JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Radiation Transport along Curved Ray Paths*

CHAU-WA LAU AND KENNETH M. WATSON
Department of Physics and Lawrence Radiation Laboratory
University of California, Berkeley, California

(Received 10 February 1970)

The transport of radiation in a turbulent, refracting medium is studied. It is shown that the conventional transport equation must be generalized. Path integrals are taken along curved ray trajectories. When these ray paths have torsion, a rotation of the polarization vectors needs to be taken into account. Two derivations of the transport equation are given. One is phenomenological and one is based on Maxwell's equations. Some discussion is given of cross polarization of radar backscatter.

1. INTRODUCTION

Earlier papers[1-3] in this series have presented a derivation from Maxwell's equations of the radiation transport equation for scattering by a turbulent medium.[4,5] A transport equation of conventional form was derived making two kinds of approximations. The first of these amounted to treating the scattering from a single "turbulent eddy," or a single correlated cluster of scatterers, in the distorted-wave Born approximation (DWBA). The second approximation made was the assumption that coherent propagation in the refracting medium could be treated in the eikonal approximation and, furthermore, that the ray paths of the eikonal approximation could be replaced by straight lines.

It was observed in I that the restriction to straight-line ray paths is quite unnecessary for deriving the transport equation. In the present paper we drop the restriction to straight ray paths. The resulting transport equation differs from the conventional one in the appearance of a rotation operator acting on the polarization indices and depending on the radius of torsion of the ray paths.

To make our discussion definite, let us consider the physical situation studied in I and illustrated in Fig. 1. The scattering medium is of finite extent and surrounded by empty space. The source of the radiation is at a great distance from the scatterer. Thus, the incident radiation at the scatterer can be considered to be a plane wave, with wavenumber vector (say) $\mathbf{k}$. The detector is also at some distance from the scatterer. Evidently, this particular choice of boundary conditions is incidental for the derivation of the transport equation.

In I the scatterer was assumed to be a plasma. This is evidently easily generalized to other scattering systems by replacing the Thomson scattering amplitudes by those appropriate for the system of interest.

The eikonal approximation, although not required for the formal derivation of a transport equation, seems necessary to obtain the conventional form with its usual geometrical interpretation. In Sec. 2 some required properties of the eikonal approximation are reviewed. A phenomenological derivation of the transport equation, generalized for curved ray paths, is given in Sec. 3. In Sec. 4 the same equation is derived from Maxwell's equations. (The corresponding quantum form can be obtained, as was done in Ref. 4.) Several applications, including a discussion of cross polarization for radar backscatter, are given in the final sections.

Before considering the generalized transport equation, let us review briefly the conventional form of this and the derivation given in I and III. In the classical theory the quantity

$$I(\mathbf{x}, \hat{\mathbf{p}}, \omega)\, d\Omega_{\hat{\mathbf{p}}}\, d\omega \tag{1.1}$$

represents the flow of radiant energy per unit area, per unit time, having angular frequency ω within $d\omega$, and propagating parallel to $\hat{\mathbf{p}}$ within the element of solid angle $d\Omega_{\hat{\mathbf{p}}}$. When the shift in frequency due to scattering may be neglected, ω appears as just a parameter in the transport equation. If, in addition, the bandwidth of the radiation is sufficiently narrow that frequency dispersion in the scattering may be neglected, the transport equation may be expressed in terms of the quantity

$$I(\mathbf{x}, \hat{\mathbf{p}}) \equiv \int_0^\infty I(\mathbf{x}, \hat{\mathbf{p}}, \omega)\, d\omega. \tag{1.2}$$

It is for this case that the derivation in I was given. The equation for $I(\mathbf{x}, \hat{\mathbf{p}}, \omega)$, including frequency shift, was obtained in III.

To describe polarization,it is necessary to generalize (1.1). As described above, we suppose that an incident plane wave with wavenumber vector $\mathbf{k}$ illuminates the scattering medium. At a point along a ray path we suppose the tangent vector is $\hat{\mathbf{p}}$. Two

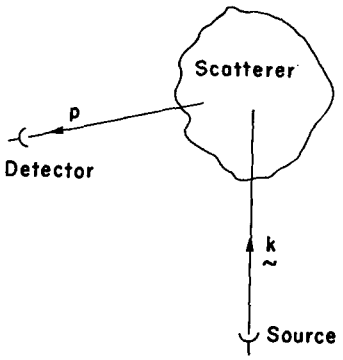


FIG. 1. Illustration of the conditions for scattering considered in this paper.

more unit vectors are then defined as

$$\hat{\mathbf{e}}_{\hat{\mathbf{p}}}(2) = C(\hat{\mathbf{p}})\hat{\mathbf{p}} \times \hat{\mathbf{k}},$$
$$\hat{\mathbf{e}}_{\hat{\mathbf{p}}}(1) = \hat{\mathbf{e}}_{\hat{\mathbf{p}}}(2) \times \hat{\mathbf{p}}. \tag{1.3}$$

For curved ray paths we take $\hat{\mathbf{p}}$ as the local tangent in Eqs. (1.3).

The electric-field vector for a wavelet propagating along the given ray path is therefore of the form

$$\mathbf{E} = E_{\hat{\mathbf{p}}}(1)\hat{\mathbf{e}}_{\hat{\mathbf{p}}}(1) + E_{\hat{\mathbf{p}}}(2)\hat{\mathbf{e}}_{\hat{\mathbf{p}}}(2), \tag{1.4}$$

and the generalized intensity is [here $(i,j) = 1, 2$]

$$I_{ij}(\mathbf{x}, \hat{\mathbf{p}}, \omega) = \text{const} \times [E^*_{\hat{\mathbf{p}}}(i)E_{\hat{\mathbf{p}}}(j)]. \tag{1.5}$$

The "const" here is chosen so that

$$I = I_{11} + I_{22}.$$

[The constant is given an explicit form in Eq. (3.6) below.]

The radiation transport equation obtained in I was of conventional form[6]:

$$\frac{d}{ds} I_{ij}(\mathbf{x}, \hat{\mathbf{p}}) + \frac{1}{l} I_{ij}(\mathbf{x}, \hat{\mathbf{p}}) = B_{ij}(\mathbf{x}, \hat{\mathbf{p}}), \tag{1.6a}$$

where

$$B_{ij}(\mathbf{x}, \hat{\mathbf{p}}) \equiv \sum_{s,t=1}^{2} \int d\Omega_{\hat{\mathbf{p}}'} (ij|\, M\, |st) I_{st}(\mathbf{x}, \hat{\mathbf{p}}'). \tag{1.7a}$$

Here the derivative in (1.6a) is taken along the straight line parallel to $\hat{\mathbf{p}}$, $l^{-1}(\mathbf{x})$ is the absorption coefficient, and $(ij|\, M(\hat{\mathbf{p}}, \hat{\mathbf{p}}')\, |st)$ is the scattering function. In more compact matrix notation we write Eqs. (1.6a) and (1.7a) in the form

$$\frac{d}{ds} \mathbf{I}(\mathbf{x}, \hat{\mathbf{p}}) + \frac{1}{l} \mathbf{I}(\mathbf{x}, \hat{\mathbf{p}}) = \mathbf{B}(\mathbf{x}, \hat{\mathbf{p}}), \tag{1.6b}$$

$$\mathbf{B}(\mathbf{x}, \hat{\mathbf{p}}) = \int d\Omega_{\hat{\mathbf{p}}'} \mathbf{M}(\hat{\mathbf{p}}, \hat{\mathbf{p}}')\mathbf{I}(\mathbf{x}, \hat{\mathbf{p}}'). \tag{1.7b}$$

Equations (1.6) and (1.7) are applicable when the frequency change with scattering may be neglected.

The case specifically considered in I was scattering by a plasma containing N electrons with coordinates $\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_N$. The probability distribution for these coordinates was written as $P_N(\mathbf{z}_1, \cdots, \mathbf{z}_N)$, normalized so that

$$\int P_N d^3z_1 \cdots d^3z_N = 1.$$

A set of distributions

$$P_1(\mathbf{z}_1), P_2(\mathbf{z}_1, \mathbf{z}_2), \cdots, P_{N-1}(\mathbf{z}_1, \cdots, \mathbf{z}_{N-1})$$

may evidently be obtained from P_N by integration, and $P_2, P_3 \cdots$ were developed in terms of 2-particle, 3-particle, $\cdots$ correlation functions. In particular, P_2 was written as

$$P_2(\mathbf{z}_1, \mathbf{z}_2) = P_1(\mathbf{z}_1)P_1(\mathbf{z}_2)[1 + g(\mathbf{z}_1, \mathbf{z}_2)]. \tag{1.8}$$

We shall here assume, as was done in I, that the pair correlation function g has the approximate form

$$g(\mathbf{z}_1, \mathbf{z}_2) \cong g(\mathbf{z}_1; |\mathbf{z}_1 - \mathbf{z}_2|). \tag{1.9}$$

This assumption is not necessary for the derivation of the transport equation, but does simplify the equations by leading to a scalar, isotropic refractive index.

The quantity $\mathbf{M}$ was then obtained in I in the form

$$\mathbf{M}(\hat{\mathbf{p}}, \hat{\mathbf{p}}') = \sigma_g(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}')\mathbf{m}, \tag{1.10}$$

where

$$(ij|\, m\, |st) = [\hat{\mathbf{e}}_{\hat{\mathbf{p}}}(i) \cdot \hat{\mathbf{e}}_{\hat{\mathbf{p}}'}(s)][\hat{\mathbf{e}}_{\hat{\mathbf{p}}}(j) \cdot \hat{\mathbf{e}}_{\hat{\mathbf{p}}'}(t)] \tag{1.11}$$

and

$$\sigma_g(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}') = \rho^2(\mathbf{x})\sigma(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}') \int d^3R g(\mathbf{x}; R) \times \exp[in_r(\mathbf{x})k(\hat{\mathbf{p}}' - \hat{\mathbf{p}}) \cdot \mathbf{R}]. \tag{1.12}$$

Here,

$$\rho(\mathbf{x}) = NP_1(\mathbf{x}) \tag{1.13}$$

is the electron density and σ is the appropriate cross section for Thomson scattering:

$$\sigma(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}') = r_0^2/(1 + \nu_c^2/\omega^2), \tag{1.14}$$

where r_0 is the classical electron radius and ν_c is the electron collision frequency. (Evidently the theory may be adapted to other elementary scatterers.) Finally, $n_r(\mathbf{x})$ is the real part of the refractive index given in first approximation by [see Eq. (4.13)]

$$n_r^2 \cong 1 - \omega_p^2/(\omega^2 + \nu_c^2), \tag{1.15}$$

with ω_p the electron plasma frequency.

The absorption coefficient has the form

$$\frac{1}{l} = \frac{1}{l_c} + \frac{1}{l_t} = 2kn_i, \tag{1.16}$$

where n_i is the imaginary part of the refractive index,

$$\sum_{s=1}^{2}\int d\Omega_{\hat{\mathbf{p}}'}(ij|\, M(\hat{\mathbf{p}},\hat{\mathbf{p}}')\,|ss) = \delta_{ij}\left(\tfrac{1}{2}\int d\Omega_{\hat{\mathbf{p}}'}\sigma_g(\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}')[1+(\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}')^2]\right) \equiv \delta_{ij}l_t^{-1}, \tag{1.17}$$

and

$$l_c^{-1} = [\omega_p^2/(\omega_2+\nu_c^2)](\nu_c/c) \tag{1.18}$$

(c is the speed of light).

As an alternative to Eq. (1.12), we may suppose the medium to be characterized by a dielectric constant, depending on certain random variables, and at time t and point $\mathbf{r}$ to have the value $\epsilon(\mathbf{r},t)$. This may be written in terms of its fluctuations $\delta\epsilon$ as

$$\epsilon = \delta\epsilon + \langle\epsilon\rangle, \tag{1.19}$$

where "$\langle\cdots\rangle$" represents an average over the random variables.

In this case Eq. (1.12) may be rewritten in the form

$$\sigma_g = \left(\frac{k^2}{4\pi}\right)^2\int d^3R\langle\delta\epsilon(\mathbf{x},0)\delta\epsilon(\mathbf{x}+\mathbf{R},0)\rangle \times \exp\,[in_r(\mathbf{x})k(\hat{\mathbf{p}}'-\hat{\mathbf{p}})\cdot\mathbf{R}]. \tag{1.20}$$

The transport equation for the case that frequency shift must be considered was obtained in III. This has the form

$$\frac{d}{ds}\mathbf{I}(\mathbf{x},\hat{\mathbf{p}},\omega)+\frac{1}{l}\mathbf{I}(\mathbf{x},\hat{\mathbf{p}},\omega) = \mathbf{B}(\mathbf{x},\hat{\mathbf{p}},\omega), \tag{1.21}$$

where now

$$\mathbf{B}(\mathbf{x},\hat{\mathbf{p}},\omega) = \int_0^\infty d\omega'\int d\Omega_{\hat{\mathbf{p}}'}\mathbf{M}(\hat{\mathbf{p}},\hat{\mathbf{p}}';\omega-\omega') \times \mathbf{I}(\mathbf{x},\hat{\mathbf{p}}',\omega'), \tag{1.22}$$

and

$$\mathbf{M}(\hat{\mathbf{p}},\hat{\mathbf{p}}';\Omega) = \sigma_g(\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}',\Omega)\mathbf{m}. \tag{1.23}$$

For the case corresponding to Eq. (1.12) we have

$$\sigma_g(\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}',\Omega) = \rho^2(\mathbf{x})\sigma(\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}')\int_{-\infty}^{\infty}d\tau\int d^3Rg(\mathbf{z};R,\tau)e^{i\Omega\tau} \times \exp\,[in_r(\mathbf{x})k(\hat{\mathbf{p}}'-\hat{\mathbf{p}})\cdot\mathbf{R}]. \tag{1.24}$$

Here $g(\mathbf{z};R,\tau)$ is the time-dependent pair correlation function for a stationary random process. Alternatively, for the case corresponding to Eq. (1.20), we have

$$\sigma_g(\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}',\Omega) = \left(\frac{k^2}{4\pi}\right)^2 \times\int_{-\infty}^{\infty}d\tau\int d^3R\langle\delta\epsilon(\mathbf{x},t)\delta\epsilon(\mathbf{x}+\mathbf{R},t+\tau)\rangle \times e^{i\Omega\tau}\exp\,[in_r(\mathbf{x})k(\hat{\mathbf{p}}'-\hat{\mathbf{p}})\cdot\mathbf{R}], \tag{1.25}$$

where $k=\omega/c$.

The approximations required to derive the transport equation from the wave equation are given in detail in I and III. We briefly review these here:

(1) $l_t \gg R_c$, where R_c is the correlation range, or the characteristic distance over which correlations contribute to σ_g. As noted in II, for plasmas this condition can often be replaced by the condition

$$\zeta(\omega_p/\omega)^4 \ll 1,$$

where ζ is the relative mean-square fluctuation in electron density.

(2) $n_i \ll n_r$. This assumption is interpreted as permitting us to neglect n_i in Eqs. (1.12), (1.20), (1.23), and (1.24), where only n_r was kept in the exponentials. That is, we assume that $(R_c/l)\ll 1$ in these equations (this is convenient, but not, of course, essential to our discussion). Assumption (2) is used also in the next section [see Eq. (2.17), for example]. In addition to the above condition, we shall require that $|\nabla n_i|$ be small, a restriction made more precise in Eq. (2.29).

(3) The eikonal approximation may be used to describe coherent wave propagation. We express the condition for validity of this approximation in the form

$$|\nabla n_r| \ll kn_r^2.$$

(4) The bending of eikonal ray paths may be neglected, and these may be considered to be straight lines.

(5) $kl \gg 1$, which we interpret as meaning that successive scatterings occur in the wave zone. This assumption is necessary if we are to use the eikonal approximation to describe wave propagation between scatterings.

The purpose of this paper is to obtain the transport equation without making assumption (4).

2. THE EIKONAL APPROXIMATION

In this section we review several aspects of the eikonal approximation which will be needed for obtaining the transport equation. We shall follow the treatment of Born and Wolf,[7] generalizing this to obtain the Green's function and to include a complex retractive index.

For an electromagnetic wave of frequency $\omega = kc$ propagating in a medium of finite extent and having a refractive index $n(\mathbf{r})$, the Maxwell equation for the electric field $\mathbf{E}(\mathbf{r})$ is

$$[k^2n^2(\mathbf{r}) - \nabla x\nabla x]\mathbf{E}(\mathbf{r}) = 0. \tag{2.1}$$

Let us suppose that in the absence of the refractive medium the field is $\mathbf{E}_0(\mathbf{r})$, where

$$[k^2 - \nabla x \nabla x]\mathbf{E}_0(\mathbf{r}) = 0. \tag{2.2}$$

As illustrated in Fig. 1, we suppose the source of $\mathbf{E}_0$ to lie far outside the refractive medium.

The field $\mathbf{E}$ may be expressed in terms of $\mathbf{E}_0$ and the dyadic Green's function $\mathbf{G}^0(\mathbf{r}, \mathbf{x})$ for infinite space[8]:

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}_0(\mathbf{r}) + \int d^3x \mathbf{G}^0(\mathbf{r}, \mathbf{x}) \cdot \mathbf{E}(\mathbf{x}) \times \{(k^2/4\pi)[n^2(\mathbf{x}) - 1]\}. \tag{2.3}$$

Here $\mathbf{G}^0$ satisfies the equation

$$(k^2 - \nabla_r x \nabla_r x)\mathbf{G}^0(\mathbf{r}, \mathbf{x}) = -4\pi \mathfrak{I}\delta(\mathbf{r} - \mathbf{x}), \tag{2.4}$$

where $\mathfrak{I}$ is the unit dyadic. That is,

$$\mathbf{G}^0(\mathbf{r}, \mathbf{x}) = [\mathfrak{I} + k^{-2}\nabla_x \nabla_x](e^{ikR}/R) - \mathfrak{I}[4\pi/(3k^2)]\delta(\mathbf{R}), \tag{2.5a}$$

with $\mathbf{R} \equiv \mathbf{r} - \mathbf{x}$. For $kR \gg 1$, we have

$$\mathbf{G}^0(\mathbf{r}, \mathbf{x}) \cong [\mathfrak{I} - \hat{\mathbf{R}}\hat{\mathbf{R}}]G^0(\mathbf{r}, \mathbf{x}), \tag{2.5b}$$

where

$$G^0(\mathbf{r}, \mathbf{x}) = e^{ikR}/R. \tag{2.6}$$

Finally, we desire the dyadic Green's function $\mathbf{G}(\mathbf{r}, x)$ for the refracting medium. This satisfies the equation

$$[k^2 n^2(\mathbf{r}) - \nabla_r x \nabla_r x]\mathbf{G}(\mathbf{r}, \mathbf{x}) = -4\pi \mathfrak{I}\delta(\mathbf{r} - \mathbf{x}). \tag{2.7}$$

Let us now write the electric- and magnetic-field vectors in the respective forms

$$\mathbf{E}(\mathbf{r}) = \mathbf{e}(\mathbf{r})e^{ikS(\mathbf{r})},$$
$$\mathbf{H}(\mathbf{r}) = \mathbf{h}(\mathbf{r})e^{ikS(\mathbf{r})}, \tag{2.8}$$

where the eikonal S is expressed in terms of its real and imaginary parts as[9]

$$S = S_r + iS_r. \tag{2.9}$$

Since $\mathbf{G}$ and $\mathbf{E}$ satisfy the same differential equation, except for boundary conditions, we may write

$$\mathbf{G}(\mathbf{r}, \mathbf{x}) = \mathbf{e}(\mathbf{r}, \mathbf{x}) \exp [ikS(\mathbf{r}, \mathbf{x})]. \tag{2.10}$$

Now we consider $\mathbf{e}$ to be a dyadic satisfying the same differential equation [Eq. (2.12)] as the vector $\mathbf{e}$.

Substitution of (2.8) into Maxwell's equations give the two differential equations of the eikonal approximation for S and $\mathbf{e}$ (we are using the notation of Born and Wolf):

$$(\nabla S)^2 = n^2, \tag{2.11}$$

$$[(\nabla^2 S) + (\nabla S)(\nabla \ln n^2) + 2(\nabla S) \cdot \nabla]\mathbf{e} = 0. \tag{2.12}$$

In addition, we have

$$\mathbf{h} = (\nabla S) \times \mathbf{e},$$
$$(\nabla S) \cdot \mathbf{e} = 0, \tag{2.13}$$
$$(\nabla S) \cdot \mathbf{h} = 0,$$

in the first order.

The time-averaged Poynting vector is (following Ref. 7, we are using unrationalized Gaussian units)

$$\mathbf{J} = \frac{c}{8\pi}[\mathrm{Re}\,(\mathbf{e} \times \mathbf{h}^*)]e^{-2kS_i}$$
$$= \frac{c}{8\pi}(\nabla S_r)\mathbf{e} \cdot \mathbf{e}^* e^{-2ikS_i}. \tag{2.14}$$

The real and imaginary parts of Eq. (2.11) lead to the two real equations

$$(\nabla S_r)^2 - (\nabla S_i)^2 = n_r^2 - n_i^2, \tag{2.15}$$

$$(\nabla S_r) \cdot (\nabla S_i) = n_r n_i. \tag{2.16}$$

Because of our assumptions that $n_i^2 \ll n_r^2$, we look for a solution to these of the form

$$(\nabla S_r)^2 \cong n_r^2, \tag{2.17}$$

with $(\nabla S_i)^2 \ll (\nabla S_r)^2$.

Following the discussion of Ref. 7, we write the equation of a given ray path as $\mathbf{r} = \mathbf{r}(s)$, where s is the path length measured from some reference point on the ray. We let

$$\hat{\mathbf{p}}(\mathbf{r}) \equiv \frac{d\mathbf{r}}{ds} \tag{2.18}$$

be the tangent vector to the given ray path at the point $\mathbf{r}$ and write

$$\nabla S_r = \hat{\mathbf{p}} n_r. \tag{2.19}$$

Equations (2.18) and (2.19) permit us to "solve" (2.17) in the form[7]

$$\frac{d\hat{\mathbf{p}}}{ds} = \nabla_\perp \ln n_r, \tag{2.20}$$

$$S_r(\mathbf{r}) = \int_p n_r\, ds, \tag{2.21}$$

$$\nabla_\perp \equiv \nabla - \hat{\mathbf{p}}\hat{\mathbf{p}} \cdot \nabla. \tag{2.22}$$

Equation (2.20) determines a given ray path, and (2.21) permits us to construct S_r, the integral being taken along the ray path passing through $\mathbf{r}$.

Equation (2.16) may now be rewritten in the form

$$\hat{\mathbf{p}} \cdot \nabla S_i = n_i. \tag{2.23}$$

This may be integrated to give

$$S_i = \int_p n_i \, ds + \delta S_i,$$

where δS_i is subject to the condition

$$\hat{\mathbf{p}} \cdot \nabla \delta S_i = 0.$$

Thus, δS_i is constant along each ray path. The boundary condition "$\mathbf{E}$ becomes equal to $\mathbf{E}_0$ if we follow a ray path backwards outside the medium" tells us then that $\delta S_i = 0$. For the Green's function the boundary condition at $\mathbf{r} \cong \mathbf{x}$, implied by Eq. (2.7), also then specifies that $\delta S_i = 0$. We therefore have

$$S_i = \int_p n_i \, ds. \tag{2.24}$$

We now turn to Eq. (2.12) for $\mathbf{e}$. Let us write $\mathbf{e}$ in the form

$$\mathbf{e} = A e^{i\alpha} \hat{\mathbf{e}}, \tag{2.25}$$

where A is the magnitude of $\hat{\mathbf{e}}$, $\hat{\mathbf{e}}$ is a unit vector, and α is a phase.

The discussion given in Born and Wolf[7] permits us to obtain directly the equation

$$\frac{dA}{ds} = -\left(\frac{\nabla^2 S_r}{2n_r}\right) A. \tag{2.26}$$

For α and $\hat{\mathbf{e}}$, we obtain the equations

$$\frac{d\alpha}{ds} = -\frac{\nabla^2 S_i}{2n_r} \tag{2.27}$$

and

$$\frac{d\hat{\mathbf{e}}}{ds} = -\hat{\mathbf{p}}(\nabla \ln n) \cdot \hat{\mathbf{e}}. \tag{2.28}$$

To check the consistency with Eqs. (2.13), we use Eqs. (2.20) and (2.28) to obtain

$$\frac{d}{ds}(\hat{\mathbf{e}} \cdot \hat{\mathbf{p}}) = \hat{\mathbf{e}} \cdot \nabla_\perp \ln n_r - \hat{\mathbf{e}} \cdot \nabla \ln n.$$

This vanishes if we can consider the contribution from n_i in Eq. (2.28) to be negligible. This will be the case if

$$\left| \int_p \frac{|\nabla n_i|}{n_r} \, ds \right| \ll 1 \tag{2.29}$$

everywhere along a given path. We shall assume that the condition (2.29) is satisfied and that we can write (2.29) in the approximate form

$$\frac{d\hat{\mathbf{e}}}{ds} = -\hat{\mathbf{p}}(\nabla_\perp \ln n_r) \cdot \hat{\mathbf{e}}. \tag{2.30}$$

The same approximation lets us set $\alpha = 0$ in Eq. (2.25).

If $\mathbf{r}$ and $\mathbf{x}$ are two points along a ray path, we may integrate (2.30) to give

$$\hat{\mathbf{e}}(\mathbf{r}) = \mathbf{P}(\mathbf{r}, \mathbf{x}) \cdot \hat{\mathbf{e}}(\mathbf{x}),$$
$$\mathbf{P}(\mathbf{r}, \mathbf{x}) = \exp\left(-\int_{\mathbf{x}}^{\mathbf{r}} \hat{\mathbf{p}}(\nabla_\perp \ln n_r) \, ds\right), \tag{2.31}$$

using a somewhat terse notation.

We are particularly interested in the eikonal representation for the Green's function $\mathbf{G}(\mathbf{r}, \mathbf{x})$. To simplify the appearance of some of our equations, we shall use interchangeably two notations for $\mathbf{G}$, A, etc. For a set of coordinates $\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_N$, we write

$$\mathbf{G}_{\alpha\beta} \equiv \mathbf{G}(\mathbf{z}_\alpha, \mathbf{z}_\beta), \tag{2.32}$$

with $\alpha \neq \beta = 1, 2, \cdots, N$. The unit tangent vectors $\hat{\mathbf{p}} = \hat{\mathbf{p}}(\mathbf{x})$ in Eqs. (1.3) satisfy Eq. (2.20). These will be written as $\hat{\mathbf{p}}_\alpha \equiv \hat{\mathbf{p}}(\mathbf{z}_\alpha)$, etc. The eikonal is written as

$$S_{\alpha\beta} = \int_{\mathbf{z}_\alpha}^{\mathbf{z}_\beta} n \, ds.$$

In matrix notation we use the representation (1.3) to write, for $i, j = 1, 2$,

$$(j| \, G_{\alpha\beta} \, |i) \equiv \hat{\mathbf{e}}_{\hat{\mathbf{p}}_\alpha}(j) \cdot \mathbf{G}_{\alpha\beta} \cdot \hat{\mathbf{e}}_{\hat{\mathbf{p}}_\beta}(i)$$
$$= A_{\alpha\beta} \exp(ikS_{\alpha\beta}) \hat{\mathbf{e}}_{\hat{\mathbf{p}}_\alpha}(j) \cdot \mathbf{P}_{\alpha\beta} \cdot \hat{\mathbf{e}}_{\hat{\mathbf{p}}_\beta}(i), \tag{2.33}$$

using Eqs. (2.25) and (2.31). We emphasize that $S_{\alpha\beta}$ and $A_{\alpha\beta}$ here are determined by Eqs. (2.21), (2.24), and (2.26). We may also write

$$(j| \, G_{\alpha\beta} \, |i) \equiv G_{\alpha\beta} \hat{\mathbf{e}}_{\hat{\mathbf{p}}_\alpha}(j) \cdot \mathbf{P}_{\alpha\beta} \cdot \hat{\mathbf{e}}_{\hat{\mathbf{p}}_\beta}(i), \tag{2.34}$$

where

$$G_{\alpha\beta} \equiv G(\mathbf{z}_\alpha, \mathbf{z}_\beta) = A_{\alpha\beta} \exp(ikS_{\alpha\beta}). \tag{2.35}$$

The expression (2.35) represents the eikonal approximation to the scalar equation

$$[\nabla_z^2 + k^2 n^2(\mathbf{z})] G(\mathbf{z}, \mathbf{x}) = -4\pi\delta(\mathbf{z} - \mathbf{x}). \tag{2.36}$$

The observation that Eq. (2.26) does not depend on n_i lets us introduce the Green's function, expressed in the eikonal approximation:

$$\hat{G}(\mathbf{r}, \mathbf{x}) = A(\mathbf{r}, \mathbf{x}) \exp[ikS_r(\mathbf{r}, \mathbf{x})]. \tag{2.37}$$

This is the eikonal approximation to the solution of

$$(\nabla_r^2 + k^2 n_r^2) \hat{G}(\mathbf{r}, \mathbf{x}) = -4\pi\delta(\mathbf{r} - \mathbf{x}). \tag{2.38}$$

It is known[10] that

$$\hat{G}(\mathbf{r}, \mathbf{x}) = \hat{G}(\mathbf{x}, \mathbf{r}), \tag{2.39}$$

from which we obtain the important symmetry relation

$$A_{\alpha\beta} = A_{\beta\alpha}. \tag{2.40}$$

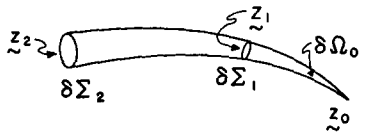


FIG. 2. A narrow tube of ray paths all passing through the point $\mathbf{z}_0$.

Let us now choose three points $\mathbf{z}_0$, $\mathbf{z}_1$, and $\mathbf{z}_2$ all on a given ray path. We construct a flux tube of ray paths all passing through $\mathbf{z}_0$ and enclosing the original path on which $\mathbf{z}_1$ and $\mathbf{z}_2$ lie. We suppose the solid angle $\delta\Omega_0$ formed by the tube at $\mathbf{z}_0$ to be very small. The respective cross-sectional areas of the flux tube at $\mathbf{z}_1$ and $\mathbf{z}_2$ are $\delta\Sigma_1$ and $\delta\Sigma_2$. This geometry is illustrated in Fig. 2.

We assume for the moment that $n_i = 0$. Then we can represent the Green's function G by $\hat{G}$, Eqs. (2.37) and (2.38). The condition that energy is conserved [i.e., $n_i = 0$] in the flux tube [see Eq. (2.14)] leads us to the relation

$$\delta\Sigma_2 A_{20}^2 n_r(2) = \delta\Sigma_1 A_{10}^2 n_r(1). \tag{2.41}$$

Here we have written $n_r(1) \equiv n_r(\mathbf{z}_1)$, etc. If we choose $\mathbf{z}_1$ to lie sufficiently close to $\mathbf{z}_0$, we have

$$A_{10} = R_{10}^{-1}$$

and

$$\delta\Sigma_{10} = R_{10}^2 \delta\Omega_0,$$

where

$$\mathbf{R}_{\alpha\beta} \equiv \mathbf{z}_\alpha - \mathbf{z}_\beta. \tag{2.42}$$

Thus, we obtain from (2.41) the result

$$A_{20} = \left[\left(\frac{n_r(0)}{n_r(2)}\right)\left(\frac{\partial\Sigma_2}{\partial\Omega_0}\right)^{-1}\right]^{\frac{1}{2}};$$

for two arbitrary points $\mathbf{z}_\alpha$ and $\mathbf{z}_\beta$ we have

$$A_{\alpha\beta} = \left[\left(\frac{n_r(\beta)}{n_r(\alpha)}\right)\left(\frac{\partial\Sigma_\alpha}{\partial\Omega_\beta}\right)^{-1}\right]^{\frac{1}{2}}. \tag{2.43}$$

Since Eq. (2.26) does not involve n_i, we can consider Eq. (2.43) to be valid even when $n_i \neq 0$ (but, of course, subject to the conditions imposed on it).

Equations (2.34), (2.35), (2.43), etc., provide the eikonal representation of the Green's function which will be required for our applications.

3. HEURISTIC DERIVATION OF THE TRANSPORT EQUATION

In this section we give a phenomenological derivation of the transport equation for the case that curvature of the ray paths must be taken into account. Consistent with our assumption [see Eq. (1.9)] that mean properties of the medium do not change significantly over a correlation distance R_c, we shall assume that the scattering strength $\mathbf{B}$ in Eq. (1.6) and the scattering function $\mathbf{M}$ are not modified.

Let us assume that the dielectric properties of the scatterer are described by the dielectric constant ϵ [see Eq. (1.19)] which depends on certain random parameters. The electric field $\mathbf{E}$, which depends parametrically on these same parameters, satisfies the equation

$$(k^2\epsilon - \nabla x \nabla x)\mathbf{E} = 0. \tag{3.1}$$

Remembering that "$\langle \cdots \rangle$" is considered to represent an average over these random variables, we introduce the *coherent* field at $\mathbf{x}$ as

$$\mathbf{E}_c(\mathbf{x}) \equiv \langle \mathbf{E}(\mathbf{x}) \rangle. \tag{3.2}$$

The coherent field satisfies the wave equation (2.1); that is, as was shown in I,[11]

$$(k^2 n^2 - \nabla x \nabla x)\mathbf{E}_c = 0. \tag{3.3}$$

The field $\mathbf{E}$ may be written as

$$\mathbf{E} = \sum_\lambda \mathbf{E}_\lambda, \tag{3.4}$$

where the sum runs over scattering from different fluctuations in ϵ and includes $\mathbf{E}_c$. We suppose that these terms are so chosen that the different $\mathbf{E}_\lambda$ are mutually incoherent. Thus,

$$\langle \mathbf{E}_\lambda \mathbf{E}_{\lambda'} \rangle = 0, \quad \text{for} \quad \lambda \neq \lambda'. \tag{3.5}$$

An expression is required for the radiation intensity $I_{ij}(\mathbf{z}, \hat{\mathbf{p}})$ [Eq. (1.5); we do not write explicitly the frequency dependence here]. Referring to Fig. 3, we consider a flux tube of ray paths all passing through a given point $\mathbf{z}_0$ within a small solid angle $\delta\Omega_0$. The electric field at $\mathbf{z}_0$ associated with these ray paths will be of the form (3.4). The index λ is assumed to refer to a scatterer at the point $\mathbf{z}_\lambda$ contained within the volume of the flux tube. At the point $\mathbf{z}_\lambda$ the area of the flux tube is $\delta\Sigma_\lambda$.

If we write each $\mathbf{E}_\lambda$ in (3.4) in the eikonal form (2.8) and use Eqs. (2.14) and (3.5), we see that

$$I_{ij}(\mathbf{z}_0, \hat{\mathbf{p}}_0) = \frac{c}{8\pi}(n_r(0)/\delta\Omega_0) \times \left\langle \sum_\lambda \exp\left[-2kS_i(\mathbf{z}_0, \mathbf{z}_\lambda)\right] \hat{\mathbf{e}}_{\hat{\mathbf{p}}_0}(i) \cdot \mathbf{e}_\lambda^* \mathbf{e}_\lambda \cdot \hat{\mathbf{e}}_{\hat{\mathbf{p}}_0}(j) \right\rangle. \tag{3.6}$$

FIG. 3. The intensity at $\mathbf{z}_0$ results from sources at points $\mathbf{z}_\lambda$ in the flux tube.

The unit vectors here are defined by Eqs. (1.3) with $\hat{\mathbf{p}}_0 \equiv \hat{\mathbf{p}}(\mathbf{z}_0)$.

Using Eq. (2.43), we have

$$\begin{aligned}\delta\Omega_0 &= \left(\frac{d\Sigma_\lambda}{d\Omega_0}\right)^{-1}\delta\Sigma_\lambda \\ &= A_{\lambda 0}^2\,\frac{n_r(\lambda)}{n_r(0)}\,\delta\Sigma_\lambda \\ &= A_{0\lambda}^2\,\frac{n_r(\lambda)}{n_r(0)}\,\delta\Sigma_\lambda . \end{aligned} \tag{3.7}$$

To obtain the final form here, we have used Eq. (2.40). Equation (2.25), with $\alpha = 0$, lets us write

$$\mathbf{e}_\lambda = A_{0\lambda}\hat{\mathbf{e}}_\lambda(\mathbf{z}_0)s_\lambda , \tag{3.8}$$

where s_λ is the source intensity at $\mathbf{z}_\lambda$. Finally, (3.7) and (3.8) let us rewrite (3.6), replacing $\mathbf{z}_0$ by a general point $\mathbf{z}$, as

$$I_{ij}(\mathbf{z}, \hat{\mathbf{p}}) = n_r^2(\mathbf{z}) \exp\,[-2kS_i(\mathbf{z})] \times \left\langle \sum_\lambda C_\lambda \hat{\mathbf{e}}_{\hat{\mathbf{p}}}(i)\cdot\hat{\mathbf{e}}_\lambda \hat{\mathbf{e}}_{\hat{\mathbf{p}}}(j)\cdot\hat{\mathbf{e}}_\lambda \right\rangle, \tag{3.9}$$

where the C_λ are independent of $\mathbf{z}$.

To obtain the transport equation, we consider the change in $\mathbf{I}$ if we displace the point z by a small distance δs along the ray path. First, let us assume that in the interval δs there is no energy added by scattering from other ray paths. (This will later be accounted for by including the scattering strength $\mathbf{B}$.) Use of Eqs. (1.16) and (2.24) gives

$$\frac{d}{ds}(2kS_i) = \frac{1}{l} .$$

From Eqs. (2.20), (2.38), and (1.3) we obtain

$$\frac{d}{ds}\,[\hat{\mathbf{e}}_{\hat{\mathbf{p}}}(i)\cdot\hat{\mathbf{e}}] = (-1)^i\phi\hat{\mathbf{e}}_{\hat{\mathbf{p}}}(i+1)\cdot\hat{\mathbf{e}}, \tag{3.10}$$

where

$$\phi = (\hat{\mathbf{p}}\cdot\hat{\mathbf{k}}/|\hat{\mathbf{p}}\times\hat{\mathbf{k}}|)\hat{\mathbf{e}}_{\hat{\mathbf{p}}}(2)\cdot\nabla \ln n_r . \tag{3.11}$$

In Eq. (3.10) we have adopted the following special notation. When $i = 1$,

$$\hat{\mathbf{e}}_{\hat{\mathbf{p}}}(i+1) = \hat{\mathbf{e}}_{\hat{\mathbf{p}}}(2);$$

when $i = 2$,

$$\hat{\mathbf{e}}_{\hat{\mathbf{p}}}(i+1) = \hat{\mathbf{e}}_{\hat{\mathbf{p}}}(1).$$

With these expressions, we obtain from (3.9) the expression

$$\begin{aligned}\frac{d}{ds}I_{ij}(\mathbf{z}, \hat{\mathbf{p}}) &= -\frac{1}{L}I_{ij}(\mathbf{z}, \hat{\mathbf{p}}) \\ &\quad + (-1)^j\phi I_{ij+1}(\mathbf{z}, \hat{\mathbf{p}}) + (-1)^i\phi I_{i+1\ j}(\mathbf{z}, \hat{\mathbf{p}}) \\ &\equiv -\frac{1}{L}I_{ij} - \sum_{s,t=1}^{2}(ij|\,R\,|st)I_{st} . \end{aligned} \tag{3.12}$$

In the first term we have introduced the notation

$$\frac{1}{L} \equiv \frac{1}{l} - \frac{d}{ds}\ln n_r^2 . \tag{3.13}$$

The second term describes the rotation of the polarization vectors as the radiation moves along a ray path.

Scattering within the interval δs was omitted in obtaining Eq. (3.12). If we include this, we must add $\mathbf{B}$ to the right-hand side of (3.12). This gives the generalized form of the transport equation (1.6):

$$\frac{d}{ds}\mathbf{I}(\mathbf{x}, \hat{\mathbf{p}}) + \frac{1}{L}\mathbf{I}(\mathbf{x}, \hat{\mathbf{p}}) + \mathbf{R}\mathbf{I}(\mathbf{x}, \hat{\mathbf{p}}) = \mathbf{B}(\mathbf{x}, \hat{\mathbf{p}}), \tag{3.14}$$

where $\mathbf{B}$ is given by Eq. (1.7). If frequency shift due to scattering must be taken into account, we obtain the generalized form of Eq. (1.21):

$$\frac{d}{ds}\mathbf{I}(\mathbf{x}, \hat{\mathbf{p}}, \omega) + \frac{1}{L}\mathbf{I}(\mathbf{x}, \hat{\mathbf{p}}, \omega) + \mathbf{R}\mathbf{I}(\mathbf{x}, \hat{\mathbf{p}}, \omega) = \mathbf{B}(\mathbf{x}, \hat{\mathbf{p}}, \omega). \tag{3.15}$$

The scattering strength $\mathbf{B}$ is now given by Eq. (1.22).

Let us write out in detail Eqs. (3.14) and (3.15). For either case, the form is

$$\begin{aligned}\frac{d}{ds}I_{11} &= -\frac{1}{L}I_{11} - \phi(I_{12} + I_{21}) + B_{11} , \\ \frac{d}{ds}I_{12} &= -\frac{1}{L}I_{12} + \phi(I_{11} - I_{22}) + B_{12} , \\ \frac{d}{ds}I_{21} &= -\frac{1}{L}I_{21} + \phi(I_{11} - I_{22}) + B_{21} , \\ \frac{d}{ds}I_{22} &= -\frac{1}{L}I_{22} + \phi(I_{12} + I_{21}) + B_{22} . \end{aligned} \tag{3.16}$$

For many applications the Stokes parameter representation is more convenient than that of Eqs. (1.3). In this representation we write

$$\begin{aligned} I_1 &= I_{11} , \\ I_2 &= I_{22} , \\ I_{12} &= \tfrac{1}{2}(I_3 - iI_4) = I_{21}^* . \end{aligned} \tag{3.17}$$

(The detailed form of $\mathbf{B}$ in this representation was given in II.) Equations (3.16), when expressed in terms of the Stokes parameters, become

$$\begin{aligned}\frac{dI_1}{ds} &= -\frac{1}{L}I_1 - \phi I_3 + B_1 , \\ \frac{dI_2}{ds} &= -\frac{1}{L}I_2 + \phi I_3 + B_2 , \\ \frac{dI_3}{ds} &= -\frac{1}{L}I_3 + 2\phi(I_1 - I_2) + B_3 , \\ \frac{dI_4}{ds} &= -\frac{1}{L}I_4 + B_4 . \end{aligned} \tag{3.18}$$

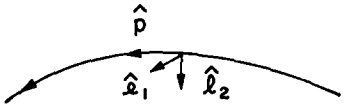


FIG. 4. Illustration of the unit vectors introduced in Eqs. (3.19) and (3.20).

A somewhat more "natural" choice than (1.3) for the unit polarization vectors is illustrated in Fig. 4. At each point on a ray path, $\hat{\mathbf{e}}_2$ lies in the osculating plane along the principal radius of curvature. Thus,

$$\frac{d\hat{\mathbf{p}}}{ds} = \frac{\hat{\mathbf{e}}_2}{R_p}, \tag{3.19}$$

where R_p is the principal radius of curvature. Then

$$\hat{\mathbf{e}}_1 = \hat{\mathbf{e}}_2 \times \hat{\mathbf{p}}, \tag{3.20}$$

and we obtain

$$\frac{d\hat{\mathbf{e}}_1}{ds} = -\frac{\hat{\mathbf{e}}_2}{R_t},$$

$$\frac{d\hat{\mathbf{e}}_2}{ds} = -\frac{\hat{\mathbf{p}}}{R_p} + \frac{\hat{\mathbf{e}}_1}{R_t}. \tag{3.21}$$

Here R_t is the radius of torsion of the ray path.

If we define $\mathbf{m}$, Eq. (1.11), and $\mathbf{I}$, Eq. (3.6), with $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$ replacing the polarization vectors (1.3), the only change in the transport equation [Eqs. (3.14) or (3.15)] occurs in the rotation matrix $\mathbf{R}$. After a simple calculation using Eqs. (3.21), we obtain

$$\frac{dI_{ij}}{ds} = -\frac{1}{L}I_{ij} + \frac{1}{R_t}[(-1)^i I_{i+1\,j} + (-1)^j I_{i\,j+1}] + B_{ij}. \tag{3.22}$$

These have the same form as do Eqs. (3.16), but with ϕ replaced by R_t^{-1}.

The form (3.22) will not be used in this paper.

4. DERIVATION OF THE TRANSPORT EQUATION

We turn now to the derivation of Eq. (3.14) from Maxwell's equations[12] [the corresponding derivation of Eq. (3.15) following the method used in III, is straightforward]. For this purpose, we can use the development given in I with only superficial changes—mostly of notation.[13] To avoid repetition of the complete development in I, we shall just indicate here these changes.

As in I, we shall suppose the scattering system (plasma) to consist of N nonrelativistic electrons confined to a finite volume and having the probability distribution $P_N(\mathbf{z}_1, \cdots, \mathbf{z}_N)$ described in Sec. 1. This is illustrated in Fig. 1, where the source and detector are shown as being a large distance from the plasma. [The extension to other scattering systems is evidently straightforward. It is only necessary to use the appropriate scattering amplitudes f_{ij} in Eqs. (4.3), etc.] The multiple scattering representation for the electric field was given by Eqs. (I.3.1) and (I.3.2)[14]:

$$\mathbf{E}(\mathbf{z}_\alpha) = \mathbf{E}_I(\mathbf{z}_\alpha) + \sum_{\beta(\neq\alpha)=1}^{N} \sum_{j=1}^{2} \hat{\mathbf{e}}_{\alpha\beta}(j) F_{\alpha\beta}(j). \tag{4.1}$$

Here $\hat{\mathbf{e}}_{\alpha\beta}(j)$, $j = 1, 2$, defined by Eqs. (1.3) with $\hat{\mathbf{p}}$ parallel to $\mathbf{z}_\alpha - \mathbf{z}_\beta$ and

$$\mathbf{E}_I(\mathbf{x}) = \hat{\mathbf{e}}(1) E_I(\mathbf{x}), \tag{4.2}$$

represents the incident radiation, taken to be a plane wave at the scatterer. The $F_{\alpha\beta}$ are defined by the equations

$$F_{\alpha\beta}(i) = G^0_{\alpha\beta} f_{i1}(\alpha\beta, \beta 0) E_I(\mathbf{z}_\beta) + \sum_{\sigma(\neq\beta)=1}^{N} \sum_{j=1}^{2} G^0_{\alpha\beta} f_{ij}(\alpha\beta, \beta\sigma) F_{\beta\sigma}(j). \tag{4.3}$$

The notation of I has been used here, with

$$G^0_{\alpha\beta} \equiv G^0(\mathbf{z}_\alpha, \mathbf{z}_\beta), \tag{4.4}$$

as defined by Eq. (2.6), and

$$f_{ij}(\alpha\beta, \beta\sigma) = f_0 \hat{\mathbf{e}}_{\alpha\beta}(i) \cdot \hat{\mathbf{e}}_{\beta\sigma}(j),$$
$$f_{i1}(\alpha\beta, \beta 0) = f_0 \hat{\mathbf{e}}_{\alpha\beta}(i) \cdot \hat{\mathbf{e}}(1). \tag{4.5}$$

The quantity $(-f_0)$ is the classical electron radius divided by $(1 - i\nu_c/\omega)$.

If we expand Eqs. (4.1) and (4.3) in a sequence of scatterings and express the result in vector form, we have

$$\mathbf{E}(\mathbf{z}_\alpha) = \mathbf{E}_I(\mathbf{z}_\alpha) + \sum_\sigma \left(f_0 \mathbf{G}^0_{\alpha\sigma} + f_0^2 \sum_\beta \mathbf{G}^0_{\alpha\beta} \cdot \mathbf{G}^0_{\beta\sigma} + f_0^3 \sum_{\beta_1,\beta_2} \mathbf{G}^0_{\alpha\beta_1} \cdot \mathbf{G}^0_{\beta_1\beta_2} \cdot \mathbf{G}^0_{\beta_2\sigma} + \cdots \right) \cdot \mathbf{E}_I(\mathbf{z}_\sigma), \tag{4.6}$$

where $\mathbf{G}^0_{\alpha\beta} \equiv \mathbf{G}^0(\mathbf{z}_\alpha, \mathbf{z}_\beta)$ is the expression (2.5b) and no two adjacent subscripts $\alpha, \beta_1, \cdots$ above are equal.

For a coherent sequence of scatterings we obtain (I.3.22) in the form

$$\mathbf{G}_{\alpha\sigma} = \mathbf{G}^0_{\alpha\sigma} + f_0 \sum_\beta \int \mathbf{G}^0_{\alpha\beta} \cdot \mathbf{G}_{\beta\sigma} P_1(\mathbf{z}_\beta)\, d^3 z_\beta + f_0^2 \sum_{\beta_1,\beta_2(\neq\beta_1)} \int \mathbf{G}^0_{\alpha\beta_1} \cdot \mathbf{G}_{\beta_1\beta_2} \cdot \mathbf{G}_{\beta_2\sigma} \times P_1(\mathbf{z}_{\beta_1}) P_1(\mathbf{z}_{\beta_2}) g(\mathbf{z}_{\beta_1}, \mathbf{z}_{\beta_2}) d^3 z_{\beta_1} d^3 z_{\beta_2}, \tag{4.7}$$

where g is the pair correlation function of Eq. (1.8), etc. We emphasize that Eq. (4.7) is just Eq. (I.3.22) rewritten in vector notation.

For $|\mathbf{z}_{\beta_1} - \mathbf{z}_{\beta_2}| \lesssim R_c$, we can rewrite Eqs. (I.3.23) and (I.3.24) as

$$\mathbf{G}_{\beta_2\sigma} \cong \{\exp[in_r(\mathbf{z}_{\beta_1})\mathbf{k}_{\beta_1\sigma} \cdot \mathbf{R}_{\beta_2\beta_1}]\}\mathbf{G}_{\beta_1\sigma}, \tag{4.8}$$

$$\mathbf{G}_{\beta_1\beta_2} \cong [\mathfrak{I} - \hat{\mathbf{R}}_{\beta_1\beta_2}\hat{\mathbf{R}}_{\beta_1\beta_2}][\exp(in_r(\mathbf{z}_{\beta_1})kR_{\beta_1\beta_2})]/R_{\beta_1\beta_2}. \tag{4.9}$$

Here $\mathbf{k}_{\beta_1\sigma}$ is the wavenumber vector at $\mathbf{z}_{\beta_1}$ for a ray originating at $\mathbf{z}_\sigma$. This represents a generalization approximate to curved ray paths of the notation used in I. We observe that [see Eqs. (2.13)]

$$\mathbf{k}_{\beta_1\sigma} \cdot \mathbf{G}_{\beta_1\sigma} = 0. \tag{4.10}$$

Use of Eqs. (4.8), (4.9), and (4.10) lets us rewrite (4.7) as the vector generalization of Eq. (I.3.25):

$$\mathbf{G}_{\alpha\sigma} = \mathbf{G}^0_{\alpha\sigma} + \int \mathbf{G}^0_{\alpha\beta} \cdot \mathbf{G}_{\beta\sigma}\gamma(\mathbf{z}_\beta)d^3z_\beta, \tag{4.11}$$

where

$$\gamma(\mathbf{z}) = \rho(\mathbf{z})f_0 + \rho^2(\mathbf{z})f_0^2\int d^3R\{\tfrac{1}{2}[1 + (\hat{\mathbf{q}} \cdot \hat{\mathbf{R}})^2]\} \times g(\mathbf{z}; R)\{\exp[in_r(\mathbf{z})k(R - \hat{\mathbf{q}} \cdot \mathbf{R})]\}/R \tag{4.12}$$

is just the expression (I.3.26b).

It follows from Eq. (2.4) that $\mathbf{G}_{\alpha\sigma}$ [defined by Eq. (4.11)] satisfies Eq. (2.7) with

$$n^2(\mathbf{z}) = 1 + (4\pi/k^2)\gamma(\mathbf{z}). \tag{4.13}$$

We may therefore use the eikonal representation (2.33) for $\mathbf{G}_{\alpha\sigma}$.

Equations (I.3.31) and (I.3.32) may now be rewritten as

$$\mathbf{E}(\mathbf{z}_\alpha) = \mathbf{E}_c(\mathbf{z}_\alpha) + \sum_{\beta(\neq\alpha)=1}^{N} \hat{\mathbf{e}}_{\hat{\mathbf{p}}_\alpha}(j)E_{\alpha\beta}(j), \tag{4.14}$$

$$E_{\alpha\beta}(i) = \sum_{l,j=1}^{2}\Big((i|\, G_{\alpha\beta}\,|l)f_{lj}(\alpha\beta, \beta 0)E_c(\mathbf{z}_\beta, j) + \sum_{\sigma(\neq\beta)=1}^{N}(i|\, G_{\alpha\beta}\,|l)f_{lj}(\alpha\beta, \beta\sigma)E_{\beta\sigma}(j)\Big). \tag{4.15}$$

In Eq. (4.14) we have written $\hat{\mathbf{e}}_{\hat{\mathbf{p}}_\alpha}(j) \equiv \hat{\mathbf{e}}_{\hat{\mathbf{p}}_\alpha}(j; \alpha, \beta)$, $j = 1, 2$, to indicate the unit vectors (1.3) on the ray path joining $\mathbf{z}_\beta$ and $\mathbf{z}_\alpha$. [When no confusion will result, we shall use the abbreviated notation of Eqs. (4.14) and (2.33).] In Eq. (4.15) we now take

$$f_{lj}(\alpha\beta, \beta\sigma) = f_0\hat{\mathbf{e}}_{\hat{\mathbf{p}}_\beta}(l; \alpha\beta) \cdot \hat{\mathbf{e}}_{\hat{\mathbf{p}}_\beta}(j; \beta\sigma), \tag{4.16}$$

etc. The *coherent field* $\mathbf{E}_c$ is just the quantity (3.2). It satisfies Eq. (3.3) with $\mathbf{E}_I$ as the incident field. At a point $\mathbf{z}$ the tangent vector to the ray path of $\mathbf{E}_c$ will be written as

$$\hat{\mathbf{p}}_c \equiv \hat{\mathbf{p}}_c(\mathbf{z}). \tag{4.17}$$

Then,

$$E_c(\mathbf{z}, j) \equiv \hat{\mathbf{e}}_{\hat{\mathbf{p}}_c}(j) \cdot \mathbf{E}_c(\mathbf{z}). \tag{4.18}$$

Finally, the Green's function $(i|\, G_{\alpha\beta}\,|l)$ in Eqs. (4.15) is defined by Eq. (2.33).

The transport equation is now derived just as in I. To describe the coherent intensity, we use the δ function (as in I) $\delta[\hat{\mathbf{p}}, \hat{\mathbf{p}}_c(\mathbf{z})]$, having the property

$$\int f(\hat{\mathbf{p}})\delta[\hat{\mathbf{p}}, \hat{\mathbf{p}}_c(\mathbf{z})]\, d\Omega_{\hat{\mathbf{p}}} = f(\hat{\mathbf{p}}_c), \tag{4.19}$$

where $f(\hat{\mathbf{p}})$ is nonsingular at $\hat{\mathbf{p}} = \hat{\mathbf{p}}_c$. Then, Eq. (I.5.8) is written in the form

$$\mathbf{I}(\mathbf{z}_\alpha, \hat{\mathbf{p}}) = \mathbf{J}_c(\mathbf{z}_\alpha)\delta[\hat{\mathbf{p}}, \hat{\mathbf{p}}_c(\mathbf{z}_\alpha)] + cn_r(\mathbf{z}_\alpha)\int\left(\frac{\partial\Sigma_\beta}{\partial\Omega_\alpha}\right) ds(\mathbf{z}_\beta)\mathbf{U}(\alpha, \beta). \tag{4.20}$$

The first term here represents the intensity of the coherent field at $\mathbf{z}_\alpha$. From Eq. (2.14) we see that

$$(J_c(\mathbf{z}))_{ij} = (c/8\pi)n_r(\mathbf{z})E_c^*(\mathbf{z}, i)E_c(\mathbf{z}, j).$$

The integral in the second term in Eq. (4.20) is taken in the reverse direction along the ray path which passes through $\mathbf{z}_\alpha$ with the local tangent $\hat{\mathbf{p}}$. The notation of Eq. (2.43) has been used, with $\delta\Sigma_\beta$ being the area at $\mathbf{z}_\beta$ of a flux tube of ray paths which pass through $\mathbf{z}_\alpha$ with tangent $\hat{\mathbf{p}}$ and lie within the small solid angle $\delta\Omega_\alpha$.

The quantity $\mathbf{U}(\alpha, \beta)$ is defined by [Eq. (I.5.14)]

$$\mathbf{U}(\alpha, \beta) = \mathbf{G}^*_{\alpha\beta} \otimes \mathbf{G}_{\alpha\beta}\Big([cn_r(\mathbf{z}_\beta)]^{-1}\mathbf{M}(\alpha\beta, \beta 0)\mathbf{J}_c(\mathbf{z}_\beta) + \int d^3z_\sigma \mathbf{M}(\alpha\beta, \beta\sigma)\mathbf{U}(\beta, \sigma)\Big). \tag{4.21}$$

Here $\mathbf{M}$ is defined by Eqs. (1.10), (1.11), and (1.12), and we have used the index pairs $(\alpha\beta)$ and $(\beta\sigma)$ to indicate the direction of the unit vectors $\hat{\mathbf{p}}$ and $\hat{\mathbf{p}}'$ at $\mathbf{z}_\beta$ in Eq. (1.10). The notation $\mathbf{G}^*_{\alpha\beta} \otimes \mathbf{G}_{\alpha\beta}$ is used to denote the open product $(i|\, G_{\alpha\beta}\,|j)^*(l|\, G_{\alpha\beta}\,|s)$.

From Eqs. (4.20) and (4.21) we obtain the transport equation

$$\mathbf{I}(\mathbf{z}, \hat{\mathbf{p}}) = \mathbf{J}_c(\mathbf{z})\delta[\hat{\mathbf{p}}, \hat{\mathbf{p}}_c(\mathbf{z})] + \int\left(\frac{\partial\Sigma_x}{\partial\Omega_z}\right) ds(\mathbf{x})\mathbf{G}^*(\mathbf{z}, \mathbf{x}) \otimes \mathbf{G}(\mathbf{z}, \mathbf{x}) \times \left(\frac{n_r(\mathbf{z})}{n_r(\mathbf{x})}\right)\int d\Omega_{\hat{\mathbf{p}}'}\mathbf{M}(\hat{\mathbf{p}}(\mathbf{x}), \hat{\mathbf{p}}')\mathbf{I}(\mathbf{x}, \hat{\mathbf{p}}'). \tag{4.22}$$

Here $\hat{\mathbf{p}}(\mathbf{x})$ is the tangent at $\mathbf{x}$ of the ray path passing through $\mathbf{z}$ with local tangent $\hat{\mathbf{p}}$.

Equation (4.22) can be simplified using Eqs. (2.33), (2.39), and (2.41). Since

$$\left(\frac{\partial\Sigma_x}{\partial\Omega_z}\right)A^2(\mathbf{z},\mathbf{x}) = \frac{n_r(\mathbf{z})}{n_r(\mathbf{x})},$$

we obtain

$$\mathbf{I}(\mathbf{z},\hat{\mathbf{p}}) = \hat{\mathbf{I}}_c(\mathbf{z})\delta[\hat{\mathbf{p}},\hat{\mathbf{p}}_c(\mathbf{z})] + \int^{\mathbf{z}} ds(\mathbf{x})\left[\exp\left(-\int_{\mathbf{x}}^{\mathbf{z}}\frac{ds}{l}\right)\right]\left(\frac{n_r^2(\mathbf{z})}{n_r^2(\mathbf{x})}\right) \times \mathbf{P}(\mathbf{z},\mathbf{x})\otimes\mathbf{P}(\mathbf{z},\mathbf{x})\int d\Omega_{\hat{\mathbf{p}}'}\mathbf{M}(\hat{\mathbf{p}}(\mathbf{x}),\hat{\mathbf{p}}')\mathbf{I}(\mathbf{x},\hat{\mathbf{p}}'). \tag{4.23}$$

Both path integrals here are taken over the ray passing through $\mathbf{z}$ with local tangent $\hat{\mathbf{p}}$. The quantities $\mathbf{P}$ here are matrices with elements

$$(i|\,P(\mathbf{z}_\alpha,\mathbf{z}_\beta)\,|s) = \hat{\mathbf{e}}_{\hat{\mathbf{p}}_\alpha}(i;\alpha,\beta)\cdot\mathbf{P}_{\alpha\beta}\cdot\hat{\mathbf{e}}_{\hat{\mathbf{p}}_\beta}(s;\alpha,\beta), \tag{4.24}$$

where $\mathbf{P}_{\alpha\beta}$ is the quantity (2.31).

To further simplify Eq. (4.23), let us define the matrix

$$\mathbf{V}(\mathbf{z},\mathbf{x}) \equiv \mathbf{P}(\mathbf{z},\mathbf{x})\otimes\mathbf{P}(\mathbf{z},\mathbf{x}). \tag{4.25}$$

This has matrix elements, for $i, j, s, t = 1, 2$,

$$(ij\,|V(\mathbf{z}_\alpha,\mathbf{z}_\beta)|\,st) = [\hat{\mathbf{e}}_{\hat{\mathbf{p}}_\alpha}(i)\cdot\mathbf{P}_{\alpha\beta}\cdot\hat{\mathbf{e}}_{\hat{\mathbf{p}}_\beta}(s)]\times[\hat{\mathbf{e}}_{\hat{\mathbf{p}}_\alpha}(j)\cdot\mathbf{P}_{\alpha\beta}\cdot\hat{\mathbf{e}}_{\hat{\mathbf{p}}_\beta}(s)]. \tag{4.26}$$

Use of Eqs. (3.10) and (3.11) lets us obtain

$$\frac{d}{ds}(ij|\,V(\mathbf{z},\mathbf{x})\,|st) = (-1)^i\phi(i+1j|\,V\,|st) + (-1)^j\phi(ij+1|\,V\,|st). \tag{4.27a}$$

This is equivalent to

$$\frac{d}{ds}\mathbf{V}(\mathbf{z},\mathbf{x}) = -\mathbf{R}(\mathbf{z})\mathbf{V}(\mathbf{z},\mathbf{x}), \tag{4.27b}$$

where $\mathbf{R}$ is the matrix introduced in Eq. (3.12).

Finally, then, we rewrite (4.23) as

$$\mathbf{I}(\mathbf{z},\hat{\mathbf{p}}) = \mathbf{I}_c(\mathbf{z},\hat{\mathbf{p}}) + \int^{\mathbf{z}} ds(\mathbf{x})\left[\exp\left(-\int_{\mathbf{x}}^{\mathbf{z}}\frac{ds}{l}\right)\right]\left(\frac{n_r^2(\mathbf{z})}{n_r^2(\mathbf{x})}\right) \times \mathbf{V}(\mathbf{z},\mathbf{x})\int d\Omega_{\hat{\mathbf{p}}'}\mathbf{M}(\hat{\mathbf{p}}(\mathbf{x}),\hat{\mathbf{p}}')\mathbf{I}(\mathbf{x},\hat{\mathbf{p}}'). \tag{4.28}$$

The coherent intensity in Eq. (4.23) has here been abbreviated by writing

$$\mathbf{I}_c(\mathbf{z},\hat{\mathbf{p}}) = \mathbf{J}_c(\mathbf{z})\delta[\hat{\mathbf{p}},\hat{\mathbf{p}}_c(\mathbf{z})]. \tag{4.29}$$

More generally, when there are several sources or *incoherent* sources, we may interpret $\mathbf{I}_c$ in Eq. (4.28) as the intensity of unscattered radiation.

Equation (4.23) represents the integral form of Eq. (3.14). Indeed, on differentiating (4.28) along the ray path with tangent $\hat{\mathbf{p}}$, we obtain the differential form (3.14). [Comparison of Eqs. (3.9) and (4.28) makes this obvious.]

5. DISCUSSION OF THE COHERENT INTENSITY

Let us first suppose that scattering can be neglected so that $\mathbf{B} = 0$ in Eq. (3.14). In this case we have only the coherent intensity $\mathbf{I}_c$ of Eq. (4.28). This satisfies the differential equation

$$\frac{d}{ds}\mathbf{I}_c(\mathbf{z},\hat{\mathbf{p}}) + \left(\frac{1}{L}+\mathbf{R}\right)\mathbf{I}_c(\mathbf{z},\hat{\mathbf{p}}) = 0. \tag{5.1}$$

We consider the case of a single coherent source, so that through any point $\mathbf{z}$ there passes only one ray path (with the exception of possible singular points).

In terms of the Stokes parameters [see Eq. (3.17)]

$$\begin{aligned} I &\equiv I_1 + I_2,\\ Q &\equiv I_1 - I_2,\\ U &= I_3,\\ V &= I_4, \end{aligned} \tag{5.2}$$

Eq. (5.1) becomes (in the remainder of this section we drop the subscript "c" from $\mathbf{I}_c$)

$$\frac{dI}{ds} + \frac{1}{L}I = 0, \tag{5.3}$$

$$\frac{dV}{ds} + \frac{1}{L}V = 0, \tag{5.4}$$

$$\frac{d\Psi}{ds} + \frac{1}{L}\Psi + \mathbf{r}\Psi = 0. \tag{5.5}$$

Here we have written

$$\Psi \equiv \begin{pmatrix} Q \\ U \end{pmatrix}, \tag{5.6}$$

$$\mathbf{r} \equiv 2i\phi\sigma_y \equiv 2\phi\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{5.7}$$

If $\mathbf{z}_2$ and $\mathbf{z}_1$ are two points on a given ray path, we may integrate Eq. (5.3):

$$I(\mathbf{z}_2,\hat{\mathbf{p}}_2) = \left[\exp\left(-\int_{\mathbf{z}_1}^{\mathbf{z}_2}\frac{ds}{l}\right)\right]\left(\frac{n_r^2(2)}{n_r^2(1)}\right)I(\mathbf{z}_1,\hat{\mathbf{p}}_1) \equiv T(2,1)I(\mathbf{z}_1,\hat{\mathbf{p}}_1). \tag{5.8}$$

In the second writing here we have introduced $T(2, 1)$ for the function which translates I at $\mathbf{z}_1$ to the point $\mathbf{z}_2$.

From Eq. (5.4) we evidently obtain

$$V(\mathbf{z}_2,\hat{\mathbf{p}}_2) = T(2,1)V(\mathbf{z}_1,\hat{\mathbf{p}}_1). \tag{5.9}$$

To integrate Eq. (5.5), we first define

$$\Phi(2, 1) \equiv 2\int_{z_1}^{z_2} \phi \, ds. \tag{5.10}$$

Then the integrated form of (5.5) is

$$\begin{aligned} Q(\mathbf{z}_2, \hat{\mathbf{p}}_2) &= T(2, 1)[\cos \Phi(2, 1)Q(\mathbf{z}_1, \hat{\mathbf{p}}_1) \\ &\quad + \sin \Phi(2, 1)U(\mathbf{z}_1, \hat{\mathbf{p}}_1)], \\ U(\mathbf{z}_2, \hat{\mathbf{p}}_2) &= T(2, 1)[\cos \Phi(2, 1)U(\mathbf{z}_1, \hat{\mathbf{p}}_1) \\ &\quad - \sin \Phi(2, 1)Q(\mathbf{z}_1, \hat{\mathbf{p}}_1)]. \end{aligned} \tag{5.11}$$

The above results let us write the solution to Eq. (5.1) in the form

$$\mathbf{I}_c(\mathbf{z}_2, \hat{\mathbf{p}}_2) = T(\mathbf{z}_2, \mathbf{z}_1)\mathbf{V}(\mathbf{z}_2, \mathbf{z}_1)\mathbf{I}_c(\mathbf{z}_1, \hat{\mathbf{p}}_1). \tag{5.12}$$

Here $T(\mathbf{z}_2, \mathbf{z}_1) \equiv T(2, 1)$ is defined by Eq. (5.8) and $\mathbf{V}$ by Eq. (4.27). From Eq. (5.11) we can extract the matrix elements of $\mathbf{V}$. These are given in Table I for the I, Q, U, and V representation and in Table II for the I_1, I_2, I_3, and I_4 representation.

The energy flux vector at a point $\mathbf{z}_1$ is

$$\mathbf{J}(1) = \int \mathbf{I}(\mathbf{z}_1, \hat{\mathbf{p}}) \, d\Omega_{\hat{\mathbf{p}}}. \tag{5.13}$$

Referring to Fig. 5, we construct a tube of ray paths passing through $\mathbf{z}_1$ with solid angle $\delta\Omega_1$. The tube has a cross section of area $\delta\Sigma_0$ at some fixed reference point $\mathbf{z}_0$. Differentiation of $\mathbf{J}/n_r$ along the ray path then leads to the equation

$$\begin{aligned} \frac{d}{ds}\left(\frac{\mathbf{J}(1)}{n_r(1)}\right) &= \frac{d}{ds}\int \frac{\mathbf{I}}{n_r} \, d\Omega_1 \\ &= \frac{d}{ds}\int \frac{\mathbf{I}}{n_r}\left(\frac{\partial\Sigma_0}{d\Omega_1}\right)^{-1} d\Sigma_0. \end{aligned} \tag{5.14}$$

Now, according to Eq. (2.43),

$$\begin{aligned} \left(\frac{\partial\Sigma_0}{\partial\Omega_1}\right) &= A_{01}^2 \frac{n_r(0)}{n_r(1)} \\ &= A_{10}^2 \frac{n_r(0)}{n_r(1)}. \end{aligned} \tag{5.15}$$

TABLE I. Matrix elements of $\mathbf{V}(\mathbf{z}_2, \mathbf{z}_1)$ in I, Q, U, and V representation. Here $\Phi \equiv \Phi(1, 2)$ in the notation of Eq. (5.10).

	$I(1)$	$Q(1)$	$U(1)$	$V(1)$
$I(2)$	1	0	0	0
$Q(2)$	0	$\cos\Phi$	$\sin\Phi$	0
$U(2)$	0	$-\sin\Phi$	$\cos\Phi$	0
$V(2)$	0	0	0	1

TABLE II. Matrix elements of $\mathbf{V}(\mathbf{z}_2, \mathbf{z}_1)$ in I_1, I_2, I_3, and I_4 representation. Here $\Phi \equiv \Phi(1, 2)$ in the notation of Eq. (5.10).

	$I_1(1)$	$I_2(1)$	$I_3(1)$	$I_4(1)$
$I_1(2)$	$\frac{1}{2}(1 + \cos\Phi)$	$\frac{1}{2}(1 - \cos\Phi)$	$\frac{1}{2}\sin\Phi$	0
$I_2(2)$	$\frac{1}{2}(1 - \cos\Phi)$	$\frac{1}{2}(1 + \cos\Phi)$	$-\frac{1}{2}\sin\Phi$	0
$I_3(2)$	$-\sin\Phi$	$\sin\Phi$	$\cos\Phi$	0
$I_4(2)$	0	0	0	1

Substitution of this into (5.14) and use of Eqs. (5.3) and (2.26) leads to the differential equation

$$\frac{d}{ds}\left(\frac{\mathbf{J}(\mathbf{z})}{n_r(\mathbf{z})}\right) = -\left(\frac{1}{l} + \frac{\nabla^2 S_r}{n_r}\right)\left(\frac{\mathbf{J}}{n_r}\right) - \mathbf{R}\left(\frac{\mathbf{J}}{n_r}\right). \tag{5.16}$$

The scalar flux

$$J(\mathbf{z}) = \int I(\mathbf{z}, \hat{\mathbf{p}}) \, d\Omega_{\hat{\mathbf{p}}} \tag{5.17}$$

is seen from (5.16) to satisfy the equation

$$\frac{d}{ds}\left(\frac{J}{n_r}\right) = -\left(\frac{1}{l} + \frac{\nabla^2 S_r}{n_r}\right)\left(\frac{J}{n_r}\right). \tag{5.18}$$

This agrees with Eq. (37), p. 116, of Born and Wolf.[7,15]

A slightly different version of the above discussion can be given as follows. We first integrate both sides of Eq. (5.8) over the solid angles of $\hat{\mathbf{p}}_1$. There results

$$\int I(\mathbf{z}_2, \hat{\mathbf{p}}_2) \, d\Omega_1 = T(2, 1)J(1). \tag{5.19}$$

Now,

$$\begin{aligned} d\Omega_1 &= d\Omega_2\left(\frac{\partial\Sigma_0}{\partial\Omega_2}\right)\left(\frac{\partial\Sigma_0}{\partial\Omega_1}\right)^{-1} \\ &= d\Omega_2\left(\frac{A_{01}^2}{A_{02}^2}\right)\left(\frac{n_r(2)}{n_r(1)}\right). \end{aligned}$$

From Eq. (2.26) we obtain

$$\frac{A_{01}^2}{A_{02}^2} = \exp\left(\int_{z_1}^{z_2} \frac{\nabla^2 S_r}{n_r} \, ds\right). \tag{5.20}$$

Substitution into Eq. (5.19) leads to

$$J(2) = \left(\frac{n_r(2)}{n_r(1)}\right)\left\{\exp\left[-\int_{z_1}^{z_2}\left(\frac{1}{l} + \frac{\nabla^2 S_r}{n_r}\right) ds\right]\right\}J(1), \tag{5.21}$$

in agreement with the similar result of Ref. 7 and an obvious integral of Eq. (5.18).

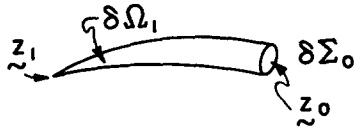


FIG. 5. Illustration of the flux tube on which Eq. (5.14) describes the intensity.

The result (5.21) permits us to integrate the set of Eqs. (5.16) in the form

$$\mathbf{J}(\mathbf{z}) = J(\mathbf{z})\mathbf{V}(\mathbf{z}, \mathbf{z}_1)\hat{\mathbf{J}}. \tag{5.22}$$

Here $J(\mathbf{z})$ is specified by Eq. (5.21), $\mathbf{V}$ by Eq. (4.27), and $\hat{\mathbf{J}}$ is a constant column matrix specifying the polarization at $\mathbf{z}_1$:

$$\hat{\mathbf{J}} = \begin{pmatrix} \hat{J}_1 \\ \hat{J}_2 \\ \hat{J}_3 \\ \hat{J}_4 \end{pmatrix}. \tag{5.23}$$

Since $J(\mathbf{z})$ is the total flux, we must have

$$\hat{J}_1 + \hat{J}_2 = 1 \tag{5.24}$$

in Eq. (5.23).

6. RADAR BACKSCATTER IN THE DWBA

In this section we discuss radar backscatter in the distorted wave Born approximation (DWBA). This is illustrated in Fig. 6, where the transmitting and receiving antenna is located a point $\mathbf{y}$ very far from a plasma scatterer. In the DWBA only a single scattering is assumed to occur. This means that the flux of energy may be obtained from Eq. (4.22) as (assuming no coherent backscatter)

$$\begin{aligned}\mathbf{J}(\mathbf{y}) &= \int \mathbf{I}(\mathbf{y}, \hat{\mathbf{p}})\, d\Omega_{\hat{\mathbf{p}}} \\ &= \int d^3x \mathbf{G}^*(\mathbf{y}, \mathbf{x}) \\ &\quad \otimes \mathbf{G}(\mathbf{y}, \mathbf{x}) \frac{n_r(\mathbf{y})}{n_r(\mathbf{x})} \int d\Omega_{\hat{\mathbf{p}}'} \mathbf{M}(\hat{\mathbf{p}}(\mathbf{x}), \hat{\mathbf{p}}')\mathbf{I}_c(\mathbf{x}, \hat{\mathbf{p}}').\end{aligned} \tag{6.1}$$

Since

$$\mathbf{I}_c(\mathbf{x}, \hat{\mathbf{p}}') = \mathbf{J}_c(\mathbf{x})\delta(\mathbf{p}', \hat{\mathbf{p}}_c(x)),$$

this may be rewritten in the form

$$\mathbf{J}(\mathbf{y}) = \int d^3x \exp\left(-\int_{\mathbf{x}}^{\mathbf{y}} \frac{ds}{l}\right) \mathbf{V}(\mathbf{y}, \mathbf{x})\mathbf{M}(-\hat{\mathbf{p}}_c(\mathbf{x}), \hat{\mathbf{p}}_c(\mathbf{x})) \times \frac{n_r(\mathbf{y})}{n_r(\mathbf{x})} A^2(\mathbf{y}, \mathbf{x})\mathbf{J}_c(\mathbf{x}). \tag{6.2}$$

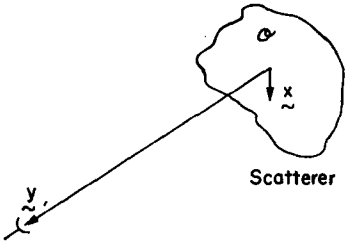


FIG. 6. Backscatter, using the same antenna for transmitting and receiving.

Now, from Eqs. (1.10) and (1.11), we see that

$$\mathbf{M}(-\hat{\mathbf{p}}_c(\mathbf{x}), \hat{\mathbf{p}}_c(\mathbf{x})) = \sigma_g(-1)\mathfrak{I}, \tag{6.3}$$

where

$$(ij|\,\mathfrak{I}\,|st) = (-1)^{i+j}\delta_{is}\delta_{jt} \tag{6.4}$$

in the representation of Eq. (1.11).

We shall assume the incident radiation to be a plane electromagnetic wave. The quantity $\mathbf{J}_c(\mathbf{x})$ has, according to Eq. (5.22), the form

$$\mathbf{J}_c(\mathbf{x}) = J_c(\mathbf{x})\mathbf{V}(\mathbf{x}, \mathbf{y})\hat{\mathbf{J}}, \tag{6.5}$$

where $\hat{\mathbf{J}}$ has the form (5.23) and specifies the polarization of the incident radiation. As the point $\mathbf{x}$ moves outside the plasma along an incident ray path, $J_c(\mathbf{x})$ assumes the *constant* value $J_c(\mathbf{y})$.

Since $\mathbf{y}$ lies outside the plasma, $n_r(\mathbf{y}) = 1$. When $\mathbf{x}$ and $\mathbf{y}$ are both outside the plasma,

$$A(\mathbf{y}, \mathbf{x}) = |\mathbf{x} - \mathbf{y}|^{-1}.$$

Thus,

$$A^2(\mathbf{x}, \mathbf{y})y^2 \to 1, \quad x \ll y,$$

as $\mathbf{x}$ moves outside the plasma along an incident ray path. Therefore, we may write

$$J_c(\mathbf{y})A^2(\mathbf{x}, \mathbf{y})y^2 \exp\left(-\int_{\mathbf{y}}^{\mathbf{x}} \frac{ds}{l}\right) = \frac{J_c(\mathbf{x})}{n_r(\mathbf{x})}. \tag{6.6}$$

That is, the quantity on the left satisfies the differential equation (5.18) and the appropriate boundary condition on an incident ray path.

The above results let us write Eq. (6.2) in the form

$$\mathbf{J}(\mathbf{y}) = (y^{-2})\int d^3x \sigma_g(-1)J_c^2(\mathbf{x})[n_r^2(\mathbf{x})J_c(\mathbf{y})]^{-1} \times [\mathbf{V}(\mathbf{y}, \mathbf{x})\mathfrak{I}\mathbf{V}(\mathbf{x}, \mathbf{y})\hat{\mathbf{J}}]. \tag{6.7}$$

In the usual Born approximation, $\mathbf{V} = \mathbf{1}$, the unit matrix, $n_r(\mathbf{x}) = 1$, and $J_c(\mathbf{x}) = J_c(\mathbf{y})$. Thus, in the Born approximation,

$$\mathbf{J}(\mathbf{y}) = [J_c(\mathbf{y})\mathfrak{I}\hat{\mathbf{J}}/y^2]\int d^3x \sigma_g(-1). \tag{6.8}$$

To further simplify Eq. (6.7), let us use the I, Q, U, and V representation of Eq. (5.2). In this representation,

$$\mathfrak{I}\begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = \begin{pmatrix} I \\ Q \\ -U \\ -V \end{pmatrix}. \tag{6.9}$$

With the grouping of terms implied by Eqs. (5.3),

(5.4), and (5.5), we may write

$$\mathbf{I} \equiv \begin{pmatrix} I \\ V \end{pmatrix} \otimes \begin{pmatrix} Q \\ U \end{pmatrix},$$
$$\mathfrak{I} = \boldsymbol{\sigma}_z \otimes \boldsymbol{\sigma}_z, \tag{6.10}$$
$$\mathbf{R} = 2i\phi\mathbf{1} \otimes \boldsymbol{\sigma}_y,$$

where

$$\boldsymbol{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Also,

$$\mathbf{V}(2, 1) = \exp\,[-i\Phi(2, 1)\mathbf{1} \otimes \boldsymbol{\sigma}_y], \tag{6.11}$$

where Φ is defined by Eq. (5.10). We see that $\mathfrak{I}^2 = \mathbf{1} \otimes \mathbf{1}$ and

$$\mathfrak{I}\mathbf{R}\mathfrak{I} = -\mathbf{R}.$$

Thus,

$$\mathfrak{I}\mathbf{V}(\mathbf{y}, \mathbf{x})\mathfrak{I} = [\mathbf{V}(\mathbf{x}, \mathbf{y})]^{-1}, \tag{6.12}$$

and we may rewrite Eq. (6.7) as

$$\mathbf{J}(\mathbf{y}) = (\mathfrak{I}\hat{\mathbf{J}}/y^2)\int d^3x\sigma_s(-1)J_c^2(\mathbf{x})[n_r^2(\mathbf{x})J_c(\mathbf{y})]^{-1}. \tag{6.13}$$

The lack of cross polarization on backscatter, characteristic of the Born approximation, is therefore also found in the DWBA. We emphasize, of course, that we have shown this only in the eikonal form of the DWBA and also when the approximation (1.9) is valid for the pair correlation function.

7. THE DIFFUSION APPROXIMATION

When the cross section $\sigma_s(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}')$ is sufficiently strongly peaked in the forward direction, the scattering function $\mathbf{B}$ can be simplified by making the diffusion approximation. The appropriate form was derived in II. We quote this here.

Using the representation (3.17) and polar coordinates for $\hat{\mathbf{p}}$ [that is, writing $\mathbf{I} = \mathbf{I}(\mathbf{x}, \theta, \phi)$] with $\hat{\mathbf{k}}$ as polar axis, we have

$$\mathbf{B}(\mathbf{x}, \theta, \phi) = 2\sigma_t\mathbf{P}\left((\sin\theta)^{-1}\frac{\partial\mathbf{I}}{\partial\phi}\right) + \sigma_t\frac{\partial}{\partial\mu}\left((1-\mu^2)\frac{\partial\mathbf{I}}{\partial\mu}\right) + \sigma_t(1-\mu^2)^{-1}\frac{\partial^2\mathbf{I}}{\partial\phi^2}. \tag{7.1}$$

Here

$$\sigma_t = \tfrac{1}{2}\int d\Omega_{\hat{\mathbf{p}}'}(1 - \hat{\mathbf{p}} \cdot \hat{\mathbf{p}}')\sigma_s(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}'), \tag{7.2}$$

and $\mathbf{P}$ is the matrix with elements

$$P_{23} = -P_{13} = \cos\theta, \qquad P_{31} = -P_{32} = 2\cos\theta, \tag{7.3}$$

all others vanishing.

* This work was supported in part by the U.S. Atomic Energy Commission and in part by a Grant from the Air Force Office of Scientific Research.

[1] K. M. Watson, J. Math. Phys. **10**, 688 (1969). This paper will henceforth be referred to as I.

[2] K. M. Watson, "Electromagnetic Wave Scattering within a Plasma in the Transport Approximation," Phys. Fluids **13**, 2514 (1970). This paper will henceforth be referred to as II.

[3] J. L. Peacher and K. M. Watson, J. Math. Phys. **11**, 1496 (1970). This paper will henceforth be referred to as III.

[4] The derivations given in I and III are an adaptation to Maxwell's equations of a general derivation for quantum systems given by K. M. Watson, Phys. Rev. **118**, 886 (1960).

[5] To our knowledge, the first published paper deriving a transport equation from a wave equation is that of L. L. Foldy, Phys. Rev. **67**, 107 (1945). Foldy assumed uncorrelated point scatterers. A transport theory for scalar waves was published by Yu. N. Barabanenkov and V. N. Finkel'berg, Zh. Eksp. Teor. Fiz **53**, 978 (1967) [Sov. Phys. JETP **26**, 587 (1968)]. These authors began with a Bethe–Salpeter equation for the product of two wave amplitudes. A different derivation was given by P. E. Stott, J. Phys. A **1**, 675 (1968). Stott used the Wigner transformation to transform the field autocorrelation function (density matrix) into a transport equation.

[6] See, for example, S. Chandrasekhar, *Radiative Transfer* (Oxford U.P., London, 1950), or V. V. Sobolev, *A Treatise on Radiative Transfer* (Van Nostrand, Princeton, N.J., 1963).

[7] M. Born and E. Wolf, *Principles of Optics* (Pergamon, New York, 1959), Chap. 3.

[8] J. Van Bladel, IRE Trans. Antennas Propagation **9**, 563 (1961). To use (2.5a) in Eq. (2.3), a spherical cavity of very small radius is excluded at $\mathbf{x} = \mathbf{r}$ for the $\nabla_x\nabla_x$ term.

[9] Born and Wolf, Ref. 7, considered only the case of a real refractive index; that is, they took $n_i = S_i = 0$.

[10] This is proven, for example, in P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), p. 873.

[11] A derivation, showing that no approximation need be made in obtaining Eq. (3.3), was given by N. Francis and K. M. Watson, Phys. Rev. **92**, 291 (1953) (see, in particular, the appendix to this paper).

[12] The reader who is not interested in following this derivation may omit Sec. 4. The principal result is the integral form of Eq. (3.14) given by Eq. (4.23) or (4.28).

[13] The corresponding quantum mechanical derivation was given in Ref. 4.

[14] Equations of I will be referred to as Eq. (I.3.1), etc.

[15] Since the refractive index was assumed to be real in Ref. 7, the l^{-1} term of Eq. (5.5) did not occur there.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Structure of the Point Spectrum of Schrödinger-Type Tridiagonal Operators

EVANGELOS K. IFANTIS
"Democritos" Nuclear Research Center, Aghia Paraskevi, Attikis, Athens, Greece

(Received 11 February 1970)

The main result in the present work concerns a criterion on the existence and the structure of proper values in a class of bounded operators (Schrödinger-type tridiagonal operators) on an abstract separable Hilbert space. The realization of these operators in the space of square summable sequences $l_2(1, \infty)$ represents a boundary-value problem of difference equations of the following form: $f(n+1) + f(n-1) + a(n)f(n) = Ef(n)$. In our case $f(n) \in l_2(1, \infty)$, and the condition $f(2) + a(1)f(1) = Ef(1)$ must hold. The approach followed is based on the reduction by Deliyannis and Ifantis [J. Math. Phys. **10**, 421 (1969)] of the above boundary-value problem to an abstract operator form, which makes possible the application of the methods of functional analysis. It is shown that for every monotonically convergent and real-valued sequence $a(n) \neq 0$, $n = 1, 2, \cdots$, there exist proper values, the greatest of which can be determined by the Ritz approximation method.

I. INTRODUCTION

One of the most important approximation methods for the evaluation of proper values is the well-known Ritz method for the determination of the greatest proper value of a bounded self-adjoint operator in Hilbert space. A well-known operator class, in which the Ritz method has been applied in the past, is the class of positive completely continuous operators. The reason for this successful application is the fact that in the aforementioned operator class the existence of proper values is well known and the value of the norm is the greatest proper value.

Another class of self-adjoint operators in which the Ritz method can be applied is the class of operators of the form

$$T = T_0 + A, \tag{1}$$

where T_0 is a bounded self-adjoint with continuous spectrum and A completely continuous and self-adjoint. Moreover,

$$\|T\| > \|T_0\| \tag{2}$$

must hold. This is because, due to the Weyl theorem,[1] the operators T_0 and T have the same essential spectrum. Therefore, if (2) holds, the extension of the spectrum of T is due to its point spectrum, i.e., the value $\|T\|$ is the absolutely greatest proper value of T.

A large class of Schrödinger-type tridiagonal operators (Sec. II) is of the form (1), and it is useful to know when (2) holds. This is due to the fact that in the case of the Schrödinger-type tridiagonal operators the Ritz method can be easily applied (Appendix A). However, before its application, the existence problem must be solved. Concerning the existence problem and the structure of the point spectrum of the Schrödinger-type self-adjoint tridiagonal operators, an important criterion is derived in the present paper.

In Secs. II and III we simply give the definition of the Schrödinger-type tridiagonal operators and state some simple propositions, which we use later.

In Sec. IV a representation of the Hardy–Lebesgue space by means of the shift operator is given. Throughout the paper, we often use certain results which follow with the help of the aforementioned representation. Besides, it is of special interest since it connects the proper value problem of tridiagonal operators with problems of analytic functions in the unit disc.

In Sec. V we derive a criterion for the existence and position of the point spectrum for a large class of bounded self-adjoint tridiagonal operators of the Schrödinger type. In Sec. VI two theorems are proved concerning the structure of the spectrum.

II. THE SCHRÖDINGER-TYPE TRIDIAGONAL OPERATORS

Let V be the shift operator on a separable Hilbert space $\mathcal{H}$, over the complex field $\mathbb{C}$, with an orthonormal basis $\{e_n\}_1^\infty$, $Ve_n = e_{n+1}$. V is an isometry and its adjoint V^* a partial isometry. We call every operator T of the form $T = T_0 + A$, where $T_0 = V + V^*$ and A is a diagonal operator, defined as

$$Ae_n = a(n)e_n, \quad n = 1, 2, \cdots, \quad a(n) \in \mathbb{C},$$

a tridiagonal operator of Schrödinger type. The operator $T_0 = V + V^*$ is self-adjoint with purely continuous spectrum (Appendix C) covering the closed interval $[-2, +2]$. In case $a(n)$ is real and $\lim a(n) = a \neq \infty$, as $n \to \infty$, we may assume A to be completely continuous and self-adjoint without restriction of the generality. Then, we have the following theorem.[2]

Theorem 1: The operator T cannot have a discrete spectrum $\{E_n\}$ such that $\lim E_n$ as $n \to \infty$ exists; i.e., it cannot have a pure point spectrum with a single limit point. In case $\|T\| > 2$, the point spectrum is not empty.

In the case of the Schrödinger-type tridiagonal operators the Ritz approximation method can be easily applied. In Appendix A two examples are given, in which we can compare the approximate values found with the exact ones.

III. SOME BASIC PROPOSITIONS

We list below the propositions concerning V^* and $T = V + V^* + A$ which we shall use later.

Proposition 1: Every point z in the interior of the unit disc in $\mathbb{C}$ belongs to the point spectrum of V^*, and the set of proper elements

$$f_z = \sum_{n=1}^{\infty} z^{n-1} e_n, \quad f_0 = e_1,$$

forms a complete system in $\mathcal{H}$ in the sense that, if f is orthogonal to f_z for every $|z| < 1$, then $f = 0$.

The first statement is obvious. The second follows from the fact that, if

$$(f_z, f) = \sum_{n=1}^{\infty} (f, e_n) z^n = 0$$

for every $|z| < 1$, then $(f, e_n) = 0$ for every n, i.e., $f = 0$.

Proposition 2: If $f \neq 0$ is a proper element of T, then $(f, e_1) \neq 0$ because, if $(f, e_1) = 0$, then we have

$$(Vf, e_1) + (V^*f, e_1) + (Af, e_1) = 0,$$

i.e., $(f, e_2) = 0$ and, consequently,

$$(f, e_3) = 0 \cdots \quad \text{and} \quad f = \sum_{n=1}^{\infty} (f, e_n) e_n = 0.$$

Proposition 3: For a Schrödinger-type tridiagonal operator the proper elements are uniquely determined if the proper values are known. That means that, if the proper values are known, the problem of finding the proper elements is straightforward.

This follows from the previous proposition because, if we normalize the proper element f by taking $(f, e_1) = 1$, then we can uniquely determine the other components $(f, e_2), (f, e_3), \cdots$.

IV. THE REALIZATION OF THE SCHRÖDINGER-TYPE TRIDIAGONAL OPERATORS IN THE HARDY-LEBESGUE SPACE

Proposition 1 allows us to establish a one-to-one correspondence between the elements $f \in \mathcal{H}$ and the analytic functions

$$\phi(z) = (f_z, f) = \sum_{n=1}^{\infty} (e_n, f) z^{n-1}, \quad |z| < 1,$$

with the additional property

$$\sum_{n=1}^{\infty} |(e_n, f)|^2 < \infty.$$

We can make the following remarks: (a) If the element f corresponds to the analytic function $\phi(z)$, i.e., if $f \to \phi(z)$, then

$$\begin{aligned} Vf &\to z\phi(z), \\ V^*f &\to z^{-1}[\phi(z) - \phi(0)] \end{aligned} \tag{3}$$

because

$$(f_z, Vf) = (V^*f_z, f) = z(f_z, f) = z\phi(z)$$

and

$$\begin{aligned} (f_z, V^*f) &= (Vf_z, f) = z^{-1}(f_z - e_1, f) \\ &= z^{-1}[(f_z, f) - (e_1, f)] = z^{-1}[\phi(z) - \phi(0)]. \end{aligned}$$

(b) For Schrödinger-type tridiagonal operators, we have $\phi(0) = (e_1, f) \neq 0$, in case that f is a proper value.

This follows from Proposition 2 of the previous section. Now let

$$\phi(z) = (f_z, f) \text{ and } (f_z, Tf) = F[\phi(z)];$$

then

$$(f_z, Tf - Ef) = F[\phi(z)] - \bar{E}\phi(z). \tag{4}$$

We observe from (4) that the problem of finding the proper values E of the operator T is equivalent to the problem of finding the values of the parameter E, for which the solutions of equation

$$F[\phi(z)] - \bar{E}\phi(z) = 0 \tag{5}$$

are analytic within the unit disc and fulfill the condition

$$\sum_{n=1}^{\infty} |a(n)|^2 < \infty, \tag{6}$$

where

$$\begin{aligned} f(z) &= \sum_{n=1}^{\infty} a(n) z^{n-1}, \\ a(n) &= \phi^{(n-1)}(0)/(n-1)!. \end{aligned}$$

Let

$$f_1 = \sum_{n=1}^{\infty} a(n) e_n, \quad f_2 = \sum_{n=1}^{\infty} b(n) e_n.$$

The analytic functions in the unit disc which correspond to the elements $f_1, f_2 \in \mathcal{H}$ are

$$\phi_1(z) = \sum_{n=1}^{\infty} \overline{a(n)} z^{n-1}, \quad \phi_2(z) = \sum_{n=1}^{\infty} \overline{b(n)} z^{n-1}.$$

The set of analytic functions within the unit disc with condition (6) obviously forms a linear space $\mathcal{H}_2$ with the usual addition and multiplication by scalars. If we define an inner product in $\mathcal{H}_2$ as

$$(\phi_1(z), \phi_2(z))_{\mathcal{H}_2} = \sum_{n=1}^{\infty} \overline{a(n)} b(n) = \overline{(f_1, f_2)}_{\mathcal{H}},$$

then $\mathcal{H}_2$ becomes a separable Hilbert space with the functions $\rho_n(z) = z^{n-1}$, $n = 1, 2, \cdots$, forming a complete orthonormal system corresponding to the basis $\{e_n\}_1^\infty$ in $\mathcal{H}$.

$\mathcal{H}_2$ is the so called Hardy–Lebesgue space, and condition (6) is equivalent[3] to the condition

$$\sup_{0<r<1} \int_0^{2\pi} |\phi(re^{i\theta})|^2 \, d\theta < \infty, \quad z = re^{i\theta}. \tag{7}$$

The proper value problem of tridiagonal operators in $\mathcal{H}$ is therefore represented as a proper value problem of operators specified by (5) in space $\mathcal{H}_2$. In the following, we derive some results which we shall use later.

We define the operator A_k as

$$A_k f = \sum_{n=1}^{k} a(n)(f, e_n)e_n, \quad f \in \mathcal{H}, \quad a(n) = \text{real numbers},$$

and consider it as perturbation of the operator $T_0 = V + V^*$. If $\phi(z) = (f_z, f)$, then the analytic function which corresponds to the element $A_k f$ is

$$\sum_{n=1}^{k} \frac{a(n)\phi^{(n-1)}(0)}{(n-1)!} z^{n-1}.$$

Thus, by virtue of (3), Eq. (5) has the form

$$z\phi(z) + z^{-1}[\phi(z) - \phi(0)] + \sum_{n=1}^{k} \frac{a(n)\phi^{(n-1)}(0)}{(n-1)!} z^{n-1} - E\phi(z) = 0,$$

i.e.,

$$\phi(z) = \phi(0) \sum_{n=1}^{k} P_n(E) z^n (z^2 - Ez + 1)^{-1}$$

or

$$\phi(z) = \phi(0)(z^2 - Ez + 1)^{-1}\Pi_k(E, z), \quad \phi(0) \neq 0, \tag{8}$$

where the $\Pi_k(E, z)$ are polynomials in z of degree k and the $P_n(E)$ are polynomials in E of degree $n - 1$. Due to the fact that A_k is nuclear and self-adjoint, the values $|E| \leq 2$ belong[1] to the continuous spectrum of $T_0 + A_k$. For $|E| > 2$ the real roots z_1 and z_2 of the equation $z^2 - Ez + 1 = 0$ are different and, by virtue of $z_1 z_2 = 1$, one of them, say z_1, is contained in the unit disc. Thus, in order for $\phi(z)$ to be analytic in the unit disc, we must have $\Pi_k(E, z_1) = 0$ or, because $z_1^2 - Ez_1 + 1 = 0$,

$$\Pi_k[(z_1^2 + 1)/z_1, z_1] = 0. \tag{9a}$$

Let $z_i(E) \neq 0$ be a root of the polynomial $\Pi_k(E, z)$, i.e.,

$$\Pi_k(E, z) = [z - z_i(E)]\Pi_{k-1}(E, z).$$

Then the values of E which satisfy the relations

$$|z_i(E)| < 1, \qquad z_i^2(E) - Ez_i(E) + 1 = 0 \tag{9b}$$

are proper values of the operator $T_0 + A_k$. In fact, from (8) we have

$$\begin{aligned}\phi(z) &= \phi(0)(z^2 - Ez + 1)^{-1}[z - z_i(E)]\Pi_{k-1}(E, z) \\ &= \phi(0)[z - 1/z_i(E)]^{-1}\Pi_{k-1}(E, z) \\ &= -\phi(0)\Pi_{k-1}(E, z)z_i(E)[1 - z_i(E)z]^{-1}\end{aligned}$$

and, because $|z_i(E)| < 1$,

$$\phi(z) = -\phi(0)\Pi_{k-1}(E, z)z_i(E) \sum_{n=1}^{\infty} [z_i(E)z]^{n-1} \in \mathcal{H}_2.$$

{Note that from the last of the conditions (9b) it follows that

$$E \pm 2 = [z_i(E) \pm 1]^2/z_i(E) \lesseqgtr 0, \quad z_i(E) \lesseqgtr 0,$$

and from this that $|E| > 2$.}

For the operator A_1,

$$A_1 f = a(1)(f, e_1)e_1, \quad f \in \mathcal{H},$$

relations (9b) give

$$|z_i| = |1/a(1)| < 1 \quad \text{or} \quad |a(1)| > 1$$

and

$$1/a^2(1) - E/a(1) + 1 = 0 \quad \text{or} \quad E = a(1) + 1/a(1).$$

The normalized proper element is

$$f_0 = \frac{(a^2(1) - 1)^{\frac{1}{2}}}{a(1)} \sum_{n=1}^{\infty} \left(\frac{1}{a(1)}\right)^{n-1} e_n.$$

For $|a(1)| \leq 1$ the point spectrum of $T_0 + A_1$ is empty.

An interesting result which follows directly from (9a) is the following:

Proposition 4: For $k < \infty$ there can exist only one finite set of proper values of $T_0 + A_k$.

V. THE EXISTENCE CRITERION

Let B be defined as $Be_n = b(n)e_n$, $n = 1, 2, \cdots$, and assume that there exists a proper value E, corresponding to the normalized proper element f of the operator $T_0 + B$, i.e.,

$$(T_0 + B)f = Ef \quad \text{with} \quad f = \sum_{m=1}^{\infty} c(n)e_n$$

and

$$\|f\|^2 = \sum_{n=1}^{\infty} |c(n)|^2 = 1. \tag{10}$$

Then we have

$$\begin{aligned}\|(T_0 + A)f\|^2 &= \|(T_0 + B)f + (A - B)f\|^2 \\ &= \|Ef + (A - B)f\|^2 \\ &= \sum_{n=1}^{\infty} |c_n|^2 |E + a(n) - b(n)|^2,\end{aligned}$$

and from this, in view of Theorem 1, it follows that, if

$$\sum_{n=1}^{\infty} |c_n|^2 |E + a(n) - b(n)|^2 > 4, \tag{11}$$

then the point spectrum of $T_0 + A$ is not empty.

Because of (10), relation (11) is satisfied if

$$|E + a(n) - b(n)| > 2 \tag{12}$$

holds for at least some n and

$$|E + a(n) - b(n)| \geq 2, \tag{13}$$

for every n. For

$$\begin{aligned} b(n) &= a(1), \quad \text{for} \quad n = 1, \\ &= 0, \quad \text{for} \quad n > 1, \end{aligned}$$

we know from the previous section that the operator $T_0 + B$ has a unique proper value E ($|E| > 2$) if and only if $|a(1)| > 1$. In other words, for real $a(1) > 1$ and $a(n) \geq 0$, relations (12) and (13) hold and, therefore, the point spectrum of $T_0 + A$ is not empty.

The above criterion, as well as some others, which was established in Ref. 2, is very strong for the case in which $a(n) > 0$ for every n, as we may see by using for the definition of the operator B the sequence

$$b(n) = 2bn^{-1}, \quad n = 1, 2, \cdots, b > 0. \tag{14}$$

The greatest proper value E of the operator $T_0 + B$, where $b(n)$ is the sequence (14), is $E = 2(1 + b^2)^{\frac{1}{2}}$ (Appendix B). Thus, if we choose $b = 2k(k^2 - 1)^{-1}$, $k > 1$, then, in case $a(n) > 0$ for every n, conditions (12) and (13) are satisfied for $n \geq k$ because the relations

$$\begin{aligned} &(1 + b^2)^{\frac{1}{2}} - bk^{-1} = 1, \\ &(1 + b^2)^{\frac{1}{2}} - b(k + \tau)^{-1} > 1, \quad \tau = 1, 2, \cdots, \end{aligned}$$

hold.

The fact that (13) is satisfied for $n < k$ implies conditions on the values of the sequence $a(n)$ ($n < k$). These conditions follow if we set $b = 2k(k^2 - 1)^{-1}$ in the relation

$$2(1 + b^2)^{-1} + a(n - 1) - 2b(n - 1)^{-1} \geq 2;$$

i.e.,

$$a(1) \geq 4(k + 1)^{-1} \to 0 \quad \text{as} \quad k \to \infty$$

and

$$\begin{aligned} a(n - 1) &\geq 2 - 2[(n - 1)k^2 - 4k + n - 1] \\ &\times [(n - 1)k^2 - n + 1]^{-1} \to 0 \quad \text{as} \quad k \to \infty. \end{aligned} \tag{15}$$

We see from (15) that for every sequence $a(n)$ with positive terms a and k, therefore, a value of b in (14) can be chosen such that condition (13) is satisfied for every $n < k$.

For a sequence $a(n)$ with negative terms we choose $b(n) = -2bn^{-1}$, $b > 0$. The least proper value E of the operator $T_0 + B$ is then $E = -2(1 + b^2)^{\frac{1}{2}}$ and, therefore, relations (12) and (13) remain the same as for the case in which $a(n) > 0$.

Finally, we note that, if $a(n) > 0$ [$a(n) < 0$] for every n, then the points inside the interval $[-\infty, -2]$ ($[2, \infty]$) belong to the resolvent set of the operator $T_0 + A$. In fact, for every $f \in \mathcal{H}$ we have

$$\begin{aligned} &\|(V + V^* + A - 2E)f\| \\ &\qquad \geq \big|\|(A - 2E)f\| - \|(V + V^*)f\|\big| \end{aligned} \tag{16}$$

and

$$\begin{aligned}\|(A - 2E)f\|^2 &= \sum_{n=1}^{\infty} |((A - 2E)f, e_n)|^2 \\ &= \sum_{n=1}^{\infty} |a(n) - 2E|^2 |(f, e_n)|^2 \\ &= 4\sum_{n=1}^{\infty} |\tfrac{1}{2}a(n) - E|^2 |(f, e_n)|^2.\end{aligned}$$

But

$$|\tfrac{1}{2}a(n) - E| \geq \rho > 1 \quad \text{if} \quad a(n) \geq 0,$$

and

$$-\infty < E < -1 \quad \text{or} \quad a(n) \leq 0 \quad \text{and} \quad 1 < E < \infty,$$

i.e.,

$$\|(A - 2E)f\|^2 \geq 4\rho^2 \|f\|^2.$$

Thus,

$$\begin{aligned}\|(A - 2E)f\| - \|(V + V^*)f\| &\geq 2\rho \|f\| \\ - \|(V + V^*)f\| &\geq 2\rho \|f\| - 2\|f\| \\ &= 2(\rho - 1)\|f\| > 0\end{aligned}$$

and, in view of (16), the following condition holds:

$$\|(T - 2E)f\| \geq 2(\rho - 1)\|f\| \ \forall f \in \mathcal{H}$$

and $2(\rho - 1) > 0$,

which is the necessary and sufficient for $2E$ to belong in the resolvent set of T.

We summarize the conclusions in the following:

Theorem 2: For every real-valued null sequence $a(n) \neq 0$, $n = 1, 2, \cdots$, with all the terms of the same sign (either positive or negative), the point spectrum of the operator $T = T_0 + A$ is not empty.

It lies in the interval $[-\|T\|, -2]$ in the case $a(n) < 0$ and in the interval $[2, \|T\|]$ in the case $a(n) > 0$ for every n.

VI. THE STRUCTURE OF THE SPECTRUM

To obtain the structure of the spectrum of $T_0 + A$, we first prove a theorem which is analogous to a well-known theorem in the theory of completely continuous operators.

Theorem 3: Let $\epsilon > 0$. For the operator $T_0 + A$ there can exist only one finite set of proper values E for which $|E| \geq 2 + \epsilon$.

Proof: Assume that $T_0 + A$ has an infinite set of proper values $\{E_n\}_1^\infty$, $|E_n| \geq 2 + \epsilon$, corresponding to the normalized proper elements $\{f_n\}_1^\infty$. [In our case, to each proper value there corresponds only one proper element (Proposition 3); i.e.,

$$(T_0 + A)f_n = E_n f_n, \quad \|f_n\| = 1 \text{ and } (f_n, f_m) = \delta_{nm}.]$$

Then we have

$$Af_n - Af_m = Ef_n - E_m f_m - T_0(f_n - f_m)$$

and

$$\|Af_n - Af_m\| \geq \big|\|E_n f_n - E_m f_m\| - \|T_0(f_n - f_m)\|\big|.$$

But

$$\|E_n f_n - E_m f_m\| = (E_n^2 + E_m^2)^{\frac{1}{2}} \geq (2 + \epsilon)\sqrt{2}$$

and

$$\|T_0(f_n - f_m)\| \leq \|T_0\| \, \|f_n - f_m\| = 2\sqrt{2}.$$

Therefore,

$$\|E_n f_n - E_m f_m\| - \|T_0(f_n - f_m)\| \geq (2 + \epsilon)\sqrt{2} - 2\sqrt{2} = \epsilon\sqrt{2} > 0.$$

Thus, $\|Af_n - Af_m\| \geq \epsilon\sqrt{2}$, which contradicts the compactness of the operator A.

Corollary: The points -2 and $+2$ are the only possible accumulation points of proper values which lie in the intervals $[-\infty, -2]$ and $[2, \infty]$.

Theorem 4: Every proper value E of the operator $T_0 + A$ is unstable in the following sense: There cannot exist a number N such that for $k > N$ the point E belongs to the point spectrum of all $T_0 + A_k$, where $A_k \to A$ as $k \to \infty$. (A_k is defined in Sec. IV.)

Proof: The proof is based on the following characteristic property of tridiagonal operators: If f is a proper element of $T_0 + A$ and at least two of its successive components (f, e_k) and (f, e_{k+1}) are equal to zero, then $(f, e_k) = 0 \; \forall \, k = 1, 2, \cdots$. The proof of this property is similar to that of Proposition 2.

We have the following three possible cases:

(a) $(T_0 + A_k)f_k = Ef_k$, $E > 2$, $\|f_k\| = 1$, and f_k infinite set, $f_k \neq f_{k+i}$, $i = 1, 2, \cdots$,

(b) $(T_0 + A_k)f = Ef$, $E > 2$, $\|f\| = 1$, $k > N$,

(c) $(T_0 + A_k)f_k = Ef_k$, $E > 2$, $\|f_k\| = 1$, and f_k finite set, $k > N$.

We can exclude case (b) by the following reasoning: Let

$$(T_0 + A_k)f = Ef, \quad (T_0 + A_{k+1})f = Ef,$$
$$(T_0 + A_{k+2})f = Ef.$$

Then

$$(A_{k+1} - A_k)f = 0 \quad \text{and} \quad (A_{k+2} - A_{k+1})f = 0,$$

i.e.,

$$a_{k+1}(f, e_{k+1})e_{k+1} = 0, \quad a_{k+2}(f, e_{k+1})e_{k+2} = 0,$$

or

$$(f, e_{k+1}) = (f, e_{k+2}) = 0.$$

Hence all the components of f are zero. For case (c) let $f_1, f_2, \cdots, f_n$ be proper elements of the operators $T_0 + A_k$, $k > N$. At least one of them, say f_1, must appear as a proper element of $T_0 + A_k$ an infinite number of times. Let

$$(T_0 + A_k)f_1 = Ef_1 \quad \text{and} \quad (T_0 + A_{k+\nu})f_1 = Ef_1;$$

then

$$(A_{k+\nu} - A_k)f_1 = a_{k+1}(f, e_{k+1})e_{k+1} + a_{k+2}(f, e_{k+2})e_{k+2} + \cdots + a_{k+\nu}(f, e_{k+\nu})e_{k+\nu} = 0.$$

Hence

$$(f_1, e_{k+1}) = (f_1, e_{k+2}) = \cdots = (f_1, e_{k+\nu}) = 0$$

and, therefore, $f_1 = 0$. Consider now case (a), and let

$$T_0 f_k + A_k f_k = Ef_k, \quad f_k \neq 0, \tag{17}$$

$$T_0 f_{k+1} + A_{k+1} f_{k+1} = Ef_{k+1}, \quad f_k \neq f_{k+1}, \tag{18}$$

$$T_0 f_{k+2} + A_{k+2} f_{k+2} = Ef_{k+2}, \quad f_k \neq f_{k+1} \neq f_{k+2}, \text{ etc.} \tag{19}$$

Observe that

$$(e_{k+1}, f_{k+1}) \neq 0,$$
$$(e_{k+2}, f_{k+2}) \neq 0 \tag{20}$$

because, if, for instance, $(e_{k+1}, f_{k+1}) = 0$, then $A_k f_{k+1} = A_{k+1} f_{k+1}$. That means [because of (17) and (18)] that E is a proper value with multiplicity, contrary to Proposition 3.

Equations (17) and (18) give

$$((A_{k+1} - A_k)f_k, f_{k+1}) = 0,$$

i.e.,

$$(e_{k+1}, f_k)(e_{k+1}, f_{k+1}) = 0,$$

from which, by virtue of (20),

$$(e_{k+1}, f_k) = 0. \tag{21}$$

Equations (17) and (19) give

$$((A_{k+2} - A_k)f_k, f_{k+2}) = 0$$

or

$$a_{k+1}(f_k, e_{k+1})(e_{k+1}, f_{k+2}) + a_{k+2}(f_k, e_{k+2})(e_{k+2}, f_{k+2}),$$

i.e.,

$$(f_k, e_{k+2})(e_{k+2}, f_{k+2}) = 0,$$

and, because of (19),

$$(f_k, e_{k+2}) = 0. \tag{22}$$

From (21) and (22) it follows that $f_k = 0$. QED

Remark: If $a(n)$ is not a null sequence but coverages monotonically to the real number a, then all the above theorems hold for the operator $T_0 + A - a$, the spectrum of which is simply a translation of the spectrum of $T_0 + A$.

APPENDIX A: THE APPLICATION OF THE RITZ METHOD

It is convenient for the application of the Ritz method to choose in the present case, as a complete system of elements in $\mathfrak{H}$, the orthonormal basis $\{e_n\}_1^\infty$. Then the n-approximate greatest proper value of the operator $T = V + V^* + A$ is given[4] from the greatest root E of the equation

$$\begin{vmatrix} (Te_1, e_1) - E & (Te_2, e_1) & \cdots & (Te_n, e_1) \\ (Te_1, e_2) & (Te_2, e_2) - E & \cdots & (Te_n, e_2) \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ (Te_1, e_n) & (Te_2, e_n) & \cdots & (Te_n, e_n) - E \end{vmatrix} = 0. \tag{A1}$$

Example 1: Here we give the case of the operator

$$T = V + V^* + A: a(n) = 12n^{-1}, \quad n = 1, 2, \cdots. \tag{A2}$$

Using the system of the two vectors e_1 and e_2, we have $Te_1 = e_2 + 12e_1$ and $Te_2 = e_1 + e_3 + 6e_2$. Therefore, Eq. (A1) takes the form

$$\begin{vmatrix} 12 - E & 1 \\ 1 & 6 - E \end{vmatrix} = 0,$$

from which we obtain the greatest root $E = 12.1622$.

Example 2: Here we give the case of the operator

$$T = V + V^* + A, \quad a(n) = \begin{cases} 10 \text{ for } n = 1 \\ 0 \text{ for } n > 1 \end{cases}. \tag{A3}$$

Then, similarly, we have $E = 10.099$.

We compare the approximate values found for the greatest proper value of the operators (A2) and (A3) with the exact ones. In the first example the exact value is $E = 2(1 + 6^2)^{\frac{1}{2}} = 12.1654$ (Appendix B); in the second, (Sec. IV) $E = 10 + 10^{-1} = 10.100$. Thus, we obtain a very satisfactory approximation by using the system of only two vectors e_1 and e_2.

APPENDIX B

Consider the proper value problem of the Schrödinger-type tridiagonal operator

$$T = V + V^* + 2bC_0^{-1}, \tag{B1}$$

where C_0 is defined as follows:

$$C_0 e_n = n e_n, \quad n = 1, 2, \cdots.$$

This problem has been solved[5] by classical methods. The proper values of (B1) for $b > 0$ are found to be

$$E_k = 2[1 + (b/k)^2]^{\frac{1}{2}}.$$

Here we give a simple solution with the help of the realization of T in the Hardy–Lebesgue space. Assume that the proper value equation

$$(V + V^* + 2bC_0^{-1})f = 2\epsilon f, \quad E = 2\epsilon, \quad b > 0,$$

or, equivalently,

$$(C_0 V + C_0 V^* + 2b)f = 2\epsilon C_0 f. \tag{B2}$$

Assume that $\phi(z) = (f_z, f)$ is the analytic function corresponding to the element $f \in \mathfrak{H}$, i.e., $f \to \phi(z)$. Then the elements f_z belong to the definition domain $D(C_0)$ of C_0, and

$$C_0 f \to z\phi'(z) + \phi(z) \tag{B3}$$

holds. In fact, we have

$$\|C_0 f_z\|^2 = \sum_{n=1}^{\infty} n^2 |z|^{2(n-1)} < \infty \quad \text{for} \quad |z| < 1$$

and

$$(f_z, C_0 f) = (C_0 f_z, f) = \sum_{n=1}^{\infty} n z^{n-1}(e_n, f) = z\phi'(z) + \phi(z)$$

because

$$\phi'(z) = z^{-1}\left(\sum_{n=1}^{\infty} n z^{n-1}(e_n, f) - \phi(z)\right).$$

Similarly, we can prove

$$C_0 V^* f \to \phi'(z),$$
$$C_0 V f \to z^2\phi'(z) + 2z\phi(z). \tag{B4}$$

By virtue of (B3) and (B4), Eq. (5), for the proper value problem (B2), has the form

$$z^2\phi'(z) + 2z\phi(z) + \phi'(z) + 2b\phi(z) - 2\epsilon[z\phi'(z) + \phi(z)] = 0,$$

from which

$$\phi(z) = \text{const} \times A(z)/B(z) \quad \text{for} \quad \epsilon \neq \pm 1, \qquad \text{(B5)}$$

where

$$A(z) = [z - \epsilon + (\epsilon^2 - 1)^{\frac{1}{2}}]^{b(\epsilon^2-1)^{-\frac{1}{2}}-1},$$
$$B(z) = [z - \epsilon - (\epsilon^2 - 1)^{\frac{1}{2}}]^{b(\epsilon^2-1)^{-\frac{1}{2}}+1},$$

and

$$\phi(z) = \text{const} \times (z \mp 1)^{-2}e^{2b/(z\mp 1)} \quad \text{for} \quad \epsilon = \pm 1.$$

For $\epsilon < 1$, $\phi(z)$ cannot be analytic. Thus, the case which the Weyl–von Neumenn theorem predicts[2] is excluded. ϵ must be contained in the interval $[1, 1 + 2b]$. For $\epsilon > 1$ we make the following remark: From the integral condition (7) it is easy to obtain the fact that, if $\phi_1(z)$ or $\phi_2(z)$ is bounded on the unit disc, then

$$\phi_1(z)\phi_2(z) \in \mathcal{H}_2.$$

Because of this statement and the fact that $\epsilon + (\epsilon^2 - 1)^{\frac{1}{2}} > 1$, we have from (B5)

$$b/(\epsilon^2 - 1)^{\frac{1}{2}} = k + 1, \quad k = 0, 1, 2, \cdots,$$

i.e. (because $b > 0$),

$$E_k = 2\epsilon = 2[1 + (b/k)^2]^{\frac{1}{2}}.$$

Similarly we can find the proper values for $b < 0$, $E_k = -2[1 + (b/k)^2]^{\frac{1}{2}}$.

APPENDIX C: THE SPECTRUM OF THE OPERATOR $V + V^*$

From the relation $\|V + V^*\| = 2$ (see Ref. 2) it follows that all E, $|E| > 2$, belong to the resolvent set of $V + V^*$. Thus, if there exists a proper value E of $V + V^*$, then $|E| \leq 2$. But then the components $\alpha(n) = (f, e_n)$ of the proper element f must satisfy the equation

$$\alpha(n + 1) + \alpha(n - 1) = E\alpha(n). \qquad \text{(C1)}$$

The solutions of the above equation for $|E| < 2$ are oscillatory, and for $E = \pm 2$ they have the general form

$$\alpha(n) = c_1 n + c_2, \quad \text{for} \quad E = 2,$$
$$\alpha(n) = c_1 n(-1)^n + c_2, \quad \text{for} \quad E = -2,$$

where c_1 and c_2 are constants. Thus, in any case, they do not belong in $l_2(1, \infty)$. The point spectrum of $V + V^*$ is therefore empty. To prove that the entire closed interval $[-2, 2]$ is the continuous spectrum of $V + V^*$, we have only to prove that every E, $|E| \leq 2$, belongs to the spectrum of $V + V^*$.

Suppose that there exists an E which does not belong to the spectrum of $V + V^*$. Then the inverse of $V + V^* - E$ exists and is bounded; i.e., the equation $(V + V^* - E)f = g$ has a unique solution $f \in \mathcal{H}$ for every $g \in \mathcal{H}$. We shall show that this is impossible for $|E| \leq 2$.

Choosing $g = e_1$, we first observe that the components of the element f in the equation

$$(V + V^* - E)f = e_1$$

cannot be finite because, otherwise, all the components must be zero. [In fact, if

$$(f, e_k) = (f, e_{k+1}) = \cdots = 0$$

from some $k \neq 1$ on, then we must have

$$(Vf, e_k) + (V^*f, e_k) - E(f, e_k) = (e_1, e_k) = 0$$

or

$$(f, e_{k-1}) = 0.$$

Consequently,

$$(Vf, e_{k-1}) + (V^*f, e_{k-1}) = 0,$$

i.e.,

$$(f, e_{k-2}) = 0 \quad \text{a.s.o.}]$$

Thus the components of the element f are infinite, and it is easy to see that they satisfy the equation (C1) with the condition

$$\alpha(2) - E\alpha(1) = 1. \qquad \text{(C2)}$$

But the solutions of the equation (C1) do not belong in $l_2(1, \infty)$ for $|E| \leq 2$. [Note that, for $|E| > 2$, we can find the unique solution of equation (C1), $\alpha(n) = -\lambda^n \in l_2(1, \infty)$, which satisfies the condition (C2). λ is the solution of the equation $\lambda^2 - E\lambda + 1 = 0$, whose absolute value is less than unity. This is consistent with the relation $\|V + V^*\| = 2$.]

[1] T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, Berlin, 1966).
[2] P. C. Deliyannis and E. K. Ifantis, J. Math. Phys. **10**, 421(1969).
[3] K. Yosida, *Functional Analysis* (Springer-Verlag, Berlin, 1965), p. 41.
[4] B. Z. Vulikh, *Introduction to Functional Analysis* (Pergamon, New York, 1963).
[5] E. K. Ifantis, Z. Angew. Math. Mech. **48**, 66 (1968).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

The Lorentz Group and the Sphere*

A. HELD, E. T. NEWMAN, AND R. POSADAS
Physics Department, University of Pittsburgh, Pittsburgh, Pennsylvania 15213

(Received 25 April 1970)

A direct connection between the spin and conformally weighted functions on the sphere and geometric objects in Minkowski space is established through the isomorphism of the conformal group of the sphere to the restricted Lorentz group. It is shown that with the use of these functions one can duplicate all the standard work on the representations of the Lorentz group. It is shown further that these functions can be used to obtain a generalization of the classical equations of motion in which internal degrees of freedom arise naturally.

1. INTRODUCTION

It is easily demonstrated that the restricted Lorentz group is isomorphic to the conformal group of the extended plane, and hence (via the stereographic projection of the sphere onto that plane) to the conformal group of the 2-dimensional sphere. Thus, an operation of an element of the Lorentz group on a geometric object defined in Minkowski space may be identified with an operation of the conformal group on an object defined on a 2-sphere. The main purpose of this paper is to identify those objects on the sphere and to establish their connection with objects in Minkowski space. We shall show that the required objects are tensor densities of arbitrary rank and weight, or equivalently the spin and conformally weighted functions on the sphere. It is these latter functions that play a dominant role in our work. With the use of these functions we can duplicate all the standard work on finite, infinite, and unitary representations of the Lorentz group. It will also be shown in Appendix A that the use of these functions arises naturally in certain physical situations. In particular, this leads to a generalization of the classical equations of motion in which internal degrees of freedom arise naturally.

In the course of this work, the operator edth, $\eth$, will be used extensively.[1,2] It is assumed that the reader is familiar with both this operator and the associated spin s spherical harmonics. In Appendix B to this paper, we demonstrate a method for obtaining a Clebsch–Gordan type angular momentum decomposition using the operator $\eth$.

2. PROPERTIES OF THE UNIT SPHERE

Starting with the standard line element on the unit sphere

$$ds^2 = d\theta^2 + \sin^2\theta\, d\phi^2 \tag{2.1}$$

and introducing the complex stereographic coordinate

$$\zeta = \cot \tfrac{1}{2}\theta e^{i\phi}, \tag{2.2}$$

the line element becomes

$$ds^2 = d\zeta\, d\bar{\zeta}/P^2, \tag{2.1'}$$

where $P = \frac{1}{2}(1 + \zeta\bar{\zeta})$.

Essential to the formalism are the two complex vectors m^μ and $\bar{m}^\mu$ which obey the normalization conditions

$$m^\mu \bar{m}_\mu = 1, \quad m^\mu m_\mu = 0, \quad \mu = 1, 2. \tag{2.3}$$

These vectors are naturally chosen to make Re m^μ and Im m^μ tangential, respectively, to the curves Im ζ = const and Re ζ = const. Thus, they take the form, in this coordinate system,

$$m^\mu = \sqrt{2}\, P\delta^\mu_\zeta, \quad \bar{m}^\mu = \sqrt{2}\, P\delta^\mu_{\bar{\zeta}}. \tag{2.4}$$

The generators of the conformal group are defined as $i\xi^\mu(\partial/\partial x^\mu)$, where ξ^μ is a solution to the conformal Killing equation

$$\xi_{\mu;\nu} + \xi_{\nu;\mu} = k g_{\mu\nu}. \tag{2.5}$$

We obtain the general solution to this equation by the following procedure.

Letting ξ_μ be a solution to (2.5), we form the scalars

$$\xi_+ = \xi_\mu m^\mu \quad \text{and} \quad \xi_- = \xi_\mu \bar{m}^\mu$$

which are of spin weight 1 and -1, respectively. Then transvecting (2.5) with $m^\mu m^\nu$, $m^\mu \bar{m}^\nu$, and $\bar{m}^\mu \bar{m}^\nu$, we obtain the three equations

$$\eth \xi_+ = 0, \tag{2.5'}$$

$$\bar{\eth}\xi_+ + \eth \xi_- = \sqrt{2}\, k(\zeta, \bar{\zeta}), \tag{2.5''}$$

$$\eth \xi_- = 0. \tag{2.5'''}$$

As ξ_μ is to be nonsingular on the sphere, ξ_+ can be expanded as a series,

$$\xi_+ = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{lm}\, {}_1Y_{lm}(\zeta, \bar{\zeta}), \tag{2.6}$$

where a_{lm} are complex constants and ${}_sY_{lm}(\zeta, \bar{\zeta})$ are the spin-s spherical harmonics.

TABLE I. Generators of the Lorentz group and their associated infinitesimal conformal and spin weight factors.

Generator	$i\xi_a^\mu \frac{\partial}{\partial x^\mu}$	k, infinitesimal conformal factor	$i\lambda$, infinitesimal spin weight factor
$N_1 = L_1$	$-\frac{1}{2}\left[(\zeta^2-1)\frac{\partial}{\partial\zeta} - (\bar\zeta^2-1)\frac{\partial}{\partial\bar\zeta}\right]$	0	$-\frac{i}{2}(\zeta+\bar\zeta)$
$N_2 = L_2$	$\frac{i}{2}\left[(\zeta^2+1)\frac{\partial}{\partial\zeta} + (\bar\zeta^2+1)\frac{\partial}{\partial\bar\zeta}\right]$	0	$-\frac{1}{2}(\zeta-\bar\zeta)$
$N_3 = L_3$	$\zeta\frac{\partial}{\partial\zeta} - \bar\zeta\frac{\partial}{\partial\bar\zeta}$	0	i
$N_4 = M_1$	$\frac{i}{2}\left[(\zeta^2-1)\frac{\partial}{\partial\zeta} + (\bar\zeta^2-1)\frac{\partial}{\partial\bar\zeta}\right]$	$\frac{\zeta+\bar\zeta}{1+\zeta\bar\zeta}$	$-\frac{1}{2}(\zeta-\bar\zeta)$
$N_5 = M_2$	$\frac{1}{2}\left[(\zeta^2+1)\frac{\partial}{\partial\zeta} - (\bar\zeta^2+1)\frac{\partial}{\partial\bar\zeta}\right]$	$-i\frac{\zeta-\bar\zeta}{1+\zeta\bar\zeta}$	$\frac{i}{2}(\zeta+\bar\zeta)$
$N_6 = M_3$	$-i\left[\zeta\frac{\partial}{\partial\zeta} + \bar\zeta\frac{\partial}{\partial\bar\zeta}\right]$	$-\frac{1-\zeta\bar\zeta}{1+\zeta\bar\zeta}$	0

From (2.5′) it follows that $l = 1$, so that

$$\xi_+ = a\,{}_1Y_{11} + b\,{}_1Y_{10} + c\,{}_1Y_{1,-1}. \tag{2.7}$$

As ξ^μ is real,

$$\xi_- = \bar\xi_+, \tag{2.8}$$

so that

$$\xi_- = \bar c\,{}_{-1}Y_{11} - \bar b\,{}_{-1}Y_{10} + \bar a\,{}_{-1}Y_{1,-1}. \tag{2.9}$$

Substituting (2.7) and (2.9) into (2.5″), we obtain

$$(a-\bar c)\,{}_0Y_{11} + (b+\bar b)\,{}_0Y_{10} + (c-\bar a)\,{}_0Y_{1,-1} = k. \tag{2.10}$$

By setting $k = 0$, we obtain the general solution for a rotation,

$$\xi_+ = p\,{}_1Y_{11} + iq\,{}_1Y_{10} + \bar p Y_{1,-1}, \tag{2.11}$$

where p is a complex constant and q a real constant. Using the relation

$$\xi^\mu = \xi_+\bar m^\mu + \xi_- m^\mu, \tag{2.12}$$

we obtain the three generators of the rotation group. These are labeled L_1, L_2, and L_3 and are shown in Table I.

Setting

$$k = 2s\,{}_0Y_{11} + 2t\,{}_0Y_{10} - 2\bar s\,{}_0Y_{1,-1}, \tag{2.13}$$

we obtain the solution for the "pure" conformal (boost) transformation

$$\xi_+ = s\,{}_1Y_{11} + t\,{}_1Y_{10} - \bar s\,{}_0Y_{1,-1}, \tag{2.14}$$

with s complex and t real.

Using (2.12) we obtain the generators of the boost transformations, together with their associated k. These are listed as M_1, M_2, and M_3 in Table I.

The normalization of the various vectors has been chosen so that the generators give rise to the canonical commutation relations

$$[L_i, L_j] = i\epsilon_{ijk}L_k, \quad [L_i, M_j] = i\epsilon_{ijk}M_k,$$
$$[M_i, M_j] = -i\epsilon_{ijk}L_k. \tag{2.15}$$

These are seen to be identical to the commutation relations of the proper Lorentz group.[3] The finite transformation associated with these generators is

$$\zeta' = \frac{a\zeta + b}{c\zeta + d}, \tag{2.16}$$

where a, b, c, and d are arbitrary complex numbers subject to the condition $ad - bc = 1$. This, the fractional linear transformation, is known to be isomorphic to the proper Lorentz group. Thus the isomorphism of the conformal group on the sphere to the proper Lorentz group is established.

Two functions which play a major role in the remainder of this paper may be defined from the fractional linear transformation as follows:

$$e^{i\lambda} = \left(\frac{\partial\bar\zeta'/\partial\bar\zeta}{\partial\zeta'/\partial\zeta}\right)^{\frac{1}{2}} = \frac{c\zeta + d}{\bar c\bar\zeta + \bar d}, \tag{2.17}$$

$$K = J^{-\frac{1}{2}}\frac{P}{P'}$$
$$= (1+\zeta\bar\zeta)[(a\zeta+b)(\bar a\bar\zeta+\bar b) + (c\zeta+d)(\bar c\bar\zeta+\bar d)]^{-1}, \tag{2.18}$$

where

$$J = \frac{\partial \zeta}{\partial \zeta'} \cdot \frac{\partial \bar{\zeta}}{\partial \bar{\zeta}'}, \quad P = \tfrac{1}{2}(1 + \zeta\bar{\zeta}), \quad P' = \tfrac{1}{2}(1 + \zeta'\bar{\zeta}').$$

λ is interpreted geometrically as the local angle of rotation of the two coordinate grids given by $\zeta =$ const and $\zeta' =$ const (after transformation). The infinitesimal form for λ is given in Table I.

K is the conformal factor associated with the transformation of the spherical metric under (2.16), i.e.,

$$\frac{d\zeta'\, d\zeta'}{P'^2} = K^2 \frac{d\zeta\, d\bar{\zeta}}{P^2}.$$

Under a pure rotation $K = 1$, while under a boost $K \neq 1$. For the form of an infinitesimal boost, i.e., $K = 1 + \epsilon k$, see Table I.

From Eqs. (2.17) and (2.18) it follows that

$$\frac{\partial \zeta}{\partial \zeta'} = e^{i\lambda} K^{-1} \frac{P}{P'}, \tag{2.19}$$

a relation that will be extensively used.

3. SPIN AND CONFORMALLY WEIGHTED FUNCTIONS

In this section we will consider functions on the sphere which transform as

$$\eta(\zeta, \bar{\zeta}) \to \eta'(\zeta'\bar{\zeta}') = K^w e^{is\lambda} \eta(\zeta, \bar{\zeta}) \tag{3.1}$$

under (2.16).

Functions of this type are said to be spin and conformally weighted, with s the spin weight and w the conformal weight. They arise naturally if we consider tensor densities on the sphere of rank s and weight $n = -\frac{1}{2}(w + s)$. (With no loss in generality the densities may be taken as symmetric and trace free–any 2-dimensional density can be decomposed into its irreducible parts and represented as a sum of terms of this type.) From the transformation properties of a tensor density, i.e.,

$$A'_{(\mu\cdots\nu)_s} = J^n \frac{\partial x^\alpha}{\partial x'^\mu} \cdots \frac{\partial x^\beta}{\partial x'^\nu} A_{(\alpha\cdots\beta)_s}, \tag{3.2}$$

where s denotes the number of indices, by contracting with $m'^\mu \cdots m'^\nu$ and using the complex stereographic coordinates of the previous section (thus $m'^\mu = \sqrt{2}\, P' \delta^\mu_{\zeta'}$), we obtain

$$A'_{\zeta'\cdots\zeta'} = J^n \left(\frac{\partial \zeta}{\partial \zeta'}\right)^s A_{\zeta\cdots\zeta}. \tag{3.3}$$

Eliminating J and $\partial\zeta/\partial\zeta'$ by Eqs. (2.18) and (2.19) results in

$$A'_{\zeta'\cdots\zeta'} = K^{-(2n+s)} \frac{P^{(2n+s)}}{P'^{(2n+s)}} e^{is\lambda} A_{\zeta\cdots\zeta} \tag{3.4a}$$

or

$$\eta'(\zeta', \bar{\zeta}') = K^w e^{is\lambda} \eta(\zeta, \bar{\zeta}), \tag{3.4b}$$

where

$$\eta = A_{\zeta\cdots\zeta} P^{(2n+s)},$$
$$\eta' = A'_{\zeta'\cdots\zeta'} P'^{(2n+s)},$$
$$w = -(2n + s).$$

We now wish to consider in particular the behavior of these functions under the infinitesimal fractional linear transformations associated with the conformal Killing vectors of the previous section.

Defining the Lie derivative $\pounds(\xi_a)$ of a function with respect to the conformal Killing vector ξ_a, $a = 1 \cdots 6$, by

$$\eta(\zeta, \bar{\zeta}) - \eta'(\zeta, \bar{\zeta}) = \pounds(\xi_a)\eta = \frac{1}{i}(N_a - iwk_a + s\lambda_a)\eta, \tag{3.5}$$

where N_a is the generator $i\xi^\mu_a(\partial/\partial x^\mu)$ and k_a and λ_a are the associated conformal and spin weight factors of Table I. The six operators $i\pounds(\xi_a)$ obey the identical commutation relations (2.15) as the associated operators N_a, and so yield a new realization of the Lie algebra of the proper Lorentz group.

Next we calculate the Casimir operators associated with this realization. They are

$$L_i^2 - M_i^2 = s^2 + w(w + 2),$$
$$L_i M_i = is(w + 1). \tag{3.6}$$

It is well known that each irreducible representation of the proper Lorentz group is determined by a pair of numbers[4] k_0, c, where k_0 is integer or half-integer and c is any complex number. If the Casimir operators are calculated in terms of these numbers, the results are

$$L_i^2 - M_i^2 = k_0^2 + c^2 - 1,$$
$$L_i M_i = -ik_0 c. \tag{3.7}$$

Since k_0 is required to be positive, and integer or half-integer, this leads to the natural identification $k_0 = |s|$, and $c = -s/|s|\,(w + 1)$ with the additional restriction that s is integer or half-integer. Thus, for every irreducible representation of the proper Lorentz group we have an associated realization of the Lie algebra in terms of $i\pounds(\xi_a)$, and conversely. The relations between k_0 and c, which determine whether the representation is finite dimensional, unitary, etc., now apply to the related s and w.

In choosing basis functions for the representation we diagonalize L_3 and L_i^2 (the Casimir operator

associated with the rotation subgroup). They are

$$L_3 = \zeta\frac{\partial}{\partial\zeta} - \bar{\zeta}\frac{\partial}{\partial\bar{\zeta}} + s,$$

$$L_i^2 \equiv L^2$$

$$= -(1+\zeta\bar{\zeta})\frac{\partial^2}{\partial\zeta\partial\bar{\zeta}} + s(1+\zeta\bar{\zeta})\left(\zeta\frac{\partial}{\partial\zeta} - \bar{\zeta}\frac{\partial}{\partial\bar{\zeta}} + s\right)$$

$$= -\eth\bar{\eth} + s(s+1). \tag{3.8}$$

The eigenfunctions of these operators ${}_sY_{lm}$ are determined by the equations

$$L_3\,{}_sY_{lm} = m\,{}_sY_{lm}$$

and

$$L^2\,{}_sY_{lm} = l(l+1)\,{}_sY_{lm}.$$

The functions ${}_sY_{lm}$ defined in this manner are (when properly normalized) just the spin s spherical harmonics.[1,5] (In order to exhibit the proper behavior under the boosts, these functions are assigned conformal weight w.)

4. PROPERTIES OF REPRESENTATIONS

At this point we wish to state a series of results concerning representations, the proof of which is postponed till the end of this section.

Consider infinitely differentiable functions of the sphere $\eta(\zeta,\bar{\zeta})$ and $\eta'(\zeta',\bar{\zeta})$ which transform with spin weight s (s integer or half-integer) and conformal weight w (w an arbitrary complex number), such that both η and η' are expandable in spin-s spherical harmonics, i.e.,

$$\eta = \sum_{l=|s|} a_{lm}\,{}_sY_{lm}(\zeta,\bar{\zeta}),$$

$$\eta' = \sum_{l=|s|} a'_{lm}\,{}_sY_{lm}(\zeta',\bar{\zeta}') = K^w e^{is\lambda}\eta. \tag{4.1}$$

These functions then form the vector space of a representation (not necessarily irreducible) of the proper Lorentz group, which will be denoted by D_χ,

$$\chi = (n_1, n_2) = (w - s + 1,\ w + s + 1). \tag{4.2}$$

The reason for this notation will be obvious later. A converse statement is also true; any irreducible representation can be realized on these D_χ spaces.[6]

To be more precise, we consider the space of two complex numbers n_1 and n_2, given by Eq. (4.2). We say that $\chi = (n_1, n_2)$ is an integer point of this space, if n_1 and n_2 are integers of the same sign and both nonzero. We then have the following results:

(a) A representation associated with a noninteger χ, is irreducible and infinite dimensional. In addition, it is equivalent to the representation associated with the point $-\chi$.

(b) A representation associated with a positive integer point (n_1 and n_2 positive integers, or equivalently w and s both integer or half-integer, with $w \geq |s|$) is neither irreducible nor totally reducible. The D_χ possesses an invariant subspace E_χ which is spanned by the finite basis vectors ${}_sY_{lm}$, $|s| \leq l \leq w$. The E_χ are the vector spaces for all the finite-dimensional representations. It is possible to obtain an infinite-dimensional representation also from D_χ by considering the factor space D_χ/E_χ. The factor space is isomorphic to the two equivalent representations $D_{n_1-n_2} \cong D_{-n_1,n_2}$. In fact, given a vector in D_{n_1,n_2}, labeled by $\eta_{(s,w)}$, the vectors

$$\eta_{(w+1,s-1)} \equiv \eth^{w-s+1}\eta_{(s,w)}$$

and

$$\eta_{(-w-1,-s-1)} \equiv \bar{\eth}^{w+s+1}\eta_{(s,w)} \tag{4.3}$$

are in D_{-n_1,n_2} and $D_{n_1,-n_2}$, respectively.

There exists one last isomorphism, namely

$$D_{n_1,n_2}/E_{n_1,n_2} \cong F_{-n_1,-n_2},$$

where $F_{-n_1,-n_2}$ (see next paragraph) is a subspace of $D_{-n_1,-n_2}$. The mapping $D_{n_1,n_2} \to F_{-n_1,-n_2}$ is explicitly given by

$$\eta_{(-s,-w-2)} = \bar{\eth}^{w+s+1}\eth^{w-s+1}\eta_{(s,w)}. \tag{4.4}$$

(c) A representation associated with a negative integer point [$-\chi = (-n_1, -n_2)$, n_1 and n_2 positive] is also neither irreducible nor totally reducible. $D_{-\chi}$ now possesses an infinite-dimensional invariant subspace spanned by ${}_{-s}Y_{lm}$, $l > w = \frac{1}{2}(n_1 + n_2) - 1$, and denoted by $F_{-n_1,-n_2}$. The factor space $D_{-\chi}/F_{-n_1,-n_2} \cong E_{n_1,n_2}$. Explicitly, the mapping from $D_{-\chi} \to E_{n_1,n_2}$ is given by

$$\eta_{(s,w)}(\zeta,\bar{\zeta}) = \int M(\zeta,\bar{\zeta},\zeta',\bar{\zeta}')\eta_{(-s,-w-2)}(\zeta',\bar{\zeta}')'\,d\Omega', \tag{4.5}$$

where $d\Omega'$ is the area element of the sphere and

$$M = \sum_{\substack{l=s \\ -l\leq m\leq l}}^{w} a_l^{(s,w)}\,{}_sY_{lm}(\zeta,\bar{\zeta})\,{}_{-s}\bar{Y}(\zeta',\bar{\zeta}')$$

with

$$a_l^{(s,w)} \equiv (-1)^{l+s}\frac{(w+|s|+1)!\,(w-|s|)!}{(w+l+1)!\,(w-l)!}. \tag{4.6}$$

The proof of these assertions is quite simple. We first show the isomorphism between the spin and conformally weighted functions and the homogeneous functions of two complex variables and then,translating into the language of spin and conformally weighted functions (via this isomorphism), known theorems which relate these homogeneous functions to the Lorentz group.

We begin by considering the infinitely differentiable, homogeneous functions of the two complex variables z_1 and z_2 and their complex conjugates $\bar{z}_1$ and $\bar{z}_2$; the degree of homogeneity is $n_1 - 1$ in z_1 and z_2 and $n_2 - 1$ in $\bar{z}_1$ and $\bar{z}_2$, i.e.,

$$f(az_1, az_2, \bar{a}\bar{z}_1, \bar{a}\bar{z}_2) = a^{n_1-1}\bar{a}^{n_2-1}f(z_1, z_2, \bar{z}_1, \bar{z}_2). \tag{4.7}$$

The linear transformation

$$\begin{aligned} z_1' &= \alpha z_1 + \beta z_2, \\ z_2' &= \gamma z_1 + \delta z_2 \end{aligned} \tag{4.8}$$

induces a transformation on the f, namely that

$$f'(z_1', z_2', \bar{z}_1', \bar{z}_2') = f(z_1, z_2, \bar{z}_1, \bar{z}_2). \tag{4.9}$$

With each $f(z_1, z_2, \bar{z}_1, \bar{z}_2)$, we can associate a function of one complex variable

$$\phi(\zeta, \bar{\zeta}) = f(\zeta, 1, \bar{\zeta}, 1).$$

The f can be recovered from ϕ by

$$f(z_1, z_2, \bar{z}_1, \bar{z}_2) = z_2^{n_1-1}\bar{z}_2^{n_2-1}\phi\left(\frac{z_1}{z_2}, \frac{\bar{z}_1}{\bar{z}_2}\right). \tag{4.10}$$

The transformation on f induced by the transformation on the z's, in turn induces a transformation[6] on the ϕ, namely,

$$\phi'(\zeta', \bar{\zeta}') = (\gamma\zeta + \delta)^{-(w-s)}(\bar{\gamma}\bar{\zeta} + \bar{\delta})^{-(w+s)}\phi(\zeta, \bar{\zeta}) \tag{4.11}$$

with

$$\zeta' = \frac{\alpha\zeta + \beta}{\gamma\zeta + \delta} \quad \text{and} \quad \begin{aligned} w - s &= n_1 - 1, \\ w + s &= n_2 - 1. \end{aligned}$$

Now by using Eqs. (2.17) and (2.19), we obtain

$$\phi'/P'^w = K^w e^{is\lambda}\phi/P^w. \tag{4.12}$$

Thus, by identifying η with ϕP^{-w}, we complete the proof of the isomorphism. The assertions then follow immediately.[6]

In addition it is easily seen, from either the isomorphism or by direct calculation, that the unitary representations are given by:

(1) principal series; s integer or half-integer, and $w = -1 + ic$, where c is an arbitrary real number.

(2) supplementary series; $s = 0$, and $w = -1 + \rho$, where ρ is an arbitrary real number, $-1 < \rho < 1$.

Given a representation labeled by (s, w), the conjugate representation is $(-s, w)$.

5. CONNECTION WITH MINKOWSKI SPACE

It is the purpose of this section to establish a direct connection between the spin and conformally weighted functions and geometric objects in Minkowski space, namely, tensors and spinors and possibly more complicated objects. More precisely, we will show a simple and direct correspondence, so that the spin and conformally weighted functions will have geometric meaning in Minkowski space.

At an arbitrary point P in Minkowski space, consider an orthogonal space–time tetrad $(t^\mu, x^\mu, y^\mu, z^\mu)$. (A Lorentz transformation will be a "rotation" of this frame.) In the 3-space orthogonal to t^μ, let S denote the unit sphere. We coordinatize S by the complex stereographic coordinates ζ, $\bar{\zeta}$, the stereographic projection being taken from S onto the (x^μ, y^μ) plane. At each point of S we introduce the complex tangent vectors $m^\mu(\zeta, \bar{\zeta})$ and $\bar{m}^\mu(\zeta, \bar{\zeta})$, tangent to the ζ and $\bar{\zeta}$ lines. These vectors are of course identical to those introduced in Sec. 2, but here they are considered as residing in Minkowski space. Relative to the chosen tetrad, any null direction can be labeled by the stereographic coordinates of S by simply projecting the null direction into the three space and noting where it intersects S. We thus write an arbitrary null vector as $l^\mu(\zeta, \bar{\zeta})$; as ζ, $\bar{\zeta}$ go over S, l^μ sweeps out the null cone. The "length" of l^μ is normalized by the condition $l_\mu t^\mu = 1$. A point to be noted is that, given the frame, by this construction there is no freedom of choice in the vectors l^μ and m^μ.

In a second orthogonal frame (obtained by a Lorentz transformation from the first) at P, the identical construction can be used to find an alternative description of the null cone by the vectors $l'^\mu(\zeta', \zeta')$, with a unit sphere S', coordinatized by ζ', $\bar{\zeta}'$, with tangent vectors m'^μ and $\bar{m}'^\mu$.

We now contend that, for any proper Lorentz transformation, there will correspond a transformation of $S \leftrightarrow S'$, given by

$$\zeta' = \frac{a\zeta + b}{c\zeta + d}. \tag{5.1}$$

Moreover,

$$l'^\mu = Kl^\mu, \quad m'^\mu = e^{i\lambda}(m^\mu + Hl^\mu), \tag{5.2}$$

where K and λ are the conformal factor and rotation angle of the previous section and H is a function of K. (See Appendix A.)

The proof can be given by direct calculation, but it is easier to infer it directly from work on the Bondi–Metzner–Sachs group[7,8] by just keeping the supertranslations all zero and restricting the considerations to flat-space.

With each pair of vectors $l^\mu(\zeta, \bar{\zeta})$ and $m^\mu(\zeta, \bar{\zeta})$, we can associate a unique pair of two component spinors $o^A(\zeta, \bar{\zeta})$ and $\iota^B(\zeta, \bar{\zeta})$ normalized so that $o_A\iota^A = 1$, by

$$l^\mu = \sigma^\mu_{AB'}o^A\bar{o}^{B'}, \quad m^\mu = \sigma^\mu_{AB'}\iota^A\bar{o}^{B'}. \tag{5.3}$$

When l^μ and m^μ transform by Eq. (5.2), it is easily seen that

$$o'^A = K^{\frac{1}{2}}e^{i(\lambda/2)}o^A.$$

With this result we can establish the direct relation of the spin and conformally weighted functions with spinors.

An arbitrary irreducible spinor, and consequently any irreducible tensor through the known association of tensors with spinors, can be written

$$\Phi_{A\cdots BA'\cdots C'},$$

with $w - s$ symmetric unprimed indices and $w + s$ symmetric primed indices. By defining

$$\eta = \Phi_{A\cdots BA'\cdots C'}o^A \cdots o^B\bar{o}^{A'} \cdots \bar{o}^{C'}, \tag{5.4}$$

we see that,under a Lorentz transformation of the Minkowski space, we have

$$\eta' = K^w e^{is\lambda}\eta. \tag{5.5}$$

(We point out that,though it may appear as if we are not performing a Lorentz transformation on the spinor $\Phi_{A\cdots BA'\cdots C'}$ itself but only on the o^A, this is not true. The following example should clarify this point.)

Consider a vector A_α in two coordinate frames, i.e.,

$$\tilde{A}_\alpha = \frac{\partial x^\mu}{\partial \tilde{x}^\alpha} A_\mu$$

with

$$\tilde{l}^\alpha = \frac{\partial \tilde{x}^\alpha}{\partial x^\mu} l'^\mu = K\frac{\partial \tilde{x}^\alpha}{\partial x^\mu} l^\mu;$$

then $\tilde{A}_\alpha\tilde{l}^\alpha = KA_\alpha l^\alpha$ or $\eta' = K\eta$, i.e., $w = 1$, $s = 0$.

What we have thus shown is that,when an irreducible spinor (or tensor) at a point is contracted with a variable 2-component spinor o^A, which spans the light cone, the result is a spin and conformally weighted function given on the local unit sphere. A local Lorentz transformation from one frame to another is equivalent to the fractional linear or conformal transformation of one unit sphere onto the other.

It is not difficult to see that this contraction can be reversed for the finite-dimensional representations, by using the orthogonality properties of the spin-s spherical harmonics.

One can generalize the above notions to a spinor or tensor field, $\phi_{A\cdots BA'\cdots C}(x^\mu)$ or $T_{\alpha\cdots\beta}(x^\mu)$. They would correspond to a spin weighted function field $\eta(\zeta, \bar{\zeta}, x^\mu)$. In this manner invariant equations for spinor or tensor fields would be written as a scalar equation in a 6-dimensional space. In addition, one need not limit oneself to just the finite-dimensional representations. The infinite-dimensional or even the partially reducible representations are just as easily described in this manner as the finite-dimensional ones.

Finally, we point out that, in some sense, one can give a geometric meaning to the infinite-dimensional representations. In a given frame,

$$\eta(\zeta, \bar{\zeta}, x^\mu) = \sum_{l=|s|}^{\infty} a_{lm}(x^\mu)\ {}_sY_{lm}(\zeta, \bar{\zeta}). \tag{5.6}$$

The $a_{lm}(x^\mu)$, $-l \leq m \leq l$, can each be identified with a 3-dimensional "tensor" (or spinor) totally symmetric and trace free in l spatial indices,

$$a_{lm} \leftrightarrow a_{(ij\cdots k)_l}.$$

Thus, η is equivalent to an infinite set of 3-dimensional tensors (or spinors), each transforming irreducibly under $O(3)$ [or $SU(2)$], but being completely mixed under a Lorentz transformation.

APPENDIX A

As an application of the results of Sec. 5, we present here a generalization of the classical relativistic equations of motion for a free particle. The generalization consists of extending the representation of the homogeneous Lorentz group associated with a vector, $s = 0$ and $w = 1$ (in particular the velocity vector of a particle), to its related infinite-dimensional, partially reducible representation, the additional components representing internal structure. For future reference, slightly more material is developed here than is needed.

We begin by taking the parametric form of an arbitrary timelike world line in Minkowski coordinates (y^μ) to be $y^\mu = \xi^\mu(u)$. The Minkowski coordinates of an arbitrary point can be expressed in terms of null coordinates $x^\mu = (u, r, x^2, x^3)$, associated with the light cones emanating from the line, by

$$y^\mu = \xi^\mu(u) + r\lambda^\mu(u, x^A), \quad A = 2, 3, \tag{A1}$$

where $x^0 = u$ is a measure of proper time on the world line, with each cone being labeled by $u = \text{const}$; $x^1 = r$ is an affine parameter along each null ray lying on the cones as well as a measure of the radius of each sphere given by $u = \text{const}$, $r = \text{const}$; x^2 and x^3 are "angular coordinates" which are related to the complex stereographic coordinates ζ and $\bar{\zeta}$ by $\zeta = x^2 + ix^3$; and λ^μ is a null vector field which sweeps

out the directions of the full cone at each $u = \text{const}$ as ζ and $\bar{\zeta}$ vary. Since λ^μ is null, it must satisfy the conditions

$$\lambda^\mu \lambda_\mu = \lambda^\mu \dot{\lambda}_\mu = \lambda^\mu \lambda_{\mu,A} = 0, \tag{A2}$$

where a dot denotes $\partial/\partial u$ and a comma-A denotes $\partial/\partial x^A$. In addition we require the normalizations

$$\xi^\mu \lambda_\mu = 1, \quad \xi^\mu \dot{\xi}_\mu = 2, \tag{A3}$$

where the 2 implies that u is $\sqrt{2}/2$ times the proper time. Finally, in order to specify how null directions are to be propagated along the world line, we let

$$\lambda^\mu = l^\mu / v, \quad l^\mu = l^\mu(\zeta, \bar{\zeta}), \quad v = v(u, \zeta, \bar{\zeta}) \tag{A4}$$

and differentiate with respect to u. This yields the propagation law

$$\dot{\lambda}^\mu = -(\dot{v}/v)\lambda^\mu, \tag{A5}$$

which simply expresses the parallel transfer of the direction of λ^μ. Note that from (A3) and (A4) we obtain the relations

$$v = \dot{\xi}^\mu l_\mu, \quad \dot{v} = \ddot{\xi}^\mu l_\mu. \tag{A6}$$

We can now calculate the metric tensor $g_{\mu\nu}(x^\alpha)$ in the null coordinates by means of (A1) and the transformation law

$$g_{\mu\nu} = \frac{\partial y^\alpha}{\partial x^\mu} \frac{\partial y^\beta}{\partial x^\nu} \eta_\alpha \ ,$$

where $\eta_{\alpha\beta}$ is the usual Minkowski metric $\eta_{\alpha\beta} = \text{diag}\,(1, -1, -1, -1)$. Using (A2)–(A6), we find that

$$g_{\mu 1} = \delta^0_\mu, \quad g_{00} = 2\left(1 - \frac{\dot{v}}{v} r\right),$$

$$g_{0A} = 0, \quad g_{AB} = \frac{l^\mu{}_{,A} l_{\mu,B}}{v^2} r^2. \tag{A7}$$

Since l^μ and v were defined in (A4) only up to an arbitrary u-independent factor, we may choose this factor such that

$$l^\mu{}_{,A} l_{\mu,B} = -\delta_{AB}/2P_0^2, \tag{A8}$$

where

$$P_0 = \tfrac{1}{2}(1 + \zeta\bar{\zeta}). \tag{A9}$$

A solution to (A8) is given by

$$l^\mu = (2\sqrt{2}\, P_0)^{-1}\left(1 + \zeta\bar{\zeta}, \zeta + \bar{\zeta}, \frac{\zeta - \bar{\zeta}}{i}, \zeta\bar{\zeta} - 1\right). \tag{A10}$$

It then follows from (A6) that

$$v = \frac{1}{\sqrt{2}}\left(\xi^0 - \frac{\zeta + \bar{\zeta}}{1 + \zeta\bar{\zeta}} \xi^1 + i \frac{\zeta - \bar{\zeta}}{1 + \zeta\bar{\zeta}} \xi^2 + \frac{1 - \zeta\bar{\zeta}}{1 + \zeta\bar{\zeta}} \xi^3\right). \tag{A11}$$

By means of the relationship $\zeta = \cot \tfrac{1}{2}\theta e^{i\phi}$, we can also express v in terms of ordinary spherical harmonics $Y_{lm}(\theta, \phi)$ as

$$v = (2\pi)^{\frac{1}{2}} \xi^0 Y_{00} - \left(\frac{2\pi}{3}\right)^{\frac{1}{2}} [(\xi^1 + i\xi^2)/\sqrt{2}\, Y_{1-1} + \xi^3 Y_{10} - (\xi^1 - i\xi^2)/\sqrt{2}\, Y_{11}]. \tag{A12}$$

The line element in the null coordinates can now be written as

$$ds^2 = 2\left(1 - \frac{\dot{v}}{v} r\right) du^2 + 2\, du\, dr - \frac{r^2}{2P_0^2 v^2} d\zeta\, d\bar{\zeta}. \tag{A13}$$

Thus, the world line uniquely determines the metric and the metric determines the world line.

Assuming that l^μ is a spin weight zero quantity, one can show by straightforward calculation from (A10) that

$$\eth_0^2 l^\mu = 0, \tag{A14}$$

where $\eth_0$ denotes the operation of $\eth$ with respect to $P_0 = \frac{1}{2}(1 + \zeta\bar{\zeta})$. It also follows from (A2) and (A8) that

$$l^\mu \eth_0 l_\mu = l^\mu \bar{\eth}_0 l_\mu = 0, \quad \eth_0 l^\mu \bar{\eth}_0 l_\mu = -1. \tag{A15}$$

With the help of (A14) and (A15), we then obtain the results

$$\eth_0 l^\mu \cdot \eth_0 l_\mu = \eth_0 l^\mu \cdot \eth_0 \bar{\eth}_0 l_\mu = 0,$$
$$l^\mu \eth_0 \bar{\eth}_0 l_\mu = -\tfrac{1}{2}\eth_0 \bar{\eth}_0 l^\mu \cdot \eth_0 \bar{\eth}_0 l_\mu = 1. \tag{A16}$$

Defining the real null vector n^μ and the complex null vectors m^μ and $\bar{m}^\mu$ by

$$n^\mu = l^\mu + \eth_0 \bar{\eth}_0 l^\mu, \tag{A17}$$

$$m^\mu = \eth_0 l^\mu, \quad \bar{m}^\mu = \bar{\eth}_0 l^\mu, \tag{A18}$$

we have constructed a null tetrad system (l^μ, n^μ, m^μ, $\bar{m}^\mu$) for Minkowski space, which satisfies the standard orthonormality conditions, namely,

$$l^\mu n_\mu = -m^\mu \bar{m}_\mu = 1, \tag{A19}$$

with all other scalar products equal to zero. In terms of this null tetrad, the Minkowski metric $\eta_{\mu\nu}$ can be expressed as

$$\eta_{\mu\nu} = l_\mu n_\nu + n_\mu l_\nu - m_\mu \bar{m}_\nu - \bar{m}_\mu m_\nu. \tag{A20}$$

Now, consider two arbitrary vectors $A^\mu(u)$ and $B^\mu(u)$ which are attached to the world line and let

$$A = A^\mu l_\mu, \quad B = B^\mu l_\mu. \tag{A21}$$

Multiplying (A20) by $A^\mu B^\nu$ and using (A17) and (A18), we obtain

$$A^\mu B_\mu = 2AB + A\eth_0 \bar{\eth}_0 B + B\eth_0 \bar{\eth}_0 A - \eth_0 A \cdot \bar{\eth}_0 B - \bar{\eth}_0 A \cdot \eth_0 B. \tag{A22}$$

In particular,

$$\dot{\xi}^\mu \dot{\xi}_\mu = 2 = 2(v^2 + v\eth_0\bar{\eth}_0 v - \eth_0 v \cdot \bar{\eth}_0 v)$$

or

$$v^2 + v\eth_0\eth_0 v - \eth_0 v \cdot \bar{\eth}_0 v = 1. \tag{A23}$$

This last equation simply states that the Gaussian curvature $\mathcal{K}$ of the 2-surface, whose metric is given by $g^\circ_{AB}dx^A\,dx^B = (d\zeta\,d\bar{\zeta}/P_0^2v^2)$, is equal to 1, i.e.,

$$\mathcal{K} = v^2\eth_0\bar{\eth}_0 \log (P_0 v) = 1. \tag{A24}$$

From the results of Sec. 5 we note that under a proper Lorentz transformation of the original Minkowski coordinates l^μ and v transform according to

$$l'^\mu = Kl^\mu, \quad v' = Kv, \tag{A25}$$

that is, v has conformal weight 1. For completeness, let us also consider the transformation of the vectors m^μ and n^μ. Putting $m'^\mu = \eth'_0 l'^\mu$ and using (A25) we have

$$\begin{aligned} m'^\mu &= 2P_0' \frac{\partial}{\partial\zeta'}(Kl^\mu) = 2\frac{P_0'}{P_0} P_0 \frac{\partial\zeta}{\partial\zeta'}\frac{\partial}{\partial\zeta}(Kl^\mu) \\ &= \frac{P_0'}{P_0}\frac{\partial\zeta}{\partial\zeta'}\eth_0(Kl^\mu). \end{aligned}$$

With the help of (2.19), we then get

$$m'^\mu = e^{i\lambda}(m^\mu + l^\mu \eth_0 \log K). \tag{A26}$$

Similarly,

$$\begin{aligned} n'^\mu &= l'^\mu + \eth'_0\bar{\eth}'_0 l'^\mu \\ &= Kl^\mu + K^{-2}\eth_0\eth_0(Kl^\mu) \\ &= K^{-1}[n^\mu + m^\mu\bar{\eth}_0 \log K + \bar{m}^\mu\eth_0 \log K \\ &\quad + l^\mu(K^2 - 1 + \eth_0\bar{\eth}_0 \log K \\ &\quad + \eth_0 \log K \cdot \bar{\eth}_0 \log K)]. \end{aligned} \tag{A27}$$

We can simplify (A27) by using the identity

$$\eth_0\bar{\eth}_0 \log K = 1 - K^2, \tag{A28}$$

which can be easily proven from the definition of K given in (2.18). Thus, under a proper Lorentz transformation the tetrad vectors transform as

$$l'^\mu = Kl^\mu, \tag{A29a}$$

$$m'^\mu = e^{i\lambda}(m^\mu + Hl^\mu), \tag{A29b}$$

$$n'^\mu = K^{-1}(n^\mu + \bar{H}m^\mu + H\bar{m}^\mu + H\bar{H}l^\mu), \tag{A29c}$$

with $H \equiv \eth_0 \log K$.

With this background, we now give a simple generalization of the relativistic equations of motion

$$\tfrac{1}{2}m\ddot{\xi}^\mu(u) = F^\mu(u). \tag{A30}$$

(The extra factor $\frac{1}{2}$ comes from the fact that our u is $\sqrt{2}/2$ times the proper time.) Contracting (A30) with l^μ, we immediately obtain the equivalent statement, namely,

$$\tfrac{1}{2}m\dot{v} = F(u, \zeta, \bar{\zeta}), \tag{A31}$$

with $F = F^\mu l_\mu$. By leaving (A31) unchanged in form but now considering v and F to be associated with the infinite-dimensional representation $s = 0$, $w = 1$

$$\left(\text{i.e.,}\quad v = \sum_{l=0}^{\infty} v_{lm}Y_{lm} \quad\text{and}\quad F = \sum_{l=0}^{\infty} F_{lm}Y_{lm}\right),$$

we have generalized (A30). The v_{00}, v_{1m}, and F_{00}, F_{1m} are related to the usual 4-velocity and 4-force as in (A12); the higher l components v_{lm} can be associated with the internal structure or moments, with the higher F_{lm} acting as the driving "force" for these moments.

Equation (A31) can be generalized further by applying the results of Sec. 4 and recalling that v has $s = 0$ and $w = 1$. It can be easily verified from (4.3) and (4.4) that the expression

$$\alpha v + \beta v^4 \eth_0^2\bar{\eth}_0^2 v + \gamma v^3 \eth_0^2 v \cdot \bar{\eth}_0^2 v, \tag{A32}$$

where α, β, and γ are constants, also has $s = 0$ and $w = 1$. Though other expressions with this property can be constructed, (A32) is the simplest. By adding this to (A31) we obtain

$$\tfrac{1}{2}m\dot{v} = F + \alpha v + \beta v^4 \eth_0^2\bar{\eth}_0^2 v + \gamma v^3 \eth_0^2 v \cdot \bar{\eth}_0^2 v \tag{A33}$$

as a nonlinear relativistically invariant equation of motion for a particle with internal degrees of freedom.

A remarkable coincidence arises if we specialize to $\alpha = 0$, $\gamma = -\beta = (12\sqrt{2}\,\kappa)^{-1}$ (κ being the gravitational constant) and treat the particle as free, i.e., $F = 0$. The resulting equation

$$m\dot{v} = (6\sqrt{2}\,\kappa)^{-1} v^3(\eth_0^2 v \cdot \bar{\eth}_0^2 v - v\eth_0^2\bar{\eth}_0^2 v) \tag{A34}$$

is identical to the Robinson–Trautman equation,[9,10] which arises in the study of algebraically special solutions of the vacuum Einstein field equations. In several recent papers,[11,12] we have argued, from a slightly different point of view, that this equation (A34) represents equations of motion.

APPENDIX B

A function f which obeys $L^2 f = l(l+1)f$ is considered to have an l value. In general the product of two such functions does not have an l value. It is well known,though, that this product may be decomposed into a sum of terms, using the Clebsch–Gordan coefficients, each of which does have an l value. In this appendix, we present a method of obtaining this decomposition which requires neither knowledge of the Clebsch–Gordan coefficients nor of the azimuthal decomposition of the individual functions.

As an example, consider two functions f and g, both with l value one. Using the operator $\eth$, we may write

$$fg = 1/6(2fg + \eth f\bar{\eth}g + \bar{\eth}f\eth g) + 1/6(4fg - \eth f\bar{\eth}g - \bar{\eth}f\eth g). \quad \text{(B1)}$$

Considered as functions, the two bracketed expressions are easily seen (using the equivalence $L^2 = -\eth\bar{\eth}$ for spin weight zero quantities) to have l values 0 and 2, respectively. We shall now prove that this decomposition can be obtained generally and give a method for its construction.

From the properties of the Wigner D matrices and their relation[2] to the spin-s spherical harmonics, it is easily shown that the product of two spin-s spherical harmonics is given by

$$ {}_{s_1}Y_{l_1m_1}(\theta, \phi)\, {}_{s_2}Y_{l_2m_2}(\theta, \phi) = \sum_l \left(\frac{(2l_1 + 1)(2l_2 + 1)}{4\pi(2l + 1)}\right)^{\frac{1}{2}} {}_sY_{lm}(\theta, \phi) \times \langle l_1, l_2; m_1, m_2 \mid l, m\rangle\langle l_1, l_2; -s_1, -s_2 \mid l, -\tilde{s}\rangle, \quad \text{(B2)}$$

where $\langle l_1, l_2; m_1, m_2 \mid l, m\rangle$ is a Clebsch–Gordan coefficient of the rotation group[13]

$$m = m_1 + m_2,$$
$$\tilde{s} = s_1 + s_2,$$
$$|l_1 - l_2| \leq l \leq |l_1 + l_2|.$$

Choosing $s_1 = -s_2 = s$ $(s > 0)$, we rewrite (B2) as

$$\eth^s Y_{l_1m_1}\bar{\eth}^s Y_{l_2m_2} = (-1)^s \sum_l \left(\frac{(l_1 + s)!\,(l_2 + s)!\,(2l_1 + 1)(2l_2 + 1)}{(l_1 - s)!\,(l_2 - s)!\,4\pi(2l + 1)}\right)^{\frac{1}{2}} \times Y_{lm}\langle l_1, l_2; m_1, m_2 \mid l, m\rangle\langle l_1, l_2; -s, s \mid l, 0\rangle. \quad \text{(B3a)}$$

For $s' < 0$, (B2) becomes

$$\bar{\eth}^{-s'} Y_{l_1m_1}\eth^{-s'} Y_{l_2m_2} = (-1)^{-s'} \sum_l \left(\frac{(l_1 - s')!\,(l_2 - s')!\,(2l_1 + 1)(2l_2 + 1)}{(l_1 + s')!\,(l_2 + s')!\,4\pi(2l + 1)}\right)^{\frac{1}{2}} \times Y_{lm}\langle l_1, l_2; m_1, m_2 \mid l, m\rangle\langle l_1, l_2; s', -s' \mid l, 0\rangle. \quad \text{(B3b)}$$

Assuming that $l_1 \geq l_2$, then (B3) represents $2l_2 + 1$ independent equations.

If f and g have l values l_1 and l_2, respectively, they may be written

$$f = \sum_{m_1} a_{m_1}Y_{l_1m_1}, \quad g = \sum_{m_2} b_{m_2}Y_{l_2m_2}, \quad -l_i \leq m_i \leq l_i. \quad \text{(B4)}$$

Multiplying (B3a) by a_{m_1} and b_{m_2} and summing over m_1 and m_2, we get

$$\eth^s f\bar{\eth}^s g = (-1)^s \sum_l \left(\frac{(l_1 + s)!\,(l_2 + s)!\,(2l_1 + 1)(2l_2 + 1)}{(l_1 - s)!\,(l_2 - s)!\,4\pi(2l + 1)}\right)^{\frac{1}{2}} \times \langle l_1, l_2; -s, s \mid l, 0\rangle \times \sum_{m_1, m_2} \langle l_1, l_2; m_1, m_2 \mid l, m\rangle a_{m_1}b_{m_2}Y_{lm} = \sum_l \alpha_{sl}D^l(\zeta, \bar{\zeta}), \quad \text{(B5)}$$

where α_{sl} are constants (see below) and D^l satisfies

$$[\eth\bar{\eth} + l(l + 1)]D^l(\zeta, \bar{\zeta}) = 0. \quad \text{(B6)}$$

Similarly, from (B3b), we obtain

$$\bar{\eth}^s f\eth^s g = \sum_l (-1)^{l_1+l_2+l}\alpha_{sl}D^l, \quad \text{(B5')}$$

where we have set $s' = -s$ in (B3b),so that $0 \leq s \leq l_2$, and have made use of the identity[13]

$$\langle l_1, l_2; s, -s \mid l, 0\rangle = (-)^{l_1+l_2+l}\langle l_1, l_2; -s, s \mid l, 0\rangle.$$

[The indices l_1 and l_2 have been suppressed in (B5).] The normalization of D^l is not uniquely determined, as any variation results in a corresponding variation of the constants α_{sl}. The expansion we seek then is

$$fg = \sum_l \alpha_{0l}D^l = \sum_l \alpha_l D^l. \quad \text{(B7)}$$

Equations (B5) and (B5') form a set of linear equations for the D^l's in terms of $\eth^s f\,\bar{\eth}^s g$ and $\bar{\eth}^s f\,\eth^s g$, which may be inverted to give expressions for the D^l's of the form

$$D^l = \sum_s a_{ls}\eth^s f\bar{\eth}^s g + b_{ls}\bar{\eth}^s f\eth^s g. \quad \text{(B8)}$$

From Eqs. (B5) it can be seen that,while the expression $\eth^s f\bar{\eth}^s g$ does not have good parity, the combination $\eth^s f\bar{\eth}^s g \pm \bar{\eth}^s f\eth^s g$ has parity $\pm(-1)^{l_1+l_2}$. Since D^l has parity $(-1)^l$, (B8) must be of the form

$$D^l = \sum_s a_{ls}[\eth^s f\bar{\eth}^s g + (-1)^{l_1+l_2+l}\bar{\eth}^s f\eth^s g]. \quad \text{(B9)}$$

In order to generate the coefficients a_{ls}, we make use of the easily established identity (valid for $0 \leq s \leq l_2 - 1$):

$$\eth\bar{\eth}(\eth^s f\bar{\eth}^s g) = \eth^{s+1}f\bar{\eth}^{s+1}g - [l_1(l_1 + 1) + l_2(l_2 + 1) - 2s^2]\eth^s f\bar{\eth}^s g + [l_1(l_1 + 1) - s(s - 1)] \times [l_2(l_2 + 1) - s(s - 1)]\eth^{s-1}f\bar{\eth}^{s-1}g. \quad \text{(B10)}$$

Substituting (B9) into (B6) and using (B10), we equate the coefficient of $\eth^{s+1}f\bar{\eth}^{s+1}g$ to 0 and obtain the recursion relation

$$\begin{aligned} a_{ls} = {} & [l_1(l_1+1) + l_2(l_2+1) \\ & - 2(s+1)^2 - l(l+1)]a_{l,s+1} \\ & - [l_1(l_1+1) - (s+1)(s+2)] \\ & \times [l_2(l_2+1) - (s+1)(s+2)]a_{l,s+2}, \end{aligned} \tag{B11}$$

where $0 \leq s \leq l_2$ and a_{ll_2} is arbitrary, reflecting the arbitrariness in the choice of D^l. Using these relations, it is a simple matter to generate all the expressions D^l of the same parity as fg and hence the final form

$$fg = \sum_l \alpha_l D^l.$$

As indicated above, expressions D^l with parity opposite to that of fg are easily calculated. A case of particular interest is that when $l_1 = l_2 = l = 1$. Let **f**, **g**, and **h** be three vectors such that

$$\mathbf{f} \times \mathbf{g} = \mathbf{h}. \tag{B12}$$

We form the function $f(\zeta, \bar{\zeta})$ associated with the vector f by

$$\begin{aligned} f(\zeta, \bar{\zeta}) = {} & \left(\frac{2\pi}{3}\right)^{\frac{1}{2}} [-(f_x - if_y)Y_{11}(\zeta, \bar{\zeta}) \\ & + (f_x + if_y)Y_{1,-1}(\zeta, \bar{\zeta}) + \sqrt{2}f_z Y_{10}(\zeta, \bar{\zeta})] \end{aligned} \tag{B13}$$

with similar expressions for $g(\zeta, \bar{\zeta})$ and $h(\zeta, \bar{\zeta})$ formed from **g** and **h**, respectively. The functions f, g, and h all have l value equal to one.

From the above it follows that $\eth f\bar{\eth}g - \bar{\eth}f\eth g$ has l value to one. In fact,

$$(i/2)(\eth f\bar{\eth}g - \bar{\eth}f\eth g) = h, \tag{B12'}$$

giving an expression for the cross product of two vectors expressed in our formalism.

ACKNOWLEDGMENT

Many of the ideas appearing in this paper arose during discussions with Professor R. Penrose,to whom the authors are deeply indebted.

* This research has been supported by the Aerospace Research Laboratory, OAR, U.S. Air Force, Contract F33615-70-C-1081.

[1] E. T. Newman and R. Penrose, J. Math. Phys. **7**, 863 (1966).
[2] J. N. Goldberg *et al.*, J. Math. Phys. **8**, 2155 (1967).
[3] See, for example, F. Gürsey, *Relativity, Groups, and Topology*, C. DeWitt and B. DeWitt, Eds. (Gordon and Breach, New York, 1964), p. 152.
[4] M. A. Naimark, *Linear Representation of the Lorentz Group* (Pergamon, Oxford, 1964).
[5] A. Staruszkiewicz, J. Math. Phys. **8**, 2221 (1967).
[6] I. M. Gel'fand, M. I. Graev, and N. Ya. Vilenkin, *Generalized Functions* (Academic, New York, 1966), Vol. V.
[7] E. T. Newman and T. Unti, J. Math. Phys. **3**, 891 (1962).
[8] R. K. Sachs, Proc. Roy. Soc. (London) **A270**, 103 (1962).
[9] I. Robinson and A. Trautman, Proc. Roy. Soc. (London) **A265**, 463 (1962).
[10] E. T. Newman and L. Tamburino, J. Math. Phys. **3**, 902 (1962).
[11] E. T. Newman and R. Posadas, Phys. Rev. Letters **22**, 1196 (1969).
[12] E. T. Newman and R. Posadas, Phys. Rev. **187**, 1784 (1969).
[13] D. Brink and G. Satchler, *Angular Momentum* (Oxford U.P., Oxford, 1962).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

New Approach to the Motion of a Pole–Dipole Particle*

E. T. NEWMAN
University of Pittsburgh, Pittsburgh, Pennsylvania 15213
AND
ROGER YOUNG
Kings College, University of London, London

(Received 29 April 1970)

A new approach to the equations of motion of the pole–dipole particle in linearized general relativity is presented. It is based on earlier work for the monopole particle [E. T. Newman and R. Posadas, Phys. Rev. Letters **22**, 1196 (1969); Phys. Rev. **187**, 1784 (1969)]. Though the results are not new, the novelty of the method holds out the serious hopes that the work can be extended to the full theory.

I. INTRODUCTION

Recently a new approach to the subject of equations of motion in general relativity was presented.[1,2] This approach, using the spin-coefficient formalism,[3,4] describes the motion of a singularity (suitably defined) in terms of the behavior of the family of null cones emanating from the singularity. The resulting description is intrinsic to the singular space–time and does not rely on an assumed regular background space.

A novel consequence of this work was that not only did it predict exact equations of motion (e.g., the radiation reaction force and the Lorentz force when

Substituting (B9) into (B6) and using (B10), we equate the coefficient of $\eth^{s+1}f\bar{\eth}^{s+1}g$ to 0 and obtain the recursion relation

$$\begin{aligned} a_{ls} = {} & [l_1(l_1+1) + l_2(l_2+1) \\ & - 2(s+1)^2 - l(l+1)]a_{l,s+1} \\ & - [l_1(l_1+1) - (s+1)(s+2)] \\ & \times [l_2(l_2+1) - (s+1)(s+2)]a_{l,s+2}, \end{aligned} \tag{B11}$$

where $0 \leq s \leq l_2$ and a_{ll_2} is arbitrary, reflecting the arbitrariness in the choice of D^l. Using these relations, it is a simple matter to generate all the expressions D^l of the same parity as fg and hence the final form

$$fg = \sum_l \alpha_l D^l.$$

As indicated above, expressions D^l with parity opposite to that of fg are easily calculated. A case of particular interest is that when $l_1 = l_2 = l = 1$. Let **f**, **g**, and **h** be three vectors such that

$$\mathbf{f} \times \mathbf{g} = \mathbf{h}. \tag{B12}$$

We form the function $f(\zeta, \bar{\zeta})$ associated with the vector f by

$$\begin{aligned} f(\zeta, \bar{\zeta}) = {} & \left(\frac{2\pi}{3}\right)^{\frac{1}{2}} [-(f_x - if_y)Y_{11}(\zeta, \bar{\zeta}) \\ & + (f_x + if_y)Y_{1,-1}(\zeta, \bar{\zeta}) + \sqrt{2}f_z Y_{10}(\zeta, \bar{\zeta})] \end{aligned} \tag{B13}$$

with similar expressions for $g(\zeta, \bar{\zeta})$ and $h(\zeta, \bar{\zeta})$ formed from **g** and **h**, respectively. The functions f, g, and h all have l value equal to one.

From the above it follows that $\eth f\bar{\eth}g - \bar{\eth}f\eth g$ has l value to one. In fact,

$$(i/2)(\eth f\bar{\eth}g - \bar{\eth}f\eth g) = h, \tag{B12'}$$

giving an expression for the cross product of two vectors expressed in our formalism.

ACKNOWLEDGMENT

Many of the ideas appearing in this paper arose during discussions with Professor R. Penrose,to whom the authors are deeply indebted.

* This research has been supported by the Aerospace Research Laboratory, OAR, U.S. Air Force, Contract F33615-70-C-1081.

[1] E. T. Newman and R. Penrose, J. Math. Phys. **7**, 863 (1966).
[2] J. N. Goldberg *et al.*, J. Math. Phys. **8**, 2155 (1967).
[3] See, for example, F. Gürsey, *Relativity, Groups, and Topology*, C. DeWitt and B. DeWitt, Eds. (Gordon and Breach, New York, 1964), p. 152.
[4] M. A. Naimark, *Linear Representation of the Lorentz Group* (Pergamon, Oxford, 1964).
[5] A. Staruszkiewicz, J. Math. Phys. **8**, 2221 (1967).
[6] I. M. Gel'fand, M. I. Graev, and N. Ya. Vilenkin, *Generalized Functions* (Academic, New York, 1966), Vol. V.
[7] E. T. Newman and T. Unti, J. Math. Phys. **3**, 891 (1962).
[8] R. K. Sachs, Proc. Roy. Soc. (London) **A270**, 103 (1962).
[9] I. Robinson and A. Trautman, Proc. Roy. Soc. (London) **A265**, 463 (1962).
[10] E. T. Newman and L. Tamburino, J. Math. Phys. **3**, 902 (1962).
[11] E. T. Newman and R. Posadas, Phys. Rev. Letters **22**, 1196 (1969).
[12] E. T. Newman and R. Posadas, Phys. Rev. **187**, 1784 (1969).
[13] D. Brink and G. Satchler, *Angular Momentum* (Oxford U.P., Oxford, 1962).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

New Approach to the Motion of a Pole–Dipole Particle*

E. T. NEWMAN
University of Pittsburgh, Pittsburgh, Pennsylvania 15213
AND
ROGER YOUNG
Kings College, University of London, London

(Received 29 April 1970)

A new approach to the equations of motion of the pole–dipole particle in linearized general relativity is presented. It is based on earlier work for the monopole particle [E. T. Newman and R. Posadas, Phys. Rev. Letters **22**, 1196 (1969); Phys. Rev. **187**, 1784 (1969)]. Though the results are not new, the novelty of the method holds out the serious hopes that the work can be extended to the full theory.

I. INTRODUCTION

Recently a new approach to the subject of equations of motion in general relativity was presented.[1,2] This approach, using the spin-coefficient formalism,[3,4] describes the motion of a singularity (suitably defined) in terms of the behavior of the family of null cones emanating from the singularity. The resulting description is intrinsic to the singular space–time and does not rely on an assumed regular background space.

A novel consequence of this work was that not only did it predict exact equations of motion (e.g., the radiation reaction force and the Lorentz force when

the electromagnetic field is coupled in), but it also predicted a type of internal structure for the singularity as well as the dynamical laws governing the time development of this structure. Though it was possible to interpret this structure as being loosely related to the mass (or "electric" type) multipole moments, there was nothing that could play the role of the spin (or "magnetic" type) moments. In particular, there was no means of treating the motion of a singularity with intrinsic angular momentum.

It is the purpose of this article to try to remedy this defect. (Though we here work only with the linearized Einstein equations, we believe that the method can be extended to the full nonlinear theory.) By first integrating the Bianchi identities in flat space (using a null coordinate system associated with an arbitrary timelike world line) with the condition that no moments[5] higher than the dipole should exist, two functions of integration, representing the mass and the complex dipole moment (mass dipole and angular momentum), are obtained. Then, by effectively using the condition that the angular behavior of the linearized metric be regular (i.e., be expandable in generalized spherical harmonics), conditions on the motion of the singularity and the time development of the moments are derived, and then shown to be identical with the well-known relativistic equations of motion for the pole–dipole particle.[6,7]

The authors believe that though the results are not new, the method of approach is sufficiently novel and holds out the strong enough hope for generalization to the full nonlinear theory to justify the present description.

2. THE NULL COORDINATE SYSTEM

As in Refs. 1–4, we introduce a null coordinate system $x^a = (u, r, x^A)$, $a = 0, 1, 2, 3$, $A = 2, 3$, and an adapted null tetrad $(l^a, n^a, m^a, \bar{m}^a)$ and apply it in flat space–time. The coordinates and tetrad are to be based on an arbitrary timelike world line $c: y^a = \xi^a(u)$, where y^a are Minkowski coordinates and the time coordinate u is normalized according to

$$\dot{\xi}^2 = 2, \tag{1}$$

and the affine parameter r chosen to satisfy $l^a\dot{\xi}_a = 1$. Further specialization results in a metric[1,2]

$$ds^2 = 2\left(1 - \frac{\dot{P}}{P}r\right) du^2 + 2\, du\, dr - \frac{r^2}{2P^2}\, d\zeta\, d\bar{\zeta}, \qquad \zeta = x^2 - ix^3, \tag{2}$$

and a null tetrad

$$l^a = \delta^a_1, \quad n^a = \delta^a_0 + \left(-1 + \frac{\dot{P}}{P}r\right)\delta^a_1, \quad m^a = -\frac{P}{r}(1, i)\delta^a_A. \tag{3}$$

The quantity $P(u, x^A)$ has the interpretation[1,2] that $\dot{P}/P$ is in one-to-one correspondence with the acceleration vector $\ddot{\xi}^a$.

The zeroth-order spin coefficients are then calculated from the field equations, with the Weyl tensor ψ_n $(n = 0 - 4)$ set equal to zero [see Ref. (2), Eqs. (B1) and (B5) with $\phi = \psi = 0$ and $K = 1$].

The following notation will be used: A superscript zero (e.g., ψ_1^0) indicates that the quantity concerned is independent of r; a subscript zero sign (e.g., $\eth_0$) will be used, where necessary, to distinguish zeroth-order from first-order quantities (this is done only in Sec. 5); if s^a and t^a are 4-vectors, then s^2, (st) denote the inner products $s^a s_a$, $s^a t_a$, respectively; and $[(s)l]$ is used to denote a quantity[8,9] whose angular behavior is like a spin-weight s, angular momentum l, spherical harmonic ${}_sY_{lm}$.

3. FIRST-ORDER CALCULATIONS

The linearized Bianchi identities for the first-order Weyl tensor depend only on zeroth-order spin coefficients and are written

$$D\psi_n + (5 - n)\frac{\psi_n}{r} = \frac{-\bar{\eth}\psi_{n+1}}{r}, \quad n = 1, 2, 3, 4, \tag{4}$$

$$\dot{\psi}_n + \left(-1 + \frac{\dot{P}}{P}r\right)D\psi_n + \left[(2 - n)\frac{\dot{P}}{P}r - 1 - n\right]\frac{\psi_n}{r} = \frac{-\eth\psi_{n+1}}{r} + n\psi_{n-1}\bar{\eth}\frac{\dot{P}}{P}, \quad n = 0, 1, 2, 3. \tag{5}$$

Here $D \equiv \partial/\partial r$ and[8,9]

$$\eth\eta = 2P^{1-s}\frac{\partial}{\partial\zeta}P^s\eta, \qquad \bar{\eth}\eta = 2P^{1+s}\frac{\partial}{\partial\bar{\zeta}}P^{-s}\eta, \tag{6}$$

where $\zeta = x^2 - ix^3$. ψ_n has spin weight $2 - n$.

In the usual multipole interpretation,[5] ψ_0 is specified by giving the quadrupole and all higher moments, so the *dipole assumption* amounts to

$$\psi_0 = 0. \tag{7}$$

The radial integration of (4) can be carried out explicitly:

$$\begin{aligned}
\psi_1 &= \frac{\psi_1^0}{r^4}, \\
\psi_2 &= \frac{\psi_2^0}{r^3} + \frac{\bar{\eth}\psi_1^0}{r^4}, \\
\psi_3 &= \frac{\psi_3^0}{r^2} + \frac{\bar{\eth}\psi_2^0}{r^3} + \frac{\bar{\eth}^2\psi_1^0}{2r^4}, \\
\psi_4 &= \frac{\psi_4^0}{r} + \frac{\bar{\eth}\psi_3^0}{r^2} + \frac{\bar{\eth}^2\psi_2^0}{2r^3} + \frac{\bar{\eth}^3\psi_1^0}{6r^4},
\end{aligned} \tag{8}$$

where the ψ_n^0, $n = 1$–4, are constants of integration. Substituting the expressions (8) for ψ_n into the propagation Eqs. (5) and equating coefficients of like powers of r, we obtain

$$\eth\psi_1^0 = 0, \tag{9a}$$

$$\eth\psi_2^0 = 3\frac{\dot{P}}{P}\psi_1^0 - \dot{\psi}_1^0, \tag{9b}$$

$$\eth\psi_3^0 = 3\frac{\dot{P}}{P}\psi_2^0 - \dot{\psi}_2^0, \tag{9c}$$

$$\eth\psi_4^0 = 3\frac{\dot{P}}{P}\psi_3^0 - \dot{\psi}_3^0. \tag{9d}$$

Since $\eth[(s)s] = 0$, Eq. (9a) states that ψ_1^0 is a $[(1)1]$ quantity—in some sense the dipole moment. The Eqs. (9b)–(9d) can be regarded as determining ψ_2^0, ψ_3^0, ψ_4^0 in terms of ψ_1^0 (and possibly arbitrary constants); Eqs. (8) then yield the complete solution for ψ_n.

The general method of solution of Eqs. (9) can be illustrated by its application to the second of these equations. [Essentially the technique is to analyze each equation according to its various angular parts (i.e., l values) and then use the raising properties ($s \to s + 1$) of $\eth$. References 6, 7, and Appendix A of Ref. 2 will prove helpful in following the manipulations through in detail.] Write

$$\eth\psi_2^0 = 3\frac{\dot{P}}{P}\psi_1^0 - \dot{\psi}_1^0 = A_2 + B_2,$$

i.e.,

$$\eth[(0)] = [(0)1] \times [(1)1] - \frac{\partial}{\partial u}[(1)1] = [(1)1] + [(1)2].$$

Then

$$\eth B_2 = 4\psi_1^0\eth\frac{\dot{P}}{P}.$$

{Note $\eth[(s)s] = 0$, but $\eth(\partial/\partial u[(s)s]) \neq 0$, or more generally, the time derivative of an $[(s)l]$ quantity is not an $[(s)l]$ quantity. We give here three useful general formulas for any (s) quantity η:

$$\begin{aligned}
\frac{\partial}{\partial u}\eth^n\eta &= n\frac{\dot{P}}{P}\eth^n\eta \\
&\quad + n(n + s - 1)\eth\frac{\dot{P}}{P}\cdot\eth^{n-1}\eta + \eth^n\dot{\eta}, \\
\frac{\partial}{\partial u}\bar{\eth}^n\eta &= n\frac{\dot{P}}{P}\bar{\eth}^n\eta \\
&\quad + n(n - s - 1)\bar{\eth}\frac{\dot{P}}{P}\cdot\bar{\eth}^{n-1}\eta + \bar{\eth}^n\dot{\eta}, \\
\frac{\partial}{\partial u}\eth\bar{\eth}\eta &= 2s\frac{\dot{P}}{P}\eta + 2\frac{\dot{P}}{P}\eth\bar{\eth}\eta \\
&\quad + s\eth\frac{\dot{P}}{P}\cdot\bar{\eth}\eta - s\eth\eta\cdot\bar{\eth}\frac{\dot{P}}{P} + \eth\bar{\eth}\dot{\eta}.
\end{aligned}\Bigg\} \tag{10}$$

Continuing the analysis:

$$\bar{\eth}\eth B_2 = -4B_2 = 4\bar{\eth}\left(\psi_1^0\cdot\eth\frac{\dot{P}}{P}\right),$$

$$B_2 = 2\psi_1^0\frac{\dot{P}}{P} - \bar{\eth}\psi_1^0\cdot\eth\frac{\dot{P}}{P},$$

$$\begin{aligned}
A_2 &= \eth\psi_2^0 - B_2 \\
&= \psi_1^0\frac{\dot{P}}{P} + \bar{\eth}\psi_1^0\cdot\eth\frac{\dot{P}}{P} - \psi_1^0.
\end{aligned}$$

Thus

$$\psi_2^0 = \quad m \quad + \quad a \quad + \quad b \quad ,$$

i.e.,

$$[(0)] = [(0)0] + [(0)1] + [(0)2], \tag{11}$$

where m is an arbitrary $[(0)0]$ function of integration (a function of u), $a \equiv \frac{1}{2}\bar{\eth}A_2$, and $b = -\frac{1}{6}\bar{\eth}B_2$. It is convenient to introduce the $[(0)1]$ quantity G defined by $\psi_1^0 \equiv \eth G$. In terms of G,

$$a = \bar{\eth}G\eth\frac{\dot{P}}{P} - \dot{G}, \tag{12}$$

$$b = \tfrac{4}{3}G\frac{\dot{P}}{P} - \tfrac{1}{3}\eth G\bar{\eth}\frac{\dot{P}}{P} - \tfrac{1}{3}\bar{\eth}G\eth\frac{\dot{P}}{P}. \tag{13}$$

Proceeding along similar lines with Eqs. (9c):

$$\eth\psi_3^0 = \quad M \quad + \quad A_3 \quad + \quad B_3 \quad + \quad C_3 \quad ,$$

i.e.,

$$\eth[(-1)] = [(0)0] + [(0)1] + [(0)2] + [(0)3], \tag{14}$$

where

$$M = -\dot{m} + \tfrac{1}{3}a\frac{\dot{P}}{P} + \tfrac{1}{6}\eth a\bar{\eth}\frac{\dot{P}}{P} + \tfrac{1}{6}\bar{\eth}a\eth\frac{\dot{P}}{P}, \tag{15a}$$

$$A_3 = 3m\frac{\dot{P}}{P} - \dot{a} + \tfrac{1}{2}\eth a\bar{\eth}\frac{\dot{P}}{P} + \tfrac{1}{2}\bar{\eth}a\eth\frac{\dot{P}}{P}, \tag{15b}$$

$$\begin{aligned}
B_3 &= \tfrac{8}{3}a\frac{\dot{P}}{P} - \tfrac{2}{3}\eth a\bar{\eth}\frac{\dot{P}}{P} - \tfrac{2}{3}\bar{\eth}a\eth\frac{\dot{P}}{P} - \dot{b} \\
&\quad + \tfrac{1}{2}\eth b\bar{\eth}\frac{\dot{P}}{P} + \tfrac{1}{2}\bar{\eth}b\eth\frac{\dot{P}}{P},
\end{aligned} \tag{15c}$$

$$C_3 = 3b\frac{\dot{P}}{P} - \tfrac{1}{2}\eth b\bar{\eth}\frac{\dot{P}}{P} - \tfrac{1}{2}\bar{\eth}b\eth\frac{\dot{P}}{P}. \tag{15d}$$

Since ψ_3^0 is a (-1) quantity, we must have

$$M = 0, \tag{16}$$

which is the first of the equations of motion. In the case of the pure mass monopole (i.e., $G = 0$), this reduces immediately to the mass conservation law $\dot{m} = 0$.

As will be shown later, further restrictions on ψ_2^0 and ψ_3^0 follow as a consequence of the first-order field

equations:

$$A_3 = 0, \tag{17}$$

$$\text{Im}\,(a) = 0, \tag{18}$$

$$\text{Im}\,(m) = 0. \tag{19}$$

Equations (16)–(18) in fact constitute the equations of motion[6,7] for a pole–dipole particle. Equation (19) forbids a "spin" monopole or "magnetic" type monopole. For a pure monopole ($G = 0$), Eq. (17) implies

$$\dot{P}/P = 0, \tag{20}$$

so that the monopole is constrained to have a time-independent mass and move along a geodesic in flat-space.

The radiation field ψ_4^0 consists of a linear combination of (−2)2, (−2)3, and (−2)4 quantities, and may be calculated from Eq. (9d):

$$-\psi_4^0 = \tfrac{1}{12}\bar{\eth}\frac{\dot{P}}{P}\bar{\eth}B_3 + \tfrac{1}{24}\bar{\eth}^2\dot{B}_3 + \tfrac{1}{60}\bar{\eth}\frac{\dot{P}}{P}\bar{\eth}C_3 + \tfrac{1}{120}\bar{\eth}^2\dot{C}_3.$$

(A similar result may be derived for the radiation field of the electromagnetic dipole moving along an arbitrary world line;

$$-\phi_2^0 = \tfrac{1}{2}\bar{\eth}\ddot{G}_e + \frac{1}{2}\frac{\dot{P}}{P}\bar{\eth}\dot{G}_e + \tfrac{1}{2}\bar{\eth}G_e\frac{\partial}{\partial u}\frac{\dot{P}}{P},$$

where now $\phi_0^0 = \eth G_e$, G_e representing the electric and magnetic dipole moments.)

4. THE POLE–DIPOLE EQUATIONS

In the present notation with the normalization $\dot{\xi}^2 = 2$, the pole–dipole equations of motion [Eqs. (5.3) and (5.7) of Ref. (6)] take the form

$$\dot{S}^{ab} + \dot{\xi}^{[a}\alpha^{b]} = 0, \tag{21}$$

$$\dot{p}^a = 0, \tag{22}$$

where S^{ab} is a skew-symmetric tensor, representing the intrinsic dipole moment and angular momentum, and

$$\alpha^a = \dot{S}^{ab}\dot{\xi}_b,$$

$$p^a = m\dot{\xi}^a + \alpha^a. \tag{23}$$

[Usually the additional condition that the dipole moment vanishes is imposed; namely

$$S^{ab}\dot{\xi}_b = 0. \tag{24}$$

Though we will not use (24), it will be mentioned later.]

Note that the full content of (21) is expressed in

$$\dot{S}^*_{ab}\dot{\xi}^b = 0, \tag{21'}$$

where the * represents the operation of taking the dual.

Before establishing the connection between (21) and (22) with the equations of the previous section, a few mathematical results are stated. If s^a and t^a are vector fields defined along c, orthogonal to the velocity vector $\dot{\xi}^a$, then contracting with Eq. (A3) of Ref. 2,we obtain

$$-(st) = 2ST + \eth S\bar{\eth}T + \eth T\bar{\eth}S, \tag{25}$$

with $S = (ls)$, $T = (lt)$. Further, it may be shown[10] that if

$$h^a = \frac{1}{\sqrt{2}}\epsilon^{abcd}\dot{\xi}_b s_c t_d \tag{26}$$

(which is the cross product of t with s in the 3-space normal to $\dot{\xi}$), then, defining $H = (lh)$,

$$H = \frac{i}{\sqrt{2}}(\eth S\bar{\eth}T - \bar{\eth}S\eth T). \tag{27}$$

Multiplying Eq. (22) with $\dot{\xi}^a$ and $l^a - \frac{1}{2}\dot{\xi}^a$, respectively, we obtain

$$2\dot{m} - (\alpha\ddot{\xi}) = 0,$$

$$m\frac{\dot{P}}{P} + (\dot{\alpha}l) + \tfrac{1}{2}(\alpha\ddot{\xi}) = 0.$$

By using (25) and the identification $a = -3(\alpha l)$, we obtain Eqs. (16) and (17).

The tensor S^*_{ab} can be represented by

$$S^*_{ab} = q_a\dot{\xi}_b - q_b\dot{\xi}_a + \epsilon_{abcd}r^c\dot{\xi}^d,$$

where $q_a = \frac{1}{2}S^*_{ab}\dot{\xi}^b$, $r_a = \frac{1}{2}S_{ab}\dot{\xi}^b$. From (21'),

$$2\dot{q}_a - \dot{q}_b\dot{\xi}^b\dot{\xi}_a + \epsilon_{abcd}r^c\ddot{\xi}^d\dot{\xi}^b = 0.$$

Multiplying by l_a and using (25), (26), and (27), we obtain

$$2i\dot{Q} + \left[\eth(R - iQ)\cdot\bar{\eth}\frac{\dot{P}}{P} - \bar{\eth}(R + iQ)\cdot\eth\frac{\dot{P}}{P}\right] = 0,$$

where $Q = (ql)$, $R = (rl)$. [Use has been made of the identity $(\dot{q}l) = \dot{Q} + Q(\dot{P}/P)$.] Finally setting $G = 6(R - iQ)$, Eq. (18) is obtained. With G so defined and $a = -3(\alpha l)$, the consistency Eq. (12) can be shown to be identically satisfied.

Note also that the vanishing of the dipole moment, i.e., Eq. 24, is equivalent to $R = 0$.

5. COMPLETE FIRST-ORDER SOLUTION

A subscript zero is now attached to all zeroth-order quantities, e.g., α_0^0, P_0, $\eth_0$. All spin coefficients and metric variables without a zero subscript are first-order quantities.

Using certain coordinate conditions,[4] the first-order radial field and metric equations can be integrated:

$$\rho = 0, \quad \kappa = \pi = \epsilon \equiv 0,$$

$$\sigma = \sigma^0 r^{-2},$$

$$\alpha = -\alpha^0 r^{-1} + \bar{\alpha}_0^0\bar{\sigma}^0 r^{-2},$$

$$\beta = \bar{\alpha}^0 r^{-1} - \alpha_0^0\sigma^0 r^{-2} - \tfrac{1}{2}\psi_1^0 r^{-3},$$

$$\tau = \bar{\alpha} + \beta,$$

$$\gamma = \gamma^0 - \tfrac{1}{2}\psi_2^0 r^{-2} + \tfrac{1}{6}(\alpha_0^0\psi_1^0 - \bar{\alpha}_0^0\bar{\psi}_1^0 - 2\bar{\eth}_0\psi_1^0)r^{-3},$$

$$\lambda = \lambda^0 r^{-1} + \bar{\sigma}^0 r^{-2},$$

$$\mu = \mu^0 r^{-1} - \psi_2^0 r^{-2} - \tfrac{1}{2}\bar{\eth}_0\psi_1^0 r^{-3},$$

$$
\begin{aligned}
\nu &= \nu^0 - \psi_3^0 r^{-1} - \tfrac{1}{2}\bar{\eth}_0\psi_2^0 r^{-2} + \tfrac{1}{6}(\bar{\psi}_1^0 - \bar{\eth}_0^2\psi_1^0)r^{-3},\\
\xi^A &= \xi^{0A} r^{-1} - \sigma^0\bar{\xi}_0^{0A} r^{-2},\\
\omega &= \omega^0 r^{-1} - \tfrac{1}{2}\psi_1^0 r^{-2},\\
X^A &= X^{0A} + \tfrac{1}{6}(\psi_1^0\bar{\xi}_0^{0A} + \bar{\psi}_1^0\xi_0^{0A})r^{-3},\\
U &= U^0 - (\gamma^0 + \bar{\gamma}^0)r\\
&\quad - \tfrac{1}{2}(\psi_2^0 + \bar{\psi}_2^0)r^{-1} - \tfrac{1}{6}(\bar{\eth}_0\psi_1^0 + \bar{\eth}\bar{\psi}_1^0)r^{-2}.
\end{aligned}
$$

Substituting in the linearized nonradial field and metric equations yields relations between the first-order constants of integration. The new first-order variable ξ^{0A} is taken as $-P_0 I(1, i)$ [see Eq. (3)]. Thus the zeroth-order P of the previous sections becomes the first (plus zeroth)-order $P_0(1 + I)$. It then follows from the field equations, using further coordinate and tetrad conditions [see e.g., Ref. (4)] that

$$
\begin{aligned}
&X^{0A} = 0, \quad \gamma^0 = -\tfrac{1}{2}I, \quad \nu^0 = I\bar{\eth}_0\frac{\dot{P}_0}{P_0} + \bar{\eth}_0 I,\\
&\alpha^0 = -\tfrac{1}{2}\eth_0 I + \tfrac{1}{2}I\bar{\eth}_0 \log P_0,\\
&\omega^0 = -\bar{\eth}_0\sigma^0, \quad \bar{\lambda}^0 = \dot{\sigma}^0 - \sigma^0\frac{\dot{P}_0}{P_0},\\
&-\mu^0 = -U^0 = K = 2I + \eth_0\bar{\eth}_0 I.
\end{aligned}
$$

In addition σ^0 and I must satisfy the first-order equations

$$\psi_2^0 - \bar{\psi}_2^0 = \bar{\eth}_0^2\sigma^0 - \eth_0^2\bar{\sigma}^0, \tag{28a}$$

$$\psi_3^0 = \eth_0(\lambda^0 + \bar{\eth}_0^2 I), \tag{28b}$$

$$\psi_4^0 = -\bar{\eth}_0^2\dot{I} - 2\bar{\eth}_0 I\bar{\eth}_0\frac{\dot{P}_0}{P_0} + 2\frac{\dot{P}_0}{P_0}\lambda^0 - \dot{\lambda}_0^0. \tag{28c}$$

Note that, though (28) satisfies (9), identically it imposes, by its structure, severe conditions on ψ_2^0 and ψ_3^0. Since σ^0 and $\lambda^0 + \bar{\eth}_0^2 I$ are spin-weight -2 quantities, it follows immediately that ψ_3^0 has no $[(-1)0]$ or $[(-1)1]$ part and that $\psi_2^0 - \bar{\psi}_2^0$ has no $[(0)0]$ and $[(0)1]$ part. These restrictions amount to nothing more than a restatement of Eqs. (16)–(19).

In addition (2b) implies

$$
\begin{aligned}
\eth\psi_3^0 &= \eth_0^2(\lambda^0 + \bar{\eth}_0^2 I)\\
&= \frac{\partial}{\partial u}\eth_0^2\bar{\sigma}^0 - 3\frac{\dot{P}_0}{P_0}\eth_0^2\bar{\sigma}^0 + \eth_0^2\bar{\eth}_0^2 I,
\end{aligned} \tag{29}
$$

and, from Eqs. (14)–(17) and (19),

$$
\begin{aligned}
\eth\psi_3^0 &= -\dot{b} + 3\frac{\dot{P}_0}{P_0}b\\
&\quad + \left(\tfrac{8}{3}a\frac{\dot{P}_0}{P_0} - \tfrac{2}{3}\eth_0 a\cdot\bar{\eth}_0\frac{\dot{P}_0}{P_0} - \tfrac{2}{3}\bar{\eth}_0 a\cdot\eth_0\frac{\dot{P}_0}{P_0}\right).
\end{aligned} \tag{30}
$$

Equating (29) and (30) yields a single equation for the determination of σ^0 and I. Though the solutions are not unique, by making the identification

$$b = -\eth_0^2\bar{\sigma}^0 \tag{31}$$

and

$$
\begin{aligned}
&\tfrac{8}{3}a\frac{\dot{P}_0}{P_0} - \tfrac{2}{3}\left(\eth_0 a\bar{\eth}_0\frac{\dot{P}_0}{P_0} + \bar{\eth}_0 a\eth_0\frac{\dot{P}}{P}\right)\\
&\qquad \equiv -\tfrac{2}{3}\eth_0\bar{\eth}_0\left(a\frac{\dot{P}_0}{P_0}\right) = \eth_0\eth_0^2 I,
\end{aligned} \tag{32}
$$

the solutions

$$
\begin{aligned}
\sigma^0 &= -\tfrac{1}{6}\eth_0 G\eth_0\frac{\dot{P}_0}{P_0},\\
I &= -\tfrac{1}{3}\tfrac{1}{6}\eth_0\bar{\eth}_0\left(a\frac{\dot{P}_0}{P_0}\right)
\end{aligned} \tag{33}
$$

are obtained. The alternate solutions can be constructed by means of infinitesimal coordinate transformations that are generalizations of supertranslations. By means of this freedom it is possible to make $I = 0$, but then the σ^0 becomes a nonlocal function of G, i.e., it depends on a time integral over G. It appears as if (33) is the only local solution.

It is possibly of interest to note that, in the case of $R = 0$, the "news" which has been broken up into two parts, σ^0 and I, is such that σ^0 contains the "magnetic" part of the news and I the "electric" part. Furthermore, it appears likely that a new congruence of null vectors can be introduced which are no longer hypersurface orthogonal but are now shear free, i.e., the "news" which was in σ^0 would then be placed in the curl or twist of the new congruence. This suggests that in the full theory one should look at the algebraically special solutions which are curling in order to study equations of motion of spinning bodies. In fact, the above was suggested by a study of the Kerr metric. In addition it appears possible that the condition $R = 0$ might be derived rather than assumed.

ACKNOWLEDGMENTS

One of the authors (E. T. N.) would like to thank Kings College, University of London, and Professor Felix Pirani for their hospitality, during the period when much of this work was done. We would also like to thank Professor Pirani for the many stimulating discussions.

* This research has been supported in part by the Aerospace Research Laboratory, OAR, U.S. Air Force, Contract F33615-70-C-1081 and through the European Office of Aerospace Research, Contract F61052-69-C-0012.

[1] E. T. Newman and R. Posadas, Phys. Rev. Letters **22**, 1196 (1969).
[2] E. T. Newman and R. Posadas, Phys. Rev. **187**, 1784 (1969).
[3] E. T. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).
[4] E. T. Newman and T. Unti, J. Math. Phys. **3**, 891 (1962).
[5] A. Janis and E. T. Newman, J. Math. Phys. **6**, 902 (1965).
[6] A. Papapetrou, Proc. Roy. Soc. (London) **A209**, 248 (1951).
[7] F. Pirani, Acta Phys. Polon. **15**, 389 (1956).
[8] E. T. Newman and R. Penrose, J. Math. Phys. **7**, 863 (1966).
[9] J. N. Goldberg, A. J. Macfarlane, E. T. Newman, and F. Rohrlich, J. Math. Phys. **8**, 2155 (1967).
[10] A. Held, E. T. Newman, and R. Posadas, J. Math. Phys. **11**, 3145 (1970).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Model of Phase Transitions for Interacting Fermi Systems*

H. L. Scott†‡
Department of Physics, Purdue University, Lafayette, Indiana 47907

(Received 19 March 1970)

The behavior of a previously discussed model of phase transitions for Fermi systems is analyzed in the region very near the critical point. It is shown that for a very general class of 2-body interactions the chemical potential is analytic in the temperature and density at the critical point, so that the model is in this sense equivalent to the classical theory of phase transitions. An extension of the model to include certain 3-body and higher interactions leaves this conclusion unchanged.

I. INTRODUCTION

In the past decade the problem of phase transitions has become widely recognized as one of the outstanding unsolved problems of statistical mechanics.[1] The question as to whether statistical mechanics is even applicable near the critical point was, of course, answered in the affirmative by Onsager[2] in 1944 with the solution of the 2-dimensional Ising model. Since that time several rigorous theorems on the behavior of the partition function near a phase transition have been derived,[3–5] but a large portion of the theoretical effort has been directed toward the analysis of various models that undergo phase transitions. With a few notable exceptions,[6] the analyses of these models have of necessity involved approximations of some sort, and, as a result, descriptions of the phase transition are generally obtained which are in agreement with the classical theory of Van der Waals.[7] As the Van der Waals theory is not in quantitative agreement with current experimental results[8] near the critical point, the basic problem of the phase transition remains unsolved and the search continues for models which on the one hand are simple enough so that the thermodynamic properties can be found via statistical mechanics and on the other hand are sufficiently complex so that nonclassical critical point behavior can be obtained.

In this work we consider a model of a phase transition, proposed in a series of papers by Gartenhaus and Stranahan,[9–11] in which the thermodynamic behavior of a system of fermions is studied near the liquid vapor condensation region. The model is defined by the Hamiltonian

$$H = \sum_{\mathbf{k}} t(\mathbf{k}) N_{\mathbf{k}} + \frac{1}{2\Omega} \sum_{\mathbf{k},\mathbf{q}} W(\mathbf{k}, \mathbf{q}) N_{\mathbf{k}} N_{\mathbf{q}}, \qquad (1)$$

where $N_{\mathbf{k}}$ is the number operator for a fermion of spin and wave vector $\mathbf{k}$, $t(\mathbf{k})$ the single-particle kinetic energy, $W(\mathbf{k}, \mathbf{q})$ the two-particle interaction, and Ω the quantization volume. In the thermodynamic limit ($\Omega \to \infty$, particle density n fixed) the partition function for the model can be evaluated exactly[12] and the resulting thermodynamic functions are summarized by the formula

$$n(\beta, \mu) = \int \frac{d\mathbf{k}}{(2\pi)^3} \rho(\mathbf{k}, \beta, \mu), \qquad (2)$$

where n is the particle density, μ the chemical potential, β the reciprocal of Boltzmann's constant times the temperature, and where we have changed the sum to an integral by the usual prescription. The quantity $\rho(\mathbf{k}, \beta, \mu)$ is the single-particle density matrix which obeys the nonlinear integral equation

$$\rho(\mathbf{k}, \beta, \mu) = \left[1 + \exp \beta\left(t(\mathbf{k}) + \int \frac{d\mathbf{q}}{(2\pi)^3} W(\mathbf{k}, \mathbf{q})\rho(\mathbf{q}) - \mu\right)\right]^{-1}. \qquad (3)$$

By considering the properties of this equation, a set of necessary and sufficient conditions on the interaction $W(\mathbf{k}, \mathbf{q})$ have been obtained which must be satisfied in order that a phase transition occur in the model.[9] When these conditions are fulfilled, it is shown that the transition in the model exhibits all the qualitative behavior observed in condensation phenomena. However, turning to a quantitative study of the model near the critical point, Gartenhaus and Stranahan found that,[11] for the attractive, factorable interaction $W(\mathbf{k}, \mathbf{q}) = -h(\mathbf{k})h(\mathbf{q})$, the critical exponents all assumed their classical values.[13] Therefore, for this special case the model is not in quantitative agreement with experiment near the critical point.

In this paper we investigate the critical-point behavior of the model for a class of more complicated interactions to determine whether it is possible in some cases to obtain nonclassical values for the critical exponents. We show that, for a very general class of interactions satisfying only certain integrability criterion, the critical exponents always retain their classical values regardless of the detailed properties of the individual interactions. We also find by extending

the model to include simple 3-body and higher interactions that these conclusions are unaltered.

Our analysis is organized in the following manner. In Sec. II we consider the behavior of the model near the critical point and show that any singular behavior at that point must be due to the vanishing of a certain function $D(\beta, \mu)$. In Sec. III the properties of the function $D(\beta, \mu)$ near the critical point are determined and the equation of state in the critical region is derived. In Sec. IV we define an extension of the model to include 3-body and higher interactions, and study the critical-point behavior of this model. Section V contains our conclusions.

II. PROPERTIES OF THE MODEL NEAR THE CRITICAL POINT

In this section we enumerate the model properties that will be needed to carry out our study of the critical region. For this purpose, we calculate the derivatives of the density with respect to temperature and chemical potential and investigate their behavior near the critical point. To facilitate this investigation, we define a function $\psi(\mathbf{k}, \beta, \mu)$ by

$$\psi(\mathbf{k}, \beta, \mu) = \int \frac{d\mathbf{q}}{(2\pi)^3} W(\mathbf{k}, \mathbf{q})\rho(\mathbf{q}, \beta, \mu), \tag{4}$$

where $\rho(\mathbf{k}, \beta, \mu)$ satisfies Eq. (3). Upon substitution of $\psi(\mathbf{k}, \beta, \mu)$ into the argument of the exponential in Eq. (3), the single-particle density matrix $\rho(\mathbf{k}, \beta, \mu)$ now becomes a functional of $\psi(\mathbf{k}, \beta, \mu)$ given by

$$\rho(\mathbf{k}, \psi, \mu) = \{1 + \exp \beta[t(\mathbf{k}) + \psi(\mathbf{k}) - \mu]\}^{-1}. \tag{5}$$

The advantage of this formulation is that the interaction $W(\mathbf{k}, \mathbf{q})$ appears as the kernel of the nonlinear integral equation (4).

In this work we consider the behavior of the model for a class of interaction $W(\mathbf{k}, \mathbf{q})$ which has the following properties: First, all interactions in the class must have the necessary behavior so that a phase transition does occur in the model[9]; secondly, all interactions must be expressible, either exactly or as an approximation in the mean,[14] by a symmetric interaction of the form

$$W(\mathbf{k}, \mathbf{q}) = \sum_{i=1}^{\mathcal{N}} \alpha_i(\mathbf{k})\beta_i(\mathbf{q}), \tag{6}$$

where $\mathcal{N}$ is a positive integer and each $\alpha_i(\mathbf{k})$ and $\beta_i(\mathbf{q})$ is a bounded function defined over a finite interval in $\mathbf{k}$ space. In the following calculations, we shall only consider interactions in this class, so that the interaction form above will always be used.

In order to calculate the derivatives of $n(\beta, \mu)$, we differentiate the density formula (2), utilizing Eq. (5) for the single-particle density matrix,and we find

$$\left(\frac{\partial n}{\partial \mu}\right)_\beta = \beta \int \frac{d\mathbf{k}}{(2\pi)^3} F(\mathbf{k}) \left[\left(\frac{\partial \psi(\mathbf{k})}{\partial \mu} \right)_\beta - 1 \right], \tag{7}$$

$$\left(\frac{\partial n}{\partial \beta}\right)_\mu = \int \frac{d\mathbf{k}}{(2\pi)^3} F(\mathbf{k}) \left[\beta \left(\frac{\partial \psi(\mathbf{k})}{\partial \beta} \right)_\mu + I(\mathbf{k}) \right], \tag{8}$$

where the functions $F(\mathbf{k})$ and $I(\mathbf{k})$ are defined by

$$F(\mathbf{k}) = -\rho(\mathbf{k})[1 - \rho(\mathbf{k})], \tag{9}$$

$$I(\mathbf{k}) = \beta^{-1} \ln \left(\frac{1 - \rho(\mathbf{k})}{\rho(\mathbf{k})} \right), \tag{10}$$

and where $\psi(\mathbf{k})$ is given by Eq. (4). Differentiating Eq. (4) successively with respect to β and μ and making use of Eq. (5), we find the derivatives of $\psi(\mathbf{k})$ satisfy the linear integral equations

$$\left(\frac{\partial \psi(\mathbf{k})}{\partial \mu} \right)_\beta = \beta \int \frac{d\mathbf{q}}{(2\pi)^3} W(\mathbf{k}, \mathbf{q}) F(\mathbf{q}) \left[\left(\frac{\partial \psi(\mathbf{q})}{\partial \mu} \right) - 1 \right], \tag{11}$$

$$\left(\frac{\partial \psi(\mathbf{k})}{\partial \beta} \right)_\mu = \int \frac{d\mathbf{q}}{(2\pi)^3} W(\mathbf{k}, \mathbf{q}) F(\mathbf{q}) \left[\beta \left(\frac{\partial \psi(\mathbf{q})}{\partial \beta} \right) + I(\mathbf{q}) \right]. \tag{12}$$

Upon substitution of the interaction defined by Eq. (6) into the above equations, both are reduced to systems of inhomogeneous algebraic equations which can be solved by Cramer's rule. The solutions of Eqs. (11) and (12) are therefore given by

$$\left(\frac{\partial \psi(\mathbf{k})}{\partial \mu} \right)_\beta = -\beta \sum_{i,j=1}^{\mathcal{N}} \alpha_i(\mathbf{k})(F\beta_j) \frac{M(i,j)}{D(\beta, \mu)}, \tag{13}$$

$$\left(\frac{\partial \psi(\mathbf{k})}{\partial \beta} \right)_\mu = \sum_{i,j=1}^{\mathcal{N}} \alpha_i(\mathbf{k})(FIB_j) \frac{M(i,j)}{D(\beta, \mu)}, \tag{14}$$

where we have used the notation

$$(FG) \equiv \int \frac{d\mathbf{k}}{(2\pi)^3} F(\mathbf{k})G(\mathbf{k}) \tag{15}$$

and $D(\beta, \mu)$ is an $\mathcal{N} \times \mathcal{N}$ determinant, the elements of which are functions of β and μ defined by

$$D_{ij} = \delta_{ij} - \beta(F\alpha_i\beta_j). \tag{16}$$

The quantity $M(i,j)$ is the minor of $D(\beta, \mu)$ formed by deleting the ith row and jth column. Now, inserting Eqs. (13) and (14) into Eqs. (7) and (8), respectively, we find

$$\left(\frac{\partial n}{\partial \mu} \right)_\beta = -\beta^2 \sum_{i,j=1}^{\mathcal{N}} (F\alpha_i)(F\beta_j) \frac{M(i,j)}{D(\beta, \mu)} - \beta(F), \tag{17}$$

$$\left(\frac{\partial n}{\partial \beta} \right)_\mu = \beta \sum_{i,j=1}^{\mathcal{N}} (F\alpha_i)(FI\beta_j) \frac{M(i,j)}{D(\beta, \mu)} + (FI). \tag{18}$$

We shall see that the function $D(\beta, \mu)$, which we subsequently refer to as the denominator function, contains all of the information needed to describe the behavior of the model very near the critical point.

We now consider the quantities $(\partial n/\partial \mu)_\beta$ and $(\partial n/\partial \beta)_\mu$ near the critical point. According to the criterion

$$0 = \left(\frac{\partial P}{\partial n}\right)_\beta = \left(\frac{\partial P}{\partial \mu}\right) \Big/ \left(\frac{\partial n}{\partial \mu}\right)_\beta = n \Big/ \left(\frac{\partial n}{\partial \mu}\right)_\beta \tag{19}$$

at the critical point, $(\partial n/\partial \mu)_\beta$ must diverge at that point. Upon examination of Eq. (17) we see that, in order to satisfy Eq. (19), the denominator function must vanish at the critical point, i.e.,

$$D(\beta_c, \mu_c) = 0, \tag{20}$$

where β_c and μ_c are the values of β and μ at the critical point. It follows then that the density is a nonanalytic function of the temperature and chemical potential at the critical point. We shall show in the next section, however, that, by choosing density and temperature as our independent thermodynamic variables, we obtain the equation of state $\mu = \mu(\beta, n)$ which is analytic in a neighborhood of the critical point. It is this equation of state which allows us to conclude that the critical exponents for the model all retain their classical values.

III. THE EQUATION OF STATE IN THE CRITICAL REGION

In this section we study the behavior of the denominator function $D(\beta, \mu)$ in the variables β and n near the critical point. By showing that all derivatives of $D(\beta, \mu(\beta, n)) \equiv D(\beta, n)$ with respect to β and n exist near the critical point, we prove that, for the class of interactions defined in Sec. II, $D(\beta, n)$ is analytic[15] in β and n at the critical point. Then, given this result, it is straightforward to derive the equation of state in the form $\mu = \mu(\beta, n)$ near the critical point. We will see that the equation of state, like the denominator function $D(\beta, n)$, is analytic at the critical point. As this analyticity is the basis for the classical theory of phase transitions,[13,16] we are able to conclude that, for the given class of interactions, the critical exponents of the model agree with those of the classical theory.

In order to determine the analytic behavior of the $\mathcal{N} \times \mathcal{N}$ determinant $D(\beta, \mu)$ [Eq. (16)], it is sufficient to consider a general element of D given by

$$D_{ij} = \delta_{ij} - \beta(F\alpha_i\beta_j).$$

We consider first the behavior of D_{ij} as a function of the density, holding the temperature constant. As D_{ij} is an implicit function of n through the function $\mu(\beta, n)$, the derivative of D_{ij} with respect to the density is given by

$$\left(\frac{\partial D_{ij}}{\partial n}\right)_\beta = \left(\frac{\partial D_{ij}}{\partial \mu}\right)_\beta \left(\frac{\partial \mu}{\partial n}\right)_\beta. \tag{21}$$

The quantity $(\partial \mu/\partial n)_\beta$ is given by the reciprocal of Eq. (17) and the derivative of D_{ij} with respect to μ is

$$\begin{aligned}\left(\frac{\partial D_{ij}}{\partial \mu}\right)_\beta &= -\beta \frac{\partial}{\partial \mu}(F\alpha_i\beta_j) \\ &= -\beta^2\left\{\alpha_i\beta_j F(1-2\rho)\left[\left(\frac{\partial \psi}{\partial \mu}\right)_\beta - 1\right]\right\},\end{aligned} \tag{22}$$

where $F(\mathbf{k})$, $\rho(\mathbf{k})$, α_i, β_j, and $\psi(\mathbf{k})$ are all defined in the previous section and where the integration notation defined in Eq. (15) is used. It will be recalled from Sec. II that the derivative of $\psi(\mathbf{k})$ with respect to μ at constant β is given by Eq. (13). Substituting this equation into Eq. (22) and keeping only terms proportional to $D^{-1}(\beta, \mu)$, we find

$$\left(\frac{\partial D_{ij}}{\partial \mu}\right)_\beta = \beta^3 \sum_{l,m=1}^{\mathcal{N}} [\alpha_i\beta_j F(1-2\rho)\alpha_l](F\beta_m)\frac{M(l, m)}{D(\beta, \mu)}. \tag{23}$$

Now, inserting the reciprocal of Eqs. (17) and (23) into Eq. (21), we see that the $D^{-1}(\beta, \mu)$ term cancels, and we are left with

$$\left(\frac{\partial D_{ij}}{\partial n}\right)_\beta = \frac{A(\beta, \mu)}{N(\beta, \mu)}, \tag{24}$$

where $A(\beta, \mu)$ is the numerator in Eq. (23) and $N(\beta, \mu)$ is the expression which multiplies $D^{-1}(\beta, \mu)$ in Eq. (17) and where terms proportional to $D(\beta, n)$ have been ignored in numerator and denominator. We shall hereafter refer to $N(\beta, \mu)$ as the numerator function. If $N(\beta, \mu)$ is nonvanishing at the critical point,[17] then the first derivative of D_{ij} with respect to n at constant β, according to Eq. (24), is finite. Applying the same argument to all higher derivatives of D_{ij} with respect to n, we conclude that D_{ij} is analytic in n near the critical point for all interactions in the given class.[15]

Turning now to the derivatives of D_{ij}, with respect to β at constant n, we make use of the theory of Jacobians[18] to write

$$\left(\frac{\partial D_{ij}}{\partial \beta}\right)_n = \left[\left(\frac{\partial D_{ij}}{\partial \beta}\right)_\mu \left(\frac{\partial n}{\partial \mu}\right)_\beta - \left(\frac{\partial D_{ij}}{\partial \mu}\right)_\beta \left(\frac{\partial n}{\partial \beta}\right)_\mu\right] \Big/ \left(\frac{\partial n}{\partial \mu}\right)_\beta. \tag{25}$$

The quantities $(\partial n/\partial\mu)_\beta$, $(\partial n/\partial\beta)_\mu$, and $(\partial D_{ij}/\partial\mu)_\beta$ are given by Eqs. (17), (18), and (23), respectively. In analogy with Eq. (22), $(\partial D_{ij}/\partial\beta)_\mu$ is given by

$$\left(\frac{\partial D_{ij}}{\partial\beta}\right)_\mu = -(F\alpha_i\beta_j) - \beta\frac{\partial}{\partial\beta}(F\alpha_i\beta_j)$$
$$= -(F\alpha_i\beta_j) - \beta^2\left\{\alpha_i\beta_j F(1-2\rho)\left[\beta\left(\frac{\partial\psi}{\partial\beta}\right)_\mu + I(\mathbf{k})\right]\right\}. \tag{26}$$

Substituting for $(\partial\psi/\partial\beta)_\mu$ and keeping only terms proportional to $D^{-1}(\beta,\mu)$, we find

$$\left(\frac{\partial D_{ij}}{\partial\beta}\right)_\mu = -\beta^3\sum_{l,m}^{\mathcal{N}}[\alpha_i\beta_j F(1-2\rho)\alpha_l](FIB_m)\frac{M(l,m)}{D(\beta,\mu)}. \tag{27}$$

Making use of the Eqs. (17), (18), (23), and (27), we find after some manipulation

$$\left(\frac{\partial D_{ij}}{\partial\beta}\right)_n = \left(\sum_{k,q,p,r}^{\mathcal{N}} H^{ij}_{k,q,p,r}(\beta,\mu)[M(k,q)M(p,r) - M(k,r)M(p,q)]\right)\Big/ D(\beta,\mu), \tag{28}$$

where

$$H^{ij}_{k,q,p,r}(\beta,\mu) = \beta^2(F\alpha_k)(FI\beta_r)[(1-2\rho)\alpha_i\beta_j\alpha_p](F\beta_q). \tag{29}$$

We now employ the identity[19]

$$M(i,j)M(k,l) - M(i,l)M(k,j) = M(i,j;k,l)D(\beta,\mu), \tag{30}$$

where $M(i,j;k,l)$ is the minor of $D(\beta,\mu)$, formed by deleting rows i and j and columns k and l. Upon substitution of Eq. (30) into Eq. (28), the $D^{-1}(\beta,\mu)$ singularity is cancelled and $(\partial D/\partial\beta)_n$ becomes

$$\left(\frac{\partial D_{ij}}{\partial\beta}\right)_n = \left(\sum_{k,q,p,r}^{\mathcal{N}} H^{ij}_{k,q,p,r}(\beta,\mu)M(k,q;p,r)\right)\Big/ N(\beta,\mu). \tag{31}$$

Assuming again that the numerator function $N(\beta,\mu)$ is nonvanishing in a neighborhood of the critical point,[17] it follows that $(\partial D_{ij}/\partial\beta)_n$ exists in this neighborhood. As was the case for the density derivatives, we can apply the same arguments to higher derivatives of D_{ij} with respect to β at constant n and we find that the cancellation always takes place. On the basis of this result and the earlier result for the density behavior, we conclude then that, for all interactions in the given class defined by Eq. (6), the associated denominator function $D(\beta,n)$ is analytic[15] in a neighborhood of the critical point.

The analyticity of $D(\beta,n)$ allows us to expand it in a Taylor series about the critical point and, very near that point, to keep only the lowest-order terms. Thus we may write

$$D(\beta,n) \cong a(n-n_c)^k + b(\beta-\beta_c) + c(n-n_c)(\beta-\beta_c) + \cdots, \tag{32}$$

where a, b, and c are constants and k is the order of the lowest nonvanishing derivative of $D(\beta,n)$ with respect to n at constant β.[20] Sufficiently close to the critical point we can replace the numerator function $N(\beta,\mu)$ by a nonzero constant N_c, and the reciprocal of Eq. (17) becomes

$$\left(\frac{\partial\mu}{\partial n}\right)_\beta \cong N_c^{-1}[a(n-n_c)^k + b(\beta-\beta_c) + c(n-n_c)(\beta-\beta_c) + \cdots]. \tag{33}$$

Integrating this expression in the critical region, we obtain

$$\mu - \mu_c \cong d(n-n_c)^{k+1} + e(\beta-\beta_c)(n-n_c) + f(\beta), \tag{34}$$

where $f(\beta)$ is a function of β only vanishing at the critical point, and d and e are constants. Examination of Eq. (34) shows that this equation of state is equivalent to the classical equation of state.[13] We are thus able to conclude immediately that the critical exponents of the model, for all interactions in the given class, are the same as those of the classical theory.

IV. EXTENSION TO MANY-BODY INTERACTIONS

In view of the results of the previous section, we now consider an extension of the model to include 3-body and higher interactions in the hope that such a model may exhibit nonclassical behavior. If such interactions are included the model Hamiltonian becomes

$$H = \sum_{\mathbf{k}} t(\mathbf{k})N_{\mathbf{k}} + \frac{1}{2\Omega}\sum_{\mathbf{k},\mathbf{q}} W_2(\mathbf{k},\mathbf{q})N_{\mathbf{k}}N_{\mathbf{q}} + \frac{1}{3!\,\Omega^2}\sum_{\mathbf{k},\mathbf{q},\mathbf{p}} W_3(\mathbf{k},\mathbf{q},\mathbf{p})N_{\mathbf{k}}N_{\mathbf{q}}N_{\mathbf{p}} + \cdots. \tag{35}$$

If we consider two through m-body interactions all of which are independent of the momentum variables, then the interaction part of H can be expressed as a polynomial of degree m in the operator $\sum_{\mathbf{k}} N_{\mathbf{k}}/\Omega$. The methods of Girardeau again allow us to calculate

the thermodynamic functions exactly in the lim $\Omega \to \infty$, and the nonlinear integral equation for the single-particle density matrix of the new model is

$$\rho(\mathbf{k}, \beta, \mu) = \{1 + \exp \beta[t(\mathbf{k}) + P(n) - \mu]\}^{-1}, \quad (36)$$

where $P(n)$ is a finite-degree polynomial in the density n.

In order to study the critical-point behavior of this model, we proceed, in the same manner as the previous sections, to determine the analytic behavior of the chemical potential $\mu(\beta, n)$ near the critical point. Substituting the expression for $\rho(\mathbf{k})$, Eq. (36), into the density formula (2) yields an implicit equation for $n(\beta, \mu)$. Differentiating this equation with respect to μ, holding β constant, we find

$$\left(\frac{\partial \mu}{\partial n}\right)_\beta = \frac{-1}{\beta(F)} + P'(n), \quad (37)$$

where

$$(F) = \int \frac{d\mathbf{k}}{(2\pi)^3} F(\mathbf{k})$$

and $F(\mathbf{k})$ is given by Eq. (9), and where $P'(n)$ is the derivative of $P(n)$ with respect to n. Since the single-particle density matrix $\rho(\mathbf{k}, \beta, \mu)$ is positive definite for finite temperature, it follows from Eq. (9) that (F) is nonzero. Thus we conclude that, because of the critical point criterion $(\partial\mu/\partial n)_\beta = 0$, Eq. (37) vanishes at the critical point and is defined in some neighborhood of that point. Furthermore, the calculation of higher derivatives shows that all derivatives of μ with respect to n at constant β exist in this neighborhood, and so the chemical potential is analytic in n. To see that μ is analytic in β, we differentiate the implicit formula for the density with respect to μ, holding n fixed, and we find

$$\left(\frac{\partial \mu}{\partial \beta}\right)_n = \frac{(FI)}{\beta(F)}, \quad (38)$$

where $I(\mathbf{k})$ is defined by Eq. (10). Again since (F) is nonzero and (I) is finite, we see that $(\partial\mu/\partial\beta)_n$ and all higher derivatives of μ with respect to β exist near the critical point. We therefore conclude that $\mu(\beta, n)$ is analytic in n and β at the critical point.[16] It now follows that the equation of state can be written in a manner similar to Eq. (34), and thus the critical exponents for our modified model are again the same as those of the classical theory.[13]

V. CONCLUSIONS

From our analysis of the model of Gartenhaus and Stranahan in the preceding sections, a distinctive feature of this model has emerged: namely, the fact that for a very general class of interactions the critical exponents associated with the model always retain their classical values. Before discussing this characteristic we note that perhaps the most important over-all property of the model is the fact, as shown by Gartenhaus and Stranahan, that all the qualitative features of the liquid vapor phase transition are exhibited. The conclusions obtained in this paper apply only to the region very near the critical point and do not effect this general result. We also remark that the interactions considered for the model generally do not have a term corresponding to a hard-core repulsion between the particles. However, for the present case of a fermion system, the catastrophic collapse which might otherwise occur is prevented by the Pauli principle.[11]

In Secs. II, III, and IV we analyze the model for a large class of 2-body interactions and for certain simple 3-body and higher interactions. Our main conclusion was that, for all interactions considered, regardless of their detailed properties, the critical exponents for the model always assume the classical values. From our analysis, we see that the reason the model exhibits classical behavior for such a wide range of interactions is the fact that, for each of these interactions, the equation of state $\mu = \mu(\beta, n)$ can be expanded in a Taylor series about the critical point. This analytic behavior in the region near the critical point was in turn derived in general for any interaction which could be expressed approximately as a finite sum of factorable interactions of the form given by Eq. (21). Thus, we conclude that this analytic behavior at the critical point is a basic property of the model, irrespective of the interaction used in the Hamiltonian (subject only to certain general integrability criterion outlined in Sec. II). In retrospect, this result is not entirely unexpected since all interactions considered have, by assumption, infinite range, and it has been shown that a large class of systems with local, infinite range interactions also have classical values for the exponents.[21,22] In this sense our conclusions confirm an earlier conjecture[11] that the theorems of Lieb[22] and Lebowitz and Penrose[21] can be generalized to nonlocal interactions as well. It appears to be a general fact, therefore, that, when the limit of an infinite interaction range is taken in order to evaluate the thermodynamic properties of a given model, while a phase transition can be produced in such a limit, any nonanalytic behavior of the thermodynamic functions near the critical point is lost.

Finally it has been shown[23] that, if we assume Girardeau's methods can be applied to infinite power series in the operator $\sum_{\mathbf{k}} N_{\mathbf{k}}/\Omega$, then by judicious choice of the coefficients in this series we can obtain

a phase transition in the model with nonclassical values for the exponents β, δ, and α'.

ACKNOWLEDGMENT

The author wishes to thank Dr. S. Gartenhaus for suggesting this investigation and his guidance, given in stimulating discussions during the course of this work.

APPENDIX: THE CASE $N(\beta_c, \mu_c) = 0$

In this appendix we investigate the consequences of the assumption that the numerator function $N(\beta, \mu)$ vanishes at the critical point. We recall from Sec. II that $N(\beta, \mu)$ is defined by

$$N(\beta, \mu) = \sum_{i,j=1}^{\mathcal{N}} (F\alpha_i)(F\beta_j)M(i, j), \tag{A1}$$

where the quantities on the right side of the equation are defined in this section. We shall show that, if Eq. (A1) vanishes at $\beta = \beta_c$, $\mu = \mu_c$, then the behavior of the model contradicts the well-known result[9] that the density fluctuations, given by

$$\beta\langle(n - \langle n\rangle)^2\rangle = \left(\frac{\partial n}{\partial \mu}\right)_\beta, \tag{A2}$$

must diverge at the critical point.

To see this, we examine the numerator function as a function of the variables β and n, $N(\beta, \mu(\beta, n)) \equiv N(\beta, n)$. It is sufficient for our analysis to consider this function in the one phase region $T > T_c$, along the critical isochore, $n = n_c$. We recall from Sec. II, Eq. (31), that, near the critical point,

$$\left(\frac{\partial D}{\partial \beta}\right)_n \cong \frac{\text{const}}{N(\beta, n)}, \tag{A3}$$

where $D(\beta, n)$ is the denominator function, also defined in Sec. II, and expressed here as a function of β and n. Now, since $N(\beta, n)$ as well as $D(\beta, n)$ consists of sums of terms of the form

$$(F\alpha_i)(F\beta_j)(F\alpha_k\beta_l)\cdots, \tag{A4}$$

the reasoning leading to Eq. (A3) can also be applied to $N(B, n)$, so that

$$\left(\frac{\partial N}{\partial \beta}\right)_n \cong \frac{\text{const}}{N(\beta, n)}. \tag{A5}$$

Setting $n = n_c$ and making use of the fact that $N(\beta_c n_c) = 0$ by hypothesis, we solve Eq. (A5) and obtain

$$N(\beta, n_c) \cong a(\beta - \beta_c)^{\frac{1}{2}}, \tag{A6}$$

where a is a constant. Substitution of Eq. (A6) into the equation for $(\partial D/\partial\beta)_n$, Eq. (A3), we find that $D(\beta, n_c)$ also can be expressed in the form

$$D(\beta, n_c) = b(\beta - \beta_c)^{\frac{1}{2}}, \tag{A7}$$

where b is a constant.

Now according to Eq. (32) of Sec. II, the density fluctuations are given by

$$\left(\frac{\partial n}{\partial \mu}\right)_\beta = c + \frac{N(\beta, n)}{D(\beta, n)}, \tag{A8}$$

where c is a constant. Substituting Eq. (A6) and (A7) into Eq. (A8), we see that $(\partial n/\partial\mu)_\beta$ remains finite as the critical point is approached along the critical isochore $n = n_c$. Since the divergence in the density fluctuations is a primary characteristic of the liquid–vapor transition, we conclude that, for all models considered in Sec. II which undergo a phase transition, the numerator function $N(\beta, n)$ does not vanish at the critical point.

* Supported in part by NASA Grant No. NG 15-005-021.

† Submitted in partial fulfillment of the requirements for the Ph.D. degree, Purdue University.

‡ Present address: Physics Department, University of Utah, Salt Lake City, Utah 84112.

[1] For a comprehensive review, see M. E. Fisher, Rept. Progr. Phys. **30**, 615 (1967).

[2] L. Onsager, Phys. Rev. **65**, 117 (1944).

[3] L. Van Hove, Physica **15**, 951 (1949).

[4] T. D. Lee and C. N. Yang, Phys. Rev. **87**, 404 (1952).

[5] See, for example, D. Ruelle, *Statistical Mechanics* (Benjamin, New York, 1969).

[6] Including the Onsager solution and the exactly soluble ferroelectric models of E. H. Lieb, Phys. Rev. **18**, 1046 (1967); **19**, 108 (1967).

[7] J. D. Van der Waals, thesis, University of Leiden, 1873. See also L. D. Landau and E. Lifshitz, *Statistical Physics* (Pergamon, London, 1958), pp. 259–68.

[8] For a review of experimental results, see P. Heller, Rept. Progr. Phys. **30**, 731 (1967).

[9] S. Gartenhaus and G. Stranahan, Phys. Rev. **138**, A1346 (1965).

[10] C. M. Andersen, S. Gartenhaus, and G. Stranahan, Phys. Rev. **146**, 101 (1966).

[11] S. Gartenhaus and G. Stranahan, Phys. Rev. **173**, 260 (1968).

[12] M. Girardeau, J. Math. Phys. **3**, 131 (1962).

[13] For the definition of these exponents, and a description of the classical theory, see, for example, M. E. Fisher, Ref. 1.

[14] An interaction $W(\mathbf{k}, \mathbf{q})$ can be approximated in the mean if, given $\epsilon > 0$, there exist a set of $2\mathcal{N}$ functions $\{\alpha_i(\mathbf{k}), \beta_i(\mathbf{k})\}_{i=1}^{\mathcal{N}}$ such that

$$\int \frac{d\mathbf{k}\, d\mathbf{q}}{(2\pi)^3} \left| W(\mathbf{k}, \mathbf{q}) - \sum_{i=1} \alpha_i(\mathbf{k})\beta_i(\mathbf{q}) \right| < \epsilon.$$

[15] In order to prove a function $f(x)$ analytic at x_0, one must show not only that all derivatives of $f(x)$ exist at x_0, but that

$$\lim_{n\to\infty} R_n(x) = 0,$$

where $R_n(x)$ is the remainder term of the nth-order Taylor expansion of x about x_0:

$$R_n(x) = \frac{1}{n!} \left.\frac{\partial^n f}{\partial x^n}\right|_{x=\xi} (x - x_0)^n$$

for $0 < \xi < |x - x_0|$. For the case under consideration we assume this condition holds in some neighborhood of the critical point.

[16] For a discussion of the analyticity requirements of the classical theory, see, for example, R. B. Griffiths, Phys. Rev. **158**, 176 (1967), Appendix B.

[17] We show in the Appendix that, if $N(\beta, \mu)$ does vanish at the critical point, then the properties of the model are inconsistent with those of known systems at the critical point. For this reason we shall always assume $N(\beta, \mu)$ is nonvanishing in a neighborhood of the critical point for interactions in the given class.

[18] L. D. Landau and E. Lifshitz, Ref. 7, p. 50.

[19] T. Muir, *Theory of Determinants* (Dover, New York, 1960), p. 134.

[20] Thermodynamic stability requires that the order of the first nonvanishing derivative of μ with respect to n be odd.

[21] J. Lebowitz and O. Penrose, J. Math. Phys. 7, 98 (1966).

[22] E. H. Lieb, J. Math. Phys. 7, 1016 (1966).

[23] S. Gartenhaus and H. L. Scott, Phys. Rev. A1, 1270 (1970).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Method of Constructing General Solutions of Certain Nonlinear Differential Equations Describing Plasma Oscillations*

E. B. BARRETT
Naval Postgraduate School, Monterey, California 93940

(Received 25 March 1970)

A general solution is derived for a system of nonlinear partial differential equations describing longitudinal plasma oscillations. The method of construction of the solution involves the imposition of arbitrary functional relationships among the integrals of an associated system of ordinary differential equations.

1. INTRODUCTION

The purpose of this paper is to discuss a mathematical technique for deriving general solutions to a subsystem of the nonlinear differential equations describing the interactions of electromagnetic fields with ionized media.

The nonlinear interaction of an electromagnetic wave with a plasma layer has been investigated theoretically and experimentally by Whitmer, Tetenbaum, and Barrett.[1-4] The mathematical model on which these studies are based consists of Maxwell's equations, together with the continuity equation and the "Navier–Stokes" equation for the electrons. The latter equations may be derived by taking the first and second moments of the Boltzmann equation. In this model, n represents the electron density; n_{i0} is the steady-state positive ion concentration; V is the electron velocity vector; e and m are the charge and mass of the electron; collisions are between electrons and neutral particles, are assumed to be elastic, and are described by a collision frequency v.

In this model, the nonlinear terms nV, VXB, and $(V \cdot \nabla)V$ appear. In Refs. 1–4, the authors discussed nonlinear effects produced by these terms, such as harmonic generation and frequency mixing. The approach taken was to expand all terms into Fourier series in time, match frequency components, and solve a sequence of boundary-value problems for the linearized equations in a plasma slab. In general, the results of the experiments gave good agreement with the theoretical predictions made by this method, within the limits of the small-signal theory.

In the present paper, a method is described for deriving general solutions to a subset of the nonlinear equations of the model, by exploiting the fact that the equations in the subset have identical principal parts.[5] From these general solutions, infinitely many special solutions may be derived, expressed in terms of the plasma parameters v and ω_p. These solutions correspond to longitudinal oscillations of the plasma. In the present paper we shall not treat the case of a strong dc magnetic field, which may produce coupling between these oscillations and transverse components of the electromagnetic field.

Let $E_x, E_y, E_z, B_x, B_y, B_z$, and V_x, V_y, V_z represent the $\hat{x}, \hat{y}$, and $\hat{z}$ components of the E, B, and V vectors, respectively. We shall assume a simple geometry in which all spatial variation is in the x direction only, and we shall begin by investigating the solution in which E_y, E_z, B_y, B_z, V_y, and V_z are $= 0$. We retain the equations

$$\frac{\partial}{\partial x} E_x = \frac{e}{\epsilon_0} n - \frac{e}{\epsilon_0} n_{i0}, \tag{1}$$

$$\frac{\partial}{\partial t} E_x = -\frac{e}{\epsilon_0} n V_x, \tag{2}$$

$$\frac{\partial}{\partial t} V_x + V_x \frac{\partial}{\partial x} V_x = \frac{e}{m} E_x - v V_x, \tag{3}$$

[16] For a discussion of the analyticity requirements of the classical theory, see, for example, R. B. Griffiths, Phys. Rev. **158**, 176 (1967), Appendix B.

[17] We show in the Appendix that, if $N(\beta, \mu)$ does vanish at the critical point, then the properties of the model are inconsistent with those of known systems at the critical point. For this reason we shall always assume $N(\beta, \mu)$ is nonvanishing in a neighborhood of the critical point for interactions in the given class.

[18] L. D. Landau and E. Lifshitz, Ref. 7, p. 50.

[19] T. Muir, *Theory of Determinants* (Dover, New York, 1960), p. 134.

[20] Thermodynamic stability requires that the order of the first nonvanishing derivative of μ with respect to n be odd.

[21] J. Lebowitz and O. Penrose, J. Math. Phys. 7, 98 (1966).

[22] E. H. Lieb, J. Math. Phys. 7, 1016 (1966).

[23] S. Gartenhaus and H. L. Scott, Phys. Rev. A1, 1270 (1970).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Method of Constructing General Solutions of Certain Nonlinear Differential Equations Describing Plasma Oscillations*

E. B. BARRETT
Naval Postgraduate School, Monterey, California 93940

(Received 25 March 1970)

A general solution is derived for a system of nonlinear partial differential equations describing longitudinal plasma oscillations. The method of construction of the solution involves the imposition of arbitrary functional relationships among the integrals of an associated system of ordinary differential equations.

1. INTRODUCTION

The purpose of this paper is to discuss a mathematical technique for deriving general solutions to a subsystem of the nonlinear differential equations describing the interactions of electromagnetic fields with ionized media.

The nonlinear interaction of an electromagnetic wave with a plasma layer has been investigated theoretically and experimentally by Whitmer, Tetenbaum, and Barrett.[1-4] The mathematical model on which these studies are based consists of Maxwell's equations, together with the continuity equation and the "Navier–Stokes" equation for the electrons. The latter equations may be derived by taking the first and second moments of the Boltzmann equation. In this model, n represents the electron density; n_{i0} is the steady-state positive ion concentration; V is the electron velocity vector; e and m are the charge and mass of the electron; collisions are between electrons and neutral particles, are assumed to be elastic, and are described by a collision frequency v.

In this model, the nonlinear terms nV, VXB, and $(V \cdot \nabla)V$ appear. In Refs. 1–4, the authors discussed nonlinear effects produced by these terms, such as harmonic generation and frequency mixing. The approach taken was to expand all terms into Fourier series in time, match frequency components, and solve a sequence of boundary-value problems for the linearized equations in a plasma slab. In general, the results of the experiments gave good agreement with the theoretical predictions made by this method, within the limits of the small-signal theory.

In the present paper, a method is described for deriving general solutions to a subset of the nonlinear equations of the model, by exploiting the fact that the equations in the subset have identical principal parts.[5] From these general solutions, infinitely many special solutions may be derived, expressed in terms of the plasma parameters v and ω_p. These solutions correspond to longitudinal oscillations of the plasma. In the present paper we shall not treat the case of a strong dc magnetic field, which may produce coupling between these oscillations and transverse components of the electromagnetic field.

Let $E_x, E_y, E_z, B_x, B_y, B_z$, and V_x, V_y, V_z represent the $\hat{x}, \hat{y}$, and $\hat{z}$ components of the E, B, and V vectors, respectively. We shall assume a simple geometry in which all spatial variation is in the x direction only, and we shall begin by investigating the solution in which E_y, E_z, B_y, B_z, V_y, and V_z are $= 0$. We retain the equations

$$\frac{\partial}{\partial x} E_x = \frac{e}{\epsilon_0} n - \frac{e}{\epsilon_0} n_{i0}, \tag{1}$$

$$\frac{\partial}{\partial t} E_x = -\frac{e}{\epsilon_0} n V_x, \tag{2}$$

$$\frac{\partial}{\partial t} V_x + V_x \frac{\partial}{\partial x} V_x = \frac{e}{m} E_x - v V_x, \tag{3}$$

describing longitudinal plasma oscillations. Multiplying (1) by V_x and adding to (2), we derive

$$\frac{\partial}{\partial t}\left(\frac{e}{m}E_x\right) + V_x\frac{\partial}{\partial x}\left(\frac{e}{m}E_x\right) = -\omega_p^2 V_x, \qquad (4)$$

where $\omega_p^2 = n_{i_0}e^2/m\epsilon_0$. The "plasma frequency" ω_p is assumed to be constant. We shall derive a general solution to the nonlinear system (3), (4).

2. INTEGRALS OF THE NONLINEAR SYSTEM

Observing that the Eqs. (3) and (4) have the same principal parts, we consider the equivalent system of nonlinear ordinary differential equations

$$\frac{dt}{1} = \frac{dx}{V_x} = \frac{d[(e/m)E_x]}{-\omega_p^2 V_x} = \frac{dV_x}{(e/m)E_x - vV_x}. \qquad (5)$$

The integration of systems of first-order partial differential equations having the same principal parts may be achieved by deriving integrals of such systems of ordinary differential equations.[5] Three integrals will be derived from (5) be standard techniques (Ref. 6, Chap. 6), and a general solution will be constructed from these integrals.

It can be easily shown that a general solution to (5) may be written in the form

$$\begin{aligned} x(t) &= C_1e^{\lambda^+t} + C_2e^{\lambda^-t} + C_3,\\ V_x(t) &= \lambda^+C_1e^{\lambda^+t} + \lambda^-C_2e^{\lambda^-t},\\ (e/m)E_x(t) &= -\omega_p^2C_1e^{\lambda^+t} - \omega_p^2C_2e^{\lambda^-t}, \end{aligned} \qquad (6)$$

where

$$\begin{aligned} \lambda^+ &= -v/2 + i(\omega_p^2 - v^2/4)^{\frac{1}{2}},\\ \lambda^- &= -v/2 - i(\omega_p^2 - v^2/4)^{\frac{1}{2}}, \end{aligned}$$

and C_1, C_2, and C_3 are arbitrary constants.

Equivalently, we have the relations

$$\begin{aligned} \frac{e^{-\lambda^+t}}{\lambda^+ - \lambda^-}\left[V_x(t) + \frac{\lambda^-}{\omega_p^2}\left(\frac{e}{m}E_x(t)\right)\right] &= C_1,\\ \frac{e^{-\lambda^-t}}{\lambda^- - \lambda^+}\left[V_x(t) + \frac{\lambda^+}{\omega_p^2}\left(\frac{e}{m}E_x(t)\right)\right] &= C_2,\\ \left[\left(\frac{e}{m}E_x(t)\right) + \omega_p^2x\right] &= C_3. \end{aligned} \qquad (7)$$

We observe that at $t = 0$, (7) provides explicit solutions for the constants of integration C_1, C_2, and C_3 in terms of the initial values V_0, x_0, and $[(e/m)E_x\ (0)]$. We also observe that for the original system of Eqs. (3), (4), appropriate initial data would be of the form

$$V_0 = V_x(x, t = 0) = F(x),$$

$$\frac{e}{m}E_x(0) = \frac{e}{m}E_x(x, t = 0) = G(x).$$

These initial data give rise to functional relations among the constants C_1, C_2, and C_3. This method of construction of general solutions to the system (3), (4), by utilizing the functional relations that the initial data impose among the constants of integration, may be stated in the form of a theorem:

Theorem: Let $\phi(x_1, x_2)$ and $\Psi(x_1, x_2)$ be any C^2 functions of two variables whose first derivatives do not vanish. A general solution to the system (3), (4) may be found by solving the equations

$$\phi(C_1, C_3) = 0 = \Psi(C_2, C_3).$$

Proof: The proof consists of calculating the derivatives:

$$\begin{aligned} \frac{\partial\phi}{\partial x} = 0 &= \frac{\partial\phi}{\partial C_1}\frac{\partial C_1}{\partial x} + \frac{\partial\phi}{\partial C_3}\frac{\partial C_3}{\partial x},\\ \frac{\partial\phi}{\partial t} = 0 &= \frac{\partial\phi}{\partial C_1}\frac{\partial C_1}{\partial t} + \frac{\partial\phi}{\partial C_3}\frac{\partial C_3}{\partial t}, \quad \text{etc.,} \end{aligned}$$

and then substituting for C_1, C_2, and C_3 the integrals derived above. The resulting equations regenerate the system (3), (4). ∎

For example, suppose that we impose the following functional relations on the integrals:

$$\begin{aligned} C_1 &= f(C_3),\\ C_2 &= h(C_3), \end{aligned}$$

where $f(x)$ and $h(x)$ are arbitrary differentiable functions $f(x)$, $h(x) \neq 0$.

It follows that we have

$$\left[\frac{i\lambda^-}{\omega_p}\left(\frac{e}{m}E_x\right) + i\omega_pV_x\right] = -2\omega_p(\omega_p^2 - \tfrac{1}{4}v^2)^{\frac{1}{2}}e^{\lambda^+t}f\left(\frac{e}{m}E_x + \omega_p^2x\right) \qquad (8)$$

and

$$\left[\frac{-i\lambda^+}{\omega_p}\left(\frac{e}{m}E_x\right) - i\omega_pV_x\right] = -2\omega_p(\omega_p^2 - \tfrac{1}{4}v^2)^{\frac{1}{2}}e^{\lambda^-t}h\left(\frac{e}{m}E_x + \omega_p^2x\right). \qquad (9)$$

Any explicit choice of the functions $f(x)$ and $h(x)$ in Eqs. (8) and (9) will result in a pair of equations defining V_x and $(e/m)E_x$ as explicit functions of (x, t).

To illustrate this idea, let $f(x) \equiv 1 \equiv h(x)$. This choice results in the spatially independent solutions

$$\left(\frac{e}{m}E_x\right) = -2\omega_p^2e^{-v/2t}\{\cos[(\omega_p^2 - v^2/4)^{\frac{1}{2}}t]\}, \qquad (10)$$

$$V_x = -e^{-v/2t}\{v\cos[(\omega_p^2 - v^2/4)^{\frac{1}{2}}t] + 2(\omega_p^2 - v^2/4)^{\frac{1}{2}}\sin[(\omega_p^2 - v^2/4)^{\frac{1}{2}}t]\}. \qquad (11)$$

Once $(e/m)E_x$ and V_x have been determined, the electron density $n(x, t)$ can be derived from either Eqs. (1) or (2):

$$\frac{\partial}{\partial x}\frac{e}{m}E_x + \omega_p^2 = \frac{e^2}{m\epsilon_0}n(x, t) = \frac{-\partial}{\partial t}\left(\frac{e}{m}E_x\right)\Big/V_x.$$

It is easy to verify that (10) and (11) and the resulting value of $n(x, t)$ are exact solutions of the original nonlinear system (1), (2), and (3). For small values of the collision frequency, $\nu/\omega_p \ll 1$, these solutions represent pure longitudinal oscillations close to the plasma frequency; the collision frequency produces exponential damping.

3. DISCUSSION

A general solution has been found for the system of equations in the nonlinear model in the special case $B_y = B_z = E_y = E_z = V_y = V_z = 0$. The method of construction of the general solution involves the imposition of arbitrary functional relationships among the integrals of the associated system of ordinary differential equations, thus displaying an infinity of possible longitudinal oscillations which satisfy the equations. The method has been illustrated by the choice $f(x) \equiv 1 \equiv h(x)$ in Eqs. (8) and (9). This particular choice gives rise to spatially independent plasma oscillations (10) and (11). An infinity of possible solutions may be generated in this way, corresponding to different initial conditions for the system (1), (2), and (3). For example, if $f(x) \equiv x \equiv h(x)$, the resulting solution represents a longitudinal oscillation that grows linearly with (x) and contains Fourier components at all integral multiples of the plasma frequency. The choice $f(x) \equiv \cos(x) \equiv h(x)$ gives rise to interesting nonseparable wavelike solutions.

It is expected that further research will provide an extension of this method of integration to construct approximate general solutions to the nonlinear model in cases where the transverse field components do not vanish, and where spatial variations in more than one dimension are allowed. A subsequent paper will treat the effect of a strong transverse magnetic field, which will introduce coupling between the longitudinal oscillations described above and the transverse components of the field.

* This research was supported in part by the U.S. Air Force Cambridge Research Laboratories.

[1] R. F. Whitmer and E. B. Barrett, Phys. Rev. **121**, 661 (1961).
[2] R. F. Whitmer and E. B. Barrett, Phys. Rev. **125**, 1478 (1962).
[3] E. B. Barrett, R. F. Whitmer, and S. J. Tetenbaum, Phys. Rev. **135**, A369 (1964).
[4] S. J. Tetenbaum, R. F. Whitmer, and E. B. Barrett, Phys. Rev. **135**, A374 (1964).
[5] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1962), Vol. II, pp. 139–47.
[6] W. Kaplan, *Ordinary Differential Equations* (Addison-Wesley, New York, 1958).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Fluids of Particles with Short-Range Repulsion and Weak Long-Range Attractive Interaction. II. The Two-Particle Distribution Function

JOHN B. JALICKEE AND ARNOLD J. F. SIEGERT
Physics Department, Northwestern University, Evanston, Illinois 60201*
AND
DAVID J. VEZZETTI
Physics Department, University of Illinois at Chicago Circle, Chicago, Illinois†

(Received 17 February 1970)

In Paper I, we presented an expansion of the pressure and density in grand canonical form and corrections to the Maxwell rule for a system of particles with short-range repulsion and weak long-range attraction. These expansions can be ordered in powers of γ, the inverse range of the attractive potential. It was assumed that the thermodynamic functions and the molecular distribution functions of the reference system, i.e., the system with only the repulsive interaction, are given. In the present paper we have calculated the γ expansion of the pair distribution function, under the same assumption. The result is obtained by functional differentiation of the series for the pressure and presented in the form of a series of diagrams. The dominant order in γ of each diagram is the same as the order of that diagram in the series for the pressure, from which it is derived.

I. STATEMENT OF THE PROBLEM

In Paper I,[1] we obtained the pressure and density of a fluid of particles with short-range repulsion and weak long-range attractive interaction as functions of the fugacity and correction terms to the Maxwell rule. These results are given in terms of diagram expansions which can be ordered in ascending powers of the reciprocal range (γ) of the long-range attractive potential.

It is known that expansions in reciprocal range have several shortcomings. The expansion fails in an obvious manner at the van der Waals critical point. It also yields a phase transition for the 1-dimensional model of Kác, Uhlenbeck, and Hemmer[2] for $\gamma > 0$ at temperatures below the van der Waals critical temperature, while the exact solution shows that there is no phase transition. A modification of the γ expansion, stated in Ref. 1, avoids the obvious failure at the van der Waals critical point, and a similar modification for the Ising model[3] is known to be successful in treating the critical region, i.e., the neighborhood of the Weiss–Curie point, for the 2- and 3-dimensional cases.

While these objections affect the results only near the coexistence curve, another objection has been raised by a previous author,[4] which, if it were true, would cast doubt on the validity of the expansion over the entire range of temperature and fugacity. This is the objection that the approximation of the 2-particle distribution function for particles with hard-core repulsion and long-range attraction does not vanish when the centers of the two particles approach each other closer than a hard-core diameter. The calculation presented here yields directly an expansion in which each diagram individually vanishes in the hard-core region.

II. FUNCTIONAL DIFFERENTIATION OF THE SERIES FOR THE PRESSURE WITH RESPECT TO THE REPULSIVE PART OF THE INTERACTION

In Paper I we obtained the pressure $P(z)$ and number density $\rho(z)$ as functions of the fugacity z for a system with a pair potential energy consisting of a strong, short-range repulsion $u(\mathbf{r})$ and a weak, long-range attractive part $-v(\mathbf{r})$ under the assumption that the pressure $P_h(y)$ and number density $\rho_h(y)$ as well as all distribution functions of a reference system with only the repulsive potential $u(\mathbf{r})$ are given as functions of its fugacity y. The results for the pressure, number density, and a corrected Maxwell rule were given in terms of a correction function $h(y)$ as follows:

$$\beta P(z) = \beta P_h(y^*) - \tfrac{1}{2}\beta v_0 \rho^2(z) + h(y^*), \tag{1}$$

$$\rho(z) = \rho_h(y^*) + \frac{\partial h(y^*)}{\partial \ln y^*}, \tag{2}$$

where $\beta = (kT)^{-1}$, v_0 is the integral of $v(\mathbf{r})$ over all space, and y^* is that root of the equation

$$\ln y - \ln (z e^{-\frac{1}{2}\beta v(0)}) = \beta v_0 \left(\rho_h(y) + \frac{\partial h(y)}{\partial \ln y} \right) \tag{3}$$

which maximizes the pressure $P(z)$. Equation (3) is the grand canonical form of the corrected Maxwell construction as discussed in I.

The appearance of the two fugacities z and y is perhaps most easily understood by analogy with the Weiss theory of ferromagnetism, with $\ln (z e^{\frac{1}{2}\beta v(0)})$

and $\ln y$ corresponding to the impressed magnetic field and the local field respectively; $\ln y^*$ then corresponds to the value of the local field obtained by the Weiss construction. Without the correction terms in Eqs. (2) and (3), $\rho_h(y)$ corresponds to the magnetization of the paramagnet in the local field and $\rho(z) = \rho_h(y^*)$ to the magnetization of the ferromagnet.

One must note that $h(y)$, in the above, is understood as the thermodynamic limit of $h(y)$ given in Paper I. [We denote the latter by $\bar{h}(y)$.] For finite separation of the pair, only terms of order V in the partition function will contribute to the pair distribution function calculated from it. Terms of order V^0, however, will contribute to the fluctuation integral. We therefore have to check that such terms do not appear in our expansion. In Appendix A this has been done explicitly for $h(y)$ taken through order $\gamma^{3\nu-1}$ and in outline for terms of higher order in γ.

The 2-particle distribution function $\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2 \mid z)$ can be obtained by functional differentiation of the logarithm of the grand canonical partition function, $\ln Q(z)$, with respect to the interaction potential. It can also be obtained by differentiating with respect to either the short-range part $u(\mathbf{r})$ or the long-range part $-v(\mathbf{r})$, alone. We have chosen to differentiate with respect to $u(\mathbf{r})$ alone because differentiation with respect to $-v(\mathbf{r})$ will mix terms of different order in reciprocal range. (See also Appendix C.)

With the abbreviation $1, 2, \cdots$ used for the coordinate vectors $\mathbf{r}_1, \mathbf{r}_2, \cdots$, we have then

$$\rho^{(2)}(1, 2 \mid z) \equiv -\frac{\delta \ln Q(z)}{\delta\beta u(1, 2)} = -\frac{\delta}{\delta\beta u(1, 2)}\Big(\ln Q_h(y^*) - \frac{V}{2\beta v_0} \times (\ln y^* - \ln z e^{-\frac{1}{2}\beta v(0)})^2 + V\bar{h}(y^*)\Big). \tag{4}$$

Since for the reference system

$$-\frac{\delta \ln Q_h(y)}{\delta\beta u(1, 2)} \equiv \rho_h^{(2)}(1, 2 \mid y), \tag{5}$$

we have

$$\rho^{(2)}(1, 2 \mid z) = \Big(\rho_h^{(2)}(1, 2 \mid y) - \frac{\delta V\bar{h}(y)}{\delta\beta u(1, 2)}\Big)_{y=y^*} - \frac{\delta \ln y^*}{\delta\beta u(1, 2)}\Big[\frac{\partial}{\partial \ln y}\Big(\ln Q_h(y) - \frac{V}{2\beta v_0} \times (\ln y - \ln z e^{-\frac{1}{2}\beta v(0)})^2 + V\bar{h}(y)\Big)\Big]_{y=y^*} \tag{6}$$

The last term vanishes because of Eq. (3). We therefore obtain

$$\rho^{(2)}(1, 2 \mid z) = \rho_h^{(2)}(1, 2 \mid y^*) - \Big[\frac{\delta V\bar{h}(y^*)}{\delta\beta u(1, 2)}\Big]_{y=y^*} \equiv \rho_h^{(2)}(1, 2 \mid y^*) + h^{(2)}(1, 2 \mid y^*). \tag{7}$$

Clearly, $\rho_h^{(2)}(1, 2 \mid y)$ vanishes when $|\mathbf{r}_1 - \mathbf{r}_2|$ is less than a hard-core diameter if $u(\mathbf{r})$ is taken to be a hard-core potential. Paper I, in Appendix A, gives the following for $\bar{h}(y)$:

$$V\bar{h}(y) = -\tfrac{1}{2}\,\mathrm{Tr}\ln(1 - \beta v\mu_2) + V\bar{h}^{(1)}(y). \tag{8}$$

The first term is in obvious operator notation with v and μ_2 considered as kernels of integral operators. Here, μ_2 is the second of the sequence of modified Ursell functions $\mu_n(1, 2, \cdots, n \mid y)$ of the reference system. $\bar{h}^{(1)}(y)$ is a sum of connected diagrams with hypervertices $\mu_n(1, 2, \cdots, n \mid y)$, $n \geq 3$, and bonds $\tilde{v}_1(1, 2)$ defined in the same operator notation by

$$\tilde{v}_1 = v/(1 - \beta v\mu_2). \tag{9}$$

We note that v and μ_2 commute as operators since $v(\mathbf{r}_1, \mathbf{r}_2)$ and $\mu_2(\mathbf{r}_1, \mathbf{r}_2 \mid y)$ depend only on the vectorial distance $\mathbf{r}_1 - \mathbf{r}_2$ under Born–von Kármán boundary conditions.

In order to perform the differentiation indicated in Eq. (7), we need thus only the functions

$$\lambda_{n+2}(1', 2', \cdots, n' \mid 1, 2 \mid y) \equiv \frac{\delta\mu_n(1', 2', \cdots, n' \mid y)}{\delta[-\beta u(1, 2)]}, \tag{10}$$

for $n \geq 2$. The modified Ursell functions are defined by

$$\mu_n(1', 2', \cdots, n' \mid y) \equiv \Big[\frac{\delta^n \ln Q_h(\{\phi\} \mid y)}{\delta\phi(1') \cdots \delta\phi(n')}\Big]_{\phi=0}, \tag{11}$$

where $Q_h(\{\phi\} \mid y)$ is the grand canonical partition function of the reference system at fugacity y in the presence of an impressed potential $-\beta^{-1}\phi(\mathbf{r})$, i.e.,

$$Q_h(\{\phi\} \mid y) = \sum_{N=0}^{\infty} \frac{y^N}{N!} \times \int \exp\Big(-\tfrac{1}{2}\beta \sum_{i,j}{}' u(\mathbf{r}_i - \mathbf{r}_j) + \sum_i \phi(\mathbf{r}_i)\Big) \times \prod_{i=1}^{N} d^\nu r_i, \tag{12}$$

where ν = number of dimensions. Therefore, with

$\rho_h^{(2)}(1, 2|\{\phi\}|y)$ denoting the pair distribution function of hard particles in a field $\phi(\mathbf{r})$,

$$\lambda_{n+2}(1', 2', \cdots, n'|1, 2|y) = \left[\frac{\delta^n \rho_h^{(2)}(1, 2|\{\phi\}|y)}{\delta\phi(1')\cdots\delta\phi(n')}\right]_{\phi=0}. \quad (13)$$

The relation of the λ functions to the modified Ursell functions and to the molecular distribution functions, together with some of their properties of interest in the present context, is given in Appendix B. If $u(\mathbf{r})$ is taken to be a hard-core potential, then, clearly, $\lambda_{n+2}(1', 2', \cdots, n'|1, 2|y)$ vanishes when $|\mathbf{r}_1 - \mathbf{r}_2|$ is less than a hard-core diameter since $\rho_h^{(2)}(1, 2|\{\phi\}|y)$ vanishes for any $\phi(\mathbf{r})$ in this case.

Taking the first term in Eq. (8), we have

$$\begin{aligned}\frac{\delta}{\delta[-\beta u(1,2)]}[-\tfrac{1}{2}\,\mathrm{Tr}\ln(1-\beta v\mu_2)] &= \tfrac{1}{2}\,\mathrm{Tr}\left(\frac{\beta v\{\delta\mu_2/\delta[-\beta u(1,2)]\}}{1-\beta v\mu_2}\right)\\ &= \tfrac{1}{2}\,\mathrm{Tr}\left(\beta\tilde{v}_1\frac{\delta\mu_2}{\delta[-\beta u(1,2)]}\right)\\ &= \tfrac{1}{2}\beta\int d^\nu r_3 d^\nu r_4 \tilde{v}_1(3,4)\lambda_4(3,4|1,2|y). \quad (14)\end{aligned}$$

This vanishes where it should. Next

$$\frac{\delta\bar{h}^{(1)}(y)}{\delta[-\beta u(1,2)]}$$

is obtained by differentiating each hypervertex and each bond of the diagram series for $h^{(1)}(y)$. Differentiating a hypervertex $\mu_n(1', 2', \cdots, n'|y)$ replaces this μ_n by $\lambda_{n+2}(1', 2', \cdots, n'|1, 2|y)$, which has the proper behavior as a function of $\mathbf{r}_1$ and $\mathbf{r}_2$. Finally, differentiating a bond $\tilde{v}_1(1', 2')$ gives

$$\begin{aligned}\frac{\delta\beta\tilde{v}_1(1',2')}{\delta[-\beta u(1,2)]} &= \frac{\delta}{\delta[-\beta u(1,2)]}\left(\frac{\beta v}{1-\beta v\mu_2}\right)_{1',2'}\\ &= \left[\frac{\beta v\{\delta\mu_2/\delta[-\beta u(1,2)]\}\beta v}{(1-\beta v\mu_2)^2}\right]_{1',2'}\\ &= \int d^\nu r_{3'} d^\nu r_{4'}\beta\tilde{v}_1(1',3')\\ &\quad\times\lambda_4(3',4'|1,2|y)\beta\tilde{v}_1(4',2'), \quad (15)\end{aligned}$$

where we use the operator notation for the intermediate steps. Again, Eq. (15) has the proper behavior. Thus, when $u(\mathbf{r})$ approaches a hard-core potential, every term in the expansion of $\rho^{(2)}(1, 2|z)$ vanishes for $|\mathbf{r}_1 - \mathbf{r}_2|$ less than a hard-core diameter. Note that no rearrangement of the original series is required to achieve this result as it would be if the differentiation had been carried out with respect to $v(\mathbf{r})$ instead.

III. DIAGRAMMATIC SERIES FOR $\rho^{(2)}(1, 2|z)$

The diagrammatic series for $\rho^{(2)}(1, 2|z)$ may now be obtained by applying the rules for differentiation to the diagrammatic series for the function $\bar{h}^{(1)}(y)$, which was given in Paper I, and adding the result to the term given by Eq. (14). For the convenience of the reader, we repeat here the rules for the construction of these diagrams. Draw S circles ($S \geq 1$) labeled $C_1, \cdots, C_j, \cdots, C_S$, each containing $n_j \geq 3$ points such that $\sum_{j=1}^{S} n_j$ is even. Connect these points by lines such that each point is connected to one and only one other point. (Lines connecting two points in the same circle are permitted.) Disconnected diagrams and diagrams which would become disconnected upon removal of one line are excluded. With any diagram G, we associate an integral I_G in the following way. Label the points in circle C_j by $\mathbf{r}_{j,1}, \mathbf{r}_{j,2}, \cdots, \mathbf{r}_{j,n_j}$. For each circle C_j write a factor $\mu_{n_j}(\mathbf{r}_{j,1}, \cdots, \mathbf{r}_{j,n_j}|y)$. For a line connecting points $\mathbf{r}_{j,k}$ and $\mathbf{r}_{j',k'}$ write a factor $\beta\tilde{v}_1(\mathbf{r}_{j,k} - \mathbf{r}_{j',k'})$, where $\tilde{v}_1$ is defined by Eq. (9). I_G is then the integral over all coordinates of the product of these factors. With any diagram G, we also associate a symmetry number S_G. Let t_{ij} be the number of lines connecting circles C_i and C_j and let S'_G be the number of permutations of the labels of the circles which leave all t_{ij} invariant. Then

$$S_G = \prod_{i\leq j}(t_{ij}!)\prod_i 2^{t_{ii}}S'_G. \quad (16)$$

We call two diagrams G and G' equivalent if the integrands of I_G and $I_{G'}$ are the same, except for the labeling of the variables of integration. Divide the class of diagrams described above into subclasses of equivalent diagrams. Select one member from each subclass to form the class $\mathfrak{G}$. Then

$$\bar{h}^{(1)}(y) = V^{-1}\sum_{G\in\mathfrak{G}}\frac{I_G}{S_G}. \quad (17)$$

In Paper I we took the potential $v(\mathbf{r})$ to be

$$v(\mathbf{r}) = \gamma^\nu\Phi(\gamma|\mathbf{r}|). \quad (18)$$

We then showed that a diagram with B bonds and S circles cannot contribute any power of γ less than $\nu(B - S + 1)$. The diagrams contributing to $\bar{h}^{(1)}(y)$ in order $\gamma^{2\nu}$ and order $\gamma^{3\nu}$ are shown in Figs. 1 and 2, respectively.

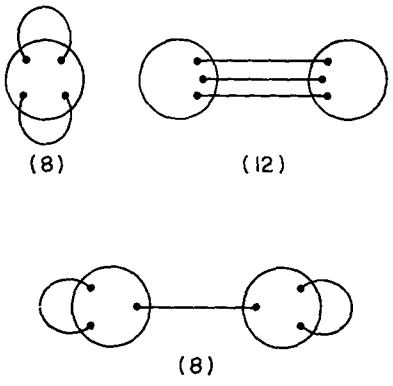
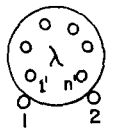


FIG. 1. Diagrams needed through order $\gamma^{3\nu-1}$. The last of these diagrams gives a contribution of zero. Symmetry numbers are shown in parentheses.

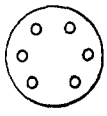
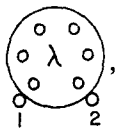
The functional differentiation introduces new hypervertices which represent the functions

$$\lambda_{n+2}(1', 2', \cdots, n' | 1, 2 | y).$$

These will be symbolized by circles with n points (field points) and two open points (root points) labeled 1 and 2, i.e.,

$$\lambda_{n+2}(1', 2', \cdots, n' | 1, 2 | y) \equiv$$  (19)

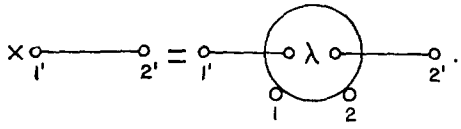
With this notation, Eqs. (10), (14), and (15), e.g., become

$$\frac{\delta}{\delta(-\beta u(1,2))}$$  $$=$$  (10′)

$$\frac{\delta}{\delta(-\beta u(1,2))}[\mathrm{Tr}\ \ln(1-\beta v\mu_2)] =$$  (14′)

and

$$\frac{\delta\beta\tilde{v}(1',2')}{\delta(-\beta u(1,2))} = \frac{\delta}{\delta(-\beta u(1,2))}$$

 (15′)

The general rule for functional differentiation of a diagram series[5] then leads to the following result for the diagram series for $\rho^{(2)}(1, 2 | z)$: Each diagram contains one λ_{n+2} circle (for some $n \geq 2$); any number (including zero) of μ_m circles ($m \geq 3$); each field point is connected to one other field point by a $\beta\tilde{v}_1$ bond. Disconnected diagrams and diagrams which would become disconnected upon removal of one bond are excluded because such diagrams do not occur in the original diagrammatic series and cannot be produced by differentiation. The integral over all field points of the product of these factors is denoted by $I_G(1, 2 | y)$. The symmetry number is calculated, as explained, before Eq. (16). In computing S'_G, it must be remembered that a circle with two root points is distinct from a circle without root points. The subclass $\mathfrak{G}$ is defined as before [text preceding Eq. (17)]. We then have

$$\rho^{(2)}(1, 2 | z) = \rho_h^{(2)}(1, 2 | y^*) + \sum_{G \in \mathfrak{G}} \frac{I_G(1, 2 | y^*)}{S_G}, \quad (20)$$

where y^* is defined by Eq. (3).

If we use the same considerations as in Paper I to determine the lowest order of γ which a given diagram may contribute, together with the properties of the λ functions discussed in Appendix B, it is an easy matter to show that each diagram of $V\bar{h}^{(1)}(y)$ retains

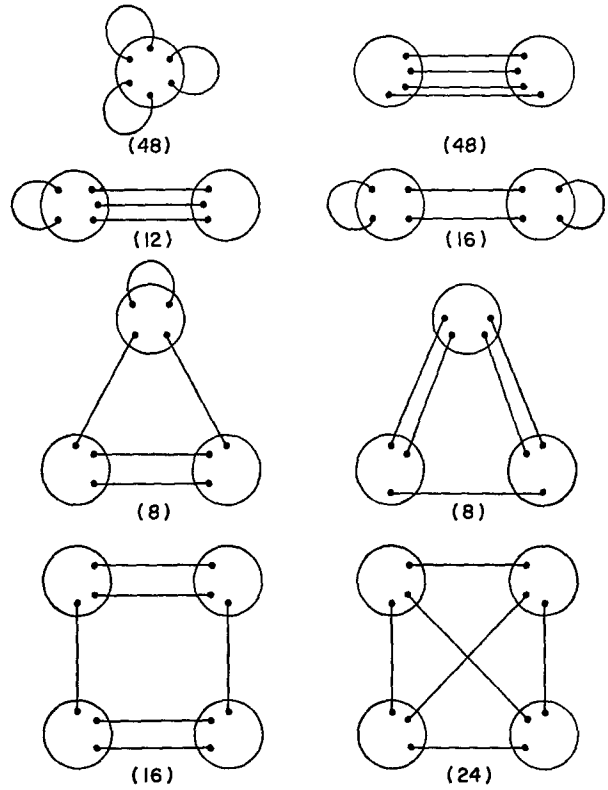
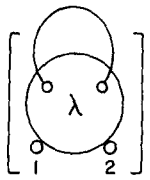


FIG. 2. Diagrams needed through order $\gamma^{4\nu-1}$. Symmetry numbers are shown in parentheses.

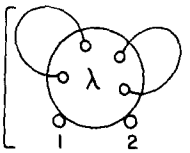
its order in γ after it is differentiated. We thus obtain

$$\rho^{(2)}(1,2\,|\,z) = \rho_h^{(2)}(1,2\,|\,y^*) + \Big[$$



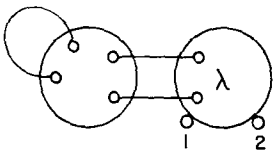
(2)

$+$



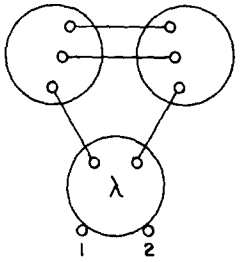
(8)

$+$



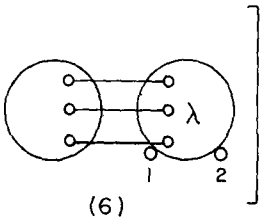
(4)

$+$



(4)

$+$



(6)

$\Big] + O(\gamma^{3\nu}).$ (21)

Here the term of $O(\gamma^{\nu})$ is separated from the four terms of $O(\gamma^{2\nu})$ by square brackets. Symmetry numbers are shown in parenthesis.

We note here that the fact that each diagram retains its order in γ after differentiation is a result of our having chosen to differentiate with respect to the function $u(1,2)$ (which does not contain γ) and the form of our diagrams. On the other hand, differentiating a diagram of $V\bar{h}^{(1)}(y)$, whose dominant order is $\gamma^{m\nu}$ with respect to $v(1,2)$, would lead to several diagrams whose dominant orders can be $\gamma^{(m-1)\nu}$ or $\gamma^{m\nu}$. This, of course, would make the ordering of the resulting $\rho^{(2)}$ series more difficult.

We have carried through the calculation of $\rho^{(2)}(1,2\,|\,z)$ to order $\gamma^{3\nu-1}$ also by differentiation with respect to $v(1,2)$. This requires terms to order $\gamma^{4\nu-1}$ in the pressure. The result to this order is identical with the result obtained by differentiation with respect to $u(1,2)$, as it should be. The calculation is given in detail in Appendix C, through order $\gamma^{2\nu-1}$.

IV. CHANGE TO DENSITY AS INDEPENDENT VARIABLE

Our expression (21) for the 2-particle distribution function, though correct, must be supplemented by Eq. (3) to determine the appropriate y^* for a given z. We may eliminate this step by writing $\rho^{(2)}$ in a 1-phase region immediately as a function of density.

In such a region, y^* and z may differ by a large amount, but $\rho(z)$ and $\rho_h(y^*)$ will differ only by quantities of order γ^{ν}. These observations follow from Eqs. (3) and (2), respectively, and the fact that $\bar{h}(y)$ is of order γ^{ν}. It is therefore convenient to change from the variables z and y^* to the variables ρ and ρ_h by the following definitions:

$$\begin{aligned}
&\rho(z) \equiv \rho, \quad \rho_h(y^*) \equiv \rho_h,\\
&\bar{h}(y^*) = \bar{h}(y^*[\rho_h]) \equiv \eta(\rho_h),\\
&\rho^{(2)}(1,2\,|\,z) \equiv \tilde{\rho}^{(2)}(1,2\,|\,\rho),\\
&\rho_h^{(2)}(1,2\,|\,y^*) \equiv \tilde{\rho}_h^{(2)}(1,2\,|\,\rho_h),\\
&\frac{\partial \bar{h}(y^*)}{\partial \ln y^*} \equiv \eta_1(y^*) \equiv \tilde{\eta}_1(\rho_h).
\end{aligned} \tag{22}$$

Using these, we obtain from (2)

$$\rho = \rho_h + \tilde{\eta}_1(\rho_h). \tag{23}$$

Calling the right-hand side of (21) $F(y^*)$ and using (22), we have

$$\tilde{\rho}^{(2)}(1,2\,|\,\rho) = F(y^*) = F(y^*[\rho_h]) \equiv \tilde{F}(\rho_h). \tag{24}$$

Using (23) in (24), we see

$$\begin{aligned}
\tilde{\rho}^{(2)}(1,2\,|\,\rho) &= \tilde{\rho}^{(2)}(1,2\,|\,\rho_h + \tilde{\eta}_1)\\
&= \left(1 + \sum_{k=1}^{\infty} \frac{(\tilde{\eta}_1)^n}{n!} \frac{\partial^n}{\partial \rho_h^n}\right) \tilde{\rho}^{(2)}(1,2\,|\,\rho_h)\\
&\equiv [1 + \hat{q}(\rho_h)]\tilde{\rho}^{(2)}(1,2\,|\,\rho_h) = \tilde{F}(\rho_h),
\end{aligned} \tag{25}$$

where the next to last equality defines the operator $\hat{q}(\rho_h)$. In the last equality we see that the variable ρ has disappeared. Therefore, ρ_h takes the status of independent variable, and we can rename it ρ. Thus, we find, finally,

$$[1 + \hat{q}(\rho)]\tilde{\rho}^{(2)}(1,2\,|\,\rho) = \tilde{F}(\rho) \tag{26}$$

and

$$\tilde{\rho}^{(2)}(1,2\,|\,\rho) = [1 + \hat{q}(\rho)]^{-1}\tilde{F}(\rho). \tag{27}$$

Here the inverse operator is defined by its power series expansion. The right side of (27) contains only functions pertaining to the hard-core system, evaluated at the true density ρ.

To evaluate (27) explicitly, we expand both $\hat{q}(\rho)$ and $\tilde{F}(\rho)$ in orders of γ. To evaluate $\gamma^{2\nu-1}$ (that is, up to but not including $\gamma^{2\nu}$), we obtain, in this way,

$$\begin{aligned}\tilde{\rho}^{(2)}(1,2\,|\,\rho) &= \tilde{\rho}_h^{(2)}(1,2\,|\,\rho) \\ &\quad + \tfrac{1}{2}\beta\int d^\nu r_{1'} d^\nu r_{2'} \tilde{\lambda}_4(1',2'|\,1,2\,|\rho)\,\tilde{v}_1(1',2') \\ &\quad - \tfrac{1}{2}\rho^2 K_h(\rho)\frac{\partial\tilde{\rho}_h^{(2)}(1,2\,|\,\rho)}{\partial\rho} \\ &\quad \times\int d^\nu r\,\frac{\partial\tilde{\mu}_2(\mathbf{r}\,|\,\rho)}{\partial\rho}\,\tilde{v}_1(\mathbf{r}). \qquad (28)\end{aligned}$$

Here $K_h(\rho)$ is the isothermal compressibility $\rho_h\partial\rho_h/\partial\beta P_h$ of a hard-core system at density ρ, and $\tilde{\mu}_2(\mathbf{r}\,|\,\rho)$ and $\tilde{\lambda}_4(1',2'|\,1,2\,|\rho)$ are the functions obtained by evaluating

$$\mu_2(\mathbf{r}\,|\,y^*) = \mu_2(\mathbf{r}\,|\,y^*(\rho_h)) \equiv \tilde{\mu}_2(\mathbf{r}\,|\,\rho_h) \qquad (29)$$

and

$$\begin{aligned}\lambda_4(1',2'|\,1,2\,|y^*) &= \lambda_4(1',2'|\,1,2\,|y^*(\rho_h)) \\ &\equiv \tilde{\lambda}_4(1',2'|\,1,2\,|\rho_h), \qquad (30)\end{aligned}$$

at $\rho_h = \rho$, in accordance with our above prescription. For example,

$$\tilde{\mu}_2(r\,|\,\rho) = \tilde{\rho}_h^{(2)}(1,2\,|\,\rho) - \rho^2 + \rho\delta(r). \qquad (31)$$

The function $\tilde{\lambda}_4$ may be obtained from Eq. (10). Carrying through this differentiation and simplifying the integrals in (28) as much as possible, we obtain

$$\begin{aligned}\tilde{\rho}^{(2)}(1,2\,|\,\rho) &= \tilde{\rho}_h^{(2)}(1,2\,|\,\rho)[1+\beta\tilde{v}_1(1,2)] \\ &\quad + \beta\int d^\nu\mathbf{r}_{1'}\tilde{\rho}_h^{(3)}(1',1,2\,|\,\rho) \\ &\quad \times[\tilde{v}_1(1,1') + \tilde{v}_1(2,1')] \\ &\quad + \tfrac{1}{2}\beta\iint d^\nu r_{1'} d^\nu r_{2'}[\tilde{\rho}_h^{(4)}(1',2',1,2\,|\,\rho) \\ &\quad - \tilde{\rho}_h^{(2)}(1',2'\,|\,\rho)\tilde{\rho}_h^{(2)}(1,2\,|\,\rho)]\tilde{v}_1(1',2') \\ &\quad - \tfrac{1}{2}\rho^2 K_h(\rho)\frac{\partial\tilde{\rho}_h^{(2)}(1,2\,|\,\rho)}{\partial\rho} \\ &\quad \times\int d^\nu r\,\frac{\partial\tilde{\rho}_h^{(2)}(\mathbf{r}\,|\,\rho)}{\partial\rho}\tilde{v}_1(\mathbf{r}) + O(\gamma^{2\nu}). \qquad (32)\end{aligned}$$

Several features of this expression are worth remarking on. First the expression is finite for $V\to\infty$, for, although

$$\iint d^\nu r_{1'} d^\nu r_{2'}\tilde{\rho}_h^{(2)}(1',2'\,|\,\rho)\tilde{v}_1(1',2')$$

is proportional to the volume, this is compensated by the integral over $\tilde{\rho}^{(4)}(1',2',1,2\,|\,\rho)$. Secondly, the value of $\tilde{v}_1(\mathbf{r})$ is never needed for $|\mathbf{r}|$ less than a hard-core diameter since all the distribution functions $\tilde{\rho}_h^{(n)}$ vanish identically in this region. This fact would permit us to use a Coulomb potential for $v(\mathbf{r})$ even though its transform $\tilde{v}_1(\mathbf{r})$ is singular at the origin.

V. SUMMARY

The pair distribution function has been obtained by functional differentiation of our diagrammatic expansion for the pressure with respect the short-range part of the pair potential. We have shown that the function obtained in this way has the necessary property of vanishing for $|\mathbf{r}_1 - \mathbf{r}_2| < a$ when the potential contains a hard core of diameter a. We have checked through order $\gamma^{3\nu-1}$ that this function is identical to that obtained by functional differentiation with respect to the long-range, attractive part of the potential. There is little doubt that this is true to all orders. In order to perform this latter differentiation, it is necessary to relax a restriction on $v(\mathbf{r})$ which had been imposed in Paper I, namely, that $v(\mathbf{r})$ considered as a kernel of an integral equation be positive definite. This restriction facilitated the estimation of the diagrams in orders of γ. It is not necessary, however, for any of the formal expansions.

In view of the above remarks, our formalism can now be applied to systems with Coulomb potentials. We can take for $v(\mathbf{r})$ the potential of a uniformly charged sphere of the size of the hard core. Outside the core, this is identical with the Coulomb potential. However, because this potential is finite inside the core, both $v(0)$ and $\tilde{v}_1(0)$ are now finite. The expansion in diagrams is now identical to that given in the text. It should be noted, however, that the diagram series will no longer be ordered in powers of γ. Thus, the proper estimation of the terms in this series requires a separate analysis from that given here.

APPENDIX A

In this appendix we will describe a procedure for showing that

$$\begin{aligned}\ln Q(z) &= \ln Q_h(y^*) - (V/2\beta v_0) \\ &\quad \times(\ln y^* - \ln z e^{-\frac{1}{2}\beta v(0)})^2 + V\hat{h}(y^*), \qquad (A1)\end{aligned}$$

where terms which explicitly depend on V ($V\to\infty$) less strongly than V^0 are neglected. We then carry out this procedure explicitly through terms of order $\gamma^{3\nu-1}$ in $h(y)$.

The expression for the partition function which we use is derived in Paper I by applying the saddle point

method to

$$Q(z) = \int_{-\infty}^{\infty} dc_0 \exp\left[-\tfrac{1}{2}c_0^2 + \ln Q_h(y) + Vh(y)\right], \quad \text{(A2)}$$

where

$$\ln y = \ln z e^{-\frac{1}{2}\beta v(0)} + c_0(\beta v_0/V)^{\frac{1}{2}}. \quad \text{(A3)}$$

Paper I gives $h(y)$, and it is identical to $\bar{h}(y)$, given in the text, if $\tilde{v}_1(\mathbf{r})$ is replaced by

$$\tilde{v}(\mathbf{r}) = \tilde{v}_1(\mathbf{r}) - \frac{1}{V}\int d^\nu r \tilde{v}_1(\mathbf{r}) = \tilde{v}_1(\mathbf{r}) - \frac{1}{V}\left(\frac{v_0}{1-\beta v_0 \mu_{20}}\right). \quad \text{(A4)}$$

The result of the integration is

$$\ln Q(z) = \ln Q_h(z) - (V/2\beta v_0) \times (\ln y^* - \ln z e^{-\frac{1}{2}\beta v(0)})^2 + Vh(y^*) - \tfrac{1}{2}\ln\left(1 - \beta v_0 \frac{\partial^2}{\partial \ln y^{*2}}[\beta\rho_h(y^*) + \bar{h}(y^*)]\right), \quad \text{(A5)}$$

where y^* is the saddle point defined in the text. The last term comes from expanding the exponent to second order and performing the integration. The problem, therefore, is to show that

$$h(y^*) - \bar{h}(y^*) = \frac{1}{2V}\ln\left(1 - \beta v_0 \frac{\partial^2}{\partial \ln y^{*2}}[\beta\rho_h(y^*) + \bar{h}(y^*)]\right) = \frac{1}{2V}\left[\ln(1-\beta v_0\mu_{20}) + \ln\left(1 - \frac{\beta v_0}{1-\beta v_0\mu_{20}}\frac{\partial^2 \bar{h}(y^*)}{\partial \ln y^{*2}}\right)\right], \quad \text{(A6)}$$

neglecting higher terms in V^{-1}.

To show this through any order in γ, one may proceed as follows. Take each term (or diagram) which contributes to $h(y)$ and replace $\tilde{v}(\mathbf{r})$ by the right-hand side of (A4). Each term may then be divided into parts which involve $\tilde{v}_1(\mathbf{r})$ and V^{-1} times its integral. The part which is simply the replacement of $\tilde{v}$ by $\tilde{v}_1$ is just the corresponding term in $\bar{h}(y^*)$; the other parts may be ordered in powers of V^{-1}. One then shows that the sum of all of the V^{-1} parts (of all the terms) is just an expansion of the right-hand side of (A6) taken through an appropriate order in γ.

For the $\gamma^{3\nu-1}$ approximation the sum of V^{-1} parts is easily shown to be

$$\frac{1}{2V}\ln(1-\beta v_0\mu_{20}) - \frac{\beta^2}{4V^2}\frac{v_0}{1-\beta v_0\mu_{20}} \times \int d^\nu r_1 \cdots \int d^\nu r_4 \tilde{v}_1(1,2)\mu_4(1,2,3,4 \mid y^*) - \frac{\beta^3}{4V^2}\frac{v_0}{1-\beta v_0\mu_{20}}\int d^\nu r_1 \cdots \int d^\nu r_6 \tilde{v}_1(1,4) \times \tilde{v}_1(2,5)\mu_3(1,2,3 \mid y^*)\mu_3(4,5,6 \mid y^*) = \frac{1}{2V}\ln(1-\beta v_0\mu_{20}) - \frac{1}{2V}\frac{\beta v_0}{1-\beta v_0\mu_{20}}\frac{\partial^2}{\partial \ln y^{*2}}\left(\frac{1}{2V}\operatorname{Tr}\ln(1-\beta v\mu_2)\right), \quad \text{(A7)}$$

where we have used the relation

$$\int d^\nu r_{n+1} \cdots \int d^\nu r_m \mu_m(1,2,\cdots,m \mid y) = \frac{\partial^{m-n}}{\partial \ln y^{m-n}}\mu_n(1,2,\cdots,n \mid y). \quad \text{(A8)}$$

Comparing (A7) with (A6) shows the former to be just the first two terms in an expansion of the latter, with $\bar{h}(y^*)$ approximated by $-1/2V \operatorname{Tr}\ln(1-\beta v\mu_2)$. We have carried out this calculation for the $\gamma^{4\nu-1}$ terms, but, since it is rather lengthy, we will not give it here.

Using diagram analysis, we may give a rigorous proof of (A6). This proof consists in showing that more than doubly connected diagrams do not contribute to $h - \bar{h}$ to order V^{-1} and that the resummation of all doubly connected diagrams contributes just the right-hand side of (A6).

APPENDIX B

If $Q\{\phi\}$ is the grand canonical partition function of a system of particles in the presence of an impressed potential $-\beta^{-1}\phi(\mathbf{r})$, the modified Ursell functions μ_n and molecular distribution functions $\rho^{(n)}$ are defined by

$$\mu_n(1,\cdots,n) \equiv \frac{\delta^n \ln Q\{\phi\}}{\delta\phi(1)\cdots\delta\phi(n)} \quad \text{(B1)}$$

and

$$\rho^{(n)}(1,\cdots,n) \equiv e^{\phi(1)}\cdots e^{\phi(n)}Q^{-1}\cdot\frac{\delta^n Q\{\phi\}}{\delta(e^{\phi(1)})\cdots\delta(e^{\phi(n)})}. \quad \text{(B2)}$$

We have suppressed the dependence on fugacity and the dependence of $\rho^{(n)}$, μ_n, and λ_n on ϕ in the notation. The functions $\rho^{(n)}$, μ_n, and λ_n when used in the main body of the paper are understood to be

specialized to $\phi = 0$. From (B1) it follows that

$$\frac{\delta\mu_n(1,\cdots,n)}{\delta\phi(n+1)} = \mu_{n+1}(1,\cdots,n,n+1) \quad \text{(B3)}$$

and from (A2)

$$\frac{\delta\rho^{(n)}(1,\cdots,n)}{\delta\phi(n+1)} = \rho^{(n+1)}(1,\cdots,n+1) - \rho^{(1)}(n+1)\rho^{(n)}(1,\cdots,n) + \rho^{(n)}(1,\cdots,n)\sum_{i=1}^{n}\delta(\mathbf{r}_i - \mathbf{r}_{n+1}). \quad \text{(B4)}$$

The functions $\lambda_{n+2}(1',\cdots,n' \mid 1,2)$ are then given, according to Eq. (13), in two equivalent forms by

$$\lambda_{n+2}(1',\cdots,n' \mid 1,2) = \frac{\delta^n\rho^{(2)}(1,2)}{\delta\phi(1')\cdots\delta\phi(n')} = \frac{\delta^n[\mu_2(1,2) + \mu_1(1)\mu_1(2) - \mu_1(1)\delta(1-2)]}{\delta\phi(1')\cdots\delta\phi(n')}, \quad \text{(B5)}$$

and satisfy the recursion formula

$$\lambda_{n+2+1}(1',\cdots,n',n'+1 \mid 1,2) = \frac{\delta\lambda_{n+2}(1',\cdots,n' \mid 1,2)}{\delta\phi(n'+1)}. \quad \text{(B6)}$$

Using (B3)–(B6), we can successively generate the λ functions in terms of either the μ_n or the $\rho^{(n)}$. For example,

$$\lambda_2(1,2) = \mu_2(1,2) + \mu_1(1)\mu_1(2) - \mu_1(1)\delta(1-2) = \rho^{(2)}(1,2), \quad \text{(B7)}$$

$$\lambda_3(1' \mid 1,2) = \mu_3(1,2,1') + \mu_1(1)\mu_2(2,1') + \mu_1(2)\mu_2(1,1') - \mu_2(1,1')\delta(1-2) = \{\rho^{(3)}(1,2,1') - \rho^{(2)}(1,2)\rho^{(1)}(1')\} + \rho^{(2)}(1,2)[\delta(1-1') + \delta(2-1')], \quad \text{(B8)}$$

$$\begin{aligned}\lambda_4(1',2' \mid 1,2) &= \mu_4(1,2,1',2') + \mu_1(1)\mu_3(2,1',2') + \mu_1(2)\mu_3(1,1',2') + \mu_2(1,1')\mu_2(2,2') + \mu_2(1,2')\mu_2(2,1') - \mu_3(1,1',2')\delta(1-2)\\ &= \{\rho^{(4)}(1,2,1',2') - \rho^{(1)}(2')\rho^{(3)}(1,2,1') - \rho^{(1)}(1')\rho^{(3)}(1,2,2') - \rho^{(2)}(1,2)\rho^{(2)}(1',2') + 2\rho^{(2)}(1,2)\rho^{(1)}(1')\rho^{(1)}(2')\}\\ &\quad + \{\rho^{(3)}(1,2,1') - \rho^{(1)}(1')\rho^{(2)}(1,2)\} \times [\delta(1-2') + \delta(2-2') + \tfrac{1}{2}\delta(1'-2')]\\ &\quad + \{\rho^{(3)}(1,2,2') - \rho^{(1)}(2')\rho^{(2)}(1,2)\} \times [\delta(1-1') + \delta(2-1') + \tfrac{1}{2}\delta(1'-2')]\\ &\quad + \rho^{(2)}(1,2)[\delta(1-2') + \delta(2-2')] \times [\delta(1-1') + \delta(2-1')]. \end{aligned} \quad \text{(B9)}$$

We note in passing that these expressions exhibit a transparent structure. The factors in braces as well as $\rho^{(2)}(1,2)$ are derived from the Ursell functions $F_{n'+1}(1,1',2',\cdots,n')$ by replacing particle 1 by the pair (1, 2) (Ref. 6).

For our purposes, the following two properties are essential: (i) From the first line of Eq. (B5), it follows that

$$\lambda_{n+2}(1',\cdots,n' \mid 1,2) = 0 \quad \text{for} \quad |\mathbf{r}_1 - \mathbf{r}_2| \le a \quad \text{(B10)}$$

if the repulsive potential has a hard core of diameter a; this is, of course, borne out by the explicit expressions; (ii) the functions $\lambda_{n+2}(1',2',\cdots,n' \mid 1,2)$ are given, by applying the recursion formula (B6) to Eq. (B7) and using Eq. (B3), as

$$\lambda_{n+2}(\mathbf{n}' \mid 1,2) = \mu_{n+2}(\mathbf{n}',1,2) + \sum_{\mathbf{a}'\subset\mathbf{n}'}\mu_{a+1}(\mathbf{a}',1)\mu_{n-a+1}(\bar{\mathbf{a}}',2) - \delta(1-2)\mu_{n+1}(\mathbf{n}',1), \quad \text{(B11)}$$

where $\mathbf{n}'$ is the set of n numbers $(1',2',\cdots,n')$ and where $\mathbf{a}'$ is a subset of $\mathbf{n}'$ and $\bar{\mathbf{a}}'$ its complement in $\mathbf{n}'$, a the number of its elements, and the sum extends over all subsets $\mathbf{a}'$ of $\mathbf{n}'$ (including the null set and $\mathbf{n}'$). Since the functions μ_n have an effective range[7] independent of γ, any μ_n with one root point can be replaced in our diagrams by

$$\mu_{n+1}(1',\cdots,n',1) \sim \frac{\partial^n\rho_h(y)}{\partial(\ln y)^n}\prod_{k=1}^{n}\delta(\mathbf{r}_1 - \mathbf{r}_k') \quad \text{(B12)}$$

for the purpose of obtaining the dominant order in γ. For the same purpose, functions μ_n with two root parts may be replaced, by the same argument, by

$$\mu_{n+2}(1',\cdots,n',1,2) \sim \frac{\partial^n\mu_2(1,2)}{\partial(\ln y)^n}\prod_{k=1}^{n}\delta(\mathbf{r}_2 - \mathbf{r}_k'). \quad \text{(B13)}$$

APPENDIX C

In this appendix we will show that the pair distribution function derived by functional differentiation of $\ln Q(z)$ with respect to the long range potential $-v(\mathbf{r})$ agrees with that given in Eq. (21) through order $\gamma^{2\nu-1}$. We begin from the expression

$$\rho^{(2)}(1,2 \mid z) \equiv \frac{\delta \ln Q(z)}{\delta\beta v(1,2)}. \quad \text{(C1)}$$

Using Eq. (1) for $\ln Q(z)$ and the stationary condition, (3), we obtain

$$\rho^{(2)}(1,2 \mid z) = \rho^2(z) - \rho(z)\delta(1-2) + \left[\frac{\delta V h(y)}{\delta\beta v(1,2)}\right]_{y=y^*}. \quad \text{(C2)}$$

Carrying out the functional differentiation of $V\bar{h}(y)$, we obtain

$$\left[\frac{\delta V\bar{h}(y)}{\delta\beta v(1,2)}\right]_{y=y^*} = \mu_2(1,2\mid y^*) + \iint d^\nu r_3 d^\nu r_4 \mu_2(1,3\mid y^*) \times \beta\tilde{v}_1(3,4)\mu_2(2,4\mid y^*) + \tfrac{1}{2}\iint d^\nu r_3 d^\nu r_4 \mu_4(1,2,3,4\mid y^*) \times \beta\tilde{v}_1(3,4) + O(\gamma^{2\nu}). \quad \text{(C3)}$$

The first two terms in this expression are obtained from the first term on the right-hand side of Eq. (8). The last term is the term of lowest order obtained by differentiation of the second term in Eq. (8). Thus, we see that in order to compute $\rho^{(2)}$ through order $\gamma^{2\nu-1}$, we must take into account terms in the pressure through order $\gamma^{3\nu-1}$. $\rho(z)$ is found, from Eqs. (2) and (8), to be

$$\rho(z) = \rho_h(y^*) + \tfrac{1}{2}\iint d^\nu r_2 d^\nu r_3 \mu_3(1,2,3\mid y^*) \times \beta\tilde{v}_1(2,3) + O(\gamma^{2\nu}). \quad \text{(C4)}$$

The integral is, of course, independent of $\mathbf{r}_1$ and is the lowest term obtained from $(\partial\bar{h}(y)/\partial \ln y)_{y=y^*}$.

Combining all these results, we obtain

$$\rho^{(2)}(1,2\mid z) = \rho_h^{(2)}(1,2\mid y^*) + \tfrac{1}{2}\iint d^\nu r_3 d^\nu r_4 \beta\tilde{v}_1(3,4) \times \{\mu_4(1,2,3,4\mid y^*) + 2\rho_h(y^*)\mu_3(1,3,4\mid y^*) + 2\mu_2(1,3\mid y^*)\mu_2(2,4\mid y^*) - \delta(1-2)\mu_3(1,3,4\mid y^*)\} + O(\gamma^{2\nu}). \quad \text{(C5)}$$

Using Eq. (B9), we see that Eq. (C5) can be written

$$\rho^{(2)}(1,2\mid z) = \rho_h^{(2)}(1,2\mid y^*) + \tfrac{1}{2}\iint d^\nu r_3 d^\nu r_4 \lambda_4(3,4\mid 1,2) \times \beta\tilde{v}_1(3,4) + O(\gamma^{2\nu}). \quad \text{(C6)}$$

This is identical through order $\gamma^{2\nu-1}$ to our Eq. (21). We have checked the agreement also to order $\gamma^{3\nu-1}$ by the same procedure. Since this calculation is straightforward and laborious, we have not presented it here.

* Program at Northwestern University supported by O.N.R. and N.S.F.

† Program at University of Illinois at Chicago Circle supported by N.S.F., Grants GP7328 and GP8734.

[1] J. B. Jalickee, A. J. F. Siegert, and D. J. Vezzetti, J. Math. Phys. **10**, 1442 (1969). This will be referred to as Paper I.

[2] M. Kác, G. E. Uhlenbeck, and P. C. Hemmer, J. Math. Phys. **4**, 216 (1963).

[3] A. J. F. Siegert and D. J. Vezzetti, J. Math. Phys. **9**, 2173 (1968).

[4] J. Zittartz, Z. Physik **180**, 219 (1964).

[5] T. Morita and K. Hiroike, Progr. Theoret. Phys. (Kyoto) **25**, 537 (1961).

[6] These functions were introduced by J. de Boer, Rept. Progr. Phys. **12**, 305 (1949), Sec. 7, III.

[7] It is assumed that, for the fugacities and temperatures of interest in the condensation of the total system, the reference system is a gas.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Complete Sets of Functions on Homogeneous Spaces with Compact Stabilizers

ROBERT LEONARD ANDERSON* AND KURT BERNARDO WOLF
Institute of Theoretical Physics, Fack, S-402 20, Göteborg 5, Sweden

(Received 12 January 1970)

We formulate and solve the problem of determining a complete set of generalized functions for a wide class of homogeneous spaces with compact stabilizers. This allows us to say precisely what unitary irreducible representations can be realized on a given homogeneous space. The techniques are applied to the n-dimensional orthogonal and unitary groups.

I. INTRODUCTION

In physics, we frequently encounter homogeneous spaces. Two well-known examples are the 3-dimensional sphere $S_3 = SO(2)\backslash SO(3)$ in angular momentum theory and Minkowski space $M = ISO(3,1)/SO(3,1)$ in Wigner's classification[1] of the unitary irreducible representations (UIR's) of the Poincaré group. With the relatively recent interest in higher-symmetry groups, homogeneous spaces have also been used, e.g., by Bég and Ruegg[2] to study $SU(3)$ and by Holland[3] to investigate some of the UIR's of $SU(n)$. Rączka *et al.* have studied the most degenerate representations of $SO(p,q)$,[4] $SU(p,q)$,[5] and $Sp(n)$.[6] Further, Lurçat,[7] Nilsson and Kihlberg[8] and others[9–11]

Carrying out the functional differentiation of $V\bar{h}(y)$, we obtain

$$\left[\frac{\delta V\bar{h}(y)}{\delta\beta v(1,2)}\right]_{y=y^*} = \mu_2(1,2\mid y^*) + \iint d^\nu r_3 d^\nu r_4 \mu_2(1,3\mid y^*) \times \beta\tilde{v}_1(3,4)\mu_2(2,4\mid y^*) + \tfrac{1}{2}\iint d^\nu r_3 d^\nu r_4 \mu_4(1,2,3,4\mid y^*) \times \beta\tilde{v}_1(3,4) + O(\gamma^{2\nu}). \quad \text{(C3)}$$

The first two terms in this expression are obtained from the first term on the right-hand side of Eq. (8). The last term is the term of lowest order obtained by differentiation of the second term in Eq. (8). Thus, we see that in order to compute $\rho^{(2)}$ through order $\gamma^{2\nu-1}$, we must take into account terms in the pressure through order $\gamma^{3\nu-1}$. $\rho(z)$ is found, from Eqs. (2) and (8), to be

$$\rho(z) = \rho_h(y^*) + \tfrac{1}{2}\iint d^\nu r_2 d^\nu r_3 \mu_3(1,2,3\mid y^*) \times \beta\tilde{v}_1(2,3) + O(\gamma^{2\nu}). \quad \text{(C4)}$$

The integral is, of course, independent of $\mathbf{r}_1$ and is the lowest term obtained from $(\partial\bar{h}(y)/\partial \ln y)_{y=y^*}$.

Combining all these results, we obtain

$$\rho^{(2)}(1,2\mid z) = \rho_h^{(2)}(1,2\mid y^*) + \tfrac{1}{2}\iint d^\nu r_3 d^\nu r_4 \beta\tilde{v}_1(3,4) \times \{\mu_4(1,2,3,4\mid y^*) + 2\rho_h(y^*)\mu_3(1,3,4\mid y^*) + 2\mu_2(1,3\mid y^*)\mu_2(2,4\mid y^*) - \delta(1-2)\mu_3(1,3,4\mid y^*)\} + O(\gamma^{2\nu}). \quad \text{(C5)}$$

Using Eq. (B9), we see that Eq. (C5) can be written

$$\rho^{(2)}(1,2\mid z) = \rho_h^{(2)}(1,2\mid y^*) + \tfrac{1}{2}\iint d^\nu r_3 d^\nu r_4 \lambda_4(3,4\mid 1,2) \times \beta\tilde{v}_1(3,4) + O(\gamma^{2\nu}). \quad \text{(C6)}$$

This is identical through order $\gamma^{2\nu-1}$ to our Eq. (21). We have checked the agreement also to order $\gamma^{3\nu-1}$ by the same procedure. Since this calculation is straightforward and laborious, we have not presented it here.

* Program at Northwestern University supported by O.N.R. and N.S.F.

† Program at University of Illinois at Chicago Circle supported by N.S.F., Grants GP7328 and GP8734.

[1] J. B. Jalickee, A. J. F. Siegert, and D. J. Vezzetti, J. Math. Phys. **10**, 1442 (1969). This will be referred to as Paper I.

[2] M. Kác, G. E. Uhlenbeck, and P. C. Hemmer, J. Math. Phys. **4**, 216 (1963).

[3] A. J. F. Siegert and D. J. Vezzetti, J. Math. Phys. **9**, 2173 (1968).

[4] J. Zittartz, Z. Physik **180**, 219 (1964).

[5] T. Morita and K. Hiroike, Progr. Theoret. Phys. (Kyoto) **25**, 537 (1961).

[6] These functions were introduced by J. de Boer, Rept. Progr. Phys. **12**, 305 (1949), Sec. 7, III.

[7] It is assumed that, for the fugacities and temperatures of interest in the condensation of the total system, the reference system is a gas.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Complete Sets of Functions on Homogeneous Spaces with Compact Stabilizers

ROBERT LEONARD ANDERSON* AND KURT BERNARDO WOLF
Institute of Theoretical Physics, Fack, S-402 20, Göteborg 5, Sweden

(Received 12 January 1970)

We formulate and solve the problem of determining a complete set of generalized functions for a wide class of homogeneous spaces with compact stabilizers. This allows us to say precisely what unitary irreducible representations can be realized on a given homogeneous space. The techniques are applied to the n-dimensional orthogonal and unitary groups.

I. INTRODUCTION

In physics, we frequently encounter homogeneous spaces. Two well-known examples are the 3-dimensional sphere $S_3 = SO(2)\backslash SO(3)$ in angular momentum theory and Minkowski space $M = ISO(3,1)/SO(3,1)$ in Wigner's classification[1] of the unitary irreducible representations (UIR's) of the Poincaré group. With the relatively recent interest in higher-symmetry groups, homogeneous spaces have also been used, e.g., by Bég and Ruegg[2] to study $SU(3)$ and by Holland[3] to investigate some of the UIR's of $SU(n)$. Rączka *et al.* have studied the most degenerate representations of $SO(p,q)$,[4] $SU(p,q)$,[5] and $Sp(n)$.[6] Further, Lurçat,[7] Nilsson and Kihlberg[8] and others[9–11]

have used some of the homogeneous spaces of the Poincaré group in order to build field theories which may provide a description in which mass and spin are treated on an equal footing.

These examples illustrate an important and well-known use of homogeneous spaces in mathematics and physics: namely, that, if there exists a right G-invariant measure on a homogeneous space $X = G_0\backslash G$ (G_0 is called the right stabilizer of the homogeneous space), then, by using the action of G on X, one can realize a unitary (in general, reducible) representation U_g of G, the Hilbert space $L_2(X)$ of square-integrable functions on X serving as a carrier space.

The central mathematical question is the decomposition (as a direct sum or integral) of $L_2(X)$ into minimal U_g-invariant subspaces, i.e., the decomposition of the regular representation U_g into irreducible representations. The elements of these minimal U_g-invariant spaces are called spherical functions.

We know, from the work of Rączka,[12] the connection between the completeness of the set of matrix elements of UIR's for the space $L_2(G)$ for locally compact, semisimple Lie groups and nuclear spectral theory: namely, that the UIR matrix elements are the generalized eigenfunctions of a complete set of operators built from the right and left universal enveloping Lie algebras. In fact, we obtain not only the spherical functions but their decomposition into 1-dimensional subspaces and, furthermore, the complete classification of all possible UIR's that appear in the regular representation of G along with their multiplicities.

Our purpose is to exploit this fact and obtain analogous results for the homogeneous spaces of the form $G_0\backslash G$, where G is an arbitrary locally compact unimodular Lie group with a complete set of UIR matrix elements. In this paper we shall discuss in detail the case when $G_0 = K$ is any (closed) compact subgroup of G.

The formalism is set up in Sec. II, where the central question "What constitutes a complete set of functions on the homogeneous space $X = K\backslash G$?" is answered. This leads immediately to the solution of the related problem in Sec. III: "What UIR's can be realized on a given homogeneous space?"

The solution is given explicitly for the orthogonal groups in Sec. IV: $SO(n)$ is treated in detail, and the extensions to $SO(n-1, 1)$ and $ISO(n)$ are indicated. In Sec. V we treat the unitary groups $U(n)$ and $SU(n)$. The parametrization of these groups and of the homogeneous space manifolds is essential for our purposes: The concrete and detailed statement of the results is simplified by the use of the Gel'fand–Tsetlin patterns.

We hope to discuss, in a future publication, homogeneous spaces where the stabilizer groups are noncompact (but locally compact).

II. COMPLETE SETS OF GENERALIZED FUNCTIONS

Let G be a locally compact Lie group and K any closed subgroup of G. We can construct the space of right (left) cosets, the homogeneous space $X = K\backslash G$ (G/K) whose points $x \in X$ are the sets $Kg(gK)$, where $g \in G$ and the topology in X is the one induced by the topology of G. The homogeneous space X is itself a group only when K is a normal subgroup of G. However, we are interested in X as a transitive manifold for G; i.e., (a) the coset Kg is mapped into the coset Kgg' under right multiplication by $g' \in G$ (gK is mapped into $g'gK$ under left multiplication by $g' \in G$), and (b) given any two cosets, there exists a $g \in G$ which maps one into the other. For the purposes of economy, we shall confine our discussion to spaces of right cosets $X = K\backslash G$.

Let $F(G)$ be the set of functions defined on the group manifold of G. Functions on the homogeneous space $X = K\backslash G$ are defined as functions which are constant on right cosets, i.e.,

$$F(X) = \{f \in F(G) \,|\, f(kg) = f(g), \quad \text{for} \quad g \in G \text{ and all } k \in K\}. \quad (1)$$

We want to point out the fact that, if $f \in L_2(X)$ and if the stabilizer K is compact, then $f \in L_2(G)$, while, if K is noncompact, $f \notin L_2(G)$. It is precisely because of this that we restrict ourselves in this paper to the case with compact stabilizers.

Since we are interested in unitary representations, we must have a right G-invariant measure on X, and we can use the following general theorem[13] to insure its existence:

Theorem: If G is a locally compact unimodular Lie group and K a subgroup of G, then there exists an invariant measure $dm(x)$ on the homogeneous space $K\backslash G$ provided that

$$|\det Ad_G(k)| = |\det Ad_K(k)| \quad \text{for all} \quad k \in K.$$

This measure $dm(x)$ is unique up to a constant factor, and

$$\int_G f(g)\, d\mu(g) = \int_{K\backslash G} dm(x) \int_K f(kg)\, d\mu(k)$$

for every $f \in C_0(G)$.

We know that $C_0(G)$, the space of all continuous functions on G with compact support, is dense in $L_2(G)$. Since all continuous functions with compact support which are constants on cosets with respect to a compact subgroup K belong to $C_0(G)$ and are dense in $L_2(X)$, the theorem is applicable to our case and covers a "sufficiently" large class of elements of $L_2(X)$.

We now restrict ourselves further to locally compact unimodular Lie groups, such that the UIR matrix elements or some subset of them constitute a complete set of generalized eigenfunctions as described by Rączka, i.e., such that they are the generalized eigenfunctions (in the sense of nuclear spectral theory) of a complete set of strongly commuting operators built from the right and left universal enveloping Lie algebras. We shall henceforth call this "the completeness requirement."

Presently, there is no general, mathematically rigorous statement characterizing all groups which satisfy the completeness requirement. The compact Lie groups are, of course, rigorously known to satisfy it by virtue of the Peter–Weyl–von Neumann theory.[14] The theory has also been developed for some of the noncompact groups, specifically $SL(2, C)$,[15] $IO(2)$,[16] and $SO(2, 1)$.[17] The work of Rączka[12] represents a generalization of the Peter–Weyl–von Neumann theorem to noncompact, semisimple Lie groups.

For groups which satisfy this completeness requirement, we can associate, with each function $f(g) \in L_2(G)$, one matrix function $F(\lambda)$ whose domain is the space of a complete set of UIR's of the group, which we denote by $\hat{G}$. The points of this space are characterized by the eigenvalues λ of a complete set of Casimir operators[18] of the group. For every λ, the rows and columns of this matrix are labeled in the same fashion as the UIR matrices $D^{\lambda}(g)$ themselves (as will be detailed below), which become the transformation kernels which relate the two functions[11,19]; i.e.,

$$F_{pq}(\lambda) = \int_G d\mu(g) f(g) D^{\lambda}_{pq}(g), \tag{2a}$$

$$f(g) = \int_{\hat{G}} d\hat{\mu}(\lambda) \sum_{p,q} F_{qp}(\lambda) D^{\lambda}_{pq}(g^{-1}), \tag{2b}$$

where $d\hat{\mu}(\lambda)$ is the Plancherel measure on $\hat{G}$. For compact groups G, the space $\hat{G}$ is a set of isolated points, and the integration in (2b) becomes a sum

$$\int_{\hat{G}} d\hat{\mu}(\lambda) \to \sum_{\lambda \in \hat{G}} \dim(\lambda) \Big/ V(G),$$

where $\dim(\lambda)$ is the dimension of the UIR labeled by λ and $V(G)$ is the volume of the group.

The norms of the two functions in (2a) and (2b) can be related by the Parseval identity

$$\int_G d\mu(g) |f(g)|^2 = \int_{\hat{G}} d\hat{\mu}(\lambda) \sum_{p,q} |F_{pq}(\lambda)|^2.$$

Let $G \supset G_r \supset G_{r-1} \supset \cdots \supset G_1$ be a chain of subgroups which includes the compact subgroup $K = G_k$ for some k, whose UIR labels can be used to classify completely the components of the UIR basis vectors [as, e.g., the canonical chains $SO(n) \supset SO(n-1) \supset \cdots \supset SO(2)$ and $U(n) \supset U(n-1) \supset \cdots \supset U(1)$ for the orthogonal and unitary groups, giving rise to the Gel'fand–Tsetlin kets; see Secs. IV and V].

It will prove convenient to regard the index labeling the rows (columns) of the representation matrices $D_{pq}(g)$, as standing for the sets $\{p_r, p_{r-1}, \cdots, p_1\}$ ($\{q_r, q_{r-1}, \cdots, q_1\}$), where p_j (q_j) is the collective label which denotes the eigenvalues of a complete set[18] of Casimir operators of the subgroup G_j along the chain. For convenience, we define $\overline{p_j} \equiv \{p_j, p_{j-1}, \cdots, p_1\}$ ($\overline{q_j} \equiv \{q_j, q_{j-1}, \cdots, q_1\}$), i.e., the row (column)-index for the UIR matrices of G_{j+1}, and

$$D^{\lambda}_{pq}(k_0) = D^{p_k}_{\bar{p}_{k-1}\bar{q}_{k-1}}(k_0) \prod_{j=k}^{r} \delta_{p_j,q_j} \tag{3}$$

for $k_0 \in K$, where the Kronecker δ in the collective labels p_j and q_j is to be regarded as a product of the Kronecker δ's in the individual indices.

Let us first examine the pair of functions (2) when $f(g)$ has the property (1) (i.e., when it is a function on $X = K\backslash G$), in order to determine the subset of $\{D^{\lambda}_{pq}(g^{-1})\}$ which constitutes a complete set in X: We find that only those D's which satisfy

$$D^{\lambda}_{pq}(g^{-1}) = D^{\lambda}_{pq}((kg)^{-1}) \quad \text{for all} \quad k \in K \quad \text{and} \quad g \in G \tag{4}$$

appear in the expansion of a $f \in L_2(X)$. We have employed the following reasoning here: $\{D^{\lambda}_{pq}(g^{-1})\}$ is a complete set for $L_2(G)$ and, since $f \in L_2(X)$ implies $f \in L_2(G)$, $\{D_{pq}(g^{-1})\}$ is also a complete set for $L_2(X)$; $f \in L_2(X)$ means $f(kg) = f(g)$ for all $k \in K$, and thus we find that the subset of $\{D_{pq}(g^{-1})\}$ which satisfies condition (4) is a complete set on X.

We can integrate (4) over $k \in K$ in both sides: The left-hand side, independent of k, will be multiplied by the volume $V(K)$ of the subgroup. The right-hand side can be written as $\sum_s D^{\lambda}_{pq}(g^{-1}) D^{\lambda}_{sq}(k^{-1})$, and the integration performed only over the last factor, where (3) can be used.

Furthermore, since the scalar representation

$$D^{0}_{00}(k) = 1$$

(which is a constant function on K) is a member of the complete set of UIR matrices for K, we have

$$\int_K d\mu(k) D^{q_k}_{\bar{s}_{k-1}\bar{q}_{k-1}}(k^{-1}) = \delta_{\bar{s}_{k-1},0}\delta_{\bar{q}_{k-1},0}\delta_{\hat{K}}(q_k, 0), \quad (5)$$

where the last factor is such that

$$\int_{\hat{K}} d\hat{\mu}(q) A(q) \delta_{\hat{K}}(q, q_0) = A(q_0)$$

for any "well-behaved" function $A(q)$ on the space $\hat{K}$. For compact groups, it can be written as

$$\delta_{\hat{K}}(q, q_0) = \delta_{q,q_0} V(K)/\dim(q_0). \quad (6)$$

As before, dim (q) is the dimension of the UIR of K labeled by q, and dim $(0) = 1$. Hence, from (3), (5), and (6) it can be seen that those D's which satisfy (4) must also satisfy

$$D^{\lambda}_{pq}(g^{-1}) = D^{\lambda}_{pq}(g^{-1})\delta_{\bar{q}_k,0} \quad (7)$$

and, hence, the complete set of functions on $X = K\backslash G$ is the set $\{D^{\lambda}_{p,q(k)}(g^{-1})\}$, where

$$q(k) \equiv \{q_r, q_{r-1}, \cdots, q_{k+1}, 0, \cdots, 0\}.$$

It is important to understand that $q(k)$ restricts the allowed values of λ to those UIR's which contain the scalar representation of K. This, in turn, restricts the allowed values of p. Concrete cases will be presented in Secs. IV and V.

There is another point of interest: namely, that we know that, since G is a locally compact unimodular Lie group, the right and left regular representations are simultaneously defined on the group and that they commute, since there exists a left and right G-invariant measure on the group manifold. The question arises as to what happens, e.g., with the action of the group (from the left) T^{L}_g for the right quasiregular representation on X. The answer is seen by acting with T^{L}_g on (2b) where the D's are restricted by (7). One finds that under T^{L}_g all elements of a given right coset are mapped into another right coset with respect to the subgroup $g^{-1}Kg$, so that functions constant on the right coset space $K\backslash G$ are mapped into functions constant on the right coset space $g^{-1}Kg\backslash G$.

III. DECOMPOSITION OF THE QUASIREGULAR REPRESENTATION

An important consequence of our knowledge of a complete set of generalized functions on X, as a subset of the UIR matrix elements of the group G, is that we automatically obtain a decomposition of the unitary quasiregular right representation into its UIR's, along with their multiplicities. Recall that the quasiregular right representation T^{R} of a locally compact Lie group G on $L_2(X)$ is defined as

$$T^{\mathrm{R}}_g f(x) = f(xg), \quad (8)$$

where $f \in L_2(X)$, $x \in X$, and $g \in G$. It is well known that, if there exists a right G-invariant measure on X, then T^{R}_g is a unitary representation.

In fact, any $f \in L_2(X)$ can be decomposed as

$$f(x) = \int_{\hat{G}\,(\text{restricted})} d\hat{\mu}(\lambda) \sum_{p,q(k)} F_{q(k),p}(\lambda) D^{\lambda}_{p,q(k)}(g^{-1}), \quad (9)$$

where $q(k)$ was defined above. But T^{R}_g acting on (9) transforms all $D^{\lambda}_{p,q(k)}(g^{-1})$ with a fixed value of $q(k)$ among themselves. Hence, for each fixed value of $q(k)$ there exists one UIR in the direct-integral decomposition of $L_2(X)$, and thus the multiplicity is exactly the number of different values of $q(k)$ constrained by a fixed (allowable) λ and $\bar{q}_k = 0$. This number may, or may not, be denumerable.

IV. APPLICATION: THE ORTHOGONAL GROUP

We shall give first a brief description of the group and representation spaces of the orthogonal groups $SO(n)$. The group manifold of $SO(n)$ can be parametrized inductively by "Euler" angles (enclosing collective variables in curly brackets) as

$$\begin{aligned} R_n(\{\vartheta\}^{(n)}) &= R_{n-1}(\{\vartheta\}^{(n-1)}) S_n(\{\vartheta^{(n)}\}), \\ S_n(\{\vartheta^{(n)}\}) &= r_{n-1,n}(\vartheta^{(n)}_{n-1,n}) \\ &\quad \times \cdots \times r_{23}(\vartheta^{(n)}_{23}) \times r_{12}(\vartheta^{(n)}_{12}), \end{aligned} \quad (10)$$

where $R_k \in SO(k)$ and $r_{ab}(\vartheta)$ is a rotation by ϑ in the (a, b) plane. The ranges of the variables are $0 \leq \vartheta_{12} < 2\pi$ and $0 \leq \vartheta_{k-1,k} \leq \pi$, $k = 2, 4, \cdots, n$. Thus, the $SO(n)$ manifold is the product of the $SO(n-1)$ manifold and the n-dimensional unit sphere: the homogeneous space $SO(n-1)\backslash SO(n)$, parametrized by the $n-1$ angles $\{\vartheta^{(n)}\}$.

Notice that, for $SO(3)$, $R_3(\alpha, \beta, \gamma) = r_{12}(\alpha) r_{23}(\beta) \times r_{12}(\gamma)$. This differs from the more general usage[20] in that the second rotation is made around the 1-axis rather than the 2-axis. This will cause no difficulty, however.

The Haar measure on $SO(n)$ can be split according to the parametrization (10) as $d\mu(R_n) = d\mu(R_{n-1})\, dS_n$, where

$$dS_n = \sin^{n-2}\vartheta_{n-1,n}\, d\vartheta_{n-1,n} \cdots \sin\vartheta_{23}\, d\vartheta_{23}\, d\vartheta_{12} \quad (11)$$

is the measure on the space $SO(n-1)\backslash SO(n)$. The volume of the group can be seen to be given by $V(SO(n)) = V(SO(n-1))A_n$ [and $V(SO(2)) = 2\pi$], where $A_n = 2\pi^{\frac{1}{2}n}/\Gamma(\frac{1}{2}n)$ is the surface of the n-dimensional sphere.

The homogeneous space $SO(k)\backslash SO(n)$ is thus the set of points $S_{k+1}(\{\vartheta^{(k+1)}\})\cdots S_n(\{\vartheta^{(n)}\})$ and is a $[\frac{1}{2}n(n-1)-\frac{1}{2}k(k-1)]$-dimensional manifold with Haar measure $dS_{k+1}\cdots dS_n$.

From the work of Gel'fand and Tsetlin[21,22] we know that the bases for UIR's of $SO(n)$ classified by the canonical chain $SO(n-1)\supset\cdots\supset SO(2)$ can be labeled as

$$\left|\begin{array}{cccc} J_{n,1} & J_{n,2} & \cdots & J_{n,[\frac{1}{2}n]} \\ J_{n-1,1} & J_{n-1,2} & \cdots & J_{n-1,\frac{1}{2}[(n-1)]} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ J_{4,1} & J_{4,2} & & \\ J_{3,1} & & & \\ J_{2,1} & & & \end{array}\right\rangle, \tag{12}$$

where $[\frac{1}{2}k]$ is the largest integer less than or equal to $\frac{1}{2}k$. This ket transforms as the $J_k=(J_{k1},J_{k2},\cdots J_{k[\frac{1}{2}k]})$ UIR of $SO(k)$. The J_{ab} are either all integer or all half-integer and are constrained by the inequalities

$$\begin{aligned} J_{2k+1,j}&\geq J_{2k,j}\geq J_{2k+1,j+1}, \\ &\qquad j=1,\cdots,k^{-1}, \\ &\qquad k=1,\cdots,[\tfrac{1}{2}(n-1)], \\ J_{2k,j}&\geq J_{2k-1,j}\geq J_{2k,j+1}, \\ &\qquad j=1,\cdots,k-1, \\ &\qquad k=1,\cdots,[\tfrac{1}{2}n], \\ J_{2k+1,k}&\geq |J_{2k,k}|, \quad k=1,\cdots,[\tfrac{1}{2}(n-1)], \\ J_{n,[\frac{1}{2}n]}&\geq 0, \quad n \text{ odd}, \\ J_{n-1,[\frac{1}{2}n]-1}&\geq |J_{n,[\frac{1}{2}n]}|, \quad n \text{ even}. \end{aligned} \tag{13}$$

The number of labels in the ket (12) is

$$L^0(n)=\sum_{m=2}^{n}[\tfrac{1}{2}m]=\tfrac{1}{4}(n^2-1), \quad n \text{ odd},$$
$$=\tfrac{1}{4}n^2, \qquad n \text{ even},$$

and the number of UIR, row, and column labels of the $D^{J_n}_{\bar{J}_{n-1}\bar{J}'_{n-1}}(R_n^{-1})$ (by using the notation of Sec. 2, namely, $\bar{J}_k\equiv\{J_k,J_{k-1},\cdots,J_2\}$) is thus $[\frac{1}{2}n]+2L^0(n-1)=\frac{1}{2}n(n-1)$, i.e., the same as the number of parameters of the group.

The scalar representation of $SO(k)$ is $J_k=(0,0,\cdots,0)\equiv 0$. Notice, then, that, if we have zeros in the J_k row, $k>2$, of (12), the inequalities (13) imply that the J_{k+1} row must consist of zeros, except for $J_{k+1,1}$, which is only constrained to be integer. In the J_{k+2} row, all except $J_{k+2,1}$ and $J_{k+2,2}$ must be zero, etc. Thus we can see that only the "most symmetric" $(J,0,\cdots,0)$ UIR'S of $SO(n)$, $n>3$, can be realized on the homogeneous space $SO(n-1)\backslash SO(n)$, a result familiar from the theory of spherical functions.[4]

For $SO(2)\backslash SO(3)$, $J_{21}=0$ implies only that J_{31} must be integer and thus the functions on the 3-dimensional sphere[23] can realize all—but only—the single-valued representations of the group $SO(3)$.

For $SO(k)\backslash SO(n)$, the ket (12) will have $\bar{J}_k=0$. These zeros will "propagate" upwards under a diagonal to the right (see Fig. 1), and there will be $[\frac{1}{2}(2k-n)]$ zeros in the row of the UIR labels. Thus, if $k\leq\frac{1}{2}(n+1)$ (n odd) or $k\leq\frac{1}{2}n$ (n even), it will be possible to realize all the single-valued UIR's of $SO(n)$ in this space.

It is not difficult to see that the number of indices $J_{q,r}$, $q\geq k+1$, which are forced to be zero is $L^0(k-1)$, so that there are $[\frac{1}{2}k]+2L^0(k-1)=\frac{1}{2}k(k-1)$ zeros in the pattern (12), and the number of remaining free labels is $\frac{1}{2}n(n-1)-\frac{1}{2}k(k-1)$, equal to the number of parameters of the space $SO(k)\backslash SO(n)$.

On the other hand, if some of the UIR labels are forced to be zero, these zeros may propagate into the row (p) indices as well, downwards along a vertical line (see Fig. 1). There will thus be $2L^0(k-1)+[\frac{1}{2}k]-L^0(2k-n)$ zeros in the column (q) pattern (12), $[\frac{1}{2}(2k-n)]$ in the UIR labels, and $L^0(2k-n-1)$ in the row (p) pattern (12). The number of remaining free labels is again equal to the number of parameters of the homogeneous space.

As far as the $SO(n-1,1)$ group is concerned, the same procedure can be applied to the chain of subgroups $SO(n-1,1)\supset SO(n-1)\supset\cdots\supset SO(2)$. The only differences lie in (10), where $r_{n-1,n}(\xi)$ is now a boost in the $(n-1)$ direction so that $0\leq\zeta<\infty$.

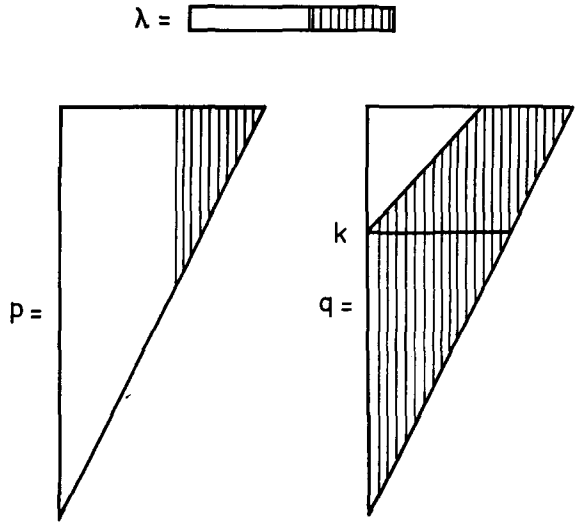


FIG. 1. Graphical representations of the zeros (shaded areas) in the UIR, row, and column labels of $D^{\lambda}_{\mathfrak{p},\mathfrak{q}(k)}(R^{-1})$ for the orthogonal group.

The homogeneous space $SO(n-1)\backslash SO(n-1,1)$ is an n-dimensional revolution hyperboloid with a measure (11), where the trigonometric function in $\vartheta_{n-1,1}$ must be replaced by a hyperbolic one in ζ. As shown by Chakrabarti,[24] the ket (12) replaces its discrete $J_{n,1}$ index by a continuous one, which is not subject to the restrictions (13). Our knowledge of these groups[25] is still unsatisfactory regarding the UIR matrix elements themselves[26]; however, the statements made regarding the complete set of functions on the homogeneous spaces considered above do not depend on their detailed knowledge.

The inhomogeneous orthogonal group $ISO(n)$ is the semidirect product of the n-dimensional translation group $T(n)$ and $SO(n)$. Its elements are the points $g=(x,R)$, $x\in T(n)$, $R\in SO(n)$, with the product $(x_2,R_2)(x_1,R_1)=(x_2+R_2x_1,R_2R_1)$.

The $ISO(n)$ manifold is thus the direct product of the $T(n)$ and $SO(n)$ manifolds, its Haar measure being $d^nx\,d\mu(R)$. Kets similar to (12) which classify the components of the UIR vectors using the chain $ISO(n)\supset SO(n)\supset\cdots\supset SO(2)$ have been set up by Chakrabarti.[24] The more common (and physically relevant) classification of the UIR's of $ISO(n)$ [and of $ISO(n-1,1)$] is the one which follows Wigner's "little group" method.[1] Harmonic analysis on these groups, with their UIR's classified by the mass-M spin-J pair of labels (J is a collective index for $n\geq 5$), has been carried out,[10,11] and the case of functions on the space of cosets of the type $ISO(n)/SO(k)$ developed in Ref. 11. This does not fall, however, within the bounds of our formalism, which so far requires the use of the "canonical" chain. This subject, then, requires further investigation.

V. APPLICATION: THE UNITARY GROUP

We shall present one parametrization of the $U(n)$ group manifold which we consider convenient because of its inductive definition, which makes it similar to the "Euler" angle parametrization of the orthogonal groups seen in the previous section.[26] It can be conveniently used to parametrize the homogeneous space $SU(k)\backslash SU(n)$.

Enclosing collective labels in curly brackets, we define

$$U_n(\{\varphi,\vartheta\}^{(n)})=U_{n-1}(\{\varphi,\vartheta\}^{(n-1)})C_n(\{\varphi^{(n)},\vartheta^{(n)}\}),$$
$$C_n(\{\varphi^{(n)},\vartheta^{(n)}\})=\Phi_n(\varphi_n^{(n)})r_{n-1,n}(\vartheta_n^{(n)})\times\cdots\times\Phi_2(\varphi_2^{(n)})r_{1,2}(\vartheta_2^{(n)})\times\Phi_1(\varphi_1^{(n)}),\quad(14)$$

where $U_k\in U(k)$; $\Phi_k(\varphi)$ is a diagonal matrix with elements $e^{[i\varphi/(1-k)]}$ in the (q,q) positions, $q=1,2,\cdots,k-1$, $e^{(i\varphi)}$ in the (k,k) position, and 1 in the remaining places on the diagonal. It is unimodular for $k\geq 2$.

As in the previous section, $r_{ab}(\vartheta)$ is a matrix with $\cos\vartheta$ in the (a,a) and (b,b) positions, 1 elsewhere on the diagonal, $-\sin\vartheta$ and $\sin\vartheta$ in the (a,b) and (b,a) positions, and 0 in the remaining places.

The number of parameters of C_n is $2n-1$, so that the parameters of $U(n)$ are n^2. The condition of unimodularity implies

$$\sum_{k=1}^{n}\varphi_1^{(k)}=0,$$

since

$$\det[C_k(\{\varphi^{(k)},\vartheta^{(k)}\})]=\exp(i\varphi_1^{(k)}),$$

and this restricts by one the number of parameters. The ranges of the "rotation" angles are $0\leq\vartheta_j^{(k)}\leq\frac{1}{2}\pi$, $j=2,\cdots,k$, and of the "phases" $0\leq\varphi_j^{(k)}<2\pi$, $j=1,\cdots,k$. The Haar measure on $U(n)$ decomposes as for the orthogonal group: $d\mu(U_n)=d\mu(U_{n-1})\,dC_n$.

We can see that $C_n^{-1}(\{\varphi^{(n)},\vartheta^{(n)}\})$ acting on the n-dimensional complex vector $(0,\cdots,0,1)$ generates $(z_1,\cdots,z_n)$, the surface of the n-dimensional complex unit sphere. The surface element of this can be found by using the fact that, for each coordinate,

$$d^2z=d\,\mathrm{Re}\,(z)\,d\,\mathrm{Im}\,(z)=r\,dr\,d\psi,$$

where $r=\mathrm{mod}\,(z)$ and $\psi=\arg(z)$. The modulus is thus only a function of the ϑ variables, the argument only of the φ variables, and

$$dC_n=(r_1\cdots r_n)(dr_1\cdots dr_n)(d\psi_1\cdots d\psi_n).$$

The first term is $\sin^{n-1}\vartheta_n\cos\vartheta_n\cdots\sin\vartheta_2\cos\vartheta_2$, the second one is (11), the measure on the n-dimensional real sphere, and the third one is just $d\varphi_1\cdots d\varphi_n$ since the Jacobian is unity.

The measure on the n-dimensional complex sphere is therefore

$$\begin{aligned}dC_n&=\sin^{2n-3}\vartheta_n^{(n)}\cos\vartheta_n^{(n)}\,d\vartheta_n^{(n)}\,d\varphi_n^{(n)}\\&\quad\times\cdots\times\sin\vartheta_2^{(n)}\cos\vartheta_2^{(n)}\,d\vartheta_2^{(n)}\,d\varphi_2^{(n)}\times d\varphi_1^{(n)}\\&=d\varphi_n^{(n)}s_n^{2n-3}\,ds_n\times\cdots\times d\varphi_2^{(n)}s_2\,ds_2\times d\varphi_1^{(n)},\end{aligned}\quad(15)$$

where $s_k=\sin_k^{(n)}$; thus, this is the measure on the homogeneous space $SU(n-1)\backslash SU(n)$. Correspondingly, the space $SU(k)\backslash SU(n)$ is parametrized by

$$C_{k+1}(\{\varphi^{(k+1)},\vartheta^{(k+1)}\})\cdots C_n(\{\varphi^{(n)},\vartheta^{(n)}\}),$$

has n^2-k^2 parameters, and its measure is $dC_{k+1}\cdots dC_n$. The volume of the group is $V(U(n))=V(U(n-1))B_n$ [and $V(U(1))=2\pi$], where $B_n=2\pi^n/\Gamma(n)=A_{2n}$, the surface of the n-dimensional complex sphere.

Using again the work of Gel'fand and Tsetlin,[27] we can write the bases for UIR's of $U(n)$, classified by the canonical chain[15] $U(n) \supset U(n-1) \supset \cdots \supset U(1)$, as

$$\left|\begin{array}{ccccccc} K_{n,1} & & K_{n,2} & & K_{n,3} & \cdots & K_{n,n} \\ & K_{n-1,1} & & K_{n-1,2} & \cdots & K_{n-1,n-1} & \\ & & \cdot & & & \cdot & \\ & & & \cdot & \cdots & \cdot & \\ & & & & \cdot & & \\ & & & K_{2,1} & & K_{2,2} & \\ & & & & K_{1,1} & & \end{array}\right\rangle . \quad (16)$$

This ket transforms as the $K_k = (K_{k1}, K_{k2}, \cdots, K_{kk})$ UIR of $U(k)$, and these labels are constrained by the inequalities

$$K_{k,j} \geq K_{k-1,j} \geq K_{k,j+1}, \quad j = 1, \cdots, k-1, \quad k = 2, \cdots, n. \quad (17)$$

The number of labels in (16) is

$$L^u(n) = \sum_{m=1}^{n} m = \tfrac{1}{2}n(n+1),$$

and the number of UIR, row, and column labels of $D^{K_n}_{\bar{K}_{n-1}\bar{K}'_{n-1}}(U)$ (again, $\bar{K}_k \equiv \{K_k, K_{k-1}, \cdots, K_1\}$) is thus $n + 2L^u(n-1) = n^2$, i.e., the same as the number of parameters of the group $U(n)$.

The representations of $SU(n)$ have the same labeling as those of $U(n)$, except that

$$(K_{n,1}, K_{n,2}, \cdots, K_{n,n}) \equiv (K_{n,1} + K, K_{n,2} + K, \cdots, K_{n,n} + K)$$

for any K, so that we can take $K = -K_{n,n}$ and thus restrict the last UIR label to zero, having thus $n^2 - 1$ UIR, row, and column labels for the (n^2-1)-parameter group $SU(n)$.

The scalar representation of $SU(k)$ is $K_k = (0, \cdots, 0)$, but this is equivalent to $(K, \cdots, K)$ in the pattern (16), forcing through the inequalities (17), all the rows below the K_k row to be $(K, \cdots, K)$ as well (see Fig. 2). This takes the place of the UIR 'O' in the case of the orthogonal groups (Sec. IV).

If the scalar representation of $SU(k)$ is to be present in (16), the inequalities (17) imply that all but two of the labels of the K_{k+1} row must be equal: $K_{k+1,2} = \cdots = K_{k+1,k} = K_{k,1}$ and, in the row above that one, $K_{k+2,3} = \cdots = K_{k+2,k} = K_{k+1,2} = K_{k,1}$, etc. Thus, we can see that only the $(J, K, \cdots, K, 0)$ UIR's of $SU(n)$ can be realized on the homogeneous space $SU(n-1)\backslash SU(n)$. In particular, this places no restriction on the $SU(3)$ UIR's which can be realized on $SU(2)\backslash SU(3)$ homogeneous space,[2,3] as can be seen also noting that every UIR of $SU(3)$ contains the scalar representation[28] of $SU(2)$.

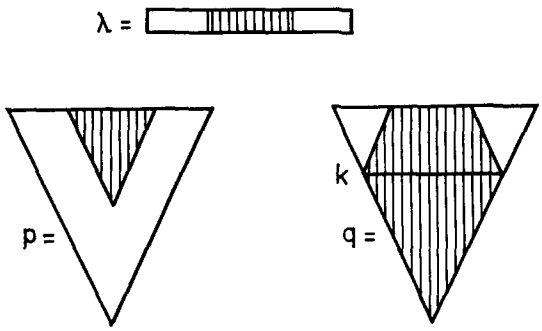


FIG. 2. Graphical representation of the regions of constant values (shaded areas) in the UIR, row, and column labels of $D^{\lambda}_{p,q(k)}(U^{-1})$ for the unitary group.

In general, all UIR's of $SU(n)$ can be realized on $SU(k)\backslash SU(n)$ when $k \leq \frac{1}{2}(n+1)$. The equality of the $2L^u(k-1) + k - 1 = k^2 - 1$ UIR labels of a scalar representation of $SU(k)$ and the triangle above it (Fig. 2) leaves $n^2 - k^2$ free parameters, i.e., the number of parameters of the space $SU(k)\backslash SU(n)$.

If $k > \frac{1}{2}(n+1)$, only the UIR's of $SU(n)$ with $K_{n,n-k+1} = K_{n,n-k+2} = \cdots = K_{n,k}$ are to be realized on $SU(k)\backslash SU(n)$. The number of free parameters can be seen (Fig. 2) to be again equal to $n^2 - k^2$.

There seems to be no fundamental difficulty in carrying out this program for the $SU(n-1, 1)$ groups[29] classified by the canonical chain $SU(n-1, 1) \supset U(n-1) \supset \cdots \supset U(1)$ nor for the inhomogeneous unitary group $ISU(n)$, whose elements are $g = (x, U)$, where $x \in T(2n)$ and $U \in SU(n)$. The kets in the Wigner "little group" chain can be constructed using the $T(2n)$ subgroup, and its representations, labeled by an n-dimensional complex vector. Again, our formalism requires that we follow a procedure parallel to Chakrabarti's,[24] in considering the "canonical" chain $ISU(n) \supset U(n) \supset U(n-1) \supset \cdots \supset U(1)$.

ACKNOWLEDGMENTS

We would like to thank Professor N. Svartholm for his hospitality at the Institute of Theoretical Physics and to thank the Carl-Bertel Nathhorsts Foundation for financial support. It is a pleasure to acknowledge many illuminating conversations with Dr. A. Kihlberg and Dr. U. Ottoson and a valuable discussion of these results with Dozent R. Raczka.

* Present address: Department of Physics, COP, University of the Pacific, Stockton, California.

[1] E. P. Wigner, Ann. Math. **40**, 39 (1939).
[2] M. A. B. Bég and H. Ruegg, J. Math. Phys. **6**, 677 (1965); T. J. Nelson, *ibid.* **8**, 857 (1967).
[3] D. F. Holland, J. Math. Phys. **10**, 531, 1903 (1969).

[4] J. Fischer, J. Niederle, and R. Raczka, J. Math. Phys. **7**, 816 (1966); R. Raczka, N. Limić, and J. Niederle, *ibid.*, 1861 (1966); B. Limić, J. Niederle, and R. Raczka, *ibid.* **7**, 2026 (1966); **8**, 1079 (1967).
[5] J. Fischer and R. Raczka, Commun. Math. Phys. **3**, 233 (1966); **4**, 8 (1967).
[6] P. Pajas and R. Raczka, J. Math. Phys. **9**, 1188 (1968); P. Pajas, *ibid.* **10**, 1777 (1969).
[7] F. Lurçat, Physics **1**, 95 (1964).
[8] J. Nilsson and A. Beskow, Arkiv Fysik **34**, 307 (1967); J. Nilsson and A. Kihlberg, Report 68-15, Institute of Theoretical Physics, Göteborg, 1968; H. Bacry and A. Kihlberg, J. Math. Phys. **10**, 2132 (1969).
[9] G. Fuchs, Ph.D. thesis, Faculté des Sciences d'Orsay, University of Paris, 1969.
[10] N. X. Hai, Commun. Math. Phys. **12**, 331 (1969).
[11] K. B. Wolf, Ph.D. thesis, Dept. of Physics and Astronomy, Tel-Aviv University, 1969; Tel-Aviv University Preprint 112-69, 1969.
[12] R. Raczka, "Operator Distributions in Group Representation Theory and Their Applications," Lecture Notes, Institute of Theoretical Physics, Göteborg, 1969.
[13] K. Maurin, *General Eigenfunction Expansions and Unitary Representations of Topological Groups* (Polish Scientific Publishers, Warsaw, 1968), p. 143.
[14] Reference 13, p. 157.
[15] M. A. Naimark, *Linear Representations of the Lorentz Group* (Pergamon, London, 1964). S. Ström, Arkiv Fysik **29**, 467 (1965); **33**, 465 (1967); **34**, 215 (1967); A. Sciarrino and M. Toller, J. Math. Phys. **8**, 1252 (1967).
[16] W. T. Sharp, Atomic Energy of Canada Ltd., Chalk River, Ontario, Report CRT-935, 1960.
[17] V. Bargmann, Ann. Math. **48**, 568 (1947); M. Andrews and J. Gunson, J. Math. Phys. **5**, 1391 (1964).
[18] L. C. Biedenharn, J. Math. Phys. **4**, 436 (1963); G. E. Baird and L. C. Biedenharn, *ibid.*, 1449 (1963); A. M. Perelemov and V. S. Popov, Yad. Fiz. **3**, 924 (1966) [Sov. J. Nucl. Phys. **3**, 676 (1966)].
[19] G. Mackey, Bull. Am. Math. Soc. **56**, 385 (1950).
[20] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton U.P., Princeton, N.J., 1957), p. 6.
[21] I. M. Gel'fand and M. L. Tsetlin, Dokl. Akad. Nauk SSSR **71**, 840 (1950).
[22] S. C. Pang and K. T. Hecht, J. Math. Phys. **8**, 1233 (1967); M. K. F. Wong, *ibid.*, 1899 (1967).
[23] Reference 20, Chap. 4.
[24] A. Chakrabarti, J. Math. Phys. **9**, 2087 (1968).
[25] U. Ottoson, Commun. Math. Phys. **8**, 228 (1968).
[26] Our parametrization is the one used in Refs. 2 and 3 and in the work currently done on the representation matrices by Ottoson. Other parametrizations are used in F. D. Murnaghan [*The Unitary and Rotation Groups* (Spartan, New York, 1962)] and L. K. Hua [*Harmonic Analysis of Functions of Several Complex Variables in the Classical Domains* (American Mathematical Society, Providence, R.I., 1963), Translations of Mathematical Monographs, Vol. 6.]
[27] I. M. Gel'fand and M. L. Tsetlin, Dokl. Akad. Nauk SSSR **71**, 825 (1950), and Ref. 18.
[28] D. L. Pursey, Proc. Roy. Soc. (London) **A275**, 284 (1963).
[29] U. Ottoson, Commun. Math. Phys. **10**, 114 (1968).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Two-Dimensional Hydrogen Bonded Crystals without the Ice Rule*

BILL SUTHERLAND
Physics Department, University of California, Berkeley, California 94720

(Received 13 April 1970)

Models of 2-dimensional hydrogen bonded crystals obeying the ice rule, which previously have been solved exactly, are generalized by removing the ice rule. Many of the peculiar and unique properties of the solutions for the constrained models are now explained by showing that these models, above critical temperature, are equivalent to new unconstrained models at critical temperature. In addition to locating the critical temperature for the general but unsolved models, we locate the singularities of the ground state energy of a related ring of interacting spins.

INTRODUCTION

Much progress has been made recently in solving exactly models for hydrogen bonded crystals in 2 dimensions.[1-4] These models are constrained by the "ice rule." The essential point for the exact solution is Lieb's observation for 2-dimensional ice[1] that the transfer matrix has the same eigenvectors as a solvable 1-dimensional quantum many-body problem, the Heisenberg–Ising ring of interacting spins.[5]

However, the properties of the models, when determined, are surprising, and totally unlike the Ising problem. One can blame this, of course, on the ice rule constraint. However, this remark really does not clarify matters.It is the purpose of this paper to point out the qualitative relationship between constraint and behavior.

1. THE GENERAL EIGHT-SITE LATTICE PROBLEM

Consider a square lattice of N^2 vertices and thus $2N^2$ edges. We assume periodic boundary conditions in both directions. Place $2N^2$ arrows one to an edge, and assign an energy to each configuration of the four arrows about a vertex. In general, therefore, there will be $2^4 = 16$ possible energy assignments.

A large class of solvable models results when the possible vertex configurations obey the "ice rule"; that is, all configurations have infinite energy except the six with two arrows in, two arrows out of a vertex. This general six-site configuration is exactly soluble.[2] The allowed sites are the first six in Fig. 1, with a possible parametrization in the language of ferroelectrics shown below. In this notation, the arrows

[4] J. Fischer, J. Niederle, and R. Raczka, J. Math. Phys. **7**, 816 (1966); R. Raczka, N. Limić, and J. Niederle, *ibid.*, 1861 (1966); B. Limić, J. Niederle, and R. Raczka, *ibid.* **7**, 2026 (1966); **8**, 1079 (1967).
[5] J. Fischer and R. Raczka, Commun. Math. Phys. **3**, 233 (1966); **4**, 8 (1967).
[6] P. Pajas and R. Raczka, J. Math. Phys. **9**, 1188 (1968); P. Pajas, *ibid.* **10**, 1777 (1969).
[7] F. Lurçat, Physics **1**, 95 (1964).
[8] J. Nilsson and A. Beskow, Arkiv Fysik **34**, 307 (1967); J. Nilsson and A. Kihlberg, Report 68-15, Institute of Theoretical Physics, Göteborg, 1968; H. Bacry and A. Kihlberg, J. Math. Phys. **10**, 2132 (1969).
[9] G. Fuchs, Ph.D. thesis, Faculté des Sciences d'Orsay, University of Paris, 1969.
[10] N. X. Hai, Commun. Math. Phys. **12**, 331 (1969).
[11] K. B. Wolf, Ph.D. thesis, Dept. of Physics and Astronomy, Tel-Aviv University, 1969; Tel-Aviv University Preprint 112-69, 1969.
[12] R. Raczka, "Operator Distributions in Group Representation Theory and Their Applications," Lecture Notes, Institute of Theoretical Physics, Göteborg, 1969.
[13] K. Maurin, *General Eigenfunction Expansions and Unitary Representations of Topological Groups* (Polish Scientific Publishers, Warsaw, 1968), p. 143.
[14] Reference 13, p. 157.
[15] M. A. Naimark, *Linear Representations of the Lorentz Group* (Pergamon, London, 1964). S. Ström, Arkiv Fysik **29**, 467 (1965); **33**, 465 (1967); **34**, 215 (1967); A. Sciarrino and M. Toller, J. Math. Phys. **8**, 1252 (1967).
[16] W. T. Sharp, Atomic Energy of Canada Ltd., Chalk River, Ontario, Report CRT-935, 1960.
[17] V. Bargmann, Ann. Math. **48**, 568 (1947); M. Andrews and J. Gunson, J. Math. Phys. **5**, 1391 (1964).
[18] L. C. Biedenharn, J. Math. Phys. **4**, 436 (1963); G. E. Baird and L. C. Biedenharn, *ibid.*, 1449 (1963); A. M. Perelemov and V. S. Popov, Yad. Fiz. **3**, 924 (1966) [Sov. J. Nucl. Phys. **3**, 676 (1966)].
[19] G. Mackey, Bull. Am. Math. Soc. **56**, 385 (1950).
[20] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton U.P., Princeton, N.J., 1957), p. 6.
[21] I. M. Gel'fand and M. L. Tsetlin, Dokl. Akad. Nauk SSSR **71**, 840 (1950).
[22] S. C. Pang and K. T. Hecht, J. Math. Phys. **8**, 1233 (1967); M. K. F. Wong, *ibid.*, 1899 (1967).
[23] Reference 20, Chap. 4.
[24] A. Chakrabarti, J. Math. Phys. **9**, 2087 (1968).
[25] U. Ottoson, Commun. Math. Phys. **8**, 228 (1968).
[26] Our parametrization is the one used in Refs. 2 and 3 and in the work currently done on the representation matrices by Ottoson. Other parametrizations are used in F. D. Murnaghan [*The Unitary and Rotation Groups* (Spartan, New York, 1962)] and L. K. Hua [*Harmonic Analysis of Functions of Several Complex Variables in the Classical Domains* (American Mathematical Society, Providence, R.I., 1963), Translations of Mathematical Monographs, Vol. 6.]
[27] I. M. Gel'fand and M. L. Tsetlin, Dokl. Akad. Nauk SSSR **71**, 825 (1950), and Ref. 18.
[28] D. L. Pursey, Proc. Roy. Soc. (London) **A275**, 284 (1963).
[29] U. Ottoson, Commun. Math. Phys. **10**, 114 (1968).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Two-Dimensional Hydrogen Bonded Crystals without the Ice Rule*

BILL SUTHERLAND
Physics Department, University of California, Berkeley, California 94720

(Received 13 April 1970)

Models of 2-dimensional hydrogen bonded crystals obeying the ice rule, which previously have been solved exactly, are generalized by removing the ice rule. Many of the peculiar and unique properties of the solutions for the constrained models are now explained by showing that these models, above critical temperature, are equivalent to new unconstrained models at critical temperature. In addition to locating the critical temperature for the general but unsolved models, we locate the singularities of the ground state energy of a related ring of interacting spins.

INTRODUCTION

Much progress has been made recently in solving exactly models for hydrogen bonded crystals in 2 dimensions.[1-4] These models are constrained by the "ice rule." The essential point for the exact solution is Lieb's observation for 2-dimensional ice[1] that the transfer matrix has the same eigenvectors as a solvable 1-dimensional quantum many-body problem, the Heisenberg–Ising ring of interacting spins.[5]

However, the properties of the models, when determined, are surprising, and totally unlike the Ising problem. One can blame this, of course, on the ice rule constraint. However, this remark really does not clarify matters.It is the purpose of this paper to point out the qualitative relationship between constraint and behavior.

1. THE GENERAL EIGHT-SITE LATTICE PROBLEM

Consider a square lattice of N^2 vertices and thus $2N^2$ edges. We assume periodic boundary conditions in both directions. Place $2N^2$ arrows one to an edge, and assign an energy to each configuration of the four arrows about a vertex. In general, therefore, there will be $2^4 = 16$ possible energy assignments.

A large class of solvable models results when the possible vertex configurations obey the "ice rule"; that is, all configurations have infinite energy except the six with two arrows in, two arrows out of a vertex. This general six-site configuration is exactly soluble.[2] The allowed sites are the first six in Fig. 1, with a possible parametrization in the language of ferroelectrics shown below. In this notation, the arrows

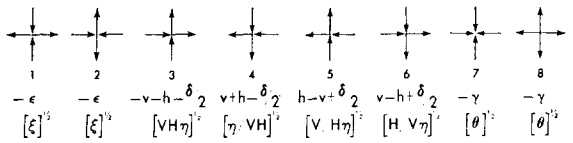


FIG. 1. Allowed configurations of arrows about a vertex, with energy assignments in the language of ferroelectrics, and the corresponding statistical weights. Here $\eta = e^{\delta/T}$, $\xi = e^{2\epsilon/T}$, $H = e^{2h/T}$, $V = e^{2v/T}$, and $\theta = e^{2\gamma/T}$.

are taken to be electric dipoles, and v and h are electric fields in the vertical and horizontal directions, respectively.

There are several peculiar properties of these models, usually blamed in a vague way on the rigid ice rule. Implicit in this criticism is the belief that the Ising model has in some way "normal" behavior. A partial list of peculiar properties: At the critical temperature all derivatives of the free energy for antiferroelectrics are continuous, while ferroelectrics have a latent heat; ferroelectrics are *completely* ordered below the critical temperature; above the critical temperature, correlation functions fall off slower than exponentially.[4,6]

To understand qualitatively the effect of the ice rule, we wish to violate it by adding two new vertex configurations, 7 and 8 in Fig. 1, and study these more general models. This new problem is completely unconstrained, as it is equivalent to an Ising model with many-body potentials. The general eight-site problem, assigning an arbitrary energy to each of the eight vertices, is equivalent to the ferroelectric parametrization of Fig. 1. This is easily seen if one realizes that vertices 7 and 8, and 1 and 2, must each occur in pairs in the crystal; thus there is no generality gained by giving them different energy. We shall, in this article, consider the problem in zero electric field, $v = h = 0$.

We treat such a lattice statistical problem by the transfer matrix method.[7] Briefly, we construct a $2^N \times 2^N$ matrix A with matrix elements between two successive rows of vertical arrows in the lattice. For two given configurations of vertical arrows φ and φ', there are always two ways to place the intervening row of horizontal arrows. Then the matrix element of A is given by

$$A(\varphi, \varphi') = \exp[-E_1(\varphi, \varphi')/T] + \exp[-E_2(\varphi, \varphi')/T],$$

where E_1 and E_2 are the energies of the N intervening vertex configurations for the two choices of horizontal arrows. Then the partition function for the lattice is

$$Z = \mathrm{Tr}\,(A^N) = \lambda^N,$$

where λ is the maximum eigenvalue of A.

2. THE XYZ HAMILTONIAN

Progress was made on solving the six-site ice rule problem, when it was realized that the eigenvectors of A were the same as the eigenvectors of a soluble Hamiltonian for a ring of spins with nearest-neighbor interactions. We shall proceed by making a similar observation—the transfer matrix for our eight-site problem has the same eigenvectors as a new Hamiltonian—which we call the XYZ model.

Consider N spins ($S = \frac{1}{2}$) on a ring interacting with nearest neighbors by the XYZ Hamiltonian:

$$\mathcal{H} = -\sum_{\text{n.n.}} (J_x\sigma_x\sigma_x' + J_y\sigma_y\sigma_y' + J_z\sigma_z\sigma_z'). \qquad (1)$$

The σ's are the Pauli spin matrices, given in a suitable basis by

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ +i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

This Hamiltonian may be exactly solved (at least for the ground state) in the following cases:

(a) Any two coefficients equal, say $A = B$, gives the Heisenberg–Ising lattice.[5]

(b) Any coefficient zero, say $C = 0$, gives the XY model.[8]

Rewriting the Hamiltonian as

$$\mathcal{H}(\Delta, \Gamma) = -(J_x + J_y)\sum_{\text{n.n.}} [\sigma_+\sigma_-' + \sigma_-\sigma_+' + \Gamma(\sigma_+\sigma_+' + \sigma_-\sigma_-') + \tfrac{1}{2}\Delta\sigma_z\sigma_z'], \qquad (2)$$

$$\Gamma = \left(\frac{J_x - J_y}{J_x + J_y}\right), \quad \Delta = \frac{2J_z}{J_x + J_y},$$

we see that $\mathcal{H}$ is exactly soluble as shown in Fig. 2.

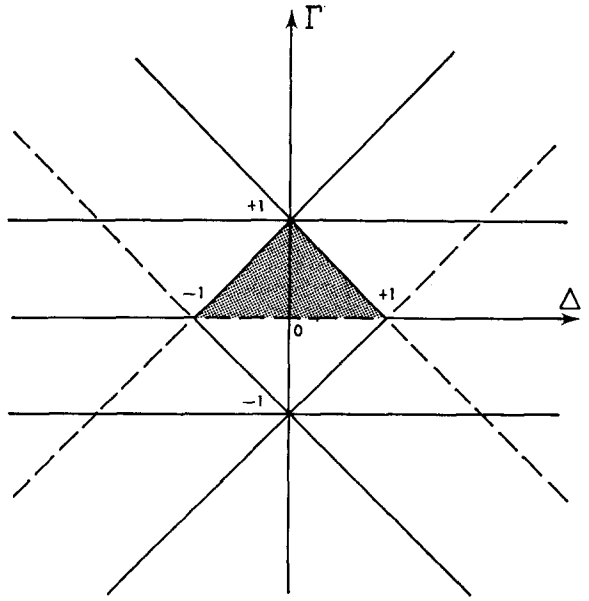


FIG. 2. $\mathcal{H}(\Delta, \Gamma)$ may be solved exactly on all lines, solid and dashed. Singularities, and hence critical points, occur on the dashed lines. Knowing the behavior in the shaded triangle is sufficient to determine the behavior in the entire (Δ, Γ) plane.

3. EIGENVECTORS OF THE TRANSFER MATRIX AND THE *XYZ* HAMILTONIAN

We show A and $\mathcal{H}$ have the same eigenvectors by showing that A and $\mathcal{H}$ commute (with, of course, a relationship between the parameters of each). We briefly outline the proof that $[A, \mathcal{H}] = 0$.[9] A may be written as $\mathrm{Tr}\,(\prod_{i=1}^{N} A(i))$, where each $A(i)$ is a 2×2 matrix with operator elements acting on the vertical arrow (or spin) at site i. The trace indicated is only over the 2×2 product matrix, *not* over the operator elements. The different choices for matrix elements simply indicate the different ways of arranging the horizontal arrows.[10] For our eight-site case, we have

$$A(i) \equiv \begin{pmatrix} A_{11}(i) & A_{12}(i) \\ A_{21}(i) & A_{22}(i) \end{pmatrix} = \begin{pmatrix} \tfrac{1}{2}[\eta^{\frac{1}{2}} + \eta^{-\frac{1}{2}} + \sigma_z(\eta^{\frac{1}{2}} - \eta^{-\frac{1}{2}})] & \xi^{\frac{1}{2}}\sigma_+ + \theta^{\frac{1}{2}}\sigma_- \\ \xi^{\frac{1}{2}}\sigma_- + \theta^{\frac{1}{2}}\sigma_+ & \tfrac{1}{2}[\eta^{\frac{1}{2}} + \eta^{-\frac{1}{2}} - \sigma_z(\eta^{\frac{1}{2}} - \eta^{-\frac{1}{2}})] \end{pmatrix}, \tag{3}$$

where the σ matrices refer to the ith site. More general cases will be treated in a later paper.

$\mathcal{H}$ is a sum of 2-body operators $\mathcal{H} = \sum_i \mathcal{H}_{i,i+1}$. Thus, when we take the commutator with A, we have

$$[A, \mathcal{H}] = \mathrm{Tr}\left(\sum_i [\cdots A(i-1)\{[A(i), \mathcal{H}_{i,i+1}]A(i+1) + A(i)[A(i+1), \mathcal{H}_{i,i+1}]\}A(i+2)\cdots]\right). \tag{4}$$

A typical commutator in the expression of Eq. (4) will again be a 2×2 matrix with operator elements of the form

$$\begin{pmatrix} [A_{11}(i), \mathcal{H}_{i,i+1}] & [A_{12}(i), \mathcal{H}_{i,i+1}] \\ [A_{21}(i), \mathcal{H}_{i,i+1}] & [A_{22}(i), \mathcal{H}_{i,i+1}] \end{pmatrix}.$$

It happens that, if we choose

$$2\Delta = \eta + \eta^{-1} - \xi - \theta, \tag{5a}$$

$$\Gamma = (\xi\theta)^{\frac{1}{2}}, \tag{5b}$$

then the quantity in parenthesis reduces to

$$A(i)Q(i+1) - Q(i)A(i+1),$$

where $Q(i)$ is a 2×2 matrix with operator elements acting *only* on site i. Thus, when we sum over i, the commutator is seen to vanish.

4. SOLUBLE CASES

We know that A and $\mathcal{H}$ have the same eigenvectors, but what does this tell us about either problem since neither, in general, is soluble? First, we can expect the lattice problem to be soluble whenever the corresponding spin problem is. In particular, for the six-site problems, $\theta = 0$ and $\Gamma = 0$, and we see that it is the Heisenberg–Ising chain, as previously known. On the other hand, the line $\Gamma = 1$, equivalent to an XY model with $\Delta = 0$, can easily be shown to correspond to an Ising model in two dimensions and zero magnetic field.[11,12] This correspondence was previously unknown, although the methods of solution were similar for the two problems.

5. DETERMINATION OF SINGULARITIES

But a much more interesting application of this correspondence makes use of the spin Hamiltonian to clearly exhibit symmetries hidden in the lattice statistical problem. Combining this observation with very plausible assumptions enables one to make exact statements about the models, although they cannot be solved.

The idea behind our reasoning is contained in Kramers and Wannier's early treatment of the Ising problem,[7] where they determined the critical temperature T_c before Onsager's exact solution.[13] T_c was determined using only an exact symmetry between high and low temperatures, with the reasonable assumption of one and only one critical temperature.

In our case, we make two assumptions, both extremely plausible. The first is that singularities in the lattice problem and in the Hamiltonian problem occur for the same values of Δ and Γ. This is very reasonable, for singularities are believed to occur at the onset of long-range order. Although the eigenvalues are not the same, and hence free energy and ground state energy are not equal, long-range order is a property of the eigenvector, and hence will be the same. Probably the qualitative nature of the singularities in energy will also be the same; this is borne out by all soluble cases, but has not been shown in general.

Secondly, we make the same assumption as Kramers and Wannier of the existence and uniqueness of the critical temperature T_c.

We now argue as follows: If we fix δ, ϵ, and γ for a particular unconstrained lattice model and then vary T, Eq. (5) generates a curve C in the (Δ, Γ) plane. These curves are always in the $\Gamma > 0$ half-plane; they begin at $\Delta = 0$, $\Gamma = 1$ for $T = \infty$ and move to infinity. Further, one can cover the $\Gamma > 0$ half-plane by varying δ, ϵ, and γ; each curve passes once through either $\Delta = \Gamma + 1$ or $\Delta = -(\Gamma + 1)$. By our second assumption, there is one singular point at the critical temperature for each curve.

If we consider these singular points as a function of

δ, ϵ, and γ, they lie on a curve D in the $\Gamma > 0$ half-plane. By our first assumption, this is also a curve of singularities of the Hamiltonian problem. We now rotate each spin $\frac{1}{2}\pi$ about the y axis; this is a unitary transformation generating a new curve of singularities D_1. Likewise, if we rotate spins on even or odd sites by $\pm\frac{1}{2}\pi$, respectively, about the y axis, this generates a curve D_2 of singularities. But by our first assumption, D_1 and D_2 will, in general, produce new singularities and hence new critical temperatures where they intersect the curves C. This, however, would contradict our second assumption, and so D_1 and D_2 must coincide with D for $\Gamma > 0$. Thus, we see that D is the curve $\Delta = \pm(\Gamma + 1)$, $\Gamma > 0$. The curve D and its images under all unitary transformations of the Hamiltonian we denote by D' and indicate by dashed lines in Fig. 2.

For the spin Hamiltonian $\mathcal{H}(A, B, C)$, we are led to propose that the ground state energy as a function of A, B, C has singularities only on the planes $|A| = |B| \geq |C|$ and the lines $|A| = |B| = 0$, and permutations of these. (The singularity at the Ising limit is not forbidden by our previous reasoning, for these points correspond to $T = 0$ or $T = \infty$.) We notice that, if we were to consider a ring of classical spins in the ground state, then they would change their orientation upon crossing one of the singular planes.

For the lattice problems, we divide the models into two types according to the nature of their order: ferroelectrics whose lowest-energy site is of type 3, 4, 5, or 6 and antiferroelectrics whose lowest-energy site is of type 1, 2, 7, or 8. Then the critical temperature is determined for ferroelectrics by

$$\Delta = \Gamma + 1 \tag{6a}$$

or

$$\eta + \eta^{-1} - \xi - \theta = 2[1 + (\xi\theta)^{\frac{1}{2}}] \tag{7a}$$

and for antiferroelectrics by

$$\Delta = -(\Gamma + 1) \tag{6b}$$

or

$$\eta + \eta^{-1} - \xi - \theta = -2[1 + (\xi\theta)^{\frac{1}{2}}]. \tag{7b}$$

These relations, of course, check for all known cases.

But we notice the surprising fact that the solvable six-site models, constrained by the ice rule, lie entirely on the line of singularities when they are above their critical temperature $|\Delta| = 1$, $\Gamma = 0$. Thus, ice rule models above T_c behave like unconstrained models at the critical point. This explains, for instance, the non-exponential decrease of the correlation function.[14] It is in this sense that we may say the ice rule forces the system to be permanently at a critical point. One expects the general unconstrained models to behave like the Ising model near critical point, although this has not been shown.

* Work supported in part by National Science Foundation grant GP13632.

[1] E. H. Lieb, Phys. Rev. Letters **18,** 692 (1967).

[2] B. Sutherland, Phys. Rev. Letters **19,** 103 (1967); C. P. Yang, *ibid.*, 586 (1967); B. Sutherland, C. N. Yang, and C. P. Yang, *ibid.*, 588 (1967).

[3] E. H. Lieb, Phys. Rev. Letters **18,** 1046 (1967); **19,** 108 (1967).

[4] B. Sutherland, Phys. Letters **26A,** 532 (1968).

[5] C. N. Yang and C. P. Yang, Phys. Rev. **150,** 321, 327 (1966). See these references for earlier work on the Heisenberg–Ising ring.

[6] The statement in Ref. 4 should be changed to indicate "no gap appears in the spectrum for $|\Delta| \leq 1$, hence decrease of correlations above critical temperature is slower than exponential; however, below critical temperature, $\Delta < -1$, indeed a gap appears, and hence we expect exponential decrease." In this regard, the work of J. des Cloizeaux and M. Gaudin, J. Math. Phys. **7,** 1384 (1966), on the spectrum of the Heisenberg–Ising ring is incorrect, as indicated by numerical studies of J. Bonner, thesis, University of London, 1968. In particular, they find no gap between the doubly degenerate ground state and excited states, while Bonner finds a gap equal to the slope of ground-state energy versus Z component of magnetization at zero magnetization. A correct calculation can reproduce Bonner's results analytically, although such a calculation has not yet been published.

[7] H. A. Kramers and G. H. Wannier, Phys. Rev. **66,** 252 (1941).

[8] E. Lieb, T. Schultz, and D. Mattis, Ann. Phys. **16,** 407 (1961); S. Katsura, Phys. Rev. **127,** 1508 (1962).

[9] I would like to thank B. McCoy for making available to the author unpublished work by T. T. Wu and himself, related to the connection between transfer matrix and Hamiltonian.

[10] This concise form for the transfer matrix was suggested to the author by C. Fan.

[11] This correspondence is given in C. Fan and F. Y. Wu, Phys. Rev. **179,** 560 (1969). We let the nearest-neighbor interactions be zero, and hence the lattice decomposes into two superimposed Ising lattices.

[12] Work concerned with relating particular hydrogen bonded models to Ising models is F. Y. Wu, Phys. Rev. Letters **18,** 605 (1967); Phys. Rev. **168,** 539 (1968). In fact, we see that Wu's condition for solvability by dimer methods is equivalent to our $\Delta = 0$.

[13] L. Onsager, Phys. Rev. **65,** 117 (1944).

[14] This possibility was anticipated in remarks to the author by M. Fisher at a Symposium on Exact Results in Statistical Mechanics held at Irvine, California, 1968.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Calculation of the Inner Multiplicity of Weights by Means of the Branching-Law Method and by Racah's Recurrence Relation*

M. K. F. WONG
Department of Physics, Fairfield University, Fairfield, Connecticut 06430
AND
B. GRUBER
Physics Department, St. Louis University, St. Louis, Missouri 63103

(Received 9 March 1970)

We discuss here in detail the validity of the branching-law method suggested in a previous paper [M. K. F. Wong, J. Math. Phys. **11**, 1489 (1970)] for the calculation of the inner multiplicity of weights in an irreducible representation of a classical group. It is found that the method works for all the irreducible representations of $SU(n)$ and $SO(2n+1)$, but that in the case of $SO(2n)$ and $Sp(2n)$ the method does not always give complete solutions except in some simple cases. It is then suggested that Racah's recurrence relation be used in these cases so that complete solutions may be obtained. It is also noted that Racah's recurrence relation alone is sufficient to obtain the inner multiplicity of all weights. This fact is utilized in the calculation of inner multiplicities in another paper [B. Gruber, J. Math. Phys. **11**, 3077 (1970)]. The method suggested in this paper is illustrated through the calculation of some typical examples of the inner multiplicity of weights in the two classical groups $SO(2n)$ and $Sp(2n)$.

INTRODUCTION

In a previous paper,[1] a method was suggested whereby the inner multiplicity of weights in an irreducible representation of a Lie group can be calculated by a formula independently derived by Straumann,[2] Klymyk,[3] and Delaney and Gruber[4] together with the branching laws of Weyl[5] (unitary), Boerner[6] (orthogonal), and Hegerfeldt[7] (symplectic). We stated there that this method, called the branching-law method, works for all the irreducible representations of $SU(n)$ and $SO(2n+1)$, but does not give complete solutions for $SO(2n)$ and $Sp(2n)$ except in some simple cases. We did not give any proofs of why this is so. It is the purpose of this paper to discuss in detail the validity of the branching-law method, i.e., under what conditions do we have complete solutions. It is then suggested that, in those cases of $SO(2n)$ and $Sp(2n)$ where complete solutions are not forthcoming from the branching-law method alone, Racah's recurrence relation[8] for the inner multiplicities be used so as to give complete solutions. It is also noted that Racah's recurrence relation alone is sufficient to obtain the inner multiplicity of all weights in all the irreducible representations of the classical groups. This fact was pointed out by Racah himself, but does not seem to have received enough attention. However, in another paper by Gruber,[9] Racah's recurrence relation has been used to obtain the inner multiplicities in all cases.

This paper is divided into two sections and two appendices. In Sec. 1, the validity of the branching-law method is discussed in detail. In Sec. 2, Racah's recurrence relation for the inner multiplicities is discussed. In the Appendix, we calculate the inner multiplicity of weights in an irreducible representation of the two classical groups $SO(2n)$ and $Sp(2n)$, by means of the branching-law method, supplemented by Racah's recurrence relation.

1. VALIDITY OF THE BRANCHING-LAW METHOD

The main question we wish to ask in this section is: Does the branching-law method give us complete solutions? In other words, are there always enough independent equations for the number of unknowns? The answer is "yes" for the $SU(n)$ and $SO(2n+1)$ groups, but "no" for $SO(2n)$ and $Sp(2n)$. Let us start by considering $SO(2n+1)$.

A. $SO(2n+1)$

For $SO(2n+1)$ the mapping $Lm = m_r = m$ onto weights m_r of its subgroup $SO(2n)$ is nonsingular, i.e., one-to-one. Moreover, it can easily be verified that dominant weights of the group are mapped onto dominant weights of the subgroup.

Let $D(M)$ denote an irreducible representation of $SO(2n+1)$. An ordering in the set of d-weights of $D(M)$ is introduced:

Let

$$m = M + k_1\beta_1 + \cdots + k_n\beta_n$$

be a dominant weight of $D(M)$ and

$$m' = M + k_1'\beta_1 + \cdots + k_n'\beta_n,$$

k_i, k_i' nonnegative integers,

β_i the simple negative roots of $SO(2n+1)$,

another dominant weight of $D(M)$. We say that the weight m is earlier than the weight m', if

$$k_n = k'_n, \quad k_{n-1} = k'_{n-1}, \cdots, k_{u+1} = k'_{u+1},$$

but

$$k_u < k'_u$$

for some value $u = n, n-1, \cdots, 1$. Thus, M is the first weight in the ordered set of d.w.'s of $D(M)$. Going from the first weight, namely M, successively to the following dominant weights of $D(M)$, we obtain through the mapping Lm a succession of weights M_r, M'_r, $M''_r, \cdots, M'''_r$ of the subgroup. Again, in the same order, the set of linear equations is written down (the suffix r referring to the subgroup):

$$\tilde{y}(M_r) = \sum_{S_r \in W_r} \sum_{m \in D(M)} \delta_{S_r} y^M(m) \delta_{Lm, M_r + R_r - S_r R_r}$$

$$\tilde{y}(M'_r) = \cdots$$

$$\vdots$$

$$\tilde{y}(M'''_r) = \cdots.$$

In this system of linear equations there are as many unknowns as there are equations. Moreover, the determinant of the system is nonzero. Thus, a unique solution exists for the unknown $y^M(m)$ for given values $\tilde{y}(M_r), \cdots, \tilde{y}(M'''_r)$.

Proof: Consider an arbitrary equation, for instance, the second equation. In this equation the dominant weight $L^{-1}(M'_r)$ will appear (on the right side of the equation). The only other weights that can appear (but do not necessarily) are the dominant weight $L^{-1}(M_r)$ and the weights $SL^{-1}(M_r)$, $S \in W$, equivalent to $L^{-1}(M_r)$. Thus, $L^{-1}(M'_r)$ is the *last* d.w. that appears. This follows from the fact that in

$$R_r - S_r R_r = k_1\beta_1 + \cdots + k_n\beta_n,$$

$$\beta_i \text{ simple negative roots of } SO(2n+1),$$

the k_i are negative integers or zero and $k_i = 0$, $i = 1, \cdots, n$ only for $S_r = 1$. From this it follows immediately that only the weights $L^{-1}(M'_r)$, $L^{-1}(M_r)$, and $SL^{-1}(M_r)$ can contribute. [If dominant weights following $L^{-1}(M'_r)$ are excluded, then more weights are equivalent to them. The dominant weight M' of a set of equivalent weights is the highest weight of the set, and from this highest weight M' any other member of the set can be reached by $M' + k_1\beta_1 + \cdots + k_n\beta_n$, k_i integers ≥ 0.]

Since the above consideration holds for any one of the equations, it follows that the matrix of this system is triangular and, moreover, that the elements in the diagonal are 1. Thus, the determinant is 1. This completes the proof.

B. $SO(2n)$

The case of $SO(2n)$ is, unfortunately, not so simple. Let the irreducible representation of $SO(2n)$ be $(M_1, M_2, \cdots, M_n)$. Now, in the decomposition of $SO(2n) \supset SO(2n-1)$, there are only $n-1$ numbers $(M'_1, M'_2, \cdots, M'_{n-1})$ representing the irreducible representations of $SO(2n-1)$ since

$$L(m_1, \cdots, m_{n-1}, m_n) = (m_1, \cdots, m_{n-1}).$$

On the other hand, the dominant weights are represented by n numbers $(m_1, m_2, \cdots, m_n)$. The question arises as to whether there are more unknowns than equations. The answer, in general, is "yes." To see this, one notes that, if two dominant weights can be written in the form $(m_1, m_2, \cdots, m_{n-1}, m_n)$ and $(m_1, m_2, \cdots, m_{n-1}, m'_n)$, where the first $n-1$ numbers are the same but the last numbers m_n and m'_n are different mod 2, then it is impossible to distinguish between them in the mapping $L(m)$, and the inner multiplicities of these two weights will occur in the same equation. But such a case is certainly possible. For example, in the irreducible representation (4, 4, 0) of $SO(6)$ we have two dominant weights (4, 3, 1) and (4, 3, −1); three dominant weights (4, 2, 2), (4, 2, −2), and (4, 2, 0), etc. The result is that the branching-law method alone is unable to give unique solutions to these weights in particular. We suggest that, in these cases, Racah's recurrence relation be used in addition to the branching formula in order to give complete solutions.

However, even in the case of $SO(2n)$ the branching-law method is not as bad as it seems. In the first place, there are simple irreducible representations where the above case does not arise. In the second place, in cases where

$$y(m_a) + y(m_b) = 2,$$

it is possible to conclude that

$$y(m_a) = 1 \quad \text{and} \quad y(m_b) = 1$$

because both $y(m_a)$ and $y(m_b)$ are positive nonzero integers. In fact, such cases occur quite often. For a numerical example, see Appendix A.

C. $Sp(2n)$

What was said about $SO(2n)$ is also true in the case of $Sp(2n)$. For example, in the irreducible representation (4, 4, 1) of $Sp(6)$, there are two dominant weights (3, 3, 3) and (3, 3, 1), or (3, 2, 2) and (3, 2, 0), which cannot be distinguished by the branching-law method alone. For a numerical example, see Appendix B.

D. $SU(n)$

The proof for $SU(n+1)$ is the same as for $SO(2n+1)$ once the mapping Lm has been shown

to be nonsingular and, moreover, to map d.w.'s of $SU(n+1)$ onto d.w.'s of the subgroup. This is, however, the case if the branching of $SU(n+1)$ with respect to its subgroup $U(n)$ is considered.

2. RACAH'S RECURRENCE RELATION

Since the branching-law method does not give complete solutions to certain type of weights in $SO(2n)$ and $Sp(2n)$, we have to look elsewhere for other equations that will furnish us with a complete set of solutions. Among the known equations for the inner multiplicity of a weight, we find Racah's recurrence relation to be the simplest. Racah has given the derivation of his recurrence relation, and has also stated that this equation alone is sufficient to determine all the inner multiplicities of weights. This point has been utilized by Gruber to calculate the inner multiplicity of weights in another paper.[9]

It is easy to show why Racah's recurrence relation is sufficient by itself. To see this, we write Racah's recurrence relation as

$$y(m) = -\sum_{\substack{S \\ S \neq 1}} \delta_S y(m + R - SR).$$

Again, we arrange the weights in a definite order: from higher to lower. Then we note that $R - SR$ is always a positive root or sum of positive roots. Therefore, any dominant weight is always related to other weights higher than itself. Now, starting with the multiplicity of the highest weight, which by Cartan's theorem is equal to unity, we can obtain the inner multiplicities of all the dominant weights one by one by means of Racah's recurrence relation.

In the appendices, we calculate numerically the inner multiplicity of weights in the two classical groups $SO(2n)$ and $Sp(2n)$. It is noted that, in the first case [$SO(6)$, with irreducible representation (4, 4, 0)], we only need to use Racah's recurrence relation twice, and, in the second case [$Sp(6)$, with irreducible representation (4, 4, 0)], we only need to use Racah's recurrence relation three times.

APPENDIX A: BRANCHING-LAW METHOD PLUS RACAH'S METHOD: $SO(6)$; IRREDUCIBLE REPRESENTATION (4, 4, 0); DIMENSION 925

For $SO(6) \supset SO(5)$:

$$\tilde{y}(4,4) = \tilde{y}(4,3) = \tilde{y}(4,2) = \tilde{y}(4,1) = \tilde{y}(4,0) = 1, \tag{A1}$$

$$\tilde{y}(4,4) = y(4,4,0) = 1, \tag{A2}$$

$$\tilde{y}(4,3) = y(4,3,1) + y(4,3,-1) - y(4,4,0) = 1,$$
$$\therefore y(4,3,1) = y(4,3,-1) \tag{A3}$$
$$= 1, \tag{A4}$$

$$\tilde{y}(4,2) = y(4,2,2) + y(4,2,-2) + y(4,2,0) - y(4,3,1) - y(4,3,-1) = 1,$$
$$\therefore y(4,2,2) = y(4,2,-2) \tag{A5}$$
$$= y(4,2,0) \tag{A6}$$
$$= 1, \tag{A7}$$

$$\tilde{y}(4,1) = y(4,1,1) + y(4,1,-1) - y(4,2,2) - y(4,2,-2) - y(4,2,0) + y(4,1,3) + y(4,1,-3)$$
$$= 1,$$
$$\therefore y(4,1,1) = y(4,1,-1) \tag{A8}$$
$$= 1, \tag{A9}$$

$$\tilde{y}(4,0) = y(4,0,0) + y(4,0,4) + y(4,0,2) + y(4,0,-4) + y(4,0,-2) - y(4,1,1) - y(4,1,-1)$$
$$- y(4,1,3) - y(4,1,-3) = 1,$$
$$\therefore y(4,0,0) = 1, \tag{A10}$$

$$\tilde{y}(3,3) = y(3,3,2) + y(3,3,-2) + y(3,3,0) - y(3,4,1) - y(3,4,-1) - y(4,2,2) - y(4,2,-2)$$
$$-y(4,2,0) = 0.$$

From Racah's recurrence relation:

$$y(3,3,2) = y(3,4,1) + y(4,2,2) - y(4,4,0) = 1, \tag{A11}$$

$$y(3,3,-2) = y(3,3,2) = 1, \tag{A12}$$

$$\therefore y(3,3,0) = 3, \tag{A13}$$

$$\tilde{y}(3,2) = y(3,2,1) + y(3,2,-1) + 2y(3,2,3) - 2y(4,1,3) - 2y(4,1,1) + y(4,4,0) - 2y(3,3,2) - y(3,3,0) = 0,$$

$$\therefore y(3,2,1) = y(3,2,-1) \tag{A14}$$

$$= 3, \tag{A15}$$

$$\tilde{y}(3,1) = y(3,1,0) + 2y(3,1,4) + 2y(3,1,2) - 2y(3,2,3) - 2y(3,2,1) - 2y(4,0,4) - 2y(4,0,2) - y(4,0,0) + 2y(4,3,1) = 0,$$

$$\therefore y(3,1,0) = 3, \tag{A16}$$

$$\tilde{y}(2,2) = y(2,2,2) + y(2,2,-2) + y(2,2,0) - 2y(3,1,4) - 2y(3,1,2) - y(3,1,0) + 2y(3,4,1) + 2y(4,1,3) + 2y(4,1,1) - y(4,4,0) - 2y(2,3,3) - 2y(2,3,1) = 0.$$

From Racah's recurrence relation:

$$y(2,2,2) = y(2,3,1) - y(2,4,2) + y(2,3,3) + y(3,1,2) - y(3,3,0) - y(4,1,1) - y(4,1,3) + y(4,2,0) + y(4,4,0) = 3, \tag{A17}$$

$$y(2,2,-2) = y(2,2,2) = 3, \tag{A18}$$

$$\therefore y(2,2,0) = 6, \tag{A19}$$

$$\tilde{y}(2,1) = 2y(2,1,3) + y(2,1,1) + y(2,1,-1) - 2y(3,0,3) - 2y(3,0,1) + 2y(3,3,2) + y(3,3,0) + 2y(4,0,4) + 2y(4,0,2) + y(4,0,0) - 2y(4,3,1) - 2y(2,2,4) - 2y(2,2,2) - y(2,2,0) = 0,$$

$$\therefore y(2,1,1) = y(2,1,-1) \tag{A20}$$

$$= 6, \tag{A21}$$

$$\tilde{y}(2,0) = 2y(2,0,4) + 2y(2,0,2) + y(2,0,0) - 2y(3,-1,4) - 2y(3,-1,2) - y(3,-1,0) + 2y(3,2,3) + 2y(3,2,1) + 2y(4,-1,3) + 2y(4,-1,1) - 2y(4,2,2) - y(4,2,0) - 2y(2,1,3) - 2y(2,1,1) = 0,$$

$$\therefore y(2,0,0) = 6, \tag{A22}$$

$$\tilde{y}(1,1) = 2y(1,1,4) + 2y(1,1,2) + y(1,1,0) - 2y(2,0,4) - 2y(2,0,2) - y(2,0,0) + 2y(2,3,3) + 2y(2,3,1) + 2y(3,0,3) + 2y(3,0,1) - 2y(3,3,2) - y(3,3,0) - 2y(4,1,3) - 2y(4,1,1) - 2y(1,2,3) - 2y(1,2,1) + 2y(4,2,2) + y(4,2,0) = 0,$$

$$\therefore y(1,1,0) = 10, \tag{A23}$$

$$\tilde{y}(0,0) = y(0,0,0) + 2y(0,0,4) + 2y(0,0,2) + 2y(1,2,3) + 2y(1,2,1) - 2y(1,-1,4) - 2y(1,-1,2) - y(1,-1,0) + 2y(2,-1,3) + 2y(2,-1,1) - 2y(2,2,4) - 2y(2,2,2) - y(2,2,0) - 2y(3,0,3) - 2y(3,0,1) - 2y(0,1,3) - 2y(0,1,1) + 2y(3,1,4) + 2y(3,1,2) + y(3,1,0) = 0,$$

$$\therefore y(0,0,0) = 15. \tag{A24}$$

APPENDIX B: BRANCHING-LAW METHOD PLUS RACAH'S METHOD: $Sp(6)$; IRREDUCIBLE REPRESENTATION (4, 4, 0); DIMENSION 1274

From $Sp(6) \supset Sp(4)$, we have:

$$\tilde{y}(4,4)=1,\quad \tilde{y}(4,3)=2,\quad \tilde{y}(4,2)=3,\quad \tilde{y}(4,1)=4,$$
$$\tilde{y}(4,0)=5,\quad \tilde{y}(3,3)=1,\quad \tilde{y}(3,2)=2,\quad \tilde{y}(3,1)=3,$$
$$\tilde{y}(3,0)=4,\quad \tilde{y}(2,2)=1,\quad \tilde{y}(2,1)=2,\quad \tilde{y}(2,0)=3,$$
$$\tilde{y}(1,1)=1,\quad \tilde{y}(1,0)=2,\quad \tilde{y}(0,0)=1, \tag{B1}$$

$$\tilde{y}(4,4)=y(4,4,0)=1, \tag{B2}$$

$$\tilde{y}(4,3)=y(4,3,1)+y(4,3,-1)=2,$$
$$\therefore y(4,3,1)=1, \tag{B3}$$

$$\tilde{y}(4,2)=2y(4,2,2)+y(4,2,0)-y(4,4,0)=3.$$

From Racah's recurrence relation:

$$y(4,2,2)=y(4,3,1)=1, \tag{B4}$$
$$\therefore y(4,2,0)=2, \tag{B5}$$
$$\tilde{y}(4,1)=2y(4,1,3)+2y(4,1,1)-2y(4,3,1)=4,$$
$$\therefore y(4,1,1)=2, \tag{B6}$$
$$\tilde{y}(4,0)=2y(4,0,4)+2y(4,0,2)+y(4,0,0)-2y(4,2,2)-y(4,2,0)=5,$$
$$\therefore y(4,0,0)=3, \tag{B7}$$
$$\tilde{y}(3,3)=2y(3,3,2)-2y(4,2,2)-y(4,2,0)+y(3,3,0)=1.$$

From Racah's recurrence relation:

$$y(3,3,2)=y(3,4,1)+y(4,2,2)-y(4,4,0)=1, \tag{B8}$$
$$\therefore y(3,3,0)=3, \tag{B9}$$
$$\tilde{y}(3,2)=2y(3,2,3)+2y(3,2,1)-2y(4,1,3)-2y(4,1,1)-2y(3,4,1)=2,$$
$$\therefore y(3,2,1)=4, \tag{B10}$$
$$\tilde{y}(3,1)=2y(3,1,4)+2y(3,1,2)+y(3,1,0)-2y(3,3,2)-y(3,3,0)-2y(4,0,4)-2y(4,0,2)$$
$$-y(4,0,0)+y(4,4,0)=3,$$
$$\therefore y(3,1,0)=6, \tag{B11}$$
$$\tilde{y}(3,0)=2y(3,0,3)+2y(3,0,1)-2y(3,2,3)-2y(3,2,1)-2y(4,-1,3)-2y(4,-1,1)$$
$$+2y(4,3,1)=4=20-16,$$
$$\tilde{y}(2,2)=2y(2,2,4)+2y(2,2,2)+y(2,2,0)-2y(2,4,2)-y(2,4,0)-2y(3,1,4)-2y(3,1,2)$$
$$-y(3,1,0)=1.$$

From Racah's recurrence relation:

$$y(2,2,2)=y(2,2,4)+y(2,3,1)+y(3,1,2)-y(3,1,4)-y(3,3,0)-y(4,1,1)+y(4,2,0)$$
$$=5, \tag{B12}$$
$$\therefore y(2,2,0)=9, \tag{B13}$$
$$\tilde{y}(2,1)=2y(2,1,3)+2y(2,1,1)-2y(2,3,3)-2y(2,3,1)-2y(3,0,3)-2y(3,0,1)+2y(3,4,1)$$
$$=2,$$
$$\therefore y(2,1,1)=10, \tag{B14}$$

$$\tilde{y}(2,0) = 2y(2,0,4) + 2y(2,0,2) + y(2,0,0) - 2y(2,2,4) - 2y(2,2,2) - y(2,2,0) - 2y(3,-1,4) - 2y(3,-1,2) - y(3,-1,0) + 2y(3,3,2) + y(3,3,0) = 3,$$

$$\therefore y(2,0,0) = 13, \tag{B15}$$

$$\tilde{y}(1,1) = 2y(1,1,4) + 2y(1,1,2) + y(1,1,0) - 2y(1,3,4) - 2y(1,3,2) - y(1,3,0) - 2y(2,0,4) - 2y(2,0,2) - y(2,0,0) + 2y(4,0,4) + 2y(4,0,2) + y(4,0,0) + 2y(2,4,2) + y(2,4,0) - y(4,4,0) = 1,$$

$$\therefore y(1,1,0) = 16, \tag{B16}$$

$$\tilde{y}(0,0) = 2y(0,0,4) + 2y(0,0,2) + y(0,0,0) - 2y(4,0,4) - 2y(4,0,2) - y(4,0,0) - 2y(0,2,4) - 2y(0,2,2) - y(0,2,0) + 2y(4,2,2) + y(4,2,0) - 2y(1,-1,4) - 2y(1,-1,2) - y(1,-1,0) + 2y(3,-1,4) + 2y(3,-1,2) + y(3,-1,0) + 2y(1,3,4) + 2y(1,3,2) + y(1,3,0) - 2y(3,3,2) - y(3,3,0) = 1,$$

$$\therefore y(0,0,0) = 22. \tag{B17}$$

* Supported in part by the National Science Foundation, Grant GP 9623.

[1] M. K. F. Wong, J. Math. Phys. **11**, 1489 (1970), henceforth referred to as I.
[2] N. Straumann, Helv. Phys. Acta **38**, 481 (1966).
[3] A. U. Klymyk, Academy of Sciences Ukranian SSR, Institute for Theoretical Physics, Kiev, No. 67-17 (1967).
[4] R. M. Delaney and B. Gruber, J. Math. Phys. **10**, 252 (1969).
[5] H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover, New York, 1931), 2nd ed.
[6] H. Boerner, *Representations of Groups* (North-Holland, Amsterdam, 1963), Chap. VIII.
[7] G. C. Hegerfeldt, J. Math. Phys. **8**, 1195 (1967).
[8] G. Racah, *Group Theoretical Concepts and Methods in Elementary Particle Physics*, F. Gursey, Ed., (Gordon and Breach, New York, 1962).
[9] B. Gruber, J. Math. Phys. **11**, 3077 (1970).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Coefficients Connecting the Stark and Field-Free Wavefunctions for Hydrogen*

C. Bruce Tarter
Lawrence Radiation Laboratory, University of California, Livermore, California 94550

(Received 8 June 1970)

A general expression for the coefficients connecting the Stark (parabolic coordinates) and field-free (spherical coordinates) wavefunctions for hydrogen is obtained. The result, which involves a generalized hypergeometric function, has been numerically evaluated through principal quantum number $n = 10$.

I. INTRODUCTION

Consider a field-free region of space with a density N_{nlm} of hydrogen atoms in the state with quantum numbers n, l, and m and wavefunction Φ_{nlm} in spherical coordinates. If at some instant a strong electric field is switched on,[1] the appropriate representation of the states is in terms of the quantum numbers n, n_1, n_2, and m (with $n = n_1 + n_2 + m + 1$) and the "Stark" wavefunctions $u_{n(n_1n_2m)}$ expressed in parabolic coordinates. This transition is described by

$$\Phi_{nlm} = \sum_{n_1,n_2} A_{nlm}^{n_1,n_2} u_{n(n_1n_2m)} \tag{1}$$

and the density of atoms in the state $u_{n(n_1n_2)m}$ is

$$N_{n(n_1n_2m)} = \sum_{l=m}^{n-1} |A_{nlm}^{n_1n_2}|^2 N_{nlm}. \tag{2}$$

The purpose of this paper is to obtain a closed expression for the coefficients $A_{nlm}^{n_1n_2}$.

II. CALCULATION OF THE COEFFICIENTS

A. The General Case

Before beginning the calculation, it is convenient to define the generalized hypergeometric function

$${}_pF_q\left[\begin{matrix}\alpha_1, \cdots, \alpha_p; \\ \beta_1, \cdots, \beta_q;\end{matrix} x\right] = \sum_{n=0}^{\infty} \frac{(\alpha_1)_n \cdots (\alpha_p)_n x^n}{(\beta_1)_n \cdots (\beta_q)_n n!}, \tag{3}$$

$$\tilde{y}(2,0) = 2y(2,0,4) + 2y(2,0,2) + y(2,0,0) - 2y(2,2,4) - 2y(2,2,2) - y(2,2,0) - 2y(3,-1,4) - 2y(3,-1,2) - y(3,-1,0) + 2y(3,3,2) + y(3,3,0) = 3,$$

$$\therefore y(2,0,0) = 13, \tag{B15}$$

$$\tilde{y}(1,1) = 2y(1,1,4) + 2y(1,1,2) + y(1,1,0) - 2y(1,3,4) - 2y(1,3,2) - y(1,3,0) - 2y(2,0,4) - 2y(2,0,2) - y(2,0,0) + 2y(4,0,4) + 2y(4,0,2) + y(4,0,0) + 2y(2,4,2) + y(2,4,0) - y(4,4,0) = 1,$$

$$\therefore y(1,1,0) = 16, \tag{B16}$$

$$\tilde{y}(0,0) = 2y(0,0,4) + 2y(0,0,2) + y(0,0,0) - 2y(4,0,4) - 2y(4,0,2) - y(4,0,0) - 2y(0,2,4) - 2y(0,2,2) - y(0,2,0) + 2y(4,2,2) + y(4,2,0) - 2y(1,-1,4) - 2y(1,-1,2) - y(1,-1,0) + 2y(3,-1,4) + 2y(3,-1,2) + y(3,-1,0) + 2y(1,3,4) + 2y(1,3,2) + y(1,3,0) - 2y(3,3,2) - y(3,3,0) = 1,$$

$$\therefore y(0,0,0) = 22. \tag{B17}$$

* Supported in part by the National Science Foundation, Grant GP 9623.

[1] M. K. F. Wong, J. Math. Phys. **11**, 1489 (1970), henceforth referred to as I.
[2] N. Straumann, Helv. Phys. Acta **38**, 481 (1966).
[3] A. U. Klymyk, Academy of Sciences Ukranian SSR, Institute for Theoretical Physics, Kiev, No. 67-17 (1967).
[4] R. M. Delaney and B. Gruber, J. Math. Phys. **10**, 252 (1969).
[5] H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover, New York, 1931), 2nd ed.
[6] H. Boerner, *Representations of Groups* (North-Holland, Amsterdam, 1963), Chap. VIII.
[7] G. C. Hegerfeldt, J. Math. Phys. **8**, 1195 (1967).
[8] G. Racah, *Group Theoretical Concepts and Methods in Elementary Particle Physics*, F. Gursey, Ed., (Gordon and Breach, New York, 1962).
[9] B. Gruber, J. Math. Phys. **11**, 3077 (1970).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Coefficients Connecting the Stark and Field-Free Wavefunctions for Hydrogen*

C. Bruce Tarter
Lawrence Radiation Laboratory, University of California, Livermore, California 94550

(Received 8 June 1970)

A general expression for the coefficients connecting the Stark (parabolic coordinates) and field-free (spherical coordinates) wavefunctions for hydrogen is obtained. The result, which involves a generalized hypergeometric function, has been numerically evaluated through principal quantum number $n = 10$.

I. INTRODUCTION

Consider a field-free region of space with a density N_{nlm} of hydrogen atoms in the state with quantum numbers n, l, and m and wavefunction Φ_{nlm} in spherical coordinates. If at some instant a strong electric field is switched on,[1] the appropriate representation of the states is in terms of the quantum numbers n, n_1, n_2, and m (with $n = n_1 + n_2 + m + 1$) and the "Stark" wavefunctions $u_{n(n_1n_2m)}$ expressed in parabolic coordinates. This transition is described by

$$\Phi_{nlm} = \sum_{n_1,n_2} A_{nlm}^{n_1,n_2} u_{n(n_1n_2m)} \tag{1}$$

and the density of atoms in the state $u_{n(n_1n_2)m}$ is

$$N_{n(n_1n_2m)} = \sum_{l=m}^{n-1} |A_{nlm}^{n_1n_2}|^2 N_{nlm}. \tag{2}$$

The purpose of this paper is to obtain a closed expression for the coefficients $A_{nlm}^{n_1n_2}$.

II. CALCULATION OF THE COEFFICIENTS

A. The General Case

Before beginning the calculation, it is convenient to define the generalized hypergeometric function

$${}_pF_q\left[\begin{matrix}\alpha_1, \cdots, \alpha_p; \\ \beta_1, \cdots, \beta_q;\end{matrix} x\right] = \sum_{n=0}^{\infty} \frac{(\alpha_1)_n \cdots (\alpha_p)_n x^n}{(\beta_1)_n \cdots (\beta_q)_n n!}, \tag{3}$$

where

$$(\alpha)_0 = 1(\alpha)_n = \alpha(\alpha+1)\cdots(\alpha+n-1).$$

The series terminates if one of the α_j is a negative integer. In the above notation the confluent hypergeometric function is written as ${}_1F_1[{}^{\alpha;}_{\beta;}{}^{x}]$ and the ordinary hypergeometric function becomes ${}_2F_1[{}^{\alpha_1,\alpha_2;}_{\beta;}{}^{x}]$. Additional properties of these functions are given elsewhere.[2,3]

The expansion coefficients defined in (1) may be written as

$$A^{n_1n_2}_{nlm} = \int u^*_{n(n_1n_2m)}\phi_{nlm}\,d\tau. \tag{4}$$

The hydrogen atom wavefunctions expressed in spherical and parabolic coordinates, respectively, are[4]

$$\phi_{nlm} = \frac{2}{n^2(2l+1)!}\left(\frac{(n+l)!}{(n-l-1)!}\right)^{\frac{1}{2}}\frac{e^{im\phi}}{\pi^{\frac{1}{2}}} \times P_l^m(\cos\theta)\left(\frac{2r}{n}\right)^l e^{-r/n} \times {}_1F_1\left[\begin{matrix}-(n-l-1);\ 2r/n\\ 2l+2;\end{matrix}\right], \tag{5}$$

$$u^*_{n(n_1n_2m)} = \frac{1}{n^2}\left(\frac{1}{m!}\right)^2\left(\frac{(n_1+m)!\,(n_2+m)!}{n_1!\,n_2!}\right)^{\frac{1}{2}}\left(\frac{r}{n}\right)^m \times e^{-r/n}\frac{e^{-im\phi}}{\pi^{\frac{1}{2}}}(\sin^m\theta) \times {}_1F_1\left[\begin{matrix}-n_1;\ r(1+\cos\theta)/n\\ m+1;\end{matrix}\right] \times {}_1F_1\left[\begin{matrix}-n_2;\ r(1-\cos\theta)/n\\ m+1;\end{matrix}\right]. \tag{6}$$

The associate Legendre polynomial may be expressed in terms of the ordinary hypergeometric function[3]:

$$P_l^m(\cos\theta) = \left(\frac{(l+m)!\,(2l+1)}{(l-m)!}\right)^{\frac{1}{2}}\frac{1}{2^m}\frac{\sin^m\theta}{m!} \times {}_2F_1\left[\begin{matrix}m+l+1;\ -(l-m);\ (1-\cos\theta)/2\\ m+1;\end{matrix}\right]. \tag{7}$$

After performing the ϕ integration and making the substitutions $y = r/2n$ and $x = \frac{1}{2}(1-\cos\theta)$, we have

$$A^{n_1n_2}_{nlm} = \frac{1}{2n}\left(\frac{1}{m!}\right)^3\frac{(2l+1)^{\frac{1}{2}}}{(2l+1)!} \times \left(\frac{(n+l)!\,(l+m)!\,(n_1+m)!\,(n_2+m)!}{(n-l-1)!\,(l-m)!\,n_1!\,n_2!}\right)^{\frac{1}{2}} I_{xy}, \tag{8}$$

where

$$I_{xy} = \int_0^\infty\int_0^1 y^{m+l+2}x^m(1-x)^m e^{-y} \times {}_2F_1\left[\begin{matrix}m+l+1,\ -(l-m);\ x\\ m+1;\end{matrix}\right] \times {}_1F_1\left[\begin{matrix}-(n-l-1);\ y\\ 2l+2;\end{matrix}\right] \times {}_1F_1\left[\begin{matrix}-n_1;\ (1-x)y\\ m+1;\end{matrix}\right]{}_1F_1\left[\begin{matrix}-n_2;\ xy\\ m+1;\end{matrix}\right] dx\,dy. \tag{9}$$

Each of the above hypergeometric functions may be expressed as a finite series:

$${}_1F_1\left[\begin{matrix}-(n-l-1);\ y\\ 2l+2;\end{matrix}\right] = \sum_{s=0}^{n-l-1}\frac{[-(n-l-1)]_s y^s}{(2l+2)_s s!}\,{}_1F_1\left[\begin{matrix}-n_1;\ (1-x)y\\ m+1;\end{matrix}\right] = \sum_{k=0}^{n_1}\frac{-(n_1)_k y^k(1-x)^k}{(m+1)_k k!}, \tag{10a}$$

$${}_1F_1\left[\begin{matrix}-n_2;\ xy\\ m+1;\end{matrix}\right] = \sum_{t=0}^{n_2}\frac{(-n_2)_t x^t y^t}{(m+1)_t t!}\,{}_2F_1\left[\begin{matrix}m+l+1;\ -(l-m);\ x\\ m+1;\end{matrix}\right] = \sum_{r=0}^{l-m}\frac{[-(l-m)]_r(m+l+1)_r x^r}{(m+1)_r r!}. \tag{10b}$$

The x and y integrals are now elementary, and we have

$$I_x = \int_0^1 x^{m+r+t}(1-x)^{m+k} = \frac{\Gamma(m+r+t+1)\Gamma(m+k+1)}{\Gamma(2m+r+t+k+2)}, \tag{11}$$

$$I_y = \int_0^\infty y^{l+m+k+t+s+2}e^{-y} = \Gamma(m+l+k+t+s+3). \tag{12}$$

Thus,

$$I_{xy} = \sum_{r=0}^{l-m}\sum_{s=0}^{n-l-1}\sum_{k=0}^{n_1}\sum_{t=0}^{n_2}\frac{(-n_1)_k(-n_2)_t[-(n-l-1)]_s[-(l-m)]_r(l+m+1)_r\Gamma(m+1)}{k!\,t!\,s!\,r!\,(m+1)_t(m+1)_r(2l+2)_s} \times \frac{\Gamma(m+r+t+1)\Gamma(k+t+s+m+l+3)}{\Gamma(2m+r+t+k+2)}. \tag{13}$$

Since the sums in r and s are uncoupled, it is convenient to evaluate these first:

$$\sum_{s=0}^{n-l-1} \frac{[-(n-l-1)]_s}{s!\,(2l+2)_s}\,\Gamma(m+l+k+t+s+3)$$
$$= \Gamma(m+l+k+t+3) \times {}_2F_1\left[\begin{matrix}-(n-l-1), m+l+k+t+3;\\ 2l+2;\end{matrix}\,1\right]$$
$$= \frac{\Gamma(m+l+k+t+3)\Gamma(2l+2)}{(n+l+1)} \times (l-m-k-t-1)_{n-l-1}, \quad (14)$$

by Vandermonde's theorem,[2] and

$$\sum_{r=0}^{l-m} \frac{(l+m+1)_r[(l-m)]_r\Gamma(m+r+t+1)}{(m+1)_r\Gamma(2m+r+t+k+2)}$$
$$= \frac{\Gamma(m+t+1)}{\Gamma(2m+k+t+2)} \times {}_3F_2\left[\begin{matrix}-(l-m), m+l+1, m+t+1;\\ m+1, 2m+k+t+2;\end{matrix}\,1\right]. \quad (15)$$

Cancelling out several terms, we are left with

$$I_{xy} = \frac{[\Gamma(m+1)]^2\Gamma(2l+2)}{\Gamma(n+l+1)} \sum_{k=0}^{n_1}\sum_{t=0}^{n_2} \frac{(-n_1)_k(-n_2)_t\Gamma(m+l+k+t+3)(l-m-k-t-1)_{n-l-1}}{k!\,t!\,\Gamma(2m+k+t+2)} \times {}_3F_2\left[\begin{matrix}-(l-m), m+l+1, m+t+1;\\ m+1, 2m+k+t+2;\end{matrix}\,1\right]. \quad (16)$$

The term $(l-m-k-t-1)_{n-l-1}$ is nonzero only if $k+t \leq l-m-1$ or $k+t \geq n-m-2 = n_1+n_2-1$. Also, utilizing a transformation[2] of the ${}_3F_2$,

$${}_3F_2\left[\begin{matrix}a, a', -N;\\ c', 1-N-c;\end{matrix}\,1\right] = \frac{(c+a)_N}{(c)_N}\,{}_3F_2\left[\begin{matrix}a, c'-a', -N;\\ c', c+a;\end{matrix}\,1\right], \quad (17)$$

one can show that our ${}_3F_2$ is zero unless $k+t \geq b-c$. Combining these three inequalities, we see that it is evident that only the three terms with $k+t \geq n_1+n_2-1$ contribute. After algebra, we have

$$I_{xy} = \frac{(-1)^{l-m}(m!)^2(2l+1)!\,(n-l-1)!}{(n+m)!} \times \{(n+l+1)(n-1)\,{}_3F_2(n_1, n_2) - (n+m) \times [n_1\,{}_3F_2(n_1-1, n_2) + n_2\,{}_3F_2(n_1, n_2-1)]\}, \quad (18)$$

where

$${}_3F_2(k, t) = {}_3F_2\left[\begin{matrix}m+t+1, m+l+1, -(l-m);\\ m+1, 2m+k+t+2;\end{matrix}\,1\right]. \quad (19)$$

After applying the transformation (17), with $c' = m+1$ and $a' = m+n_2+1$, and noting the relation of Rainville,[5]

$$(\alpha_1-\beta_2+1)\,{}_3F_2\left[\begin{matrix}\alpha_1, \alpha_2, \alpha_3;\\ \beta_1, \beta_2;\end{matrix}\,1\right] = \alpha_1\,{}_3F_2\left[\begin{matrix}\alpha_1+1, \alpha_2, \alpha_3;\\ \beta_1, \beta_2;\end{matrix}\,1\right] - (\beta_2-1)\,{}_3F_2\left[\begin{matrix}\alpha_1, \alpha_2, \alpha_3;\\ \beta_1, \beta_2-1;\end{matrix}\,1\right], \quad (20)$$

we finally have

$$I_{xy} = \frac{m!\,(2l+1)!\,2n(n-m-1)!\,(-1)^{l-m}}{(n+l)!} \times {}_3F_2\left[\begin{matrix}m+l+1, -(l-m), -n_2;\\ m+1, -(n-m-1);\end{matrix}\,1\right]. \quad (21)$$

Combining this with (8), we have that our final expression for the coefficients becomes

$$A_{nlm}^{n_1n_2} = (-1)^{l-m}\frac{(n-m-1)!}{m!} \times \left(\frac{(2l+1)(l+m)!\,(n_1+m)!\,(n_2+m)!}{(n+l)!\,(l-m)!\,n_1!\,n_2!\,(n-l-1)!}\right)^{\frac{1}{2}} \times {}_3F_2\left[\begin{matrix}m+l+1, -(l-m), -n_2;\\ m+1, -(n-m-1);\end{matrix}\,1\right]. \quad (22)$$

B. Special Properties

Utilizing Eq. (17) with $N = l-m$, $c = -l$, $c' = -(n-m-1) = -n_1-n_2$, $a = m+l+1$, and $a' = -n_2$, one has

$${}_3F_2\left[\begin{matrix}m+l+1, -(l-m), -n_2;\\ m+1, -(n-m-1);\end{matrix}\,1\right] = (-1)^{l-m}\,{}_3F_2\left[\begin{matrix}m+l+1, -(l-m), -n_1;\\ m+1, -(n-m-1);\end{matrix}\,1\right], \quad (23)$$

which yields the symmetry relation

$$A_{nlm}^{n_1,n_2} = (-1)^{l-m}A_{nlm}^{n_2,n_1}. \quad (24)$$

Thus, for example, if $n_1 = n_2$ and $l-m$ is odd, $A_{nlm}^{n_1,n_2} = 0$.

In some cases the formula (8) is reducible, viz., the hypergeometric function may be summed explicitly in terms of gamma functions. For the case $l = m$ one

may easily show that

$$A_{nlm}^{n_1,n_2} = \left(\frac{\binom{n_1+l}{l}\binom{n_2+l}{l}}{\binom{n+l}{2l+1}} \right)^{\frac{1}{2}}. \tag{25}$$

Using Vandermonde's theorem,[3] we have that the case $l = n - 1$ may be reduced to

$$A_{nlm}^{n_1,n_2} = (-1)^{n_2+l+m} \left[\frac{\binom{n-m-1}{n_1}\binom{n+m-1}{n_1+m}}{\binom{2n-2}{n-1}} \right]^{\frac{1}{2}}. \tag{26}$$

In Eqs. (25) and (26) the binomial coefficients are defined as usual by

$$\binom{\lambda}{\mu} = \frac{\lambda!}{\mu!\,(\lambda-\mu)!}. \tag{27}$$

III. TABULATION OF THE COEFFICIENTS

In Table I, the coefficients $A_{nlm}^{n_1,n_2}$ are tabulated for $n \leq 3$. A more extensive tabulation through $n = 10$ is given elsewhere.[6] The accuracy of these tables has been checked using the relations

$$\sum_{l=m}^{n-1} |A_{nlm}^{n_1,n_2}|^2 = \sum_{n_1,n_2} |A_{nlm}^{n_1,n_2}|^2 = 1. \tag{28}$$

TABLE I. The expansion coefficients $A_{nlm}^{n_1,n_2}$ for $n \leq 3$.

$(n, l, m; n_1, n_2)$	$A_{nlm}^{n_1,n_2}$	$(n, l, m; n_1, n_2)$	$A_{nlm}^{n_1,n_2}$
(1, 0, 0; 0, 0)	1.000	(3, 1, 0; 1, 1)	0.000
(2, 0, 0; 0, 1)	0.707	(3, 1, 0; 2, 0)	−0.707
(2, 0, 0; 1, 0)	0.707	(3, 1, 1; 0, 1)	0.707
(2, 1, 0; 0, 1)	0.707	(3, 1, 1; 1, 0)	0.707
(2, 1, 0; 1, 0)	−0.707	(3, 2, 0; 0, 2)	0.408
(2, 1, 1; 0, 0)	1.000	(3, 2, 0; 1, 1)	−0.816
(3, 0, 0; 0, 2)	0.577	(3, 2, 0; 2, 0)	0.408
(3, 0, 0; 1, 1)	0.577	(3, 2, 1; 0, 1)	0.707
(3, 0, 0; 2, 0)	0.577	(3, 2, 1; 1, 0)	−0.707
(3, 1, 0; 0, 2)	0.707	(3, 2, 2; 0, 0)	1.000

* Work performed under the auspices of the U.S. Atomic Energy Commission.

[1] Alternatively, if the atom is moving, a magnetic field will generate an effective Lorentz electric field. The original motivation for the calculation was an experimental situation in which an energetic beam of hydrogen atoms moved from a field-free region into one in which there was a strong magnetic field (see Ref. 6).

[2] W. N. Bailey, *Generalized Hypergeometric Series* (Cambridge Tracts No. 32, Cambridge, 1935).

[3] A. Erdélyi *et al.*, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. I.

[4] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One and Two Electron Atoms* (Academic, New York, 1957).

[5] E. D. Rainville, Bull. Am. Math. Soc. **22**, 370 (1945).

[6] C. B. Tarter, Lawrence Radiation Laboratory Rept. UCRL-7493, 1963. Additional information and applications of the present work are also given in this unpublished report.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

The Proper Vibration of the Space

J. KULHÁNEK
Physics Department, University of Windsor, Windsor, Ontario, Canada

(Received 29 September 1969; Revised Manuscript Received 3 March 1970)

The proper vibrations of homogeneous and isotropic space are examined on the basis of the equations of the de Broglie wave field (field equations). The time-dependent part of the wavefunction, which is a solution of the Klein–Gordon equation, satisfies the differential equation which coincides with the differential equation derived from field equations for the time-dependent part of the Robertson–Walker metric.

1. INTRODUCTION

The wave processes in homogeneous and isotropic closed space are studied in Refs. 1–6. In the case when the space is closed, D'Alembert's, Maxwell's, Klein–Gordon's, and Dirac's equations provide eigenvalue problems. Our analysis here seems to point toward a possibility which differs considerably from the ones mentioned above. We are considering a possibility of proper vibration of the homogeneous and isotropic space itself. In the case when the space is closed, its proper vibration might be discontinuous and thus provide an adequate description of the observed atomicity of matter and light.

We now give a brief review of the conclusions reached in this paper. We start from field equations[7] which are a system of simultaneous second-order nonlinear partial-differential equations for the components of metric tensor g_{ab}. They have real characteristic surfaces,[8] which are identical with the 3-dimensional wavesurfaces of the de Broglie waves.[7,9]

may easily show that

$$A_{nlm}^{n_1,n_2} = \left(\frac{\binom{n_1+l}{l}\binom{n_2+l}{l}}{\binom{n+l}{2l+1}} \right)^{\frac{1}{2}}. \tag{25}$$

Using Vandermonde's theorem,[3] we have that the case $l = n - 1$ may be reduced to

$$A_{nlm}^{n_1,n_2} = (-1)^{n_2+l+m} \left[\frac{\binom{n-m-1}{n_1}\binom{n+m-1}{n_1+m}}{\binom{2n-2}{n-1}} \right]^{\frac{1}{2}}. \tag{26}$$

In Eqs. (25) and (26) the binomial coefficients are defined as usual by

$$\binom{\lambda}{\mu} = \frac{\lambda!}{\mu!\,(\lambda-\mu)!}. \tag{27}$$

III. TABULATION OF THE COEFFICIENTS

In Table I, the coefficients $A_{nlm}^{n_1,n_2}$ are tabulated for $n \leq 3$. A more extensive tabulation through $n = 10$ is given elsewhere.[6] The accuracy of these tables has been checked using the relations

$$\sum_{l=m}^{n-1} |A_{nlm}^{n_1,n_2}|^2 = \sum_{n_1,n_2} |A_{nlm}^{n_1,n_2}|^2 = 1. \tag{28}$$

TABLE I. The expansion coefficients $A_{nlm}^{n_1,n_2}$ for $n \leq 3$.

$(n, l, m; n_1, n_2)$	$A_{nlm}^{n_1,n_2}$	$(n, l, m; n_1, n_2)$	$A_{nlm}^{n_1,n_2}$
(1, 0, 0; 0, 0)	1.000	(3, 1, 0; 1, 1)	0.000
(2, 0, 0; 0, 1)	0.707	(3, 1, 0; 2, 0)	−0.707
(2, 0, 0; 1, 0)	0.707	(3, 1, 1; 0, 1)	0.707
(2, 1, 0; 0, 1)	0.707	(3, 1, 1; 1, 0)	0.707
(2, 1, 0; 1, 0)	−0.707	(3, 2, 0; 0, 2)	0.408
(2, 1, 1; 0, 0)	1.000	(3, 2, 0; 1, 1)	−0.816
(3, 0, 0; 0, 2)	0.577	(3, 2, 0; 2, 0)	0.408
(3, 0, 0; 1, 1)	0.577	(3, 2, 1; 0, 1)	0.707
(3, 0, 0; 2, 0)	0.577	(3, 2, 1; 1, 0)	−0.707
(3, 1, 0; 0, 2)	0.707	(3, 2, 2; 0, 0)	1.000

* Work performed under the auspices of the U.S. Atomic Energy Commission.

[1] Alternatively, if the atom is moving, a magnetic field will generate an effective Lorentz electric field. The original motivation for the calculation was an experimental situation in which an energetic beam of hydrogen atoms moved from a field-free region into one in which there was a strong magnetic field (see Ref. 6).

[2] W. N. Bailey, *Generalized Hypergeometric Series* (Cambridge Tracts No. 32, Cambridge, 1935).

[3] A. Erdélyi *et al.*, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. I.

[4] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One and Two Electron Atoms* (Academic, New York, 1957).

[5] E. D. Rainville, Bull. Am. Math. Soc. **22**, 370 (1945).

[6] C. B. Tarter, Lawrence Radiation Laboratory Rept. UCRL-7493, 1963. Additional information and applications of the present work are also given in this unpublished report.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

The Proper Vibration of the Space

J. KULHÁNEK
Physics Department, University of Windsor, Windsor, Ontario, Canada

(Received 29 September 1969; Revised Manuscript Received 3 March 1970)

The proper vibrations of homogeneous and isotropic space are examined on the basis of the equations of the de Broglie wave field (field equations). The time-dependent part of the wavefunction, which is a solution of the Klein–Gordon equation, satisfies the differential equation which coincides with the differential equation derived from field equations for the time-dependent part of the Robertson–Walker metric.

1. INTRODUCTION

The wave processes in homogeneous and isotropic closed space are studied in Refs. 1–6. In the case when the space is closed, D'Alembert's, Maxwell's, Klein–Gordon's, and Dirac's equations provide eigenvalue problems. Our analysis here seems to point toward a possibility which differs considerably from the ones mentioned above. We are considering a possibility of proper vibration of the homogeneous and isotropic space itself. In the case when the space is closed, its proper vibration might be discontinuous and thus provide an adequate description of the observed atomicity of matter and light.

We now give a brief review of the conclusions reached in this paper. We start from field equations[7] which are a system of simultaneous second-order nonlinear partial-differential equations for the components of metric tensor g_{ab}. They have real characteristic surfaces,[8] which are identical with the 3-dimensional wavesurfaces of the de Broglie waves.[7,9]

Under the assumption that the space $x^0 = \text{const}$ is homogeneous and isotropic, we can solve the field equations using the Robertson–Walker metric. The Robertson–Walker metric is a relatively simple metric which can be associated with the particle at rest in the customary comoving coordinate system.

The field equations yield a single differential equation for the scalar factor which describes the time development of the metric of space. This differential equation coincides with the differential equation for the time-dependent part of the wavefunction, which is the solution of Klein–Gordon equation for the Robertson–Walker metric.[4]

It is a property of the Robertson–Walker metric that the distance of two points in space is proportional to the scalar factor and thus it is a function of time. The time dependence of distance is determined by the nonlinear differential equation which represents an autonomous (not explicitly containing the independent variable), conservative (not containing a first derivative) oscillation.

The volume of the space shares the common oscillation. The function which describes the time dependence of the volume is a solution of a nonlinear differential equation which also represents an autonomous, conservative oscillation.

The equation for the distance oscillation and the equation for the volume oscillation are coupled. The nonlinear term in the equation of the distance oscillation is proportional to the conserved energy of volume oscillator, and the nonlinear term in the equation of the volume oscillator is proportional to the conserved energy of the distance oscillator.

The conserved energy of the distance oscillator is identical, within the sign, to the constant curvature of space. This means that in the case of Euclidean space the distance oscillator has zero energy. When the space is pseudospherical or spherical (the spherical space is closed), the energy of the distance oscillator is positive or negative, respectively.

Generally, the frequency of the distance oscillation and the frequency of the volume oscillation vary with the time. For a sufficiently large volume both frequencies are constant and the distance oscillation frequency is one-third of the volume-oscillator freqeuncy.

2. FIELD EQUATIONS

The field equations[7] are

$$R_{ab} - \tfrac{1}{2}Rg_{ab} - \frac{\mathcal{K}^2}{h^2}\,g_{ab} = -\left(R + 4\frac{\mathcal{K}^2}{h^2}\right)(\tfrac{1}{2}g_{ab} - \dot{x}_a\dot{x}_b), \quad (1)$$

where $g_{ab} = g_{ba}$, $\mathcal{K}^2$ is the rest mass,[10] h is Planck's constant and the $\dot{x}^a$ are components of the unit normal 4-vector to the 3-surface of the de Broglie wave. On the left-hand side of (1) is a well-known tensor whose covariant divergence vanishes. Thus, from (1) we have the conservation equations

$$g^{bl}\left[\left(R + 4\frac{\mathcal{K}^2}{h^2}\right)(\tfrac{1}{2}g_{ab} - \dot{x}_a\dot{x}_b)\right]_{|l} = 0. \quad (2)$$

Because of

$$\dot{x}^a\dot{x}^b g_{ab} = 1, \quad (3)$$

we have from (2) that

$$\frac{\partial}{\partial x^a}\left[(-g)^{\frac{1}{2}}\left(R + 4\frac{\mathcal{K}^2}{h^2}\right)^{\frac{1}{2}}\dot{x}^a\right] = 0. \quad (4)$$

For any given set of four functions $\dot{x}^a(x^k)$, sufficiently smooth (for simplicity let us suppose them to be of class C^2) for which (3) is valid, the system (1) is a set of 10 nonlinear partial-differential equations to be satisfied by 10 unknowns g_{ab}. The four conservation equations (2) are a consequence of (1) and imply no restrictions on the chosen $\dot{x}^a(x^k)$ since they contain the unknowns g_{ab} not only in the coefficients but also in the derivatives. Since the given $\dot{x}^a(x^k)$ do not uniquely determine the coordinate system we have to add coordinate conditions. There are only three of these because of (3) which represents the fourth coordinate condition.

The Eqs. (1) do not contain sources; hence, we are dealing with continuous and nondualistic field theory. There is no reason to distinguish between exterior and interior solutions of Eqs. (1).

3. THE PARTICLE AT REST IN 3-SPACE

For the particle at rest in 3-space we will assume that the components of the unit 4-vector $\dot{x}^a(x^k)$ are $(0, 0, 0, 1)$. Since the 4-vector $\dot{x}^a$ is unitary, we have from (3) that

$$g_{44} = 1. \quad (5)$$

This equation is one coordinate condition. To solve the field equations, we need add any further three coordinate conditions.

Let us assume as these coordinate conditions

$$g_{\alpha 4} = 0, \quad \alpha = 1, 2, 3, \quad (6)$$

and hence we use the Gauss normal coordinates. Since the particle is at rest in the observer's 3-space, we will further assume that we have no reason to prefer any direction in 3-space. Consequently, we will consider that the geometry of space–time admits coordinates in which the first metric form is

$$ds^2 = (dx^0)^2 - U^2(x^0)\,d\omega^2, \quad (7)$$

where $U(x^0)$ is an arbitrary function and

$$d\omega^2 = h_{\alpha\beta}(x^1, x^2, x^3)\, dx^\alpha\, dx^\beta \tag{8}$$

defines the 3-space of constant curvature k in which $d\omega^2 > 0$ for any two neighboring points. Since the signature of (7) is -2, we will put $\mathcal{K}^2 = \varkappa^2$ and consider $\mathcal{K}$ as the rest mass of the particle.[10] From the field equations we now have

$$\begin{aligned} R_{\alpha\beta} + \frac{\mathcal{K}^2}{h^2}\, g_{\alpha\beta} &= 0, \\ R_{44} - R - 3\frac{\mathcal{K}^2}{h^2} &= 0, \\ \frac{d}{dx^0}(-g)^{\frac{1}{2}}\left(R + 4\frac{\mathcal{K}^2}{h^2}\right)^{\frac{1}{2}} &= 0, \end{aligned} \tag{9}$$

where

$$g_{\alpha\beta} = -U^2 h_{\alpha\beta}. \tag{10}$$

For the Ricci tensor and the scalar curvature R of space–time we get[11]

$$\begin{aligned} R_{\alpha\beta} &= -(U\ddot{U} + 2k + 2\dot{U}^2)h_{\alpha\beta}, \\ R_{44} &= 3\ddot{U}U^{-1}, \\ R &= 6(U\ddot{U} + \dot{U}^2 + k)U^{-2}. \end{aligned} \tag{11}$$

Substituting from (11) into (9) we find that the first two Eqs. (9) are identical and we obtain the following equation:

$$U\ddot{U} + 2\dot{U}^2 + \frac{\mathcal{K}^2}{h^2}\, U^2 + 2k = 0. \tag{12}$$

From the last of Eqs. (9) we get

$$\frac{d}{dx^0}(6)^{\frac{1}{2}}U^2\left(U\ddot{U} + \dot{U}^2 + \frac{2}{3}\frac{\mathcal{K}^2}{h^2}\, U^2 + k\right)^{\frac{1}{2}} = 0. \tag{13}$$

Thus, for the function U we have two simultaneous equations

$$\begin{aligned} U\ddot{U} + 2\dot{U}^2 + \frac{\mathcal{K}^2}{h^2}\, U^2 + 2k &= 0, \\ U^4\left(U\ddot{U} + \dot{U}^2 + \frac{2}{3}\frac{\mathcal{K}^2}{h^2}\, U^2 + k\right) &= \tfrac{1}{6}\beta^2, \end{aligned} \tag{14}$$

where β is a constant of integration. The compactibility condition of these equations is

$$\ddot{U} + \frac{1}{3}\frac{\mathcal{K}^2}{h^2}\, U = \frac{1}{3}\frac{\beta^2}{U^5}. \tag{15}$$

The nonlinear differential equation (15) represents an autonomous (not explicitly containing the independent variable), conservative (not containing a first derivative) oscillation. Hence, simultaneous solutions of Eqs. (14) are all solutions of equation

$$U^4\left(\dot{U}^2 + \frac{1}{3}\frac{\mathcal{K}^2}{h^2}\, U^2 + k\right) = -\tfrac{1}{6}\beta^2. \tag{16}$$

For further considerations the constant β^2 is no longer convenient. We will write $\beta^2 = -\frac{4}{3}W$ where W is a constant, the meaning of which will be discussed later.

4. TIME-DEPENDENT DISTANCE AND VOLUME OF SPACE

From the line element (7) it follows that the distance of two points in space is proportional to U and the volume of space is proportional to U^3. Let us denote $V \equiv U^3$. Then, we can transcribe Eq. (16) as

$$\tfrac{1}{2}\dot{V}^2 + \frac{3}{2}\frac{\mathcal{K}^2}{h^2}\, V + \tfrac{9}{2}kV^{\frac{4}{3}} = W. \tag{17}$$

Differentiation of (17) gives

$$\ddot{V} + 3\frac{\mathcal{K}^2}{h^2}\, V + 6kV^{\frac{1}{3}} = 0. \tag{18}$$

Equation (18) describes the volume oscillations. The constant W, introduced in connection with the integration of the conservation law (4) which, in our special case, is given by the last Eq. (9), is the energy of the oscillator described by Eq. (18). Now, if we rewrite Eq. (16) in the form

$$\tfrac{1}{2}\dot{U}^2 + \frac{1}{6}\frac{\mathcal{K}^2}{h^2}\, U^2 - \frac{1}{9}\frac{W}{U^4} = -\tfrac{1}{2}k \tag{19}$$

and differentiate it, we get

$$\ddot{U} + \frac{1}{3}\frac{\mathcal{K}^2}{h^2}\, U + \frac{4}{9}\frac{W}{U^5} = 0. \tag{20}$$

From Eq. (19), we see that the quantity $-\frac{1}{2}k$, where k is a constant curvature of space, is the energy of distance oscillator described by Eq. (20). Equation (18), which describes the volume oscillation, and Eq. (10), which describes the distance oscillations, are coupled. The nonlinear term in (18) depends on the curvature k, which essentially gives the energy $-\frac{1}{2}k$ of the distance oscillator (20). On the other hand, the nonlinear term in (20) depends on W, which is the energy volume oscillator (18).

The time varying frequency of the oscillator (18) is

$$\nu = \frac{3}{(2\sqrt{3})\pi}\left(\frac{\mathcal{K}^2}{h^2} + \frac{2k}{V^{\frac{2}{3}}}\right)^{\frac{1}{2}} \tag{21}$$

and the oscillator (20) has varying frequency

$$\tilde{\nu} = \frac{1}{(2\sqrt{3})\pi}\left(\frac{\mathcal{K}^2}{h^2} + \frac{4}{3}\frac{W}{V^2}\right)^{\frac{1}{2}}. \tag{22}$$

In Eqs. (21) and (22), as before, $V \equiv U^3$. When the volume V is very large, we learn from (22) and (21) that

$$\tilde{\nu} = \tfrac{1}{3}\nu. \tag{23}$$

The frequency of the distance oscillator is one-third the frequency of the volume oscillator. There are three cases when Eq. (16) can be integrated immediately by means of elementary functions.

As a first case let us consider $k = W = 0$. The energy of the distance and volume oscillator is zero and the space is Euclidean ($k = 0$). However, there is no real solution for U as a function of x^0. This violates our basic assumption that U is a real function of the time x^0. Let us remark that the frequency relation (23) is exactly valid in this case.

In the second case let us consider $k = 0$ and $W \neq 0$. The energy $-\frac{1}{2}k$ of the distance oscillator is zero and the space is Euclidean ($k = 0$). The energy W of a volume oscillator is a constant. For the time dependence of volume V, we obtain from Eq. (18)

$$V = A \sin\left(\sqrt{3}\,\frac{\mathcal{K}}{h}\,x^0 + B\right), \tag{24}$$

where A and B are constants of integration. We can write $W = \frac{3}{2}A^2(\mathcal{K}^2/h^2)$. The frequency (21) of the volume oscillator is a constant but the frequency of the distance oscillator varies with time.

The third case, $k \neq 0$ and $W = 0$. The energy W of the volume oscillator is zero. The energy $-\frac{1}{2}k$ of the distance oscillator is positive for $k < 0$ (pseudospherical space). In this case, we have

$$U = A \sin\left(\frac{1}{\sqrt{3}}\,\frac{\mathcal{K}}{h}\,x^0 + B\right), \tag{25}$$

where A and B are integrating constants. One easily verifies that

$$\frac{1}{6}\,\frac{\mathcal{K}^2}{h^2}\,A^2 = -\tfrac{1}{2}k. \tag{26}$$

Thus, for real U, we have necessarily $k < 0$ (pseudospherical space) and the energy $-\frac{1}{2}k$ of distance oscillator is positive. The frequency (21) of the volume oscillations varies with time and the frequency (22) of the distance oscillations is constant.

We observe that, under the cases introduced above, the space with positive curvature k, i.e., spherical space which is finite,does not appear.

To obtain the solution of Eq. (16) for $k \neq 0$ and $W \neq 0$, the variable x^0 is no longer convenient. Using new independent variable $d\tau = U^{-1}\,dx^0$ and writing $U^2 \equiv y$, we obtain from (16), after differentiation,

$$\frac{d^2y}{d\tau^2} + 4ky + 2\,\frac{\mathcal{K}^2}{h^2}\,y^2 = 0.$$

The foregoing equation has general solution given by Jacobi elliptic function.

5. KLEIN-GORDON EQUATION

Schrödinger[4] solved the familiar wave equation of the second order,

$$\frac{1}{\sqrt{-g}}\,\frac{\partial}{\partial x^a}\left(g^{ab}\sqrt{-g}\,\frac{\partial \psi}{\partial x^b}\right) + \mu^2\psi = 0, \tag{27}$$

where ψ is the wavefunction and $\mu = 2\pi m/h$ (m is rest mass and h is Planck's constant) for the line element (7). For the time-dependent part $f(x^0)$ of the wavefunction ψ, Schrödinger obtained the equation

$$\frac{d^2f}{d\tau^2} + [m(m+2)U^4 + \mu^2U^6]f = 0. \tag{28}$$

(The symbol U is used for the scale-factor function and not R as was used by Schrödinger.)

Here

$$d\tau = U^{-3}\,dx^0 \tag{29}$$

and m is a nonnegative integer, and $m(m+2)$ is a constant of separation. Equation (16) can be rewritten, with the help of (29), as

$$\left(\frac{dU}{d\tau}\right)^2 = -\frac{1}{3}\,\frac{\mathcal{K}^2}{h^2}\,U^8 - kU^6 + \tfrac{2}{9}WU^2 \tag{30}$$

and hence,

$$\frac{d^2U}{d\tau^2} + \left(\tfrac{2}{9}W + 3kU^4 + \frac{4}{3}\,\frac{\mathcal{K}^2}{h^2}\,U^6\right)U = 0. \tag{31}$$

Now, if we put in (31),

$$W = 0, \quad 3k = m(m+2), \quad \frac{4}{3}\,\frac{\mathcal{K}^2}{h^2} = \mu^2, \tag{32}$$

then from (31) and (28) we have that

$$\frac{d^2f}{d\tau^2}\,U - \frac{d^2U}{d\tau^2}\,f = 0. \tag{33}$$

The relation

$$f_1\,\frac{df_2}{d\tau} - f_2\,\frac{df_1}{d\tau} = \text{const} \tag{34}$$

holds for any two solutions f_1 and f_2 of Eq. (28) and thus, because of (33), U is a solution of (28). In this way we arrived at the result that if the scalar factor function U is given as a solution of Eq. (30) then, by replacing the constants therein and with the help of relations (32), we directly obtain the time-dependent part of the wavefunction ψ.

CONCLUSION

It is well known that the 3-space, with $k > 0$ is an analog of the surface of an ordinary sphere. The

vibrations of a spherical surface[12] are characterized by discrete eigenfrequencies. The discreteness is a consequence of the fitting of wavelengths into a finite span.

The vibrations of 3-space with $k > 0$, which are considered in this paper, are similar to a pulsating sphere. This particular motion of a spherical surface is characterized by eigenfrequency which is zero.[12] The assumed isotropy of 3-space ruled out the higher eigenfrequencies. However, it is possible that the eigenfrequencies of nonisotropic and finite 3-space might provide an adequate description of the observed atomicity of matter and light.

[1] E. Schrödinger, Proc. Roy. Irish Acad. **46A**, 25 (1940).
[2] B. C. Mukerji, Z. Physik **101**, 270 (1936).
[3] E. Schrödinger, Commentatones Pontificiae Acad. Sci. **2**, 321 (1938).
[4] E. Schrödinger, Physica **6**, 899 (1939).
[5] P. O. Müller, Physik. Z. **40**, 366 (1939).
[6] L. Infeld and A. E. Shild, Phys. Rev. **70**, 410 (1946).
[7] J. Kulhánek, Nuovo Cimento Supp. **4**, 172 (1966).
[8] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1962), Vol. II.
[9] J. Kulhánek, Nuovo Cimento **63B**, 497 (1969).
[10] Explanatory footnote: The constant $\mathcal{K}$ is defined as $\mathcal{K}^2 = g^{\alpha\delta}\gamma_\alpha\gamma_\delta$ where γ_a are components of 4-momentum. When the signature of first metric form is $+2$, we put $\mathcal{K}^2 = -\mathcal{K}^2$ and consider $\mathcal{K}$ as the rest mass. When the signature is -2, we put $\mathcal{K}^2 = \mathcal{K}^2$ and consider $\mathcal{K}$ as the rest mass.
[11] H. P. Robertson, Rev. Mod. Phys. **5**, 62 (1933).
[12] E. J. Konopinski, *Classical Descriptions of Motion* (Freeman, San Francisco, 1969), p. 407.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Equations of the de Broglie Wave Field and Their Relationship to Riemann's Curvature Tensor

J. KULHÁNEK
Physics Department, University of Windsor, Windsor, Ontario, Canada

(Received 10 April 1970)

The equations of the de Broglie wave field (field equations) [J. Kulhanek, Nuovo Cimento Supp. **4**, 172 (1966)] under special conditions require a very particular geometry together with a specific interpretation of the curvature scalar. The purpose of the present paper is to show that the same condition turns the conservation law (which is a consequence of the field equations) into an identity and that the Rainich [Nature **115**, 498 (1925)] decomposition of Riemann's curvature tensor gives only one component.

A SPECIAL CASE OF THE FIELD EQUATIONS

The general form of the field equations[1] is

$$R_{ab} - \tfrac{1}{2}Rg_{ab} - \frac{\mathcal{K}^2}{h^2}\,g_{ab} = -\left(R + 4\frac{\mathcal{K}^2}{h^2}\right)(\tfrac{1}{2}g_{ab} - \dot{x}_a\dot{x}_b), \quad (1)$$

where $g_{ab} = g_{ba}$, $\mathcal{K}^2$ is the rest mass, h is Planck's constant, and the $\dot{x}^a$ are components of the unit 4-vector normal to the 3-wavesurface of the de Broglie wave. On the left-hand side of (1) is a well-known tensor, whose covariant divergence vanishes. Thus, from (1) we have the conservation equations,

$$g^{bl}\left[\left(R + 4\frac{\mathcal{K}^2}{h^2}\right)(\tfrac{1}{2}g_{ab} - \dot{x}_a\dot{x}_b)\right]_{|l} = 0. \quad (2)$$

It is well known that the quantity $R_{ik}\xi^i\xi^k$ is the scalar curvature of a 3-dimensional space which is perpendicular to ξ^a. If we put $\xi^a \equiv \dot{x}^a$, then from (1) it follows that

$$R_{ab}\dot{x}^a\dot{x}^b = R + 3(\mathcal{K}^2/h^2). \quad (3)$$

Thus quantity $R + 3(\mathcal{K}^2/h^2)$ represents scalar curvature of the 3-wavesurface of the de Broglie wave.

In the case when we assume that

$$R + 4(\mathcal{K}^2/h^2) = 0, \quad (4)$$

then the conservation law (2) is trivial and Eqs. (1) are reduced to

$$R_{ab} - \tfrac{1}{4}Rg_{ab} = 0. \quad (5)$$

From Eqs. (3) and (4) we have that

$$R_{ab}\dot{x}^a\dot{x}^b = -(\mathcal{K}^2/h^2) \quad (6)$$

or

$$R_{ab}\dot{x}^a\dot{x}^b = \tfrac{1}{4}R. \quad (7)$$

The scalar curvature of the 3-wavesurface of the de Broglie wave is given as $-\mathcal{K}^2/h^2$ or $\frac{1}{4}R$.

Rainich[2] showed that Riemann's curvature tensor $R_{ab\,ek}$ in the 4-dimensional space can be decomposed into two parts which have different properties of symmetry. Considering the Riemannian curvature in the 2-direction defined by the unit bivector V^{ab} and

vibrations of a spherical surface[12] are characterized by discrete eigenfrequencies. The discreteness is a consequence of the fitting of wavelengths into a finite span.

The vibrations of 3-space with $k > 0$, which are considered in this paper, are similar to a pulsating sphere. This particular motion of a spherical surface is characterized by eigenfrequency which is zero.[12] The assumed isotropy of 3-space ruled out the higher eigenfrequencies. However, it is possible that the eigenfrequencies of nonisotropic and finite 3-space might provide an adequate description of the observed atomicity of matter and light.

[1] E. Schrödinger, Proc. Roy. Irish Acad. **46A**, 25 (1940).
[2] B. C. Mukerji, Z. Physik **101**, 270 (1936).
[3] E. Schrödinger, Commentatones Pontificiae Acad. Sci. **2**, 321 (1938).
[4] E. Schrödinger, Physica **6**, 899 (1939).
[5] P. O. Müller, Physik. Z. **40**, 366 (1939).
[6] L. Infeld and A. E. Shild, Phys. Rev. **70**, 410 (1946).
[7] J. Kulhánek, Nuovo Cimento Supp. **4**, 172 (1966).
[8] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1962), Vol. II.
[9] J. Kulhánek, Nuovo Cimento **63B**, 497 (1969).
[10] Explanatory footnote: The constant $\mathcal{K}$ is defined as $\mathcal{K}^2 = g^{\alpha\delta}\gamma_\alpha\gamma_\delta$ where γ_a are components of 4-momentum. When the signature of first metric form is $+2$, we put $\mathcal{K}^2 = -\mathcal{K}^2$ and consider $\mathcal{K}$ as the rest mass. When the signature is -2, we put $\mathcal{K}^2 = \mathcal{K}^2$ and consider $\mathcal{K}$ as the rest mass.
[11] H. P. Robertson, Rev. Mod. Phys. **5**, 62 (1933).
[12] E. J. Konopinski, *Classical Descriptions of Motion* (Freeman, San Francisco, 1969), p. 407.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Equations of the de Broglie Wave Field and Their Relationship to Riemann's Curvature Tensor

J. KULHÁNEK
Physics Department, University of Windsor, Windsor, Ontario, Canada

(Received 10 April 1970)

The equations of the de Broglie wave field (field equations) [J. Kulhanek, Nuovo Cimento Supp. **4**, 172 (1966)] under special conditions require a very particular geometry together with a specific interpretation of the curvature scalar. The purpose of the present paper is to show that the same condition turns the conservation law (which is a consequence of the field equations) into an identity and that the Rainich [Nature **115**, 498 (1925)] decomposition of Riemann's curvature tensor gives only one component.

A SPECIAL CASE OF THE FIELD EQUATIONS

The general form of the field equations[1] is

$$R_{ab} - \tfrac{1}{2}Rg_{ab} - \frac{\mathcal{K}^2}{h^2}\,g_{ab} = -\left(R + 4\frac{\mathcal{K}^2}{h^2}\right)(\tfrac{1}{2}g_{ab} - \dot{x}_a\dot{x}_b), \quad (1)$$

where $g_{ab} = g_{ba}$, $\mathcal{K}^2$ is the rest mass, h is Planck's constant, and the $\dot{x}^a$ are components of the unit 4-vector normal to the 3-wavesurface of the de Broglie wave. On the left-hand side of (1) is a well-known tensor, whose covariant divergence vanishes. Thus, from (1) we have the conservation equations,

$$g^{bl}\left[\left(R + 4\frac{\mathcal{K}^2}{h^2}\right)(\tfrac{1}{2}g_{ab} - \dot{x}_a\dot{x}_b)\right]_{|l} = 0. \quad (2)$$

It is well known that the quantity $R_{ik}\xi^i\xi^k$ is the scalar curvature of a 3-dimensional space which is perpendicular to ξ^a. If we put $\xi^a \equiv \dot{x}^a$, then from (1) it follows that

$$R_{ab}\dot{x}^a\dot{x}^b = R + 3(\mathcal{K}^2/h^2). \quad (3)$$

Thus quantity $R + 3(\mathcal{K}^2/h^2)$ represents scalar curvature of the 3-wavesurface of the de Broglie wave.

In the case when we assume that

$$R + 4(\mathcal{K}^2/h^2) = 0, \quad (4)$$

then the conservation law (2) is trivial and Eqs. (1) are reduced to

$$R_{ab} - \tfrac{1}{4}Rg_{ab} = 0. \quad (5)$$

From Eqs. (3) and (4) we have that

$$R_{ab}\dot{x}^a\dot{x}^b = -(\mathcal{K}^2/h^2) \quad (6)$$

or

$$R_{ab}\dot{x}^a\dot{x}^b = \tfrac{1}{4}R. \quad (7)$$

The scalar curvature of the 3-wavesurface of the de Broglie wave is given as $-\mathcal{K}^2/h^2$ or $\frac{1}{4}R$.

Rainich[2] showed that Riemann's curvature tensor $R_{ab\,ek}$ in the 4-dimensional space can be decomposed into two parts which have different properties of symmetry. Considering the Riemannian curvature in the 2-direction defined by the unit bivector V^{ab} and

the Riemannian curvature in the perpendicular 2-direction defined by the unit bivector $\bar{V}^{ab}$, we can then write[2]

$$R_{ik\ lm} = S_{ik\ lm} + A_{ik\ lm}. \tag{8}$$

For the 2-directions V^{ab} and $\bar{V}^{ab}$ the component $S_{ik\ lm}$ defines the same sign curvature while $A_{ik\ lm}$ defines curvature of the opposite sign. We can write that

$$S_{ik\ lm}V^{ik}V^{lm} = S_{ik\ lm}\bar{V}^{ik}\bar{V}^{lm} \tag{9}$$

and

$$A_{ik\ lm}V^{ik}V^{lm} = -A_{ik\ lm}\bar{V}^{ik}\bar{V}^{lm}. \tag{10}$$

Now, in Ref. 2 it is shown that

$$A_{ik\ lm} = \tfrac{1}{2}(g_{il}K_{km} + g_{km}K_{il} - g_{im}K_{kl} - g_{kl}K_{im}), \tag{11}$$

where we denote

$$K_{lm} \equiv R_{lm} - \tfrac{1}{4}Rg_{lm}. \tag{12}$$

From (11), (12), and (5) it follows that

$$A_{ik\ lm} = 0. \tag{13}$$

We see from (11) that (13) and (5) are equivalent. The field equations (1) with condition (2) imply that Riemann's tensor is equal to the component $S_{ik\ lm}$ of the Rainich decomposition (8).

[1] J. Kulhánek, Nuovo Cimento Suppl. **4**, 172 (1966).
[2] C. Y. Rainich, Nature **115**, 498 (1925).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Calculation of Correlation Functions of Solutions of a Stochastic Ordinary Differential Equation

J. A. MORRISON
Bell Telephone Laboratories, Inc., Murray Hill, New Jersey 07974

(Received 27 May 1970)

In this paper we use the "smoothing method" to calculate the correlation functions of the solutions of the equation

$$\frac{d^2u}{dz^2} + \beta_0^2[1 + \eta N(z)]u = 0,$$

satisfying nonstochastic initial conditions, where $N(z)$ is a real, wide-sense stationary stochastic process with zero mean and β_0 and $\eta \ll 1$ are positive constants. It is shown that an appropriate application of the smoothing method leads to the exact results in the case when $N(z)$ is the random telegraph process. Moreover, under appropriate conditions on the general process $N(z)$, approximate expressions are obtained for the correlation functions in terms of the first- and second-order moments of the solutions, and approximate expressions are given for these moments.

1. INTRODUCTION AND SUMMARY

In a previous paper[1] we used the "smoothing method" to calculate approximately the first- and second-order moments of the solutions of the stochastic differential equation (1.2). Our interest was in the propagation of an electromagnetic wave through a randomly stratified dielectric slab, but Eq. (1.2), in which the stochastic process $N(z)$ satisfies (1.1), also corresponds to a harmonic oscillator with a random spring and arises in many other contexts. In this paper we apply the smoothing method to calculate approximately the correlation functions of the solutions of (1.2). We also show that the results are in fact exact in the particular case in which $N(z)$ is the random telegraph process $T(z)$, defined in Sec. 3. We derived the exact results in this case in an earlier paper.[2]

The smoothing method[3] for calculating the expected value of the solution of a linear stochastic equation has been developed extensively by Keller[4–10] and Bourret.[11–16] In Sec. 2 we give an outline of the method in a form which is appropriate for our purposes, and is close to that given by Keller. However, our application of the smoothing approximation to the calculation of the correlation functions of the solutions of Eq. (1.2) differs considerably from that proposed by Keller.[8–10] In our approach it is necessary to solve only ordinary differential and integro-differential equations, whereas the scheme proposed by Keller leads to partial differential and integro-differential equations.

Let $N(z)$ be a real, wide-sense stationary stochastic process with

$$\langle N(z)\rangle = 0, \quad \langle N(y)N(z)\rangle = \Gamma(y - z), \tag{1.1}$$

the Riemannian curvature in the perpendicular 2-direction defined by the unit bivector $\bar{V}^{ab}$, we can then write[2]

$$R_{ik\ lm} = S_{ik\ lm} + A_{ik\ lm}. \tag{8}$$

For the 2-directions V^{ab} and $\bar{V}^{ab}$ the component $S_{ik\ lm}$ defines the same sign curvature while $A_{ik\ lm}$ defines curvature of the opposite sign. We can write that

$$S_{ik\ lm}V^{ik}V^{lm} = S_{ik\ lm}\bar{V}^{ik}\bar{V}^{lm} \tag{9}$$

and

$$A_{ik\ lm}V^{ik}V^{lm} = -A_{ik\ lm}\bar{V}^{ik}\bar{V}^{lm}. \tag{10}$$

Now, in Ref. 2 it is shown that

$$A_{ik\ lm} = \tfrac{1}{2}(g_{il}K_{km} + g_{km}K_{il} - g_{im}K_{kl} - g_{kl}K_{im}), \tag{11}$$

where we denote

$$K_{lm} \equiv R_{lm} - \tfrac{1}{4}Rg_{lm}. \tag{12}$$

From (11), (12), and (5) it follows that

$$A_{ik\ lm} = 0. \tag{13}$$

We see from (11) that (13) and (5) are equivalent. The field equations (1) with condition (2) imply that Riemann's tensor is equal to the component $S_{ik\ lm}$ of the Rainich decomposition (8).

[1] J. Kulhánek, Nuovo Cimento Suppl. **4**, 172 (1966).
[2] C. Y. Rainich, Nature **115**, 498 (1925).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Calculation of Correlation Functions of Solutions of a Stochastic Ordinary Differential Equation

J. A. MORRISON
Bell Telephone Laboratories, Inc., Murray Hill, New Jersey 07974

(Received 27 May 1970)

In this paper we use the "smoothing method" to calculate the correlation functions of the solutions of the equation

$$\frac{d^2u}{dz^2} + \beta_0^2[1 + \eta N(z)]u = 0,$$

satisfying nonstochastic initial conditions, where $N(z)$ is a real, wide-sense stationary stochastic process with zero mean and β_0 and $\eta \ll 1$ are positive constants. It is shown that an appropriate application of the smoothing method leads to the exact results in the case when $N(z)$ is the random telegraph process. Moreover, under appropriate conditions on the general process $N(z)$, approximate expressions are obtained for the correlation functions in terms of the first- and second-order moments of the solutions, and approximate expressions are given for these moments.

1. INTRODUCTION AND SUMMARY

In a previous paper[1] we used the "smoothing method" to calculate approximately the first- and second-order moments of the solutions of the stochastic differential equation (1.2). Our interest was in the propagation of an electromagnetic wave through a randomly stratified dielectric slab, but Eq. (1.2), in which the stochastic process $N(z)$ satisfies (1.1), also corresponds to a harmonic oscillator with a random spring and arises in many other contexts. In this paper we apply the smoothing method to calculate approximately the correlation functions of the solutions of (1.2). We also show that the results are in fact exact in the particular case in which $N(z)$ is the random telegraph process $T(z)$, defined in Sec. 3. We derived the exact results in this case in an earlier paper.[2]

The smoothing method[3] for calculating the expected value of the solution of a linear stochastic equation has been developed extensively by Keller[4–10] and Bourret.[11–16] In Sec. 2 we give an outline of the method in a form which is appropriate for our purposes, and is close to that given by Keller. However, our application of the smoothing approximation to the calculation of the correlation functions of the solutions of Eq. (1.2) differs considerably from that proposed by Keller.[8–10] In our approach it is necessary to solve only ordinary differential and integro-differential equations, whereas the scheme proposed by Keller leads to partial differential and integro-differential equations.

Let $N(z)$ be a real, wide-sense stationary stochastic process with

$$\langle N(z)\rangle = 0, \quad \langle N(y)N(z)\rangle = \Gamma(y - z), \tag{1.1}$$

where $\langle\ \rangle$ denotes the ensemble average. Each sample function $N(z)$ defines two real functions $u_m(z)$, $m = 1, 2$, on $0 \leq z < \infty$ which are the linearly independent solutions of

$$\frac{d^2u_m}{dz^2} + \beta_0^2[1 + \eta N(z)]u_m = 0, \tag{1.2}$$

satisfying the initial conditions

$$u_1(0) = 1 = u_2'(0), \quad u_1'(0) = 0 = u_2(0), \tag{1.3}$$

where β_0 and $\eta \ll 1$ are positive constants. The ensemble of functions $\{u_m(z)\}$, $m = 1, 2$, forms two real random processes. We assume that almost all the sample functions $N(z)$ are smooth enough so that the solutions $u_m(z)$ of (1.2) exist. Throughout the paper we adopt the notation

$$v_m(z) = \frac{du_m}{dz} = u_m'(z), \quad m = 1, 2. \tag{1.4}$$

We introduce the matrix of solutions

$$\mathbf{F}(z) = \begin{bmatrix} u_1(z) & u_2(z) \\ v_1(z) & v_2(z) \end{bmatrix}, \tag{1.5}$$

with initial conditions, from (1.3),

$$\mathbf{F}(0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}. \tag{1.6}$$

The procedure we adopt is to write down the equation satisfied by the Kronecker product[17]

$$\mathbf{W}(z, \zeta) = \mathbf{F}(z + \zeta) \times \mathbf{F}(z), \quad \zeta \geq 0, \tag{1.7}$$

regarded as a function of ζ. [$\mathbf{A} \times \mathbf{B} = (a_{ij}) \times \mathbf{B} = (a_{ij}\mathbf{B})$.] The initial condition is

$$\mathbf{W}(z, 0) = \mathbf{F}(z) \times \mathbf{F}(z), \tag{1.8}$$

and hence is stochastic for $z > 0$. Thus, in the application of the smoothing method to the equation for $\mathbf{W}(z, \zeta)$, a knowledge of the incoherent part of $\mathbf{W}(z, 0)$ is required. But, by application of the smoothing approximation to the equation satisfied by $\mathbf{W}(z, 0)$, this quantity may be represented as an integral involving $\langle \mathbf{F}(z) \times \mathbf{F}(z) \rangle$, which we have already calculated by the smoothing method.[1]

The equation obtained for $\langle \mathbf{W}(z, \zeta) \rangle$ is solved by means of Laplace transforms, and the case in which $\Gamma(z) = \exp(-2b|z|)$ is considered in detail in Sec. 3. It is shown that the results are exact in the case $N(z) = T(z)$, where $T(z)$ is the random telegraph process, so that η is not restricted to be small in this particular case. In showing this, more explicit expressions than those which we gave previously[2] are obtained for the Laplace transforms of some of the quantities that arise. We previously[1] showed that, when $N(z) = T(z)$, the smoothing method gives the exact results for the first- and second-order moments of the solutions of (1.2), thus verifying the indirect proof given by Bourret.[15,16] However, it is not clear that Bourret's proof can be extended to the correlation functions. In fact, in Sec. 5 we consider another application of the smoothing method to the calculation of the correlation functions which does not appear to give the exact result in the random telegraph case.

In Sec. 4 the general case is considered. It is assumed that $\gamma(\sigma)$, the Laplace transform of $\Gamma(\zeta)$, as defined by (3.5) and (4.7), is analytic for Re $(\sigma) \geq -a$, where $a > 0$ is independent of η. Then, it is shown that

$$\langle \mathbf{F}(z + \zeta) \times \mathbf{F}(z) \rangle \approx [\langle \mathbf{F}(\zeta) \rangle \times \mathbf{I}] \langle \mathbf{F}(z) \times \mathbf{F}(z) \rangle, \tag{1.9}$$

on the intervals $0 \leq \eta^2\beta_0\zeta \leq \Xi$ and $0 \leq \eta^2\beta_0 z \leq Z$, where Ξ and Z are $O(1)$. Moreover, from the expressions for the Laplace transforms of the first- and second-order moments, which we obtained previously[1] by the smoothing method, we derive approximate expressions for the quantities on the right-hand side of (1.9) [see (4.8)–(4.17)]. These expressions are shown in Appendix D to be consistent with those derived from the results of Papanicolaou and Keller,[18] who applied a 2-variable method to calculate the first- and second-order moments of the solutions of (1.2). The approximations to the first- and second-order moments, and the approximation (1.9) to the correlation functions, may also be derived from earlier results of Khas'minskii[19] and Stratonovich,[20] as we will show in another paper.

2. FORMULATION AND APPLICATION OF THE SMOOTHING METHOD

We first give a formulation of the smoothing method, in the context in which it is needed here. Thus, consider a linear stochastic equation for the matrix $\mathbf{U}$,

$$(\mathbf{L}_0 + \eta\mathbf{L}_1)\mathbf{U} = 0, \tag{2.1}$$

where $\mathbf{L}_0$ is a nonstochastic matrix differential operator, but $\mathbf{L}_1$ is a stochastic one, and η is a small positive parameter. It is supposed that the stochastic average of $\mathbf{L}_1$ is zero, that is,

$$\langle \mathbf{L}_1 \rangle = 0. \tag{2.2}$$

We define the incoherent part of $\mathbf{U}$ as

$$\subset \mathbf{U} \supset = \mathbf{U} - \langle \mathbf{U} \rangle, \tag{2.3}$$

so that

$$\langle \subset \mathbf{U} \supset \rangle = 0. \tag{2.4}$$

If we substitute for $\mathbf{U}$ from (2.3) into (2.1) and take the stochastic average, we obtain

$$\mathbf{L}_0\langle\mathbf{U}\rangle + \eta\langle\mathbf{L}_1\subset\mathbf{U}\supset\rangle = 0. \tag{2.5}$$

We next subtract (2.5) from (2.1) and obtain

$$\mathbf{L}_0\subset\mathbf{U}\supset + \eta\mathbf{L}_1\langle\mathbf{U}\rangle + \eta[\mathbf{L}_1\subset\mathbf{U}\supset - \langle\mathbf{L}_1\subset\mathbf{U}\supset\rangle] = 0. \tag{2.6}$$

So far no approximations have been made. However, neglecting terms of order η^2 in (2.6), we obtain the smoothing approximation

$$\mathbf{L}_0\subset\mathbf{U}\supset = -\eta\mathbf{L}_1\langle\mathbf{U}\rangle. \tag{2.7}$$

If the inverse operator $\mathbf{L}_0^{-1}$ is applied to this equation and the resulting expression for $\subset\mathbf{U}\supset$ is substituted into (2.5), the equation for $\langle\mathbf{U}\rangle$ as given by Keller[5] is obtained, with terms of order η^3 and higher being neglected. However, the above form is more convenient for our purposes.

We proceed to derive an equation for the quantity $\mathbf{W}(z, \zeta)$ defined in (1.7). We first obtain an equation for $\mathbf{F}(z)$ from (1.2), (1.4), and (1.5). Let

$$\mathbf{A} = \begin{bmatrix} 0 & -1 \\ \beta_0^2 & 0 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}. \tag{2.8}$$

Then,

$$\frac{d\mathbf{F}}{dz} + [\mathbf{A} + \eta\beta_0^2 N(z)\mathbf{C}]\mathbf{F}(z) = 0. \tag{2.9}$$

It follows from (2.9) that, for $\zeta > 0$,

$$\frac{\partial}{\partial\zeta}\mathbf{F}(z + \zeta) + [\mathbf{A} + \eta\beta_0^2 N(z + \zeta)\mathbf{C}]\mathbf{F}(z + \zeta) = 0. \tag{2.10}$$

We now take the Kronecker product of the expression in (2.10) with $\mathbf{F}(z)$ and obtain, from (1.7),

$$\frac{\partial\mathbf{W}}{\partial\zeta} + \{[\mathbf{A} + \eta\beta_0^2 N(z + \zeta)\mathbf{C}] \times \mathbf{I}\}\mathbf{W}(z, \zeta) = 0, \quad \zeta > 0. \tag{2.11}$$

The initial value $\mathbf{W}(z, 0)$ is given by (1.8).

The equation for $\mathbf{W}$ has the form of the equation for $\mathbf{U}$ in (2.1), with

$$\mathbf{L}_0 = \left(\mathbf{I}\frac{\partial}{\partial\zeta} + \mathbf{A}\right) \times \mathbf{I}, \quad \mathbf{L}_1 = \beta_0^2 N(z + \zeta)(\mathbf{C} \times \mathbf{I}). \tag{2.12}$$

Note that (2.2) is satisfied by virtue of (1.1). Thus, in the smoothing approximation, from (2.7), and from (1.8),

$$\mathbf{L}_0\subset\mathbf{W}\supset = -\eta\mathbf{L}_1\langle\mathbf{W}\rangle,$$
$$\subset\mathbf{W}(z, 0)\supset = \subset\mathbf{F}(z) \times \mathbf{F}(z)\supset. \tag{2.13}$$

Also, from (2.5) and (1.8),

$$\mathbf{L}_0\langle\mathbf{W}\rangle = -\eta\langle\mathbf{L}_1\subset\mathbf{W}\supset\rangle, \quad \langle\mathbf{W}(z, 0)\rangle = \langle\mathbf{F}(z) \times \mathbf{F}(z)\rangle. \tag{2.14}$$

Now let

$$\mathbf{P}(\zeta) = \begin{bmatrix} \cos\beta_0\zeta & \frac{1}{\beta_0}\sin\beta_0\zeta \\ -\beta_0\sin\beta_0\zeta & \cos\beta_0\zeta \end{bmatrix}. \tag{2.15}$$

It may be verified, using (2.8), that

$$\frac{d\mathbf{P}}{d\zeta} + \mathbf{A}\mathbf{P} = 0, \quad \mathbf{P}(0) = \mathbf{I}, \tag{2.16}$$

so that $\mathbf{P}$ is a fundamental matrix. Thus, with $\mathbf{L}_0$ given by (2.12),

$$\mathbf{L}_0[\mathbf{P}(\zeta) \times \mathbf{I}] = 0, \quad \mathbf{P}(0) \times \mathbf{I} = \mathbf{I} \times \mathbf{I}. \tag{2.17}$$

Hence, from (2.12) and (2.13),

$$\subset\mathbf{W}(z, \zeta)\supset = [\mathbf{P}(\zeta) \times \mathbf{I}]\subset\mathbf{F}(z) \times \mathbf{F}(z)\supset - \eta\beta_0^2\int_0^\zeta [\mathbf{P}(\zeta - \xi) \times \mathbf{I}]N(z + \xi)(\mathbf{C} \times \mathbf{I}) \times \langle\mathbf{W}(z, \xi)\rangle\, d\xi. \tag{2.18}$$

Before writing down an equation for $\langle\mathbf{W}\rangle$, we turn to the calculation of $\subset\mathbf{F}(z) \times \mathbf{F}(z)\supset$.

Now, from (2.9) it follows that

$$\frac{d}{dz}[\mathbf{F}(z) \times \mathbf{F}(z)] + [(\mathbf{A} \times \mathbf{I}) + (\mathbf{I} \times \mathbf{A})][\mathbf{F}(z) \times \mathbf{F}(z)] + \eta\beta_0^2 N(z)[(\mathbf{C} \times \mathbf{I}) + (\mathbf{I} \times \mathbf{C})][\mathbf{F}(z) \times \mathbf{F}(z)] = 0. \tag{2.19}$$

But previously[1] we analyzed this equation by the smoothing method and derived an expression for the Laplace transform of $\langle\mathbf{F}(z) \times \mathbf{F}(z)\rangle$. Let

$$\mathbf{\Psi}(z) = [\mathbf{P}(z) \times \mathbf{P}(z)][(\mathbf{C} \times \mathbf{I}) + (\mathbf{I} \times \mathbf{C})]. \tag{2.20}$$

Then, from the equation corresponding to (2.7), since $\subset\mathbf{F}(0) \times \mathbf{F}(0)\supset = 0$ from (1.6), we obtain, from (2.16) and (2.19),

$$\subset\mathbf{F}(z) \times \mathbf{F}(z)\supset = -\eta\beta_0^2\int_0^z \mathbf{\Psi}(z - \xi)N(\xi)\langle\mathbf{F}(\xi) \times \mathbf{F}(\xi)\rangle\, d\xi. \tag{2.21}$$

Thus, from (1.1), (2.12), (2.14), (2.18), and (2.21), we have

$$\frac{\partial\langle\mathbf{W}\rangle}{\partial\zeta} + (\mathbf{A} \times \mathbf{I})\langle\mathbf{W}(z, \zeta)\rangle = \eta^2\beta_0^4\int_0^\zeta \Gamma(\zeta - \xi)\{[\mathbf{C}\mathbf{P}(\zeta - \xi)\mathbf{C}] \times \mathbf{I}\}\langle\mathbf{W}(z, \xi)\rangle\, d\xi + \eta^2\beta_0^4\{[\mathbf{C}\mathbf{P}(\zeta)] \times \mathbf{I}\}\int_0^z \Gamma(z + \zeta - \xi)\mathbf{\Psi}(z - \xi) \times \langle\mathbf{F}(\xi) \times \mathbf{F}(\xi)\rangle\, d\xi. \tag{2.22}$$

The initial condition is as in (2.14). We consider the solution of Eq. (2.22) in the next two sections.

3. PROCESS WITH AN EXPONENTIAL CORRELATION FUNCTION

In this section we consider the particular case in which

$$\Gamma(z) = e^{-2b|z|}. \tag{3.1}$$

In an earlier paper[2] we obtained exact solutions for the first- and second-order moments, and the correlation functions, of the solutions of (1.2), in the case in which $N(z) = T(z)$, where $T(z)$ is the random telegraph process, defined in the following way.[21] A given function of the ensemble $\{T(z)\}$ can assume only the values ± 1, and as a function of z it makes independent random traversals from one value to the other. For fixed z, a sample function chosen at random will equal 1 or -1 with probability 1/2. The probability that a given sample function makes n traversals in an interval of length z is given by the Poisson distribution

$$p(n, z) = \frac{(bz)^n}{n!} e^{-bz}, \quad n = 0, 1, 2, \cdots, \tag{3.2}$$

where b is the average number of traversals per unit length. A straightforward calculation yields[21]

$$\langle T(z)\rangle = 0, \quad \langle T(y)T(z)\rangle = e^{-2b|y-z|}. \tag{3.3}$$

We showed[1] that the smoothing method leads to the exact results for the first- and second-order moments of the solutions of (1.2), subject to (1.3), when $N(z) = T(z)$. Here we show that the smoothing method, as outlined in the previous section, leads to the exact results also for the correlation functions, in this case. We relegate the details to Appendix A, but give the results in this section in a more explicit form than we gave earlier. Previously[2] we obtained, in the case $N(z) = T(z)$, the exact result

$$\langle \mathbf{F}(z+\zeta) \times \mathbf{F}(z)\rangle = \tfrac{1}{2}\sum_{k=1}^{2}\left[\left(\sum_{j=1}^{2}\langle \mathbf{\Phi}(\zeta) \mid 0\rangle_{jk} \times \mathbf{I}\right)\langle \mathbf{F}(z) \times \mathbf{F}(z)\rangle_k\right]. \tag{3.4}$$

The reader is referred to Ref. 2 for the meanings of the functions occurring in (3.4).

Expressions were given for the Laplace transform of $\langle \mathbf{\Phi}(\zeta) \mid 0\rangle_{jk}$. We define

$$\Lambda(\mathbf{H}) = \int_0^\infty e^{-\sigma\zeta}\mathbf{H}(\zeta)\, d\zeta. \tag{3.5}$$

Let

$$d(\sigma) = \{(\sigma^2 + \beta_0^2)[(\sigma + 2b)^2 + \beta_0^2] - \eta^2\beta_0^4\}. \tag{3.6}$$

Then

$$\Lambda\left(\sum_{k=1}^{2}\sum_{j=1}^{2}\langle \mathbf{\Phi} \mid 0\rangle_{jk}\right) = \frac{2}{d(\sigma)}\begin{pmatrix} \sigma[(\sigma+2b)^2+\beta_0^2] & [(\sigma+2b)^2+\beta_0^2] \\ -\beta_0^2\{[(\sigma+2b)^2+\beta_0^2] - \eta^2\beta_0^2\} & \sigma[(\sigma+2b)^2+\beta_0^2]\end{pmatrix}, \tag{3.7}$$

and

$$\Lambda\left(\sum_{k=1}^{2}(-1)^k\sum_{j=1}^{2}\langle \mathbf{\Phi} \mid 0\rangle_{jk}\right) = \frac{2\eta\beta_0^2}{d(\sigma)}\begin{pmatrix} (\sigma+2b) & 1 \\ \sigma(\sigma+2b) & \sigma\end{pmatrix}. \tag{3.8}$$

We remark that[2]

$$\langle \mathbf{F}(\zeta)\rangle = \tfrac{1}{2}\sum_{j=1}^{2}\sum_{k=1}^{2}\langle \mathbf{\Phi}(\zeta) \mid 0\rangle_{jk}. \tag{3.9}$$

Expressions were also given for the Laplace transforms of the elements of $\langle \mathbf{F}(z) \times \mathbf{F}(z)\rangle$, and we remark that[2]

$$\tfrac{1}{2}\sum_{k=1}^{2}\langle \mathbf{F}(z) \times F(z)\rangle_k = \langle F(z) \times F(z)\rangle. \tag{3.10}$$

We define

$$\mathfrak{L}(\mathbf{G}) = \int_0^\infty e^{-sz}\mathbf{G}(z)\, dz. \tag{3.11}$$

Let

$$\Delta = \Delta(s) \equiv \{s(s+2b)(s^2+4\beta_0^2)[(s+2b)^2+4\beta_0^2] - 16\eta^2\beta_0^4(s+b)^2\}. \tag{3.12}$$

Then[2]

$$\mathfrak{L}(\langle u_1^2\rangle) = (1/\Delta)\{(s+2b)(s^2+2\beta_0^2)[(s+2b)^2+4\beta_0^2] - 8\eta^2\beta_0^4(s+b)\}, \tag{3.13}$$

$$\mathfrak{L}(\langle v_1^2\rangle) = (2/\Delta)\beta_0^4(s+2b)\{[(s+2b)^2+4\beta_0^2] + \eta^2[s(s+2b) - 4\beta_0^2]\}, \tag{3.14}$$

and

$$\mathfrak{L}(\langle u_2^2\rangle) = (2/\Delta)(s+2b)[(s+2b)^2+4\beta_0^2]. \tag{3.15}$$

The remaining elements of $\mathfrak{L}(\langle \mathbf{F} \times \mathbf{F}\rangle)$ are given by

$$\mathfrak{L}(\langle v_2^2\rangle) = \mathfrak{L}(\langle u_1^2\rangle), \tag{3.16}$$

$$\mathfrak{L}(\langle u_1u_2\rangle) = \mathfrak{L}(\langle u_2v_2\rangle) = \tfrac{1}{2}s\mathfrak{L}(\langle u_2^2\rangle), \tag{3.17}$$

$$\mathfrak{L}(\langle u_1v_1\rangle) = \mathfrak{L}(\langle v_1v_2\rangle) = \tfrac{1}{2}[s\mathfrak{L}(\langle u_1^2\rangle) - 1] \tag{3.18}$$

and

$$\mathfrak{L}(\langle u_1v_2\rangle + \langle v_1u_2\rangle) = \tfrac{1}{2}s^2\mathfrak{L}(\langle u_2^2\rangle),$$
$$\mathfrak{L}(\langle u_1v_2\rangle - \langle v_1u_2\rangle) = 1/s. \tag{3.19}$$

Finally, expressions were given[2] for the Laplace transforms of the elements of $\langle \mathbf{F}(z) \times \mathbf{F}(z)\rangle_k$, but these involved the products and inverses of certain 2×2 matrices. The evaluation of these expressions is somewhat tedious, and only some of the details

are given in Appendix B. It is found that

$$\Delta\mathfrak{L}(\langle u_1^2\rangle_2 - \langle u_1^2\rangle_1) = 4\eta\beta_0^2[s^2(s+2b) + 4\beta_0^2 b], \quad (3.20)$$

$$\Delta\mathfrak{L}(\langle u_1v_1\rangle_2 - \langle u_1v_1\rangle_1) = 2\eta\beta_0^2(s+2b)[s^2(s+2b) + 4\beta_0^2 b], \quad (3.21)$$

$$\Delta\mathfrak{L}(\langle v_1^2\rangle_2 - \langle v_1^2\rangle_1) = -8\eta\beta_0^4(s+b)[s(s+2b) + 2(1-\eta^2)\beta_0^2], \quad (3.22)$$

$$\Delta\mathfrak{L}(\langle u_2^2\rangle_2 - \langle u_2^2\rangle_1) = 16\eta\beta_0^2(s+b), \quad (3.23)$$

$$\Delta\mathfrak{L}(\langle u_2v_2\rangle_2 - \langle u_2v_2\rangle_1) = 8\eta\beta_0^2(s+b)(s+2b), \quad (3.24)$$

$$\Delta\mathfrak{L}(\langle v_2^2\rangle_2 - \langle v_2^2\rangle_1) = 4\eta\beta_0^2[s(s+2b)^2 - 4\beta_0^2 b], \quad (3.25)$$

$$\Delta\mathfrak{L}(\langle u_1u_2\rangle_2 - \langle u_1u_2\rangle_1) = 8\eta\beta_0^2 s(s+b), \quad (3.26)$$

$$\Delta\mathfrak{L}(\langle v_1v_2\rangle_2 - \langle v_1v_2\rangle_1) = 2\eta\beta_0^2 s[s(s+2b)^2 - 4\beta_0^2 b], \quad (3.27)$$

and

$$\Delta\mathfrak{L}(\langle u_1v_2\rangle_2 - \langle u_1v_2\rangle_1) = 4\eta\beta_0^2 s(s+b)(s+2b) = \Delta\mathfrak{L}(\langle v_1u_2\rangle_2 - \langle v_1u_2\rangle_1). \quad (3.28)$$

With the aid of the above results, we establish in Appendix A that, when $\Gamma(z)$ is given by (3.1), the smoothing method leads to the expression given in (3.4) for $\langle \mathbf{F}(z+\zeta) \times \mathbf{F}(z)\rangle$, so that the results are exact in the case $N(z) = T(z)$. Now, Eqs. (3.13)–(3.19), together with (3.20)–(3.28), determine the elements of $\mathfrak{L}(\langle \mathbf{F} \times \mathbf{F}\rangle)_k$, $k = 1, 2$, in view of (3.10). Denote the roots of the equation $\Delta(s) = 0$, as given by (3.12), by s_p, $p = 1, \cdots, 6$, and let $s_0 = 0$. Then, inverting the Laplace transforms, we obtain

$$\langle \mathbf{F}(z) \times \mathbf{F}(z)\rangle_k = \sum_{p=0}^{6} \mathbf{C}_{p,k} e^{s_p z}, \quad k = 1, 2, \quad (3.29)$$

where the $\mathbf{C}_{p,k}$ are constant 4×4 matrices, which are calculated from the residues at $s = s_p$. It is remarked that for small η one of the roots of $\Delta(s) = 0$ has a positive real part. Similarly, the elements of $\Lambda(\sum_{j=1}^{2} \langle \mathbf{\Phi} \mid 0\rangle_{jk})$ are given by (3.7) and (3.8). Denote the roots of the equation $d(\sigma) = 0$, as given by (3.6), by σ_r, $r = 1, \cdots, 4$. Then, inverting the Laplace transforms, we have

$$\sum_{j=1}^{2} \langle \mathbf{\Phi}(\zeta) \mid 0\rangle_{jk} = \sum_{r=1}^{4} \mathbf{B}_{r,k} e^{\sigma_r \zeta}, \quad k = 1, 2, \quad (3.30)$$

where the $\mathbf{B}_{r,k}$ are constant 2×2 matrices. It can be shown that $\mathrm{Re}\,(\sigma_r) < 0$ for $r = 1, 2, 3, 4$, as long as $\eta < 1$. Combining (3.4), (3.29), and (3.30), we have

$$\langle \mathbf{F}(z+\zeta) \times \mathbf{F}(z)\rangle = \tfrac{1}{2} \sum_{k=1}^{2} \sum_{r=1}^{4} \sum_{p=0}^{6} (\mathbf{B}_{r,k} \times \mathbf{I}) \mathbf{C}_{p,k} e^{\sigma_r \zeta + s_p z}. \quad (3.31)$$

4. THE GENERAL PROCESS

We return now to the general case, and consider first the calculation of the first-order moments by the smoothing method. We did this previously[1] in scalar form, but it is desirable here to have the results in matrix form. Applying the smoothing method to Eq. (2.9), with the variable z replaced by ζ, and solving the equation corresponding to (2.7), using (1.6) and (2.16), we obtain

$$\subset \mathbf{F}(\zeta) \supset = -\eta\beta_0^2 \int_0^{\zeta} \mathbf{P}(\zeta - \xi) N(\xi) \mathbf{C} \langle \mathbf{F}(\xi)\rangle \, d\xi. \quad (4.1)$$

Then, from (1.1) and the equation corresponding to (2.5),

$$\frac{d\langle \mathbf{F}\rangle}{d\zeta} + \mathbf{A}\langle \mathbf{F}(\zeta)\rangle = \eta^2\beta_0^4 \int_0^{\zeta} \Gamma(\zeta - \xi) \mathbf{CP}(\zeta - \xi) \mathbf{C} \langle \mathbf{F}(\xi)\rangle \, d\xi, \quad (4.2)$$

with initial condition, from (1.6), $\langle \mathbf{F}(0)\rangle = \mathbf{I}$. Taking Laplace transforms, as in (3.5), we obtain

$$[\sigma\mathbf{I} + \mathbf{A} - \eta^2\beta_0^4 \Lambda(\Gamma \mathbf{CPC})] \Lambda(\langle \mathbf{F}\rangle) = \mathbf{I}. \quad (4.3)$$

An explicit formula for $\Lambda(\langle \mathbf{F}\rangle)$ is given in Appendix C, but for the moment (4.3) suffices. Note, from (1.1), (3.3), (3.9), (A3), and (A5), that (4.3) is exact in the case $N(z) = T(z)$.

We now consider Eq. (2.22) for $\langle \mathbf{W}(z, \zeta)\rangle$. Taking Laplace transforms with respect to ζ and using the initial condition in (2.14), we obtain

$$\{[\sigma\mathbf{I} + \mathbf{A} - \eta^2\beta_0^4 \Lambda(\Gamma \mathbf{CPC})] \times \mathbf{I}\} \Lambda(\langle \mathbf{W}\rangle) - \langle \mathbf{F}(z) \times \mathbf{F}(z)\rangle = \eta^2\beta_0^4 \Lambda\left(\{[\mathbf{CP}(\zeta)] \times \mathbf{I}\} \int_0^{z} \Gamma(z + \zeta - \xi) \mathbf{\Psi}(z - \xi) \times \langle \mathbf{F}(\xi) \times \mathbf{F}(\xi)\rangle \, d\xi \right). \quad (4.4)$$

But, from (4.3),

$$\{[\sigma\mathbf{I} + \mathbf{A} - \eta^2\beta_0^4 \Lambda(\Gamma \mathbf{CPC})] \times \mathbf{I}\}^{-1} = \Lambda(\langle \mathbf{F}\rangle) \times \mathbf{I} = \Lambda(\langle \mathbf{F}\rangle \times \mathbf{I}). \quad (4.5)$$

Hence, from (1.7), (4.4), and (4.5), solving for $\Lambda(\langle \mathbf{W}\rangle)$ and inverting the Laplace transforms, we get

$$\langle \mathbf{F}(z+\zeta) \times \mathbf{F}(z)\rangle - [\langle \mathbf{F}(\zeta)\rangle \times \mathbf{I}] \langle \mathbf{F}(z) \times \mathbf{F}(z)\rangle = \eta^2\beta_0^4 \int_0^{\zeta} \{[\langle \mathbf{F}(\zeta - \theta)\rangle \mathbf{CP}(\theta)] \times \mathbf{I}\} \times \int_0^{z} \Gamma(z + \theta - \xi) \mathbf{\Psi}(z - \xi) \langle \mathbf{F}(\xi) \times \mathbf{F}(\xi)\rangle \, d\xi \, d\theta, \quad (4.6)$$

for $\zeta \geq 0$. It follows from the results of the previous section that (4.6) is exact in the case $N(z) = T(z)$.

Now, we previously[1] applied the smoothing method to determine the first- and second-order moments of $\mathbf{F}$, and expressions were obtained for the Laplace transforms of these quantities. It is shown in Appendix C that the expression for $\Lambda(\langle\mathbf{F}\rangle)$ obtained from (4.3) is consistent with our previous results. Let

$$\gamma(\sigma) = \Lambda(\Gamma). \tag{4.7}$$

The expression for $\Lambda(\langle\mathbf{F}\rangle)$ is inverted approximately in Appendix C, under the assumption that $\gamma(\sigma)$ is analytic for $\text{Re}\,(\sigma) \geq -a$, where $a > 0$ is independent of η. This condition implies that $\Gamma(\zeta)$ is exponentially small for large ζ. It is found that, for sufficiently small η and $\eta^2\beta_0\zeta \leq O(1)$,

$$\langle u_1(\zeta)\rangle = \langle v_2(\zeta)\rangle \approx \tfrac{1}{2}(e^{\sigma_1\zeta} + e^{\sigma_2\zeta}) \tag{4.8}$$

and

$$\langle v_1(\zeta)\rangle \approx \tfrac{1}{2}i\beta_0(e^{\sigma_1\zeta} - e^{\sigma_2\zeta}),$$

$$\langle u_2(\zeta)\rangle \approx \frac{-i}{2\beta_0}(e^{\sigma_1\zeta} - e^{\sigma_2\zeta}), \tag{4.9}$$

where

$$\sigma_1 = \sigma_2^* \approx i\beta_0 + \tfrac{1}{4}\eta^2\beta_0^2[\gamma(2i\beta_0) - \gamma(0)]. \tag{4.10}$$

Now consider the second-order moments, the Laplace transforms of which we obtained previously[1] and now give in Appendix C. The transforms are inverted approximately, under the above assumption on $\gamma(\sigma)$. The expression in (C7) has simple poles at s_1, s_2, and s_3 where

$$s_1 \approx \tfrac{1}{2}\eta^2\beta_0^2[\gamma(2i\beta_0) + \gamma(-2i\beta_0)] \tag{4.11}$$

and

$$s_2 = s_3^* \approx 2i\beta_0 + \tfrac{1}{2}\eta^2\beta_0^2[\gamma(2i\beta_0) - 2\gamma(0)]. \tag{4.12}$$

As noted previously,[1] the expression on the right-hand side of Eq. (4.11) is nonnegative, a consequence of the fact that $\Gamma(z)$ is the correlation function of a real process. It is found that, for sufficiently small η, and $\eta^2\beta_0 z \leq O(1)$,

$$\langle u_1^2(z)\rangle = \langle v_2^2(z)\rangle \approx \tfrac{1}{4}(e^{s_2 z} + 2e^{s_1 z} + e^{s_3 z}), \tag{4.13}$$

$$\langle v_1^2(z)\rangle \approx -\tfrac{1}{4}\beta_0^2(e^{s_2 z} - 2e^{s_1 z} + e^{s_3 z}) \approx \beta_0^4\langle u_2^2(z)\rangle, \tag{4.14}$$

$$\langle u_1(z)u_2(z)\rangle = \langle u_2(z)v_2(z)\rangle \approx \frac{-i}{4\beta_0}(e^{s_2 z} - e^{s_3 z}), \tag{4.15}$$

$$\langle u_1(z)v_1(z)\rangle = \langle v_1(z)v_2(z)\rangle \approx \tfrac{1}{4}i\beta_0(e^{s_2 z} - e^{s_3 z}), \tag{4.16}$$

and

$$\langle u_1(z)v_2(z)\rangle \approx \tfrac{1}{4}(e^{s_2 z} + 2 + e^{s_3 z}),$$

$$\langle v_1(z)u_2(z)\rangle \approx \tfrac{1}{4}(e^{s_2 z} - 2 + e^{s_3 z}). \tag{4.17}$$

We remark that Papanicolaou and Keller[18] investigated the solution of (1.2), in the case of $m = 1$, by means of a 2-variable expansion procedure. The above approximations to the first- and second-order moments may be obtained from their results, and we give a few of the details in Appendix D.

Let us now return to (4.6), and consider the double integral therein. From (2.15) and (2.20), $\mathbf{P}(\theta)$ and $\mathbf{\Psi}(z - \xi)$ are bounded. From the above approximations, for $\eta^2 \ll 1$, to the elements of $\langle\mathbf{F}\rangle$ and $\langle\mathbf{F} \times \mathbf{F}\rangle$, these quantities are bounded on fixed intervals $0 \leq \eta^2\beta_0\zeta \leq \Xi$ and $0 \leq \eta^2\beta_0 z \leq Z$, where Ξ and Z are $O(1)$. Moreover,

$$\int_0^\zeta\int_0^z |\Gamma(z + \theta - \xi)|\, d\xi\, d\theta = \int_0^\zeta\int_0^z |\Gamma(\xi + \theta)|\, d\xi\, d\theta \tag{4.18}$$

is clearly bounded for $\zeta \geq 0$ and $z \geq 0$, under our assumption on $\gamma(\sigma)$, which implies that $\Gamma(\zeta)$ is exponentially small for large ζ. It follows that the right-hand side of (4.6) is $O(\eta^2)$ for $0 \leq \eta^2\beta_0\zeta \leq \Xi$ and $0 \leq \eta^2\beta_0 z \leq Z$, so that (1.9) holds.

5. ANOTHER APPLICATION OF THE SMOOTHING METHOD

In conclusion, we consider another application of the smoothing method to the calculation of the expectation of the matrix $\mathbf{W}(z, \zeta)$ defined in (1.7). In Sec. 2 we derived an equation for $\mathbf{W}$ regarded as a function of ζ, but here we derive one for $\mathbf{W}$ regarded as a function of z. Thus, from (2.9),

$$\frac{\partial}{\partial z}\mathbf{F}(z + \zeta) + [\mathbf{A} + \eta\beta_0^2 N(z + \zeta)\mathbf{C}]\mathbf{F}(z + \zeta) = 0. \tag{5.1}$$

We now take the Kronecker product of the expression in (5.1) with $\mathbf{F}(z)$ and of $\mathbf{F}(z + \zeta)$ with the expression in (2.9), and add, to obtain

$$\frac{\partial\mathbf{W}}{\partial z} + [(\mathbf{A} \times \mathbf{I}) + (\mathbf{I} \times \mathbf{A})]\mathbf{W}(z, \zeta) + \eta\beta_0^2[N(z + \zeta)(\mathbf{C} \times \mathbf{I}) + N(z)(\mathbf{I} \times \mathbf{C})]\mathbf{W}(z, \zeta) = 0. \tag{5.2}$$

From (1.6) and (1.7), the initial condition is

$$\mathbf{W}(0, \zeta) = \mathbf{F}(\zeta) \times \mathbf{I}. \tag{5.3}$$

Note that, if $\zeta = 0$, then (5.2) reduces to (2.19).

Applying the smoothing method to Eq. (5.2) and solving the equation corresponding to (2.7), using (2.16) and (5.3), we obtain

$$\subset\mathbf{W}(z, \zeta)\supset = [\mathbf{P}(z) \times \mathbf{P}(z)][\subset\mathbf{F}(\zeta)\supset \times \mathbf{I}] - \eta\beta_0^2\int_0^z [\mathbf{P}(z - \xi) \times \mathbf{P}(z - \xi)][N(\xi + \zeta)(\mathbf{C} \times \mathbf{I}) + N(\xi)(\mathbf{I} \times \mathbf{C})]\langle\mathbf{W}(\xi, \zeta)\rangle\, d\xi. \tag{5.4}$$

But, in the smoothing approximation, $\subset\mathbf{F}(\zeta)\supset$ is given by (4.1), in terms of $\langle\mathbf{F}\rangle$, which has already been determined. From (5.2) and the equation corresponding to (2.5), it follows that

$$\frac{\partial\langle\mathbf{W}\rangle}{\partial z} + [(\mathbf{A}\times\mathbf{I}) + (\mathbf{I}\times\mathbf{A})]\langle\mathbf{W}(z,\zeta)\rangle = -\eta\beta_0^2\langle[N(z+\zeta)(\mathbf{C}\times\mathbf{I}) + N(z)(\mathbf{I}\times\mathbf{C})]\subset\mathbf{W}(z,\zeta)\supset\rangle. \quad (5.5)$$

Combining Eqs. (4.1), (5.4), and (5.5) and using (1.1), we obtain an equation for $\langle\mathbf{W}\rangle$, which we refrain from writing down. From (5.3), the initial condition is

$$\langle\mathbf{W}(0,\zeta)\rangle = \langle\mathbf{F}(\zeta)\rangle\times\mathbf{I}. \quad (5.6)$$

The equation for $\langle\mathbf{W}\rangle$ may be solved by means of Laplace transforms. However, it appears that the exact result is not obtained in the case $N(z) = T(z)$, whereas it is obtained from Eq. (2.22) in this case. This could be explained by the occurrence of both $N(z+\zeta)$ and $N(z)$ in (5.2), whereas only the process $N(z+\zeta)$ occurs in (2.11).

APPENDIX A

We here consider the solution of Eq. (2.22), subject to the initial condition given in (2.14), in the particular case in which the correlation function of the process $N(z)$ is given by (3.1). Thus, in (2.22) we have

$$\Gamma(\zeta-\xi) = e^{-2b(\zeta-\xi)}, \quad 0\le\xi\le\zeta, \quad (A1)$$

and

$$\Gamma(z+\zeta-\xi) = e^{-2b\zeta}e^{-2b(z-\xi)}, \quad \zeta\ge 0, \quad 0\le\xi\le z. \quad (A2)$$

We will take Laplace transforms with respect to ζ, as in (3.5). Note, from (2.16), that

$$\Lambda(\mathbf{P}) = (\sigma\mathbf{I}+\mathbf{A})^{-1} = \frac{1}{(\sigma^2+\beta_0^2)}\begin{pmatrix}\sigma & 1\\ -\beta_0^2 & \sigma\end{pmatrix}, \quad (A3)$$

from (2.8), which is consistent with (2.15). Then, from (2.22), using (2.14) and (A1)–(A3), we obtain

$$(\{\sigma\mathbf{I}+\mathbf{A}-\eta^2\beta_0^4\mathbf{C}[(\sigma+2b)\mathbf{I}+\mathbf{A}]^{-1}\mathbf{C}\}\times\mathbf{I})\Lambda(\langle\mathbf{W}\rangle) - \langle\mathbf{F}(z)\times\mathbf{F}(z)\rangle = \eta^2\beta_0^4(\{\mathbf{C}[(\sigma+2b)\mathbf{I}+\mathbf{A}]^{-1}\}\times\mathbf{I}) \times\int_0^z e^{-2b(z-\xi)}\boldsymbol{\Psi}(z-\xi)\langle\mathbf{F}(\xi)\times\mathbf{F}(\xi)\rangle\,d\xi. \quad (A4)$$

Now, from (2.8), (3.6), (3.7), and (A3), it is found, in a straightforward manner, that

$$\{\sigma\mathbf{I}+\mathbf{A}-\eta^2\beta_0^4\mathbf{C}[(\sigma+2b)\mathbf{I}+\mathbf{A}]^{-1}\mathbf{C}\}^{-1} = \tfrac{1}{2}\Lambda\Big(\sum_{k=1}^{2}\sum_{j=1}^{2}\langle\boldsymbol{\Phi}\mid 0\rangle_{jk}\Big), \quad (A5)$$

and, from (3.8), that

$$\eta\beta_0^2\{\sigma\mathbf{I}+\mathbf{A}-\eta^2\beta_0^4\mathbf{C}[(\sigma+2b)\mathbf{I}+\mathbf{A}]^{-1}\mathbf{C}\}^{-1}\times\mathbf{C}[(\sigma+2b)\mathbf{I}+\mathbf{A}]^{-1} = \tfrac{1}{2}\Lambda\Big(\sum_{k=1}^{2}(-1)^k\sum_{j=1}^{2}\langle\boldsymbol{\Phi}\mid 0\rangle_{jk}\Big). \quad (A6)$$

Hence, solving (A4) for $\Lambda(\langle\mathbf{W}\rangle)$ and inverting the Laplace transforms with the help of (A5) and (A6), we obtain

$$\langle\mathbf{W}(z,\zeta)\rangle = \tfrac{1}{2}\sum_{k=1}^{2}\Big[\Big(\sum_{j=1}^{2}\langle\boldsymbol{\Phi}(\zeta)\mid 0\rangle_{jk}\times\mathbf{I}\Big)\boldsymbol{\Theta}_k(z)\Big], \quad (A7)$$

where

$$\boldsymbol{\Theta}_k(z) = \langle\mathbf{F}(z)\times\mathbf{F}(z)\rangle + (-1)^k\eta\beta_0^2\int_0^z e^{-2b(z-\zeta)}\boldsymbol{\Psi}(z-\zeta)\times\langle\mathbf{F}(\xi)\times\mathbf{F}(\xi)\rangle\,d\xi. \quad (A8)$$

We will show that

$$\boldsymbol{\Theta}_k(z) = \langle\mathbf{F}(z)\times\mathbf{F}(z)\rangle_k, \quad (A9)$$

whence, from (1.7) and (3.4), the smoothing method leading to Eq. (2.22) gives the correct result for the random telegraph process $T(z)$.

Now, taking Laplace transforms with respect to z, as in (3.11), and using the convolution theorem, it follows from (A8) that

$$\mathfrak{L}(\boldsymbol{\Theta}_k) = \{(\mathbf{I}\times\mathbf{I}) + (-1)^k\eta\beta_0^2\mathfrak{L}[e^{-2bz}\boldsymbol{\Psi}(z)]\}\mathfrak{L}(\langle\mathbf{F}\times\mathbf{F}\rangle). \quad (A10)$$

But, from (2.8), (2.15), and (2.20), it is found that

$$\boldsymbol{\Psi}(z) = \begin{pmatrix} \beta_0^{-1}\sin 2\beta_0 z & \beta_0^{-2}\sin^2\beta_0 z & \beta_0^{-2}\sin^2\beta_0 z & 0\\ \cos 2\beta_0 z & (2\beta_0)^{-1}\sin 2\beta_0 z & (2\beta_0)^{-1}\sin 2\beta_0 z & 0\\ \cos 2\beta_0 z & (2\beta_0)^{-1}\sin 2\beta_0 z & (2\beta_0)^{-1}\sin 2\beta_0 z & 0\\ -\beta_0\sin 2\beta_0 z & \cos^2\beta_0 z & \cos^2\beta_0 z & 0\end{pmatrix}. \quad (A11)$$

Thus,

$$(s+2b)[(s+2b)^2+4\beta_0^2]\mathfrak{L}[e^{-2bz}\Psi(z)] = \begin{pmatrix} 2(s+2b) & 2 & 2 & 0 \\ (s+2b)^2 & (s+2b) & (s+2b) & 0 \\ (s+2b)^2 & (s+2b) & (s+2b) & 0 \\ -2\beta_0^2(s+2b) & [(s+2b)^2+2\beta_0^2] & [(s+2b)^2+2\beta_0^2] & 0 \end{pmatrix}. \tag{A12}$$

Now, in view of (1.5), the elements of $\mathfrak{L}(\langle \mathbf{F} \times \mathbf{F}\rangle)$ are given by (3.13)–(3.19). Then the elements of $\mathfrak{L}(\boldsymbol{\Theta}_k)$ may be calculated from (A10) and (A12). But, the elements of $\mathfrak{L}(\langle \mathbf{F} \times \mathbf{F}\rangle_k)$, as we show in Appendix B, are given by (3.20)–(3.28), in conjunction with (3.10) and the expressions for the elements of $\mathfrak{L}(\langle \mathbf{F} \times \mathbf{F}\rangle)$. It may be verified that

$$\mathfrak{L}(\boldsymbol{\Theta}_k) = \mathfrak{L}(\langle \mathbf{F} \times \mathbf{F}\rangle_k), \quad k = 1, 2, \tag{A13}$$

so that (A9) holds, as we wanted to show.

APPENDIX B

We here simplify the expressions we obtained previously[2] for the elements of $\mathfrak{L}(\langle \mathbf{F} \times \mathbf{F}\rangle_k)$. First, for $m = 1, 2$, let

$$\mathbf{X}_m = \mathfrak{L}\begin{bmatrix}\langle u_m^2\rangle_1 \\ \langle u_m^2\rangle_2\end{bmatrix}, \quad \mathbf{Y}_m = \mathfrak{L}\begin{bmatrix}\langle u_m v_m\rangle_1 \\ \langle u_m v_m\rangle_2\end{bmatrix},$$
$$\mathbf{Z}_m = \mathfrak{L}\begin{bmatrix}\langle v_m^2\rangle_1 \\ \langle v_m^2\rangle_2\end{bmatrix}. \tag{B1}$$

Define

$$\mathbf{S} = \begin{bmatrix}(s+b) & -b \\ -b & (s+b)\end{bmatrix},$$
$$\mathbf{B} = \beta_0^2\begin{bmatrix}(1+\eta) & 0 \\ 0 & (1-\eta)\end{bmatrix}, \quad \mathbf{E} = \begin{bmatrix}1 \\ 1\end{bmatrix}. \tag{B2}$$

Then, it was found that

$$\mathbf{X}_1 = (\mathbf{BS} + \mathbf{SB} + \tfrac{1}{2}\mathbf{S}^3)^{-1}(\mathbf{B} + \tfrac{1}{2}\mathbf{S}^2)\mathbf{E},$$
$$\mathbf{X}_2 = (\mathbf{BS} + \mathbf{SB} + \tfrac{1}{2}\mathbf{S}^3)^{-1}\mathbf{E}, \tag{B3}$$

$$\mathbf{Y}_1 = \tfrac{1}{2}(\mathbf{SX}_1 - \mathbf{E}), \quad \mathbf{Y}_2 = \tfrac{1}{2}\mathbf{SX}_2, \tag{B4}$$

and

$$\mathbf{Z}_1 = [(\mathbf{B} + \tfrac{1}{2}\mathbf{S}^2)\mathbf{X}_1 - \tfrac{1}{2}\mathbf{SE}], \quad \mathbf{Z}_2 = (\mathbf{B} + \tfrac{1}{2}\mathbf{S}^2)\mathbf{X}_2. \tag{B5}$$

Now introduce the row vector

$$\mathbf{D}' = [-1 \quad 1]. \tag{B6}$$

Note, from (B2), the relationship

$$\mathbf{D}'\mathbf{E} = 0, \quad \mathbf{D}'\mathbf{S} = (s+2b)\mathbf{D}'. \tag{B7}$$

Hence, from (B4),

$$\mathbf{D}'\mathbf{Y}_1 = \tfrac{1}{2}(s+2b)\mathbf{D}'\mathbf{X}_1, \quad \mathbf{D}'\mathbf{Y}_2 = \tfrac{1}{2}(s+2b)\mathbf{D}'\mathbf{X}_2, \tag{B8}$$

and, from (B5),

$$\mathbf{D}'\mathbf{Z}_1 = \mathbf{D}'(\mathbf{B} + \tfrac{1}{2}\mathbf{S}^2)\mathbf{X}_1, \quad \mathbf{D}'\mathbf{Z}_2 = \mathbf{D}'(\mathbf{B} + \tfrac{1}{2}\mathbf{S}^2)\mathbf{X}_2. \tag{B9}$$

But, from (B2) and (B7),

$$\mathbf{D}'(\mathbf{B} + \tfrac{1}{2}\mathbf{S}^2) = \tfrac{1}{2}[(s+2b)^2 + 2\beta_0^2]\mathbf{D}' - \eta\beta_0^2\mathbf{E}'. \tag{B10}$$

However, we previously[2] calculated $\mathbf{E}'\mathbf{X}_m$, and in fact, from (3.10) and (B1),

$$\tfrac{1}{2}\mathbf{E}'\mathbf{X}_m = \mathfrak{L}(\langle u_m^2\rangle), \quad m = 1, 2. \tag{B11}$$

Hence it suffices to calculate $\mathbf{D}'\mathbf{X}_m$ and then to use (3.13), (3.15), and (B8)–(B11). After some tedious calculations, the details of which we omit, it is found from (B3) that

$$\mathbf{D}'\mathbf{X}_1 = 4\eta\beta_0^2[s^2(s+2b) + 4\beta_0^2 b]/\Delta,$$
$$\mathbf{D}'\mathbf{X}_2 = 16\eta\beta_0^2(s+b)/\Delta, \tag{B12}$$

where

$$\Delta = \det\,[2(\mathbf{BS} + \mathbf{SB}) + \mathbf{S}^3], \tag{B13}$$

and is given explicitly by (3.12). We thus obtain (3.20)–(3.25).

Secondly, let

$$\mathbf{G} = \mathfrak{L}\begin{bmatrix}\langle u_1 u_2\rangle_1 \\ \langle u_1 u_2\rangle_2\end{bmatrix}, \quad \mathbf{H} = \mathfrak{L}\begin{bmatrix}\langle u_1 v_2\rangle_1 \\ \langle u_1 v_2\rangle_2\end{bmatrix}, \tag{B14}$$

$$\mathbf{J} = \mathfrak{L}\begin{bmatrix}\langle v_1 u_2\rangle_1 \\ \langle v_1 u_2\rangle_2\end{bmatrix}, \quad \mathbf{K} = \mathfrak{L}\begin{bmatrix}\langle v_1 v_2\rangle_1 \\ \langle v_1 v_2\rangle_2\end{bmatrix}. \tag{B15}$$

Then, it was found that[2]

$$\mathbf{G} = \tfrac{1}{2}(\mathbf{BS} + \mathbf{SB} + \tfrac{1}{2}\mathbf{S}^3)^{-1}\mathbf{SE}, \quad \mathbf{K} = -\mathbf{S}^{-1}\mathbf{BSG}, \tag{B16}$$

and

$$\mathbf{H} + \mathbf{J} = \mathbf{SG}, \quad (\mathbf{H} - \mathbf{J}) = \mathbf{S}^{-1}\mathbf{E} = s^{-1}\mathbf{E}. \tag{B17}$$

But, since $\mathbf{SE} = s\mathbf{E}$, from (B3) and (B16) we have

$$\mathbf{G} = \tfrac{1}{2}s\mathbf{X}_2, \quad \mathbf{D}'\mathbf{G} = \tfrac{1}{2}s\mathbf{D}'\mathbf{X}_2. \tag{B18}$$

Also, from (B3), (B5), and (B16)–(B18),

$$\tfrac{1}{2}s\mathbf{Z}_2 - \mathbf{K} = \tfrac{1}{2}s\mathbf{S}^{-1}(\mathbf{BS} + \mathbf{SB} + \tfrac{1}{2}\mathbf{S}^3)\mathbf{X}_2$$
$$= \tfrac{1}{2}s\mathbf{S}^{-1}\mathbf{E} = \tfrac{1}{2}\mathbf{E}. \tag{B19}$$

Hence, since $\mathbf{D}'\mathbf{E} = 0$,

$$\mathbf{D}'\mathbf{K} = \tfrac{1}{2}s\mathbf{D}'\mathbf{Z}_2. \tag{B20}$$

Finally, from (B7) and (B17),

$$\mathbf{D}'(\mathbf{H} + \mathbf{J}) = (s + 2b)\mathbf{D}'\mathbf{G}, \quad \mathbf{D}'(\mathbf{H} - \mathbf{J}) = 0. \tag{B21}$$

Equations (3.26)–(3.28) follow from (B18), (B20), and (B21).

APPENDIX C

We consider here the approximate inversion of the Laplace transforms of the first- and second-order moments. We first consider the expression for $\Lambda(\langle\mathbf{F}\rangle)$ obtained from (4.3). Now, from (2.8) and (2.15),

$$\mathbf{CP}(\zeta)\mathbf{C} = \begin{pmatrix} 0 & 0 \\ \beta_0^{-1}\sin\beta_0\zeta & 0 \end{pmatrix}. \tag{C1}$$

Hence, from (4.7),

$$\eta^2\beta_0^4\Lambda(\Gamma\mathbf{CPC}) = \begin{pmatrix} 0 & 0 \\ \delta & 0 \end{pmatrix}, \tag{C2}$$

where

$$\begin{aligned}\delta &= \eta^2\beta_0^3\int_0^\infty e^{-\sigma\zeta}\sin\beta_0\zeta\Gamma(\zeta)\,d\zeta \\ &= \frac{\eta^2\beta_0^3}{2i}[\gamma(\sigma - i\beta_0) - \gamma(\sigma + i\beta_0)].\end{aligned} \tag{C3}$$

Thus, from (2.8),

$$[\sigma\mathbf{I} + \mathbf{A} - \eta^2\beta_0^4\Lambda(\Gamma\mathbf{CPC})] = \begin{pmatrix} \sigma & -1 \\ (\beta_0^2 - \delta) & \sigma \end{pmatrix}. \tag{C4}$$

Then, from (4.3),

$$\begin{aligned}\Lambda(\langle\mathbf{F}\rangle) &= [\sigma\mathbf{I} + \mathbf{A} - \eta^2\beta_0^4\Lambda(\Gamma\mathbf{CPC})]^{-1} \\ &= \frac{1}{(\sigma^2 + \beta_0^2 - \delta)}\begin{bmatrix} \sigma & 1 \\ (\delta - \beta_0^2) & \sigma \end{bmatrix},\end{aligned} \tag{C5}$$

which is consistent with our previous results.[1]

Now, according to the complex inversion formula,[22]

$$\langle\mathbf{F}(\zeta)\rangle = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} e^{\sigma\zeta}\Lambda(\langle\mathbf{F}\rangle)\,d\sigma, \tag{C6}$$

for sufficiently large c. We assume that $\gamma(\sigma)$ is analytic for $\mathrm{Re}\,(\sigma) \geq -a$, where $a > 0$ is independent of η. Then, for sufficiently small η, $\Lambda(\langle\mathbf{F}\rangle)$ has simple poles in the neighborhood of $\sigma = \pm i\beta_0$, which we denote by σ_1 and $\sigma_2 = \sigma_1^*$. The approximate values of σ_1 and σ_2, using (C3), are given by (4.10). Moreover, for sufficiently small η, the remaining singularities of $\Lambda(\langle\mathbf{F}\rangle)$ lie in $\mathrm{Re}\,(\sigma) < -\frac{1}{2}a$. Accordingly, we shift the contour of integration in (C6) to the line $\mathrm{Re}\,(\sigma) = -\frac{1}{2}a$, and include the contributions from the poles at $\sigma = \sigma_1$ and $\sigma = \sigma_2$. But, the integral along $\mathrm{Re}\,(\sigma) = -\frac{1}{2}a$ vanishes for $\eta = 0$. Hence, for small η, the main contributions to $\langle\mathbf{F}(\zeta)\rangle$ arise from the poles at $\sigma = \sigma_1$ and $\sigma = \sigma_2$. Thus, from (C5), we obtain the approximations given in (4.8) and (4.9).

Consider next the second-order moments. We found previously[1] that

$$\begin{aligned}\tfrac{1}{2}\mathfrak{L}(\langle u_2^2\rangle) = \Omega &\equiv \{s(s^2 + 4\beta_0^2) - \tfrac{1}{2}\eta^2\beta_0^2[2s^2\gamma(s) \\ &\quad - (s + 2i\beta_0)^2\gamma(s - 2i\beta_0) \\ &\quad - (s - 2i\beta_0)^2\gamma(s + 2i\beta_0)]\}^{-1},\end{aligned} \tag{C7}$$

$$\begin{aligned}\mathfrak{L}(\langle u_1^2\rangle) = \Omega\{&(s^2 + 2\beta_0^2) \\ &- \tfrac{1}{2}\eta^2\beta_0^2[2s\gamma(s) - (s + 2i\beta_0)\gamma(s - 2i\beta_0) \\ &- (s - 2i\beta_0)\gamma(s + 2i\beta_0)]\},\end{aligned} \tag{C8}$$

and

$$\begin{aligned}\mathfrak{L}(\langle v_1^2\rangle) = \Omega\beta_0^4(&2 + \eta^2[(s + 2i\beta_0)\gamma(s - 2i\beta_0) \\ &+ (s - 2i\beta_0)\gamma(s + 2i\beta_0)] \\ &+ \eta^2\beta_0^2\{2\gamma(s - 2i\beta_0)\gamma(s + 2i\beta_0) \\ &- \gamma(s)[\gamma(s - 2i\beta_0) + \gamma(s + 2i\beta_0)]\}).\end{aligned} \tag{C9}$$

The remaining elements of $\mathfrak{L}(\langle\mathbf{F} \times \mathbf{F}\rangle)$ are given by (3.16)–(3.19). Under the above assumptions on $\gamma(\sigma)$ it follows that, for sufficiently small η, the expression for Ω has simple poles in the neighborhood of $s = 0$ and $s = \pm 2i\beta_0$, which we denote by s_1, s_2, and $s_3 = s_2^*$. Their approximate values are given by (4.11) and (4.12). Proceeding as above, it follows that the main contributions to $\langle u_1^2(z)\rangle$, $\langle v_1^2(z)\rangle$, and $\langle u_2^2(z)\rangle$ arise from these three poles, and from (3.16) and (C7)–(C9) we obtain the approximations given in (4.13) and (4.14). The approximations in (4.15)–(4.17) may be obtained by inverting the relationships in (3.17)–(3.19), with the help of the initial conditions

$$\langle\mathbf{F}(0) \times \mathbf{F}(0)\rangle = \mathbf{I} \times \mathbf{I},$$

and using (4.13), (4.14),and the zero-order approximations to s_1, s_2, and s_3.

APPENDIX D

We give here some details of the calculation of the first- and second-order moments of the solutions of (1.2) from the results of Papanicolaou and Keller.[18] With some changes in notation and generalization to the matrix solution in (1.5), the first-order moments are given by

$$\langle\mathbf{F}(\zeta)\rangle \approx \mathbf{R}e^{(\mathbf{Q}+\eta^2\mathbf{M}_1)\zeta}\mathbf{R}^{-1}, \tag{D1}$$

where

$$\mathbf{Q} = \begin{pmatrix} i\beta_0 & 0 \\ 0 & -i\beta_0 \end{pmatrix}, \quad \mathbf{R} = \begin{pmatrix} 1 & 1 \\ i\beta_0 & -i\beta_0 \end{pmatrix}, \tag{D2}$$

and

$$\mathbf{M}_1 = \tfrac{1}{4}\beta_0^2\begin{pmatrix} [\gamma(2i\beta_0) - \gamma(0)] & 0 \\ 0 & [\gamma(-2i\beta_0) - \gamma(0)] \end{pmatrix}. \tag{D3}$$

Hence, from (4.10),

$$(\mathbf{Q} + \eta^2\mathbf{M}_1) \approx \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}, \quad e^{(\mathbf{Q}+\eta^2\mathbf{M}_1)\zeta} \approx \begin{pmatrix} e^{\sigma_1\zeta} & 0 \\ 0 & e^{\sigma_2\zeta} \end{pmatrix}, \tag{D4}$$

for $\eta^2\beta_0\zeta \leq O(1)$. Equations (4.8) and (4.9) follow, after some further straightforward algebra.

The second-order moments are given by[18]

$$\langle \mathbf{F}(z) \times \mathbf{F}(z) \rangle \approx (\mathbf{R} \times \mathbf{R}) \exp\{[(\mathbf{Q} \times \mathbf{I}) + (\mathbf{I} \times \mathbf{Q}) + \eta^2\mathbf{N}_1]z\}(\mathbf{R} \times \mathbf{R})^{-1}, \tag{D5}$$

where

$$\mathbf{N}_1 = \tfrac{1}{2}\beta_0^2 \begin{pmatrix} [\gamma(2i\beta_0) - 2\gamma(0)] & 0 & 0 & 0 \\ 0 & \mathrm{Re}\,[\gamma(-2i\beta_0)] & \mathrm{Re}\,[\gamma(-2i\beta_0)] & 0 \\ 0 & \mathrm{Re}\,[\gamma(-2i\beta_0)] & \mathrm{Re}\,[\gamma(-2i\beta_0)] & 0 \\ 0 & 0 & 0 & [\gamma(-2i\beta_0) - 2\gamma(0)] \end{pmatrix}. \tag{D6}$$

Hence, from (4.11), (4.12), (D2), and (D6),

$$[(\mathbf{Q} \times \mathbf{I}) + (\mathbf{I} \times \mathbf{Q}) + \eta^2\mathbf{N}_1] \approx \begin{bmatrix} s_2 & 0 & 0 & 0 \\ 0 & \frac{1}{2}s_1 & \frac{1}{2}s_1 & 0 \\ 0 & \frac{1}{2}s_1 & \frac{1}{2}s_1 & 0 \\ 0 & 0 & 0 & s_3 \end{bmatrix}. \tag{D7}$$

Equations (4.13)–(4.17) follow after some further straightforward calculations.

[1] J. A. Morrison and J. McKenna, J. Math. Phys. **11**, 2361 (1970).

[2] J. McKenna and J. A. Morrison, J. Math. Phys. **11**, 2348 (1970).

[3] U. Frisch, "Wave Propagation in Random Media," in *Probabilistic Methods in Applied Mathematics*, A. T. Bharucha-Reid, Ed. (Academic, New York, 1968), p. 114. What throughout the paper we call the smoothing method, Frisch calls the first-order smoothing approximation.

[4] J. B. Keller, Proc. Symp. Appl. Math. **13**, 227 (1962).

[5] J. B. Keller, Proc. Symp. Appl. Math. **16**, 145 (1964).

[6] F. C. Karal and J. B. Keller, J. Math. Phys. **5**, 537 (1964).

[7] J. B. Keller and F. C. Karal, J. Math. Phys. **7**, 661 (1966).

[8] J. B. Keller, "Stochastic Equations and the Theory of Turbulence" in *Geophysical Fluid Dynamics*, Course Lectures and Seminars (Woods Hole Oceanographic Institute, Woods Hole, Mass., 1966), pp. 131–36.

[9] J. B. Keller, *Proceeding Symposium on Turbulence in Fluids and Plasmas* (Brooklyn Polytechnical Institute, Brooklyn, N.Y., 1968), pp. 131–42.

[10] J. B. Keller, *Contemporary Physics*, Trieste Symposium, 1968 (International Atomic Energy Agency, Vienna, 1969), Vol. I, pp. 257–72.

[11] R. C. Bourret, Nuovo Cimento **26**, 1 (1962).

[12] R. C. Bourret, Can. J. Phys. **40**, 782 (1962).

[13] R. C. Bourret, Phys. Letters **13**, 323 (1964).

[14] R. C. Bourret, Can. J. Phys. **43**, 619 (1965).

[15] R. C. Bourret, Can. J. Phys. **44**, 2519 (1966).

[16] U. Frisch and R. C. Bourret, *Parastochastics* (Institute of Astrophysics, Paris, 1967).

[17] R. Bellman, *Introduction to Matrix Analysis* (McGraw-Hill, New York, 1960), p. 227.

[18] G. C. Papanicolaou and J. B. Keller, "Stochastic Differential Equations with Applications to Random Harmonic Oscillators and Wave Propagation in Random Media," submitted to SIAM J. Appl. Math.; see G. C. Papanicolaou, Ph.D. thesis, New York University, 1969.

[19] R. Z. Khas'minskii, Teoriya Veroyatnostei i ee Primeneniya **11**, 444 (1966) [Theoret. Probability Appl. (USSR) **11**, 390 (1966)].

[20] R. L. Stratonovich, *Topics in the Theory of Random Noise* (Gordon and Breach, New York, 1963 and 1967), Vols. I and II.

[21] W. B. Davenport, Jr., and W. L. Root, *Random Signals and Noise* (McGraw-Hill, New York, 1958), p. 61.

[22] W. Magnus and F. Oberhettinger, *Formulas and Theorems for the Functions of Mathematical Physics* (Chelsea, New York, 1954), p. 121.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Mixed-Basis D Functions and Clebsch–Gordan Coefficients of Noncompact Groups

WILLIAM H. KLINK*
Institute for Advanced Study, Princeton, New Jersey 08540

(Received 17 November 1969; Revised Manuscript Received 6 April 1970)

Mixed-basis D functions are introduced as a tool for deriving Clebsch–Gordan coefficients of induced representations of semisimple groups. The Clebsch–Gordan coefficients of $SU(1,1)$ and $SL(2,C)$ are computed as examples.

1. INTRODUCTION

In previous papers[1] Mackey's induced representation theory[2] has been used to compute Clebsch–Gordan coefficients for those noncompact groups which can be written as semidirect product groups. In these papers it was conjectured that it might be possible to use similar techniques for the semisimple noncompact groups. The point of this paper is to show that mixed-basis D functions, obtained from integrals over D functions of certain subgroups, provide a tool by which the Clebsch–Gordan coefficients for some representations of the semisimple groups can be computed.

The motivation for computing Clebsch–Gordan coefficients of induced representations of semisimple groups arises from a model in which the energy dependence of reduced amplitudes of 2-body reactions is given in terms of such Clebsch–Gordan coefficients.[3] But even aside from such a physical motivation, there is the interesting mathematical question of the functional form and analytic behavior of Clebsch–Gordan coefficients of noncompact groups.

Two semisimple groups of current interest in high-energy physics are $SU(1,1)$ and $SL(2,C)$[4]; the tensor product reduction for both of these groups is given in the mathematical literature.[5] The Clebsch–Gordan coefficients of $SU(1,1)$ have been computed in various degrees of generality by a number of authors[6]; recently, the Clebsch–Gordan coefficients for the principal series of $SL(2,C)$ were also computed.[7] Now the techniques to be presented in Sec. 2 for computing Clebsch–Gordan coefficients are different from any of those used in Refs. 6 and 7; hence, after the Clebsch–Gordan coefficients of $SU(1,1)$ and $SL(2,C)$ are computed in Secs. 3 and 4, a comparison between the methods will be given in the conclusion.

2. GENERAL ANALYSIS OF D FUNCTIONS AND CLEBSCH–GORDAN COEFFICIENTS

D functions are usually defined in terms of the action of the unitary operator $U(g)$ of a group element $g \in G$ acting on an (in general improper) vector $|[\chi]x\rangle$:

$$U(g)\,|[\chi]x\rangle = \sum_{x'} \int D^{\chi}_{x'x}(g)\,|[\chi]x'\rangle. \tag{2.1}$$

$[\chi]$ denotes the set of labels specifying a unitary irreducible representation of the group G. x denotes a set of "eigenvalues" arising from a complete set of commuting elements of the Lie algebra of G; the sum and integral sign in (2.1) indicates that x may have both discrete and continuous parts.

The vectors $|[\chi]x\rangle$ are generally not elements of the Hilbert space on which the unitary operators $U(g)$ act; rather, they are generalized functions on a suitably defined rigged Hilbert space.[8] In this work the complication of treating the vectors $|[\chi]x\rangle$ (or for that matter the D functions or Clebsch–Gordan coefficients[9]) as generalized functions will not be considered.

However, it should be pointed out that, when dealing with the orthogonality and completeness properties of the D functions, the fact that the D functions are generalized functions becomes important. In contrast, for compact groups, where the D functions are finite dimensional, the orthogonality and completeness relations are readily obtained.[10] Only the invariance of Haar measure and Schur's lemma are needed.

Equation (2.1) can be transformed into a more useful form by taking the matrix elements of the operator $U(g)$:

$$\langle[\chi]x'|\,U(g)\,|[\chi]x\rangle = D^{\chi}_{x'x}(g). \tag{2.2}$$

For those representations of G which can be written as induced representations,[1,2] the D functions are readily computed. To see this, we denote by H the subgroup of G which induces irreducible representations of G. Let L be a 1-dimensional[11] representation of H acting on the vector space $\mathcal{H}(L)$ and $f(g)$ a function from G to $\mathcal{H}(L)$ satisfying $f(hg) = L(h)f(g)$; the class of such functions, suitably restricted, forms a Hilbert space denoted by $\mathcal{H}(U^L)$ on which the

induced representation acts:

$$U(g')f(g) = [m(g, g')]^{\frac{1}{2}} f(gg'), \quad g' \in G, \quad f \in \mathcal{H}(U^L). \tag{2.3}$$

$m(g', g)$ is a multiplier, so chosen as to make $U(g)$ unitary.

Before showing how $m(g', g)$ is calculated, note that the natural framework for induced representations is in terms of vector bundles over the homogeneous space G/H.[12] However, this setting is rather awkward when defining the mixed-basis D function, and for that reason it is not used here.

Rather, $f(g)$ will be considered as a function over cosets $f(g_c)$ [since $f(g) = L(h)f(g_c)$], where $\{g_c\}$ is a set of elements of G labeling the right cosets:

$$G = \bigcup_c Hg_c, \quad g = h(g)g_c(g). \tag{2.4}$$

Then

$$\begin{aligned} U(g)f(g_c) &= [m(h(g_c g))]^{\frac{1}{2}} f(g_c g) \\ &= [m(h(g_c g))]^{\frac{1}{2}} L(h(g_c g)) f(g_{c'}(g_c g)), \end{aligned} \tag{2.5}$$

where $h(g_c g)$ and $g_{c'}(g_c g)$ are defined by

$$g_c g = h(g_c g) g_{c'}(g_c g)$$

and Eq. (2.4).

The multiplier $m(h)$ is given by

$$m(h) = \frac{\det {}^{\text{adj}}H(h)}{\det {}^{\text{adj}}G(h)} \tag{2.6}$$

where ${}^{\text{adj}}G(h)$ is the adjoint representation of G defined in matrix form by

$${}^{\text{adj}}G(g)(X_i) \equiv hX_i h^{-1} = \sum_j ({}^{\text{adj}}G(h))_{ji} X_j, \tag{2.7}$$

with $\{X_i\}$ a basis in the Lie algebra of G.

With the multiplier given by (2.6), it can be shown that $U(g)$ is a unitary operator:

$$\begin{aligned} \|U(g)f\|^2 &= \int_{G/H} dg_c \, |U(g)f(g_c)|^2 \\ &= \int_{G/H} dg_c m(h(g_c g)) \, |f(g_{c'}(g_c g))|^2 \\ &= \int_{G/H} dg_{c'} \left| \frac{\partial g_c}{\partial g_{c'}} \right| m(h(g_c g)) \, |f(g_{c'})|^2 \\ &= \int_{G/H} dg_{c'} \, |f(g_{c'})|^2 \\ &= \|f\|^2, \end{aligned} \tag{2.8}$$

where the multiplier—by definition—has been so chosen as to cancel the Jacobian resulting in the change of variable from g_c to $g_{c'}$.

Consider now, as (in general improper) realizations of the functions $f(g_c)$, D functions arising from a subgroup of G whose elements include $\{g_c\}$. It will be shown how the D functions of G can be built from those D functions of appropriate subgroups of G. The D functions of the subgroups of G, when restricted to the coset labels g_c, can be thought of as "unit" vectors spanning the representation space of G labeled by χ. We follow the mathematical literature in denoting such vectors as $e_x(g_c)$; although defined in general only for coset labels g_c, it is possible to extend $e_x(g_c)$ to all elements g in G by setting

$$e_x(g) = L(h)e_x(g_c), \tag{2.9}$$

just as $f(g) = L(h)f(g_c)$ [$L(h)$ is the (1-dimensional) representation of the inducing subgroup H]. D functions over G are then

$$D^{\chi}_{x'x}(g) = \int_{G/H} dg_c e^*_{x'}(g_c)[U(g)e_x(g_c)], \tag{2.10}$$

which is seen to be a concrete realization of the defining Eq. (2.2) for D functions.

By defining an operator which translates g to the right, it can further be seen that $D^{\chi}_{x'x}(g)$ is a concrete realization of $|[\chi]x\rangle$ for x' held fixed:

$$\begin{aligned} O(g_0)D^{\chi}_{x'x}(g) &\equiv D^{\chi}_{x'x}(gg_0) \\ &= \sum_{x''} \int D^{\chi}_{x'x''}(g) D^{\chi}_{x''x}(g_0) \\ &= \sum_{x''} \int D^{\chi}_{x''x}(g_0) D^{\chi}_{x'x''}(g). \end{aligned} \tag{2.11}$$

It will prove necessary in the following development to broaden the definition of D functions. In Eq. (2.10) the "eigenvalues" x' and x both came from the same complete set of commuting elements of the Lie algebra of G. Consider, however, the possibility of allowing the two sets of "eigenvalues" to arise from two different complete sets of commuting observables of the Lie algebra of G.

In order to obtain Clebsch–Gordan coefficients, the left set of "eigenvalues" of such "mixed-basis" D functions must arise out of a special set of commuting elements of the Lie algebra of G. In the following paragraphs attention will be focused on how to choose this set for the principal series of representations of the semisimple group G, although the ideas probably generalize to the discrete and exceptional series.

For the principal series of unitary irreducible representations of G, the inducing subgroup H is obtained from the Iwasawa decomposition to be NAM, where N is nilpotent, A is Abelian, and M is the centralizer of A in G.[12] We choose the set of "eigenvalues" y—the left index in the mixed-basis D function—to include the eigenvalues of A and M, and then extend this set to be a complete set by

choosing any other convenient elements from the Lie algebra of G. This special complete set will generate D functions and, hence, vectors denoted by $e_y(g_c)$, which, as before, can be extended to

$$e_y(g) \equiv L(h)e_y(g_c).$$

The mixed-basis D functions are then

$$D^{\chi}_{yx}(g) = \int_{G/H} dg_c e^*_y(g_c)[U(g)e_x(g_c)]. \quad (2.12)$$

Such a D function has the virtue that under the operator $O(g_0)$, defined in Eq. (2.11), it transforms to the right as a vector $|[\chi]x\rangle$, while to the left, in the y variable, it transforms like an induced representation element $f(\tilde{g}_c)$, where $\tilde{g}_c \in G$ and $\tilde{H}$ are defined below:

$$\begin{aligned} O(g_0)D^{\chi}_{yx}(\tilde{g}_c) &\equiv D^{\chi}_{yx}(\tilde{g}_c g_0) \\ &= D^{\chi}_{yx}(h(\tilde{g}_c g_0)g_{c'}(\tilde{g}_c g_0)) \\ &= L(h(\tilde{g}_c g_0))D^{\chi}_{yx}(\tilde{g}_{c'}(\tilde{g}_c g_0)). \end{aligned} \quad (2.13)$$

The mixed-basis D functions are thus seen to play a dual role, serving, on the one hand, as concrete realizations of $|[\chi]x\rangle$ and as (improper) functions transforming properly as required by induced representation theory.

It is precisely this dual role which is exploited in computing Clebsch–Gordan coefficients, for Mackey has shown how the tensor product of two (or more) induced representations can be decomposed into a direct integral over double cosets of induced representations. Such a decomposition does not, in general, lead to irreducible representations, but this problem is readily handled when computing Clebsch–Gordan coefficients. It is merely necessary to know the inducing subgroups appearing in the double coset decomposition. If these subgroups are denoted by $\tilde{H}$, the Clebsch–Gordan coefficients, as shown in Ref. 1, can be written as

$$\begin{aligned} &\langle[\chi]x;\eta \mid [\chi_1]x_1;[\chi_2]x_2\rangle \\ &= N(\chi\chi_1\chi_2)\int_{G/\tilde{H}} d\tilde{g}_c D^{\chi*}_{\tilde{y}x}(\tilde{g}_c)D^{\chi_1}_{y_1x_1}(g_{D_1}\tilde{g}_c)D^{\chi_2}_{y_2x_2}(g_{D_2}\tilde{g}_c), \end{aligned} \quad (2.14)$$

where $\{\tilde{g}_c\}$ is a set of coset representatives of G relative to $\tilde{H}$. g_{D_1} and g_{D_2} are elements of G labeling the double cosets of the outer product group $\{(g, g')\}$, g, $g' \in G$, relative to the outer product inducing subgroups (H_1, H_2) and the diagonal subgroup $\{(g, g)\} = G$. $\tilde{H}$, as shown in Ref. 1, is

$$(g_{D_1}g_{D_2})^{-1}(H, H)(g_{D_1}g_{D_2}) \cap (G, G),$$

while $N(\chi\chi_1\chi_2)$ is a normalization factor depending only on the irreducible representation labels. The left indices y_1 and y_2 are chosen so that $D^{\chi_1}_{y_1x_1}$ and $D^{\chi_2}_{y_2x_2}$ transform properly to the left relative to the irreducible representations χ_1 and χ_2. The $\tilde{y}$ appearing on $D^{\chi}_{\tilde{y}x}$ is chosen to transform to the left like the tensor product of χ_1 and χ_2. Finally, η refers to a set of degeneracy parameters (having to do with multiplicity) arising from both the labels y_1, y_2 and $\tilde{y}$, and the double coset labels g_{D_1} and g_{D_2}.

3. CLEBSCH–GORDAN COEFFICIENTS OF $SU(1, 1)$

The classic analysis of a semisimple noncompact group was carried out by Bargmann[13] on $SU(1, 1)$. The tensor product decomposition of various classes of irreducible representations of $SU(1, 1)$ has been carried out by Pukanszky.[5] In this section, as an example of the formalism developed in Sec. 2, the Clebsch–Gordan coefficients arising from the tensor product decomposition of the principal series will be computed, for comparison with the results of Ref. 6.

A general element g of $SU(1, 1)$ can be written

$$\begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix}, \quad |\alpha|^2 - |\beta|^2 = 1.$$

Its Iwasawa decomposition is $g = sk$ with

$$k \in K = \begin{pmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{pmatrix},$$

$$s \in S = \begin{pmatrix} \cosh\theta + in & i\sinh\theta + n \\ -i\sinh\theta + n & \cosh\theta - in \end{pmatrix}; \quad (3.1)$$

K is the maximal compact subgroup of $SU(1, 1)$, while S, the solvable subgroup, has an Abelian part A, obtained by setting $n = 0$ in (3.1), and a nilpotent part N, obtained by setting $\theta = 0$; thus, $a(\theta) = s(\theta, n = 0)$, $n = s(\theta = 0, n)$.

The principal series of $SU(1, 1)$ is induced from the subgroup $S = NA$, with representations

$$s \in S \to L^{\rho}(\theta) = e^{i\rho\theta}, \quad \rho \text{ real.} \quad (3.2)$$

The function space on which the induced representation acts is generally chosen to contain functions over K, so that the Hilbert space $\mathcal{H}(U^{\rho})$ has norm

$$\|f\|^2 = (2\pi)^{-1}\int_0^{2\pi} d\varphi\, |f(\varphi)|^2 < \infty. \quad (3.3)$$

To compute the action of an arbitrary element of $SU(1, 1)$ on $f(\varphi)$, it is necessary to see how group

elements of $SU(1,1)$ move coset labels [viz., (2.4)]:

$$\begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} = \begin{pmatrix} \cosh\theta + in & i\sinh\theta + n \\ -i\sinh\theta + n & \cosh\theta - in \end{pmatrix} \times \begin{pmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{pmatrix},$$
$$\alpha = e^{i\varphi}(\cosh\theta + in), \qquad (3.4)$$
$$\beta = e^{-i\varphi}(i\sinh\theta + n),$$
$$\varphi(\alpha,\beta): \quad e^{i\varphi} = (\alpha - i\beta^*)/|\alpha - i\beta^*|,$$
$$\theta(\alpha,\beta): \quad e^{-\theta} = |\alpha - i\beta^*|.$$

$n(\alpha,\beta)$ is not needed since, as can be seen from (3.2), it gets sent into the identity.

The multiplier $m(s)$, according to (2.6), is

$$\det[{}^{\mathrm{adj}}S(s)]/\det[{}^{\mathrm{adj}}SU(1,1)(s)].$$

Since $SU(1,1)$ is semisimple, $\det[{}^{\mathrm{adj}}SU(1,1)(s)] = 1$. Using the definitions of ${}^{\mathrm{adj}}S(s)$ given in (2.7), we readily see that $\det[{}^{\mathrm{adj}}S(s)] = e^{2\theta}$.

Hence, the principal series of representations of $SU(1,1)$, labeled by ρ, can be written as

$$U(\alpha_0,\beta_0)f(\varphi) = e^{\theta(\varphi,\alpha_0,\beta_0)}e^{i\rho\theta(\varphi,\alpha_0,\beta_0)}f(\varphi') = [e^{-\theta(\varphi,\alpha_0,\beta_0)}]^{-1-i\rho}f(\varphi') = |\alpha_0 e^{i\varphi} - i\beta^* e^{-i\varphi}|^{-1-i\rho}f(\varphi') \qquad (3.5)$$

with $\varphi'(\varphi,\alpha_0,\beta_0)$ given by

$$e^{i\varphi'} = \frac{\alpha_0 e^{i\varphi} - i\beta_0^* e^{-i\varphi}}{|\alpha_0 e^{i\varphi} - i\beta_0^* e^{-i\varphi}|}. \qquad (3.6)$$

To obtain D functions for $SU(1,1)$, we choose $e_x(g_c)$ [Eq. (2.9)] to be the D functions of K, namely $e^{im\varphi}$, so that $e_x(g) = e^{i\rho\theta}e^{im\varphi}$. Then

$$D^{\rho}_{m'm}(\alpha_0,\beta_0) = \int_{SU(1,1)/H} dg_c e^{-im'\varphi}[U(\alpha_0,\beta_0)e^{im\varphi}] = (2\pi)^{-1}\int_0^{2\pi} d\varphi e^{-im'\varphi}[U(\alpha_0,\beta_0)e^{im\varphi}], \qquad (3.7)$$

where the action of (α_0,β_0) on φ is given in Eq. (3.6). The D functions [Eq. (3.7)] were first obtained by Bargmann.[13]

However, as was pointed out in Sec. 2, in order to calculate the Clebsch–Gordan coefficients of $SU(1,1)$, it is necessary to compute mixed-basis D functions. For the group $SU(1,1)$ it is sufficient to choose, as a complete set containing the Lie algebra element of A, just this element itself. Then the S–A double cosets must be computed. This is done in the Appendix; the basis vector $e_\nu(g_c)$ is chosen to be $e^{i\rho'\theta}$ and extended to

$$e_{\rho'}(g) = L(h)e_{\rho'}(g_c) = e^{i\rho\theta_1(g)}e^{i\rho'\theta_2(g)}. \qquad (3.8)$$

Then the mixed-basis D functions are

$$D^{\rho}_{\rho'm}(g) = \int_{SU(1,1)/H} dg_c e_{\rho'}(g_c)[U(g)e_m(g_c)] = (2\pi)^{-1}\int_0^{2\pi} d\varphi\,|(\sin 2\varphi)^{-1}|^{-\frac{1}{2}i\rho} \times |\tan\varphi|^{-\frac{1}{2}i\rho'}[U(g)e^{im\varphi}]. \qquad (3.9)$$

As will be seen below, only $D^{\rho}_{\rho'm}(nk(\varphi))$ and $D^{\rho}_{\rho'm}(g_D nk(\varphi))$ are actually needed in computing Clebsch–Gordan coefficients. Thus,

$$D^{\rho}_{\rho'm}(g_D nk(\varphi)) = D^{\rho}_{\rho'm}(g_D n)e^{im\varphi}, \qquad (3.10)$$

and letting $z_\pm(n,\varphi) = (1+in)e^{i\varphi} \pm ine^{-i\varphi}$ gives

$$D^{\rho}_{\rho'm}(n) = (2\pi)^{-1}\int d\varphi e_{\rho'}(\varphi)\,|z_-|^{-1-i\rho}\left(\frac{z_-}{|z_-|}\right)^m,$$
$$D^{\rho}_{\rho'm}(g_D n) = (2\pi)^{-1}\int d\varphi e_{\rho'}(\varphi)\,|z_+|^{-1-i\rho}\left(\frac{z_+}{|z_+|}\right)^m e^{\frac{1}{2}im\pi}. \qquad (3.11)$$

To compute the Clebsch–Gordan coefficients for the principal series of $SU(1,1)$, it is necessary to compute the double cosets arising in the $SU(1,1)$ outer product group. It is not difficult to see that this is equivalent to computing the S–S double cosets. The Appendix shows [Eq. (A6)] that the double coset $S\left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)S$ is dense in $SU(1,1)$ and, further, the inducing subgroup for the diagonal group is

$$\tilde{H} = \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)^{-1}S\left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right) \cap S = A. \qquad (3.12)$$

Hence the Clebsch–Gordan coefficients will be given by an integral over $SU(1,1)/A$:

$$\begin{aligned}\langle[\rho]m \mid [\rho_1]m_1;[\rho_2]m_2\rangle &= N(\rho\rho_1\rho_2)\int_{SU(1,1)/A} d\tilde{g}_c D^{\rho^*}_{-\rho_1+\rho_2 m}(\tilde{g}_c) \times D^{\rho_1}_{-\rho_1 m_1}(\tilde{g}_c)D^{\rho_2}_{-\rho_2 m_2}(g_D\tilde{g}_c) \\ &= N(\rho\rho_1\rho_2)\int_N dn \quad 2\pi^{-1}\int d\varphi \\ &\quad \times \langle[\rho] - \rho_1 + \rho_2|\ U(n)\ |[\rho]m\rangle^*\, e^{-im\varphi} \\ &\quad \times \langle[\rho_1] - \rho_1|\ U(n)\ |[\rho_1]m_1\rangle \\ &\quad \times e^{im_1\varphi}\,\langle[\rho_2] - \rho_2|\ U(g_D n)\ |[\rho_2]m_2\rangle\, e^{im_2\varphi} \\ &= \delta_{m\,m_1+m_2}N(\rho\rho_1\rho_2) \\ &\quad \times \int_N dn\,\langle[\rho] - \rho_1 + \rho_2|\ U(n)\ |[\rho]m\rangle^* \\ &\quad \times \langle[\rho_1] - \rho_1|\ U(n)\ |[\rho_1]m_1\rangle \\ &\quad \times \langle[\rho_2] - \rho_2|\ U(g_D n)\ |[\rho_2]m_2\rangle; \end{aligned} \qquad (3.13)$$

the matrix elements are given in Eq. (3.11). The factors $-\rho_1+\rho_2$, $-\rho_1$, and $-\rho_2$ appearing as the left indices on the three D functions are so chosen that the D functions will transform properly to the left. g_D is the double coset representative $\left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$.

4. D FUNCTIONS AND CLEBSCH-GORDAN COEFFICIENTS FOR $SL(2, C)$

Both the D functions and Clebsch–Gordan coefficients of $SL(2, C)$ have been previously derived,[7] using techniques different than those given in Sec. 2. For that reason the results for $SL(2, C)$ will only be sketched.

A general element of $SL(2, C)$ can be written $\left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$, $\alpha\beta-\beta\gamma=1$; the inducing subgroup for the principal series of unitary irreducible representations of $SL(2, C)$ is

$$H=NAM=\left\{\begin{pmatrix}\lambda^{-1} & \mu\\ 0 & \lambda\end{pmatrix}\right\} \tag{4.1}$$

with representations

$$h\in H\rightarrow L^{m\rho}=|\lambda|^{i\rho}(\lambda/|\lambda|)^m,\quad \rho \text{ real},$$
$$m \text{ an integer or half integer.}^{14} \tag{4.2}$$

As far as applications to physics are concerned, the representations of $SL(2, C)$ are of most interest in a spherical basis. The D functions and Clebsch–Gordan coefficients will also be computed in this basis. Hence we choose, as right coset representatives, elements of $SL(2, C)$ over the sphere:

$$\begin{pmatrix}\cos\frac{1}{2}\theta, & -e^{-i\varphi}\sin\frac{1}{2}\theta\\ e^{i\varphi}\sin\frac{1}{2}\theta, & \cos\frac{1}{2}\theta\end{pmatrix}\in SL(2, C)$$

gives

$$\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}=\begin{pmatrix}\lambda^{-1} & \mu\\ 0 & \lambda\end{pmatrix}\begin{pmatrix}\cos\frac{1}{2}\theta, & -e^{-i\varphi}\sin\frac{1}{2}\theta\\ e^{i\varphi}\sin\frac{1}{2}\theta, & \cos\frac{1}{2}\theta\end{pmatrix},$$
$$g=hg_c,$$
$$\delta=\lambda\cos\tfrac{1}{2}\theta,\quad \gamma=\lambda e^{i\varphi}\sin\tfrac{1}{2}\theta,$$
$$\beta=\lambda^{-1}\cos\tfrac{1}{2}\theta-\mu e^{-i\varphi}\sin\tfrac{1}{2}\theta,\quad \tan\tfrac{1}{2}\theta'=|\gamma/\delta|,$$
$$\varphi=\arg\gamma-\arg\delta,\quad |\lambda|^2=|\gamma|^2+|\delta|^2,$$
$$\arg\lambda=\arg\delta. \tag{4.3}$$

Then

$$U(g)f(\theta,\varphi)=|\lambda(\theta,\varphi,g)|^{i\rho-2}\times[\lambda(\theta,\varphi,g)/|\lambda|]^m f(\theta',\varphi'), \tag{4.4}$$

where $f(\theta, \varphi)$ has norm

$$\|f\|^2=(4\pi)^{-1}\int d\Omega\,|f(\theta,\varphi)|^2<\infty. \tag{4.5}$$

The coset labels (θ, φ) are moved to (θ', φ') under the action of $g\in SL(2, C)$, where (θ', φ') are defined by $g_c(\theta, \varphi)g=hg_c(\theta', \varphi')$, which can be solved using (4.3). The term $|\lambda|^{-2}$ in Eq. (4.4) is the multiplier, making $U(g)$ unitary in the Hilbert space of functions [Eq. (4.5)].

Now we choose as functions $f(\theta, \varphi)$ the D functions of $SU(2)$, namely $D^J_{mM}(\theta, \varphi)$, $J\geq|m|$, and extend $D^J_{mM}(\theta, \varphi)$ to all elements $g\in SL(2, C)$ by writing

$$e_{JM}(g)=|\lambda(g)|^{i\rho}[\lambda(g)/|\lambda|]^m D^J_{mM}(\theta(g),\varphi(g)). \tag{4.6}$$

The usual D functions of $SL(2, C)$ in a spherical basis are then given by

$$D^{m\rho}_{J'M'JM}(g)=\int_{SL(2,C)/H}dg_c e^*_{J'M'}(g_c)[U(g)e_{JM}(g_c)]$$
$$=(4\pi)^{-1}\int d\Omega e^*_{J'M'}(\theta,\varphi)[U(g)e_{JM}(\theta,\varphi)]. \tag{4.7}$$

However, it is more to the point to find the mixed-basis D functions in order to calculate the Clebsch–Gordan coefficients of $SL(2, C)$. To obtain the mixed-basis D functions, it is necessary to know the H–AM double cosets, where AM consists of matrices of the form

$$\begin{pmatrix}\lambda^{-1} & 0\\ 0 & \lambda\end{pmatrix}.$$

These double cosets are found in the Appendix [Eq. (A8)]. The relevant basis vector is $e_{m'\rho'}(g)$, which is of the form

$$e_{m'\rho'}(g)=|\lambda_1(g)|^{i\rho}[\lambda_1(g)/|\lambda_1|]^m|\lambda_2(g)|^{i\rho'}[\lambda_2(g)/|\lambda_2|]^{m'}, \tag{4.8}$$

with the dependence of λ_1 and λ_2 on g given in Eq. (A9).

Then the mixed-basis D function is

$$D^{m\rho}_{m'\rho'JM}(g)$$
$$=\langle[m\rho]m'\rho'|\,U(g)\,|[m\rho]JM\rangle$$
$$=\int_{SL(2,C)/H}dg_c e^*_{m'\rho'}(g_c)[U(g)e_{JM}(g_c)]$$
$$=(4\pi)^{-1}\int_{\text{sphere}}d\varphi d(\cos\theta)e^*_{m'\rho'}(\theta,\varphi)[U(g)D^J_{mM}(\theta,\varphi)]$$
$$=(4\pi)^{-1}\int_{\text{sphere}}d\varphi d(\cos\theta)\,|\tfrac{1}{2}\sin\theta|^{-\frac{1}{2}i\rho}\,|\cot\tfrac{1}{2}\theta|^{-\frac{1}{2}i\rho'}\times e^{i\varphi/2(m'-m)}[U(g)D^J_{mM}(\theta,\varphi)]. \tag{4.9}$$

Actually, as will be seen, only $D^{m\rho}_{m'\rho'JM}(g_D\mu\Omega)$, where $g_D=\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, $\mu=\left(\begin{smallmatrix} 1 & \mu \\ 0 & 1 \end{smallmatrix}\right)$, and Ω a point on the unit sphere, need be computed. Thus,

$$D^{m\rho}_{m'\rho'JM}(g_D\mu\Omega)=\sum_{M'}(4\pi)^{-1}\int d\varphi d(\cos\theta)e^*_{m'\rho'}(\theta,\varphi)U(g_D)U(\mu)\times D^J_{mM'}(\theta,\varphi)D^J_{M'M}(\Omega). \tag{4.10}$$

With the help of the H–H double cosets [Eq. (A10)], the inducing subgroup occurring in the tensor product decomposition of two irreducible representations, $[m_1\rho_1]$ and $[m_2\rho_2]$, is seen (using Ref. 1) to be $\tilde{H} = AM$. Hence the Clebsch–Gordan coefficients are

$$\begin{aligned}
&\langle [m\rho]JM \mid [m_1\rho_1]J_1M_1;\, [m_2\rho_2]J_2M_2\rangle \\
&= N(m\rho, m_1\rho_1, m_2\rho_2) \\
&\quad\times \int_{SL(2,C)/AM} d\tilde{g}_c D^{m\rho *}_{-m_1+m_2,-\rho_1+\rho_2 JM}(\tilde{g}_c) \\
&\quad\times D^{m_1\rho_1}_{-m_1-\rho_1 J_1 M_1}(\tilde{g}_c) D^{m_2\rho_2}_{-m_2-\rho_2 J_2 M_2}(g_D\tilde{g}_c) \\
&= N\int_N d\mu(4\pi)^{-1}\int d\Omega D^{m\rho *}_{-m_1+m_2-\rho_1+\rho_2 JM}(\mu\Omega) \\
&\quad\times D^{m_1\rho_1}_{-m_1-\rho_1 J_1 M_1}(\mu\Omega) D^{m_2\rho_2}_{-m_2-\rho_2 J_2 M_2}(g_D\mu\Omega) \\
&= N\sum_{M'M_1'M_2'}(4\pi)^{-1}\int d\Omega D^{J*}_{M'M}(\Omega) D^{J_1}_{M_1'M_1}(\Omega) \\
&\quad\times D^{J_2}_{M_2'M_2}(\Omega)\int_N d\mu D^{m\rho *}_{-m_1+m_2-\rho_1+\rho_2 JM'}(\mu) \\
&\quad\times D^{m_1\rho_1}_{-m_1-\rho_1 J_1 M_1'}(\mu) D^{m_2\rho_2}_{-m_2-\rho_2 J_2 M_2'}(g_D\mu) \\
&= N\delta_{M\, M_1+M_2} \\
&\quad\times \sum_{M'M_1'M_2'}\langle [J]M'M_1+M_2 \mid [J_1]M_1'M_1;\, [J_2]M_2'M_2\rangle^2 \\
&\quad\times \int_N d\mu D^{m\rho *}_{-m_1+m_2-\rho_1+\rho_2 JM'}(\mu) \\
&\quad\times D^{m_1\rho_1}_{-m_1-\rho_1 J_1 M_1'}(\mu) D^{m_2\rho_2}_{-m_2-\rho_2 J_2 M_2'}(g_D\mu), \qquad (4.11)
\end{aligned}$$

where $\langle J, M_1'+M_2', M_1+M_2 \mid J_1M_1'M_1; J_2M_2'M_2\rangle$ is an $SU(2)$ Clebsch–Gordan coefficient[15] and N is a normalization factor depending on how the basis states $|[m\rho]JM\rangle$ are normalized.

5. CONCLUSION

In the process of computing D functions and Clebsch–Gordan coefficients a number of assumptions were made restricting the generality of the results. It is interesting to see to what extent the assumptions can be relaxed.

This paper has dealt only with semisimple groups. Presumably most of the nonsemisimple groups have unitary irreducible representations which can be written as induced representations, according to a theorem of Mackey.[16] For example, the theorem has been applied to the nilpotent class of Lie groups[17]; but, for semisimple noncompact groups, it is well known that there exist classes of unitary irreducible representations, such as the discrete and exceptional series, which cannot be written as induced representations. However, such classes of representations can be generated as induced representations by choosing appropriate subgroups. Thus the discrete series can be generated by representations of K, the maximal compact subgroup of G; such representations are realized on the G/K homogeneous spaces.

For these classes of representations the method for computing D functions would not work. However, it may be possible to analytically continue the D functions obtained for the principal series to the discrete series. For $SU(1,1)$ it has been shown that such analytic continuation not only gives the D functions for the discrete series of $SU(1,1)$, but also the D functions of the compact group $SU(2)$.[18] Once the D functions for classes of representations like the discrete and exceptional series are known, it should be possible to calculate the Clebsch–Gordan coefficients.

Other classes of representations of semisimple groups present another difficulty. Classes of representations such as the supplementary series cannot be generated by induced representations acting on homogeneous spaces; rather, the appropriate function spaces are defined in terms of bilinear kernels.[14] For such classes of representations, the Clebsch–Gordan coefficients could not be obtained using the techniques discussed in this paper.

Finally, the restriction to 1-dimensional representations of the inducing subgroup was made. It should be clear that such a restriction is for convenience only, and easily lifted. The Poincaré group probably provides the best example of a group whose inducing representations are not 1 dimensional—whenever the spin for the positive mass representations is nonzero.

With these restrictions in mind, how are the techniques for computing Clebsch–Gordan coefficients presented in Sec. 2 to be compared with the standard method for computing Clebsch–Gordan coefficients, as given, for example, in Ref. 7? The usual formula for Clebsch–Gordan coefficients can be written

$$\begin{aligned}
&\langle [\chi]x \mid [\chi_1]x_1;\, [\chi_2]x_2\rangle \\
&\qquad = N\begin{pmatrix}\chi & \chi_1 & \chi_2\\ x' & x_1' & x_2'\end{pmatrix}\int_G dg D^{\chi *}_{x'x}(g) D^{\chi_1}_{x_1'x_1}(g) D^{\chi_2}_{x_2'x_2}(g),
\end{aligned}$$

where the D functions under the integral are generally chosen to be the ordinary (nonmixed basis) D functions as given, for example, in Eq. (2.10). The important thing to notice about the above formula is that the normalization factor N now depends not only on the irreducible representation labels, but also on the primed indices. If these primed indices are set equal to their respective unprimed ones, it is possible to

show that the above equation can be written

$$|\langle[\chi]x \mid [\chi_1]x_1; [\chi_2]x_2\rangle|^2 = \int_G dg D^{\chi *}_{xx}(g) D^{\chi_1}_{x_1 x_1}(g) D^{\chi_2}_{x_2 x_2}(g),$$

so that, after a phase convention has been chosen, the Clebsch–Gordan coefficient itself is given by a square root of the above integral.

In contrast, the normalization factor appearing in Eq. (2.14) depends only on the irreducible representation labels, and thus, after a phase convention has been chosen, the Clebsch–Gordan coefficients are given by an integral over $G/\tilde{H}$. The formulas are then simpler than those given by the standard formula for Clebsch–Gordan coefficients, but, on the other hand, less general in that Eq. (2.14) can be written only for representations which can be written as induced representations.

Finally, if in the tensor product reduction a representation appears more than once (multiplicity greater than one), the standard Clebsch–Gordan formula provides no means for distinguishing between these representations, and hence an ambiguity appears. In contrast, Eq. (2.14) contains a multiplicity parameter η which generally will handle the multiplicity problem. In fact, as shown in Refs. 1, it is possible to compute the Clebsch–Gordan coefficients arising from the n-fold tensor product reduction, where the multiplicity generally is greater than one. For example, in the Poincaré group, a spin multiplicity already appears in coupling two positive mass nonzero spin representations together. When more than two representations are coupled together, enormous mass multiplicity can also occur. All such multiplicity is readily handled in Eq. (2.14).

ACKNOWLEDGMENTS

The author wishes to thank Dr. E. Thieleker of Argonne National Laboratory for many helpful discussions. He also wishes to thank Dr. C. Kaysen for the hospitality of the Institute for Advanced Study.

APPENDIX

Various double coset decompositions of $SU(1, 1)$ and $SL(2, C)$ are required to compute the mixed-basis D functions and Clebsch–Gordan coefficients. This appendix will show how such decompositions are obtained.

Consider first $SU(1, 1)$. What is required are the S–A double cosets, in order to compute the mixed-basis D functions [Eq. (3.8)]. It seems easiest to construct the homogeneous space $SU(1, 1)/A$ and then compute the action of S on points of $SU(1, 1)/A$.

$SU(1, 1)/A$ can be realized as the following manifold. Consider the transformations

$$H(x') = gH(x)g^\dagger, \quad g \in SU(1, 1),$$

and

$$H(x) = \begin{pmatrix} z & x - iy \\ x + iy & z \end{pmatrix};$$

then it is readily seen that g carries a point $H(x)$ into $H(x')$. Further, the point $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ can be chosen as a stability point since it is left invariant by A [recall that A consists of matrices of the form $\left(\begin{smallmatrix} \cosh\theta & i\sinh\theta \\ -i\sinh\theta & \cosh\theta \end{smallmatrix}\right)$]. The space is, of course, nothing other than the hyperboloid $z^2 - (x^2 + y^2) = -1$.

An arbitrary element of S moves a fixed but arbitrary point $H(x_0)$ of $SU(1, 1)/A$ into $H(x)$, where

$$x = x_0 + (z_0 + y_0)2ne^{-\theta},$$

$$y = y_0 e^{2\theta} - x_0 2ne^{\theta} - (z_0 + y_0)(2n^2 + \sinh 2\theta). \quad \text{(A1)}$$

Using (A1), we can show that all points $H(x)$, $x^2 + y^2 \geq 1$, can be reached from three points (which will fix the double cosets), one of which is the stability point. It is the quantity $(z_0 + y_0)$ which distinguishes the various regions, chosen in the following way:

Region I, $z_0 + y_0 = 0$, corresponds to the point $\begin{pmatrix} 0 & \pm 1 \\ \pm 1 & 0 \end{pmatrix}$;

Region II, $z_0 + y_0 > 0$, generates $H(x)$ such that $y > 0$, x arbitrary; (A2)

Region III, $z_0 + y_0 < 0$, generates $H(x)$ such that $y < 0$, x arbitrary.

Region I corresponds to the identity double coset. A convenient choice of points generating Regions II and III is $z_0 = x_0 = 0$, $y_0 = 1$ (Region II), $y_0 = -1$ (Region III). An element of $SU(1, 1)$ which carries the stability point to the generating points for Regions II and III [that is, $\pm\left(\begin{smallmatrix} 0 & -i \\ -i & 0 \end{smallmatrix}\right)$] can be chosen to be

$$\begin{pmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{pmatrix} \quad \text{with} \quad \varphi = \pm\tfrac{3}{4}\pi.$$

Thus, we have three double cosets

$$S\begin{pmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{pmatrix} A,$$

with $\varphi = 0$ corresponding to Region I, $\varphi = \frac{3}{4}\pi$ corresponding to Region II, and $\varphi = -\frac{3}{4}\pi$ corresponding to Region III.

In order to be able to write an arbitrary element of $SU(1,1)$ in terms of the S–A double coset decomposition

$$\begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} = s(\theta_1, n)\begin{pmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{pmatrix} s(\theta_2, 0), \quad \text{(A3)}$$

with φ taking on the three values corresponding to the three double cosets, it is necessary to make sure that each element of $SU(1,1)$ is uniquely determined by θ_1, n, θ_2, and φ. From (A1) it is readily shown that, if $\cosh 2\theta > 2n^2$, uniqueness holds. Which of the three double cosets are to be used in (A2) depends on the value of α and β; in particular,

$$\begin{aligned} z_0 + y_0 &= 2\,\mathrm{Re}\,\alpha\beta^* + \mathrm{Im}\,(\alpha^{2*} + \beta^{2*}) \\ &= -2(\alpha_R - \beta_I)(\alpha_I - \beta_R), \end{aligned} \quad \text{(A4)}$$

which, according to (A2), determines the region of the hyperboloid and hence the double coset.

When (A3) is written out, for the two nontrivial double cosets one obtains

$$e^{2\theta_1} = \mp \frac{1}{2(\alpha_R - \beta_I)(\alpha_I - \beta_R)}, \quad e^{2\theta_2} = \mp \frac{\alpha_I - \beta_R}{\alpha_R - \beta_I}, \quad \text{(A5)}$$

where the upper sign refers to Region II ($z_0 + y_0 > 0$), the lower sign to Region III ($z_0 + y_0 < 0$). From Eqs. (A4) and (A5) it is thus possible to compute $\langle [\rho]\rho' \mid [\rho]m \rangle$ [Eq. (3.10)].

To compute the Clebsch–Gordan coefficients of $SU(1,1)$, it is also necessary to know the S–S double coset decomposition. It has been shown[19] that there are a finite number of such double cosets, the number being equal to the order of the Weyl group, defined as the quotient of the normalizer of K divided by the centralizer of K.[12] Further, one of these double cosets is dense in G. For $SU(1,1)$ the order of the Weyl group is 2, so that besides the identity double coset, which is obviously never dense in G, the Weyl group yields the element $\left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$ as a suitable choice of double coset representative; then

$$S\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} S \quad \text{(A6)}$$

is dense in $SU(1,1)$.

Since Gel'fand[20] has listed many of the homogeneous spaces of $SL(2,C)$, the relevant double coset decompositions are easier to work out than for those of $SU(1,1)$. In particular, it is possible to realize the homogeneous space $SL(2,C)/H$, where H is the inducing subgroup for the principal series [Eq. (4.1)], as the set of complex numbers k (including ∞) with the group action given by

$$k' = (\alpha k + \gamma)/(\beta k + \delta). \quad \text{(A7)}$$

With $k = 0$ chosen as the stability point, it is clear that the stability group is H.

Then the H–AM double coset decomposition can be given by the following representatives:

Region I:

when acted on by AM the point $k = 1$ is sent into λ^{-2}, which covers all values of the manifold except $k = 0$ and $k = \infty$

Region II:

the identity element (Region III, with $k = \infty$ will be ignored). (A8)

Since the point $k = 1$ can be reached from $k = 0$ by the element $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$, $H\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)AM$ is dense in $SL(2,C)$. Finally for $g \in SL(2,C)$, $h_1 \in H$, $h_2 \in AM$, and $g_D = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$ ($g \notin H$, $\delta \neq 0$),

$$\begin{gathered} g = h_1(g) g_D(g) h_2(g), \\ \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} \lambda_1^{-1} & \mu_1 \\ 0 & \lambda_1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \lambda_2^{-1} & 0 \\ 0 & \lambda_2 \end{pmatrix}, \\ \delta = \lambda_1\lambda_2, \quad \gamma = \lambda_1\lambda_2^{-1}, \quad \beta = \mu\lambda_2, \\ \lambda_1^2 = \delta\gamma, \quad \lambda_2^2 = \delta/\gamma. \end{gathered} \quad \text{(A9)}$$

The H–H double cosets, required for the Clebsch–Gordan coefficients, can be obtained either from the Weyl group (again of order 2) or from the $SL(2,C)/H$ homogeneous space. The two double coset representatives can be chosen to be the identity element and $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, so that

$$H\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} H \quad \text{(A10)}$$

is dense in $SL(2,C)$.

* Research sponsored by the National Science Foundation, Grant No. GP-16147.

[1] W. H. Klink, J. Math. Phys. **10**, 1477 (1969); *Lectures in Theoretical Physics* (Gordon and Breach, New York, 1969), Vol. XID, pp. 43–76.

[2] See, for example, G. W. Mackey, "The Theory of Group Representations," University of Chicago, Department of Mathematics Report, 1956 (unpublished).

[3] W. H. Klink, Phys. Rev. **181**, 2023, 2027 (1969).

[4] A small sampling of the current literature would include A. Sciarrino and M. Toller, J. Math. Phys. **8**, 252 (1967); M. Toller, Nuovo Cimento **53A**, 671 (1968); F. H. Hadjioannou, Nuovo Cimento **44**, 185 (1966); D. Z. Freedman and J. M. Wang, Phys. Rev. **153**, 1596 (1967); Y. Nambu, Phys. Rev. **160**, 1171 (1967).

[5] L. Pukanszky, Trans. Am. Math. Soc. **100**, 116 (1961); M. A. Naimark, Am. Math. Soc. Transl. (2) **36**, 101 (1964).

[6] I. Ferretti and M. Verde, Nuovo Cimento **55**, 110 (1968); S. S. Sannikov, Dokl. Akad. Nauk SSSR **171**, 1058 (1966) [Sov. Phys. Dokl. **11**, 1045 (1967)]; M. Andrews and J. Gunson, J. Math. Phys. **5**, 1391 (1964); W. J. Holman III and L. C. Biedenharn, Jr., Ann. Phys. **47**, 205 (1968); **39**, 1 (1966); K. Wang, UCRL Preprint No. UCRL-19306, 1969.

[7] R. L. Anderson, R. Rączka, M. A. Rashid, and P. Winternitz, J. Math. Phys. **11**, 1050 (1970).

[8] I. M. Gel'fand and G. E. Shilov, *Generalized Functions* (Academic, New York, 1964), Vol. 1.

[9] G. Rideau, Ann. Inst. Henri Poincaré **3**, 339 (1965).

[10] L. C. Biedenharn, "Group Theoretical Approaches to Nuclear Spectroscopy," in *Lectures in Theoretical Physics Vol. V* (Interscience, New York, 1963).

[11] The restriction to 1-dimensional representations is for convenience only, since both $SU(1, 1)$ and $SL(2, C)$ have inducing subgroups whose representations are 1 dimensional.

[12] See, for example, R. Hermann, *Lie Groups for Physicists* (Benjamin, New York, 1966).

[13] V. Bargmann, Ann. Math. **48**, 568 (1947).

[14] In general, we are following the notation of M. A. Naimark, *Linear Representations of the Lorentz Group* (Macmillan, New York, 1964); the current physics literature seems to use $[k_0, c]$ to label the irreducible representations of $SL(2, C)$ rather than $[m, \rho]$. Naimark shows that the two labels are related by $k_0 = |\frac{1}{2}m|, c = -i(\text{sgn } m)\frac{1}{2}\theta$.

[15] See, for example, N. J. Vilenkin, *Special Functions and the Theory of Group Representations* (American Mathematical Society, Providence, R.I., 1968).

[16] The relevant theorem is what Mackey calls the imprimitivity theorem, extended to projective representations; see Ref. 2.

[17] A. A. Kirillov, Usp. Mat. Nauk **106**, 57 (1962) [Russ. Math. Surv. **17**, 53 (1962)]; for solvable groups see L. Auslander and C. Moore, Mem. Amer. Math. Soc. **62**, 1 (1966).

[18] See, for example, J. F. Boyce, J. Math. Phys. **8**, 675 (1967).

[19] Private communication with Dr. E. Thieleker, Argonne National Laboratory; F. Bruhat, Bull. Soc. Math. France **84**, 97 (1956).

[20] I. M. Gel'fand, M. I. Graev, and N. Y. Vilenkin, *Generalized Functions* (Academic, New York, 1966), Vol. 5.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

N-Representability Problem: Particle-Hole Equivalence*

MARY BETH RUSKAI
Institut de Physique Théorique, Université de Genève and Battelle Institute Advanced Studies Center, Geneva, Switzerland

(Received 17 March 1970)

It is shown that every p matrix has a dual matrix which describes the p-hole properties of the system. A general procedure is given for using particle-hole equivalence to obtain new N-representability conditions. In particular, necessary and sufficient conditions are given for both pure and ensemble N-representability of p matrices whose 1-rank is $N + p$.

I. INTRODUCTION

Recently, there has been considerable interest in the so-called N-representability problem.[1–3] Many necessary conditions are known, but sufficient conditions have been given in only a few very special cases. In this paper, particle–hole equivalence is used to obtain a general procedure for deriving new N-representability conditions from known conditions. The procedure is applicable to both necessary conditions and to sufficient conditions. In particular, one can obtain necessary and sufficient conditions for both pure and ensemble N-representability of p matrices whose 1-rank is $N + p$.

It is well known that by considering 1-particle states to be occupied by either particles or holes, one can obtain two completely equivalent descriptions of a quantum mechanical system. However, this equivalence has never been used explicitly to study the N-representability problem. Even the interpretation of Garrod's Q-matrix condition[4] as a dual hole condition does not seem to have been given previously. Because the pth-order density matrix is explicitly constructed to describe p-particle properties, the particle–hole equivalence is obscured when one uses reduced density matrices to describe the properties of a system. To describe p-hole properties, one must consider a different, but related, kernel called the p-hole matrix.[5] The second quantization formalism provides an elegant language for discussing particle–hole equivalence and for defining the dual p-particle and p-hole matrices. Nevertheless, we prefer to begin with the more pedestrian Slater determinant approach; and then make the simple connection with the second quantization approach.

A review of relevant terminology is given below. An N-particle *ensemble density matrix* ρ is defined as

$$\rho = \sum_m \alpha_m \Psi^s{}_m(1 \cdots N)\Psi^s{}^*_m(1' \cdots N'), \tag{1}$$

where

$$0 \leq \alpha_m \leq 1, \tag{2a}$$

$$\sum_m \alpha_m = 1, \tag{2b}$$

and $\Psi^s{}_m$ is an antisymmetric N-particle function. ρ describes a *pure state* when the sum in (1) consists of a single term, i.e.,

$$\rho = \Psi^s\Psi^s{}^*. \tag{3}$$

The *pth-order reduced density matrix* of ρ is defined[7] as

$$D^p(x; x') = \int \rho(x, y; x', y)\, dy, \tag{4}$$

where x stands for all space and spin coordinates of particles $1 \cdots p$ and y for those of particles $p + 1 \cdots N$. ρ (or Ψ^s if ρ is pure) is called the pre-image of

[10] L. C. Biedenharn, "Group Theoretical Approaches to Nuclear Spectroscopy," in *Lectures in Theoretical Physics Vol. V* (Interscience, New York, 1963).

[11] The restriction to 1-dimensional representations is for convenience only, since both $SU(1, 1)$ and $SL(2, C)$ have inducing subgroups whose representations are 1 dimensional.

[12] See, for example, R. Hermann, *Lie Groups for Physicists* (Benjamin, New York, 1966).

[13] V. Bargmann, Ann. Math. **48**, 568 (1947).

[14] In general, we are following the notation of M. A. Naimark, *Linear Representations of the Lorentz Group* (Macmillan, New York, 1964); the current physics literature seems to use $[k_0, c]$ to label the irreducible representations of $SL(2, C)$ rather than $[m, \rho]$. Naimark shows that the two labels are related by $k_0 = |\frac{1}{2}m|, c = -i(\text{sgn } m)\frac{1}{2}\theta$.

[15] See, for example, N. J. Vilenkin, *Special Functions and the Theory of Group Representations* (American Mathematical Society, Providence, R.I., 1968).

[16] The relevant theorem is what Mackey calls the imprimitivity theorem, extended to projective representations; see Ref. 2.

[17] A. A. Kirillov, Usp. Mat. Nauk **106**, 57 (1962) [Russ. Math. Surv. **17**, 53 (1962)]; for solvable groups see L. Auslander and C. Moore, Mem. Amer. Math. Soc. **62**, 1 (1966).

[18] See, for example, J. F. Boyce, J. Math. Phys. **8**, 675 (1967).

[19] Private communication with Dr. E. Thieleker, Argonne National Laboratory; F. Bruhat, Bull. Soc. Math. France **84**, 97 (1956).

[20] I. M. Gel'fand, M. I. Graev, and N. Y. Vilenkin, *Generalized Functions* (Academic, New York, 1966), Vol. 5.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

N-Representability Problem: Particle-Hole Equivalence*

MARY BETH RUSKAI
Institut de Physique Théorique, Université de Genève and Battelle Institute Advanced Studies Center, Geneva, Switzerland

(Received 17 March 1970)

It is shown that every p matrix has a dual matrix which describes the p-hole properties of the system. A general procedure is given for using particle-hole equivalence to obtain new N-representability conditions. In particular, necessary and sufficient conditions are given for both pure and ensemble N-representability of p matrices whose 1-rank is $N + p$.

I. INTRODUCTION

Recently, there has been considerable interest in the so-called N-representability problem.[1–3] Many necessary conditions are known, but sufficient conditions have been given in only a few very special cases. In this paper, particle–hole equivalence is used to obtain a general procedure for deriving new N-representability conditions from known conditions. The procedure is applicable to both necessary conditions and to sufficient conditions. In particular, one can obtain necessary and sufficient conditions for both pure and ensemble N-representability of p matrices whose 1-rank is $N + p$.

It is well known that by considering 1-particle states to be occupied by either particles or holes, one can obtain two completely equivalent descriptions of a quantum mechanical system. However, this equivalence has never been used explicitly to study the N-representability problem. Even the interpretation of Garrod's Q-matrix condition[4] as a dual hole condition does not seem to have been given previously. Because the pth-order density matrix is explicitly constructed to describe p-particle properties, the particle–hole equivalence is obscured when one uses reduced density matrices to describe the properties of a system. To describe p-hole properties, one must consider a different, but related, kernel called the p-hole matrix.[5] The second quantization formalism provides an elegant language for discussing particle–hole equivalence and for defining the dual p-particle and p-hole matrices. Nevertheless, we prefer to begin with the more pedestrian Slater determinant approach; and then make the simple connection with the second quantization approach.

A review of relevant terminology is given below. An N-particle *ensemble density matrix* ρ is defined as

$$\rho = \sum_m \alpha_m \Psi^s{}_m(1 \cdots N)\Psi^s{}^*_m(1' \cdots N'), \tag{1}$$

where

$$0 \leq \alpha_m \leq 1, \tag{2a}$$

$$\sum_m \alpha_m = 1, \tag{2b}$$

and $\Psi^s{}_m$ is an antisymmetric N-particle function. ρ describes a *pure state* when the sum in (1) consists of a single term, i.e.,

$$\rho = \Psi^s\Psi^s{}^*. \tag{3}$$

The *pth-order reduced density matrix* of ρ is defined[7] as

$$D^p(x; x') = \int \rho(x, y; x', y)\, dy, \tag{4}$$

where x stands for all space and spin coordinates of particles $1 \cdots p$ and y for those of particles $p + 1 \cdots N$. ρ (or Ψ^s if ρ is pure) is called the pre-image of

D^p. A *p matrix*, also denoted $D^p(x; x')$, is a function of $2p$ particles which is Hermitian, nonnegative, antisymmetric in each set of indices, and satisfies Tr $D^p = 1$. A p matrix is said to be *pure or ensemble N-representable* if it can be derived from an antisymmetric N-particle pure state or ensemble according to (4). The 1-*rank* R of a p matrix is the number of nonzero eigenvalues of its 1-matrix. The eigenfunctions of the 1-matrix are called *natural spin orbitals (NSO)*.

The hole matrix is discussed in Sec. II; p matrices with $R = N + p$ in Sec. III; and general N-representability conditions in Sec. IV.

II. THE HOLE MATRIX

Before defining the hole matrix, it will be useful to derive a new expression for the expansion coefficients of D^p. The hole matrix is then discussed in the second part of this section. Finally, the hole matrix is considered in the second quantization formalism.

A Useful Formula

One can expand D^p in a set $\{f_k\}$ of R orthonormal 1-particle states as

$$D = \binom{N}{p}^{-1} \sum_{k_1<k_2\cdots<k_p} \sum_{l_1<l_2\cdots<l_p} d^p_{k_1\cdots k_p, l_1\cdots l_p} \times [k_1\cdots k_p][l_1\cdots l_p]'^*, \quad (5)$$

where $[k_1\cdots k_p]$ is the Slater determinant[8] formed from $f_{k_1}\cdots f_{k_p}$. The expansion of Ψ in 1-particle states can be written as

$$\Psi = \sum_{k_1<k_2\cdots<k_N} b_{k_1\cdots k_N}[k_1\cdots k_N], \quad (6)$$

but it will be more useful to write it in the form

$$\Psi = \sum_{k_1<k_2\cdots<k_M} c_{k_1\cdots k_M}\Phi_{k_1\cdots k_M}, \quad (7)$$

where $R = M + N$ and $\Phi_{k_1\cdots k_M}$ is the Slater determinant[8] which does *not* contain the states $f_{k_1}\cdots f_{k_M}$. It will be convenient to use, instead of $c_{k_1\cdots k_M}$, the antisymmetric tensor $x_{k_1\cdots k_M}$ defined by

$$x_{k_1\cdots k_M} = (-)^{k_1+k_2+\cdots k_M} c^*_{k_1\cdots k_M} \quad \text{when} \quad k_1 < k_2 < \cdots < k_M. \quad (8)$$

To further simplify the notation, capital letters will be used to indicate ordered sets of indices. In particular:

$$\begin{aligned} I_\mu &= i_1\cdots i_\mu, & i_1 < i_2 < \cdots < i_\mu,\\ J_\tau &= j_1\cdots j_\tau, & j_1 < j_2 < \cdots < j_\tau,\\ K &= k_1\cdots k_\nu, & k_1 < k_2 < \cdots < k_\nu,\\ L &= l_1\cdots l_\nu, & l_1 < l_2 < \cdots < l_\nu,\\ M_\sigma &= m_1\cdots m_\sigma, & m_1 < m_2 < \cdots < m_\sigma,\\ N_{\mu-\sigma} &= n_1\cdots n_{\mu-\sigma}, & n_1 < n_2 < \cdots < n_{\mu-\sigma}. \end{aligned}$$

One writes $I \subset J$ if every i in I is equal to some j in J, $I \cap J \neq 0$ if there is at least one i in I and one j in J with $i = j$, and $I \cap J = 0$ if $i \neq j$ for all i in I and j in J. Finally, define $D^0 = 1$.

First, let D^p be the pth-order reduced density matrix whose pre-image is given by (7), and consider only those elements of D^p which have the first τ indices equal (and all others unequal). Then

$$d^p_{J_\tau K, J_\tau L} = d^p_{j_1\cdots j_\tau k_1\cdots k_\nu, j_1\cdots j_\tau l_1\cdots l_\nu} = (-)^\nu \sum_{\substack{I_\mu \\ I_\mu\cap J_\tau = 0}} x_{I_\mu K} x^*_{I_\mu L} \quad (9)$$

$$= (-)^\nu \sum_{I_\mu} x_{I_\mu K} x^*_{I_\mu L} - (-)^\nu \sum_{\substack{I_\mu \\ I_\mu\cap J_\tau \neq 0}} x_{I_\mu K} x^*_{I_\mu L} \quad (10)$$

$$= (-)^\nu \Big(d^\nu_{K,L} - \sum_{m\subset J_\tau} S[m] + \sum_{M_2\subset J_\tau} S[M_2] - \cdots + \cdots (-)^\tau S[J_\tau] \Big), \quad (11)$$

where $\tau + \nu = p$, $\mu + \nu = M = R - N$, and

$$S[M_\sigma] = \sum_{N_{\mu-\sigma}} x_{M_\sigma N_{\mu-\sigma} K} x^*_{M_\sigma N_{\mu-\sigma} L}. \quad (12)$$

To derive (11) from (10), one must verify that, when $I_\mu \cap J_\tau \neq 0$, the term $x_{I_\mu K} x^*_{I_\mu L}$ has been subtracted exactly once. If I_μ has exactly κ elements in J_τ, the term $x_{I_\mu K} x^*_{I_\mu L}$ occurs in the sum over $S[M_\sigma]$ exactly $\binom{\kappa}{\sigma}$ times. Since[9]

$$\sum_{\sigma=1}^{\kappa} (-)^\sigma \binom{\kappa}{\sigma} = -1,$$

each term has indeed been counted exactly once. Equation (11) can be inverted to obtain an expression for $S[J_\sigma]$, i.e.,

$$(-)^\sigma S[J_\sigma] = (-)^\nu d^q_{J_\sigma K, J_\sigma L} - d^\nu_{K,L} + \sum_{m\subset J_\sigma} S[m] - \cdots + \cdots (-)^\sigma \sum_{M_{\sigma-1}\subset J_\sigma} S[M_{\sigma-1}], \quad (13)$$

where $q = \sigma + \nu$. Now one can use (13) to successively eliminate $S[M_\sigma]$ terms from (11). One substitutes for all $S[M_{\tau-1}]$ terms, then for $S[M_{\tau-2}]$, etc. Each substitution leaves only lower-order $S[M_\sigma]$ terms. Finally, $d^p_{J_\tau K, J_\tau L}$ will be written as a linear combination of $S[J_\tau]$ and elements of D^q with $\nu \leq q < p$. In fact,

$$d^p_{J_\tau K, J_\tau L} = \sum_{I_{\tau-1}\subset J_\tau} d^{p-1}_{I_{\tau-1}K, I_{\tau-1}L} - \sum_{I_{\tau-2}\subset J_\tau} d^{p-2}_{I_{\tau-2}K, I_{\tau-2}L} + \cdots - \cdots (-)^{\tau+1} d^\nu_{K,L} + (-)^{\tau+\nu} S[J_\tau]. \quad (14)$$

If $d^p_{m_1\cdots m_p,n_1\cdots n_p}$ has any arbitrary μ indices equal, one can derive an expression differing from (14) in only + or − signs. Simply define

$$\hat{d}^q_{I_\sigma K,I_\sigma L} = \epsilon_1\epsilon_2 d^q_{m_1\cdots m_q,n_1\cdots n_q}, \tag{15}$$

where

$$I_\sigma K = \mathfrak{I}_1(m_1\cdots m_q), \quad m_1 < m_2\cdots < m_q,$$
$$I_\sigma L = \mathfrak{I}_2(n_1\cdots n_q), \quad n_1 < n_2\cdots < n_q,$$

$\mathfrak{I}_1$ and $\mathfrak{I}_2$ being permutations, and where ϵ_j is the parity of $\mathfrak{I}_j$. Then (14) is correct if $d^q_{I_\sigma K,I_\sigma L}$ is replaced by $\hat{d}^q_{I_\sigma K,I_\sigma L}$.

Properties of the Hole Matrix

The general hole matrix obtained from an arbitrary p matrix is now considered.

Definition 1: The *pth-order hole matrix* η^p corresponding to a p matrix D^p is defined by

$$\eta^p(1\cdots p;1'\cdots p') = \binom{R-N}{p}^{-1} \sum_{m_1<m_2\cdots<m_p} \sum_{n_1<n_2\cdots<n_p} h^{p*}_{m_1\cdots m_p,n_1\cdots n_p} \times [m_1\cdots m_p][n_1\cdots n_p]'^*, \tag{16}$$

when $N+p \le R < \infty$ and

$$h^p_{m_1\cdots m_p,n_1\cdots n_p} = \epsilon_1\epsilon_2 \hat{h}^p_{J_\tau K,J_\tau L} \tag{17}$$

$$= \epsilon_1\epsilon_2 \sum_{\sigma=0}^{\tau}\left((-)^{\sigma+\nu}\sum_{I_\sigma\subset J_\tau} \hat{d}^{\sigma+\nu}_{I_\sigma K,I_\sigma L}\right), \tag{18}$$

where ϵ_1, ϵ_2, J_τ, K, and L are related to $m_1\cdots m_p$ and $n_1\cdots n_p$ as in (15). The matrix of expansion coefficients, $\mathbf{H}^p = \{h^p_{m_1\cdots m_p,n_1\cdots n_p}\}$, will also be called the hole matrix. Although η^p is well defined only when $N+p \le R < \infty$, $\mathbf{H}^p$ is well defined for all R.

For the important cases $p=1$ and $p=2$, (18) becomes

$$h^1_{i,j} = \delta_{ij} - d^1_{i,j}, \tag{19}$$

$$h^2_{ij,kl} = d^2_{ij,kl} - \delta_{ik}d^1_{jl} + \delta_{jk}d^1_{il} - \delta_{jl}d^1_{ik} + \delta_{il}d^1_{jk} + \delta_{ik}\delta_{jl}. \tag{20}$$

It is clear from (20) that η^2 is identical (except for normalization) to the Q matrix discussed by Garrod.[4,10]

It is not immediately obvious from its definition that η^p describes the hole properties of a system. The next few theorems show that η^p has properties which are analogous to those of D^p. Theorem 4 then shows why η^p really describes hole properties.

Theorem 1: The hole matrix η^p has the following properties:

(a) η^p is Hermitian;

(b) $\operatorname{Tr}\eta^p = 1$ or $\operatorname{Tr}\mathbf{H}^p = \binom{R-N}{p}$; (21)

(c) η^p is antisymmetric in each set of indices;

(d) $$\int \eta^p(1\cdots p-1,p;1'\cdots p-1',p)\,dp = \eta^{p-1}. \tag{22}$$

Proof: Proofs of (a) and (c) are trivial. To prove (b), note that

$$\sum_{J_p}\left(\sum_{I_q\subset J_p}\hat{d}^q_{I_q,I_q}\right) = \binom{N}{q}\binom{R-q}{p-q}.$$

Some simple manipulations with Gould's formulas[9] give

$$\sum_{\sigma=0}^{p}(-)^q\binom{N}{q}\binom{R-q}{p-q} = \binom{R-N}{p},$$

which proves (b). Part (d) can be proved similarly. An easier way is to use Theorem 3 to first prove (d) when D^p has a pure pre-image and then extend this to arbitrary p matrices.

Theorem 2: Every hole matrix corresponds to a unique p matrix, i.e., if $D^p_1 \neq D^p_2$, then $\eta^p_1 \neq \eta^p_2$.

Proof: Using (22), one can determine all lower-order η^q, $q<p$, from η^p. Therefore, one can determine all D^q with $q<p$ so that (18) can be inverted to give D^p from η^p.

Theorem 3: If D^p has a pure state pre-image, then every element of its hole matrix $\mathbf{H}^p$ satisfies

$$h^p_{k_1\cdots k_p,l_1\cdots l_p} = \epsilon_1\epsilon_2 S[J_\tau] = \sum_I x_{KI}x^*_{LI}, \tag{23}$$

where

$$K = k_1\cdots k_p,$$
$$L = l_1\cdots l_p,$$
$$I = i_1\cdots i_{M-p}, \quad M = R-N,$$
$$J_\tau = K\cap L.$$

In particular, if $M=p$,

$$h^p_{k_1\cdots k_p,l_1\cdots l_p} = x_{k_1\cdots k_p}x^*_{l_1\cdots l_p}. \tag{24}$$

Note that Eq. (23) is completely analogous to the equation

$$d^p_{k_1\cdots k_p,l_1\cdots l_p} = \sum_{i_1\cdots i_{N-p}} b_{k_1\cdots k_p i_1\cdots i_{N-p}}b^*_{l_1\cdots l_p i_1\cdots i_{N-p}}, \tag{25}$$

where $b_{j_1\cdots j_N}$ is defined by (6) and antisymmetry.

The Second Quantization Approach

One can obtain the elements of any reduced-density matrix D^p as expectation values of second quantization operators. Then, in the state $|\alpha\rangle$,

$$d^p_{k_1\cdots k_p, l_1\cdots l_p} = \langle\alpha|\, a^\dagger_{k_p}\cdots a^\dagger_{k_1} a_{l_1}\cdots a_{l_p}\,|\alpha\rangle, \quad (26)$$

where $a^\dagger_i$ and a_j are the usual fermion creation and annihilation operators. Since D^p describes all p-particle properties of the state $|\alpha\rangle$, one expects the p-hole properties to be given by something of the form

$$\langle\alpha|\, b^\dagger_{k_p}\cdots b^\dagger_{k_1} b_{l_1}\cdots b_{l_p}\,|\alpha\rangle, \quad (27)$$

where $b^\dagger_j$ creates a hole in the jth state and b_i annihilates a hole. The next theorem states that $\mathbf{H}^p$ does indeed have the form (27).

Theorem 4: If D^p has a pure state pre-image $|\alpha\rangle$, then every element of its hole matrix $\mathbf{H}^p$ satisfies

$$h^p_{k_1\cdots k_p, l_1\cdots l_p} = \langle\alpha|\, a_{k_p}\cdots a_{k_1} a^\dagger_{l_1}\cdots a^\dagger_{l_p}\,|\alpha\rangle. \quad (28)$$

Proof: This is really just Theorem 3 in a new notation. One can also obtain a more direct proof by using the anticommutation relation

$$a^\dagger_i a_j + a_j a^\dagger_i = \delta_{ij}$$

to change (26) to (28); this is simply equivalent to deriving (14) in a different notation. Since a_j can be considered as either annihilating a particle or creating a hole, one can identify $b^\dagger_j$ with a_j. Thus (28) shows that η^p (or $\mathbf{H}^p$) has exactly the form required to represent hole properties.

III. N-REPRESENTABILITY: $R \leq N + p$

In this section N-representability conditions are given for p matrices with $R = N + p$.

For $R < N + p$ one can always formally add unoccupied states (completely filled holes) to the set of 1-particle states. Then

$$\hat{\mathbf{H}}^p = \begin{pmatrix} \mathbf{H}^p & 0 \\ 0 & \mathbf{\Lambda} \end{pmatrix}, \quad (29)$$

where

$\hat{\mathbf{H}}^p$ has $R = N + p$,

$\mathbf{H}^p$ has $R < N + p$,

$\mathbf{\Lambda}$ is a nonnegative matrix.

Theorems 5–8 can then be applied to $\hat{\mathbf{H}}^p$. In fact, one can show that a necessary condition for ensemble N-representability when $R < N + p$ is that $\mathbf{H}^p = \mathbf{0}$, so that Theorems 5–8 must actually be valid for $\mathbf{\Lambda}$.

Pure States

The solution to the pure N-representability problem can be stated quite simply in terms of the hole matrix.

Theorem 5: A p matrix D^p with $R = N + p$ is pure N-representable if and only if its hole matrix is idempotent, i.e., $(\eta^p)^2 = \eta^p$.

Proof: Necessity follows directly from (24). For sufficiency, note that the matrix $\mathbf{H}^p = \{h_{k_1\cdots k_p, l_1\cdots l_p}\}$ satisfies

$$\mathrm{Tr}\,\mathbf{H}^p = \binom{R-N}{p} = \binom{p}{p} = 1. \quad (30)$$

Then idempotency implies that $\mathbf{H}^p$ can be written as

$$\mathbf{H}^p = \mathbf{w}\mathbf{w}^\dagger, \quad (31)$$

where $\mathbf{w}$ is a normalized column vector

$$\mathbf{w} = \{w_{k_1\cdots k_p}(k_1 < k_2 \cdots < k_p)\}.$$

Now let Ψ be given by (7) with

$$c_{k_1\cdots k_p} = (-)^{k_1+k_2+\cdots+k_p} w^*_{k_1\cdots k_p}. \quad (32)$$

Then η^p is the p-hole matrix corresponding to Ψ, and Theorem 2 implies that Ψ is the pre-image of $\mathbf{D}^p$.

Recently, Yoseloff and Kuhn[11] have given a necessary condition for pure N-representability when $R = N + p$ in the form of an inequality on the diagonal elements of D^p. Theorem 5 implies that this condition needs not be checked explicitly when testing for N-representability; when η^p is idempotent, Yoseloff's inequality is automatically satisfied. In particular, when $p = 2$, Yoseloff's inequality was previously shown[4] to follow directly from the fact that, when η^2 is idempotent, it will also be nonnegative.

It is interesting to note that the case $p = 2$, $R = N + 2$ has been solved previously. The solution depends on the fact that the eigenvalues of the 1-matrix must satisfy a paired degeneracy condition which is reflected in the natural expansion of the pre-image.[12] In the absence of extra degeneracy, degenerate NSO's will always be paired in the pre-image.[13] Then the solution of the pure N-representability problem for D^2 can be obtained as a special case of a known result for a class of functions with special pairing properties.[14,15]

Ensembles

The connection between the theory of convex sets and the N-representability problem is well known[1] and has been discussed in considerable detail by Kummer.[16] It provides a means of relating the pure and ensemble N-representability problems. The space of ensemble N-representable p matrices $\mathfrak{P}^p_N$ is convex, and the extreme points of $\mathfrak{P}^p_N$ are contained in the set of pure N-representable p matrices. There are some topological difficulties involved in applying the Krein–Milman theorem to $\mathfrak{P}^p_N$, but for all practical purposes

$\mathfrak{P}^p_N$ is the convex hull of its extreme points. Therefore, any ensemble N-representable p matrix D^p can be written as

$$D^p = \sum_m \alpha_m D^p_m, \tag{33}$$

where

$$0 \leq \alpha_m \leq 1, \tag{34a}$$

$$\sum_m \alpha_m = 1, \tag{34b}$$

and D^p_m is extreme and therefore pure N-representable.

With this result it is a simple matter to use Theorem 5 to solve the ensemble N-representability problem for $R = N + p$.

Theorem 6: An ensemble N-representable p matrix with $R = N + p$ is an extreme point of $\mathfrak{P}^p_N$ if and only if it has a pure state pre-image.

Proof: D^p is ensemble N-representable if and only if it can be written as

$$D^p = \sum_{m=1} \alpha_m D^p_m,$$

where α_m and D^p_m are as in (33). The corresponding hole matrix is

$$\eta^p = \sum_m \alpha_m \eta^p_m, \tag{35}$$

where η^p_m is the unique hole matrix corresponding to D^p_m. By assumption,

$$(\eta^p_m)^2 = \eta^p_m.$$

Thus, if D^p has $R = N + p$, Theorem 5 implies that D^p is pure N-representable:

$$\Leftrightarrow \quad (\eta^p)^2 = \eta^p$$

$$\Leftrightarrow \quad \eta^p_1 = \eta^p_2 = \cdots = \eta^p_M$$

$$\Leftrightarrow \quad D^p_1 = D^p_2 = \cdots = D^p_M$$

$$\Leftrightarrow \quad D^p \text{ is extreme.}$$

Theorem 7: A p matrix with $R = N + p$ is ensemble N-representable if and only if its hole matrix η^p is nonnegative.

Proof: Theorem 6 implies that a p matrix D^p with $R = N + p$ is ensemble N-representable if and only if it can be written as

$$D^p = \sum_m \alpha_m D^p_m,$$

where α_m and D^p_m are as in (33) and (34). D^p_m is pure N-representable, and one can assume, without loss of generality, that the pre-images of D^p_m are orthonormal. Then Theorem 5 implies that $\mathbf{H}^p_m = \mathbf{w}_m \mathbf{w}^\dagger_m$ and $\mathbf{w}^\dagger_m \mathbf{w}_n = \delta_{mn}$. Since the correspondence between $\mathbf{H}^p_m$ and D^p_m is unique, D^p is ensemble N-representable if and only if

$$\mathbf{H}^p = \sum_m \alpha_m \mathbf{H}^p_m \tag{36}$$

$$= \sum_m \alpha_m \mathbf{w}_m \mathbf{w}^\dagger_m. \tag{37}$$

Since the vectors $\mathbf{w}_m$ were assumed to be orthonormal, (37) is just the eigenvector expansion of $\mathbf{H}^p$. The nonzero eigenvalues α_m are > 0 by definition. Thus, η^p is nonnegative if and only if D^p is ensemble N-representable.

Davidson[17] has given a procedure for obtaining inequalities on the elements of the 2-matrix which are necessary for ensemble N-representability and which are independent of Garrod's condition[4] that the Q matrix (20) is nonnegative. However, Theorem 7 implies that, when $R \leq N + p$, these inequalities are reduced to a finite set which can be derived from the Q matrix condition.

Theorem 8: The pre-image ρ of any p matrix with $R = N + p$ is unique, and $(\mathbf{H}^p)^*$ is the matrix of expansion coefficients of ρ in a Slater determinant basis.

Proof: The proof of Theorem 7 implies that the pre-image of D^p is

$$\rho = \sum_m \alpha_m \Psi'_m \Psi'^*_m,$$

where Ψ'_m is given by (7) with expansion coefficients $c_{k_1 \cdots k_p}$ determined by $\mathbf{w}_m$ according to (32). Clearly, ρ is unique if α_m is nondegenerate. But any linear transformation on $\mathbf{w}_m$ induces the same linear transformation on $\{\Psi'_m\}$ so that ρ is always unique. Expanding Ψ'_m as in (7), one gets

$$\rho = \sum_{k_1 < \cdots k_p} \sum_{l_1 < \cdots l_p} h^*_{k_1 \cdots k_p, l_1 \cdots l_p} \hat{\Phi}_{k_1 \cdots k_p} \Phi^*_{l_1 \cdots l_p}, \tag{38}$$

where

$$\hat{\Phi}_{k_1 \cdots k_p} = (-)^{k_1 + k_2 \cdots k_p} \Phi_{k_1 \cdots k_p}$$

and $\Phi_{k_1 \cdots k_p}$ is the Slater determinant defined previously.[8]

Theorem 8 implies that, in order to test D^p for N-representability, one must generate its pre-image. The unfortunate connection between N-representability tests and finding the pre-image occurs in all known sets of sufficient conditions for N-representability. It has been discussed in general elsewhere.[2,18,19]

When $R = N + p$, $\mathbf{H}^p$ describes an ensemble p-hole state. Theorems 5–8 are simply the hole versions of the well-known results on p-representability of p matrices.

For purposes of comparison, these results are:

Theorem 5: A p matrix is pure p-representable if and only if it is idempotent.

Theorem 6: The extreme points of $\mathfrak{P}_p^p$ are the pure states.

Theorem 7: A p matrix is ensemble p-representable if and only if it is nonnegative.

Theorem 8: The p-particle pre-image of a p matrix is itself.

IV. N-REPRESENTABILITY: GENERAL HOLE CONDITIONS

The Basic Theorem

The results of the previous section are just a special case of a more general result which we discuss now.

Definition 2: A hole matrix η^p is said to be *pure M-representable* if there is an antisymmetric tensor $x_{k_1\cdots k_M}$ satisfying (23); it is *ensemble M-representable* if there are J such tensors satisfying

$$h^p_{k_1\cdots k_p, l_1\cdots l_p} = \sum_{j=1}^{J} \alpha_j \sum_{I} x^j_{KI} x^{j*}_{LI}, \tag{39}$$

where I, K, and L are as in (23) and $\{\alpha_j\}$ is as in (2). Now, if η^p is M-representable, each tensor $x^j_{k_1\cdots k_M}$ defines an antisymmetric $(R-M)$-particle function Ψ_j according to (7). Then consider the $(R-M)$-particle ensemble density matrix

$$\rho = \sum_j \alpha_j \Psi_j \Psi_j^*$$

and the unique p matrix D^p corresponding to η^p. Then it is clear from Theorem 2 that ρ must be the pre-image of D^p. Conversely, whenever D^p is ensemble N-representable, tensors $x^j_{k_1\cdots k_M}$ satisfying (39) will exist. Thus one can conclude that N-representability of D^p and η^p are equivalent problems.

Theorem 9: A p matrix D^p with $R = M + N$ is pure (ensemble) N-representable if and only if its hole matrix η^p is pure (ensemble) M-representable. In addition, for every condition on D^p, there is a corresponding condition on η^p and vice versa; corresponding conditions are obtained by making the exchanges

$$\text{(a)} \quad N \leftrightarrow M, \quad R = M + N, \tag{40}$$

$$\text{(b)} \quad d^p_{k_1\cdots k_p, l_1\cdots l_p} \leftrightarrow h^p_{k_1\cdots k_p, l_1\cdots l_p}, \tag{41}$$

$$\text{(c)} \quad b_{k_1\cdots k_N} \leftrightarrow x_{k_1\cdots k_M}. \tag{42}$$

Examples: Necessary Conditions

Using Theorem 9, one can extend the nonnegativity of reduced density matrices to hole matrices.

Theorem 10: A necessary condition for ensemble N-representability of a p matrix D^p is that the corresponding hole matrix η^p is nonnegative. When $p = 1$, this is equivalent to the well-known fact that the eigenvalues of D^1 must be bounded above by $1/N$. When $p = 2$, it is simply Garrod's Q matrix condition.[4]

One can try to generate new necessary conditions by applying Theorem 9 to all known conditions.[3] We mention, as examples, only two such results here. In the Slater determinant basis, the diagonal elements $h^p_{k_1\cdots k_p, k_1\cdots k_p}$ are bounded above by 1. The rank of η^p must be greater than $\binom{R-N}{p}$.

The application of Theorem 9 to known conditions does not always generate new conditions. For example, if one makes the substitutions $d^1_{ii} \leftrightarrow h^1_{ii}$ and $d^2_{ij,ij} \leftrightarrow h^2_{ij,ij}$ in one of the Davidson inequalities,[17] one merely generates another inequality in the same set.

If $\hat{G}$ is the hole analog of the G matrix,[6] one finds that

$$\hat{g}_{ab,cd} = g_{ba,dc}. \tag{43}$$

Therefore,

$$\hat{\mathbf{G}} = \mathbf{PGP}^{-1},$$

where $\mathbf{P}$ is a matrix which permutes the rows of $\mathbf{G}$. Thus, although $\hat{G} \neq G$, they are unitarily equivalent, and one cannot obtain any new conditions from $\hat{G}$.

Bounds on Eigenvalues of η^2

We now try to use Theorem 9 to obtain an upper bound on the eigenvalues of η^2. Let λ and ω be the maximum eigenvalues of D^2 and η^2, respectively. Recall that

$$0 \leq \lambda \leq (N-1)^{-1} \tag{44}$$

and that one can obtain D^2 with λ arbitrarily close to the upper bound by considering extreme antisymmetrized general power (EAGP) functions.[1,20-22] An EAGP function has an R-fold degenerate 1-matrix, and the largest eigenvalue of its 2-matrix is

$$\lambda = (N-1)^{-1}[(R-N+2)/R]. \tag{45}$$

Now Theorem 9 implies that the eigenvalues of any 2-hole matrix are bounded by

$$0 \leq \omega \leq (R-N-1)^{-1}. \tag{46}$$

The interesting question is whether or not (46) is the best possible bound on ω. It is not difficult to see that pre-image of η^p will be the hole analog of an EAGP function if and only if the pre-image of D^p is. It is then easy to show that the largest eigenvalue of an EAGP η^2 is

$$\omega = (R-N-1)^{-1}[(N+2)/R]. \tag{47}$$

Clearly (47) cannot be made arbitrarily close to the upper bound $(R-N-1)^{-1}$. This appears to contradict the duality principle which we have stated. However, the contradiction is easily resolved. To make λ large, one must take $R \gg M$. But, if N is fixed, one cannot make $R \gg M = R - N$. In order that ω reach its bound, one must fix M and increase N with R. Thus, for fixed N, (46) may not be the best possible bound. Indeed, one can obtain a better bound on ω as follows. First, define

$$\hat{\omega} = \binom{R-N}{2}\omega,$$

$$\hat{\lambda} = \binom{N}{2}\lambda.$$

Then one can show from (20) that

$$\hat{\omega} \leq \hat{\lambda} + 1. \tag{48}$$

This clearly implies that

$$\omega \leq (R-N-1)^{-1}[(N+2)/(R-N)]. \tag{49}$$

By taking $R \gg N$ for an EAGP function, one can make (47) arbitrarily close to (49). Thus, (49) gives the best possible bound on ω.

Theorem 11: The eigenvalues of η^p are bounded above by $(R-N-1)^{-1}[(N+2)/(R-N)]$.

It is interesting to note that the analog of (48),

$$\hat{\lambda} \leq \hat{\omega} + 1, \tag{50}$$

also holds. Combining (48) and (50), one gets

$$\hat{\lambda} - 1 \leq \hat{\omega} \leq \lambda + 1, \tag{51}$$

$$\hat{\omega} - 1 \leq \hat{\lambda} \leq \hat{\omega} + 1. \tag{52}$$

Sufficient Conditions

One can also use Theorem 9 to obtain new sufficient conditions for N-representability. As mentioned previously, the results of Sec. III can be easily derived in this way. Unfortunately, there are not many known sufficient conditions to which one can apply Theorem 9.

When $p = 2$, sufficient conditions for pure N-representability are known in the following cases:

(a) $N = 3$ (see Ref. 23);

(b) N odd, with a very restricted class of pre-images with triple excitations and pairing properties[2,15];

(c) N odd, with $(N+1)/2$ pairs of degenerate NSO's[2,15];

(d) Single excitation pre-images[24];

(e) the $\mathbf{\Omega}^1$ matrix defined from the N-completeness conditions has a sufficiently small degeneracy.[2,19]

Case (a) can be used to obtain sufficient conditions when $R = N + 3$. Cases (b) and (d) involve very specialized pre-images which do not have interesting hole analogs. Case (c) is itself the hole analog of a special class of functions with $N = 3$. The N-completeness conditions can certainly be extended to η^2, but this is not a particularly useful procedure.

* Much of this work was done at the Theoretical Chemistry Institute of the University of Wisconsin, where the author held an NSF predoctoral fellowship.

[1] A. J. Coleman, Rev. Mod. Phys. **35**, 668 (1963).
[2] M. B. Ruskai, Ph.D. Thesis; Technical Note No. 346, Theoretical Chemistry Institute, University of Wisconsin (July, 1969).
[3] For a recent review of N-representability, see Ref. 2, Chap. 1.
[4] C. Garrod and J. K. Percus, J. Math. Phys. **5**, 1756 (1964).
[5] This should not be confused with the G matrix, or particle–hole matrix, which Garrod and Rosina[6] have used to discuss particle–hole correlations.
[6] C. Garrod and M. Rosina, J. Math. Phys. **10**, 1855 (1969).
[7] Other normalizations of D^p are also in common use.
[8] In order to be sure of the correct signs in later equations, we choose the phase of the Slater determinant so that the coefficient of $f_{i_1}(1)f_{i_2}(2)\cdots f_{i_N}(N)$ is $+1$ when $j_1 < j_2 \cdots < j_N$.
[9] H. W. Gould, J. Math. Phys. **10**, 49 (1969).
[10] Garrod's expression[4] differs from (20) by the term $\delta_{il}\delta_{jk}$ because he uses a simple product basis while (16) uses a Slater determinant basis.
[11] M. L. Yoseloff and H. W. Kuhn, J. Math. Phys. **10**, 703 (1969).
[12] These results on the 1-matrix and its pre-image are clearly just the hole analogs of what occurs when $N = 2$.
[13] Reference 2, Theorem 2.4.
[14] Reference 15, Theorem 5.2. If N is odd, $\xi_1 = 1/N$. If N is even, define σ_r from z_{rr} instead of $\mathbf{T}^r$.
[15] M. B. Ruskai and J. E. Harriman, Phys. Rev. **169**, 101 (1968).
[16] H. Kummer, J. Math. Phys. **8**, 2063 (1967).
[17] E. R. Davidson, J. Math. Phys. **10**, 725 (1969).
[18] D. W. Smith, in the Report of a Seminar at Queen's University, Kingston, Ontario, Canada, A. J. Coleman and R. M. Erdahl, Eds., July, 1968 (unpublished).
[19] M. B. Ruskai, Phys. Rev. **183**, 129 (1969).
[20] A. J. Coleman, J. Math. Phys. **6**, 1425 (1965).
[21] F. Sasaki, Phys. Rev. **138**, B1338 (1965).
[22] C. N. Yang, Rev. Mod. Phys. **34**, 694 (1962).
[23] D. W. Smith, J. Chem. Phys. **43**, S258 (1965).
[24] R. McWeeny, in Ref. 18.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

$SU(4) \supset SU(2) \otimes SU(2)$ Projection Techniques*

J. P. DRAAYER†
Physics Department, University of Michigan, Ann Arbor, Michigan 48104

(Received 13 February 1970)

The twofold multiplicity problem associated with the Wigner supermultiplet reduction $SU(4) \supset SU(2) \otimes SU(2)$ is resolved by spin–isospin projection techniques analogous to the angular momentum projection technique introduced by Elliott to resolve the $SU(3) \supset R(3)$ multiplicity problem. The projection quantum numbers, which furnish either an integer or half-integer characterization of the multiplicity, are assigned according to an (ST)-multiplicity formula derived from a consideration of the symmetry properties of spin–isospin degeneracy diagrams. An expression is obtained for the coefficients which relate the $SU(4) \supset SU(2) \otimes SU(2)$ projected basis states to states labeled according to the natural $U(4) \supset U(3) \supset U(2) \supset U(1)$ chain. General expressions for $SU(4) \supset SU(2) \otimes SU(2)$ coupling coefficients and tensorial matrix elements are given in terms of the corresponding $U(4) \supset U(3) \supset U(2) \supset U(1)$ quantities.

1. INTRODUCTION

In 1937 Wigner[1] pioneered work that established $SU(4)$ as a group of major importance in nuclear structure studies. Its basis, the charge independence of nuclear forces, followed from an observed approximate fourfold degeneracy of nuclear energy levels. The result was the introduction of a nucleon distinguishing isospin quantum number which was combined with that of ordinary spin in the development of a spin–isospin supermultiplet theory. Group-theoretically, it corresponds to a state labeling scheme based upon the spin–isospin reduction $SU(4) \supset SU(2) \otimes SU(2)$.

In general, a complete specification of states in the supermultiplet scheme requires six labels in addition to those of the irreducible representation (IR) of $SU(4)$. The direct product $SU(2) \otimes SU(2)$ provides only four; two additional labels are needed. Techniques that can be used to resolve the multiplicity have been proposed by several authors.[2] In particular, Moshinsky and Nagel[2] have given a recipe for the construction of two operators whose eigenvalues may be used to complete the labeling. Labels obtained in this manner do not, however, exhibit any obvious symmetry properties, nor do they correspond in any way to know quantities of physical interest. In addition, the labels are not necessarily rational numbers.

A mathematically more convenient reduction is the natural or Gel'fand[3] chain $U(4) \supset U(3) \supset U(2) \supset U(1)$. In this case, the IR labels of $U(3)$, $U(2)$, and $U(1)$ provide the required six labels. Unfortunately, the reduction is unphysical. Nevertheless, since calculations are simpler within such a framework, the scheme has been used to calculate quantities of physical interest which depend only upon the IR labels of $SU(4)$. An example in point is that of the $SU(4)$ unitary recoupling coefficients (U functions) given by Hecht and Pang.[4]

The purpose of the present paper is to state and prove the existence of another solution to the $SU(4) \supset SU(2) \otimes SU(2)$ multiplicity problem, one in which the two additional labels are chosen so as to furnish an integer or half-integer characterization of the multiplicity that exhibits spin–isospin symmetry properties. The technique used is one of spin–isospin projection; it parallels closely Elliott's[5] resolution of the multiplicity problem in the $SU(3) \supset R(3)$ reduction. The simplifications associated with the $U(4) \supset U(3) \supset U(2) \supset U(1)$ reduction are incorporated into the scheme via coefficients which relate the projected $SU(4) \supset SU(2) \otimes SU(2)$ basis states to those labeled according to the $U(4) \supset U(3) \supset U(2) \supset U(1)$ chain.

To establish notation, Sec. 2 is devoted to a brief review of $SU(4)$ operator and state labeling techniques. In Sec. 3 a discussion of $SU(4)$ spin–isospin degeneracy diagrams is given, and a new rule for determining the number of occurrences of a spin–isospin pair (ST) in a given IR of $SU(4)$ is derived. In Sec. 4 the projection hypothesis is stated, and the completeness of the states so defined is proved. In Sec. 5 an expression is obtained for the coefficients which relate the projected basis states to those labeled according to the canonical $U(4) \supset U(3) \supset U(2) \supset U(1)$ reduction; general expressions for $SU(4) \supset SU(2) \otimes SU(2)$ coupling coefficients and tensorial matrix elements in terms of the corresponding $U(4) \supset U(3) \supset U(2) \supset U(1)$ quantities are also given.

2. BASIC NOTATION

A. Infinitesimal Generators

The 16 infinitesimal generators of $U(4)$ are given in terms of nucleon spin-charge creation and annihilation operators by

$$A_{\alpha\beta} = \sum_s a_\alpha^{s\dagger} a_\beta^s, \tag{2.1}$$

where s denotes the full set of space quantum numbers. The $A_{\alpha\beta}$ satisfy the $U(4)$ commutation relations

$$[A_{\alpha\beta}, A_{\rho\sigma}] = \delta_{\beta\rho}A_{\alpha\sigma} - \delta_{\sigma\alpha}A_{\rho\beta}. \tag{2.2}$$

Deletion of the operator $N = \sum_\alpha A_{\alpha\alpha}$ which commutes with the $A_{\alpha\beta}$ leads to a set of 15 infinitesimal generators for the group $SU(4)$. If $\alpha = 1, 2, 3$, and 4 represent the spin–isospin quantum numbers m_s and m_t in the sense

$$\begin{aligned} |1\rangle &= |+\tfrac{1}{2}, +\tfrac{1}{2}\rangle, \quad |2\rangle = |+\tfrac{1}{2}, -\tfrac{1}{2}\rangle, \\ |3\rangle &= |-\tfrac{1}{2}, +\tfrac{1}{2}\rangle, \quad |4\rangle = |-\tfrac{1}{2}, -\tfrac{1}{2}\rangle, \end{aligned} \tag{2.3}$$

then the $SU(4)$ generators can be expressed in terms of $SU(4) \supset SU(2) \otimes SU(2)$ tensors as[4]

$$\begin{aligned} S_0 &= \tfrac{1}{2}(A_{11} - A_{33} + A_{22} - A_{44}), \\ T_0 &= \tfrac{1}{2}(A_{11} - A_{22} + A_{33} - A_{44}), \\ E_{00} &= \tfrac{1}{2}(A_{11} - A_{22} - A_{33} + A_{44}), \\ S_+ &= A_{13} + A_{24}, \qquad S_- = A_{31} + A_{42}, \\ T_+ &= A_{12} + A_{34}, \qquad T_- = A_{21} + A_{43}, \\ E_{10} &= A_{13} - A_{24}, \qquad E_{-1\,0} = A_{31} - A_{42}, \\ E_{01} &= A_{12} - A_{34}, \qquad E_{0\,-1} = A_{21} - A_{43}, \\ E_{11} &= A_{14}, \qquad E_{-1\,-1} = A_{41}, \\ E_{1\,-1} &= A_{23}, \qquad E_{-1\,1} = A_{32}. \end{aligned} \tag{2.4}$$

The commutation properties of $\bar{S}$, $\bar{T}$, and $\bar{E}$ follow from the commutation properties of the $A_{\alpha\beta}$ given by Eq. (2.2).

B. Irreducible Representations

Gel'fand patterns of the type

$$|G\rangle = \left| \begin{array}{cccccc} h_{14} & & h_{24} & & h_{34} & & h_{44} \\ & h_{13} & & h_{23} & & h_{33} & \\ & & h_{12} & & h_{22} & & \\ & & & h_{11} & & & \end{array} \right\rangle \tag{2.5}$$

furnish a complete set of labels for the basis states of an IR of $U(4)$. The $h_{\alpha\beta}$, $1 \le \alpha \le \beta \le 4$, specify the IR's of $U(\beta)$ in the canonical chain $U(4) \supset U(3) \supset U(2) \supset U(1)$ to which the state belongs. The $h_{\alpha\beta}$ are integral and satisfy the Young tableau or betweenness conditions

$$h_{\alpha\beta} \ge h_{\alpha,\beta-1} \ge h_{\alpha+1,\beta} \ge 0. \tag{2.6}$$

Replacing each $h_{\alpha\beta}$ by $h_{\alpha\beta} - h_{44}$ leads to the corresponding basis state for $SU(4)$; it differs from the $U(4)$ state by at most an h_{44}-dependent phase factor.

Other characterizations for the IR's of $SU(4)$ include the set of three numbers $(\lambda_1\lambda_2\lambda_3)$ given by $\lambda_1 = h_{14} - h_{24}$, $\lambda_2 = h_{24} - h_{34}$, and $\lambda_3 = h_{34} - h_{44}$. $SU(4)$ conjugation properties can then be expressed as relating the $(\lambda_1\lambda_2\lambda_3)$ and $(\lambda_3\lambda_2\lambda_1)$ IR's.[6] Wigner[1] introduced the triplet of numbers $(PP'P'')$ given by $P = \frac{1}{2}(\lambda_1 + 2\lambda_2 + \lambda_3)$, $P' = \frac{1}{2}(\lambda_1 + \lambda_3)$, and $P'' = \frac{1}{2}(\lambda_1 - \lambda_3)$. They are associated with the maximum eigenvalues for the operators E_{00}, S_0, and T_0 (e.g., P = maximum eigenvalue of E_{00} contained in the IR, P' = maximum eigenvalue of S_0 for states with $E_{00} = P$, and P'' = maximum eigenvalue of T_0 for states with $E_{00} = P$ and $S_0 = P'$).[7] In what follows, simplicity of formulation will determine which labels are used. In all cases the relationships as given above apply.

The states $|G\rangle$ are eigenstates of the operators $A_{\alpha\alpha}$ with eigenvalues w_α,

$$\begin{aligned} A_{\alpha\alpha}|G\rangle &= w_\alpha |G\rangle, \\ w_\alpha &= \sum \text{row } \alpha - \sum \text{row } (\alpha - 1) \\ &= \sum_\beta h_{\beta\alpha} - \sum_\beta h_{\beta,\alpha-1}. \end{aligned} \tag{2.7}$$

States of particular interest in the present development are those for which the operator $E_{00} = \frac{1}{2}(A_{11} - A_{22} - A_{33} + A_{44})$ assumes either its (a) maximum ($E_{00}^{\max} = P$) or (b) minimum ($E_{00}^{\min} = -P$) eigenvalue. The $h_{\alpha\beta}$ for such states are uniquely specified by K_S and K_T, the eigenvalues of $\frac{1}{2}(A_{11} + A_{22} - A_{33} - A_{44}) = S_0$ and $\frac{1}{2}(A_{11} - A_{22} + A_{33} - A_{44}) = T_0$, respectively. Explicitly,

$$|G_{E\uparrow}\{K_S K_T\}\rangle = \left| \begin{array}{ccccccc} h_1 & & h_2 & & h_3 & & h_4 \\ & h_1-p & & h_3 & & h_4 & \\ & & h_1-p & & h_3-q & & \\ & & & h_1-p & & & \end{array} \right\rangle,$$
$$0 \le p \le \lambda_1, \; 0 \le q \le \lambda_3, \tag{2.8a}$$

$$|G_{E\downarrow}\{K_S K_T\}\rangle = \left| \begin{array}{ccccccc} h_1 & & h_2 & & h_3 & & h_4 \\ & h_1 & & h_2 & & h_3-q & \\ & & h_1-p & & h_3-q & & \\ & & & h_3-q & & & \end{array} \right\rangle,$$
$$0 \le p \le \lambda_1, \; 0 \le q \le \lambda_3, \tag{2.8b}$$

where

$$\begin{aligned} K_S + K_T &= h_1 - h_2 - 2p = \lambda_1 - 2p, \\ K_S - K_T &= h_2 - h_3 - 2q = \lambda_3 - 2q, \end{aligned} \tag{2.9a}$$

$$\begin{aligned} K_S + K_T &= h_2 - h_3 - 2q = \lambda_3 - 2q, \\ K_S - K_T &= h_1 - h_2 - 2p = \lambda_1 - 2p \end{aligned} \tag{2.9b}$$

for $|G_{E\uparrow}\rangle$ and $|G_{E\downarrow}\rangle$, respectively.[8] The solid curves in Fig. 2 of Sec. 4 illustrate the result schematically.

Note that for (a) $(\lambda_1\lambda_3)$-(odd, even) K_S and K_T are half-integral with K_S differing from K_T by twice an integer, for (b) $(\lambda_1\lambda_3)$-(odd, odd) K_S and K_T are integral with K_S differing from K_T by twice an integer plus one, for (c) $(\lambda_1\lambda_3)$-(even, odd) K_S and K_T are half-integral with K_S differing from K_T by twice an integer plus one, and for (d) $(\lambda_1\lambda_3)$-(even, even) K_S and K_T are integral with K_S differing from K_T by twice an integer. That is, the odd–even characteristics of λ_1 and λ_3 furnish a complete characterization of distinct symmetry types for the $\{K_S K_T\}$-values associated with the $|G_E\rangle$.

3. SPIN-ISOSPIN MULTIPLICITIES

Racah[9] has given a relatively simple algebraic formula for determining the multiplicity $N_{(ST)}(\lambda_1\lambda_2\lambda_3)$ of (ST)-values in an IR $(\lambda_1\lambda_2\lambda_3)$ of $SU(4)$. Some simplifications in his result follow from the investigations of Kretzschmar[10] and Perelomov and Popov.[11] In each case the expressions given are based upon the Littlewood rules[12] which allow $N_{(ST)}(\lambda_1\lambda_2\lambda_3)$ to be related to a sum over terms of the type $N_{(ST)}(\lambda_1'\lambda_2'\lambda_3')$, where the IR's $(\lambda_1'\lambda_2'\lambda_3')$ have particularly simple multiplicity structures. In this section an expression for $N_{(ST)}(\lambda_1\lambda_2\lambda_3)$ is given which involves a sum over terms of the type $N_{(S'T')}(\lambda_1 0\lambda_3)$ where the $(S'T')$-values are related to the (ST)-values in a very simple way. Since Racah's expression for $N_{ST}(\lambda_1 0\lambda_3)$ is quite transparent, the result is particularly convenient for a study of the origin of (ST)-multiplicities and leads quite naturally to a rule for the projection numbers of Sec. 4.

A. Degeneracy Diagrams

A spin–isospin degeneracy diagram for the IR $(\lambda_1\lambda_2\lambda_3)$ of $SU(4)$ is a regular lattice of points (ST) each of which is labeled by the numerical value of $N_{(ST)}(\lambda_1\lambda_2\lambda_3)$, the multiplicity of the pair (ST) in the IR $(\lambda_1\lambda_2\lambda_3)$. Figure 4 of Sec. 4 gives examples. The spin–isospin symmetry property $N_{(ST)}(\lambda_1\lambda_2\lambda_3) = N_{(TS)}(\lambda_1\lambda_2\lambda_3)$ corresponds to reflection symmetry in the $S = T$ plane. The conjugation properties of $SU(4)$ imply that $N_{(ST)}(\lambda_1\lambda_2\lambda_3) = N_{(ST)}(\lambda_3\lambda_2\lambda_1)$. A systematic study of $SU(4)$ spin–isospin degeneracy diagrams can therefore be limited to a consideration of those IR's of $SU(4)$ for which $\lambda_1 \geq \lambda_3$ and within such IR's those (ST)-values for which $S \leq T$.

Figure 1 illustrates features common to all $SU(4)$ spin–isospin degeneracy diagrams. The heavy solid curve $EP(\lambda_1\lambda_2\lambda_3)$ is, in the terminology of Perelomov and Popov,[11] the enveloping polygon for the spin–isospin degeneracy diagram associated with the $(\lambda_1\lambda_2\lambda_3)$ IR of $SU(4)$. It circumscribes all (ST)-values

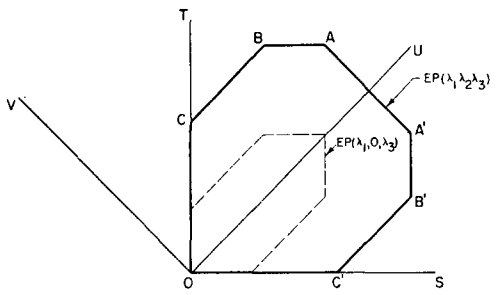


FIG. 1. General features of an $SU(4)$ spin–isospin degeneracy diagram. The heavy solid curve $EP(\lambda_1\lambda_2\lambda_3)$ is the enveloping polygon for the spin–isospin degeneracy diagram associated with the $(\lambda_1\lambda_2\lambda_3)$ IR of $SU(4)$. The (ST)- and $[UV]$-coordinates of the boundary points are given by

$$A:(P',P),[Q,+\lambda_2];\ B:(P'',P),[Q',+Q''];\ C:(0,Q''),[Q'',+Q''],$$
$$A':(P,P'),[Q,-\lambda_2];\ B':(P,P''),[Q',-Q''];\ C':(Q'',0),[Q'',-Q''],$$

where $(PP'P'')$ are the Wigner supermultiplet quantum numbers

$$P = \tfrac{1}{2}(\lambda_1 + 2\lambda_2 + \lambda_3),\quad P' = \tfrac{1}{2}(\lambda_1 + \lambda_3),\quad P'' = \tfrac{1}{2}(\lambda_1 - \lambda_3)$$

and the $(QQ'Q'')$ triplet of numbers is given by

$$Q = \lambda_1 + \lambda_2 + \lambda_3,\quad Q' = \lambda_1 + \lambda_2,\quad Q'' = \lambda_2 + \lambda_3.$$

The dashed curve $EP(\lambda_1 0\lambda_3)$ is the corresponding result for $\lambda_2 = 0$.

for which $N_{(ST)}(\lambda_1\lambda_2\lambda_3)$ is nonzero. The boundary points for the polygon are as given in the figure. The axes $U = T + S$ and $V = T - S$ have been included as a simplifying feature for the discussion that is to follow. The dashed curve $EP(\lambda_1 0\lambda_3)$ is the corresponding result for $\lambda_2 = 0$. As shown, the figure corresponds to $\lambda_1 + \lambda_3$ even and hence integral (ST)-values. For $\lambda_1 + \lambda_3$ odd and hence half-integral (ST)-values, the schematics are identical, the only difference being that the lines OC and OC' are shifted one-half unit from the coordinate axes.

As can be seen from Fig. 1, $EP(\lambda_1\lambda_2\lambda_3)$ and $EP(\lambda_1 0\lambda_3)$ are simply related; for $S \leq T$, $EP(\lambda_1\lambda_2\lambda_3)$ corresponds to $EP(\lambda_1 0\lambda_3)$ shifted λ_2 units along the T axis, and, for $S > T$, $EP(\lambda_1\lambda_2\lambda_3)$ corresponds to $EP(\lambda_1 0\lambda_3)$ shifted λ_2 units along the S axis. More precisely, $EP(\lambda_1\lambda_2\lambda_3)$ is the envelope of all isosceles right triangles built by λ_2 regular lattice displacements[13] upon the (ST)-values of $EP(\lambda_1 0\lambda_3)$. Therefore, $EP(\lambda_1 0\lambda_3)$ is a characteristic structure common to all IR $(\lambda_1\lambda_2\lambda_3)$ (λ_1 and λ_3 fixed; λ_2 arbitrary) of $SU(4)$. Furthermore, note that for $\lambda_2 = 0$ the boundary points B and B' coincide with the boundary points $\{P'', P'\}$ and $\{P', P''\}$ of Fig. 2a (Sec. 4). Therefore, like rule (2.9) for the $\{K_S K_T\}$-values associated with $|G_E\rangle$, a classification scheme based on the odd–even characteristics of the fundamental lengths $U_A - U_C = \lambda_1$ and $V_B - V_A = \lambda_3$ furnishes a complete characterization of distinct $EP(\lambda_1 0\lambda_3)$ and hence $EP(\lambda_1\lambda_2\lambda_3)$ symmetry types.

The results for $EP(\lambda_1\lambda_2\lambda_3)$ suggest that $N_{(ST)}(\lambda_1\lambda_2\lambda_3)$ may be simply related to $N_{(S'T')}(\lambda_1 0\lambda_3)$ and, furthermore, that the classification scheme (a) $(\lambda_1\lambda_3)$-(odd, even), (b) $(\lambda_1\lambda_3)$-(odd, odd), (c) $(\lambda_1\lambda_3)$-(even, odd), and (d) $(\lambda_1\lambda_3)$-(even, even) may furnish a complete characterization of distinct $N_{(ST)}(\lambda_1 0\lambda_3)$ and hence $N_{(ST)}(\lambda_1\lambda_2\lambda_3)$ symmetry types. To test the hypothesis, a quantitative study of the numerology of related degeneracy diagrams was made (e.g., see Fig. 4 in Sec. 4). In terms of $N_{[UV]}(\lambda_1\lambda_2\lambda_3) \equiv N_{(ST)}(\lambda_1\lambda_2\lambda_3)$, $U = T + S$, and $V = T - S$, the result of the investigation, with $V \geq 0$, is that

$$\begin{aligned} N_{[UV]}(\lambda_1\lambda_2\lambda_3) &= N_{[UV]}(\lambda_1, \lambda_2 - 1, \lambda_3) \\ &\quad + N_{[U'V']}(\lambda_1 0\lambda_3) + \delta_{[UV]}(\lambda_1\lambda_2\lambda_3), \\ U' &= U - \lambda_2, \\ V' &= \text{map}\,[V - \lambda_2, |\text{mod}\,(V - \lambda_2, 2)|], \end{aligned} \tag{3.1}$$

where $\delta_{[UV]}(\lambda_1\lambda_2\lambda_3) = 0$ for cases (a), (b), and (c) and, for case (d),

$$\begin{aligned} \delta_{[UV]}(\lambda_1\lambda_2\lambda_3) &= \quad 1, \quad \lambda_2 > U \geq V,\ U - \lambda_2 \text{ even}, \\ &= -1, \quad \lambda_2 > U \geq V,\ U - \lambda_2 \text{ odd}, \\ &= \quad 0, \quad \text{otherwise}. \end{aligned} \tag{3.2}$$

The formula is recursive and therefore may be iterated to yield

$$\begin{aligned} N_{[UV]}(\lambda_1\lambda_2\lambda_3) &= \sum_m N_{[U'V']}(\lambda_1 0\lambda_3), \\ U' &= U - m, \\ V' &= \max\,[V - m, |\text{mod}\,(V - m, 2)|], \\ 0 \leq m \leq \lambda_2, &\quad m \neq U \quad \text{if} \quad U - \lambda_2 \text{ odd}, \end{aligned} \tag{3.3}$$

which is applicable to all four cases (a)–(d). In terms of $N_{(ST)}(\lambda_1 0\lambda_3)$, Eq. (3.3) has the form

$S > T$:

$$\begin{aligned} N_{(ST)}(\lambda_1\lambda_2\lambda_3) &= N_{(ST)}(\lambda_1 0\lambda_3) \\ &\quad + N_{(S-1,T)}(\lambda_1 0\lambda_3) \\ &\quad + \cdots \\ &\quad + N_{(TT)}(\lambda_1 0\lambda_3) \\ &\quad + N_{(T,T-1)}(\lambda_1 0\lambda_3) \\ &\quad + N_{(T-1,T-1)}(\lambda_1 0\lambda_3) \\ &\quad + \cdots \\ &\quad + N_{(S'T')}(\lambda_1 0\lambda_3), \\ S' + T' &= S + T - \lambda_2, \end{aligned} \tag{3.4a}$$

$S \leq T$:

$$\begin{aligned} N_{(ST)}(\lambda_1\lambda_2\lambda_3) &= N_{(ST)}(\lambda_1 0\lambda_3) \\ &\quad + N_{(S,T-1)}(\lambda_1 0\lambda_3) \\ &\quad + \cdots \\ &\quad + N_{(SS)}(\lambda_1 0\lambda_3) \\ &\quad + N_{(S,S-1)}(\lambda_1 0\lambda_3) \\ &\quad + N_{(S-1,S-1)}(\lambda_1 0\lambda_3) \\ &\quad + \cdots \\ &\quad + N_{(S'T')}(\lambda_1 0\lambda_3), \\ S' + T' &= S + T - \lambda_2, \end{aligned} \tag{3.4b}$$

where $N_{(00)}(\lambda_1 0\lambda_3)$ is *not* included if $S + T - \lambda_2$ is odd. The next section is devoted to an analytic proof of this result.

B. Proof of the Multiplicity Formula

Racah[9] has shown that

$$\begin{aligned} N_{[UV]}(\lambda_1\lambda_2\lambda_3) &= \omega_{[UV]}(\lambda_1 + \lambda_2, \lambda_2 + \lambda_3) \\ &\quad - \omega_{[UV]}(\lambda_1 + \lambda_2 + \lambda_3 + 1, \lambda_2 - 1) \\ &\quad - \omega_{[UV]}(\lambda_1 - 1, \lambda_3 - 1), \end{aligned} \tag{3.5}$$

where $\omega_{[UV]}(xy)$ vanishes unless

$$\begin{aligned} x + y &\geq \max\,(U + V, U - V), \\ x + y &\equiv U + V \equiv U - V\ (\text{mod } 2), \end{aligned}$$

and that, if these conditions are satisfied and $x \geq y$,

$$\begin{aligned} \omega_{[UV]}(xy) &= \omega_{[UV]}(yx) \\ &= \varphi(y + 2 - |V|) \\ &\quad - \varphi(y + 1 - U) \\ &\quad + \varphi(U - x + 1) \\ &\quad - \tfrac{1}{2}\varphi(U - |V| - x + y + 1). \end{aligned} \tag{3.6}$$

The function $\varphi(z)$ is given by

$$\begin{aligned} \varphi(z) &= [z^2/4], \quad z \geq 0, \\ &= 0, \qquad z < 0, \end{aligned} \tag{3.7}$$

where the boldface brackets indicate the greatest integer contained in the argument.

Define

$$\begin{aligned} \Delta N_{[UV]}(\lambda_1\lambda_2\lambda_3) &= N_{[UV]}(\lambda_1\lambda_2\lambda_3) \\ &\quad - N_{[UV]}(\lambda_1, \lambda_2 - 1, \lambda_3), \end{aligned} \tag{3.8a}$$

$$\begin{aligned} \Delta\omega_{[UV]}(xy) &= \omega_{[UV]}(xy) \\ &\quad - \omega_{[UV]}(x - 1, y - 1), \end{aligned} \tag{3.8b}$$

$$\Delta\varphi(z) = \varphi(z) - \varphi(z - 1). \tag{3.8c}$$

Then, to prove Eq. (3.1), it is sufficient to demonstrate the equivalence of

$$\begin{aligned}\Delta N_{[UV]}(\lambda_1\lambda_2\lambda_3) &= \Delta\omega_{[UV]}(\lambda_1+\lambda_2, \lambda_2+\lambda_3) \\ &\quad - \Delta\omega_{[UV]}(\lambda_1+\lambda_2+\lambda_3+1, \lambda_2-1) \\ &= \Delta\varphi(\lambda_2+\lambda_3+2-V) - \Delta\varphi(\lambda_2+\lambda_3+1-U) \\ &\quad + \Delta\varphi(U-\lambda_1-\lambda_2-1) - \Delta\varphi(\lambda_2+1-V) \\ &\quad + \Delta\varphi(\lambda_2-U) - \Delta\varphi(U-\lambda_1-\lambda_2-\lambda_3) \quad (3.9)\end{aligned}$$

and

$$\begin{aligned}N_{[U'V']}(\lambda_1 0\lambda_3) &= \Delta\omega_{[U'V']}(\lambda_1\lambda_3) \\ &= \Delta\varphi(\lambda_3+2-V') - \Delta\varphi(\lambda_3+1-U') \\ &\quad + \Delta\varphi(U'-\lambda_1-1). \quad (3.10)\end{aligned}$$

For $(\lambda_1\lambda_3)$-(even, even) and $\lambda_2 > U \geq V$, the factor $\delta_{[UV]}(\lambda_1\lambda_2\lambda_3)$ must, of course, be added to

$$N_{[U'V']}(\lambda_1 0\lambda_3).$$

Consider the following special cases:

Case 1: $U \geq V \geq \lambda_2$.

Case 2: $U \geq \lambda_2 \geq V$:

(a) $V - \lambda_2 = -2n$,

(b) $V - \lambda_2 = -2n - 1$.

Case 3: $\lambda_2 > U \geq V$:

(a) $U = V + 2n + 1$:

(1) $(\lambda_1\lambda_3)$-(odd, even),

(2) $(\lambda_1\lambda_3)$-(even, odd);

(b) $U = V + 2n$:

(1) $(\lambda_1\lambda_3)$-(odd, odd),

(2) $(\lambda_1\lambda_3)$-(even, even).

For case 1 the result is trivial since $U' = U - \lambda_2$, $V' = V - \lambda_2$ makes $\Delta N_{[UV]}(\lambda_1\lambda_2\lambda_3)$ and $N_{[U'V']}(\lambda_1 0\lambda_3)$ identical functions in φ. In both (a) and (b) of case 2 an application of the result $\Delta\varphi(m + 2n) = \Delta\varphi(m) + n$, m, n integer, leads to the desired conclusion. Case 3 is somewhat more complicated because $U' = U - \lambda_2 < 0$ implies that $N_{[U'V']}(\lambda_1 0\lambda_3) = 0$. In this case it is therefore necessary to demonstrate the equivalence of $\Delta N_{[UV]}(\lambda_1\lambda_2\lambda_3)$ and $\delta_{[UV]}(\lambda_1\lambda_2\lambda_3)$. The substitution $\lambda_2 - U = 2m + \delta$ and $\lambda_2 + 1 - V = 2n + \nu$, m, n integer and μ, ν being 0 or 1, simplifies $\Delta N_{[UV]}(\lambda_1\lambda_2\lambda_3)$ to

$$\begin{aligned}\Delta N_{[UV]}(\lambda_1\lambda_2\lambda_3) &= \Delta\varphi(\lambda_3+1+\nu) \\ &\quad - \Delta\varphi(\lambda_3+1+\mu). \quad (3.11)\end{aligned}$$

For 3(a) $\mu = \nu$ so that $\Delta N_{[UV]}(\lambda_1\lambda_2\lambda_3) = 0$. For 3(b) $\mu \neq \nu$, but the substitution $\lambda_3 + 1 = 2k + \kappa$, k integer and κ being 0 or 1, leads to

$$\begin{aligned}\Delta N_{[UV]}(\lambda_1\lambda_2\lambda_3) &= \Delta\varphi(\nu) - \Delta\varphi(\mu) \\ &= 0 \quad (3.12)\end{aligned}$$

for (b1) and

$$\begin{aligned}\Delta N_{[UV]}(\lambda_1\lambda_2\lambda_3) &= \Delta\varphi(\nu+1) - \Delta\varphi(\mu+1) \\ &= \begin{cases} 1, & U-\lambda_2 \text{ even}, \quad \mu=0, \nu=1 \\ -1, & U-\lambda_2 \text{ odd}, \quad \mu=1, \nu=0 \end{cases}, \quad (3.13)\end{aligned}$$

for (b2), which is the desired result.

4. SPIN-ISOSPIN PROJECTION

The additional quantum numbers that are required to resolve the twofold multiplicity associated with the reduction $SU(4) \supset SU(2) \otimes SU(2)$ may be chosen in a variety of ways. The solution proposed by Moshinsky and Nagel[2] is not necessarily the most convenient because of the algebraic diffculties inherent with the corresponding eigenvalue problem. In this section the existence of another solution to the multiplicity problem is stated and proved. It is based upon spin–isospin projection techniques in which the $\{K_S K_T\}$-pairs associated with the states $|G_E\rangle$ furnish the required labels.

A. Projection Hypothesis

A projection operator for a state of total angular momentum J with projection M may be expressed in Hill–Wheeler integral form[14] as

$$P^J_{MK} = (2J+1)\int d\Omega D^{J*}_{MK}(\Omega) R_J(\Omega), \quad (4.1)$$

where $D^J_{MK}(\Omega)$ is an $R(3)$ rotation matrix and $R_J(\Omega)$ is an $R(3)$ rotation operator,

$$\begin{aligned}R_J(\Omega) &= e^{-i\alpha J_3} e^{-i\beta J_2} e^{-i\gamma J_3}, \\ [J_1, J_2] &= iJ_3 \quad \text{(cyclic)}. \quad (4.2)\end{aligned}$$

The integration is over Euler angles $(\alpha\beta\gamma)$. From this definition it follows that

$$P^{J'}_{M'K'} P^J_{MK} = \delta_{J'J}\delta_{K'M} P^J_{M'K}, \quad (4.3)$$

$$P^{J\dagger}_{MK} = P^J_{KM}, \quad (4.4)$$

where $P^{J\dagger}_{MK}$ indicates the Hermitian conjugate of P^J_{MK}. Cases of interest in the present analysis are those for which J is either the spin S or the isospin T of Eq. (2.4).

Since eigenstates of the total spin and isospin operators may be obtained from a state $|G\rangle$ by simply applying the projection operators $P^S_{M_SK_S}$ and $P^T_{M_TK_T}$, we define

$$|GK_SSM_SK_TTM_T\rangle \equiv P^S_{M_SK_S}P^T_{M_TK_T}|G\rangle. \quad (4.5)$$

The complete G symbol has been retained in the projected ket as a reminder of the Gel'fand state from which it was derived; only the IR labels $h_{\alpha 4}$, however, remain valid state labels. In many cases the $|GK_SSM_SK_TTM_T\rangle$ will turn out to be identically zero. It remains to specify the $|G\rangle$ and pairs $\{K_SK_T\}$ with their corresponding (ST)-values for which projected states span the IR space.

The Projection Hypothesis

The projected states

$$|G_EK_SSM_SK_TTM_T\rangle \equiv P^S_{M_SK_S}P^T_{M_TK_T}|G_E\rangle, \quad (4.6)$$

with $|G_E\rangle$ the Gel'fand states for which the operator E_{00} assumes its maximum ($\lambda_1 \geq \lambda_3$) or minimum ($\lambda_1 < \lambda_3$) eigenvalue, span the $(\lambda_1\lambda_2\lambda_3)$ IR space of $SU(4)$ if with each integer ($\lambda_1 + \lambda_3$ even) or half-integer ($\lambda_1 + \lambda_3$ odd) pair $\{K_SK_T\}$ satisfying

$$\begin{aligned} &K_S + K_T = \max(\lambda_1\lambda_3) - 2p,\\ &K_S - K_T = \min(\lambda_1\lambda_3) - 2q,\\ &\quad 0 \leq p \leq [\max(\lambda_1\lambda_3)/2],\\ &\quad \kappa \leq q \leq \min(\lambda_1\lambda_3),\\ &\qquad \kappa = 0, \quad K_S + K_T \neq 0,\\ &\qquad \kappa = [\min(\lambda_1\lambda_3)/2], \quad K_S + K_T = 0, \end{aligned} \quad (4.7)$$

is associated the (ST)-values

$$\sigma > \tau: \quad (ST) = (\sigma + \mu, \tau + \nu), \quad 0 \leq \mu \leq \lambda_2, \quad 0 \leq \nu < \sigma - \tau + \lambda_2 - \mu, \quad (4.8a)$$

$$\sigma \leq \tau \neq 0: \quad (ST) = (\sigma + \mu, \tau + \nu), \quad 0 \leq \nu \leq \lambda_2, \quad 0 \leq \mu \leq \tau - \sigma + \lambda_2 - \nu, \quad (4.8b)$$

$$\sigma = \tau = 0: \quad (ST) = (\lambda_2 - 2\mu - \nu, \nu), \quad 0 \leq \mu \leq [\lambda_2/2], \quad 0 \leq \nu \leq \lambda_2 - 2, \quad (4.8c)$$

where $\sigma = |K_S|$ and $\tau = |K_T|$. The projections M_S and M_T assume the usual values $-S \leq M_S \leq S$ and $-T \leq M_T \leq T$.

The proof of the hypothesis will be made in two steps. First, the value of $N_{(ST)}(\lambda_1\lambda_2\lambda_3)$ predicted by the rule will be shown to be precisely that derived in Sec. 3. And, secondly, the assumption that there exists a function belonging to the IR space but orthogonal to the projected states will be shown to lead to a contradiction. Before proceeding, however, we first consider in more detail the structure of the rule as given by Eqs. (4.7) and (4.8).

Since the Gel'fand states $|G_E\rangle$ are eigenstates of S_0 and T_0, the $\{K_SK_T\}$-pairs of Eq. (4.7) are necessarily a subset of the allowed $\{K_SK_T\}$-pairs given by Eq. (2.9). The choice made (see Fig. 2) is not, however, unique; other possibilities exist. For example, simply replacing each $\{K_SK_T\}$-pair of Eq. (4.7) by $\{-K_S, -K_T\}$ (inversion in the $\{K_SK_T\}$-plane) provides an equally acceptable set of projection numbers. It is also true that any partial inversion in the $\{K_SK_T\}$-plane provides an acceptable set of projection numbers. The essential feature of any such choice is that only one of the pairs, $\{K_SK_T\}$ or its inversion $\{-K_S, -K_T\}$, be included. Inclusion of both pairs leads to states which are not linearly independent. The choice made by Eq. (4.7) is therefore one of convention; its simplifying feature is that it maximizes the number of $\{K_SK_T\}$-pairs contained within $EP(\lambda_1 0 \lambda_3)$.

In some applications it is convenient to know the rule corresponding to Eq. (4.7) for projection from $|G_{E\uparrow}\rangle$ if $\lambda_1 < \lambda_3$ and from $|G_{E\downarrow}\rangle$ if $\lambda_1 \geq \lambda_3$. It can be obtained from Eq. (4.7) by simply interchanging the max–min specifications. It follows that the rules for determining the $\{K_SK_T\}$-pairs for projection from $|G_{E\uparrow}\rangle$ and $|G_{E\downarrow}\rangle$ without regard to the relationship of λ_1 and λ_3 are given by the following:

projection from $|G_{E\uparrow}\rangle$:

$$\begin{aligned} &K_S + K_T = \lambda_1 - 2p,\\ &K_S - K_T = \lambda_3 - 2q,\\ &\quad 0 \leq p \leq [\lambda_1/2],\\ &\quad \kappa \leq q \leq \lambda_3,\\ &\qquad \kappa = 0, K_S + K_T \neq 0,\\ &\qquad \kappa = [\lambda_3/2], K_S + K_T = 0; \end{aligned} \quad (4.9a)$$

projection from $|G_{E\downarrow}\rangle$:

$$\begin{aligned} &K_S + K_T = \lambda_3 - 2q,\\ &K_S - K_T = \lambda_1 - 2p,\\ &\quad 0 \leq q \leq [\lambda_3/2],\\ &\quad \kappa \leq p \leq \lambda_1,\\ &\qquad \kappa = 0, K_S + K_T \neq 0,\\ &\qquad \kappa = [\lambda_1/2], K_S + K_T = 0. \end{aligned} \quad (4.9b)$$

Figure 2 illustrates the result schematically. The dashed curves ($K_S + K_T = 0$ not allowed) and the broken curves ($K_S + K_T = 0$ allowed) divide the $\{K_SK_T\}$-pairs of Eq. (2.9) into two sets equivalent under

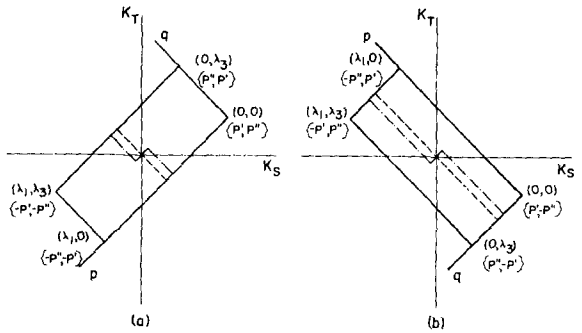


FIG. 2. The envelope of $\{K_S K_T\}$-pairs associated with $|G_E\rangle$. (a) $|G_E\rangle = |G_{E\uparrow}\rangle$, (b) $|G_E\rangle = |G_{E\downarrow}\rangle$. The boundaries are denoted by their (pq)- and $\{K_S K_T\}$-values. The dashed curves ($K_S + K_T = 0$ not allowed) and the broken curves ($K_S + K_T = 0$ allowed) divide the $\{K_S K_T\}$-pairs into two sets equivalent under inversion; the pairs for which $K_S + K_T \geq 0$ are by convention the projection numbers of Eq. (4.9).

inversion; the pairs for which $K_S + K_T \geq 0$ are by convention the projection numbers of Eq. (4.9). In any case the spectrum of (ST)-values given by Eq. (4.8) depends only upon σ and τ and is therefore independent of the $\{K_S K_T\}$-rule chosen as long as all $\{K_S K_T\}$-pairs belonging to the Gel'fand state $|G_E\rangle$ under consideration, but not equivalent under inversion, are included in the rule specification.

Figure 3 illustrates Eq. (4.8) by giving the spectrum of (ST)-values associated with a given $\{K_S K_T\}$-pair for the cases $\sigma < \tau$, $\sigma = \tau = \sigma'$, and $\sigma = \tau = 0$. The schematics of the figure are such that the (ST)-values labeled by the same symbol are those derived from the same $\{K_S K_T\}$-pair. In the examples shown, $\lambda_2 = 4$. For $\sigma < \tau$, both $\{K_S K_T\} = \{\sigma\tau\}$ and $\{K_S K_T\} = \{\tau\sigma\}$ have been given. In the case $\sigma < \tau$, note that except for $(ST) = (\tau + \lambda_2 - \nu, \tau + \nu)$, $0 \leq \nu \leq \lambda_2$, for each $(ST)_{\{K_S K_T\}}$ (labeled by +) there exists the transpose set $(TS)_{\{K_T K_S\}}$ (labeled by ○). The asymmetry can be removed for λ_2 odd by relating $(ST) = (\tau + \lambda_2 - \nu, \tau + \nu)$, $0 \leq \nu \leq [\lambda_2/2]$, to $\{\sigma\tau\}$ and $(ST) = (\tau + \lambda_2 - \nu, \tau + \nu)$, $[\lambda_2/2] + 1 \leq \nu \leq \lambda_2$, to $\{\tau\sigma\}$. For λ_2 even, however, the asymmetry associated with $(ST) = (\tau + \frac{1}{2}\lambda_2, \tau + \frac{1}{2}\lambda_2)$ cannot be removed. The choice made by Eqs. (4.8) is therefore again one of convention. Its simplifying feature is manifest in the form of Eqs. (4.8a) and (4.8b). For $\sigma = \tau = \sigma'$, an asymmetry only exists if $\{K_S K_T\} = \{-\sigma', \sigma'\}$. It is related to the fact that the transpose of $(ST)_{\{K_S, -K_S\}}$ is not allowed because $\{-K_S, K_S\}$ is related to $\{K_S, -K_S\}$ by inversion. The singularity of the point $\{K_S K_T\} = \{00\}$ is manifest in the form of Eq. (4.8c).

The eight degeneracy diagrams of Fig. 4 illustrate in complete detail the result of associating (ST)-values as prescribed by Eqs. (4.8) with the $\{K_S K_T\}$-pairs defined by Eqs. (4.7). The examples shown correspond to symmetry types (a) $(\lambda_1\lambda_3)$-(odd, even), (b) $(\lambda_1\lambda_3)$-(odd, odd), (c) $(\lambda_1\lambda_3)$-(even, odd), (d) $(\lambda_1\lambda_3)$-(even, even) for two cases, λ_2 zero and λ_2 such that the degeneracy of $S = T = P'$ is a maximum. On each degeneracy diagram the $\{K_S K_T\}$-lattice corresponding to Eqs. (4.7) is given in outline form. Note that for symmetry types (a) and (b) the $\{K_S K_T\}$-lattices are rectangular ($K_S + K_T = 0$ not allowed). The corresponding degeneracy diagrams reflect a maximum degree of regularity. For symmetry types (c) and (d) the $\{K_S K_T\}$-lattices are not rectangular ($K_S + K_T = 0$ allowed). Nevertheless, since symmetry type (c) is equivalent to symmetry type (a) under conjugation ($\lambda_1\lambda_3$ interchange), degeneracy diagrams of type (c) also possess a maximum degree of regularity. For symmetry type (d), however, the singularity of the point $\{K_S K_T\} = \{00\}$ is an inherent feature which propagates an irregularity into the multiplicities of the (ST)-values associated with $(ST) = (00)$ by $\lambda \leq \lambda_2$ regular lattice displacements.

B. Completeness of the Projected States

First of all, consider the multiplicity $N^R_{(ST)}(\lambda_1\lambda_2\lambda_3)$ of (ST)-values predicted by Eqs. (4.8). As can be seen from Fig. 3, the basic structure of the rule is one of triangulation. That is, the (ST)-values associated with each $\{K_S K_T\}$-pair for $\lambda_2 \geq 0$ are simply those

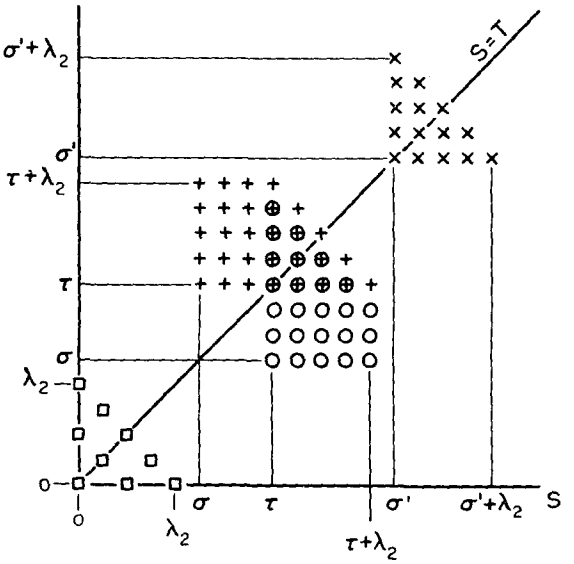


FIG. 3. Spectrum of (ST)-values associated with the projection numbers $\{K_S K_T\}$:

$\{K_S K_T\} = \{\sigma\tau\}$: +, $\{K_S K_T\} = \{\tau\sigma\}$: ○,

$\{K_S K_T\} = \{\sigma'\sigma'\}$: ×, $\{K_S K_T\} = \{0, 0\}$: □.

In the examples shown, $\lambda_2 = 4$.

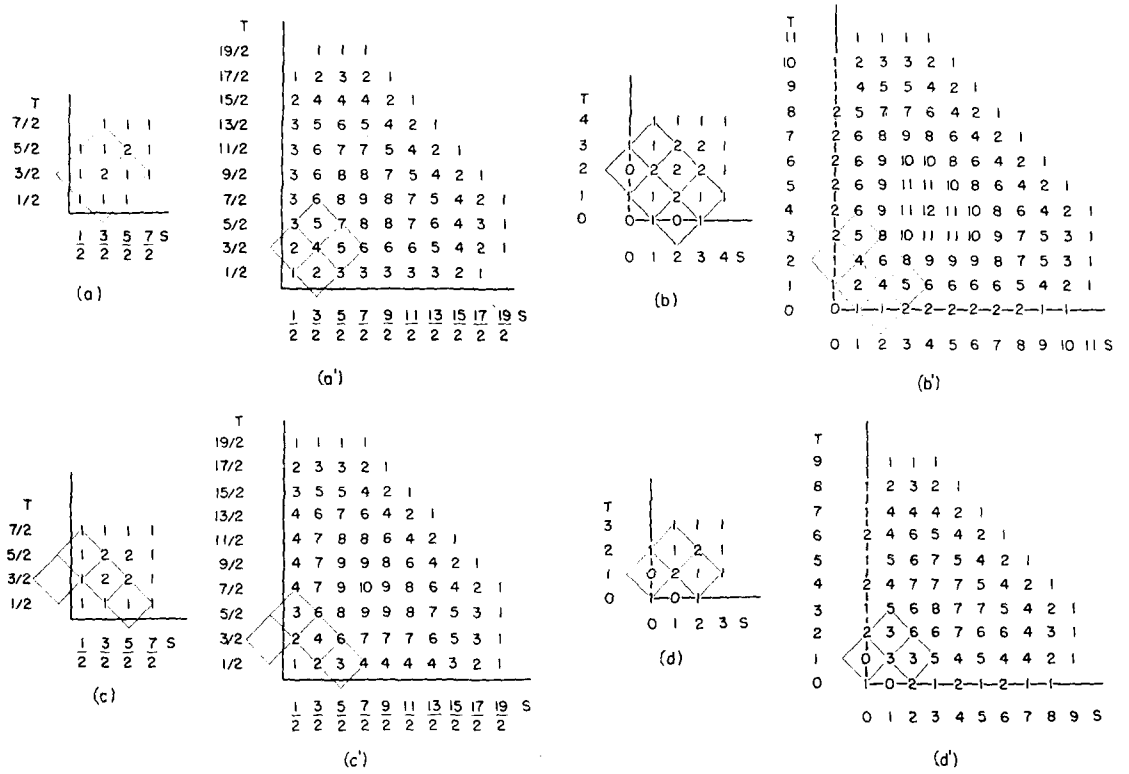


FIG. 4. Spin–isospin degeneracy diagrams for the $(\lambda_1\lambda_2\lambda_3)$ IR of $SU(4)$. (a) $N_{(ST)}(5, 0, 2)$, (b) $N_{(ST)}(5, 0, 3)$, (c) $N_{(ST)}(4, 0, 3)$, (d) $N_{(ST)} \times (4, 0, 2)$, (a′) $N_{(ST)}(5, 6, 2)$, (b′) $N_{(ST)}(5, 7, 3)$, (c′) $N_{(ST)}(4, 6, 3)$, (d′) $N_{(ST)}(4, 6, 2)$. The $\{K_SK_T\}$-lattices given by Eq. (4.7) are included in outline form. The value of λ_2 in (a′), (b′), (c′), and (d′) corresponds to a maximum value for the degeneracy of $S = T = P'$.

(ST)-values contained within the envelope of isosceles right triangles built by λ_2 regular lattice displacements from the (ST)-values associated with $\{K_SK_T\}$ for $\lambda_2 = 0$. The one exception, $\{K_SK_T\} = \{00\}$, admits only the subset of these (ST)-values for which $S + T$ differs from λ_2 by twice an integer ($U - \lambda_2$ even). It therefore follows that the $\{K_SK_T\}$-pairs that contribute to $N^R_{(ST)}(\lambda_1\lambda_2\lambda_3)$ are the $\{K_SK_T\}$-pairs that contribute to the $N^R_{(S'T')}(\lambda_1 0\lambda_3)$ related to $N^R_{(ST)}(\lambda_1\lambda_2\lambda_3)$ in the same way as the $N_{(S'T')}(\lambda_1 0\lambda_3)$ are related to $N_{(ST)}(\lambda_1\lambda_2\lambda_3)$. That is, $N^R_{(ST)}(\lambda_1\lambda_2\lambda_3)$ satisfies Eqs. (3.4). It remains to prove that

$$N^R_{(ST)}(\lambda_1 0\lambda_3) = N_{(ST)}(\lambda_1 0\lambda_3).$$

Consider Eqs. (4.8) for the special case $\lambda_2 = 0$:

$$\sigma > \tau\colon\ (ST) = (\sigma, \tau + \nu),\quad 0 \le \nu < \sigma - \tau; \tag{4.10a}$$

$$\sigma \le \tau\colon\ (ST) = (\sigma + \mu, \tau),\quad 0 \le \mu \le \tau - \sigma. \tag{4.10b}$$

Then $N^R_{(ST)}(\lambda_1 0\lambda_3)$ is equal to the number of $\{K_SK_T\}$-pairs given by Eqs. (4.7) for which (ST) is contained in the set given by Eqs. (4.10):

$$S > T\colon\quad N^R_{(ST)}(\lambda_1 0\lambda_3) = \text{number of } \{K_SK_T\}\text{-pairs for which } \sigma = S,\ \tau \le T; \tag{4.11a}$$

$$S \le T\colon\quad N^R_{(ST)}(\lambda_1 0\lambda_3) = \text{number of } \{K_SK_T\}\text{-pairs for which } \sigma \le S,\ \tau = T. \tag{4.11b}$$

The algebraic formulation is straightforward; it leads directly to the result that $N^R_{(ST)}(\lambda_1 0\lambda_3) = N_{(ST)}(\lambda_1 0\lambda_3)$ and hence $N^R_{(ST)}(\lambda_1\lambda_2\lambda_3) = N_{(ST)}(\lambda_1\lambda_2\lambda_3)$. On the degeneracy diagrams of Fig. 4 the $\{K_SK_T\}$-lattices corresponding to Eqs. (4.7) have been included. By using Eqs. (4.11) the result can be verified for each of the four cases (a) $(\lambda_1\lambda_3)$-(odd, even), (b) $(\lambda_1\lambda_3)$-(odd, odd), (c) $(\lambda_1\lambda_3)$-(even, odd), and (d) $(\lambda_1\lambda_3)$-(even, even).

To complete the proof of the projection hypothesis, an adaptation of the method first given by Elliott[5] for the $SU(3) \supset R(3)$ reduction and subsequently used by Williams and Pursey[15] in considering the $R(5) \supset R(3)$ reduction problem will be used. It proceeds by *reductio ad absurdum*. That is, the consequence of assuming that the projected states do not

span the IR space is shown to be a contradiction. Explicitly, suppose there exists a function $|\varphi(S'M'_S T'M'_T)\rangle$ belonging to the IR but orthogonal to all the $|G_E K_S S M_S K_T T M_T\rangle$,

$$\langle \varphi(S'M'_S T'M'_T) \mid G_E K_S S M_S K_T T M_T\rangle = 0. \quad (4.12)$$

Since $N^R_{(ST)}(\lambda_1\lambda_2\lambda_3) = N_{(ST)}(\lambda_1\lambda_2\lambda_3)$, the only nontrivial implications of such an assumption are those which follow for $S' = S$, $M'_S = M_S$, $T' = T$, and $M'_T = M_T$, namely,

$$\begin{aligned}
&\langle \varphi(SM_S TM_T) \mid G_E\{K_S K_T\}\rangle \\
&\quad = \langle P^S_{M_S M_S} P^T_{M_T M_T} \varphi(SM_S TM_T) \mid G_E\{K_S K_T\}\rangle \\
&\quad = \langle \varphi(SM_S TM_T| \, P^{S\dagger}_{M_S M_S} P^{T\dagger}_{M_T M_T} |G_E\{K_S K_T\}\rangle \\
&\quad = \langle \varphi(SM_S TM_T)| \, P^S_{M_S M_S} P^T_{M_T M_T} |G_E\{K_S K_T\}\rangle \\
&\quad = \delta_{M_S K_S}\delta_{M_T K_T} \\
&\qquad \times \langle \varphi(SM_S TM_T) \mid G_E K_S S M_S K_T T M_T\rangle = 0.
\end{aligned} \quad (4.13)$$

As is shown below, Eq. (4.13) implies that

$$\langle \varphi(SM_S TM_T)| \, \mathcal{O} \, |G_E\{K_S K_T\}\rangle = 0, \quad (4.14)$$

where $\mathcal{O}$ is an arbitrary element of $SU(4)$. But, by definition of irreducibility, functions of the type $\mathcal{O}\,|G_E\{K_S K_T\}\rangle$ span the IR space. Hence a contradiction exists; the hypothesis that there exists a function $|\varphi(SM_S TM_T)\rangle$ belonging to the IR which is orthogonal to all the $|G_E K_S S M_S K_T T M_T\rangle$ is false. It follows that the $|G_E K_S S M_S K_T T M_T\rangle$ span the IR space.

The argument given above hinges upon a proof that Eq. (4.13) implies Eq. (4.14). For this, note that the operator $\mathcal{O}$ being an element of $SU(4)$ implies that it can be expressed as a power series in the generators of the group. Furthermore, note that the commutation properties of the generators imply that the order of the generators within each term of such an expansion can be chosen in any desired manner. Then we define

$$\begin{aligned}
\xi^1_\pm &= \tfrac{1}{2}(S_\pm + E_{\pm 1\,0}), \\
\xi^2_\pm &= \tfrac{1}{2}(S_\pm - E_{\pm 1\,0}), \\
\eta^1_\pm &= \tfrac{1}{2}(T_\pm + E_{0\,\pm 1}), \\
\eta^2_\pm &= \tfrac{1}{2}(T_\pm - E_{0\,\pm 1}),
\end{aligned} \quad (4.15)$$

and consider the case of projection from $|G_{E\uparrow}\rangle$. It is convenient to divide the generators into the two sets

$$\begin{aligned}
\text{A:}\quad E_{00} &= \tfrac{1}{2}(A_{11} - A_{22} - A_{33} + A_{44}), \\
S_0 &= \tfrac{1}{2}(A_{11} + A_{22} - A_{33} - A_{44}), \\
T_0 &= \tfrac{1}{2}(A_{11} - A_{22} + A_{33} - A_{44}), \\
E_{11} &= A_{14}, \quad E_{-1\,-1} = A_{41}, \\
E_{1\,-1} &= A_{23}, \quad E_{-1\,1} = A_{32}, \\
\xi^1_+ &= A_{13}, \quad \xi^2_- = A_{42}, \\
\eta^1_+ &= A_{12}, \quad \eta^2_- = A_{43}, \\
\text{B:}\quad S_+ &= A_{13} + A_{24}, \quad S_- = A_{31} + A_{42}, \\
T_+ &= A_{12} + A_{34}, \quad T_- = A_{21} + A_{43}.
\end{aligned} \quad (4.16a,b)$$

When a generator of the set A operates on $|G_{E\uparrow}\rangle$, the result is either another intrinsic state of the same type ($E_{00}, S_0, T_0, E_{11}, E_{-1\,-1}, E_{1\,-1}, E_{-1\,1}$) or zero ($\xi^1_+$, ξ^2_-, η^1_+, η^2_-). Generators of the set B do not reproduce intrinsic states but are operators which act only in the direct product space $SU(2) \otimes SU(2)$. Express $\mathcal{O}$ in the form

$$\mathcal{O} = \sum_\alpha C_\alpha \pi_{B_\alpha} \pi_{A_\alpha}, \quad (4.17)$$

where the C_α are constants and π_{A_α} and π_{B_α} are products of generators of the type A and B, respectively. Then consider

$$\begin{aligned}
&\langle \varphi(SM_S TM_T)| \, \mathcal{O} \, |G_E\{K_S K_T\}\rangle \\
&\quad = \sum_\alpha C_\alpha \langle \varphi(SM_S TM_T)| \, \pi_{B_\alpha}\pi_{A_\alpha} |G_E\{K_S K_T\}\rangle.
\end{aligned} \quad (4.18)$$

Each factor π_{A_α} acting to the right changes at most K_S and K_T, and the π_{B_α} factors acting to the left change at most M_S and M_T. Therefore,

$$\begin{aligned}
&\langle \varphi(SM_S TM_T)| \, \mathcal{O} \, |G_E\{K_S K_T\}\rangle \\
&\quad = \sum_\alpha C_\alpha \langle \varphi(SM_S TM_T)| \, \pi_{B_\alpha}\pi_{A_\alpha} |G_E\{K_S K_T\}\rangle \\
&\quad = \sum_{\alpha M'_S M'_T K'_S K'_T} C'_\alpha \langle \varphi(SM'_S TM'_T \mid G_E\{K'_S K'_T\}\rangle = 0.
\end{aligned} \quad (4.19)$$

The equivalent proof for the case of projection from $|G_{E\downarrow}\rangle$ follows by merely replacing the ξ^1_+, ξ^2_-, η^1_+, η^2_- operators of set A by the operators ξ^1_-, ξ^2_+, η^1_-, η^2_+.

5. TRANSFORMATION BRACKETS

Although the projection numbers $\{K_S K_T\}$ furnish an integral or half-integral solution exhibiting spin–isospin symmetry properties for the $SU(4) \supset SU(2) \otimes SU(2)$ multiplicity problem, the projected states are not normalized nor are they necessarily orthogonal on the K_S and K_T labels. The difficulties associated with the nonorthonormality of the projected states can be resolved, however, if an expression for the coefficients (transformation brackets) which relate the projected states to the orthonormal Gel'fand basis vectors is known. This section is devoted to deriving such an expression. The method used is similar to that developed in Ref. 16, where the analogous problem in the $SU(3) \supset R(3)$ reduction was considered; it is based on the results of Moshinsky and Chacón[17] for the matrix elements of the permutations (1, 2), (2, 3), and (3, 4) between the $U(4)$ basis states $|G\rangle$.

A. The Expression

Since the Gel'fand basis vectors $|G\rangle$ for a given IR of $U(4)$ form an orthonormal set which spans the representation space, an arbitrary projected state $|GK_SSM_SK_TTM_T\rangle$ belonging to the IR may be expanded in terms of the $|G\rangle$ as

$$|GK_SSM_SK_TTM_T\rangle = \sum_{G'} \langle G' \,|\, GK_SSM_SK_TTM_T\rangle\, |G'\rangle, \quad (5.1)$$

where it is to be understood that $h'_{\alpha 4} = h_{\alpha 4}$. The $\langle G' \,|\, GK_SSM_SK_TTM_T\rangle$ in Eq. (5.1) are the transformation brackets which relate the $U(4) \supset SU(2) \otimes SU(2)$ scheme of Sec. 4 to the Gel'fand $U(4) \supset U(3) \supset U(2) \supset U(1)$ scheme. By definition of the projected states, we have

$$\begin{aligned}\langle G' \,|\, GK_SSM_SK_TTM_T\rangle &= \langle G'|\, P^S_{M_SK_S} P^T_{M_TK_T}\, |G\rangle \\ &= (2S+1)\int d\Omega_S D^{S*}_{M_SK_S}(\Omega)(2T+1) \\ &\quad \times \int d\Omega_T D^{T*}_{M_TK_T}(\Omega_T)\langle G'|\, R_S(\Omega_S)R_T(\Omega_T)\, |G\rangle.\end{aligned} \quad (5.2)$$

Therefore, an expression for the

$$\langle G' \,|\, GK_SSM_SK_TTM_T\rangle$$

can be obtained if the matrix elements

$$\langle G'|\, R_S(\Omega_S)R_T(\Omega_T)\, |G\rangle$$

are known. Note that the inverse of the transformation matrix defined by Eq. (5.1) is only guaranteed to exist if the $|GK_SSM_SK_TTM_T\rangle$ are restricted to the projected basis vectors $|G_EK_SSM_SK_TTM_T\rangle$ defined in Sec. 4 by the projection hypothesis. An expression for the $\langle G' \,|\, G_EK_SSM_SK_TTM_T\rangle$ follows as a special case of the general result for $\langle G' \,|\, GK_SSM_SK_TTM_T\rangle$.

For notational convenience let

$$|G\rangle = \left| \begin{matrix} h_1 & h_2 & h_3 & h_4 \\ & x & y & z \\ & & p & q \\ & & & r \end{matrix} \right\rangle. \quad (5.3)$$

The infinitesimal generators of $SU(2)$ corresponding to $U(2)$ in the chain $U(4) \supset U(3) \supset U(2) \supset U(1)$ are given by

$$J_+ = A_{12}, \quad J_- = A_{21}, \quad J_0 = \tfrac{1}{2}(A_{11} - A_{22}), \quad (5.4)$$

where

$$J_\pm = J_1 \pm iJ_2. \quad (5.5)$$

Then, for

$$\mathcal{R}(\Omega) = e^{-i\alpha J_0}e^{-i\beta J_2}e^{-i\gamma J_0}, \quad (5.6)$$

it follows that

$$\begin{aligned}\langle G'|\, \mathcal{R}(\Omega)\, |G\rangle &= \delta_{x'x}\delta_{y'y}\delta_{z'z}\delta_{p'p}\delta_{q'q}D^j_{m'm}(\Omega), \\ j = \tfrac{1}{2}(p-q), &\quad m = r - \tfrac{1}{2}(p+q), \\ m' &= r' - \tfrac{1}{2}(p+q).\end{aligned} \quad (5.7)$$

To relate $R_S(\Omega)$ and $R_T(\Omega)$ to operators of the type $\mathcal{R}(\Omega)$, the permutation operators (1, 2), (2, 3), and (3, 4) can be used. For example, consider $R_S(\Omega)$. Let

$$\begin{aligned}S_0 &= S_0^1 + S_0^2, \\ S_0^1 &= \tfrac{1}{2}(A_{11} - A_{33}) = (2,3)J_0(2,3), \\ S_0^2 &= \tfrac{1}{2}(A_{22} - A_{44}) \\ &= (1,2)(3,4)(2,3)J_0(2,3)(3,4)(1,2), \\ &[S_0^1, S_0^2] = 0,\end{aligned} \quad (5.8a)$$

$$\begin{aligned}S_2 &= S_2^1 + S_2^2, \\ S_2^1 &= (2i)^{-1}(A_{13} - A_{31}) = (2,3)J_2(2,3), \\ S_2^2 &= (2i)^{-1}(A_{24} - A_{42}) \\ &= (1,2)(3,4)(2,3)J_2(2,3)(3,4)(1,2), \\ &[S_2^1, S_2^2] = [S_0^1, S_2^2] = [S_0^2, S_2^1] = 0.\end{aligned} \quad (5.8b)$$

Then

$$\begin{aligned}R_S(\Omega) &= e^{-i\alpha S_0}e^{-i\beta S_2}e^{-i\gamma S_0} \\ &= e^{-i\alpha S_0^1}e^{-i\alpha S_0^2}e^{-i\beta S_2^1}e^{-i\beta S_2^2}e^{-i\gamma S_0^1}e^{-i\gamma S_0^2} \\ &= e^{-i\alpha S_0^1}e^{-i\beta S_2^1}e^{-i\gamma S_0^1}e^{-i\alpha S_0^2}e^{-i\beta S_2^2}e^{-i\gamma S_0^2} \\ &= (2,3)\mathcal{R}(\Omega)(2,3)(1,2)(3,4)(2,3) \\ &\quad \times \mathcal{R}(\Omega)(2,3)(3,4)(1,2).\end{aligned} \quad (5.9)$$

In a similar fashion it can be shown that

$$\begin{aligned}R_T(\Omega) &= \mathcal{R}(\Omega)(2,3)(1,2)(3,4)(2,3) \\ &\quad \times \mathcal{R}(\Omega)(2,3)(3,4)(1,2)(2,3).\end{aligned} \quad (5.10)$$

From Eqs. (5.9) and (5.10) it follows that

$$\begin{aligned}R_S(\Omega_S)R_T(\Omega_T) &= R_T(\Omega_T)R_S(\Omega_S) \\ &= \mathcal{R}(\Omega_T)(2,3)(1,2)(3,4)(2,3) \\ &\quad \times \mathcal{R}(\Omega_T)(2,3)(1,2)(3,4) \\ &\quad \times \mathcal{R}(\Omega_S)(2,3)(1,2)(3,4)(2,3) \\ &\quad \times \mathcal{R}(\Omega_S)(2,3)(1,2)(3,4).\end{aligned} \quad (5.11)$$

Define

$$\begin{aligned}M_{G''G}(\Omega) = \langle G'|\, &\mathcal{R}(\Omega)(2,3)(1,2)(3,4)(2,3) \\ &\times \mathcal{R}(\Omega)(2,3)(1,2)(3,4)\, |G\rangle\end{aligned} \quad (5.12)$$

so that

$$\langle G'|\, R_S(\Omega_S)R_T(\Omega_T)\, |G\rangle = \sum_{G''} M_{G'G''}(\Omega_T)M_{G''G}(\Omega_S). \quad (5.13)$$

Let

$$\mathcal{M}_{G'G}(KJM) = (2J+1)\int d\Omega D^{J*}_{MK}(\Omega)M_{G'G}(\Omega). \quad (5.14)$$

The transformation brackets of Eq. (5.2) are then given by

$$\langle G' \mid GK_S SM_S K_T TM_T\rangle = \sum_{G''} \mathcal{M}_{G'G''}(K_T TM_T)\mathcal{M}_{G''G}(K_S SM_S). \quad (5.15)$$

An expression for the matrix $\mathcal{M}_{G'G}(KJM)$ can be obtained by using the completeness of the orthonormal set of states $|G\rangle$ and Eq. (5.7) to put $M_{G'G}(\Omega)$ into the form

$$\begin{aligned} M_{G'G}(\Omega) &= \sum_{G_\alpha} \langle G'|\, \mathfrak{R}(\Omega)\, |G_1\rangle\langle G_1|\, (2,3)\, |G_2\rangle\langle G_2|\, (1,2)\, |G_3\rangle \\ &\quad \times \langle G_3|\, (3,4)\, |G_4\rangle\langle G_4|\, (2,3)\, |G_5\rangle\langle G_5|\, \mathfrak{R}(\Omega)\, |G_6\rangle \\ &\quad \times \langle G_6|\, (2,3)\, |G_7\rangle\langle G_7|\, (1,2)\, |G_8\rangle\langle G_8|\, (3,4)\, |G\rangle \\ &= \sum_{\substack{G_\alpha(\alpha\neq 1,5)\\ K'M''}} D^{J'}_{M'K'}(\Omega) D^{J''}_{M''K''}(\Omega)\langle G_1|\, (2,3)\, |G_2\rangle \\ &\quad \times \langle G_2|\, (1,2)\, |G_3\rangle\langle G_3|\, (3,4)\, |G_4\rangle\langle G_4|\, (2,3)\, |G_5\rangle \\ &\quad \times \langle G_6|\, (2,3)\, |G_7\rangle\langle G_7|\, (1,2)\, |G_8\rangle\langle G_8|\, (3,4)\, |G\rangle, \\ J' &= \tfrac{1}{2}(p'-q'), \quad M' = r' - \tfrac{1}{2}(p'+q'), \\ K' &= r_1 - \tfrac{1}{2}(p'+q'), \\ J'' &= \tfrac{1}{2}(p_6 - q_6), \quad M'' = r_5 - \tfrac{1}{2}(p_6+q_6), \\ K'' &= r_6 - \tfrac{1}{2}(p_6+q_6), \end{aligned} \quad (5.16)$$

where, except for r_1 (determined by K') and r_5 (determined by M''), the elements of G_1 and G_5 are equal to the corresponding elements in the G' and G_6, respectively. Then, by using the well-known result expressing the integral of three rotation matrices in terms of a product of two $SU(2)$ Wigner (Clebsch–Gordan) coefficients, it follows that

$$\begin{aligned} \mathcal{M}_{G'G}(KJM) &= \sum_{G_\alpha(\alpha\neq 1,5)} \langle J'M'; J''M'' \mid JM\rangle\langle J'K'; J''K'' \mid JK\rangle \\ &\quad \times \langle G_1|\, (2,3)\, |G_2\rangle\langle G_2|\, (1,2)\, |G_3\rangle\langle G_3|\, (3,4)\, |G_4\rangle \\ &\quad \times \langle G_4|\, (2,3)\, |G_5\rangle\langle G_6|\, (2,3)\, |G_7\rangle\langle G_7|\, (1,2)\, |G_8\rangle \\ &\quad \times \langle G_8|\, (3,4)\, |G\rangle. \end{aligned} \quad (5.17)$$

The permutation matrices $\langle G'|\, (n-1,n)\, |G\rangle$, $n = 2, 3, 4$, required for an evaluation of Eq. (5.17), have been given by Moshinsky and Chacón[17]; they are equivalent to special unitary recoupling coefficients for the groups $U(1)$, $U(2)$, and $U(3)$, respectively. Note that $(n-1, n)$ operating on $|G\rangle$ changes only the $h_{\alpha\beta}$ for which $\beta = n-1$ and these in such a manner that the result is zero unless $w'_{n-1} = w_n$. The apparent $6 \times 6 = 36$-fold sum over the G_α in Eq. (5.17) is therefore in actual fact at worst a sixfold sum. The result as given by Eq. (5.17) may, however, be the most convenient for the purposes of machine coding since the summations over G_2, G_3, G_4 and G_7, G_8 are matrix multiplications involving the permutation matrices. The remaining summation over G_6 then involves simply the product of two Clebsch–Gordan coefficients and one element from each of the matrix products.

It is to be noted that the transformation brackets are equivalent to normalization and overlap integrals of the projected states. This may be seen by considering

$$\begin{aligned} \langle G'K'_S SM_S K'_T TM_T \mid GK_S SM_S K_T TM_T\rangle &= \langle G'|\, P^{S\dagger}_{M_S K'_S} P^{T\dagger}_{M_T K'_T} |GK_S SM_S K_T TM_T\rangle \\ &= \langle G'|\, P^{S}_{K'_S M_S} P^{T}_{K'_T M_T} |GK_S SM_S K_T TM_T\rangle \\ &= \langle G' \mid GK_S SK'_S K_T TK'_T\rangle. \end{aligned} \quad (5.18)$$

B. The Application

In general, the transformation brackets[18]

$$A(G' \mid G_E K_S SM_S K_T TM_T)$$

relate the set of nonorthogonal basis vectors $|G_E K_S SM_S K_T TM_T\rangle$ to the set of orthonormal basis vectors $|G'\rangle$ and are therefore the elements of a nonorthogonal matrix A. The inverse expansion of the $|G\rangle$ in terms of the $|G_E K'_S S'M'_S K'_T T'M'_T\rangle$ exists, and the coefficients $B(G_E K'_S S'M'_S K'_T T'M'_T \mid G)$ can be obtained by inverting the appropriate A matrix. An equivalent but perhaps somewhat simpler evaluation of these coefficients can be obtained by considering directly the expansion

$$|G\rangle = \sum_{\substack{K'_S S'M'_S\\ K'_T T'M'_T}} B(G_E K'_S S'M'_S K'_T T'M'_T \mid G) \times |G_E K'_S S'M'_S K'_T T'M'_T\rangle. \quad (5.19)$$

Then

$$\begin{aligned} |GK_S SM_S K_T TM_T\rangle &= P^S_{M_S K_S} P^T_{M_T K_T} |G\rangle \\ &= \sum_{\substack{K'_S S'M'_S\\ K'_T T'M'_T}} B(G_E K'_S S'M'_S K'_T T'M'_T \mid G) \\ &\quad \times P^S_{M_S K_S} P^T_{M_T K_T} |G_E K'_S S'M'_S K'_T T'M'_T\rangle \\ &= \sum_{\substack{K'_S S'M'_S\\ K'_T T'M'_T}} B(G_E K'_S S'M'_S K'_T T'M'_T \mid G) \\ &\quad \times \delta_{S'S}\delta_{M'_S K_S}\delta_{T'T}\delta_{M'_T K_T} |G_E K'_S SM_S K'_T TM_T\rangle \\ &= \sum_{K'_S K'_T} B(G_E K'_S SK_S K'_T TK_T \mid G) \\ &\quad \times |G_E K'_S SM_S K'_T TM_T\rangle. \end{aligned} \quad (5.20)$$

That is, the $B(G_E K'_S S'M'_S K'_T T'M'_T \mid G)$ are not only the coefficients in the expansion of the $|G\rangle$ in terms of the $|G_E K'_S S'M'_S K'_T T'M'_T\rangle$, but they are also the

coefficients in the expansion of $|GK_SSM_SK_TTM_T\rangle$ in terms of the $|G_EK'_SSM_SK'_TTM_T\rangle$. Using this result, we can determine a unique solution for the

$$B(G_EK'_SS'M'_SK'_TT'M'_T \mid G)$$

from the set of simultaneous equations

$$A(G' \mid GK_SSM_SK_TTM_T) = \sum_{K'_SK'_T} B(G_EK'_SSK_SK'_TTK_T \mid G) \times A(G' \mid G_EK'_SSM_SK'_TTM_T). \quad (5.21)$$

In those cases for which the $\{K_SK_T\}$-labels are redundant, it follows that the $B(G_EK'_SS'M'_SK'_TT'M'_T \mid G)$ are simply given as the ratio of two transformation brackets. Since B is the inverse of A, Eq. (5.21) also shows that

$$\sum_{SM_STM_T} A(G' \mid GM_SSM_SM_TTM_T) = \delta_{G'G} \quad (5.22)$$

and

$$\sum_G A(G \mid GK_SSM_SK_TTM_T) = \delta_{M_SK_S}\delta_{M_TK_T} \sum_{K'_SK'_T} = \delta_{M_SK_S}\delta_{M_TK_T}N_{(ST)}(\lambda_1\lambda_2\lambda_3). \quad (5.23)$$

In a fashion similar to that demonstrated in detail in Ref. 16 for the $SU(3) \supset R(3)$ case, quantities of physical interest which depend upon the $SU(4) \supset SU(2) \otimes SU(2)$ labels can be expressed in terms of the corresponding quantities labeled according to the canonical $U(4) \supset U(3) \supset U(2) \supset U(1)$ scheme by means of the A's and B's. For example, for the $SU(4) \supset SU(2) \otimes SU(2)$ coupling coefficients defined by

$$|\rho G_{3E}K_{S_3}S_3M_{S_3}K_{T_3}T_3M_{T_3}\rangle = \sum_{\substack{K_{S_1}S_1M_{S_1}K_{T_1}T_1M_{T_1} \\ K_{S_2}S_2M_{S_2}K_{T_2}T_2M_{T_2}}} C_1(G_{1E}K_{S_1}S_1M_{S_1}K_{T_1}T_1M_{T_1}; G_{2E}K_{S_2}S_2M_{S_2}K_{T_2}T_2M_{T_2} \mid \rho G_{3E}K_{S_3}S_3M_{S_3}K_{T_3}T_3M_{T_3}) \times |G_{1E}K_{S_1}S_1M_{S_1}K_{T_1}T_1M_{T_1}\rangle |G_{2E}K_{S_2}S_2M_{S_2}K_{T_2}T_2M_{T_2}\rangle, \quad (5.24a)$$

$$|G_{1E}K_{S_1}S_1M_{S_1}K_{T1}T_1M_{T_1}\rangle |G_{2E}K_{S_2}S_2M_{S_2}K_{T_2}T_2M_{T_2}\rangle = \sum_{\rho G_{3E}K_{S_3}S_3M_{S_3}K_{T_3}T_3M_{T_3}} C_2(\rho G_{3E}K_{S_3}S_3M_{S_3}K_{T_3}T_3M_{T_3} \mid G_{1E}K_{S_1}S_1M_{S_1}K_{T_1}T_1M_{T_1}; G_{2E}K_{S_2}S_2M_{S_2}K_{T_2}T_2M_{T_2}) \times |\rho G_{3E}K_{S_3}S_3M_{S_3}K_{T_3}T_3M_{T_3}\rangle, \quad (5.24b)$$

it can be shown that

$$C_1(G_{1E}K_{S_1}S_1M_{S_1}K_{T_1}T_1M_{T_1}; G_{2E}K_{S_2}S_2M_{S_2}K_{T_2}T_2M_{T_2} \mid \rho G_{3E}K_{S_3}S_3M_{S_3}K_{T_3}T_3M_{T_3}) = \langle S_1M_{S_1}; S_2M_{S_2} \mid S_3M_{S_3}\rangle\langle T_1M_{T_1}; T_2M_{T_2} \mid T_3M_{T_3}\rangle \sum_{\substack{G'_1M'_{S_1}M'_{T_1} \\ G'_2M'_{S_2}M'_{T_2}}} \langle S_1M'_{S_1}; S_2M'_{S_2} \mid S_3K_{S_3}\rangle \times \langle T_1M'_{T_1}; T_2M'_{T_2} \mid T_3K_{T_3}\rangle B(G_{1E}K_{S_1}S_1M'_{S_1}K_{T_1}T_1M'_{T_1} \mid G'_1) \times B(G_{2E}K_{S_2}S_2M'_{S_2}K_{T_2}T_2M'_{T_2} \mid G'_2)\langle G'_1; G'_2 \mid \rho G_{3E}\rangle, \quad (5.25a)$$

$$C_2(\rho G_{3E}K_{S_3}S_3M_{S_3}K_{T_3}T_3M_{T_3} \mid G_{1E}K_{S_1}S_1M_{S_1}K_{T_1}T_1M_{T_1}; G_{2E}K_{S_2}S_2M_{S_2}K_{T_2}T_2M_{T_2}) = \langle S_1M_{S_1}; S_2M_{S_2} \mid S_3M_{S_3}\rangle\langle T_1M_{T_1}; T_2M_{T_2} \mid T_3M_{T_3}\rangle \frac{(2S_1+1)(2T_1+1)}{(2S_3+1)(2T_3+1)} \times \sum_{\substack{G'_3M'_{S_2}M'_{T_2} \\ G'_3M'_{S_3}M'_{T_3}}} \langle S_1K_{S_1}; S_2M'_{S_2} \mid S_3M'_{S_3}\rangle\langle T_1K_{T_1}; T_2M'_{T_2} \mid T_3M'_{T_3}\rangle B(G_{3E}K_{S_3}S_3M'_{S_3}K_{T_3}T_3M'_{T_3} \mid G'_3) \times \langle \rho G'_3 \mid G_{1E}; G'_2\rangle A(G'_2 \mid G_{2E}K_{S_2}S_2M'_{S_2}K_{T_2}T_2M'_{T_2}), \quad (5.25b)$$

where ρ is a label that distinguishes multiple occurrences of a given IR of G_3 in the reduction of the direct product $G_1 \otimes G_2$. In Eqs. (5.25), $\langle G_1; G_2 \mid \rho G_3\rangle$ and $\langle \rho G_3 \mid G_1; G_2\rangle$ are $U(4) \supset U(3) \supset U(2) \supset U(1)$ Wigner coefficients, and the $\langle J_1M_1; J_2M_2 \mid J_3M_3\rangle$ are ordinary $SU(2)$ Wigner coefficients.

Similarly, consider the $SU(4) \supset SU(2) \otimes SU(2)$ tensors defined by

$$T(GK_SSM_SK_TTM_T) = (2S+1)\int d\Omega_S D^{S*}_{M_SK_S}(\Omega_S)(2T+1) \times \int d\Omega_T D^{T*}_{M_TK_T}(\Omega_T) \times R_S(\Omega_S)R_T(\Omega_T)T(G)R_T^{-1}(\Omega_T)R_S^{-1}(\Omega_S), \quad (5.26)$$

where $T(G)$ is the corresponding $U(4) \supset U(3) \supset U(2) \supset U(1)$ tensor defined by

$$[A_{\alpha\beta}, T(G)] = \sum_{G'} \langle G' | A_{\alpha\beta} | G \rangle T(G'). \tag{5.27}$$

The $\{K_S K_T\}$-quantum-numbers resolve the $SU(4) \supset SU(2) \otimes SU(2)$ tensorial multiplicity in precisely the same manner as described in Sec. 3 for the $SU(4) \supset SU(2) \otimes SU(2)$ basis states. It can then be shown that

$$\begin{aligned}\langle \rho G_{3E} K_{S_3} S_3 M_{S_3} K_{T_3} T_3 M_{T_3} | \, T(G_{1E} K_{S_1} S_1 M_{S_1} K_{T_1} T_1 M_{T_1}) \, | G_{2E} K_{S_2} S_2 M_{S_2} K_{T_2} T_2 M_{T_2} \rangle \\ = \langle G_3 \| \, T(G_1) \, \| G_2 \rangle_\rho \sum_{K'_{S_3} K'_{T_3}} C_2(\rho G_{3E} K'_{S_3} S_3 M_{S_3} K'_{T_3} T_3 M_{T_3} \, | \, G_{1E} K_{S_1} S_1 M_{S_1} K_{T_1} T_1 M_{T_1}; \, G_{2E} K_{S_2} S_2 M_{S_2} K_{T_2} T_2 M_{T_2}) \\ \times A(G_{3E} \, | \, \rho G_{3E} K'_{S_3} S_3 K_{S_3} K'_{T_3} T_3 K_{T_3}),\end{aligned} \tag{5.28}$$

where $\langle G_3 \| \, T(G_1) \, \| G_2 \rangle_\rho$ is the reduced matrix element of $T(G_1)$ corresponding to the state $|_\rho G_3\rangle$.

The particularly elegant feature of all such relationships is that a knowledge of the A's and B's allows completely general expressions for $SU(4) \supset SU(2) \otimes SU(2)$ quantities to be expressed in terms of a subset of the corresponding $U(4) \supset U(3) \supset U(2) \supset U(1)$ quantities [e.g., all $SU(4) \supset SU(2) \otimes SU(2)$ coupling coefficients are determined in terms of $U(4) \supset U(3) \supset U(2) \supset U(1)$ Wigner coefficients for which one set of labels corresponds to the operator E_{00} having either its maximum or minimum eigenvalue]. Furthermore, the problems associated with phase conventions and multiplicity relate simply and directly to the corresponding problems in the canonical scheme.

6. DISCUSSION

The fact that a many-nucleon wavefunction can be decomposed into a product of its space and its spin–isospin parts allows the techniques developed in this paper to be applied quite independently of any special spatial considerations. A case of particular interest, however, is that dealing with shell-model calculations up to and through the first half of the $2s$-$1d$ shell. For such nuclei the most promising theoretical tool for the spatial part of the wavefunction is the Elliott $SU(3) \supset R(3)$ classification. For this reason the techniques developed in Ref. 16 together with those of the present paper furnish expressions which can be used to simplify as well as extend such theoretical investigations.

The simplifications are, of course, in calculational technique in that the $SU(3) \supset R(3)$ and $SU(4) \supset SU(2) \otimes SU(2)$ transformation brackets reduce the difficulties inherent in the physically significant labeling schemes, but not present in the corresponding canonical labeling schemes, to forms which can be machine coded. Nevertheless, the solution furnished by the transformation brackets to the problems associated with the nonorthonormality of the projected states is indirect and not necessarily the most convenient for purposes of machine-coding matrix element calculations. The difficulty is that the $SU(3) \supset R(3)$ coupling coefficients of Ref. 16 and the $SU(4) \supset SU(2) \otimes SU(2)$ coupling coefficients of the present paper are not Wigner coefficients. That is, the coupling coefficients do not represent the scalar product of orthonormalized coupled and uncoupled basis states.

By orthonormalizing separately within each L and (ST)-multiplet according to a symmetric recipe (e.g., see Ref. 19), the transformations which orthonormalize the $SU(3) \supset R(3)$ and $SU(4) \supset SU(2) \otimes SU(2)$ basis states can be given in simple algebraic form as the ratio of normalization and overlap integrals. Since such integrals are equivalent to transformation brackets, the problems associated with the nonorthonormality of the projected states can be resolved. And, in particular, they can be resolved in a form convenient for machine coding while still maintaining all the simplifications associated with the projective processes. In fact, the $SU(3) \supset R(3)$ and $SU(4) \supset SU(2) \otimes SU(2)$ orthonormalizing transformations can be incorporated directly into programs which calculate the transformation brackets. The result is then $SU(3) \supset R(3)$ and $SU(4) \supset SU(2) \otimes SU(2)$ transformation brackets which relate physically significant orthonormal basis states to the corresponding canonical basis states. Within such a framework the $SU(3) \supset R(3)$ coupling coefficients of Ref. 13 and the $SU(4) \supset SU(2) \otimes SU(2)$ coupling coefficients of the present paper become Wigner coefficients, and hence standard algebraic techniques introduced by Racah[20] can be applied to simplify matrix element calculations.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge helpful discussions with D. L. Pursey and S. A. Williams, Iowa State University, and to thank K. T. Hecht, University of Michigan, for his careful reading and constructive comments on the manuscript. The author would also like to thank Professor A. Bohr, Niels Bohr Institute,

for the hospitality and interest shown during the early stages of the research reported in this paper.

* Publication fees supported by the U.S. Office of Naval Research, Contract Nonr-1224(59).

† National Science Foundation Postdoctoral Fellow.

[1] E. P. Wigner, Phys. Rev. **51**, 106 (1937).

[2] M. Moshinsky and J. G. Nagel, Phys. Letters **5**, 173 (1963); R. T. Sharp and C. S. Lam, J. Math. Phys. **10**, 2033 (1969); M. Brunet and M. Resnikoff, J. Math. Phys. **11**, 1471, 1474 (1970).

[3] I. M. Gel'fand and M. L. Zetlin, Dokl. Akad. Nauk SSSR **71**, 825 (1950).

[4] K. T. Hecht and S. C. Pang, J. Math. Phys. **10**, 1571 (1969).

[5] J. P. Elliott, Proc. Roy. Soc. (London) **A245**, 128, 562 (1958).

[6] $SU(4)$ conjugation properties are defined and discussed more fully by Brunet and Resnikoff, Ref. 2.

[7] The supermultiplet quantum numbers $(PP'P'')$ specify the IR of $O(6)$ which is locally isomorphic to $SU(4)$ according to the standard Gel'fand labeling scheme.

[8] For notational convenience, states of the type (2.8) will be denoted simply by $|G_E\rangle$, an arrow being added to the E when a distinction between types (a) and (b) is required. The $\{K_S K_T\}$-label will only be included when it is necessary to distinguish between $|G_E\rangle$ states of different spin-isospin projections.

[9] G. Racah, Rev. Mod. Phys. **21**, 494 (1949).

[10] M. Kretzschmar, Z. Physik **157**, 558 (1960).

[11] A. M. Perelomov and V. S. Popov, Yad. Fiz. **2**, 738 (1965) [Sov. J. Nucl. Phys. **2**, 528 (1966)].

[12] D. E. Littlewood, *The Theory of Group Characters* (Oxford U.P., London, 1940).

[13] A regular lattice displacement is one which increases either S or T by one unit.

[14] The Löwdin–Shapiro form for the projection operator may also prove to be useful in some applications; for example, see Yu. F. Smirnov, "Projection Operators for Semisimple Lie Groups", *Proceedings on the International Conference on Clustering Phenomena in Nuclei, Bochum, Germany, 1969* (International Atomic Energy Agency, Vienna, 1969).

[15] S. A. Williams and D. L. Pursey, J. Math. Phys. **9**, 1230 (1968).

[16] J. P. Draayer and S. A. Williams, Nucl. Phys. **A129**, 647 (1969).

[17] M. Moshinsky and E. Chacón, *Spectroscopic and Group Theoretical Methods in Physics*, edited by F. Bloch, S. G. Cohen, A. de-Shalit, S. Sambursky, and I. Talmi (North-Holland, Amsterdam, 1968).

[18] To avoid possible confusion, the notational change

$$A(G' \mid GK_S SM_S K_T TM_T) \equiv \langle G' \mid GK_S SM_S K_T TM_T \rangle,$$

expressing the transformation brackets as the elements of a matrix A, with rows labeled by the $h'_{\alpha\beta}$ and columns labeled by the $\{K_S SM_S K_T TM_T\}$-values, will be adopted in this section.

[19] J. D. Vergados, Nucl. Phys. **A111**, 681 (1968).

[20] U. Fano and G. Racah, *Irreducible Tensorial Sets* (Academic, New York, 1959).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Theorem of Uniqueness and Local Stability for Liouville–Einstein Equations*

YVONNE CHOQUET-BRUHAT
Département de Mathématiques, Faculté des Sciences, Université de Paris
11 rue Pierre-Curie, 75 Paris 5e, France

(Received 24 March 1970)

We prove, by use of energy inequalities, a theorem of uniqueness and local (i.e., for finite time) stability for the solution of Cauchy problem relative to the integro-differential system of Einstein and Liouville. A global theorem of geometrical uniqueness follows from a general method, previously given. We will prove elsewhere an existence theorem.

INTRODUCTION

The aim of this paper is to prove a uniqueness theorem for the solution of the Cauchy problem for the coupled Liouville–Einstein equations, i.e., for a collisionless relativistic gas under its own gravitational field. Such a gas provides a model reasonably appropriate for physical systems like systems of galaxies or some systems of stars (which are then the "particles" of the gas) and certain plasmas or radiations (in this last case the particles have a zero rest mass).

With the uniqueness theorem we prove a local stability theorem; i.e., we prove that the solution (metric and distribution function) depends continuously on the initial data: such a theorem, which states that a small initial perturbation gives rise to a small perturbation during some finite time, seems the first necessary step to be assured of before any more elaborate research on stability.

The plan of this paper is the following:

In Sec. I, I give a brief review of the fundamental concepts of relativistic kinetic theory, and I recall the equations governing the motion of a self-gravitating collisionless, relativistic gas: the coupled Einstein and Liouville equations. I also recall, or establish, a few general properties of these equations which will be used in the following (i.e., local equivalence of Einstein equations in harmonic coordinates and tensorial Einstein equations, and use of bounded parameters for the momenta in the Liouville equation).

In Sec. II, I establish some inequalities satisfied by the difference of two solutions of the Cauchy problem

for the hospitality and interest shown during the early stages of the research reported in this paper.

* Publication fees supported by the U.S. Office of Naval Research, Contract Nonr-1224(59).

† National Science Foundation Postdoctoral Fellow.

[1] E. P. Wigner, Phys. Rev. **51**, 106 (1937).

[2] M. Moshinsky and J. G. Nagel, Phys. Letters **5**, 173 (1963); R. T. Sharp and C. S. Lam, J. Math. Phys. **10**, 2033 (1969); M. Brunet and M. Resnikoff, J. Math. Phys. **11**, 1471, 1474 (1970).

[3] I. M. Gel'fand and M. L. Zetlin, Dokl. Akad. Nauk SSSR **71**, 825 (1950).

[4] K. T. Hecht and S. C. Pang, J. Math. Phys. **10**, 1571 (1969).

[5] J. P. Elliott, Proc. Roy. Soc. (London) **A245**, 128, 562 (1958).

[6] $SU(4)$ conjugation properties are defined and discussed more fully by Brunet and Resnikoff, Ref. 2.

[7] The supermultiplet quantum numbers $(PP'P'')$ specify the IR of $O(6)$ which is locally isomorphic to $SU(4)$ according to the standard Gel'fand labeling scheme.

[8] For notational convenience, states of the type (2.8) will be denoted simply by $|G_E\rangle$, an arrow being added to the E when a distinction between types (a) and (b) is required. The $\{K_S K_T\}$-label will only be included when it is necessary to distinguish between $|G_E\rangle$ states of different spin-isospin projections.

[9] G. Racah, Rev. Mod. Phys. **21**, 494 (1949).

[10] M. Kretzschmar, Z. Physik **157**, 558 (1960).

[11] A. M. Perelomov and V. S. Popov, Yad. Fiz. **2**, 738 (1965) [Sov. J. Nucl. Phys. **2**, 528 (1966)].

[12] D. E. Littlewood, *The Theory of Group Characters* (Oxford U.P., London, 1940).

[13] A regular lattice displacement is one which increases either S or T by one unit.

[14] The Löwdin–Shapiro form for the projection operator may also prove to be useful in some applications; for example, see Yu. F. Smirnov, "Projection Operators for Semisimple Lie Groups", *Proceedings on the International Conference on Clustering Phenomena in Nuclei, Bochum, Germany, 1969* (International Atomic Energy Agency, Vienna, 1969).

[15] S. A. Williams and D. L. Pursey, J. Math. Phys. **9**, 1230 (1968).

[16] J. P. Draayer and S. A. Williams, Nucl. Phys. **A129**, 647 (1969).

[17] M. Moshinsky and E. Chacón, *Spectroscopic and Group Theoretical Methods in Physics*, edited by F. Bloch, S. G. Cohen, A. de-Shalit, S. Sambursky, and I. Talmi (North-Holland, Amsterdam, 1968).

[18] To avoid possible confusion, the notational change

$$A(G' \mid GK_S SM_S K_T TM_T) \equiv \langle G' \mid GK_S SM_S K_T TM_T \rangle,$$

expressing the transformation brackets as the elements of a matrix A, with rows labeled by the $h'_{\alpha\beta}$ and columns labeled by the $\{K_S SM_S K_T TM_T\}$-values, will be adopted in this section.

[19] J. D. Vergados, Nucl. Phys. **A111**, 681 (1968).

[20] U. Fano and G. Racah, *Irreducible Tensorial Sets* (Academic, New York, 1959).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Theorem of Uniqueness and Local Stability for Liouville–Einstein Equations*

YVONNE CHOQUET-BRUHAT
Département de Mathématiques, Faculté des Sciences, Université de Paris
11 rue Pierre-Curie, 75 Paris 5e, France

(Received 24 March 1970)

We prove, by use of energy inequalities, a theorem of uniqueness and local (i.e., for finite time) stability for the solution of Cauchy problem relative to the integro-differential system of Einstein and Liouville. A global theorem of geometrical uniqueness follows from a general method, previously given. We will prove elsewhere an existence theorem.

INTRODUCTION

The aim of this paper is to prove a uniqueness theorem for the solution of the Cauchy problem for the coupled Liouville–Einstein equations, i.e., for a collisionless relativistic gas under its own gravitational field. Such a gas provides a model reasonably appropriate for physical systems like systems of galaxies or some systems of stars (which are then the "particles" of the gas) and certain plasmas or radiations (in this last case the particles have a zero rest mass).

With the uniqueness theorem we prove a local stability theorem; i.e., we prove that the solution (metric and distribution function) depends continuously on the initial data: such a theorem, which states that a small initial perturbation gives rise to a small perturbation during some finite time, seems the first necessary step to be assured of before any more elaborate research on stability.

The plan of this paper is the following:

In Sec. I, I give a brief review of the fundamental concepts of relativistic kinetic theory, and I recall the equations governing the motion of a self-gravitating collisionless, relativistic gas: the coupled Einstein and Liouville equations. I also recall, or establish, a few general properties of these equations which will be used in the following (i.e., local equivalence of Einstein equations in harmonic coordinates and tensorial Einstein equations, and use of bounded parameters for the momenta in the Liouville equation).

In Sec. II, I establish some inequalities satisfied by the difference of two solutions of the Cauchy problem

relative to the coupled Einstein–Liouville equations by using energy integrals related to these equations.

In Sec. III, I deduce from the preceding inequalities a local uniqueness and stability theorem for the solutions of such a Cauchy problem. The local uniqueness implies global geometric uniqueness.[1]

In the Appendix, I treat the case of continuous masses (star clusters).

I. EQUATIONS AND GENERAL PROPERTIES

We will in this part recall briefly the origin and the form of the Liouville equation for the distribution function f on a space-time of general relativity. In order that the stress–energy tensor corresponding to the particle flow associated with this distribution function be bounded, we will have to make some hypothesis about the decrease at infinity (in momentum space) of this function ("weighted" distribution function). For all metrics verifying, in a local chart, the usual hypothesis of boundedness, differentiability, and hyperbolicity, we will show that it is possible to choose (we treat, for simplicity, the case of particles with a given rest-mass) parameters in momentum space which take their values in a bounded domain of R^3: This choice will simplify the establishment of inequalities in Sec. III.

We finally write Einstein equations and give a form for the stress–energy tensor adapted to the further study.

A. Phase Space and Distribution Function

Within the framework of general relativity the phase space is the tangent bundle $T(M)$ of a 4-dimensional differentiable manifold M, which has a hyperbolic[2] metric g. If x^α are local coordinates in M, we call (x^α, p^α) the local coordinates in $T(M)$, where the p^α are the components of a vector p tangent to M at x in the natural frame associated with the local coordinates x^α.

A particle is a path in phase space $(x(t), p(t))$, where $x(t)$ describes the position of the particle and $p(t)$ its 4-momentum. If m is the proper mass of the particle, the length of p in the metric g is m, and the path of the particle in the phase space lies in the subbundle

$$g_{\alpha\beta}p^\alpha p^\beta = m^2. \tag{1}$$

We will suppose, moreover, that M is time oriented and that p points toward the future. Then p, for given x, is on the future sheet P_x of the hypersurface of Eq. (1). P_x is called the mass-m hyperboloid.

The volume element in M and $T_x(M)$ being, respectively, in the coordinates (x^α) and (p^α),

$$\begin{aligned} &\eta = |g|^{\frac{1}{2}}\, d^4x, \quad \omega = |g|^{\frac{1}{2}}\, d^4p, \\ &d^4x = dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3, \\ &d^4p = dp^0 \wedge dp^1 \wedge dp^2 \wedge dp^3, \end{aligned} \tag{2}$$

the volume element in P_x is given by the Leray form such that

$$\varpi \wedge d[\tfrac{1}{2}(g_{\alpha\beta}p^\alpha p^\beta - m^2)] = \omega,$$

which may be written

$$\varpi = (|g|^{\frac{1}{2}}/p_0)d^3p, \quad d^3p = dp^1 \wedge dp^2 \wedge dp^3. \tag{3}$$

In the absence of external forces (which we suppose[3]), the particle paths are geodesics, i.e., trajectories of the vector field X on $T(M)$:

$$X = [p^\alpha, -\Gamma^\alpha_{\lambda\mu}p^\lambda p^\mu].$$

For particles of mass m, these trajectories lie on $P(M)$.

The distribution function[4] $f(x, p)$ is a scalar function on $T(M)$. Its interpretation is that the 7-form on $T(M)$,

$$\theta = f(x, p)i \quad (\eta \wedge \omega),$$

induces, on each 7-dimensional submanifold Σ of $T(M)$, the volume element for the number of particle paths crossing Σ.

In the case where particles are microscopic (relativistic gas or plasma), the proper masses take on a finite range of discrete values m_j. The distribution function is then the finite sum of distribution functions corresponding to each of these values

$$f = \sum_j f_j,$$

f_j being defined on $P_j(M)$.

In the case where the particles are macroscopic (star or galactic clusters), the proper masses take on continuous values, and the distribution function could be defined on $T^+(M)$ (fiber $g_{\alpha\beta}p^\alpha p^\beta \geq 0$, p pointing towards the future). The astronomical data prove, however, that the masses of the stars of a cluster are bounded from above and from below[5]: $f(x, p)$ is then to be defined only on the submanifold

$$m_1^2 \leq g_{\alpha\beta}p^\alpha p^\beta \leq m_2^2.$$

B. Liouville Equation

If we suppose that particle paths are differentiable geodesics (no collisions), the conservation law of the number of particles imposes the fact that θ is invariant under the trajectories of X, i.e., that

$$i_X\theta = 0, \tag{4}$$

$$i_X\, d\theta = 0. \tag{5}$$

Equation (4) is, for any f, a consequence of the identity

$$(i_X)^2 = 0.$$

(5) gives a differential equation for f, called the Liouville equation:

$$i_X\, df = (df)(X) = p^\alpha \frac{\partial f}{\partial x^\alpha} - \Gamma^\alpha_{\lambda\mu} p^\lambda p^\mu \frac{\partial f}{\partial p^\alpha} = 0. \quad (6)$$

If we denote by U a vector field on M, strictly timelike for the metric g [i.e., $U \cdot p \equiv U_\alpha p^\alpha \geq k > 0$, $\forall\, p \in T^+_x(M)$] and set (φ is the weighted distribution function)

$$\varphi(x, p) = (U \cdot p)^N f(x, p), \quad (7)$$

we deduce from (6) that φ satisfies the differential equation

$$p^\alpha \frac{\partial \varphi}{\partial x^\alpha} - \Gamma^\alpha_{\lambda\mu} p^\lambda p^\mu \frac{\partial \varphi}{\partial p^\alpha} - N\varphi(p_\alpha U^\alpha)^{-1} p^\lambda p^\mu \nabla_\lambda U_\mu = 0; \quad (8)$$

let us remark that

$$p^\lambda p^\mu \nabla_\lambda U_\mu = \tfrac{1}{2} p^\lambda p^\mu (\nabla_\lambda U_\mu + \nabla_\mu U_\lambda)$$

vanishes for all $p \in T_x(M)$ if and only if U_μ is a Killing vector field of the metric g.

On the other hand, the inequality

$$p^\lambda p^\mu \nabla_\lambda U_\mu = i_X d(U \cdot p) \leq 0$$

means that $U \cdot p$ is nonincreasing on particle paths. The existence of a timelike vector field verifying this assumption, for a given metric, has been proved by Bitcheler[6] and will be called the Bitcheler lemma. We will not use this lemma here.

Remark: The vector field X is tangent to $P(M)$; thus Eqs. (6) or (8) may be restricted to partial differential equations on $P(M)$.

C. Bounded Parameters on P_x

In a coordinate chart, it is physically reasonable to suppose that the metric of the space-time is differentiable enough for the geodesics to be well defined, and uniformly hyperbolic, and to choose adapted coordinates; more precisely, we state the following:

Hypothesis H_1, Definition: A metric, given in a local chart by ten functions $g_{\alpha\beta}$ defined on an open set U of R^4, satisfies hypothesis H_1, if on U:

(a) There exist positive constants a, b, and M such that

$$-g_{ij}X^iX^j \geq b^2 \sum (X^i)^2, \quad b > 0,$$
$$g_{00} \geq a^2 > 0, \quad g^{00} \geq a^2 > 0, \quad |g^{\alpha\beta}| \leq M$$

(uniform hyperbolicity of g, with $x^0 = c^{te}$ uniformly spacelike, and boundedness of g);

(b) $g_{\alpha\beta}$ is C^2 and

$$|\partial_\lambda g^{\alpha\beta}| \leq M', \quad |\partial^2_{\lambda\mu} g^{\alpha\beta}| \leq M'.$$

Remark: The metrics satisfying H_1 form a bounded, closed convex subset of the space $(\times C^2_{(U)})$[10].

If[7] $m \neq 0$, we may take, as parameters on the m hyperboloid P_x,

$$v^i = p^i/p^0. \quad (9)$$

We denote by $\mathcal{M}_x$ the image in R^3 of P_x by the mapping $p \to (v^i)$. For all $x \in M$ and all metrics g satisfying H_1 on U, the domains $\mathcal{M}_x$ are contained in a fixed bounded domain $\bar{\mathcal{M}} \subset R^3$: Eq. (1) may indeed be written

$$g_{00}(p^0)^2 + g_{ij}(p^i + \bar{g}^{0i}p^0)(p^j + \bar{g}^{0j}p^0) - g_{ij}\bar{g}^{0i}\bar{g}^{0j}(p^0)^2 = m^2, \quad (10)$$

where $\bar{g}^{0i} = \bar{g}^{ih}g_{0h}$, with $\bar{g}^{ih}$ being elements of the matrix inverse of g_{ih}. We deduce from (10) and hypothesis H_1

$$p^0 \geq \frac{m}{(g_{00} - g_{ij}\bar{g}^{0i}\bar{g}^{0j})^{\frac{1}{2}}} = m\,|g^{00}|^{\frac{1}{2}} \geq ma \quad (11)$$

and

$$\begin{aligned} b^2 \sum_{i=1}^3 (v^i)^2 &\leq 2b^2 \sum_{i=1}^3 (v^i + \bar{g}^{0i})^2 + 2b^2 \sum_{i=1}^3 (\bar{g}^{0i})^2 \\ &\leq -2g_{ij} \sum_{i=1}^3 (v^i + \bar{g}^{0i})(v^j + \bar{g}^{0j}) + 2b^2 \sum_{i=1}^3 (\bar{g}^{0i})^2 \\ &\leq 2g_{00} - g_{ij}\bar{g}^{0j}\bar{g}^{0i} + 2b^2 \sum_{i=1}^3 (\bar{g}^{0i})^2 \\ &= \frac{2}{g^{00}} + 2b^2 \sum_{i=1}^3 (\bar{g}^{0i})^2. \end{aligned} \quad (12)$$

Hence

$$\sum_{i=1}^3 (v^i)^2 \leq K(a, b, M). \quad (13)$$

A straightforward computation gives the following for the Leray form (3) in the parameters v^i on the fiber P_x:

$$\varpi = |g|^{\frac{1}{2}}[(p^0)^4/m^2]d^3v. \quad (14)$$

If we define the functions $\nu^{\alpha\beta}(x, v)$ on $P(M)$ by

$$\nu^{\alpha\beta}(x, v) = p^\alpha p^\beta (p^0)^4/(U_\alpha p^\alpha)^N, \quad (15)$$

with U_α strictly timelike for all metrics satisfying H_1, chosen such that

$$U_0 + U_i v^i \geq k > 0, \quad \forall\, \{v^i\} \in \bar{\mathcal{M}}, \quad (16)$$

then $\nu^{\alpha\beta}(x, v)$ is bounded on $\bar{\mathcal{M}}$ if

$$N \geq 6.$$

If $\varphi(x, v)$ is the expression of φ on P_x in the parameters v^i, Eq. (8) reads

$$\frac{\partial \varphi}{\partial x^0} + v^i \frac{\partial \varphi}{\partial x^i} - A^i \frac{\partial \varphi}{\partial v^i} - NF\varphi = 0, \tag{17}$$

with

$$A^i = \Gamma^i_{00} + 2\Gamma^i_{j0}v^j + \Gamma^i_{jh}v^j v^h - v^i\Gamma^0_{00} - 2v^i\Gamma_{j0}v^j - v^i\Gamma^0_{jh}v^j v^h, \tag{18}$$

$$F = (U_0 + U_i v^i)^{-1} \times [\nabla_0 U_0 + v^i(\nabla_0 U_i + \nabla_i U_0) + v^i v^j \nabla_i U_j]. \tag{19}$$

Under the hypothesis H_1, the functions A^i and F are C^1 on $U \times \bar{\mathcal{M}}$, with bounded derivatives of order ≤ 1 or, if U_α has been chosen[8] C^2, with bounded derivatives of order ≤ 2.

Hypothesis H_2, Definition: We will say that a function $\varphi(x, v)$ on $U \times \bar{\mathcal{M}}$ satisfies the hypothesis H_2 if, on $U \times \bar{\mathcal{M}}$, φ is C^1 and there exists a constant M'' such that

$$|\varphi| < M'', \quad |\partial_\lambda \varphi| \leq M'', \quad \left|\frac{\partial \varphi}{\partial v^i}\right| \leq M''.$$

D. Einstein Equations

We will now suppose that the particles are the sources of the gravitational field, i.e., that Einstein equations are satisfied:

$$S^{\alpha\beta} = R^{\alpha\beta} - \tfrac{1}{2}g^{\alpha\beta}R = T^{\alpha\beta}, \tag{20}$$

with $T^{\alpha\beta}$, the stress–energy tensor due to the particles, given at each point $x \in M$ by

$$T^{\alpha\beta} = \int_{\Pi_x(M)} f(x, p)p^\alpha p^\beta \varpi. \tag{21}$$

$\Pi_x(M)$ denotes the fiber appropriate to the problem at hand [i.e., $P_x(M)$, $\sum_j P_{jx}(M)$, $T^+_x(M)$, $\cdots$].

To study the solutions of (20) from the point of view of analysis, it will be convenient to use local coordinates which are harmonic (i.e., $\partial_\lambda[|g|^{\frac{1}{2}} g^{\lambda\mu}] = 0$). We know that, in such coordinates, the Einstein tensor reads

$${}^{(h)}S^{\alpha\beta} = -\tfrac{1}{2}g^{\lambda\mu}\partial^2_{\lambda\mu}g^{\alpha\beta} + H^{\alpha\beta}(g^{\lambda\mu}, \partial_\gamma g^{\lambda\mu}), \quad \partial_\lambda = \frac{\partial}{\partial x^\lambda}, \tag{22}$$

where $H^{\alpha\beta}$ is a rational function (with denominator a power of $|g|$) of $g^{\lambda\mu}$ and $\partial_\gamma g^{\lambda\mu}$.

On the other hand, it is known that, if f satisfies the Liouville equation, the stress–energy tensor (21) satisfies the conservation laws

$$\nabla_\alpha T^{\alpha\beta} = 0. \tag{23}$$

By standard arguments, it then follows that any metric g solution of

$${}^{(h)}S^{\alpha\beta} = T^{\alpha\beta}$$

verifying the constraints on an initial spacelike hypersurface will verify the tensorial Einstein equations (20).

If we suppose, for simplicity, that all particles have the same mass[9] $m \neq 0$, the tensor $T^{\alpha\beta}$ reads

$$T^{\alpha\beta} = m^{-2}\int_{\mathcal{M}_x} \varphi v^{\alpha\beta} |g|^{\frac{1}{2}} d^3v,$$

where $\mathcal{M}_x$ is a bounded domain and $v^{\alpha\beta}$ [cf. Eq. (15)] a bounded function on $\mathcal{M}_x$.

II. ENERGY INEQUALITIES

We define, in a coordinate system, norms for the difference of two metrics ${}^1g$ and ${}^2g$ and of the corresponding weighted distribution functions ${}^1\varphi$ and ${}^2\varphi$, at a given "time" $x^0 = t$, and then establish inequalities between the sum of norms and the sum of the corresponding norms at an initial time ($x^0 = 0$). We set ${}^1g^{\lambda\mu} - {}^2g^{\lambda\mu} = \gamma^{\lambda\mu}$, ${}^1\varphi - {}^2\varphi = \psi$, where ${}^1g$ and ${}^2g$ are two metrics satisfying the hypothesis H_1, with ${}^1\varphi$ and ${}^2\varphi$ satisfying H_2 and the couples $({}^1g, {}^1\varphi)$ and $({}^2g, {}^2\varphi)$ satisfying both of the coupled Einstein–Liouville equations.

By subtracting equations satisfied respectively by $({}^1g, {}^1\varphi)$ and $({}^2g, {}^2\varphi)$, we get

$$-\tfrac{1}{2}\,{}^2g^{\alpha\beta}\partial^2_{\alpha\beta}\gamma^{\lambda\mu} - \tfrac{1}{2}\gamma^{\alpha\beta}\partial^2_{\alpha\beta}\,{}^1g^{\lambda\mu} + {}^2H^{\lambda\mu} - {}^1H^{\lambda\mu} = v^{\lambda\mu}, \tag{24}$$

$$\partial_0\psi + v^i\partial_i\psi - {}^2A^i\frac{\partial \psi}{\partial v^i} - N\,{}^2F\psi = V, \tag{25}$$

where

$$V = ({}^2A^i - {}^1A^i)\frac{\partial\,{}^1\varphi}{\partial v^i} + N({}^2F - {}^1F)\,{}^1\varphi \tag{26}$$

and

$$v^{\lambda\mu} = \int_{{}^2\mathcal{M}_x} {}^2\varphi\gamma^{\alpha\beta}\,|{}^2g|^{\frac{1}{2}}\,d^3v - \int_{{}^1\mathcal{M}_x} {}^1\varphi v^{\alpha\beta}\,|{}^1g|^{\frac{1}{2}}\,d^3v, \tag{27}$$

with ${}^2\mathcal{M}_x$ (resp. ${}^1\mathcal{M}_x$) denoting the range of the variables $\{v^i\}$ corresponding to the fiber P_x for the metric ${}^2g$ (resp. ${}^1g$). Equation (27) may be written

$$v^{\lambda\mu} = \int_{{}^2\mathcal{M}_x \cap {}^1\mathcal{M}_x} [({}^2\varphi - {}^1\varphi)\,|{}^2g|^{\frac{1}{2}} + (|{}^2g|^{\frac{1}{2}} - |{}^1g|^{\frac{1}{2}})\,{}^1\varphi]v^{\alpha\beta}d^3v + \int_{{}^2\mathcal{M}_x - ({}^2\mathcal{M}_x \cap {}^1\mathcal{M}_x)} {}^2\varphi v^{\alpha\beta}\,|{}^2g|^{\frac{1}{2}}\,d^3v - \int_{{}^1\mathcal{M}_x - ({}^2\mathcal{M}_x \cap {}^1\mathcal{M}_x)} {}^1\varphi v^{\alpha\beta}\,|{}^1g|^{\frac{1}{2}}\,d^3v.$$

Our uniqueness will then follow from energy inequalities applied to (24) and (25). We will also deduce from these inequalities a stability theorem.

A. Norms and Energy Inequalities for Differences in Metrics

We denote by X the part in the future ($x^0 \geq 0$) of a subset of V, globally hyperbolic for all metrics satisfying H_1, with Cauchy surface ω_0, a relatively compact subset of S_0 (i.e., all timelike or null curves issuing from $x \in X$, towards the past, cut ω_0 once and only once; the existence of such domains X is easy to prove); we denote by ω_t the hypersurface $x^0 = t$ of X.

We set[10]

$$\|\gamma^{\lambda\mu}\|_t^1 = \int_{\omega_t} [|D^1\gamma^{\lambda\mu}|^2 + |\gamma^{\lambda\mu}|^2] d^3x, \tag{28}$$

where

$$|D^1\gamma^{\lambda\mu}|^2 = \sum_\alpha |\partial_\alpha \gamma^{\lambda\mu}|^2 \tag{29}$$

and

$$\|\gamma\|_t^1 = \sum_{\lambda,\mu} \|\gamma^{\lambda\mu}\|_t^1, \tag{30}$$

and analogously

$$\|v\|_t = \int_{\omega_t} \sum_{\lambda,\mu} |v^{\lambda\mu}|^2 \, d^3x.$$

We obtain bounds for (28), as usual,[11] for second-order partial differential equations: We multiply by $\partial_0\gamma^{\alpha\beta}$ and integrate over $\tilde{X}$ (intersection of X with the past of ω_t for the metric ${}^2g$), using Stoke's formula (the "lateral" boundary of X is characteristic and gives rise to a nonnegative integral) and using, moreover, the fact that, for any differentiable function u on X,

$$|u(x^\alpha)|^2 \leq 2x^0 \int_0^{x^0} |\partial_0 u(x^i, \tau)|^2 \, d\tau + 2\,|u(x^i, 0)|;$$

we thus find that there exist constants C_1, C_2, and C_3 depending only on bounds H_1 such that

$$\|\gamma\|_t^1 \leq C_1 \|\gamma\|_0^1 + (C_2 + 2t)\int_0^t \|\gamma\|_\tau^1 d\tau + C_3 \int_0^t \|v\|_\tau | \, d\tau.$$

B. Norm and Energy Inequality for Difference in Distribution Function

Let us denote by Ω the subbundle of $P(V)$ with basis $\tilde{X}$:

$$(x, p) \in \tilde{\Omega} \rightleftarrows x \in \tilde{X}, \quad v = \{v^i = p^i/p^0\} \in {}^2\mathcal{M}_x,$$

a point (x, p) in the boundary $\partial\tilde{\Omega}$ of $\tilde{\Omega}$ is either

$$x \in \partial\tilde{X} \quad \text{or} \quad p \in \partial P_x, \quad \text{i.e.,} \quad v \in \partial\, {}^2\mathcal{M}_x;$$

we see that $\partial\tilde{\Omega}$ is the union of a "lateral" boundary $\tilde{L}$, generated by null geodesics of ${}^2g$, and of subsets where x belongs to ω_t or to ω_0.

If we multiply (25) by ψ and make some obvious transformations, we get

$$\partial_0 |\psi|^2 + \partial_i(v^i |\psi|^2) - \frac{\partial}{\partial v^i}({}^2A^i |\psi|^2) = \left(N\, {}^2F - \frac{\partial}{\partial v^i}\, {}^2A^i\right) |\psi|^2 + 2V\psi. \tag{31}$$

If we integrate (31) over $\tilde{\Omega}$ and apply Stoke's formula on the one hand, the integral over $\tilde{L}$ vanishes, and we get (C_4 and C_5 depend only on bounds H_1)

$$|||\psi|||_t \leq |||\psi|||_0 + C_4 \int_0^t |||\psi|||_\tau \, d\tau + C_5 \int_0^t |||V|||_\tau \, d\tau,$$

where we have set (recall that $\tilde{\omega}_\tau \subset \omega_\tau$)

$$|||\psi|||_\tau = \int_{x\in\omega_\tau} \int_{v\in {}^2\mathcal{M}_x} |\psi(x, v)|^2 \, d^3v d^3x$$

and analogous definition for $|||V|||_\tau$.

C. Bounds for $\|V\|$ and $|||v|||$

We deduce from (26) and bounds H_1 and H_2 that there exists a constant K, depending only on these bounds, such that

$$|||V|||_t \leq K \|\gamma\|_t^1. \tag{32}$$

We consider now (27). We remark that ${}^2\mathcal{M}_x$ (resp. ${}^1\mathcal{M}_x$) is defined by

$${}^2\phi(v^i) \leq {}^2g_{00}^{-1} \quad [\text{resp. } {}^1\phi(v^i) \leq {}^1g_{00}^{-1}],$$

where [cf.(9)]

$$\phi(v^i) \overset{\text{DEF}}{=} -g_{ij}(v^i + g^{-0i})(v^j + g^{-0j});$$

it results from the definition of $\gamma^{\lambda\mu}$ and hypothesis H_1 that the points v^i of ${}^2\mathcal{M}_x$ which are not in ${}^1\mathcal{M}_x$ satisfy the following inequalities ($C_{\lambda\mu}$ constants depending on bounds H_1),

$$({}^2g^{00})^{-1} - C_{\lambda\mu}|\gamma^{\lambda\mu}| \leq \phi^2(v^i) \leq ({}^2g^{00})^{-1}, \quad C_{\lambda\mu} \geq 0, \tag{33}$$

from which there results the existence of constants $D_{\lambda\mu} \geq 0$ such that the measure of the set ${}^2\mathcal{M}_x - ({}^1\mathcal{M}_x \cap {}^2\mathcal{M}_x)$ is bounded by $D_{\lambda\mu}|\gamma^{\lambda\mu}|$.

It is then easy to obtain the majoration, where the constants K_2 and K_3 depend only on the bounds H_1 and H_2:

$$\|v\|_t \leq K_2 |||\psi|||_t + K_3 \|\gamma\|_t^1. \tag{34}$$

We then deduce from inequalities (32) and (34) that, if

$$t \leq T \quad \text{on } X,$$

there exists a constant C, depending only on T, H_1,

and H_2, such that

$$\|\gamma\|_t^1 + |||\psi|||_t \leq C_1 \|\gamma\|_0^1 + |||\psi|||_0 + C\int_0^t (\|\gamma\|_\tau^1 + |||\psi|||_\tau)\, d\tau. \quad (35)$$

An easy consequence of this integral inequality is

$$\|\gamma\|_t^1 + |||\psi|||_t \leq (C_1 \|\gamma\|_0^1 + |||\psi|||_0)e^{ct}. \quad (36)$$

III. UNIQUENESS AND STABILITY THEOREMS

We will deduce easily from the inequality (36) the following:

Local uniqueness theorem (discrete masses):

Hypothesis:

(1) ${}^1g$ and ${}^2g$ are two metrics satisfying H_1. On $S\ (x^0 = 0)$:

$${}^1g_{\alpha\beta} = {}^2g_{\alpha\beta}, \quad \partial_\lambda {}^1g_{\alpha\beta} = \partial_\lambda {}^2g_{\alpha\beta}, \quad \text{for} \quad x^0 = 0. \quad (37)$$

(2) ${}^1\varphi$ and ${}^2\varphi$ satisfy H_2 and

$${}^1\varphi = {}^2\varphi, \quad \text{for} \quad x^0 = 0, \quad (38)$$

which implies [under hypothesis (37)]

$${}^1f = {}^2f \quad \text{for} \quad x^0 = 0.$$

Conclusion: In a neighborhood X of S,

$${}^1g = {}^2g,$$

and, on the corresponding bundle $P(X)$,

$${}^1f = {}^2f.$$

Proof: Equations (37) and (38) imply

$$\|\gamma\|_0^1 = 0, \quad |||\psi|||_0 = 0;$$

thus, by (36),

$$\|\gamma\|_t^1 = 0 \quad \text{and} \quad |||\psi|||_t = 0, \quad t \leq T.$$

Under the hypotheses H_1 and H_2, these equations imply

$$\gamma^{\lambda\mu} = 0 \quad \text{on } X, \text{ i.e., } \quad {}^1g^{\lambda\mu} = {}^2g^{\lambda\mu},$$

and $\psi = 0$ on the corresponding bundle $P(X)$; thus ${}^1f = {}^2f$. From this local uniqueness, that we have obtained in harmonic coordinates, one deduces by standard arguments (cf. above) *a geometric (i.e., up to isometry) global uniqueness theorem for the solution of Cauchy problem in the class of smooth globally hyperbolic metrics.*

Stability: The inequality (36) proves that, if the norms at time $x^0 = 0$, $\{\|\gamma\|_0^1\}^{\frac{1}{2}}$ and $\{|||\psi|||_0\}^{\frac{1}{2}}$, are less than some small number ϵ, the corresponding norms at time $x^0 = 0$ will be less than $K\epsilon$, where K is bounded for bounded t. This property proves what could be called a local or *weak stability theorem*[11]: Such a theorem states that a small perturbation in the metric and the distribution function gives rise to a small perturbation, at least during some finite time. The strong stability,[12] in a rigorous sense, would be obtained if K could be proved to be a bounded function of t: In general, K increases infinitely (and even exponentially) with t. It would be interesting to study its order of magnitude, in special cases, after a proper choice of coordinates geometrically (or physically) meaningful.

ACKNOWLEDGMENTS

I thank Professor J. Ehlers for pointing out this open problem to me and Professor C. DeWitt and Professor B. DeWitt for interesting comments and discussions.

APPENDIX: CONTINUOUS MASS (STAR OR GALACTIE CLUSTERS)

It is very easy to extend the local uniqueness and stability theorem to the case of proper masses ranging continuously from $m_1 > 0$ to $m_2 > 0$. The fiber Π_x is then given by the manifold with boundary

$$m_1^2 \leq g_{\alpha\beta}p^\alpha p^\beta \leq m_2^2.$$

We can take on Π_x the parameters $v^i = p^i/p^0$, $t = g_{\alpha\beta}p^\alpha p^\beta$. For a given metric g, the range of $\{v^i\}$ is the bounded domain $\mathcal{M}_x$ of R^3 (cf. Sec. ID), independent of t, and, if we set

$$f(x, p) = \varphi(x, v, t)(U_\alpha p^\alpha)^{-6},$$

we have

$$T^{\alpha\beta} = \int_{t_1}^{t_2}\int_{\mathcal{M}_x} \varphi\nu^{\alpha\beta}\, |g|^{\frac{1}{2}}\, t^{-1}\, d^3v\, dt,$$

where $\nu^{\alpha\beta}$ is the bounded function (15) (with $N = 6$).

The function $\varphi(x, v, t)$ satisfies the same Liouville equation (17) as in Sec. ID (the derivative $\partial\varphi/\partial t$ does not appear, due to the fact that the absolute derivative of g vanishes). It is then straightforward to apply arguments analogous to those of Sec. II.

* This work was done while the author was at the University of North Carolina in Chapel Hill, supported by the National Science Foundation.

[1] J. L. Synge, *The Relativistic Gas* (North-Holland, Amsterdam, 1957).

[2] I.e., pseudo-Riemannian, with signature $(+---)$.

[3] There will be complications in writing, but no essential difficulty if an electromagnetic field is also present.

[4] For a detailed exposition, see J. Ehlers, "General Relativity and Kinetic Theory," lecture notes, Varenna, 1969.

[5] About by $\frac{1}{10}$ and 10 solar masses, respectively.

[6] K. Bitcheler, Commun. Math. Phys. **4**, 352 (1967).

[7] For $m = 0$, a modified version of the following arguments may be used.

[8] Such a choice is always possible: Take, for instance, $U_0 = 1$, $U_i = 0$.

[9] It is easy to extend the following proofs to the case of a finite number of rest masses m_i.

[10] $\|\gamma\|_t^1$ is the square of a norm (Sobolev space);see Y. Leray, *Hyperbolic Differential Equations* (Institute for Advanced Study, Princeton N.J., 1952).

[11] There are physical phenomena which are not weakly stable.

[12] A study of stability in the spherically symmetric case is done by Y. Ipser and K. Thorne, Astrophys. J. **154**, 251 (1968).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Realization of the Lie Group $G(0, 1)$ by the Function of Landau Levels

S. YANAGAWA AND T. MORIYA
Department of Applied Physics, Faculty of Engineering, University of Tokyo, Tokyo, Japan

(Received 16 March 1970)

The irreducible basis of the Lie group $G(0, 1)$ are obtained in connection with the quantum physical problem: a motion of a free electron in a magnetic field. The differential operators are shown to be the infinitesimal operators defined by Brown.

Much progress on the relation between Lie group and special functions has recently been developed. The most familiar example is the connection between the rotation group and spherical harmonics, which gives fundamental knowledge in quantum mechanics. Now we will point out that the representation of the 4-dimensional Lie algebra $G(0, b)$ is realized by the functions of Landau levels which express states for the motion of a free electron in a magnetic field H. The 4-dimensional Lie algebra $G(0, b)$ with basis A_+, A_-, A_3, and E is defined by the following commutation relations:

$$[A_+, A_-] = -bE, \quad [A_3, A_+] = A_+, \quad [A_3, A_-] = -A_-. \tag{1}$$

Now we associate following differential operators to each A (in a polar coordinate system):

$$A_\pm = \pm e^{\pm i\theta}\left(2\beta\rho \pm \frac{\partial}{\partial\rho} + \frac{i}{\rho}\frac{\partial}{\partial\theta}\right), \qquad A_3 = \frac{1}{i}\frac{\partial}{\partial\theta}, \quad E = 1. \tag{2}$$

It is easily verified that they satisfy the commutation relations (1), $b = 2\beta$. The Casimir operator is given by

$$C = A_+A_- - bA_3 = \frac{\partial^2}{\partial\rho^2} + \frac{1}{\rho}\frac{\partial}{\partial\rho} + \frac{1}{\rho^2}\frac{\partial^2}{\partial\theta^2} + i\beta\frac{\partial}{\partial\theta} - \frac{\beta^2}{4}\rho^2 - \beta. \tag{3}$$

If we put $\beta = m\omega_c/\hbar$ and $\omega_c = eH/mc$, it is found that the Casimir operators are simply connected to the Schrödinger equation of a free electron in a magnetic field H directed to the Z direction, apart from the Z component; that is,

$$\left[\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2}{\partial\theta^2} + \frac{im\omega_c}{\hbar}\frac{\partial}{\partial\theta} + \frac{2m}{\hbar^2}(E - \tfrac{1}{8}m\omega_c^2\rho^2)\right]\psi = 0. \tag{4}$$

The eigenfunction $\psi^j_{\pm n}$ of (4), for zero or any positive integer n, is

$$e^{\pm in\theta}e^{-\frac{1}{4}\beta\rho^2}\rho^n L^n_{n+j}(\tfrac{1}{2}\beta\rho^2) \tag{5}$$

and has the eigenvalue $(n + j + \frac{1}{2})\hbar\omega_c$ (upper sign) and $(j + \frac{1}{2})\hbar\omega_c$ (lower sign), respectively. The degenerate eigenfunctions of semi-infinite numbers (upper bounded) with a constant eigenvalue $(l + \frac{1}{2})\hbar\omega_c$ will be obtained by operating the $A_\pm$ operators on any function with the same eigenvalue; for $\psi^0_l, n = l, j = 0$,

$$A_+\psi^0_l = 0, \quad A_-\psi^0_l = 2\psi^1_{l-1}, \cdots,$$

and, in general, we have the following recurrence relations:

$$A_-\psi^j_n = 2(j + 1)\psi^{j+1}_{n-1}, \quad n + j = l,$$

and

$$A_-\psi^l_{-n} = -\frac{\beta}{n + l + 1}\psi^l_{-(n+1)}. \tag{6}$$

Then these functions will form the irreducible basis of $G(0, b)$, isomorph to $G(0, 1)$. It is to be noted that the operators $A_\pm$ are to be called infinitesimal magnetic translational operators in a polar coordinate. These are simply related to the magnetic translation operators $T_x \pm iT_y$ introduced by Brown[1] and are easily derived from them.

[1] E. Brown, Phys. Rev. **133**, A1038 (1964).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Regge Trajectories for the Inverse Square Potential*

ERASMO FERREIRA
Instituto de Física, Pontifícia Universidade Católica, Rio de Janeiro, Brazil
AND
JAVIER SESMA
Departamento de Física Teórica, Universidad de Valencia, Valencia, Spain

(Received 4 September 1969; Revised Manuscript Received 29 April 1970)

The analytic properties of the S function in the complex angular momentum plane for regular potentials with inverse square tails are discussed. Special attention is given to the determination of the poles of S in the limits of low and high energies. Two soluble examples are considered in detail.

1. INTRODUCTION

The purpose of this paper is to study some features of the r^{-2} potential. Our interest is mainly on the consequences of the long-range tail.

As repulsive singular potentials do not present difficulties or features of special interest, we restrict our attention to attractive potentials. In order to study in a practical way the effect of the singularity, we perform a regularization of the form of the potential within a range b about the origin. In Sec. 2 we study some properties of the S function for a given range b and look for general results which remain meaningful as b becomes small.

In Sec. 3 we discuss two particular examples. In one case, we put $V(r) = V(b) = \text{const}$ for $r \leq b$; in the other, we introduce a repulsive hard core of radius b. These two special cases have in common the interesting property that the forms of their Regge trajectories are independent of the range b.

2. REGULAR POTENTIAL WITH INVERSE SQUARE TAIL

Let us consider the Schrödinger equation for a particle of mass m and energy $E = \hbar^2k^2/2m$ in a potential of the form

$$\begin{aligned} V(r) &= v(r), && r < b, \\ &= -(\hbar^2/2m)\mu r^{-2}, && r > b, \end{aligned} \tag{2.1}$$

where $v(r)$ is a regular function in the interval $[0, b]$ and is such that $V(r)$ does not present an infinite discontinuity at $r = b$. We define the variables

$$\lambda = l + \tfrac{1}{2}, \tag{2.2}$$

$$\nu = (\lambda^2 - \mu)^{\frac{1}{2}}, \tag{2.3}$$

and

$$z = kb, \tag{2.4}$$

where l is the angular momentum.

Let us call

$$D(b, k, \lambda) = bL(b, k, \lambda) - \tfrac{1}{2}, \tag{2.5}$$

where $L(b, k, \lambda)$ is the logarithmic derivative at $r = b$ of the l wave reduced radial wavefunction for the inner region. The solution in the outer region is a linear combination of Bessel functions $H_\nu^{(1)}(kr)$ and $H_\nu^{(2)}(kr)$. Following the usual procedure of imposing continuity in the logarithmic derivative of the wavefunction, we obtain for the S function

$$S(b, k, \lambda) = -\exp[i\pi(\lambda - \nu)] \times \frac{D(b, k, \lambda)H_\nu^{(2)}(z) - zH_\nu^{(2)\prime}(z)}{D(b, k, \lambda)H_\nu^{(1)}(z) - zH_\nu^{(1)\prime}(z)}. \tag{2.6}$$

A. General Remarks

The analytical properties of $S(b, k, \lambda)$ have been discussed by Barut and Calogero.[1] We wish to add a few remarks to their work.

The first remark concerns the possibility of S presenting nonessential singularities of the second kind, namely, indeterminacy points of the form 0/0. The simultaneous vanishing of numerator and denominator in Eq. (2.6) would imply in the vanishing of the Wronskian of $H_\nu^{(1)}(z)$ and $H_\nu^{(2)}(z)$. Since this Wronskian is given by

$$W[H_\nu^{(1)}(z), H_\nu^{(2)}(z)] = -4i(\pi z)^{-1}, \tag{2.7}$$

it does not vanish for any given z. We conclude that such indeterminacy points are never present in the S function for a potential with an inverse square tail.

The second remark concerns the analytic continuation of S to the half-plane $\text{Re}\,\lambda < 0$. This continuation depends on the form of the function $D(b, k, \lambda)$ and, consequently, on the form of the potential $v(r)$.[2] By inspection of Eq. (2.6) and by making use of properties of the Hankel functions, we obtain that the Mandelstam reflection property[3]

$$S(b, k, -\lambda) = S(b, k, \lambda), \quad \lambda = \text{integer}, \tag{2.8}$$

is satisfied if

$$D(b, k, -\lambda) = D(b, k, \lambda), \quad \lambda = \text{integer}.$$

We thus see that this result, obtained previously[1] for potentials of finite range, is also valid in the presence of an $1/r^2$ tail. In what follows we consider only the half-plane Re $\lambda > 0$.

We now proceed to study the location of the poles of S in the complex angular momentum plane. We first study the general problem of a regular potential of the form given in Eq. (2.1) and discuss the pole distribution in the limits of low and high energies. The pole locations are determined by the solutions of

$$D(b, k, \lambda)H_\nu^{(1)}(z) - zH_\nu^{(1)\prime}(z) = 0. \tag{2.9}$$

B. Low Energies

It is known[4] that $H_\nu^{(1)}(z)$ and $H_\nu^{(1)\prime}(z)$, considered as functions of ν and z, both vanish at $\nu = 0$, $z = 0$. Thus, for zero energy the zeros of the denominator in Eq. (2.6) are located at

$$\lambda = \lambda_0 \equiv \mu^{\frac{1}{2}}, \tag{2.10}$$

a result which does not depend on the form of the potential $v(r)$ in the inner region. At the threshold, the positions of the poles are entirely determined by the inverse square tail, and all poles are at the same point of the complex angular momentum plane.

Besides the poles located at $\lambda = \lambda_0$, it is possible to have poles at different values of λ. These values are the possible solutions of the equation for λ,

$$D(b, 0, \lambda) + \nu = 0, \tag{2.10'}$$

which is obtained from Eq. (2.9) dividing by $H_\nu^{(1)}(z)$ and taking the limit $k \to 0$. Of course, the existence of solutions of Eq. (2.10') depends both on the form of the potential in the inner region $v(r)$, through the value of $D(b, 0, \lambda)$, and on the intensity of the inverse tail, through the value of ν. Special attention should be paid to Eq. (2.10') for a given $v(r)$. However, since our interest is on the effects of a r^{-2} tail and not on the details of a specific potential in the inner region $v(r)$, we shall consider that $v(r)$ has been chosen in such a manner that Eq. (2.10') has no solutions. Of course, this restricts in some way the potential $v(r)$, but it is not a very severe restriction. It is not difficult to find forms of $v(r)$ regular at the origin and for which Eq. (2.10') has no solutions. Although not explicitly specified, we shall assume through all this section that $v(r)$ satisfies those conditions. This assumption must be kept in mind below, particularly when we draw conclusions about bound states and resonances.

To see what happens for small values of $|z|$, we first expand the Hankel functions in Eq. (2.9) in power series, retaining only the first few terms. We obtain

$$(\tfrac{1}{2}z)^{2\nu}\{1 - \nu[D(b, k, \lambda)]^{-1} + O(z^2)\} = e^{i\pi\nu}\{1 + \nu[D(b, k, \lambda)]^{-1} + O(z^2)\} \times \Gamma(1 + \nu)/\Gamma(1 - \nu). \tag{2.11}$$

We follow the procedure of Keller *et al.*[4] and take the logarithm of the two members of the above equation. Noting that

$$\log\left[\frac{\Gamma(1 + \nu)}{\Gamma(1 - \nu)}\right] = -2\gamma\nu - 2\sum_{m=1}^{\infty}\frac{\zeta(2m + 1)\nu^{2m+1}}{2m + 1}, \tag{2.12}$$

where γ is the Euler constant and ζ indicates the Riemann zeta function, we obtain, for the equation determining the poles,

$$\log(\tfrac{1}{2}z) = -in\pi/\nu + \tfrac{1}{2}i\pi - \gamma + [D(b, k, \lambda)]^{-1} + O(\nu^2) + O(z^2), \tag{2.13}$$

where $n = 1, 2, 3, \cdots$. The above Eqs. (2.11) and (2.13) are meaningless when $v(r)$ is such that $D(b, k, \lambda) = 0$ for $z = 0$. If this be the case, we have to discuss the solutions of Eq. (2.9), taking into account the behavior of $D(b, k, \lambda)$ for small $|z|$. If $D(b, k, \lambda) \to 0$ as any power of z when $z \to 0$, we can easily verify that the location of the poles is determined by an equation similar to Eq. (2.13), with the only change that the term $[D(b, k, \lambda)]^{-1}$ does not appear and n is replaced by $n - \frac{1}{2}$.

Let us call

$$D_0 = D(b, 0, \lambda_0) \tag{2.14}$$

the value of $D(b, k, \lambda)$ at $z = 0$, $\nu = 0$, which we assume to be different from zero. We expand:

$$D(b, k, \lambda) = D_0 + (\lambda - \lambda_0)\left[\frac{\partial D}{\partial \lambda}\right]_{\lambda_0} + z\left[\frac{\partial D}{\partial z}\right]_0 + \cdots. \tag{2.15}$$

The derivatives of D are well behaved for small ν and z, so that Eq. (2.13) gives

$$\log(\tfrac{1}{2}z) = -in\pi/\nu + \tfrac{1}{2}i\pi - \gamma + D_0^{-1} + O(\nu^2) + O(z). \tag{2.16}$$

Writing

$$z = \rho e^{i\varphi} \tag{2.17}$$

and

$$\delta = -[\log(\tfrac{1}{2}\rho)]^{-1}, \tag{2.18}$$

we obtain

$$\nu = in\pi\delta\{1 - [i(\tfrac{1}{2}\pi - \varphi) - \gamma + D_0^{-1}]\delta + [i(\tfrac{1}{2}\pi - \varphi) - \gamma + D_0^{-1}]^2\delta^2 + O(\delta^3)\}. \tag{2.19}$$

The above equation gives explicitly the position of the poles of $S(b, k, \lambda)$ in the complex angular momentum plane, for a given b and small complex

values of k. The dependence on the shape of the potential $v(r)$ in the inner region appears only through the value of D_0.

Now we draw some conclusions from the previous results. Equation (2.10) says that at zero energy all poles are on the positive real axis for $\mu > 0$ and on the imaginary axis for $\mu < 0$. It is well known[2] and can be seen directly from Eq. (2.19) that, for negative energies, poles in the half-plane $\mathrm{Re}\,\lambda > 0$ can be only on the real axis. For $\mu > 0$ (attractive tail), as the energy approaches zero from negative values, the poles move along the segment of the real axis between the origin and the point $\lambda = \mu^{\frac{1}{2}}$. For $\mu > \frac{1}{4}$ there is an infinite number of bound states for every angular momentum satisfying $l = 0, 1, 2, \cdots, \leq (\mu^{\frac{1}{2}} - \frac{1}{2})$. There are no bound states for values of angular momentum larger than $\mu^{\frac{1}{2}} - \frac{1}{2}$. These results do not depend on the form $v(r)$ of the potential in the inner region.

As D_0 is real, Eq. (2.19) tells us that, for

$$[\log(\tfrac{1}{2}\rho)]^{-1} \ll 1, \tag{2.20}$$

the Regge trajectories are independent of the form of $v(r)$. The above condition is satisfied for small absolute values of the energy or for small values of the range b of the regularizing potential $v(r)$.

For $\mu > 0$, Eq. (2.19) says that, for pure imaginary values of z (negative energies), ν is pure imaginary. As z increases from zero toward real values, the poles leave the real axis in the λ plane along different trajectories. To obtain the trajectories of the poles for small real z, we first write Eq. (2.19) in the form

$$\lambda = \mu^{\frac{1}{2}}\Big\{1 - \frac{n^2\pi^2}{2\mu}\delta^2\Big[1 - 2[i(\tfrac{1}{2}\pi - \varphi) - \gamma + D_0^{-1}]\delta + \Big(3[i(\tfrac{1}{2}\pi - \varphi) - \gamma + D_0^{-1}]^2 + \frac{n^2\pi^2}{4\mu}\Big)\delta^2 + O(\delta^3)\Big]\Big\}, \tag{2.21}$$

which is obtained by substituting ν given by Eq. (2.19) into

$$\lambda \approx \mu^{\frac{1}{2}}(1 + \nu^2/2\mu - \nu^4/8\mu^2), \quad |\nu| \ll \mu. \tag{2.22}$$

We then obtain

$$\mathrm{Im}\,\lambda = (2^{\frac{1}{2}}\mu^{\frac{1}{4}}/n)(\lambda_0 - \mathrm{Re}\,\lambda)^{\frac{3}{2}}. \tag{2.23}$$

Equation (2.23) shows that, as the energy increases from the value zero, the poles leave the point λ_0 moving to the left along curves tangent to the real axis (see Fig. 1). This form of Regge trajectories, which occurs whenever an attractive inverse square tail is present, is peculiar and essentially different from what is observed for short-range potentials.[5]

The curve corresponding to $n = 1$ has the highest value of Im λ for a given Re λ. The poles labeled with

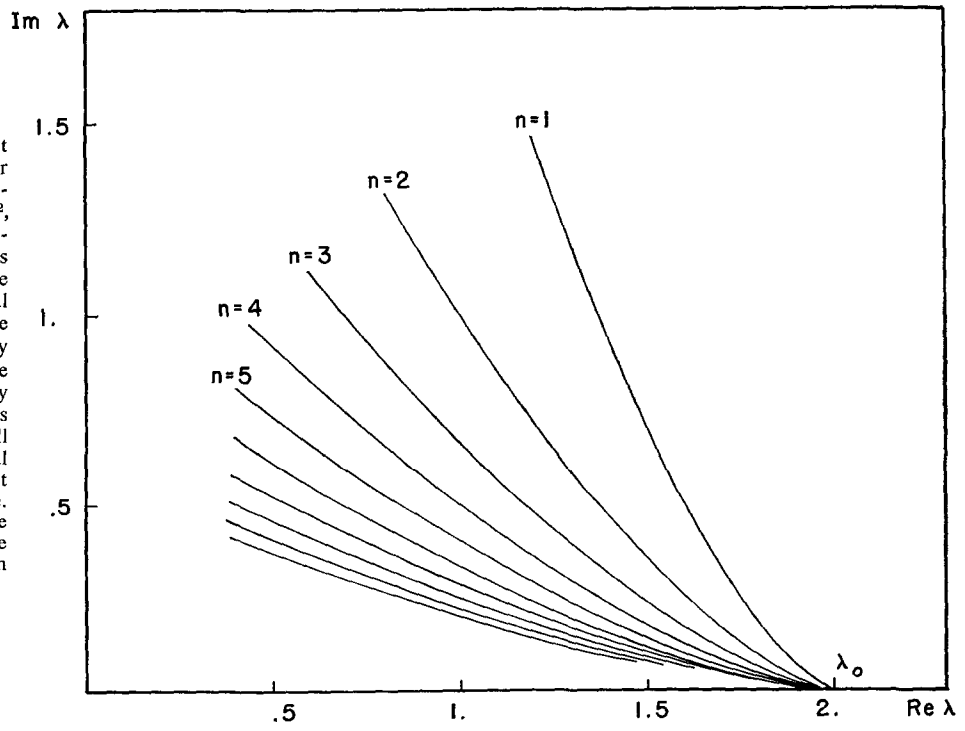


FIG. 1. Regge trajectories at small energies for a regular potential with an attractive inverse square tail $-(\mu\hbar^2/2m)/r^2$, for $\mu = 4$. As the energy increases from negative values towards zero, the poles move along the segment of the real axis between the origin and the point $\lambda_0 = \mu^{\frac{1}{2}}$. At zero energy an infinite number of poles are located at λ_0. As the energy increases from zero, the poles move along separate curves, all of which are tangent to the real axis at λ_0. The few highest curves are shown in the figure. An infinite set of curves are located between these and the segment of the real axis between λ_0 and the origin.

higher values of n move along trajectories located between this first curve and the real axis. The segment of the real axis between λ_0 and the origin is an accumulation segment for the set of trajectories.

Resonances at low energies do not occur for potentials with inverse square tail. Regge trajectories for positive energies can pass near points corresponding to half-integer values of λ, but this does not apply in the occurrence of resonances because the poles move toward the left as the energy increases.

It should be mentioned that the observed fact of infinitely many trajectories approaching the point $\lambda = \lambda_0$ for a r^{-2} tail is the exact analog of the well-known fact that, without such a tail, infinitely many trajectories approach $\lambda = 0$ as $k \to 0$.

C. High Energies

At large positive energies the poles tend to infinity in the λ plane. We introduce the Taylor expansion

$$H_\nu^{(1)}(z) = H_\lambda^{(1)}(z) + (\nu - \lambda)\frac{\partial}{\partial\lambda} H_\lambda^{(1)}(z) + \tfrac{1}{2}(\nu - \lambda)^2 \frac{\partial^2}{\partial\lambda^2} H_\lambda^{(1)}(z) + \cdots \quad (2.24)$$

and use the relation

$$\frac{\partial}{\partial\lambda} H_\lambda(z) = -\frac{\partial}{\partial z} H_\lambda(z),$$

which is valid for large λ and z, to obtain

$$H_\nu^{(1)}(z) = H_\lambda^{(1)}(z) + (\lambda - \nu)H_\lambda^{(1)\prime}(z) + \cdots. \quad (2.25)$$

Taking this expansion into the equation determining the poles, we find

$$D(b, k, \lambda)H_\nu^{(1)}(z) - zH_\lambda^{(1)\prime}(z) + (\lambda - \nu) \times [D(b, k, \lambda)H_\lambda^{(1)\prime}(z) + zH_\lambda^{(1)\prime\prime}(z)] + \cdots = 0. \quad (2.26)$$

Since

$$\lambda - \nu = O(\lambda^{-1}),$$

the leading terms in the above equation give

$$D(b, k, \lambda)H_\lambda^{(1)}(z) - zH_\lambda^{(1)\prime}(z) = 0,$$

which is the equation for the poles in a finite range potential of the form $v(r)$, without tail.

3. TWO EXAMPLES

We now choose particular convenient forms for $v(r)$. We treat two cases which present the interesting property that the form of their Regge trajectories does not depend on the range b.

A. Constant Potential Near the Origin

We first discuss the continuous potential

$$V(r) = -(\hbar^2/2m)\mu/b^2, \quad r < b,$$
$$= -(\hbar^2/2m)\mu/r^2, \quad r > b. \quad (3.1)$$

In this case

$$D(b, k, \lambda) = \xi J_\lambda'(\xi)/J_\lambda(\xi), \quad (3.2)$$

where

$$\xi = (z^2 + \mu)^{\frac{1}{2}}.$$

The S function

$$S(b, k, \lambda) = S(z, \lambda) = -\exp[i\pi(\lambda - \nu)] \times \frac{\xi J_\lambda'(\xi)H_\nu^{(2)}(z) - zJ_\lambda(\xi)H_\nu^{(2)\prime}(z)}{\xi J_\lambda'(\xi)H_\nu^{(1)}(z) - zJ_\lambda(\xi)H_\nu^{(1)\prime}(z)} \quad (3.3)$$

depends on k and b through the product $kb = z$. Owing to this fact, the trajectories, described by the poles as the energy varies, are the same for all values of b.

Using the well-known formula for the Wronskian of Bessel functions

$$W(J_\lambda(\xi), J_{-\lambda}(\xi)) = -2\sin(\lambda\pi)/\pi\xi,$$

we derive from Eq. (3.2)

$$D(b, k, -\lambda) = D(b, k, \lambda), \quad \lambda = \text{integer},$$

and so the Mandelstam reflection property (2.8) is satisfied.

In Sec. 2 we mentioned the possibility of $D(b, k, \lambda)$ becoming zero for $z = 0$, $\lambda = \lambda_0$. In the present example this possibility is excluded, as follows from the fact that $\lambda J_\lambda'(\lambda)/J_\lambda(\lambda)$ is a monotonously increasing function of λ for real positive λ.[6]

B. Hard Core

The general treatment given in Sec. 2 to potentials with an inverse square tail was limited to cases in which $V(r)$ is regular everywhere. We now consider an example which is not included in these cases: Namely, $v(r)$ is taken as a hard core of radius b.

The S function

$$S(b, k, \lambda) = S(z, \lambda) = -\exp[i\pi(\lambda - \nu)]H_\nu^{(2)}(z)/H_\nu^{(1)}(z) \quad (3.4)$$

admits continuation into the half-plane $\operatorname{Re}\lambda < 0$, through the equation

$$S(b, k, -\lambda) = \exp(-i2\pi\lambda)S(b, k, \lambda). \quad (3.5)$$

This last relation shows that Eq. (2.8) is valid.

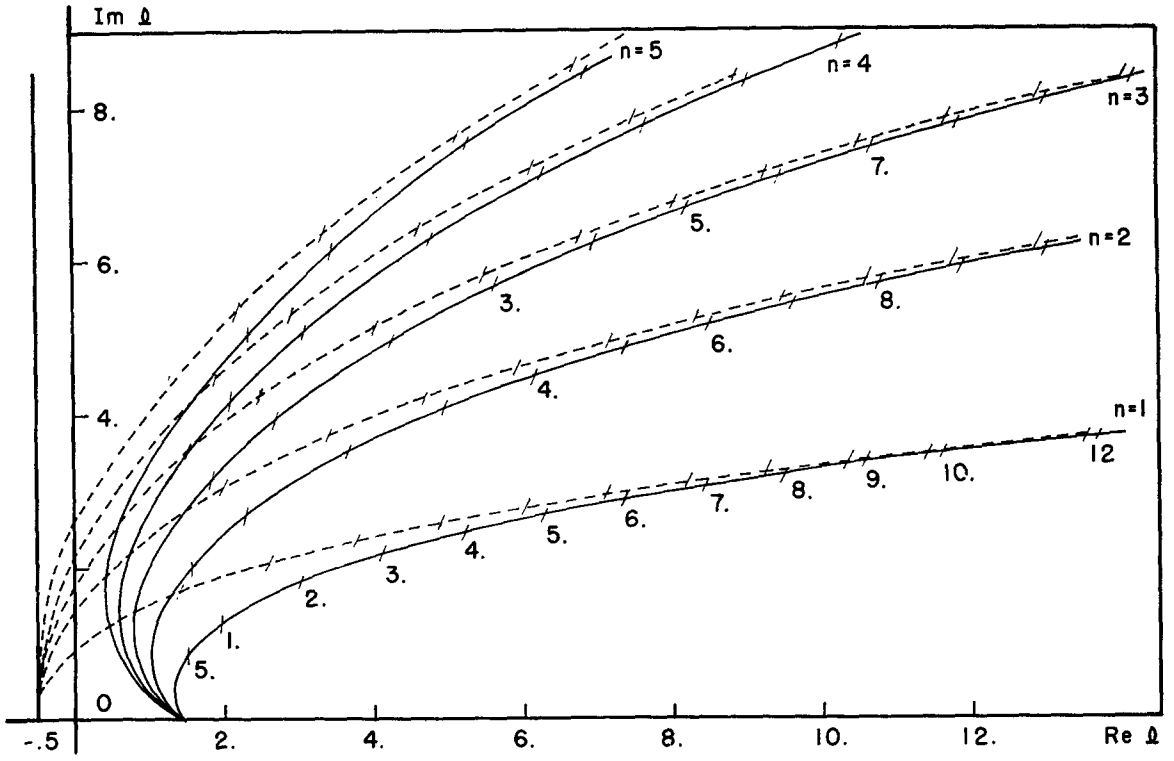


FIG. 2. Regge trajectories for a potential formed by a repulsive core with radius b and an attractive inverse square tail $-(\hbar^2/2m)\mu/r^2$. The solid lines represent the first trajectories for $\mu = 4$. For reference we show in dashed lines the corresponding trajectories for a pure hard core potential, $\mu = 0$. The form of the curves does not depend on the core radius b. For infinite negative energy there is an infinite number of poles at $(-\frac{1}{2}, \infty)$. As the energy increases the poles move down the line $\mathrm{Re}\, l = -\frac{1}{2}$, pass the origin for a nonpositive value of the energy and move along the positive real axis, all poles reaching the point $(\mu^{\frac{1}{2}} - \frac{1}{2}, 0)$ when the energy becomes zero. As the energy takes positive values, the poles pass to the complex plane, moving along different trajectories, all of which are tangent to the real axis at the point $(\mu^{\frac{1}{2}} - \frac{1}{2}, 0)$. For large energies the curves tend asymptotically to the trajectories corresponding to the pure hard core case. The numbers on the curves indicate the values of the parameter z.

As happened in the previous example, the S function depends on k and b only through the product $z = kb$. This implies that the trajectory followed in the angular momentum plane by a Regge pole when the energy varies from $-\infty$ to $+\infty$ is the same whatever the value of b. There is only a change of scale in the values of the parameter k when we pass from a given value of the core radius to another value.

The location of the poles of S is determined by the roots of $H_\nu^{(1)}(z)$, considered as a function of ν, for different values of the parameter z. These roots have been discussed by Keller *et al.*[4] for real z and by Ferreira *et al.*[7] for imaginary z. All results concerning the occurrence of bound states, and the behavior of poles, which were obtained in Sec. 2, remain valid in the present example. For illustration, we have represented in Figs. 2 and 3 the Regge trajectories for a potential $V(r)$ formed by an attractive inverse square tail $r > b$ and a repulsive hard core $r < b$.

For infinite negative energies all poles are on the positive imaginary axis of the λ plane. As the energy increases, the poles move down the $\mathrm{Re}\, l = -\frac{1}{2}$ line and reach the real axis at an energy which is different for each pole. As the energy is further increased, the poles move along the real axis. All poles reach the point $\mathrm{Re}\, l = \mu^{\frac{1}{2}} - \frac{1}{2}$ when the energy becomes zero. If $\mu^{\frac{1}{2}} > \frac{1}{2}$, the poles pass through the points $l = 0, 1, \cdots, < (\mu^{\frac{1}{2}} - \frac{1}{2})$ which correspond to physical bound states. The values of $\mathrm{Im}\, l$ at the poles are shown as a function of the energy in Fig. 3, for a few Regge trajectories. For comparison, we show in the same figure the corresponding five trajectories for the hard core potential without the inverse square tail. We see that at low energies the trajectories depend essentially on the presence of the tail. As $E \to -\infty$ (and so $\mathrm{Im}\, z \to +\infty$), the poles tend to ignore the tail and depend only on the core. This is in agreement with what we have seen in Sec. 2.

For positive energies the poles pass to the complex l plane, leaving the real axis at the point $\mu^{\frac{1}{2}} - \frac{1}{2}$. An approximate expression for the location of the poles at small $|z|$ can be obtained with a procedure similar

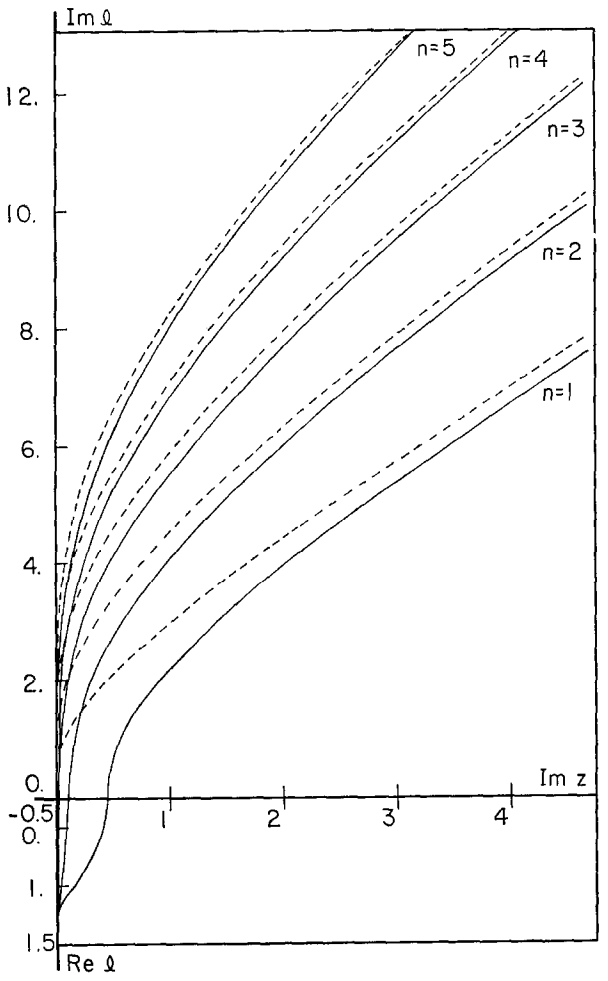


FIG. 3. Displacement of the Regge poles for negative energies for a potential formed by a repulsive core with b and an attractive inverse square tail $-(\hbar^2/2m)\mu/r^2$. The solid lines represent the first few trajectories for $\mu = 4$. For reference we show in dashed lines the trajectories for the pure hard core case, $\mu = 0$. The upper part of the vertical axis represents Im l, while the lower part indicates the values of Re l. In the example of the figure above, there is an infinite number of physical bound states for $l = 0$ and $l = 1$. At Re $l = \mu^{\frac{1}{2}} - \frac{1}{2}$ the poles pass to the complex plane. In the case of pure hard core, the poles do not move along the real axis and pass to the complex plane at Re $l = -\frac{1}{2}$.

to that used to derive Eq. (2.21). It is

$$\lambda = \mu^{\frac{1}{2}}\left\{1 - \frac{n^2\pi^2}{2\mu}\delta^2\left[1 - 2[i(\tfrac{1}{2}\pi - \varphi) - \gamma]\delta + \left(3[i(\tfrac{1}{2}\pi - \varphi) - \gamma]^2 + \frac{n^2\pi^2}{4\mu}\right)\delta^2 + O(\delta^3)\right]\right\}, \quad |z| \ll 1. \quad (3.6)$$

In first order we again obtain Eq. (2.23).

For large values of $|z|$ we can use, for the zeros of $H_\nu^{(1)}(z)$, the approximate expression given by Keller *et al.*[4]:

$$\nu_n = z + 6^{-\frac{1}{3}}e^{i\pi/3}q_n z^{\frac{1}{3}} + (180)^{-1}6^{\frac{1}{3}}e^{2\pi/3}q_n^2 z^{-\frac{1}{3}} + O(z^{-1}), \quad |z| \gg n, \quad (3.7)$$

where q_n is the nth zero of the Airy function,

$$\mathrm{Ai}\,(q_n) = \int_0^\infty \cos\,(t^3 - q_n t)\,dt = 0.$$

Using the equation

$$\lambda = \nu + O(\nu^{-1}),$$

we obtain

$$\lambda_n = z + 6^{-\frac{1}{3}}e^{i\pi/3}q_n z^{\frac{1}{3}} + (180)^{-1}6^{\frac{1}{3}}e^{i2\pi/3}q_n^2 z^{-\frac{1}{3}} + O(z^{-1}), \quad |z| \gg n, \quad (3.8)$$

where the strength μ of the tail does not appear.

The coincidence between the expressions in the right-hand side of Eqs. (3.7) and (3.8) was to be expected in view of the fact that the addition of a r^{-2} term in the hard core region is of no consequence. So, the Regge trajectories for a hard core potential with a r^{-2} tail are obtainable by the simple transformation Eq. (2.3) from those of a pure hard core potential. This makes our previous result quite self-evident.

In Fig. 2 we show the trajectories of the poles for positive energies in the angular momentum plane. All trajectories leave the real axis at Re $l = \mu^{\frac{1}{2}} - \frac{1}{2}$ moving towards the left along curves tangent to the real axis.

The trajectories corresponding to a pure hard core are also shown in Fig. 2 for comparison. We see that for large energies the poles tend to ignore the presence of the tail.

ACKNOWLEDGMENTS

We are grateful to the staff of the Centro de Cálculo de la Universidad de Madrid for the use of their equipment. One of the authors (J. S.) wishes to thank the Conselho Nacional de Pesquisas, Brazil, for financial assistance, and to Instituto de Física, Universidade Católica, for the hospitality during the period when this work was completed.

* This work was supported by Conselho Nacional de Pesquisas, Brazil, and by G.I.F.T., Spain.

[1] A. O. Barut and F. Calogero, Phys. Rev. **128**, 1383 (1962).
[2] R. G. Newton, J. Math. Phys. **3**, 867 (1962).
[3] S. Mandelstam, Ann. Phys. (N.Y.) **19**, 254 (1962).
[4] J. B. Keller, S. I. Rubinow, and M. Goldstein, J. Math. Phys. **4**, 829 (1963).
[5] R. G. Newton, *The Complex j-Plane* (Benjamin, New York, 1964), Chap. 9.
[6] G. N. Watson, *A Treatise on the Theory of Bessel Functions* (Cambridge U.P., Cambridge, 1966), 2nd ed., p. 260.
[7] E. M. Ferreira, J. Sesma, and P. L. Torres, "Zeros of the Modified Hankel Function," Numerische Math., to be published.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Electromagnetic Sources in Moving Simple Medium

GÜNTHER JOHANNSEN
Danish Engineering Academy, Electrical Department (formally Laboratory of Electromagnetic Field Theory)
The Technical University of Denmark, Lyngby, Denmark

(Received 4 May 1970)

A retarded potential tensor (4-vector) is derived in an arbitrary system of inertia for an arbitrary electromagnetic source in a moving homogeneous, isotropic, nondispersive, lossless dielectric. The velocity is uniform, and the result is relativistic correct.

I. INTRODUCTION

The differential equation for the potential tensor is simplest in the system of inertia K_m where the medium is at rest. A 4-dimensional integral representation for the potential tensor can readily be obtained by using the 4-dimensional Green's theorem. One may carry out the integration with respect to the time coordinate to get the space integral representation for the (retarded) potential tensor.

As pointed out in Ref. 1, the space integral representation in K_m is not very utilizable in another system of inertia K which is in uniform motion relative to K_m. This is due to the fact that space–time coordinates mix under the Lorentz transformation from K to K_m.

It is shown that the 4-dimensional integral representation for the potential tensor in K_m can be transformed to an arbitrary system of inertia K. Integrating with respect to the time coordinate then leads to the space integral representation for the (retarded) potential tensor in K. The results found here are in agreement with those in Refs. 2 and 3, where the pertinent differential equation is integrated by using Fourier transformation and an operational method, respectively.

II. INTEGRAL REPRESENTATIONS OF THE POTENTIAL TENSOR

We use Cartesian tensor notation. By a tensor we understand a tensor defined on the Lorentz transformation group. Latin subscripts run from 1 to 4; Greek subscripts run from 1 to 3. The coordinate x_4 is equal to ict, where t is the time and c the speed of light in vacuum; therefore, the metric tensor in 4-space is equal to the Kronecker symbol δ_{ij} (when Cartesian spatial coordinates are used), and we do not distinguish between contravariant and covariant tensors. Repeated subscripts obey the summation convention, and commas in subscripts denote partial differentiation with respect to coordinates (or covariant differentiation, since the metric tensor is independent of the coordinates).

The potential tensor A_i satisfies a differential equation,[1,4] which in an arbitrary system of inertia K can be written as a tensor equation:

$$A_{i,nn} - \kappa A_{i,st} U_s U_t = -S_i,$$
$$\kappa \equiv \frac{n^2 - 1}{c^2}, \quad n \equiv \frac{c}{c'}, \quad c' \equiv \frac{1}{(\mu\epsilon)^{\frac{1}{2}}}, \tag{1}$$
$$S_i \equiv \mu\left(J_i + \frac{\kappa}{n^2} J_r U_r U_i\right).$$

μ and ϵ are the permeability and the dielectric constant of the medium. U_i is the velocity tensor of K_m, and J_i is the current density tensor.[1,5]

Let primed quantities refer to K_m. Since $U_i' = (0, 0, 0, ic)$, we get from (1)

$$A'_{i,nn} + c^2\kappa A'_{i,44} = A'_{i,\nu\nu} + n^2 A'_{i,44} = -S'_i. \tag{2}$$

Introducing new variables by $y'_\rho \equiv x'_\rho$ and $y'_4 \equiv n^{-1}x'_4$, we get from (2)

$$A'_{i,nn}(y'_r) = -S'_i(y'_r), \tag{3}$$

where $(y'_r) \equiv (y'_1, y'_2, y'_3, y'_4)$.

For the solution of (3) see Ref. 5, p. 146. The result is

$$A'_i(y'_r) = (4\pi^2)^{-1} \int_{-\infty}^{\infty}\int\int\int\int \left(S'_i(z'_r) \Big/ \sum_{r=1}^{4}(z'_r - y'_r)^2\right) dV(z'_r). \tag{4}$$

Substituting back to variables x'_r, we obtain

$$A'_i(x'_r) = \int_{-\infty}^{\infty}\int\int\int\int S'_i(z'_r) G'(z'_r - x'_r)\, dV(z'_r), \tag{5}$$

where

$$G'(z'_r - x'_r) \equiv (4\pi^2 n)^{-1}[u'_r u'_r + (\kappa/n^2)(u'_r U'_r)^2]^{-1}$$

and

$$u'_r \equiv z'_r - x'_r.$$

The poles of the function G' are given by $u'_4 = \tau'_\pm \equiv \pm in(u'_\rho u'_\rho)^{\frac{1}{2}}$. As in Ref. 5 we can deform the contour for the z'_4 integration to circumvent one of the poles. Since advanced potentials are not of interest, we consider the pole τ'_-.

From (5) we get

$$A'_i(x'_r) = \int_{-\infty}^{\infty}\int\int\int \oint_{L'} S'_i(z'_r) G'(u'_r)\, du'_4\, du'_3\, du'_1\, du'_2, \tag{6}$$

where L' is a closed path surrounding τ'_-.

Equation (6) is valid in K_m; we now transform (6) to an arbitrary system of inertia K. Let $x_r = a_{rs}x'_s$ be the proper Lorentz transformation connecting coordinates in K with coordinates in K_m. Multiplying (6) by a_{ji} and substituting x'_r by x_r, we see that the left side in (6) (which is a tensor) becomes $A_j(x_r)$. a_{ji} is independent of the coordinates, and, since S'_i is a tensor, we have $a_{ji}S'_i(z'_r) = S_j(z_r)$ if z_r and z'_r are connected by the proper Lorentz transformation, i.e., $z_r = a_{rs}z'_s$. Furthermore, G' transforms like an invariant, i.e.,

$$G(u_r) = (4\pi^2 n)^{-1}[u_r u_r + (\kappa/n^2)(u_r U_r)^2]^{-1}. \quad (7)$$

Without loss of generality, we choose a_{rs} so that

$$x^1 = x'_1, \qquad x_2 = x'_2,$$
$$x_3 = \gamma(x'_3 + i\beta x'_4), \quad x_4 = \gamma(x'_4 - i\beta x'_3),$$

where $\gamma \equiv (1 - \beta^2)^{-\frac{1}{2}}$, $\beta \equiv v/c$, and v is the velocity of K relative to K_m (Fig. 1).

To see how the poles to be circumvented are transformed under the Lorentz transformation, we consider the tensor equation

$$u_r u_r + (\kappa/n^2)(u_r U_r)^2 = 0, \quad (8)$$

which is a 3-dimensional hypersurface in Minkowski 4-space. In K_m(8) reduces to $(u'_4/i)^2 - (nu'_3)^2 = (n\rho')^2$ and $\rho' \equiv (u'^2_1 + u'^2_2)^{\frac{1}{2}}$ which in Euclidean 2-space is the equation of a hyperbola (Cartesian coordinates u'_4/i, u'_3). In Fig. 2 the situation is illustrated in a way due to Minkowski.[6]

When $n\beta < 1$ and u_3 is given, there is one pole to be circumvented. When $n\beta > 1$ (this situation is shown in Fig. 2), there are two poles to be circumvented if $u_3 > |a|^{\frac{1}{2}}\rho$, $a \equiv [1 - (n\beta)^2]/(1 - \beta^2)$, and no poles are to be circumvented if $u_3 < |a|^{\frac{1}{2}}\rho$ (i.e., A_i is equal to zero).

The roots in (8) are given by

$$\frac{u_4}{i} = \tau_\pm \equiv \frac{n\beta}{1 - (n\beta)^2} \times [(n - n^{-1})u_3 \pm (\beta^{-1} - \beta)(u_3^2 + a\rho^2)^{\frac{1}{2}}]. \quad (9)$$

Since $n\beta < 1$ implies that $a > 0$ and $(\beta^{-1} - \beta) > n - n^{-1} > 0$, we actually see that $\tau_+ > 0$ and $\tau_- < 0$. In this case, from (6), we can deduce

$$A_i(x_r) = \iiint_{-\infty}^{\infty} \oint_L S_i(z_r)G(u_r)\, du_4\, du_3\, du_1\, du_2, \quad (10)$$

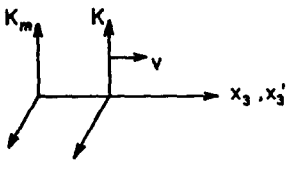


FIG. 1. The motion of K relative to K_m.

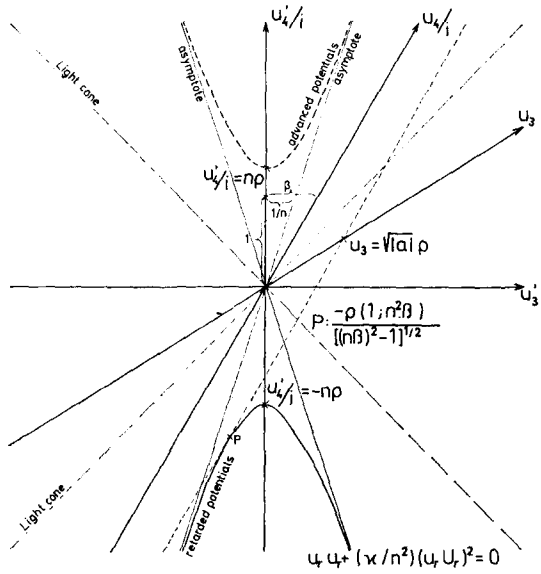


FIG. 2. Location of poles of the function G in Minkowski space.

where L circumvents τ_-. By the method of residues, we get

$$A_i(x_r) = \frac{1}{4\pi} \iiint_{-\infty}^{\infty} \frac{S_i(z_\rho, x_4 + \tau_-)}{[(z_3 - x_3)^2 + a\rho^2]^{\frac{1}{2}}}\, dV(z_\rho), \quad n\beta < 1. \quad (11)$$

In the case of $n\beta > 1$, we actually see from (9) that $\tau_\pm > 0$ when $u_3 < -|a|^{\frac{1}{2}}\rho$, that $\tau_\pm$ are imaginary when $-|a|^{\frac{1}{2}}\rho < u_3 < |a|^{\frac{1}{2}}\rho$, and finally that $\tau_\pm < 0$ when $u_3 > |a|^{\frac{1}{2}}\rho$. From (6) we get

$$A_i(x_r) = \iiint_{-\infty}^{\infty} \oint_L S_i(z_r)G(u_r)\, du_4\, du_3\, du_1\, du_2, \quad (12)$$

where the u_3 integration is to be taken from ∞ to $|a|^{\frac{1}{2}}\rho$ (L surrounding τ_-) and back to ∞ (L surrounding τ_+). It turns out that[7]

$$A_i(x_r) = -\frac{1}{4\pi} \iiint_{-\infty}^{\infty} \frac{S_i(z_\rho, x_4 + \tau_+) + S_i(z_\rho, x_4 + \tau_-)}{[(z_3 - x_3)^2 + a\rho^2]^{\frac{1}{2}}} \times \theta[z_3 - (x_3 + |a|^{\frac{1}{2}}\rho)]\, dV(z_\rho), \quad n\beta > 1, \quad (13)$$

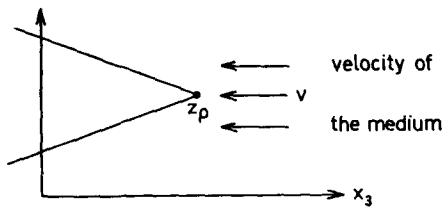


FIG. 3. Conical region in K, where a source point contributes to the field, $n\beta > 1$.

where θ is the unit step function, i.e.,

$$\theta(t) = 0 \quad \text{when} \quad t < 0,$$
$$= 1 \quad \text{when} \quad t \geq 0.$$

Equations (11) and (13) hold in an arbitrary system of inertia. The step function in (13) implies that a source point (z_ρ) contributes to the field only at points inside a conical region given by $(x_3 - z_3) \leq -|a|^{\frac{1}{2}}\rho$.

III. TIME HARMONIC SOURCES, 3-DIMENSIONAL REPRESENTATION

Let the source be time harmonic in K: $S_i(x_\tau) = S_i(x_\rho)e^{-kx_4}$, where $k \equiv \omega/c$ and ω is the frequency. Omitting the time factor e^{-kx_4}, we derive from (9)

$$S_i(z_\rho, x_4 + \tau_-) = S_i(z_\rho)\exp[-ikb(z_3 - x_3)] \times \exp\{ik(n/a)[(z_3 - x_3)^2 + a\rho^2]^{\frac{1}{2}}\}, \quad (14)$$

where

$$b \equiv \frac{n\beta}{1-(n\beta)^2}\left(n - \frac{1}{n}\right) = \frac{\beta\kappa c^2}{1-(n\beta)^2}.$$

Furthermore, we derive

$$S_i(z_\rho, x_4 + \tau_-) + S_i(z_\rho, x_4 + \tau_+) = 2S_i(z_\rho)e^{-ikb(z_3-x_3)}\cos\{(kn/a)[(z_3 - x_3)^2 + a_\rho^2]^{\frac{1}{2}}\}. \quad (15)$$

From the definition of S_i, and making use of the continuity equation $J_{\rho,\rho} = -J_{4,4} = i\omega\rho = kJ_4$, we get, for the spatial components of S_i,

$$S_\lambda(z_\rho) = \mu\left(\delta_{\lambda\nu} + \frac{\kappa}{n^2}U_\lambda U_\nu\right)J_\nu(z_\rho) + i\frac{\mu\kappa c^2}{n^2}\gamma U_\lambda J_{\nu,\nu}(z_\rho). \quad (16)$$

Similarly,

$$\frac{c}{i}S_4(z_\rho) = \frac{\mu c^2}{i\omega}\left(1 - \frac{\kappa c^2}{n^2}\gamma^2\right)J_{\nu,\nu} + \frac{\mu\kappa c^2}{n^2}\gamma U_\nu J_\nu. \quad (17)$$

Substituting (14)–(17) into (11) and (13) leads to expressions which are in agreement with Ref. 1.

The field vector $\bar{E}$ may be obtained by using the equation $\bar{E} = -\nabla\Phi + i\omega\bar{A}$ (Ref. 1; it can be shown that $\bar{E} = -\nabla\Phi - \partial\bar{A}/\partial t$ holds in any system of inertia). By some calculation the results may be transformed to an expression as given in Ref. 8.

[1] K. S. H. Lee and C. H. Papas, J. Math. Phys. **5**, 1668 (1964).
[2] R. T. Compton, Jr., J. Math. Phys. **7**, 2145 (1966).
[3] C. T. Tai, J. Math. Phys. **8**, 646 (1967).
[4] J. M. Jauch and K. M. Watson, Phys. Rev. **74**, 951 (1948).
[5] C. Møller, *The Theory of Relativity* (Oxford U.P., London, 1952).
[6] J. L. Synge, *Relativity: The Special Theory* (North-Holland, Amsterdam, 1965).
[7] Equation (10) in Ref. 3 should contain a sum of two δ functions.
[8] C. T. Tai, IEEE Trans. Antennas Propagation **13**, 322 (1965).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Expansion of an n-Point Function as a Sum of Commutators*

DON B. ZAGIER†
Laboratory for Nuclear Science, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received 22 October 1969; Revised Manuscript Received 15 April 1970)

We write the n-point function of currents as a sum over nested commutators, a form more suitable for certain current algebra calculations.

I. INTRODUCTION AND RESULT

In this paper we consider the n-point function, defined as

$$T(q_1, \cdots, q_{n-1})_{\mu_1\cdots\mu_n} = \int\int\cdots\int d^4x_1 \cdots d^4x_{n-1}e^{-iq_1x_1 - \cdots - iq_{n-1}x_{n-1}} \times \langle 0|\, T(j_{\mu_1}(x_1)\cdots j_{\mu_{n-1}}(x_{n-1})j_{\mu_n}(0))\,|0\rangle, \quad (1)$$

where $T(j_1\cdots j_n)$ is the product of the n current operators $j_1, \cdots, j_n$, in the order of the time components of their points of evaluation:

$$T(j_1(x_1)\cdots j_n(x_n)) = \sum_{\pi\in S_n}\theta(x^0_{\pi(1)} - x^0_{\pi(2)})\cdots\theta(x^0_{\pi(n-1)} - x^0_{\pi(n)}) \times j_{\pi(1)}(x_{\pi(1)})\cdots j_{\pi(n)}(x_{\pi(n)}), \quad (2)$$

where the sum is over all permutations π in S_n, the symmetric group of order n. The function $\theta(t)$ is the usual step function. In Eq. (1) we only integrate over $n-1$ space variables because we are using the translation invariance of the theory to work in a coordinate system where $x_n = 0$ and $q_1 + \cdots + q_n = 0$.

It is a straightforward calculation, which is given below, to rewrite Eq. (1) as a linear combination of products of n current operators (or, rather, of their Fourier transforms in momentum space) not involving the step functions θ. However, current algebra treats only commutators of operators rather than arbitrary products, so that it is desirable to express the n-point function as a linear combination of commutators of operators.

where θ is the unit step function, i.e.,

$$\theta(t) = 0 \quad \text{when} \quad t < 0,$$
$$= 1 \quad \text{when} \quad t \geq 0.$$

Equations (11) and (13) hold in an arbitrary system of inertia. The step function in (13) implies that a source point (z_ρ) contributes to the field only at points inside a conical region given by $(x_3 - z_3) \leq -|a|^{\frac{1}{2}}\rho$.

III. TIME HARMONIC SOURCES, 3-DIMENSIONAL REPRESENTATION

Let the source be time harmonic in K: $S_i(x_\tau) = S_i(x_\rho)e^{-kx_4}$, where $k \equiv \omega/c$ and ω is the frequency. Omitting the time factor e^{-kx_4}, we derive from (9)

$$S_i(z_\rho, x_4 + \tau_-) = S_i(z_\rho)\exp[-ikb(z_3 - x_3)] \times \exp\{ik(n/a)[(z_3 - x_3)^2 + a\rho^2]^{\frac{1}{2}}\}, \quad (14)$$

where

$$b \equiv \frac{n\beta}{1-(n\beta)^2}\left(n - \frac{1}{n}\right) = \frac{\beta\kappa c^2}{1-(n\beta)^2}.$$

Furthermore, we derive

$$S_i(z_\rho, x_4 + \tau_-) + S_i(z_\rho, x_4 + \tau_+) = 2S_i(z_\rho)e^{-ikb(z_3-x_3)}\cos\{(kn/a)[(z_3 - x_3)^2 + a_\rho^2]^{\frac{1}{2}}\}. \quad (15)$$

From the definition of S_i, and making use of the continuity equation $J_{\rho,\rho} = -J_{4,4} = i\omega\rho = kJ_4$, we get, for the spatial components of S_i,

$$S_\lambda(z_\rho) = \mu\left(\delta_{\lambda\nu} + \frac{\kappa}{n^2}U_\lambda U_\nu\right)J_\nu(z_\rho) + i\frac{\mu\kappa c^2}{n^2}\gamma U_\lambda J_{\nu,\nu}(z_\rho). \quad (16)$$

Similarly,

$$\frac{c}{i}S_4(z_\rho) = \frac{\mu c^2}{i\omega}\left(1 - \frac{\kappa c^2}{n^2}\gamma^2\right)J_{\nu,\nu} + \frac{\mu\kappa c^2}{n^2}\gamma U_\nu J_\nu. \quad (17)$$

Substituting (14)–(17) into (11) and (13) leads to expressions which are in agreement with Ref. 1.

The field vector $\bar{E}$ may be obtained by using the equation $\bar{E} = -\nabla\Phi + i\omega\bar{A}$ (Ref. 1; it can be shown that $\bar{E} = -\nabla\Phi - \partial\bar{A}/\partial t$ holds in any system of inertia). By some calculation the results may be transformed to an expression as given in Ref. 8.

[1] K. S. H. Lee and C. H. Papas, J. Math. Phys. **5**, 1668 (1964).
[2] R. T. Compton, Jr., J. Math. Phys. **7**, 2145 (1966).
[3] C. T. Tai, J. Math. Phys. **8**, 646 (1967).
[4] J. M. Jauch and K. M. Watson, Phys. Rev. **74**, 951 (1948).
[5] C. Møller, *The Theory of Relativity* (Oxford U.P., London, 1952).
[6] J. L. Synge, *Relativity: The Special Theory* (North-Holland, Amsterdam, 1965).
[7] Equation (10) in Ref. 3 should contain a sum of two δ functions.
[8] C. T. Tai, IEEE Trans. Antennas Propagation **13**, 322 (1965).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Expansion of an n-Point Function as a Sum of Commutators*

DON B. ZAGIER†
Laboratory for Nuclear Science, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received 22 October 1969; Revised Manuscript Received 15 April 1970)

We write the n-point function of currents as a sum over nested commutators, a form more suitable for certain current algebra calculations.

I. INTRODUCTION AND RESULT

In this paper we consider the n-point function, defined as

$$T(q_1, \cdots, q_{n-1})_{\mu_1\cdots\mu_n} = \int\int\cdots\int d^4x_1 \cdots d^4x_{n-1}e^{-iq_1x_1 - \cdots - iq_{n-1}x_{n-1}} \times \langle 0|\, T(j_{\mu_1}(x_1)\cdots j_{\mu_{n-1}}(x_{n-1})j_{\mu_n}(0))\,|0\rangle, \quad (1)$$

where $T(j_1\cdots j_n)$ is the product of the n current operators $j_1, \cdots, j_n$, in the order of the time components of their points of evaluation:

$$T(j_1(x_1)\cdots j_n(x_n)) = \sum_{\pi\in S_n}\theta(x^0_{\pi(1)} - x^0_{\pi(2)})\cdots\theta(x^0_{\pi(n-1)} - x^0_{\pi(n)}) \times j_{\pi(1)}(x_{\pi(1)})\cdots j_{\pi(n)}(x_{\pi(n)}), \quad (2)$$

where the sum is over all permutations π in S_n, the symmetric group of order n. The function $\theta(t)$ is the usual step function. In Eq. (1) we only integrate over $n-1$ space variables because we are using the translation invariance of the theory to work in a coordinate system where $x_n = 0$ and $q_1 + \cdots + q_n = 0$.

It is a straightforward calculation, which is given below, to rewrite Eq. (1) as a linear combination of products of n current operators (or, rather, of their Fourier transforms in momentum space) not involving the step functions θ. However, current algebra treats only commutators of operators rather than arbitrary products, so that it is desirable to express the n-point function as a linear combination of commutators of operators.

Bjorken[1] and Johnson and Low[2] pointed out that in the case $n = 2$ the 2-point function is asymptotically equal to a commutator term. They showed, specifically, that the leading term in the asymptotic expansion of

$$M_{\mu\nu}(q, \cdots) = -\int d^4x e^{-iq\cdot x}\langle A|\ T(j_\mu(x) j_\nu(0))\ |B\rangle, \quad (3)$$

as $q_0 \to \infty$ with $\mathbf{q}$ fixed, is

$$\frac{1}{q_0}\int d^3\mathbf{x} e^{-i\mathbf{q}\cdot\mathbf{x}}\langle A|\ [j_\mu(0, \mathbf{x}), j_\nu(0, \mathbf{0})]\ |B\rangle \quad (4)$$

and that the higher terms involve time derivatives of the currents; thus the next term is

$$\frac{1}{q_0^2} i\int d^3\mathbf{x} e^{-i\mathbf{q}\cdot\mathbf{x}}\left\langle A\left|\left[\frac{dj_\mu}{dt}(0, \mathbf{x}), j_\nu(0, \mathbf{0})\right]\right|B\right\rangle. \quad (5)$$

Not only the leading term, but all subsequent terms in the expansion of the n-point function, as the energies q_i^0 become infinite, can, in fact, be written as sums of equal-time commutators.[3] The expression we obtain is

$$\begin{aligned} &T(q_1, \cdots, q_{n-1})_{\mu_1\cdots\mu_n} \\ &= \frac{i^{n-1}}{n}\int\int\cdots\int d^3\mathbf{x}_1 \cdots d^3\mathbf{x}_{n-1} e^{-i\mathbf{q}_1\cdot\mathbf{x}_1 - \cdots - i\mathbf{q}_{n-1}\cdot\mathbf{x}_{n-1}} \\ &\quad \times \sum_{\pi\in S_n}\left[\prod_{s=1}^{n-1}\left(E_{\pi(1)} + \cdots + E_{\pi(s)} \right.\right. \\ &\quad \left.\left. + i\frac{\partial}{\partial t_{\pi(1)}} + \cdots + i\frac{\partial}{\partial t_{\pi(s)}}\right)^{-1}\right] \\ &\quad \times \langle 0|\ [[\cdots [j_{\mu_{\pi(1)}}, j_{\mu_{\pi(2)}}], \cdots], j_{\mu_{\pi(n)}}]\ |0\rangle|_{t_1=\cdots=t_n=0}, \end{aligned} \quad (6)$$

where the meaning of the right-hand side is that for each permutation π we expand each factor

$$\left(E_{\pi(1)} + \cdots + E_{\pi(s)} + i\frac{\partial}{\partial t_{\pi(1)}} + \cdots + i\frac{\partial}{\partial t_{\pi(s)}}\right)^{-1}$$

as a series

$$\sum_{r=0}^{\infty}\left(-i\frac{\partial}{\partial t_{\pi(1)}} - \cdots - i\frac{\partial}{\partial t_{\pi(s)}}\right)^r \Big/ (E_{\pi(1)} + \cdots + E_{\pi(s)})^{r+1},$$

formally multiply these differential operators, apply the product operator to the commutator

$$[[\cdots [j_{\mu_{\pi(1)}}(t_{\pi(1)}, \mathbf{x}_{\pi(1)}), j_{\mu_{\pi(2)}}(t_{\pi(2)}, \mathbf{x}_{\pi(2)})], \cdots], j_{\mu_{\pi(n)}}(t_{\pi(n)}, \mathbf{x}_{\pi(n)})],$$

evaluate at $t_1 = \cdots = t_n = 0$, and finally Fourier-transform the space part of the result and divide by n. We have made the convention that

$$\frac{\partial}{\partial t_1} + \cdots + \frac{\partial}{\partial t_n} = 0$$

(since x_n and hence t_n are identically zero, $\partial/\partial t_n$ is undefined), and t_i and E_i are the time components of x_i^μ and q_i^μ, respectively. Written out in full, our result is

$$\begin{aligned} &T(E_1, \cdots, E_{n-1})_{1\cdots n} \\ &= i^{n-1}\int\int\cdots\int dt_1 \cdots dt_{n-1} \\ &\quad \times e^{-iE_1t_1 - \cdots - iE_{n-1}t_{n-1}}\langle 0|\ T(j_{\mu_1}(x_1)\cdots j_{\mu_n}(0))\ |0\rangle \\ &= \frac{1}{n}\sum_{\pi\in S_n}\sum_{r_1=0}^{\infty}\cdots\sum_{r_{n-1}=0}^{\infty} \\ &\quad \times \left(\prod_{s=1}^{\pi^{-1}(n)-1}\frac{(-i\partial/\partial t_{\pi(1)} - \cdots - i\partial/\partial t_{\pi(s)})^{r_s}}{(E_{\pi(1)} + \cdots + E_{\pi(s)})^{r_s+1}}\right) \\ &\quad \times \left(\prod_{s=\pi^{-1}(n)}^{n-1}\frac{(i\partial/\partial t_{\pi(s+1)} + \cdots + i\partial/\partial t_{\pi(n)})^{r_s}}{(E_{\pi(1)} + \cdots + E_{\pi(s)})^{r_s+1}}\right) \\ &\quad \times \langle 0|\ [[\cdots [j_{\mu_{\pi(1)}}(t_{\pi(1)}, \mathbf{x}_{\pi(1)}), j_{\mu_{\pi(2)}}(t_{\pi(2)}, \mathbf{x}_{\pi(2)})], \cdots], \\ &\qquad\qquad j_{\mu_{\pi(n)}}(t_{\pi(n)}, \mathbf{x}_{\pi(n)})]\ |0\rangle. \end{aligned}$$

Here we have omitted the integration over the space variables.

Before deriving this result, we will state it in a different form. Since the space variables and integrations do not affect the problem, we will cease to write them; similarly, we omit the brackets $\langle 0|\cdots|0\rangle$ denoting the vacuum state. Although the j_{μ_i} are, in fact, components of a single current, we do not use this, but treat them as separate functions; since the subscripts μ_i do not change, we omit them. Thus $j_{\mu_i}(t_i, \mathbf{x}_i)$ will be denoted $j_i(t_i)$ for $1 \le i \le n-1$. For convenience, we define j_n by $j_n(t) = j_{\mu_n}(0)\delta(t)$. We use the following Fourier transform:

$$\tilde{f}(E) = \frac{1}{2\pi}\int e^{-iEt} f(t)\, dt, \quad f(t) = \int e^{+iEt}\tilde{f}(E)\, dE. \quad (7)$$

With this definition of the Fourier transform, we have

$$\begin{aligned} (-i)^r \frac{d^r}{dt^r} f(t)|_{t=0} &= \int (-i)^r \frac{\partial^r}{\partial t^r} e^{iEt}\tilde{f}(E)\, dE|_{t=0} \\ &= \int E^r \tilde{f}(E)\, dE. \end{aligned} \quad (8)$$

Making all of these changes and substituting the definition (2) for the time-ordered product, we obtain as the theorem to be demonstrated,[4]

$$\begin{aligned} &i^{n-1}\int\int\cdots\int dt_1 \cdots dt_n e^{-iE_1t_1 - \cdots - iE_nt_n} \\ &\quad \times \sum_{\pi\in S_n}\theta(t_{\pi(1)} - t_{\pi(2)})\cdots\theta(t_{\pi(n-1)} - t_{\pi(n)}) \\ &\quad \times j_{\pi(1)}(t_{\pi(1)})\cdots j_{\pi(n)}(t_{\pi(n)}) \\ &= \frac{2\pi}{n}\int\int\cdots\int dE_1' \cdots dE_n'\delta(E_1' + \cdots + E_n') \\ &\quad \times \sum_{\pi\in S_n}\frac{[[\cdots [\tilde{j}_{\pi(1)}(E_1'), \tilde{j}_{\pi(2)}(E_2')], \cdots], \tilde{j}_{\pi(n)}(E_n')]}{\prod_{s=1}^{n-1}(E_{\pi(1)} + \cdots + E_{\pi(s)} - E_1' - \cdots - E_s')}. \end{aligned} \quad (9)$$

This is the real result, and the previous form, in which integrations had been done by using Eq. (8), is simply its expansion for large energies [which may not be valid; it is only the identity (9) that will be proved rigorously].

The proof of the theorem, Eq. (9), is purely algebraic in nature and so makes no reference to the actual existence of the commutators and T products which we consider. It has been shown[5] that, in fact, they do not exist in certain perturbation-theoretical models. We regard the question of their existence in general as being open at the present time. In the following, we assume that there is a theory for which the problems found in Ref. 5 do not exist.

II. PROOF

The proof of Eq. (9) will proceed in two stages: first, transforming the left-hand side to an expression involving a linear combination of products of n current operators, and then rewriting this as a sum of commutators. With the normalization of Eqs. (7), the Fourier transform of $\theta(t-t_0)$ is $e^{-iEt_0}/2\pi iE$, and the rule for transforming a product is

$$\widetilde{fg}(E)=\frac{1}{2\pi}\int f(t)g(t)e^{-iEt}\,dt=\int\tilde{f}(E')\tilde{g}(E-E')\,dE', \tag{10}$$

so that

$$\int\theta(t-t_0)f(t)e^{-iEt}\,dt=\int\frac{\tilde{f}(E')e^{-i(E-E')t_0}}{i(E-E')}\,dE'. \tag{11}$$

Applying this repeatedly, we obtain [abbreviating $\theta(t_i-t_j)$ to θ_{ij}]

$$\begin{aligned}
&\iint\cdots\int dt_1\cdots dt_n\\
&\quad\times e^{-iE_{\pi(1)}t_1-\cdots-iE_{\pi(n)}t_n}\theta_{12}\cdots\theta_{n-1,n}j_{\pi(1)}(t_1)\cdots j_{\pi(n)}(t_n)\\
&=\iint\cdots\int dt_2\cdots dt_n\\
&\quad\times e^{-iE_{\pi(2)}t_2-\cdots-iE_{\pi(n)}t_n}\theta_{23}\theta_{34}\cdots\theta_{n-1,n}\\
&\quad\times\int\frac{\tilde{j}_{\pi(1)}(E_1')e^{-i(E_{\pi(1)}-E_1')t_2}}{i(E_{\pi(1)}-E_1')}\,dE_1'j_{\pi(2)}(t_2)\cdots j_{\pi(n)}(t_n)\\
&=\iint\cdots\int dE_1'\,dE_2'\,dt_3\cdots dt_n\\
&\quad\times e^{-iE_{\pi(3)}t_3-\cdots-iE_{\pi(n)}t_n}e^{-i(E_{\pi(1)}+E_{\pi(2)}-E_1'-E_2')t_3}\\
&\quad\times\theta_{34}\cdots\theta_{n-1,n}\frac{\tilde{j}_{\pi(1)}(E_1')}{i(E_{\pi(1)}-E_1')}\\
&\quad\times\frac{\tilde{j}_{\pi(2)}(E_2')}{i(E_{\pi(1)}+E_{\pi(2)}-E_1'-E_2')}j_{\pi(3)}(t_3)\cdots j_{\pi(n)}(t_n)\\
&=\cdots\\
&=\frac{1}{i^{n-1}}\iint\cdots\int dE_1'\cdots dE_{n-1}'\,dt_n\\
&\quad\times e^{-i(E_{\pi(1)}+\cdots+E_{\pi(n)}-E_1'-\cdots-E_{n-1}')t_n}\\
&\quad\times\frac{\tilde{j}_{\pi(1)}(E_1')}{E_{\pi(1)}-E_1'}\,\frac{\tilde{j}_{\pi(2)}(E_2')}{E_{\pi(1)}+E_{\pi(2)}-E_1'-E_2'}\cdots\\
&\quad\times\frac{\tilde{j}_{\pi(n-1)}(E_{n-1}')}{E_{\pi(1)}+\cdots+E_{\pi(n-1)}-E_1'-\cdots-E_{n-1}'}j_{\pi(n)}(t_n).
\end{aligned}$$

The final expression contains no θ functions and, therefore, can be evaluated by the direct substitution of the first of Eqs. (7). We notice that

$$E_{\pi(1)}+\cdots+E_{\pi(n)}=E_1+\cdots+E_n=0$$

(since $q_1+\cdots+q_n$ was zero), so that the result is

$$\frac{2\pi}{i^{n-1}}\iint\cdots\int dE_1'\cdots dE_{n-1}'\times\frac{\tilde{j}_{\pi(1)}(E_1')\cdots\tilde{j}_{\pi(n-1)}(E_{n-1}')\tilde{j}_{\pi(n)}(-E_1'-\cdots-E_{n-1}')}{\prod_{s=1}^{n-1}(E_{\pi(1)}+\cdots+E_{\pi(s)}-E_1'-\cdots-E_s')},$$

which may be written more symmetrically as

$$\frac{2\pi}{i^{n-1}}\iint\cdots\int dE_1'\cdots dE_n'\delta(E_1'+\cdots+E_n')\times\frac{\tilde{j}_{\pi(1)}(E_1')\cdots\tilde{j}_{\pi(n)}(E_n')}{\prod_{s=1}^{n-1}(E_{\pi(1)}+\cdots+E_{\pi(s)}-E_1'-\cdots-E_s')}.$$

This completes the first stage. Multiplying our last equation by i^{n-1} and summing over all permutations π of $1, 2, \cdots, n$, for the left-hand side of Eq. (9),we obtain the expression

$$2\pi\iint\cdots\int dE_1'\cdots dE_n'\delta(E_1'+\cdots+E_n')\times\sum_{\pi\in S_n}\frac{\tilde{j}_{\pi(1)}(E_1')\cdots\tilde{j}_{\pi(n)}(E_n')}{\prod_{s=1}^{n-1}(E_{\pi(1)}+\cdots+E_{\pi(s)}-E_1'-\cdots-E_s')},$$

the desired expression as a sum of products of the $\tilde{j}_i$. Both in this expression and in Eq. (9) we could just as well have written $E_{\pi(i)}'$ for E_i' in the denominators, since for each permutation π one could relabel the symmetric expression

$$\iint\cdots\int dE_1'\cdots dE_n'\delta(E_1'+\cdots+E_n').$$

Therefore, it suffices to prove the purely algebraic identity

$$n\sum_{\pi\in S_n}\frac{\tilde{j}_{\pi(1)}(E_{\pi(1)}')\cdots\tilde{j}_{\pi(n)}(E_{\pi(n)}')}{\prod_{s=1}^{n-1}(E_{\pi(1)}+\cdots+E_{\pi(s)}-E_{\pi(1)}'-\cdots-E_{\pi(s)}')}=\sum_{\pi\in S_n}\frac{[[\cdots[\tilde{j}_{\pi(1)}(E_{\pi(1)}'),\tilde{j}_{\pi(2)}(E_{\pi(2)}')],\cdots],\tilde{j}_{\pi(n)}(E_{\pi(n)}')]}{\prod_{s=1}^{n-1}(E_{\pi(1)}+\cdots+E_{\pi(s)}-E_{\pi(1)}'-\cdots-E_{\pi(s)}')}. \tag{12}$$

Having simplified the problem, we will again simplify our notation. Since the character of j as a function of energy no longer interests us, we will omit the argument and tilde and write simply j_i for $j_i(E_i')$. Next, we have $E_1' + \cdots + E_n' = 0$. [Because of the presence of the Dirac δ, it does not matter whether or not the sum in Eq. (9) can be written as a sum of commutators off the subspace $E_1' + \cdots + E_n' = 0$; indeed, simple examples show that it cannot.] Of course, $E_1 + \cdots + E_n$ is always zero. Furthermore, because we were able to group together E's and E''s with the same subscript in Eq. (12), we can now define a new set of numbers $F_i = E_i - E_i'$, and the identity to be demonstrated becomes

$$n \sum_{\pi \in S_n} \frac{j_{\pi(1)} \cdots j_{\pi(n)}}{F_{\pi(1)}(F_{\pi(1)} + F_{\pi(2)}) \cdots (F_{\pi(1)} + \cdots + F_{\pi(n-1)})} = \sum_{\pi \in S_n} \frac{[[\cdots [j_{\pi(1)}, j_{\pi(2)}], \cdots], j_{\pi(n)}]}{F_{\pi(1)}(F_{\pi(1)} + F_{\pi(2)}) \cdots (F_{\pi(1)} + \cdots + F_{\pi(n-1)})}, \tag{13}$$

where the F_i are n numbers such that their sum is zero, but no subsum is zero [the identity is meaningless if any subsum vanishes, but it suffices to prove it in the converse case since the region where some subsum vanishes has zero measure in the $(n-1)$-dimensional space $E_1' + \cdots + E_n' = 0$], the j_i are noncommuting quantities, and the sums extend over all permutations of $1, 2, \cdots, n$.

To prove Eq. (13), we evidently have to expand the commutators on the right and then rearrange the sum so that we can pick out the coefficient of a given product $j_{\pi(1)} \cdots j_{\pi(n)}$ and check that it is indeed

$$n/F_{\pi(1)}(F_{\pi(1)} + F_{\pi(2)}) \cdots (F_{\pi(1)} + \cdots + F_{\pi(n-1)}).$$

A commutator of n operators has 2^{n-1} terms, half of them positive and half negative, and we must start by finding the rule which determines which of the $n!$ possible permutations appear and with what sign. If we expand a commutator such as $[[[[h_1, h_2], h_3], h_4], h_5]$, then typical terms are $h_5h_4h_1h_2h_3$ and $-h_3h_1h_2h_4h_5$. Inspecting the terms, we see that each one has descending subscripts up to h_1 and then ascending, so that it is in the form $h_{\sigma(1)}h_{\sigma(2)} \cdots h_{\sigma(n)}$, where $\sigma(1) > \cdots > \sigma(k) = 1 < \sigma(k+1) < \cdots < \sigma(n)$ for some k, and that the sign of such a term is $(-1)^{k-1}$. Now, for a given value of k, we can choose any $k-1$ of the $n-1$ numbers $2, 3, \cdots, n$ to precede $\sigma(k) = 1$, but then their order is determined; thus the set $S_{n,k}$ of permutations σ with $\sigma(1) > \cdots > \sigma(k) = 1 < \sigma(k+1) < \cdots < \sigma(n)$ has $\binom{n-1}{k-1}$ members and, since

$$\sum_{k=1}^{n} \binom{n-1}{k-1} = 2^{n-1},$$

if all the terms of the commutator are of the special form considered, then all terms of this form appear in the commutator. That this is, in fact, the case can be seen easily by induction: We have to prove

$$[[\cdots [h_1, h_2], \cdots], h_n] = \sum_{k=1}^{n} (-1)^{k-1} \sum_{\sigma \in S_{n,k}} h_{\sigma(1)} \cdots h_{\sigma(n)}, \tag{14}$$

an identity plainly valid for n equal to one or two. If it is valid for n, then $[[[\cdots [h_1, h_2], \cdots], h_n], h_{n+1}]$ is given by

$$\begin{aligned}
&\sum_{k=1}^{n} (-1)^{k-1} \sum_{\sigma \in S_{n,k}} (h_{\sigma(1)}h_{\sigma(2)} \cdots h_{\sigma(n)}h_{n+1} - h_{n+1}h_{\sigma(1)}h_{\sigma(2)} \cdots h_{\sigma(n)}) \\
&= \sum_{k=1}^{n} (-1)^{k-1} \sum_{\substack{\pi \in S_{n+1} \\ \pi(1) > \cdots > \pi(k) \\ = 1 < \cdots < \pi(n+1) = n+1}} h_{\pi(1)} \cdots h_{\pi(n+1)} \\
&\quad + \sum_{k=2}^{n+1} (-1)^{k-1} \sum_{\substack{\pi \in S_{n+1} \\ n+1 = \pi(1) > \cdots > \pi(k) \\ = 1 < \cdots < \pi(n+1)}} h_{\pi(1)} \cdots h_{\pi(n+1)} \\
&= \sum_{k=1}^{n+1} (-1)^{k-1} \sum_{\pi \in S_{n+1,k}} h_{\pi(1)} \cdots h_{\pi(n+1)},
\end{aligned}$$

there being the last equality because any permutation π of $S_{n+1,k}$ must have either $\pi(1)$ or $\pi(n+1)$ equal to $n+1$ [since any other $\pi(i)$ is smaller than one of these]. This completes the proof of Eq. (14).

Now we can expand the right-hand side of Eq. (13) to obtain

$$\begin{aligned}
&\sum_{\pi \in S_n} \frac{[[\cdots [j_{\pi(1)}, j_{\pi(2)}], \cdots], j_{\pi(n)}]}{F_{\pi(1)}(F_{\pi(1)} + F_{\pi(2)}) \cdots (F_{\pi(1)} + \cdots + F_{\pi(n-1)})} \\
&= \sum_{\pi \in S_n} \sum_{k=1}^{n} \sum_{\sigma \in S_{n,k}} (-1)^{k-1} \times \frac{j_{\pi\sigma(1)} \cdots j_{\pi\sigma(n)}}{F_{\pi(1)}(F_{\pi(1)} + F_{\pi(2)}) \cdots (F_{\pi(1)} + \cdots + F_{\pi(n-1)})}.
\end{aligned}$$

The coefficient of $j_{\tau(1)}j_{\tau(2)} \cdots j_{\tau(n)}$ in this is obtained by noting that for each σ in $S_{n,k}$ there is a unique π in S_n (namely, $\pi = \tau\sigma^{-1}$) such that $\pi\sigma = \tau$; thus, the coefficient is

$$\sum_{k=1}^{n} \sum_{\sigma \in S_{n,k}} \frac{(-1)^{k-1}}{F_{\tau\sigma^{-1}(1)}(F_{\tau\sigma^{-1}(1)} + F_{\tau\sigma^{-1}(2)}) \cdots (F_{\tau\sigma^{-1}(1)} + \cdots + F_{\tau\sigma^{-1}(n-1)})}.$$

To prove Eq. (13), we must show that this sum equals $n\prod_{s=1}^{n-1}(F_{\tau(1)}+\cdots+F_{\tau(s)})^{-1}$. By expanding the commutator, we have gotten rid of the operators and have reduced the problem to an algebraic identity among ordinary numbers. Having fastened our attention on a single permutation τ, we need no longer carry it as a subscript, but set $G_i = F_{\tau(i)}$. If we can prove

$$\sum_{\sigma\in S_{n,k}} \frac{1}{G_{\sigma^{-1}(1)}(G_{\sigma^{-1}(1)}+G_{\sigma^{-1}(2)})\cdots(G_{\sigma^{-1}(1)}+\cdots+G_{\sigma^{-1}(n-1)})} = \frac{(-1)^{k-1}}{G_1(G_1+G_2)\cdots(G_1+\cdots+G_{n-1})} \tag{15}$$

for $1 \le k \le n$, then the desired equality will follow on summation from 1 to n.

We will prove Eq. (15) by another simple induction. For $n = 2$, it reduces to $+1/G_1 = +1/G_1$ or $+1/G_2 = -1/G_1$, depending on whether k is one or two, and both are true since $G_1 + G_2$ is zero. Our previous induction hinged on the fact that, for σ in $S_{n,k}$, either $\sigma(1)$ or $\sigma(n)$ must be n, since each $\sigma(i)$ is smaller than one of them; this one depends on the fact that $\sigma(k-1)$ or $\sigma(k+1)$ must be 2 since every $\sigma(i)$, except $\sigma(k) = 1$, is greater than one of them. Hence the left-hand side of Eq. (15) is

$$\frac{1}{G_k}\sum_{\substack{\sigma\in S_{n,k}\\ \sigma(k-1)=2}} \frac{1}{(G_k+G_{k-1})(G_k+G_{k-1}+G_{\sigma^{-1}(3)})\cdots(G_k+G_{k-1}+\cdots+G_{\sigma^{-1}(n-1)})} + \frac{1}{G_k}\sum_{\substack{\sigma\in S_{n,k}\\ \sigma(k+1)=2}} \frac{1}{(G_k+G_{k+1})(G_k+G_{k+1}+G_{\sigma^{-1}(3)})\cdots(G_k+G_{k+1}+\cdots+G_{\sigma^{-1}(n-1)})}. \tag{16}$$

If k is 1, the first sum is empty and, if k is n, the second is also empty; but if this is kept in mind, the following proof still is applicable. In any case Eq. (15) is almost trivial for $k = 1$ or $k = n$. To evaluate the two sums in Eq. (16), we use the fact that both can be transformed to special cases of Eq. (15) for $n-1$. Thus, if we define numbers H_i for $1 \le i \le n-1$ and [for each σ in the first sum in Eq. (16)] a permutation π of S_{n-1} by

$$\begin{aligned} H_i &= G_i, && 1 \le i \le k-2, && \pi(i) = \sigma(i) - 1, && 1 \le i \le k-2,\\ &= G_{k-1}+G_k, && i = k-1, && = 1, && i = k-1,\\ &= G_{i+1}, && k \le i \le n-1, && = \sigma(i+1) - 1, && k \le i \le n-1, \end{aligned}$$

and notice that

$$H_1 + \cdots + H_{n-1} = G_1 + \cdots + (G_{k-1}+G_k) + \cdots + G_n = 0,$$

we can rewrite the first sum in Eq. (16) as

$$\sum_{\pi\in S_{n-1,k-1}} \frac{1}{(H_{\pi^{-1}(1)})(H_{\pi^{-1}(1)}+H_{\pi^{-1}(2)})\cdots(H_{\pi^{-1}(1)}+\cdots+H_{\pi^{-1}(n-2)})} = \frac{(-1)^{k-2}}{H_1(H_1+H_2)\cdots(H_1+\cdots+H_{n-2})},$$

the equality following from the induction hypothesis. Hence the first term in Eq. (16) is

$$(-1)^{k-2}/G_kG_1(G_1+G_2)\cdots(G_1+\cdots+G_{k-2}) \times (G_1+\cdots+G_{k-1}+G_k)\cdots(G_1+\cdots+G_{n-1}).$$

Exactly similarly, the second term in Eq. (16) is

$$(-1)^{k-1}/G_kG_1(G_1+G_2)\cdots(G_1+\cdots+G_{k-1}) \times (G_1+\cdots+G_k+G_{k+1})\cdots(G_1+\cdots+G_{n-1}).$$

Adding these, we see that the expression (16), which represents the left-hand side of Eq. (15), equals

$$\frac{(-1)^{k-1}(G_1+\cdots+G_k)+(-1)^{k-2}(G_1+\cdots+G_{k-1})}{G_kG_1(G_1+G_2)\cdots(G_1+\cdots+G_{n-1})} = \frac{(-1)^{k-1}}{G_1(G_1+G_2)\cdots(G_1+\cdots+G_{n-1})},$$

which is the desired right-hand side of Eq. (15). It is interesting that the crucial hypothesis $G_1+\cdots+G_n = 0$ did not enter the proof except to establish the case $n = 2$, and for larger n it was only needed to be able to apply the induction hypothesis.

ACKNOWLEDGMENTS

I would like to thank Professor I. Gerstein for suggesting this problem and for his guidance and encouragement in preparing the manuscript. Part of this paper was written while I held a fellowship from the National Science Foundation.

* Based on a thesis submitted to the Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts, in partial fulfillment of the requirements for the B.S. Degree. This work was partially supported by the U.S. Atomic Energy Commission under Contract No. AT (30-1)-2098.

† Present address: New College, Oxford, England.

[1] J. D. Bjorken, Phys. Rev. **148**, 1467 (1966).

[2] K. Johnson and F. Low, Progr. Theoret. Phys. (Kyoto), Suppl. Nos. **37–38**, 74 (1966).

[3] Some partial results of this type have been used by P. Oleson [Phys. Rev. **175**, 2165 (1968)] and A. Sirlin [Phys. Rev. **176**, 1875 (1968)]. Our main result, Eq. (9), agrees with the specific examples discussed by these authors.

[4] The reason for the factor 2π in the right-hand side of Eq. (9) is that, with the Fourier transform of Eq. (7), the transform $\tilde{j}_n(E)$ is not $j_{\mu_n}(0)$ but $(2\pi)^{-1}j_{\mu_n}(0)$. [$j_n(t)$ was defined as $j_{\mu_n}(0)\delta(t)$.]

[5] R. Jackiw and G. Preparata, Phys. Rev. Letters **22**, 975 (1969); S. Adler and W.-K. Tung, Phys. Rev. Letters **22**, 978 (1969).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Existence of Charged States of the Yang–Mills Field*

HENDRICUS G. LOOS
Douglas Advanced Research Laboratories, Huntington Beach, California
and
Department of Physics, University of California at Riverside

(Received 5 March 1970)

For a tentative choice of configuration space Ω, it is proved that the Yang–Mills field, self-interacting but not coupled to other fields, has states with a nonvanishing isospin component if gauge-invariant quantization is used. This is shown by proving existence of a solution for the elliptic boundary-value problem $\nabla_\beta \nabla^\beta \zeta^i(\mathbf{x}) = 0$ on all of 3-dimensional Euclidean space, subject to the asymptotic condition $\zeta^i = c^i + O(r^{-1})$, $\partial_\beta \partial^\beta \zeta^i = O(r^{-4})$ as $r \to \infty$, where c^i are constants; ∇_β is the covariant derivative belonging to the spatial Yang–Mills potentials $b_\beta{}^i(\mathbf{x})$. The existence proof is a modification of Schauder's proof to an unbounded domain. Ω consists of all numerical real multiplet functions $b_\beta{}^i(\mathbf{x})$ which are of order $O(r^{-2})$ as $r \to \infty$, have $\partial_\beta b^{\beta i} = O(r^{-4})$, and satisfy certain smoothness conditions. Also, for this configuration space, the problem of existence of equivalent transverse potentials is reduced to a simpler uniqueness problem. In the classical theory, the existence of solutions ζ^i implies that the constraint equation can be satisfied for any choice of the "covariant-transverse" part of $B^{0\beta i}$ within a very large class, by a unique "covariant-longitudinal" part of $B^{0\beta i}$, if the potentials $b_\alpha{}^i(\mathbf{x})$ have the full $SU(2)$ as holonomy group.

1. INTRODUCTION

The present paper is concerned with certain clarifications needed for a further development of the gauge-invariant quantization[1] of the Yang–Mills field. In that quantization method, use is made of the Schrödinger representation, in which the spatial Yang–Mills potentials $b_\beta{}^i(\mathbf{x})$, $\beta = 1, 2, 3$ and $i = 1, 2, 3$, are diagonal, and states are represented by complex-valued functionals $\Psi[b]$ of the real numerical potential functions $b_\beta{}^i(\mathbf{x})$. By an obvious generalization of concepts for quantum mechanical systems with a finite number of degrees of freedom, the $b_\beta{}^i(\mathbf{x})$ are seen as generalized coordinates which specify the *configuration* of the Yang–Mills field. The configuration space Ω for the Yang–Mills field consists of all "kinematically" possible configurations. So far, it has not been possible to give a complete and final specification of the configuration space Ω. There are restrictive demands on Ω which come from mathematical aspects of the theory, while Ω should be general enough to include all "physically" significant configurations. Among the mathematical aspects of the theory which force restrictions on Ω are the existence of certain derivatives and integrals, which set conditions of smoothness and asymptotic behavior on the functions $b_\beta{}^i(\mathbf{x})$. If transverse potentials are to be used in the theory, Ω must be restricted such that all potentials $b \in \Omega$ can be transformed into transverse potentials by gauge transformations, and this amounts to an existence problem for an elliptic boundary value problem on all of 3-dimensional Euclidean space. The physical demands on the configuration space Ω depend on our expectations for the physical capabilities of the Yang–Mills field, and one can easily go wrong here.

One of our physical requirements is that the Yang–Mills field is capable of states with a nonvanishing component of isospin. The main purpose of the present paper is to prove existence of such charged states for the Yang–Mills field, *for a tentative choice of the configuration space* Ω. This choice of Ω seems physically attractive, and as far as we know meets the essential mathematical requirements of the theory, as it stands at the present time. The second purpose of the paper is to show that, for this configuration space Ω, the existence problem for gauge transformations which change potentials $b \in \Omega$ into transverse potentials can be reduced to a simpler uniqueness problem.

In Sec. II it is shown that existence of charged states of the Yang–Mills field, for the tentative choice of the configuration space Ω, is implied by the existence of solutions of an elliptic boundary-value problem on all of 3-dimensional Euclidean space. The unboundedness of the domain causes a difficulty in solving this existence problem. If the domain were bounded, existence would be assured by a theorem of Schauder[2]; little work appears to have been done on existence theorems for elliptic boundary-value problems on unbounded domains, and the few theorems of this type which we found in the literature do not apply to our system of differential equations. We present here an existence proof which is essentially a modification of Schauder's proof to an unbounded domain, and which goes through on account of the asymptotic properties of the potentials belonging to Ω.

In Sec. 3 the boundary-value problem is rewritten as a functional equation, so that Riesz's determinant-free form of the Fredholm alternatives[3,4] can be

applied. The relevant part of that theorem states that the linear functional equation $f - L[f] = g$ has a solution f in a Banach space $\mathfrak{B}$ for every $g \in \mathfrak{B}$ if $L[f]$ is a completely continuous operator[5] on $\mathfrak{B}$ and if the homogeneous equation $f - L[f] = 0$ has only the null solution. In Sec. 4 it is shown that, indeed, our homogeneous problem has only the null solution, if the spatial holonomy group[6] is simple and compact, as is required for physical reasons,[7] and if the potentials are nondegenerate, i.e., if their holonomy group is the full $SU(2)$. The remaining task is then to show complete continuity of the functional operator $L[f]$ on a judiciously chosen Banach space $\mathfrak{B}$ of functions f. This is carried out in Sec. 5; $\mathfrak{B}$ is chosen to consist of functions f which, besides being bounded and Hölder continuous, also have an asymptotic bound and satisfy asymptotically a certain modified Hölder condition. The norm for $\mathfrak{B}$ is chosen such that it depends adequately on the asymptotic behavior of f. The unboundedness of the domain also requires a modification of the argument showing that the integral expression $\xi[f]$ for the solution of Poisson's equation with Hölder-continuous source f may be differentiated twice under the integral (after a partial integration); this is done by replacing the Weierstrass polynomials used in the approximation procedure[8] by integral transforms employing a Gaussian kernel. General as well as asymptotic estimates are needed for ξ, its first, and its second derivatives utilizing the asymptotic behavior of the potentials; these estimates are found by replacing integrals by larger integrals which can be evaluated.

The existence proof also makes it possible to reduce the problem of existence of equivalent transverse potentials to a simpler uniqueness question, for the tentative choice of Ω. This is carried out in Sec. 6. Boulware[9] has presented an argument purporting to show that any potentials can be gauge transformed to transverse potentials, but he assumes existence of a certain Green's function; of course, this existence remains to be shown. However, Boulware's method which reduces the nonlinear problem to a linear one is valuable and is used here; the resulting linear equations, when thrown in the form of a functional equation, involve a functional operator which is similar to the operator encountered in the existence problem for charged states, and complete continuity of this operator in the same Banach space readily follows. Existence of a solution would follow from nonexistence of a nonnull solution for the associated homogeneous problem.

The tentative choice of the configuration space Ω is reviewed in Sec. 7, using the results found here for the existence problems, and physical considerations. In Sec. 8, the results obtained are applied to the classical constraint equation.

β, γ, and δ range from 1 to 3, and the x^β are Cartesian inertial coordinates in 3-dimensional Euclidean space. α denotes the Hölder exponent. As an argument of a function, $\mathbf{x}$ stands for x^β. The norm $|\mathbf{x}|$ is written as x. ∂_β denotes $\partial/\partial x^\beta$, and the summation convention is used. The real numerical functions $b_\beta{}^i(\mathbf{x})$, $i = 1, 2, 3$, are the spatial Yang–Mills potentials, which express the configuration of the quantized Yang–Mills field in the Schrödinger representation.[1] The group indices i, j, k, l, and m range from 1 to 3, and the structure constants for $SU(2)$ are denoted by $c_{ij}{}^k$. If desired, throughout this paper the isospin group $SU(2)$ may be replaced by any compact semisimple group, by appropriate change of group index range and structure constants. We make use of the covariant derivative

$$\nabla_\beta v^i = \partial_\beta v^i - c_{jk}{}^i b_\beta{}^j v^k, \tag{1.1}$$

where v^i is any vector in the Lie algebra space of $SU(2)$. Group indices are lowered and raised by means of the group metric

$$g_{ij} = c_{ik}{}^l c_{jl}{}^k \tag{1.2}$$

and its inverse. For compact semisimple groups, g_{ij} is *negative* definite. The metric tensor in the Minkowskian event space of special relativity is taken with signature $+ - - -$; this results in a negative-definite spatial part $g_{\beta\gamma}$.

2. CHARGED STATES

In the gauge-invariant quantization[1] of the Yang–Mills field (self-interacting but not coupled to other fields), using the Schrödinger picture and the Schrödinger representation, states are represented by complex-valued functionals $\Psi[b_\beta{}^i(\mathbf{x})]$ of the numerical spatial Yang–Mills potentials $b_\beta{}^i(\mathbf{x})$. The configuration space Ω over which the $\Psi[b_\beta{}^i(\mathbf{x})]$ are defined is *tentatively* taken as the space of all real-valued functions $b_\beta{}^i(\mathbf{x})$ subject to the condition that, for every function $b_\beta{}^i(\mathbf{x})$, a constant B exists, such that

$$\begin{gathered} |b| \leq B, \quad |D_1 b| \leq B, \\ |D_1 b(\mathbf{x}) - D_1 b(\mathbf{y})| \leq B\,|\mathbf{x} - \mathbf{y}|^\alpha, \quad 0 < \alpha < 1, \\ x > R, \quad x^2\,|b| \leq B, \quad x^3\,|D_1 b| \leq B, \\ x^4\,|\partial_\beta b^{\beta i}| \leq B, \\ R < x \leq y, \\ x^4\,|\partial_\beta b^{\beta i}(\mathbf{x}) - \partial_\beta b^{\beta i}(\mathbf{y})| \leq B\,|\mathbf{x} - \mathbf{y}|^\alpha; \end{gathered} \tag{2.1}$$

in (2.1), β and i have been suppressed, and the partial derivative $\partial_\gamma b_\beta{}^i(\mathbf{x})$ for any γ, β, and i has been

written as $D_1 b$. Conditions (2.1) are reviewed in Sec. 7. Ψ needs to be independent of the time component of the Yang–Mills potentials in order to satisfy the primary constraint.[1] States $\Psi[b_\beta]$ are called good or bad as they do or do not satisfy the secondary constraint.[1] Bad states $\Psi[b_\beta]$ are included in the state space for the sake of simplicity of equal-time commutators. Among the good states are the physical states, for which we at least require that Ψ falls off appropriately for large $b_\beta{}^i(\mathbf{x})$ and that at large spatial distances x the $B_{0\beta}{}^i(\mathbf{x})\Psi$ fall off (componentwise) as x^{-2} or faster.

Physical state functionals $\Psi[b_\beta]$ are invariant under *local gauge transformations*, the infinitesimal members of which are changes of the potentials $b_\beta{}^i(\mathbf{x})$ by the amounts

$$\delta b_\beta{}^i(\mathbf{x}) = -\nabla_\beta \eta^i(\mathbf{x}), \tag{2.2}$$

where the $\eta^i(\mathbf{x})$ are smooth real infinitesimal functions, here subject to the conditions at infinity,

$$\eta^i(\mathbf{x}) = O(x^{-1}) \tag{2.3}$$

and

$$\partial_\beta \eta^i = O(x^{-2}), \quad \partial_\beta \partial^\beta \eta^i = O(x^{-4}). \tag{2.4}$$

To show that this invariance statement is true,[1] we calculate the change in $\Psi[b_\beta]$ caused by the infinitesimal transformation (2.2); the result is

$$\begin{aligned}\delta\Psi &= -\int d^3x \frac{\delta\Psi}{\delta b_\beta{}^i} \nabla_\beta \eta^i \\ &= -i\int (d^2x)_\beta B^{0\beta}{}_i \Psi \eta^i + i\int d^3x (\nabla_\beta B^{0\beta}{}_i)\Psi\eta^i. \end{aligned} \tag{2.5}$$

Since Ψ is good, it satisfies the secondary constraint[1]

$$\nabla_\beta B^{0\beta}{}_i \Psi = 0, \tag{2.6}$$

so that the last term in (2.5) vanishes. Since for physical states $B^{0\beta}{}_i(\mathbf{x})\Psi$ is at most of order x^{-2} and $\eta^i(\mathbf{x})$ is restricted by (2.3), the surface term in (2.5) also vanishes. Hence, $\delta\Psi = 0$. Inversely, any state $\Psi[b_\beta]$ for which $B^{0\beta}{}_i(\mathbf{x})\Psi$ is at most of order x^{-2} and which is invariant under local gauge transformations satisfies the secondary constraint, as follows from (2.5) in view of the arbitrariness of $\eta^i(\mathbf{x})$ up to (2.3) and (2.4). The argument given above would go through if, instead of (2.3) and (2.4), one would just have $\eta^i(\mathbf{x}) \to 0$ as $x \to \infty$, but $\eta^i(\mathbf{x}) = O(x^{-1})$ is required in order that δb, given by (2.2), is of order $O(x^{-2})$, which, in turn, is required to keep the local gauge transformations from leading out of Ω. Similarly, the last of conditions (2.4) is required in order that the transformed potentials satisfy the condition $\partial_\beta b^{\beta i} = O(x^{-4})$, which is part of (2.1). A local gauge transformation takes a point b of Ω into the point b'. Applying all possible local gauge transformations to the point b, the transformed points b' form a gauge-invariant manifold χ. Through every point b passes one χ; all the χ's together form a set of equivalence classes in Ω. A physical state functional $\Psi[b]$ has a constant value on any χ, but the value Ψ usually differs from one χ to another.

An *infinitesimal isospin transformation* is a transformation of the type (2.2):

$$\delta b_\beta{}^i = -\nabla_\beta \zeta^i, \tag{2.7}$$

where the $\zeta^i(\mathbf{x})$ are smooth real infinitesimal functions subject to the conditions

$$\zeta^i = \xi^i + c^i, \tag{2.8}$$

$$x \to \infty, \quad \xi^i = O(x^{-1}), \quad \partial_\beta \xi^i = O(x^{-2}),$$
$$\partial_\beta \partial^\beta \xi^i = O(x^{-4}), \tag{2.9}$$

but are otherwise arbitrary. The c^i are infinitesimal constants which do not all vanish. Condition (2.8) and the first condition (2.9) assure that at spatial infinity an isospin rotation occurs; the second and third conditions (2.9) are necessary in order that δb of (2.7) does not lead out of Ω.

If there existed functions $\zeta^i(\mathbf{x})$, subject to (2.8) and (2.9), which produce, on a set of b's with nonvanishing measure, a displacement δb given by (2.7), which leads from one manifold χ to another, then such an isospin transformation would cause a noticeable[10] change in some physical state functional; this would assure existence of physical states with a nonvanishing isospin component. Since there are no $\eta^i(\mathbf{x})$ for which the displacement (2.2) coincides with the displacement (2.7), the latter displacement does not lie in χ. However, this is not enough; since Ω is infinite dimensional, it must be shown that the angle between the isospin displacement (2.7) and χ is *nonzero* on a set of b's of nonzero measure.

Suppose that, on a set $\mathfrak{N} \subset \Omega$ of nonzero measure, there exist infinitesimal functions[11] $\zeta^i(b, \mathbf{x})$, subject to (2.8) and (2.9), which produce a displacement (2.7) orthogonal to $\chi(b)$ in the sense of Ref. 1, i.e., at b,

$$\int d^3x (\nabla^\beta \zeta^i) \delta b_{\beta i} = 0 \tag{2.10}$$

for all $\delta b_{\beta i}(\mathbf{x})$ tangent to $\chi(b)$. If b_0 and b both belong to $\mathfrak{N}$, then one has

$$\zeta^i(b, \mathbf{x}) = \zeta^i(b_0, \mathbf{x}) + \eta^i(b_0, b, \mathbf{x}), \tag{2.11}$$

where $\eta^i(b_0, b, \mathbf{x})$ are functions satisfying (2.3) and (2.4), because the c^i of (2.8) are the same for b as for

b_0. (2.11) implies that, at b,

$$\delta b_\beta{}^i(\mathbf{x}) = -\nabla_\beta \zeta^i(b_0, \mathbf{x}) = -\nabla_\beta \zeta^i(b, \mathbf{x}) + \nabla_\beta \eta^i(b_0, b, \mathbf{x}), \quad (2.12)$$

showing that the displacement δb at b is the sum of a displacement orthogonal to $\chi(b)$ and a displacement along $\chi(b)$. In (2.12), $\delta b_\beta{}^i$, ζ^i, and η^i are infinitesimal. The displacement vector $\delta b_\beta{}^i(\mathbf{x})$ at b makes a nonzero angle with $\chi(b)$ at all $b \in \mathfrak{N}$ except for points b at which

$$\nabla_\beta \zeta^i(b, \mathbf{x}) = 0. \quad (2.13)$$

(2.13) implies that the vector ζ^i is invariant under all equivalence displacements[6] along closed loops in 3-dimensional space. Hence, ζ^i is invariant under the spatial holonomy group.[6,7] This implies that, in the Lie algebra space of $SU(2)$, $\zeta^i L_i$ commutes with the Lie algebra $\mathfrak{L}_s$ of the spatial holonomy group. Since, on account of (2.8) and (2.9), ζ^i cannot vanish for all $\mathbf{x}$, it follows that $\mathfrak{L}_s$ cannot be the whole Lie algebra of $SU(2)$. Hence, if a solution $\zeta^i(b, \mathbf{x})$ exists for (2.13), then the potentials $b_\beta{}^i(\mathbf{x})$ must be *degenerate*, i.e., their spatial holonomy group is not the full $SU(2)$. The set $\mathfrak{D}$ of degenerate potentials in Ω forms a manifold of potentials which are equivalent to potentials $b_\beta{}^i$ which vanish for $i = 2$ and 3, so that the spatial holonomy group is either $SU(1)$ or the identity; hence, the dimension of $\mathfrak{D}$ is $\frac{1}{3}$ of the dimension of Ω, and the set $\mathfrak{D}$ has zero measure in Ω. The intersection of $\mathfrak{D}$ and $\mathfrak{N}$ then also has zero measure, and it follows that the vector $\delta b_\beta{}^i(\mathbf{x})$ of (2.12) makes a nonzero angle with $\chi(b)$ at all $b \in \mathfrak{N}$ except possibly for a set of measure zero. Therefore, *if the set $\mathfrak{N}$ has a nonzero measure in Ω, the Yang–Mills field has physical states with a nonvanishing isospin component.* The $\delta b_{\beta i}$ in (2.10) are tangent to $\chi(b)$, i.e., they must be infinitesimal local gauge transformations (2.2). Inserting (2.2) in (2.10) gives, at b,

$$\int d^3x (\nabla^\beta \zeta^i) \nabla_\beta \eta_i = 0 \quad (2.14)$$

for all $\eta_i(\mathbf{x})$ subject to (2.3) and (2.4). Partial integration of (2.14) and use of Gauss' theorem gives

$$\int (d^2x)_\beta (\nabla^\beta \zeta^i) \eta_i - \int d^3x (\nabla_\beta \nabla^\beta \zeta^i) \eta_i = 0. \quad (2.15)$$

For a spherical surface with infinite radius, the surface integral vanishes on account of (2.3), (2.8), (2.9), and (2.1). The volume integral in (2.15) vanishes for all $\eta_i(\mathbf{x})$ subject to (2.3) and (2.4) if for all $\mathbf{x}$, at b,

$$\nabla_\beta \nabla^\beta \zeta^i = 0, \quad (2.16)$$

with ζ^i subject to (2.8) and (2.9). $\mathfrak{N}$ is the set of b's in Ω for which there exists a solution $\zeta^i(\mathbf{x})$ of this elliptic boundary-value problem. In Secs. 3, 4, and 5 we will prove the following:

Orthogonal Isospin Transformation Theorem: The set $\mathfrak{N}$ of b's for which there exists a unique infinitesimal isospin transformation $\delta b \perp \chi$ includes all non-degenerate potentials $b \in \Omega$.

Since the set $\mathfrak{D}$ of degenerate potentials has zero measure in Ω, the set $\mathfrak{N}$ has nonvanishing measure in Ω if the orthogonal isospin transformation theorem holds; therefore, *if the orthogonal isospin transformation theorem holds, the Yang–Mills field with configuration space Ω has physical states with a nonvanishing isospin component.*

3. EQUIVALENT FUNCTIONAL EQUATION

Writing out the covariant derivatives, we may write the boundary-value problem (2.16), (2.8), and (2.9) as

$$\partial_\beta \partial^\beta \xi^i - 2c_{jk}{}^i b^{\beta j} \partial_\beta \xi^k - c_{jk}{}^i (\partial_\beta b^{\beta j}) \xi^k + c_{jk}{}^i c_{lm}{}^k b_\beta{}^j b^{\beta l} \xi^m = c_{jk}{}^i (\partial_\beta b^{\beta j}) c^k - c_{jk}{}^i c_{lm}{}^k b_\beta{}^j b^{\beta l} c^m, \quad (3.1)$$

$$x \to \infty, \quad \xi^i = O(x^{-1}),$$

$$x \to \infty, \quad D_1 \xi^i = O(x^{-2}), \quad \partial_\beta \partial^\beta \xi^i = O(x^{-4}). \quad (3.2)$$

We will first ignore conditions (3.2) and prove existence of a solution for the problem (3.1); afterwards, we show that this solution satisfies conditions (3.2). We introduce[2] the functions $f^i(\mathbf{x})$ defined by

$$\partial_\beta \partial^\beta \xi^i = f^i. \quad (3.3)$$

Using the condition at infinity for ξ^i, we may write the solution of (3.3) in terms of f^i as

$$\xi^i[f^j] = -\frac{1}{4\pi} \int d^3y f^i(\mathbf{y}) \frac{1}{\rho}, \quad \rho = |\mathbf{x} - \mathbf{y}|. \quad (3.4)$$

With (3.4), the boundary-value problem (3.1) may be expressed as the functional equation[2]

$$f^i - L^i[f^j] = g^i, \quad (3.5)$$

where

$$L^i[f] = 2c_{jk}{}^i b^{\beta j} \partial_\beta \xi^k[f] + c_{jk}{}^i (\partial_\beta b^{\beta j}) \xi^k[f] - c_{jk}{}^i c_{lm}{}^k b^{\beta j} b_\beta{}^l \xi^m[f], \quad (3.6)$$

and

$$g^i = c_{jk}{}^i (\partial_\beta b^{\beta j}) c^k - c_{jk}{}^i c_{lm}{}^k b^{\beta j} b_\beta{}^l c^m. \quad (3.7)$$

We should like to apply to the functional equation (3.5) part of the determinant-free form of the Fredholm alternatives[3,4] which states that, if $L^i[f^j]$ is a completely continuous linear operator[5] on a Banach space

$\mathfrak{B}$ of functions f^j and if the homogeneous problem

$$f^i - L^i[f^j] = 0 \tag{3.8}$$

has only the null solution, then Eq. (3.5) has a unique solution for every $g^i \in \mathfrak{B}$. In preparation, we discuss first our choice of Banach space $\mathfrak{B}$ of functions f^i. Since the space $\mathfrak{B}$ must contain the g^i, we choose $\mathfrak{B}$ as the space of the real multiplet functions $f^i(\mathbf{x})$, subject to the same conditions as $g^i(\mathbf{x})$, i.e., the $f^i(\mathbf{x})$ are bounded, and for every $f^i(\mathbf{x})$ there exist constants H_α, K, and G_α such that

$$\begin{aligned} |f^i(\mathbf{x}) - f^i(\mathbf{y})| &\leq H_\alpha |\mathbf{x} - \mathbf{y}|^\alpha, \quad 0 < \alpha < 1, \\ x^4 |f^i(\mathbf{x})| &\leq K, \quad x > R, \\ x^4 |f^i(\mathbf{x}) - f^i(\mathbf{y})| &\leq G_\alpha |\mathbf{x} - \mathbf{y}|^\alpha, \quad R < x \leq y. \end{aligned} \tag{3.9}$$

H_α, K, and G_α are taken as the smallest values for which (3.9) holds. In fact, one may take

$$\begin{aligned} H_\alpha &= \sup \frac{|f^i(\mathbf{x}) - f^i(\mathbf{y})|}{|\mathbf{x} - \mathbf{y}|^\alpha}, \\ G_\alpha &= \sup_{R<x\leq y} \frac{x^4 |f^i(\mathbf{x}) - f^i(\mathbf{y})|}{|\mathbf{x} - \mathbf{y}|^\alpha}, \\ K &= \sup_{x>R} x^4 |f^i(\mathbf{x})|. \end{aligned} \tag{3.10}$$

As the norm in $\mathfrak{B}$ we take

$$\|f^i\| = \sup |f^i| + H_\alpha[f^i] + K[f^i] + G_\alpha[f^i]. \tag{3.11}$$

The norm (3.11) differs from Schauder's norm[2] by the last two terms, which involve the asymptotic behavior of f^i.

That $\mathfrak{B}$ is indeed a Banach space may be seen as follows. First, it is clear that the norm (3.11) satisfies the three conditions

(a) $\|f^i\| \geq 0$, $\|f^i\| = 0$ only if $f^i = 0$,
(b) $\|cf^i\| = c\,\|f^i\|$ for c constant,
(c) $\|f^i + g^i\| \leq \|f^i\| + \|g^i\|$.

To show completeness of $\mathfrak{B}$, we consider a Cauchy sequence $f_n{}^i(\mathbf{x})$, $n = 1, 2, \cdots$, i.e.,

$$\|f_n{}^i(\mathbf{x}) - f_m{}^i(\mathbf{x})\| < \epsilon,$$

for all n, $m > N(\epsilon)$, for any positive ϵ. This implies with (3.11)

$$|f_n{}^i(\mathbf{x}) - f_m{}^i(\mathbf{x})| < \epsilon, \tag{3.12}$$

$$H_\alpha[f_n{}^i - f_m{}^i] < \epsilon, \tag{3.13}$$

$$K[f_n{}^i - f_m{}^i] < \epsilon, \tag{3.14}$$

$$G_\alpha[f_n{}^i - f_m{}^i] < \epsilon. \tag{3.15}$$

Condition (3.12) for fixed i and $\mathbf{x}$ implies the existence of a limit. Taking this limit for every i and $\mathbf{x}$ defines a function $f^i(\mathbf{x})$. Then,[12] for every $\mathbf{x}$ and i there exists an integer $p(\mathbf{x}, i) > N(\epsilon)$ such that

$$|f^i(\mathbf{x}) - f_p{}^i(\mathbf{x})| < \epsilon;$$

hence

$$\begin{aligned} |f^i(\mathbf{x}) - f_n{}^i(\mathbf{x})| &\leq |f^i(\mathbf{x}) - f_p{}^i(\mathbf{x})| + |f_p{}^i(\mathbf{x}) - f_n{}^i(\mathbf{x})| \\ &< 2\epsilon. \end{aligned}$$

Defining

$$H_n{}^i(\mathbf{x}, \mathbf{y}) = \frac{f_n{}^i(\mathbf{x}) - f_n{}^i(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|^\alpha},$$

we see that (3.13) implies that

$$|H_n{}^i(\mathbf{x}, \mathbf{y}) - H_m{}^i(\mathbf{x}, \mathbf{y})| < \epsilon.$$

Using the same argument as before, but now for values of i, $\mathbf{x}$, and $\mathbf{y}$, one finds

$$\begin{aligned} H^i(\mathbf{x}, \mathbf{y}) &= \lim_{n\to\infty} H_n(\mathbf{x}, \mathbf{y}), \\ |H^i(\mathbf{x}, \mathbf{y}) - H_n{}^i(\mathbf{x}, \mathbf{y})| &< 2\epsilon, \end{aligned}$$

which, on account of (3.10), implies

$$H_\alpha[f^i(\mathbf{x}) - f_n{}^i(\mathbf{x})] < 2\epsilon.$$

Similarly, we find

$$G_\alpha[f^i(\mathbf{x}) - f_n{}^i(\mathbf{x})] < 2\epsilon$$

and

$$K[f^i(\mathbf{x}) - f_n{}^i(\mathbf{x})] < 2\epsilon.$$

Therefore, (3.11) gives

$$\|f^i(\mathbf{x}) - f_n{}^i(\mathbf{x})\| < 8\epsilon, \quad \text{all} \quad n > N(\epsilon),$$

for any positive number ϵ. $f^i(\mathbf{x})$ belongs to $\mathfrak{B}$ because $f^i(\mathbf{x})$ is bounded;

$$H_\alpha[f^i] = \sup |H^i(\mathbf{x}, \mathbf{y})| < \infty,$$

and similar statements hold for $K[f^i]$ and $G_\alpha[f^i]$.

4. THE HOMOGENEOUS PROBLEM

In this section we show that the homogeneous problem (3.8) has only null solutions[13] f^i in $\mathfrak{B}$, for any nondegenerate $b \in \Omega$. This is done by going back to the partial differential equation form of (3.8):

$$\nabla_\beta \nabla^\beta \xi^i = 0, \tag{4.1}$$

$$\xi^i = O(x^{-1}), \quad x \to \infty. \tag{4.2}$$

Transvecting (4.1) with ξ_i and integrating the result over all of the 3-dimensional space gives

$$0 = \int d^3x \xi_i \nabla_\beta \nabla^\beta \xi^i. \tag{4.3}$$

Partial integration of (4.3) and use of Gauss' theorem gives

$$0 = \int (d^2x)_\beta \xi_i \nabla^\beta \xi^i - \int d^3x (\nabla_\beta \xi_i) \nabla^\beta \xi^i. \quad (4.4)$$

From the asymptotic condition (4.2) and (2.1) it follows that the surface integral in (4.4), taken over the infinite sphere, vanishes. For a compact semi-simple group, the group metric is negative definite. Since the spatial metric tensor $g_{\beta\gamma}$ is here also negative definite, it follows that the spatial integral in (4.4) is positive definite, unless

$$\nabla_\beta \xi_i = 0. \quad (4.5)$$

Hence, if a solution of (4.1) and (4.2) exists, then (4.5) must hold. But we have seen in Sec. 2 [discussion following Eq. (2.13)] that existence of a nonnull solution ξ_i of (4.5) for potentials $b_\beta{}^i(\mathbf{x})$ implies that the $b_\beta{}^i(\mathbf{x})$ are degenerate. It follows that for nondegenerate potentials $b \in \Omega$ the problem (4.1), (4.2) has only the null solution.

5. COMPLETE CONTINUITY OF THE FUNCTIONAL OPERATOR $L^i[f^j]$

In this section we prove that $L^i[f^j]$ is a mapping in $\mathfrak{B}$ and that this mapping is completely continuous,[4] i.e., that $L^i[f^j]$ takes any bounded set of $\mathfrak{B}$ into a compact set. The proof is a modification of Schauder's proof[2] to an unbounded domain. For a bounded domain, the main concern is the effect of the singularity of the kernel $1/\rho$ in (3.4); for an unbounded domain, we must in addition avoid divergences or inadequate asymptotic behavior of estimating integrals due to the infinite integration domain, and this makes the procedure of deriving sufficiently sharp estimates more delicate. The unboundedness of the domain makes itself felt in two other places as well. First, on an unbounded domain, boundedness of the first derivatives of a function does not imply Hölder continuity of that function on the whole domain. Second, the calculation of estimates for the second derivatives $D_2\xi$ requires differentiating (3.4) once under the integral, integrating by parts, and differentiating the resulting surface and volume integrals under the integrals. This last step requires that $D_1 f$ is bounded, and this is not implied by the conditions (3.9) for f. A well-known method of dealing with this difficulty is to find functions $f_n(\mathbf{x}), n = 1, 2, \cdots$, which belong to $\mathfrak{B}$, uniformly approximate $f(\mathbf{x})$, and have bounded first derivatives; the second differentiation under the integrals discussed above can be executed for $f_n(\mathbf{x})$, and it is then proved that the difference of $D_2\xi_n$ and $D_2\xi$, where the latter is expressed as the result of *formal* differentiation under the integral, can be made arbitrarily small by choosing n large enough. For a bounded domain, Weierstrass polynomials may be taken[8] for the $f_n(\mathbf{x})$. For an unbounded domain, this does not work; instead, we use as approximating functions

$$f_n(\mathbf{x}) = \left(\frac{n^3}{\pi^3}\right)^{\frac{1}{2}} \int d^3u f(\mathbf{u}) e^{-n|\mathbf{x}-\mathbf{u}|^2}, \quad n = 1, 2, \cdots. \quad (5.1)$$

We start out with a calculation of estimates for $\xi(\mathbf{x})$ and $D_1\xi(\mathbf{x})$, separating the cases $x \leq 2R$ and $x > 2R$.

Theorem 1: For $x \leq 2R$, $|\xi(\mathbf{x})|$ and $|D_1\xi(\mathbf{x})|$ are bounded by a number[14] $A(\|f\|, R)$.

Proof: We write (3.4) as the sum of integrals over the regions $y \leq 3R$ and $y > 3R$, take absolute values, replace in the first integral $|f(\mathbf{y})|$ by $\|f\|$, and in the second integral $|f(\mathbf{y})|$ by K/y^4 and ρ by R. The result is

$$|\xi(\mathbf{x})| \leq \frac{25R^2 \|f\|}{2} + \frac{K}{3R^2}, \quad x \leq 2R.$$

Since f is bounded, $D_1\xi$ may be calculated by differentiating (3.4) under the integral.[15] Using the procedure indicated above, one finds

$$|D_1\xi(\mathbf{x})| \leq 5R \|f\| + \frac{K}{3R^3}, \quad x \leq 2R.$$

Theorem 2: For $x > 2R$, $x|\xi(\mathbf{x})|$ and $x^2|D_1\xi(\mathbf{x})|$ are bounded by a number[14] $A(\|f\|, R)$.

Proof: The integral (3.4) is written as the sum of integrals over the regions $x \leq R$ and $x > R$. After taking absolute values, one replaces in the first integral $|f(\mathbf{y})|$ by $\|f\|$ and ρ by $x - R$, and in the second integral $|f(\mathbf{y})|$ by K/y^4; the result is

$$|\xi(\mathbf{x})| \leq \frac{\|f\| R^3}{3(x-R)} + \frac{K}{4\pi} \int_{y>R} \frac{d^3y}{\rho y^4}, \quad x > 2R. \quad (5.2)$$

If β is the angle between the vectors $\mathbf{x}$ and $\mathbf{y}$, one has

$$\rho^2 = x^2 + y^2 - 2xy \cos\beta, \quad (5.3)$$

and, for fixed x and y,

$$\rho\, d\rho = xy \sin\beta\, d\beta. \quad (5.4)$$

Rotation of the volume element d^3y around $\mathbf{x}$ gives a volume $2\pi y^2\, dy \sin\beta\, d\beta$; using (5.4), one finds from (5.2)

$$|\xi(\mathbf{x})| \leq \frac{\|f\| R^3}{3(x-R)} + \frac{K}{x}\left(\frac{1}{R} - \frac{1}{2x}\right), \quad x > 2R. \quad (5.5)$$

Since f is bounded, $D_1\xi$ may be calculated by differentiating (3.4) under the integral.[15] Using the same procedure as the one indicated above for $|\xi|$, one finds

$$|D_1\xi(\mathbf{x})| \leq \frac{R^3\|f\|}{3(x-R)^2} + \frac{K}{4x^3}\int_{R/x}^{\infty}\frac{d\zeta}{\zeta^3}\ln\left(\frac{1+\zeta}{1-\zeta}\right)^2, \quad x > 2R. \tag{5.6}$$

Inspection of the integral in (5.6) shows that no divergences occur due to the infinite upper integration limit or due to the singularity $\zeta = 1$ of the integrand. For large x/R the integral is bounded by a constant C' times x/R. This gives for (5.6)

$$|D_1\xi(\mathbf{x})| \leq \frac{R^3\|f\|}{3(x-R)^2} + \frac{KC'}{4Rx^2}, \quad x > 2R. \quad \blacksquare \tag{5.7}$$

Before calculating estimates for $D_2\xi$, we need some properties of the functions $f_n(\mathbf{x})$ defined by (5.1).

Lemma 1: If $f(\mathbf{x})$ satisfies α-H and α-G conditions, so does $f_n(\mathbf{x})$, for the same α.

Proof: Writing $\mathbf{v} = \mathbf{y} - \mathbf{x}$, one has

$$f_n(\mathbf{x}) - f_n(\mathbf{y}) = \left(\frac{n^3}{\pi^3}\right)^{\frac{1}{2}}\int d^3u[f(\mathbf{u}) - f(\mathbf{u}+\mathbf{v})]e^{-n|\mathbf{u}-\mathbf{v}|^2}. \tag{5.8}$$

Using the α-H condition on $f(\mathbf{x})$,

$$|f(\mathbf{u}) - f(\mathbf{u}+\mathbf{v})| \leq H_\alpha v^\alpha,$$

one finds from (5.8)

$$|f_n(\mathbf{x}) - f_n(\mathbf{y})| \leq H_\alpha|\mathbf{x}-\mathbf{y}|^\alpha.$$

The Hölder coefficient of $f_n(\mathbf{x})$ is not larger than the Hölder coefficient of $f(\mathbf{x})$. To show that $f_n(\mathbf{x})$ has the α-G property, we take $\mathbf{x}$ and $\mathbf{y}$ such that $2R < x \leq y$, and find from (5.8)

$$|f_n(\mathbf{x}) - f_n(\mathbf{y})| \leq \left(\frac{n^3}{\pi^3}\right)^{\frac{1}{2}}\int d^3uC_\alpha\frac{v^\alpha}{M^4}e^{-n|\mathbf{u}-\mathbf{x}|^2}, \tag{5.9}$$

where

$$C_\alpha = \sup(R^4H_\alpha, G_\alpha)$$

and

$$M = \sup[R, \inf(u, |\mathbf{u}+\mathbf{v}|)]. \tag{5.10}$$

For fixed $\mathbf{x}$ and $\mathbf{y}$ the space of integration in (5.9) falls into two half-spaces a and b in which, respectively, $u < |\mathbf{u}+\mathbf{v}|$ and $u > |\mathbf{u}+\mathbf{v}|$. We have

$$\text{in } a, \qquad M = M_a = \sup(R, u),$$
$$\text{in } b, \qquad M = M_b = \sup(R, |\mathbf{u}+\mathbf{v}|). \tag{5.11}$$

The integral (5.9) over a and over b may be replaced by a larger number, which results from taking the integral (5.9) over all of space, using M_a, plus the integral (5.9) over all of space, using M_b:

$$|f_n(\mathbf{x}) - f_n(\mathbf{y})| \leq I_1 + I_2,$$
$$I_1 = \left(\frac{n^3}{\pi^3}\right)^{\frac{1}{2}}C_\alpha v^\alpha\int\frac{d^3u}{M_a^4}e^{-n|\mathbf{u}-\mathbf{x}|^2}, \tag{5.12}$$
$$I_2 = \left(\frac{n^3}{\pi^3}\right)^{\frac{1}{2}}C_\alpha v^\alpha\int\frac{d^3u}{M_b^4}e^{-n|\mathbf{u}-\mathbf{x}|^2}.$$

Calling $|\mathbf{u}-\mathbf{x}| = \rho$ and using a relation like (5.4), one finds

$$I_1 = C_\alpha\frac{v^\alpha}{x}\left(\frac{n}{\pi}\right)^{\frac{1}{2}}\left(\frac{1}{R^4}\int_0^R u\,du(e^{-n(x+u)^2} - e^{-n(x-u)^2}) + \int_R^\infty\frac{du}{u^3}(e^{-n(x+u)^2} - e^{-n(x-u)^2})\right). \tag{5.13}$$

Upper bounds for the integrals in (5.13) are found by making various replacements in the integrands, such as replacing in the first integral $(x+u)^2$ by x^2 and $(x-u)^2$ by $(x-R)^2$ and by replacing, in the first part of the last integral, $1/u^3$ by $(u+x)/R^4$, executing a number of partial integrations, and making changes in the regions of integration. The result found is

$$I_1 \leq \frac{C_{1\alpha}}{x^4}v^\alpha, \tag{5.14}$$

where the constant $C_{1\alpha}$ depends only on $\|f\|$ and R. The difference between I_1 and I_2 is due to the function M, which is M_a and M_b, respectively, for I_1 and I_2. Using this fact, it is easy to show that I_2 is obtained by changing, in I_1, the $\mathbf{x}$ into $\mathbf{y}$; hence, by (5.14), one must have

$$I_2 \leq \frac{C_{1\alpha}v^\alpha}{y^4} \leq \frac{C_{1\alpha}v^\alpha}{x^4}, \tag{5.15}$$

since $x \leq y$.

Lemma 2: If $|f(\mathbf{x})| \leq A$ for all $\mathbf{x}$ and if $|f(\mathbf{x})| \leq K/x^4$ for $x > R$, then $f_n(\mathbf{x})$ also satisfies these conditions, for the same A and for a K' which depends only on A, K, and R.

Proof: Taking the absolute value of (5.1) and replacing $|f(\mathbf{u})|$ by A, one obtains $|f(\mathbf{x})| \leq A$. To prove the second part of the theorem, $|f(\mathbf{x})|$ for $x > 2R$ is written as the sum of an integral over $u \leq R$ and an integral over $u > R$. In the first integral, $|f(\mathbf{x})|$ is replaced by A and in the second integral $|f(\mathbf{x})|$ is replaced by K/u^4.

As a result one finds

$$|f_n(\mathbf{x})| \leq \tfrac{4}{3}\pi R^3 A\left(\frac{n^3}{\pi^3}\right)^{\frac{1}{2}} e^{-n(x-R)^2} + K\left(\frac{n^3}{\pi^3}\right)^{\frac{1}{2}} \int_{u>R} \frac{d^3u}{u^4}\, e^{-n|\mathbf{u}-\mathbf{x}|^2}, \quad x > 2R. \tag{5.16}$$

The first term on the right-hand side of (5.16) is at most $C_1/x^4(n)^{\frac{1}{2}}$, where the constant C_1 depends only on A and R. The second term on the right has already been estimated in the proof of Lemma 1, with the result

$$\left(\frac{n^3}{\pi^3}\right)^{\frac{1}{2}} \int_{u>R} \frac{d^3u}{u^4}\, e^{-n|\mathbf{u}-\mathbf{x}|^2} \leq \frac{C_2}{x^4},$$

where the constant C_2 depends only on R. It follows that $|f(\mathbf{x})| \leq K'(A, R, K)/x^4$ for $x > 2R$. Since $f(\mathbf{x})$ is bounded by A, this inequality holds also for $x > R$ with a different constant K', which depends only on A, R, and K.

Lemma 3: For $f(\mathbf{x}) \in \mathfrak{B}$, $|D_1 f_n(\mathbf{x})| \leq 2(n/\pi)^{\frac{1}{2}} \|f\|$.

Proof: Differentiating (5.1) under the integral and replacing $|f(\mathbf{u})|$ by $\|f\|$ and $|u_\alpha - x_\alpha|$ by $\rho = |\mathbf{u} - \mathbf{x}|$, one finds

$$|D_1 f_n(\mathbf{x})| \leq 4\pi n\left(\frac{n^3}{\pi^3}\right)^{\frac{1}{2}} \|f\| \int_0^\infty d\rho \rho^3 e^{-n\rho^2} = 2\left(\frac{n}{\pi}\right)^{\frac{1}{2}} \|f\|.$$

Lemma 4: If $f(\mathbf{x}) \in \mathfrak{B}$, one has, for all $\mathbf{x}$,

$$|f_n(\mathbf{x}) - f(\mathbf{x})| \leq C_\alpha / n^{\frac{1}{2}\alpha},$$

where the constant C_α only depends on H_α, and, for $x > R$, $|f_n(\mathbf{x}) - f(\mathbf{x})| \leq A/n^{\frac{1}{2}\alpha} x^4$, where the constant A depends only on $\|f\|$ and R.

Proof: For any $\mathbf{x}$,

$$|f_n(\mathbf{x}) - f(\mathbf{x})| \leq \left(\frac{n^3}{\pi^3}\right)^{\frac{1}{2}} \int d^3u\, |f(\mathbf{x}) - f(\mathbf{u})|\, e^{-n\rho^2}, \tag{5.17}$$

where $\rho = |\mathbf{u} - \mathbf{x}|$. Replacing $|f(\mathbf{u}) - f(\mathbf{x})|$ by $H_\alpha \rho^\alpha$, one finds

$$|f_n(\mathbf{x}) - f(\mathbf{x})| \leq 2H_\alpha\left(\frac{n}{\pi}\right)^{\frac{1}{2}} \int_0^\infty \rho^{1+\alpha}\, d e^{-n\rho^2} = 2H_\alpha(1+\alpha)\left(\frac{n}{\pi}\right)^{\frac{1}{2}} \int_0^\infty d\rho \rho^\alpha e^{-n\rho^2}.$$

Replacing ρ^α by 1 for $0 \leq \rho \leq 1$ and by ρ for $\rho > 1$, one finds

$$|f_n(\mathbf{x}) - f(\mathbf{x})| \leq H_\alpha(1+\alpha)[1 + (2/e\pi)^{\frac{1}{2}}]/(2n)^{\frac{1}{2}\alpha}. \tag{5.18}$$

For $x \geq 2R$, one has from (5.17) and (3.9)

$$|f_n(\mathbf{x}) - f(\mathbf{x})| \leq E_\alpha\left(\frac{n^3}{\pi^3}\right)^{\frac{1}{2}} \int d^3u\, \frac{\rho^\alpha}{P^4}\, e^{-n\rho^2}, \tag{5.19}$$

where

$$E_\alpha = \sup\,(R^4 H_\alpha, G_\alpha)$$

and

$$P = \sup\,[R, \inf\,(u, x)].$$

Splitting the integration region for $\mathbf{u}$ into the parts $u \leq R$, $R < u \leq \frac{1}{2}x$, and $u > \frac{1}{2}x$ and using, in the first two integrals, polar coordinates as shown by (5.3) and (5.4), we may write (5.19) as

$$|f_n(\mathbf{x}) - f(\mathbf{x})| \leq I_1 + I_2 + I_3,$$

$$I_1 = \frac{2E_\alpha}{x}\left(\frac{n^3}{\pi}\right)^{\frac{1}{2}} \frac{1}{R^4} \int_0^R du u \int_{x-u}^{x+u} d\rho \rho^{1+\alpha} e^{-n\rho^2},$$

$$I_2 = \frac{2E_\alpha}{x}\left(\frac{n^3}{\pi}\right)^{\frac{1}{2}} \int_R^{\frac{1}{2}x} \frac{du}{u^3} \int_{x-u}^{x+u} d\rho \rho^{1+\alpha} e^{-n\rho^2},$$

$$I_3 = E_\alpha\left(\frac{n^3}{\pi^3}\right)^{\frac{1}{2}} \int_{u>\frac{1}{2}x} d^3u\, \frac{\rho^\alpha}{P^4}\, e^{-n\rho^2}.$$

I_1 is estimated by replacing ρ^α by $R^\alpha(\rho/R)^2$, executing the integration over ρ, replacing $\exp\,[-n(x+u)^2]$ by $\exp\,(-nx^2)$, and replacing $\exp\,[-n(x-u)^2]$ by $\exp\,[-n(x-R)^2]$; the result is

$$I_1 \leq (n)^{\frac{1}{2}}(e^{-nx^2} + e^{-n(x-R)^2})O(x) \leq C_3/n^{\frac{1}{2}}x^4, \quad x > 2R,$$

where C_3 depends only on $\|f\|$ and R. Replacing, in I_2, $\exp\,(-n\rho^2)$ by $\exp\,(-\frac{1}{4}nx^2)$ gives

$$I_2 \leq E_\alpha(n^3)^{\frac{1}{2}} e^{-\frac{1}{4}nx^2} O(x^{\alpha-1}) \leq \frac{C_4(\|f\|, R)}{x^4(n)^{\frac{1}{2}}}.$$

Replacing, in I_3, P by $\frac{1}{2}x$ gives $I_3 \leq C_5(\|f\|, R)/x^4 n^{\frac{1}{2}\alpha}$. Hence,

$$|f_n(\mathbf{x}) - f(\mathbf{x})| \leq \frac{A(\|f\|, R)}{x^4 n^{\frac{1}{2}\alpha}}, \quad \text{for} \quad x \geq 2R; \tag{5.20}$$

on account of (5.18), the inequality (5.20) also holds for $x > R$, with a different coefficient $A(\|f\|, R)$. ∎

Since $f \in \mathfrak{B}$ does not necessarily have a bounded derivative, the second derivatives $D_2\xi$ cannot be obtained by differentiating $D_1\xi$ under the integral without further argument. The standard method of coping with this situation[8] is to approximate f by differentiable functions f_n, for which the above-mentioned procedure is allowed and leads to $D_2\xi[f_n]$, and to show that the difference $|D_2\xi[f_n] - D_2\xi[f]|$ (where $D_2\xi[f]$ is the same expression in f as $D_2\xi[f_n]$

is in f_n) can be made arbitrarily small by choosing n large enough. It turns out that the functions $f_n(\mathbf{x})$ defined by (5.1) are adequate for this purpose, if the expression for $D_2\xi$ is derived by a judicious split of the integration region in the expression for $D_1\xi$, and partial integrations. The next two theorems deal with this matter.

Theorem 3: For $x \leq 2R$,

$$\partial_\gamma\partial_\beta\xi = -\frac{1}{4\pi}(I_1 + I_2 + I_3 + I_4),$$

where

$$\begin{aligned} I_1 &= \int_{y=3R} (d^2y)_\beta [f(\mathbf{y}) - f(\mathbf{x})]\partial_\gamma \frac{1}{\rho}, \\ I_2 &= -\int_{y=3R} (d^2y)_\beta f(\mathbf{y})\partial_\gamma \frac{1}{\rho}, \\ I_3 &= \int_{y\leq 3R} d^3y[f(\mathbf{y}) - f(\mathbf{x})]\partial_\beta\partial_\gamma \frac{1}{\rho}, \\ I_4 &= \int_{y>3R} d^3y f(\mathbf{y})\partial_\beta\partial_\gamma \frac{1}{\rho}. \end{aligned} \tag{5.21}$$

Proof: By Lemma 3, $D_1 f_n(\mathbf{x})$ is bounded. Therefore, one may take

$$\begin{aligned} -4\pi\partial_\gamma\partial_\beta\xi[f_n] &= \int d^3y \left(\frac{\partial}{\partial y^\beta} f_n(\mathbf{y})\right)\partial_\gamma \frac{1}{\rho} \\ &= \int_{y\leq 3R} d^3y \left(\frac{\partial}{\partial y^\beta}(f_n(\mathbf{y}) - f_n(\mathbf{x}))\right)\partial_\gamma \frac{1}{\rho} \\ &\quad + \int_{y>3R} d^3y \left(\frac{\partial}{\partial y^\beta} f_n(\mathbf{y})\right)\partial_\gamma \frac{1}{\rho} \\ &= \int_{y=3R} (d^2y)_\beta [f_n(\mathbf{y}) - f_n(\mathbf{x})]\partial_\gamma \frac{1}{\rho} \\ &\quad + \int_{y\leq 3R} d^3y [f_n(\mathbf{y}) - f_n(\mathbf{x})]\partial_\beta\partial_\gamma \frac{1}{\rho} \\ &\quad - \int_{y=3R} (d^2y)_\beta f_n(\mathbf{y})\partial_\gamma \frac{1}{\rho} \\ &\quad + \int_{y>3R} d^3y f_n(\mathbf{y})\partial_\beta\partial_\gamma \frac{1}{\rho}. \end{aligned} \tag{5.22}$$

Denoting (5.22) by $D_2\xi[f_n]$ and (5.21) by $D_2\xi[f]$, we will prove that $D_2\xi[f_n] - D_2\xi[f]$ can be made arbitrarily small by choosing n large enough. We have

$$\begin{aligned} &4\pi\,|D_2\xi[f_n] - D_2\xi[f]| \\ &\leq \int_{y=3R} (d^2y)_\beta\, |f_n(\mathbf{y}) - f(\mathbf{y}) - f_n(\mathbf{x}) + f(\mathbf{x})| \frac{1}{\rho^2} \\ &\quad + \int_{y=3R} (d^2y)_\beta\, |f_n(\mathbf{y}) - f(\mathbf{y})| \frac{1}{\rho^2} \\ &\quad + 2\int_{y\leq 3R} d^3y\, |f_n(\mathbf{y}) - f(\mathbf{y}) - f_n(\mathbf{x}) + f(\mathbf{x})| \frac{1}{\rho^3} \\ &\quad + 2\int_{y>3R} d^3y\, |f_n(\mathbf{y}) - f(\mathbf{y})| \frac{1}{\rho^3}. \end{aligned} \tag{5.23}$$

The surface integrals in (5.23) can be made arbitrarily small, by Lemma 4, the boundedness of $1/\rho$, and the finiteness of the integration area. Replacing, in the last volume integral of (5.23), $1/\rho^3$ by $1/R^3$ and using Lemma 4, we have

$$2\int_{y>3R} d^3y\, |f_n(\mathbf{y}) - f(\mathbf{y})| \frac{1}{\rho^3} \leq \frac{4\pi A}{R^3 n^{\frac{1}{2}\alpha}} \int_{3R}^{\infty} \frac{dy}{y^2} \leq \frac{4\pi A}{3R^4 n^{\frac{1}{2}\alpha}}.$$

For the first volume integral in (5.23) the integration volume is split[16] into the interior of a small sphere with radius R' centered at $\mathbf{x}$ and the remaining volume. Using the α-H continuity of $f(\mathbf{y})$ and of $f_n(\mathbf{y})$ (Lemma 1) in the first integral, and using Lemma 4 and $1/\rho^3 \leq 1/(R')^3$ in the second integral, one finds

$$\begin{aligned} &\int_{y\leq 3R} d^3y\, |f_n(\mathbf{y}) - f(\mathbf{y}) - f_n(\mathbf{x}) + f(\mathbf{x})| \frac{1}{\rho^3} \\ &\qquad \leq 8\pi H_\alpha \int_0^{R'} d\rho\rho^{\alpha-1} + \frac{2C_\alpha}{(R')^3 n^{\frac{1}{2}\alpha}} \int_{y\leq 3R} d^3y \\ &\qquad = \frac{8\pi H_\alpha}{\alpha}(R')^\alpha + \frac{72\pi C_\alpha R^3}{(R')^3 n^{\frac{1}{2}\alpha}}. \end{aligned} \tag{5.24}$$

The first term on the right-hand side of (5.24) may be made as small as desired, by choosing R' small enough; with that value of R', the second term on the right-hand side of (5.24) can be made as small as desired by choosing n large enough.

Theorem 4: For $x > 2R$ and $a = \frac{1}{2}R$,

$$-4\pi\partial_\gamma\partial_\beta\xi[f] = J_1 + J_2 + J_3 + J_4 + J_5 + J_6 + J_7, \tag{5.25a}$$

where

$$J_1 = \int_{y=x-a} (d^2y)_\beta f(\mathbf{y})\partial_\gamma \frac{1}{\rho}, \tag{5.25b}$$

$$J_2 = -\int_{y=x-a} (d^2y)_\beta [f(\mathbf{y}) - f(\mathbf{x})]\partial_\gamma \frac{1}{\rho}, \tag{5.25c}$$

$$J_3 = \int_{y=x+a} (d^2y)_\beta [f(\mathbf{y}) - f(\mathbf{x})]\partial_\gamma \frac{1}{\rho}, \tag{5.25d}$$

$$J_4 = -\int_{y=x+a} (d^2y)_\beta f(\mathbf{y})\partial_\gamma \frac{1}{\rho}, \tag{5.25e}$$

$$J_5 = \int_{y\leq x-a} d^3y f(\mathbf{y})\partial_\beta\partial_\gamma \frac{1}{\rho}, \tag{5.25f}$$

$$J_6 = \int_{x-a\leq y\leq x+a} d^3y [f(\mathbf{y}) - f(\mathbf{x})]\partial_\beta\partial_\gamma \frac{1}{\rho}, \tag{5.25g}$$

$$J_7 = \int_{y\geq x+a} d^3y f(\mathbf{y})\partial_\beta\partial_\gamma \frac{1}{\rho}. \tag{5.25h}$$

Proof: Since $D_1 f_n(\mathbf{x})$ is bounded, $D_2\xi[f_n]$ may be obtained by differentiating $D_1\xi[f_n]$ under the integral:

$$-4\pi\partial_\gamma\partial_\beta\xi[f_n] = \int d^3y \left(\frac{\partial}{\partial y^\beta} f(\mathbf{y})\right)\partial_\gamma \frac{1}{\rho}. \quad (5.26)$$

Expressions (5.25), written in terms of $f_n(\mathbf{x})$ instead of $f(\mathbf{x})$, follow from (5.26) by splitting the integration region into the parts $y \le x - a$, $x - a < y \le x + a$, and $y > x + a$, by replacing, in the second region, $(\partial/\partial y^\beta) f(\mathbf{y})$ by $(\partial/\partial y^\beta)[f(\mathbf{y}) - f(\mathbf{x})]$ and by executing partial integrations. It will be shown that, for $x > 2R$, $x^3\,|D_2\xi[f_n(\mathbf{x})] - D_2\xi[f(\mathbf{x})]|$ can be made arbitrarily small by choosing n large enough. Using the expression (5.25) for $D_2\xi[f]$ and the statement (5.25) modified by replacing f by f_n, one finds[17]

$$4\pi\,|D_2\xi[f_n] - D_2\xi[f]| \le I_1 + I_2 + I_3 + I_4 + I_5 + I_6 + I_7, \quad (5.27a)$$

where

$$I_1 = \int_{y=x-a} d^2y\,|f_n(\mathbf{y}) - f(\mathbf{y})|\frac{1}{\rho^2}, \quad (5.27b)$$

$$I_2 = \int_{y=x-a} d^2y\,|f_n(\mathbf{y}) - f(\mathbf{y}) - f_n(\mathbf{x}) + f(\mathbf{x})|\frac{1}{\rho^2}, \quad (5.27c)$$

$$I_3 = \int_{y=x+a} d^2y\,|f_n(\mathbf{y}) - f(\mathbf{y}) - f_n(\mathbf{x}) + f(\mathbf{x})|\frac{1}{\rho^2}, \quad (5.27d)$$

$$I_4 = \int_{y=x+a} d^2y\,|f_n(\mathbf{y}) - f(\mathbf{y})|\frac{1}{\rho^2}, \quad (5.27e)$$

$$I_5 = 2\int_{y\le x-a} d^3y\,|f_n(\mathbf{y}) - f(\mathbf{y})|\frac{1}{\rho^3}, \quad (5.27f)$$

$$I_6 = 2\int_{x-a\le y\le x+a} d^3y\,|f_n(\mathbf{y}) - f(\mathbf{y}) - f_n(\mathbf{x}) + f(\mathbf{x})|\frac{1}{\rho^3}, \quad (5.27g)$$

$$I_7 = 2\int_{y\ge x+a} d^3y\,|f_n(\mathbf{y}) - f(\mathbf{y})|\frac{1}{\rho^3}. \quad (5.27h)$$

I_1 is estimated by using the second part of Lemma 4 and (5.4); the result is

$$I_1 \le \frac{2\pi A(\|f\|, R)}{n^{\frac{1}{2}\alpha}x(x-a)^3}\ln\left(\frac{2x-a}{a}\right) \le \frac{A_1(\|f\|, R)}{n^{\frac{1}{2}\alpha}x^3}. \quad (5.28)$$

A similar procedure gives the result (5.28) for I_2, I_3, and I_4. I_5 is written as the sum of an integral I_{51} over the interior of the sphere with radius R and remainder I_{52}. By Lemma 4, we have $I_{51} \le 8\pi R^3 C_\alpha/3n^{\frac{1}{2}\alpha}(x-R)^3$. Again by Lemma 4, one has for the remainder I_{52}

$$I_{52} \le \frac{2A}{n^{\frac{1}{2}\alpha}}\int_{R\le y\le x-a} d^3y\,\frac{1}{y^4\rho^3};$$

using polar coordinates and (5.4), we find

$$I_{52} \le \frac{8\pi A}{n^{\frac{1}{2}\alpha}x^3}\left(\frac{1}{y} + \frac{1}{2x}\ln(x+y) - \frac{1}{2x}\ln(x-y)\right)\Bigg|_{y=R}^{y=x-a} \le \frac{A_2(\|f\|, R)}{n^{\frac{1}{2}\alpha}x^3}.$$

I_6 is written as an integral I_{61} over a sphere with radius $R' < a$ centered at $\mathbf{x}$ and a remainder I_{62}. Using the $\alpha - G$ continuity of $f(\mathbf{y})$ and of $f_n(\mathbf{y})$ (Lemma 1), one finds

$$x^3 I_{61} \le \frac{A_3(\|f\|, R)}{\alpha}(R')^\alpha. \quad (5.29)$$

An estimate for the remainder I_{62} is obtained by using the second part of Lemma 4 and polar coordinates centered at $\mathbf{x}$; the result is

$$x^3 I_{62} \le \frac{A_4(\|f\|, R)}{n^{\frac{1}{2}\alpha}x}\left(\ln\frac{a}{R'} + 1 - \frac{a}{2x+a}\right). \quad (5.30)$$

$x^3 I_{61}$ of (5.29) can be made arbitrarily small by choosing R' small enough. For that value of R', $x^3 I_{62}$ of (5.30) can be made arbitrarily small by choosing n large enough. Hence, $x^3 I_6 = x^3 I_{61} + x^3 I_{62}$ can be made arbitrarily small by choosing n large enough. I_7 is estimated by using the second part of Lemma 4 and polar coordinates centered at the origin; the result is

$$I_7 \le \frac{8\pi A}{n^{\frac{1}{2}\alpha}x^4}\int_{1-a/x}^{\infty}\frac{d\zeta}{\zeta^3(\zeta-1)(\zeta+1)}. \quad (5.31)$$

The integral in (5.31) is convergent and is smaller than a constant times $\ln(a/x)$. Hence, we have

$$I_7 \le \frac{A_5(\|f\|, R)}{n^{\frac{1}{2}\alpha}x^3}. \quad \blacksquare$$

After this preparation, we can proceed to estimate $D_2\xi$ for $x \le 2R$ and for $x > 2R$.

Theorem 5: For $x \le 2R$, $|D_2\xi| \le A(\|f\|, R)$.

Proof: Using (5.21), we have the following estimates. Replacing, in $|I_1|$, $|f(\mathbf{x})|$ and $|f(\mathbf{y})|$ by $\|f\|$ and ρ by R, one finds

$$|I_1| \le 36\pi\,\|f\|;$$

a similar result holds for $|I_2|$. Using the Hölder continuity of $f(\mathbf{x})$ in $|I_3|$ gives $|I_3| \le 4\pi H_\alpha(5R)^\alpha/\alpha$.

Replacing, in $|I_4|$, $|f(\mathbf{y})|$ by K/y^4 and ρ by R gives

$$|I_4| \leq 4\pi K/3R^2 .$$

Theorem 6: For $x > 2R$, $x^3\,|D_2\xi(\mathbf{x})| \leq A(\|f\|, R)$.

Proof: In the proof of Theorem 4 we have shown that

$$x^3\,|D_2\xi[f_n(\mathbf{x})] - D_2\xi[f(\mathbf{x})]|$$

can be made arbitrarily small by choosing n large enough; since the expression for $D_2\xi[f_n]$ is valid, the expression for $D_2\xi[f]$ given by Theorem 4 may be used for the purpose of proving the present theorem. In the surface integrals of (5.25), $|f(\mathbf{y})|$ is replaced by $K/(x-a)^4$, and polar coordinates are introduced. As a result one finds that $|J_1|$, $|J_2|$, $|J_3|$, and $|J_4|$ are smaller than $A_1(\|f\|, R)/x^3$. The integral J_5 is written as the sum of an integral J_{51} over the interior of the sphere with radius R, and a remainder J_{52}. By replacing, in $|J_{51}|$, $|f(\mathbf{y})|$ by $\|f\|$, and ρ by $x - R$, one finds

$$|J_{51}| \leq 4\pi R^3\,\|f\|/3(x-R)^3.$$

The remaining integral $|J_{52}|$ is estimated by replacing $|f(\mathbf{y})|$ by K/y^4 and using polar coordinates; the result is

$$|J_{52}| \leq A_2(\|f\|, R)/x^3.$$

J_6 is estimated by using the $\alpha - G$ property of $f(\mathbf{x})$ and replacing the integration domain by the region $y \leq 2x + a$; the result is

$$|I_6| \leq A_3(\|f\|, R)/\alpha x^3.$$

Replacing, in $|J_7|$, $|f(\mathbf{y})|$ by K/y^4, and using polar coordinates, we find

$$|J_7| \leq \frac{4\pi K}{x^4}\int_{1-a/x}^{\infty} \frac{d\zeta}{\zeta^3(\zeta-1)(\zeta+1)}. \qquad (5.32)$$

The integral in (5.23) is convergent and is smaller than a constant times $\ln(a/x)$. Hence, we have

$$|J_7| \leq A_4(\|f\|, R)/x^3. \quad \blacksquare$$

Equipped with the estimates for ξ, $D_1\xi$, and $D_2\xi$, we can now proceed to show that the operator L^i maps $\mathfrak{B}$ into $\mathfrak{B}$. This is done in the following theorems.

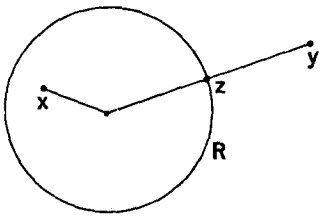


FIG. 1. Choice of point z for the case $x \leq R$.

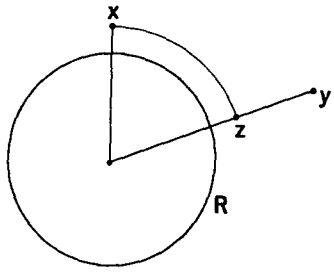


FIG. 2. Choice of point z for the case $x > R$.

Theorem 7: $\xi[f(\mathbf{x})]$ is α-Hölder continuous.

Proof: We will show that

$$|\xi(\mathbf{x}) - \xi(\mathbf{y})| \leq A(\|f\|, R)\,|\mathbf{x} - \mathbf{y}|^{\alpha}$$

separately for the three cases: (1) $x \leq R$, $y \leq R$, (2) $x \leq R$, $y > R$, and (3) $x > R$, $y > R$. (1) Using Theorems 1 and 2, we have

$$\begin{aligned}|\xi(\mathbf{x}) - \xi(\mathbf{y})| &\leq |\mathbf{x} - \mathbf{y}|\sup|D_1\xi| \\ &\leq A\,|\mathbf{x} - \mathbf{y}|^{1-\alpha}\,|\mathbf{x} - \mathbf{y}|^{\alpha} \\ &\leq A(2R)^{\alpha}\,|\mathbf{x} - \mathbf{y}|^{\alpha} \quad \text{for} \quad x \leq R,\ y \leq R.\end{aligned}$$

(2) For $x \leq R$, $y > R$, let $\mathbf{z}$ be the intersection of the straight line through $\mathbf{y}$ and the origin, with the sphere R (Fig. 1). Using Theorems 1 and 2, one has

$$\begin{aligned}|\xi(\mathbf{x}) - \xi(\mathbf{y})| &\leq |\xi(\mathbf{x}) - \xi(\mathbf{z})| + |\xi(\mathbf{z}) - \xi(\mathbf{y})| \\ &\leq A\,|\mathbf{x} - \mathbf{z}| + A\int_z^y \frac{dr}{r^2} \\ &= A\,|\mathbf{x} - \mathbf{z}| + \frac{A\,|\mathbf{y} - \mathbf{z}|}{yR} \\ &\leq A\,|\mathbf{x} - \mathbf{z}|^{\alpha}(2R)^{1-\alpha} + \frac{A\,|\mathbf{y} - \mathbf{z}|^{\alpha}}{R^{1+\alpha}},\end{aligned} \qquad (5.33)$$

since $|\mathbf{x} - \mathbf{z}| \leq 2R$ and $|\mathbf{y} - \mathbf{z}| \leq y$. Because $|\mathbf{x} - \mathbf{z}| \leq |\mathbf{x} - \mathbf{y}|$ and $|\mathbf{y} - \mathbf{z}| \leq |\mathbf{x} - \mathbf{y}|$, (5.33) implies

$$|\xi(\mathbf{x}) - \xi(\mathbf{y})| \leq A_1(\|f\|, R, \alpha)\,|\mathbf{x} - \mathbf{y}|^{\alpha}.$$

(3) For $R < x \leq y$, let $\mathbf{z}$ be the intersection of the straight line through $\mathbf{y}$ and the origin, with the sphere of radius x (Fig. 2). Writing $\xi(\mathbf{x}) - \xi(\mathbf{y})$ as the line integral of $D_1\xi$ along the circular arc from $\mathbf{x}$ to $\mathbf{z}$ and along the straight line from $\mathbf{z}$ to $\mathbf{y}$ and using Theorem 2, one has

$$\begin{aligned}|\xi(\mathbf{x}) - \xi(\mathbf{y})| &\leq \frac{\pi A}{2x^2}\,|\mathbf{x} - \mathbf{z}| + \frac{A\,|\mathbf{y} - \mathbf{z}|}{xy} \\ &= \frac{\pi A}{2x^2}\,|\mathbf{x} - \mathbf{z}|^{\alpha}(2x)^{1-\alpha} + \frac{A}{x^{1+\alpha}}\,|\mathbf{y} - \mathbf{z}|^{\alpha},\end{aligned} \qquad (5.34)$$

since $|\mathbf{x} - \mathbf{z}| \leq 2x$ and $|\mathbf{y} - \mathbf{z}| \leq y$. Because

$$|\mathbf{x} - \mathbf{z}| \leq |\mathbf{x} - \mathbf{y}| \quad \text{and} \quad |\mathbf{y} - \mathbf{z}| \leq |\mathbf{x} - \mathbf{y}|,$$

(5.34) implies

$$|\xi(\mathbf{x}) - \xi(\mathbf{y})| \leq A_2(\|f\|, R, \alpha)\,|\mathbf{x} - \mathbf{y}|^\alpha.$$

Theorem 8: $D_1\xi[f(\mathbf{x})]$ is α-Hölder continuous, and

$$x^3\,|D_1\xi(\mathbf{x}) - D_1\xi(\mathbf{y})| \leq A(\|f\|, R, \alpha)\,|\mathbf{x} - \mathbf{y}|^\alpha, \qquad R < x \leq y.$$

Proof: (1) $x \leq R$, $y \leq R$; using Theorem 5, we have

$$|D_1\xi(\mathbf{x}) - D_1\xi(\mathbf{y})| \leq |\mathbf{x} - \mathbf{y}| \sup |D_2\xi| \leq A\,|\mathbf{x} - \mathbf{y}| \leq (2R)^{1-\alpha}A\,|\mathbf{x} - \mathbf{y}|^\alpha.$$

(2) $x \leq R$, $y > R$; let $\mathbf{z}$ be the point shown in Fig. 1. Using Theorem 5 and 6, we have

$$|D_1\xi(\mathbf{x}) - D_1\xi(\mathbf{y})| \leq A\,|\mathbf{x} - \mathbf{z}| + A\int_R^y \frac{dr}{r^3} \leq A\,|\mathbf{x} - \mathbf{z}| + \frac{A}{yR^2}\,|\mathbf{z} - \mathbf{y}| \leq A_1(\|f\|, R, \alpha)\,|\mathbf{x} - \mathbf{z}|^\alpha.$$

(3) $R < x \leq y$; let $\mathbf{z}$ be the point shown in Fig. 2. Using Theorem 6, we have

$$|D_1\xi(\mathbf{x}) - D_1\xi(\mathbf{y})| \leq \frac{\pi A}{2x^3}\,|\mathbf{x} - \mathbf{z}| + A\int_z^y \frac{dr}{r^3} \leq \frac{\pi A(2x)^\alpha}{2x^3}\,|\mathbf{x} - \mathbf{z}|^\alpha + \frac{A}{y^\alpha x^2}\,|\mathbf{y} - \mathbf{z}|^\alpha \leq \frac{A_2(\|f\|, R, \alpha)}{x^2}\,|\mathbf{x} - \mathbf{z}|^\alpha.$$

Lemma 5: The $b_\beta{}^i(\mathbf{x})$ are α-Hölder continuous and have an α-G property: $R < x \leq y$,

$$x^2\,|b(\mathbf{x}) - b(\mathbf{y})| \leq A(B, R, \alpha)\,|\mathbf{x} - \mathbf{y}|^\alpha.$$

Proof: The proof is constructed in the same way as the proof of Theorem 8.

Theorem 9: The operator L^i maps $\mathfrak{B}$ into $\mathfrak{B}$; $\|L^i[f^j(\mathbf{x})]\|$ is bounded by a number which depends only on $\|f\|$, R, and B.

Proof: The functions $L^i[f^j(\mathbf{x})]$ of (3.6) are sums of terms of the type $bD_1\xi$, $(\partial_\beta b^\beta)\xi$, and $b^2\xi$. On account of (2.1) and Theorems 1 and 2, $|b|$, $(\partial_\beta b^\beta)|\xi|$, and $|D_1\xi|$ are bounded by a number depending only on $\|f\|$, R, and B. Hence, the same holds for $L^i[f^j(\mathbf{x})]$. Since $b(\mathbf{x})$ is bounded and α-Hölder continuous and since, by Theorem 1 and 8, $D_1\xi$ is bounded and α-Hölder continuous, $bD_1\xi$ is α-Hölder continuous. Similarly, the α-Hölder continuity of terms of the type $(\partial_\beta b^\beta)\xi$ and $b^2\xi$ follows from the boundedness and α-Hölder continuity of b, $\partial_\beta b^\beta$, and ξ. Since all the bounds and Hölder coefficients referred to above are bounded by a number depending only on $\|f\|$, R, and B, the Hölder coefficient of $L^i[f^j(\mathbf{x})]$ is bounded by such a number.

The α-G property of L^i is proved by showing that this property holds for $L_1 = bD_1\xi$, $L_2 = (\partial_\beta b^\beta)\xi$, and $L_3 = b^2\xi$. We have

$$|L_1(\mathbf{x}) - L_1(\mathbf{y})| \leq |b(\mathbf{x})|\,|D_1\xi(\mathbf{x}) - D_1\xi(\mathbf{y})| + |D_1\xi(\mathbf{y})|\,|b(\mathbf{x}) - b(\mathbf{y})|. \quad (5.35)$$

Using Theorems 1, 2, and 8 and Eq. (2.1), we conclude from (5.35) that for $R < x \leq y$

$$|L_1(\mathbf{x}) - L_1(\mathbf{y})| \leq \frac{A(\|f\|, R, B)}{x^4}\,|\mathbf{x} - \mathbf{y}|^\alpha.$$

Similarly, we have

$$|L_2(\mathbf{x}) - L_2(\mathbf{y})| \leq |\partial_\beta b^\beta(\mathbf{x})|\,|\xi(\mathbf{x}) - \xi(\mathbf{y})| + |\xi(\mathbf{y})|\,|\partial_\beta b^\beta(\mathbf{x}) - \partial_\beta b^\beta(\mathbf{y})|;$$

using (2.1) and Theorems 2 and 7, we conclude that

$$|L_2(\mathbf{x}) - L_2(\mathbf{y})| \leq \frac{A(\|f\|, R, B)}{x^4}\,|\mathbf{x} - \mathbf{y}|^\alpha. \quad (5.36)$$

Finally, we have

$$|L_3(\mathbf{x}) - L_3(\mathbf{y})| \leq |b(\mathbf{x})|^2\,|\xi(\mathbf{x}) - \xi(\mathbf{y})| + |b(\mathbf{x})\xi(\mathbf{y})|\,|b(\mathbf{x}) - b(\mathbf{y})| + |b(\mathbf{y})\xi(\mathbf{y})|\,|b(\mathbf{x}) - b(\mathbf{y})| \leq \frac{A(\|f\|, R, B)}{x^4}\,|\mathbf{x} - \mathbf{y}|^\alpha, \quad (5.37)$$

on account of Theorems 1, 2, and 7.

The asymptotic property

$$x^4\,|L[f(\mathbf{x})]| \leq A(\|f\|, R, B), \quad x > R,$$

follows from the asymptotic properties of b, $(\partial_\beta b^\beta)\xi$, and $D_1\xi$ [(2.1) and Theorem 2]. ∎

We are now prepared to prove the following theorem:

Theorem 10: The operator L^i is completely continuous.

Proof: Consider any infinite set[18] $\{f_n{}^i\} \subset \mathfrak{B}$ for which $\|f_n{}^i\| \leq C$. We must show that the set $\{L^i[f_n{}^j]\}$ is compact. Since we already know that $\mathfrak{B}$ is complete, it suffices to show that every infinite subsequence of $\{L^i[f_n{}^j]\}$ contains a Cauchy sequence with respect to the norm (3.11).

From $\|f_n{}^i\| \leq C$ and Theorem 9 it follows that $|L^i[f_n{}^j(\mathbf{x})]| \leq A(C, R, B)$. Since the 3-dimensional Euclidean space is separable, the bounded sequence

of functions $\{L^i[f_n{}^j(\mathbf{x})]\}$ contains a weak Cauchy sequence[19] $\{L^i[f_{n_1}{}^j(\mathbf{x})]\}$:

$$|L^i[f_{n_1}{}^j(\mathbf{x})] - L^i[f_{m_1}{}^j(\mathbf{x})]| < \epsilon, \quad n_1, m_1 > N_1(\epsilon). \tag{5.38}$$

For Theorems 1 and 2 it follows that

$$|\xi[f_{n_1}{}^i(\mathbf{x})]| < A(C, R)$$

and

$$x\,|\xi[f_{n_1}{}^i(\mathbf{x})]| < A(C, R), \quad x > R.$$

In the same way as we arrived at (5.38), it follows that there exists a subsequence

$$\{f_{n_2}{}^i\} \subset \{f_{n_1}{}^i\}$$

such that

$$|\xi[f_{n_2}{}^i] - \xi[f_{m_2}{}^i]| < \epsilon,$$

$$x\,|\xi[f_{n_2}{}^i] - \xi[f_{m_2}{}^i]| < \epsilon, \quad n_2, m_2 > N_2(\epsilon), \quad x > R. \tag{5.39}$$

From Theorems 1 and 2 we have

$$|D_1\xi[f_{n_2}{}^i]| < A(C, R)$$

and

$$x^2\,|D_1\xi[f_{n_2}{}^i]| < A(C, R), \quad x > R.$$

Therefore, a subsequence $\{f_{n_3}{}^i\} \subset \{f_{n_2}{}^i\}$ exists such that

$$|D_1\xi[f_{n_3}{}^i] - D_1\xi[f_{m_3}{}^i]| < \epsilon,$$

$$x^2\,|D_1\xi[f_{n_3}{}^i] - D_1\xi[f_{m_3}{}^i]| < \epsilon, \quad n_3, m_3 > N_3(\epsilon), \quad x > R. \tag{5.40}$$

Similarly, from Theorems 5 and 6 it follows that there exists a further subsequence $\{f_{n_4}{}^i\} \subset \{f_{n_3}{}^i\}$ such that

$$|D_2\xi[f_{n_4}{}^i] - D_2\xi[f_{m_4}{}^i]| < \epsilon,$$

$$x^3\,|D_2\xi[f_{n_4}{}^i] - D_2\xi[f_{m_4}{}^i]| < \epsilon, \quad n_4, m_4 > N_4(\epsilon), \quad x > R. \tag{5.41}$$

By the method used in the proof of Theorem 7, (5.39), (5.40), and (5.41) have the consequence

$$H_\alpha[\xi[f_{n_4}{}^i] - \xi[f_{m_4}{}^i]] < \epsilon A(R, \alpha) \tag{5.42}$$

and

$$H_\alpha[D_1\xi[f_{n_4}{}^i] - D_1\xi[f_{m_4}{}^i]] < \epsilon A(R, \alpha). \tag{5.43}$$

From (5.39), (5.40), (5.42), (5.43), Lemma 5, and the fact that b and $\partial_\beta b^\beta$ are bounded by a number which depends only on B, it follows that

$$H_\alpha[L^i[f_{n_4}{}^i] - L^i[f_{m_4}{}^j]] < \epsilon A_3(B, C, R, \alpha). \tag{5.44}$$

In a similar manner it follows that

$$G_\alpha[L^i[f_{n_4}{}^j] - L^i[f_{m_4}{}^j]] < \epsilon A_4(B, C, R, \alpha) \tag{5.45}$$

and

$$K[L^i[f_{n_4}{}^j] - L^i[f_{m_4}{}^j]] < \epsilon A_5(B, C, R, \alpha). \tag{5.46}$$

From (5.38), (5.44), (5.45), and (5.46) it follows that

$$\|L^i[f_{n_4}{}^j] - L^i[f_{m_4}{}^j]\| < \epsilon A_6(B, C, R, \alpha), \quad n_4, m_4 > N_4(\epsilon). \qquad \blacksquare$$

The existence of solutions $f(\mathbf{x}) \in \mathfrak{B}$ of the functional equations (3.5), for nondegenerate b and for any values c^m in (3.7), follows from the complete continuity of the operator L^i on $\mathfrak{B}$, and the nonexistence of nonnull solutions of the associated homogeneous problem (4.1), (4.2), for nondegenerate b, by application of the determinant-free Fredholm alternatives.[3,4] The solution $\xi^i(\mathbf{x})$ has the property $D_1\xi^i(\mathbf{x}) = O(x^{-2})$, on account of Theorem 2. Furthermore, the solution ξ^i satisfies (3.3), which together with (3.9) shows that $\partial_\beta\partial^\beta\xi^i = O(x^{-4})$. Hence, the solution $\xi^i(\mathbf{x})$ of the boundary-value problem (3.1) satisfies the asymptotic conditions (3.2). Hence the *orthogonal isospin transformation theorem* (Sec. II) is proved, *and it follows that the Yang–Mills field with configuration space* Ω *has physical states with a nonvanishing isospin component.*

6. TRANSVERSE POTENTIALS

The second existence problem which has bearing on the configuration space concerns potentials b which can be transformed to transverse potentials by a local gauge transformation

$$S(\mathbf{x}) = \exp \eta^i(\mathbf{x})L_i; \tag{6.1}$$

such potentials are here called *normal* potentials. The question arises whether all potentials $b \in \Omega$ are normal; this would be desirable if one wants to use transverse potentials in calculations. In the present paper, we do not solve this existence problem, but we reduce it to a simpler uniqueness problem. This is done by using a method of Boulware[9] which reduces the nonlinear problem to a linear problem and by proving the complete continuity of the relevant operator occurring in the equivalent functional equation. Boulware's method[9] amounts to finding the conditions that in Ω all potentials $b + \delta b$ infinitesimally close to normal potentials b are normal. For a finite local gauge transformation (6.1) one has

$$\Gamma'_\beta = S^{-1}\Gamma_\beta S - S^{-1}\partial_\beta S, \tag{6.2}$$

where

$$\Gamma_\beta = b_\beta{}^i L_i, \quad \Gamma'_\beta = b'_\beta{}^i L_i, \tag{6.3}$$

and the L^i are constant basis elements of the Lie algebra of $SU(2)$. For transverse potentials $b_T{}^{\beta i}$ one has

$$\partial_\beta b_T{}^{\beta i} = 0 \tag{6.4}$$

and, therefore, also

$$\partial_\beta \Gamma_T{}^\beta = 0. \tag{6.5}$$

When considering local gauge transformations which are not infinitesimal, it is easier to work with the Γ_β than with the $b_\beta{}^i$. Suppose that the $\Gamma_\beta \in \Omega$ are normal, i.e., there exists a transformation S of the form (6.1) with $\eta^i(\mathbf{x})$ subject to the asymptotic conditions (2.3) and (2.4), such that

$$\Gamma_T{}^\beta = S^{-1}\Gamma^\beta S - S^{-1}\partial^\beta S, \tag{6.6}$$

where the $\Gamma_T{}^\beta$ satisfy (6.5); i.e.,

$$\partial_\beta(S^{-1}\Gamma^\beta S - S^{-1}\partial^\beta S) = 0. \tag{6.7}$$

Let $\delta\Gamma^\beta \in \Omega$ be an arbitrary infinitesimal change of Γ^β. Let the associated change in S which keeps (6.7) valid be $\delta S = SX$; X then must satisfy the equation

$$\partial_\beta \nabla_T{}^\beta X = \partial_\beta(S^{-1}\delta\Gamma^\beta S), \tag{6.8}$$

where $\nabla_T{}^\beta X$ is the covariant derivative of X, belonging to the potentials $\Gamma_T{}^\beta$:

$$\nabla_T{}^\beta X = \partial^\beta X - [\Gamma_T{}^\beta, X]. \tag{6.9}$$

On account of (6.2) one may write

$$S^{-1}\delta\Gamma^\beta S = \delta\Gamma'^\beta, \tag{6.10}$$

where $\delta\Gamma'^\beta$ is the change in $\Gamma_T{}^\beta$ due to $\delta 1^\beta$, *keeping S fixed.* Clearly, if Γ and $\delta\Gamma$ belong to Ω, and S is a local gauge transformation, then $\delta\Gamma'$ belongs to Ω as well. Therefore, with

$$\delta\Gamma'^\beta = \delta b'^{\beta i} L_i, \tag{6.11}$$

the $\delta b'^{\beta i}(\mathbf{x})$ satisfy conditions (2.1). $I + X$ must be a local gauge transformation, and since X is infinitesimal, we may write

$$X = \eta^i(\mathbf{x}) L_i, \tag{6.12}$$

where the η^i are subject to conditions (2.3) and (2.4). In terms of $\eta^i(\mathbf{x})$ and $\delta b'^{\beta i}$, (6.9) reads

$$\partial_\beta \nabla_T{}^\beta \eta^i = \partial_\beta \delta b'^{\beta i}. \tag{6.13}$$

Since $\delta b'^{\beta i} \in \Omega$, the right-hand side of (6.13) satisfies the condition (3.9) and therefore lies in the Banach space $\mathfrak{B}$ defined in Sec. 3. Writing out the covariant derivative in (6.13) gives

$$\partial_\beta \partial^\beta \eta^i - c_{jk}{}^i b_T{}^{\beta j} \partial_\beta \eta^k = \partial_\beta \delta b'^{\beta i}. \tag{6.14}$$

We will introduce the functions h^i defined by

$$\partial_\beta \partial^\beta \eta^i = h^i; \tag{6.15}$$

using the asymptotic condition (2.3), we may write the solution η^i of (6.15) in terms of h^i as

$$\eta^i[h^j] = -\frac{1}{4\pi}\int d^3y h^i(\mathbf{y})\frac{1}{\rho}, \quad \rho = |\mathbf{x} - \mathbf{y}|. \tag{6.16}$$

With (6.16), the boundary-value problem (6.14), (2.3) may be expressed as the functional equation

$$h^i - M^i[h] = \partial_\beta \delta b'^{\beta i}, \tag{6.17}$$

where

$$M^i[h] = c_{jk}{}^i b_T{}^{\beta j} \partial_\beta \eta^k[h]. \tag{6.18}$$

As shown in Sec. 5, the operator L^i of (3.6) is a completely continuous operator on $\mathfrak{B}$. M^i of (6.18) is proportional to the first term of L^i, and from the proof of complete continuity of L^i it can be seen that M^i is also a completely continuous operator on $\mathfrak{B}$. Since $\partial_\beta \delta b'^{\beta i}$ belongs to $\mathfrak{B}$, it follows by application of determinant free form of the Fredholm alternatives[3,4] that the boundary-value problem (6.13), (2.3) has a solution in $\mathfrak{B}$ if the associated homogeneous problem

$$\partial_\beta \nabla_T{}^\beta \eta^i = 0, \tag{6.19}$$

with asymptotic condition (2.3), only has the null solution in $\mathfrak{B}$. Moreover, if the boundary-value problem (6.13), (2.3) has a solution in $\mathfrak{B}$, it follows from (6.15), Theorem 2, and (3.9) that property (2.4) is satisfied for the solution. Hence, if for all $b_T \in \Omega$, (6.19) and (2.3) has only the null solution in $\mathfrak{B}$, the boundary-value problem (6.13), (2.3), (2.4) has a unique solution in $\mathfrak{B}$, and then all potentials $b \in \Omega$ infinitesimally close to normal potentials are normal. Since there certainly exist *some* normal potentials in Ω, it would follow that all potentials $b \in \Omega$ are normal. Hence we have the result:

Reduction Theorem: If the homogeneous problem (6.19) and (2.3) only has the null solution, for all $b_T \in \Omega$, then for every $b \in \Omega$ there exists a unique local gauge transformation $b \to b'$ such that b' is transverse.

7. TENTATIVE CONFIGURATION SPACE

The configuration space considered in this paper is the metric space of all real-valued spatial Yang-Mills potential functions $b_\beta{}^i(\mathbf{x})$, subject to conditions (2.1), taking as the distance[1] between $b_\beta{}^i(\mathbf{x})$ and $b_\beta{}^i(\mathbf{x}) + \Delta b_\beta{}^i(\mathbf{x})$

$$\left(\int d^3x\, \Delta b_\beta{}^i(\mathbf{x})\, \Delta b^\beta{}_i(\mathbf{x})\right)^{\frac{1}{2}}. \tag{7.1}$$

For ease of reference, we separately show the conditions which make up (2.1):

$$|b| \leq B, \tag{7.2}$$

$$|D_1 b| \leq B, \tag{7.3}$$

$$|D_1 b(\mathbf{x}) - D_1 b(\mathbf{y})| \leq B\,|\mathbf{x} - \mathbf{y}|^{\alpha}, \quad 0 < \alpha < 1; \tag{7.4}$$

$$x > R,$$

$$x^2\,|b| \leq B, \tag{7.5}$$

$$x^3\,|D_1 b| \leq B, \tag{7.6}$$

$$x^4\,|\partial_\beta b^{\beta i}| \leq B; \tag{7.7}$$

$$R < x \leq y,$$

$$x^4\,|\partial_\beta b^{\beta i}(\mathbf{x}) - \partial_\beta b^{\beta i}(\mathbf{y})| \leq B\,|\mathbf{x} - \mathbf{y}|^{\alpha}. \tag{7.8}$$

Conditions (7.2) and (7.3) ensure that the space–space components of the Yang–Mills field

$$B_{\beta\gamma}{}^{i} = \partial_\beta b_\gamma{}^{i} - \partial_\gamma b_\beta{}^{i} - c_{jk}{}^{i} b_\beta{}^{j} b_\gamma{}^{k} \tag{7.9}$$

are bounded. The asymptotic condition (7.5) ensures that the distance (7.1) between any two points b and $b + \Delta b$ of Ω is finite.

Adopting the asymptotic electromagnetic identification,[7] which amounts to identifying, in the world where there exists only the Yang–Mills field, the far-away Yang–Mills field as electromagnetic, we allow a nonvanishing total magnetic moment by the asymptotic condition (7.5) for b. Photons at spatial infinity are excluded by (7.5), and this may seem objectionable. However, in considering scattering problems which involve photons in the in and out states, one can take these states at times such that the photons do not appear at spatial infinity, if the photons are taken as localized packets. This amounts to considering the S matrix between large negative and positive times $\pm T$, instead of $\pm\infty$. In a calculation based on absence of photons at infinity, a definite S matrix may be obtained by letting T go to infinity as a final step. In the asymptotic electromagnetic identification, the different i components of the expectation values for $B_{0\beta}{}^{i}(\mathbf{x})$ and $B_{\beta\gamma}{}^{i}(\mathbf{x})$ must be proportional for all values of β, γ, and $\mathbf{x}/x$, as $x \to \infty$, since there is only a single electromagnetic field. To achieve this, further restrictions may have to be imposed on the configuration space, but this problem must be studied considering the state space and the development of states in time as well.

In preparation for comments on the asymptotic condition (7.7) for $\partial_\beta b^{\beta i}$, we consider in Ω the decomposition of an arbitrary infinitesimal vector $\delta b^{\alpha i}$ at b into a vector $\delta_\perp b$ orthogonal to the gauge-invariant manifold χ at b and a vector $\delta_\parallel b$ along χ:

$$\delta b^{\beta i} = \delta_\perp b^{\beta i} + \delta_\parallel b^{\beta i}. \tag{7.10}$$

In the same way as we derived (2.16), it can be shown that

$$\nabla_\beta \delta_\perp b^{\beta i} = 0. \tag{7.11}$$

Since $\delta_\parallel b^{\beta i}$ is due to a local gauge transformation, we have

$$\delta_\parallel b^{\beta i} = -\nabla^\beta \eta^i. \tag{7.12}$$

Taking the covariant divergence of (7.10) and using (7.11) and (7.12) gives

$$\nabla_\beta \nabla^\beta \eta^i = -\nabla^\beta \delta b_\beta{}^{i}. \tag{7.13}$$

The existence proof given in Secs. 3, 4, and 5 applies to the solution $\eta^i(\mathbf{x})$ of Eq. (7.13) subject to the asymptotic conditions (2.3) and (2.4). Also, the considerations of Sec. 4 apply to the homogeneous problem associated with (7.13). Consequently, we have in Ω existence and uniqueness of the decomposition (7.10) at every nondegenerate $b \in \Omega$.

If the asymptotic condition (7.7) for $\partial_\beta b^{\beta i}$ were dropped, then (7.6) would set the asymptotic behavior of $\partial_\beta b^{\beta i}$, which would be $O(x^{-3})$ instead of $O(x^{-4})$. Since

$$\partial^\beta \partial_\beta (x^{-1} \ln x) = O(x^{-3}),$$

there would then exist a displacement $\delta b_\beta{}^{i}(\mathbf{x})$ in Ω and a $\eta^i(\mathbf{x}) = O(x^{-1} \ln x)$, for which (7.13) is satisfied. For nondegenerate $b \in \Omega$, the solution $\eta^i(\mathbf{x})$ of (7.13) is unique; this follows from the argument of Sec. 4, which goes through here because the surface integral in (4.4) vanishes for $x \to \infty$, if $\eta^i = O(x^{-1} \ln x)$ is taken instead of ξ^i. Hence, if condition (7.7) were dropped, there would exist a displacement $\delta b^{\beta i}(\mathbf{x})$ in Ω such that the solution η^i of (7.13) is of order $O(x^{-1} \ln x)$; for that solution η^i the $\delta_\parallel b^{\beta i}$ calculated from (7.12) would be of order $O(x^{-1} \ln x)$, and therefore the displacement $\delta_\parallel b^{\beta i}$ would not lie in Ω.

Dropping (7.7) would also have consequences for the reduction theorem and for the orthogonal isospin transformation theorem. The right-hand side of (6.13) would be of order $O(x^{-3})$, which is outside $\mathfrak{B}$. If $\mathfrak{B}$ were modified to include functions $h^i(\mathbf{x})$ of order $O(x^{-3})$ instead of $O(x^{-4})$, then $\eta^i[h^j]$ calculated from (6.16) would be of order $O(x^{-1} \ln x)$ for certain $h^i(\mathbf{x})$; the vector δb calculated from (2.2) would then be of order $O(x^{-2} \ln x)$ and therefore would not lie in Ω. The second term on the left-hand side of (6.14) cannot bring relief because it is of order $O(x^{-4})$, if η is of order $O(x^{-1})$ as desired. Hence, if (7.7) would be dropped, there would exist

potentials $b \in \Omega$, which cannot be transformed to transverse potentials. Dropping (7.7) would have a similar effect on the considerations regarding the existence of a solution of Eqs. (3.1) and (3.2). As a result, the orthogonal isospin transformation theorem would read: The set $\mathfrak{N}$ includes all nondegenerate potentials $b \in \Omega$ for which $\partial_\beta b^{\beta i}$ is of order $O(x^{-4})$. Since this set has a nonvanishing measure in Ω, it still would follow that the Yang–Mills field (with this modified configuration space) has physical states with a nonvanishing isospin component.

The undesirable factor which would result from dropping (7.7) is only logarithmic; hence, a restriction weaker than (7.7) may be sufficient to avoid the problems discussed. However, if one allows only asymptotic restrictions expressed by an integral power of x, then (7.7) is the weakest restriction with the desired effect. Instead of including a restriction like (7.7), one could attempt to enlarge Ω such that potentials of order $O(x^{-2} \ln x)$ are allowed. However, there does not seem to be a physical need to include such configurations; on the other hand, the condition (7.7) does not seem to constitute a physical restriction. Hence, we see no merit in such a modification of the configuration space.

The α-Hölder continuity condition (7.4) is a minimum smoothness condition. In practice, we need second derivatives of $b_\beta{}^i$ in the theory, and we would demand

$$|D_2 b| \leq B; \quad x \to \infty, \quad x^4 |D_2 b| \leq B. \tag{7.14}$$

If still higher derivatives of b are required, similar conditions for the higher derivatives would be imposed on the configuration space. Conditions (7.14) imply the Hölder continuity (7.4); the purpose of imposing (7.4) instead of the stronger condition (7.14) is to state the orthogonal isospin transformation theorem and the reduction theorem in as sharp a form as we know, i.e., with the weakest conditions on the $b_\beta{}^i(\mathbf{x})$. A similar comment applies to the α-G condition (7.8) for $\partial_\beta b^{\beta i}$.

8. THE CONSTRAINT EQUATION FOR THE CLASSICAL THEORY

The existence of solutions ζ^i of (2.16), subject to the boundary conditions (2.8) and (2.9), has bearing on the constraint equation

$$\nabla_\beta B^{0\beta i} = 0 \tag{8.1}$$

of the classical theory. The "electric" part $B^{0\beta i}$ of the Yang–Mills field may be written[7]

$$B^{0\beta i} = \nabla^\beta \varphi^i + \nabla_\gamma \psi^{\beta\gamma i}, \tag{8.2}$$

where $\psi^{\beta\gamma i}$ is antisymmetric in β and γ. The decomposition (8.2) gives, for the constraint equation (8.1),

$$\nabla_\beta \nabla^\beta \varphi^i = \tfrac{1}{2} c^i{}_{jk} B_{\beta\gamma}{}^j \psi^{\beta\gamma k}, \tag{8.3}$$

if use is made of the identity[6]

$$\nabla_{[\beta} \nabla_{\gamma]} v^i = -\tfrac{1}{2} c^i{}_{jk} B_{\beta\gamma}{}^j v^k. \tag{8.4}$$

Equation (8.3) for φ^i is of similar form as (2.16), which with (2.8) may be written

$$\nabla_\beta \nabla^\beta \xi^i = -\nabla_\beta \nabla^\beta c^i; \tag{8.5}$$

only the source terms on the right are different. For $B^{0\beta i}$ in (8.2) to be of order $O(x^{-2})$ we take

$$\varphi^i = O(x^{-1}), \tag{8.6}$$

and $\psi^{\beta\gamma i} = O(x^{-1})$; then, $\varphi^i(\mathbf{x})$ is restricted asymptotically by the same condition as $\xi^i(\mathbf{x})$. In order that our existence theorem applies to (8.3), the right-hand side $\frac{1}{2} c^i{}_{jk} B_{\beta\gamma}{}^j \psi^{\beta\gamma k}$ must lie in the Banach space $\mathfrak{B}$. It can easily be seen that this is the case if the $\psi^{\beta\gamma i}(\mathbf{x})$ are bounded and α-Hölder continuous, $\psi^{\beta\gamma i} = O(x^{-1})$, and if $\psi^{\beta\gamma i}(\mathbf{x})$ and $b_\alpha{}^i(\mathbf{x}) \in \Omega$ satisfy the modified asymptotic Hölder conditions

$$\begin{aligned} x\,|\psi(\mathbf{x}) - \psi(\mathbf{y})| &\leq G_\alpha |\mathbf{x} - \mathbf{y}|^\alpha, \\ x^3 |D_1 b(\mathbf{x}) - D_1 b(\mathbf{y})| &\leq G_\alpha |\mathbf{x} - \mathbf{y}|^\alpha, \quad R < x \leq y. \end{aligned} \tag{8.7}$$

Then, Eq. (8.3) subject to the asymptotic condition (8.6) has a (unique) solution $\varphi^i \in \mathfrak{B}$, if the homogeneous equation associated with (8.3) has only the null solution; in Sec. 4 it is shown that this is the case if the holonomy group of the $b_\alpha{}^i(\mathbf{x})$ is the full $SU(2)$. Only the covariant-longitudinal part $\nabla^\beta \varphi^i$ of (8.2) contributes to the isospin.[7]

* Work conducted at the Douglas Advanced Research Laboratories under company-sponsored Independent Research and Development funds.

[1] H. G. Loos, Phys. Rev. **188**, 2342 (1969).
[2] J. Schauder, Math. Z. **38**, 2, 257 (1934).
[3] F. Reisz, Acta Math. **41**, 71 (1918).
[4] A. N. Kolmogorov and S. V. Fomin, *Elements of the Theory of Functions and Functional Analysis* (Graylock, Rochester, N.Y., 1957), Vol. 1.
[5] A linear operator on a Banach space is completely continuous, if it turns any bounded set into a compact set.
[6] H. G. Loos, J. Math. Phys. **8**, 2114 (1967).
[7] H. G. Loos, Nuovo Cimento **58A**, 365 (1968).

[8] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1962), Vol. II, p. 250.

[9] D. G. Boulware, Ann. Phys. (N.Y.) **56**, 140 (1970).

[10] A change in a physical-state functional is noticeable, if it effects a physical scalar product.

[11] $\zeta^i(b, \mathbf{x})$ here means that ζ^i is a functional of $b(\mathbf{x})$ and a function of $\mathbf{x}$.

[12] N. Dunford and J. T. Schwartz, *Linear Operators* (Interscience, New York, 1966), Part I, p. 258.

[13] H. G. Loos, Nuovo Cimento **53A**, 201 (1968).

[14] So many bounds depending only on $\|f\|$ and R arise in this paper, that we denote them by a common symbol, such as $A(\|f\|, R)$, if no confusion can arise.

[15] Ref. 8, p. 247.

[16] Ref. 8, p. 250.

[17] Of course, in $D_2\xi[f_n] - D_2\xi[f]$, D_2 is meant to be the partial derivative with respect to the same variables.

[18] The functions f_n used in the proof of Theorem 10 have nothing to do with the functions defined by (5.1).

[19] Ref. 4, p. 94, Theorem 1.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

General Formula to Derive Branching Rules

A. VAN DAELE*
Institute of Theoretical Physics, University of Leuven, Leuven, Belgium†

(Received 25 February 1970)

We consider an irreducible representation of a semisimple Lie algebra L. When restricted to a semisimple subalgebra K of L, this representation can be reduced with respect to K. We derive a general formula for the multiplicity of a certain irreducible representation of K, which occurs in it. The result is an extension of Kostant's formula for the multiplicity of a weight, where the subalgebra K is the Cartan subalgebra of L. Using Kostant's formula, we write down a set of equations, containing the required multiplicity, completely analogous with the usual formula involving the characters. We rewrite these equations using some properties of the partition function (used in Kostant's formula) and of the Weyl groups. Finally we solve them with the help of an "orthogonality property." We illustrate the applicability by working out two nontrivial examples.

I. INTRODUCTION

Some of the most important rules in the application of semisimple Lie groups and their Lie algebras in particle, atomic, and nuclear physics are the so-called branching rules. In the reduction of an irreducible representation of a semisimple Lie algebra L with respect to a semisimple subalgebra K, the different irreducible representations of the subalgebra K which occur, together with their multiplicities, are then given by these rules.

A general closed formula, valid for finite-dimensional representations, was obtained by Straumann[1] and by Delaney and Gruber.[2] This formula involves the knowledge of the multiplicities of the different weights in the considered irreducible representation of L.

In this paper we present another closed formula which does not make use of these multiplicities. It contains a formula of Mandel'tsveig[3] as a special case. The result is similar to Kostant's formula[4] for the multiplicity of a weight. Moreover, it is an extension of this formula because, in the special case where the subalgebra K is the Cartan subalgebra, the different irreducible representations of the subalgebra (which are all 1-dimensional) are given by the different weight, together with their multiplicities.

In fact, our starting point is Kostant's formula. We consider the multiplicity of a weight in an irreducible representation of the semisimple Lie algebra L. We restrict this representation to the subalgebra K, and we derive a formula for the multiplicity of a weight (with respect to K) in this (usually) reducible representation of K.

On the other hand, this multiplicity equals the sum over the irreducible representations of K of the product of the multiplicity of a weight in an irreducible representation of K times the multiplicity of this representation in the restriction to K of the considered representation of L. Modifying the two results, we can use a property of the considered partition functions—closely related to the orthogonality relations of the characters—to extract the required multiplicity.

But, before we can start, we are forced to give a few introductory notions and some properties of the considered partition functions. We also prove two lemmas, related to the commutant of the Cartan subalgebra J of K in L. These two lemmas are essential in the interpretation of an important factor in the formula. In the Appendix we treat two examples.

II. BASIC CONCEPTS

We consider a semisimple subalgebra K of a semisimple Lie algebra L over the field of real numbers. If J is a Cartan subalgebra of K, we can find a Cartan subalgebra H of L such that H contains J (see Ref. 5, p. 149).

We denote the real vector spaces, generated by the roots of L, by H^* and of K by J^*. Thus, the elements of $H^*(J^*)$ are linear functions on $H(J)$. We denote by A and C the set of positive roots of L and K, respectively. (Positivity in H^* and J^* will be introduced later.)

The partition function $P_B(\lambda)$, where $\lambda \in H^*$ (respectively J^*) and B is a finite subset of H^* (respectively J^*), is defined as the number of nonnegative integer solutions of the equation in x_α, $\alpha \in B$,

$$\lambda = \sum_{\alpha \in B} \alpha \cdot x_\alpha .$$

The special function δ_B is defined as $\frac{1}{2}\sum_{\alpha \in B} \alpha$, where B is as before. The Weyl groups of K and L are denoted by W_K and W.

We can now formulate the problem in mathematical terms. The multiplicity $n(\Lambda, \lambda)$ of a weight $\lambda \in H^*$ in the irreducible L-module with highest weight $\Lambda \in H^*$ is given by Kostant's formula

$$n(\Lambda, \lambda) = \sum_{S \in W} (\det S) P_A((\Lambda + \delta_A)S - (\lambda + \delta_A)),$$

so that the multiplicity $\bar{n}(\Lambda, m)$ of a weight $m \in J^*$ in the restriction to K of the irreducible L-module of highest weight $\Lambda \in H^*$ equals

$$\bar{n}(\Lambda, m) = \sum_{\lambda, \lambda(J)=m} n(\Lambda, \lambda), \tag{1}$$

where we denote by $\lambda(J)$ the restriction of a function $\lambda \in H^*$ to J.

On the other hand, if we denote by $N[\Lambda, M]$ the multiplicity of the irreducible K-module with highest weight M, which occurs in the restriction to K of the irreducible L-module with highest weight Λ, we have the equality

$$\bar{n}(\Lambda, m) = \sum_{M \in J^*} N[\Lambda, M] n(M, m),$$

where the summation runs over those weights $M \in J^*$ which can be the highest weight of an irreducible K-module.

These two expressions of $\bar{n}(\Lambda, m)$ yield an equation for every $m \in J^*$. In order to solve this set, we rewrite both sides, and we use a property of the partition function to get rid of the summation over M. The result is an expression for $N[\Lambda, M]$.

III. PROPERTIES OF THE PARTITION FUNCTIONS

We first prove three properties of the partition function.

Property 1: If A_1, A_2, and A are three finite subsets of H^* such that $A_1 \cup A_2 = A$ and $A_1 \cap A_2 = \varnothing$, then we have for all $\lambda \in H^*$ that

$$P_A(\lambda) = \sum_{\mu \in H^*} P_{A_1}(\lambda - \mu) P_{A_2}(\mu).$$

(This property is closely related to the recursion formula for the partition function of the algebra A_l of Radhakrishnan.[6])

Proof: Consider the two equations in x_α, $x_\alpha \geq 0$:

$$\lambda = \mu + \sum_{\alpha \in A_1} \alpha \cdot x_\alpha,$$

$$\mu = \sum_{\alpha \in A_2} \alpha \cdot x_\alpha.$$

The number of solutions of the set equals the product of the number of solutions of the first times that of the second. So, it is

$$P_{A_1}(\lambda - \mu) P_{A_2}(\mu).$$

Eliminating the μ in the equations, we get

$$\lambda = \sum_{\alpha \in A} \alpha \cdot x_\alpha.$$

So, we have here $\sum_\mu P_{A_1}(\lambda - \mu) P_{A_2}(\mu)$ solutions. On the other hand, this number is $P_A(\lambda)$ and the property is proved. Clearly, the same result holds if H^* is replaced by J^*.

Property 2: If A is a finite subset of H^* and B the set of restrictions of the elements of A to the subspace J of H, then for every $m \in J^*$ we have that

$$P_B(m) = \sum_{\substack{\lambda \in H^* \\ \lambda(J)=m}} P_A(\lambda).$$

Proof: Consider the two equations in x_α, $x_\alpha \geq 0$, and λ:

$$\lambda = \sum_{\alpha \in A} \alpha \cdot x_\alpha,$$

$$\lambda(J) = m.$$

For every λ such that $\lambda(J) = m$, we have $P_A(\lambda)$ solutions. Thus the number of solutions is

$$\sum_{\lambda, \lambda(J)=m} P_A(\lambda).$$

Eliminating λ, we get the equation

$$m = \sum_{\alpha \in A} x_\alpha \cdot \alpha(J).$$

By the definition of B, the number of solutions of this equation is $P_B(m)$, and it follows that

$$P_B(m) = \sum_{\lambda, \lambda(J)=m} P_A(\lambda).$$

Property 3: If C is a finite subset of the positive elements of J^* and if, for every $m \in J^*$,

$$\sum_{n \in J^*} P_C(n - m) X_n = 0, \quad X_n \in \mathbb{Z},$$

then it follows that either all $X_n = 0$ or that there exists an n_0 such that $X_{n_0} \neq 0$ and there are infinitely many elements of the form

$$n = n_0 + \sum_{\alpha \in C} \alpha \cdot x_\alpha, \quad x_\alpha \geq 0,$$

with $X_n \neq 0$.

Proof: Suppose that there exists an n_0 with $X_{n_0} \neq 0$. Consider the following equation with $m = n_0$:

$$\sum_{n \in J^*} P_C(n - n_0) X_n = 0.$$

Since $P_C(n - n_0) = 0$, if $n < n_0$, and $P_C(n - n_0) = 1$, if $n = n_0$, we can rewrite the equation as

$$\sum_{n > n_0} P_C(n - n_0) X_n = -X_{n_0}.$$

The left-hand side must be different from zero since $X_{n_0} \neq 0$. Thus there exists an $n > n_0$ of the form

$$n = n_0 + \sum_{\alpha \in C} \alpha \cdot x_\alpha, \quad x_\alpha \geq 0,$$

with $X_n \neq 0$. It follows immediately that there are infinitely many elements of this form with $X_n \neq 0$.

IV. THE COMMUTANT IN L OF THE CARTAN SUBALGEBRA J OF K

Let us consider Property 2 of the partition function, and suppose that there exists an element $\alpha \in A$ such that its restriction to J, $\alpha(J)$ is zero. It follows that B contains zero as an element, and so $P_B(m)$ is always infinite or zero.

Indeed, $P_B(m)$ is the number of solutions of

$$m = \sum_{\beta \in B} \beta \cdot x_\beta .$$

Clearly, if $0 \in B$ and if there exists a solution $\{x_\beta\}$, then, for every positive integer n, we have that

$$m = \sum_{\beta \in B} \beta \cdot x_\beta + n \cdot 0,$$

and so $\{x_\beta + n\delta(\beta, 0)\}$ is also a solution.

Now we want to use Property 2 to work out the sum over λ with $\lambda(J) = m$ in formula (1). Since $n(\Lambda, \lambda)$ contains the partition function P_A, where A is the set of positive roots of L, it is clear from these considerations that the subset of A, $\{\alpha \in A, \alpha(J) = 0\}$, will play an important role in our treatment, and since $\alpha(J) = 0$ means that the corresponding generator E_α in L commutes with J, E_α belongs to the commutant of J in L if α is an element of this subset.

Therefore, we prove two lemmas concerning this commutant and its Weyl group.

Lemma 1: If J is a subalgebra of a Cartan subalgebra H of L, then the subalgebra L_1 of L, consisting of elements commuting with H, is the direct sum of a commutative and a semisimple part.

Proof: Consider the set of roots in L, according to the Cartan subalgebra H, and the corresponding basis $\{E_\alpha, \alpha \text{ a root}\}$ in L. Suppose that an element $\sum_\alpha \lambda_\alpha \cdot E_\alpha \in L_1$ and so commutes with all elements $h \in J$. Thus,

$$\left(\sum_\alpha \lambda_\alpha E_\alpha, h\right) = \sum_\alpha \lambda_\alpha [E_\alpha, h] = \sum_\alpha \lambda_\alpha \alpha(h) E_\alpha = 0.$$

Since the $\{E_\alpha\}$ constitute a basis, we have for every h and α that

$$\lambda_\alpha \alpha(h) = 0,$$

which is equivalent to saying that

$$\lambda_\alpha = 0 \quad \text{if} \quad \alpha(h) \neq 0,$$

and so $\{E_\alpha, \alpha \text{ a root}, \alpha(J) = 0\}$ is a basis in L_1.

We now introduce an ordered basis in J^* and H^* such that the corresponding sets of positive elements have the following properties, $\alpha \in H^*$:

$$\alpha > 0 \Rightarrow \alpha(J) \geq 0$$

and

$$\alpha(J) > 0 \Rightarrow \alpha > 0.$$

Consider a basis in H^*, $\{\lambda_i, i = 1, l\}$, and consider the restrictions to J, $\{\lambda_i(J)\}$. We can rearrange the indices so that the first restrictions constitute a basis in J^*. If k is the dimension of J, then we may suppose that $\{\lambda_i(J), i = 1, k\}$ is a basis in J^*.

We can now write the restrictions $\lambda_q(J)$ with $q = k + 1, l$ in terms of the basis in J^*,

$$\lambda_q(J) = \sum_{i=1}^{k} a_i^q \lambda_i(J), \quad k < q \leq l,$$

and redefine the basis in H^*,

$$\lambda'_q = \lambda_q - \sum_{i=1}^{k} a_i^q \lambda_i, \quad \text{if} \quad k < q \leq l,$$

$$\lambda'_q = \lambda_q, \quad \text{if} \quad 0 < q \leq k.$$

An element of J^*, written as $\sum_{i=1}^{k} p_i \lambda_i(J)$, is called positive if the first nonzero coefficient is positive; an element of H^*, written as $\sum_{i=1}^{l} p_i \lambda'_i$, is called positive if the first nonzero coefficient is positive. The inclusion properties can easily be verified. Having defined the positive roots in L, we can associate the corresponding simple system of roots in L, and we denote it by $\pi = \{\alpha_1, \alpha_2, \cdots, \alpha_l\}$.

Let us consider a positive root β such that $\beta(J) = 0$. Since $\{\alpha_i\}$ is a simple system, we can write

$$\beta = \sum k_i \alpha_i, \quad \forall\, k_i \geq 0.$$

Restricting to J, we have

$$\beta(J) = 0 = \sum k_i \alpha_i(J).$$

Since $\alpha_i(J) \geq 0$, we have that $k_i = 0$ if $\alpha_i(J) > 0$.

If we call π' the subset of π for which $\alpha_i(J) = 0$, then every positive root β, $\beta(J) = 0$, can be written as a sum with nonnegative coefficients of the basis roots π'. We may conclude that π' is a simple system for the roots with zero restriction to J (see Ref. 5, p. 121).

Next we consider the canonical generators $\{h_i, e_i, f_i\}$ according to the system π, and we define L_2 to be the subalgebra of L generated by the subset $\{h_i, e_i, f_i$ with $\alpha_i \in \pi'\}$. It is easy to see that L_2 must be semisimple. If $\alpha_i \in \pi'$, then the e_i and f_i commute with J and so do the h_i, and it follows that L_2 is contained in L_1.

Next we consider the subalgebra H_2 of H defined as the set

$$\{h, h \in H, \forall\, \alpha_i \in \pi' \Rightarrow \alpha_i(h) = 0\}.$$

It is now possible to show that L_1 is the direct sum of H_2 and L_2. Indeed, if α is a nonzero root and if the

corresponding generator E_α commutes with J, we can write

$$\alpha = \sum_{\alpha_i \in \pi'} k_i \alpha_i,$$

and E_α is generated by $\{e_i, f_i \text{ with } \alpha_i \in \pi'\}$ and hence belongs to L_2. On the other side, if $h \in H$, then we can solve the set of equations ($\alpha_i \in \pi'$ and summation over those indices such that $\alpha_j \in \pi'$)

$$\alpha_i(h) = \sum k_j \alpha_i(h_j)$$

because $\alpha_i(h_j)$ is the Cartan matrix of L_2, which is nonsingular. It follows that $\alpha_i(h - \sum k_j h_j) = 0$ for $\alpha_i \in \pi'$, and so $h - \sum k_j h_j \in H_2$. Furthermore, it is clear that H_2 commutes with L_2.

Lemma 2: The subgroup W_0 of the Weyl group W of L, defined as the set

$$\{S, S \in W, \forall\, \lambda \in H \Rightarrow \lambda S(J) = \lambda(J)\},$$

equals the Weyl group W_{L_2} of the algebra L_2 (as defined in Lemma 1).

Proof: It is easy to show that $W_{L_2} \subseteq W_0$. Suppose that α is a nonzero root of L_2; then the mapping $S_\alpha : H^* \to H^*$ is defined as

$$\lambda S_\alpha = \lambda + [(\lambda, \alpha)/(\alpha, \alpha)]\alpha \quad \text{(see Ref. 5).}$$

If we restrict to J, we get

$$\lambda S_\alpha(J) = \lambda(J) \quad \text{since} \quad \alpha(J) = 0.$$

This holds for all $\lambda \in H^*$ and all the nonzero roots of L_2, and, because W_{L_2} is generated by the reflexions S_α, where α is a nonzero root of L_2, we may conclude that

$$W_{L_2} \subseteq W_0.$$

To prove that $W_0 \subseteq W_{L_2}$, we show that, for every $S \in W_0$, there exists a $S_1 \in W_{L_2}$ such that SS_1 leaves the simple system π of L fixed, and so (see Ref. 5, p. 242) $SS_1 = 1$ and $S \in W_{L_2}$. Given $S \in W_0$, we first construct such an element S_1 of W_{L_2}. We consider the simple system π' in L_2, contained in π (as in Lemma 1) and we show that $\pi' S$ is again a simple system in L_2.

Take an arbitrary nonzero root β of L_2; βS^{-1} is again a nonzero root of L_2 because $S \in W_0$. Thus it can be written in terms of the simple system π':

$$\beta S^{-1} = \sum_{\beta_i \in \pi'} k_i \beta_i,$$

with $\forall\, k_i \geq 0$ or $\forall\, k_i \leq 0$ (see Ref. 5, p. 121).

It follows that $\beta = \sum k_i(\beta_i S)$ with the same condition for the coefficients, and, since β was arbitrary, the set $\pi'' = \pi' S$ is a simple system for L_2. Hence, by a property of the Weyl group, there exists an element $S_1 \in W_{L_2}$ such that

$$\pi'' S_1 = \pi'.$$

We can now show that $\pi S S_1 = \pi$. Clearly, $\pi S S_1$ is a simple system because $SS_1 \in W$.

Consider a simple root $\alpha_i \in \pi \setminus \pi'$. We have that $\alpha_i(J) > 0$ and

$$\alpha_i S S_1(J) = \alpha_i(J) > 0,$$

and by the construction of the positive roots $\alpha_i S S_1 > 0$. On the other hand, we have that $\pi' S S_1 = \pi'$, so that the set $\pi S S_1$ consists of positive simple roots. It follows that $\pi S S_1 = \pi$ by a property of simple systems. QED

V. A FORMULA TO DERIVE BRANCHING RULES

Theorem 1: The multiplicity of a weight $m \in J^*$ in a K-module, which is the restriction to K of an irreducible L-module with highest weight Λ ($\in H^*$) is

$$\bar{n}(\Lambda, m) = \sum_{S \in W} (\det S) P_B((\Lambda + \delta_A)S(J) - (m + \delta_B)) \times \dim [(\Lambda + \delta_A)S - \delta_{A_1}], \quad (2)$$

where A is the set of positive roots of L, B the set of nonzero restrictions of A to J, A_1 the subset of A, having zero restriction to J, $\dim [\lambda] = 0$ if λ is not dominant (with respect to L_2),[7] and $\dim [\lambda] =$ dimension of the L_2-module with highest weight equal to the restriction of λ to L_2, if λ is dominant.

Proof: Clearly, the considered multiplicity

$$\bar{n}(\Lambda, m) = \sum_{\lambda, \lambda(J) = m} n(\Lambda, \lambda),$$

where $n(\Lambda, \lambda)$ is the multiplicity of the weight $\lambda \in H^*$ in the irreducible L-module with highest weight Λ. Thus by Kostant's formula

$$\bar{n}(\Lambda, m) = \sum_{\lambda, \lambda(J) = m} \sum_{S \in W} (\det S) \times P_A((\Lambda + \delta_A)S - (\lambda + \delta_A)).$$

To prove the theorem, we must evaluate the sum over λ. However, we are not allowed to commute both sums since, if e.g., $\alpha \in A$ and $\alpha(J) = 0$, then

$$\sum_{\lambda, \lambda(J) = m} P_A((\Lambda + \delta_A)S - (\lambda + \delta_A))$$

might be infinite, as we already mentioned in Sec. IV. Therefore, we introduce the subsets of A,

$$A_1 = \{\alpha \in A, \alpha(J) = 0\},$$
$$A_2 = \{\alpha \in A, \alpha(J) > 0\}.$$

We already saw in Lemma 1 that $A = A_1 \cup A_2$. Clearly, $A_1 \cap A_2 = \phi$, and so we can use Property 1 to get

$$\bar{n}(\Lambda, m) = \sum_{\lambda, \lambda(J)=m} \sum_{S \in W} (\det S) \times \sum_{\mu \in H^*} P_{A_2}(\mu - \lambda) P_{A_1}((\Lambda + \delta_A)S - (\mu + \delta_A))$$
$$= \sum_{\lambda, \lambda(J)=m} \sum_{\mu} P_{A_2}(\mu - \lambda) \times \sum_{S \in W} (\det S) P_{A_1}((\Lambda + \delta_A)S - (\mu + \delta_A)).$$

Now

$$\sum_{\lambda, \lambda(J)=m} P_{A_2}(\mu - \lambda) = P_B(\mu(J) - m)$$

by Property 2, and, because B is a set of positive elements, the result is finite. Thus we may commute the sum over λ and μ and use Property 2.

Hence,

$$\bar{n}(\Lambda, m) = \sum_{\mu} P_B(\mu(J) - m) \times \sum_{S \in W} (\det S) P_{A_1}((\Lambda + \delta_A)S - (\mu + \delta_A)).$$

We have that $P_{A_1}(\lambda) = 0$ or $\lambda = \sum_{\alpha \in A_1} x_\alpha \alpha$, and so $\lambda(J) = 0$.

It follows that we may replace $\mu(J)$ by

$$(\Lambda + \delta_A)S(J) - \delta_A(J),$$

and then we find

$$\bar{n}(\Lambda, m) = \sum_{\mu} \sum_{S \in W} (\det S) P_B((\Lambda + \delta_A)S(J) - (m + \delta_B)) P_{A_1}((\Lambda + \delta_A)S - (\mu + \delta_A)).$$

The last factor is a partition function over the elements A_1, which are the positive roots of the algebra L_2; since μ occurs only in this term and we sum over μ, we only miss a summation over the Weyl group of L_2 in order to have a dimension of an irreducible L_2-module. But we know from Lemma 2 that this Weyl group is the subgroup W_0 of W.

So we introduce this subgroup, and we split the sum over W, into a sum over the left cosets and over W_0 itself.

Since $(\Lambda + \delta_A)SS_0(J) = (\Lambda + \delta_A)S(J)$ if $S_0 \in W_0$, we get

$$\bar{n}(\Lambda, m) = \sum_{\mu} \sum_{S \in W/W_0} (\det S) P_B((\Lambda + \delta_A)S(J) - (m + \delta_B)) \times \sum_{S_0 \in W_0} (\det S_0) P_{A_1}((\Lambda + \delta_A)SS_0 - (\mu + \delta_A)).$$

W_0 is the Weyl group of L_2 by Lemma 2, and the A_1 are the positive roots of L_2; thus we have that

$$\sum_{\mu} \sum_{W_0} (\det S_0) P_{A_1}((\Lambda + \delta_A)SS_0 - (\mu + \delta_A)) = \sum_{\mu} \sum_{W_{L_2}} (\det S_0) P_{A_1}[((\Lambda + \delta_A)S - \delta_{A_1} + \delta_{A_1})S_0 - (\mu + \delta_{A_1})]$$

(we may replace δ_A by δ_{A_1} because we sum over μ)

= dimension of the L_2-module with the restriction to the Cartan subalgebra of L_2 of $(\Lambda + \delta_A)S - \delta_{A_1}$ as highest weight, if this weight is dominant with respect to L_2.

It is easy to see that there is only one S in every coset such that $(\Lambda + \delta_A)S - \delta_{A_1}$ is dominant with respect to L_2. Suppose that

$$(\Lambda + \delta_A)S - \delta_{A_1} \quad \text{and} \quad (\Lambda + \delta_A)SS_0 - \delta_{A_1}$$

are both dominant; then so are $(\Lambda + \delta_A)S$ and $(\Lambda + \delta_A)SS_0$, and it follows that $S_0 = 1$ (Ref. 5, p. 262).

As a result of the preceding considerations, we may write

$$\bar{n}(\Lambda, m) = \sum_{S \in W} (\det S) P_B((\Lambda + \delta_A)S(J) - (m + \delta_B)) \times \dim [(\Lambda + \delta_A)S - \delta_{A_1}]$$

provided that we define $\dim [\lambda] = 0$ if λ is not dominant: otherwise,

$$\dim [\lambda] = \sum_{\mu} \sum_{W_{L_2}} (\det S_0) P_{A_1}((\lambda + \delta_{A_1})S_0 - (\mu + \delta_{A_1})).$$

QED

Theorem 2: The multiplicity of the irreducible K-submodules with highest weight $M \in J$ in the restriction to K of the irreducible L-module with highest weight Λ equals

$$N[\Lambda, M] = \sum_{S \in W} (\det S) P_D((\Lambda + \delta_A)S(J) - (M + \delta_B)) \times \dim [(\Lambda + \delta_A)S - \delta_{A_1}], \quad (3)$$

where $D = B \setminus C$, C the positive roots of K, and B, A, A_1, and $\dim [\lambda]$ are as in Theorem 1.

Proof: As we already mentioned in Sec. II, the equation

$$\bar{n}(\Lambda, m) = \sum_{M \text{ dominant}} N[\Lambda, M] n(M, m), \quad (4)$$

with

$$n(M, m) = \sum_{S_K \in W_K} (\det S_K) P_C((M + \delta_C)S_K - (m + \delta_C)),$$

holds for all $m \in J^*$. Theorem 1 provides us another formula for $\bar{n}(\Lambda, m)$, but we are not yet able to use

Property 3: On the one hand, we have a partition function over B and, on the other hand, a partition function over C. Obviously, C is contained in B because an element β of B is a positive solution of the set of eigenvalue equations

$$\forall\, h \in J, \quad [h, x] = \beta(h)x, \quad x \in L,$$

and an element of C is a similar solution, but with x in the subspace K of L.

This suggests that we should use Property 1 again, to arrive at the same partition function on both sides, and then use Property 3. So let $D = B \setminus C$ and let us rewrite formula (2) using Property 1:

$$\tilde{n}(\Lambda, m) = \sum_{S \in W} (\det S) \dim [(\Lambda + \delta_A)S - \delta_{A_1}] \times \sum_{m_1} P_D((\Lambda + \delta_A)S(J) - m_1)P_C(m_1 - (m + \delta_B)).$$

By a redefinition of m_1 and a re-ordering of the terms, we get

$$\sum_{m_1} P_C(m_1 - (m + \delta_C)) \times \sum_{S \in W} (\det S) \dim [(\Lambda + \delta_A)S - \delta_{A_1}] \times P_D((\Lambda + \delta_A)S(J) - (m_1 + \delta_D)).$$

Here we have a summation over $m_1 \in J^*$, in formula (4) we have a summation over M dominant and over $S_K \in W_K$. In fact, both are summations over the whole space J^*: Consider the subset

$$\mathfrak{D} = \{(M + \delta_C)S_K, S_K \in W_K, M \text{ dominant}\}.$$

All these weights are different; suppose that two of them are equal:

$$(M + \delta_C)S_K = (M' + \delta_C)S'_K.$$

Then

$$(M + \delta_C)S_K S'^{-1}_K = (M' + \delta_C)$$

and, because M and M' are dominant,

$$S_K S'^{-1}_K = 1 \quad \text{(Ref. 5, p. 262)}$$

so that

$$M + \delta_C = M' + \delta_C.$$

It follows that with every element $m_1 \in \mathfrak{D}$ we can associate a unique $S_K(m_1)$ and $M(m_1)$ such that $M(m_1)$ is dominant and

$$m_1 = (M(m_1) + \delta_C)S_K(m_1).$$

Now we can rewrite formula (4):

$$\tilde{n}(\Lambda, m) = \sum_{m_1 \in \mathfrak{D}} P_C(m_1 - (m + \delta_C)) \det S_K(m_1) N(\Lambda, M(m_1)),$$

and we can extend the summation over the whole space, because, if $m_1 \notin \mathfrak{D}$, then we may put

$$N(\Lambda, M(m_1)) = 0.$$

Equating the two results for $\tilde{n}(\Lambda, m)$, we have

$$\sum_{m_1} P_C(m_1 - (m + \delta_C))X(m_1) = 0,$$

with

$$X(m_1) = [\det S_K(m_1)]N(\Lambda, M(m_1)) - \sum_{S \in W} (\det S) \dim [(\Lambda + \delta_A)S - \delta_{A_1}] \times P_D((\Lambda + \delta_A)S(J) - (m_1 + \delta_D)).$$

We can now use Property 3. Therefore, we first suppose that $X(m_1) \neq 0$ for an element $m_1 = n_0$, and so there are infinitely many points of the form $n = n_0 + \sum_{\alpha \in C} x_\alpha \alpha$, $x_\alpha \geq 0$, such that $X(n) \neq 0$.

On the other side, we know that $N(\Lambda, M(m_1))$ is only different from zero for a finite set of points M and so for a finite set of points m_1. Therefore, there should be infinitely many points such that the second part of the sum is different from zero, and, since W is finite, this should be true for some element S of W. But $P_D((\Lambda + \delta_A)S(J) - (m_1 + \delta_D))$ can only be different from zero if

$$(\Lambda + \delta_A)S(J) - (m_1 + \delta_D) = \sum_{\alpha \in D} x_\alpha \alpha.$$

It follows that there are infinitely many points m_1 of the form

$$m_1 = n_0 + \sum_{\alpha \in C} x_\alpha \alpha = (\Lambda + \delta_A)S(J) - \delta_D - \sum_{\alpha \in D} x_\alpha \alpha,$$

and so we have infinitely many solutions of

$$n_0 + \sum_{\alpha \in C} x_\alpha \alpha + \sum_{\alpha \in D} x_\alpha \alpha = (\Lambda + \delta_A)S(J) - \delta_D.$$

Thus, $P_B((\Lambda + \delta_A)S(J) - (n_0 + \delta_D))$ would be infinite, which is impossible since B is a finite set of positive roots.

We conclude by Property 3 that $X(m_1) = 0$ for all m_1, and, if we take a point m_1 with $S_K(m_1) = 1$, we get

$$N(\Lambda, M) = \sum_{S \in W} (\det S) \dim [(\Lambda + \delta_A)S - \delta_{A_1}] \times P_D((\Lambda + \delta_A)S(J) - (M + \delta_C + \delta_D)),$$

and, since $\delta_C + \delta_D = \delta_B$, we have proved formula (3).

Remarks:

(i) In formula (3) the sum over $S \in W$ runs effectively over those elements such that

$(\Lambda + \delta_A)S - \delta_{A_1}$ is dominant with respect to L_2

and

$$(\Lambda + \delta_A)S(J) - \delta_B \geq M.$$

Because $\Lambda + \delta_A \geq (\Lambda + \delta_A)S$, we have that

$$\Lambda(J) \geq (\Lambda + \delta_A)S(J) - \delta_B.$$

Therefore, there is, in fact, only summed over a subset of the cosets W/W_0 a subset, which is, in many cases, small with respect to the whole W. This is, of course, a great advantage.

(ii) Applying formula (3), one should take care of the fact that A, the set of positive roots of L, when restricted to J must yield positive or zero weights with respect to the ordering in J^*.

ACKNOWLEDGMENTS

I am indebted to Dr. A. Verbeure for helpful discussions. Thanks are also due to Professor L. Bouckaert, Professor F. Cerulus, and Professor D. Speiser for their interest in this work.

APPENDIX

As a first illustration of the applicability, we treat the nontrivial case of the subalgebra A_1 of the simple Lie algebra G_2, encountered in nuclear and atomic physics (see, e.g., Ref. 8, p. 106).

We first determine the various elements that are to be used in the formula. We consider the canonical generators of G_2, $\{h_i, e_i, f_i;\ i = 1, 2\}$, and we denote the elements Λ of H^* by

$$(\lambda_1, \lambda_2) = (\Lambda(h_1), \Lambda(h_2)).$$

In this notation we have that $\delta_A = (1, 1)$, and the different possibilities for $(\lambda_1, \lambda_2)S$ and $\det S$ are (see Ref. 5, p. 235)

$(\lambda_1, \lambda_2)^+$, $(-\lambda_1, \lambda_2 + 3\lambda_1)^-$, $(\lambda_1 + \lambda_2, -\lambda_2)^-$,
$(2\lambda_1 + \lambda_2, -\lambda_2 - 3\lambda_1)^+$, $(-\lambda_1 - \lambda_2, 3\lambda_1 + 2\lambda_2)^+$,
$(-2\lambda_1 - \lambda_2, 3\lambda_1 + 2\lambda_2)^-$, $(2\lambda_1 + \lambda_2, -3\lambda_1 - 2\lambda_2)^-$,
$(\lambda_1 + \lambda_2, -3\lambda_1 - 2\lambda_2)^+$, $(-2\lambda_1 - \lambda_2, +\lambda_2 + 3\lambda_1)^+$,
$(-\lambda_1 - \lambda_2, \lambda_2)^-$, $(\lambda_1, -\lambda_2 - 3\lambda_1)^-$, $(-\lambda_1, -\lambda_2)^+$.

(The plus or minus sign indicates whether $\det S$ is $+1$ or -1.)

In this special case, the adjoint representation of G_2 reduces with respect to A_1 into a 3- and 11-dimensional representation. It follows immediately that the commutant of the Cartan subalgebra J of A_1 contains only one element apart from J itself, so that it is commutative, and that the factor

$$\dim [(\Lambda + \delta_A)S - \delta_{A_1}] \text{ is always equal to 1.}$$

Knowing the different roots of G_2, one can find that the canonical element h of J is $10h_1 + 6h_2$, and, if we denote the elements M of J^* by $l = \frac{1}{2}M(h)$, we find that the restriction of a weight (λ_1, λ_2) of H^* to J is

$$(\lambda_1, \lambda_2)(J) = 5\lambda_1 + 3\lambda_2.$$

In this notation we also find that

$$B = \{5, 4, 3, 2, 1, 1\}, \quad D = \{5, 4, 3, 2, 1\},$$

so that $\delta_B = 8$.

The different possibilities for the weight

$$((\Lambda + \delta_A)S(J) - \delta_B)$$

and $\det S$, such that this weight can be positive, are

$$(5\lambda_1 + 3\lambda_2)^+, \quad (4\lambda_1 + 3\lambda_2 - 1)^-, \quad (5\lambda_1 + 2\lambda_2 - 1)^-,$$
$$(\lambda_1 + 2\lambda_2 - 5)^+, \quad (4\lambda_1 + \lambda_2 - 3)^+,$$
$$(\lambda_2 - \lambda_1 - 8)^-, \quad (|\lambda_1 - \lambda_2| - 8)^-.$$

Using all these results and the fact that $P_D(n) = 0$ if $n < 0$, we are able to write down our formula as

$$\begin{aligned} n[(\lambda_1, \lambda_2), l] = {} & P_D(5\lambda_1 + 3\lambda_2 - l) \\ & - P_D(4\lambda_1 + 3\lambda_2 - 1 - l) \\ & - P_D(5\lambda_1 + 2\lambda_2 - 1 - l) \\ & + P_D(\lambda_1 + 2\lambda_2 - 5 - l) \\ & + P_D(4\lambda_1 + \lambda_2 - 3 - l) \\ & - P_D(|\lambda_1 - \lambda_2| - 8 - l), \end{aligned}$$

and the first few values of P_D are as follows:

n:	0	1	2	3	4	5	6	7	$8 \cdots$;
$P_D(n)$:	1	1	2	3	5	7	10	13	$18 \cdots$.

Having done all this preliminary work, we now find it easy to decompose an arbitrary irreducible G_2-module with respect to A_1. Consider, e.g., the G_2-module of highest weight $\Lambda = (1, 1)$ and dimension 64. Our formula becomes

$$n[(1, 1), l] = P_D(8 - l) - 2P_D(6 - l) + P_D(2 - l).$$

The result is

$l =$	0	1	2	3	4	5	6	7	8,
$n[(1, 1), l] =$	0	0	1	1	1	1	0	1	1,

in agreement with Ref. 8, p. 147, where the irreducible representations of G_2 are labeled by $(\lambda_1 + \lambda_2, \lambda_1)$.

As a second example, we take the chain $A_3 \supset A_1 \oplus A_1$ ($SU_4 \supset SO_4$ in group theoretical notations). If $\{h_i, e_i, f_i, i = 1, 2, 3\}$ is a set of canonical generators of A_3, we denote the elements Λ of H^* by[9]

$$[\lambda_1\, \lambda_2\, \lambda_3\, 0] = [\Lambda(h_1 + h_2 + h_3)\Lambda(h_2 + h_3)\Lambda(h_3)0].$$

In this notation we have that $\delta_A = [3, 2, 1, 0]$, and, if we define $[\mu_1\, \mu_2\, \mu_3\, \mu_4] = \Lambda + \delta_A$, we find that the Weyl group acts by permutations on the indices of the μ's.

For every permutation (P_1, P_2, P_3, P_4) of $(1, 2, 3, 4)$ we have a possibility for $(\Lambda + \delta_A)S$, namely

$$[\mu_{P_1} - \mu_{P_4}, \mu_{P_2} - \mu_{P_4}, \mu_{P_3} - \mu_{P_4}, 0].$$

In order to satisfy the relations between the positive elements of H^* and J^* (see Lemma 1), we choose the canonical elements of J, the Cartan subalgebra of $A_1 \oplus A_1$, as

$$h_1 + 2h_2 + h_3 \quad \text{and} \quad h_1 + h_3,$$

and we denote the elements M of J^* as

$$(m_1, m_2) = (M(h_1 + 2h_2 + h_3), M(h_1 + h_3)),$$

so that the restriction of an element Λ of H^* to J becomes

$$(\lambda_1 + \lambda_2 - \lambda_3, \lambda_1 - \lambda_2 + \lambda_3).$$

The sets B and D are

$$B = \{(2, 2), (2, 0), (2, 0), (2, -2), (0, 2), (0, 2)\},$$
$$D = \{(2, 2), (2, 0), (2, -2), (0, 2)\},$$

and we have that $\delta_B = (4, 2)$.

Clearly, the dimension factor is always equal to 1, so that our formula becomes

$$N[[\lambda_1\lambda_2\lambda_3 0], (m_1, m_2)] = \sum_P (-1)^P P_D((\mu_{P_1} + \mu_{P_2} - \mu_{P_3} - \mu_{P_4} - m_1 - 4, \mu_{P_1} - \mu_{P_2} + \mu_{P_3} - \mu_{P_4} - m_2 - 2)),$$

where $(-1)^P$ is $+1$ or -1 whether the permutation P is even or odd.

We only require results for those multiplicities $N[\Lambda, M]$ such that M is dominant, which means that m_1 and m_2 are not negative, and from the knowledge of D we may conclude that only those permutations must be taken into account for which either

$$\mu_{P_1} + \mu_{P_2} - \mu_{P_3} - \mu_{P_4} - 4 = k > 0$$

and

$$\mu_{P_1} - \mu_{P_2} + \mu_{P_3} - \mu_{P_4} - 2 \geq -k$$

or

$$\mu_{P_1} + \mu_{P_2} - \mu_{P_3} - \mu_{P_4} = 4$$

and

$$\mu_{P_1} - \mu_{P_2} + \mu_{P_3} - \mu_{P_4} - 2 \geq 0.$$

Let us now turn to an explicit example: Consider the regular representation of A_3 with $\Lambda = [2\ 1\ 1\ 0]$. We have that $\Lambda + \delta_A = [5\ 3\ 2\ 0]$. The only possible permutations satisfying the above restrictions are

$$[5\ 3\ 2\ 0], \quad [3\ 5\ 2\ 0], \quad [5\ 3\ 0\ 2], \quad [5\ 2\ 3\ 0],$$

so that the result is

$$N[[2\ 1\ 1\ 0], (m_1, m_2)] = P_D((2 - m_1, 2 - m_2)) - P_D((-m_1, 4 - m_2)) - 2P_D((2 - m_1, -2 - m_2)).$$

And we get by a simple calculation the following table.

TABLE AI.

(m_1, m_2)	$P_D((2 - m_1, 2 - m_2))$	$P_D((-m_1, 4 - m_2))$	$P_D((2 - m_1, -2 - m_2))$	N
(2, 2)	1	0	0	1
(2, 0)	1	0	0	1
(0, 4)	1	1	0	0
(0, 2)	2	1	0	1
(0, 0)	3	1	1	0

We can easily check the relation between the dimensions:

$$15 = \sum (m_1 + 1)(m_2 + 1).$$

* Aspirant N.F.W.O.
† Postal address: Instituut voor Theoretische Fysica, Departement Natuurkunde, Celestijnenlaan, 200 D, 3030 Heverlee, Belgium.

[1] N. Straumann, Helv. Phys. Acta **38**, 481 (1965).
[2] R. M. Delaney and B. Gruber, J. Math. Phys. **10**, 252 (1969).
[3] V. B. Mandel'tsveig, Yad. Fiz. **1**, 1106 (1965) [Sov. J. Nucl. Phys. **1**, 787 (1965)].
[4] B. Kostant, Trans. Am. Math. Soc. **93**, 53 (1959).
[5] N. Jacobson, *Lie Algebras*, Tracts in Mathematics N. 10 (Interscience, New York, 1962).
[6] D. Radhakrishnan, J. Math. Phys. **10**, 2129 (1969).
[7] For the definition of a dominant weight, see Ref. 5.
[8] B. R. Judd, *Operator Techniques in Atomic Spectroscopy*, Advanced Physics Monograph Series (McGraw-Hill, New York, 1963).
[9] This notation is the usual Young partition for SU_4.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Cell Model Field Theories

HANS GRIMMER
Advanced Studies Center, Battelle Institute, Carouge-Geneva, Switzerland

(Received 31 March 1970)

We investigate a class of nonrelativistic, self-interacting model field theories, assuming the existence of field and momentum operators that satisfy the canonical commutation relations (CCR's). The large symmetry that we assume permits us to determine explicitly all the relevant CCR representations and Hamiltonians. For each irreducible representation there exists only one Hamiltonian, which describes no interaction. Hence every model with interaction requires a reducible CCR representation. We expect a nontrivial S operator to exist for many of these models. We examine how close our models can come to relativistic theories.

1. INTRODUCTION

No models describing interacting particles and satisfying all the Wightman axioms are known at the present time. Recently Glimm and Jaffe constructed a model that fulfills many of the axioms but violates a fundamental one: Their space has one dimension, not three. There is not much hope that this shortcoming can be overcome with their techniques. To get a better insight into the difficulties and to find the best approach to our goal of finding realistic theories of interacting fields, it is imperative to look also for interacting models in three space dimensions that satisfy some of the Wightman axioms. This paper is a contribution to this program.[1]

We start by writing down a set of basic assumptions containing a symmetry requirement that we illustrate before by considering a class of classical Hamiltonians. This symmetry is the only manifestly unrealistic assumption that we shall make. It permits us to find all models that satisfy the assumptions, and this is so even though our list of axioms is much shorter than Wightman's. The solutions are found by mathematical techniques that avoid boxes, momentum space cutoffs, and divergencies. They are simple compared to the sophisticated ones employed by Glimm and Jaffe in the construction of their model.

This paper is based on work by Klauder on "rotationally symmetric" models.[2,3] It avoids a number of their unphysical features while preserving the possibility of deriving strong and explicit results by simple methods. Rotationally symmetric models have Hamiltonians with a discrete spectrum only (leading to a trivial S matrix) and with interactions involving the relation between the values of the field at points with equal time, yet arbitrarily large spatial separation. Our models will not show these undesirable features.

The noninteracting ones among our models are similar to some that have been considered in statistical mechanics.[4,5]

In Sec. 2, after stating the basic assumptions, we give our main results, grouped into two theorems. We discuss their physical interpretation and a special model, which shows how close solutions satisfying our assumptions can come to realistic theories. Sections 3 and 4 are devoted to the proofs of the two theorems.

2. FORMULATION OF THE PROBLEM, SUMMARY AND DISCUSSION OF THE RESULTS

A. The Classical Models

We want to find the quantum theories corresponding in some sense to the classical theories with Hamiltonians of the form

$$H(f, g) = \tfrac{1}{2}[(f, f) + V\{(g_i, g_j)\}], \tag{1}$$

where the classical field g and momentum f are elements of some linear manifold, which we shall finally choose as $L_r^2(R_3)$, the space of real-valued square-integrable functions with arguments in the 3-dimensional Euclidean space R_3. (f, f) denotes the scalar product in $L^2(R_3)$. V depends on infinitely many arguments, labeled by two triples of integers, $\mathbf{i}$, $\mathbf{j} \in Z_3$. We have

$$(h_i', h_j) = \int_c h'^*(\mathbf{i} + \mathbf{x}) h(\mathbf{j} + \mathbf{x})\, d\mathbf{x}, \quad h', h \in L^2(R_3), \tag{2}$$

where c denotes the integration over the unit cube whose center is at the origin of a Cartesian coordinate system in R_3 and whose edges are parallel to the coordinate axes.

As an illustration, consider how close such a Hamiltonian can be to the one for the relativistic scalar field with quartic self-interaction,

$$H_r(f, g) = \tfrac{1}{2}\int d\mathbf{x}[f^2(\mathbf{x}) + (\nabla g)^2(\mathbf{x}) + m^2 g^2(\mathbf{x}) + \lambda g^4(\mathbf{x})]. \tag{3}$$

We have to approximate $(\nabla g)(\mathbf{x})$ by $(g(\mathbf{x}+\mathbf{e}_1)-g(\mathbf{x}), g(\mathbf{x}+\mathbf{e}_2)-g(\mathbf{x}), g(\mathbf{x}+\mathbf{e}_3)-g(\mathbf{x}))$, where $\mathbf{e}_i$ is a unit vector parallel to the ith coordinate axis. Choosing a very small length unit, we should get a good approximation:

$$H_a(f,g) = \tfrac{1}{2}\Big[(f,f) + 2\Big(\sum_i (g_i, 3g_i - g_{i+e_1} - g_{i+e_2} - g_{i+e_3})\Big) + m^2(g,g) + \lambda \sum_i (g_i, g_i)^2\Big]. \quad (4)$$

Notice that the interaction is no longer local, but, by choosing a length unit of 10^{-16} cm, relations between the values of the field g at points more than $\sqrt{3}\times 10^{-16}$ cm apart will have no effect, in agreement with experiment. Later I shall return to this example.

We denote by $\mathfrak{T}'$ the group of real linear transformations T that satisfy $H(f,g) = H(Tf, Tg)$ for all H of the form (1). $\mathfrak{T}'$ does not contain translations. In agreement with our division of space into cubes, we content ourselves with invariance of H under the group $\mathfrak{T}''$ of "lattice" translations, whose elements $T_{\mathbf{n}}$, $\mathbf{n}\in Z_3$, are defined by

$$f(\mathbf{x}) = (T_{\mathbf{n}}f)(\mathbf{x}+\mathbf{n}).$$

This requires

$$V\{(g_i, g_j)\} = V\{((T_{\mathbf{n}}g)_i, (T_{\mathbf{n}}g)_j)\} \quad \text{for all } \mathbf{n}\in Z_3. \quad (5)$$

If we restrict the set of $H(f,g)$ by imposing (5), then $H(f,g) = H(Tf,Tg)$ will remain true for $T\in\mathfrak{T} = \mathfrak{T}'\otimes\mathfrak{T}''$. $H(f,g)$ is also invariant under "time reversal," i.e., under the replacement $f\to -f$, $g\to g$.

B. The Basic Assumptions

1. Cyclic Representations of the Canonical Commutation Relations

For all $f\in L_r^2(R_3)$, there exist two self-adjoint operators $\phi(f)$ and $\pi(f)$ acting on a separable Hilbert space $\mathfrak{H}$ with positive-definite metric and satisfying

$$\phi(cf) = c\phi(f), \quad \pi(cf) = c\pi(f), \quad \text{for all } c\in R, \quad (6)$$

such that

$$U[f,g] \equiv \exp\{i[\phi(f) - \pi(g)]\} \quad (7)$$

fulfills

$$U[f',g']U[f,g] = \exp\{i\tfrac{1}{2}[(f',g) - (g',f)]\} \times U[f'+f, g'+g]. \quad (8)$$

There exists a vector $\Phi_0\in\mathfrak{H}$, $\|\Phi_0\| = 1$, such that the set σ of vectors

$$\Phi[f,g] \equiv U[f,g]\Phi_0$$

is total, i.e., the finite linear combinations of vectors in σ are dense in $\mathfrak{H}$.

Since the $U[f,g]$ are unitary, no domain questions arise in (8), which is the Weyl form of the CCR's. One easily deduces from (8) the Heisenberg form of the CCR's, $[\phi(f), \pi(g)] = i(f,g)$. We shall avoid the use of $\phi(\mathbf{x})$, which can formally be introduced by writing $\phi(f) = \int \phi(\mathbf{x})f(\mathbf{x})\,d\mathbf{x}$, because only the smeared quantities $\phi(f)$ and $\pi(g)$ are operators, a term that we shall reserve for densely defined linear transformations. The last assumption does not imply that every bounded operator that commutes with all the $U[f,g]$ is a multiple of the identity; i.e., besides the irreducible representations of the CCR's, some reducible ones are also cyclic.

2. Existence and Uniqueness of an Invariant Vector

For all $T\in\mathfrak{T}$ there exists a unitary transformation $U[T]$ with the property

$$U[T]\Phi[f,g] = \Phi[Tf, Tg], \quad (9)$$

and there exists an antiunitary transformation $\mathfrak{J}$ such that

$$\mathfrak{J}\Phi[f,g] = \Phi[-f,g]. \quad (10)$$

Up to a constant, there is only one vector that is invariant under all the $U[T]$ with $T\in\mathfrak{T}'$.

3. Properties of the Hamiltonian

There exists a self-adjoint operator $\mathcal{H}\geq 0$ such that under suitable domain conditions

$$[U[T], \mathcal{H}] = 0, \quad \text{for all} \quad T\in\mathfrak{T}, \quad (11)$$

and

$$[\phi(f), \mathcal{H}] = i\pi(f). \quad (12)$$

Up to a constant,

$$\mathcal{H}\Phi = 0 \quad (13)$$

has a unique solution.

Equation (11) expresses the assumed symmetry of $\mathcal{H}$, and Eq. (12) was gained by the usual replacement of the Poisson brackets of the classical theory by commutators. The ground state of $\mathcal{H}$, which we introduced in (13), coincides with the vacuum state Φ_0.

Proof: $0 = \mathcal{H}\Phi = U[T]\mathcal{H}\Phi = \mathcal{H}U[T]\Phi$. Hence

$$U[T]\Phi = \alpha(T)\Phi,$$

where $\alpha(T)$ is a complex number. Equation (7) requires $U[T]U[T'] = U[TT']$. The $\alpha(T)$ form therefore a 1-dimensional representation of $\mathfrak{T}'$. However, $\mathfrak{T}'$ has only the trivial 1-dimensional representation. According to the last of the assumptions in subsection 2B2, this implies $\Phi = c\Phi_0$.

The above axioms suffice to determine all solutions without too much difficulty. Instead of adding other, physically motivated assumptions, we shall first search for the solutions and then single out special cases with additional welcome properties.

C. The Main Results

The first theorem deals with the vacuum expectation value of $U[f,g]$, $(\Phi_0, U[f,g]\Phi_0) = (\Phi_0, \Phi[f,g])$, which we shall call the reproducing kernel. It determines $(\Phi[f',g'], \Phi[f,g])$ because of (8). For $f(\mathbf{x}) \in L^2(R_3)$, we define a partial Fourier transform $\tilde{f}(\mathbf{x},\mathbf{k})$ by

$$f(\mathbf{x}+\mathbf{l}) = \int_D \hat{d}\mathbf{k} \tilde{f}(\mathbf{x},\mathbf{k}) e^{i\mathbf{k}\mathbf{l}}, \tag{14}$$

where $\mathbf{x} \in c$, $\mathbf{l} \in Z_3$, $\int_D = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}$, and $\hat{d}\mathbf{k} = (2\pi)^{-3}\,d\mathbf{k}$.

Theorem 1: For a solution of assumptions 1 and 2 (Sec. 2B), it is necessary and sufficient that

$$(\Phi_0, \Phi[f,g]) = \exp\left(-\tfrac{1}{4}\int_D \hat{d}\mathbf{k}\int_c d\mathbf{x}[\xi(\mathbf{k})m^{-1}(\mathbf{k})\,|\tilde{f}(\mathbf{x},\mathbf{k})|^2 + m(\mathbf{k})\,|\tilde{g}(\mathbf{x},\mathbf{k})|^2]\right), \tag{15}$$

where[6] $C^{-1} < m(\mathbf{k}) = m(-\mathbf{k}) < C$ and $1 \leq \xi(\mathbf{k}) = \xi(-\mathbf{k}) < C$ for some constant C. Two CCR representations are equivalent if and only if their corresponding pairs $\{m(\mathbf{k}), \xi(\mathbf{k})\}$ are equal. The representation is irreducible if $\xi(\mathbf{k}) \equiv 1$ and reducible otherwise.

The second theorem deals with the Hamiltonian. I have found that, for the irreducible representations, there exists one and only one $\mathcal{H}$ satisfying assumptions 3 (Sec. 2B); for the reducible ones, there are infinitely many. In general, an equivalence class of representations will therefore not uniquely determine the functional $(\Phi[f',g'], \mathcal{H}\Phi[f,g])$. For its form, I have derived only a necessary and a sufficient condition, which do not coincide. In formulating the results, we shall make use of the following abbreviations:

$$(\tilde{h}'(\mathbf{k}'), \tilde{h}(\mathbf{k})) = \int_c d\mathbf{x}\tilde{h}'^*(\mathbf{x},\mathbf{k}')\tilde{h}(\mathbf{x},\mathbf{k}) \tag{16}$$

and

$$\tilde{G}(\mathbf{x},\mathbf{k}) = \{\tfrac{1}{2}m(\mathbf{k})[\xi(\mathbf{k})-1]\xi^{-1}(\mathbf{k})\}^{\frac{1}{2}}\tilde{g}(\mathbf{x},\mathbf{k}).$$

Notice that $\tilde{G}(\mathbf{x},\mathbf{k}) \equiv 0$ for irreducible representations.

Theorem 2: The σ matrix elements of the Hamiltonians $\mathcal{H}$ belonging to a reproducing kernel (15) and satisfying assumptions 3 have the form

$$(\Phi[f',g'], \mathcal{H}\Phi[f,g]) = \tfrac{1}{2}(\Phi[f',g'], \Phi[f,g]) \times \left(\int_D [(\tilde{f}'(\mathbf{k}) - im(\mathbf{k})\tilde{g}'(\mathbf{k}), \tilde{f}(\mathbf{k}) - im(\mathbf{k})\tilde{g}(\mathbf{k})) + 2e(\mathbf{k})(\tilde{G}'(\mathbf{k}), \tilde{G}(\mathbf{k}))]\,\hat{d}\mathbf{k} + F\{(\tilde{G}'(\mathbf{k}'), \tilde{G}(\mathbf{k}))\}\right) \tag{17}$$

with

$$F\{(\tilde{G}'(\mathbf{k}'), \tilde{G}(\mathbf{k}))\} = \sum_{n=2}^{\infty}\frac{2}{n!}\left(\prod_{i=1}^{n}\int_D\int_D \hat{d}\mathbf{k}_i'\,\hat{d}\mathbf{k}_i(\tilde{G}'(\mathbf{k}_i'), \tilde{G}(\mathbf{k}_i))\right) \cdot b(\mathbf{k}_1', \mathbf{k}_1, \cdots, \mathbf{k}_n', \mathbf{k}_n), \tag{18}$$

$$b(\mathbf{k}_1', \mathbf{k}_1, \cdots, \mathbf{k}_n', \mathbf{k}_n) = b^*(\mathbf{k}_1, \mathbf{k}_1', \cdots, \mathbf{k}_n, \mathbf{k}_n') \tag{19}$$

$$= \hat{b}(\mathbf{k}_1', \mathbf{k}_1, \cdots, \mathbf{k}_n') \cdot (2\pi)^3 \sum_{\mathbf{l}\in Z_3}\delta\left(2\pi\mathbf{l} + \sum_{i=1}^{n}\mathbf{k}_i' - \mathbf{k}_i\right), \tag{20}$$

and

$$e(\mathbf{k}) > 0 \quad \text{for all } \mathbf{k} \text{ with } \quad \xi(\mathbf{k}) > 1.$$

I can prove that such $\mathcal{H}$ exist if $e(\mathbf{k}) < C$ and if

$$F\{(\tilde{G}'(\mathbf{k}'), \tilde{G}(\mathbf{k}))\} = \sum_{\mathbf{M}} b_{\mathbf{M}}\int_D\int_D \hat{d}\mathbf{k}_1'\,\hat{d}\mathbf{k}_1\beta(\mathbf{k}_1')\beta^*(\mathbf{k}_1) \times \exp[i\mathbf{M}(\mathbf{k}_1' - \mathbf{k}_1)](\tilde{G}'(\mathbf{k}_1'), \tilde{G}(\mathbf{k}_1)) \times \int_D \hat{d}\mathbf{k}_2'\beta(\mathbf{k}_2')\beta^*(\mathbf{k}_1' + \mathbf{k}_2' - \mathbf{k}_1) \times (\tilde{G}'(\mathbf{k}_2'), \tilde{G}(\mathbf{k}_1' + \mathbf{k}_2' - \mathbf{k}_1)) \tag{21}$$

with $|\beta(\mathbf{k})| < C$ and $0 \leq b_{\mathbf{M}} = b_{-\mathbf{M}} < C$ for some C.

The necessary conditions above do not ensure $\mathcal{H} \geq 0$. On the other hand, it can easily be shown that we may have terms of higher than fourth order in G' and G.

Given a classical Hamiltonian $H(f,g)$, one usually replaces f and g by operators π and ϕ satisfying the CCR's and imposes normal ordering. We shall call this prescription the correspondence principle (CP). Then[7]

$$(\Phi[f,g], \mathcal{H}\Phi[f,g]) = H(f,g). \tag{22}$$

An argument due to Araki[7] shows that, for reducible representations, the Hamiltonians of Theorem 2 cannot be constructed as functions of ϕ and π alone. The CP gives $\mathcal{H}$ expressed in terms of ϕ and π, thus

restricting us to the rather trivial irreducible representations. We abandon therefore the CP despite its important role in the historical development of quantum mechanics. That the CP is often a too strong requirement is indicated also by the work of Arens and Babbitt.[8] Our basic assumptions show that we maintain the connection between commutators and Poisson brackets of ϕ and π and of ϕ and $\mathcal{H}$. We maintain also (22). The right side of (17) has, for diagonal matrix elements, the form

$$\tfrac{1}{2}[(f,f) + V\{(\tilde{g}(\mathbf{k}'), \tilde{g}(\mathbf{k}))\}], \tag{23}$$

which is equivalent to the right side of (1). The restriction (20) ensures that (5) is satisfied. Not all classical theories satisfying (1) and (5) have a quantum counterpart that fulfills assumptions 1–3 because these assumptions lead to further properties of V, e.g., $V = V^*$. We shall not investigate further the relation between classical and corresponding quantum theories. This could be done along the lines described by Klauder in an article on the weak correspondence principle (WCP).[9] The relation (22) explains why, for the smearing functions of the field (and momentum) operators, we found it convenient to use the same letters f (and g) as for the classical momentum (and field).

D. The Two-Point Function

If $\mathcal{H}$ has the form (17) with $F \equiv 0$, one finds

$$(\Phi[f', g'], e^{-it\mathcal{H}}\Phi[f, g]) = N'N \exp\left(\int_D \hat{d}\mathbf{k} \sum_{\alpha=-,+} e^{-itm_\alpha(\mathbf{k})}(\tilde{h}'_\alpha(\mathbf{k}), \tilde{h}_\alpha(\mathbf{k}))\right), \tag{24}$$

where

$$N' = (\Phi[f', g'], \Phi_0), \quad N = (\Phi_0, \Phi[f, g]),$$

$$m_\pm(\mathbf{k}) = \tfrac{1}{2}(m(\mathbf{k}) + e(\mathbf{k}) \pm \{[m(\mathbf{k}) + e(\mathbf{k})]^2 - 4m(\mathbf{k})e(\mathbf{k})\xi^{-1}(\mathbf{k})\}^{\frac{1}{2}}), \tag{25}$$

$$\tilde{h}_\pm(\mathbf{x}, \mathbf{k}) = [\tfrac{1}{2}\rho_\pm(\mathbf{k})m_\pm^{-1}(\mathbf{k})]^{\frac{1}{2}}[\tilde{f}(\mathbf{x}, \mathbf{k}) - im_\pm(\mathbf{k})\tilde{g}(\mathbf{x}, \mathbf{k})], \tag{26}$$

and

$$\rho_\pm(\mathbf{k}) = \pm m_\pm(\mathbf{k})[m_+(\mathbf{k}) - m_-(\mathbf{k})]^{-1} \times [1 - \xi(\mathbf{k})m_\mp(\mathbf{k})m^{-1}(\mathbf{k})]. \tag{27}$$

We use this result to calculate the two-point function

$$(\Phi_0, \phi(f', t' + t)\phi(f, t')\Phi_0) = (\Phi_0, \phi(f')e^{-it\mathcal{H}}\phi(f)\Phi_0) = (\Phi_0, \phi(f', t)\phi(f)\Phi_0), \tag{28}$$

where

$$\phi(f, t) = e^{it\mathcal{H}}\phi(f)e^{-it\mathcal{H}}. \tag{29}$$

Setting $g \equiv g' \equiv 0$ and equating terms that are linear in f' and f, we conclude from (24)

$$(\Phi_0, \phi(f')e^{-it\mathcal{H}}\phi(f)\Phi_0) = \int_D \hat{d}\mathbf{k} \sum_{\alpha=-,+} \tfrac{1}{2}\rho_\alpha(\mathbf{k})m_\alpha^{-1}(\mathbf{k})e^{-itm_\alpha(\mathbf{k})}(\tilde{f}'(\mathbf{k}), \tilde{f}(\mathbf{k})). \tag{30}$$

This holds even if $F \not\equiv 0$ because F does not affect the two-point function. To bring (30) into a more familiar form, we formally calculate it for $f'(\mathbf{x}') = \delta(\mathbf{x}' - \mathbf{x})$ and $f(\mathbf{x}') = \delta(\mathbf{x}' - \mathbf{y})$:

$$\langle 0|\, \phi(\mathbf{x}, t)\phi(\mathbf{y})\, |0\rangle \equiv (\Phi_0, \phi(\mathbf{x}, t)\phi(\mathbf{y})\Phi_0) = \int \hat{d}\mathbf{k} e^{i\mathbf{k}(\mathbf{x}-\mathbf{y})} \sum_{\alpha=-,+} \tfrac{1}{2}\rho_\alpha(\mathbf{k})m_\alpha^{-1}(\mathbf{k})e^{-itm_\alpha(\mathbf{k})}. \tag{31}$$

Here the integration extends over all of R_3; $\rho_\alpha(\mathbf{k})$ and $m_\alpha(\mathbf{k})$ have been extended to functions with period D defined for all $\mathbf{k} \in R_3$. In the case of an irreducible representation of the CCR's, (31) reduces to

$$\langle 0|\, \phi(\mathbf{x}, t)\phi(\mathbf{y})\, |0\rangle = \int \hat{d}\mathbf{k} \tfrac{1}{2}m^{-1}(\mathbf{k})e^{i\mathbf{k}(\mathbf{x}-\mathbf{y})-im(\mathbf{k})t}.$$

If we replaced $m(\mathbf{k})$, which is periodic in k_1, k_2, and k_3, by $\omega_\mathbf{k} = (m^2 + |\mathbf{k}|^2)^{\frac{1}{2}}$, we would get the two-point function of a relativistic free field.

E. Physical Interpretation

The way in which $\mathbf{k}$ and $m_\pm(\mathbf{k})$ enter (31) allows us to consider them as momentum and energy. Consider a state with definite energy and a definite momentum $\mathbf{k}_0$. Two such states with different energies $m_\pm(\mathbf{k}_0)$ are created by the field with relative strengths $\rho_+(\mathbf{k}_0) > 0$ and $0 < \rho_-(\mathbf{k}_0) = 1 - \rho_+(\mathbf{k}_0)$ if $\xi(\mathbf{k}_0) > 1$, but only one is created if $\xi(\mathbf{k}_0) = 1$. One can write

$$(\Phi[f', g'], \Phi[f, g]) = N'N \exp\left(\int_D \hat{d}\mathbf{k}[(\tilde{h}'_+(\mathbf{k}), \tilde{h}_+(\mathbf{k})) + (\tilde{h}'_-(\mathbf{k}), \tilde{h}_-(\mathbf{k}))]\right). \tag{32}$$

As we shall discuss later in more detail, this implies that the abstract Hilbert space $\mathfrak{H}$ can be realized as follows:

$$\mathfrak{H} = \bigoplus_{n=0}^{\infty} \mathfrak{H}_n,$$

where $\mathfrak{H}_n$ is the symmetrized direct product of n factors $\mathfrak{H}_1$; $\mathfrak{H}_0 = C$ is spanned by Φ_0. The component of $\Phi[f, g]$ in $\mathfrak{H}_1$ is $N(\tilde{h}_+(\mathbf{x}, \mathbf{k}) \oplus \tilde{h}_-(\mathbf{x}, \mathbf{k}))$. The two-point function, $(\phi(f', t)\Phi_0, \phi(f)\Phi_0)$, is the scalar product of two vectors in $\mathfrak{H}_1$. We can consider $\mathfrak{H}_n$ as the n-particle subspace of $\mathfrak{H}$. Our models describe two kinds of particle for those values of $\mathbf{k}$ for which

$\xi(\mathbf{k}) > 1$. A wavefunction describing two or more particles will be left invariant if we exchange simultaneously the coordinates and the labels indicating the type of any two particles.

A problem of great interest, which is still under investigation, concerns the S operator. We write $\mathcal{H} = \mathcal{H}_0 + \mathcal{U}$, where $\mathcal{H}_0$ is the operator whose σ matrix elements are given by (17) with $F \equiv 0$. Note that $\mathcal{H}_0$ and $\mathcal{U}$ act in the same Hilbert space as $\mathcal{H}$. Therefore, we may try to define

$$S = \lim_{t_1 \to +\infty} \lim_{t_2 \to -\infty} e^{i\mathcal{H}_0 t_1} e^{-i\mathcal{H} t_1} e^{i\mathcal{H} t_2} e^{-i\mathcal{H}_0 t_2}.$$

Preliminary results indicate that for many nonzero $\mathcal{U}$'s the limit S exists in the strong sense, is unitary, and describes scattering and exchange of particles. Our models cannot describe production because $\mathcal{H}$ leaves each $\mathfrak{H}_n$ invariant.

F. A Special Model

We want to discuss the models satisfying assumptions 1–3 that correspond to $H_a(f, g)$ of the form (4). Before looking at the general case, we take $\lambda = 0$. Then H_a approximates the Hamiltonian of a relativistic free field. We shall show that, as in the relativistic case, there exists an irreducible CCR-representation whose Hamiltonian has the required diagonal σ matrix elements. Theorem 2 gives, for irreducible representations,

$$(\Phi[f, g], \mathcal{H}\Phi[f, g]) = \tfrac{1}{2}\left((f,f) + \int_D \hat{d}\mathbf{k} m^2(\mathbf{k})(\tilde{g}(\mathbf{k}), \tilde{g}(\mathbf{k}))\right).$$

This becomes equal to (4) if we choose

$$m^2(\mathbf{k}) = m^2 + \mathrm{f}(\mathbf{k})$$

with

$$\mathrm{f}(\mathbf{k}) \equiv 2(3 - \cos k_1 - \cos k_2 - \cos k_3),$$

where k_1, k_2, and k_3 are the components of $\mathbf{k}$. $m(\mathbf{k})$ has the properties required in Theorem 1 if $m^2 > 0$. For $|\mathbf{k}| \ll 1$, we obtain

$$m^2(\mathbf{k}) \approx m^2 + |\mathbf{k}|^2,$$

i.e., the relativistic energy–momentum relationship. We may take the momentum unit so large that $|\mathbf{k}| \ll 1$ is satisfied for all momenta in the experimentally accessible range just by choosing a very small length unit.

If we take a reducible representation, we shall desire ρ_+ to be independent of $\mathbf{k}$ and $m_\pm(\mathbf{k})$ to have the relativistic form for $|\mathbf{k}| \ll 1$. In the relativistic case, the energy is uniquely determined by the mass and the momentum. In analogy, we require $m_\pm^2(\mathbf{k}) = m_\pm^2 + \mathrm{f}(\mathbf{k})$. Elementary calculations yield the following beautiful results. Theorems 1 and 2 guarantee the existence and uniqueness of a quantum theory with $F \equiv 0$ for given values m_-, m_+, and ρ_+ that satisfy $0 < m_- < m_+$ and $0 < \rho_+ < 1$. The corresponding classical model has the form (4) with $m^2 = \rho_+ m_+^2 + \rho_- m_-^2$ and $\lambda = 0$. More specifically, the quantum theory is given by

$$m(\mathbf{k}) = \rho_+ m_+(\mathbf{k}) + \rho_- m_-(\mathbf{k}),$$
$$e(\mathbf{k}) = m_+(\mathbf{k}) + m_-(\mathbf{k}) - m(\mathbf{k}),$$

and

$$\xi(\mathbf{k}) = m(\mathbf{k})e(\mathbf{k})[m_+(\mathbf{k})m_-(\mathbf{k})]^{-1}.$$

This result holds also for other forms of $\mathrm{f}(\mathbf{k})$, except that the classical model will no longer have the form (4).

Let us turn to the case of arbitrary λ. We can write

$$\sum_{\mathbf{j}} (g_{\mathbf{j}}, g_{\mathbf{j}})^2 = \int_D \int_D \hat{d}\mathbf{k}_1' \, \hat{d}\mathbf{k}_1 \beta(\mathbf{k}_1')\beta(\mathbf{k}_1)(\tilde{G}(\mathbf{k}_1'), \tilde{G}(\mathbf{k}_1)) \times \int_D \hat{d}\mathbf{k}_2' \beta(\mathbf{k}_2')\beta(\mathbf{k}_1' + \mathbf{k}_2' - \mathbf{k}_1) \times (\tilde{G}(\mathbf{k}_2'), \tilde{G}(\mathbf{k}_1' + \mathbf{k}_2' - \mathbf{k}_1)),$$

where

$$\beta(\mathbf{k}) = \{\tfrac{1}{2} m(\mathbf{k})[\xi(\mathbf{k}) - 1]\xi^{-1}(\mathbf{k})\}^{-\frac{1}{2}}.$$

$|\beta(\mathbf{k})| < C$ is easily verified. Therefore, if $\lambda > 0$, Theorem 2 ensures the existence of a quantum theory corresponding to $H_a(f, g)$ with $m(\mathbf{k})$, $e(\mathbf{k})$, $\xi(\mathbf{k})$, and $\beta(\mathbf{k})$ given above, and $\beta_{\mathbf{M}} = \lambda\delta_{\mathbf{M},\mathbf{0}}$.

3. THE REPRODUCING KERNEL

In this section we prove Theorem 1 as follows. We begin by deriving some of the restrictions on the functional $(\Phi_0, \Phi[f, g])$ that follow from our basic assumptions. Many of the methods we use have been suggested by Ref. 2. A theorem due to Araki, which we shall state as Lemma 5, will help us to find further necessary conditions, strong enough that, for each functional that fulfills them, there exists a representation of the CCR satisfying assumptions 1 whose reproducing kernel is equal to the functional in question. It remains then to check whether the assumptions 2 are satisfied by these representations.

Functional Form and Continuity of the Reproducing Kernel

In the following we shall often label the unit cubes by elements of Z_+, the set of positive integers. (h_i, h_j) tells us to replace i and j by the corresponding $\mathbf{i}$ and $\mathbf{j}$ and to apply (2).

Lemma 1: $\Phi[f + \alpha f', g + \beta g']$ is strongly continuous in the real parameters α and β for all f, g, f', and g' in $L_r^2(R_3)$.

It suffices to prove continuity at $\alpha = \beta = 0$. We introduce the abbreviations $W[f] = U[f, 0]$ and $V[g] = U[0, g]$. Using the unitarity of the $U[f, g]$, which follows from assumptions 1, we find

$$\begin{aligned}\|\Phi[f + \alpha f', g + \beta g'] - \Phi[f, g]\| &\leq \|(W[f + \alpha f'] - W[f])\Phi_0\| \\ &\quad + \|(V[g + \beta g'] - V[g])\Phi[f, 0]\| \\ &\quad + |e^{i\frac{1}{2}(f+\alpha f', g+\beta g')} - e^{i\frac{1}{2}(f,g)}| \\ &= T_3 + T_2 + T_1.\end{aligned}$$

There exists, for each given $\epsilon > 0$, a $\delta_1(f, g, f'g') > 0$ such that $T_1 < \frac{1}{3}\epsilon$ if $|\alpha| + |\beta| < \delta_1$. We have

$$T_2 = \|(V[\beta g'] - 1)V[g]\Phi[f, 0]\|.$$

Because $V[cg] = e^{-ic\pi(g)}$ implies that $V[cg]$ is strongly continuous in the real parameter c, there exists $\delta_2(f, g, g') > 0$ such that $T_2 < \frac{1}{3}\epsilon$ if $|\beta| < \delta_2$. Similarly, one shows that there exists $\delta_3(f, f') > 0$ such that $T_3 < \frac{1}{3}\epsilon$ if $|\alpha| < \delta_3$, which completes the proof of the lemma.

We shall apply this lemma to the investigation of $(\Phi[f', g'], \Phi[f, g])$. Since

$$(\Phi[f', g'], \Phi[f, g]) = e^{-i\frac{1}{2}[(f', g) - (g', f)]}(\Phi_0, \Phi[f - f', g - g']), \quad (33)$$

it is sufficient to consider the following:

$$(\Phi_0, \Phi[f, g]) = (\Phi_0, \Phi[Tf, Tg]), \quad \text{for all} \quad T \in \mathfrak{T}, \quad (34)$$

because $U[T]\Phi_0 = \Phi_0$. It follows, even if we restrict our attention to $T \in \mathfrak{T}'$, that $(\Phi_0, \Phi[f, g])$ can depend only on (f_i, f_j), (f_i, g_j), and (g_i, g_j), $i, j \in Z_+$. We write

$$(\Phi_0, \Phi[f, g]) = K\{\alpha_{ij} = (f_i, f_j), \beta_{ij} = (f_i, g_j), \gamma_{ij} = (g_i, g_j)\}.$$

Let us prove that the reproducing kernel is continuous in each of the arguments appearing in K. We shall write $f^1, g^1 \sim f^2, g^2$ if there exists a $T \in \mathfrak{T}'$ such that $f^2 = Tf^1$ and $g^2 = Tg^1$. Consider f, g and $\hat{f}, \hat{g}$ such that $(f_i, f_j) = (\hat{f}_i, \hat{f}_j)$ and $(g_i, g_j) = (\hat{g}_i, \hat{g}_j)$ for $i \leq j$ and $(f_i, g_j) = (\hat{f}_i, \hat{g}_j)$ for all i and j with the exception that one of the scalar products with carats is ϵ^2 larger than the corresponding one without a carat. I shall prove that there exist $L_r^2(R_3)$-functions with the properties

$$\|f^1 - f^2\| + \|g^1 - g^2\| + \|\hat{f}^1 - \hat{f}^2\| + \|\hat{g}^1 - \hat{g}^2\| \leq 4\epsilon$$

and

$$f, g \sim f^1, g^1, \quad f^2, g^2 \sim \hat{f}^2, \hat{g}^2, \quad \hat{f}^1, \hat{g}^1 \sim \hat{f}, \hat{g}.$$

It follows from Lemma 1 that $(\Phi_0, \Phi[\hat{f}, \hat{g}])$ comes as close to $(\Phi_0, \Phi[f, g])$ as we like if we choose $\epsilon > 0$ small enough. If $(\hat{f}_k, \hat{f}_k) = (f_k, f_k) + \epsilon^2$ for a certain k, then we can take $f^2 = f^1 + \epsilon h_k$, $g^2 = g^1$ and $\hat{f}^2 = \hat{f}^1 = \hat{f}$, $\hat{g}^2 = \hat{g}^1 = \hat{g}$, where h_k lies in the subspace $L_r^2(k)$ of $L_r^2(R_3)$ whose elements vanish outside cell k, $(h_k, h_k) = 1$, and $(h_k, f_i) = (h_k, g_i) = 0$ for $i \in Z_+$. If we insisted on the fact that $f^1 = f$ and $g^1 = g$, h_k would not always exist. The situation is similar if $(\hat{g}_k, \hat{g}_k) = (g_k, g_k) + \epsilon^2$, but a little more complicated in the other cases. Take, e.g., $(\hat{f}_k, \hat{g}_l) = (f_k, g_l) + \epsilon^2$; consider h_k, h_l, $\hat{h}_k$, and $\hat{h}_l$ such that

$$\begin{gathered}(h_k, h_k) = (h_l, h_l) = 1, \quad (h_k, h_l) = 1, \\ (h_i, f_j) = (h_i, g_j) = 0, \\ (\hat{h}_k, \hat{h}_k) = (\hat{h}_l, \hat{h}_l) = 1, \quad (\hat{h}_k, \hat{h}_l) = 0, \\ (\hat{h}_i, f_j) = (\hat{h}_i, g_j) = 0,\end{gathered}$$

for $i = k, l$ and $j \in Z_+$. $f^2 = f^1 + \epsilon h_k$, $g^2 = g^1 + \epsilon h_l$, $\hat{f}^2 = \hat{f}^1 + \epsilon \hat{h}_k$, and $\hat{g}^2 = \hat{g}^1 + \epsilon \hat{h}_l$ have the required properties. The argument is similar in the remaining cases.

The reproducing kernel has the further properties

$$\begin{aligned}(\Phi_0, U[f, g]\Phi_0)^* &= (\Phi_0, U^{-1}[f, g]\Phi_0) \\ &= (\Phi_0, U[-f, -g]\Phi_0)\end{aligned}$$

and

$$\begin{aligned}(\Phi_0, U[f, g]\Phi_0)^* &= (\mathfrak{J}\Phi_0, \mathfrak{J}U[f, g]\Phi_0) \\ &= (\Phi_0, U[-f, g]\Phi_0).\end{aligned}$$

We collect the results in the following.

Lemma 2: Every reproducing kernel $(\Phi_0, \Phi[f, g])$ satisfying the assumptions 1 and 2 has the functional form

$$K\{\alpha_{ij} = (f_i, f_j), \beta_{ij} = (f_i, g_j), \gamma_{ij} = (g_i, g_j)\}.$$

K is a continuous function in each of its arguments and satisfies

$$\begin{aligned}K^*\{\alpha_{ij}, \beta_{ij}, \gamma_{ij}\} &= K\{\alpha_{ij}, \beta_{ij}, \gamma_{ij}\} \\ &= K\{\alpha_{ij}, -\beta_{ij}, \gamma_{ij}\}. \quad (35)\end{aligned}$$

Determination of All Reproducing Kernels

In the beginning of this subsection, we shall consider the reproducing kernel for the slightly different set of assumptions 1′ and 2′ obtained by replacing $L_r^2(R_3)$ by L everywhere in 1 and 2. L is the subspace of $L_r^2(R_3)$ whose elements have compact support. Lemma 2 can be proved as before.

Consider two sequences f^k and g^k, $k \in Z_+$, satisfying

(a) $(f^k, g) \to 0$ and $(g^k, f) \to 0$ as $k \to \infty$ for all $f, g \in L$;

(b) $U[f^k, g^k]$ converges weakly to a bounded operator A.

It follows from these two assumptions and

$$U[f^k, g^k]U[f, g] = \exp\,[i(f^k, g) - i(g^k, f)] \times U[f, g]U[f^k, g^k]$$

that

$$AU[f, g] = U[f, g]A. \tag{36}$$

We call A, which depends on the sequences f^k and g^k, a tag operator, $\|A\| \leq 1$.

In the sequel we shall repeatedly use the following theorem[10]:

Lemma 3: Let B_n be a sequence of uniformly bounded operators. If $\lim (\Phi, B_n\Psi)$ as $n \to \infty$ exists for all Φ, Ψ of a total set, the B_n will converge weakly towards an operator B and

$$(\Phi, B\Psi) = \lim_{n\to\infty} (\Phi, B_n\Psi).$$

This tells us that it is sufficient for establishing (b) to prove the existence of $\lim (\Phi[f', g'], U[f^k, g^k] \times \Phi[f, g])$ as $k \to \infty$ for all $f', g', f, g \in L$. Because of (a) and (8), we only have to look at

$$(\Phi_0, U[f + f^k, g + g^k]\Phi_0) = K\{(f_i + f_i^k, f_j + f_j^k), (f_i + f_i^k, g_j + g_j^k), (g_i + g_i^k, g_j + g_j^k)\}. \tag{37}$$

To define appropriate sequences, choose f', $g' \in L$. Let $i_1 < i_2 < \cdots < i_m$ and $i_{m+1} < \cdots < i_{m+n}$ be the numbers of the cells in which f' and g', respectively, do not vanish identically, and let $\mathbf{i}_r$ be the coordinate of the center of cell i_r. We can write

$$f'_{i_r}(\mathbf{x}) = \sum_{j=1}^{r} p_j^r u^j(\mathbf{x} - \mathbf{i}_r)$$

and

$$g'_{i_s}(\mathbf{x}) = \sum_{j=1}^{s} p_j^s u^j(\mathbf{x} - \mathbf{i}_s),$$

where $u^j(\mathbf{x}) \in L_r^2(c)$ and $(u^i, u^j) = \delta_{i,j}$ for $i, j = 1, 2, \cdots, m + n$. Complete this set of vectors to an orthonormal basis in $L_r^2(c)$. We choose the sequences

$$f_{i_r}^k(\mathbf{x}) = \sum_{j=1}^{r} p_j^r u^{j+k}(\mathbf{x} - \mathbf{i}_r)$$

and

$$g_{i_s}^k(\mathbf{x}) = \sum_{j=1}^{s} p_j^s u^{j+k}(\mathbf{x} - \mathbf{i}_s). \tag{38}$$

It is obvious that they satisfy (a), but they satisfy also (b), since the finite number of arguments in (37) that do not vanish for all k all converge:

$$(\Phi_0, U[f + f^k, g + g^k]\Phi_0) \to K\{(f_i, f_j) + (f'_i, f'_j), (f_i, g_j) + (f'_i, g'_j), (g_i, g_j) + (g'_i, g'_j)\} = (\Phi_0, U[f, g]A\Phi_0). \tag{39}$$

Because the replacement $f \to Tf$, $g \to Tg$ does not change the arguments in K, we get

$$(\Phi_0, U[-f, -g]A\Phi_0) = (\Phi[f, g], A\Phi_0) = (\Phi[Tf, Tg], A\Phi_0) = (\Phi[f, g], U^{-1}[T]A\Phi_0).$$

Since the $\Phi[f, g]$ form a total set and since only multiples of Φ_0 are invariant under the $U[T]$ for all $T \in \mathfrak{T}'$, we conclude that

$$A\Phi_0 = a\Phi_0, \tag{40}$$

where a is a number, called the tag, which can easily be determined by putting $f = g = 0$ in (39):

$$a = K\{(f'_i, f'_j), (f'_i, g'_j), (g'_i, g'_j)\}.$$

Using (39) again, we find

$$K\{(f_i, f_j) + (f'_i, f'_j), (f_i, g_j) + (f'_i, g'_j), (g_i, g_j) + (g'_i, g'_j)\} = K\{(f_i, f_j), (f_i, g_j), (g_i, g_j)\} \times K\{(f'_i, f'_j), (f'_i, g'_j), (g'_i, g'_j)\}.$$

By a lengthy but elementary argument, one shows[11] first that K is always positive and then, by making use of (35), that

$$K\{(f_i, f_j), (f_i, g_j), (g_i, g_j)\} = K\{(f_i, f_j), (g_i, g_j)\} = \exp\left(-\sum_{m,n} [A_{mn}(f_m, f_n) + C_{mn}(g_m, g_n)]\right),$$

with $A_{mn} = A_{nm} = A_{mn}^*$ and $C_{mn} = C_{nm} = C_{mn}^*$. So far, we have exploited our assumptions for $T \in \mathfrak{T}'$ only. Using (34) also for $T \in \mathfrak{T}''$ and labeling the cubes by the coordinates of their centers, we find

$$K\{(f_\mathfrak{l}, f_\mathfrak{j}), (g_\mathfrak{l}, g_\mathfrak{j})\} = \exp\,(-\sum [A_\mathfrak{m}(f_\mathfrak{n}, f_{\mathfrak{m}+\mathfrak{n}}) + C_\mathfrak{m}(g_\mathfrak{n}, g_{\mathfrak{m}+\mathfrak{n}})]), \tag{41}$$

with

$$A_\mathfrak{m} = A_{-\mathfrak{m}} = A_\mathfrak{m}^* \quad \text{and} \quad C_\mathfrak{m} = C_{-\mathfrak{m}} = C_\mathfrak{m}^*.$$

$\sum A_\mathfrak{m}(f_\mathfrak{n}, f_{\mathfrak{m}+\mathfrak{n}}) \geq 0$ for all $f \in L$ because $K \leq 1$. Therefore, $\sum A_{\mathfrak{m}-\mathfrak{n}} c_\mathfrak{m}^* c_\mathfrak{n} \geq 0$ for each sequence $\{c_\mathfrak{n} \mid \mathfrak{n} \in Z_3,\ c_\mathfrak{n} = 0$ for all but a finite set of $\mathfrak{n} \in Z_3\}$. Hence, we can apply the following variant of Bochner's theorem.

Lemma 4[12]: If $\sum B_{m-n}c_m^*c_n \geq 0$ for each sequence $\{c_n \mid n \in Z_3,\ c_n = 0$ for all but a finite set of $n \in Z_3\}$, then $B_{-n} = B_n^*$, $|B_n| \leq B_0$ for $n \in Z_3$, and the B_n are the Fourier coefficients of some positive measure on D: B_n can be written as a Riemann–Stieltjes integral,

$$B_n = \int_D \hat{d}\mu(\mathbf{x}) e^{inx},$$

where $\int_D \hat{d}\mu(\mathbf{x})$ is finite.

It follows that the reproducing kernel must have the form

$$\exp\left(-\int_D [(\tilde{f}(\mathbf{k}), \tilde{f}(\mathbf{k}))\, \hat{d}\mu(\mathbf{k}) + (\tilde{g}(\mathbf{k}), \tilde{g}(\mathbf{k}))\, \hat{d}\nu(\mathbf{k})]\right). \tag{42}$$

Using (33), we see that this gives for $(\Phi[f', g'], \Phi[f, g])$ the form

$$\begin{aligned} &K(f', g'; f, g) \\ &= N'N \exp\left(-i\tfrac{1}{2}\int_D \hat{d}\mathbf{k}[(\tilde{f}'(\mathbf{k}), \tilde{g}(\mathbf{k})) - (\tilde{g}'(\mathbf{k}), \tilde{f}(\mathbf{k}))]\right) \\ &\quad \times \exp\left(2\int_D [(\tilde{f}'(\mathbf{k}), \tilde{f}(\mathbf{k}))\, \hat{d}\mu(\mathbf{k}) \right. \\ &\quad \left. + (\tilde{g}'(\mathbf{k}), \tilde{g}(\mathbf{k}))\, \hat{d}\nu(\mathbf{k})]\right). \end{aligned} \tag{43}$$

To see under what conditions this is, in fact, equal to $(\Phi[f', g'], \Phi[f, g])$ in a theory satisfying our assumptions, we shall make use of a theorem due to Araki,[13] which can be stated as follows:

Lemma 5: The necessary and sufficient condition that a functional $K(0, 0; f, g)$, $f, g \in L$ [or $L_r^2(R_3)$], equal $(\Phi_0, \Phi[f, g])$ of a theory satisfying (1)′ [or (1)] is that, for all $f_i, g_i \in L$ [or $L_r^2(R_3)$],

$$K(0, 0; f, g)^* = K(0, 0; -f, -g), \quad K(0, 0; 0, 0) = 1,$$

$K(0, 0; sf_1 + f_2, tg_1 + g_2)$ is separately continuous in s, t, and

$$\sum_{i,j=1}^{n} c_i^* c_j K(f_i, g_i; f_j, g_j) \geq 0$$

for all integers n and for all sets of complex numbers c_i, where

$$\begin{aligned} K(f_i, g_i; f_j, g_j) = \exp\{-i\tfrac{1}{2}[(f_i, g_j) - (g_i, f_j)]\} \\ \times K(0, 0; f_j - f_i, g_j - g_i). \end{aligned}$$

Expression (42) satisfies the first three conditions. For a large class of $\mu(\mathbf{k})$, $\nu(\mathbf{k})$, it will also satisfy the last one.

We shall now return to our original assumptions. If σ is total in $\mathfrak{H}$, the set σ_L consisting of the $\Phi[f, g]$ with $f, g \in L$ will be total in some subspace $\mathfrak{H}_L \subset \mathfrak{H}$. Because all the other assumptions 1 and 2 imply the corresponding assumptions 1′ and 2′, the restrictions we have found for the reproducing kernel will also apply if we use 1 and 2. It remains to find the additional restrictions that have to be imposed on $\mu(\mathbf{k})$ and $\nu(\mathbf{k})$.

The requirement that $K(0, 0; sf, tg)$ be separately continuous in s and t for all $f, g \in L_r^2(R_3)$ requires that the exponent which appears in (42) be finite for all $f, g \in L_r^2(R_3)$. This implies that

$$\int_D \hat{d}\mu(\mathbf{k}) \int_c d\mathbf{x}\, |\tilde{f}(\mathbf{x}, \mathbf{k})|^2 < C \int_D \hat{d}\mathbf{k} \int_c d\mathbf{x}\, |\tilde{f}(\mathbf{x}, \mathbf{k})|^2$$

for some $C \in R$ and for all $f \in L_r^2(R_3)$. We can apply the Radon–Nikodym theorem[14] because for any $F(\mathbf{k}) \geq 0$ in $L^1(D)$ there exists $f \in L_r^2(R_3)$ such that

$$F(\mathbf{k}) = \int_c d\mathbf{x}\, |\tilde{f}(\mathbf{x}, \mathbf{k})|^2.$$

The theorem states that $d\mu(\mathbf{k}) = A(\mathbf{k})\, d\mathbf{k}$, where $A(\mathbf{k})$ is positive and bounded. A similar result holds for $d\nu(\mathbf{k}) = C(\mathbf{k})\, d\mathbf{k}$.

The first three conditions of Lemma 5 are now satisfied for all $f, g \in L_r^2(R_3)$. The fourth one is easy to discuss in this case. Putting $A(\mathbf{k}) = \frac{1}{4}n(\mathbf{k})$ and $C(\mathbf{k}) = \frac{1}{4}m(\mathbf{k})$, we can write (43) as

$$\begin{aligned} K(f', g'; f, g) = N'N \exp\left(\tfrac{1}{2}\int_D \hat{d}\mathbf{k}[n(\mathbf{k})(\tilde{f}'(\mathbf{k}), \tilde{f}(\mathbf{k})) \right. \\ + m(\mathbf{k})(\tilde{g}'(\mathbf{k}), \tilde{g}(\mathbf{k})) - i(\tilde{f}'(\mathbf{k}), \tilde{g}(\mathbf{k})) \\ \left. + i(\tilde{g}'(\mathbf{k}), \tilde{f}(\mathbf{k}))]\right) \end{aligned}$$

and the fourth condition as

$$\sum_n \frac{1}{n!} \sum_{i,j=1}^{N} (c_i^* N_i)(c_j N_j)(A_{ij})^n \geq 0$$

with

$$\begin{aligned} A_{ij} = \tfrac{1}{2}\int_D \hat{d}\mathbf{k}[n(\mathbf{k})(\tilde{f}_i(\mathbf{k}), \tilde{f}_j(\mathbf{k})) + m(\mathbf{k})(\tilde{g}_i(\mathbf{k}), \tilde{g}_j(\mathbf{k})) \\ - i(\tilde{f}_i(\mathbf{k}), \tilde{g}_j(\mathbf{k})) + i(\tilde{g}_i(\mathbf{k}), \tilde{f}_j(\mathbf{k}))]. \end{aligned}$$

If two Hermitian matrices A and B are positive semidefinite, then the matrix C, whose matrix elements are $C_{ij} = A_{ij}B_{ij}$ in some fixed orthonormal basis, is also positive semidefinite.[15] The assumptions of Lemma 5

are therefore satisfied if $\Sigma \geq 0$, where

$$\Sigma = \sum_{i,j=1}^{N} (c_i^* N_i)(c_j N_j) A_{ij}$$
$$= \int_D d\hat{\mathbf{k}} \int_c d\mathbf{x} \times \left| \sum_i c_i N_i [2m(\mathbf{k})]^{-\frac{1}{2}} [\tilde{f}_i(\mathbf{x}, \mathbf{k}) + m(\mathbf{k}) \tilde{g}_i(\mathbf{x}, \mathbf{k})] \right|^2 + [n(\mathbf{k})m(\mathbf{k}) - 1] \left| \sum_i c_i N_i [2m(\mathbf{k})]^{-\frac{1}{2}} \tilde{f}_i(\mathbf{x}, \mathbf{k}) \right|^2.$$

This shows that $\Sigma \geq 0$ is satisfied if $m(\mathbf{k})n(\mathbf{k}) \geq 1$. On the other hand, it is not difficult to prove that the fourth condition is violated if $n(\mathbf{k})m(\mathbf{k}) \geq 1$ does not hold. Because there exists C such that $m(\mathbf{k}) < C$ and $n(\mathbf{k}) < C$, it follows that $C^{-1} < m(\mathbf{k}) < C$ and $1 \leq \xi(\mathbf{k}) < C^2$, where $\xi(\mathbf{k}) = m(\mathbf{k})n(\mathbf{k})$.

We have shown that the reproducing kernel of a theory satisfying 1 and 2 has necessarily the form given in Theorem 1. To complete its proof, it remains to verify that the assumptions 2 and the reducibility and equivalence properties are satisfied.

Verification of the Required Properties

We can write

$$(\Phi[f', g'], \Phi[f, g]) = N'N \exp\left(\int_D d\hat{\mathbf{k}} \int_c d\hat{\mathbf{x}} [\tilde{h}_1'^*(\mathbf{x}, \mathbf{k}) \tilde{h}_1(\mathbf{x}, \mathbf{k}) + \tilde{h}_2'^*(\mathbf{x}, \mathbf{k}) \tilde{h}_2(\mathbf{x}, \mathbf{k})] \right), \quad (44)$$

where

$$\tilde{h}_1(\mathbf{x}, \mathbf{k}) = [2m(\mathbf{k})]^{-\frac{1}{2}} [\tilde{f}(\mathbf{x}, \mathbf{k}) - im(\mathbf{k}) \tilde{g}(\mathbf{x}, \mathbf{k})],$$
$$\tilde{h}_2(\mathbf{x}, \mathbf{k}) = [2m(\mathbf{k})]^{-\frac{1}{2}} \eta(\mathbf{k}) \tilde{f}(\mathbf{x}, \mathbf{k}), \quad (45)$$
$$\eta^2(\mathbf{k}) = \xi(\mathbf{k}) - 1.$$

Because of (9), it follows from this that

$$(\Phi[f', g'], U[T]\Phi[f, g]) = (U[T^{-1}]\Phi[f', g'], \Phi[f, g]), \quad (46)$$

which can hold only if $U[T]$ is linear. Equation (9) determines therefore how $U[T]$ acts on a dense set in $\mathfrak{H}$, i.e., $U[T]$ is an operator. Equation (46) tells us that $U^+[T] = U[T^{-1}]$ and (9) that

$$\Phi[f, g] = U[T^{-1}]U[T]\Phi[f, g] = U[T]U[T^{-1}]\Phi[f, g].$$

Hence $U^+[T]U[T] = U[T]U^+[T] = 1$. One shows in very much the same way that (10) defines an antiunitary operator $\mathfrak{I}$.

Let us show that only multiples of Φ_0 are invariant under the $U[T]$ for all $T \in \mathfrak{T}'$. We expand $f \in L_r^2(R_3)$ in terms of the orthonormal functions $u^j(\mathbf{x})$ that we used in (38):

$$f(\mathbf{x}) = \sum_{\mathbf{i}} \sum_j p_j^{\mathbf{i}} u^j(\mathbf{x} - \mathbf{i}).$$

Define $T_k \in \mathfrak{T}'$ by

$$(T_k f)(\mathbf{x}) = \sum_{\mathbf{i}} \left(\sum_{j=1}^{k} p_{k+1-j}^{\mathbf{i}} u^j(\mathbf{x} - \mathbf{i}) + \sum_{j=k+1}^{\infty} p_j^{\mathbf{i}} u^j(\mathbf{x} - \mathbf{i}) \right).$$

Calculating the quantities $(\Phi[f', g'], U[T_k]\Phi[f, g])$, we find that they converge towards $(\Phi[f', g'], \Phi_0) \times (\Phi_0, \Phi[f, g])$ because T_k converges weakly to zero for $k \to \infty$. By Lemma 3, we have

$$(\Lambda, \Psi) = (\Lambda, U[T_k]\Psi) \to (\Lambda, \Phi_0)(\Phi_0, \Psi)$$

for arbitrary $\Lambda \in \mathfrak{H}$ and Ψ invariant under all the $U[T_k]$. Therefore $\Psi = (\Phi_0, \Psi)\Phi_0$.

Next we shall prove that the representation of the CCR is irreducible if $\xi(\mathbf{k}) \equiv 1$. Define

$$\tilde{F}(\mathbf{x}, \mathbf{k}) = m^{-\frac{1}{2}}(\mathbf{k}) \tilde{f}(\mathbf{x}, \mathbf{k})$$

and

$$\tilde{G}(\mathbf{x}, \mathbf{k}) = m^{\frac{1}{2}}(\mathbf{k}) \tilde{g}(\mathbf{x}, \mathbf{k}).$$

The sets of all $\tilde{f}(\mathbf{x}, \mathbf{k})$ and of all $\tilde{F}(\mathbf{x}, \mathbf{k})$ coincide. If the $U[f, g]$ satisfy 1, the $U_0[F, G]$, defined by $U_0[F, G] = U[f, g]$, will fulfill 1, too. Equation (15) yields

$$(\Phi_0, U_0[F, G]\Phi_0) = \exp\left(-\tfrac{1}{4} \int_D d\hat{\mathbf{k}} \int_c d\mathbf{x} [\xi(\mathbf{k}) |\tilde{F}(\mathbf{x}, \mathbf{k})|^2 + |\tilde{G}(\mathbf{x}, \mathbf{k})|^2] \right).$$

For $\xi(\mathbf{k}) \equiv 1$, this becomes $e^{-\frac{1}{4}[(F,F)+(G,G)]}$, the reproducing kernel of the Fock representation.[16] The groups of all $U_0[F, G]$ and of all the $U_F[f, g]$ for the Fock representation must therefore be isomorphic. However, the sets of all $U[f, g]$ and of all $U_0[F, G]$ are the same. Hence, every bounded operator that commutes with all the $U[f, g]$ is a multiple of the identity. This is what irreducibility means. If $\xi(\mathbf{k}) \not\equiv 1$, the representation will be reducible.

The proof that two representations with the same reproducing kernel are equivalent has been given by Naimark.[17] To show that two representations with different reproducing kernels are inequivalent, we use the tag operators. By (36) and (40) we have $AU[f, g]\Phi_0 = U[f, g]A\Phi_0 = aU[f, g]\Phi_0$ for all $f, g \in L$. Therefore $A = aI$ on $\mathfrak{H}_L$. Assume there exists a unitary operator V such that

$$U'[f, g] = VU[f, g]V^{-1} \quad \text{for all} \quad f, g \in L_r^2(R_3).$$

The existence of a weak limit A for $U[f^k, g^k]$ implies that the $U'[f^k, g^k]$ converge weakly towards an operator A' in the image of $\mathfrak{H}_L$ under V and that

$$A' = VAV^{-1} = a.$$

Looking at

$$a = \exp\left(-\tfrac{1}{4}\int_D \hat{d}\mathbf{k}[\xi(\mathbf{k})m^{-1}(\mathbf{k})(\tilde{f}(\mathbf{k}), \tilde{f}(\mathbf{k})) + m(\mathbf{k})(\tilde{g}(\mathbf{k}), \tilde{g}(\mathbf{k}))]\right),$$

we recognize that it will always be possible to find f, $g \in L$ such that a and a' are different unless $m(\mathbf{k}) \equiv m'(\mathbf{k})$ and $\xi(\mathbf{k}) \equiv \xi'(\mathbf{k})$.

4. THE HAMILTONIAN

Exponential Hilbert Space

Consider (44) and (45). It is convenient to put

$$\tilde{h}_1(\mathbf{x}, \mathbf{k}) = \tilde{h}(\mathbf{x}, \mathbf{k}), \quad \tilde{h}_2(\mathbf{x}, \mathbf{k}) = \tilde{h}(\mathbf{x}, \mathbf{k} + (2\pi, 0, 0))$$

and to define a domain D' by $\mathbf{k} \in D'$ if $\mathbf{k} \in D$ or if $(\mathbf{k} - (2\pi, 0, 0)) \in D$ and $\xi(\mathbf{k} - (2\pi, 0, 0)) > 1$. Hence

$$(\Phi[f', g'], \Phi[f, g]) = N'N \exp\left(\int_{D'} \hat{d}\mathbf{k}(h'(\mathbf{k}), h(\mathbf{k}))\right). \tag{47}$$

This suggests a realization of $\mathfrak{H}$. We introduce the abbreviation $\mathfrak{h}$ for the Hilbert space $L^2(c, D')$ with elements $\tilde{\phi}(\mathbf{x}, \mathbf{k})$. Call $\mathfrak{H}_n$ the symmetrized direct product of n factors $\mathfrak{h}$, and let

$$\bar{\mathfrak{H}} = \bigoplus_{n=0}^{\infty} \mathfrak{H}_n.$$

$\bar{\mathfrak{H}}$ is called the exponential space of $\mathfrak{h}$.[18] For $\phi, \psi \in \mathfrak{h}$, we have

$$e^{\phi} \equiv \bigoplus_{n=0}^{\infty} (n!)^{-\frac{1}{2}}(\otimes\phi)^n \in \bar{\mathfrak{H}}$$

and $(e^{\phi}, e^{\psi}) = e^{(\phi,\psi)}$. $\mathfrak{H}$ is realizable as a subspace of $\bar{\mathfrak{H}}$, $\mathfrak{H} \subset \bar{\mathfrak{H}}$, if we put

$$\Phi[f, g] = Ne^{h}.$$

We claim $\mathfrak{H} = \bar{\mathfrak{H}}$. It is not difficult to see that the vectors of the form $(\otimes h)^n$ span $\mathfrak{H}_n$. It remains to prove that $(\otimes h)^n \in \mathfrak{H}$ for all n, which we do by induction. Φ_0 spans $\mathfrak{H}_0$. We have

$$\lim_{s\to 0} s^{-n}(n!)^{\frac{1}{2}}\left(N^{-1}\Phi[sf, sg] - \bigoplus_{m=0}^{n-1}(m!)^{-\frac{1}{2}}(\otimes sh)^m\right) = (\otimes h)^n.$$

If the sequence ψ_n converges weakly to $\psi \in \mathfrak{h}$, we have for all $\phi \in \mathfrak{h}$

$$(e^{\phi}, e^{\psi_n}) = e^{(\phi,\psi_n)} \to e^{(\phi,\psi)}.$$

Since the e^{ϕ} span $\mathfrak{H}$, this implies that the e^{ψ_n} converge weakly to $e^{\psi} \in \mathfrak{H}$.

Operators that Commute with the $U[T]$ for $T \in \mathfrak{T}'$ or $T \in \mathfrak{T}$

Using the realization of $\mathfrak{H}$ that we have just introduced, we shall now consider classes of operators of which the Hamiltonians are special cases. From

$$(\Phi[f', g'], U[T]\Phi[f, g]) = N'N \sum (n!)^{-1}(h', Th)^n,$$

it follows that each $\mathfrak{H}_n$ is invariant under $U[T]$ for all $T \in \mathfrak{T}$,

$$U[T] = \bigoplus_{n=0}^{\infty} U_n[T].$$

The representations $U_n[\mathfrak{T}']$ are disjoint. By this we mean that, for $p \neq q$, no representation of $\mathfrak{T}'$ in a subspace of $\mathfrak{H}_p$ can be equivalent to a representation in a subspace of $\mathfrak{H}_q$; i.e., $PU_p[T]P = VQU_q[T]QV^{-1}$ for all $T \in \mathfrak{T}'$, where P and Q are projections and V is unitary, implies $P = Q = 0$.

Proof[19]: Consider a sequence T_m, $T_m \in \mathfrak{T}'$, that converges weakly to cI. To see that such sequences exist we choose an orthonormal basis in $L^2(c)$. The matrix shown in Fig. 1 determines how T_m acts in $L^2(c)$ and, because elements of $\mathfrak{T}'$ act similarly in each cube, how it acts in $L^2(R_3)$. We have

$$(e^{h'}, U[T_m]e^{h}) = \sum (n!)^{-1}(h', T_m h)^n \to (e^{h'}, e^{ch}).$$

Lemma 3 states that $U[T_m]$ converges weakly to $\bigoplus_{n=0}^{\infty} c^n I_n$, where I_n is the identity operator in $\mathfrak{H}_n$. $Pc^p I_p P = VQc^q I_q QV^{-1}$ holds for $p \neq q$ and $|c| \neq 1$ only if $P = Q = 0$.

The set of all bounded operators that commute with $U[T]$ for all $T \in \mathfrak{T}$ is called the commutator of the $U[T]$ and is written as $\{U[\mathfrak{T}]\}'$.

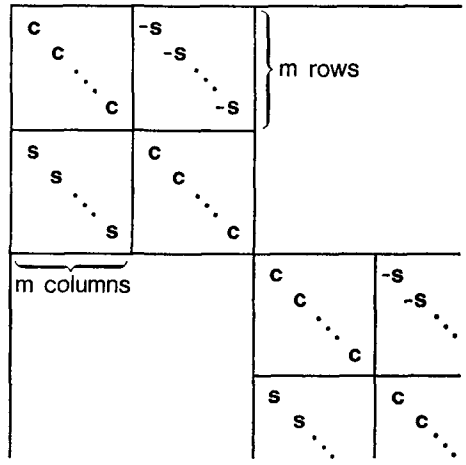


FIG. 1. The matrix determining T_m. c stands for $\cos\theta$ and s for $\sin\theta$.

Disjointness implies that $\mathcal{B} \in \{U[\mathfrak{T}']\}'$ has the structure

$$\mathcal{B} = \bigoplus_{n=0}^{\infty} \mathcal{B}_n .$$

$\mathfrak{H}_1$ can be considered as a direct integral:

$$\mathfrak{H}_1 = \int_{D'}^{\oplus} d\mathbf{k}\mathfrak{h}(\mathbf{k}).$$

Because each Hilbert space $\mathfrak{h}(\mathbf{k})$ carries an equivalent irreducible representation of $\mathfrak{T}'$, $\mathcal{B}_1\tilde{h}(\mathbf{x}, \mathbf{k})$ must have the form

$$[\mathcal{B}_1\tilde{h}(\mathbf{x}, \mathbf{k})](\mathbf{x}', \mathbf{k}') = \delta(\mathbf{x}' - \mathbf{x})b(\mathbf{k}', \mathbf{k})\tilde{h}(\mathbf{x}, \mathbf{k}),$$

which we abbreviate to

$$[\mathcal{B}_1\tilde{h}(\mathbf{x}, \mathbf{k})](\mathbf{k}') = b(\mathbf{k}', \mathbf{k})\tilde{h}(\mathbf{x}, \mathbf{k}).$$

$b(\mathbf{k}', \mathbf{k})$ must be understood in the sense of distributions, e.g., to $\mathcal{B} = I$ corresponds $b(\mathbf{k}', \mathbf{k}) = \delta(\mathbf{k}' - \mathbf{k})$. $\mathcal{B}_n$ acts as follows[20]:

$$\left(\mathcal{B}_n \prod_{i=1}^{n} \tilde{h}(\mathbf{x}_i, \mathbf{k}_i)\right)(\mathbf{k}_1', \cdots, \mathbf{k}_n') = b(\mathbf{k}_1', \mathbf{k}_1, \cdots, \mathbf{k}_n', \mathbf{k}_n) \prod_{i=1}^{n} \tilde{h}(\mathbf{x}_i, \mathbf{k}_i).$$

Therefore

$$(\Phi[f', g'], \mathcal{B}\Phi[f, g]) = N'N \sum_n \frac{1}{n!}\left(\prod_{i=1} \int_{D'}\int_{D'} \hat{d}\mathbf{k}_i' \, \hat{d}\mathbf{k}_i (\tilde{h}'(\mathbf{k}_i'), \tilde{h}(\mathbf{k}_i))\right) \times b(\mathbf{k}_1', \mathbf{k}_1, \cdots, \mathbf{k}_n', \mathbf{k}_n). \quad (48)$$

We may require

$$b(\mathbf{k}_1', \mathbf{k}_1, \cdots, \mathbf{k}_n', \mathbf{k}_n) = b(\mathbf{k}_{i_1}', \mathbf{k}_{i_1}, \cdots, \mathbf{k}_{i_n}', \mathbf{k}_{i_n})$$

for all permutations $i_1, \cdots, i_n$ of $1, \cdots, n$ without restricting the generality. A straightforward calculation shows that for $\mathcal{B} \in \{U[\mathfrak{T}]\}'$, the b must satisfy (20); (19) is fulfilled for symmetric $\mathcal{B}$.

Construction of Hamiltonians

In this subsection we shall prove Theorem 2, beginning with the part that gives necessary conditions. The sufficient conditions will be verified by constructing the operators in question first for irreducible and then for reducible CCR representations.

We could replace h_1 and h_2 in the preceding discussions of this section by

$$\tilde{H}_1(\mathbf{x}, \mathbf{k}) = \tilde{h}_1(\mathbf{x}, \mathbf{k}) \cos\theta(\mathbf{k}) - \tilde{h}_2(\mathbf{x}, \mathbf{k}) \sin\theta(\mathbf{k}),$$
$$\tilde{H}_2(\mathbf{x}, \mathbf{k}) = \tilde{h}_1(\mathbf{x}, \mathbf{k}) \sin\theta(\mathbf{k}) + \tilde{h}_2(\mathbf{x}, \mathbf{k}) \cos\theta(\mathbf{k}) \quad (49)$$

for any choice of $\theta(\mathbf{k})$ with $\sin\theta(\mathbf{k}) = 0$ or $\cos\theta(\mathbf{k}) = 0$ if $\xi(\mathbf{k}) = 1$, because

$$\sum_{i=1,2} (\tilde{h}_i'(\mathbf{k}), \tilde{h}_i(\mathbf{k})) = \sum_{i=1,2} (\tilde{H}_i'(\mathbf{k}), \tilde{H}_i(\mathbf{k})).$$

The following choice of $\theta(\mathbf{k})$ will prove useful for the construction of Hamiltonians:

$$\sin\theta(\mathbf{k}) = \xi^{-\frac{1}{2}}(\mathbf{k}),$$
$$\cos\theta(\mathbf{k}) = [\xi(\mathbf{k}) - 1]^{\frac{1}{2}}\xi^{-\frac{1}{2}}(\mathbf{k}) \equiv \zeta(\mathbf{k}).$$

It leads to

$$\tilde{d}(\mathbf{x}, \mathbf{k}) \equiv \tilde{d}_1(\mathbf{x}, \mathbf{k}) \equiv -i[\tfrac{1}{2}m(\mathbf{k})]^{\frac{1}{2}}\zeta(\mathbf{k})\tilde{g}(\mathbf{x}, \mathbf{k})$$

and

$$\tilde{d}(\mathbf{x}, \mathbf{k} + (2\pi, 0, 0)) \equiv \tilde{d}_2(\mathbf{x}, \mathbf{k}) \equiv [2m(\mathbf{k})\xi(\mathbf{k})]^{-\frac{1}{2}} \times [\xi(\mathbf{k})\tilde{f}(\mathbf{x}, \mathbf{k}) - im(\mathbf{k})\tilde{g}(\mathbf{x}, \mathbf{k})]. \quad (50)$$

Since $e^{-it\mathcal{H}} \in \{U[\mathfrak{T}]\}'$, we may put

$$(\Phi[f', g'], e^{-it\mathcal{H}}\Phi[f, g]) = N'N\, A\{(\tilde{d}'(\mathbf{k}'), \tilde{d}(\mathbf{k}))\} = (\Phi[f', g'], \Phi[f, g])C\{(\tilde{d}'(\mathbf{k}'), \tilde{d}(\mathbf{k}))\}.$$

If $\Phi[f, g]$ belongs to the domain where $\mathcal{H}$ is defined, we shall therefore be able to write

$$(\Phi[f', g'], \mathcal{H}\Phi[f, g]) = (\Phi[f', g'], \Phi[f, g])G\{(\tilde{d}'(\mathbf{k}'), \tilde{d}(\mathbf{k}))\}.$$

We use an adaptation of the method used in Ref. 2 to derive restrictions on G that follow from (12). Computing $\partial(\Phi[f', g'], \Phi[f, g + \tau e])/\partial\tau$ at $\tau = 0$, where $e \in L_r^2(R_3)$ and $\tau \in R$, we find

$$M = (\Phi[f', g'], \pi(e)\Phi[f, g]) = \tfrac{1}{2}(\Phi[f', g'], \Phi[f, g]) \int_D \hat{d}\mathbf{k}[\tfrac{1}{2}m(\mathbf{k})]^{\frac{1}{2}} \times [\zeta(\mathbf{k})(\tilde{e}(\mathbf{k}), \tilde{d}_1(\mathbf{k}) + \tilde{d}_1'(\mathbf{k})) + \xi^{-\frac{1}{2}}(\mathbf{k})(\tilde{e}(\mathbf{k}), \tilde{d}_2(\mathbf{k}) + \tilde{d}_2'(\mathbf{k}))].$$

Because of (12) we have

$$M = \frac{\partial}{\partial\tau}(\Phi[f', g'], e^{-i\tau\phi(e)}\mathcal{H}e^{i\tau\phi(e)}\Phi[f, g])\Big|_{\tau=0}.$$

Calculating this and comparing the two expressions for M, one finds

$$(\Phi[f', g'], \mathfrak{H}\Phi[f, g]) = \tfrac{1}{2}(\Phi[f', g'], \Phi[f, g]) \times \left(\int_D (\tilde{f}'(\mathbf{k}) - im(\mathbf{k})\tilde{g}'(\mathbf{k}), \tilde{f}(\mathbf{k}) - im(\mathbf{k})\tilde{g}(\mathbf{k}))\, \hat{d}\mathbf{k} + E\{(\tilde{d}_1'(\mathbf{k}'), \tilde{d}_1(\mathbf{k}))\}\right). \quad (51)$$

The domain conditions that validate the above calculations and guarantee that (12) holds will be satisfied for all $\mathcal{H}$ that we shall construct. $\mathcal{H}\Phi_0 = 0$ implies

$$E\{(\tilde{d}_1'(\mathbf{k}'), \tilde{d}_1(\mathbf{k})) \equiv 0\} = 0.$$

Using $[U[T], \mathcal{H}] = 0$ for all $T \in \mathfrak{T}$, we find

$$\tfrac{1}{2}E\{(\tilde{d}_1'(\mathbf{k}'), \tilde{d}_1(\mathbf{k}))\} = \sum_{n=1}^{\infty} \frac{1}{n!}\left(\prod_{i=1}^{n} \int_D \int_D \hat{d}\mathbf{k}_i' \, \hat{d}\mathbf{k}_i (\tilde{d}_1'(\mathbf{k}_i'), \tilde{d}_1(\mathbf{k}_i))\right) \times b(\mathbf{k}_1', \mathbf{k}_1, \cdots, \mathbf{k}_n', \mathbf{k}_n).$$

We complete the proof of the necessary conditions given in Theorem 2 by remarking that $e(\mathbf{k}) > 0$ for all $\mathbf{k}$ with $\xi(\mathbf{k}) > 1$ is necessary for a nonnegative Hamiltonian with a unique ground state.

Next we shall construct the Hamiltonian for the irreducible representations of the CCR. We define[21]

$$\mathcal{H}'(\otimes h_1)^p(\otimes h_2)^{n-p} = \sum_{i=1}^{p} m(\mathbf{k}_i) \overset{p}{\underset{i=1}{\otimes}} \tilde{h}_1(\mathbf{x}_i, \mathbf{k}_i) \overset{n}{\underset{i=p+1}{\otimes}} \tilde{h}_2(\mathbf{x}_i, \mathbf{k}_i). \quad (52)$$

By linearity and continuity we can extend the domain of definition to $\mathfrak{H}_n$. $\mathcal{H}'$ is a positive bounded symmetric operator on each $\mathfrak{H}_n$. We may consider it as a positive symmetric operator on $\mathfrak{H}$ and extend it by the method of Friedrichs[22] to a self-adjoint operator.

$$\mathcal{H}'\Phi[f, g] = N \overset{\infty}{\underset{n=0}{\oplus}} (n!)^{-\frac{1}{2}} \overset{n}{\underset{p=0}{\oplus}} \binom{n}{p}^{\frac{1}{2}} \times \sum_{i=1}^{p} m(\mathbf{k}_i) \overset{p}{\underset{i=1}{\otimes}} \tilde{h}_1(\mathbf{x}_i, \mathbf{k}_i) \overset{n}{\underset{i=p+1}{\otimes}} \tilde{h}_2(\mathbf{x}_i, \mathbf{k}_i)$$

lies in $\mathfrak{H}$,

$$(\Phi[f', g'], \mathcal{H}'\Phi[f, g]) = (\Phi[f', g'], \Phi[f, g]) \int_D \hat{d}\mathbf{k} m(\mathbf{k}) (\tilde{h}_1'(\mathbf{k}), \tilde{h}_1(\mathbf{k})).$$

Similarly, one determines the σ matrix elements of $e^{-it\mathcal{H}'}$:

$$(\Phi[f', g'], e^{-it\mathcal{H}'}\Phi[f, g]) = (\Phi[f', g'], \Phi[f, g]) \times \exp\left(\int_D \hat{d}\mathbf{k}(e^{-itm(\mathbf{k})} - 1)(\tilde{h}_1'(\mathbf{k}), \tilde{h}_1(\mathbf{k}))\right).$$

In the case of irreducible representations, $\mathcal{H}'$ has the σ matrix elements required for the Hamiltonian $\mathcal{H}$. Equation (11) is satisfied because

$$U[T](\otimes h_1)^p(\otimes h_2)^{n-p} = (\otimes Th_1)^p(\otimes Th_2)^{n-p},$$

and (12) because $\mathcal{H}'$ has the σ matrix elements of the form (51). That $\mathcal{H}'\Phi = 0$ has only the solution $\Phi = c\Phi_0$ follows for irreducible representations from (52).

There remains to construct the Hamiltonians for reducible representations that fulfill the sufficient conditions. Let

$$\mathcal{H}(e)(\otimes d_1)^p(\otimes d_2)^{n-p} = \sum_{i=1}^{p} e(\mathbf{k}_i) \overset{p}{\underset{i=1}{\otimes}} \tilde{d}_1(\mathbf{x}_i, \mathbf{k}_i) \overset{n}{\underset{i=p+1}{\otimes}} \tilde{d}_2(\mathbf{x}_i, \mathbf{k}_i),$$

with $e(\mathbf{k}) < C$ and $e(\mathbf{k}) > 0$ for all $\mathbf{k}$ with $\xi(\mathbf{k}) > 1$. This determines, similarly as (52), a positive self-adjoint operator on $\mathfrak{H}$ with

$$(\Phi[f', g'], \mathcal{H}(e)\Phi[f, g]) = (\Phi[f', g'], \Phi[f, g]) \int_D \hat{d}\mathbf{k} e(\mathbf{k}) (\tilde{d}_1'(\mathbf{k}), \tilde{d}_1(\mathbf{k})).$$

$\mathcal{H}_0 = \mathcal{H}' + \mathcal{H}(e)$ satisfies assumptions 3. A straightforward calculation yields that the operator $\mathcal{O}$ defined by

$$\mathcal{O}(\otimes h_+)^p(\otimes h_-)^{n-p} = \left(\sum_{i=1}^{p} m_+(\mathbf{k}_i) + \sum_{i=p+1}^{n} m_-(\mathbf{k}_i)\right) \times \overset{p}{\underset{i=1}{\otimes}} \tilde{h}_+(\mathbf{x}_i, \mathbf{k}_i) \overset{n}{\underset{i=p+1}{\otimes}} \tilde{h}_-(\mathbf{x}_i, \mathbf{k}_i)$$

coincides with $\mathcal{H}_0$. This result can be used to derive (24).

To construct operators $\mathcal{U}$ with σ matrix elements

$$(\Phi[f', g'], \mathcal{U}\Phi[f, g]) = \tfrac{1}{2}(\Phi[f', g'], \Phi[f, g]) F\{(\tilde{G}'(\mathbf{k}'), \tilde{G}(\mathbf{k}))\}, \quad (53)$$

where F satisfies (21), it is useful to introduce smeared "creation" and "annihilation" operators. Let

$$A(\alpha)(\otimes d_1)^p(\otimes d_2)^{n-p} = ip^{\frac{1}{2}}(\alpha, d_1) \overset{p-1}{\underset{i=1}{\otimes}} \tilde{d}_1(\mathbf{x}_i, \mathbf{k}_i) \overset{n-1}{\underset{i=p}{\otimes}} \tilde{d}_2(\mathbf{x}_i, \mathbf{k}_i),$$

where $\alpha \in L^2(R_3)$. This determines a linear operator with bound $N^{\frac{1}{2}} \|\alpha\|$ on $\mathfrak{H}^{(N)} = \oplus_{n=0}^{N} \mathfrak{H}_n$ and therefore a linear operator on $\mathcal{H}$. $A(\alpha)$ is closable: If we take a sequence Ψ_n such that each Ψ_n is in some $\mathfrak{H}^{(N)}$ and that $\Psi_n \to 0$ and $A(\alpha)\Psi_n \to \Phi$, then $\Phi = 0$.

Indirect proof: Assume $\Phi \neq 0$; let l be the smallest integer such that the projection P_l of Φ on $\mathfrak{H}_l$ is different from zero, $\|P_l\Phi\| = c > 0$. Then there exists N such that $\|P_l A(\alpha)\Psi_n\| > \frac{1}{2}c$ for all $n > N$. Hence $\|P_{l+1}\Psi_n\| > \frac{1}{2}c(l+1)^{-\frac{1}{2}} \|\alpha\|^{-1}$ for all n, which contradicts $\Psi_n \to 0$. Because $\Psi \in \mathfrak{H}^{(n)}$ implies $A(\alpha)\Psi \in \mathfrak{H}^{(n)}$, it follows that $\mathcal{O} = A(\alpha_1) \cdot \cdots \cdot A(\alpha_n)$ is defined for each vector in any $\mathfrak{H}^{(n)}$, $\mathfrak{H}^{(n)} \subset D(\mathcal{O})$. An argument similar to the one for $A(\alpha)$ shows that $\mathcal{O}$ is closable, too; hence $\mathcal{O}^+$ exists. One proves then

$\mathfrak{H}^{(n)} \subset D(\mathcal{O}^+)$. Hence $\mathcal{O}' = A^+(\alpha_1)\cdot\cdots\cdot A^+(\alpha_m)\cdot A(\alpha_{m+1})\cdot\cdots\cdot A(\alpha_n)$ is an operator with $\mathfrak{H}^{(n)} \subset D(\mathcal{O}')$.[23] Because

$$A(\alpha)\Phi[f,g] = (\alpha, G)\Phi[f,g],$$

we have, e.g.,

$$(\Phi[f',g'], A^+(\alpha_1)A^+(\alpha_2)A(\alpha_2)A(\alpha_1)\Phi[f,g]) = (G',\alpha_1)(G',\alpha_2)(\alpha_2,G)(\alpha_1,G)(\Phi[f',g'],\Phi[f,g]).$$

Notice that the closure of the operator in this equation is positive and self-adjoint because it is of the form $\mathcal{O}^+\mathcal{O}$ and $\mathcal{O} = A(\alpha_2)A(\alpha_1)$ is closable.[24] We introduce an orthonormal basis $\alpha_{\mathbf{n}}(\mathbf{x})$, $\mathbf{n}\in Z_3$ in $L^2(c)$. Parseval's theorem states for g', $g\in L^2(c)$ that

$$\sum_{\mathbf{n}}\int_c d\mathbf{x}\, g'^*(\mathbf{x}')\alpha_{\mathbf{n}}(\mathbf{x}')\int_c d\mathbf{x} g(\mathbf{x})\alpha_{\mathbf{n}}^*(\mathbf{x}) = \int_c d\mathbf{x} g'^*(\mathbf{x})g(\mathbf{x}).$$

For the $\epsilon_{\mathbf{m}}(\mathbf{k}) = e^{i\mathbf{m}\mathbf{k}}$, $\mathbf{m}\in Z_3$, one has also such a theorem. Let $\beta_{\mathbf{m}}(\mathbf{k}) = \beta(\mathbf{k})\epsilon_{\mathbf{m}}(\mathbf{k})$, where $|\beta(\mathbf{k})| < C$. We have

$$\begin{aligned}
&\tfrac{1}{2}\sum_{\mathbf{M}} b_{\mathbf{M}} \sum_{\mathbf{m},\mathbf{n}_1,\mathbf{n}_2} (\Phi[f',g'], A^+(\alpha_{\mathbf{n}_1}\beta_{\mathbf{m}+\mathbf{M}})A^+(\alpha_{\mathbf{n}_2}\beta_{\mathbf{m}}) \\
&\quad\times A(\alpha_{\mathbf{n}_2}\beta_{\mathbf{m}})A(\alpha_{\mathbf{n}_1}\beta_{\mathbf{m}+\mathbf{M}})\Phi[f,g]) \\
&= \tfrac{1}{2}(\Phi[f',g'],\Phi[f,g]) \\
&\quad\times \sum_{\mathbf{M}} b_{\mathbf{M}} \sum_{\mathbf{m}}\int_D\int_D \hat{d}\mathbf{k}_1'\,\hat{d}\mathbf{k}_1\beta(\mathbf{k}_1')\beta^*(\mathbf{k}_1)(\tilde{G}'(\mathbf{k}_1'),\tilde{G}(\mathbf{k}_1)) \\
&\quad\times \int_D\int_D \hat{d}\mathbf{k}_2'\,\hat{d}\mathbf{k}_2\beta(\mathbf{k}_2')\beta^*(\mathbf{k}_2)(\tilde{G}'(\mathbf{k}_2'),\tilde{G}(\mathbf{k}_2)) \\
&\quad\times \exp\{i[\mathbf{m}(\mathbf{k}_1'+\mathbf{k}_2'-\mathbf{k}_1-\mathbf{k}_2)+\mathbf{M}(\mathbf{k}_1'-\mathbf{k}_1)]\}.
\end{aligned}$$

This coincides with the right side of (53). If $b_{\mathbf{M}} = C > 0$ for all $\mathbf{M}$, it becomes

$$\begin{aligned}
&\tfrac{1}{2}C(\Phi[f',g'],\Phi[f,g])\int_D \hat{d}\mathbf{k}_1\,|\beta(\mathbf{k}_1)|^2\,(\tilde{G}'(\mathbf{k}_1),\tilde{G}(\mathbf{k}_1)) \\
&\qquad\times\int_D \hat{d}\mathbf{k}_2\,|\beta(\mathbf{k}_2)|^2\,(\tilde{G}'(\mathbf{k}_2),\tilde{G}(\mathbf{k}_2)) \\
&\qquad = \tfrac{1}{2}C(\Phi[f',g'],[\mathcal{K}(|\beta|^2)]^2\Phi[f,g]).
\end{aligned}$$

$\frac{1}{2}C[\mathcal{K}(|\beta|^2)]^2$ is positive and self-adjoint because $\mathcal{K}(|\beta|^2)$ is self-adjoint, and it has σ in its domain. To prove the existence of $\mathcal{U}$ for the sequence $b_{\mathbf{M}}$ mentioned in Theorem 2, we shall make use of the following lemma, which is a weakened form of a theorem given by Kato.[25]

Lemma 6: Let $\mathcal{O}_n$ be a nondecreasing sequence of positive self-adjoint operators that is majorized by a self-adjoint operator $\mathcal{O}_0$. Then there exists a positive self-adjoint operator $\mathcal{O}$ with the property

$$(\Phi,\mathcal{O}\Psi) = \lim_{n\to\infty}(\Phi,\mathcal{O}_n\Psi)$$

for all Φ and Ψ in the domain of $\mathcal{O}_0$.

The assumptions of this theorem are satisfied by

$$\mathcal{O}_0 = \tfrac{1}{2}C[\mathcal{K}(|\beta|^2)]^2$$

and

$$\mathcal{O}_n = \tfrac{1}{2}\sum_{(n)} b_{\mathbf{M}}A^+(\alpha_{\mathbf{n}_1}\beta_{\mathbf{m}+\mathbf{M}}) \times A^+(\alpha_{\mathbf{n}_2}\beta_{\mathbf{m}})A(\alpha_{\mathbf{n}_2}\beta_{\mathbf{m}})A(\alpha_{\mathbf{n}_1}\beta_{\mathbf{m}+\mathbf{M}}),$$

with $0 \le b_{\mathbf{M}} = b_{-\mathbf{M}} < C$, where (n) stands for $\mathbf{M}, \mathbf{m}, \mathbf{n}_1, \mathbf{n}_2$ with $|\mathbf{M}|, |\mathbf{m}|, |\mathbf{n}_1|, |\mathbf{n}_2| \le n$. Because the domains of the three positive, self-adjoint operators $\mathcal{K}'$, $\mathcal{K}(e)$, and $\mathcal{U}$ have a dense set in common, their sum will be symmetric and can be extended to a self-adjoint operator $\mathcal{K}$. It is easy to verify that $\mathcal{K}$ satisfies our assumptions.

ACKNOWLEDGMENTS

It is a pleasure to thank Professor J. R. Klauder for help and encouragement at all stages of this work. It also benefited from discussions with Professor C. McCarthy, Dr. P. W. Higgs, Dr. M. Reeken, and Professor R. F. Streater. I am grateful to Janggen-Pöhn Stiftung and to Edinburgh University for financial support when I began this work at the Universities of Syracuse and Edinburgh.

[1] J. R. Klauder, Lectures, Achte Internationale Universitätswochen für Kernphysik, Schladming, 1969.
[2] J. R. Klauder, J. Math. Phys. **6**, 1666 (1965).
[3] H. D. I. Abarbanel, J. R. Klauder, and J. G. Taylor, Phys. Rev. **152**, 1198 (1966).
[4] E. J. Verboven, Physica **32**, 2081 (1966).
[5] H. Araki and E. J. Woods, J. Math. Phys. **4**, 637 (1963).
[6] The relations that follow must be understood in the sense of almost everywhere.
[7] See Ref. 2, p. 1669.
[8] R. Arens and D. Babbitt, J. Math. Phys. **6**, 1071 (1965).
[9] J. R. Klauder, J. Math. Phys. **8**, 2392 (1967).
[10] See Ref. 2, p. 1672.
[11] The argument is an adaptation of Klauder's procedure in Ref. 2 and has been given in detail in H. K. Grimmer, thesis, Edinburgh University, 1968.
[12] A proof of this theorem for Z instead of Z_3 can be found in R. E. Edwards, *Fourier Series* (Holt, Rinehart and Winston, New York, 1967), Vol. II, p. 116.
[13] H. Araki, J. Math. Phys. **1**, 492 (1960), Theorem 4.3.
[14] See, e.g., p. 139 of F. Riesz and B. Sz.-Nagy, *Functional Analysis* (Ungar, New York, 1955).
[15] This theorem, due to Schur, is proved, e.g., in Ref. 13, p. 500.
[16] See, e.g., Ref. 5, p. 642.
[17] See M. A. Naimark, *Normed Rings* (Noordhoff, Groningen, 1959), p. 242.
[18] Such constructions have recently been reviewed by J. R. Klauder, J. Math. Phys. **11**, 609 (1970).
[19] A similar argument for rotationally symmetric models has been given by J. R. Klauder, Lectures, *1967 Brandeis Summer Institute in Theoretical Physics* (Gordon and Breach, New York) (in press).
[20] A detailed proof has been given in Ref. 11.
[21] The left side of Eq. (52) is a simplified notation for
$$[\mathfrak{H}'(\otimes h_1)^p(\otimes h_2)^{n-p}](\mathbf{x}_1,\mathbf{k}_1,\cdots,\mathbf{x}_n,\mathbf{k}_n).$$
[22] See, e.g., Ref. 14, p. 329.
[23] We use the same symbol for a closable operator and for its closure.
[24] See, e.g., Ref. 14, p. 312.
[25] See T. Kato, *Perturbation Theory for Linear Operators* (Springer, Berlin, 1966), p. 459.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Hamiltonian Operators via Feynman Path Integrals

LEON COHEN
Hunter College of the City University, New York, New York 10021

(Received 21 April 1970)

It is shown that Kerner and Sutcliffe's derivation leading to a unique quantum mechanical Hamiltonian from its classical counterpart is indeed not unique. This is done by using the same method as Kerner and Sutcliffe to derive different quantum Hamiltonians from the same classical Hamiltonian. It is also shown that other ordering rules besides that of Born and Jordan can be derived with the same technique.

INTRODUCTION

In a recent paper[1] the Feynman path integral formulation of quantum mechanics was used to derive a rule for obtaining the quantum mechanical Hamiltonian operator from its classical counterpart. It is further shown that the rule of Born and Jordan is a consequence. It is claimed that, since the derivation is unambiguous, the ordering rule derived should hence have a fundamental position in quantum theory. There have been many rules proposed since the beginnings of quantum theory, e.g., the Weyl rule, the rule of Born and Jordan, and the symmetrization rule. The author[2,3] has obtained a method of generating all possible rules and has shown that no rules can be consistently used to derive quantum operators from their classical functions. We demonstrate below that the derivation of Kerner and Sutcliffe does not lead to a unique rule and indeed that their method can be used to derive other rules.

In Feynman's path integral formulation the propagator $K(X'', X', t'' - t')$ is given by

$$\sum_{\zeta} e^{i/\hbar A(\zeta)}, \tag{1}$$

where $\sum_{\zeta}$ signifies summation over all paths between X'' and X' and $A(\zeta)$ is the action calculated along a given path. Kerner and Sutcliffe consider the case where the time difference $t'' - t'$ is infinitesimal. They calculate the action

$$A = \int p\, dX - H\, dt \tag{2}$$

by taking the set of paths between X' and X'' such that X' and X'' are connected linearly with time

$$X = X' + \frac{X'' - X'}{t'' - t'}(t - t'). \tag{3}$$

This relationship is used in the calculation of the action. It is then shown that from

$$\psi(X'', t'') = \int K(X'', X', t'' - t')\psi(X', t')\, dX' \tag{4}$$

follows

$$i\hbar \frac{\partial \psi(X'', t)}{\partial t} = \int k(X'', X')\psi(X')\, dX', \tag{5}$$

where

$$k(X'', X') = \frac{1}{2\pi\hbar} \int dp \bar{H} e^{i/\hbar p(X'' - X')} \tag{6}$$

and $\bar{H}$ is the averaged Hamiltonian

$$\bar{H} = \frac{1}{t'' - t'} \int_{t'}^{t''} H(X, p)\, dt. \tag{7}$$

Therefore,

$$\hat{H}\psi = \int k(X'', X')\psi(X')\, dX', \tag{8}$$

where $\hat{H}$ is the quantum mechanical operator corresponding to the classical Hamiltonian. Since ψ is arbitrary, it is then straightforward to derive the ordering rule between H and $\hat{H}$ which for the above yields the rule of Born and Jordan. By taking two examples we show that there are other k's which satisfy Eq. (5) and yield different quantum mechanical operators. This relies on the freedom of choice in the calculation of the action for infinitesimal time differences. The examples were chosen to produce the two other well-known rules, the Weyl rule and the symmetrization rule.

EXAMPLES

Example 1: Suppose in calculating the action we estimate the Hamiltonian by $H(\frac{1}{2}(X' + X''), p)$. The action is then

$$A = p(X'' - X') - (t'' - t')H(\tfrac{1}{2}(X' + X''), p). \tag{9}$$

By use of the momentum as a parameter to define the set of all paths, the propagator becomes

$$K_1 = \frac{1}{2\pi\hbar} \int dp e^{i/\hbar[p(X'' - X') - (t'' - t')H]}.$$

By expanding in powers of $t'' - t'$ and using (4), we have in the limit

$$i\hbar \frac{\partial \psi}{\partial t} = \int k_1(X'', X')\psi(X', t')\, dX' \tag{10}$$

and

$$\hat{H}\psi(X'') = \int k_1(X'', X')\psi(X', t)\, dX', \tag{11}$$

where now

$$k_1(X'', X') = \frac{1}{2\pi\hbar}\int H(\tfrac{1}{2}(X' + X''), p)e^{i/\hbar p(X''-X')}\, dp. \tag{12}$$

Equations (11) and (12) [also (18) below] correspond to Eqs. (6) and (5) of Kerner and Sutcliffe's paper. In general, they will give different answers for the quantum mechanical Hamiltonian operator even though the same classical Hamiltonian is used.

If we take H in the form $f(X)p^k$, then

$$\hat{H}\psi(X'', t) = \frac{1}{2\pi\hbar}\iint e^{i/\hbar p(X''-X')}f(\tfrac{1}{2}(X' + X''))p^k\, dX'\, dp \tag{13}$$

$$= (-i\hbar)^k\int \delta^k(X'' - X')f(\tfrac{1}{2}(X' + X''))\, dX', \tag{14}$$

$$\hat{H}\psi(X, t) = (-i\hbar)^k\sum_{l=0}^{k}\binom{k}{l}\frac{f^l(X)}{2^l}\left(\frac{\partial}{\partial X}\right)^{k-l}\psi(X, t). \tag{15}$$

Specializing further to the case $f(X) = X^m$, the quantum mechanical operator becomes

$$\hat{H} = \sum_l (-i\hbar)^l\binom{k}{l}\binom{m}{l}\frac{l!}{2^l}X^{m-l}\hat{p}^{k-l}, \tag{16}$$

which is the Weyl rule of ordering.

Example 2: Approximating the Hamiltonian by $\frac{1}{2}[H(X') + H(X'')]$ and following the same procedure, we have

$$i\hbar\frac{\partial\psi}{\partial t} = \int k_2(X'', X')\psi(X')\, dX' \tag{17}$$

with

$$k_2(X'', X') = \int \tfrac{1}{2}[H(X') + H(X'')]e^{i/\hbar p(X''-X')}\, dp. \tag{18}$$

For a classical Hamiltonian of the form $X^m p^k$, this yields

$$\sum_l (-i\hbar)^l\binom{k}{l}\binom{m}{l}\frac{l!}{2}(\delta_{l0} + 1)X^{m-l}\hat{p}^{k-l}, \tag{19}$$

which is the rule of symmetrization.

CONCLUSION

One can find other approximations to the action which would give other rules of ordering besides these well-known ones. The whole set of rules given in Ref. 2 can be so derived. The reason for this vast choice is that the value of the Hamiltonian between the infinitesimally close points X' and X'' is not critical since the limit will eventually be taken. The only requirement is that the $H(X', X'', p)$ chosen must be such that, when X'' approaches X', H goes to $H(X', p)$. It is worth noticing that, if the classical Hamiltonian is of the form $f(X) + g(p)$, then all the rules give the same answer.

As is well known, the Schrödinger and Heisenberg formulation does not force a unique H or a unique ordering rule from the classical function. We have shown that the Feynman formulation does not either.

ACKNOWLEDGMENT

This work was done during tenure of a grant by the National Science Foundation for a different project. We are indebted for the support.

[1] E. Kerner and W. Sutcliffe, J. Math. Phys. **11**, 391 (1970).
[2] L. Cohen, J. Math. Phys. **7**, 781 (1966).
[3] L. Cohen, Phil. Sci. **33**, 317 (1966).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Adiabatic Response to an Oscillatory Field*

RALPH H. YOUNG
Department of Chemistry, Jackson State College, Jackson, Mississippi 39217
AND
WALTER J. DEAL, JR.
Department of Chemistry, University of California at Riverside, Riverside, California 92502

(Received 16 March 1970)

The behavior of a quantum-mechanical system with a slowly modulated oscillatory Hamiltonian is characterized by an adiabatic theorem similar to that for a system with a slowly changing "static" Hamiltonian. Quasiperiodic states—solutions of the instantaneous Schrödinger equation with an oscillatory Hamiltonian—play the same role as eigenfunctions of the instantaneous Hamiltonian do in an adiabatic theorem for a nearly static Hamiltonian. As an example, the theorem is used to establish the correct wavefunction to be used in computing the refractive index of atomic hydrogen.

1. INTRODUCTION

Consider an isolated system in a stationary state. If an off-resonance oscillatory field is slowly turned on, into what state does the system go? In a different context, Nozières and Pines[1] have dealt with this question through first order in time-dependent perturbation theory, using a particularly simple form for the modulated field. Since the conventional (variation-of-constants) time-dependent perturbation theory converges rapidly only for short times, it does not seem an appropriate tool for investigating the behavior of the system when the field is turned on very slowly.[2] Furthermore, it would be useful to have an answer to this question which does not depend on the exact way in which the field is turned on.

This paper is concerned with the general question of the time development of a wavefunction for a system whose Hamiltonian is a slowly modulated oscillatory function of time. Separate consideration is given to finite-basis (e.g., spin) problems and to infinite-basis (e.g., atomic) problems. Modifications needed to adapt the present treatment to the resonant absorption of energy are outlined in Sec. 5.

An oscillatory Hamiltonian may depend on a number of parameters, such as the strength, direction, or frequency of an applied field. When the Hamiltonian is modulated, these parameters become functions of time; the course of the modulation may be described by their dependence on a single variable σ which specifies a point on the modulation path. The effect of the modulation rate may be studied by traversing a fixed segment of the path, say from $\sigma = 0$ to $\sigma = 1$, at a rate inversely proportional to a time scale T,

$$\sigma = t/T, \tag{1}$$

and by observing the limiting behavior of the system as T becomes large.

An oscillatory quantity is characterized by its fundamental frequency ω and the values it assumes during one cycle of its oscillation. Thus a modulated Hamiltonian can be written as

$$\mathcal{H} = \mathcal{H}(\varphi; \sigma), \tag{2}$$

where φ is the phase

$$\varphi(t) = \int_0^t \omega'\, dt'. \tag{3}$$

{For convenience, complicated expressions like $\omega[\sigma(t')]$ are abbreviated throughout this discussion. For example, $\omega[\sigma(t)]$ is written ω; $\omega[\sigma(t')]$ is written ω'. That is, ω is given the value it has at time t; ω' is given the value taken on by ω at time t'. Other quantities are abbreviated similarly.}

2. ADIABATIC THEOREM FOR A FINITE-BASIS SYSTEM

Consider first a system whose description involves only a finite basis set. If σ were constant, there would be a complete set of solutions to the Schrödinger equation (quasiperiodic states) of the "normal-mode" form[3,4] ${}^k\psi(\varphi; \sigma) \exp[-i\,{}^k\mathcal{E}(\sigma)t]$, where the "Bloch state" ${}^k\psi$ satisfies the equation

$$(\mathcal{H} - {}^k\mathcal{E} - i\omega\partial_\varphi)\,{}^k\psi(\varphi; \sigma) = 0. \tag{4}$$

A modulated Hamiltonian is not exactly periodic and, strictly speaking, has no quasiperiodic states. Over a short period of time, however, such a Hamiltonian may be very nearly periodic, so that it makes sense to discuss the problem in terms of "instantaneous quasiperiodic states"

$${}^k\psi(\varphi; \sigma) \exp\left(-i\int_0^t {}^k\mathcal{E}'\, dt'\right).$$

From the totality of such states, one may select a subset S (which might consist of only one state) to serve as a basis for an approximate solution of the

time-dependent Schrödinger equation. A general time-dependent linear combination of these,

$$\tilde{\Psi} = \sum_{k\in S} {}^{k}c(t) \exp\left(-i\int_0^t {}^{k}\mathcal{E}'\,dt'\right) {}^{k}\psi(\varphi;\sigma), \quad (5)$$

is most nearly a solution of the Schrödinger equation if the coefficients are obtained by "solving the Schrödinger equation within the limited basis"[5]:

$$\langle {}^{l}\psi(\varphi;\sigma)|\,\mathcal{H}(t) - id_t\,|\tilde{\Psi}(t)\rangle = 0, \quad 0 < t < T, \text{ all } l \in S, \quad (6a)$$

$$d_t\,{}^{l}c = -T^{-1}\sum_{k\in S} \exp\left(i\int_0^t ({}^{l}\mathcal{E}' - {}^{k}\mathcal{E}')\,dt'\right) {}^{k}c\langle {}^{l}\psi \mid \partial_\sigma\, {}^{k}\psi\rangle. \quad (6b)$$

Equation (6b) preserves the norm of $\tilde{\Psi}(t)$. From (6b) it follows that

$$(\mathcal{H} - id_t)\tilde{\Psi} = -iT^{-1}(1-P)\sum_{k\in S}\exp\left(-i\int_0^t {}^{k}\mathcal{E}'\,dt'\right) {}^{k}c\,\partial_\sigma\,{}^{k}\psi, \quad (7a)$$

$$P = \sum_{l\in S} |{}^{l}\psi(\varphi;\sigma)\rangle\langle {}^{l}\psi(\varphi;\sigma)|. \quad (7b)$$

The mixing coefficients vary slowly when T is long:

$$|d_t\,{}^{l}c| \le T^{-1}\sum_{k\in S} |\langle {}^{l}\psi \mid \partial_\sigma\,{}^{k}\psi\rangle| \equiv {}^{l}B(\varphi;\sigma)T^{-1}. \quad (8)$$

Theorem: Let the system be initially ($t = \sigma = 0$) in an arbitrary linear combination $\tilde{\Psi}(0)$ of quasiperiodic states belonging to S:

$$\tilde{\Psi}(0) = \sum_{k\in S} {}^{k}c(0)\,{}^{k}\psi(0;0). \quad (9)$$

Then, when σ reaches 1 ($t = T$), the system will be in the state $\tilde{\Psi}(T)$ as given by (5) and (6b), with an error that vanishes as T^{-1} when T becomes large, provided that, for all $t \in [0, T]$:

(i) $\mathcal{H}(\varphi;\sigma)$ is continuous in both variables; $\omega(\sigma)$ and the ${}^{k}\mathcal{E}(\sigma)$, $k \in S$, are continuous in σ;

(ii) the ${}^{k}\psi(\varphi;\sigma)$ and the auxiliary functions ${}^{k}\zeta(\varphi;\sigma)$, $k \in S$, defined by

$$[\mathcal{H}(\varphi;\sigma) - {}^{k}\mathcal{E}(\sigma) - i\omega(\sigma)\partial_\varphi]\,{}^{k}\zeta(\varphi;\sigma) = i(1-P)\partial_\sigma\,{}^{k}\psi, \quad (10a)$$

$${}^{k}\zeta(\varphi + 2\pi;\sigma) = {}^{k}\zeta(\varphi;\sigma), \quad \text{all } \varphi, \quad (10b)$$

$$\langle {}^{l}\psi \mid {}^{k}\zeta\rangle = 0, \quad \text{all } l \in S, \quad (10c)$$

and all their partial derivatives are continuous in both variables. (See below.)

Since the end points of the interval are arbitrary, this theorem amounts to the assertion that, as the time scale becomes long, the exact solution to the Schrödinger equation tends to $\tilde{\Psi}(t)$.

Proof: Let

$$\tilde{\Phi}(t) = \tilde{\Psi}(t) + T^{-1}\sum_{k\in S} {}^{k}c\exp\left(-i\int_0^t {}^{k}\mathcal{E}'\,dt'\right) {}^{k}\zeta, \quad (11a)$$

$$(\mathcal{H} - id_t)\tilde{\Phi} = -iT^{-2}\sum_{k\in S}\exp\left(-i\int_0^t {}^{k}\mathcal{E}'\,dt'\right)\partial_\sigma({}^{k}c\,{}^{k}\zeta), \quad (11b)$$

and let $\Psi(t)$ and $\Phi(t)$ be true solutions to the Schrödinger equation based on $\tilde{\Psi}(0)$ and $\tilde{\Phi}(0)$. Then, according to Appendix C,

$$\|\Phi(T) - \tilde{\Phi}(T)\| \le \int_0^T \|(\mathcal{H}' - id_{t'})\tilde{\Phi}'\|\,dt', \quad (12)$$

and by use of the triangle inequality it follows that

$$\|\Psi(T) - \tilde{\Psi}(T)\| \le T^{-1}\sum_{k\in S}\left(\|{}^{k}c(0)\,{}^{k}\zeta(0;0)\| + \|{}^{k}c(T)\,{}^{k}\zeta[\varphi(T);1]\| + \int_0^1 \|\partial_\sigma({}^{k}c\,{}^{k}\zeta)\|\,d\sigma\right) \quad (13a)$$

$$\le T^{-1}\sum_{k\in S}\left(\|{}^{k}c(0)\,{}^{k}\zeta(0;0)\| + \sup_{0\le\varphi<2\pi}\|{}^{k}\zeta(\varphi;1)\| + \int_0^1 \sup_{0\le\varphi<2\pi}[\|{}^{k}B(\varphi;\sigma)\,{}^{k}\zeta(\varphi;\sigma)\| + \|\partial_\sigma\,{}^{k}\zeta(\varphi;\sigma)\|]\,d\sigma\right). \quad (13b)$$

It is a straightforward matter to see that the continuity conditions advanced are indeed sufficient. Thus $\|\Psi(T) - \tilde{\Psi}(T)\|$ is bounded by a constant multiple of T^{-1}.

The existence and continuity properties of the ${}^{k}\zeta$ may be based on the properties of the ${}^{k}\psi$ themselves. The following conditions are sufficient (but perhaps more restrictive than necessary) to establish that the ${}^{k}\zeta$ exist, are continuous, and have continuous first partial derivatives:

(a) $\omega(\sigma)$ and *all* the ${}^{k}\mathcal{E}(\sigma)$ are continuous and have continuous σ-derivatives,

(b) the ${}^{l}\psi$, $l \notin S$, are continuous and have continuous derivatives with respect to both variables,

(c) $\partial_\sigma\,{}^{k}\psi$ and $\partial_\sigma^2\,{}^{k}\psi$, all $k \in S$, are continuous in both variables,

(d) no ${}^{k}\mathcal{E}(\sigma)$, $k \in S$, differs by only an integral multiple of ω from an ${}^{l}\mathcal{E}(\sigma)$, $l \notin S$,

(e) $\omega(\sigma)$ is never zero, for all $t \in [0, T]$.

If ${}^{k}\zeta$ is expanded in terms of the ${}^{l}\psi$, $l \notin S$, the expansion coefficients are readily obtained; again, it is not difficult to see where each of conditions (a)–(e) is applied.

3. ADIABATIC THEOREM FOR AN INFINITE-BASIS SYSTEM

As discussed previously, it is not expected that an infinite-basis system subjected to an oscillatory perturbation will possess true quasiperiodic states.[3] Nevertheless, it is generally possible to obtain oscillatory functions, involving truncated perturbation series for the Bloch states and quasi-energies, which are good approximate solutions to the time-dependent Schrödinger equation for an oscillatory Hamiltonian:

$$ {}^k\psi^{[L]}(\varphi;\sigma) = \sum_{l=0}^{L} \lambda^l\, {}^k\psi^{(l)}(\varphi;\sigma), \tag{14a} $$

$$ {}^k\psi^{(l)}(\varphi;\sigma) = \sum_{n=A^{(l)}}^{B^{(l)}} {}^k\psi_n{}^{(l)}(\sigma)\exp(in\varphi), \tag{14b} $$

$$ {}^k\mathcal{E}^{[L]}(\sigma) = \sum_{l=0}^{L} \lambda^l\, {}^k\mathcal{E}^{(l)}(\sigma) \tag{14c} $$

(see Appendix B). A modified version of the preceding adiabatic theorem holds for such states. For simplicity the present treatment is restricted to a Hamiltonian having the form

$$ \mathcal{H}(\varphi;\sigma) = \sum_{l=0}^{2} \lambda^l \mathcal{H}^{(l)}(\varphi;\sigma), \tag{15a} $$

$$ \mathcal{H}^{(l)}(\varphi;\sigma) = \sum_{n=-l}^{l} H_n{}^{(l)}(\sigma)\exp(in\varphi), \tag{15b} $$

such as appears in a treatment of a system under the influence of nearly monochromatic radiation. It is further assumed that $H_0{}^{(0)}$ is not modulated. (These restrictions are easily removed.)

Consider the approximate wavefunction

$$ \tilde{\Psi}(t) = \sum_{k\in S} {}^kc(t)\exp\left(-i\int_0^t {}^k\mathcal{E}^{[L]}(\sigma')\,dt'\right) {}^k\psi^{[L]}. \tag{16a} $$

It turns out to be convenient to choose the ${}^kc$ somewhat differently from (6b):

$$ d_t\, {}^kc = -T^{-1}\sum_{l\in S} {}^lc\langle {}^k\psi \mid \partial_\sigma\, {}^l\psi\rangle_0^{[L]} \times \exp\left(i\int_0^t [{}^k\mathcal{E}^{[L]}(\sigma') - {}^l\mathcal{E}^{[L]}(\sigma')]\,dt'\right), \tag{16b} $$

where $\langle {}^k\psi \mid \partial_\sigma\, {}^l\psi\rangle_0^{[L]}$ is the zero-frequency Fourier component of the overlap expression, truncated at Lth order. Thus

$$ [\mathcal{H}(\varphi;\sigma) - id_t]\tilde{\Psi} = \sum_{k\in S} {}^kc\exp\left(-i\int_0^t {}^k\mathcal{E}^{[L]}(\sigma')\,dt'\right)({}^kR^{[L]} - T^{-1}\, {}^k\chi^{[L]}), \tag{17a} $$

$$ {}^kR^{[L]} = (\mathcal{H} - {}^k\mathcal{E}^{[L]} - i\omega\partial_\varphi)\, {}^k\psi^{[L]} = O(\lambda^{L+1}), \tag{17b} $$

$$ -i\, {}^k\chi^{[L]} = \partial_\sigma\, {}^k\psi^{[L]} - \sum_{l\in S} {}^l\psi^{[L]}\langle {}^l\psi \mid \partial_\sigma\, {}^k\psi\rangle_0^{[L]}, \tag{17c} $$

$$ \langle {}^j\psi^{[L]} \mid {}^k\chi^{[L]}\rangle_0 = O(\lambda^{L+1}), \quad \text{all} \quad j,k\in S. \tag{17d} $$

To obtain bounds on ${}^kc(t)$ and $d_\sigma\, {}^kc$, we regard (16b) as a Schrödinger equation for the vector $\mathbf{c}(t)$ whose elements are the ${}^kc(t)$:

$$ id_t\mathbf{c} = \mathbf{M}\mathbf{c}, \tag{18a} $$

$$ {}^{kl}M = -iT^{-1}\langle {}^k\psi \mid \partial_\sigma\, {}^l\psi\rangle_0^{[L]} \times \exp\left(i\int_0^t [{}^k\mathcal{E}^{[L]}(\sigma') - {}^l\mathcal{E}^{[L]}(\sigma')]\,dt'\right). \tag{18b} $$

It is possible to arrange for the ${}^k\psi^{[L]}$ to be zero-frequency orthonormal through Lth order [cf. Eq. (B3)]:

$$ \langle {}^j\psi \mid {}^k\psi\rangle_0^{[L]} = \delta_{jk}, $$

from which fact it follows readily that $\mathbf{M}$ is Hermitian and that $\|\mathbf{c}\|$ is conserved. Equation (16b) can again be used to obtain bounds on $d_\sigma\, {}^kc$; for instance,

$$ |d_\sigma\, {}^kc| \le \|\mathbf{c}(0)\| \left(\sum_{l\in S} |\langle {}^l\psi \mid \partial_\sigma\, {}^k\psi\rangle_0^{[L]}|^2\right)^{\frac{1}{2}} \equiv {}^kB(\sigma). \tag{19} $$

It will be necessary to introduce quantities analogous to the ${}^k\zeta$ of the preceding discussion:

$$ {}^k\zeta^{[L]}(\varphi;\sigma) = \sum_{l=1}^{L} \lambda^l\, {}^k\zeta^{(l)}(\varphi;\sigma), \tag{20a} $$

$$ {}^k\zeta^{(l)}(\varphi;\sigma) = \sum_{n=A^{(l)}}^{B^{(l)}} {}^k\zeta_n{}^{(l)}(\sigma)\exp(in\varphi), \tag{20b} $$

$$ (\mathcal{H} - {}^k\mathcal{E}^{[L]} - i\omega\partial_\varphi)\, {}^k\zeta^{[L]} = {}^k\chi^{[L]} + {}^kV^{[L]}, \qquad {}^kV^{[L]} = O(\lambda^{L+1}). \tag{21} $$

Their existence and properties will be discussed after the modified adiabatic theorem is discussed.

Theorem: Let the system be initially in an arbitrary linear combination $\tilde{\Psi}(0)$ of the ${}^k\psi^{[L]}$, $k\in S$:

$$ \tilde{\Psi}(0) = \sum_{k\in S} {}^kc(0)\, {}^k\psi^{[L]}(0;0). \tag{22} $$

Then when σ reaches 1, the true state of the system, $\Psi(T)$, will be approximated by $\tilde{\Psi}(T)$ [Eqs. (16)] according to

$$ \|\Psi(T) - \tilde{\Psi}(T)\| \le XT^{-1} + Y + ZT \tag{23} $$

[X, Y, and Z are constants given in (30)], provided that[6]:

(i) the quantities $H_n{}^{(l)}\, {}^k\psi_{n'}{}^{(l')}$, $H_n{}^{(l)}\, {}^k\zeta_{n'}{}^{(l')}$, $\partial_\sigma\, {}^k\psi_n{}^{(l)}$, and $\partial_\sigma\, {}^k\zeta_n{}^{(l)}$, $k\in S$, are continuous functions of σ,

(ii) ω and the ${}^k\mathcal{E}^{[L]}$ are continuous functions of σ,

(iii) the exact solutions to the Schrödinger equation, $\Psi(t)$ and $\Phi(t)$, based on the initial conditions (22) and (24),

$$ \tilde{\Phi}(0) = \tilde{\Psi}(0) + T^{-1}\sum_{k\in S} {}^kc\, {}^k\zeta^{[L]}(0;0), \tag{24} $$

have continuous total derivatives $d_t\Psi$ and $d_t\Phi$.

Proof: Let

$$\tilde{\Phi}(t) = \tilde{\Psi}(t) + T^{-1}\sum_{k\in S} {}^{k}c \exp\left(-i\int_0^t {}^{k}\mathcal{E}^{[L]}(\sigma')\,dt'\right) {}^{k}\zeta^{[L]}, \tag{25}$$

$$\begin{aligned}(\mathcal{H} - id_t)\tilde{\Phi} = \sum_{k\in S} &\exp\left(-i\int_0^t {}^{k}\mathcal{E}^{[L]}(\sigma')\,dt'\right) \\ &\times [{}^{k}c({}^{k}R^{[L]} - T^{-1}\,{}^{k}V^{[L]} - iT^{-2}\partial_\sigma\,{}^{k}\zeta^{[L]}) \\ &- iT^{-2}(\partial_\sigma\,{}^{k}c)\,{}^{k}\zeta^{[L]}].\end{aligned} \tag{26}$$

Then, according to Appendix C,

$$\|\Phi(T) - \tilde{\Phi}(T)\| \leq \int_0^t \|(\mathcal{H}' - id_{t'})\tilde{\Phi}'\|\,dt', \tag{27}$$

$$\begin{aligned}\|\Psi(T) - \tilde{\Psi}(T)\| &\leq \|\Psi(T) - \Phi(T)\| \\ &\quad + \|\Phi(T) - \tilde{\Phi}(T)\| + \|\tilde{\Phi}(T) - \tilde{\Psi}(T)\| \qquad (28)\\ &\leq \|\tilde{\Phi}(0) - \tilde{\Psi}(0)\| + \|\tilde{\Phi}(T) - \tilde{\Psi}(T)\| \\ &\quad + \int_0^t \|(\mathcal{H}' - id_{t'})\tilde{\Phi}'\|\,dt' \qquad (29)\\ &\leq XT^{-1} + Y + ZT, \qquad (30a)\end{aligned}$$

$$\begin{aligned}X = \sum_{k\in S}\Bigg[&\|{}^{k}c(0)\,{}^{k}\zeta^{[L]}(0;0)\| \\ &+ \|\mathbf{c}(0)\| \sup_{0\leq\varphi<2\pi}\|{}^{k}\zeta^{[L]}(\varphi;1)\| \\ &+ \int_0^1 \Bigg(\|\mathbf{c}(0)\| \sup_{0\leq\varphi<2\pi}\|\partial_\sigma\,{}^{k}\zeta^{[L]}(\varphi;\sigma)\| \\ &+ {}^{k}B(\sigma) \sup_{0\leq\varphi<2\pi}\|{}^{k}\zeta^{[L]}(\varphi;\sigma)\|\Bigg)\,d\sigma\Bigg],\end{aligned} \tag{30b}$$

$$Y = \sum_{k\in S}\|\mathbf{c}(0)\| \int_0^1 \sup_{0\leq\varphi<2\pi}\|{}^{k}V^{[L]}(\varphi;\sigma)\|\,d\sigma = O(\lambda^{L+1}), \tag{30c}$$

$$Z = \sum_{k\in S}\|\mathbf{c}(0)\| \int_0^1 \sup_{0\leq\varphi<2\pi}\|{}^{k}R^{[L]}(\varphi;\sigma)\|\,d\sigma = O(\lambda^{L+1}). \tag{30d}$$

Tighter bounds which depend on T are readily obtained.

The crucial question, of course, concerns the existence and continuity properties of the ${}^{k}\zeta^{[L]}$. Consider the perturbation equations for the ${}^{k}\psi^{[L]}$ themselves:

$$(H_0{}^{(0)} - i\omega\partial_\varphi - {}^{k}\mathcal{E}^{(0)})\,{}^{k}\psi^{(0)} = 0, \tag{31a}$$

$$\begin{aligned}(H_0{}^{(0)} - i\omega\partial_\varphi - {}^{k}\mathcal{E}^{(0)})\,{}^{k}\psi^{(l)} \\ = -\sum_{l'=1}^{l}(\mathcal{H}^{(l')} - {}^{k}\mathcal{E}^{(l')})\,{}^{k}\psi^{(l-l')}.\end{aligned} \tag{31b}$$

These equations are Rayleigh–Schrödinger perturbation theory in a Hilbert space $\mathfrak{h}(A^{(L)}, B^{(L)})$ composed of periodic functions of the form[7]

$$\sum_{n=A^{(L)}}^{B^{(L)}} f_n \exp(in\varphi), \tag{32}$$

where the f_n are elements in the Hilbert space $\mathfrak{h}$ of phase-independent functions conventionally used to describe the physical system. The appropriate inner product in this space is the zero-frequency Fourier component of the overlap, $\langle\,|\,\rangle_0$. The unperturbed operator $H_0{}^{(0)} - i\omega\partial_\varphi$ is self-adjoint in this space, and the perturbations $\mathcal{H}^{(1)}$ and $\mathcal{H}^{(2)}$ are symmetric (assuming that $H_0{}^{(0)}$ is self-adjoint and the $\mathcal{H}^{(l)}$, $l = 1, 2$, are symmetric in $\mathfrak{h}$). The unperturbed eigenfunctions and eigenvalues are the ${}^{k}\psi^{(0)}$ and ${}^{k}\mathcal{E}^{(0)}$.

The perturbation theory of the ${}^{k}\psi^{[L]}$ is discussed in more detail in Appendix B. Suffice it to say that the ${}^{k}\psi^{[L]}$ come in families which are degenerate in $\mathfrak{h}(A^{(L)}, B^{(L)})$ at zeroth order (that is, that have the same ${}^{k}\mathcal{E}^{(0)}$).

The perturbation equations for the ${}^{k}\zeta^{[L]}$,

$$(H_0{}^{(0)} - i\omega\partial_\varphi - {}^{k}\mathcal{E}^{(0)})\,{}^{k}\zeta^{(0)} = {}^{k}\chi^{(0)}, \tag{33a}$$

$$\begin{aligned}(H_0{}^{(0)} - i\omega\partial_\varphi - {}^{k}\mathcal{E}^{(0)})\,{}^{k}\zeta^{(l)} \\ = {}^{k}\chi^{(l)} - \sum_{l'=1}^{l}(\mathcal{H}^{(l')} - {}^{k}\mathcal{E}^{(l')})\,{}^{k}\zeta^{(l-l')},\end{aligned} \tag{33b}$$

$$\langle {}^{j}\psi^{[L]}\,|\,{}^{k}\chi^{[L]}\rangle_0 = O(\lambda^{L+1}), \quad \text{all } j \in S, \tag{33c}$$

$${}^{k}\chi^{(l)} \in \mathfrak{h}(A^{(l)}, B^{(l)}) \subseteq \mathfrak{h}(A^{(L)}, B^{(L)}), \tag{33d}$$

are of the same form as those that have recently been discussed in connection with the adiabatic theorem for pseudo-eigenfunctions of a nearly static Hamiltonian. It is safe to assume that the individual Fourier components of the rhs of (33a) and (33b) are square integrable through as many orders $l - 1$ as are possible to carry the perturbation theory; the entire rhs then has finite norm and also satisfies the frequency criterion for belonging to $\mathfrak{h}(A^{(L)}, B^{(L)})$. It follows from the discussion referred to that these equations possess solutions through Lth order *provided that the set S contains all members of a family corresponding to a single ${}^{k}\mathcal{E}^{(0)}$, if it contains any one, and that none of the numbers ${}^{k}\mathcal{E}^{(0)} - n\omega$, $A^{(L)} \leq n \leq B^{(L)}$, lies in the continuous spectrum of $H_0{}^{(0)}$.* It is possible for there to be solutions even if the second condition is not met.[8]

The more difficult investigation of the continuity properties of the ${}^{k}\zeta^{[L]}$ will not be attempted here. For real physical cases involving the interaction of a material system with a modulated-monochromatic light beam, there is every reason to believe that the provisos of the theorem are satisfied. A preliminary investigation suggests that it should be possible to establish these continuity properties directly on the form and continuity properties of the Hamiltonian.

4. COROLLARIES

The following corollaries apply to the finite-basis case.

Corollary 1: When the conditions on the ${}^k\zeta$ are violated at a constant number of points during the interval $[0, T]$, then $\|\Psi(T) - \tilde{\Psi}(T)\|$ vanishes as T becomes large (although the exact manner in which it does so may vary from case to case). The proof is virtually identical to that of Corollary 1 in Ref. 5.

According to this corollary, adiabatic behavior is not precluded by the presence of points where condition (d) is violated and points where ω goes through zero (e.g., a point at which an oscillatory perturbation is turned on). In order to insure that nonadiabatic terms vanish at T^{-1}, it still seems necessary to require that ω not pass through zero. A corresponding experiment might begin at zero intensity and nonzero frequency, rather than vice versa.

Corollary 2: Suppose that the ${}^k\mathcal{E}$, $k \in S$, are identical or differ by a constant multiple of ω throughout the modulation. Then the mixing coefficients ${}^kc(t)$ become independent of T, with an error which vanishes as T^{-1} provided that the following (probably unduly stringent) conditions hold:

(i) ${}^k\psi$, $\partial_\sigma {}^k\psi$, and $\partial_\sigma{}^2 {}^k\psi$, all $k \in S$, are continuous in both variables,

(ii) $\partial_\sigma\omega$ is continuous in σ,

(iii) $\omega(\sigma)$ is never zero.

Since an arbitrary multiple of ω may be added to a quasi-energy, with a compensating redefinition of the corresponding Bloch function, the case of quasi-energies differing by a multiple of ω throughout the modulation readily reduces to the case of identical quasienergies; without loss of generality, one may restrict consideration to the latter case.

Equation (6b) may be rewritten in the form

$$(\mathbf{h} - id_\sigma)\mathbf{c} = 0, \tag{34a}$$

where $\mathbf{c}$ is a vector whose components are the ${}^kc(t)$ and

$${}^{lk}h(\varphi; \sigma) = -i\langle {}^l\psi \mid \partial_\sigma {}^k\psi\rangle. \tag{34b}$$

The average of $\mathbf{h}(\varphi; \sigma)$ over one cycle (with σ held constant) defines its zero-frequency Fourier component $\mathbf{h}_0(\sigma)$, which figures in another "Schrödinger equation"

$$(\mathbf{h}_0 - id_\sigma)\mathbf{f}(\sigma) = 0. \tag{35}$$

A more precise statement of the corollary is that, if

$$\mathbf{f}(0) = \mathbf{c}(0), \tag{36}$$

then $\mathbf{c}(T) - \mathbf{f}(1)$ vanishes as T^{-1} when T becomes long.

The matrices $\mathbf{h}$ and $\mathbf{h}_0$ are Hermitian. Thus, if $\mathbf{z}(\varphi; \sigma)$ is defined by

$$\mathbf{z}(\varphi; \sigma) = -\omega^{-1}\int_0^\varphi [\mathbf{h}(\varphi'; \sigma) - \mathbf{h}_0(\sigma)]\mathbf{f}(\sigma)\, d\varphi', \qquad \sigma = \text{const}, \tag{37}$$

one may use the error bound in Appendix C to obtain

$$\|\mathbf{c}(T) - \mathbf{f}(1)\| \leq T^{-1}\left(\|\mathbf{z}[\varphi(T); 1]\| + \|\mathbf{z}(0; 0)\| + \int_0^1 \|[\mathbf{h}(\varphi; \sigma) - i\partial_\sigma]\mathbf{z}(\varphi; \sigma)\|\, d\sigma\right), \tag{38a}$$

$$\leq T^{-1}\left(\sup_{0\leq\varphi<2\pi}\|\mathbf{z}(\varphi; 1)\| + \|\mathbf{z}(0; 0)\| + \int_0^1 \sup_{0\leq\varphi<2\pi}\|[\mathbf{h}(\varphi; \sigma) - i\partial_\sigma]\mathbf{z}(\varphi; \sigma)\|\, d\sigma\right). \tag{38b}$$

If ω passes through zero, $\|\mathbf{c}(T) - \mathbf{f}(1)\|$ can be shown to vanish, but not as T^{-1}. (See Corollary 1 above and Corollary 1 of Ref. 5.)

Corollary 3: The following simple but rather restrictive conditions are sufficient for the validity of the theorem and Corollary 2:

(i) the dependence of $\mathcal{H}$ and ω on σ is

$$\mathcal{H} = \mathcal{H}(\varphi; \mu), \quad \omega = \omega(\mu), \quad \mu = \mu(\sigma), \tag{39}$$

(ii) $\mathcal{H}(\varphi; \mu)$ and $\omega(\mu)$ are analytic functions of their (real-valued) arguments,

(iii) $\mu(\sigma)$ possesses a continuous first and piecewise continuous second derivative, $0 \leq \sigma \leq 1$,

(iv) at no time does a ${}^k\mathcal{E}(\mu)$, $k \in S$, differ from an ${}^l\mathcal{E}(\mu)$, $l \notin S$, by an integral multiple of ω,

(v) $\omega(\mu)$ is nonzero at all times,

(vi) the ${}^kc(T)$ are determined by integrating (6b).

Corollary 1 holds if conditions (4) and/or (5) are violated at a constant number of σ values.

With a *fixed* value of μ, one can solve the Schrödinger equation to obtain an evolution operator $U(t, 0; \mu)$. This operator is analytic in both variables [conditions (i), (ii), and (v)].[9] It follows that the ${}^k\psi(\varphi; \mu)$ and the ${}^k\mathcal{E}(\mu)$ are analytic functions of their arguments.[10]

From this analyticity follow conditions (a), (b), and (c), with the exception that $\partial_\sigma{}^2 {}^k\psi$ is only piecewise continuous. It is readily seen that this degree of discontinuity does not prevent the ${}^k\zeta$ from satisfying condition (ii) of the theorem. For details, see Ref. 11, where a rather different approach is followed.

5. DISCUSSION

The finite-basis adiabatic theorem, if its requirements are met, provides a complete characterization of the behavior of a finite-basis system with a modulated oscillatory Hamiltonian in the limit of slow modulation:

(a) Quasiperiodic states whose quasi-energies differ by a multiple of ω at most a fixed number of times during the modulation do not mix, and

(b) the mixing of quasiperiodic states whose quasi-energies are identical throughout the modulation depends only on the fraction of the modulation path traversed.

In particular, when the modulation path is such that one quasi-energy never differs from any other by an integer multiple of ω, the corresponding quasiperiodic state becomes an exact solution of the time-dependent Schrödinger equation (provided that the phase of the Bloch function is appropriately chosen as a function of σ). At intermediate time scales, the theorem can be used to describe the way in which the states in one set mix with one another while adiabatically not mixing with those in another set.

The conclusions of the infinite-basis theorem must be expressed in a somewhat different fashion. Their major application is to the interaction of matter with a classical radiation field; in that context, one may not consider a variable time scale because a vector potential of the form $\mathbf{A}_+(\mathbf{r};\sigma)\exp(i\varphi) + \mathbf{A}_-(\mathbf{r};\sigma)\times\exp(-i\varphi)$ can generally satisfy Maxwell's equations for only one value of T. The theorem nevertheless establishes an "adiabatic approximation" to the solution of the Schrödinger equation for matter subjected to a slowly modulated classical radiation field. The error bound (30) contains terms of order T^{-1}, λ^{L+1}, and $\lambda^{L+1}T$, where L is the order at which the perturbation series terminate. The first represents deviation from adiabaticity; the second and third result from using truncated perturbation series. When no two ${}^k\mathcal{E}^{(0)}$, $k \in S$, differ by an integer multiple of ω, the first is actually of order λ/T, rather than T^{-1}, because the $\partial_\sigma\,{}^k\psi^{(0)}$ may be arranged to be zero. In applications, one must choose L on the basis of the physical situation: that L which yields the smallest error estimate is probably the best approximate expression for the time-dependent wavefunction.

The quality of the "adiabatic approximation" $\tilde{\Psi}(t)$ may be illustrated by the example of a hydrogen atom irradiated with blue light (4860 Å). Let the amplitude of the oscillatory vector potential grow from zero at $t = 0$ to a maximum at $T/2 = 1$ nsec (10^6 cycles of the field) and decrease back to zero at T. (This "rise time" would be necessary if the frequency spectrum were to be pure within 1 ppm.) Let the maximum amplitude correspond to a moderate intensity, 1 W/cm^2. The perturbation equations admit of a solution through third order based on the unperturbed ${}^{1s}\psi^{(0)}$; the four-photon resonance at 4ω precludes a fourth-order solution. (A higher-order perturbation treatment would have to treat $1s$, $2s$, and $2p$ as degenerate in a larger Hilbert space of periodic functions.) If the system is initially in the $1s$ state, ${}^{1s}\psi^{[3]}$ is a reasonable approximate wavefunction through the period of irradiation. One may estimate the orders of magnitude of X/T, Y, and ZT as 10^{-14}, 10^{-22}, and 10^{-14} at times between 0 and T. Thus the error bound on the wavefunction is considerably smaller than $\|\lambda^2\,{}^{1s}\psi^{(2)}\|$ (10^{-12}) but is larger than $\|\lambda^3\,{}^{1s}\psi^{(3)}\|$; ${}^{1s}\psi^{[2]}$ is thus a demonstrably good approximate wavefunction. Note that it is necessary to include $\lambda^3\,{}^{1s}\psi^{(3)}$ in the approximate wavefunction to obtain this estimate.

There is a limitation on the scope of the theorem in cases involving resonance. Suppose S contains the members of a family of ${}^k\psi^{[L]}$ corresponding to a particular ${}^k\mathcal{E}^{(0)}$ and that there is an eigenfunction of $H_0^{(0)}$, ϕ, which has energy ${}^k\mathcal{E}^{(0)} - m\omega_0$, $m = A^{(L)}$, $A^{(L+1)}, \cdots$, or $B^{(L)}$. If ω passes through ω_0 at some point in the modulation path, $\phi\exp(im\varphi)$ there becomes degenerate in $\mathfrak{h}(A^{(L)}, B^{(L)})$ with the ${}^k\psi^{(0)}$ in that family. As a rule the perturbation series ${}^k\psi^{[L]}$ will diverge as ω approaches ω_0, and the theorem becomes inapplicable. The corresponding problem in static perturbation theory appears when the splitting between two manifolds of degenerate unperturbed functions is small compared to a perturbation—a zero-order eigenfunction "should" be a linear combination of functions belonging to both manifolds, but perturbation theory constraints it to belong to one or the other; large higher-order corrections are needed to compensate for this weakness in the perturbation theory. The difficulty in the time-dependent problem might be resolved by slightly altering the unperturbed operator so as to render the two manifolds exactly degenerate throughout the modulation path and compensating for this change by adding corrective terms to the perturbation. In the present example, the new unperturbed Hamiltonian would be $\tilde{H}_0^{(0)}$,

$$\tilde{H}_0^{(0)} = H_0^{(0)} + m[\omega_0 - \omega(\sigma)]\,|\phi\rangle\langle\phi|, \qquad (40)$$

under which $\phi\exp(im\varphi)$ would always be exactly degenerate with the former ${}^k\psi^{(0)}$. The compensating term might be added to $\mathcal{H}^{(1)}$:

$$\tilde{H}_0^{(1)} = \lambda^{-1}m[\omega(\sigma) - \omega_0]\,|\phi\rangle\langle\phi|; \qquad (41)$$

or to $\mathcal{H}^{(2)}$:

$$\tilde{H}_0^{(2)} = H_0^{(2)} + \lambda^{-2}m[\omega(\sigma) - \omega_0]\,|\phi\rangle\langle\phi|. \qquad (42)$$

The choice between (41) and (42) would be based on comparison of $m[\omega(\sigma) - \omega_0]$ with the effects of $\lambda\mathcal{H}^{(1)}$ and $\lambda^2\mathcal{H}^{(2)}$. [It is no great problem to render the appearance of λ in (41) or (42) consistent with the use of λ as a formal perturbation parameter.] With the unperturbed Hamiltonian thus modified, there would be no more ${}^k\mathcal{E}^{(0)}$-crossing points, and the present adiabatic theorem would be applicable. It should be mentioned that no modification of the Hamiltonian is needed for a modulation process which can be described in terms of modulated field amplitude, phase, and direction but constant frequency (exactly at resonance). Further treatment of the resonant case is planned.

APPENDIX A: EXISTENCE OF SOLUTIONS TO RAYLEIGH-SCHRÖDINGER PERTURBATION THEORY[12]

The conclusions of this appendix are prerequisite to the treatment, in Appendix B, of the existence of the functions ${}^k\psi^{[L]}$.

Consider the Rayleigh–Schrödinger perturbation problem associated with an operator H,

$$H = \sum_{l=0}^{\Lambda} \lambda^l H^{(l)}; \tag{A1}$$

let truncated perturbation expansions of an eigenfunction and eigenvalue of H be

$$\psi_k^{[L]} = \sum_{l=0}^{L} \lambda^l \psi_k^{(l)}, \tag{A2a}$$

$$E_k^{[L]} = \sum_{l=0}^{L} \lambda^l E_k^{(l)}. \tag{A2b}$$

Then, through as many orders l as possess solutions, $\psi_k^{(l)}$ and $E_k^{(l)}$ satisfy

$$(H^{(0)} - E_k^{(0)})\psi_k^{(0)} = 0, \tag{A3a}$$

$$(H^{(0)} - E_k^{(0)})\psi_k^{(l)} = -[(H^{[l]} - E_k^{[l]})\psi_k^{[l-1]}]^{(l)}, \quad l \geq 1, \tag{A3b}$$

where the superscript l on the rhs denotes the lth-order part of the bracketed quantity and $H^{[L]}$ is H truncated at lth order. It is assumed that $H^{(0)}$ is self-adjoint and the $H^{(l)}$, $l \geq 1$, are symmetric.

Suppose Eqs. (A3) possess solutions through some particular order $l = m - 1$; is there a solution through mth order? Let $\mathcal{S}$ be the set of all indices j such that $\psi_j^{(0)}$ is a discrete eigenfunction of $H^{(0)}$ with the same eigenvalue as $\psi_k^{(0)}$, $E_j^{(0)} = E_k^{(0)}$, and let $\mathcal{M}$ be the linear manifold spanned by the $\psi_j^{(0)}$, $j \in \mathcal{S}$. Then (A3b) definitely possesses solutions at order $l = m$ if (i) the rhs is square integrable and (ii) orthogonal to $\mathcal{M}$ and if (iii) $E_k^{(0)}$ is not embedded in the continuous spectrum of $H^{(0)}$. [If $E_k^{(0)}$ is embedded in the continuous spectrum, (A3b) may still possess a solution.[8]] It is safe to assume that condition (i) is met; the remainder of this appendix concerns condition (ii).

If $E_k^{(0)}$ is nondegenerate, the requisite orthogonality is trivially achieved by appropriately choosing $E_k^{(m)}$. If $E_k^{(0)}$ is degenerate, there are more general solutions based on $\mathcal{M}$, $\tilde{\psi}_j^{[m-1]}$, of the form

$$\tilde{\psi}_j^{[m-1]} = \psi_j^{[m-1]} + \sum_{p\in\mathcal{S}} \sum_{l=m-M_{pj}-1}^{m-1} \lambda^l a_{pj}^{(l)} \psi_p^{[m-l-1]}, \tag{A4}$$

where M_{pj} is either the highest-order s at which $E_p^{(s)} = E_j^{(s)}$ or $m - 1$ if E_p and E_j are identical through $(m - 1)$th order, for

$$(H^{[m-1]} - E_j^{[m-1]})\tilde{\psi}_j^{[m-1]} = O(\lambda^m). \tag{A5}$$

It is desired to choose the coefficients $a_{pj}^{(l)}$ and the $E_j^{(m)}$ so that

$$\langle \psi_i^{(0)} | H^{[m]} - E_j^{[m]} | \tilde{\psi}_j^{[m-1]} \rangle^{(m)} = 0 \tag{A6}$$

for all $i, j \in \mathcal{S}$: these $\tilde{\psi}_j^{[m-1]}$ can then become the starting point for an mth-order solution of (A3b).

After some algebra involving the fact that the analogs of (A6) at lower orders must have been satisfied, it is possible to convert (A6) to the form

$$\sum_{p\in\mathcal{S}'} \langle \psi_i^{(0)} | H^{[m]} - E_p^{[m-1]} | \psi_p^{[m-1]} \rangle^{(m)} (a_{pj}^{(0)} + \delta_{pj})$$

$$= \begin{cases} E_j^{(m)}(a_{ij}^{(0)} + \delta_{ij}) & \text{for } i \in \mathcal{S}', \quad \text{(A7a)} \\ (E_j^{(M_{ij}+1)} - E_i^{(M_{ij}+1)}) a_{ij}^{(m-M_{ij}-1)} & \text{for } i \notin \mathcal{S}', \end{cases}$$

(A7b)

$$\mathcal{S}' = \{p \mid p \in \mathcal{S}, M_{pj} = m - 1\}. \tag{A7c}$$

For each subset $\mathcal{S}'$ for which the $E_j^{[m-1]}$ are identical (some such subsets consisting of just one member), (A7a) is an eigenvalue equation for the matrix

$$\{\langle \psi_i^{(0)} | H^{[m]} - E_p^{[m-1]} | \psi_p^{[m-1]} \rangle^{(m)}\}.$$

Because $(H^{[m]} - E_{j \text{ or } p}^{[m-1]})\psi_p^{[m-1]}$ is already of mth order, one may add higher-order terms to $\psi_i^{(0)}$ without affecting the value of the matrix element; written in the form

$$\{\langle \psi_i^{[m-1]} | H^{[m]} - E_{i,j, \text{ or } p}^{[m-1]} | \psi_p^{[m-1]} \rangle^{(m)} \mid i \in \mathcal{S}'\}, \tag{A8}$$

the matrix is manifestly Hermitian and therefore possesses a complete set of eigenvectors.[13] After all the $E_j^{(m)}$ and $a_{pj}^{(0)}$ $(p \in \mathcal{S}')$ have been determined by (A7a), the $a_{ij}^{[m-M_{ij}-1]}$ can be chosen to satisfy (A7b); (A7a) remains satisfied since these coefficients do not appear in it.

Thus condition (ii) can always be met. If $E_k^{(0)}$ is not embedded in the continuous spectrum of $H^{(0)}$, the

Rayleigh–Schrödinger perturbation equations (A3) have solutions through as many orders l as the quantities $[(H^{[l]} - E_j^{[l-1]})\psi_j^{[l-1]}]^{(l)}$ are square integrable for all j such that $E_j^{(0)} = E_k^{(0)}$. According to arguments presented previously,[5] it may further be arranged that these solutions be orthonormal through as many orders as exist at all.

APPENDIX B: EXISTENCE OF THE ${}^k\psi^{[L]}$

As mentioned earlier, Eqs. (31) are Rayleigh–Schrödinger theory in $\mathfrak{h}(A^{(L)}, B^{(L)})$. The ${}^k\psi_n^{(0)}$ are eigenfunctions of $H_0^{(0)}$:

$$[H_0^{(0)} + n\omega - {}^k\mathcal{E}^{(0)}]\,{}^k\psi_n^{(0)} = 0, \quad \text{for } A^{(0)} \leq n \leq B^{(0)}. \tag{B1}$$

The set $\mathcal{S}$ of functions ${}^k\psi_n^{(0)} \exp(in\varphi)$ corresponding to solutions of (B1) defines a linear manifold $\mathcal{M}$ in $\mathfrak{h}(A^{(L)}, B^{(L)})$ which is degenerate w.r.t. the unperturbed operator $H_0^{(0)} - i\omega\partial_\varphi$. One may base a set of solutions ${}^k\psi^{[L]} \in \mathfrak{h}(A^{(L)}, B^{(L)})$ on $\mathcal{M}$ only if $\mathfrak{h}(A^{(L)}, B^{(L)})$ contains no other eigenfunctions of $H_0^{(0)} - i\omega\partial_\varphi$, that is, only if (B1) has no solutions with $A^{(L)} \leq n < A^{(0)}$ or $B^{(0)} < n \leq B^{(L)}$.

The rhs of the higher-order equation (31b) has finite norm if all its Fourier components are square integrable. The number ${}^k\mathcal{E}^{(0)}$ does not lie in the continuous spectrum of $H_0^{(0)} - i\omega\partial_\varphi$ in $\mathfrak{h}(A^{(L)}, B^{(L)})$ unless one of the numbers ${}^k\mathcal{E}^{(0)} - n\omega$, $A^{(L)} \leq n \leq B^{(L)}$, lies in the continuous spectrum of $H_0^{(0)}$. Thus the arguments in Appendix A guarantee the existence of solutions through as many orders as the Fourier components mentioned are square integrable if none of the numbers ${}^k\mathcal{E}^{(0)} - n\omega$, $A^{(L)} \leq n \leq B^{(L)}$, lies in the continuous spectrum of $H_0^{(0)}$; and there *may* yet be solutions if this second condition is violated.

In consequence of the particular form for $\mathcal{H}$ given in (15), it is easily seen that $A^{(L)}$ and $B^{(L)}$ may be taken to be

$$A^{(L)} = A^{(0)} - L, \quad B^{(L)} = B^{(0)} + L. \tag{B2}$$

The condition on square integrability of the rhs of (31b) can probably be taken for granted; thus, if one desires to construct perturbational quasiperiodic states on the basis of $H_0^{(0)}$-eigenfunctions with eigenvalues ${}^k\mathcal{E}^{(0)} - n\omega$ in a particular range, $A^{(0)} \leq n \leq B^{(0)}$ (none of these eigenvalues lying in the continuous spectrum of $H_0^{(0)}$), there is assurance that such ${}^k\psi^{[L]}$ exist if the spectrum of $H_0^{(0)}$ includes none of the numbers ${}^k\mathcal{E}^{(0)} - n\omega$ with $n = A^{(0)} - L, A^{(0)} - L + 1, \cdots, A^{(0)} - 1, B^{(0)} + 1, B^{(0)} + 2, \cdots, B^{(0)} + L$.

Again, if the ${}^k\psi^{[L]}$ exist, it may be arranged that they all be orthonormal through Lth order. In the present case, this orthonormality reads

$$\langle {}^j\psi^{[L]} \,|\, {}^k\psi^{[L]} \rangle_0 = \delta_{jk} + O(\lambda^{L+1}). \tag{B3}$$

APPENDIX C: ERROR BOUND FORMULA

Let $\Phi(t)$ be an exact solution to the Schrödinger equation and $\tilde{\Phi}(t)$ an approximate one. If $\tilde{\Phi}(0) = \Phi(0)$, it has previously been shown that[5]

$$\Phi(t) - \tilde{\Phi}(t) = -i\int_0^t U(t, t')[\mathcal{H}(t') - id_{t'}]\tilde{\Phi}(t')\,dt'. \tag{C1}$$

The derivation was rigorous only for the case of a finite-dimensional space. In the previous application of (C1),[5] the weaker formula

$$\|\Phi(t) - \tilde{\Phi}(t)\| \leq \int_0^t \|[\mathcal{H}(t') - id_{t'}]\tilde{\Phi}(t')\|\,dt' \tag{C2}$$

could have been used instead. This appendix offers a rigorous proof of a somewhat stronger version of (C2).

Theorem: Assume throughout the time span of interest, $[0, T]$, that $\mathcal{H}(t)$ is self-adjoint with a domain that includes $\Phi(t)$ and $\tilde{\Phi}(t)$ and that $d_t\Phi$ and $d_t\tilde{\Phi}$ are continuous[6] and satisfy

$$d_t\Phi = -i\mathcal{H}\Phi, \tag{C3a}$$

$$d_t\tilde{\Phi} = -i\mathcal{H}\tilde{\Phi} - i\chi, \tag{C3b}$$

with χ again a continuous function of time. [It is not required that $\Phi(0) = \tilde{\Phi}(0)$.] Then

$$\|\Phi(T) - \tilde{\Phi}(T)\| \leq \|\Phi(0) - \tilde{\Phi}(0)\| + \int_0^T \|[\mathcal{H}(t) - id_t]\tilde{\Phi}(t)\|\,dt. \tag{C4}$$

Proof: When $\Phi(t) \neq \tilde{\Phi}(t)$, application of the chain rule gives

$$d_t\,\|\Phi - \tilde{\Phi}\| = (2\,\|\Phi - \tilde{\Phi}\|)^{-1} d_t(\Phi - \tilde{\Phi} \,|\, \Phi - \tilde{\Phi}) \tag{C5}$$

$$= \|\Phi - \tilde{\Phi}\|^{-1}\,\mathrm{Im}\,(\chi \,|\, \Phi - \tilde{\Phi}), \tag{C6}$$

$$|d_t\,\|\Phi - \tilde{\Phi}\|\,| \leq \|\chi\|. \tag{C7}$$

If $\Phi \neq \tilde{\Phi}$ *throughout* $[0, T]$, then

$$|\,\|\Phi(T) - \tilde{\Phi}(T)\| - \|\Phi(0) - \tilde{\Phi}(0)\|\,| \leq \int_0^T \|\chi(t)\|\,dt, \tag{C8}$$

$$\|\Phi(T) - \tilde{\Phi}(T)\| \leq \|\Phi(0) - \tilde{\Phi}(0)\| + \int_0^T \|[\mathcal{H}(t) - id_t]\tilde{\Phi}(t)\|\,dt. \tag{C9}$$

If $\Phi(t) = \tilde{\Phi}(t)$ somewhere in the interval, (C6) breaks down there; even so, $\|\Phi - \tilde{\Phi}\|$ is still continuous there. Let t' be the largest t-value at which $\Phi = \tilde{\Phi}$. Then it is easily shown that

$$\|\Phi(T) - \tilde{\Phi}(T)\| \leq \int^{T} \|\chi(t)\|\, dt \leq \|\Phi(0) - \tilde{\Phi}(0)\| + \int_0^T \|[\mathcal{K}(t) - id_t]\tilde{\Phi}(t)\|\, dt. \quad \text{(C10)}$$

An identical result holds in unitary space.

* Acknowledgment is made to the donors of the Petroleum Research Fund, administered by the American Chemical Society, for partial support of this research. This research was sponsored in part by the National Science Foundation under Grant No. GP-6346. It has benefited from facilities made available by the Advanced Research Projects Agency through the Center for Materials Research at Stanford University, where this work was begun.

[1] P. Nozières and D. Pines, Nuovo Cimento **9**, 470 (1958); see also M. Born and P. Jordan, *Elementare Quantenmechanik* (Springer-Verlag, Berlin, 1930), pp. 236ff.
[2] For a critique of the conventional time-dependent perturbation theory, see J. H. Shirley, Ph.D. thesis, California Institute of Technology, 1963.
[3] R. H. Young, W. J. Deal, Jr., and N. R. Kestner, Mol. Phys. **17**, 369 (1969).
[4] J. H. Shirley, Phys. Rev. **138**, B979 (1965).
[5] R. H. Young and W. J. Deal, Jr., Phys. Rev. A **1**, 419 (1970).
[6] Here convergence in the mean is used in the definitions of continuity and differentiation.
[7] See, for instance, M. H. Stone, *Linear Transformations in Hilbert Space* (American Mathematical Society, New York, 1966), p. 30.
[8] Reference 7, p. 128.
[9] S. Lefschetz, *Differential Equations: Geometric Theory* (Interscience, New York, 1963), p. 44.
[10] T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1966), pp. 71, 99–102; for a simpler treatment which can be adapted, see F. Rellich, Math. Ann. (Leipzig) **113**, 600 (1939).
[11] Ralph H. Young, Ph.D. thesis, Stanford University, 1968.
[12] Sufficient conditions for the existence *and convergence* of the Rayleigh-Schrödinger theory are well known. See, for instance, T. Kato, J. Fac. Sci., Univ. Tokyo, Sect. I, **6**, 145 (1949).
[13] E. Wigner, *Group Theory* (Academic, New York, 1959), p. 26.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Erratum: Lattice Dynamics of Cubic Lattices with Long-Range Interactions

[J. Math. Phys. **11**, 1513 (1970)]

JOHN A. DAVIES
Department of Physics, Clark University, Worcester, Massachusetts 01610

(Received 15 June 1970)

Equation (19) is valid only under the condition that the short-range forces extend to nearest neighbors only or that, if they extend beyond nearest neighbors, the short-range forces between like particles are independent of k. The above condition is not mentioned in the paper. The validity of Eq. (22) and of the second-order terms in Eqs. (30), (34), and (36) are subject to this condition. The form of Eq. (20) does not depend upon this condition. None of the other results in the paper are subject to this condition except those in Sec. VI, where the condition is explicitly stated.

Equations (22), (30), (34), and (36) are easily modified to become valid for short-range forces arising from central potentials not subject to the above condition. In the lower right submatrix in Eq. (22), replace the constants b, c, and d by b', c', and d', respectively. The long-range contributions to corresponding primed and unprimed constants are the same, but the short-range contributions may differ. In Eqs. (30) and (36) replace $g(b, \beta)$ with $g(b, b', \beta)$ and $g(c + d, \gamma + \delta)$ with $g(c + d, c' + d', \gamma + \delta)$, where $g(x, y, z) = x + \mu^2 y + 2\mu z$. Corresponding replacements are to be made in Eq. (34).

If $\Phi(t) = \tilde{\Phi}(t)$ somewhere in the interval, (C6) breaks down there; even so, $\|\Phi - \tilde{\Phi}\|$ is still continuous there. Let t' be the largest t-value at which $\Phi = \tilde{\Phi}$. Then it is easily shown that

$$\|\Phi(T) - \tilde{\Phi}(T)\| \leq \int^{T} \|\chi(t)\|\, dt \leq \|\Phi(0) - \tilde{\Phi}(0)\| + \int_0^T \|[\mathcal{K}(t) - id_t]\tilde{\Phi}(t)\|\, dt. \quad \text{(C10)}$$

An identical result holds in unitary space.

* Acknowledgment is made to the donors of the Petroleum Research Fund, administered by the American Chemical Society, for partial support of this research. This research was sponsored in part by the National Science Foundation under Grant No. GP-6346. It has benefited from facilities made available by the Advanced Research Projects Agency through the Center for Materials Research at Stanford University, where this work was begun.

[1] P. Nozières and D. Pines, Nuovo Cimento **9**, 470 (1958); see also M. Born and P. Jordan, *Elementare Quantenmechanik* (Springer-Verlag, Berlin, 1930), pp. 236ff.
[2] For a critique of the conventional time-dependent perturbation theory, see J. H. Shirley, Ph.D. thesis, California Institute of Technology, 1963.
[3] R. H. Young, W. J. Deal, Jr., and N. R. Kestner, Mol. Phys. **17**, 369 (1969).
[4] J. H. Shirley, Phys. Rev. **138**, B979 (1965).
[5] R. H. Young and W. J. Deal, Jr., Phys. Rev. A **1**, 419 (1970).
[6] Here convergence in the mean is used in the definitions of continuity and differentiation.
[7] See, for instance, M. H. Stone, *Linear Transformations in Hilbert Space* (American Mathematical Society, New York, 1966), p. 30.
[8] Reference 7, p. 128.
[9] S. Lefschetz, *Differential Equations: Geometric Theory* (Interscience, New York, 1963), p. 44.
[10] T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1966), pp. 71, 99–102; for a simpler treatment which can be adapted, see F. Rellich, Math. Ann. (Leipzig) **113**, 600 (1939).
[11] Ralph H. Young, Ph.D. thesis, Stanford University, 1968.
[12] Sufficient conditions for the existence *and convergence* of the Rayleigh-Schrödinger theory are well known. See, for instance, T. Kato, J. Fac. Sci., Univ. Tokyo, Sect. I, **6**, 145 (1949).
[13] E. Wigner, *Group Theory* (Academic, New York, 1959), p. 26.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 11 NOVEMBER 1970

Erratum: Lattice Dynamics of Cubic Lattices with Long-Range Interactions

[J. Math. Phys. **11**, 1513 (1970)]

JOHN A. DAVIES
Department of Physics, Clark University, Worcester, Massachusetts 01610

(Received 15 June 1970)

Equation (19) is valid only under the condition that the short-range forces extend to nearest neighbors only or that, if they extend beyond nearest neighbors, the short-range forces between like particles are independent of k. The above condition is not mentioned in the paper. The validity of Eq. (22) and of the second-order terms in Eqs. (30), (34), and (36) are subject to this condition. The form of Eq. (20) does not depend upon this condition. None of the other results in the paper are subject to this condition except those in Sec. VI, where the condition is explicitly stated.

Equations (22), (30), (34), and (36) are easily modified to become valid for short-range forces arising from central potentials not subject to the above condition. In the lower right submatrix in Eq. (22), replace the constants b, c, and d by b', c', and d', respectively. The long-range contributions to corresponding primed and unprimed constants are the same, but the short-range contributions may differ. In Eqs. (30) and (36) replace $g(b, \beta)$ with $g(b, b', \beta)$ and $g(c + d, \gamma + \delta)$ with $g(c + d, c' + d', \gamma + \delta)$, where $g(x, y, z) = x + \mu^2 y + 2\mu z$. Corresponding replacements are to be made in Eq. (34).