\Phi_2).$$

Let $\mathcal{D}=\mathrm{Span}\{\exp[i\rho(f)]\,\Omega;\ \forall f\in\mathcal{S}$ and $\Omega=$ the ground state$\}$.

$\mathcal{D}$ is a dense linear manifold in $\mathcal{H}$. For the free Bose gas it can be shown $\mathbf{K}(\mathbf{x})\mathcal{D}\subset\nu$. As a result

$$(\Phi_1,\ H_0\Phi_2)=\tfrac{1}{8}\int d\mathbf{x}\left(\mathbf{K}(\mathbf{x})\Phi_1,\ \frac{1}{\rho(\mathbf{x})}\mathbf{K}(\mathbf{x})\Phi_2\right).$$

is a well-defined Hermitian form for all Φ_1 and $\Phi_2\in\mathcal{D}$.

Remark: The seemingly natural operation of $1/\rho(\mathbf{x})$ on $v=w_1(\mathbf{x})\rho(\mathbf{x})\Phi_1$, $[1/\rho(\mathbf{x})]v=w_1(\mathbf{x})\Phi_1$, is not in fact well defined, since, if v can also be written as $v=w_2(\mathbf{x})\rho(\mathbf{x})\Phi_2$, it does not necessarily follow that $w_1(\mathbf{x})\Phi_1=w_2(\mathbf{x})\Phi_2$.

By generalizing the form of H_0 we will show for an interacting system that:

(1) H is defined as a bilinear form on the dense domain,

$$\mathcal{D}=\mathrm{Span}\{U(f)\Omega; f\in\mathcal{S}\text{ and }\Omega=\text{the ground state}\},$$

by

$$H=\tfrac{1}{8}\int d\mathbf{x}\,\tilde{\mathbf{K}}(\mathbf{x})^\dagger\frac{1}{\rho(\mathbf{x})}\tilde{\mathbf{K}}(\mathbf{x}),\quad \text{where } \tilde{\mathbf{K}}(\mathbf{x})=\mathbf{K}(\mathbf{x})-\mathbf{A}(\mathbf{x},\rho).$$

The operator $\mathbf{A}(\mathbf{x},\rho)$ will be defined precisely later.

(2) H is both Hermitian and positive.

(3) $(\Phi, H\Omega)=0$, $\Phi\in\mathcal{D}$, where $\Omega=$ the ground state.

We start by assuming there is a representation of $U(f)$ and $V(\varphi)$ on a Hilbert space $\mathcal{H}$ along with a Hamiltonian H satisfying the following conditions:

(i) There is a normalized state of lowest energy; the ground state Ω. We require $H\geq 0$. Thus the zero of energy is chosen such that

$$H\Omega=0. \tag{3.5}$$

(ii) $\mathcal{D}=\mathrm{Span}\{U(f)\Omega; f\in\mathcal{S}\}$ is dense in $\mathcal{H}$ and $\mathcal{D}\subset$ the domain of H.

(iii) Current conservation

$$[H,\rho(f)]=-iJ(\nabla f). \tag{3.6}$$

(iv) There is an antiunitary time reversal operator T such that

$$T\rho(f)T^{-1}=\rho(f),\quad TJ(\mathrm{g})T^{-1}=-J(\mathrm{g}),\quad\text{and}\quad T\Omega=\Omega. \tag{3.7}$$

We will also make use of the relation

$$e^A B e^{-A}=\sum_{n=0}^{\infty}\frac{1}{n!}(\mathrm{ad}^n A)B, \tag{3.8}$$

where $(\mathrm{ad}^0 A)B=B$ and $(\mathrm{ad}^n A)B=[A,(\mathrm{ad}^{n-1}A)B]$.

Two simple results we will need can easily be derived from Eqs. (2.3), (3.6), and 3.8. These are

$$[e^{i\rho(f)},J(\mathrm{g})]=-\tfrac{1}{2}i[\exp(i\rho(f)),K(\mathrm{g})]=-\rho(\mathrm{g}\cdot\nabla f)\times\exp[i\rho(f)] \tag{3.9}$$

and

$$[\exp(i\rho(f)),H]=[-J(\nabla f)+\tfrac{1}{2}\rho(\nabla f\cdot\nabla f)]\exp[i\rho(f)]. \tag{3.10}$$

Our first theorem shows time reversal invariance and current conservation are sufficient to determine the matrix elements of $J(\mathrm{g})$ and H in terms of those for ρ.

Theorem 1: Suppose there is a representation of $U(f)$ and $V(\varphi)$ satisfying conditions (i)–(iv) above. Let $|f\rangle=\exp[i\rho(f)]\Omega$. Then,

$$\langle f_1|J(\mathrm{g})|f_2\rangle=\tfrac{1}{2}\langle f_1|\rho(\mathrm{g}\cdot\nabla(f_1+f_2))|f_2\rangle \tag{3.11}$$

and

$$\langle f_1|H|f_2\rangle=\tfrac{1}{2}\langle f_1|\rho(\nabla f_1\cdot\nabla f_2)|f_2\rangle. \tag{3.12}$$

Proof: Using time reversal invariance [Eq. (3.7)], we have

$$\begin{aligned}\langle f_1|J(\mathrm{g})|f_2\rangle&=(TJ(\mathrm{g})\exp[i\rho(f_2)]\,\Omega,\ Te^{i\rho(f_1)}\Omega)\\&=-(\Omega,\ \exp[i\rho(f_2)]J(\mathrm{g})\exp[-i\rho(f_1)]\Omega).\end{aligned}$$

Substituting in Eq. (3.9) twice and also using Eq. (2.6), we obtain

$$\begin{aligned}\langle f_1|J(\mathrm{g})|f_2\rangle=-(\Omega,\ &\exp[-i\rho(f_1)]\\&\times[J(\mathrm{g})-\rho(\mathrm{g}\cdot\nabla(f_1+f_2))]\exp[i\rho(f_2)]\Omega).\end{aligned}$$

Therefore,

$$\langle f_1|J(\mathrm{g})|f_2\rangle=\tfrac{1}{2}\langle f_1|\rho(\mathrm{g}\cdot\nabla(f_1+f_2))|f_2\rangle.$$

Next, by applying current conservation and using Eqs. (3.5), (3.10), and (3.11) we have

$$\begin{aligned}\langle f_1|H|f_2\rangle&=(\exp[i\rho(f_1)]\Omega,\ [H,\ \exp(i\rho(f_2))]\Omega)\\&=(\exp[i\rho(f_1)]\Omega,\ [J(\nabla f_2)-\tfrac{1}{2}\rho(\nabla f_2\cdot\nabla f_2)]\\&\qquad\times\exp[i\rho(f_2)]\Omega)\\&=\tfrac{1}{2}\langle f_1|\rho(\nabla f_1\cdot\nabla f_2)|f_2\rangle.\end{aligned}$$

Remarks: (1) A Hermitian form on a dense set of states does not necessarily determine an unbounded operator. If the form determines a Hermitian operator it may have many (or no) self-adjoint extensions depending on the choice of its domain. Therefore, Eqs. (3.11) and (3.12) are not sufficient to determine J and H as operators.

(2) As a result of Eq. (3.12)

$$(\rho(f_1)\Omega,\ H\rho(f_2)\Omega)=\tfrac{1}{2}(\Omega,\rho(\nabla f_1\cdot\nabla f_2)\Omega)$$

In the N/V limit, for a translational invariant system

$$(\Omega, \rho(\mathbf{x})\Omega) = \bar{\rho}, \quad \text{the average density.}$$

Therefore, $(\rho(f_1)\Omega,\ H\rho(f_2)\Omega) = \frac{1}{2}\bar{\rho}\int d\mathbf{x}\,\nabla f_1 \cdot \nabla f_2$.

Now, using the matrix elements of H given by Eq. (3.12) an expression for H will be derived in terms of $\rho(\mathbf{x})$ and $\mathbf{J}(\mathbf{x})$. For this purpose, we first determine an operator $\mathbf{A}(\mathbf{x},\rho)$ having the property that $K(\mathrm{g})\Omega = A(\mathrm{g},\rho)\Omega$. Consider a representation with Hilbert space $\mathcal{H} = L^2_\mu(\mathcal{S}')$ and ground state $\Omega(F) = 1$. Let $A(\mathrm{g})$ be the operator of multiplication by $(K(\mathrm{g})\Omega)(F)$ defined by

$$(A(\mathrm{g})\Phi)(F) = (K(\mathrm{g})\Omega)(F)\Phi(F)$$

and

$$\text{Domain } A(\mathbf{g}) = \{\Phi(F) \in \mathcal{H};\ \int d\mu(F)\,|(A(\mathrm{g})\Phi)(F)|^2 < \infty\}.$$

Since $\exp[i\rho(f)]$ is multiplication by $\exp[i(F,f)]$ we have $[A(g),\ \exp(i\rho(f))] = 0$. Also $A(g)\exp[i\rho(f)]\Omega = \exp[i\rho(f)]K(\mathrm{g})\Omega$. As a result the domain of $A(\mathrm{g})$ includes the set $\mathcal{D}$, and therefore it is a dense set. By time reversal invariance it follows that $(K(\mathrm{g})\Omega)(F)^* = (K(\mathrm{g})\Omega)(F)$. Thus $A(\mathrm{g})$ is Hermitian. Moreover, $A(\mathrm{g})$ is self-adjoint. To prove this, it is sufficient to show that $[\mathrm{Range}(A \pm i)]^\perp = \{0\}$. Let $\Phi(F) \in [\mathrm{Range}\,(A \pm i)]^\perp$. Then

$$\int_{\mathcal{S}'} d\mu(F)\Phi(F)(A(\mathrm{g}) \pm i)\,\Psi(F) = 0, \quad \forall \Psi(F) \in \text{Domain } A(\mathrm{g}).$$

Pick $\Psi(F) = \chi_c(F) =$ the characteristic function for the set $C \subset \mathcal{S}'$. Then

$$\int_c d\mu(F)\,\Phi(F)\,[(K(\mathrm{g})\Omega)(F) \pm i] = 0, \quad \forall C \subset \mathcal{S}'$$

and therefore

$$\Phi(F)\,[(K(\mathrm{g})\Omega)(F) \pm i] = 0,$$

Since $[(K(\mathrm{g})\Omega)(F) \pm i] \neq 0$, we have $\Phi(F) = 0$. Therefore, $A(\mathrm{g})$ is self-adjoint.

It will be useful to express $A(\mathrm{g})$ as a function of ρ. This is possible since the ρ's are multiplication operators and polynomials in ρ applied to the ground state are dense. We proceed as follows: Let $\mathcal{J} = \{f_j;\ j = 1, 2, \cdots\}$ be a countable dense set of test functions (e.g., in Schwartz's space, finite linear combinations with rational coefficients of the Hermite functions).

Let $\mathcal{D}' = \mathrm{Span}\{\exp[i\rho(f)\Omega];\ f \in \mathcal{J}\}$. Since $\mathcal{D}$ is dense in $\mathcal{H}$, by the continuity of the representation it follows that $\mathcal{D}'$ is also dense. However, the states $\{\exp[i\rho(f)]\Omega;\ f \in \mathcal{J}\}$ are neither orthogonal nor linearly independent. It is therefore convenient to orthogonalize them using the Gram–Schmit procedure. Let

$$|h_1\rangle = U(f_1)\Omega,$$

$$|h_n\rangle = \sum_{j=1}^{n} a_j^{(n)} U(f_j)\Omega, \quad \text{such that } (h_i, h_j) = \delta_{i,j}.$$

Clearly, Span $\{h_j;\ j = 1, 2, \cdots\} = \mathcal{D}'$. Since this set is dense, we can write

$$K(\mathrm{g})\Omega = \sum_{n=1}^{\infty} b_n(\mathrm{g})\left(\sum_{j=1}^{n} a_j^{(n)} U(f_j)\right)\Omega.$$

The desired operator $A(\mathrm{g},\rho)$ is defined by

$$A(\mathrm{g},\rho) = \sum_{n=1}^{\infty} b_n(\mathrm{g})\left(\sum_{j=1}^{n} a_j^{(n)} U(f_j)\right). \tag{3.13}$$

Furthermore, $K(\mathrm{g})\Omega$ depends linearly on g. As a result $b_n(\mathrm{g})$ is a linear distribution; $b_n(\mathrm{g}) = \int d\mathbf{x}\, b_n(\mathbf{x}) \cdot \mathrm{g}(\mathbf{x})$. Therefore we can write, $A(\mathrm{g},\rho) = \int d\mathbf{x}\,\mathrm{g}(\mathbf{x}) \cdot \mathbf{A}(\mathbf{x},\rho)$, where

$$\mathbf{A}(\mathbf{x},\rho) = \sum_{n=1}^{\infty} \mathbf{b}_n(\mathbf{x})\left(\sum_{j=1}^{n} a_j^{(n)} U(f_j)\right).$$

Next, define $\tilde{\mathbf{K}}(\mathbf{x}) = \mathbf{K}(\mathbf{x}) - \mathbf{A}(\mathbf{x},\rho)$. By construction we have

$$\tilde{\mathbf{K}}(\mathbf{x})\Omega = 0 \tag{3.14}$$

and

$$[\exp(i\rho(f)), \tilde{\mathbf{K}}(\mathbf{x})] = [\exp(i\rho(f)), \mathbf{K}(\mathbf{x})] = -2i\nabla f(\mathbf{x})\rho(\mathbf{x}) \times \exp[i\rho(f)]$$

Theorem 2: $\frac{1}{8}\int d\mathbf{x}\,\tilde{\mathbf{K}}(\mathbf{x})^\dagger\,[1/\rho(\mathbf{x})]\tilde{K}(\mathbf{x})$ is a well-defined Hermitian form with domain $\mathcal{D}$. Furthermore,

$$\left\langle \Phi_1 \left| \frac{1}{8}\int d\mathbf{x}\,\tilde{\mathbf{K}}(\mathbf{x})^\dagger \frac{1}{\rho(\mathbf{x})}\tilde{K}(\mathbf{x}) \right| \Phi_2 \right\rangle = (\Phi_1, H\,\Phi_2), \quad \forall \Phi_1, \Phi_2 \in \mathcal{D}. \tag{3.15}$$

Proof: Observe that

$$\left\langle f_1 \left| \frac{1}{8}\int d\mathbf{x}\,\tilde{\mathbf{K}}(\mathbf{x})^\dagger \frac{1}{\rho(\mathbf{x})}\tilde{\mathbf{K}}(\mathbf{x}) \right| f_2 \right\rangle$$

$$= \frac{1}{8}\int d\mathbf{x} \left\langle \tilde{\mathbf{K}}(\mathbf{x})\exp[i\rho(f_1)]\Omega,\ \frac{1}{\rho(\mathbf{x})}\tilde{\mathbf{K}}(\mathbf{x})\exp[i\rho(f_2)]\Omega \right\rangle$$

$$= \frac{1}{8}\int d\mathbf{x} \Big\langle -2i\nabla f_1(\mathbf{x})\rho(\mathbf{x})\exp[i\rho(f_1)]\Omega,$$

$$\times \frac{1}{\rho(\mathbf{x})}(-2i)\nabla f_2(\mathbf{x})\rho(\mathbf{x})\exp[i\rho(f_2)]\Omega \Big\rangle$$

$$= \tfrac{1}{2}(\exp[i\rho(f_1)]\Omega,\ \rho(\nabla f_1 \cdot \nabla f_2)\exp[i\rho(f_2)]\Omega)$$

$$= \langle f_1 | H | f_2 \rangle.$$

This can be extended by linearity to the domain $\mathcal{D}$.

Formal manipulations can easily be performed with this form of H. For example, we can verify current conservation:

$$\left[\frac{1}{8}\int d\mathbf{x}\,\tilde{\mathbf{K}}(\mathbf{x})^\dagger \frac{1}{\rho(\mathbf{x})}\tilde{\mathbf{K}}(\mathbf{x}), \rho(f)\right]$$

$$= \frac{1}{8}\int d\mathbf{x}\Big(\tilde{\mathbf{K}}(\mathbf{x})^\dagger \frac{1}{\rho(\mathbf{x})}[\tilde{\mathbf{K}}(\mathbf{x}), \rho(f)]$$

$$+ [\tilde{\mathbf{K}}(\mathbf{x})^\dagger,\ \rho(f)]\frac{1}{\rho(\mathbf{x})}\tilde{\mathbf{K}}(\mathbf{x})\Big)$$

$$= \frac{1}{4}\int d\mathbf{x}\left(\tilde{\mathbf{K}}(\mathbf{x})^\dagger \frac{1}{\rho(\mathbf{x})}\rho(\mathbf{x})\nabla f(\mathbf{x}) - \nabla f(\mathbf{x})\rho(\mathbf{x})\frac{1}{\rho(\mathbf{x})}\tilde{\mathbf{K}}(\mathbf{x})\right)$$

$$= \frac{1}{4}\int d\mathbf{x}\,\nabla f(\mathbf{x})\,[\tilde{\mathbf{K}}(\mathbf{x})^\dagger - \tilde{\mathbf{K}}(\mathbf{x})] = -iJ(\nabla f).$$

In the last step we used $\mathbf{A}(\mathbf{x},\rho)^\dagger = \mathbf{A}(\mathbf{x},\rho)$, which follows from time reversal invariance.

These manipulations can be cast into a rigorous form by showing that $(\Phi_1, H\rho(f)\Phi_2) - (\rho(f)\Phi_1, H\Phi_2) = -i(\Phi_1, J(\nabla f)\Phi_2)$, $\forall\ \Phi_1, \Phi_2 \in \mathcal{D}$ follows from Eqs. (3.11) and (3.12).

In an alternative approach, only a representation of $U(f)$ and $V(\varphi)$ is assumed. Then the Hermitian form in Eq. (3.15) is used to define an operator with all the properties of a Hamiltonian. It is necessary to show the Hermitian form is positive. This can be done if one assumes $(\Phi, \rho(f)\Phi) \geq 0, \forall\ f \in \mathcal{S}$ such that $f(\mathbf{x}) \geq 0$ and $\Phi \in$ Domain of $\rho(f)$. This is physically necessary since the expectation value of the density in any state must be positive. In the representation with the Hilbert space

$$\mathcal{H} = L^2_\mu(\mathcal{S}'),$$
$$(\Phi, \rho(f)\Phi) = \int d\mu(F)(F,f)|\Phi(F)|^2.$$

Therefore, the measure is concentrated on functionals $F \in \mathcal{S}'$ such that $(F, f) \geq 0, \forall\ f(\mathbf{x}) \geq 0$.

Theorem 3:

$$\tfrac{1}{8}\int d\mathbf{x}\,\tilde{\mathbf{K}}(\mathbf{x})^\dagger \frac{1}{\rho(\mathbf{x})} \tilde{\mathbf{K}}(\mathbf{x})$$

is a positive Hermitian form.

Proof: Let

$$\Phi = \sum_{j=1}^{n} a_j U(f_j)\Omega,$$

$$\left\langle \Phi, \tfrac{1}{8}\int d\mathbf{x}\,\tilde{\mathbf{K}}(\mathbf{x})^\dagger \frac{1}{\rho(\mathbf{x})} \tilde{\mathbf{K}}(\mathbf{x})\Phi \right\rangle$$

$$= \tfrac{1}{2}\sum_{j,k=1}^{n} a_k^* a_j \langle f_k | \rho(\nabla f_k \cdot \nabla f_j) | f_j \rangle$$

$$= \tfrac{1}{2}\int d\mu(F) \left(F, \left|\sum_{j=1}^{n} a_j \nabla f_j \exp[i(F, f_j)]\right|^2\right) \geq 0.$$

The following theorem of Friedrichs[17] tells us that the Hermitian form in Eq. (3.15) defines a positive self-adjoint operator.

Friedrichs' theorem: A positive semidefinite Hermitian form $\{\psi_1, \psi_2\}$ defined on a dense linear set R in a Hilbert space $\mathcal{H}$ can be extended by continuity to a positive semidefinite Hermitian form on a larger linear set $R' > R$ which consists of elements $\psi \in \mathcal{H}$ such that, for some sequence $\psi_n \in R$, $\|\psi - \psi_n\| \to 0$ and $\{\psi_n - \psi_m, \psi_n - \psi_m\} \to 0$. Furthermore, there exists a unique positive self-adjoint operator A such that $\mathcal{D}(A) \subset R'$ and $\{\psi_1, \psi_2\} = (\psi_1, A\psi_2)$, $\forall \psi_1 \in R'$ and $\psi_2 \in \mathcal{D}(A)$.

Therefore, the expression $\frac{1}{8}\int d\mathbf{x}\,\tilde{\mathbf{K}}(\mathbf{x})^\dagger [1/\rho(\mathbf{x})]\,\tilde{\mathbf{K}}(\mathbf{x})$ can be used to define an operator with all the properties of a Hamiltonian. If we had begun with a Hamiltonian, it would not be clear whether this would be the same as the one constructed from Friedrichs' theorem due to the technical question concerning the domain of H. We will not pursue this matter further here.

Remarks: (1) Eqs. (3.11) and (3.12) and the result that Eq. (3.12) defines a positive Hermitian form have been obtained independently by Aref'eva[18] using different methods.

(2) Coester and Haag[19] have discussed a similar form for the Hamiltonian in terms of the canonical relativistic scalar fields $\varphi(\mathbf{x})$ and $\pi(\mathbf{x})$.

(3) There is an interesting similarity between the form of the Hamiltonian derived above and the Hamiltonian for a particle in a magnetic field:

$$H_0 = \tfrac{1}{8}\int d\mathbf{x}\,\mathbf{K}(\mathbf{x})^\dagger \frac{1}{\rho(\mathbf{x})}\mathbf{K}(\mathbf{x}) \longleftrightarrow H_0 = \frac{p^2}{2m},$$

$$H = \tfrac{1}{8}\int d\mathbf{x}\,[\mathbf{K}(\mathbf{x}) - \mathbf{A}(\mathbf{x}, \rho)]^\dagger \frac{1}{\rho(\mathbf{x})}[\mathbf{K}(\mathbf{x}) - \mathbf{A}(\mathbf{x}, \rho)]$$

$$\longleftrightarrow H = \frac{1}{2m}\left(p - \frac{e}{c}A\right)^2.$$

In our case, for an interaction the free Hamiltonian is modified by $\mathbf{K}(\mathbf{x}) \to \mathbf{K}(\mathbf{x}) - \mathbf{A}(\mathbf{x}, \rho)$ while in quantum mechanics the free Hamiltonian is modified by $p \to p - (e/c)A$. There is also a difference. In quantum mechanics $\dot{x}_{\text{free}} = p/m \to \dot{x} = (1/m)[p - (e/c)A]$ while in our case $\dot{\rho} = -\nabla\cdot\mathbf{J}$ remains true for both the free case and the interaction.

(4) In terms of the canonical fields both the currents [Eq. (2.2)] and the Hamiltonian [Eq. (3.1)] have the same form for both bosons and fermions. In terms of the currents (as we will see in the following paper) the free Hamiltonian has a different form for bosons and fermions. This is not as surprising as it might appear at first sight. In quantum mechanics the free Hamiltonian for bosons and fermions is formally the same; $H = -\frac{1}{2}\sum_{j=1}^{n} \partial^2/\partial x_j^2$. However, the domains are different; symmetric functions for bosons and antisymmetric functions for fermions. As a result the free Bose Hamiltonian and the free Fermi Hamiltonian are different operators with distinct spectra.[13]

(5) Hopefully there will be a systematic method for determining $\mathbf{A}(\mathbf{x}, \rho)$ for a given potential. Equation (3.4) might be used as a guide towards this end.

4. FUNCTIONAL DIFFERENTIAL EQUATION FOR $L(f)$

Using the results of the previous section we will derive a functional differential equation for the generating functional $L(f)$. When supplemented by the appropriate boundary conditions this equation can be used to determine $L(f)$ and hence a representation corresponding to a given physical system. This has been done in great detail for the free Bose gas in Ref. 20 (see also Goldin and Sharp[21]).

We start with the ground state condition [Eq. (3.14)], $\tilde{\mathbf{K}}(\mathbf{x})\Omega = 0$. Forming the inner product of $\tilde{\mathbf{K}}(\mathbf{x})\Omega$ with $\exp[-i\rho(f)]\Omega$, we find $0 = (\Omega, \exp[i\rho(f)]\tilde{\mathbf{K}}(\mathbf{x})\Omega)$. Using the definition of $\tilde{\mathbf{K}}(\mathbf{x})$ and Eq. (3.11), we then have

$$0 = (\Omega, \exp[i\rho(f)]\,[\nabla\rho(\mathbf{x}) - i\nabla f(\mathbf{x})\rho(\mathbf{x})]\Omega) - (\Omega, \exp[i\rho(f)]\,\mathbf{A}(\mathbf{x}, \rho)\Omega). \tag{4.1}$$

Both terms can be evaluated using functional derivatives of $L(f)$. Since

$$\frac{1}{i}\frac{\delta}{\delta f(\mathbf{x})}L(f) = (\Omega, \exp[i\rho(f)]\rho(\mathbf{x})\Omega),$$

Eq. (4.1) can be written

$$[\nabla - i\nabla f(\mathbf{x})]\,\frac{1}{i}\,\frac{\delta}{\delta f(\mathbf{x})}\,L(f) = \mathbf{A}\left(\mathbf{x},\ \frac{1}{i}\,\frac{\delta}{\delta f}\right) L(f). \qquad (4.2)$$

The solutions of this equation which are physically admissable are restricted by several conditions. These include the general properties [Eqs. (2.11)—(2.14)] of a generating functional, namely:

(1) $L(f) = L(-f)^*$.

(2) $L(0) = 1$.

(3) $|L(f)| \leq 1$.

(4) $L(f)$ is a positive functional.

Other conditions may include:

(5) $L(f)$ is an extremal solution in the sense that it cannot be written as a convex linear combination of two other solutions. This has the effect of requiring the representation of $U(f)$ and $V(\varphi)$ to be irreducible (see Ref. 20, Theorem 3.4).

In the N/V limit we can also use translational invariance or the cluster decomposition property. (These will be explained further in the next section.)

(6) $\frac{1}{i}\,\frac{\delta}{\delta f(\mathbf{x})}\,L(f)\Big|_{f=0} = (\Omega, \rho(\mathbf{x})\Omega) = \bar{\rho}$.

(7) $L(f) = L(f_{\mathbf{a}})$, where $f_{\mathbf{a}}(\mathbf{x}) = f(\mathbf{x} - \mathbf{a})$.

(8) $\lim_{\mathbf{a}\to\infty} L(f + h_{\mathbf{a}}) = L(f)\,L(h)$, where $h_{\mathbf{a}}(\mathbf{x}) = h(\mathbf{x} - \mathbf{a})$.

For the free Bose gas, Eqs. (4.2) becomes

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

In this case it is known[20] that conditions (2)—(6) uniquely determine $L(f)$. It is not known whether these conditions are sufficient in other cases. Furthermore, it is not yet known how to determine the $\mathbf{A}(\mathbf{x}, \rho)$ corresponding to a specific interaction. However, in the following paper, $\mathbf{A}(\mathbf{x}, \rho)$ and $L(f)$ are given explicitly in the N/V limit, and Eq. (4.2) is verified for three additional cases:

(1) Bosons in an external potential,

$$\mathbf{A}(\mathbf{x}, \rho) = \rho(\mathbf{x})\nabla \ln \bar{\rho}(\mathbf{x}).$$

(2) Free Fermi gas in one dimension,

$$A(x, \rho) = 2\rho(x)\int \frac{dy}{x - y}\,\rho(y)$$

(3) $2/x^2$ interaction in one dimension,

$$A(x, \rho) = 4\rho(x)\int \frac{dy}{x - y}\,\rho(y).$$

5. *L(f)* IN THE *N/V* LIMIT

In this section we discuss some general properties of the generating functional $L(f)$ in the N/V limit. First, for an N-particle representation we find an expression for $L(f)$ in terms of correlation functions. This form of $L(f)$ is extended to the N/V limit when the correlation functions satisfy appropriate bounds. Next, we consider the consequences of translational invariance and the cluster decomposition property. It is shown that different generating functionals give rise to unitarily inequivalent representations of $U(f)$. Finally, the particle nature of the N/V limit representation is examined.

A. Expansion of *L(f)* in terms of correlation functions

The N-particle representations of the current algebra [Eq. (2.3)] have been studied by Grodnik and Sharp,[11] and Goldin.[6] We will use the correspondence between these representations and conventional quantum mechanics to obtain an expression for $L(f)$ in terms of correlation functions. An N-particle representation is defined on the Hilbert space:

$$\mathcal{H} = \begin{cases} L^2{}_S(R^N), & \text{the totally symmetric functions for bosons,} \\ L^2{}_A(R^N), & \text{the totally antisymmetric functions for fermions,} \end{cases}$$

Acting on $\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N) \in \mathcal{H}$, we have

$$\rho(\mathbf{x})\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \sum_{k=1}^{N} \delta(\mathbf{x} - \mathbf{x}_k)\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N)$$

or

$$\rho(f)\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \sum_{k=1}^{N} f(\mathbf{x}_k)\,\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N) \qquad (5.1)$$

and

$$\mathbf{J}(\mathbf{x})\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \frac{1}{2i}\sum_{k=1}^{N} [-\nabla_x\delta(\mathbf{x} - \mathbf{x}_k) + 2\delta(\mathbf{x} - \mathbf{x}_k)\nabla_{x_k}]\,\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N)$$

or $\qquad (5.2)$

$$J(\mathbf{g})\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \frac{1}{2i}\sum_{k=1}^{N} [2\mathbf{g}(\mathbf{x}_k)\cdot\nabla_{x_k} + (\nabla\cdot\mathbf{g})(\mathbf{x}_k)]\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N).$$

The generating functionals are given by

$$L(f) = (\Omega,\ \exp[i\rho(f)]\Omega) \qquad (5.3)$$

$$= \int d\mathbf{x}_1 \cdots \int d\mathbf{x}_N \exp[if(\mathbf{x}_1)] \cdots \exp[if(\mathbf{x}_N)]\Omega^*\Omega(\mathbf{x}_1, \ldots, \mathbf{x}_N),$$

where $\Omega(\mathbf{x}_1, \ldots, x_N) =$ the ground state wavefunction, and

$$L(f, \mathbf{g}) = (\Omega,\ \exp[i\rho(f)]\exp[iJ(\mathbf{g})]\Omega) = \int d\mathbf{x}_1 \cdots \int d\mathbf{x}_n\,\Omega^*(\mathbf{x}_1 \cdots \mathbf{x}_N) \times \prod_{k=1}^{N} \exp[if(\mathbf{x}_k)]\exp[ij(\mathbf{x}_k, \mathbf{g})]\,\Omega(\mathbf{x}_1 \cdots \mathbf{x}_N), \qquad (5.4)$$

where $j(\mathbf{x}, \mathbf{g}) = (1/2i)[2\mathbf{g}(\mathbf{x})\cdot\nabla + (\nabla\cdot\mathbf{g})(\mathbf{x})]$.

Remarks: (1) One can write

$$\exp[ij(\mathbf{x}, \mathbf{g})]\,\psi(\mathbf{x}) = \psi \circ \varphi(\mathbf{x})\left(\det \frac{\partial}{\partial x_m}\,\varphi_n(\mathbf{x})\right)^{1/2},$$

where φ is the flow corresponding to the vector field $\mathbf{g}$. The factor $[\det(\partial/\partial x_m)\varphi_n(\mathbf{x})]$ is the Jacobian of the transformation $\mathbf{x} \to \varphi(\mathbf{x})$, and is necessary in order for $\exp[iJ(\mathbf{g})]$ to be unitary. (See Ref. 6.)

(2) $\mathcal{H}$ is unitarily equivalent to $L^2_\mu(\mathcal{S}')$ where the measure is concentrated on $\{F \in \mathcal{S}';\ F = \sum_{k=1}^{N} \delta(\mathbf{x} - \mathbf{x}_k)\}$ and $d\mu(F) = d\Omega^*\Omega(\mathbf{x}_1 \cdots \mathbf{x}_N)$. Furthermore, the ground state is given by $\Omega(F) = 1$. Boson and fermion representations are distinguished by the multipliers $\chi_\varphi(F)$.

For a representation defined by $L(f, \mathbf{g})$ it is convenient to think in terms of the n-point functions, $(\Omega, \rho(\mathbf{x}_1) \cdots \rho(\mathbf{x}_m)\mathbf{J}(\mathbf{x}_{m+1}) \cdots \mathbf{J}(\mathbf{x}_n)\Omega)$, instead of the measure and multipliers on $\mathcal{H}=L^2_\mu(\mathcal{S}')$. By the reconstruction theorem (see Ref. 22) the n-point functions determine a representation of the current algebra. All the n-point functions can be obtained by taking functional derivatives of $L(f, \mathbf{g})$. Therefore, $L(f, \mathbf{g})$ determines a representation of the current algebra.

Remarks: (1) There is a slight complication in determining the n-point functions from $L(f, \mathbf{g})$. The ρ's are obtained directly by taking functional derivatives:

$$\frac{1}{i}\frac{\delta}{\delta f(\mathbf{x}_1)}\cdots\frac{1}{i}\frac{\delta}{\delta f(\mathbf{x}_m)}L(f,0)\Big|_{f=0}$$

$$=(\Omega,\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_m)\Omega).$$

Since the J's do not commute,

$$\frac{1}{i}\frac{\delta}{\delta \mathbf{g}(\mathbf{x}_1)}\cdots\frac{1}{i}\frac{\delta}{\delta \mathbf{g}(\mathbf{x}_m)}L(0,\mathbf{g})\Big|_{\mathbf{g}=0}$$

$$=\frac{1}{m!}\sum_{\tau}(\Omega,\mathbf{J}(\mathbf{x}_{\tau_1})\cdots\mathbf{J}(\mathbf{x}_{\tau_m})\Omega),$$

where $\sum_\tau=$ the sum over all permutations of $(1, 2, \ldots, m)$.

However, by using the commutation relations (2.3), $(\Omega, \mathbf{J}(\mathbf{x}_1)\ldots\mathbf{J}(\mathbf{x}_m)\Omega)$ can be obtained inductively from

$$\frac{1}{i}\frac{\delta}{\delta \mathbf{g}(\mathbf{x}_1)}\cdots\frac{1}{i}\frac{\delta}{\delta \mathbf{g}(\mathbf{x}_m)}L(0,\mathbf{g})\Big|_{\mathbf{g}=0}$$

plus the n-point functions of lower order ($n<m$).

(2) The J's (in the n-point functions) can be replaced by ρ's using the operator $\mathbf{A}(\mathbf{x},\rho)$ defined in Sec. 3:

$$J(\mathbf{x}_1)\Omega=-\tfrac{1}{2}i[A(\mathbf{x}_1,\rho)-\nabla\rho(\mathbf{x}_1)]\Omega,$$

$$\mathbf{J}(\mathbf{x}_1)\mathbf{J}(\mathbf{x}_2)\Omega=-\tfrac{1}{2}i\{[\mathbf{J}(\mathbf{x}_1),(\mathbf{A}(\mathbf{x}_2,\rho)-\nabla\rho(\mathbf{x}_2))]+(\mathbf{A}(\mathbf{x}_2,\rho)-\nabla\rho(\mathbf{x}_2))\mathbf{J}(\mathbf{x}_1)\}\Omega.$$

Using the functional representation[11] $\mathbf{J}(\mathbf{x})=\rho(\mathbf{x})(1/i)\nabla\,\delta/\delta\rho(\mathbf{x})+F(\rho(\mathbf{x}))$, we have

$$[\mathbf{J}(\mathbf{x}_1),\mathbf{A}(\mathbf{x}_2,\rho)]=\rho(\mathbf{x}_1)\frac{1}{i}\nabla_{\mathbf{x}_1}\frac{\delta}{\delta\rho(\mathbf{x}_1)}\mathbf{A}(\mathbf{x}_2,\rho).$$

Thus $\mathbf{J}(\mathbf{x}_1)\mathbf{J}(\mathbf{x}_2)\Omega$ can be obtained from a function of ρ on Ω. This procedure can be extended to $\mathbf{J}(\mathbf{x}_1)\cdots\mathbf{J}(\mathbf{x}_n)\Omega$. Therefore, a representation of the current algebra is determined by $\mathbf{A}(\mathbf{x},\rho)$ and $L(f)$, provided the derivatives of $\mathbf{A}(\mathbf{x},\rho)$ are well behaved. Goldin[6] used an expression similar to $\mathbf{A}(\mathbf{x},\rho)$ to give rigorous sufficient conditions for recovering a representation of the current algebra from that of the exponentiated currents.

The n-point functions of ρ can be related to the correlation functions, which are defined as follows (for the N-particle representation):

$$R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)=\begin{cases}1, & \text{for } n=0,\\ N!/(N-n)!\int d\mathbf{x}_{n+1}\cdots\int d\mathbf{x}_N\,\Omega^*\Omega(\mathbf{x}_1,\ldots,\mathbf{x}_N), & 0<n\le N,\\ 0, & N<n.\end{cases}$$

Using the symmetry of the wavefunction and Eq. (5.1), we obtain

$$(\Omega,\rho(\mathbf{x}_1)\Omega)=R_1(\mathbf{x}_1),$$

$$(\Omega,\rho(\mathbf{x}_1)\rho(\mathbf{x}_2)\Omega)=R_2(\mathbf{x}_1,\mathbf{x}_2)+\delta(\mathbf{x}_1-\mathbf{x}_2)R_1(\mathbf{x}_1),$$

$$(\Omega,\rho(\mathbf{x}_1)\rho(\mathbf{x}_2)\rho(\mathbf{x}_3)\Omega)=R_3(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3)+\sum_{\text{perm}}\delta(\mathbf{x}_1-\mathbf{x}_2)R_2(\mathbf{x}_2,\mathbf{x}_3)+\delta(\mathbf{x}_1-\mathbf{x}_2)\,\delta(\mathbf{x}_1-\mathbf{x}_3)R_1(\mathbf{x}_1).$$

Thus $(\Omega,\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_n)\Omega)$ is the sum of n terms, each term being the sum over permutations of the variables $\mathbf{x}_1,\ldots,\mathbf{x}_n$ of the product of m delta functions multiplied by R_{n-m}.

Remark: The above expressions are independent of the number of particles in the representation. As we will see they are also true in the N/V limit. If $\rho(\mathbf{x})$ can be written in terms of the canonical field operators as, $\rho(\mathbf{x})=\psi(\mathbf{x})^\dagger\psi(\mathbf{x})$ (Eq. 2.2), the correlation functions are the n-point functions for the canonical fields:

$$R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)=(\Omega,\psi^\dagger(\mathbf{x}_1)\cdots\psi^\dagger(\mathbf{x}_n)\,\psi(\mathbf{x}_n)\cdots\psi(\mathbf{x}_1)\Omega).$$

The correlation functions have the physical interpretation,

$$(1/n!)R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)=\left(\begin{array}{l}\text{The probability of finding } n\\ \text{particles at the points } \mathbf{x}_1,\ldots,\mathbf{x}_n\\ \text{regardless of the positions of the}\\ \text{remaining particles}\end{array}\right).$$

We can now obtain an expression for $L(f)$ in terms of the correlation functions. Let $F(\mathbf{x})=\exp[if(\mathbf{x})]-1$, and note that

$$\exp[if(\mathbf{x}_1)]=F(\mathbf{x}_1)+1,$$

$$\exp[if(\mathbf{x}_1)]\exp[if(\mathbf{x}_2)]=F(\mathbf{x}_1)F(\mathbf{x}_2)+F(\mathbf{x}_1)+F(\mathbf{x}_2)+1,$$

$$\exp[if(\mathbf{x}_1)]\cdots\exp[if(\mathbf{x}_n)]=\sum_{\text{perm}}\sum_{j=0}^{n}\frac{1}{[j!(n-j)!]}\prod_{k=1}^{j}F(\mathbf{x}_{\tau_k}).\tag{5.5}$$

Substituting Eq. (5.5) into Eq. (5.3) for $L(f)$ and using the symmetry of the wavefunction Ω and appropriate change of variable labels in the integrals, we obtain

$$L(f)=\sum_{n=0}^{N}\frac{1}{n!}\int d\mathbf{x}_1\cdots\int d\mathbf{x}_n F(\mathbf{x}_1)\cdots F(\mathbf{x}_n)R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n).\tag{5.6}$$

[As a check notice the leading term, the one without any δ functions, in the n-point function

$$(\Omega,\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_n)\Omega)=\frac{1}{i}\frac{\delta}{\delta f(\mathbf{x}_1)}\cdots\frac{1}{i}\frac{\delta}{\delta f(\mathbf{x}_n)}L(f)\Big|_{f=0}$$

is just $R_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)$.]

In order to carry out the N/V limit, we introduce the following notation: Let $R_n^{(N)}=$ The nth correlation function for N particles in a box of volume V, and let

$$a_n^{(N)}=\int_v d\mathbf{x}_1\cdots\int_v d\mathbf{x}_n F(\mathbf{x}_1)\cdots F(\mathbf{x}_n)R_n^{(N)}(\mathbf{x}_1,\ldots,\mathbf{x}_n).$$

The generating functional for the N-particle representation can now be written as $L_N(f)=\sum_{n=0}^\infty(1/n!)a_n^{(N)}$.

If the N/V limit is to exist, we might expect $R_n^{(N)} \to R_n \; \forall \, n$ and

$$L_N(f) \to L(f) = \sum_{n=0}^{\infty} \frac{1}{n!} a_n$$

where

$$a_n = \int_{-\infty}^{\infty} d\mathbf{x}_1 \cdots \int_{-\infty}^{\infty} d\mathbf{x}_n F(\mathbf{x}_1) \cdots F(\mathbf{x}_n) R_n(\mathbf{x}_1 \cdots \mathbf{x}_n).$$

In the next theorem we give sufficient conditions for the N/V limit of $L(f)$ to exist. These conditions are probably adequate for most physical systems. [They will be used in the following paper to explicitly calculate $L(f)$ in the N/V limit for several examples.]

Theorem 4: If $R_n^{(N)}(\mathbf{x}_1, \ldots, \mathbf{x}_n) \to R_n(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ and $|R_n^{(N)}| \leq c^n n^{n/2} \; \forall n, N$ for some constant c, then $L_N(f) \to L(f)$.

Remark: Girard[23] used an expression similar to Eq. (5.6) in studying the thermodynamics of a free Bose gas in terms of the local current algebra. The proof given below is essentially the same as the one he used.

Lemma: The series $S(c) = \sum_{n=0}^{\infty} (1/n!) c^n n^{n/2}$ converges for all c.

Proof: We use the ratio test. Let S_n = the ratio of the $(n+1)$th term to the nth term. Then

$$\begin{aligned} S_n &= \frac{[1/(n+1)!]\, c^{n+1} (n+1)^{(n+1)/2}}{(1/n!)\, c^n n^{n/2}} \\ &= c(n+1)^{(n-1)/2}/n^{n/2} \\ &= c(n+1)^{-1/2}(1+1/n)^{n/2} \\ &\to c\, 0\, e^{1/2} = 0. \end{aligned}$$

Therefore, the series for S converges.

Proof of theorem 4: Since $R_n^{(N)} \to R_n$ and $|R_n^{(N)}| \leq c^n n^{n/2}$, it follows that $|R_n| \leq c^n n^{n/2}$. As a result,

$$\begin{aligned} |a_n| &\leq \int d\mathbf{x}_1 \cdots \int d\mathbf{x}_n |F(\mathbf{x}_1) \cdots F(\mathbf{x}_n) R_n| \\ &\leq \left(c \int d\mathbf{x} \,|\exp[if(\mathbf{x})] - 1|\right)^n n^{n/2}. \end{aligned}$$

Let $\bar{c} = c \int d\mathbf{x}\, |\exp[if(\mathbf{x})] - 1|$. The series for $L(f)$ is bounded term by term by the series for $S(\bar{c})$. Therefore $L(f)$ converges. Furthermore, the series for $L_N(f)$ and $L(f)$ converge uniformly. We now show $L_N(f) \to L(f)$. First notice that there exists an n_0 such that, for $N > n_0$, $|S(\bar{c}) - \sum_{n=0}^{N} (1/n!) \bar{c}^n n^{n/2}| < \epsilon/4$. Furthermore, there exists an N_0 such that, for $N > N_0$, $(1/n!)|a_n^{(N)} - a_n| < \epsilon/2n_0$ for $n \leq n_0$.

Then, for $N > n_0$ and N_0, we have

$$\begin{aligned} \left| L_N(f) - L(f) \right| &\leq \left| L_N(f) - \sum_{n<n_0} \frac{1}{n!} a_n^{(N)} \right| + \left| L(f) - \sum_{n<n_0} \frac{1}{n!} a_n \right| + \sum_{n<n_0} \frac{1}{n!} \left| a_n^{(N)} - a_n \right| \\ &\leq \epsilon/4 + \epsilon/4 + n_0(\epsilon/2n_0) = \epsilon. \end{aligned}$$

Since ϵ is arbitrary, $L_N(f) \to L(f)$.

Remark: In order for $L(f)$ to be a generating functional for a representation of $U(f)$, it must satisfy Eqs. (2.11)–(2.14). These equations are preserved when limits are taken. Since the $L_N(f)$ satisfy them, it follows that $L(f)$ also satisfies them. Therefore, $L(f)$ defines a generating functional.

An alternative expression for $L(f)$ can be obtained in terms of the cluster functions[24] of the correlation functions. These are defined as

$$T_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \sum_G (-)^{m-n}(m-1)! \prod_{j=1}^{m} R_{G_j}(\mathbf{x}_k \in G_j),$$

where G = a partition of $(1, 2, \ldots, n)$ into subsets $(G_1, G_2, \ldots, G_m)$. $L(f)$ can be expressed in terms of T_n as follows:

$$L(f) = \exp\left(\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n!} \int d\mathbf{x}_1 \cdots \int d\mathbf{x}_n \times F(\mathbf{x}_1) \cdots F(\mathbf{x}_n) T_n(\mathbf{x}_1 \cdots \mathbf{x}_n)\right). \tag{5.7}$$

Remark: T_n is the nonrelativistic analogue of the truncated n-point functions[25] in relativistic field theory.

B. Translational invariance and the cluster decomposition property

Translational invariance and the cluster decomposition property play an important role in determining representation of the local currents in the N/V limit. A representation of $U(f)$ and $V(\varphi)$ is translational invariant if there is a set of unitary operators $Q(\mathbf{a})$, continuous in $\mathbf{a}$, such that

(i) $$Q(\mathbf{a}_1)Q(\mathbf{a}_2) = Q(\mathbf{a}_1 + \mathbf{a}_2), \tag{5.8}$$

(ii) $$Q(\mathbf{a})U(f)Q(\mathbf{a})^{-1} = U(f_{\mathbf{a}}), \quad \text{where } f_{\mathbf{a}}(\mathbf{x}) = f(\mathbf{x} - \mathbf{a}), \tag{5.9}$$

(iii) $$Q(\mathbf{a})V(\varphi)Q(\mathbf{a})^{-1} = V(\varphi_{\mathbf{a}}), \quad \text{where } \varphi_{\mathbf{a}}(\mathbf{x}) = \varphi(\mathbf{x} - \mathbf{a}) + \mathbf{a}, \tag{5.10}$$

(iv) $$Q(\mathbf{a})\Omega = \Omega. \tag{5.11}$$

These conditions are equivalent to the requirement that the generating functional is translational invariant, i.e.,

$$L(f_{\mathbf{a}}, \varphi_{\mathbf{a}}) = L(f, \varphi). \tag{5.12}$$

Also, the correlation functions are translational invariant, i.e.,

$$R_n(\mathbf{x}_1 + \mathbf{a}, \mathbf{x}_2 + \mathbf{a}, \ldots, \mathbf{x}_n + \mathbf{a}) = R_n(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n). \tag{5.13}$$

Furthermore, $R_1(\mathbf{x}) = (\Omega, \rho(\mathbf{x})\Omega) = \bar{\rho}$, the average density.

The cluster decomposition property is based on the physical idea that as particles get far apart their interaction becomes negligible. This condition can be expressed in terms of correlation functions by requiring

$$\lim_{\lambda \to \infty} \{R_{n+m}(\mathbf{x}_1 \cdots \mathbf{x}_n, \mathbf{y}_1 + \lambda\mathbf{a}, \cdots, \mathbf{y}_m + \lambda\mathbf{a}) - R_n(\mathbf{x}_1 \cdots \mathbf{x}_n) R_m(\mathbf{y}_1 + \lambda\mathbf{a}, \cdots, \mathbf{y}_m + \lambda\mathbf{a})\} = 0. \tag{5.14}$$

Combined with translational invariance we then have

$$\lim_{\lambda\to\infty} R_{n+m}(\mathbf{x}_1\cdots\mathbf{x}_n, \mathbf{y}_1+\lambda\mathbf{a}, \cdots, \mathbf{y}_m+\lambda\mathbf{a}) = R_n(\mathbf{x}_1\cdots\mathbf{x}_n)R_m(\mathbf{y}_1\cdots\mathbf{y}_m). \tag{5.15}$$

By using Eq. (5.6) this implies

$$\lim_{\lambda\to\infty} L(f+h_{\lambda\mathbf{a}}) = L(f)L(h), \quad \text{where } h_{\lambda\mathbf{a}}(\mathbf{x}) = h(\mathbf{x}-\lambda\mathbf{a}). \tag{5.16}$$

This relation can be used as a boundary condition in determining physical solutions of the functional equation (4.2) for $L(f)$. (See Appendix A for an example.)

Remark: The cluster decomposition property can also be expressed in terms of the cluster functions of the correlation functions as follows: Let

$$r(\mathbf{x}_1,\ldots,\mathbf{x}_n) = \begin{pmatrix}\text{the radius of the smallest ball containing} \\ \text{the points } \mathbf{x}_1,\ldots,\mathbf{x}_n\end{pmatrix}.$$

Then $T_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)\to 0$ as $r(\mathbf{x}_1,\ldots,\mathbf{x}_n)\to\infty$.

Translational invariance and the cluster decomposition property have important consequences in relativistic quantum field theory. We will discuss the corresponding results for the nonrelativistic local current algebra. This discussion is greatly facilitated by the application of some results of Araki.[8] The next theorem shows that the ground state is unique.

Theorem: Suppose the generating functional $L(f) = (\Omega, U(f)\Omega)$ defines a continuous unitary representation of $U(f)$ satisfying the cluster decomposition property and translational invariance. Then any state Ω' invariant under $Q(\mathbf{a})$ up to a factor [i.e., $Q(\mathbf{a})\Omega' = W(\mathbf{a})\Omega'$, where $W(\mathbf{a})$ is a complex number] is a multiple of Ω.

Proof: (See Araki,[8] Theorem 6.1)

Thus the ground state is the only translational invariant state. The generators of the translation operators are the momentum operators; i.e., $Q(\mathbf{a}) = \exp(i\mathbf{a}\cdot\mathbf{P})$, where $\mathbf{P}=$ the total momentum operator. Suppose the state $|\mathbf{p}\rangle$ is a momentum eigenstate, then $Q(\mathbf{a})|\mathbf{p}\rangle = \exp(i\mathbf{a}\cdot\mathbf{p})|\mathbf{p}\rangle$. By the above theorem $|\mathbf{p}\rangle$ is a multiple of Ω. Therefore Ω is the only momentum eigenstate. (Furthermore, $\mathbf{P}\Omega = 0$.)

The above theorem has an additional consequence.

Corollary: Suppose the generating functional $L(f) = (\Omega, U(f)\Omega)$ determines a continuous unitary representation of $U(f)$ satisfying the cluster decomposition property and translational invariance. Then the set of operators $B = \{U(f), Q(\mathbf{a})\}$ is irreducible. (I.e., any bounded operator that commutes with every operator in the set B is a multiple of the identity.)

Proof: (See Araki,[8] Sec. 6.)

In bounded regions the translation operators are similar to the operators $V(\varphi)$. In fact, if the flow $\xi_{\mathbf{a}}(\mathbf{x}) = \mathbf{x}+\mathbf{a}$, is a valid test function, then it follows from the multiplication law [Eq. (2.6)] that $V(\xi_{\mathbf{a}})V(\xi_{\mathbf{b}}) = V(\xi_{\mathbf{a}+\mathbf{b}})$, $V(\xi_{\mathbf{a}})U(f)V(\xi_{\mathbf{a}})^{-1} = U(f_{\mathbf{a}})$, and $V(\xi_{\mathbf{a}})V(\varphi)V(\xi_{\mathbf{a}})^{-1} = V(\varphi_{\mathbf{a}})$. Thus $V(\xi)$ behaves like a translation operator [except for $V(\xi)\Omega = \Omega$]. However, we have been considering only continuous representations. Therefore, it is necessary to impose a topology on the set of flows. Goldin[26] has discussed this point. He suggests a topology on a restricted set of flows φ for which $\varphi(\mathbf{x})\to\mathbf{x}$ as $|\mathbf{x}|\to\infty$. Thus $\xi_{\mathbf{a}}$ would not be in the set of test functions. In order to obtain the translation operators from $V(\varphi)$ we are led to consider a sequence of flows φ_n converging to $\xi_{\mathbf{a}}$. The next theorem gives a sufficient condition for $V(\varphi_n)\to Q(\mathbf{a})$.

Theorem 5: Let φ_n be a sequence of flows such that $f\circ\varphi_n\to f_{\mathbf{a}}$, $\forall f\in\mathcal{S}$, and $\varphi_n^{-1}\circ\varphi\circ\varphi_n\to\varphi_{\mathbf{a}}$ for all flows φ. If $(\Omega, V(\varphi_n)\Omega)\to 1$, then $V(\varphi_n)\to Q(\mathbf{a})$.

Proof: Since $(\Omega, V(\varphi_n)\Omega)\to 1$, it follows $V(\varphi_n)\Omega\to\Omega$. Let $D = \text{Span}\{U(f)V(\varphi)\Omega;\ f\in\mathcal{S} \text{ and } \varphi\in\text{flows}\}$. D is a dense set for any representation defined from a generating functional $L(f,\varphi)$. Let $\psi\in D$. We will show that $V(\varphi_n)\psi\to Q(\mathbf{a})\psi$,

$$\|V(\varphi_n)\psi - Q(\mathbf{a})\psi\|^2 = \|V(\varphi_n)\psi\|^2 + \|Q(\mathbf{a})\psi\|^2 - (V(\varphi_n)\psi, Q(\mathbf{a})\psi) - (Q(\mathbf{a})\psi, V(\varphi_n)\psi)$$

Since $V(\varphi)$ and $Q(\mathbf{a})$ are unitary, $\|V(\varphi_n)\psi\| = \|\psi\| = \|Q(\mathbf{a})\psi\|$. Since $\psi\in D$ we can write $\psi = \sum_{j=1}^m b_j U(f_j)V(\varphi_j)\Omega$. Then,

$$\begin{aligned}(V(\varphi_n)\psi, Q(\mathbf{a})\psi) &= \sum_{j=1}^m b_j^*(V(\varphi_n)U(f_j)V(\varphi_j)\Omega, Q(\mathbf{a})\psi) \\ &= \sum_{j=1}^m b_j^*(V(\varphi_n)\Omega, V(\varphi_n^{-1}\circ\varphi_j\circ\varphi_n)^{-1} \\ &\qquad\times U(f_j\circ\varphi_n)^{-1}Q(\mathbf{a})\psi).\end{aligned}$$

Since the representations we are considering are strongly continuous,

$$V(\varphi_n^{-1}\circ\varphi_j\circ\varphi_n)^{-1}U(f\circ\varphi_n)^{-1}Q(\mathbf{a})\psi \to V(\varphi_{j\mathbf{a}})^{-1} \times U(f_{j\mathbf{a}})^{-1}Q(\mathbf{a})\psi$$

and, since $V(\varphi_n)\Omega\to\Omega$, we have

$$\begin{aligned}(V(\varphi_n)\psi, Q(\mathbf{a})\psi) &\to \sum_{j=1}^n b_j^*(\Omega, V(\varphi_{j\mathbf{a}})^{-1}U(f_{j\mathbf{a}})^{-1}Q(\mathbf{a})\psi) \\ &= \sum_{j=1}^n b_j^*(U(f_{j\mathbf{a}})V(\varphi_{j\mathbf{a}})\Omega, Q(\mathbf{a})\psi) \\ &= (Q(\mathbf{a})\psi, Q(\mathbf{a})\psi) = \|\psi\|^2.\end{aligned}$$

Therefore $\|V(\varphi_n)\psi - Q(\mathbf{a})\psi\|\to 0$. Since D is dense it follows $V(\varphi_n)\to Q(\mathbf{a})$.

Remark: Theorem 5 has a physical interpretation. Since $\mathbf{J}(\mathbf{x})$ is the momentum density, we expect $\int_{-\infty}^{\infty}\mathbf{J}(\mathbf{x})\cdot\mathbf{a} = \mathbf{a}\cdot\mathbf{P}$, where $\mathbf{P}=$ the total momentum operator. Thus, $\exp(\text{it}\int\mathbf{J}(\mathbf{x})\cdot\mathbf{a}d\mathbf{x}) = \exp(\text{it}\,\mathbf{a}\cdot\mathbf{P}) = Q(t\mathbf{a})$. But $\exp(\text{it}\,\mathbf{J}(\mathbf{g})) = V(\varphi_t)$, where φ_t is the flow corresponding to the vector field g; i.e., $(d/dt)\varphi_t(\mathbf{x}) = \mathbf{g}\circ\varphi_t(\mathbf{x})$ and $\varphi_{t=0}(\mathbf{x}) = \mathbf{x}$. For $\mathbf{g}(\mathbf{x}) = \mathbf{a}$, $\varphi_t(\mathbf{x}) = \mathbf{x}+t\mathbf{a}$. Thus we expect $Q(\mathbf{a}) = V(\xi_{\mathbf{a}})$, where $\xi_{\mathbf{a}}(\mathbf{x}) = \mathbf{x}+\mathbf{a}$. However, $\int_{-\infty}^{\infty}\mathbf{J}(\mathbf{x})\cdot\mathbf{a}\,d\mathbf{x}$ may not be well defined since it is an integral over all space. Thus we must take an appropriate limit to make the integral well defined.

In Appendix B it will be shown for the representation of $U(f)$ and $V(\varphi)$ corresponding to a free Bose gas, there is a sequence φ_n satisfying the conditions of Theorem 5. Therefore the translation operators are in the closure of the algebra generated by the set $\{V(\varphi)\}$.

Then by the previous corollary it follows the set of operators $\{U(f), V(\varphi)\}$ are irreducible. (This result was proved by different means in Ref. 20.) It is not yet known whether this result is true for other representations of physical interest.

Next, we will show that different Hamiltonians give rise to unitarily inequivalent representations of the local current algebra. In order to do this, we need the following theorem.

Theorem: Suppose the generating functional $L(f) = (\Omega, U(f)\Omega)$ determines a continuous unitary representation of $U(f)$ satisfying the cluster decomposition property and translational invariance. If there is a set of unitary operators $Q'(\mathbf{a})$ and a cyclic vector Ω' [i.e., Span$\{U(f)\Omega'; f\in \mathcal{S}\}$ is dense] satisfying Eq. (5.8), (5.9), and (5.11), then there exists a unitary operator S such that $SU(f)S^{-1} = U(f)$, $SQ(\mathbf{a})S^{-1} = Q'(\mathbf{a})$, and $S\Omega = \Omega'$.

Proof: (See Araki,[8] Theorem 6.2).

Corollary 1: Suppose the generating functionals $L_1(f) = (\Omega_1, U_1(f)\Omega_1)$ and $L_2(f) = (\Omega_2, U_2(f)\Omega_2)$, each satisfying translational invariance, define two continuous unitary representations of $U(f)$. Furthermore, suppose $L_2(f)$ satisfies the cluster decomposition property. Then the representations are unitarily equivalent iff $L_1(f) = L_2(f)$.

Proof: Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be the Hilbert spaces and Ω_1 and Ω_2 the cyclic vectors for the two representations. Suppose the representations are unitarily equivalent. Then there exists a unitary operator S_1 such that $S_1: \mathcal{H}_1 \to \mathcal{H}_2$ and $S_1U_1(f)S_1^{-1} = U_2(f)$. Let $\Omega_2' = S_1\Omega_1$ and $Q_2'(\mathbf{a}) = S_1Q_1(\mathbf{a})S_1^{-1}$. It is easily shown that Ω_2' is cyclic in $\mathcal{H}_2$ and Eqs. (5.8), (5.9), and (5.11) are satisfied for Ω_2' and $Q_2'(\mathbf{a})$. By the above theorem there exists a unitary operator S_2 such that; $S_2:\mathcal{H}_2 \to \mathcal{H}_2$, $S_2U_2(f)S_2^{-1} = U_2(f)$, and $S_2\Omega_2 = \Omega_2'$. Let $S = S_2^{-1}S_1$: $\mathcal{H}_1 \to \mathcal{H}_2$. Then $S\Omega_1 = \Omega_2$ and $SU_1(f)S^{-1} = U_2(f)$. Therefore

$$\begin{aligned} L_1(f) &= (\Omega_1, U_1(f)\Omega_1)_1 \\ &= (S\Omega_1, SU_1(f)\Omega_1)_2 \\ &= (\Omega_2, U_2(f)\Omega_2)_2 = L_2(f). \end{aligned}$$

Conversely, if $L_1(f) = L_2(f)$, the representations are clearly unitarily equivalent.

Remark: The last two theorems have used only $L(f)$. They are important for representations of $U(f)$ and $V(\varphi)$ in which Span$\{U(f)\Omega\}$ is dense. Furthermore, they can be generalized using $L(f, \varphi)$ for representations in which Span$\{U(f)V(\varphi)\Omega\}$ is dense.

Now suppose there are two representations of $U(f)$ and $V(\varphi)$ with Hamiltonians H_1 and H_2 of the form $H = \frac{1}{8}\int d\mathbf{x}\tilde{\mathbf{K}}(\mathbf{x})^\dagger [1/\rho(\mathbf{x})]\tilde{\mathbf{K}}(\mathbf{x})$ with $\tilde{\mathbf{K}}^{(1)}(\mathbf{x}) = \mathbf{K}(\mathbf{x}) - \mathbf{A}^{(1)}(\mathbf{x}, \rho)$ and $\tilde{\mathbf{K}}^{(2)}(\mathbf{x}) = \mathbf{K}(\mathbf{x}) - \mathbf{A}^{(2)}(\mathbf{x}, \rho)$. If the representations are unitarily equivalent, then, by Corollary 1, $L_1(f) = L_2(f)$. Therefore we may take $\mathcal{H}_1 = \mathcal{H}_2$. Consider the following identity:

$$(\Omega, \exp[i\rho(f)]\,[\rho(-\nabla\cdot \mathbf{g}) - i\rho(\nabla f\cdot \mathbf{g})]\Omega) = (\Omega, \exp[i\rho(f)]K(\mathbf{g})\Omega).$$

Since $K(\mathbf{g})\Omega = A^{(1)}(\mathbf{g}, \rho)\Omega = A^{(2)}(\mathbf{g}, \rho)\Omega$, we have

$$(\Omega, \exp[i\rho(f)]A^{(1)}(\mathbf{g}, \rho)\Omega) = (\Omega, \exp[i\rho(f)]A^{(2)}(\mathbf{g}, \rho)\Omega); \quad \text{for all } f\in \mathcal{S}.$$

Therefore, $A^{(1)}(\mathbf{g}, \rho)\Omega = A^{(2)}(\mathbf{g}, \rho)\Omega$. Since $[A(\mathbf{g}, \rho), \exp[i\rho(f)]] = 0$ and Span$\{\exp[i\rho(f)]\Omega; f\in \mathcal{S}\}$ is dense, it follows that $A^{(1)}(\mathbf{g}, \rho) = A^{(2)}(\mathbf{g}, \rho)$. We have proved the following theorem.

Theorem 6: Suppose there are two continuous unitary representations of $U(f)$ and $V(\varphi)$ (denoted by $i = 1, 2$) with Hamiltonians

$$H_i = \tfrac{1}{8}\int d\mathbf{x}\tilde{\mathbf{K}}^{(i)}(\mathbf{x})^\dagger [1/\rho(\mathbf{x})]\tilde{\mathbf{K}}^{(i)}(\mathbf{x}), \quad \text{where } \tilde{\mathbf{K}}^{(i)}(\mathbf{x}) = \mathbf{K}(\mathbf{x}) - \mathbf{A}^{(i)}(\mathbf{x}, \rho)$$

and satisfying the cluster decomposition property, translational invariance, and time reversal invariance. If $\mathbf{A}^{(1)}(\mathbf{x}, \rho) \neq \mathbf{A}^{(2)}(\mathbf{x}, \rho)$, then the representations are unitarily inequivalent.

Remarks: (1) Roughly speaking, Theorem 6 states that different Hamiltonians correspond to inequivalent representations. Two important questions remain unanswered at this time. First, given a system of particles (boson or fermion) with an interaction potential $V(\mathbf{x})$, is $\mathbf{A}(\mathbf{x}, \rho)$ uniquely determined? Second, does a Hamiltonian with a given $\mathbf{A}(\mathbf{x}, \rho)$ uniquely determine the representation? The second question is equivalent to asking whether the functional equation (4.2) for $L(f)$ has a unique solution. This is known to be the case for a free Bose gas[20] [$\mathbf{A}(\mathbf{x}, \rho) = 0$], but uniqueness has not been established for other cases.

(2) Since $\mathbf{A}(\mathbf{x}, \rho)$ considered as a function of ρ may be an unbounded operator, its definition is representation dependent. For some representations it may not even be defined. In some representations two operators $\mathbf{A}^{(1)}(\mathbf{x}, \rho)$ and $\mathbf{A}^{(2)}(\mathbf{x}, \rho)$ may be equal while in others they may be unequal.

(3) Consider the N/V limit of interacting physical systems characterized by a coupling constant λ and for which the assumptions in Theorem 6 are valid. If the systems are described by unitarily equivalent representations, then by Theorem 6 and Eq. (3.12) the Hamiltonians H_λ are identical as Hermitian forms. Therefore, the Hamiltonian operators would be different self-adjoint extensions of the same Hermitian form. Furthermore, the ground states are the same since there is a unique translational invariant state. On the other hand, if the systems are described by unitarily inequivalent representations, then solving the systems by perturbation theory is more difficult since it is no longer possible to express the ground states as a convergent series in λ. This point will be discussed in the next paper in connection with a specific example.

C. Restriction of the N/V limit representation to a finite volume

We can gain further insight about $L(f)$ in the N/V limit by restricting the test functions to have support in a bounded set ν.

If supp $f \subset \nu$, then

$$L(f)=\sum_{n=0}^{\infty}(1/n!)\int_{\nu}d\mathbf{x}_1\cdots\int_{\nu}d\mathbf{x}_nF(\mathbf{x}_1)\cdots F(\mathbf{x}_n)R_n(\mathbf{x}_1\cdots\mathbf{x}_n).$$

Since

$$F(\mathbf{x}_1)\cdots F(\mathbf{x}_n)=\sum_{\text{perm}}\sum_{j=0}^{n}\frac{(-)^{n-j}}{j!(n-j)!}\prod_{k=1}^{j}\exp[if(\mathbf{x}_{r_k})]$$

and R_n is a symmetric function, we have

$$L(f)=\sum_{n=0}^{\infty}\frac{1}{n!}\int_{\nu}d\mathbf{x}_1\cdots\int_{\nu}d\mathbf{x}_n\exp[if(\mathbf{x}_1)]\cdots\exp[if(\mathbf{x}_n)] \times P_n(\nu;\mathbf{x}_1\cdots\mathbf{x}_n), \tag{5.17}$$

where

$$P_n(\nu;\mathbf{x}_1\cdots\mathbf{x}_n)=\sum_{j=0}^{\infty}\frac{(-)^j}{j!}\int_{\nu}d\mathbf{x}_{n+1}\cdots\int_{\nu}d\mathbf{x}_{n+j}R_{n+j}(\mathbf{x}_1\cdots\mathbf{x}_{n+j}). \tag{5.18}$$

P_n has the physical interpretation,

$$(1/n!)P_n(\nu;\mathbf{x}_1\cdots\mathbf{x}_n)=\begin{pmatrix}\text{the probability for finding } n\\ \text{particles at points}\\ \mathbf{x}_1,\cdots,\mathbf{x}_n \text{ and the remaining}\\ \text{particles outside } \nu\end{pmatrix}.$$

To prove this, we consider N particles in a box of volume V, in this case

$$\begin{aligned}
\frac{1}{n!}P_n^{(N)}(\nu;\mathbf{x}_1\cdots\mathbf{x}_n)
&=\frac{N!}{n!(N-n)!}\int_{V-\nu}d\mathbf{x}_{n+1}\cdots\int_{V-\nu}d\mathbf{x}_N\Omega^*\Omega(\mathbf{x}_1\cdots\mathbf{x}_N)\\
&=\frac{N!}{n!(N-n)!}\sum_{j=0}^{N-n}(-)^j\frac{(N-n)!}{j!(N-j-n)!}\\
&\quad\times\int_{\nu}d\mathbf{x}_{n+1}\cdots\int_{\nu}d\mathbf{x}_{n+j}\int_{V}d\mathbf{x}_{n+j+1}\cdots\int_{V}d\mathbf{x}_N\Omega^*\Omega\\
&=\frac{N!}{n!(N-n)!}\sum_{j=0}^{N-n}(-)^j\frac{(N-n)!}{j!(N-j-n)!}\\
&\quad\times\int_{\nu}d\mathbf{x}_{n+1}\cdots\int_{\nu}d\mathbf{x}_{n+j}\frac{(N-j-n)!}{N!}R_{j+n}\\
&=\frac{1}{n!}\sum_{j=0}^{N-n}\frac{(-)^j}{j!}\int_{\nu}d\mathbf{x}_{n+1}\cdots\int_{\nu}d\mathbf{x}_{n+j}R_{n+j}(\mathbf{x}_1\cdots\mathbf{x}_{n+j}).
\end{aligned}$$

As $N\to\infty$, we obtain the expression in Eq. (5.18).

Remarks: (1) Formally Eq. (5.18) can be inverted. The R's are given in terms of the P's by the equation: For $\mathbf{x}_1,\ldots,\mathbf{x}_n\in\nu$

$$R_n(\mathbf{x}_1\cdots\mathbf{x}_n)=\sum_{j=0}^{\infty}\frac{1}{j!}\int_{\nu}d\mathbf{x}_{n+1}\cdots\int_{\nu}d\mathbf{x}_{n+j}P_{n+j}(\nu;\mathbf{x}_1\cdots\mathbf{x}_{n+j}) \tag{5.19}$$

If the sum in Eq. (5.19) converges and is consistent (i.e., the same value of R_n is obtained for points $\mathbf{x}_1,\ldots,\mathbf{x}_n$ in overlapping volumes), then the R's can be determined from the local probability distributions. Since the R's determine $L(f)$, this implies that $L(f)$ can be determined by its local behavior.

(2) If the volume ν is not bounded, each term in the expansion for P_n [Eq. (5.18)] will be infinite.

(3) As a result of the probability interpretation for P_n,

(i) $P_n(\nu;\mathbf{x}_1\cdots\mathbf{x}_n)\geq 0$

and

(ii) $\sum_{n=0}^{\infty}\frac{1}{n!}\int_{\nu}d\mathbf{x}_1\cdots\int_{\nu}d\mathbf{x}_nP_n(\nu;\mathbf{x}_1\cdots\mathbf{x}_n)=1.$

Property (ii) also follows from $L(0)=1$.

(4) If $R_n\leq c^n n^{n/2}\,\forall n$, then the lemma to Theorem 4 can be extended to show that P_n exists (i.e., the series for P_n converges). However, this is not sufficient to imply $P_n\geq 0$.

From Eq. (5.17) we see that, for supp $f\subset\nu$, $L(f)$ is the sum of terms which have the form of N-particle generating functions [Eq. (5.3)] with ground state given by $\Omega_N^*\Omega_N(\mathbf{x}_1\cdots\mathbf{x}_N)=(1/N!)P_N(\nu;\mathbf{x}_1\cdots\mathbf{x}_N)$. As a result the N/V limit representation restricted to finite volumes (this is a representation of the subalgebra formed by restricting the test functions) can be represented in the Hilbert space formed by the direct sum of N-particle spaces (Fock space). However, the ground state for this restriction would not have a definite number of particles. Thus, locally the N/V limit can be considered as "Fock space." This is the "particlelike" nature of the N/V limit.

For a free Bose gas $P_n(\nu;\mathbf{x}_1\cdots\mathbf{x}_n)$ can be calculated exactly. It has been shown for this case that[7]

$$\begin{aligned}L(f)&=\exp\left(\bar{\rho}\int d\mathbf{x}(\exp[if(\mathbf{x})]-1)\right)\\&=\sum_{n=0}^{\infty}\frac{1}{n!}\int d\mathbf{x}_1\cdots\int d\mathbf{x}_nF(\mathbf{x}_1)\cdots F(\mathbf{x}_n)\bar{\rho}^n.\end{aligned}$$

Therefore, $R_n(\mathbf{x}_1\cdots\mathbf{x}_n)=\bar{\rho}^n$. As a result,

$$\begin{aligned}P_n(\nu;\mathbf{x}_1\cdots\mathbf{x}_n)&=\frac{1}{n!}\sum_{j=0}^{\infty}\frac{(-)^j}{j!}\int_{\nu}d\mathbf{x}_{n+1}\cdots\int_{\nu}d\mathbf{x}_{n+j}\bar{\rho}^{n+j}\\&=\frac{1}{n!}\sum_{j=0}^{\infty}\frac{(-)^j}{j!}\nu^j\bar{\rho}^{n+j}=\frac{\bar{\rho}^n}{n!}\exp(-\bar{\rho}\nu).\end{aligned}$$

This is a Poisson distribution with mean equal to $\bar{\rho}\,\nu$. This is to be expected since we have taken the limit of a large number of noninteracting particles ($N\to\infty$) with the probability of finding a given particle in a given unit volume (prob $=1/V$) approaching zero such that the product ($N\cdot$ prob $=N/V=\bar{\rho}$) is a constant.

Remark: The Hilbert space, $\mathcal{H}=L_\mu^2(\mathcal{S}')$, can be used to represent the N/V limit. The measures for the N-particle representations and Eq. (5.17) suggest the measure in the N/V limit is concentrated on functionals consisting of a countably infinite number of delta functions; $F=\sum_{j=1}^{\infty}\delta(\mathbf{x}-\mathbf{x}_j)$ such that if $n_F(\nu)$ is the number of delta functions with support in volume ν then $\lim_{\nu\to\infty}n_F(\nu)/\nu=\bar{\rho}$. The functionals can be characterized by the sequence of points $\{\mathbf{x}_1,\mathbf{x}_2\cdots\}$ which can be interpreted as the positions of the particles. The measure μ can be considered as a measure on these sequences. In this context there are similarities with recent work of Lenard[27] in which he discussed the state in classical statistical mechanics in terms of correlation functions. The present formalism becomes distinctly quantum mechanical in nature only when the J's are considered.

Also, representations corresponding to different average densities $\bar{\rho}_1$ and $\bar{\rho}_2$ will have measures μ_1 and μ_2 with different sets of measure zero (in $\mathcal{S}'$). As a result, representations corresponding to different average densities are unitarily inequivalent.

The same methods may be used to obtain expressions for $L(f, \mathbf{g})$ similar to those for $L(f)$. The results are

$$L(f, \mathbf{g}) = \sum_{n=0}^{\infty} \frac{1}{n!} \int d\mathbf{x}_1 \int d\mathbf{y}_1 \cdots \int d\mathbf{x}_n \int d\mathbf{y}_n \, \delta(\mathbf{x}_1 - \mathbf{y}_1) \cdots \delta(\mathbf{x}_n - \mathbf{y}_n) \tag{5.20}$$
$$\times \prod_{k=1}^{n} [\exp(if(\mathbf{x}_k)) \exp(ij(\mathbf{x}_k, \mathbf{g})) - 1] R_n(\mathbf{y}_1 \cdots \mathbf{y}_n; \mathbf{x}_1 \cdots \mathbf{x}_n),$$

where $j(\mathbf{x}, \mathbf{g}) = -\frac{1}{2} i [2\mathbf{g}(\mathbf{x}) \cdot \nabla + (\nabla \cdot \mathbf{g})(\mathbf{x})]$. In the N-particle representations $R_n(\;;\;)$ is given by

$$R_n(\mathbf{y}_1 \cdots \mathbf{y}_n; \mathbf{x}_1 \cdots \mathbf{x}_n) = \frac{N!}{(N-n)!} \int d\mathbf{x}_{n+1} \cdots \int d\mathbf{x}_N \times \Omega^*(\mathbf{y}_1 \cdots \mathbf{y}_n, \mathbf{x}_{n+1}, \ldots, \mathbf{x}_N) \times \Omega(\mathbf{x}_1 \cdots \mathbf{x}_N).$$

In terms of the canonical field operators,

$$R_n(\mathbf{y}_1 \cdots \mathbf{y}_n; \mathbf{x}_1 \cdots \mathbf{x}_n) = (\Omega, \psi^\dagger(\mathbf{y}_n) \cdots \psi^\dagger(\mathbf{y}_1) \psi(\mathbf{x}_1) \cdots \psi(\mathbf{x}_n) \Omega).$$

Clearly $R_n(\mathbf{x}_1 \cdots \mathbf{x}_n; \mathbf{x}_1 \cdots \mathbf{x}_n) = R_n(\mathbf{x}_1 \cdots \mathbf{x}_n)$. Also, as a consequence of Schwartz's inequality, $|R_n(\mathbf{y}_1 \cdots \mathbf{y}_n; \mathbf{x}_1 \cdots \mathbf{x}_n)|^2 \le R_n(\mathbf{x}_1 \cdots \mathbf{x}_n) R_n(\mathbf{y}_1 \cdots \mathbf{y}_n)$.

An alternative expression for $L(f, \mathbf{g})$ can be obtained in terms of the cluster functions defined as

$$T_n(\mathbf{y}_1 \cdots \mathbf{y}_n; \mathbf{x}_1 \cdots \mathbf{x}_n) = \sum_G (-)^{m-n} (m-1)! \prod_{j=1}^{m} R_{G_j}(\mathbf{y}_k \in G_j; \mathbf{x}_k \in G_j),$$

where G = a partition of $(1, 2, \ldots, n)$ into subsets $(G_1, \ldots, G_m)$. $L(f, \mathbf{g})$ can now be expressed as

$$L(f, \mathbf{g}) = \exp \sum_{n=1}^{\infty} \frac{(-)^{n-1}}{n!} \int d\mathbf{x}_1 \int d\mathbf{y}_1 \cdots \int d\mathbf{x}_n \int d\mathbf{y}_n \, \delta(\mathbf{x}_1 - \mathbf{y}_1) \cdots \delta(\mathbf{x}_n - \mathbf{y}_n)$$
$$\times \prod_{k=1}^{n} [\exp(if(\mathbf{x}_k)) \exp(ij(\mathbf{x}_k, \mathbf{g})) - 1] T_n(\mathbf{y}_1 \cdots \mathbf{y}_n; \mathbf{x}_1 \cdots \mathbf{x}_n). \tag{5.21}$$

Finally, if supp $f \subset \nu$ and supp $\mathbf{g} \subset \nu$

$$L(f, \mathbf{g}) = \sum_{n=0}^{\infty} \frac{1}{n!} \int_\nu d\mathbf{x}_1 \cdots \int_\nu d\mathbf{x}_n \int d\mathbf{y}_1 \cdots \int d\mathbf{y}_n \, \delta(\mathbf{x}_1 - \mathbf{y}_1) \cdots \delta(\mathbf{x}_n - \mathbf{y}_n) \tag{5.22}$$
$$\times \prod_{k=1}^{n} [\exp(if(\mathbf{x}_k)) \exp(ij(\mathbf{x}_k, \mathbf{g}))] P_n(\nu; \mathbf{y}_1 \cdots \mathbf{y}_n; \mathbf{x}_1 \cdots \mathbf{x}_n),$$

where

$$P_n(\nu; \mathbf{y}_1 \cdots \mathbf{y}_n; \mathbf{x}_1 \cdots \mathbf{x}_n) = \sum_{k=0}^{\infty} \frac{(-)^k}{k!} \int_\nu d\mathbf{x}_{n+1} \cdots \int_\nu d\mathbf{x}_{n+k} \times R_{n+k}(\mathbf{y}_1 \cdots \mathbf{y}_n, \mathbf{x}_{n+1} \cdots \mathbf{x}_{n+k}; \mathbf{x}_1 \cdots \mathbf{x}_{n+k}). \tag{5.23}$$

Thus the generating functional for a representation of $U(f)$ and $V(\varphi)$ in the N/V limit restricted to a finite volume is the sum of terms similar to $L(f, \mathbf{g})$ for an N-particle representation [Eq. (5.4)]. [If $P_n(\nu; \mathbf{y}_1 \cdots \mathbf{y}_n; \mathbf{x}_1 \cdots \mathbf{x}_n) = W_n(\mathbf{y}_1 \cdots \mathbf{y}_n) W_n(\mathbf{x}_1 \cdots \mathbf{x}_n)$, then the restriction is the direct sum of N-particle representations.]

6. SUMMARY

We have shown that the Hamiltonian, considered as a densely defined Hermitian form, can be written

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

where

$$\tilde{\mathbf{K}}(\mathbf{x}) = [\nabla \rho(\mathbf{x}) + 2i\mathbf{J}(\mathbf{x})] - \mathbf{A}(\mathbf{x}, \rho).$$

The generating functional in the N/V limit can be expressed as

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

where $F(\mathbf{x}) = \exp[if(\mathbf{x})] - 1$ and R_n = the nth correlation function. $L(f)$ satisfies the functional differential equation

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

Furthermore, under the assumption of translational invariance and the cluster decomposition property, inequivalent representations are needed for different Hamiltonians.

There remains two problems in determining representations of physical interest:

(1) Given a potential $V(\mathbf{x})$, determine $\mathbf{A}(\mathbf{x}, \rho)$.

(2) Given $\mathbf{A}(\mathbf{x}, \rho)$ solve the functional equation (subject to the appropriate boundary conditions) for $L(f)$.

Undoubtably these tasks can be accomplished in general only by using approximation methods. Once a representation for a given system has been determined, its dynamics can be studied. By extending this approach to study the thermodynamics of a system is also of interest.

ACKNOWLEDGMENTS

The author would like to thank Professor D. H. Sharp for his advice, continual encouragement, and careful proofreading of this article. The author would also like to thank Arnold Dicke for many helpful suggestions.

APPENDIX A

In this section we will show how the cluster decom-

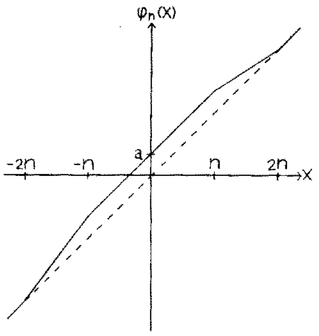


FIG. 1. The flow $\varphi_n(x)$ vs x, in one dimension.

position property can be used as a boundary condition for the functional differential equation (4.2) to uniquely determine the generating functional for a free Bose gas in the N/V limit. (In Ref. 20 other boundary conditions were used for this purpose.) We will assume we already know that the generating functional for a free Bose gas satisfies the equation

$$(\nabla - i\nabla f)\frac{1}{i}\frac{\delta}{\delta f(\mathbf{x})}L(f)=0. \tag{A1}$$

The first method for solving this equation is based on the use of integrating factors. Equation (A1) can be rewritten as

$$\nabla\left(\exp[-if(\mathbf{x})]\frac{1}{i}\frac{\delta}{\delta f(\mathbf{x})}L(f)\right)=0. \tag{A2}$$

Integrating between point $\mathbf{x}$ and $\infty(\int_{\mathbf{x}}^{\infty} d\mathbf{r}\cdot)$, we obtain

$$\exp[-if(\mathbf{x})]\frac{1}{i}\frac{\delta}{\delta f(\mathbf{x})}L(f)\Bigg|_{\mathbf{x}=\infty} - \exp[-if(\mathbf{x})]\frac{1}{i}\frac{\delta}{\delta f(\mathbf{x})}L(f)=0. \tag{A3}$$

By using the cluster decomposition property,

$$\frac{1}{i}\frac{\delta}{\delta f(\mathbf{x})}L(f)=(\Omega,\rho(\mathbf{x})\exp[if(\mathbf{x})]\Omega) \to (\Omega,\rho(\mathbf{x})\Omega)(\Omega,\exp[i\rho(f)]\Omega) \text{ as } |\mathbf{x}|\to\infty,$$

translational invariance, $(\Omega,\rho(\mathbf{x})\Omega)=\bar{\rho}=$ the average density, and the fact $f(\mathbf{x})\to 0$ as $|\mathbf{x}|\to\infty$, Eq. (A3) becomes

$$\bar{\rho}L(f)-\exp[-if(\mathbf{x})]\frac{1}{i}\frac{\delta}{\delta f(\mathbf{x})}L(f)=0. \tag{A4}$$

This can be written as

$$\frac{1}{i}\frac{\delta}{\delta f(\mathbf{x})}\{\exp[-\bar{\rho}\textstyle\int(\exp[if(\mathbf{x})]-1)\,d\mathbf{x}]L(f)\}=0. \tag{A5}$$

Therefore, $\exp[-\bar{\rho}\int(e^{if(\mathbf{x})}-1)d\mathbf{x}]L(f)=\text{const.}$ (A6)

The constant can be determined from the requirement $L(0)=1$. The result is

$$L(f)=\exp[\bar{\rho}\textstyle\int(e^{if(\mathbf{x})}-1)d\mathbf{x}]. \tag{A7}$$

An alternative method for solving Eq. (A1) uses the cluster decomposition property for the correlation functions. We have shown in the N/V limit that $L(f)$ has the form [Eq. (5.6)]

$$L(f)=\sum_{n=0}^{\infty}\frac{1}{n!}\int d\mathbf{x}_1\cdots\int d\mathbf{x}_n(\exp[if(\mathbf{x}_1)]-1)\cdots(\exp[if(\mathbf{x}_n)]-1)\times R_n(\mathbf{x}_1\cdots\mathbf{x}_n). \tag{A8}$$

Substituting Eq. (A8) into Eq. (A1), we obtain

$$\sum_{n=0}^{\infty}\frac{\exp[if(\mathbf{x}_1)]}{(n-1)!}\int d\mathbf{x}_2\cdots\int d\mathbf{x}_n(\exp[if(\mathbf{x}_2)]-1)\cdots(\exp[if(\mathbf{x}_n)]-1)\times\nabla_1 R_n(\mathbf{x}_1\cdots\mathbf{x}_n)=0. \tag{A9}$$

Since Eq. (A9) is true for all f, each term separately must be zero,

$$\nabla_1 R_n(\mathbf{x}_1\cdots\mathbf{x}_n)=0\ \forall\ n. \tag{A10}$$

Furthermore, R_n is a symmetric function. Therefore, $R_n=$ const. The cluster decomposition property can be used to relate the different constants as follows:

$$R_1(\mathbf{x})=(\Omega,\rho(\mathbf{x})\Omega)=\bar{\rho},$$

$$\lim_{\mathbf{a}\to\infty}R_n(\mathbf{x}_1\cdots\mathbf{x}_{n-1},\mathbf{x}_n+\mathbf{a})=R_1(\mathbf{x}_n)R_{n-1}(\mathbf{x}_1\cdots\mathbf{x}_{n-1}) =\bar{\rho}R_{n-1}(\mathbf{x}_1\cdots\mathbf{x}_{n-1}).$$

By induction we have $R_n=\bar{\rho}^n$. Therefore,

$$L(f)=\sum_{n=0}^{\infty}\frac{1}{n!}\int d\mathbf{x}_1\cdots\int d\mathbf{x}_n(\exp[if(\mathbf{x}_1)]-1)\cdots(\exp[if(\mathbf{x}_n)]-1)\bar{\rho}^n =\exp[\bar{\rho}\textstyle\int d\mathbf{x}(e^{if(\mathbf{x})}-1)].$$

APPENDIX B

In Ref. 20 it was shown the generating functional for the representation of $U(f)$ and $V(\varphi)$ corresponding to the free Bose gas is given by

$$L(f,\varphi)=(\Omega,U(f)\,V(\varphi)\Omega) =\exp\left\{\bar{\rho}\int d\mathbf{x}\left[\exp[if(\mathbf{x})]\left(\det\frac{\partial\varphi(\mathbf{x})_r}{\partial x_s}\right)^{1/2}-1\right]\right\}.$$

In this section we will show there is a sequence of test functions φ_n such that $V(\varphi_n)\to Q(\mathbf{a})$, the translation operator. First, it is necessary to define which flows are to be used as test tunctions. Goldin[26] has suggested a topology on the flows in analogy to the topology on Schwartz's space. His topology is defined by a countable number of metrics,

$$\langle\langle\varphi,\psi\rangle\rangle_n=\max_{0\le|m|\le n}\ \sup_{\mathbf{x}}\left|(1+|\mathbf{x}|^2)^n[\varphi^{(m)}(\mathbf{x})-\psi^{(m)}(\mathbf{x})]\right|.$$

Since we want the test functions to include the identity flow $\varphi_0(\mathbf{x})=\mathbf{x}$ and to have an inverse, we will take the test functions to be the set of flows φ such that $\langle\langle\varphi,\varphi_0\rangle\rangle_n<\infty$ and $\langle\langle\varphi^{-1},\varphi_0\rangle\rangle_n<\infty$ for all n.

By Theorem 5, if $\varphi_n(\mathbf{x})\to\mathbf{x}+\mathbf{a}$ and $(\Omega,V(\varphi_n)\Omega)\to 1$, then $V(\varphi_n)\to Q(\mathbf{a})$. We will first consider the one-dimensional case. Let

$$\varphi_n(x)=\begin{cases}x, & 2n<x,\\ x+a(2n-x)/n, & n<x<2n,\\ x+a, & |x|<n,\\ x+a(2n+x)/n, & -2n<x<-n,\\ x, & x<-2n.\end{cases}$$

(See Fig. 1.)

φ_n would be a test function except that its derivative is discontinuous at four points. By changing φ_n in a small region about each discontinuity it can be made into a smooth function (and hence a test function) without changing the subsequent arguments.

Clearly $\varphi_n(x) \to x + a$ as $n \to \infty$.

In order to verify $(\Omega, V(\varphi_n)\Omega) \to 1$, we must show

$$\int_{-\infty}^{\infty} dx \left[\left(\frac{d\varphi_n}{dx}\right)^{1/2} - 1\right] \to 0 \quad \text{as } n \to \infty.$$

Let

$$\begin{aligned} I_n &= \int_{-\infty}^{\infty} \left[\left(\frac{d\varphi_n}{dx}\right)^{1/2} - 1\right] dx \\ &= \int_{-2n}^{-n} \left[\left(1 - \frac{a}{n}\right)^{1/2} - 1\right] dx + \int_{n}^{2n} \left[\left(1 + \frac{a}{n}\right)^{1/2} - 1\right] dx. \end{aligned}$$

For n large $(1 \pm a/n)^{1/2} = 1 \pm \frac{1}{2} a/n + O(1/n^2)$,

$$\begin{aligned} I_n &= \int_{-2n}^{-n} \left[-\frac{1}{2}\frac{a}{n} + O\left(\frac{1}{n^2}\right)\right] dx + \int_{n}^{2n} \left[\frac{1}{2}\frac{a}{n} + O\left(\frac{1}{n^2}\right)\right] dx \\ &= -\tfrac{1}{2} a + n\, O\left(\frac{1}{n^2}\right) + \tfrac{1}{2} a + n\, O\left(\frac{1}{n^2}\right) \\ &= O\left(\frac{1}{n}\right) \to 0 \quad \text{as } n \to \infty. \end{aligned}$$

Therefore $V(\varphi_n) \to Q(a)$ by Theorem 5.

In two dimensions consider a translation in the x direction by a distance a. Let

$$\varphi_n(x, y)_y = y,$$

$$\varphi_n(x, y)_x = x + a\, \alpha_n(x)\, \beta_n(y),$$

where

$$\alpha_n(x) = \begin{cases} (2n - x)/n, & n < x < 2n, \\ 1, & -n < x < n, \\ (2n + x)/n, & -2n < x < -n, \\ 0, & 2n < |x|, \end{cases}$$

and

$$\beta_n(y) = \begin{cases} [(n + \Delta) - y]/\Delta, & n < y < n + \Delta, \\ 1, & -n < y < n, \\ [(n + \Delta) + y]/\Delta, & -(n + \Delta) < y < -n, \\ 0, & n + \Delta < |y|, \end{cases}$$

$\Delta =$ an arbitrary positive constant.

To prove $(\Omega, V(\varphi_n)\Omega) \to 1$, it is necessary to show

$$\int dx \int dy \left[\left(\det \frac{\partial \varphi_n(x, y)_r}{\partial x_s}\right)^{1/2} - 1\right] \to 0.$$

This can be verified by a calculation similar to the one-dimensional case. In fact a similar argument works for any number of space dimensions.

Therefore, for the free Bose gas representation there is a sequence of test functions φ_n such that $Q(\mathbf{a}) = \lim V(\varphi_n)$.

*Part of the work reported here is included in a thesis to be submitted to the University of Pennsylvania in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

†Work supported in part by U.S. Atomic Energy Commission under Contract No. AT(30-1)-3071.

[1]R.F. Dashen and D.H. Sharp, Phys. Rev. **165**, 1857 (1968).
[2]D.H. Sharp, Phys. Rev. **165**, 1867 (1968).
[3]C.G. Callen, R.F. Dashen, and D.H. Sharp, Phys. Rev. **165**, 1883 (1968).
[4]H. Sugawara, Phys. Rev. **170**, 1659 (1968).
[5]A.A. Dicke and G.A. Goldin, Phys. Rev. D **5**, 845 (1972).
[6]G.A. Goldin, J. Math. Phys. **12**, 462 (1971).
[7]G.A. Goldin and D.H. Sharp, "Lie Algebras of Local Currents and their Representations," in the 1969 *Battelle Recontres: Group Representations*, edited by V. Bargmann (Springer-Verlag, Berlin, 1970), p. 300.
[8]H. Araki, J. Math. Phys. **1**, 492 (1960).
[9]D.H. Sharp, "What We Have Learned about Representing Local Nonrelativistic Current Algebras," in *Local Currents and Their Applications*, edited by D.H. Sharp and A.S. Wightman (North-Holland, Amsterdam, 1974), p. 85.
[10]I. Gel'fand and N. Vilenkin, *Generalized Functions* (Academic, New York, 1964), Vol. 4.
[11]J. Grodnik and D.H. Sharp, Phys. Rev. D **1**, 1531 (1970).
[12]A.A. Dicke and G.A. Goldin (to be published).
[13]J. Grodnik and D.H. Sharp, Phys. Rev. D **1**, 1546 (1970).
[14]A.A. Dicke (to be published).
[15]H. Araki and E.J. Woods, J. Math. Phys. **4**, 637 (1963).
[16]H. Araki and W. Wyss, Helv. Phys. Acta. **37**, 136 (1964).
[17]R. Riesz and B. Sz. Nagy, *Functional Analysis* (Ungar, New York, 1955), p. 329.
[18]Y. Aref'eva, Theoretical and Mathematical Physics (Consultants Bureau, N.Y.) transl. of Theor. Math. Fiz. **10**, 233 (1972).
[19]F. Coester and R. Haag, Phys. Rev. **117**, 1137 (1960).
[20]G.A. Goldin, J. Grodnik, R. Powers, and D.H. Sharp, J. Math. Phys. **15**, 88 (1974).
[21]G.A. Goldin and D.H. Sharp, "Functional Differential Equations Determining Representations of Local Current Algebras" in *Magic without Magic: John Archibald Wheeler*, edited by J. Klauder (Freeman, San Francisco, 1972), p. 171.
[22]D. Ruelle, *Statistical Mechanics* (Benjamin, New York, 1969), p. 100.
[23]A. Girard, J. Math. Phys. **14**, 353 (1973).
[24]M.L. Metha, *Random Matrices* (Academic, New York, 1967), Chap. 5 and Appendix 6.
[25]R. Haag, Phys. Rev. **112**, 669 (1958).
[26]G. Goldin, Ph.D. thesis, Princeton University, 1968.
[27]A. Lenard, Commun. Math. Phys. **30**, 35 (1973).

Addendum: The unitarity equation for scattering in the absence of spherical symmetry [J. Math. Phys. 15, 745 (1974)]

Michael Tortorella

Department of Mathematics, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201
(Received 20 February 1974)

We have recently discovered an improved version of theorem 3.3 of our paper. We exploit the fact that, in general, $m(g)=\min\{Q(H)(\hat{n}_1,\hat{n}_2):\hat{n}_1,\hat{n}_2\in S\}$ is not zero, but is positive. A similar result appears in Ref. 1, Theorem 1A, but this theorem improves on that one as well because we establish a better sufficient condition for the uniqueness of the solution of $\varphi=\mathcal{M}(\varphi)$ in the whole function space (instead of only in the set $\tilde{T}$ of Ref. 1).

We will henceforth write $m(G)\equiv m$ and $M(G)\equiv M$ wherever there is no possibility of confusion.

Lemma 1: Suppose $M(G)<1$ and φ is a square-integrable solution of $\varphi=\mathcal{M}(\varphi)$. Then

$$\sin[\varphi(\hat{n}_1,\hat{n}_2)]\geq m[(1-M^2)/(1-2mM+m^2)]^{1/2}$$

for almost every $[d_2\Omega]$ $\hat{n}_1$ and $\hat{n}_2$ in S.

Proof: Let $x=\text{ess inf}(\sin\varphi)$. Then from Lemma 3.1,

$$0\leq\sin^{-1}x\leq\varphi(\hat{n}_1,\hat{n}_2)\leq\sin^{-1}M<\pi/2 \text{ a.e.}$$

Since $\varphi=\mathcal{M}(\varphi)$,

$$\begin{aligned}\sin[\varphi(\hat{n}_1,\hat{n}_2)]&\geq Q(H)(\hat{n}_1,\hat{n}_2)\,[(1-M^2)^{1/2}(1-x^2)^{1/2}\\&\quad+Mx]\ \text{ a.e.}\\&\geq m[(1-M^2)^{1/2}(1-x^2)^{1/2}+Mx].\end{aligned}$$

Therefore, $x=\text{ess inf}(\sin\varphi)\geq m(1-M^2)^{1/2}(1-x^2)^{1/2}+Mx)$, and the obvious computation then shows that $x\geq m[(1-M^2)/(1-2mM+m^2)]^{1/2}$.

Hence, if $M<1$, it is enough to look for solutions of $\varphi=\mathcal{M}(\varphi)$ in the set

$$D=\{\varphi\in X:\sin^{-1}m[(1-M^2)/(1-2mM+m^2)]^{1/2}\leq\varphi(\hat{n}_1,\hat{n}_2)\leq\sin^{-1}M\ \text{ a.e. } [d_2\Omega]\}.$$

Theorem 2: Let $G: S\times S\rightarrow\mathbb{R}^+$ be a positive continuous function, and let $0\leq m(G)\leq Q(H)(\hat{n}_1,\hat{n}_2)\leq M(G)<1$ for every $\hat{n}_1$, $\hat{n}_2$ in S. Suppose also that $M_2(G)<1$, where

$$M_2(G)=(1/2\pi)M(M-m)^2[1-2mM+m^2-M^2+M^4]^{-1}\times\sup\{\textstyle\int_S G(\hat{n}_1,\hat{n})\,d\Omega:\hat{n}_1\in S\}.$$

Then there is a unique square-integrable solution of the equation $\varphi=\mathcal{M}(\varphi)$. The solution is the limit of a sequence of successive approximations which converges in the norm of $L^2(S\times S,\ d_2\Omega)$.

Proof: The proof is like that of Theorem 3.3. By Lemma 1, it is enough to show that $\mathcal{M}$ has a unique fixed point in D. We establish this using the Banach contraction mapping principle. To show $\mathcal{M}: D\rightarrow D$, it is enough to show that if $\varphi\in D$, then

$$\sin\mathcal{M}(\varphi)\geq m[(1-M^2)/(1-2mM+m^2)]^{1/2}\ \text{ a.e.}$$

This follows because $m[(1-M^2)/(1-2mM+m^2)]^{1/2}$ is a solution of the equation $x=[(1-M^2)^{1/2}(1-x^2)^{1/2}+Mx]m$. We next show that

$$\|\mathcal{M}(\varphi_1)-\mathcal{M}(\varphi_2)\|\leq M_2(G)^{1/2}\|\varphi_1-\varphi_2\|.$$

To do this, use the method of Martin as in Ref. 5, pp. 137–39, except note now that the minimum of $\sin\varphi$ which makes the right-hand side of (25) of Ref. 5 a maximum is $Q(n_1,n_2)\,[(1-M^2)/(1-2mM+m^2)]^{1/2}$. In the language of Theorem 3.3, this gives $\|\delta\mathcal{M}(\varphi;x)\|\leq M_2(G)^{1/2}\|x\|$ for every $x\in M$ and $\varphi\in D$. We now apply the mean value theorem (Ref. 8, Proposition 2.3) to get the above estimate. The end of the proof is just as in Theorem 3.3.

Corollary 3: If $G\equiv c<1$, then $\text{Im}F=c^2$, and $F==(c-c^2)^{1/2}+ic$ is the (essentially) unique solution of (2.1). That is, if the differential cross section is constant and less than one, there is only one square-integrable scattering amplitude which satisfies unitarity and yields this cross section.

Proof: $M-m=0$. Note that $c\leq1$ is necessary because necessarily $|\text{Im}F|\leq|F|=G$.

Remark: $M_2(G)\leq M_1(G)$, and equality holds if and only if $m(G)=0$. Thus this theorem improves Theorem 3.3. Also, this result is sharper than that of Theorem 1A of Ref. 1. Not only do we obtain a sufficient condition for uniqueness in the whole space, but also this condition is weaker than that in Ref. 1. For example, in the spherically symmetric case, if $M(G)=0.70$, to use Theorem 1A of Ref. 1, we must have $m(G)\gtrsim0.22$ (see illustration, Ref. 11, p. 157). However, Theorem 2 above works with $m(G)=0$, because if $m=0$, then $M_2(G)=M_1(G)$, and by the remark (1) following Theorem 3.3, $M(G)<0.79$ implies $M_1(G)<1$.