JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 13, NUMBER 6 JUNE 1972

N-Particle Noninteracting Green's Function*

W. Grant Cooper†
Department of Biophysical Sciences, University of Houston, Houston, Texas 77004

and

Donald J. Kouri
Department of Chemistry, University of Houston, Houston, Texas 77004
(Received 26 February 1971)

A prescription is given for obtaining the Green's function for N free particles which can have different masses. The approach is systematic and straightforward. A coordinate transformation of the Fourier integral representation of the N-particle noninteracting Green's function facilitates the integration over $3N$-1 angular variables of wavenumber space. A single radial integral can then be evaluated. The resulting Green's function representation may be of use in applying the integral form of Schrödinger's equation to calculate the ground and excited states of atoms.

For certain problems in quantum mechanics, it is advantageous to reduce the number of continuous variables, replacing them with sums over discrete variables. Many investigators[1-6] have utilized some form of hyperspherical coordinates to express a $3N$-dimensional problem, $N = 2, 3, 4, \ldots$, in terms of a single continuous radial variable with the remaining $3N$-1 coordinates being angles. Functions of the $3N$ coordinates may then be expanded in terms of a set of "generalized" spherical harmonics labeled by discrete indices. If N is 1, ordinary spherical polar coordinates result.

In this paper, a well-known hyperspherical coordinate system[7,8] is employed to obtain the noninteracting Green's function for $N + 1$ distinguishable particles. Although other forms of the many-particle Green's functions are available,[9,10] the representation developed here permits one to employ the integral solution precedure of Sams and Kouri,[11] which is a numerical method for noniteratively solving integral equations. It has been used successfully in carrying out numerous scattering calculations[12] and bound state investigations for model potentials such as the Lennard-Jones.[12,6] Since such integral equations are constructed using Green's functions, their numerical method can be employed to calculate energy eigenstates of systems composed of several particles if the relevant Green's functions are known. The present representation for the N-particle Green's function is computationally convenient for solving the integral form of Schrödinger's equation for atoms,[13] and it therefore may be of some benefit in analyzing other three- or more-body problems.[14]

In center of mass coordinates, r_1 is the vector between particles 1 and 2 with masses m_1 and m_2, respectively, r_2 is the vector from their center of mass to a third particle, and r_3 is a vector from the center of mass of all three to a fourth particle. The ith vector r_i, $i = 1, 2, \ldots, N$, is from the center of mass of all the particles 1, 2, 3, ..., i to the center of mass particle $i + 1$. The ith reduced mass, μ_i, is

$$\mu_i = m_{i+1}\left(\sum_{j=1}^{i} m_j\right)\Big/\sum_{j=1}^{i+1} m_j, \tag{1}$$

and η_i is defined by

$$\eta_i^2 = \mu_i. \tag{2}$$

The Green's function $G(\mathfrak{R}|\mathfrak{R}')$ for N free particles having reduced masses η_i^2 satisfies the $3N$-dimensional inhomogeneous differential equation

$$\left(\sum_{i=1}^{N}\left(\frac{1}{\eta_i}\right)^2\nabla_i^2 + K^2\right) G(\mathfrak{R}|\mathfrak{R}') = -\delta(\mathbf{r}_1 - \mathbf{r}_1')\delta(\mathbf{r}_2 - \mathbf{r}_2')\ldots\delta(\mathbf{r}_N - \mathbf{r}_N'), \tag{3}$$

where the vector $\mathfrak{R}$ possesses $3N$ components and the right side of Eq. (3) is a product of N three-dimensional Dirac δ functions. The Fourier integral representation of $G(\mathfrak{R}|\mathfrak{R}')$ may be written as

$$G(\mathfrak{R}|\mathfrak{R}') = \left(\frac{1}{2\pi}\right)^{3N}\int d\mathbf{u}_1\int d\mathbf{u}_2\cdots\int d\mathbf{u}_N \times\left(\frac{e^{i\mathbf{u}_1\cdot(\mathbf{r}_1-\mathbf{r}_1')}e^{i\mathbf{u}_2\cdot(\mathbf{r}_2-\mathbf{r}_2')}\cdots e^{i\mathbf{u}_N\cdot(\mathbf{r}_N-\mathbf{r}_N')}}{(u_1/\eta_1)^2 + (u_2/\eta_2)^2 + \cdots + (u_N/\eta_N)^2 - K^2}\right), \tag{4}$$

using the Fourier representation of the Dirac δ functions.[15] The method for treating the singulari-

ties in Eq. (4) follow from the boundary conditions imposed on $G(\mathfrak{R})$ as $\mathfrak{R} \to \infty$. It is convenient to let

$$\mathbf{g}_i = \mathbf{u}_i/\eta_i \tag{5}$$

and

$$\mathbf{R}_i = \eta_i \mathbf{r}_i \tag{6}$$

in order to obtain the equation

$$G(\mathfrak{R}|\mathfrak{R}') = \left(\frac{1}{2\pi}\right)^{3N} \int \eta_1 d\mathbf{g}_1 \int \eta_2 d\mathbf{g}_2 \cdots \int \eta_N d\mathbf{g}_N \times \left(\frac{e^{i\mathbf{g}_1 \cdot (\mathbf{R}_1 - \mathbf{R}_1')} e^{i\mathbf{g}_2 \cdot (\mathbf{R}_2 - \mathbf{R}_2')} \cdots e^{i\mathbf{g}_N \cdot (\mathbf{R}_N - \mathbf{R}_N')}}{g_1^2 + g_2^2 + \cdots + g_N^2 - K^2} \right). \tag{7}$$

By employing a partial wave decomposition[16] and integrating Eq. (7) over the $2N$ spherical polar angles $\vartheta_{g_1}, \ldots, \vartheta_{g_N}, \varphi_{g_1}, \ldots, \varphi_{g_N}$, the expression for $G(\mathfrak{R}|\mathfrak{R}')$ becomes

$$G(\mathfrak{R}|\mathfrak{R}') = \left(\frac{2}{\pi}\right)^N \int_0^\infty \eta_1^3 g_1^2 dg_1 \int_0^\infty \eta_2^3 g_2^2 dg_2 \cdots \int_0^\infty \eta_N^3 g_N dg_N \sum_{l_1=0}^{\infty} \sum_{l_2=0}^{\infty} \cdots \sum_{l_N=0}^{\infty} \sum_{m_1=-l_1}^{l_1} \sum_{m_2=-l_2}^{l_2} \cdots \sum_{m_N=-l_N}^{l_N}$$
$$\left\{ \left(\frac{j_{l_1}(g_1 R_1) j_{l_1}(g_1 R_1') j_{l_2}(g_2 R_2) j_{l_2}(g_2 R_2') \cdots j_{l_N}(g_N R_N) j_{l_N}(g_N R_N')}{g_1^2 + g_2^2 + \cdots + g_N^2 - K^2} \right) \right.$$
$$\left. \times Y_{l_1}^{m_1}(\hat{R}_1) Y_{l_1}^{m_1 *}(\hat{R}_1') Y_{l_2}^{m_2}(\hat{R}_2) Y_{l_2}^{m_2 *}(\hat{R}_2') \cdots Y_{l_N}^{m_N}(\hat{R}_N) Y_{l_N}^{m_N *}(\hat{R}_N') \right\}, \tag{8}$$

where the orthonormality condition

$$\prod_{i=1}^{N} \int_0^{2\pi} d\varphi_{g_i} \int_0^{\pi} d\vartheta_{g_i} \sin\vartheta_{g_i} Y_{l_i}^{m_i}(\hat{g}_i) Y_{l_i'}^{m_i' *}(\hat{g}_i) = \prod_{i=1}^{N} \delta_{l_i l_i'} \delta_{m_i m_i'} \tag{9}$$

has been employed in performing the integration.

The orbital angular momentum of the ith particle, $i = 1, 2, \ldots, N$, is designated l_i and the z component is m_i; so the spherical Bessel functions written in the form $j_{l_i}(g_i R_i)$ represent "single-particle functions." The l_ith spherical Bessel function $j_{l_i}(g_i R_i)$ is defined in terms of the cylindrical Bessel function $J_{l_i+1/2}(g_i R_i)$ by the equation

$$j_{l_i}(g_i R_i) = (\pi/2 g_i R_i)^{1/2} J_{l_i+1/2}(g_i R_i) \tag{10}$$

and the $Y_{l_i}^{m_i}(\hat{R}_i)$ $(m_i = -l_i, -l_i + 1, \ldots, l_i)$ are ordinary spherical harmonics defined by Edmonds.[17]

After Eq. (8) is expressed in hyperspherical coordinates, the many-particle representation of the Green's function will be explicitly given by executing the remaining integrals, which are tedious but not difficult. The hyperspherical coordinates include $4N$ spherical polar angles $\theta_i, \theta_i', \phi_i, \phi_i'$ and hyperspherical distances ρ and ρ' defined by

$$\rho^2 = \sum_{i=1}^{N} (r_i)^2 \tag{11}$$

and

$$\rho'^2 = \sum_{i=1}^{N} (r_i')^2 \tag{12}$$

where r_i is the ith radial spherical polar coordinate for the observer; primes denote source coordinates. In addition there are $2N-2$ hyperspherical angles α_j and α_j', $j = 1, 2 \ldots, N-1$, given by the relations

$$\begin{aligned} r_1 &= \rho \cos\alpha_1, \\ r_2 &= \rho \sin\alpha_1 \cos\alpha_2, \\ &\cdots, \\ r_{N-1} &= \rho \sin\alpha_1 \sin\alpha_2 \cdots \cos\alpha_{N-1}, \\ r_N &= \rho \sin\alpha_1 \sin\alpha_2 \cdots \sin\alpha_{N-1}, \quad 0 \le \alpha_j \le \pi/2. \end{aligned} \tag{13}$$

and

$$\begin{aligned} r_1' &= \rho' \cos\alpha_1', \\ r_2' &= \rho' \sin\alpha_1' \cos\alpha_2', \\ &\cdots, \\ r_{N-1}' &= \rho' \sin\alpha_1' \sin\alpha_2' \cdots \cos\alpha_{N-1}', \\ r_N' &= \rho' \sin\alpha_1' \sin\alpha_2' \cdots \sin\alpha_{N-1}', \quad 0 \le \alpha_j' \le \pi/2. \end{aligned} \tag{14}$$

Hyperspherical coordinates of wavenumber space have radius Q, given by

$$Q^2 = \sum_{i=1}^{N} (g_i)^2, \tag{15}$$

and the hyperspherical angles β_j, $j = 1, 2, \ldots, N-1$, are introduced by the relations

$$\begin{aligned} g_1 &= Q \cos\beta_1, \\ g_2 &= Q \sin\beta_1 \cos\beta_2, \\ &\cdots, \\ g_{N-1} &= Q \sin\beta_1 \sin\beta_2 \cdots \cos\beta_{N-1}, \\ g_N &= Q \sin\beta_1 \sin\beta_2 \cdots \sin\beta_{N-1}, \quad 0 \le \beta_j \le \pi/2. \end{aligned} \tag{16}$$

The other $2N$ coordinates, ϑ_{g_i} and φ_{g_i}, are the spherical polar angles over which integration was carried out to obtain Eq. (8).

With this choice of coordinates, the products of spherical Bessel functions $\prod_{i=1}^{N} j_{l_i}(g_i R_i)$ and $\prod_{i=1}^{N} j_{l_i}(g_i R_i')$ in Eq. (8) can be separately coupled, using the product expansion relation[18]:

$$J_\nu(z \sin\Theta \sin\Phi) J_\mu(z \cos\Theta \cos\Phi) = 2z^{-1} \sum_{\lambda=0}^{\infty} J_{\nu+\mu+2\lambda+1}(z) \cos^\mu\Theta \sin^\nu\Phi \times {}_2F_1(-\lambda, \nu+\mu+\lambda+1; \nu+1; \sin^2\Theta) \times {}_2F_1(-\lambda, \nu+\mu+\lambda+1; \nu+1; \sin^2\Phi)$$

$$\times \cos^{\mu}\Phi \sin^{\nu}\Theta (-1)^{\lambda}(\nu+\mu+2\lambda+1) \times \left(\frac{\Gamma(\nu+\mu+\lambda+1)\Gamma(\nu+\lambda+1)}{\lambda![\Gamma(\nu+1)]^2\Gamma(\mu+\lambda+1)}\right). \quad (17)$$

The orders ν, μ of the Bessel functions can assume any value except negative integers, and λ is zero or a positive integer. Thus, ${}_2F_1$ is the Jacobi polynomial of $\lambda+1$ terms, and Γ is the usual gamma function. Equation (8) may be represented in hyperspherical coordinates exclusively after $2(N-1)$ coupling operations are performed; however, a definite pattern is followed to implement this transformation. By using Eq. (17), the functions $j_{l_N}(g_N R_N)$ and $j_{l_N}(g_{N-1}R_{N-1})$ are initially coupled to yield a two-particle "cluster function," where the hyperspherical variable ρ is a parameter. Likewise the functions $j_{l_N}(g_N R'_N)$ and $j_{l_{N-1}}(g_{N-1}R'_{N-1})$ are coupled producing the two-particle "cluster function" with ρ' entering as a parameter, i.e.,

$$j_{l_N}(g_N R_N)\, j_{l_{N-1}}(g_{N-1}R_{N-1}) = \left[\left(\frac{\pi}{2g_N R_N}\right)\left(\frac{\pi}{2g_{N-1}R_{N-1}}\right)\right]^{1/2} \times J_{l_N+1/2}(g_N R_N)\, J_{l_{N-1}+1/2}(g_{N-1}R_{N-1})$$
$$= \frac{\pi}{2Q\rho}\left(\frac{1}{\sin\beta_1 \dots \sin\beta_{N-2}\sin\alpha_1 \dots \sin\alpha_{N-2}}\right) \times \sum_{\lambda_1=0}^{\infty}\left(\frac{1}{Q\rho}\right) J_{l_N+l_{N-1}+2\lambda_1+2}(Q\rho \sin\beta_1 \dots \sin\beta_{N-2}$$
$$\times \sin\alpha_1 \dots \sin\alpha_{N-2}) \cos^{l_{N-1}}\alpha_{N-1} \sin^{l_N}\alpha_{N-1} \times F(-\lambda_1, l_N + l_{N-1} + \lambda_1 + 2; l_N + \tfrac{3}{2}; \sin^2\alpha_{N-1}) \times \cos^{l_{N-1}}\beta_{N-1} \sin^{l_N}\beta_{N-1} \times F(-\lambda_1, l_N + l_{N-1} + \lambda_1 + 2; l_N + \tfrac{3}{2}; \sin^2\beta_{N-1}) \times (-1)^{\lambda_1}\mathfrak{N}_1^{-2} \quad (18)$$

and

$$j_{l_N}(g_N R'_N)\, j_{l_{N-1}}(g_{N-1}R'_{N-1}) = \left[\left(\frac{\pi}{2g_N R'_N}\right)\left(\frac{\pi}{2g_{N-1}R'_{N-1}}\right)\right]^{1/2} \times J_{l_N+1/2}(g_N R'_N)\, J_{l_{N-1}+1/2}(g_{N-1}R'_{N-1})$$
$$= \frac{\pi}{2Q\rho'}\left(\frac{1}{\sin\beta_1 \dots \sin\beta_{N-2}\sin\alpha'_1 \dots \sin\alpha'_{N-2}}\right) \times \sum_{\lambda'_1=0}^{\infty}\left(\frac{1}{Q\rho'}\right) J_{l_N+l_{N-1}+2\lambda'_1+2} \times (Q\rho' \sin\beta_1 \dots \sin\beta_{N-2}\sin\alpha'_1 \dots \sin\alpha'_{N-2}) \times \cos^{l_{N-1}}\alpha'_{N-1} \sin^{l_N}\alpha'_{N-1} \times F(-\lambda'_1, l_N + l_{N-1} + \lambda'_1 + 2; l_N + \tfrac{3}{2}; \sin^2\alpha'_{N-1}) \times \cos^{l_{N-1}}\beta_{N-1} \sin^{l_N}\beta_{N-1}) \times F(-\lambda_1, l_N + l_{N-1} + \lambda'_1 + 2; l_N + \tfrac{3}{2}; \sin^2\beta_{N-1}) \times (-1)^{\lambda'_1}(\mathfrak{N}'_1)^{-2}, \quad (19)$$

where

$$\mathfrak{N}_1^2 = \frac{\lambda_1![\Gamma(l_N+\frac{3}{2})]^2\Gamma(l_{N-1}+\lambda_1+\frac{3}{2})}{2(l_N+l_{N-1}+2\lambda_1+2)\Gamma(l_N+l_{N-1}+\lambda_1+2)\Gamma(l_N+\lambda_1+\frac{3}{2})} \quad (20)$$

and

$$\mathfrak{N}'^2_1 = \frac{\lambda'_1![\Gamma(l_N+\frac{3}{2})]^2\Gamma(l_{N-1}+\lambda'_1+\frac{3}{2})}{2(l_N+l_{N-1}+2\lambda'_1+2)\Gamma(l_N+l_{N-1}+\lambda'_1+2)\Gamma(l_N+\lambda'_1+\frac{3}{2})}. \quad (21)$$

Thus the two-particle "cluster functions" are shown in Eqs. (18) and (19).

The second step in transforming the integrand of Eq. (8) entails coupling the single-particle functions $j_{l_{N-2}}(g_{N-2}R_{N-2})$ and $j_{l_{N-2}}(g_{N-2}R'_{N-2})$ to the two-particle "cluster functions" of Eqs. (18) and (19), respectively, using the product expansion relation of Eq. (17). To these expanded results, the functions $j_{l_{N-3}}(g_{N-3}R_{N-3})$ and $j_{l_{N-3}}(g_{N-3}R'_{N-3})$ are, respectively, coupled which completes the third transformation step. Finally, the N-particle "cluster function" is formed by coupling the remaining single-particle function, j_{l_1}, to the $(N-1)$-particle "cluster function."

However, the integral representation for the noninteracting N-particle Green's function $G(\mathfrak{R}|\mathfrak{R}')$ can be generated from a generalized expression for the product expansion operation that will couple M particles, $M = 2, 3, 4, \dots, N$, within a hypersphere of $3M$ dimensions and will thus yield an M-particle "cluster function." The $(M-1)$th product expansion operation is now considered where an M_a-particle function is coupled to an M_b-particle product expansion, and it follows that $M = M_a + M_b$. (Note that M_a is equal to one in the present treatment.) Via this coupling prescription with the hyperspherical coordinates introduced, the $(M-1)$th expansion operation couples the $(N-M+1)$th single-particle expression to the $(M-2)$th M_b-particle expression. By identifying the relations

$$\tilde{Q}\tilde{\rho}\cos\beta_{N-M+1}\cos\alpha_{N-M+1} = Q\rho \sin\beta_1 \dots \sin\beta_{N-M} \times \cos\beta_{N-M+1}\sin\alpha_1 \dots \sin\alpha_{N-M}\cos\alpha_{N-M+1}, \quad (22)$$

$$\tilde{Q}\tilde{\rho}\sin\alpha_{N-M+1}\sin\beta_{N-M+1} = Q\rho \sin\beta_1 \dots \sin\beta_{N-M} \times \sin\beta_{N-M+1}\sin\alpha_1 \dots \sin\alpha_{N-M}\sin\alpha_{N-M+1}, \quad (23)$$

and

$$\sigma = N - M + 1, \quad (24)$$

then the $(M-1)$th product expansion in the sequence of $(N-1)$ expansions may be expressed as

$$\left(\frac{\pi}{2\tilde{Q}\tilde{\rho}\cos\beta_\sigma\cos\alpha_\sigma}\right)^{M_a/2} J_{l_\sigma+1/2}(\tilde{Q}\tilde{\rho}\cos\beta_\sigma\cos\alpha_\sigma) \times \left(\frac{\pi}{2}\right)^{M_b/2}\left(\frac{1}{\tilde{Q}\tilde{\rho}\sin\beta_\sigma\sin\alpha_\sigma}\right)^{(3M_b/2-1)}$$

$$\times J_{\nu_{M-2}}(\tilde{Q}\tilde{\rho}\sin\beta_\sigma \sin\alpha_\sigma)$$
$$= \left(\frac{\pi}{2}\right)^{M/2} \sum_{\aleph_{M-1}} \left[\left(\frac{1}{\tilde{Q}\tilde{\rho}}\right)^{(3M/2-1)} J_{\nu_{M-1}}(\tilde{Q}\tilde{\rho}) \cos^{l_\sigma}\alpha_\sigma \right.$$
$$\times \sin^{\{\nu_{M-2}-[(3M_b/2)-1]\}}\alpha_\sigma$$
$$\times F(-\aleph_{M-1}, \nu_{M-1}-\aleph_{M-1}; \nu_{M-2}+1; \sin^2\alpha_\sigma)$$
$$\times \cos^{l_\sigma}\beta_\sigma \sin^{\{\nu_{M-2}-[(3M_b/2)-1]\}}\beta_\sigma$$
$$\times F(-\aleph_{M-1}, \nu_{M-1}-\aleph_{M-1}; \nu_{M-2}+1; \sin^2\beta_\sigma)$$
$$\left.\times (-1)^{\aleph_{M-1}} (\mathfrak{N}_{M-1})^{-2}\right]$$
$$= \left(\frac{\pi}{2}\right)^{M/2} \left(\frac{1}{\tilde{Q}\tilde{\rho}}\right)^{(3M/2-1)}$$
$$\times \sum_{\aleph_{M-1}=0}^{\infty} \left[J_{\nu_{M-1}}(\tilde{Q}\tilde{\rho}) \cos^{\lambda a(\sigma)}\alpha_\sigma \sin^{\lambda b(M-2)}\alpha_\sigma \right.$$
$$\times \cos^{\lambda a(\sigma)}\beta_\sigma \sin^{\lambda b(M-2)}\beta_\sigma$$
$$\times F(-\aleph_{M-1}, \nu_{M-1}-\aleph_{M-1}; \nu_{M-2}+1; \sin^2\alpha_\sigma)$$
$$\times F(-\aleph_{M-1}, \nu_{M-1}-\aleph_{M-1}; \nu_{M-2}+1; \sin^2\beta_\sigma)$$
$$\left.\times (-1)^{\aleph_{M-1}} (\mathfrak{N}_{M-1})^{-2}, \right. \tag{25}$$

where

$$(\mathfrak{N}_{M-1})^2 = \frac{(\aleph_{M-1})![\Gamma(\nu_{M-2}+1)]^2\Gamma(\aleph_{M-1}+\lambda_a(\sigma)+3M_a/2)}{2(\nu_{M-1})\Gamma(\nu_{M-1}-\aleph_{M-1})\Gamma(\aleph_{M-1}+\lambda_{b(M-2)}+3M_b/2)}. \tag{26}$$

The M-particle "cluster functions" constructed in Eq. (25) by the $(M-1)$-product expansion operation contains a cylindrical Bessel function order ν_{M-1}, which is given by

$$\nu_{M-1} = \sum_{i=0}^{M-1} l_{N-i} + 2\sum_{i=1}^{M-1}\aleph_i + \frac{3M}{2} - 1, \tag{27}$$

where the l_{N-i} are individual particle orbital angular momentum quantum numbers and the $\aleph_i$ are defined in Eq. (17). The order ν_{M-2} of the cylindrical function belonging to the M_b-particle "cluster function" is obtained from the relation

$$\nu_{M-2} = \sum_{i=0}^{M-2} l_{N-i} + 2\sum_{i=1}^{M-2}\aleph_i + \frac{3M_b}{2} - 1. \tag{28}$$

It is also apparent that the powers $\lambda a(\sigma)$ of the cosine functions produced by the $(M-1)$th product expansion are numbers related to the number of particles M_a; i.e., the $\lambda a(\sigma)$ are quantum numbers given by

$$\lambda a(\sigma) = l_{N-M+1}. \tag{29}$$

Similarly the powers $\lambda b(M-2)$ of the sine functions are connected with the number of particles M_b included by the $(M-2)$th product expansion.

The relevant relation here is

$$\lambda b(M-2) = \sum_{i=0}^{M-2} l_{N-i} + 2\sum_{i=1}^{M-2}\aleph_i. \tag{30}$$

It also may be seen that

$$\nu_{M-2} = \lambda b(M-2) + (3M_b/2) - 1 \tag{31}$$

and

$$\nu_{M-1} = 2\aleph_{M-1} + \lambda a(\sigma) + \lambda b(M-2) + (3M/2) - 1. \tag{32}$$

Now hyperspherical coordinates and Eqs. (27)–(30) are used in order to write the N-particle Green's function integral of Eq. (8) as

$$G(\mathfrak{R}|\mathfrak{R}') = \prod_{i=1}^{N}(\eta_i)^3 \left(\frac{1}{\rho\rho'}\right)^{(3N/2-1)} \int_0^\infty \frac{Q J_{\nu_{M-1}}(Q\rho) J_{\nu_{M-1}}(Q\rho')\,dQ}{Q^2-K^2}$$
$$\times \int_0^{\pi/2} \sin^2\beta_{N-1}\cos^2\beta_{N-1}\,d\beta_{N-1} \int_0^{\pi/2} \sin^5\beta_{N-2}\cos^2\beta_{N-2}$$
$$\times d\beta_{N-2}\cdots \int_0^{\pi/2} \sin^{3(N-1)-1}\beta_1 \cos^2\beta_1\,d\beta_1$$
$$\times \sum_{l_N}\sum_{l_{N-1}}\cdots\sum_{l_1}\sum_{m_N}\sum_{m_{N-1}}\cdots\sum_{m_1}\sum_{\aleph_1=0}^{\infty}\sum_{\aleph_2=0}^{\infty}\cdots$$
$$\times \sum_{\aleph_{N-1}=0}^{\infty}\sum_{\aleph'_1=0}^{\infty}\sum_{\aleph'_2=0}^{\infty}\cdots\sum_{\aleph'_{N-1}=0}^{\infty}$$
$$\times \left(\frac{(-1)^{\aleph_1}}{\mathfrak{N}_1^2}\right) \cos^{l_{N-1}}\alpha_{N-1} \sin^{\lambda b(0)}\alpha_{N-1}$$
$$\times F(-\aleph_1, \nu_1-\aleph_1; \nu_0+1; \sin^2\alpha_{N-1})$$
$$\times \cos^{l_{N-1}}\beta_{N-1} \sin^{\lambda b(0)}\beta_{N-1}$$
$$\times F(-\aleph_1, \nu_1-\aleph_1; \nu_0+1; \sin^2\beta_{N-1})$$
$$\times \left(\frac{(-1)^{\aleph'_1}}{\mathfrak{N}'^2_1}\right) \cos^{l_{N-1}}\alpha'_{N-1} \sin^{\lambda b(0)}\alpha'_{N-1}$$
$$\times F(-\aleph'_1, \nu_1-\aleph'_1; \nu_0+1; \sin^2\alpha'_{N-1})$$
$$\times \cos^{l_{N-1}}\beta_{N-1} \sin^{\lambda b(0)}\beta_{N-1}$$
$$\times F(-\aleph'_1, \nu_1-\aleph'_1; \nu_0+1; \sin^2\beta_{N-1})$$
$$\times \left(\frac{(-1)^{\aleph_2}}{\mathfrak{N}_2^2}\right) \cos^{l_{N-2}}\alpha_{N-2} \sin^{\lambda b(1)}\alpha_{N-2}$$
$$\times F(-\aleph_2, \nu_2-\aleph_2; \nu_1+1; \sin^2\alpha_{N-2})$$
$$\times \cos^{l_{N-2}}\beta_{N-2} \sin^{\lambda b(1)}\beta_{N-2}$$
$$\times F(-\aleph_2, \nu_2-\aleph_2; \nu_1+1; \sin^2\beta_{N-2})$$
$$\times \left(\frac{(-1)^{\aleph'_2}}{\mathfrak{N}'^2_2}\right) \cos^{l_{N-2}}\alpha'_{N-2} \sin^{\lambda b(1)}\alpha'_{N-2}$$
$$\times F(-\aleph'_2, \nu_2-\aleph'_2; \nu_1+1; \sin^2\alpha'_{N-2})$$
$$\times \cos^{l_{N-2}}\beta_{N-2} \sin^{\lambda b(1)}\beta_{N-2}$$
$$\times F(-\aleph'_2, \nu_2-\aleph'_2; \nu_1+1; \sin^2\beta_{N-2})$$
$$\cdots$$
$$\times \left(\frac{(-1)^{\aleph_{N-1}}}{\mathfrak{N}_{N-1}^2}\right) \cos^{l_1}\alpha_1 \sin^{\lambda b(N-2)}\alpha_1$$
$$\times F(-\aleph_{N-1}, \nu_{N-1}-\aleph_{N-1}; \nu_{N-2}+1; \sin^2\alpha_1)$$
$$\times \cos^{l_1}\beta_1 \sin^{\lambda b(N-2)}\beta_1$$
$$\times F(-\aleph_{N-1}, \nu_{N-1}-\aleph_{N-1}; \nu_{N-2}+1; \sin^2\beta_1)$$
$$\times \left(\frac{(-1)^{\aleph'_{N-1}}}{\mathfrak{N}'^2_{N-1}}\right) \cos^{l_1}\alpha'_1 \sin^{\lambda b(N-2)}\alpha'_1$$
$$\times F(-\aleph'_{N-1}, \nu_{N-1}-\aleph'_{N-1}; \nu_{N-2}+1; \sin^2\alpha'_1)$$
$$\times \cos^{l_1}\beta_1 \sin^{\lambda b(N-2)}\beta_1$$
$$\times F(-\aleph'_{N-1}, \nu_{N-1}-\aleph'_{N-1}; \nu_{M-2}+1; \sin^2\beta_1)$$
$$\times Y_{l_N}^{m_N}(\hat{R}_N) Y_{l_N}^{m_N *}(\hat{R}'_N) Y_{l_{N-1}}^{m_{N-1}}(\hat{R}_{N-1}) Y_{l_{N-1}}^{m_{N-1} *}$$
$$\times (\hat{R}'_{N-1})\cdots Y_{l_1}^{m_1}(\hat{R}_1) Y_{l_1}^{m_1 *}(\hat{R}'_1). \tag{33}$$

The orthogonality condition[19]

$$\int_0^{\pi/2} (\sin\beta_\sigma)^{(2\lambda b(M-2)+3M_b-1)} (\cos\beta_\sigma)^{(2\lambda a(\sigma)+3M_a-1)} \times F(-\aleph_{M-1}, \nu_{M-1}-\aleph_{M-1}; \nu_{M-2}+1; \sin^2\beta_\sigma)$$

$$\times F(-\lambda'_{M-1}, \nu_{M-1} - \lambda'_{M-1}; \nu_{M-2} + 1; \sin^2\beta_\sigma)\, d\beta_\sigma$$

$$= \delta_{\lambda_{M-1}\lambda'_{M-1}} \times \left(\frac{(\lambda_{M-1})![\Gamma(\nu_{M-2}+1)]^2 \Gamma(\lambda_{M-1} + \lambda a(\sigma) + 3M_a/2)}{2\nu_{M-1}\Gamma(\nu_{M-1} - \lambda_{M-1})\Gamma(\lambda_{M-1} + \lambda b(M-2) + 3M_b/2)} \right) \tag{34}$$

is utilized to perform the $(N-1)$-fold integration over the hyperspherical angles β_j, $j = 1, 2, \ldots, N-1$, in Eq. (33), and the resulting expression is

$$\begin{aligned} G(\mathfrak{R}|\mathfrak{R}') = {} & \prod_{i=1}^{N} (\eta_i)^3 \left(\frac{1}{\rho\rho'}\right)^{(3N/2-1)} \\ & \times \sum_{l_N}\sum_{l_{N-1}} \cdots \sum_{l_1}\sum_{m_N m_{N-1}} \cdots \sum_{m_1}\sum_{\lambda_1}\sum_{\lambda_2}\cdots\sum_{\lambda_{N-1}} \\ & \times \int_0^\infty \frac{Q J_{\nu_{N-1}}(Q\rho) J_{\nu_{N-1}}(Q\rho')\, dQ}{Q^2 - K^2} \\ & \times Y_{l_N}^{m_N}(\hat{R}_N)\, Y_{l_N}^{m_N *}(\hat{R}'_N)\, Y_{l_{N-1}}^{m_{N-1}}(\hat{R}_{N-1}) \\ & \times Y_{l_{N-1}}^{m_{N-1} *}(\hat{R}'_{N-1}) \cdots Y_{l_1}^{m_1}(\hat{R}_1)\, Y_{l_1}^{m_1 *}(\hat{R}'_1) \\ & \times \mathfrak{N}_1^{-2} \cos^{l_{N-1}}\alpha_{N-1} \sin^{\lambda b(0)}\alpha_{N-1} \\ & \times F(-\lambda_1, \nu_1 - \lambda_1; \nu_0 + 1; \sin^2\alpha_{N-1}) \\ & \times \cos^{l_{N-1}}\alpha'_{N-1} \sin^{\lambda b(0)}\alpha'_{N-1} \\ & \times F(-\lambda_1, \nu_1 - \lambda_1; \nu_0 + 1; \sin^2\alpha'_{N-1}) \\ & \times \mathfrak{N}_2^{-2} \cos^{l_{N-2}}\alpha_{N-2} \sin^{\lambda b(1)}\alpha_{N-2} \\ & \times F(-\lambda_2, \nu_2 - \lambda_2; \nu_1 + 1; \sin^2\alpha_{N-2}) \\ & \times \cos^{l_{N-2}}\alpha'_{N-2} \sin^{\lambda b(1)}\alpha'_{N-2} \\ & \times F(-\lambda_2, \nu_2 - \lambda_2; \nu_1 + 1; \sin^2\alpha'_{N-2}) \\ & \cdots \\ & \times \mathfrak{N}_{N-1}^{-2} \cos^{l_1}\alpha_1 \sin^{\lambda b(N-2)}\alpha_1 \\ & \times F(-\lambda_{N-1}, \nu_{N-1} - \lambda_{N-1}; \nu_{N-2} + 1; \sin^2\alpha_1) \\ & \times \cos^{l_1}\alpha'_1 \sin^{\lambda b(N-2)}\alpha'_1 \\ & \times F(-\lambda_{N-1}, \nu_{N-1} - \lambda_{N-1}; \nu_{N-2} + 1; \sin^2\alpha'_1) \end{aligned} \tag{35}$$

The summations in the above equation collapse into finite sums for a fixed value of $N-1$. Thus, only a finite number of terms are associated with any one hyperradial Green's function.

The hyperradial integral in Eq. (34) can be evaluated to yield[20]

$$\tfrac{1}{2}\pi i J_{\nu_{N-1}}(K\rho_<)\, H^{(1)}_{\nu_{N-1}}(K\rho_>),$$

and the outgoing wave solution $G(\mathfrak{R}|\mathfrak{R}')$ of Eq. (2) is thus determined where $\rho_<, \rho_>$ are the lesser and greater, respectively, of ρ and ρ'. The cylindrical Hankel function of the first kind is designated $H^{(1)}$, and the $\mathfrak{N}_j^{-2}$, $j = 1, 2, \ldots, N-1$, are obtained from Eq. (25).

This representation of the many-particle Green's function appears to be useful in obtaining non-Hartree–Fock wavefunctions and energy levels for many electron atoms.[13]

ACKNOWLEDGEMENTS

The authors gratefully acknowledge support of this research by the Robert A. Welch Foundation, Houston, Texas under Grant D-395 and by the National Science Foundation. Useful suggestions by the referee are appreciated. The author W. G. C. is indebted to W. N. Sams for very helpful discussions.

* This research will constitute a portion of a thesis to be presented by W.G.C. to the Graduate School of the University of Houston in partial fulfillment of the requirements for the Ph.D. degree in Biophysical Sciences.
† Robert A. Welch Foundation Predoctoral Fellow, 1970-71 and trainee under NIH Graduate Training Grant GM 942, 1966–70.
[1] L. M. Delves, Nucl. Phys. **9**, 391 (1958–1959).
[2] L. M. Delves, Nucl. Phys. **20**, 275 (1960).
[3] F. T. Smith, Phys. Rev. **120**, 1058 (1960). Additional references are given in this paper.
[4] F. T. Smith, J. Math. Phys. **3**, 735 (1962).
[5] A. J. Dragt, J. Math. Phys. **6**, 533 (1965).
[6] R. D. Amado, in *Elementary Particle Physics and Scattering Theory*, edited by M. Chrétien and S. Schweber (Gordon and Breach, New York, 1970), Vol. II, pp. 5–123.
[7] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), p. 1730.
[8] R. J. White and F. H. Stillinger, Jr., J. Chem. Phys. **52**, 5800 (1970).
[9] A. Sommerfeld, Math. Ann. **119**, 1 (1943).
[10] L. C. Hostler, J. Math. Phys. **11**, 2966 (1970).
[11] W. N. Sams and D. J. Kouri, J. Chem. Phys. **51**, 4809, 4815 (1969).
[12] W. N. Sams and D. J. Kouri, J. Chem. Phys. **53**, 496 (1970).
[13] W. G. Cooper and D. J. Kouri, J. Chem. Phys. (to be published).
[14] Reference 6, p. 29.
[15] A. Messiah, *Quantum Mechanics* (North-Holland, Amsterdam, 1965), p. 470.
[16] Reference 15, p. 497, where $e^{i\mathbf{g}\cdot(\mathbf{R}-\mathbf{R}')}$ is written $(e^{i\mathbf{g}\cdot\mathbf{R}})(e^{i\mathbf{g}\cdot\mathbf{R}'})^*$.
[17] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton U. P., Princeton, N.J., 1957), p. 21.
[18] G. N. Watson, *A Treatise on the Theory of Bessel Functions* (Cambridge U.P., Cambridge, 1966), p. 370.
[19] Reference 7, p. 1755.
[20] Reference 18, p. 429.

Another Set of Axioms for Classical Gas Dynamics

Michael Schilder

*Department of Mathematics, Newark College of Engineering, Newark, New Jersey 07102**

(Received 6 March 1970; Revised Manuscript Received 8 February 1971)

Another set of axioms is given for a nonequilibrium classical gas composed of hard, spherical, nonattracting molecules. It is shown that the thermodynamic functions of this type of gas can be expressed as multiple integrals. It is also shown in a certain limiting case that the thermodynamic functions of this type of gas can be expressed as Wiener integrals.

1. INTRODUCTION

A new model is given for a nonequilibrium gas composed of hard, spherical, nonattracting molecules. It is shown that with this model such nonequilibrium parameters as average velocity at a point, pressure at a point, density, and the correlation functions may be exactly expressed as multiple integrals. Thus a theory of gas dynamics is derived in terms of inte-

$$\times F(-\lambda'_{M-1}, \nu_{M-1} - \lambda'_{M-1}; \nu_{M-2} + 1; \sin^2\beta_\sigma)\, d\beta_\sigma$$

$$= \delta_{\lambda_{M-1}\lambda'_{M-1}} \times \left(\frac{(\lambda_{M-1})![\Gamma(\nu_{M-2}+1)]^2 \Gamma(\lambda_{M-1} + \lambda a(\sigma) + 3M_a/2)}{2\nu_{M-1}\Gamma(\nu_{M-1} - \lambda_{M-1})\Gamma(\lambda_{M-1} + \lambda b(M-2) + 3M_b/2)} \right) \tag{34}$$

is utilized to perform the $(N-1)$-fold integration over the hyperspherical angles β_j, $j = 1, 2, \ldots, N-1$, in Eq. (33), and the resulting expression is

$$\begin{aligned} G(\mathfrak{R}|\mathfrak{R}') = {} & \prod_{i=1}^{N} (\eta_i)^3 \left(\frac{1}{\rho\rho'}\right)^{(3N/2-1)} \\ & \times \sum_{l_N}\sum_{l_{N-1}} \cdots \sum_{l_1}\sum_{m_N m_{N-1}} \cdots \sum_{m_1}\sum_{\lambda_1}\sum_{\lambda_2}\cdots\sum_{\lambda_{N-1}} \\ & \times \int_0^\infty \frac{Q J_{\nu_{N-1}}(Q\rho) J_{\nu_{N-1}}(Q\rho')\, dQ}{Q^2 - K^2} \\ & \times Y_{l_N}^{m_N}(\hat{R}_N)\, Y_{l_N}^{m_N *}(\hat{R}'_N)\, Y_{l_{N-1}}^{m_{N-1}}(\hat{R}_{N-1}) \\ & \times Y_{l_{N-1}}^{m_{N-1} *}(\hat{R}'_{N-1}) \cdots Y_{l_1}^{m_1}(\hat{R}_1)\, Y_{l_1}^{m_1 *}(\hat{R}'_1) \\ & \times \mathfrak{N}_1^{-2} \cos^{l_{N-1}}\alpha_{N-1} \sin^{\lambda b(0)}\alpha_{N-1} \\ & \times F(-\lambda_1, \nu_1 - \lambda_1; \nu_0 + 1; \sin^2\alpha_{N-1}) \\ & \times \cos^{l_{N-1}}\alpha'_{N-1} \sin^{\lambda b(0)}\alpha'_{N-1} \\ & \times F(-\lambda_1, \nu_1 - \lambda_1; \nu_0 + 1; \sin^2\alpha'_{N-1}) \\ & \times \mathfrak{N}_2^{-2} \cos^{l_{N-2}}\alpha_{N-2} \sin^{\lambda b(1)}\alpha_{N-2} \\ & \times F(-\lambda_2, \nu_2 - \lambda_2; \nu_1 + 1; \sin^2\alpha_{N-2}) \\ & \times \cos^{l_{N-2}}\alpha'_{N-2} \sin^{\lambda b(1)}\alpha'_{N-2} \\ & \times F(-\lambda_2, \nu_2 - \lambda_2; \nu_1 + 1; \sin^2\alpha'_{N-2}) \\ & \cdots \\ & \times \mathfrak{N}_{N-1}^{-2} \cos^{l_1}\alpha_1 \sin^{\lambda b(N-2)}\alpha_1 \\ & \times F(-\lambda_{N-1}, \nu_{N-1} - \lambda_{N-1}; \nu_{N-2} + 1; \sin^2\alpha_1) \\ & \times \cos^{l_1}\alpha'_1 \sin^{\lambda b(N-2)}\alpha'_1 \\ & \times F(-\lambda_{N-1}, \nu_{N-1} - \lambda_{N-1}; \nu_{N-2} + 1; \sin^2\alpha'_1) \end{aligned} \tag{35}$$

The summations in the above equation collapse into finite sums for a fixed value of $N-1$. Thus, only a finite number of terms are associated with any one hyperradial Green's function.

The hyperradial integral in Eq. (34) can be evaluated to yield[20]

$$\tfrac{1}{2}\pi i J_{\nu_{N-1}}(K\rho_<)\, H^{(1)}_{\nu_{N-1}}(K\rho_>),$$

and the outgoing wave solution $G(\mathfrak{R}|\mathfrak{R}')$ of Eq. (2) is thus determined where $\rho_<, \rho_>$ are the lesser and greater, respectively, of ρ and ρ'. The cylindrical Hankel function of the first kind is designated $H^{(1)}$, and the $\mathfrak{N}_j^{-2}$, $j = 1, 2, \ldots, N-1$, are obtained from Eq. (25).

This representation of the many-particle Green's function appears to be useful in obtaining non-Hartree–Fock wavefunctions and energy levels for many electron atoms.[13]

ACKNOWLEDGEMENTS

The authors gratefully acknowledge support of this research by the Robert A. Welch Foundation, Houston, Texas under Grant D-395 and by the National Science Foundation. Useful suggestions by the referee are appreciated. The author W. G. C. is indebted to W. N. Sams for very helpful discussions.

* This research will constitute a portion of a thesis to be presented by W.G.C. to the Graduate School of the University of Houston in partial fulfillment of the requirements for the Ph.D. degree in Biophysical Sciences.
† Robert A. Welch Foundation Predoctoral Fellow, 1970-71 and trainee under NIH Graduate Training Grant GM 942, 1966–70.
[1] L. M. Delves, Nucl. Phys. **9**, 391 (1958–1959).
[2] L. M. Delves, Nucl. Phys. **20**, 275 (1960).
[3] F. T. Smith, Phys. Rev. **120**, 1058 (1960). Additional references are given in this paper.
[4] F. T. Smith, J. Math. Phys. **3**, 735 (1962).
[5] A. J. Dragt, J. Math. Phys. **6**, 533 (1965).
[6] R. D. Amado, in *Elementary Particle Physics and Scattering Theory*, edited by M. Chrétien and S. Schweber (Gordon and Breach, New York, 1970), Vol. II, pp. 5–123.
[7] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), p. 1730.
[8] R. J. White and F. H. Stillinger, Jr., J. Chem. Phys. **52**, 5800 (1970).
[9] A. Sommerfeld, Math. Ann. **119**, 1 (1943).
[10] L. C. Hostler, J. Math. Phys. **11**, 2966 (1970).
[11] W. N. Sams and D. J. Kouri, J. Chem. Phys. **51**, 4809, 4815 (1969).
[12] W. N. Sams and D. J. Kouri, J. Chem. Phys. **53**, 496 (1970).
[13] W. G. Cooper and D. J. Kouri, J. Chem. Phys. (to be published).
[14] Reference 6, p. 29.
[15] A. Messiah, *Quantum Mechanics* (North-Holland, Amsterdam, 1965), p. 470.
[16] Reference 15, p. 497, where $e^{i\mathbf{g}\cdot(\mathbf{R}-\mathbf{R}')}$ is written $(e^{i\mathbf{g}\cdot\mathbf{R}})(e^{i\mathbf{g}\cdot\mathbf{R}'})^*$.
[17] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton U. P., Princeton, N.J., 1957), p. 21.
[18] G. N. Watson, *A Treatise on the Theory of Bessel Functions* (Cambridge U.P., Cambridge, 1966), p. 370.
[19] Reference 7, p. 1755.
[20] Reference 18, p. 429.

Another Set of Axioms for Classical Gas Dynamics

Michael Schilder

*Department of Mathematics, Newark College of Engineering, Newark, New Jersey 07102**

(Received 6 March 1970; Revised Manuscript Received 8 February 1971)

Another set of axioms is given for a nonequilibrium classical gas composed of hard, spherical, nonattracting molecules. It is shown that the thermodynamic functions of this type of gas can be expressed as multiple integrals. It is also shown in a certain limiting case that the thermodynamic functions of this type of gas can be expressed as Wiener integrals.

1. INTRODUCTION

A new model is given for a nonequilibrium gas composed of hard, spherical, nonattracting molecules. It is shown that with this model such nonequilibrium parameters as average velocity at a point, pressure at a point, density, and the correlation functions may be exactly expressed as multiple integrals. Thus a theory of gas dynamics is derived in terms of inte-

grals, rather than in terms of partial differential equations.

Since the average velocity at a point of the gas modeled in this paper satisfies the hydrodynamic equations, and since its total energy is preserved with time independent potentials, the integral formulation given here is essentially equivalent to the differential formulations.

While these integrals are difficult to approximate numerically, their numerical approximation seems to be no more difficult than the present methods used to solve the Maxwell–Boltzmann equation or the Navier–Stokes equations.

It will further be shown in the limiting case as the temperature of the gas approaches infinity that the typical path of a molecule in the gas has the same distribution as Brownian motion with a drift coefficient. In this case it will be shown that all parameters of the gas may be expressed as Wiener integrals. The author has recently shown how this type of Wiener integral can be numerically evaluated.

There is a connection between the type of Wiener integral derived in this paper and the path integral of quantum mechanics. This connection will be given in another paper.

The approach taken is qualitatively as follows. At each point the initial local temperature, mean velocity, and density are given for the gas. It is assumed that when the molecules of the gas collide, the *average* velocity and energy of the molecules at the point where the collision takes place are conserved. It is shown that the preservation of these quantities implies preservation of local temperature also. It is also assumed that the paths of the molecules are continuous. A further assumption is made that the distribution of the velocities of the molecules just after a collision has maximum entropy, subject to the constraints of preservation of energy and average velocity. It is shown that, with these assumptions, the distribution of the velocities at a given point just after a collision is completely determined and that this distribution is normal with mean the preserved average velocity and with variance proportional to the preserved local temperature. Since the paths of the molecules are assumed to be continuous, the distribution of the positions of the molecules just after a collision is the same as just before. Therefore, it is shown that, with these assumptions, it is possible to calculate the distribution of both the position and velocity of the molecules just after a collision, given that it is known just before a collision. Since it is assumed that the molecules obey the laws of classical mechanics, the distribution of their positions and velocities can be determined in the times between collisions since their position paths obey a given differential equation with initial conditions the random position and velocity of their last collision.

It is then shown that the distribution of the positions and velocities of the molecules at any time after the initial time and at any place can be calculated. This distribution is determined by computing the expected value of any function of the paths or their velocities with respect to it. This expected mean, or average, is shown to be a multiple integral, or a ratio of multiple integrals.

While these multiple integrals can be explicitly written down, their actual numerical calculation appears to present serious difficulties.

Since the mean intermolecular time is very small for most gases, a natural approximation for the above integrals is to let this parameter tend to zero. An analysis of what happens in this case is given. In order that the variance of the positions of the molecules stay positive, it is necessary that the temperature of the gas approach infinity as the mean intermolecular collision time approaches zero.

If these assumptions are made, it is shown that the above multiple integrals become Wiener integrals (real path integrals), which are very close in form to the Feynman path integrals which solve the Schrödinger equation.

These real path integrals depend on only the external potential and the initial density and average velocity.

Thus with the above assumptions *any* parameter of a nonequilibrium gas can be expressed as a Wiener integral.

While the collisions in a gas without long range interaction forces occur at random time intervals, the assumption is made that the collisions occur at regular time intervals. This is justified on the ground that the intercollision time is very small, so that the error introduced is negligible.

In the case of the high temperature approximation, the equation of the position paths of the molecules, $x(t)$, formally becomes a stochastic equation of the form

$$\dot{x}(t) = u(t, x(t)) + \lambda \dot{z}(t),$$

where $\dot{z}(t)$ is the derivative of Brownian motion and u satisfies the partial differential equation

$$u_t + uu_x + \tfrac{1}{2}u_{xx} + V_x = 0,$$

where V is the external potential for the gas. This contrasts with the Ornstein–Uhlenbeck model

$$\ddot{x} = -V_x + \lambda \dot{z},$$

described, for example, in Nelson.[1]

2. THE MODEL AND SOME DEFINITIONS

The following axioms are assumed for the gas:

(a) The molecules of the gas obey the laws of Newtonian mechanics.

(b) The position vectors of all molecules, the energy, and the average velocity of the gas at each point of the space in which the gas is enclosed are continuous functions of time.

(c) The molecules are nonattracting, and the collisions between them are elastic.

(d) At collisions the entropy is maximized.
A fifth axiom is added for the sake of mathematical convenience.

(e) The collisions occur at all points of the space at regular intervals $t_i = i\Delta$, $i = 0, 1, 2, \cdots$.

It will also be assumed for mathematical convenience

that the mass of the molecules to be dealt with is unity and that they move in only one dimension. A considerable amount of notational and mathematical clarity is achieved by this last assumption with no loss in generality since all results are invariant with respect to the dimension of the space in which the molecules move, if the external force is the negative of the gradient of a potential, and the initial mean velocity is irrotational.

It is assumed that the gas occupies all of space. A later paper will shown how boundary conditions can be included in the model.

It is also assumed that if a function is differentiated, its derivative exists and, that if the function is integrated, its integral exists.

We assume now a gas satisfying axioms (a), (b), (c), (d), and (e) listed previously. At each point x and for all times $t \geq 0$ a potential $V(t,x)$ is assumed given, with $-V_x = f$.

At time 0, it is assumed that at each point x of space the average velocity $u(0,x)$, the density $\rho(0,x)$, and the variance of the velocity $k(0,x)$ are given.

It will be shown how, with the help of axioms (a)–(e), the quantities u, ρ, and k can be determined inductively for all time.

At times t_i, by axiom (e), a collision takes place. We will denote by t_i^- quantities at time t_i *just before the collision takes place* and by t_i^+ quantities *just after the collision takes place*. By axiom (b), the quantities u, k, and ρ are continuous in time at collisions (ρ is continuous since the paths of the molecules are), and therefore

$$u(t_i^-,x) = u(t_i^+,x), \quad k(t_i^-,x) = k(t_i^+,x), \quad \text{and}$$

$$\rho(t_i^-,x) = \rho(t_i^+,x) \quad \text{for each } x.$$

We use the notation

$$E\{F(x,\dot{x})\}$$

to denote the expected, average, or mean value of the function $F(x,\dot{x})$ of the position paths of the molecules x and their velocities $\dot{x}$ and the notation

$$E\{F(x,\dot{x}) \mid x(t) = x\}$$

to denote the expected, average, or mean value of $F(x,\dot{x})$ given $x(t) = x$, or, in other words, the expected, mean, or average value of $F(x,\dot{x})$ over all paths whose position at time t is x. Thus

$$u(t,x) = E\{\dot{x}(t) \mid x(t) = x\}$$

and

$$k(t,x) = E\{(\dot{x}(t) - u(t,x))^2 \mid x(t) = x\}.$$

At time $0^- = t_0^-$ at each point x, the velocities have a certain distribution whose exact form does not concern us; we know by assumption, however, that their mean is $u(0,x)$ and that their variance is $k(0,x)$.

Let N be total number of molecules in the gas. Then $\rho(0,x)dxN$ is defined to be the number of molecules at x at time 0.

Let $\rho(0^-,\dot{x} \mid x(0) = x)$ be the unknown conditional distribution of the velocities of the molecules at x at time 0^-. It can be seen that the total number of molecules at x with velocity $\dot{x}$ at time 0^- is $\rho(0^-,\dot{x} \mid x(0) = x)\rho(0,x)Ndxd\dot{x}$, and thus the total kinetic energy of the molecules at x at time 0 is

$$\int_{-\infty}^{\infty} \tfrac{1}{2}\dot{x}^2\rho(0^-,\dot{x} \mid x(0) = x)d\dot{x}\ \rho(0,x)dx\,N = E\{\tfrac{1}{2}[\dot{x}(0)]^2 \mid x(0) = x\}\rho(0,x)dx\,N$$

The energy at $0^-, x, E(0^-,x)$, is now seen to be

$$\begin{aligned} E(0^-,x) &= (E\{[\dot{x}(0)]^2/2 \mid x(0) = x\} + V(0,x))\rho(0,x)dx\,N \\ &= (\tfrac{1}{2}E\{[\dot{x}(0) - u(0,x)]^2 \mid x(0) = x\} + \tfrac{1}{2}u^2(0,x) \\ &\quad + V(0,x))\rho(0,x)dx\,N \\ &= (\tfrac{1}{2}k(0,x) + \tfrac{1}{2}u^2(0,x) + V(0,x))\rho(0,x)Ndx \end{aligned} \tag{2.1}$$

Equation (2.1) shows that $k(0,x)$ must be continuous at time 0, since all the other members of (2.1) are.

By assumption, collisions occur at $0, x$. By axiom (b), the paths of the molecules are continuous, and therefore the position of a molecule just after a collision is the same as just before it. Thus the density function is the same just after a collision as just before it. This is not true, of course, of the velocities which by axiom (c) change instantaneously.

To obtain the distribution function for the velocities just after a collision, axiom (d) is used. By axiom (b) $E(0,x)$ and the average velocity $u(0,x)$ are the same just after the collision as just before, and therefore, as is well known, we must maximize

$$-\int_{-\infty}^{\infty}\rho(0^+,\dot{x} \mid x(0) = x)\ln[\rho(0^+,\dot{x} \mid x(0) = x)]d\dot{x} \tag{2.2}$$

subject to the restrictions

$$\int_{-\infty}^{\infty}\rho(0^+,\dot{x} \mid x(0) = x)d\dot{x} = 1, \tag{2.3}$$

$$\int_{-\infty}^{\infty}\dot{x}\rho(0^+,\dot{x} \mid x(0) = x)dx = u(0,x), \tag{2.4}$$

$$\int_{-\infty}^{\infty}\tfrac{1}{2}\dot{x}^2\rho(0^+,\dot{x} \mid x(0) = x)d\dot{x} + V(0,x)\rho(0,x)dxN = E(0,x). \tag{2.5}$$

Equation (2.4) is, of course, just the condition of preservation of average velocity, and (2.5) is the preservation of energy. Via (2.1), (2.3) and (2.4), it can be seen that (2.5) can be written as

$$\int_{-\infty}^{\infty}[\dot{x} - u(0,x)]^2\rho(0^+,\dot{x} \mid x(0) = x)d\dot{x} = k(0,x). \tag{2.5'}$$

The variational problem (2.2), (2.3), (2.4), and (2.5) is solved by Lagrange multipliers. Its solution is easily shown to be

$$\rho(0^+,\dot{x}) \mid x(0) = x) = \frac{\exp\{-[\dot{x} - u(0,x)]^2/2c\}}{\int_{-\infty}^{\infty}\exp\{-[\dot{x} - u(0,x)]^2/2c\}d\dot{x}}, \tag{2.6}$$

where c is determined by (2.5')

Since

$$(2\pi c)^{-1/2}\int_{-\infty}^{\infty}[\dot{x} - u(0,x)]^2 \times \exp\{-[\dot{x} - u(0,x)]^2/2c\}d\dot{x} = c,$$

it follows that

$$c = k(0,x)$$

and from (2.6) it follows that

$$\rho(0^+,\dot{x}\,|\,x) = [2\pi k(0,x)]^{-1/2} \times \exp\{-[\dot{x}-u(0,x)]^2/2k(0,x)\}. \quad (2.7)$$

If the gas considered were in equilibrium, then the c defined above would be RT, where R is the gas constant and T is the temperature of the gas (see Jeans,[2] for example). In the case considered, therefore, $k(0,x)$ has interpretation of being $RT(0,x)$, where R is the gas constant and $T(0,x)$ is the local temperature.

The density function $\rho(0^+,\dot{x}\,|\,x(0)=x)$ gives the distribution $\dot{x}$ at 0^+ given $x(0)=x$. The joint distribution of x and $\dot{x}$ is therefore

$$\rho(0^+,\dot{x},x) = [2\pi k(0,x)]^{-1/2}\rho(0,x) \times \exp\{[\dot{x}-u(0,x)]^2/2k(0,x)\}. \quad (2.8)$$

If $0 \le t \le t_1$ and F is an integrable function of $x(t)$ and $\dot{x}(t)$, it follows that

$$\begin{aligned} &E\{F(\dot{x}(t),x(t))\} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(\dot{x}(t),x(t))\rho(0,\dot{x}_0,x_0)d\dot{x}_0dx_0 \\ &= \int_{-\infty}^{\infty}[2\pi k(0,x_0)]^{-1/2}\int_{-\infty}^{\infty} F(\dot{x}(t),x(t))\rho(0,x_0) \\ &\quad\times \exp\{-[\dot{x}_0-u(0,x_0)]^2/2k(0,x_0)\}d\dot{x}_0dx_0. \end{aligned} \quad (2.9)$$

$x(t)$ is the unique solution to the equation $\dot{x}'(t) = f(t,x(t))$ with initial conditions $x(0)=x_0$ and $\dot{x}(0)=\dot{x}_0$.

It is to be emphasized that the differential equation $\ddot{x}(t) = f(t,x(t))$ must be solved for all values of its initial conditions, which are the variables of integration. This can, of course, be done explicitly only in special cases such as linear or quadratic potentials, $V(t,x)$.

Let $\chi_B(x) = 1$ if x is in a set B which is in the space in which the molecules are moving, and let $\chi_B(x) = 0$ if x is not in B. It follows then that the normalized density at t,x is

$$\rho(t,x) = \lim_{B\to x} E\{\chi_B(x(t))\}/|B|,$$

where $|B|$ is the length (in one dimension) of B; $\lim_{B\to x}$ means that B contains x and shrinks down to it. We define

$$E\{F(\dot{x}(t),x(t))\,|\,x(t)=x\} \text{ as } \lim_{B\to x}\frac{E\{F(\dot{x}(t),x(t))\chi_B(x(t))\}}{E\{\chi_B(x(t))\}}, \quad (2.10)$$

where the $\lim_{B\to x}$ is the same as before. See Kac.[3]

It can be seen if $0 \le t \le t_1$

$$\rho(t,x) = \int^{x(t)=x}[2\pi k(0,x_0)]^{-1/2}\rho(0,x_0)\exp\{-[\dot{x}_0-u(0,x_0)]^2/2k(0,x_0)\}d\dot{x}_0dx_0, \quad (2.11)$$

$$u(t,x) = \frac{\int^{x(t)=x}[2\pi k(0,x_0)]^{-1/2}\dot{x}(t)\rho(0,x_0)\exp\{-[\dot{x}_0-u(0,x_0)]^2/2k(0,x_0)\}d\dot{x}_0dx_0}{\rho(t,x)}, \quad (2.12)$$

and

$$k(t,x) = \frac{\int^{x(t)=x}[2\pi k(0,x)]^{-1/2}[\dot{x}(t)-u(0,x)]^2\rho(0,x_0)\exp\{-[\dot{x}_0-u(0,x_0)]^2/2k(0,x_0)\}d\dot{x}dx}{\rho(t,x)}. \quad (2.13)$$

As before, $\dot{x}(t)$ is obtained by solving the differential equation $\ddot{x}(t) = f(t,x(t))$ with initial conditions the variables of integration $\dot{x}_0, x_0$. The notation $\int^{x(t)=x}d\dot{x}_0dx_0$ means to integrate over all values of $\dot{x}_0$ and x_0 such that $x(t)=x$.

It follows that if the initial density, temperature [since $RT(0,x) = k(0,x)$], and average velocity of the gas are known for each initial point x and if the gas considered satisfies axioms (a)–(e) that these parameters can be calculated (in principle anyway) at time t_1^-, for all values of x. t_1 is, as defined before, the time of the second collision. Equation (2.8) shows that if u, k, and ρ are known just before the first collision, and if axioms (a)–(e) are assumed, the distribution of x and $\dot{x}$ just after the first collision can be determined. We apply (2.8) to determine the distribution of x and $\dot{x}$ just after the second collision, using the values of u, ρ, and k at t_1 obtained from (2.11), (2.12), and (2.13).

Therefore, by induction, it follows that any thermodynamic function [see (2.9)] of the gas can be calculated, since in fact the general distribution of x and $\dot{x}$ is now given.

If $t_n \le t \le t_{n+1}$, the following formula can be obtained by induction:

$$\begin{aligned} &E\{F(\dot{x}(t),x(t))\} \\ &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left(\prod_{i=0}^{i=n}2\pi k(t_i,x_i)\right)^{-1/2}F(\dot{x}(t),x(t))\rho(0,x_0) \\ &\quad\times\exp\left(-\sum_{i=0}^{i=n}\frac{[\dot{x}_i-u(t_i,x)]^2}{2k(t_i,x_i)}\right)dx_0d\dot{x}_0d\dot{x}_1\cdots d\dot{x}_n \end{aligned} \quad (2.14)$$

$u(t_i,x)$ and $k(t_i,x)$ for $i \ge 1$ can be calculated as before or by the recursive formulas

$$u(t_i,x) = \frac{\int\cdots\int^{x(t_i)=x}[\prod_{j=0}^{j=i-1}2\pi k(t_j,x_j)]^{-1/2}\rho(0,x_0)[\dot{x}(t_i)]\exp[-\sum^{i-1}[\dot{x}_j-u(t_j,x_j)]^2/2k(t_j,x_j)]dx_0d\dot{x}_0\cdots d\dot{x}_{i-1}}{\int\cdots\int^{x(t_i)=x}[\prod_{j=0}^{j=i-1}2\pi k(t_j,x_j)]^{-1/2}\rho(0,x_0)\exp\{-\sum[\dot{x}_j-u(t_j,x_j)]^2/2k(t_j,x_j)\}dx_0d\dot{x}_0\cdots d\dot{x}_{i-1}}, \quad (2.15)$$

$$k(t_i,x) = \frac{\int\cdots\int^{x(t_i)=x} [\Pi_{j=0}^{j=i-1} 2\pi k(t_j,x_j)]^{-1/2}\rho(0,x_0)\{\dot{x}(t_i)-u(t_i,x_i)]^2 \exp\{-\sum^{i-1}[\dot{x}_j-u(t_j,x_j)]^2/2k(t_j,x_j)\}dx_0 d\dot{x}_0\cdots d\dot{x}_{i-1}}{\int\cdots\int^{x(t_i)=x}[\Pi\pi k(t_j,x_j)]^{-1/2}\rho(0,x_0)\exp\{-\sum[\dot{x}_j-u(t_j,x_j)]^2/2k(t_j,x_j)\}dx_0 d\dot{x}_0\cdots d\dot{x}_{i-1}} \tag{2.16}$$

The notation $\int\cdots\int^{x(t_i)=x}$ means to integrate over those values of $x_0, \dot{x}_0, \ldots, \dot{x}_{i-1}$ such that $x(t_i) = x$. The $x_0, \dot{x}_0, \ldots, \dot{x}_n$ are independent quantities (i.e., they are variables of integration). The variables x_i are calculated inductively with $x_1 = x(t_1)$, where $x(t)$ satisfies the differential equation $\ddot{x}(t) = f(t,x(t))$ with initial conditions $x(0) = x_0$, $\dot{x}(0) = \dot{x}_0$, $x_2 = x(t_2)$, where $\ddot{x}(t) = f(t,x(t))$ with initial conditions $x(t_1) = x_1$ and $\dot{x}(t_1) = \dot{x}_1$ and so on for all the other x_i. $x(t)$ is the solution to $\ddot{x}(t) = f(t,x(t))$ with initial conditions $x(t_n) = x_n$, $\dot{x}(t_n) = \dot{x}_n$. $u(0,x)$ and $k(0,x)$ are, as before, given initially.

Another way of interpreting the probability distribution of $x(t)$ and $\dot{x}(t)$ is possible. This representation will be needed later. If $t_n < t \le t_{n+1}$, then $\dot{x}(t), x(t)$ is the solution of the differential equation $\ddot{x}(t) = f(t,x(t))$ with initial conditions

$$x(t_n^+) = x(t_n^-) \tag{2.17}$$

and

$$\dot{x}(t_n^+) = u(t_n, x(t_n^-)) + \dot{z}_n(k(t_n, x(t_n^-)))^{1/2}. \tag{2.18}$$

$\dot{z}_n$ is a Gaussian (normal) random variable with mean zero and variance one. u and k are, as before, computed by the formulas

$$u(t_n,x) = E\{\dot{x}(t_n^-) \mid x(t_n) = x\}$$

$$k(t_n,x) = E\{(\dot{x}(t_n^-) - u(t,x(t_n)))^2 \mid x(t_n) = x\}$$

$x(0)$ has distribution $\rho(0,x(0))$, which is assumed given as are $u(0,x)$ and $k(0,x)$.

This can be seen as follows.

We change variables in (2.14) by letting

$$\dot{z}_i = [\dot{x}_i - u(t_i,x_i)]/[k(t_i,x_i)]^{1/2}, \text{ or}$$

$$\dot{x}_i = u(t_i,x_i) + \dot{z}_i(k(t_i,x_i))^{1/2}. \tag{2.19}$$

The Jacobian of this transformation is $\Pi_{i=0}^{i=n}[k(t_i,x_i)]^{1/2}$ and so (2.14) can be rewritten as

$$E\{F(\dot{x}(t),x(t))\} = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}(2\pi)^{-n/2}F(\dot{x}(t),x(t))\rho(0,x_0) \times \exp[-\sum \dot{z}_i^2/2]dx_0 d\dot{z}_0\cdots d\dot{z}_n. \tag{2.20}$$

Since the functions $(2\pi)^{-1/2}\exp[-(\dot{z}_i^2/2)^2]$ are the distribution functions of normal variables with mean zero and variance one and, since it can be seen that $\dot{x}(t), x(t)$ is actually a function of $\dot{z}_0, \ldots, \dot{z}_n$ as described by (2.17) and (2.18), it follows that the right-hand side of (2.20) is the expected value of the arbitrary function $F(\dot{x}(t),x(t))$ with respect to the Markov process described by (2.17) and (2.18). Since the values of u and k are computed exactly as before, this proves (2.17) and (2.18). (2.17) and (2.18) give another interpretation of the motion of the molecules of the gas. The molecules move according to the laws of Newtonian mechanics until a collision. The position remains the same at the collision, while the velocity is changed to a normal random variable whose mean and variance are the same as the mean and variance of the velocity of the fluid at the point of collision. The phrase "at the point of collision" is the root of most of the mathematical complications of this paper. However, the approach taken seems to be the only way that the gas may be modeled so that energy is preserved.

In a real fluid, of course, the collisions do not occur at regular time intervals but have a probability distribution with respect to both time and space. The most natural distribution of collisions in both time and space of a gas satisfying axioms (a)–(d) is Poisson, and axiom (e) could now be replaced with this assumption. However, it does not seem that the increase in accuracy derived by doing this would justify the additional mathematical complication.

3. SOME CONSEQUENCES OF THE MODEL

If the mathematical model of the proceeding section is to be considered physically valid, consequences of it must agree with known physical phenomena. In this section it will be explained why the total energy of the gas is preserved if the external potential is time independent and in what manner its parameters satisfy the hydrodynamic equations.

It is well known that while a gas is freely flowing between collisions, it satisfies the hydrodynamic equations, which are

$$\rho_t + (u\rho)_x = 0, \tag{3.1}$$

$$u_t + uu_x = -P_x/\rho + f. \tag{3.2}$$

ρ, u, and f are defined in Sec. 2. P is pressure and according to Jeans[4] may be defined, in the notation of this paper, as $P = \rho k$.

At the collision times $\{t_j\}$, all the above quantities are continuous by assumption, although the derivatives ρ_t, u_x, and P_x might not be continuous as functions of time over the times t_j. It should be remembered, however, that any derivation of the equations of motion of a gas with discrete, hard collisions must involve some kind of time smoothing.

That ρ, u, f, and P satisfy (3.1) and (3.2) may also be derived from the theory of the last section.

By definition, the energy of the gas is preserved across collisions at each point and therefore for the entire gas at the collision times. In between collisions, energy is preserved along each path, if the external potential is independent of time, and therefore the total energy of the gas is preserved in these times also. Thus the mean energy of the gas is preserved for all time if the external potential is independent of time.

Since the total energy of the gas is preserved and since the entropy of the gas is always increasing, one

would expect from classical statistical mechanics that the distribution of the velocities and the positions of the molecules in the limit as time approaches infinity would be the Maxwell–Boltzman distribution. A direct proof of this fact seems very difficult and has not been found.

4. THE APPROXIMATION TO BROWNIAN MOTION

While exact expressions for the parameters of the gas considered in this paper are given by (2.14), they are, as has already been pointed out, very difficult to use for computational purposes. Since the intermolecular collision time Δ is extremely small, it seems that a good approximation for these integrals would be to let this parameter tend to zero. This approximation will be considered in this section.

If, however, we let Δ, which in this paper has the interpretation as the mean length of time of a free path of a molecule, tend to zero, without changing the other parameters of the gas, then it can be seen that the variance of $x(t)$ tends to zero which is clearly not desirable.

From another point of view, it is shown in books on gas dynamics (see Jeans[2]) that the mean length of time for a free path is const/$\sqrt{\text{temp}}$. Therefore, if Δ is to tend to zero, we must let the temperature tend to infinity.

In this paper $k(t,x)$ has the dimensions of temperature. It is, therefore, multiplied by λ^2/Δ, where λ is a new constant parameter which gives still some measure of the temperature or activity of the gas.

Representation of the gas as a stochastic process now has the form

$$\dot{x}(t_n^+) = u(t_n, x(t_n)) + \lambda(\dot{z}_n/\sqrt{\Delta})\sqrt{k(t_n, x(t_n))}, \tag{4.1}$$

$$x(t_n^+) = x(t_n^-). \tag{4.2}$$

Consider now a process

$$z(t) = \sum_i^{[t/\Delta]} \frac{\Delta\,\dot{z}_i}{\sqrt{\Delta}} = \sum_i^{[t/\Delta]} \dot{z}_i\sqrt{\Delta}$$

where the $\dot{z}_i$ are, as before, independent random variables with mean zero and variance one. $[t/\Delta]$ is the greatest integer less than t/Δ.

A t is now fixed. The number of collisions n from time 0 to time t is $[t/\Delta]$. We now let $\Delta \to 0$. By definition $[z(t+\Delta) - z(t)]/\Delta$ tends the $\dot{z}(t)$, the derivative of Brownian motion. (4.1) formally becomes

$$\dot{x}(t) = u(t, x(t)) + \lambda\dot{z}(t)[k(t,x)]^{1/2} \tag{4.3}$$

The stochastic equation (4.3) is called an Ito equation and has been extensively studied; see Ito[5] or Doob[6] for the theory of this equation.

We now make another assumption: that the $k(t, x(t))$ of (4.3) is in fact 1. This assumption is necessary since, while it is possible to derive a partial differential equation for u, it has not yet been possible to derive an expression for $k(t,x)$ in the limiting case of $\Delta \to 0$. The assumption that $k(t,x)$ is 1 in the limit case can be justified in a number of ways. It might be assumed that the temperature becomes uniformly infinite throughout the gas, in which case it is justified. The derivation of the Navier–Stokes equation (which is what is really desired for the case at hand) usually includes the assumption that temperature variations are negligible. Since, as has already been pointed out, $k(t,x)$ has the dimensions of temperature, it is again justifiable to set $k(t,x) = 1$. A third possibility of justification is to change the original model. $k(t,x)$ is included in the model so that energy will be conserved across collisions. If we drop this assumption and replace it with the assumption that only average velocity is preserved across collisions and that the impact force $\dot{z}_n$ is of constant variance, then it is again justified to set $k(t,x) = 1$. This assumption will be made for the remainder of the paper.

Equation (4.3) now reads

$$\dot{x}(t) = u(t, x(t)) + \lambda\,\dot{z}(t). \tag{4.4}$$

A rigorous justification that the system described by (4.1) and (4.2) has a limit as $\Delta \to 0$ and that it is of the form (4.3) seems very difficult and has not yet been obtained.

On the other hand, the fact that the stochastic differential equation (4.3) describes a system which is formally the limit of the model of Sec. 2 is fairly clear. Thus (4.3) is introduced to describe such a limit.[7]

Let

$$Dx(t) = \lim_{h \to 0^+} E\{[x(t+h) - x(t)]/h \mid \mathcal{O}_t\}$$

and

$$D_* x(t) = \lim_{h \to 0^+} E\{[x(t+h) - x(t)]/h \mid \mathcal{F}_t\},$$

where $\mathcal{O}_t$ and $\mathcal{F}_t$ are respectively the σ-algebras generated by $x(\alpha)$ with $0 \le \alpha \le t$ and $t \le \alpha < \infty$.

The operators D and D_* are introduced and studied by Nelson.[8]

Since it follows easily from the theory of Sec. 2, for $\Delta > 0$, that

$$D_*^2 x(t) = D^2 x(t) = f(t, x(t)),$$

if $x(t)$ is defined by (4.1) and (4.2), it will be assumed that the $x(t)$ defined by (4.4) has the following properties:

$$D^2 x(t) = D(u(t, x(t))) = f(t, x(t)) \tag{4.5}$$

and

$$D_*^2 x(t) = D_* u_*(t, x(t))) = f(t, x(t)) \tag{4.6}$$

u_* is the drift coefficient of the Markov process (4.4) with time reversed. See Nelson.[9]

Since Nelson[10] shows

$$Du(t, x(t)) = u_t + uu_x + (\lambda^2/2)u_{xx}$$

and

$$D_* u_*(t, x(t)) = u_{*t} + u_* u_{*x} - (\lambda^2/2)u_{*xx},$$

it follows from (4.5) and (4.6), if certain differentiability and boundedness assumptions are also included in the model (4.4), that

$$u_t + uu_{xx} + (\lambda^2/2)u_{xx} = f(t, x) \tag{4.7}$$

and

$$u_{*t} + u_* u_{*x} - (\lambda^2/2)U_{*xx} = f(t, x). \tag{4.8}$$

We proceed to study the models (4.4) and

$$\dot{x}_*(t) = u_*(t, x(t)) + \lambda \dot{z}_*(t), \tag{4.9}$$

where u and u_* satisfy respectively the partial differential equations (4.7) and (4.8) and where $\dot{z}_*(t)$ of (4.9) is the derivative of Brownian motion with time reversed. See Nelson.[10]

It will now be shown that the expected value of any function of the paths $x(t)$, where $x(t)$ is the solution of the stochastic equation $\dot{x}(t) = u(t, x(t)) + \lambda \dot{z}(t)$ and where $u(t, x)$ is the solution of the partial differential equation (4.7), can be expressed as a Wiener integral, or as a real path integral.

Lemma 4.1: Suppose $u(t, x)$ satisfies (4.7) and $U_x(t, x) = u(t, x)$. Then

$$U_t + \tfrac{1}{2} U_x^2 + \tfrac{1}{2}\lambda^2 U_{xx} + V(t, x) + c(t) = 0. \tag{4.10}$$

Proof: Lemma 4.1 is proved by differentiating (4.10) with respect to x, and using (4.7).

By suitable choice of V or U, the function $c(t)$ in (4.10) can be made zero. It will be assumed that this is the case for the remainder of this paper.

Equation (4.10) is known as the stochastic Hamilton–Jacobi equation, or as Bellman's equation, and has been extensively studied by engineers; see Wonham[11] or Flemming.[12]

Let $E^W\{F(z) | z(0) = x\}$ be the expected value of the function $F(z)$ with respect to the Wiener measure that has variance parameter λ^2, and with probability one, has all its paths starting at x at time zero.

The main theorem of this paper can now be proved.

Theorem 4.1: Suppose u satisfies (4.7), that $|u| \leq K$, a constant,[13] and that (4.4) has a unique solution.[14]

Then,

$$E\{F(x(\cdot)) \mid x(0) = x\} = E^w\left\{F(x(\cdot)) \exp\left[\frac{1}{\lambda^2}\left(\int_0^T V(t, x(t))dt + U(T, x(T)) - U(0, x(0))\right)\right] \mid x(0) = x\right\}. \tag{4.11}$$

$V(t, x)$ is the potential for the gas, $U(t, x)$ is a function such that $U_x = u$, where u satisfies (4.7), and F is an arbitrary integrable function of the Brownian motion paths $x(\cdot)$.

Proof: Since $\dot{x} = u + \lambda \dot{z}$, it follows that $x(\cdot) = \int_0^{(\cdot)} u(\alpha, x(\alpha))d\alpha + x(0) + \lambda(z(\cdot) - z(0))$ and therefore $x(\cdot)$ can be written as a function of $z(\cdot)$ and $x(0) = x$: $x(\cdot, z(\cdot), x(0))$. See Schilder[15] for a more thorough discussion of this point.

$E\{F(x(\cdot))\}$ is really by definition

$$E^w\{F(x(\cdot), z(\cdot), x(0)) \mid z(0) = 0\}.$$

[Since only $\dot{z}$ appears in (4.4), $z(0)$ is arbitrary, and it is therefore set equal to zero.]

In the last Wiener integral, we change variables by letting $\dot{z}(t) = (1/\lambda)(\dot{x}(t) - u(t, x(t))$. See (4.4), where $x(\cdot)$ is the new variable of integration and $x(0)$ is fixed at x. Girsanov[16] shows rigorously that $E\{F(x(\cdot))\}$ is

$$E^u\left\{F(x(\cdot)) \exp\left[\frac{1}{\lambda^2}\left(\int_0^T u(t, x(t))dx(t) - \tfrac{1}{2}\int_0^T u^2(t, x(t))dt\right)\right] \mid x(0) = x\right\}.$$

The integral $\int_0^T u(t, x(t))dx(t)$ is called an Ito integral. See Doob[6] or Ito[5] for its properties. [It should be remembered that $\dot{x}(t)$ has infinite variance and is extremely discontinuous, and therefore an Ito integral is not quite the same as an ordinary Stieltjes integral.] In the above works it is shown that if $x(t)$ is Brownian motion, which it may be considered in the above integral, then

$$U(T, x(T)) - U(0, x(0)) = \int_0^T U_t(t, x(t))dt + \int_0^T U_x(t, x(t))dx(t) + \frac{\lambda^2}{2}\int_0^T U_{xx}(t, x(t))dt.$$

Since $U_x = u(t, x)$ by hypothesis, we have

$$\int_0^T u(t, x(t))dx(t) = -\int_0^T U_t(t, x(t))dt - \frac{\lambda^2}{2}\int_0^T U_{xx}(t, x(t))dt + U(T, x(T)) - U(0, x(0)).$$

Making these substitutions, we get

$$E\{F(x(\cdot)) \mid x(0) = x\} = E^w\left\{F(x(\cdot)) \exp\left[-\left(\frac{1}{\lambda^2}\right)\int_0^T\left(U_t + \frac{\lambda^2}{2} U_{xx} + \tfrac{1}{2}(U_x)^2\right)dt + U(T, x(T))/\lambda^2 - U(0, x(0))/\lambda^2\right] \mid x(0) = x\right\}.$$

By Lemma 4.1, the last expression is

$$E^w\left\{F(x(\cdot)) \exp\left[\left(\frac{1}{\lambda^2}\right)\int_0^T V(t, x(t))dt + U(T, x(T)) - U(0, x(0))\right] \mid x(0) = x\right\},$$

as was to be shown.

Since $\frac{1}{2}\dot{x}^2(t) - V(t, x(t))$ is actually the Lagrangian for the problem $\ddot{x} = -V_x$, and since $\int_0^T(\frac{1}{2}\dot{x}^2 - V)dt - U(T, x(T)) + U(0, x(0))$ can be defined as its action (see Feynman and Hibbs[17]), (4.11) can be rewritten as

$$E\{F(x(\cdot))\} = \int F(x(\cdot)) \exp[-(1/\lambda^2)\,(\text{action})]\mathfrak{D}x(t),$$

using Feynman and Hibbs integral notation.

As $\lambda \to 0$, then $1/\lambda^2 \to \infty$ and the above integral becomes singular, putting all its mass on the path $x(t)$ which minimizes the action, and $E\{F(x(\cdot))\}$ becomes just $F(y(\cdot))$, where $y(\cdot)$ minimizes the action. This is shown rigorously in Schilder.[18] As $\lambda \to 0$, it can be seen from (4.4) that the random motion in the gas is going away. Another proof is therefore given, for the simple case considered, of Hamilton's variational principle. The theory presented in this paper can most likely be generalized to other types of Lagrangians by using the theory developed by Varadhan[19] or Flemming.[12]

Corollary 4.1:

$$E\{F(x(\cdot))\} = \int_{-\infty}^{\infty} E^w\left\{F(x(\cdot)\exp\left[-\frac{1}{\lambda^2}\left(\int_0^T V(t,x(t))dt + U(T,x(T)) - U(0,x(0))\right)\right] \mid x(0) = x\right\}\rho(0,x)dx.$$

Proof: The proof follows directly from the definition of conditional expectation [that $x(0) = x$].

Let $U_{*x} = u_*$.

Theorem 4.2: Suppose u_* satisfies (4.8), that $|u_*| \leq K$, a constant, that (4.9) has a unique solution, and that F is an arbitrary integrable function of the paths $x_*(t)$. Then

$$E\{F(x(\cdot)) \mid x(T) = y\} = E^w\left\{F(x(\cdot))\exp\left[\frac{1}{\lambda^2}\left(\int_0^T V(t,x(t))dt - U_*(0,x(0))\right)\right] \mid x(T) = y\right\}\exp[U_*(T,y)/\lambda^2]. \tag{4.12}$$

Proof: The proof is carried out in the same manner as Theorem 4.1.

Corollary 4.2:

$$\exp[U_*(T,y)/\lambda^2] = \left(E^w\left\{\exp\left[(1/\lambda^2)\left(\int_0^T V(t,x(t))dt - U_*(0,x(0))\right)\right] \times \mid x(T) = y\right\}\right)^{-1}.$$

Proof: Corollary 4.2 is proved by setting the F of Theorem 4.2 equal to one, and by transposing $\exp[1/\lambda^2)U_*(T,y)$ to the left-hand side of (4.12).

Theorem 4.3: Suppose F and u_* satisfy the hypothesis of Theorem 4.2; then

$$E\{F(x(\cdot)) \mid x(T) = y\} = \frac{E^w\left\{F(x(\cdot))\exp\left[(1/\lambda^2)\left(\int_0^T V(t,x(t))dt - U_*(0,x(0))\right)\right] \mid x(T) = y\right\}}{E^w\left\{\exp\left[(1/\lambda^2)\left(\int_0^T V(t,x(t))dt - U_*(0,x(0))\right)\right] \mid x(T) = y\right\}} \tag{4.13}$$

Proof: The proof is immediate from Theorem 4.2 and Corollary 4.2.

Theorems 4.1 and 4.3 show that any macroscopic parameter of the gas considered can be expressed as a Wiener integral or as the ratio of Wiener integrals that depend only on V, $U(0,x)$, and $U(T,x)$, or on V and $U_*(0,x)$. Schilder[15] shows how this type of Wiener integral can be expanded in a power series in λ. Nelson[20] shows that u, u_*, and ρ are related by the interesting formula

$$u_*(t,x) = u(t,x) - \lambda^2\rho_x(t,x)/\rho(t,x), \tag{4.14}$$

where ρ again is the density. This formula can also be derived from Corollary 4.2 of this paper. If (4.14) is integrated with respect to x and raised to the eth power, it is found that

$$\exp[U_*(t,x)/\lambda^2]\rho(t,x) = \exp[U(t,x)/\lambda^2],$$

whence the term $\exp[-U(0,x(0))/\lambda^2]$ of (4.11) can be replaced by $\exp[-U_*(0,x(0))/\lambda^2]\rho(0,x(0))$.

It will now be shown that $u_*(T,y)$ can be expressed as the ratio of two Wiener integrals. This theorem is due to Professor M. Donsker. See Varadhan.[19]

Theorem 4.4: Suppose u_* satisfies the conditions of Theorem 4.2 and that $U_*(0,x) = \varphi(x)$. Then

$$u_*(T,y) = \frac{E^w\left[\left(-\int_0^T V_x(t,x(t))dt + \varphi_x(x(0))\right)\exp\left[(1/\lambda^2)\left(\int_0^T V(t,x(t))dt - \varphi(x(0))\right)\right] \mid x(T) = y\right]}{E^w\left\{\exp\left[(1/\lambda^2)\left(\int_0^T V(t,x(t))dt - \varphi(x(0))\right)\right] \mid x(T) = y\right\}}.$$

Proof: From Corollary 4.2, it follows that

$$u_*(T,y) = -\frac{d}{dy}\left(\ln E^w\left\{\exp\left[\frac{1}{\lambda^2}\left(\int_0^T V(t,x(t))dt - \varphi(x(0))\right)\right] \mid x(T) = y\right\}\right);$$

the proof is obtained by changing variables in the above Wiener integral by letting $x(t) = y(t) + y$, by differentiating with respect to y, and then by changing back to the $x(t)$ variable of integration.

The partial differential equation (4.8) is known as Burger's equation and is used in the study of turbulence. See Meecham.[21]

If the $F(x(\cdot))$ of Theorem 4.3 is $\lim_{h\to 0}\{[x(t+h) - x(t)]/h\}$, then the left-hand side of (4.13) is by definition $u_*(T,y)$. That the right-hand side of (4.13) satisfies Burger's equation is shown in Schilder.[13]

By construction it follows as $\Delta \to 0$ that the variance of $\dot{x}(t)$ tends to infinity. Thus what has been called pressure in Sec. 3 tends to infinity, and the equation of Sec. 3, $u_t + uu_x = -P_x/\rho + f$, becomes meaningless as $\Delta \to 0$. On the other hand, it has been in reality assumed that the u defined in the limit as $\Delta \to 0$ satisfies $u_t + uu_x + (\lambda^2/2)u_{xx} = f$. The explanation for this phenomenon is that the t derivative limit of u_t and the Δ limit cannot be interchanged. If one is going to consider a gas with an infinite number of collisions in every finite interval, then the natural order of limits is clearly to let the number of collisions tend to infinity and then take the time derivative of velocity. This is the order considered in this paper.

Doob[6] shows that $\rho(t, x)$ satisfies

$$\rho_t + (u\rho)_x = \tfrac{1}{2}\lambda^2\rho_{xx},$$

that is, the Fokker–Planck equation. This equation and the equation $u_t + uu_x + \lambda^2/2\, u_{xx} = f$ coupled with it, are the continuity and motion equations for the gas considered here.

Nelson[1] has developed a different theory for the kinematics of stochastic motion.

Kirkwood[22] has shown that random motion for "liquids and other condensed systems" can be modeled on an Ito equation of the form $\ddot{x}(t) = f(t, x(t)) + \beta\dot{x}(t) + \lambda\dot{z}(t)$ (see also Onsager and Machlup[23]), where, as before in this paper, f is external force, x is the position of the particle, and z is Brownian motion. It is easy to show, however, that a fluid modeled on this type of equation does not conserve energy (it is sufficient to let $f = 0$ and that its average velocity vector at a point does not satisfy the usual hydrodynamical equations. See Schilder[15] or DeGroot and Mazur.[24]

While the theory of this paper was developed under the assumption that f has derived from a potential, some of the results go through without this assumption if the paper transpositions are made.

A later paper will show how the concepts developed here can be applied to quantum mechanics.

---

* Present address: 381 Broad St., Newark, N.J. 07104.

1 E. Nelson, *Dynamical Theories of Brownian Motion* (Princeton U.P., Princeton, N.J., 1968), Chap 9.
2 J.H. Jeans, *The Dynamical Theory of Gases* (Dover, New York, 1954).
3 M. Kac, *Probability and Related Topics in the Physical Sciences* (Interscience, London, 1959), p. 172.
4 Reference 2, p. 114.
5 K. Ito, Mem. Am. Math. Soc. **4**, (1951).
6 J. L. Doob, *Stochastic Processes* (Wiley, New York, 1953), Chap. 6, Sec. 3.
7 This approach was suggested by the referee.
8 Reference 1, pp. 80, 95.
9 Reference 1, p. 103.
10 Reference 1, p. 104.
11 M. Wonham, "Lecture Notes on Stochastic Control," *Lecture Notes*, 67-2 (Brown University, 1967).
12 F. H. Fleming, J. Differential Eqs. 5, 515 (1969).
13 For one-dimensional motion this hypothesis can be weakened by using results found in M. Schilder, Pacific J. Math. **25**, 77 (1967).
14 See Doob, Ref. 6, Chap. 6, Sec. 3, for a definition and sufficient conditions that (4. 4) have a unique solution.
15 M. Schilder, J. Stat. Phys. **1**, 497 (1970).
16 I. V. Girsanov, Theory Probability Appl. **5**, 285 (1960).
17 R. T. Feynman and R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
18 M. Schilder, Trans. Am. Math. Soc. **125**, 63 (1966).
19 R. S. Varadhan, Commun. Pure Appl. Math. **19**, 118 (1966).
20 Reference 1, p. 105.
21 W. C. Meecham, J. Stat. Phys. 1, 211 (1969).
22 J. G. Kirkwood, J Chem. Phys. 14, 180 (1946).
23 L. Onsager and S. Machlup, Phys. Rev. **91**, 817 (1953).
24 S. R. DeGroot and P. Mazer, *Non-Equilibrium Thermodynamics* (North-Holland, Amsterdam, 1960), pp. 187–94.

Gårding Domains and Analytic Vectors for Quantum Fields

Gerhard C. Hegerfeldt

Institut für Theoretische Physik, Universität Göttingen, Göttingen, Germany

(Received 22 February 1971)

If one studies the canonical commutation relations (CCR's) of quantum field theory in the unitary Weyl form, one does not know if one can find a common dense domain for the field operators since their domain of definition depends on the test function. We consider here a general class of test function spaces including the spaces S and $\mathfrak{D}$ of Schwartz and the space $\mathfrak{U}_0 \simeq \mathbb{R}^{(\infty)}$ of all finite linear combinations of a countable basis. It is shown that there exists an invariant Gårding domain D on which all fields are defined and strongly continuous. D consists of analytic vectors for the fields. It turns out that the test function space can be enlarged by continuity. For irreducible or factor representations it becomes even a Hilbert space. The basic idea of the proof is the same as in the Schrödinger representation for one degree of freedom and very transparent. We simply use rapidly decreasing functions in "Q-space" and "P-space" as smoothing factors. That this can be done in the infinite case also is due to a new and interesting measure theoretic result derived here. As an application of our results, we mention that the renormalized fields (after removing the cutoff) of the Φ^{2n}_2 model of Glimm and Jaffe possess a Gårding domain for test functions in S or $\mathfrak{D}$ for each time.

1. INTRODUCTION

The canonical, equal time, commutation relations of quantum field theory (CCR's),

$$[\Phi(\mathbf{x}, t_0), \Pi(\mathbf{x}', t_0)] = i\delta^{(3)}(\mathbf{x} - \mathbf{x}'),$$
$$\Phi(\mathbf{x}, t_0), \Phi(\mathbf{x}', t_0) = [\Pi(\mathbf{x}, t_0), \Pi(\mathbf{x}', t_0)] = 0, \tag{1.1}$$

are usually studied in the Weyl form. Heuristically, it is obtained by first smearing the fields with real test functions,

$$\Phi(f) \equiv \Phi(f, t_0) = \int d^3\mathbf{x}\Phi(\mathbf{x}, t_0)f(\mathbf{x}),$$
$$\Pi(g) \equiv \Pi(g, t_0) = \int d^3\mathbf{x}\Pi(\mathbf{x}, t_0)g(\mathbf{x}) \tag{1.2}$$

so that

$$[\Phi(f), \Pi(g)] = i\int f(\mathbf{x})g(\mathbf{x})d^3x \equiv i(f, g). \tag{1.3}$$

Defining unitary operators

$$U(f) = e^{i\Phi(f)}, \qquad V(g) = e^{i\Pi(g)}, \tag{1.4}$$

one obtains, in a formal way, the Weyl commutation relations

$$U(f_1)U(f_2) = U(f_1 + f_2),$$
$$V(g_1)V(g_2) = V(g_1 + g_2), \tag{1.5}$$
$$V(g)U(f) = e^{i(f,g)}U(f)V(g).$$

Taking Eqs. (1. 5) as a starting point for a rigorous investigation of the representations of the CCR's, one avoids all sorts of domain questions connected with the unbounded operators $\Phi(f)$ and $\Pi(g)$. So one defines a representation of the CCR's to be a family $\{U(f), V(g), f \in \mathfrak{U}_\Phi; g \in \mathfrak{U}_\Pi\}$ of unitary operators satisfying Eqs. (1. 5), where $\mathfrak{U}_\Phi$ and $\mathfrak{U}_\Pi$ are real linear spaces and where (f, g) is a nondegenerate bilinear form on $\mathfrak{U}_\Phi \times \mathfrak{U}_\Pi$; furthermore, one demands as minimal con-

Doob[6] shows that $\rho(t, x)$ satisfies

$$\rho_t + (u\rho)_x = \tfrac{1}{2}\lambda^2\rho_{xx},$$

that is, the Fokker–Planck equation. This equation and the equation $u_t + uu_x + \lambda^2/2\, u_{xx} = f$ coupled with it, are the continuity and motion equations for the gas considered here.

Nelson[1] has developed a different theory for the kinematics of stochastic motion.

Kirkwood[22] has shown that random motion for "liquids and other condensed systems" can be modeled on an Ito equation of the form $\ddot{x}(t) = f(t, x(t)) + \beta\dot{x}(t) + \lambda\dot{z}(t)$ (see also Onsager and Machlup[23]), where, as before in this paper, f is external force, x is the position of the particle, and z is Brownian motion. It is easy to show, however, that a fluid modeled on this type of equation does not conserve energy (it is sufficient to let $f = 0$ and that its average velocity vector at a point does not satisfy the usual hydrodynamical equations. See Schilder[15] or DeGroot and Mazur.[24]

While the theory of this paper was developed under the assumption that f has derived from a potential, some of the results go through without this assumption if the paper transpositions are made.

A later paper will show how the concepts developed here can be applied to quantum mechanics.

---

* Present address: 381 Broad St., Newark, N.J. 07104.

1 E. Nelson, *Dynamical Theories of Brownian Motion* (Princeton U.P., Princeton, N.J., 1968), Chap 9.
2 J.H. Jeans, *The Dynamical Theory of Gases* (Dover, New York, 1954).
3 M. Kac, *Probability and Related Topics in the Physical Sciences* (Interscience, London, 1959), p. 172.
4 Reference 2, p. 114.
5 K. Ito, Mem. Am. Math. Soc. **4**, (1951).
6 J. L. Doob, *Stochastic Processes* (Wiley, New York, 1953), Chap. 6, Sec. 3.
7 This approach was suggested by the referee.
8 Reference 1, pp. 80, 95.
9 Reference 1, p. 103.
10 Reference 1, p. 104.
11 M. Wonham, "Lecture Notes on Stochastic Control," *Lecture Notes*, 67-2 (Brown University, 1967).
12 F. H. Fleming, J. Differential Eqs. 5, 515 (1969).
13 For one-dimensional motion this hypothesis can be weakened by using results found in M. Schilder, Pacific J. Math. **25**, 77 (1967).
14 See Doob, Ref. 6, Chap. 6, Sec. 3, for a definition and sufficient conditions that (4. 4) have a unique solution.
15 M. Schilder, J. Stat. Phys. **1**, 497 (1970).
16 I. V. Girsanov, Theory Probability Appl. **5**, 285 (1960).
17 R. T. Feynman and R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
18 M. Schilder, Trans. Am. Math. Soc. **125**, 63 (1966).
19 R. S. Varadhan, Commun. Pure Appl. Math. **19**, 118 (1966).
20 Reference 1, p. 105.
21 W. C. Meecham, J. Stat. Phys. 1, 211 (1969).
22 J. G. Kirkwood, J Chem. Phys. 14, 180 (1946).
23 L. Onsager and S. Machlup, Phys. Rev. **91**, 817 (1953).
24 S. R. DeGroot and P. Mazer, *Non-Equilibrium Thermodynamics* (North-Holland, Amsterdam, 1960), pp. 187–94.

Gårding Domains and Analytic Vectors for Quantum Fields

Gerhard C. Hegerfeldt

Institut für Theoretische Physik, Universität Göttingen, Göttingen, Germany

(Received 22 February 1971)

If one studies the canonical commutation relations (CCR's) of quantum field theory in the unitary Weyl form, one does not know if one can find a common dense domain for the field operators since their domain of definition depends on the test function. We consider here a general class of test function spaces including the spaces S and $\mathfrak{D}$ of Schwartz and the space $\mathfrak{U}_0 \simeq \mathbb{R}^{(\infty)}$ of all finite linear combinations of a countable basis. It is shown that there exists an invariant Gårding domain D on which all fields are defined and strongly continuous. D consists of analytic vectors for the fields. It turns out that the test function space can be enlarged by continuity. For irreducible or factor representations it becomes even a Hilbert space. The basic idea of the proof is the same as in the Schrödinger representation for one degree of freedom and very transparent. We simply use rapidly decreasing functions in "Q-space" and "P-space" as smoothing factors. That this can be done in the infinite case also is due to a new and interesting measure theoretic result derived here. As an application of our results, we mention that the renormalized fields (after removing the cutoff) of the Φ^{2n}_2 model of Glimm and Jaffe possess a Gårding domain for test functions in S or $\mathfrak{D}$ for each time.

1. INTRODUCTION

The canonical, equal time, commutation relations of quantum field theory (CCR's),

$$[\Phi(\mathbf{x}, t_0), \Pi(\mathbf{x}', t_0)] = i\delta^{(3)}(\mathbf{x} - \mathbf{x}'),$$
$$\Phi(\mathbf{x}, t_0), \Phi(\mathbf{x}', t_0) = [\Pi(\mathbf{x}, t_0), \Pi(\mathbf{x}', t_0)] = 0, \tag{1.1}$$

are usually studied in the Weyl form. Heuristically, it is obtained by first smearing the fields with real test functions,

$$\Phi(f) \equiv \Phi(f, t_0) = \int d^3\mathbf{x}\Phi(\mathbf{x}, t_0)f(\mathbf{x}),$$
$$\Pi(g) \equiv \Pi(g, t_0) = \int d^3\mathbf{x}\Pi(\mathbf{x}, t_0)g(\mathbf{x}) \tag{1.2}$$

so that

$$[\Phi(f), \Pi(g)] = i\int f(\mathbf{x})g(\mathbf{x})d^3x \equiv i(f, g). \tag{1.3}$$

Defining unitary operators

$$U(f) = e^{i\Phi(f)}, \qquad V(g) = e^{i\Pi(g)}, \tag{1.4}$$

one obtains, in a formal way, the Weyl commutation relations

$$U(f_1)U(f_2) = U(f_1 + f_2),$$
$$V(g_1)V(g_2) = V(g_1 + g_2), \tag{1.5}$$
$$V(g)U(f) = e^{i(f,g)}U(f)V(g).$$

Taking Eqs. (1. 5) as a starting point for a rigorous investigation of the representations of the CCR's, one avoids all sorts of domain questions connected with the unbounded operators $\Phi(f)$ and $\Pi(g)$. So one defines a representation of the CCR's to be a family $\{U(f), V(g), f \in \mathfrak{U}_\Phi; g \in \mathfrak{U}_\Pi\}$ of unitary operators satisfying Eqs. (1. 5), where $\mathfrak{U}_\Phi$ and $\mathfrak{U}_\Pi$ are real linear spaces and where (f, g) is a nondegenerate bilinear form on $\mathfrak{U}_\Phi \times \mathfrak{U}_\Pi$; furthermore, one demands as minimal con-

tinuity condition ray continuity, i.e., that $U(\lambda f)$ and $V(\lambda g)$ are weakly continuous in λ.

Ray continuity allows one to recover the smeared fields as infinitesimal generators by Stone's theorem, and the fields satisfy Eq. (1.3) on a dense domain. However, all domains depend on f and g, and one does not know if there is a *common* domain on which every $\Phi(f)$ and $\Pi(g)$ is defined. This is unfortunate since in applications it is often advantageous to work with the fields which are linear in f.

In the present paper, a common domain for the fields $\{\Phi(f)\}$ and $\{\Pi(g)\}$ is constructed under additional continuity requirements which are in general stronger than ray continuity. If the test function space were finite dimensional, this would follow immediately from a well-known theorem in group theory by Gårding.[1]

Following usage, we call a *Gårding domain* a domain on which all infinitesimal generators—in our case the fields—are defined and essentially self-adjoint and which is left invariant by the infinitesimal generators and the associated unitary groups. If A is an operator in a Hilbert (or Banach) space $\mathfrak{H}$, φ is called an analytic vector for A if $A^n\varphi$ is defined for each n and if, for some $t > 0$,

$$\sum_n \frac{t^n}{n!} \|A^n\varphi\| < \infty. \tag{1.6}$$

φ is called *entire for* A if Eq. (1.6) holds for all $t > 0$. Analytic vectors are very useful since on them one can consider power series in A which converge absolutely. Furthermore, a result by Nelson[2] states that a symmetric operator A is essentially self-adjoint on a domain D if D contains a dense set of analytic vectors.

In the following we take the same test function space for Φ and Π, $\mathcal{U} = \mathcal{U}_\Phi = \mathcal{U}_\Pi$ and, moreover, assume that $\mathcal{U}$ is a nuclear space such as S, the space of Schwartz, that $U(f)$ and $V(f)$ are weakly continuous in f for the topology of $\mathcal{U}$, and that (f,g) is also continuous. We note that the Weyl operators can then be regarded as a continuous representation of an *infinite-dimensional* nuclear Lie group in the sense of Gel'fand.[3]

In particular we have in mind the spaces S and $\mathcal{D}$ of Schwartz and the space $\mathcal{U}_0 \simeq \mathbb{R}^{(\infty)}$ which consists of all finite linear combinations of a countable basis and hence is isomorphic to the space $\mathbb{R}^{(\infty)}$ of all finite sequences. For $\mathcal{U}_0$ the continuity requirement reduces to ray continuity since $\mathcal{U}_0$, regarded as a union of increasing finite-dimensional subspaces, is a nuclear space (as a strict inductive limit). The spaces S, $\mathcal{D}$, and $\mathcal{U}_0$ are separable barreled nuclear spaces,[4] and only this last fact will be used.

The basic idea underlying our construction of a Gårding domain and analytic vectors is astonishingly simple and taken from the corresponding problem in the Schrödinger representation for one degree of freedom. There the obvious thing to do is to consider functions which decrease rapidly enough in x-space and p-space to control the growth of any power of x and p, e.g., one could consider hermite functions or something similar,

$$\varphi(x) = h(x)e^{-|x|^2}, \tag{1.7}$$

where $h(x)$ is a polynomial or a function not increasing too fast. The very same idea will be applied here, only that instead of the single norm $|\cdot|$ we will have to deal with a countable number of norms.

In case of the metrizable space S, this idea can be applied directly if one uses the form of representations of the CCR's given by Gel'fand[3] for cyclic $U(f)$ and by Hegerfeldt and Melsheimer[5] for the general case ("Q-space"; cf. Sec. 3). For spaces such as $\mathcal{D}$ and $\mathcal{U}_0$, which are not metrizable, an additional measure-theoretic property is needed. This result (Proposition 2.1), which seems to be new and quite interesting, is proved in Sec. 2. A Gårding domain is then constructed in Sec. 3 along the ideas outlined above.

It turns out that in doing so one obtains actually a stronger result which shows that the test function space can in general be chosen larger than the original $\mathcal{U}$. This is spelled out in Sec. 4.

At the end of Sec. 4 an interesting extension is made. It is shown that for *irreducible* or factor representations one can extend the test function space to a *Hilbert space*.

As an application of our results we note that for fixed time the *renormalized* fields of Glimm and Jaffe[6] in the Φ^{2n}_2 model (with the cutoff removed) possess a Gårding domain on which the fields are strongly continuous if the test functions are taken to lie in S or $\mathcal{D}$. The renormalized Weyl operators are continuous in S and $\mathcal{D}$ due to the locally Fock property of the limit state.

For infinitely many Q_i and P_i (i.e., for a test function space $\mathcal{U}_0$) similar results have already been obtained by Reed[7] using the realization of the CCR's of Gårding and Wightman.[8] In particular, it was shown that $\mathcal{U}_0$ can always be extended to a Hilbert space. Reed's results are recovered from ours as a special case in Sec. 4. Reed's original derivation could not be carried over to test function spaces like S or $\mathcal{D}$; this is an indication that the "Q-space" realization of the CCR's may be more advantageous.

2. SUPPORT PROPERTIES OF MEASURES ON INFINITE-DIMENSIONAL SPACES

The following auxiliary measure-theoretic results are of interest in themselves. The main result needed for the next section is Proposition 2.1. For a space like S it is an immediate consequence of Lemma 2.1, and the reader may omit the technical details.

Throughout, $\mathcal{U}$ denotes a locally convex vector space and $\mathcal{U}'$ its dual. Let $f \in \mathcal{U}$ and $F \in \mathcal{U}'$. We write (f, F) for the value of the linear functional F applied to f. We consider the σ-algebra $\mathfrak{B}(\mathcal{U}', \mathcal{U})$ of subsets of $\mathcal{U}'$ generated by all cylinder sets over finite-dimensional subspaces of $\mathcal{U}$. It is generated by all sets of the form

$$\mathcal{O}_{f,B} = \{F \in \mathcal{U}' : (f, F) \in B \subset \mathbb{R}\}, \tag{2.1}$$

where $f \in \mathcal{U}$ and $B \subset \mathbb{R}$ is an open or a Borel subset of the real line. Note that each f, regarded as a function on $\mathcal{U}'$, is measurable with respect to $\mathfrak{B}(\mathcal{U}', \mathcal{U})$.

The next lemma is a simple variant of Lemma 3.4 of Ref. 9.

Lemma 2.1: Let p be a continuous seminorm on a separable space $\mathcal{U}$. Then the space $\mathcal{U}'_p$ of all p-continuous functionals on $\mathcal{U}$ is a measurable subspace of $\mathcal{U}'$.

Proof: The fact that $\mathcal{U}'_p \subset \mathcal{U}'$ follows from the continuity of $p(f)$ by $|(f,F)| \leq c_F \cdot p(f)$ for $F \in \mathcal{U}'_p$. Putting $0/0 = 0$ and $a/0 = \infty$ for $a \neq 0$, we now define, for $F \in \mathcal{U}'$,

$$p'(F) \equiv \sup_f |(f,F)|\, p(f)^{-1}. \tag{2.2}$$

Then $F \in \mathcal{U}'_p$ if and only if $p'(F) < \infty$. If $\{f_r\}$ is a countable dense subset of $\mathcal{U}$, then, by continuity,

$$p'(F) = \sup_r |(f_r,F)|\, p(f_r)^{-1}.$$

Since each (f_r,F) is measurable, so is $p'(F)$, and hence the set $\{F: p'(F) < \infty\}$ is measurable. QED

Corollary 2.1: The dual norm $p'(F)$ in Eq. (2.2) is a measurable function on $\mathcal{U}'$.

In the next lemma, $\mathcal{U}$ is assumed to be a barreled space. This is only to insure that each convergent generalized sequence (net) $\{f_\alpha\}$ in $\mathcal{U}$ converges uniformly on weakly compact subsets K of $\mathcal{U}'$, i.e., that (f_α,F) is uniformly convergent for $F \in K$. This is an immediate consequence of the Banach–Steinhaus theorem,[10] and therefore a detailed definition is omitted.

Remark: Products, inductive limits, and quotient spaces of barreled spaces are barreled.[10] In particular all F-spaces (i.e., complete metrizable locally convex spaces) and strict inductive limits (unions) of F-spaces are barreled, e.g., S, $\mathcal{D}$, and $\mathbb{R}^{(\infty)} \simeq \mathcal{U}_0$. We also note that a nuclear F-space is separable,[11] as well as a strict inductive limit of such spaces.

Lemma 2.2: Let $\mathcal{U}$ be barreled and nuclear, and let K be a weakly compact subset of $\mathcal{U}'$. Define a seminorm on $\mathcal{U}$ by

$$q_K(f) = \sup_{F \in K} |(f,F)|. \tag{2.3}$$

Then this seminorm is continuous, and there is a separable continuous *Hilbertian* seminorm p on $\mathcal{U}$ and a constant c such that $q_K \leq cp$ and $K \subset \mathcal{U}'_p \subset \mathcal{U}'$.

Proof: Continuity of q_K follows from uniform convergence on K. Now, the nuclear topology is given by a basis family of separable Hilbertian seminorms on $\mathcal{U}$.[12] Hence one of these norms, p say, satisfies $q_K \leq cp$ for some $c > 0$. It follows that each q_K-continuous linear functional on $\mathcal{U}$ is also p-continuous, and so $K \subset \mathcal{U}'_p \subset \mathcal{U}'$. QED

Remark: We note that the dual norm $p'(F)$ as defined in Eq. (2.2) is a separable Hilbertian norm on $\mathcal{U}'_p$ since p induces a Hilbertian norm on $\mathcal{U}/\mathfrak{N}$ where $\mathfrak{N}$ is the null space of p. If the nuclearity assumption is dropped, then p need not be Hilbertian, which will be needed in Sec. 3. Nuclearity, however, is crucial for the next lemma.

Lemma 2.3: Let $\mathcal{U}$ be nuclear and μ be a normed positive measure on $\mathcal{U}'$. For each $\epsilon > 0$ there is a weakly compact set $K \subset \mathcal{U}'$ with outer measure $\mu^*(K) \geq 1 - \epsilon$.

Proof: Since $\mathcal{U}$ is nuclear, there is[13] a weakly compact set $K \subset \mathcal{U}'$ such that for every cylinder set Z containing K one has $\mu(Z) \geq 1 - \epsilon$. Now $\mu^*(K) = \inf\{\sum_1^\infty \mu(Z_\nu)\}$, where Z_ν are cylinder sets with $K \subset \cup Z_\nu$. For each $\eta > 0$ and each Z_ν there is an open cylinder set $\mathcal{O}_\nu$ with base in the same subspace of $\mathcal{U}$ such that $Z_\nu \subset \mathcal{O}_\nu$ and $\mu(\mathcal{O}_\nu) \leq \mu(Z_\nu) + \eta 2^{-\nu}$, by the regularity of normed measures on the Borel sets of finite-dimensional spaces. Since $K \subset \cup_1^\infty \mathcal{O}_\nu$, there is a $\nu_0 < \infty$ such that $K \subset \cup_1^{\nu_0} \mathcal{O}_\nu$, by compactness. Since a finite union of cylinder sets is again a cylinder set, one has

$$1 - \epsilon \leq \mu(\bigcup_1^{\nu_0} \mathcal{O}_\nu) \leq \sum_1^\infty \mu(\mathcal{O}_\nu) \leq \sum_1^\infty \mu(Z_\nu) + \eta$$

for each η. This implies $\mu^*(K) \geq 1 - \epsilon$. QED

Now we can prove the main objective of this section, namely that a finite (or σ-finite) measure on $\mathcal{U}'$ is concentrated on a countable union of normed subspaces $\mathcal{U}'_{p_n}$ of $\mathcal{U}'$.

Proposition 2.1. Let $\mathcal{U}$ be a separable barreled nuclear space, such as S, $\mathcal{D}$, or $\mathcal{U}_0 \simeq \mathbb{R}^{(\infty)}$, and let μ be a finite positive measure on $\mathfrak{B}(\mathcal{U}', \mathcal{U})$. Then there exists a sequence $\{p_n\}$ of separable continuous Hilbertian seminorms on $\mathcal{U}$, $p_1 \leq p_2 \leq \ldots$, with associated dual spaces $\mathcal{U}'_{p_n}$ such that

(i) p'_n as defined by Eq. (2.2) is a measurable nonnegative extended real-valued function on $\mathcal{U}'$, and $p'_1 \geq p'_2 \geq \cdots$;

(ii) each $\mathcal{U}'_{p_n}$ is a measurable subspace of $\mathcal{U}'$. $\mathcal{U}'_{p_1} \subset \mathcal{U}'_{p_2} \subset \cdots$ $\mathcal{U}'_{p_n} = \{F \in \mathcal{U}': p'_n(F) < \infty\}$, and p'_n restricted to $\mathcal{U}'_{p_n}$ is the dual (separable Hilbertian) norm to p_n;

(iii) $\mu(\bigcup_n \mathcal{U}'_{p_n}) = \mu(\mathcal{U}')$.

Proof: We can assume $\mu(\mathcal{U}') = 1$. For each positive integer n there exists a weakly compact subset K_n of $\mathcal{U}'$ with outer measure $\mu^*(K_n) \geq 1 - 2^{-n}$. By Lemma 2.2, there is a separable continuous Hilbertian seminorm p_n on $\mathcal{U}$ such that $K_n \subset \mathcal{U}'_{p_n} \subset \mathcal{U}'$. By Lemma 2.1, $\mathcal{U}'_{p_n}$ is measurable and so $\mu(\mathcal{U}'_{p_n}) \geq 1 - 2^{-n}$. By induction, one can choose p_n such that $p_1 \leq p_2 \leq \ldots$. Then one has $p'_1 \geq p'_2 \geq \ldots$ and $\mathcal{U}'_{p_1} \subset \mathcal{U}'_{p_2} \subset \ldots$, and $\mu(\cup_n \mathcal{U}'_{p_n}) = 1$. The remaining statements then follow from the preceding lemmas. QED

For S or for a countably normed nuclear space of Gel'fand and Vilenkin[3] the proposition follows directly from Lemma 2.1 since in this case $\mathcal{U}' = \cup_n \mathcal{U}'_{p_n}$ for suitable Hilbertian norms.

Let $\mathcal{W}$ be a subspace of $\mathcal{U}'$. A measure μ is called $\mathcal{W}$-quasi-invariant if $\mu(A) = 0$ implies $\mu(A + g) = 0$ and conversely, for all $g \in \mathcal{W}$ and all $A \in \mathfrak{B}(\mathcal{U}', \mathcal{U})$. μ is called $\mathcal{W}$-ergodic if there is no nontrivial $\mathcal{W}$-quasi-invariant measure which is absolutely continuous with respect to μ. In particular, if $\mathcal{U}$ has been embedded in $\mathcal{U}'$ by means of a nondegenerate bilinear form on $\mathcal{U}$, one may consider $\mathcal{W} = \mathcal{U}$. We have the following additional result.

Corollary 2.2: Let μ be as in Proposition 2.1 and $\mathcal{W}$-quasi-invariant. If $\mu(\mathcal{U}'_{p_n}) \neq 0$, then $\mathcal{W} \subset \mathcal{U}'_{p_n}$.

Proof: Assume $g \notin \mathcal{U}'_{p_n}, g \in \mathcal{W}$. Then $(\mathcal{U}'_{p_n} + \lambda g) \cap (\mathcal{U}'_{p_n} + \lambda' g) = \emptyset$ for $\lambda \neq \lambda'$ since otherwise $(\lambda - \lambda')g \in \mathcal{U}'_{p_n}$ and so $g \in \mathcal{U}'_{p_n}$. By quasi-invariance one has $\mu(\mathcal{U}'_{p_n} + \lambda g) > 0$, and so $\mathcal{U}'$ contains uncountably many disjoint sets of positive measure. This contradicts $\mu(\mathcal{U}') < \infty$. QED

In passing we note the following simple corollary.

Corollary 2.5 (zero-one-property): If μ is $\mathcal{W}$-ergodic and finite, then every measurable linear subspace $\mathcal{L} \subset \mathcal{U}'$ has either zero or full measure.

Proof: If $\mathcal{W} \subset \mathcal{L}$, then $\mathcal{L}$ is invariant under $\mathcal{W}$, and so $\mu(\mathcal{L}) = \mu(\mathcal{U}')$ or 0. If $g \in \mathcal{W}$ and $g \notin \mathcal{L}$, then $\mu(\mathcal{L}) = 0$, by the preceding proof. QED

3. THE MAIN THEOREM

Every representation of the CCR's in a separable Hilbert space $\mathfrak{H}$ can be realized by means of a direct integral with a quasi-invariant measure on the algebraic dual of the test function space.[5] If one deals with a continuous representation of the CCR's over a nuclear test function space $\mathcal{U}$, one can perform the construction with a measure on $\mathcal{U}'$, or more precisely on the σ-algebra $\mathfrak{B}(\mathcal{U}', \mathcal{U})$. Since the bilinear form (f, g) is nondegenerate and continuous, there is a natural embedding of $\mathcal{U}$ in $\mathcal{U}'$ so that we can assume $\mathcal{U}$ to be a subspace of $\mathcal{U}'$.

There is a $\mathcal{U}$-quasi-invariant measure μ on $\mathcal{U}'$ with $\mu(\mathcal{U}') = 1$ and a direct integral decomposition of $\mathfrak{H}$,

$$\mathfrak{H} = \int_{\mathcal{U}'}^{\oplus} \mathfrak{H}(F)\, d\mu(F) \tag{3.1}$$

such that $U(f)$ becomes multiplication by $e^{i(f,F)}$,

$$(U(f)\varphi)(F) = e^{i(f,F)}\varphi(F) \tag{3.2}$$

and

$$(V(g)\varphi)(F) = \left(\frac{d\mu(F+g)}{d\mu(F)}\right)^{1/2} A_g(F)\varphi(F+g), \tag{3.3}$$

where $A_g(F)$ is a unitary map from $\mathfrak{H}(F+g)$ onto $\mathfrak{H}(F)$ and satisfies, for μ-almost all F,

$$A_g(F)A_{g'}(F+g) = A_{g+g'}(F) \tag{3.4}$$

Since $U(f)$ and $V(f)$ are, up to a sign, on equal footing, a similar realization holds in which $V(g)$ becomes multiplication by $e^{i(g,F)}$. Indeed, if we put $\tilde{U}(f) = V(f)$ and $\tilde{V}(g) = U(-g)$, then these operators fulfill the Weyl relations, and hence there is a $\mathcal{U}$-quasi-invariant measure $\tilde{\mu}$, Hilbert spaces $\tilde{\mathfrak{H}}(F)$ and operators $\tilde{A}_g(F)$ which yield a realization for $\tilde{U}(f)$ and $\tilde{V}(g)$ analogous to Eqs. (3.1)–(3.4).

The realization which diagonalizes $U(f)$ may be called "Q-space" realization, in analogy to the quantum mechanical situation, and "P-space" realization diagonalizes $V(g)$. In Q-space, the action of the field operators $\Phi(f)$ is given by multiplication by (f, F) on vectors in the domain of $\Phi(f)$,

$$(\Phi(f)\varphi)(F) = (f, F)\varphi(F). \tag{3.5}$$

The same holds for $\Pi(g)$ in P-space.

In the following lemmas we exploit the Hilbertian norms $\{p'_n\}$ of Proposition 2.1. Let $(F_1, F_2)_n$ be the scalar product on $\mathcal{U}'_{p_n}$ such that

$$p'_n(F)^2 = (F, F)_n, \qquad F \in \mathcal{U}'_{p_n}. \tag{3.6}$$

Note that since the scalar product can be expressed by p'_n, $(F_1, F)_n$ is a measurable function on $\mathcal{U}'_{p_n}$ in F for fixed F_1.

Now let D_1 be a dense subset of $\mathfrak{H}$, and let D_2 be the linear span of vectors obtained from D_1 by multiplication with smoothing factors of exponential decrease,

$$D_2\colon \varphi(F) = \left\{\prod_1^M (f_k, F)\right\}\left\{\prod_1^N (g_k, F)_n\right\} e^{-p'_n(F)^2}\,\hat{\varphi}(F), \tag{3.7}$$

and finite linear combinations of such vectors; here $\hat{\varphi} \in D_1$ and $f_k \in \mathcal{U}, g_k \in \mathcal{U}, n = 1, 2, \ldots$, and $M, N = 0, 1, 2, \ldots$, where for $M = 0$ or $N = 0$ the corresponding factors are to be omitted. The factors in curly brackets are only inserted to obtain invariance properties of D_2. Note that $\mathcal{U} \subset \mathcal{U}'_{p_n}$, by Corollary 2.2, so that $(g, F)_n$ is defined for $g \in \mathcal{U}$ and $F \in \mathcal{U}'_{p_n}$. If $F \notin \mathcal{U}'_{p_n}$ then $e^{-p'_n(F)^2} = 0$ and we can define $\varphi(F) = 0$.

Lemma 3.1: Let D_1 be a dense subset of $\mathfrak{H}$ and let D_2 be defined as in Eq. (3.7). Then D_2 is dense in $\mathfrak{H}$, it consists of analytic, or rather entire, vectors for each $\Phi(f)$, and is invariant under $\{\Phi(f), f \in \mathcal{U}\}$.

Proof: Note that, for $F \in \mathcal{U}'_{p_n}$, one has $|(f, F)| \leq p_n(f)p'_n(F)$ and $|(g, F)_n| \leq p'_n(g)p'_n(F)$. Thus, by elementary calculus,

$$\left\{\prod_{k=1}^{\nu_1} |(f_k, F)|\right\}\left\{\prod_{j=\nu_1+1}^{\nu} |(f_j, F)_n|\right\} e^{-\alpha p'_n(F)^2} \leq \left\{\prod_{k=1}^{\nu_1} p_n(f_k)\right\}\left\{\prod_{j=\nu_1+1}^{\nu} p'_n(f_j)\right\}\alpha^{-\nu/2}\left(\frac{\nu}{2}\right)^{\nu/2} e^{-\nu/2} \tag{3.8}$$

for $\alpha > 0$. By Proposition 2.1, $\varphi(F)$ in Eq. (3.7) is measurable and, by Eq. (3.8), square-integrable and thus in $\mathfrak{H}$. Also $\Phi(f)D_2 \subset D_2$ for $f \in \mathcal{U}$. From Eq. (3.8) one finds, with a suitable constant c,

$$\|\Phi(f)^\nu \hat{\varphi}\| \leq \|\hat{\varphi}\|\, C^\nu\, \nu^{\nu/2}, \tag{3.9}$$

and hence

$$\sum \frac{t^\nu}{\nu!}\, \|\Phi(f)^\nu \varphi\| < \infty$$

for all $t > 0$, by Stirling's formula. By the triangle inequality, the sum of two analytic vectors is analytic.

Finally, to show that D_2 is dense in $\mathfrak{H}$, assume that $\langle \psi_0, \varphi \rangle = 0$ for all φ in Eq. (3.7) with $M, N = 0$. Let $\{\hat{\varphi}_\nu\}$ be a sequence in D_1 with $\psi_0 = \text{s-lim}\ \hat{\varphi}_\nu$. Putting $\varphi_\nu(F) = e^{-p'_n(F)^2}\,\hat{\varphi}_\nu(F)$, one then has

$$0 = \langle \psi_0, \varphi_\nu \rangle \to \int \|\psi_0(F)\|^2\, e^{-p'_n(F)^2}\, d\mu = 0.$$

Hence $\psi_0(F) = 0$ a.e. on $\mathcal{U}'_{p_n}$ for each n, and thus on $\mathcal{U}'$, by Proposition 2.1. QED

Corollary 3.1: For $\varphi \in D_2$, $\Phi(f)\varphi$ is strongly con-

tinuous in f for the nuclear topology of $\mathcal{U}$ and for the topology generated by the seminorms $\{p_n\}$.

Proof: $f_\alpha \to f_0\ (\mathcal{U})$ implies $p_n(f_\alpha - f_0) \to 0$ for all n. Let φ be as in Eq. (3.6). Then for suitable n,

$$\|\Phi(f_\alpha) - \Phi(f_0)\|^2 \leq p_n(f_\alpha - f_0)^2 \int p_n'(F)^2 \|\varphi(F)\|^2 d\mu \to 0.$$

QED

It follows by duality in P-space that $\{\Pi(g); g \in \mathcal{U}\}$ possesses also a dense invariant subspace of analytic vectors, with analogous properties as D_2 above. Hence we may take D_1 in the previous lemma to be such a set, and we shall take in particular the linear span of vectors which, in P-space, look similar as in Eq. (3.7),

$$D_1\colon\ \tilde{\hat{\varphi}}(F) = \left\{\prod_{k=1}^{M} (g_K, F)\right\} e^{-\tilde{p}_n'(F)^2} \tilde{\psi}(F) \qquad (3.10)$$

where the tilde denotes P-space realization; we take $\psi \in \mathfrak{H}$ and $\tilde{p}_n'$ to be the norms of Proposition 2.1 belonging to the measure $\tilde{\mu}$ of the P-space realization. Then one has estimates for $\|\Pi(g)^\nu \hat{\varphi}\|$ similar to those for $\|\Phi(f)^\nu \varphi\|$ above.

Lemma 3.2: Let D_2 be defined as in Eq. (3.7), with D_1 given by Eq. (3.10). Then D_2 consist of entire vectors for $\{\Phi(f), \Pi(f); f \in \mathcal{U}\}$ and is invariant under these operators. Furthermore, for $\varphi \in D_2$, $\Phi(f)\varphi$, and $\Pi(f)\varphi$ are strongly continuous in f for the nuclear topology on $\mathcal{U}$ and for the topology generated by the seminorms p_n or $\tilde{p}_n$, respectively.

Proof: Lemma 3.1 and Corollary 3.2 apply to $\Phi(f)$. To deal with $\Pi(g)$, we write $\rho(F)$ for the multiplicative factors in Eq. (3.7) such that

$$\varphi(F) = \rho(F)\hat{\varphi}(F). \qquad (3.11)$$

From the Q-space realization of $V(g)$, Eq. (3.3), we have

$$\begin{aligned} &-i\tau^{-1}(\{V(\tau g) - \mathbb{1}\}\varphi)(F) \\ &\quad = -i\tau^{-1}\{\rho(F + \tau g) - \rho(F)\}(V(\tau g)\hat{\varphi})(F) \\ &\quad + \rho(F)(-i\tau^{-1})(\{V(\tau g) - \mathbb{1}\}\hat{\varphi})(F). \end{aligned} \qquad (3.12)$$

By definition of $\Pi(g)$, the last term converges strongly to $\rho(F)(\Pi(g)\hat{\varphi})(F)$ since $\rho(F)$ is bounded. To deal with the first term, the *Hilbertian* nature of the norms $\{p_n'\}$ is essential. Using the scalar product $(F_1, F_2)_n$, we see that $\rho(F + \tau g)$ is differentiable in τ, and the mean value theorem of calculus shows that the first term on the right-hand side of Eq. (3.12) converges strongly.

Thus

$$\begin{aligned} (\Pi(g)\varphi)(F) = &-i\frac{\partial}{\partial\tau}\rho(F + \tau g)\Big|_{\tau=0} \hat{\varphi}(F) \\ &+ \rho(F)(\Pi(g)\hat{\varphi})(F). \end{aligned} \qquad (3.13)$$

The differentiation of the exponential brings down a factor $(g, F)_n$. Hence, by Eq. (3.7), $\Pi(g)\varphi$ is again in D_2 and depends continuously on g since both terms in Eq. (3.13) do.

Using the Leibniz rule and the estimate for $\|\Pi(g)^\nu \hat{\varphi}\|$ analogous to Eq. (3.9), one finds after some calculation

$$\|\Pi(g)^\nu \varphi\| \leq C^\nu \nu^{\nu/2} \qquad (3.14)$$

for some constant c. Hence φ is entire for $\Pi(g)$. QED

Our main result on Gårding domains and analytic vectors is now an easy consequence of the forgoing lemma.

Theorem: Let $\{U(f), V(g)\}$ be a continuous representation of the CCR's with a separable barreled nuclear test function space $\mathcal{U}$, such as S, $\mathfrak{D}$, or $\mathcal{U}_0$, in a Hilbert space $\mathfrak{H}$. Then there exists a dense domain $D \subset \mathfrak{H}$, a Gårding domain, with the following properties:

(i) The fields $\{\Phi(f), \Pi(g)\}$ are essentially self-adjoint on D, and D is invariant under $\{U(f), V(g)\}$ as well as under all polynomials in the fields.

(ii) On D, any product of field operators $\{\Phi(f_i)\}$ and $\{\Pi(g_j)\}$, in any order, is jointly strongly continuous in the f_i and g_j; in particular, any matrix element of the form

$$\langle\varphi, \ldots \Phi(f_i) \ldots \Pi(g_j) \ldots \psi\rangle \qquad (3.15)$$

is a jointly continuous function if $\varphi, \psi \in D$.

(iii) D can be chosen to consist of analytic and even entire vectors for the fields $\{\Phi(f), \Pi(g)\}$.

Proof: Since $\mathcal{U}$ is separable, the representation is a direct sum of representations in separable Hilbert spaces. So we can assume $\mathfrak{H}$ as separable. Let D_2 be the domain of Lemma 3.3. We show that D defined as the linear span of $\{U(f)V(g)\}D_2$ has the required properties, where f and g run through $\mathcal{U}$. The procedure is standard.

From the definition of the fields as strong derivatives of the Weyl operators it follows that

$$\Pi(g')U(f)V(g)\varphi = U(f)V(g)\Pi(g)\varphi + (f, g')U(f)V(g)\varphi \qquad (3.16)$$

and a similar expression for $\Phi(f')$. Hence the fields are defined on D and leave it invariant since D_2 is invariant. Clearly D is also invariant under $U(f)V(g)$.

Since $\Pi(g')\varphi$ is strongly continuous in g' for $\varphi \in D_2$, so is $\Pi(g')U(f)V(g)\varphi$, by Eq. (3.15). The same holds for Π replaced by Φ. If one has a product of field operators applied to $U(f)V(g)\varphi$ the resulting vector is a continuous function in each field separately because each field can be brought to the left and the additional terms originating from the CCR's are continuous. By the kernel theorem for nuclear spaces this implies joint continuity.

Turning to analyticity, we note that for $\varphi \in D_2$ one obtains from Eq. (3.15)

$$\Pi(g')^\nu U(f)V(g)\varphi = U(f)V(g)\{\Pi(g') + (f, g')\}^\nu \varphi. \qquad (3.17)$$

From

$$\begin{aligned} \|\{\Pi(g') + (f, g')\}^\nu \varphi\| &\leq \|\{|\Pi(g')| + |(f, g')|\}^\nu \varphi\| \\ &\leq 2^\nu \|\{|\Pi(g')|^\nu + |(f, g')|^\nu\}\varphi\| \\ &\leq 2^\nu \|\Pi(g')^\nu \varphi\| + 2^\nu |(f, g')|^\nu \|\varphi\| \end{aligned}$$

it then follows that $U(f)V(g)\varphi$, and thus every vector of D, is entire for $\Pi(g')$. The same applies to $\Phi(f')$. By Nelson's theorem, the fields are essentially self-adjoint on D. QED

Remarks: The domain D constructed above is by no means the largest Gårding domain. One obtains a larger domain if one uses smoothing factors $h(p'(F)^2)$, with $h \in S(\mathbb{R}^1)$, and forms the union of the resulting domains for all $h \in S$ and for all continuous seminorms. For specific applications it may be necessary to enlarge D in this way. For example, in the Fock representation the vacuum Ω need not lie in D, but it does lie in the so enlarged domain if $\Omega \to \varphi_0(F) \equiv 1$ and if μ is the Gaussian measure, for then there are nontrivial functions of the form $\exp\{p'(F)\}^2$, with a suitable norm p on $\mathfrak{U}$, which are integrable; hence one obtains $\varphi_0(F) \equiv 1$ by the smoothing operation.

For possible applications we note that $\Phi(f)^2 + \Pi(f)^2$ is essentially self-adjoint on D_2; in fact, a direct computation shows that the vectors in D_2 are analytic. To treat $\Phi(f)^m$ and $\Pi(g)^m$, one can use smoothing factors of faster decrease.

It may be worthwhile to point out the role of the norms $\{p_n\}$ and of Proposition 2.1. This result is used in the construction mainly to ensure that sufficiently many factors $\exp\{-p'(F)^2\}$ do not vanish, μ almost everywhere. It is at least conceivable that for more general spaces the sets $\mathfrak{U}'_p = \{F \in \mathfrak{U}'; p'(F) < \infty\}$ have all measure zero for every continuous seminorm p on $\mathfrak{U}$.

4. EXTENSION AND DISCUSSION OF RESULTS

In the last section, we have proved slightly more than the main theorem. It is an easy consequence of the previous lemmas that the representation of the CCR's can be extended to a larger test function space and that the extended fields are still essentially self-adjoint on the above Gårding domain D.

To be more precise, the seminorms $\{p_n\}$ and $\{\tilde{p}_n\}$ are actually norms since $\mathfrak{U} \subset \mathfrak{U}'_{p_n}$ and thus $|(f,g)| \leq p_n(f)p'_n(g) = 0$ for all g implies $f = 0$. Hence we can define metrics d_1 and d_2 on $\mathfrak{U}$ by

$$d_1(f) = \sum_n 2^{-n} \frac{p_n(f)}{1 + p_n(f)} \tag{4.1}$$

and d_2 with p_n replaced by $\tilde{p}_n$. The topology on $\mathfrak{U}$ induced by d_1 is just the one generated by the norms and analogously for d_2. We also consider the metric

$$d = d_1 + d_2. \tag{4.2}$$

Corollary 4.1: Let the assumptions be as in the main theorem. Then the representation can be extended by continuity to the completion $\overline{\mathfrak{U}}(d)$ of $\mathfrak{U}$ in the metric d. The extended representation is strongly continuous in the metric d. There is a Gårding domain D for the fields satisfying (i)–(iii) of the above Theorem, with $\mathfrak{U}$ replaced by $\overline{\mathfrak{U}}(d)$.

Proof: Convergence of a sequence $\{f_\nu\}$ in d implies convergence in the norms p_n and $\tilde{p}_n$. By Proposition 2.1, (f_ν, F) then converges, μ and $\tilde{\mu}$ almost everywhere. Hence $U(f_\nu)$ and $V(f_\nu)$ are strongly continuous, by Lebesgue's bounded convergence. It follows[14] that the representation can be extended to $\overline{\mathfrak{U}}(d)$. The action of $U(\bar{f})$ for $\bar{f} \in \overline{\mathfrak{U}}(d)$ is given by multiplication by $e^{i(\bar{f},F)}$, where $(\bar{f}, F) = \lim (f_\nu, F)$ for $F \in \cup_n \mathfrak{U}'_{p_n}$. Note that the elements of $\cup_n \mathfrak{U}'_{p_n}$ are d-continuous functionals on $\overline{\mathfrak{U}}(d)$.

The action of $\Phi(\bar{f})$ in Q-space is given, as before, by multiplication by $(\bar{f}, F)$ with $F \in \cup_n \mathfrak{U}'_{p_n}$, and similarly for $\Pi(\bar{g})$ in P-space. One can now apply directly the arguments of the last section to obtain a Gårding domain as in (i)–(iii) on which, to begin with, products of fields are separately strongly continuous in the test functions. Since on complete metric spaces multilinear functionals are jointly continuous if they are separately continuous,[15] this implies joint weak continuity and thus also joint strong continuity. QED

We remark that the norms can be chosen such that $p_n = \tilde{p}_n$ and such that $\overline{\mathfrak{U}}(d)$ becomes a nuclear metric space. For $\mathfrak{U} = S$, the above extension may be trivial since S is already countably normed. For $\mathfrak{D}$, however, this is not true. Further below we will also obtain an important extension of S if the representation is irreducible.

The result of Reed[7] follows from the last corollary. Reed considers as test function space the finite linear span of an orthonormal set, i.e., $\mathfrak{U} = \mathfrak{U}_0$. It is easy to see that there is a space $\mathfrak{U}_p$, complete with respect to a norm p, such that $\mathfrak{U}_0 \subset \mathfrak{U}'_p \subset \overline{\mathfrak{U}}(d)$ and such that $p(f_\nu) \to 0$ implies $d(f_\nu) \to 0$. By corollary 4.1, the fields with test functions in $\mathfrak{U}_p$ have a Gårding domain and are strongly continuous on it with respect to the norm p. This is just Reed's result. It is also instructive to apply the techniques of Secs. 2 and 3 directly to $\mathfrak{U}_0$ since in this case one can see directly where the norms $\{p_n\}$ come from and what they look like.

Irreducible Representations

Denote by $\chi_n(F)$ the characteristic function of $\mathfrak{U}'_{p_n} \subset \mathfrak{U}'$ and by P_n the projection operator given as multiplication by $\chi_n(F)$ in Q-space. Since $\mathfrak{U} \subset \mathfrak{U}'_{p_n}$, one has $\chi_n(F + g) = \chi_n(F)$, and so P_n commutes with $\{U(f) V(g)\}$. Thus $P_n \mathfrak{H}$ is an invariant subspace. Noting that $P_n \in \{U(f); f \in \mathfrak{U}\}''$, one obtains $P_n \in \{U(f), V(g)\}' \cap \{U(f), V(g)\}''$, and hence $P_n = 1$ or 0 for a factor representation. This means $\mu(\mathfrak{U}'_{p_n}) = 1$ or 0, and therefore one can choose all p_n to be equal, $p_1 = p_2 = \ldots$, so that a single Hilbertian norm p_1 suffices. In the same way one sees that all $\tilde{p}_n$ can be chosen to be equal. Defining a Hilbertian norm p on $\mathfrak{U}$ by

$$p(f)^2 = p_1(f)^2 + \tilde{p}_1(f)^2 \tag{4.3}$$

one arrives as in Corollary 4.1 at the following interesting extension.

Theorem 4.1: Let $\{U(f), V(g)\}$ be a continuous irreducible or factor representation of the CCR's with a separable barreled nuclear test function space $\mathfrak{U}$, such as S, $\mathfrak{D}$, or $\mathfrak{U}_0 \simeq \mathbb{R}^{(\infty)}$, in a separable Hilbert space $\mathfrak{H}$. Then there is a Hilbertian norm p on $\mathfrak{U}$ such that the representation can be extended by continuity to a *Hilbertian* test function space $\overline{\mathfrak{U}}_p$ obtained from $\mathfrak{U}$ as the Hilbert space completion of $\mathfrak{U}$ in the norm p.

The extended representation is strongly continuous in the norm p, and there exists a Gårding domain D for the fields satisfying (i)–(iii) of the main theorem, with $\mathfrak{U}$ replaced by $\overline{\mathfrak{U}}_p$.

If the representation is not irreducible or not a factor representation, we learn from the above argument that it decomposes into a direct sum of representations in $P_1\mathfrak{H}$, $(P_2 - P_1)\mathfrak{H}$, $(P_3 - P_2)\mathfrak{H}$ etc. such that $U(f)$ is p_n-continuous in $(P_n - P_{n-1})\mathfrak{H}$. For fixed subrepresentation one can apply the same argument to $V(g)$, and thus we find that *the representation is a direct sum of subrepresentations in each of which $U(f)$ and $V(g)$ are continuous with respect to some Hilbertian norm* (depending on the subrepresentation).

ACKNOWLEDGMENTS

The author would like to thank H. J. Borchers, J. Challifour, and J. Slawny for helpful comments.

[1] L. Gårding, Proc. Nat. Acad. Sci. (U.S.) **33**, 331 (1947).
[2] E. Nelson, Ann. Math. **70**, 572 (1959).
[3] I. M. Gel'fand and N. Y. Vilenkin, *Generalized Functions* (Academic, New York, 1964), Vol. 4 (last section).
[4] F. Trèves, *Topological Vector Spaces, Distributions and Kernels* (Academic, New York, 1967).
[5] G. C. Hegerfeldt and O. Melsheimer, Commun. Math. Phys. **12** 304 (1969).
[6] J. Glimm and A. Jaffe, Acta Math, **125**, 203 (1970)
[7] M. C. Reed, Commun. Math. Phys. **14**, 336 (1969).
[8] L. Gårding and A. S. Wightman, Proc. Nat. Acad. Sci. (U.S.) **40**, 622 (1954).
[9] G. C. Hegerfeldt, Nuovo Cimento **4B**, 225 (1971).
[10] Cf. Ref. 4, Theorem 33. 2 and Proposition 36. 1.
[11] A. Pietsch, *Nukleare Lokalkonvexe Räume. 2. Auflage* (Akademie-Verlag, Berlin, 1969), Sec. 4. 4. 10.
[12] Reference 11, Secs. 4. 4. 1 and 4. 4. 9. See also Y. Umemura, Publ. Res. Inst. Math. Sci. A **1**, 1 (1966).
[13] Y. Umemura, Ref. 12, Secs. 10 and 12.
[14] G. C. Hegerfeldt and J. R. Klauder, Commun. Math. Phys. **16**, 329 (1970).
[15] G. Choquet, *Lectures on Analysis*, edited by J. Marsden, T. Lance, and S. Gelbart (Benjamin, New York, 1969), Vol. II, Sec. 23. 16.

Clebsch-Gordan Coefficients and Special Function Identities. II. The Rotation and Lorentz Groups in 3-Space

Willard Miller, Jr.
Mathematics Department, University of Minnesota, Minneapolis, Minnesota 55455

(Received 17 February 1971)

It is shown that the construction of concrete models of Clebsch–Gordan decompositions for tensor products of irreducible group representations leads to a wide variety of special function identities. In this paper the representation theory of the rotation and Lorentz groups in 3-space is used to give elegant derivations of identities involving Laguerre, Gegenbauer, hypergeometric, and generalized hypergeometric functions. Some of these identities may be new in this general form.

INTRODUCTION

In Ref. 1, which we refer to as I, a method was described whereby a knowledge of the Clebsch–Gordan decomposition for the tensor product of two representations of a group G, could be used to derive special function identities. The idea is easy to describe. Suppose G has a family of irreducible representations $\{D_u\}$ with Clebsch–Gordan series

$$D_u \otimes D_v \cong \sum_w \oplus D_w,$$

such that each irreducible representation D_w occurs at most once in the tensor product. If $\{j_n^{(u)}\}$ is a canonical basis for D_u, then there exists a relation of the form

$$\text{(a)} \quad \tilde{j}_h^{(w)} = \sum_{n,m} C(u, n; v, m \,|\, w, h) j_n^{(u)} \otimes j_m^{(v)},$$

where the constants $C(\cdot\,|\,\cdot)$ are Clebsch–Gordan coefficients. Suppose we have an explicit function-space model of the representation $D_u \otimes D_v$. Then the vectors $j_n^{(u)} \otimes j_m^{(v)}$ will be special functions and if the model is simple enough, the special functions $\tilde{j}_h^{(w)}$ can be computed directly. In this case, expression (a) becomes an identity relating the special functions $j_n^{(u)} \otimes j_m^{(v)}$ and $\tilde{j}_h^{(w)}$. This identity can be inverted since the coefficients $C(\cdot\,|\,\cdot)$ satisfy orthogonality relations.

The above method is useful for a given group if there is a procedure for constructing a variety of models of the group representations. In Refs. 2–4, a number of such models are cataloged for groups of common occurence in physics. Here we use these models to give elegant derivations of identities associated with the rotation and homogeneous Lorentz groups in 3-space. Some of these identities may be new in this general form; certainly their close relationship to one another and to group theory is new.

Most of the following explicit examples are associated with the Lorentz group G_3 but the analogous examples for $SO(3)$ are usually self-evident.

In physical applications, integral forms of these identities appear when one computes matrix elements corresponding to a quantum mechanical system with symmetry group $SO(3)$ or G_3.[5,6] However, the group theoretic method has validity independent of the computation of matrix elements, so the results of this paper are not presented in integral form. The reader can write most of the following identities in various integral forms by using well-known orthogonality relations for the Laguerre, Gegenbauer, and hypergeometric functions.

1. THE GROUPS $SU(2)$ AND G_3

The group $SU(2)$ consists of all 2×2 unitary unimodular matrices. In Euler angles, every $A \in SU(2)$ can be written as

$$A(\varphi_1, \theta, \varphi_2) = \begin{pmatrix} e^{-i(\varphi_1+\varphi_2)/2}\cos(\theta/2), & ie^{-i(\varphi_1-\varphi_2)/2}\sin(\theta/2) \\ ie^{i(\varphi_1-\varphi_2)/2}\sin(\theta/2), & e^{i(\varphi_1+\varphi_2)/2}\cos(\theta/2) \end{pmatrix} = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}, \qquad |a|^2 + |b|^2 = 1. \quad (1.1)$$

If $ab \neq 0$ the Euler angles can be defined uniquely by

$$\cos(\theta/2) = |a|, \quad \sin(\theta/2) = |b|, \quad 0 \leq \theta \leq \pi,$$

If the representation is not irreducible or not a factor representation, we learn from the above argument that it decomposes into a direct sum of representations in $P_1\mathfrak{H}$, $(P_2 - P_1)\mathfrak{H}$, $(P_3 - P_2)\mathfrak{H}$ etc. such that $U(f)$ is p_n-continuous in $(P_n - P_{n-1})\mathfrak{H}$. For fixed subrepresentation one can apply the same argument to $V(g)$, and thus we find that *the representation is a direct sum of subrepresentations in each of which $U(f)$ and $V(g)$ are continuous with respect to some Hilbertian norm* (depending on the subrepresentation).

ACKNOWLEDGMENTS

The author would like to thank H. J. Borchers, J. Challifour, and J. Slawny for helpful comments.

[1] L. Gårding, Proc. Nat. Acad. Sci. (U.S.) **33**, 331 (1947).
[2] E. Nelson, Ann. Math. **70**, 572 (1959).
[3] I. M. Gel'fand and N. Y. Vilenkin, *Generalized Functions* (Academic, New York, 1964), Vol. 4 (last section).
[4] F. Trèves, *Topological Vector Spaces, Distributions and Kernels* (Academic, New York, 1967).
[5] G. C. Hegerfeldt and O. Melsheimer, Commun. Math. Phys. **12** 304 (1969).
[6] J. Glimm and A. Jaffe, Acta Math, **125**, 203 (1970)
[7] M. C. Reed, Commun. Math. Phys. **14**, 336 (1969).
[8] L. Gårding and A. S. Wightman, Proc. Nat. Acad. Sci. (U.S.) **40**, 622 (1954).
[9] G. C. Hegerfeldt, Nuovo Cimento **4B**, 225 (1971).
[10] Cf. Ref. 4, Theorem 33. 2 and Proposition 36. 1.
[11] A. Pietsch, *Nukleare Lokalkonvexe Räume. 2. Auflage* (Akademie-Verlag, Berlin, 1969), Sec. 4. 4. 10.
[12] Reference 11, Secs. 4. 4. 1 and 4. 4. 9. See also Y. Umemura, Publ. Res. Inst. Math. Sci. A **1**, 1 (1966).
[13] Y. Umemura, Ref. 12, Secs. 10 and 12.
[14] G. C. Hegerfeldt and J. R. Klauder, Commun. Math. Phys. **16**, 329 (1970).
[15] G. Choquet, *Lectures on Analysis*, edited by J. Marsden, T. Lance, and S. Gelbart (Benjamin, New York, 1969), Vol. II, Sec. 23. 16.

Clebsch-Gordan Coefficients and Special Function Identities. II. The Rotation and Lorentz Groups in 3-Space

Willard Miller, Jr.
Mathematics Department, University of Minnesota, Minneapolis, Minnesota 55455

(Received 17 February 1971)

It is shown that the construction of concrete models of Clebsch–Gordan decompositions for tensor products of irreducible group representations leads to a wide variety of special function identities. In this paper the representation theory of the rotation and Lorentz groups in 3-space is used to give elegant derivations of identities involving Laguerre, Gegenbauer, hypergeometric, and generalized hypergeometric functions. Some of these identities may be new in this general form.

INTRODUCTION

In Ref. 1, which we refer to as I, a method was described whereby a knowledge of the Clebsch–Gordan decomposition for the tensor product of two representations of a group G, could be used to derive special function identities. The idea is easy to describe. Suppose G has a family of irreducible representations $\{D_u\}$ with Clebsch–Gordan series

$$D_u \otimes D_v \cong \sum_w \oplus D_w,$$

such that each irreducible representation D_w occurs at most once in the tensor product. If $\{j_n^{(u)}\}$ is a canonical basis for D_u, then there exists a relation of the form

$$\text{(a)} \quad \tilde{j}_h^{(w)} = \sum_{n,m} C(u, n; v, m \,|\, w, h) j_n^{(u)} \otimes j_m^{(v)},$$

where the constants $C(\cdot\,|\,\cdot)$ are Clebsch–Gordan coefficients. Suppose we have an explicit function-space model of the representation $D_u \otimes D_v$. Then the vectors $j_n^{(u)} \otimes j_m^{(v)}$ will be special functions and if the model is simple enough, the special functions $\tilde{j}_h^{(w)}$ can be computed directly. In this case, expression (a) becomes an identity relating the special functions $j_n^{(u)} \otimes j_m^{(v)}$ and $\tilde{j}_h^{(w)}$. This identity can be inverted since the coefficients $C(\cdot\,|\,\cdot)$ satisfy orthogonality relations.

The above method is useful for a given group if there is a procedure for constructing a variety of models of the group representations. In Refs. 2–4, a number of such models are cataloged for groups of common occurence in physics. Here we use these models to give elegant derivations of identities associated with the rotation and homogeneous Lorentz groups in 3-space. Some of these identities may be new in this general form; certainly their close relationship to one another and to group theory is new.

Most of the following explicit examples are associated with the Lorentz group G_3 but the analogous examples for $SO(3)$ are usually self-evident.

In physical applications, integral forms of these identities appear when one computes matrix elements corresponding to a quantum mechanical system with symmetry group $SO(3)$ or G_3.[5,6] However, the group theoretic method has validity independent of the computation of matrix elements, so the results of this paper are not presented in integral form. The reader can write most of the following identities in various integral forms by using well-known orthogonality relations for the Laguerre, Gegenbauer, and hypergeometric functions.

1. THE GROUPS $SU(2)$ AND G_3

The group $SU(2)$ consists of all 2×2 unitary unimodular matrices. In Euler angles, every $A \in SU(2)$ can be written as

$$A(\varphi_1, \theta, \varphi_2) = \begin{pmatrix} e^{-i(\varphi_1+\varphi_2)/2}\cos(\theta/2), & ie^{-i(\varphi_1-\varphi_2)/2}\sin(\theta/2) \\ ie^{i(\varphi_1-\varphi_2)/2}\sin(\theta/2), & e^{i(\varphi_1+\varphi_2)/2}\cos(\theta/2) \end{pmatrix} = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}, \qquad |a|^2 + |b|^2 = 1. \quad (1.1)$$

If $ab \neq 0$ the Euler angles can be defined uniquely by

$$\cos(\theta/2) = |a|, \quad \sin(\theta/2) = |b|, \quad 0 \leq \theta \leq \pi,$$

$$-\tfrac{1}{2}(\varphi_1+\varphi_2)=\arg a,\quad \tfrac{1}{2}(\varphi_2-\varphi_1)+\tfrac{1}{2}\pi=\arg b,$$
$$-2\pi\le\varphi_1,\varphi_2<2\pi. \tag{1.2}$$

However, if $ab=0$ these angles are not unique. As a basis for the Lie algebra $su(2)$ we choose the matrices $\mathcal{J}_1,\mathcal{J}_2,\mathcal{J}_3$, such that

$$A(0,\theta,0)=\exp\theta\mathcal{J}_1\quad A\left(\frac{\pi}{2},\theta,-\frac{\pi}{2}\right)=\exp\theta\mathcal{J}_2,$$
$$A(\varphi,0,0)=A(0,0,\varphi)=\exp\varphi\mathcal{J}_3. \tag{1.3}$$

These matrices satisfy the commutation relations

$$[\mathcal{J}_1,\mathcal{J}_2]=\mathcal{J}_3,\quad [\mathcal{J}_3,\mathcal{J}_1]=\mathcal{J}_2,\quad [\mathcal{J}_2,\mathcal{J}_3]=\mathcal{J}_1. \tag{1.4}$$

Another convenient basis is given by

$$\mathcal{J}^{\pm}=\mp\mathcal{J}_2+i\mathcal{J}_1,\qquad \mathcal{J}^3=i\mathcal{J}_3, \tag{1.5}$$

which belong to the complexification of $su(2)$. Here,

$$[\mathcal{J}^3,\mathcal{J}^{\pm}]=\pm\mathcal{J}^{\pm},\quad [\mathcal{J}^+,\mathcal{J}^-]=2\mathcal{J}^3. \tag{1.6}$$

The irreducible unitary representations of $SU(2)$ are D_u, $2u=0,1,2,\ldots$, each defined on a $(2u+1)$-dimensional Hilbert space $\mathcal{H}_u$ with ON basis $\{p_m: m=-u,-u+1,\ldots,u-1,u\}$. The defining relations are

$$J^3p_m=mp_m,\qquad J^{\pm}p_m=[(u\mp m)(u\pm m+1)]^{1/2}p_{m\pm1},$$
$$m=-u,\ldots,u, \tag{1.7}$$

where $J^{\pm},J^3$ are the linear operators corresponding to $\mathcal{J}^{\pm},\mathcal{J}^3$, respectively, in the Lie algebra representation induced by D_u. The matrix elements $U_{n,m}(A)$ of the unitary operators $\mathbf{U}(A)$ on $\mathcal{H}_u$ which determine this representation are

$$U_{n,m}(A)=\langle p_n,U(A)p_m\rangle=\left(\frac{(u+m)!\,(u-n)!}{(u+n)!\,(u-m)!}\right)^{1/2}$$
$$\times a^{u+n}\bar{a}^{u-m}\bar{b}^{m-n}\frac{1}{\Gamma(m-n+1)}$$
$$\times F(-u-n,m-u,m-n+1;-|b/a|^2)$$
$$=(i)^{n-m}\left(\frac{(u+m)!\,(u-n)!}{(u+n)!\,(u-m)!}\right)^{1/2}$$
$$\times e^{-i(n\varphi_1+m\varphi_2)}P_u^{-n,m}(\cos\theta), \tag{1.8}$$

where

$$P_u^{r,m}(x)=\left(\frac{1+x}{2}\right)^{(m-r)/2}\left(\frac{1-x}{2}\right)^{(m+r)/2}$$
$$\times\frac{1}{\Gamma(m+r+1)}F(u+m+1,-u+m;$$
$$m+r+1;\tfrac{1}{2}(1-x)) \tag{1.9}$$

and A is given by (1.1). Here, $\langle\cdot,\cdot\rangle$ is the inner product on $\mathcal{H}_u$, linear in the second argument and $F(a,b;c;z)$ is the hypergeometric function, see Ref. 7, Vol. 1.

The group G_3 consists of all 2×2 complex matrices of the form

$$A=\begin{pmatrix}a & b\\ \bar{b} & \bar{a}\end{pmatrix},\qquad a,b,\in\mathbb{C},\qquad \det A=|a|^2-|b|^2=1. \tag{1.10}$$

This is a real 3-parameter matrix group isomorphic to $SL(2,R)$.[2] Furthermore, G_3 is the twofold covering group of the homogeneous Lorentz group in 3-space.[8] We can choose real coordinates (μ,ρ,ν) for A so that

$$A(\mu,\rho,\nu)=\begin{pmatrix}e^{-i(\mu+\nu)/2}\cosh(\rho/2) & e^{i(\nu-\mu)/2}\sinh(\rho/2)\\ e^{i(\mu-\nu)/2}\sinh(\rho/2) & e^{i(\mu+\nu)/2}\cosh(\rho/2)\end{pmatrix}. \tag{1.11}$$

Here we require,

$$|a|=\cosh\rho,\qquad |b|=\sinh\rho,\qquad 0\le\rho<\infty,$$
$$\mu=-\arg a-\arg b,\qquad \nu=\arg b-\arg a. \tag{1.12}$$

The matrices $\mathcal{J}_1,\mathcal{J}_2,\mathcal{J}_3$, such that

$$A(0,\rho,\pi)=\exp\rho\mathcal{J}_1,\quad A(0,\rho,0)=\exp\rho\mathcal{J}_2,$$
$$A(\mu,0,0)=A(0,0,\mu)=\exp\mu\mathcal{J}_3, \tag{1.13}$$

form a basis for the Lie algebra $\mathcal{G}_3$ of G_3. The commutation relations are

$$[\mathcal{J}_1,\mathcal{J}_2]=-\mathcal{J}_3,\quad [\mathcal{J}_3,\mathcal{J}_1]=\mathcal{J}_2,\quad [\mathcal{J}_3,\mathcal{J}_2]=-\mathcal{J}_1. \tag{1.14}$$

A more convenient basis for many purposes is $\mathcal{J}^{\pm}=-\mathcal{J}_2\pm i\mathcal{J}_1$, $\mathcal{J}^3=i\mathcal{J}_3$ in the complexification of $\mathcal{G}_3$. Here the commutation relations are

$$[\mathcal{J}^+,\mathcal{J}^-]=2\mathcal{J}^3,\ [\mathcal{J}^3,\mathcal{J}^{\pm}]=\pm\mathcal{J}^{\pm}, \tag{1.15}$$

identical with (1.6).

We consider a class D_u^+ of irreducible unitary representations of G_3, defined for $u>0$ (discrete series). Here, D_u^+ can be realized on the Hilbert space $\mathcal{H}$ with ON basis $\{j_n: n=0,1,2,\ldots\}$. The defining relations are

$$J^3j_n=(u+n)j_n,\qquad J^+j_n=[(2u+n)(n+1)]^{1/2}j_{n+1},$$
$$J^-j_n=-[n(2u+n-1)]^{1/2}j_{n-1},\qquad n=0,1,\ldots, \tag{1.16}$$

where $J^{\pm},J^3$ are the representation operators corresponding to $\mathcal{J}^{\pm},\mathcal{J}^3$, respectively. (To be more precise, D_u^+ is a global representation of G_3 only for $2u$, an integer. For $2u$ not an integer, D_u^+ is a local representation of G_3 and a global irreducible representation of the simply connected covering group of G_3 (see Refs. 2, 8, and 9). The matrix elements of D_u^+ are

$$V_{n,m}(A)=\langle j_n,V(A)j_m\rangle$$
$$=\left(\frac{\Gamma(2u+n)m!}{\Gamma(2u+m)n!}\right)^{1/2}\cdot a^n\bar{a}^{-2u-m}\bar{b}^{m-n}$$
$$\times\frac{F(-n,2u+m;m-n+1;|b/a|^2)}{\Gamma(m-n+1)}$$
$$=\left(\frac{\Gamma(2u+n)m!}{\Gamma(2u+m)n!}\right)^{1/2}e^{-i[\mu(u+n)+\nu(u+m)]}\mathfrak{P}_{-u(\cosh\rho)}^{-u-n,u+m}, \tag{1.17}$$

where the coordinates of A are given by (1.10), (1.11), and

$$\mathfrak{P}_{\nu}^{\mu,\xi}(z)=\frac{1}{\Gamma(\xi+\mu+1)}\left(\frac{z+1}{2}\right)^{(\xi-\mu)/2}\left(\frac{z-1}{2}\right)^{(\xi+\mu)/2}$$
$$\times F(\nu+\xi+1,\xi-\nu;\mu+\xi+1;\tfrac{1}{2}(1-z)). \tag{1.18}$$

As is well known, the Clebsch–Gordan series for $SU(2)$ is

$$D_u \otimes D_v \cong \sum_{w=|u-v|}^{u+v} \oplus D_w. \tag{1.19}$$

The vectors $\{p_{n,m} = p_n \otimes p'_m : n = -u, -u+1, \ldots, m = -v, -v+1, \ldots, v\}$ form a natural basis for the representation space $\mathcal{K}_u \otimes \mathcal{K}'_v$, while a canonical basis for the subspace transforming according to D_w can be denoted $\{p_k^w : k = -w+1, \ldots, w\}$. The Clebsch–Gordan (CG) coefficients relating these two bases are

$$C(u, n; v, m \mid w, k) = \langle p_{n,m}, p_k^w \rangle', \tag{1.20}$$

where $\langle \cdot, \cdot \rangle'$ is the inner product on $\mathcal{K}_u \otimes \mathcal{K}'_v$. If the basis vectors are chosen appropriately, the CG coefficients are defined by the generating function

$$\exp[\alpha(x_2 - x_3) + \beta(x_3 - x_1) + \gamma(x_1 - x_2)]$$
$$= \sum_{j_1+j_2+j_3=0}^{\infty} \sum_{m_i=-j_i}^{j_i} (j_1 + j_2 + j_3 + 1)^{1/2} \frac{\alpha^{-j_1+j_2+j_3} \beta^{j_1-j_2+j_3}}{[(-j_1 + j_2 + j_3)!(j_1 - j_2 + j_3)!}$$
$$\times \frac{\gamma^{j_1+j_2-j_3} x_1^{j_1+m_1} x_2^{j_2+m_2} x_3^{j_3+m_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}}{(j_1 + j_2 - j_3)!\,(j_1 + m_1)!\,(j_1 - m_1)!\,(j_2 + m_2)!(j_2 - m_2)!(j_3 + m_3)!(j_3 - m_3)!]^{1/2}}, \tag{1.21}$$

where the 3-j coefficients are

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \frac{(-1)^{j_3-m_3}}{\sqrt{2j_3+1}} C(j_1, m_1; j_2, m_2 \mid j_3 - m_3) \tag{1.22}$$

and the sum is taken over all j_i, m_i for which (1.21) makes sense. In particular, $C(u, n; v, m \mid w, k) = 0$ unless $k = n + m$ and $|u - v| \le w \le u + v$. The various symmetries and explicit formulas for the CG coefficients which abound in the literature can all be obtained from (1.21), see Refs. 10 and 11.

The Clebsch–Gordan series for the tensor product $D_u^+ \otimes D_v^+$ of G_3 representations is[2,6]

$$D_u^+ \otimes D_v^+ \cong \sum_{s=0}^{\infty} \oplus D_{u+v+s}^+. \tag{1.23}$$

The vectors $\{j_{n,m} = j_n \otimes j'_m : n, m = 0, 1, \cdots\}$ form a natural ON basis for the representation space $\mathcal{K} \otimes \mathcal{K}'$. A canonical basis for the subspace of $\mathcal{K} \otimes \mathcal{K}'$ transforming according to D_{u+v+s}^+ can be denoted $\{j_h^s : h = 0, 1, \cdots\}$. The CG coefficients are

$$E(u, n; v, m \mid s, h) = \langle j_{n,m}, j_h^s \rangle', \tag{1.24}$$

where $\langle \cdot, \cdot \rangle'$ is the inner product on $\mathcal{K} \otimes \mathcal{K}'$. With an appropriate choice of basis vectors, the CG coefficients are given by the generating function.

$$\left(\frac{(2u + 2v + 2s - 1)\Gamma(2u + 2v + s - 1)\Gamma(2v + s)}{s!\,\Gamma(2u)\Gamma(2v)}\right)^{1/2}$$
$$\times (1 - by)^{-2u-s} (1 - bx)^{-2v-s} (y - x)^s$$
$$= \sum_{h,n,m=0}^{\infty} \left(\frac{\Gamma(2u + 2v + 2s + h)}{h!}\right)^{1/2}$$
$$\times E(u, n; v, m \mid s, h) y^n x^m b^h, \quad |bx| < 1, \ |by| < 1. \tag{1.25}$$

We can expand the left-hand side of (1.25) to obtain explicit expressions for the CG coefficients. In general they are rather complicated finite sums. However in the special cases $s = 0$ or $h = 0$, the sum contains only one term and the CG coefficient reduces to the square root of a quotient of gamma functions, as the reader can easily verify.

From the definitions (1.20), (1.24) it follows that the CG coefficients satisfy orthogonality relations. Indeed the coefficients $E(\cdot)$ are real and satisfy

$$\sum_{x,h=0} E(u, n_1; v, m_1 \mid s, h) E(u, n_2, v, m_2 \mid s, h) = \delta_{n_1 n_2} \delta_{m_1 m_2}, \tag{1.26}$$
$$\sum_{n,m=0}^{\infty} E(u, n; v, m \mid s_1, h_1) E(u, n; v, m \mid s_2, h_2) = \delta_{s_1 s_2} \delta_{h_1 h_2}.$$

The coefficients $C(\cdot)$ satisfy similar relations except that the sums are finite.

2. IDENTITIES FOR THE MATRIX ELEMENTS OF $SO(3)$ AND G_3

Just as in I, Sec. 2, we can use products of matrix elements of the representations D_u, D_u^+ to construct new models of these representations. Since the methods are identical with I we present only the results.

For fixed b and c, the functions

$$p_h^{(u+v+s)}(A, A') = \sum_{n,m=0}^{\infty} E(u, n; v, m \mid s, h)$$
$$\times V_{b,n}^{(u)}(A) V_{c,m}^{(v)}(A'), \quad h = 0, 1, 2, \cdots, \tag{2.1}$$

form a canonical basis for a model of D_{u+v+s}^+ under the group action

$$[P(B)f](A, A') = f(AB, A'B), \quad A, A', B \in G_3 \tag{2.2}$$

on functions defined on $G_3 \times G_3$. Hence, $V_{b,n}^{(u)}(A)$ is the matrix element (1.17) corresponding to the representation D_u^+. Note that the sum on the right-hand side of (2.1) is finite since $E(u, n; v, m \mid s, h) = 0$ unless $n + m = s + h$.

Using the transformation properties of the basis $p_h^{(u+v+s)}$, we can also show

$$p_h^{(u+v+s)}(A, A') = \sum_{j=0}^{\infty} E(u, s + j - c; v, c \mid s, j)$$
$$\times V_{j,h}^{(u+v+s)}(A') V_{b,s+j-c}^{(u)}(A(A')^{-1}). \tag{2.3}$$

Equating (2.1) and (2.3) we obtain a family of identi-

ties obeyed by the matrix elements. In particular, for $A = A'$ the identity reduces to the formula

$$\sum_{n,m} E(u,n;v,m\,|\,s,h)V^{(u)}_{b,n}(A)V^{(v)}_{c,m}(A) = E(u,b;v,c\,|\,s,b+c-s)V^{(u+v+s)}_{b+c-s,h}(A), \quad (2.4)$$

since $V^{(u)}_{n,m}(E) = \delta_{n,m}$ for E the identity matrix.

The construction of models of the representations D_u of $SU(2)$ is analogous to that given above, and formulas (2.1) and (2.3) can easily be modified for this case. Of special interest is the case where the basis contains only one element $p^{(0)}_0(A,A')$, i.e., this function transforms according to the identity representation D_0. Nonzero functions $p^{(0)}_0(A,A')$ can be constructed only if $u = v$, in which case the analogy of (2.1) is

$$p^{(0)}_0(A,A') = \sum_{n=-u}^{u} C(u,n;u,-n\,|\,0,0)\,U^{(u)}_{b,n}(A)\,U^{(u)}_{c,-n}(A') \quad (2.5)$$

with fixed b, c. The analogy of (2.3) is

$$p^{(0)}_0(A,A') = C(u,-c;u,c\,|\,0,0)U^{(u)}_{b,-c}(A(A')^{-1}). \quad (2.6)$$

Equating (2.5) and (2.6) we obtain a family of addition theorems for the matrix elements. The simplest case, $b = c = 0, u = l$, yields the well-known addition theorem

$$P_l[\cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\varphi-\varphi')] = \frac{4\pi}{2l+1}\sum_{m=-l}^{l} Y_{lm}(\theta',\varphi')Y_{lm}(\theta,\varphi) \quad (2.7)$$

for the Legendre polynomials (see Ref. 12, p. 68).

3. DIFFERENTIAL OPERATOR MODELS

In this section we construct new models of the representations D^+_u as classified in Ref. 2, Chap. 5, and use these models and the results of Sec. 1 to obtain special function identities.

The Type B operators

$$J^+ = e^{i\theta}(x\frac{\partial}{\partial x} - i\frac{\partial}{\partial\theta} - x), \qquad J^- = e^{-i\theta}(x\frac{\partial}{\partial x} + i\frac{\partial}{\partial\theta}), \quad J^3 = -i\frac{\partial}{\partial\theta} \quad (3.1)$$

and basis functions

$$j_n(x,\theta) = \left(\frac{n!}{\Gamma(n+2u)}\right)^{1/2} \times x^u L^{(2u-1)}_n(x)\,e^{i(u+n)\theta}, \qquad n = 0,1,2,\ldots, \quad (3.2)$$

form a model of D^+_u, i.e., they satisfy expressions (1.16). Here $L^{(\alpha)}_n(x)$ is a generalized Laguerre polynomial (see Ref. 7, Vol. 1).

It follows that the functions

$$j_{n,m}(x,\theta) = j^{(1)}_n(x,\theta)\,j^{(2)}_m(x,\theta) = \left(\frac{n!}{\Gamma(n+2u)}\right)^{1/2} \times x^u L^{(2u-1)}_n(ax)e^{i(u+n)\theta}\left(\frac{m!}{\Gamma(m+2v)}\right)^{1/2} \times x^v_m L^{(2v-1)}([1-a]x)\,e^{i(v+m)\theta}, \quad n,m = 0,1,2,\ldots, \quad (3.3)$$

and the operators (3.1) define a model of $D^+_u \otimes D^+_v$ where a, u, v are real constants such that $u > 0, v > 0$. Indeed,

$$\begin{aligned} J^+(j^{(1)}_n j^{(2)}_m) &= j^{(1)}_n e^{i\theta}\left(x\frac{\partial}{\partial x} - i\frac{\partial}{\partial\theta} - (1-a)x\right)j^{(2)}_m \\ &\quad + j^{(2)}_m e^{i\theta}\left(x\frac{\partial}{\partial x} - i\frac{\partial}{\partial\theta} - ax\right)j^{(1)}_n = [(2u+n) \\ &\quad \times (n+1)]^{1/2} j^{(1)}_{n+1}\, j^{(2)}_m \\ &\quad + [(2v+m)(m+1)]^{1/2} j^{(1)}_n j^{(2)}_{m+1} \\ &= e^{i\theta}\left(x\frac{\partial}{\partial x} - i\frac{\partial}{\partial\theta} - x\right)j^{(1)}_n j^{(2)}_m, \end{aligned} \quad (3.4)$$

with similar interpretations of J^- and J^3.

We now compute the basis vectors j^s_h, $s, h = 0,1,2,\ldots$, corresponding to the Clebsch–Gordan series (1.23). From (1.24) we have

$$j^s_h(x,\theta) = \sum_{n,m=0}^{\infty} E(u,n;v,m\,|\,s,h)\,j^{(1)}_n(x,\theta)\,j^{(2)}_m(x,\theta). \quad (3.5)$$

[Recall that $E(\cdot) = 0$ unless $n + m = s + h$.] On the other hand, we can compute the j^s_h directly for this model by using the fact that they satisfy (1.16) with $n = h$, $u = u + v + s$. Indeed, from (3.2),

$$j^s_h(x,\theta) = c_s\left(\frac{h!}{\Gamma(h+2u+2v+2s)}\right)^{1/2} \times x^{u+v+s}L^{(2u+2v+2s-1)}_h(x)\,e^{i(u+v+s+h)\theta}, \quad (3.6)$$

where c_s is a constant. To determine c_s, we equate (3.5) and (3.6) in the case $h = 0$. In this special case, (3.5) simplifies to

$$\begin{aligned} j^s_0 &= \left(\frac{s!\Gamma(2u+2v+s-1)\,\Gamma(2u+s)\Gamma(2v+s)}{\Gamma(2u+2v+2s-1)}\right)^{1/2} \\ &\quad \times \sum_{n=0}^{\infty}(-1)^n[(s-n)!n!\,\Gamma(2u+n)\Gamma(2v+s-n)]^{-1/2} \\ &\quad \times j^{(1)}_n j^{(2)}_{s-n}. \end{aligned} \quad (3.7)$$

Substituting (3.3) and (3.6) into this expression, comparing coefficients of x^{u+v+s} on both sides of the resulting equation we find

$$\begin{aligned} c_s &= \frac{(a-1)^s}{\Gamma(2u)}\left(\frac{\Gamma(2u+s)\Gamma(2u+2v+s-1)}{s!\Gamma(2v+s)}\right. \\ &\quad \left.\times(2u+2v+2s-1)\right)^{1/2} F\left(1-s-2v, \right. \\ &\quad \left. -s;2u;\frac{a}{a-1}\right), \qquad s = 0,1,2,\ldots, \end{aligned} \quad (3.8)$$

where $F(\alpha,\beta;\gamma;z)$ is the hypergeometric function (see Ref. 7, Vol. 1) Note that c_s is a polynomial of order s in a. The final identity is obtained by substituting (3.3), (3.6), and (3.8) into (3.5).

For $a = 0$ this identity simplifies to

$$x^s L^{(2u+2v+2s-1)}_h(x)(-1)^s \times\left(\frac{h!\Gamma(2u+s)\Gamma(2u+2v+s-1)(2u+2v+2s-1)}{x!\Gamma(2v+s)\Gamma(2u+2v+2s+h)}\right)^{1/2}$$

$$= \sum_{n,m} E(u,n;v,m\,|\,s,h)\left(\frac{m!\Gamma(2u+n)}{n!\Gamma(2v+m)}\right)^{1/2} L_m^{(2v-1)}(x). \tag{3.9}$$

A second model of D_u^+ is defined by the operators

$$J^{\pm} = e^{\pm i\theta}\left((z^2-1)^{1/2}\frac{\partial}{\partial z} \pm \frac{iz}{(z^2-1)^{1/2}}\frac{\partial}{\partial\theta} \mp \frac{r}{(z^2-1)^{1/2}}\right), \qquad J^3 = -i\frac{\partial}{\partial\theta}, \tag{3.10}$$

and basis functions

$$j_n^{(z,\theta)} = [\Gamma(2u+n)n!]^{-1/2}\mathfrak{P}_{-u}^{-r,-u-n}(z)e^{i(u+n)\theta}, \qquad n = 0,1,2,\ldots,\ z = \cosh\rho, \tag{3.11}$$

where r is an arbitrary constant. It follows that the functions

$$\begin{aligned} j_{n,m}^{(z,\theta)} &= j_n(z,\theta)j_m(z,\theta) = [\Gamma(2u+n)n!]^{-1/2} \\ &\times \mathfrak{P}_{-u}^{-r_1,-u-n}(z)\,e^{i(u+n)\theta}[\Gamma(2v+m)m!]^{-1/2} \\ &\times \mathfrak{P}_{-v}^{-r_2,-v-m}(z)\,e^{i(v+m)\theta}, \\ r_1 + r_2 &= r, \qquad n,m = 0,1,2,\cdots, \end{aligned} \tag{3.12}$$

and the operators (3.10) define a realization of $D_u^+ \otimes D_v^+$. Indeed, writing $J^+ = \tilde{J}^+ - e^{i\theta}r(z^2-1)^{-1/2}$, we have

$$J^+(j_nj_m) = j_n(\tilde{J}^+ - e^{i\theta}r_2(z^2-1)^{-1/2})j_m + j_m(\tilde{J}^+ - e^{i\theta}r_1(z^2-1)^{-1/2})j_n$$

with a similar interpretation of J^- and J^3. From (3.11) we see that the basis functions j_h^s corresponding to the Clebsch–Gordan series (1.23) must be

$$j_h^s = c_s[\Gamma(2u+2v+2s+h)h!]^{-1/2} \times \mathfrak{P}_{-u-v-s}^{-r,-u-v-s-h}(z)\,e^{i(u+v+s+h)\theta}. \tag{3.13}$$

To compute the constant c_s we substitute (3.12) and (3.13), $h = 0$, into (3.7). Canceling the common factor

$$\left(\frac{z-1}{2}\right)^{-(u+v+s+r)/2}\left(\frac{z+1}{2}\right)^{-(u+v+s-r)/2}$$

on both sides of the equation and setting $z = 1$, we obtain

$$\begin{aligned} c_s &= \frac{\Gamma(-u-v-s-r+1)}{\Gamma(2u)\Gamma(-r_1-u+1)\Gamma(-r_2-v-s-1)} \\ &\times \left(\frac{\Gamma(2u+s)\Gamma(2u+2v+s-1)}{s!\Gamma(2v+s)}\right. \\ &\times \left.(2u+2v+2s-1)\right)^{1/2} \\ &\times {}_3F_2(-s,-2v-s+1,u+r_1; \\ &\quad 2u,-r_2-v-s-1;1). \end{aligned} \tag{3.14}$$

Our final identity is obtained by substituting (3.12), (3.13), and (3.14) into (3.5). In the very special case $r_1 = r_2 = s = 0$ this identity reduces to

$$\frac{\Gamma(-u-v+1)}{\Gamma(1-u)\Gamma(-v-1)}\left(\frac{\Gamma(2u+2v)}{\Gamma(2u)\Gamma(2v)\Gamma(2u+2v+h)h!}\right)^{1/2} \times \mathfrak{P}_{-u-v-s}^{u+v+h}(z) = \sum_{n=0}^{h} E(u,n;v,h-n\,|\,0,h) \times \frac{\mathfrak{P}_{-u}^{u+n}(z)\ \mathfrak{P}_{-v}^{v+s-n}(z)}{[\Gamma(2u+n)\Gamma(2v+s-n)n!(s-n)!]^{1/2}}, \tag{3.15}$$

where $\mathfrak{P}_{\nu}^{\mu}(z)$ is a Legendre function of the first kind (see Ref. 7, Vol. 1). [In this special case $s = 0$, the coefficients $E(\cdot)$ are easy to evaluate explicitly.]

For our next model of D_u^+ we choose operators

$$J^{\pm} = e^{\pm i\theta}\left((x^2-1)\frac{\partial}{\partial x} \mp ix\frac{\partial}{\partial\theta}\right), \qquad J^3 = -\frac{\partial}{\partial\theta}, \tag{3.16}$$

and basis functions

$$j_n(x,\theta) = \left(\frac{n!}{\Gamma(2u+n)}\right)^{1/2}(x^2-1)^{u/2}C_n^u(x)e^{i(u+n)\theta}, \qquad n = 0,1,2,\ldots, \tag{3.17}$$

where $C_n^u(x)$ is a Gegenbauer polynomial (see Ref. 7, Vol. 2). It follows that a model of $D_u^+ \otimes D_v^+$ is determined by the operators (3.16) and basis functions

$$j_{n,m}(x,\theta) = j_n(x,\theta)\,j_m(x,\theta) = \left(\frac{n!m!}{\Gamma(2u+n)\Gamma(2v+m)}\right)^{1/2} \times (x^2-1)^{(u+v)/2}C_n^u(x)C_m^v(x)\,e^{i(u+v+n+m)\theta}, \qquad n,m = 0,1,2,\cdots. \tag{3.18}$$

The basis functions j_h^s transforming according to D_{u+v+s}^+ can be obtained directly from (3.17):

$$j_h^s(x,\theta) = c_s\left(\frac{h!}{\Gamma(2u+2v+2s+h)}\right)^{1/2}(x^2-1)^{(u+v+s)/2} \times C^{u+v+s}(x)e^{i(u+v+s+h)\theta}, \qquad s,h = 0,1,2,\cdots. \tag{3.19}$$

To determine the constants c_s we substitute (3.18) and (3.19) into (3.7) and divide through by the common factor $(x^2-1)^{(u+v)/2}$. If s is odd, the right-hand side of the resulting expression is odd and the left-hand side is even. Thus

$$c_s = 0, \qquad s \text{ odd}. \tag{3.20}$$

If s is even, we compare coefficients of x^s on both sides of the equation to obtain

$$\begin{aligned} c_s &= \left(\frac{\Gamma(2u+s)(2u+2v+2s-1)}{s!\Gamma(2v+s)}\right)^{1/2}\frac{2^s\Gamma(v+s)}{\Gamma(2u)\Gamma(v)} \\ &\times {}_3F_2(u,-s,-2v-s+1;2u,-v-s+1;1), \\ &s \text{ even}. \end{aligned} \tag{3.21}$$

Substituting (3.18)–(3.21) into (3.5), we obtain our general identity. In the special case $s = 0$, this formula reduces to

$$\left(\frac{h!(2u+2v-1)}{\Gamma(2u+2v+h)\Gamma(2u)\Gamma(2v)}\right)^{1/2}C_h^{u+v}(x) = \sum_{n=0}^{h} E(u,n;v,h-n\,|\,0,h) \times \left(\frac{n!(h-n)!}{\Gamma(2u+n)\Gamma(2v+h-n)}\right)^{1/2}C_n^u(x)C_{h-n}^v(x), \tag{3.22}$$

where the coefficients $E(\cdot)$ can be simply evaluated.

Our next model of D_u^+ is defined by operators

$$J^+ = e^{i\theta}\left(x(1-x)\frac{\partial}{\partial x} - i\frac{\partial}{\partial\theta} - qx\right),$$
$$J^- = e^{i\theta}\left(x\frac{\partial}{\partial x} + i\frac{\partial}{\partial\theta}\right), \qquad (3.23)$$
$$J^3 = -i\frac{\partial}{\partial\theta},$$

and basis functions

$$j_n(x,\theta) = \left(\frac{\Gamma(2u+n)}{n!}\right)^{1/2} x^u F(-n, u+q; 2u; x) e^{i(u+n)\theta}, \quad n = 0, 1, 2, \ldots, \qquad (3.24)$$

where q is a constant and $F(\alpha,\beta;\gamma,x)$ is a hypergeometric function (see Ref. 7, Vol. 1). It follows easily that the operators (3.23) and basis functions

$$j_{n,m}(x,\theta) = \left(\frac{\Gamma(2u+n)\Gamma(2v+m)}{n!m!}\right)^{1/2} x^{u+v} \times F(-n, u+q_1; 2u; x) \times F(-m, v+q_2; 2v; x)\, e^{i(u+v+n+m)\theta}, \quad n, m = 0, 1, 2, \ldots, \qquad (3.25)$$

define a model of $D_u^+ \otimes D_v^+$ where $q = q_1 + q_2$. From (3.24) we see that the basis vectors j_h^s transforming according to D^+_{u+v+s} are given by

$$j_h^s(x,\theta) = c_s\left(\frac{\Gamma(2u+2v+2s+h)}{h!}\right)^{1/2} x^{u+v+s} \times F(-h, u+v+s+q; 2u+2v+2s; x) \times e^{i(u+v+s+h)\theta}, \quad s, h = 0, 1, 2, \cdots. \qquad (3.26)$$

To compute the constants c_s we substitute (3.25) and (3.26) into (3.7) and equate coefficients of x^{u+v+s} on both sides of the resulting expression. We find

$$c_s = \left(\frac{\Gamma(2u+2v+s-1)\Gamma(2u+s)}{s!\,(2u+2v+2s-1)\Gamma(2v+s)}\right)^{1/2} \times \frac{\Gamma(2v)\Gamma(v+q_2+s)}{\Gamma(2u+2v+2s-1)\Gamma(v+q_2)}\, {}_3F_2(-s, u+q_1, -2v-s+1; 2u, -v-q_2-s+1; 1). \qquad (3.27)$$

Substituting (3.25)–(3.27) into (3.5), we obtain our general identity. In the special case $s = 0$, it reduces to

$$\left(\frac{\Gamma(2u)\Gamma(2v)\Gamma(2u+2v+h)}{\Gamma(2u+2v)\,h!}\right)^{1/2} \times F(-h, u+v+q; 2u+2v; x) = \sum_{n=0}^{h} E(u, n, ; v, h-n\,|\,0, h) \times \left(\frac{\Gamma(2u+n)\Gamma(2v+h-n)}{n!(h-n)!}\right)^{1/2} \times F(-n, u+q_1; 2u; x) \times F(-h+n, v+q_2; 2v; x). \qquad (3.28)$$

The reader can discover other interesting special cases of this general identity by varying q_1 and q_2, e.g., set $q_1 = -u$.

4. A DIFFERENCE OPERATOR MODEL

As shown in Refs. 3 or 13, the operators

$$J^+ = e^{i\theta}\left((x-1)L - x - i\frac{\partial}{\partial\theta} + q\right),$$
$$J^- = e^{-i\theta}\left(-(x+r)E + x + i\frac{\partial}{\partial\theta} + r + q - 1\right), \qquad (4.1)$$
$$J^3 = -i\frac{\partial}{\partial\theta}$$

and basis functions

$$j_n(x,\theta) = \left(\frac{\Gamma(2u+n)}{n!}\right)^{1/2} \times \frac{\Gamma(u-q+1)\Gamma(x+r+q-u-n-1)}{\Gamma(x+r-n)} \times {}_3F_2(-n, u-q-r+1, u-q+1; 2u, x+r-n; 1)\, e^{i(u+n)\theta}, \quad n = 0, 1, 2, \cdots \qquad (4.2)$$

form a model of D_u^+, where r, q are constants and $Ef(x,\theta) = f(x+1,\theta)$, $Lf(x,\theta) = f(x-1,\theta)$. Furthermore, the operators

$$J^+ = e^{i\theta}\left(-i\frac{\partial}{\partial\theta} + v\right), \quad J^- = e^{-i\theta}\, i\frac{\partial}{\partial\theta} + v, \quad J^3 = -i\frac{\partial}{\partial\theta}, \qquad (4.3)$$

and basis functions

$$j'_m(\theta) = \sqrt{\frac{\Gamma(2v+m)}{m!}}\, e^{i(v+m)\theta}, \quad m = 0, 1, 2, \ldots, \qquad (4.4)$$

form a model of D_v^+. Thus the operators (4.1) and basis functions

$$j_{n,m}(x,\theta) = j_n(x,\theta)j'_m(\theta) = \left(\frac{\Gamma(2u+n)\Gamma(2v+m)}{n!m!}\right)^{1/2} \times \frac{\Gamma(u-q+v+1)\Gamma(x+r+q-u-v-n-1)}{\Gamma(x+r-n)} \times {}_3F_2(-n, u+v-q-r+1, u+v-q+1; 2u, x+r-n; 1)\, e^{i(u+v+n+m)\theta}, \quad n, m = 0, 1, 2, \cdots, \qquad (4.5)$$

define a model of $D_u^+ \otimes D_v^+$. The basis functions $j_h^s(x,\theta)$ transforming according to D^+_{u+v+s} can be obtained immediately from (4.2) with u replaced by $u+v+s$ and $n = h$:

$$j_h^s(x,\theta) = c_s\left(\frac{\Gamma(2u+2v+2s+h)}{h!}\right)^{1/2} \times \frac{\Gamma(u+v+s-q+1)}{\Gamma(x+r-h)} \times \Gamma(x+r+q-u-v-s-h-1) \times {}_3F_2(-h, u+v+s-q-r+1, u+v+s-q+1; 2u+2v+2s, x+r-h; 1) \times e^{i(u+v+s+h)\theta}. \qquad (4.6)$$

To compute the constants c_s, we substitute (4.5) and (4.6) into (3.7) and set $x = -r + 1$. We can then sum the right-hand side to obtain

$$c_s = \left(\frac{\Gamma(2u+2v+s-1)\Gamma(2u+s)\Gamma(2v+s)}{s!\,(2u+2v+2s-1)}\right)^{1/2} \times \frac{(-1)^s}{\Gamma(2u+2v+2s-1)} \times {}_3F_2(-s, u+v-q-r+1, u+v-q+1; \times 2u, u+v-q+1; 1). \quad (4.7)$$

[1] W. Miller, J. Math Phys. **13**, 648 (1972).
[2] W. Miller, *Lie Theory and Special Functions* (Academic, New York, 1968).
[3] W. Miller, J. Math. Anal. Appl. **28**, 383 (1969).
[4] W. Miller, SIAM J. Math. Anal. **2**, 307 (1971).
[5] L. Armstrong, J. Math. Phys. **12**, 953 (1971).
[6] M. Cunningham, J. Math. Phys. **13**, 33 (1972).
[7] A. Erdelyi *et al.*, *Higher Transcendental Functions* (Bateman Manuscript Project (McGraw-Hill, New York, 1953), Vols. I, II.
[8] V. Bargmann, Ann. Math. **48**, 568 (1947).
[9] P. Sally, "Analytic Continuation of the Irreducible Unitary Representations of the Universal Covering Group of SL(2R)," AMS Mem., No. 69, Providence, R. I. (1967).
[10] T. Regge, Nuovo Cimento **10**, 544 (1958).
[11] V. Bargmann, Rev. Mod. Phys. **34**, 829 (1962).
[12] J. Jackson, *Classical Electrodynamics* (Wiley, New York, 1962).
[13] W. Miller, "Lie Theory and Difference Equations. II," J. Math. Anal. Appl. (to be published).

Spectrum Generating Algebras and Symmetries in Mechanics. I

Robert Hermann*

Department of Mathematics, Rutgers University, New Brunswick, New Jersey

(Received 17 January 1971)

Certain differential-geometric and Lie group theoretic facts that are useful in the systematic study and search for spectrum generating algebras are presented.

1. INTRODUCTION

Dothan has pointed out[1] several ideas and possible directions of research in connection with the "spectrum generating Lie algebras" of quantum mechanics. Typically, such ideas also have analogs in classical mechanics. Since the problems in classical mechanics often have a geometric foundation, one finds interconnections between geometry and Lie group theory, quantum mechanics and elementary particle physics. The aim of this paper is to survey more extensively some of these links than was possible in Dothan's paper.

2. POISSON BRACKETS STRUCTURES AND CANONICAL TRANSFORMATIONS ON MANIFOLD

We adopt Ref. 2 as a basic reference for the ideas and notations of differential geometry on manifolds. Let M be a manifold of even dimension, with a closed two-differential form ω of maximal rank on M. A diffeomorphism $\phi: M \to M$ is a *canonical transformation* if ϕ preserves the form ω, i.e., $\phi^*(\omega) = \omega$. A vector field $X \in V(M)$ defines an *infinitesimal canonical* transformation if

$$X(\omega) = 0 \quad (2.1)$$

[$X(\omega)$ denotes the Lie derivative[2] of the form ω by the vector field X]. The set of vector fields X satisfying (2.1) forms a Lie algebra [under Jacobi bracket $(X, Y) \to [X, Y]$] of vector fields, that we denote by $V(\omega)$. It may be thought of as the "Lie algebra" of the group of canonical transformations.

Let $F(M)$ denote the C^∞, real-valued functions on M. The form ω defines a Lie algebra structure $(f_1, f_2) \to \{f_1, f_2\}$ called the *Poisson bracket*. To define it, for $f \in F(M)$, let X_f be the vector field such that

$$df = X_f \lrcorner \omega. \quad (2.2)$$

Set

$$\{f_1, f_2\} = -X_{f_1}(f_2) \quad (2.3)$$

for $f_1, f_2 \in F(M)$.

Then, one can prove the following results.

The bracket $\{\ ,\ \}$ defined by 2.3 makes $F(M)$ into a Lie algebra. (2.4)

The mapping $f \to X_f$ is a Lie algebra homomorphism of $F(M)$ into $V(\omega)$. (2.5)

The kernel of this homomorphism consists of the constant functions on M.

To recover the classical expression for Poisson bracket to be found in all mechanics books, suppose (p_i, q_i), $1 \le i, j \le m$ is a coordinate system for M such that

$$\omega = dp_i \wedge dq_i.$$

Then, for $f \in F(M)$,

$$X_f = -\frac{\partial f}{\partial p_i}\frac{\partial}{\partial q_i} + \frac{\partial f}{\partial q_i}\frac{\partial}{\partial p_i}. \quad (2.6)$$

Given $h \in F(M)$, the integral curves of the vector field X_h are the solutions of Hamilton's equations, with h the Hamiltonian.[2] Thus, if h is the function that represents the total energy of the mechanical system, a basic problem is to study these integral curves, i.e., to study the one-parameter group of canonical transformations generated by X_h. Now, in Ref. 2 certain general insights of the "Lie theory" of ordinary differential equations have been explained. In particular, they apply to the problem of finding the integral curves of X_h.

Definition: A function $f \in F(M)$ is a *symmetry* of h if

$$\{f, h\} = 0. \quad (2.7)$$

If f satisfies (2.7), then it follows from (2.5) that $[X_f, X_h] = 0$, hence that the one-parameter group generated by X_f and X_h commute. The aim of the theory of "spectrum generating algebras", stated in rather vague terms, is to study Lie subalgebras of $F(M)$, whose elements f satisfy commutation relations

To compute the constants c_s, we substitute (4.5) and (4.6) into (3.7) and set $x = -r + 1$. We can then sum the right-hand side to obtain

$$c_s = \left(\frac{\Gamma(2u+2v+s-1)\Gamma(2u+s)\Gamma(2v+s)}{s!\,(2u+2v+2s-1)}\right)^{1/2} \times \frac{(-1)^s}{\Gamma(2u+2v+2s-1)} \times {}_3F_2(-s, u+v-q-r+1, u+v-q+1; \times 2u, u+v-q+1; 1). \quad (4.7)$$

[1] W. Miller, J. Math Phys. **13**, 648 (1972).
[2] W. Miller, *Lie Theory and Special Functions* (Academic, New York, 1968).
[3] W. Miller, J. Math. Anal. Appl. **28**, 383 (1969).
[4] W. Miller, SIAM J. Math. Anal. **2**, 307 (1971).
[5] L. Armstrong, J. Math. Phys. **12**, 953 (1971).
[6] M. Cunningham, J. Math. Phys. **13**, 33 (1972).
[7] A. Erdelyi *et al.*, *Higher Transcendental Functions* (Bateman Manuscript Project (McGraw-Hill, New York, 1953), Vols. I, II.
[8] V. Bargmann, Ann. Math. **48**, 568 (1947).
[9] P. Sally, "Analytic Continuation of the Irreducible Unitary Representations of the Universal Covering Group of SL(2R)," AMS Mem., No. 69, Providence, R. I. (1967).
[10] T. Regge, Nuovo Cimento **10**, 544 (1958).
[11] V. Bargmann, Rev. Mod. Phys. **34**, 829 (1962).
[12] J. Jackson, *Classical Electrodynamics* (Wiley, New York, 1962).
[13] W. Miller, "Lie Theory and Difference Equations. II," J. Math. Anal. Appl. (to be published).

Spectrum Generating Algebras and Symmetries in Mechanics. I

Robert Hermann*

Department of Mathematics, Rutgers University, New Brunswick, New Jersey

(Received 17 January 1971)

Certain differential-geometric and Lie group theoretic facts that are useful in the systematic study and search for spectrum generating algebras are presented.

1. INTRODUCTION

Dothan has pointed out[1] several ideas and possible directions of research in connection with the "spectrum generating Lie algebras" of quantum mechanics. Typically, such ideas also have analogs in classical mechanics. Since the problems in classical mechanics often have a geometric foundation, one finds interconnections between geometry and Lie group theory, quantum mechanics and elementary particle physics. The aim of this paper is to survey more extensively some of these links than was possible in Dothan's paper.

2. POISSON BRACKETS STRUCTURES AND CANONICAL TRANSFORMATIONS ON MANIFOLD

We adopt Ref. 2 as a basic reference for the ideas and notations of differential geometry on manifolds. Let M be a manifold of even dimension, with a closed two-differential form ω of maximal rank on M. A diffeomorphism $\phi: M \to M$ is a *canonical transformation* if ϕ preserves the form ω, i.e., $\phi^*(\omega) = \omega$. A vector field $X \in V(M)$ defines an *infinitesimal canonical* transformation if

$$X(\omega) = 0 \quad (2.1)$$

[$X(\omega)$ denotes the Lie derivative[2] of the form ω by the vector field X]. The set of vector fields X satisfying (2.1) forms a Lie algebra [under Jacobi bracket $(X, Y) \to [X, Y]$] of vector fields, that we denote by $V(\omega)$. It may be thought of as the "Lie algebra" of the group of canonical transformations.

Let $F(M)$ denote the C^∞, real-valued functions on M. The form ω defines a Lie algebra structure $(f_1, f_2) \to \{f_1, f_2\}$ called the *Poisson bracket*. To define it, for $f \in F(M)$, let X_f be the vector field such that

$$df = X_f \lrcorner \omega. \quad (2.2)$$

Set

$$\{f_1, f_2\} = -X_{f_1}(f_2) \quad (2.3)$$

for $f_1, f_2 \in F(M)$.

Then, one can prove the following results.

The bracket $\{\ ,\ \}$ defined by 2.3 makes $F(M)$ into a Lie algebra. (2.4)

The mapping $f \to X_f$ is a Lie algebra homomorphism of $F(M)$ into $V(\omega)$. (2.5)

The kernel of this homomorphism consists of the constant functions on M.

To recover the classical expression for Poisson bracket to be found in all mechanics books, suppose (p_i, q_i), $1 \le i, j \le m$ is a coordinate system for M such that

$$\omega = dp_i \wedge dq_i.$$

Then, for $f \in F(M)$,

$$X_f = -\frac{\partial f}{\partial p_i}\frac{\partial}{\partial q_i} + \frac{\partial f}{\partial q_i}\frac{\partial}{\partial p_i}. \quad (2.6)$$

Given $h \in F(M)$, the integral curves of the vector field X_h are the solutions of Hamilton's equations, with h the Hamiltonian.[2] Thus, if h is the function that represents the total energy of the mechanical system, a basic problem is to study these integral curves, i.e., to study the one-parameter group of canonical transformations generated by X_h. Now, in Ref. 2 certain general insights of the "Lie theory" of ordinary differential equations have been explained. In particular, they apply to the problem of finding the integral curves of X_h.

Definition: A function $f \in F(M)$ is a *symmetry* of h if

$$\{f, h\} = 0. \quad (2.7)$$

If f satisfies (2.7), then it follows from (2.5) that $[X_f, X_h] = 0$, hence that the one-parameter group generated by X_f and X_h commute. The aim of the theory of "spectrum generating algebras", stated in rather vague terms, is to study Lie subalgebras of $F(M)$, whose elements f satisfy commutation relations

with h which are more complicated than (2.7), yet still sufficiently simple to be able to derive useful information about the physical system. Certain useful general remarks about this program have been made by Dothan.[1] Hence, we will not attempt to further formalize the general features of the program, but will instead proceed to study certain more specific facts of a differential geometric nature. In later papers in this series, we will return to more general speculations.

3. TRANSITIVE LIE GROUPS OF CANONICAL TRANSFORMATIONS

Let us suppose now that **G** is a finite dimensional Lie subalgebra of $F(M)$. Let G be the simply connected Lie group, whose Lie algebra is **G**. Let us suppose that **G** arises as the infinitesimal version of an action of G on M as a group of canonical transformations. Further, we will suppose that G acts transitively on M. (The "spectrum generating algebras" of such physical systems as the harmonic oscillator and the hydrogen atom seem to satisfy this transitivity condition, so it seems to be a reasonable speculation.) Let L be the isotropy subgroup of G at a point of M. Then, by the general principles of Lie group transformation theory, M can be identified with the coset space G/L. Our aim in this section is to investigate the general conditions of the existence of a closed two-form ω on M that is invariant under G imposes on **G** and **L**. We will suppose, for simplicity, that L is a connected subgroup of G. Further, in this section we will not suppose *a priori* that ω is a maximal rank form nor that M is even dimensional.

Theorem 3.1: The coset space G/L admits a two-differential form invariant under G if and only if there is a skew-symmetric bilinear form ω: $\mathbf{G} \times \mathbf{G} \to R$ such that

$$\omega(\mathbf{L}, \mathbf{G}) = 0, \tag{3.1}$$

$$\omega([Y, X], Z) - \omega([Z, X], Y) - \omega(X, [Z, Y)] = 0 \quad \text{for } X, Y, Z \in \mathbf{G}, \tag{3.2}$$

$$\omega([X, Y], Z) + \omega(Y, [X, Z]) = 0 \quad \text{for } X \in \mathbf{L}, \; Y, Z \in \mathbf{G}. \tag{3.3}$$

Proof: Since G acts as a transformation group on M, by the general principles of Lie theory **G** can be identified with a Lie subalgebra of $V(M)$. Let p be the point of M for which L is the isotropy subgroup. Then,

$$\mathbf{L} = \{X \in \mathbf{G}: X(p) = 0\}. \tag{3.4}$$

Suppose first that a closed two-form ω is given on M, such that

$$g^*(\omega) = \omega \quad \text{for all } g \in G.$$

Then, also,

$$X(\omega) = 0 \quad \text{for } X \in \mathbf{G}. \tag{3.5}$$

Define ω by the following formula:

$$\omega(X, Y) = \omega(X, Y)(p) \quad \text{for } X, Y \in \mathbf{G}. \tag{3.6}$$

Then, it is readily verified that (3.1)–(3.3) follow from (3.4)–(3.6) and the condition that $d\omega = 0$. The steps are reversible, and show that such an ω gives rise to an ω satisfying (3.6).

Kostant has given[3] a general procedure for the construction of a form ω satisfying (3.1)–(3.3). Let $\mathbf{G}^d$ denote the dual space to the vector space **G**, i.e., the space of linear mappings α: $\mathbf{G} \to R$.

Theorem 3.2: Suppose that **L** is a Lie subalgebra of **G** and α is an element of $\mathbf{G}^d$ which satisfies the following condition:

$$\alpha([\mathbf{L}, \mathbf{G}]) = 0. \tag{3.7}$$

Then, ω defined by the following formula satisfies conditions (3.1)–(3.3), hence defines a G-invariant, closed two-form ω on G/L:

$$\omega(X, Y) = \alpha([X, Y]) \quad \text{for } X, Y \in \mathbf{G}. \tag{3.8}$$

Proof: Notice that (3.7) is equivalent to (3.1). The other two conditions, (3.2)–(3.3), follow from the Jacobi identity for the Lie algebra **G**.

Kostant's formula (3.8) gives one class of homogeneous spaces G/L with invariant closed 2-forms. We shall now prove that, if **G** is semisimple, all structures are of this type.

Let **G** be a Lie algebra, with ω a form satisfying (3.1)–(3.3). Suppose further that B is a symmetric nonsingular bilinear form $\mathbf{G} \times \mathbf{G} \to R$ that is invariant under the adjoint representation, i.e., satisfies the following condition:

$$B([X, Y], Z) + B(Y, [X, Z]) = 0 \quad \text{for } X, Y, Z \in \mathbf{G}. \tag{3.9}$$

(For example, if **G** is semisimple, the Killing form[4] will satisfy this condition.) Then, there is a linear transformation A: $\mathbf{G} \to \mathbf{G}$, such that

$$\omega(X, Y) = B(AX, Y) \quad \text{for } X, Y \in \mathbf{G}. \tag{3.10}$$

Theorem 3.3: If A satisfies (3.10), then it is a derivation of **G** that commutes with $\mathbf{Ad}\mathbf{L}$, i.e., it satisfies the following conditions:

$$[X, AY] = A[X, Y] \quad \text{for } X \in \mathbf{L}, \; Y \in \mathbf{G}, \tag{3.11}$$

$$A([X, Y]) = [AX, Y] + [X, AY] \quad \text{for } X, Y \in \mathbf{G}. \tag{3.12}$$

Proof: Let us first prove (3.11). Given $X \in \mathbf{L}$, $Y \in \mathbf{G}$, set

$$A'(X, Y) = [X, AY] - A([X, Y]). \tag{3.13}$$

Then, for $Z \in \mathbf{G}$,

$$\begin{aligned} B(A'(X, Y), Z) &= B([X, AY], Z) - B(A[X, Y], Z) \\ &= -B(AY, [X, Z]) - B(A[X, Y], Z), \text{ using (3.9)}, \\ &= -\omega(Y, [X, Z]) - \omega([X, Y], Z), \text{ using (3.10)}, \\ &= 0, \end{aligned}$$

using (3.3). Since B is nonsingular, (3.11) follows.

To prove (3.12), set

$$A''(X, Y) = A([X, Y]) - [AX, Y] - [X, AY] \quad \text{for } X, Y \in \mathbf{G}.$$

Given $Z \in \mathbf{G}$,

$$\begin{aligned} B(A''(X,Z),Z) &= B(A[X,Y],Z) - B([AX,Y],Z) \\ &\quad - B([X,AY],Z) \\ &= \omega([X,Y],Z) + B(AX,[Z,Y]) + B(AY,[X,Z]) \\ &= \omega([X,Y],Z) + \omega(X,[Z,Y]) + \omega(Y,[X,Z]) \\ &= 0, \end{aligned}$$

using (3.2). Again, (3.12) now follows from the fact that B is nonsingular.

Corollary: If $\mathbf{G}$ is semisimple, there is an element $Z \in G$, such that

$$[Z, \mathbf{L}] = 0, \tag{3.14}$$

$$\omega(X,Y) = -B(Z,[X,Y]) \quad \text{for } X, Y \in \mathbf{G}. \tag{3.15}$$

In particular, ω is of the type given by Kostant's formula. Further, the form ω on G/L defined by ω is of maximal rank, with G/L even dimensional, if and only if $\mathbf{L}$ is the centralizer of Z in $\mathbf{G}$, i.e.,

$$\mathbf{L} = \{X \in \mathbf{G}: [X,Z] = 0\}. \tag{3.16}$$

Proof: A well-known[4] theorem of Cartan asserts that any derivation of a semisimple Lie algebra is inner. In particular, since A is a derivation, there is a $Z \in \mathbf{G}$, such that

$$A(Y) = \mathrm{Ad}Z(Y) = [Z,Y] \quad \text{for } Y \in \mathbf{G}.$$

Equations (3.14)–(3.16) now follow readily from this fact.

Remark: The referee has remarked that the above argument can be sharpened to cover the more general hypothesis (than semisimplicity) that G has vanishing second real-coefficient cohomology groups. In the case that w is a maximal rank form on an even-dimensional M, this result is indeed known.[3] Of course, our aim is to apply it to the further case where M is odd dimensional, e.g., the "energy surface" of the phase space of mechanical system.

We can now apply these results to various classification problems that are associated with Dothan's general ideas concerning spectrum generating algebras.

4. GENERAL FACTS ABOUT SPECTRUM GENERATING ALGEBRAS

Let us now turn to the case where M is the phase space of a classical mechanical system consisting of a single particle. Then,

$$\dim M = 6. \tag{4.1}$$

Suppose that G is a semisimple group of canonical transformations that acts transitively on M. Let $\mathbf{G}$ be the Lie algebra of G, identified with a subalgebra of $V(\omega)$ and $F(M)$. Let L be the isotropy subgroup of G at a point $p \in M$. Then $\mathbf{G}$ and $\mathbf{L}$ satisfy (3.14)–(3.16). Further, let us suppose that h is a function on M that is the Hamiltonian of the physical system. Suppose that K is a connected subgroup of G that consists of symmetries of h, i.e., such that

$$[\mathbf{K}, h] = 0. \tag{4.2}$$

Let N be the "energy surface" at p, i.e., the set of points $p' \in M$, such that

$$h(p') = h(p). \tag{4.3}$$

Then, in view of (4.2), K maps N into itself. Let us suppose—as part of what is meant by "spectrum generating algebra"—that the following condition is satisfied:

$$K \text{ acts transitively on } N. \tag{4.4}$$

Let B be the Killing form of the Lie algebra $\mathbf{G}$. Let ω be the closed, maximal rank two-form on M defining the Poisson bracket structure. [In terms of the usual canonical coordinates (p_i, q_i), $1 \le i \le 3$, ω is the form $dp_i \wedge dq_i$]. Then, there exists an element $Z \in \mathbf{G}$, such that $\mathbf{L}$ is the centralizer of Z in $\mathbf{G}$ and

$$\begin{aligned} \omega(X,Y) &= -B(Z,[X,Y]) \\ &= B([X,Z],Y) \quad \text{for } X, Z \in \mathbf{L}. \end{aligned} \tag{4.5}$$

Further, we have the following result.

Theorem 4.1: If $\mathbf{A}$ is an Abelian subalgebra of $\mathbf{G}$, then

$$\dim\mathbf{A} - \dim(\mathbf{A} \cap \mathbf{L}) \le 3. \tag{4.6}$$

If $\mathbf{A}'$ is an Abelian subalgebra of $\mathbf{K}$, then

$$\dim\mathbf{A}' - \dim(\mathbf{A}' \cap \mathbf{L}) \le 2. \tag{4.7}$$

Proof: By (4.5) if X, Y are elements of an Abelian subalgebra of $\mathbf{G}$, then

$$\omega(X,Y) = 0. \tag{4.8}$$

Now, ω passes to the quotient to define a nonsingular bilinear, skew-symmetric form on $\mathbf{G/L}$, which is of dimension six, which we may denote by ω'. Now, a linear subspace $V \subset \mathbf{G/L}$ is said to be *isotropic* if $\omega'(V,V) = 0$. It is a fact[5] of linear algebra that the dimension of such an isotropic subspace is of dimension no greater than three. Now, the image of $\mathbf{A}$ in $\mathbf{G/L}$ is such an isotropic subspace, by (4.8). Combining these two remarks proves (4.6). (4.7) is proved similarly.

5. MAXIMAL ORBITS

We will now abstract from the physicist's intuitive idea of what is meant by a "spectrum generating algebra," a precise problem in Lie group theory. Suppose, for the moment, that G, L, K are Lie groups, with L and K subgroups of G, and with $\mathbf{G}$ a "spectrum generating algebra". We have seen in Sec. 4 that one of the conditions going into the "spectrum generating" condition is $\dim(G/L) = 6$.

The dimension of at least one orbit of K is 5. This suggests the following general problem:

> Classify all triples (G, L, K), with G, L, K connected Lie groups, L and K closed subgroups of G, such that the following condition is satisfied: There is at least one orbit of K on G/L, whose dimension is one less than the dimension of G/L. (5.1)

In this section, we will make some general remarks about these possibilities, in the case where K is compact. (As Dothan remarks,[1] this case seems to be the plausible one when one seeks "spectrum generating algebras" whose representations hopfully define the "bound-states" of a quantum mechanical physical system.) In this case, one can use the known general facts[2] concerning the structure of compact Lie groups of transformations on manifolds. In the rest of this section, we will summarize what can be said using these facts.

Let $M = G/L$ and suppose that K is compact. Then, there is a positive Riemannian metric on M, such that K acts as a group of isometries of this metric. We will suppose that one such metric is fixed on M and that it is complete. For $p \in M$, let K^p be the isotropic subgroup of K at p i.e., the set of elements $g \in K$, such that $gp = p$. Recall[2] the definition of *maximal orbit:* A point $p \in M$ is on such an orbit if the following conditions are satisfied:

$$\dim K^p \leq \dim K^{p'} \quad \text{for all points } p' \in M. \tag{5.2}$$

Let p be a point of M, and let N be the orbit of K through the point p. Let N_p be the tangent space to N at p. Let $N_p^\perp$ be the orthogonal complement to N_p in M_p (with respect to the fixed Riemannian metric on M). Then, K^p acts on tangent vectors to M at p.

$$g_v = g_*(v) \quad \text{for } v \in M_p,\ g \in G. \tag{5.3}$$

The action (5.3) of K^p defines a representation of K by linear tranformations on M_p, called the *tangent linear isotropy representation.* The infinitesimal version of this defines a linear representation of $\mathbf{K}^p$, the Lie algebra of K^p, by linear transformations on M_p.

The following result can be proved, using the techniques for the study of orbits of Riemannian transformation groups given in Ref. 2.

Theorem 5.1: p lies on a *maximal orbit* of K if and only if

$$Xv = 0 \quad \text{for all } X \in K^p,\ \text{all } v \in N_p^\perp. \tag{5.4}$$

Let us now translate condition (5.4) into an algebraic condition involving G. Let us suppose that G is a semisimple Lie algebra, and that K and L are linear subspaces that are nonsingular with respect to the Killing form of G. Let $\mathbf{K}^\perp$ and $\mathbf{L}^\perp$ denote the orthogonal complement of K and L in G, with respect to the Killing form of G. Then, one has the following relations:

$$\mathbf{G} = \mathbf{K} \oplus \mathbf{K}^\perp = \mathbf{L} \oplus \mathbf{L}^\perp. \tag{5.5}$$

(The direct sum in (5.5) is only a vector space direct sum.)

$$[\mathbf{K}, \mathbf{K}^\perp] \subset \mathbf{K}^\perp, \quad [\mathbf{L}, \mathbf{L}^\perp] \subset \mathbf{L}^\perp.$$

Suppose further that p is the identity coset element of G/K, so that the isotropy subgroup of G at p is L. Then, one sees that condition (5.4) is equivalent to the following condition:

$$[\mathbf{K} \cap \mathbf{L}, \mathbf{K}^\perp] \subset \mathbf{L}. \tag{5.6}$$

Now, the question arises of translating (5.6) into a more familiar algebraic condition on the triple (G, K, L). We will attempt to do this here only in the case where K and L are both symmetric subalgebras of G, so that the following conditions are satisfied:

$$[\mathbf{K}^\perp, \mathbf{K}^\perp] \subset \mathbf{K}, \quad [\mathbf{L}^\perp, \mathbf{L}^\perp] \subset \mathbf{L}. \tag{5.7}$$

In addition, suppose that

$$\mathbf{K} \cap \mathbf{L} \text{ is a nonsingular subspace of G with respect to the Killing form of G.} \tag{5.8}$$

Theorem 5.2: If conditions (5.7), (5.8) are satisfied, then condition (5.6) implies the following two conditions:

$$[\mathbf{K}^\perp \cap \mathbf{L}^\perp, \mathbf{K}^\perp \cap \mathbf{L}^\perp] = 0, \tag{5.9}$$

$$[\mathbf{K} \cap \mathbf{L}, \mathbf{K}^\perp \cap \mathbf{L}^\perp] = 0. \tag{5.10}$$

Proof: Let $B(\ ,\ )$ be the Killing form of the semisimple Lie algebra G. Assume (5.6); 5.10 then follows automatically. To prove (5.9), proceed as follows:

Suppose $X, Y \in \mathbf{K}^\perp \cap \mathbf{L}^\perp$, $Z \in \mathbf{K} \cap \mathbf{L}$. Then,

$$B([X, Y], Z) = -B(Y, [X, Z]) = 0, \quad \text{by (5.6).}$$

Then, $B([\mathbf{K}^\perp \cap \mathbf{L}^\perp, \mathbf{K}^\perp \cap \mathbf{L}^\perp], \mathbf{K} \cap \mathbf{L}) = 0.$ (5.11)

But, using (5.7), we have

$$[\mathbf{K}^\perp \cap \mathbf{L}^\perp, \mathbf{K}^\perp \cap \mathbf{L}^\perp] \subset \mathbf{K} \cap \mathbf{L}. \tag{5.12}$$

Because of (5.8), (5.11) and (5.12) combined to prove (5.9).

Finally, in case (5.1) is satisfied, we have the following condition:

$$\dim (\mathbf{K}^\perp \cap \mathbf{L}^\perp) = 1. \tag{5.13}$$

These results give us the tools to investigate examples of "spectrum generating algebras."

6. SPECTRUM GENERATING ALGEBRAS ASSOCIATED WITH COMPACT SYMMETRIC PAIRS

Let us first recall some ideas from the theory of symmetric spaces.[4,6,7] Let G be a Lie algebra, and let σ be an automorphism of G, such that

$$\sigma^2 = (\text{identity}). \tag{6.1}$$

Let $\mathbf{K} = \{X \in \mathbf{G}: \sigma|X| = X\}$. Then K is called a *symmetric subalgebra* of G. A pair (K, L) of such symmetric subalgebras is called a *symmetric pair.* Such a symmetric pair will be called a *compact symmetric pair* if G is a semisimple Lie algebra of compact type, i.e., G is the Lie algebra of a compact, connected, semisimple Lie group G. (In Sec. 7 we will extend some of the results to certain noncompact groups G.)

Let us suppose that such a compact symmetric pair (K, L) is fixed. Let K and L be the connected sub-

groups of G generated by **K** and **L**. In Refs. 7–9, certain general results are proved concerning the orbit structure of K acting on G/L. In this section, we will apply some of this work to the problem of classifying spectrum generating algebras.

Let σ and σ' be the automorphisms of G which satisfy (6.1), such that

$$\mathbf{K} = \{X \in \mathbf{G}\colon \sigma(X) = X\}, \qquad \mathbf{L} = \{X \in \mathbf{G}\colon \sigma'(X) = X\}.$$

Set

$$\mathbf{H} = \{X \in \mathbf{G}\colon \sigma(X) = \sigma'(X)\}. \tag{6.2}$$

Then, **H** is a subalgebra of **G**, such that

$$\sigma(\mathbf{H}) = \mathbf{H}.$$

Also, $\mathbf{K} \cap \mathbf{L} = \{X \in \mathbf{H}\colon \sigma(X) = X\}$. Thus, $\mathbf{K} \cap \mathbf{L}$ is a symmetric subalgebra of **H**. Set

$$\mathbf{P} = \{X \in \mathbf{H}\colon \sigma(X) = -X\}. \tag{6.3}$$

Let **A** be a maximal Abelian subalgebra of **P**, i.e., **A** is a Cartan Subalgebra[4,6] associated with the symmetric space $H/K \cap L$. Let A be the connected subgroup of **G** corresponding to the subalgebra **A**. We will now state a basic result, referring to Ref. 8 for the proof. (Alternately, a proof could be given using Theorem 5 and general facts about Riemannian transformation groups.)

Theorem 6.2: Let p_0 be the identity coset of $M = G/L$. Then, the orbit Ap_0 touches each orbit of K at least once and is transversal to each such orbit at the point of contact. At each maximal orbit, Ap_0 fills up the perpendicular space to the orbits. In particular,

$$\dim\mathbf{A} = \dim M - (\text{dimension of each maximal orbit of } K). \tag{6.4}$$

Corollary: The maximal orbits of K are of one less dimension than M if and only if $H/K \cap L$ is a symmetric space of rank one.

[The corollary is of obvious interest in connection with finding examples which satisfy (5.1).]

Now, we can made further remarks about the possible structure of **L** in case the following condition is satisfied:

$$\sigma\sigma' = \sigma'\sigma. \tag{6.5}$$

In this case, set $\sigma'' = \sigma\sigma'$. Notice that $(\sigma'')^2 = \text{identity}$, and

$$\mathbf{H} = \{X \in \mathbf{G}\colon \sigma''(X) = X\}. \tag{6.6}$$

In particular, **H** is also a symmetric subalgebra of **G**, hence the possibilities for $\dim A = 1$ can be classified using Cartan's classification of compact symmetric spaces. However, we will not carry out this program here.

Now, **G** is a candidate for a "spectrum generating algebra" if G/L has a σ-invariant Poisson bracket structure. By Theorem 3.3, this will happen if and only if **L** is the centralizer of an element of $X \in \mathbf{G}$.

In this case, **L** will be a maximal rank subalgebra of **G**. If **G** is simple, the possibilities for **L** may be read off from the table on page 128 of Ref. 4.

For example, if $G = SO(5, R)$, the only possibility for L is the $SO(3, R) \times SO(2, R)$ subgroup of G. Now, as Dothan remarks,[1] the spectrum generating algebra for the hydrogen atom is $G = SO(4, 1)$, which is a noncompact real form of $SO(5, R)$. (As we will see in the next section, it is often possible to pass back and forth from the "compact" to the "noncompact" situations using Cartan's construction of noncompact Lie algebras.) Now, in the hydrogen atom, K is the subgroup $SO(4, R)$. It may now be readily verified, with the aid of (6.6), that H is also an $SO(4, R)$ subgroup of $SO(5, R)$ with $K \cap L$ equal to $SO(3, R)$. Notice now that $H/K \cap L$ is indeed a rank one symmetric space, hence to maximal orbits of K on G/L are five-dimensional, as they are in the hydrogen atom.

Another example is provided by the harmonic oscillator, where $G = SU(4)$, and where K and L are both $U(3)$ subgroups. In this case also it is readily seen that the maximal orbits of K on G/L are five-dimensional, while $\dim(G/L) = 6$, and G/L as a G-invariant Poisson bracket structure. Now we turn to the description of a method for generating noncompact spectrum generating algebras.

7. CONSTRUCTION OF NONCOMPACT SPECTRUM GENERATING ALGEBRAS

We have just given a method for generating compact semisimple spectrum generating algebras. However, the examples cited by Dothan[1] involve noncompact, semisimple Lie algebras. We will now work out a method for passing back-and-forth between these two types.

First, let us present Cartan's method for defining different "real forms" of the same complex Lie algebra. Let **G** be a Lie algebra over the real numbers as a field of scalars, and let $\alpha\colon \mathbf{G} \to \mathbf{G}$ be an automorphism of **G**, such that

$$\alpha^2 = (\text{identity}). \tag{7.1}$$

Let us now define a "new" bracket structure on the underlying vector space of **G**, to be denoted by $[\ ,\ ]'$:

$$[X, Y]' = \tfrac{1}{2}\{[X, \alpha(Y)] + [\alpha(X), Y] + [X, Y] - \alpha([X, Y])\}, \tag{7.2}$$

Now, we have the following result, whose proof is easy, but which we will not give here:

Theorem 7.1: Formula (7.2) defines $[\ ,\ ]'$ as a real Lie algebra structure on the underlying vector space of **G**. If we denote this Lie algebra by **G'**, then the complexifications of **G** and **G'** are isomorphic as *complex Lie algebras*, i.e., **G** and **G'** are "real forms" of the same complex Lie algebra.

The following result is obvious from formula (7.2).

Theorem 7.2: If **L** is a subspace of **G** such that $[\mathbf{L}, \mathbf{L}] \subset \mathbf{L}$, i.e., **L** is a Lie subalgebra of **G** with respect to the Lie algebra $[\ ,\ ]$, and if $\alpha(\mathbf{L}) = \mathbf{L}$, then **L** is also a Lie subalgebra with respect to the Lie algebra structure $[\ ,\ ]'$ on **G** defined by formula (7.2).

We can now apply these remarks to the problem of constructing examples of noncompact spectrum generating algebras from the compact ones. Let **G** be a compact semisimple Lie algebra; and suppose that

K, L are symmetric subalgebras of **G**, such that (G, K, L) satisfy the conditions (5.3) needed to say that **G** is a "spectrum generating algebra". Let α be an automorphism of **G** which satisfies (7.1), such that

$$\alpha(\mathbf{L}) = \mathbf{L}, \qquad \alpha(\mathbf{K}) = \mathbf{K}.$$

Form the noncompact Lie algebra **G'** which is a real form of **G** via formula (7.2). Then, by Theorem (7.2), **L** and **K** form Lie subalgebras of **G'**, that we call **L'**, **K'**. Then, one sees readily that **L'**, **K'** are symmetric subalgebras of **G'**. It seems likely that (**G'**, **L'**, **K'**) satisfies the conditions needed to specify G' as a "spectrum generating algebra." At least, the two examples of most interest for physics,

$$(G = SU(4), G' = SU(1,3); G = SO(5,R), G' = SO(4,1)),$$

do indeed define **G'** as a spectrum generating algebra. (They correspond physically to the hydrogen atom and harmonic oscillator). However, we will not go into the Lie group theoretic details needed to settle this conjecture, but will leave it as a starting point for a possible later paper.

* Supported by the Office of Naval Research under Contract NR 041-378. Reproduction in whole or in part is permitted for any purpose of the United States Government.

1 Y. Dothan, Phys. Rev. D **2**, 2944 (1970).
2 R. Hermann, *Differential Geometry and the Calculus of Variations* (Academic New York, 1968).
3 B. Kostant, *Lectures in Modern Analysis*, III, Vol. 170 in *Lectures in Mathematics* (Springer-Verlag, Heidelberg, 1970).
4 R. Hermann, *Lie Groups for Physicists* (Benjamin, New York, 1966).
5 E. Artin, *Geometric Algebra* (Interscience, New York, 1957).
6 S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic, New York, 1962).
7 L. Conlon, Bull. Am. Math. Soc. **75**, 610 (1969).
8 R. Hermann, Proc. Am. Math. Soc. **11**, 544 (1960).
9 R. Hermann, Proc. Ned. Akad. Wet. **65**, 291 (1962).

Spectrum Generating Algebras and Symmetries in Mechanics. II

Robert Hermann*
Department of Mathematics, Rutgers College, New Brunswick, New Jersey 08903
(Received 14 January 1971)

In case a group of symmetries of a classical mechanics system acts on a locally transitive way on the energy surface, it is shown how the time evolution of the system is related to orbits of one-parameter subgroups of the symmetry group. In particular, this provides a group-theoretic interpretation of certain types of regularization of collisions. Generalizations of this phenomenon to quantum mechanics are also discussed.

1. INTRODUCTION

In Paper I[1] we have discussed certain situations which are "maximally symmetric" in the sense that the group of symmetries act transitively on the underlying space. In this paper we will pursue further the study of such transitivity situations, concentrating attention on the case where there is a "local group" of canonical transformations which commutes with the Hamiltonian and which acts in a "local transitive" way on the energy surfaces. We will show how these symmetries enable one to discuss the "collision" properties of the classical equations of motion, generalizing material that is already known[2] for the hydrogen atom–Kepler motion situation. We shall also discuss the question of extending a locally transitive group of Hamiltonian symmetries to a group of canonical transformations that act in a locally transitive way on classical state space. Here we encounter, in a slightly different form, the phenomena discussed in Ref. 3 of a one-parameter group of symmetries of a system of differential equations, for which certain orbits are also solutions of the differential equations themselves. (The most classical version of this phenomenon is the Lagrange rotating equilateral triangle solution of the three-body problem.) We will also briefly discuss in Sec. 4 this phenomenon in a quantum-mechanical context.

2. LOCALLY TRANSITIVE HAMILTONIAN SYMMETRIES

Again, we will refer to Ref. 3 for the differential-geometric concepts and notations to be used here. Let M be an even dimensional manifold, and let ω be a maximal-rank, closed two-differential form on M. (For example, M could be[4,5] the cotangent bundle of a "configuration space" manifold N. In this case, there is a natural choice of two-form on M, whose associated "Hamiltonian–Jacobi theory" gives the usual material of classical mechanics.)

A diffeomorphism $\phi: M \to M$, such that

$$\phi^*(\omega) = \omega \tag{2.1}$$

is called a *canonical transformation*. A *local canonical transformation* is an ordered triple (U, U', ϕ) consisting of two open subsets U, U' of M, together with a diffeomorphism $\phi: U \to U'$ which satisfies (2.1). A vector field $X \in V(m)$ is an *infinitesimal canonical transformation* if

$$X(\omega) = 0. \tag{2.2}$$

The set of all such vector fields is denoted by $V(\omega)$ and forms a Lie algebra with the Lie algebra bracket $[,]$ just the Jacobi bracket[3] of vector fields. Given an $X \in V(\omega)$, the integral curves[3] of X define a family of local canonical transformations on M. Let **G** be a Lie subalgebra of $V(\omega)$, and let M' be a subset of M. **G** is said to act in a *locally transitive* way on M' if the following condition is satisfied:

> Given points $p_1', p_2' \in M'$ which are sufficiently close together, there is a local canonical transformation generated by integral curves of an $X \in \mathbf{G}$ which maps p_1' into p_2'. (2.3)

Suppose, for example, that M' is a submanifold of M, and that **G** is tangent to M', i.e.,

K, L are symmetric subalgebras of **G**, such that (G, K, L) satisfy the conditions (5.3) needed to say that **G** is a "spectrum generating algebra". Let α be an automorphism of **G** which satisfies (7.1), such that

$$\alpha(\mathbf{L}) = \mathbf{L}, \qquad \alpha(\mathbf{K}) = \mathbf{K}.$$

Form the noncompact Lie algebra **G'** which is a real form of **G** via formula (7.2). Then, by Theorem (7.2), **L** and **K** form Lie subalgebras of **G'**, that we call **L'**, **K'**. Then, one sees readily that **L'**, **K'** are symmetric subalgebras of **G'**. It seems likely that (**G'**, **L'**, **K'**) satisfies the conditions needed to specify G' as a "spectrum generating algebra." At least, the two examples of most interest for physics,

$$(G = SU(4), G' = SU(1,3); G = SO(5,R), G' = SO(4,1)),$$

do indeed define **G'** as a spectrum generating algebra. (They correspond physically to the hydrogen atom and harmonic oscillator). However, we will not go into the Lie group theoretic details needed to settle this conjecture, but will leave it as a starting point for a possible later paper.

* Supported by the Office of Naval Research under Contract NR 041-378. Reproduction in whole or in part is permitted for any purpose of the United States Government.

1 Y. Dothan, Phys. Rev. D **2**, 2944 (1970).
2 R. Hermann, *Differential Geometry and the Calculus of Variations* (Academic New York, 1968).
3 B. Kostant, *Lectures in Modern Analysis*, III, Vol. 170 in *Lectures in Mathematics* (Springer-Verlag, Heidelberg, 1970).
4 R. Hermann, *Lie Groups for Physicists* (Benjamin, New York, 1966).
5 E. Artin, *Geometric Algebra* (Interscience, New York, 1957).
6 S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic, New York, 1962).
7 L. Conlon, Bull. Am. Math. Soc. **75**, 610 (1969).
8 R. Hermann, Proc. Am. Math. Soc. **11**, 544 (1960).
9 R. Hermann, Proc. Ned. Akad. Wet. **65**, 291 (1962).

Spectrum Generating Algebras and Symmetries in Mechanics. II

Robert Hermann*
Department of Mathematics, Rutgers College, New Brunswick, New Jersey 08903
(Received 14 January 1971)

In case a group of symmetries of a classical mechanics system acts on a locally transitive way on the energy surface, it is shown how the time evolution of the system is related to orbits of one-parameter subgroups of the symmetry group. In particular, this provides a group-theoretic interpretation of certain types of regularization of collisions. Generalizations of this phenomenon to quantum mechanics are also discussed.

1. INTRODUCTION

In Paper I[1] we have discussed certain situations which are "maximally symmetric" in the sense that the group of symmetries act transitively on the underlying space. In this paper we will pursue further the study of such transitivity situations, concentrating attention on the case where there is a "local group" of canonical transformations which commutes with the Hamiltonian and which acts in a "local transitive" way on the energy surfaces. We will show how these symmetries enable one to discuss the "collision" properties of the classical equations of motion, generalizing material that is already known[2] for the hydrogen atom–Kepler motion situation. We shall also discuss the question of extending a locally transitive group of Hamiltonian symmetries to a group of canonical transformations that act in a locally transitive way on classical state space. Here we encounter, in a slightly different form, the phenomena discussed in Ref. 3 of a one-parameter group of symmetries of a system of differential equations, for which certain orbits are also solutions of the differential equations themselves. (The most classical version of this phenomenon is the Lagrange rotating equilateral triangle solution of the three-body problem.) We will also briefly discuss in Sec. 4 this phenomenon in a quantum-mechanical context.

2. LOCALLY TRANSITIVE HAMILTONIAN SYMMETRIES

Again, we will refer to Ref. 3 for the differential-geometric concepts and notations to be used here. Let M be an even dimensional manifold, and let ω be a maximal-rank, closed two-differential form on M. (For example, M could be[4,5] the cotangent bundle of a "configuration space" manifold N. In this case, there is a natural choice of two-form on M, whose associated "Hamiltonian–Jacobi theory" gives the usual material of classical mechanics.)

A diffeomorphism $\phi: M \to M$, such that

$$\phi^*(\omega) = \omega \tag{2.1}$$

is called a *canonical transformation*. A *local canonical transformation* is an ordered triple (U, U', ϕ) consisting of two open subsets U, U' of M, together with a diffeomorphism $\phi: U \to U'$ which satisfies (2.1). A vector field $X \in V(m)$ is an *infinitesimal canonical transformation* if

$$X(\omega) = 0. \tag{2.2}$$

The set of all such vector fields is denoted by $V(\omega)$ and forms a Lie algebra with the Lie algebra bracket $[,]$ just the Jacobi bracket[3] of vector fields. Given an $X \in V(\omega)$, the integral curves[3] of X define a family of local canonical transformations on M. Let **G** be a Lie subalgebra of $V(\omega)$, and let M' be a subset of M. **G** is said to act in a *locally transitive* way on M' if the following condition is satisfied:

> Given points $p_1', p_2' \in M'$ which are sufficiently close together, there is a local canonical transformation generated by integral curves of an $X \in \mathbf{G}$ which maps p_1' into p_2'. (2.3)

Suppose, for example, that M' is a submanifold of M, and that **G** is tangent to M', i.e.,

For each $p' \in M$, the set $\{X(p'): X \in \mathbf{G}\} = \mathbf{G}(p')$ generated by $\mathbf{G}$ is contained in $M'_{p'}$, the tangent space to M' at p'. (2.4)

Then, one sees readily that, if the following condition is satisfied, then $\mathbf{G}$ acts on a locally transitive manner on M':

$$\mathbf{G}(p') = M'_{p'} \quad \text{for all } p' \in M'. \tag{2.5}$$

Example: Suppose that M is the state space of a classical mechanical system, with ω the canonical 2-form. Use ω to define a Poisson bracket structure on $F(M)$. Given $f \in F(M)$, let $X_f \in V(M)$ be the vector field, such that

$$\begin{aligned} &df = X_f \lrcorner \omega, \\ &X_f(\omega) = 0, \quad \text{i.e., } X_f \text{ is an infinitesimal canonical transformation,} \\ &\{f_1, f_2\} = -X_{f_1}(f_2) \quad \text{for } f_1, f_2 \in F(M). \end{aligned} \tag{2.6}$$

Let h be a fixed element of $F(M)$, considered as the *Hamiltonian* of the system. Let $f_1, \ldots, f_n$ be functions on M, such that $\{h, f_a\} = 0$ for $1 \le a \le n$. These functions are called *constants of motion.* The vector field X_{f_a} then satisfies

$$0 = [X_h, X_{f_a}] = X_{f_a}(h),$$

which indicates that the local canonical transformations generated by the X_{f_a} are symmetries of the equations of motion defined by h. Using (2.6), we now have the following result:

Theorem 2.1: Suppose that $\dim M = 2m$. If the values at each point of M of $df_1, \ldots, df_n$ spans a space of covectors of dimension $(2m - 1)$, then the Lie algebra of canonical transformations generated by the $X_{f_1}, \ldots, X_{f_n}$ acts on a locally transitive manner on the energy surfaces $h^{-1}(E)$, $E \in R$ of M.

For example, if h is the Hamiltonian of the hydrogen atom or the harmonic oscillator, then it is well known how to choose such constants of motion, which form (under Poisson bracket) finite-dimensional Lie algebras [$SO(4, R)$ for the hydrogen atom, $U(3)$ for the oscillator].

Suppose now that, in general, $\mathbf{K}$ is a Lie algebra of infinitesimal canonical transformations on M, each element of which generates a symmetry of h. For $E \in R$, set

$$M_E = h^{-1}(E).$$

Let us assume that the "energy surfaces" M_E are submanifolds of M. Then, each $X \in \mathbf{K}$ is tangent to M_E. Let us suppose also that $\mathbf{K}$ acts in a locally transitive manner on M_E. Let X_h be the vector field on M such that $dh = X_h \lrcorner \omega$. Then, the integral curves of X_h are the solutions of the Hamilton equations, with Hamiltonian h. Since $X_h(h) = \{h, f\} = 0$, each integral curves also lies completely on one energy surface M_E.

Theorem 2.2: The integral curves of X_h that lie on an energy surface M_E are characteristic curves of the 2-form ω restricted to M_E. (See Ref. 3 for the definition of a characteristic curve of a 2-form.)

Proof: Since $dh = X_h \lrcorner \omega$, we see that $X_h \lrcorner \omega = 0$ when restricted to M_E (since $dh = 0$ on M_E), which is the condition needed[3] to assert that the integral curves of X_h are characteristic curves.

Theorem 2.3: If $t \to \sigma(t)$ is an integral curve of X_h, which lies on M_E and if $\mathbf{K}$ acts locally transitively on M_E, then there is an $X \in \mathbf{K}$ such that σ is an integral curve of X. Further, if $f \in F(M)$ is such that $df = X \lrcorner \omega$, and if f_E denotes the function f restricted to M_E, then each point of σ is a critical point of f_E.

Proof: To say that $\mathbf{K}$ consists of symmetries of h is to say that

$$X(h) = 0 \quad \text{for all } X \in \mathbf{K}. \tag{2.7}$$

Hence, $[X, X_h] = 0$ for all $X \in \mathbf{K}$.

Then, as was shown in Ref. 3, the one-parameter group generated by X maps an integral curve of X_h into another integral curve.

Let p be the point of M_E such that $\sigma(0) = p$. Then, our assumption that $\mathbf{K}$ acts in a locally transitive way on M_E implies that there is an $X \in \mathbf{K}$ whose value at p is precisely the tangent vector to σ at p. Let us suppose that this tangent vector is nonzero. (The result is trivial in the case it is zero.) Then, as shown in Ref. 3, there is a coordinate system $x_1, \ldots, x_m$ for M, valid in a neighborhood of p, such that

$$0 = x_1(p) = \cdots = x_m(p), \qquad X_h = \frac{\partial}{\partial x_1}. \tag{2.8}$$

Suppose that, in this coordinate system,

$$X = A_1 \frac{\partial}{\partial x_1} + A_2 \frac{\partial}{\partial x_2} + \cdots + A_m \frac{\partial}{\partial x_m}. \tag{2.9}$$

The coefficients $A_1, \ldots, A_m$ in (2.9) are then functions of $x_1, \ldots, x_m$. Condition (2.7) implies that these functions depend only on $x_2, \ldots, x_m$. The condition that $X_h(p) = X(p)$ implies that

$$\begin{aligned} &A_1(0, \ldots, 0) = 1, \\ &A_2(0, \ldots, 0) = 0 = \cdots = A_m(0, \ldots 0). \end{aligned} \tag{2.10}$$

The curve $t \to \sigma(t)$, as the integral curve of X_h starting at p, is given explicitly as follows, in these coordinates:

$$t \to (t, 0, \cdots, 0) = (x_1(t), x_2(t), \cdots, x_m(t)).$$

Then

$$\frac{dx_1}{dt} = 1 = A_2(0, \ldots, 0) = A_2(x_2(t), \ldots, x_m(t)),$$

$$\frac{dx_2}{dt} = 0 = A_2(x_2(t), \ldots, x_m(t)),$$

$$\frac{dx_m}{dt} = 0 = A_m(x_2(t), \ldots, x_m(t)).$$

The relations then show, explicitly, that the curve $t \to \sigma(t)$ is also an integral curve of X, which proves the first part.

Recall that a point $p \in M_E$ is a *critical point* of a function f_e if $df_E(p) = 0$. Suppose that $df = X/\omega$ and that $t \to \sigma(t)$ is a curve in M_E, which is an integral

curve of *both* X and X_h. Then, by Theorem 2.2, the curve $t \to \sigma(t)$ is a characteristic curve of ω restricted to M_E.

Hence,

$$\begin{aligned} df_E(\sigma(t)) &= X(\sigma(t))\lrcorner\omega \quad \text{restricted to } M_E \\ &= Xh(\sigma(t))\lrcorner\omega \quad \text{restricted to } M_E \\ &= 0, \end{aligned}$$

which proves that each point of the curve $t \to \sigma(t)$ is, indeed, a critical point of the function f_E.

Remark: The fact that f, restricted to the submanifold M_E, has a critical point can be phrased in a more symmetric way, using Lagrange multipliers. In fact, one sees that f_E has a critical point at the point $p \in M_E$ if and only if there is a real number λ such that

$$df + \lambda dh = 0 \quad \text{at } p. \tag{2.11}$$

(λ is, of course, the Lagrange multiplier.)

Another way of looking at this has been pointed out by Smale.[6] Consider the map $M \to R^2$ which assigns to the point p the pair $(f(p), h(p))$ of real numbers. Then, the points $p \in N$ which satisfy (2.11) are obviously the points at which this map is not of maximal rank, i.e., the singular points of the map.

Finally, the general phenomenon encountered here—solutions of differential equations which are also orbits of groups of symmetries of the equations—has been studied in various geometric contexts by the author.[3,7,8]

3. REGULARIZATION OF SINGULARITIES OF DIFFERENTIAL EQUATIONS BY MEANS OF GROUP THEORY

Let M continue to be an even-dimensional manifold with a closed, maximal-rank 2-form ω. Let h be a Hamiltonian function on M, and let X_h be the Hamiltonian vector field, such that

$$dh = X_h\lrcorner\omega.$$

An integral curve $t \to \sigma(t)$ of X_h, defined over the interval $0 \le t < a$, is said to have a *singularity* at $t = a$ if it cannot be continued beyond the point $t = a$. For example, if M is the state space of a system of particles, a "collision" at $t = a$ would be such a singularity, since some of the momenta of the particles would go to infinity as $t \to a$.

Now, we must explain what is meant by "regularization" of the singularity. Among the many possible meanings of this concept, we choose a "geometric" one: The energy surface M_E on which the curve σ lies is to be embedded as a submanifold of a manifold M', such that X_h arises by restriction to M_E of a vector field on M', and such that $\lim\sigma(t)$ exists in M' as $t \to a$.

For example, Moser[2] has constructed such an embedding for the case where h is the Hamiltonian of the hydrogen atom. In this section, we will discuss certain features of this problem that can be looked at in a group-theoretical way.

Let us suppose that $\mathbf{K}$ is a finite dimensional Lie algebra of infinitesimal transformations on M, consisting of symmetries of h, i.e., such that $X(h) = 0$ for all $X \in \mathbf{K}$.

Then, $\mathbf{K}$ acts on the energy surface $M_E = h^{-1}(E)$. Let us further suppose that it acts in a locally transitive way on M_E. Then, as we have seen, if $t \to \sigma(t)$, $0 \le t < a$, is an integral curve of X_h which lies on M_E, there is an $X \in \mathbf{K}$ such that σ is also an integral curve of X. We shall now describe a situation where this fact may be utilized to extend σ beyond $t = a$.

Definition: Let $\mathbf{K}$ be a finite-dimensional Lie algebra of vector fields which acts in a locally transitive way on a manifold N. Let K be a connected Lie group whose Lie algebra is isomorphic to $\mathbf{K}$. Then, the integration of the vector fields in $\mathbf{K}$ defines a "local" Lie group action of K on N. This action is said to be *globalizable* if there is a closed subgroup L of K and an embedding of N as an open subset of $N = K/L$, such that the action of $\mathbf{K}$ on N results from the restriction to N of the vector fields in $\mathbf{K}$ acting on the coset space K/L.

Remark: If the action of $\mathbf{K}$ is globalizable in this way, then the geometric situation implies that an integral curve $t \to \sigma(t)$, $0 \le t < a$, of an $X \in \mathbf{K}$, which lies in N, can always be continued beyond a (in fact, to $-\infty < t < \infty$), into a curve lying in $N' = K/L$. In fact, this curve is just the orbit $t \to g(t)\sigma(0)$ of the one-parameter subgroup $g(t) = \exp(tX)$ of K generated by the element $X \in \mathbf{K}$.

Let us return now to the case where $N = M_E$, and $\mathbf{K}$ results from the action of a Lie algebra of symmetries of h. Then, if N' is such a globalization of N, it follows that the integral curves of X_h can be continued into N'. For example, for the case of the hydrogen atom, Moser has constructed N' explicitly as $SO(4,R)/SO(2,R)$.

Now, unfortunately, there are as yet no general theorems in the mathematical literature asserting when a locally transitive action of a Lie algebra on a manifold N may be globalized in this way. (This would be analogous to the question—equally unknown—of the possible "completion" of an incomplete Riemannian metric.) However, at least these geometric remarks do provide one with a geometric framework in which to think of the problem, and provide us with a "geometric" set of sufficient conditions for the "regularization" of singularities.

4. THE QUANTUM MECHANICAL ANALOGS

Part of our motivation for this work in classical mechanics was to develop ideas that had an analog in quantum mechanics. We will now develop this analogy explicitly, focusing attention on the phenomenon investigated above in a classical setting of orbits of one-parameter groups of symmetries that are also solutions of the Schrödinger equation.

Let H be a complex vector space, with a real-bilinear form $(\psi_1, \psi_2) \to \langle\psi_1 | \psi_2\rangle$ that is Hermitian symmetric, in the sense that it satisfies the following conditions:

$$\langle\psi_1 | \psi_2\rangle = \langle\psi_2 | \psi_1\rangle^* \quad \text{for } \psi_1, \psi_2 \in H,$$
$$\langle c\psi_1 | \psi_2\rangle = c^*\langle\psi_1 | \psi_2\rangle \quad \text{for } c \in C,\ \psi_1, \psi_2 \in H,$$
$$\langle\psi | \psi\rangle > 0 \text{ if } \psi \in H \text{ is} \neq 0.$$

Such a structure will define H as a *Hilbert space*. (This is the terminology in physics; in mathematics one adds completeness.)

Now let M be the projective space constructed from H, i.e., the space of one-dimensional linear subspaces of H. M should be considered as the quantum mechanical *state space*. We will now discuss some of its differential-geometric properties. (It is one of a type of "infinite-dimensional manifolds," whose differential-geometric properties will be treated according to general lines discussed in Ref. 4.)

Let π be the map: $H - (0) \to M$ which assigns to each nonzero $\psi \in H$ the one-dimensional complex-linear subspace to which it belongs. For each $\psi \in H - (0)$, denote by $\psi^\perp$ the orthogonal complement subspace to ψ, i.e., the set of $\psi_1 \in H$ such that $\langle \psi | \psi_1 \rangle = 0$. One can then regard the assignment $\psi \to \psi^\perp$ as defining a vector bundle over $H - (0)$. Now, the fibers of this vector bundle are clearly constant on the fibers of π (since $(c\psi)^\perp = \psi^\perp$ for $c \in C$), hence there is defined a vector bundle, which we denote by $T(M)$ on M.

As the notation "$T(M)$" should indicate, $T(M)$ can be identified with the "tangent bundle" to the "infinite-dimensional manifold M." [For $\pi(\psi) \in M$, denote by $M_{\pi(\psi)} = \psi$ the fiber of $T(M)$ over the point $\pi(\psi)$.] To see this, suppose that $t \to \psi(t)$ is a curve in $H - (0)$, with $\psi(0) = \psi$. Assign to this projected curve $t \to \pi\psi(t)$ in M the "tangent vector," which is the orthogonal projection on ψ of the derivative $(dx/dt)(0)$. This element of $\psi^\perp = M_{\pi(\psi)}$ is then independent of the lifting $t \to \psi(t)$ chosen for the curve $t \to \pi(\psi(t))$. Clearly, two such curves have "first order contact" at $t = 0$ if their "tangent vectors" in $M_{\pi(\psi)} = \psi^\perp$, assigned in this way, coincide. We will denote this orthogonal projection in $\psi^\perp$ by the notation

$$\pi_* \left(\frac{d\psi}{dt} (0) \right) .$$

Let $A: H \to H$ be a Hermitian operator on H, i.e., $\langle \psi_1 | A\psi_2 \rangle = \langle A\psi_1 | \psi_2 \rangle$ for $\psi_1, \psi_2 \in H$. Let us suppose that A generates a one-parameter unitary group acting on H, denoted by $U(t) = \exp[-(it/\hbar)A]$. In other words, for $\psi \in H$

$$\frac{d}{dt} U(t)\psi = -\frac{iA}{\hbar} (U(t)\psi). \tag{4.1}$$

$U(t)$ permutes the one-dimensional linear subspaces of H, hence acts as a one-parameter transformation group on M. Let X_A be the vector field on M defined by this one-parameter group. Using (4.1), we see that for $\psi \in S$,

$$X_A(\pi(\psi)) = \pi_*(-(iA/\hbar)\psi). \tag{4.2}$$

In particular, note that $X_A(\pi(\psi)) = 0$, i.e., $\pi(\psi)$ is a "singular point" for the vector field X_A, if and only if ψ is an eigenvector of A.

Each such Hermitian operator A also determines a real-valued function $f_A: M \to R$ as follows:

$$f_A(\pi(\psi)) = \langle A\psi | A\psi \rangle / \langle \psi | \psi \rangle \quad \text{for } \psi \in H - (0). \tag{4.3}$$

[Note that the right-hand side of (4.3) is invariant under the transformation $\psi \to c\psi$ for $c \in C$, hence (4.3) really does define f_A as a genuine function on M.] Let us compute the "differential" df_A of f_A: Suppose $t \to \psi(t)$ is a curve in H, with $\psi(0) = \psi$. Then,

$$\begin{aligned} &\frac{d}{dt} f_A(\pi(\psi(t)))|_{t=0} \\ &= \frac{1}{\langle \psi | \psi \rangle^2} \left[\left(\left\langle \frac{d\psi}{dt} (0) \middle| A(\psi) \right\rangle + \left\langle \frac{d\psi}{dt} (0) \middle| A\psi \right\rangle^* \right) \right. \\ &\quad \left. \times (\langle \psi | \psi \rangle + \langle \psi | A\psi \rangle) \left(\left\langle \frac{d\psi}{dt} (0) \middle| \psi \right\rangle + \left\langle \frac{d\psi}{dt} (0) \middle| \psi \right\rangle^* \right) \right]. \end{aligned}$$

Now, in this computation, we can suppose that the curve $t \to \psi(t)$ has been normalized so that

$$\langle \psi | \psi \rangle = 1, \quad \frac{d\psi}{dt} (0) \in \psi^\perp. \tag{4.4}$$

With these normalizations, we then have

$$\begin{aligned} \frac{d}{dt} f_A(\pi(\psi(t)))|_{t=0} &= df_A \left[\pi_* \left(\frac{d\psi}{dt} (0) \right) \right] \\ &= \left\langle \frac{d\psi}{dt} (0) \middle| A\psi \right\rangle + \left\langle \frac{d\psi}{dt} \middle| A\psi \right\rangle^* \\ &= \left\langle \frac{d\psi}{dt} (0) \middle| i\hbar X_A(\pi(\psi)) \right\rangle + \left\langle \frac{d\psi}{dt} (0) \middle| i\hbar X_A(\pi|(\psi)) \right\rangle^* \\ &\quad \text{[using (4.2)]} \\ &= i\hbar \left(\left\langle \frac{d\psi}{dt} (0) \middle| X_A(\pi(\psi)) \right\rangle - \left\langle \frac{d\psi}{dt} (0) \middle| X_A(\pi(\psi)) \right\rangle^* \right). \end{aligned} \tag{4.5}$$

To interpret this formula in an interesting differential-geometric way, let us define ω as a real, skew-symmetric bilinear form: $M_{\pi(\psi)} \times M_{\pi(\psi)} \to R$, as follows:

$$\omega(\psi_1, \psi_2) = i\hbar(\langle \psi_1 | \psi_2 \rangle^* - \langle \psi_1 | \psi_2 \rangle) \quad \text{for } \psi_1, \psi_2 \in M_{\pi(\psi)} = \psi^\perp. \tag{4.6}$$

Then, (4.5) can be rewritten in the following form:

$$df_A \left[\pi_* \left(\frac{d\psi}{dt} (0) \right) \right] = \omega \left(X_A(\psi), \frac{d\psi}{dt} (0) \right). \tag{4.7}$$

Now, as ψ varies, ω defines a "2-differential form" on M. One can, in fact, prove that it is a "closed 2-form" on the infinite-dimensional manifold M, when the "exterior derivative" is defined as explained in Ref. 4. As ψ varies, $\psi \to X_A(\psi)$ defines a "vector field" over M. The function f_A, the vector field X_A, and the 2-form ω are then related—using Eqs. (4.7), via the relation

$$df_A = X_A \lrcorner \omega. \tag{4.8}$$

This relation is then a typical "symplectic manifold" situation.[3-5] What we have done is to show that the reciprocity between "observables," i.e., "Hermitian operators" and "one-parameter unitary groups," i.e., "quantum symmetries," which is a typical feature of quantum mechanics, can be put into a "classical mechanics—symplectic manifold" framework by analyzing the differential geometric structure of M, the quantum mechanical state space.

After these general remarks, let us turn to the quantum-mechanical analog of Theorem 2.4 and relation (2.8). Let A and A_k be two Hermitian operators on H,

which commute and which generate one-parameter unitary groups, $t \to U(t)$, $t \to U_h(t)$, on H. (A_h may be thought of as the "Hamiltonian" of the system.) Let $\pi(\psi)$ be a point of M which is a critical point of $f_A + \lambda f_{A_h}$ where λ is a real constant, i.e., $d(f_A + \lambda f_{A_h}) = 0$ at ψ. Relation (4.7) now implies, as it does in classical mechanics, that $X_A + \lambda X_{A_h}$ is zero at $\pi(\psi)$. Formula (11.2) now implies that ψ is an eigenvector of $A + \lambda A_h$. In turn, this implies, using the relation $[A, A_h] = 0$, that the orbits $t \to U(t)\pi(\psi)$, $U_h(t)\pi(\psi)$ coincide, just as in classical mechanics.

Remark: This "symplectic manifold" approach to quantum mechanics has been emphasized by Kostant.[9] His ideas were presented at the 1969 Battelle Rencontres on Mathematical Physics in Seattle, and the presentation here is influenced by his presentation there. I have also developed an alternate and related approach in Ref. 10.

* This work was supported in part by the office of Naval Research under contract No. NR-041-378. Reproduction in whole or part is permitted for any purpose by the United States Government.

1 R. Hermann, J. Math. Phys. **13**, 833 (1972).
2 J. Moser, Commun. Pure Appl. Math. **23**, 609 (1970).
3 R. Hermann, *Differential Geometry and the Calculus Variations* (Academic, New York, 1968).
4 R. Hermann, *Vector Bundles in Mathematical Physics* (Benjamin, New York, 1970), Vol. II.
5 R. Abraham and J. Marsden, *Foundations of Mechanics* (Benjamin, New York, 1967).
6 S. Smale, Inventions Math. **10**, 305 (1970).
7 R. Hermann, Ill. J. Math. **6**, 634 (1962).
8 R. Hermann, Bull. Am. Math. Soc. **66**, 91 (1960).
9 B. Kostant, "Quantization and unitary representations," in *Lectures in Modern Analysis*, III, Vol. 170 in *Lectures in Mathematics* (Springer-Verlag, Heidelberg, 1970).
10 R. Hermann, J. Math. Phys. **6**, 1768 (1965).

Expansion of the Master Equation for One-Dimensional Random Walks with Boundary*

N. G. van Kampen† and Irwin Oppenheim
Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received 3 September 1971)

In order to understand the behavior of coarse-grained equations in the presence of a boundary, the following model is investigated. A homogeneous one-dimensional random walk is bounded on one side by some boundary conditions of rather arbitrary form. The corresponding master equation is approximated by the Fokker–Planck equation plus partial differential equations for the higher orders. The boundary condition for the Fokker–Planck approximation is well known; but those for the higher order terms are here derived. To the second order they amount to a virtual displacement of the boundary. The case of a two-step random walk, however, gives rise to an unexpected complication, inasmuch as nonpropagating solutions of the master equation cannot be ignored in the boundary condition, although they do not contribute to the differential equations themselves.

1. INTRODUCTION

In recent years much effort in nonequilibrium statistical mechanics has been devoted to the derivation of equations describing the time dependence of reduced distribution functions and macroscopic variables. These equations have been derived from the exact equations of motion and are particularly simple when some pertinent physical parameter is quite small. Examples are the Boltzmann equation for a low density gas, the Langevin and Fokker–Planck equations for heavy particles in a bath of light particles, and the hydrodynamic equations in systems with small spatial gradients. All of these equations can be extended to situations where the parameter (e.g., the density) is still small, but large enough so that the lowest order equations are no longer valid. These extensions have resulted in the Choh–Uhlenbeck equation[1] for dilute gases and the generalized hydrodynamic equations.[2]

All of the derivations mentioned above were carried out in the thermodynamic limit and in the absence of boundaries. It would be of great interest to derive the transport equations for a system with boundaries, but little progress has been made in this direction.[3(a)] Phenomenological considerations do supply a clue to the expected results of such analyses, at least to the lowest order. Thus, for example, the hydrodynamic equations describing the bulk properties for nonequilibrium fluids have the same form whether a boundary is present or not. The presence of the boundary merely imposes boundary conditions on these equations. We emphasize that these boundary conditions are often not derived from molecular theory but imposed on intuitive phenomenological grounds. While the intuitive approach may work for the lowest order equations, there is no assurance that it will, and even less assurance that it will be valid for the extended equations.

In this paper we study a number of discrete random walk models in the presence of boundaries. These models are of interest because they are fairly easy to analyze and because they have been widely used to describe stochastic processes in physics and chemistry. We expand the exact differential-difference equation (master equation) and find to lowest order the Fokker–Planck equation, and similar partial differential equations for the successive higher orders. The boundary conditions for these differential equations are also derived from the exact master equation. The form of the differential equations is not affected by the presence of boundaries. However, the boundary conditions which their solutions must satisfy are not intuitive and can be determined only from a detailed study of the solutions to the master equation itself.

2. THE RANDOM WALK MODEL

By "random walk" we mean a stationary, continuous time Markov process whose range of possible values consists of integral numbers n. The random walk is *unbounded* when n ranges from $-\infty$ to $+\infty$, and is *bounded* (on the left) when n takes on only the values

which commute and which generate one-parameter unitary groups, $t \to U(t)$, $t \to U_h(t)$, on H. (A_h may be thought of as the "Hamiltonian" of the system.) Let $\pi(\psi)$ be a point of M which is a critical point of $f_A + \lambda f_{A_h}$ where λ is a real constant, i.e., $d(f_A + \lambda f_{A_h}) = 0$ at ψ. Relation (4.7) now implies, as it does in classical mechanics, that $X_A + \lambda X_{A_h}$ is zero at $\pi(\psi)$. Formula (11.2) now implies that ψ is an eigenvector of $A + \lambda A_h$. In turn, this implies, using the relation $[A, A_h] = 0$, that the orbits $t \to U(t)\pi(\psi)$, $U_h(t)\pi(\psi)$ coincide, just as in classical mechanics.

Remark: This "symplectic manifold" approach to quantum mechanics has been emphasized by Kostant.[9] His ideas were presented at the 1969 Battelle Rencontres on Mathematical Physics in Seattle, and the presentation here is influenced by his presentation there. I have also developed an alternate and related approach in Ref. 10.

* This work was supported in part by the office of Naval Research under contract No. NR-041-378. Reproduction in whole or part is permitted for any purpose by the United States Government.

1 R. Hermann, J. Math. Phys. **13**, 833 (1972).
2 J. Moser, Commun. Pure Appl. Math. **23**, 609 (1970).
3 R. Hermann, *Differential Geometry and the Calculus Variations* (Academic, New York, 1968).
4 R. Hermann, *Vector Bundles in Mathematical Physics* (Benjamin, New York, 1970), Vol. II.
5 R. Abraham and J. Marsden, *Foundations of Mechanics* (Benjamin, New York, 1967).
6 S. Smale, Inventions Math. **10**, 305 (1970).
7 R. Hermann, Ill. J. Math. **6**, 634 (1962).
8 R. Hermann, Bull. Am. Math. Soc. **66**, 91 (1960).
9 B. Kostant, "Quantization and unitary representations," in *Lectures in Modern Analysis*, III, Vol. 170 in *Lectures in Mathematics* (Springer-Verlag, Heidelberg, 1970).
10 R. Hermann, J. Math. Phys. **6**, 1768 (1965).

Expansion of the Master Equation for One-Dimensional Random Walks with Boundary*

N. G. van Kampen† and Irwin Oppenheim
Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received 3 September 1971)

In order to understand the behavior of coarse-grained equations in the presence of a boundary, the following model is investigated. A homogeneous one-dimensional random walk is bounded on one side by some boundary conditions of rather arbitrary form. The corresponding master equation is approximated by the Fokker–Planck equation plus partial differential equations for the higher orders. The boundary condition for the Fokker–Planck approximation is well known; but those for the higher order terms are here derived. To the second order they amount to a virtual displacement of the boundary. The case of a two-step random walk, however, gives rise to an unexpected complication, inasmuch as nonpropagating solutions of the master equation cannot be ignored in the boundary condition, although they do not contribute to the differential equations themselves.

1. INTRODUCTION

In recent years much effort in nonequilibrium statistical mechanics has been devoted to the derivation of equations describing the time dependence of reduced distribution functions and macroscopic variables. These equations have been derived from the exact equations of motion and are particularly simple when some pertinent physical parameter is quite small. Examples are the Boltzmann equation for a low density gas, the Langevin and Fokker–Planck equations for heavy particles in a bath of light particles, and the hydrodynamic equations in systems with small spatial gradients. All of these equations can be extended to situations where the parameter (e.g., the density) is still small, but large enough so that the lowest order equations are no longer valid. These extensions have resulted in the Choh–Uhlenbeck equation[1] for dilute gases and the generalized hydrodynamic equations.[2]

All of the derivations mentioned above were carried out in the thermodynamic limit and in the absence of boundaries. It would be of great interest to derive the transport equations for a system with boundaries, but little progress has been made in this direction.[3(a)] Phenomenological considerations do supply a clue to the expected results of such analyses, at least to the lowest order. Thus, for example, the hydrodynamic equations describing the bulk properties for nonequilibrium fluids have the same form whether a boundary is present or not. The presence of the boundary merely imposes boundary conditions on these equations. We emphasize that these boundary conditions are often not derived from molecular theory but imposed on intuitive phenomenological grounds. While the intuitive approach may work for the lowest order equations, there is no assurance that it will, and even less assurance that it will be valid for the extended equations.

In this paper we study a number of discrete random walk models in the presence of boundaries. These models are of interest because they are fairly easy to analyze and because they have been widely used to describe stochastic processes in physics and chemistry. We expand the exact differential-difference equation (master equation) and find to lowest order the Fokker–Planck equation, and similar partial differential equations for the successive higher orders. The boundary conditions for these differential equations are also derived from the exact master equation. The form of the differential equations is not affected by the presence of boundaries. However, the boundary conditions which their solutions must satisfy are not intuitive and can be determined only from a detailed study of the solutions to the master equation itself.

2. THE RANDOM WALK MODEL

By "random walk" we mean a stationary, continuous time Markov process whose range of possible values consists of integral numbers n. The random walk is *unbounded* when n ranges from $-\infty$ to $+\infty$, and is *bounded* (on the left) when n takes on only the values

1, 2, 3, $\cdots$. For the unbounded case we assume *homogeneity*, i.e., the probability per unit time a_r, for making a jump of r units is independent of n. If the largest possible single jump is s units, we call it an s-step random walk. Then the master equation for the probability $p_n(t)$ to be at site n at time t is

$$\dot{p}_n(t) = \sum_{r=-s}^{s} a_r p_{n+r}(t). \tag{1}$$

If $a_r = a_{-r}$, the random walk is *symmetric*. We shall here restrict ourselves to the symmetric one-step and two-step cases. We assume $a_1 \neq 0$ to ensure that all sites can be reached. [See Ref. 3(b).]

For the bounded case we take the same Eq. (1) for all positive n excepting a finite number of them near the boundary. For the exceptional p_n special equations are required, whose precise form is descriptive of the physical properties of the boundary. Clearly an infinite variety of boundaries is possible. We shall call a boundary *reflecting* when the boundary equations are such that total probability is conserved, and *absorbing* if total probability decreases. Among all varieties of absorbing boundaries one may single out a special case, to be called purely absorbing, which is the kind of absorbing boundary usually considered. [See Ref. 3(c).]

It is not hard to choose the boundary equations such that the total probability *increases*, at a rate proportional to the probability already present at the neighboring sites; this kind of "stimulated emission," however, will be excluded. Moreover, for nonhomogeneous master equations, i.e., a_r depending on n, it often happens that the process is automatically bounded due to the vanishing of a_{-r} $(r = 1, 2, \ldots, s)$ at $n = 0$; such "natural boundaries" are much easier to treat and are not the subject of this article.

The bounded s-step random walk problem can be solved exactly, in the sense that all $p_n(t)$ can be found for $t > 0$ when their initial values are prescribed. In particular, the reflecting and the purely absorbing case have been amply treated in the literature,[5] the latter often in connection with first passage problems. However, for reasons explained in the Introduction, we are interested in developing an approximation scheme based on the smallness of the individual jumps. The first step of this approximation is the Fokker–Planck or diffusion equation with the well-known boundary conditions, viz., the probability density vanishes on an absorbing boundary and has zero slope on a reflecting boundary. Higher order corrections to the Fokker–Planck equation have previously been obtained for the unbounded case.[6] This article is concerned with the higher orders in the presence of a boundary. It will appear that a nontrivial complication arises for multiple step random walks. Rather than trying to provide an exhaustive treatment of all possibilities, we shall confine ourselves to three special paradigms.

It is convenient to define a *pure boundary*, meaning that the recursion relation (1) is valid for *all* $n \geq s + 1$ (supposing that a_{-s} does not vanish), so that only the equations for $p_1, p_2, \ldots, p_s$ are modified. In the case of a pure boundary the following equivalent but more convenient way of formulating the boundary conditions is possible. Introduce s auxiliary variables $p_0, p_{-1}, p_{-2}, \ldots, p_{-s+1}$ and stipulate that (1) is valid for *all* $n \geq 1$, but that the auxiliary variables are connected with the actual variables $p_1, p_2, \cdots$ by a set of s suitably chosen linear relations. How to choose these linear relations will become clear in the applications. *Pure absorption* will be defined as the special case that these linear relations have the simple form $p_0 = p_{-1} = \cdots = p_{-s+1} = 0$.

3. ONE-STEP RANDOM WALK WITH PURE BOUNDARY—EXACT SOLUTION

With suitable choice of time unit, Eq. (1) reduces, in the symmetric one-step case, to

$$\dot{p}_n = p_{n+1} + p_{n-1} - 2p_n. \tag{2}$$

In the case of a pure boundary this is valid for $n = 2, 3, \cdots$. Inasmuch as only one-step jumps are permitted, the special equation for p_1 must have the form

$$\dot{p}_1 = p_2 - (1 + c)p_1, \tag{3}$$

with a single parameter c. For the total probability W one finds

$$\frac{dW}{dt} = \frac{d}{dt}\sum_{n=1}^{\infty} p_n = -cp_1.$$

Hence the boundary is reflecting for $c = 0$ and absorbing for $c > 0$. The case $c < 0$ describes stimulated emission and will therefore be excluded.

Equation (2) for $n \geq 2$ together with (3) for $n = 1$ may be expressed in an equivalent way by declaring (2) valid for $n = 1$ as well and putting

$$p_0 = (1 - c)p_1. \tag{4}$$

(Note that p_0 is merely an auxiliary quantity and is *not* equal to the probability that the particle has been absorbed.)

According to the definition in Sec. 2 the boundary is called purely absorbing when $c = 1$.

To solve this bounded random walk problem exactly first note that (2) is obeyed by

$$e^{-\lambda t} z^n \tag{5}$$

provided that z and λ are connected by the characteristic equation

$$\lambda = 2 - z - 1/z. \tag{6}$$

For fixed time constant λ there are two roots z_1, z_2 and hence two solutions of the form (5); thus the general solution of (2) with time factor $e^{-\lambda t}$ is

$$p_n^{(\lambda)}(t) = e^{-\lambda t}(C_1 z_1^n + C_2 z_2^n). \tag{7}$$

We require that $p_n^{(\lambda)}$ is bounded for $n \to \infty$; it will be shown that that is sufficient for obtaining a complete set of normal modes. Hence one must have $|z_1| \leq 1$ and $|z_2| \leq 1$ (see, however, Appendix A). As $z_1 z_2 = 1$ according to (6), we may write

$$z_1 = e^{i\vartheta}, \quad z_2 = e^{-i\vartheta}, \quad 0 \leq \vartheta \leq \pi. \tag{8}$$

Note that in order to obey (2) for $n = 2, 3, \cdots$, it is

necessary that (7) holds for $n = 1, 2, \cdots$. Thus *owing to the fact that the boundary is pure, all* $p_n(t)$ *have to be of this form.* ϑ is related to λ by (6), or

$$\lambda = 2 - 2\cos\vartheta = 4\sin^2\tfrac{1}{2}\vartheta. \tag{9}$$

This covers the values $0 \leq \lambda \leq 4$, which determines the eigenvalue spectrum of our bounded problem.

The constants C_1, C_2 are now chosen such that (3) is satisfied:

$$(1 + c - \lambda)(C_1 z_1 + C_2 z_2) = C_1 z_1^2 + C_2 z_2^2. \tag{10a}$$

Equivalently one may use (4) to obtain

$$C_1 + C_2 = (1 - c)(C_1 z_1 + C_2 z_2). \tag{10b}$$

It follows that

$$-C_1/C_2 = [1 - (1 - c)e^{-i\vartheta}]/[1 - (1 - c)e^{i\vartheta}]. \tag{10c}$$

The normal mode solutions of the bounded random walk are now fully determined by (7), (9), and (10).

It is instructive to write the result in a more familiar form. Write $S(\vartheta) = e^{2i\eta(\vartheta)}$ for the right-hand side of (10c); note that $|S(\vartheta)| = 1$ and η is real. With suitable normalization the normal mode solution (7) may then be written

$$p_n^{(\lambda)}(t) = e^{-\lambda t}(2/\pi)^{1/2}\sin[\vartheta n + \eta(\vartheta)]. \tag{11}$$

Thus $\eta(\vartheta)$ is the phase shift due to the boundary and $S(\vartheta)$ is the "S matrix". Note, however, that this S-matrix is always unitary, regardless of whether total probability is conserved or not! The orthogonality and completeness of these normal modes is shown in Appendix A. Consequently the final solution may be written in terms of an evolution operator

$$p_n(t) = \sum_{m=1}^{\infty} U_{nm}(t)p_m(0),$$

$$U_{nm}(t) = \frac{2}{\pi}\int_0^{\pi} \sin[\vartheta n + \eta(\vartheta)] \times \sin[\vartheta m + \eta(\vartheta)]e^{-4t\sin^2(\vartheta/2)}d\vartheta. \tag{12}$$

4. ONE-STEP RANDOM WALK WITH PURE BOUNDARY—APPROXIMATE TREATMENT

Following the program outlined in the Introduction we now investigate the approximations that are applicable when the individual jumps may be treated as small compared to the distances that one is interested in. Accordingly we set

$$\epsilon n = x, \qquad p_n(t) = \epsilon P(x, t), \tag{13}$$

and expand in ϵ. In other words, we are interested in an approximation method for solutions that vary slowly compared to the size of the jumps. Of course the result can be found directly by expanding the exact result (12); but our aim is to find an independent method for solving the problem in this approximation.

The master equation (2) gives for $P(x,t)$,

$$\dot{P}(x, t) = P(x + \epsilon, t) + P(x - \epsilon, t) - 2P(x, t) = \epsilon^2\frac{\partial^2 P}{\partial x^2} + \frac{\epsilon^4}{12}\frac{\partial^4 P}{\partial x^4} + \cdots.$$

Using the new time variable $\tau = \epsilon^2 t$ and expanding P,

$$P = P^{(0)} + \epsilon P^{(1)} + \epsilon^2 P^{(2)} + \cdots,$$

one finds successively

$$\frac{\partial P^{(0)}}{\partial \tau} = \frac{\partial^2 P^{(0)}}{\partial x^2}, \tag{14a}$$

$$\frac{\partial P^{(1)}}{\partial \tau} = \frac{\partial^2 P^{(1)}}{\partial x^2}, \tag{14b}$$

$$\frac{\partial P^{(2)}}{\partial \tau} = \frac{\partial^2 P^{(2)}}{\partial x^2} + \frac{1}{12}\frac{\partial^4 P^{(0)}}{\partial x^4}, \quad \cdots \tag{14c}$$

These equations apply to all $x > 0$.

The boundary condition (4) translates into

$$P(0, \tau) = (1 - c)P(\epsilon, \tau),$$

which, on expanding in ϵ, amounts to

$$P^{(0)} = (1 - c)P^{(0)}, \tag{15a}$$

$$P^{(1)} = (1 - c)P^{(1)} + (1 - c)P^{(0)\prime}, \tag{15b}$$

$$P^{(2)} = (1 - c)P^{(2)} + (1 - c)P^{(1)\prime} + \tfrac{1}{2}(1 - c)P^{(0)\prime\prime}. \tag{15c}$$

(Here and in the future we abbreviate the notation by writing $P^{(0)}$ for $P^{(0)}(0, \tau)$, etc., and $P^{(0)\prime}$ for $\partial P^{(0)}(x, \tau)/\partial x$ evaluated at $x = 0$.) The conclusion from these equations depends on whether or not c vanishes.

First for $c \neq 0$, one concludes

$$P^{(0)} = 0, \qquad P^{(1)} = [(1 - c)/c]P^{(0)\prime},$$

$$P^{(2)} = [(1 - c)/c]P^{(1)\prime}, \cdots. \tag{16a}$$

In the last equation we have used the fact that $P^{(0)\prime\prime} = 0$ as a consequence of $P^{(0)} = 0$ and Eq. (14a). Thus we find that *to lowest order,* $P(x, \tau)$ *obeys the familiar boundary condition for an absorbing boundary whenever* c *does not vanish.* The reason is that any absorption, however small, is sufficient on our slow time scale to absorb everything that reaches the boundary. The next two orders in (16a) may be written in the form of the boundary condition

$$P(0, \tau) = \epsilon[(1 - c)/c]P'(0, \tau) + \mathcal{O}(\epsilon^3).$$

That is, to first and second order the value of $P(x, \tau)$ on the boundary differs from zero by an amount proportional to its slope. The proportionality constant $l = \epsilon(1 - c)/c$ has been called the slip coefficient.[3,7] In the next order, however, the boundary condition can no longer be expressed in terms of a single slip coefficient; rather one finds

$$P(0, \tau) = \epsilon[(1 - c)/c]P'(0, \tau) + \epsilon^3[(1 - c)(3 - 2c)/6c^2] \times P'''(0, \tau) + \cdots.$$

For the reflecting boundary $c = 0$, Eq. (15a) is moot, while (15b) and (15c) yield

$$P^{(0)\prime} = 0, \qquad P^{(1)\prime} = -\tfrac{1}{2}P^{(0)\prime\prime}. \tag{16b}$$

The same pattern continues in higher orders, e.g.,

$$P^{(2)\prime} = -\tfrac{1}{2}P^{(1)\prime\prime} - \tfrac{1}{6}P^{(0)\prime\prime\prime} = -\tfrac{1}{2}P^{(1)\prime\prime}. \qquad (16c)$$

Equations (14), together with the boundary conditions (16) constitute the desired approximate treatment in successive powers of ϵ. In order to compare this with the exact solution in Sec. 3 we write (11) in terms of the continuous variable x, setting $\vartheta = \epsilon k$,

$$\epsilon P_\lambda(x, \tau) = e^{-\lambda t}(2/\pi)^{1/2} \sin(kx + \eta). \qquad (17)$$

If one now expands λ as given in (9),

$$\lambda = \epsilon^2 k^2 - \tfrac{1}{12}\epsilon^2 k^4 + \cdots,$$

one finds that, indeed, (17) obeys (14) in successive orders.

To verify the boundary conditions, expand η as given by (10):

$$\eta = \text{Im}\log\{1 - (1 - c)e^{-i\epsilon k}\}$$

$$= \epsilon[(1 - c)/c]k + \mathcal{O}(\epsilon^3).$$

Substituting this in (17), one finds again that the successive orders of (17) obey the boundary conditions (16a). Note that, to second order, η is proportional to k, which according to (17) simply amounts to a displacement of the point where $P(x, \tau)$ vanishes from $x = 0$ to $x = -l$. This is equivalent with the existence of a slip coefficient and does not go through in higher orders.

In the reflecting case one finds by taking $c = 0$ in (10),

$$\eta = \tfrac{1}{2}\pi - \tfrac{1}{2}\epsilon k + \mathcal{O}(\epsilon^3).$$

On substituting this in (17) one easily finds agreement with (16b). Here again one may represent the effect of the boundary in first and second order as a shift of the boundary. The probability density $P(x, \tau)$ has no longer zero slope at $x = 0$, but at $x = \frac{1}{2}\epsilon$. The concept of a slip coefficient does not apply to the reflecting boundary.

In spite of this agreement it is not true that the approximate treatment leads to a *convergent* power series expansion of the evolution operator (12). The solutions of (14) involve all wave numbers k, whereas the integral in (12) contains values of k only up to π/ϵ. The solutions that we have added incorrectly are functions that vary appreciably within a distance ϵ and are therefore meaningless; they decay in time like e^{-t} rather than $e^{-\tau}$. Of course this kind of error is inevitable in any approximation in terms of continuous functions, and presumably in any approximation based on coarse graining. It demonstrates that the expansion in powers of ϵ can only be an asymptotic expansion.

5. EXAMPLE OF ONE-STEP RANDOM WALK WITH IMPURE BOUNDARY

Suppose a particle jumps randomly between neighboring points of a one-dimensional lattice, but the transition probabilities between the endpoint $n = 1$ and its neighbor $n = 2$ differ from the others:

$$\dot{p}_n = p_{n+1} + p_{n-1} - 2p_n, \qquad n = 3, 4, \ldots, \qquad (18a)$$

$$\dot{p}_2 = p_3 + ap_1 - (1 + b)p_2, \qquad (18b)$$

$$\dot{p}_1 = bp_2 - ap_1. \qquad (18c)$$

Note that the total probability W is conserved.

To find the normal mode solutions we first solve (18a) by setting $p_n^{(\lambda)}$ equal to (7) *for all* $n \geq 2$. As to $p_1^{(\lambda)}$, we know its time dependence,

$$p_1^{\{\lambda\}}(t) = e^{-\lambda t}q;$$

but its amplitude q is not determined by (18a). The two boundary equations (18b) and (18c), however, state

$$(1 + b - \lambda)(C_1 z_1^2 + C_2 z_2^2) = C_1 z_1^3 + C_2 z_2^3 + aq, \qquad (19a)$$

$$(a - \lambda)q = b(C_1 z_1^2 + C_2 z_2^2). \qquad (19b)$$

These equations determine C_1, C_2, and q uniquely up to an arbitrary normalization constant.

In order to make the transition to the continuous limit we set

$$\epsilon n = x, \qquad \epsilon^2 t = \tau, \qquad p_n(t) = \epsilon P(x, \tau) \qquad \text{for } n \geq 2.$$

Equation (18a) leads again to (14). Equations (18b) and (18c) serve to determine $p_1(\tau)$ and also to find the boundary condition for $P(x, \tau)$. We shall now use the dot for differentiation with respect to τ, so that the time derivatives pick up a factor ϵ^2. Thus (18b) and (18c) become

$$\epsilon^2 \dot{P}(2\epsilon) = P(3\epsilon) + (a/\epsilon)p_1 - (1 + b)P(2\epsilon),$$
$$\epsilon \dot{p}_1 = bP(2\epsilon) - (a/\epsilon)p_1.$$

Order ϵ^{-1} merely states $p_1^{\{0\}} = 0$. Order ϵ^0 yields

$$ap_1^{(1)} = bP^{(0)}, \qquad (20)$$

but does not yet lead to a boundary condition for $P^{(0)}$. Order ϵ^1 yields

$$0 = -bP^{(1)} + (1 - 2b)P^{(0)\prime} + ap_1^{\{2\}},$$
$$0 = bP^{(1)} + 2bP^{(0)\prime} - ap_1^{\{2\}}. \qquad (21)$$

Hence $P^{(0)\prime} = 0$ and $ap_1^{\{2\}} = bP^{(1)}$. Thus *to zeroth order P obeys the diffusion equation with the boundary condition for a reflecting wall.* According to (20) the probability for occupying site 1 is proportional to the value of the continuous probability density at the wall. The proportionality factor contains a factor ϵ owing to the fact that site 1 is only one state among a dense aggregate of other states. To the next order one finds

$$P^{(1)\prime} = (b/a - \tfrac{3}{2})P^{(0)\prime\prime},$$
$$ap_1^{(3)} = bP^{(2)} + b(2b/a - b - 1/a)P^{(0)\prime\prime}.$$

6. TWO-STEP RANDOM WALK WITH PURE BOUNDARY

The symmetric two-step random walk is described by the master equation

$$\dot{p}_n = \alpha p_{n+2} + p_{n+1} + p_{n-1} + \alpha p_{n-2} - (2 + 2\alpha)p_n. \qquad (22)$$

In the presence of a pure boundary this recursion relation applies to $n = 3, 4, \cdots$ and has to be supplemented by two special equations for p_1 and p_2. We take

$$\dot{p}_2 = \alpha p_4 + p_3 + p_1 - bp_2, \tag{23a}$$

$$\dot{p}_1 = \alpha p_3 + p_2 - ap_1, \tag{23b}$$

with two positive constants a and b.

For the total probability one finds from (22) and (23),

$$\dot{W} = (\alpha + 2 - b)p_2 + (\alpha + 1 - a)p_1. \tag{24}$$

Hence the boundary is reflecting when

$$a = 1 + \alpha, \qquad b = 2 + \alpha. \tag{25}$$

Smaller values of a or b will not be considered as they correspond to stimulated emission. Larger values correspond to an absorbing boundary. An equivalent formulation of the boundary conditions is obtained by declaring (22) valid for all $n \geq 1$, and setting

$$p_0 = [(2 + 2\alpha - b)/\alpha]p_2 \equiv Bp_2, \tag{26a}$$

$$p_{-1} = [(2 + 2\alpha - a)/\alpha]p_1 - [(2 + 2\alpha - b)/\alpha^2]p_2 \equiv Ap_1 - (B/\alpha)p_2. \tag{26b}$$

The constants A and B are simply abbreviations. Note that one must have

$$B \leq 1, \qquad A \leq 1 + 1/\alpha.$$

The equality sign corresponds to the reflecting case. The purely absorbing boundary is characterized by $A = B = 0$.

Solutions of (23a) are obtained by taking $p_n(t)$ equal to (5), where λ and z are now related to each other by

$$\lambda = 2 + 2\alpha - z - (1/z) - \alpha z^2 - (\alpha/z^2). \tag{27}$$

To each λ correspond four roots z_1, z_2, z_3, z_4, so that the general solution is

$$e^{-\lambda t}(C_1 z_1^n + C_2 z_2^n + C_3 z_3^n + C_4 z_4^n).$$

When one of the roots has absolute value greater than unity this solution grows exponentially for $n \to \infty$, unless the corresponding coefficient is zero. On the other hand, there are two boundary conditions (23), or equivalently (26), which can be satisfied only if at least three constants C are available.

The full discussion of the roots of (27) is elementary but laborious (Appendix C). When studying the continuous limit, however, one is interested only in the low values of λ. In that case it is easy to see that, in addition to the roots $z_1 = e^{i\vartheta}$, $z_2 = e^{-i\vartheta}$ given by

$$\lambda = 2 + 2\alpha - 2\cos\vartheta - 2\alpha\cos 2\vartheta, \tag{28}$$

there is another pair of roots z_3, z_4 with $|z_3| < 1$, and $|z_4| = 1/|z_3| > 1$ (see Eq. 39 in Appendix D). Hence

$$p_n^{(\lambda)}(t) = e^{-\lambda t}(C_1 e^{i\vartheta n} + C_2 e^{-i\vartheta n} + C_3 z_3^n). \tag{29}$$

The term with z_3^n decreases exponentially with increasing n, and is therefore nonpropagating. The phase shift η is still determined by $e^{2i\eta} = -C_1/C_2$.

Going to the continuous limit as before one obtains from (23a)

$$\frac{\partial P^{(0)}}{\partial \tau} = (1 + 4\alpha)\frac{\partial^2 P^{(0)}}{\partial x^2}, \tag{30a}$$

$$\frac{\partial P^{(1)}}{\partial \tau} = (1 + 4\alpha)\frac{\partial^2 P^{(1)}}{\partial x^2}, \tag{30b}$$

$$\frac{\partial P^{(2)}}{\partial \tau} = (1 + 4\alpha)\frac{\partial^2 P^{(2)}}{\partial x^2} + \frac{1 + 16\alpha}{12}\frac{\partial^4 P^{(0)}}{\partial x^4}. \tag{30c}$$

The reason why the exponentially decreasing term does not show up here is that $z_3^n = z_3^{x/\epsilon}$ collapses into the origin. In this connection it is essential to note that $|z_3|$ remains less than unity for $\epsilon \to 0$. In fact, one has from (27) with $\vartheta = \epsilon k$,

$$\lambda = \epsilon^2 k^2 + \cdots,$$

so that

$$z_3(\lambda) = z_3^{(0)} + \epsilon^2 z_3^{(2)} + \cdots.$$

Here $z_0^{(0)} = z_3(0)$ is the third root of (27) for $\lambda = 0$, which is readily seen to lie between -1 and 0.

On the other hand, it is clear that one does not have the freedom to impose two boundary conditions at $x = 0$ on the solutions of (30). The nonpropagating mode associated with z_3 is necessary in order to satisfy these boundary conditions. Thus even in the limit one needs more detailed information concerning the recursion relation (22) than just the form of its continuous approximation (30).[8]

To find the boundary condition for $P(x, \tau)$ at $x = 0$, it is convenient to study the individual modes separately for each time factor $e^{-\lambda t}$. Any factor λ that may arise in the boundary condition will ultimately be replaced with $(-\partial/\partial t)$, so that the boundary condition applies to all normal modes and therefore to $P(x, \tau)$ itself. Accordingly we substitute in (26),

$$p_n^{(\lambda)}(t) = e^{-\lambda t}(\epsilon P_\lambda(n\epsilon) + C_3 z_3^n).$$

To lowest order this leads to two equations for $C_3^{(0)}$ alone:

$$C_3^{(0)} = BC_3^{(0)} z_3^{(0)2},$$
$$C_3^{(0)} z_3^{(0)-1} = AC_3^{(0)} z_3^{(0)} - (B/\alpha)C_3^{(0)} z_3^{(0)2}.$$

It follows that $C_3^{(0)} = 0$, because $Bz_3^{(0)2} < 1$.

The next order yields two equations for $P_\lambda^{(0)}$ and $C_3^{(1)}$:

$$(1 - B)P_\lambda^{(0)} + \{1 - Bz_3^{(0)2}\}C_3^{(1)} = 0, \tag{31a}$$

$$(1 - A + B/\alpha)P_\lambda^{(0)} + \{z_3^{(0)-1} - Az_3^{(0)} + (B/\alpha)z_3^{(0)2}\}C_3^{(1)} = 0. \tag{31b}$$

Unless the determinant vanishes, the only solution is $P_\lambda^{(0)} = C_3^{(1)} = 0$. Thus to this order the exponentially decreasing solution does not come in, and $P^{(0)}(x, \tau)$ obeys the diffusion equation with the usual boundary condition for an absorbing boundary. In the reflecting case (25), however, one has $1 - B = 0$ and $1 - A + B/\alpha = 0$, so that the determinant of (31) does vanish. Hence there is a nontrivial solution

$$P_\lambda^{(0)} \text{ arbitrary}, \qquad C_3^{(1)} = 0.$$

In this case, therefore, (31) does not lead to a boundary condition. It is shown in Appendix D that the reflect-

ing boundary is the only case in which the determinant is zero.

To second order (26) gives

$$(1-B)P_\lambda^{(1)} + \{1 - Bz_3^{(0)2}\}C_3^{(2)} = 2BP_\lambda^{(0)\prime}, \tag{32a}$$

$$(1 - A + B/\alpha)P_\lambda^{(1)} + [z_3^{(0)-1} - Az_3^{(0)} + (B/\alpha)z_3^{(0)2}]C_3^{(2)} = (1 + A - 2B/\alpha)P_\lambda^{(0)\prime}. \tag{32b}$$

By solving these equations one obtains the boundary value of $P^{(1)}(x, \tau)$ in terms of the slope of $P^{(0)}(x, \tau)$ at $x = 0$. For instance, in the case of pure absorption $A = B = 0$, one finds successively

$$P^{(1)} = -[z_3^{(0)}/(1 - z_3^{(0)})]P^{(0)\prime},$$
$$C_3^{(2)} = [z_3^{(0)}/(1 - z_3^{(0)})]P_\lambda^{(0)\prime},$$
$$P^{(2)} = [z_3^{(0)}/(1 - z_3^{(0)})]P^{(1)\prime},$$
$$C_3^{(3)} = [z_3^{(0)}/(1 - z_3^{(0)})]P_\lambda^{(1)\prime}.$$

This may again be interpreted in terms of a slip coefficient

$$l = -\epsilon[z_3^{(0)}/(1 - z_3^{(0)})],$$

or alternatively as a displacement of the boundary point to

$$x = -l = \epsilon[z_3^{(0)}/(1 - z_3^{(0)})].$$

In the reflecting case both coefficients of $P_\lambda^{(1)}$ in (32) are zero and the only solution is $C_3^{(2)} = P_\lambda^{(0)} = 0$ (see Appendix D). After somewhat lengthy computations one finds in the next order

$$P^{(1)\prime} = -\tfrac{1}{2}P^{(0)\prime\prime}, \quad C_3^{(3)} = P_\lambda^{(0)\prime\prime}[1 - z_3^{(0)2}]^{-1}. \tag{33}$$

7. THE TOTAL PROBABILITY

In this section we investigate the behavior of the total probability in the successive orders of ϵ. Provided that the boundary is pure, one may use the Euler–Maclaurin formula

$$W = \sum_{n=1}^{\infty} p_n(t) = \epsilon \sum_{n=1}^{\infty} P(n\epsilon, \tau)$$
$$= \int_0^\infty P(x, \tau)dx - \tfrac{1}{2}\epsilon P(0, \tau)$$
$$- \tfrac{1}{12}\epsilon^2 P'(0, \tau) + \tfrac{1}{720}\epsilon^4 P'''(0, \tau) + \cdots.$$

Hence the successive orders of W are

$$W^{(0)} = \int_0^\infty P^{(0)}(x, \tau)dx, \tag{34a}$$

$$W^{(1)} = \int_0^\infty P^{(1)}(x, \tau)dx - \tfrac{1}{2}P^{(0)}, \tag{34b}$$

$$W^{(2)} = \int_0^\infty P^{(2)}(x, \tau)dx - \tfrac{1}{2}P^{(1)} - \tfrac{1}{12}P^{(0)\prime}. \tag{34c}$$

Consider the *one-step random walk* treated in Secs. 3 and 4. Using (14) one obtains, for the derivatives with respect to τ,

$$\dot{W}^{(0)} = -P^{(0)\prime},$$
$$\dot{W}^{(1)} = -P^{(1)\prime} - \tfrac{1}{2}P^{(0)\prime\prime},$$
$$\dot{W}^{(2)} = -P^{(2)\prime} - \tfrac{1}{2}P^{(1)\prime\prime} - \tfrac{1}{6}P^{(0)\prime\prime\prime}.$$

For the reflecting case each term duly vanishes when the boundary conditions (16b) and (16c) are inserted. For the absorbing case the equations reduce to

$$\dot{W}^{(0)} = -P^{(0)\prime}, \quad \dot{W}^{(1)} = -P^{(1)\prime},$$
$$\dot{W}^{(2)} = -P^{(2)\prime} - [(3 - 2c)/6c]P^{(0)\prime\prime\prime}.$$

These equations may be combined into

$$\dot{W} = -P'(0, \tau) - \epsilon^2[(3 - 2c)/6c]P'''(0, \tau) + \mathcal{O}(\epsilon^3).$$

Consider the *two-step random walk* with pure boundary treated in Sec. 6. One now has, for each normal mode, using $\lambda = \epsilon^2 k^2$,

$$W_\lambda(\tau) = e^{-k^2\tau}\left(\epsilon\sum_{n=1}^{\infty} P_\lambda(n\epsilon) + C_3\frac{z_3}{1 - z_3}\right).$$

From this one finds, taking into account that $C_3^{(0)} = C_3^{(1)} = 0$,

$$\dot{W}^{(0)} = -(1 + 4\alpha)P^{(0)\prime},$$
$$\dot{W}^{(1)} = -(1 + 4\alpha)P^{(1)} - \tfrac{1}{2}(1 + 4\alpha)P^{(0)\prime\prime},$$
$$\dot{W}_\lambda^{(2)} = e^{-k^2\tau}[-(1 + 4\alpha)P_\lambda^{(2)\prime} - \tfrac{1}{2}(1 + 4\alpha)P_\lambda^{(1)\prime\prime}$$
$$- \tfrac{1}{6}(1 + 10\alpha)P_\lambda^{(0)\prime\prime\prime} - k^2C_3^{(2)}z_3^{(0)}/(1 - z_3^{(0)})].$$

For the *reflecting* boundary it is clear that $\dot{W}^{(0)} = 0$ and $\dot{W}^{(1)} = 0$, see (33). It is now easier to find the next higher order boundary condition for the reflecting case from the conservation of probability, since we found already $C_3^{(2)} = 0$,

$$P^{(2)\prime} = -\tfrac{1}{2}P^{(1)\prime\prime}.$$

Again this can be interpreted by saying that to second order in ϵ the slope of $P(x, \tau)$ vanishes at $x = \frac{1}{2}\epsilon$ rather than at $x = 0$.

For the two-step random walk with *purely absorbing boundary*, one finds for the total probability

$$\dot{W}^{(0)} = -(1 + 4\alpha)P^{(0)\prime}, \quad \dot{W}^{(1)} = -(1 + 4\alpha)P^{(1)\prime},$$

and to second order

$$\dot{W}_\lambda^{(2)} = e^{-k^2\tau}\left[-(1 + 4\alpha)P_\lambda^{(2)\prime} + \frac{1 + 4\alpha}{2}\frac{z_3^{(0)}}{1 - z_3^{(0)}}P_\lambda^{(0)\prime\prime\prime} - \frac{1 + 10\alpha}{6}P_\lambda^{(0)\prime\prime\prime} - k^2\left(\frac{z_3^{(0)}}{1 - z_3^{(0)}}\right)^2 P_\lambda^{(0)\prime}\right].$$

Hence,

$$\dot{W}^{(2)} = -(1 + 4\alpha)P^{(2)} + \left(\frac{1 + 4\alpha}{2}\frac{z_3^{(0)}(1 + z_3^{(0)})}{(1 - z_3^{(0)})^2} - \frac{1 + 10\alpha}{6}\right)P^{(0)\prime\prime\prime}.$$

Finally consider the *impure boundary* case of Sec. 5. One now has

$$W = p_1 + \epsilon\sum_{n=2}^{\infty} P(n\epsilon)$$
$$= p_1 - \epsilon P(\epsilon) + \int_0^\infty P(x, \tau)dx$$
$$- \tfrac{1}{2}\epsilon P(0, \tau) - \tfrac{1}{12}\epsilon^2 P'(0, \tau).$$

From this one finds that, indeed, the total probability is conserved in successive orders of ϵ:

$$\dot{W}^{(0)} = \int_0^\infty \dot{P}^{(0)}(x, \tau)dx = -P^{(0)\prime} = 0,$$
$$\dot{W}^{(1)} = \dot{p}_1^{\{1\}} - P^{(1)\prime} - \tfrac{3}{2}\dot{P}^{(0)}$$
$$= (b/a)P^{(0)\prime\prime} - (b/a - \tfrac{3}{2})P^{(0)\prime\prime} - \tfrac{3}{2}P^{(0)\prime\prime} = 0.$$

8. CONCLUSIONS

(i) Expansion of the master equation of a random walk leads to the familiar diffusion equation for the zeroth and first order, and to diffusion equations with inhomogeneous term for the higher orders. These differential equations are valid at all interior points of the accessible interval.

(ii) In the case of a pure boundary, the continuous probability density $P(x, \tau)$ obeying these equations constitutes a full description of all probabilities; in the case of an impure boundary one or more separate quantities $p_1(\tau), p_2(\tau), \cdots$ are needed for the probabilities at the sites near the boundary.

(iii) For one-step processes the boundary conditions on $P(x, \tau)$ are obtained by substituting in the discrete master equation simply $p_n(\tau) = \epsilon P(n\epsilon, \tau)$, possibly with separate values for $p_1(\tau), p_2(\tau), \cdots$. For s-step processes, however, one has to add linear combinations of $s - 1$ additional nonpropagating modes of the homogeneous master equation.

(iv) To zeroth order the result is: The slope of $P(x, \tau)$ vanishes at any boundary that conserves probability, but $P(x, \tau)$ itself vanishes as soon as there is some absorption. To second order the same boundary conditions apply but at a slightly displaced point. In higher orders the boundary condition cannot be stated in such simple terms.

ACKNOWLEDGMENTS

The authors would like to thank Professor K. E. Shuler and Dr. R. Trimble for helpful discussions.

APPENDIX A

In addition to the solutions (7) where z_1, z_2 are given by (8), there may be solutions with $|z_1| < 1$, $|z_2| > 1$, and $C_2 = 0$. According to (10) this requires that $z_1 = (1 - c)^{-1}$, which is consistent with $|z_1| < 1$ only when $c > 0$. The corresponding time constant λ is

$$\lambda_* = c + 1 + 1/(c - 1).$$

This is an isolated point of the spectrum because $\lambda_* > 4$. The corresponding normal mode solution is

$$p_n^{(*)}(t) = e^{-\lambda_* t}(1 - c)^{-n}.$$

Note that one has $z_1 = e^{i\vartheta_*}$, where ϑ_* is a pole of the S matrix in the upper half of the complex ϑ plane. More explicitly, $\vartheta_* = \pi + i\kappa$ with $e^{\kappa} = c - 1$. This additional isolated solution has to be included when $c > 2$, in order that the solutions are complete. However, this fact is not relevant when studying the continuous limit, because that limit involves only low values of λ.

To investigate the completeness of the set of solutions (11) with $0 \leq \vartheta \leq \pi$, we verify the completeness relation, or to put it differently, we compute $U_{nm}(0)$:

$$\frac{2}{\pi}\int_0^{\pi} \sin[\vartheta n + \eta(\vartheta)]\sin[\vartheta m + \eta(\vartheta)]d\vartheta$$
$$= \frac{1}{2\pi}\int_0^{2\pi}(e^{i\vartheta(n-m)} - e^{i\vartheta(n+m)+2i\eta(\vartheta)})d\vartheta$$
$$= \delta_{nm} - \frac{1}{2\pi}\int_0^{2\pi} S(\vartheta)e^{i(n+m)\vartheta}\,d\vartheta.$$

The integration path may be extended by adding the line from $i\infty$ to 0, and the line from 2π to $2\pi + i\infty$, because their contributions cancel owing to the periodicity. Furthermore the integrand vanishes at $i\infty$ since $S(\vartheta) \sim e^{|\vartheta|}$, whereas $n + m \geq 2$. For $0 \leq c \leq 2$, there are no poles in the upper half plane, so that the integral vanishes and the desired completeness relation is obtained. For $c > 2$ there is a pole $\vartheta_* = \pi + i\kappa$, whose residue yields for the integral

$$c(c - 2)e^{-\kappa(n+m)}.$$

This shows that for $c > 2$ one has to include the isolated solution

$$p_n^{(*)}(t) = e^{-\lambda_* t}[c(c - 2)]^{1/2}(c - 1)^{-n}.$$

APPENDIX B

From (19) follows on eliminating q and expressing λ through (9),

$$-\frac{C_1}{C_2} = -e^{-i\vartheta}\,\frac{1 + (a + b - 2)e^{-i\vartheta} + (1 - b)e^{-2i\vartheta}}{1 + (a + b - 2)e^{i\vartheta} + (1 - b)e^{2i\vartheta}}. \quad \text{(B1)}$$

Again denoting this quantity by $e^{2i\eta(\vartheta)}$, one sees that the normal mode solutions are given by (11) for $n \geq 2$, and in addition one has

$$p_1^{(\lambda)} = e^{-\lambda t}(2/\pi)^{1/2}\,\frac{b\,\sin[2\vartheta + \eta(\vartheta)]}{a - \lambda}.$$

The denominator vanishes for $\lambda = a$, but so does the numerator.

The denominator of (B1) has one zero in the upper half plane when $a + 2b > 4$ and otherwise none. The zero is given by

$$e^{i\vartheta} = \{2 - a - b + [(a + b)^2 - 4a]^{1/2}\}/2(1 - b).$$

When going to the limit one has $\vartheta = \epsilon k$, so that in terms of k the pole moves to $+i\infty$ when ϵ tends to zero. This is the reason why the corresponding normal mode does not appear in the expansion in powers of ϵ; but it also demonstrates that the expansion can only be an asymptotic one.

APPENDIX C

To solve (27) for z put $z + z^{-1} = 2w$:

$$4\alpha w^2 + 2w + (\lambda - 2 - 4\alpha) = 0. \quad \text{(C1)}$$

Either solution of this quadratic equation for w gives rise to two roots z to be found from

$$z^2 - 2wz + 1 = 0.$$

Thus the four roots of (27) consist of two pairs z_1, z_2 and z_3, z_4, such that $z_1 z_2 = 1$ and $z_3 z_4 = 1$. In order that there are at least three roots with $|z_j| \leq 1$, it is necessary that at least one pair has the form $e^{i\vartheta}, e^{-i\vartheta}$ with real ϑ. Hence all admissible values of λ are of the form (28), and lie therefore between 0 and

$$2 + 4\alpha + 1/4\alpha = [(4\alpha)^{1/2} + 1/\sqrt{4\alpha})]^2. \quad \text{(C2)}$$

This upper bound is also the condition that the two solutions of (C1)

$$w_{\pm} = \{-1 \pm [(4\alpha + 1)^2 - 4\alpha\lambda]^{1/2}\}/4\alpha$$

are real. In addition it is necessary that at least one of these has the form $\cos\vartheta$, i.e., that it lies between -1 and $+1$.

To investigate this condition we have to distinguish between the two cases $\alpha \leq \frac{1}{4}$ and $\alpha > \frac{1}{4}$. In the case $\alpha \leq \frac{1}{4}$, one readily finds that $w_- < -1$, while $|w_+| \leq 1$ holds only for $0 \leq \lambda \leq 4$. This determines the spectrum of the bounded two-step random walk with $\alpha \leq \frac{1}{4}$; to each λ in $(0, 4)$ belongs a single normal mode solution.

In the case $\alpha > \frac{1}{4}$, the same nondegenerate spectrum exists; but in addition one finds that for λ between 4 and the upper bound (C2) both w_+ and w_- lie in the interval $(-1, +1)$. Hence for these values of λ there are *four* admissible roots z_j, and therefore *two* linearly independent solutions of the bounded two-step random walk with $\alpha > \frac{1}{4}$. For the low values of λ relevant for the continuous limit this complication cannot occur.

APPENDIX D

The determinant of (31) is

$$\begin{vmatrix} 1-B & 1-Bz^2 \\ 1-A+B/\alpha & z^{-1}-Az+(B/\alpha)z^2 \end{vmatrix},$$

where z stands for $z_3^{(0)}$. After subtracting the left column from the column on the right, a factor $1-z$ splits off. Subsequently multiplying the top row with $1/\alpha$ and adding it to the bottom row, and adding the left column to the right one, we get

$$(1-z)\begin{vmatrix} 1-B & 1+Bz \\ 1+(1/\alpha)-A & (1/z)+1+1/\alpha \end{vmatrix}.$$

The two elements of the left column are nonnegative and will be denoted by $R = 1-B$, $S = 1+1/\alpha-A$. In terms of R and S the value of the determinant is found to be

$$(1-z)\{RSz-(1+z)S+[(1/z)+1+1/\alpha]R\}.$$

The first and second terms in $\{\ \}$ are negative (or zero when $S=0$). Hence, if we show that the coefficient of the third term is negative, it proves that the determinant cannot vanish unless $R = S = 0$, i.e., unless the boundary is reflecting. It suffices to show

$$1+[1+(1/\alpha)]z < 0, \tag{D1}$$

where the explicit value of z is found from solving (27) for $\lambda = 0$,

$$z = z_3^{(0)} = -(1+1/2\alpha)+[(1/\alpha)+1/4\alpha^2]^{1/2}. \tag{D2}$$

On substituting this z in (D1), one easily finds that the inequality is true. This completes the proof.

APPENDIX E

Equations (32) reduce in the reflecting case to

$$(1-z^2)C_3^{(2)} = 2P_\lambda^{(0)\prime},$$

$$\{z^{-1}-[1+(1/\alpha)]z+(1/\alpha)]z^2\}C_3^{(2)} = [2-(1/\alpha)]P_\lambda^{(0)\prime}.$$

The determinant is

$$(1-z)\begin{vmatrix} 1+z & -2 \\ (1/z)+1-z/\alpha & -2+1/\alpha \end{vmatrix} = (2/z)(1-z)^2\{1+(1+1/2\alpha)z\}.$$

The first two factors are not zero and for the factor $\{\ \}$ one has, using (D1),

$$1+z+(z/2\alpha) > 1+z+(z/\alpha) > 0.$$

This justifies the statement that the only solution of (32) in the reflecting case is $P^{(0)\prime} = C_3^{(2)} = 0$.

---

* Supported in part by the National Science Foundation.

† Permanent address: Institute for Theoretical Physics, Rijksuniversiteit, Utrecht, Netherlands.

1 S. T. Choh and G. E. Uhlenbeck, *The Kinetic Theory of Dense Gases* (University of Michigan 1958); E. G. D. Cohen, "The Boltzmann Equation and its Generalization to Higher Densities," in *Fundamental Problems in Statistical Mechanics* edited by E. G. D. Cohen (North-Holland, Amsterdam, 1962).

2 For literature see P. A. Selwyn and I. Oppenheim, Physica **54**, 161 (1971).

3 (a) J. C. Maxwell, Phil. Trans. [A] **170**, 231 (1879); H. A. Kramers, Nuovo Cimento Suppl. **6**, 297 (1949). (b) For the general equation (1) the precise condition is: Those r for which a_r does not vanish must not have a common divisor other than unity.[4] (c) Those who feel that total probability should always be equal to unity may either read "number of particles" instead of "probability," or add a limbo state containing all absorbed probability.

4 F. L. Spitzer and C. J. Stone, Illinois J. Math. **4**, 253 (1960).

5 (a) See e.g. W. Feller, *An Introduction to Probability Theory and Its Applications* (Wiley, New York, 1957), 2nd ed., Vol. 1, Chaps. 14 and 16; F. L. Spitzer, *Principles of Random Walk* (Van Nostrand, Princeton, N.J. 1964), Chap. 4; D. R. Cox and H. D. Miller, *The Theory of Stochastic Processes* (Wiley, New York, 1965), Chap. 2; K. Lakatos-Lindenberg and K. E. Shuler J. Math. Phys. **12**, 633 (1971). (b) This expansion is more elementary than the one in Ref. 6, and is only possible for a homogeneous case, which has no equilibrium distribution with finite variance.

6 N. G. van Kampen, Can. J. Phys. **39**, 551 (1961); also in *Fluctuation Phenomena in Solids*, edited by R. E. Burgess (Academic, New York, 1965).

7 H. Helmholtz and G. v. Piotrowski, Wien. Sitzungsber. **40**, 607 (1860)

8 This statement needs to be qualified inasmuch as for an s-step random walk the coefficients in the continuous approximation (30) together uniquely fix the a_r and thereby the recursion relation from which one started. In principle, therefore, they contain all the information needed for computing z_3, but in a rather recondite form.

Topological Analysis of Self-Generating Interactions in Lagrangian Field Theory

E. Gunzig
Faculté des Sciences, Université Libre de Bruxelles, Bruxelles, Belgium
(Received 16 February 1971)

We determine entirely the n-body nonlocal potentials, parameters of a given Lagrangian field theory, in terms of a set of independent functionals of the many-body propagators. These functionals are more suitable for the description of self-generating interactions than the many-body propagators themselves.

1. INTRODUCTION

In previous papers,[1,2] De Dominicis and Englert set up a formalism for a systematic study of self-generating interactions in Lagrangian field theory.

The following procedure was adopted: One considers the limiting case in which the set of all n-body nonlocal interaction potentials ($n \geq 2$), the parameters of a given nonlocal Lagrangian field theory, vanishes.[2-5] One then analyzes the structure of the n-body propagators in this limit. The possible existence of nontrivial (i.e., nonfree field), n-body propagators would give rise to self-generating interactions.

In other words, starting from a free Lagrangian field theory, one first simulates an interaction by a set of nonlocal potentials v_n, $n \geq 2$. Then, in the limiting case of the whole set vanishing, one searches an unconventional solution for the n-body propagators, which exhibits some residue of the original interaction; this would imply self-generating interactions.

In order to make such an idea useful, the authors of Ref 1 were led to consider the following problem:

Is it possible to characterize *ab initio* a Lagrangian field theory by all its Green's functions, namely to deduce from them the set of all n-body potentials? This problem was solved[1] in a formally simple manner:

The n-body potentials were expressed by the set of all corresponding connected Feynman graphs, with connected Green's functions as vertices and fully dressed propagators, with reversed sign, as internal lines.

Putting these potentials to zero enables one, in principle, to search for self-generating interactions.

However, the inversion formula,[1,6] inspite of its great formal simplicity, is not useful in realizing this program, essentially due to the enormous complexity of the set of all connected Feynmann graphs contributing to the n-body potentials. For this reason one introduces a set of independent Γ variables,[2] equally suitable to characterize a given field theory as the original Green's functions, and closely related to them.

They are obtained from the corresponding Green's functions, by undressing their external lines with suitable functionals of lower order propagators (a precise definition will be given in Sec. 2).

The crucial point is that the inversion formula, when expressed in terms of these new variables, takes a much simpler form than the original one. Indeed, the potentials may now be expressed in terms of a very restricted family of diagrams, thus making it feasible to search for self-generating interactions.

Our aim in this work, is to determine explicitly the n-body potentials in terms of the Γ variables. In Sec. 2 we give a set of definitions which will be used throughout this work and in particular a precise definition of the Γ variables.

Together with the inversion formula, this yields the functional relations relating the potentials to the Γ variables.

If the equations obtained by putting the whole set of potentials v_n, $n \geq 2$ to zero in these functional relations, have a nontrivial solution for Γ, then these Γ will characterise the self-generating interactions.

These equations are derivable[2] from a stationarity principle on a functional, which involves only the n-body potentials expressed in terms of the first n Γ variables namely $v_n(\Gamma_2, \ldots, \Gamma_n)$, $\Gamma_m = 0$, $m > n$. Hence, the structure of the equations, characterizing the self-generating interactions, depends crucially on these particular functionals; their explicit form has been announced in Ref. (2) and we shall concentrate on rigorously determining these.

In other words, we shall express diagrammatically the functional relations connecting the potentials to the Γ variables, taking into account the condition $\Gamma_m = 0$, $m > n$, denoted from now on by $(\mathrm{mod}\Gamma_n)$.

As we have mentioned, we shall find that only a very restricted set of diagrams will contribute to v_n. The set of all other graphs will be eliminated by two distinct mechanisms.

The first is related to the particular functional dependence of the potentials in terms of the Γ variables, which determines the numerical coefficient with which a given graph appears in the expression for v_n; this coefficient, called the weight of the given graph, may be zero. In that case, of course, this particular diagram disappears from the set of contributions to the potential.

The second mechanism is completely independent of the functional relation and is of a purely topological nature; namely, we shall prove that a large class of diagrams is topologically forbidden.

The weight of an arbitrary graph is obtained in Sec. 3, where a general equation is given.

In Sec. 4, a systematic investigation is made of the weights of graphs belonging to various families.

Section 5 is devoted to the topological problem.

2. DEFINITIONS

Among the following definitions, some are of secondary importance for the development of subsequent work; since these were completely defined in Ref. 1, we shall herein merely give a brief indication to make this paper self-contained. Only the rigourously given definitions will be numbered.

Let $\Phi(x)$ be a spinless charged boson field (Heisenberg operator).

Let $G_{2n}(x_1, \ldots, x_n; y_1, \ldots, y_n) = \langle \Omega | T\phi(x_1) \cdots \phi^+(y_1) \cdots \phi^+(y_n) | \Omega \rangle$, $n \geq 2$ denote the n-body propagator or Green's function; $|\Omega\rangle$ is the normalized physical vacuum and T the time-ordered product operator.

$$G_1(x; y) = \langle \Omega | T\Phi(x); \Phi^+(y) | \Omega \rangle$$

is the fully dressed one-body propagator.

Let $\tilde{G}_{2n}(x_1, \ldots, x_n; y_1, \ldots, y_n)$ be that part of G_{2n}, not containing independent propagations of any subset of $m < n$ particles; if G_{2n} is expressible in terms of Feynman diagrams, $\tilde{G}_{2n}$ is the completely connected part of G_{2n}.

Let $C_{2n}(x_1, \ldots, x_n; y_1, \ldots, y_n)$, $n \geq 2$, denote $\tilde{G}_{2n}$ with the $2n$ external lines removed.

Let $W_{2n}(x_1, \ldots, x_n; y_1, \ldots, y_n)$, $n \geq 2$, denote any set of nonlocal interaction potentials, symmetrized in both x and y variables.

Represent by $\tilde{\mathfrak{F}}^{\Sigma}_{2n}(G_1; W_{2m})$, $m \geq 2$, the functional of the potentials represented diagrammatically by the sum of all connected Feynman graphs drawn with W's as vertices and G_1 as internal oriented (from an x to a y variable) lines, having $2n$ external (oriented) lines. Furthermore, each diagram is accompanied by the numerical coefficient $1/g$, where g denotes the order of its symmetry group. This group is, as usual, defined as the set of all permutations of the internal lines of a given graph which leaves it topologically unchanged.

Represent by $\bar{\mathfrak{F}}^{\Sigma}_{2n}(G_1; W_{2m})$ the same functional as the previous one, but with the $2n$ external lines amputated from all graphs.

To denote the corresponding families of diagrams (and not the sum of all its members), we use the same symbols but without the superscript Σ.

Denote by $\tilde{\mathfrak{F}} \equiv \cup_n \tilde{\mathfrak{F}}_{2n}$ the set of all connected Feynman diagrams (of any order).

In the same way, $\bar{\mathfrak{F}} \equiv \cup_n \bar{\mathfrak{F}}_{2n}$.

Definition 1: The dimension dim$G$ of a given graph $G \in \tilde{\mathfrak{F}}$ is the number of external lines of G; if $G \in \bar{\mathfrak{F}}$, then dim$G$ is the number of external stumps of G.

Definition 2: A $2p$-subdiagram or a $2p$-part of a given $G \in \tilde{\mathfrak{F}}$ is a member of $\tilde{\mathfrak{F}}_{2p}$ obtained from G by cutting $2p$ (internal or external) lines; if $G \in \bar{\mathfrak{F}}$, the same definition holds but after redressing G by external lines.

We now define the three following classes of connected diagrams[2]:

Definition 3: $\tilde{g}_{2n}(G_1; W_{2m}) \subset \tilde{\mathfrak{F}}_{2n}(G_1; W_{2m})$, $m \geq 2$, satisfies the following requirements:

(i) Each member of $\tilde{g}_{2n}$ contains at most one W_{2n} and no W_{2m}, $m \geq n$;

(ii) Let G_{2p} be an arbitrary $2p$-part of a diagram belonging to $\tilde{g}_{2n}$, $p < n$, then $G_{2p} \in \tilde{g}_{2p}$.

Examples of graphs belonging and not belonging to $\tilde{g}$ are given in Fig. 1.

Definition 4: $\tilde{I}_{2n}(G_1; W_{2m}) \subset \tilde{\mathfrak{F}}_{2n}(G_1; W_{2m})$, $m \geq 2$. The members of this set are called irreducible diagrams and are defined by the following requirements: Let G_{2p} be an arbitrary $2p$ part of a diagram belonging to $\tilde{I}_{2n}$, $p \leq n$, then $G_{2p} \in \tilde{g}_{2p}$, except for the whole diagram (Fig. 2).

Definition 5: $\tilde{R}_{2n}(G_1; W_{2m}) \subset \tilde{\mathfrak{F}}_{2n}(G_1; W_{2m})$, $m \geq 2$. The members of this set are called reducible diagrams and are defined by (Fig. 2)

$$\tilde{R}_{2n} \equiv \mathcal{C}_{\tilde{\mathfrak{F}}_{2n}} \tilde{g}_{2n} \cup \tilde{I}_{2n},$$

where we will always use the symbol $\mathcal{C}$ for the complement (in the usual sense).

All these definitions hold for the corresponding graphs with external lines cut. In the latter case we use the same symbols with superscript—instead of ~.

Let $G^i \in \tilde{\mathfrak{F}}$, $\forall\, i$. If the G^i have a common subdiagram, then they are said to intersect.

With these definitions in mind, we can now give an exact definition of the Γ variables mentioned in the introduction.

$\Gamma_{2n}(x_1 \cdots x_n; y_1 \cdots y_n)$, $n \geq 2$, are the functions symmetrized in both x and y variables which satisfy the functional relations

$$\tilde{G}_{2n} = \tilde{g}^{\Sigma}_{2n}(G_1; \Gamma_{2m}), \qquad m \geq 2, \tag{2.1}$$

or, with external lines removed

$$C_{2n} = \bar{g}^{\Sigma}_{2n}(G_1; \Gamma_{2m}), \qquad m \geq 2, \tag{2.2}$$

On the other hand, the inversion formula, expressing the n-body potential in terms of the connected Green's functions, reads

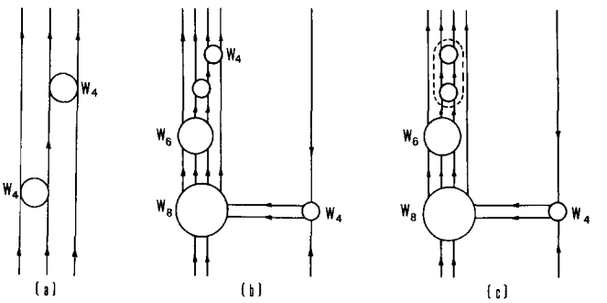


FIG. 1. (a) Graph belonging to $\tilde{g}_6$. (b) Graph belonging to $\tilde{g}_8$. (c) Graph not belonging to $\tilde{g}_8$; the dotted loop indicates the 4-part not contained in $\tilde{g}_4$.

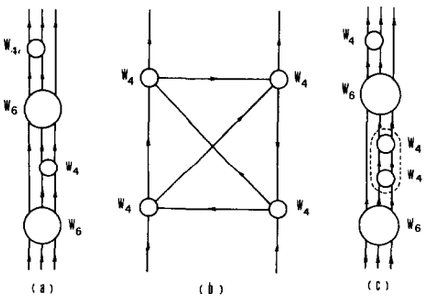


FIG. 2. (a) Graph belonging to $\tilde{I}_6$. (b) Graph belonging to $\tilde{I}_4$. (c) Graph belonging to $\tilde{R}_6$; the dotted loop indicates the 4-part not contained in $\tilde{g}_4$.

$$v_{2n} = \bar{\mathfrak{F}}^{\Sigma}_{2n}(-G_1; C_{2m}), \qquad m \geq 2, \tag{2.3}$$

the notation $-G_1$ indicates that the evaluation of the diagrams is done in the conventional way, but attributing the value $(-G_1)$ instead of $(+G_1)$ to each internal line.

Combining (2.2) with (2.3) yields

$$v_{2n} = \bar{\mathfrak{F}}^{\Sigma}_{2n}(-G_1; \bar{g}^{\Sigma}_{2m}(G_1; \Gamma_{2l})), \qquad m, l \geq 2. \tag{2.4}$$

This is the basic functional relation connecting the n-body potential v_{2n} to the set of Γ variables.

We shall now make this relation explicit diagrammatically (modΓ_{2n}), by determining the weight of an arbitrary graph (drawn with the Γ as vertices) in the expression for the potential.

3. THE WEIGHT OF AN ARBITRARY GRAPH

Let G^N be an arbitrarily chosen member of the family $\bar{\mathfrak{F}}_{2n}(G_1; \Gamma_{2m})$; N stands for the number of internal lines of the graph G. To alleviate our notation, we shall henceforth drop indices whenever there is no danger of confusion.

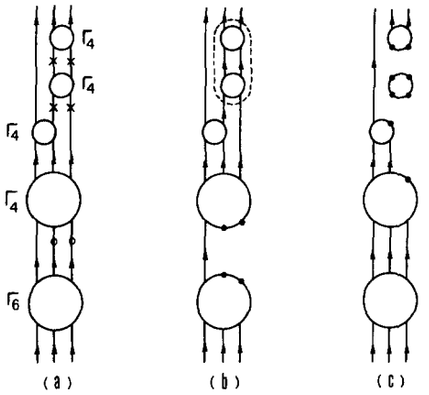


FIG. 3. (a) The graph G; the two lines belonging to a particular $e^i(e = 2)$ are marked by ○ and the four lines belonging to another $e_i(e = 4)$ are marked by ×. (b) A connected 2-support (but not $\bar{2}$-support). · denotes a stump. The dotted loop indicates the 4-part of the 2-support not contained in $\bar{g}_4$. (c) A nonconnected 4-support.

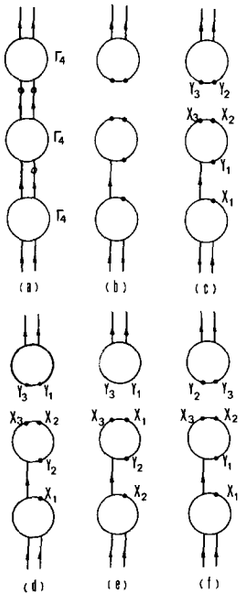


FIG. 4. (a) the graph G; the three lines belonging to a particular $e^i(e = 3)$ are marked by ○. · denotes a stump. (b) A nonconnected $\bar{3}$-support of G. (c) a 3-skeleton attached to the $\bar{3}$-support in (b). (d) A structure which is not a 3-skeleton. (e) A different 3-skeleton attached to the 3-support in (b). (f) A 3-skeleton identical to the one in (c), by symmetry properties of Γ_4.

Our task consists in determining the weight of the given diagram.

With G we associate a set E_G defined as follows: E_G is the set whose elements e^i_G are the various subsets of the N internal lines of G, including the empty subset $\emptyset$; e denotes the number of constituent lines and i fixes the particular choice of these lines. Thus

$$\begin{cases} 0 \leq e \leq N \\ 1 \leq i \leq \binom{N}{e}. \end{cases}$$

We shall also need to consider the family $\check{\mathfrak{F}}$ whose elements are the subsets of $\bar{\mathfrak{F}}$.

The dimension $2m$ of an element $\check{f}_{2m} \in \check{\mathfrak{F}}$ is defined to be the sum of the dimensions of all graphs in $\check{f}$.

We define a mapping $F_G: E_G \to \check{\mathfrak{F}}: e^i_G \mapsto F_G(e^i_G)$ defined in the following manner: $F_G(e^i_G)$ is an element $\check{f}_{2(n+e)} \in \check{\mathfrak{F}}$; $\check{f}_{2(n+e)}$ is obtained from G by amputation of the e lines of e^i_G; $2e$ is the number of stumps resulting from the above mentioned amputation.

A given element $\check{f} \in \check{\mathfrak{F}}$, image by F_G of one or more $e^i_G \in E_G$ will be called e-support (e_s) of G, unless each graph of $\check{f}$ is contained in a family $\bar{g}(G_1; \Gamma)$ (cf. Def. 3). (Clearly all the elements in $F^{-1}_G(\check{f})$ contains the same number of lines so e is the same). In the latter case, it will be called $\bar{e}$-support of G, $(\bar{e}_s)$, Fig. 3. A particular $e(\bar{e})$-support of G will be denoted $e^\alpha_s(\bar{e}^\alpha_s)$.

Let $\bar{E}_G \equiv F^{-1}_G \{ \bigcup_{\text{all } e,\alpha} \bar{e}^\alpha_s \}$.

Clearly $\bar{E}_G \subseteq E_G$

Each $e_s(\bar{e}_s)$ defines a set of e-skeletons ($\bar{e}$-skeletons) of G obtained from $e_s(\bar{e}_s)$ by attaching to its $2e$ stumps, $2e$ variables $\{x_i, y_i\}$, $1 \leq i \leq e$, in all possible ways, subject to the following conditions:

(i) The insertion of e oriented lines joining x_i to y_i for each i must yield back the graph G.

Notice therefore that each e-skeleton has the same topological structure as the corresponding e_s but its stumps are labeled with $2e$ variables according to the above prescription.

(ii) Two e-skeletons are considered to be distinct if their respective assignments of the $2e$ variables to stumps are distinct, modulo the symmetry properties of the Γ (Fig. 4).

It follows from these definitions and from the nature of the functional relation (2.4) that the number of distinct mechanisms leading to G, in the expression for the n-body potential v_{2n} is the number of distinct $\bar{e}_s$; each one of these mechanisms can be realized as many times as there are distinct e-skeletons attached to a given $\bar{e}_s$.

We define the partial weight $P_{\bar{e}_s}(G)$ relative to a given $\bar{e}_s$ to be the contribution to the weight of G, due to this $\bar{e}_s$; clearly the weight of G equals the sum of all partial weights.

We now proposed to determine $P_{\bar{e}_s}(G)$.

From (2.4), it follows that

$$P_{\bar{e}_s}(G) = N_{\bar{e}} \frac{1}{g_{\bar{e}}} \frac{1}{g_{v[\bar{g}]}} (-1)^e, \tag{3.1}$$

where $N_{\bar{e}}$ = number of distinct $\bar{e}$-skeletons attached to the given $\bar{e}_s$; $g_{\bar{e}}$ = order of the symmetry group of $\bar{e}_s$; and $g_{v[\bar{g}]}$ = order of the symmetry group of the graph, representing the functional which achieves the insertion operation defined previously. We shall determine successively each of these quantities.

The following two situations may arise:

(1) $\bar{e}_s$ is connected.

(2) $\bar{e}_s$ is not connected.

We first consider (1).

(i) Suppose that each element of $F_G^{-1}(\bar{e}_s)$ contains l_i lines, picked out from a set of k_i lines on which are defined g_i symmetry operations of G;

$$1 \le l_i \le k_i, \quad 1 \le i \le m, \quad \sum_i l_i = e.$$

We clearly see here that for each element in $F_G^{-1}(\bar{e}_s)$ the numbers l_i, k_i, g_i are the same.
It is clear that

$$N_{\bar{e}} = \binom{e}{l_1}\binom{e-l_1}{l_2}\binom{e-l_1-l_1}{l_3}\cdots\binom{e-l_1-l_2-\cdots-l_{m-1}}{l_m} = \frac{e!}{\prod_i l_i!}; \tag{3.2}$$

$$\text{(ii)}\quad g_{\bar{e}} = \frac{g}{[k_1(k_1-1)\cdots(k_1-l_1+1)]\cdots[k_m(k_m-1)\cdots(k_m-l_m+1)]}. \tag{3.3}$$

(iii) $\bar{e}_s$ being connected, the functional $g_{v[\bar{g}]}$ is represented by the graph of Fig. 5.

Hence

$$g_{v[\bar{g}]} = e!. \tag{3.4}$$

Inserting (3.2), (3.3), and (3.4) in (3.1), we obtain

$$P_{\bar{e}_s}(G) = \frac{e!}{\prod_i l_i!}\frac{[k_1(k_1-1)\cdots(k_1-l_1+1)]\cdots[k_m(k_m-1)\cdots(k_m-l_m+1)]}{g}\frac{1}{e!}(-1)^e$$
$$= \frac{1}{g}(-1)^e\left(\frac{k_1!}{(k_1-l_1)!}\frac{1}{l_1!}\right)\left(\frac{k_2!}{(k_2-l_2)!}\frac{1}{l_2!}\right)\cdots\left(\frac{k_m!}{(k_m-l_m)!}\frac{1}{l_m!}\right) = \frac{(-1)^e}{g}\prod_i\binom{k_i}{l_i}. \tag{3.5}$$

Since $\prod_i\binom{k_i}{l_i}$ is the number of elements of $F_G^{-1}(\bar{e}_s)$, we attribute to each one of these elements the value $(-1)^e$, and the partial weight in question is the sum of these values:

$$P_{\bar{e}_s}(G) = \sum_{\bar{e}\in F_G^{-1}(\bar{e}_s)}(-1)^e. \tag{3.6}$$

We now consider Case (2). Let r be the number of connected components of $\bar{e}_s$; the latter are $\bar{m}_i$-supports of r distinct $2p$-parts of G, joined in G by c lines; hence (Fig. 6)

$$\sum_{i=1}^{r} m_i = e - c.$$

(i) The number $N_{\bar{e}}$ is clearly the product of the $N_{\bar{m}_i}$. Denote by $l_j^i, k_j^i, 1 \le j \le n_i, 1 \le i \le r$, the analogs of l_i, k_i in Case 1. Then

$$N_{\bar{m}_i} = m_i!\Big/\prod_{j=1}^{n_i} l_j^i!,$$

so that

$$N_{\bar{e}} = \prod_{i=1}^{r} N_{\bar{m}_i} = \prod_{i=1}^{r}\frac{m_i!}{\prod_{j=1}^{n_i} l_j^i!}. \tag{3.7}$$

(ii) Clearly,

$$g_{\bar{e}} = \frac{1}{K}\frac{g}{\prod_{i=1}^{r}[k_1^i(k_1^i-1)\cdots(k_1^i-l_1^i+1)]\cdots[k_{n_i}^i(k_{n_i}^i-1)\cdots(k_{n_i}^i-l_{n_i}^i+1)]}. \tag{3.8}$$

If the c lines, joining the r connected components, define K symmetry operation of G:

(iii) $\bar{e}_s$ being nonconnected, the functional $g_{v[\bar{g}]}$ is represented by the graph in Fig. 7 and

$$g_{v[\bar{g}]} = K\prod_{i=1}^{r} m_i!. \tag{3.9}$$

Inserting (3.9), (3.8), and (3.7) into (3.1), we obtain

$$P_{\bar{e}_s}(G) = \prod_{i=1}^{r} \frac{m_i!}{n_i \prod_{j=1} l_j^i!} K \frac{[k_1^i(k_1^i-1)\cdots(k_1^i-l_1^i+1)]\cdots[k_{n_i}^i\cdots(k_{n_i}^i-l_{n_i}^i+1)]}{g} \frac{1}{K}\frac{1}{\prod_i m_i!}(-1)^e$$

$$= \frac{(-1)^e}{g}\prod_{i=1}^{r}\frac{[k_1^i(k_1^i-1)\cdots(k_1^i-l_1^i+1)}{l_1^i!}\cdots\frac{[k_{n_i}^i\cdots(k_{n_i}^i-l_{n_i}^i+1)]}{l_{n_i}^i!} = \frac{(-1)^e}{g}\prod_{i=1}^{r}\prod_{j=1}^{n_i}\binom{k_j^i}{l_j^i}. \quad (3.10)$$

$$\prod_{i=1}^{r}\prod_{j=1}^{n_i}\binom{k_j^i}{l_j^i}$$

being the number of elements contained in $F_G^{-1}(\bar{e}_s)$, the result is identical to the analogous case in Case 1.

Since the total weight $P(G)$ is the sum of all partial weights, we obtain the basic relation

$$P(G) = \sum_{e\in\bar{E}_G}(-1)^e. \quad (3.11)$$

Therefore, the weight of G is entirely determined by the structure of $\bar{E}_G$.

The basic formula (3.11) leads to the following.

Theorem: If G_N, $N \neq 0$, is such that $\bar{E}_G \cong E_G$, then $P(G) = 0$.

Indeed, the number of elements of E_G being $\sum_{e=0}^{N}\binom{N}{e}$, it follows from (3.11) that

$$P(G) = \sum_{e=0}^{N}\binom{N}{e}(-1)^e = 0.$$

The case $N = 0$ corresponds to the diagram constructed with only one Γ_{2n}; in this case, $E_{\Gamma_{2n}}$ contains just one element viz. ϕ, so that $e = 0$; trivially,

$$E_{\Gamma_{2n}} \cong \bar{E}_{\Gamma_{2n}}$$

and

$$P(\Gamma_{2n}) = (-1)^0 = 1.$$

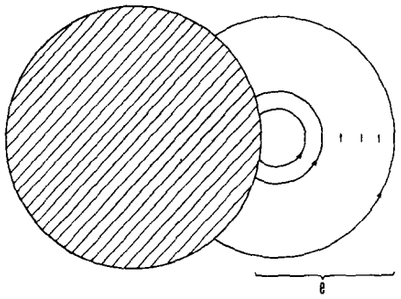


FIG. 5. The graph representing the functional which achieves the insertion operation (cf. Sec. 3). The hatched circle stands for the connected $\bar{e}$-support.

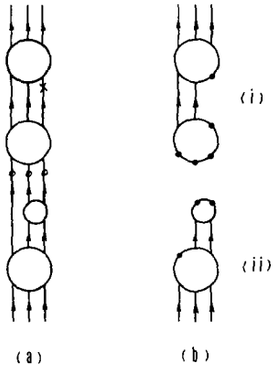


FIG. 6. (a) the graph G; (b) A nonconnected $\bar{4}$-support. The c = three lines disconnecting the graph G (by amputation) are marked by ○.

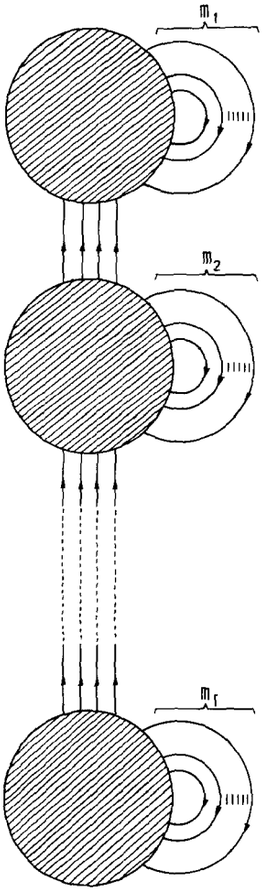


FIG. 7. Insertion operation (cf. Sec. 3) for a nonconnected e-support. The hatched circles stand for the connected components of the $\bar{e}$-support. The total number is vertical lines is c.

4. THE WEIGHT OF VARIOUS FAMILIES OF DIAGRAMS

A. The Weight of a Diagram Belonging to $\bar{g}$

Theorem 4.1: If $G \in \bar{g}$, then

$$P(G) = 0. \quad (4.1)$$

To prove this theorem we need the following lemmas:

Lemma 1: If $G_{2n} \in \bar{g}_{2n}$, each $2p$-part (even if $p \geq n$) is contained in $\bar{g}_{2p}$.

Indeed, let P_{2n} be a $2n$-part of G:

(i) P_{2n} contains no Γ_{2m}, $m > n$, and at most one Γ_{2n};

(ii) Each $2l$-part of P_{2n}, $l < n$, being a $2l$-part of G, is contained in $\bar{g}_{2l}$; hence

$$P_{2n} \in \bar{g}_{2n}. \quad (a)$$

Consider now a $2(n+1)$ part $P_{2(n+1)}$ of G:

(i') $P_{2(n+1)}$ contains no Γ_{2m}, $m \geq n+1$.

(ii') Every $2l$ part of $P_{2(n+1)}$, $l < n+1$ is contained in $\bar{g}_{2l}$ by (a). Hence

$$P_{2(n+1)} \in \bar{g}_{2(n+1)}$$

and so on by induction.

We note that this property is independent of modΓ_{2n}.

Lemma 2: If $G_{2n} \in \bar{g}_{2n}$, then $\overline{E}_G \equiv E_G$. Consider a 1-support of G, which is the image under the mapping F_G, of the element $e \in E_G$ containing a single internal line l of G.

This line defines two types of subdiagrams of $G: H^a$ and H^b, which, respectively, do or do not contain l as an internal line.

For an arbitrary subdiagram H of G, let $F_G(H)$ be the part of the 1-support obtained from H by amputating the line l in G.

$F_G(H)$ is either a subdiagram of the 1-support, or a pair of subdiagrams (according to whether $F_G(H)$ is connected or not).

Clearly, $F_G(H^b_{2m}) \equiv H^b_{2m}$ so that by Lemma 1, we have

$$F_G(H^b_{2m}) \in \bar{g}_{2m}, \quad \forall\ m.$$

Consider further the set $\cap_a H^a$.

Each subdiagram of $\cap_a H^a$ is, by construction, and H^b subdiagram of G. Hence, each subdiagram S_{2m} of $F_G(\cap_a H^a)$ belongs to $\bar{g}_{2m}$, $\forall\ m$

Thus

$$F_G(\cap_a H^a) \in \bar{g}.$$

For an arbitrary H^a, $F_G(H^a)$ contains one or more subdiagrams containing $F_G(\cap_a H^a)$ and hence is contained in $\bar{g}$; it follows that the 1-support of G, associated to the line l, is contained in $\bar{g}$.

Each e-support of G being a 1-support of an $(e-1)$-support of G, the above property is valid by induction for all e_s. This completes the proof of Lemma 2.

Theorem (4.1) is an immediate corollary of the two lemmas.

B. The Weight of an Irreducible Diagram I_{2n}

Our first step will be to show that Lemma 1 also applies to I_{2n}, modΓ_{2n}.

To achieve this, we need the following:

Lemma 3: Given an I_{2n}, each $2p$-part (even when $p > n$) is contained in $\bar{g}_{2p}$ (modΓ_{2n}).

Indeed, let $P_{2(n+1)}$ be a $2(n+1)$-part of I_{2n}. Then:

(i) $P_{2(n+1)}$ contains no Γ_{2m}, $m \geq n+1$ by modΓ_{2n}.

(ii) Every $2l$-part of $P_{2(n+1)}$, $l < n+1$, being a $2l$-part of I_{2n}, is contained in $\bar{g}_{2l}$; hence

$$P_{2(n+1)} \in \bar{g}_{2(n+1)}. \qquad \text{(a)}$$

Consider now a $2(n+2)$-part $P_{2(n+2)}$ of I_{2n}:

(i') $P_{2(n+2)}$ contains no Γ_{2m}, $m \geq n+2$, by (modΓ_{2n}).

(ii') Every $2l$-part of $P_{2(n+2)}$, $l < n+2$ is contained in $\bar{g}_{2l}$ by (a) and so on by induction.

Therefore the structure of subdiagrams of I and $\bar{g}$ is identical (mod Γ_{2n}). By a reasoning already used in Lemma 2, we find that each $2p$-part (p arbitrary) of an e_s of I_{2n}, is contained in $\bar{g}_{2p}$; since moreover, $\dim e_s = 2(n+e)$, $e_s \in \bar{g}_{2(n+e)}$ (modΓ_{2n}) and $e > 0$; (if e_s is not connected, each one of its connected components is contained in $\bar{g}$).

Thus,

Lemma 4: Each e_s, $e > 0$, of an irreducible graph I_{2n}, is an $\bar{e}_s$ (modΓ_{2n}).

Only $F_G(\emptyset) \notin \bar{g}_{2n}$ and therefore

$$\overline{E}_I \equiv \underset{e \neq 0}{E_I}.$$

From (3.11),

$$P(I_{2n}) = \sum_{e \in \overline{E}_I} (-1)^e = \sum_{e=1}^{N} \binom{N}{e} (-1)^e = -1. \qquad (4.2)$$

C. Reducible Graphs

An important class of reducible diagrams consists of what we shall call the coverings R^N.

Definition: We shall say, that a graph R^N is a covering, if it is made up of a family $\{I^i\}$, $1 \leq i \leq N$, of irreducible graphs, such that $\forall\ i\ \exists\ j$, with $I^i \cap I^j \neq \phi$.

The I^i are called irreducible components of R^N (Fig. 8).

An N convolution C^N is a particular R^N subject to the following restriction:

Each I^i intersects exactly two others, except for two of them, called endpoints, each one of which intersects exactly one I^i.

A linear N convolution L^N is an N convolution satisfying: $\dim I^i = \dim I^j$, $\forall\ i, j$.

Our next task is to determine the weight of an L^N. Consider a particular L^N subjected to the following condition: $\cap(I^i, I^{i+1}) \equiv \Gamma_{2n}$ for each i. Let I be an irreducible endpoint component of L^N.

Let $\{l_i\}$, $1 \leq i \leq L$ be the L internal lines of I. Consider the two sets $\overline{E}_{L^N}$ and $\overline{E}_{L^N-1}$; to each element $\bar{e} \in \overline{E}_{L^N-1}$, we associate a subset $\mathcal{S}_{\bar{e}} \subset E_{L^N}$ consisting of:

(a) the L elements $(\bar{e}, l_i)$, $\quad 1 \leq i \leq L$;

(b) The $\binom{L}{2}$ elements $(\bar{e}, l_i, l_j)$, $\quad i \neq j$, $1 \leq i, j \leq L$;

⋮

( ) The element $(\bar{e}, l_1, l_2, \ldots, l_L)$.

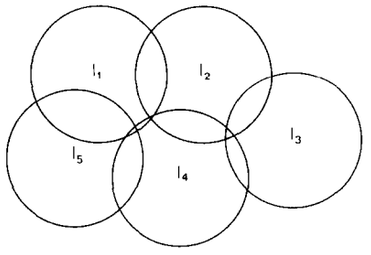


FIG. 8. A possible R^5; the circles stand for irreducible components of R^5.

By Lemma 4, $\mathcal{E}_{\bar{e}} \subset \bar{E}_{L^N}$; furthermore, $\overline{E}_{L^N} \equiv \bigcup_{\bar{e} \in \overline{E}_{L^{N-1}}} \mathcal{E}_{\bar{e}}$ so that, by (3.11)

$$\begin{aligned} P(L^N) &= \sum_{e \in \overline{E}_{L^{N-1}}} \left[L(-1)^{e+1} + \binom{L}{2}(-1)^{e+2} \right. \\ &\quad \left. + \cdots + \binom{L}{L}(-1)^{e+L} \right] \\ &= \sum_{e \in \overline{E}_{L^{N-1}}} (-1)^e \left[L(-1) + \binom{L}{2}(-1)^2 \right. \\ &\quad \left. + \cdots + \binom{L}{L}(-1)^L \right] \\ &= -P(L^{N-1}). \end{aligned}$$

On the other hand, we have shown before that $P(I) = -1$; the solution of the above recurrence relation is therefore

$$P(L^N) = (-1)^N. \tag{4.3}$$

Let us now consider an arbitrary reducible G^N and a set of $0 \leq K < N$ internal lines of G; to this set we associate the set E_K in the manner that E_G is associated with the N internal lines of G.

Let us introduce a subset $E_G^* \subset E_K$, satisfying the following conditions:

To each element $e^* \in E_G^*$, we associate a subset $\mathcal{E}_{e^*} \subset E_G$ containing the following elements:

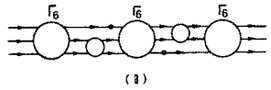
(a) the element e^*;

(b) the M elements (e^*, l_i), $\quad 1 \leq i \leq M$;

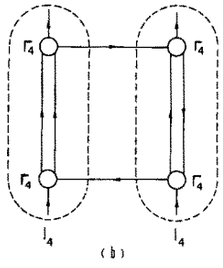
⋮

( ) the element $(e^*, l_1, l_2, \ldots, l_M)$;

where l_i are the $M = N - K$ internal lines of G, not contained in the above mentioned set of K lines.



(a)



(b)

FIG. 9. (a) Graph for which E_G^* does not exist; the element containing the lines marked 0 satisfies (i) but not (ii) in the definition of E_G^* (cf. Sec. 4). (b) Graph for which E_G^* exists.

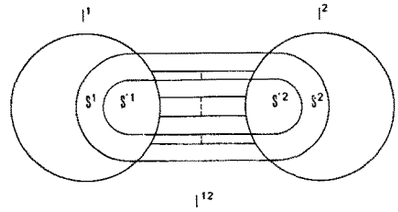


FIG. 10. A diagram of type I^{12}; I^1, I^2 are the two irreducible graphs; S^1, S'^1 and S^2, S'^2 are subdiagrams of I^1 and I^2, respectively. Figure 9(b) is a particular case of this general configuration.

The sets $\mathcal{E}_{e^*}$ are restricted by the following two conditions:

(i) $\quad \mathcal{E}_{e^*} \subset \overline{E}_G, \quad \forall\, e^* \in E_G^*$;

(ii) $\quad \bigcup_{e^* \in E_G^*} \mathcal{E}_{e^*} \equiv \overline{E}_G$.

For a given diagram G, the set E_G^* may or may not exist, according to the structure of G (Fig. 9) (If it exists it is in general not unique).

Theorem 4.2: If the reducible graph G is such that a set E_G^* exists, then $P(G) = 0$.

Indeed, it follows from the definition of E_G^* and the basic relation (3.11), that

$$\begin{aligned} P(G) &= \sum_{e^* \in E_G^*} \left[(-1)^{e^*} + M(-1)^{e^*+1} + \binom{M}{2}(-1)^{e^*+2} \right. \\ &\quad \left. + \cdots + \binom{M}{M}(-1)^{e^*+M} \right] \\ &= \sum_{e^* \in E_G^*} (-1)^{e^*} \left[\sum_{m=0}^{M} \binom{M}{m}(-1)^m \right] = 0. \end{aligned}$$

In the case $K = 0$, the existence of E_G^* implies by construction $\mathcal{E}_\phi = E_G = E_G$, so that we recover a previously obtained result.

We shall now apply this result to particular reducible graphs.

Consider first a reducible diagram, obtained by dressing an irreducible graph either (a) by a second irreducible diagram, or (b) by a member of $\bar{g}$.

(a) The general structure of such a graph (denoted I^{12}) is shown in Fig. 10, and a particular example is given Fig. 9(b).

The two irreducible graphs I^1 and I^2 have no common subdiagram, since otherwise they would form a convolution whose weight has already been calculated.

We shall first demonstrate that every subdiagram of I^{12} of type S^{12} (Fig. 10) is necessarily contained in $\bar{g}$; indeed, suppose this were not the case: two cases may then arise:

(i) All subdiagrams of S^{12} (which are either entirely contained in I^1, or in I^2, or are of type S'^{12}) are contained in $\bar{g}$; in this case, S^{12} is by definition a irreducible diagram and I^{12} would be a convolution.

(ii) There exists at least one subdiagram of S^{12}, not contained in $\bar{g}$, which by Lemma 3 is necessarily of type S'^{12}, $S'^1 \subset S^1$, $S'^2 \subset S^2$; S'^{12} must in turn necessarily contain a subdiagram of type S''^{12}, since otherwise I^{12} would be a convolution, and so on;

This procedure must clearly come to an end; indeed, from a certain stage onwards, the subdiagrams are entirely contained either in I^1 or in I^2, so that, by lemma 3, they are members of $\bar{g}$.

Thus the existence of a subdiagram S^{12}, not contained in $\bar{g}$, is incompatible with the fact that I^{12} is not a convolution. It follows that $E_{I^{12}}^*$ exists and

$$E_{I^{12}}^* \equiv (\overline{E}_{I^1}, \overline{E}_{I^2}),$$

where $(\overline{E}_{I^1}, \overline{E}_{I^2})$ stands for the totality of couples $(\bar{e}_1, \bar{e}_2)$, $\bar{e}_1 \in \overline{E}_{I^1}$, $\bar{e}_2 \in \overline{E}_{I^2}$.

Hence by Theorem 4.2,

$$P(I^{12}) = 0.$$

This property can be immediately extended to a more general reducible diagram $I^{12\cdots m}$ shown in Fig. 11; indeed,

$$E^*_{I^{12\cdots m}} \equiv (\bar{E}_{I^1}, \bar{E}_{I^2}, \ldots, \bar{E}_{I^m}),$$

hence

$$P(I^{12\cdots m}) = 0. \tag{4.4}$$

(b) The general structure of such a graph denoted $I\bar{g}$ is indicated in Fig. 11(b) and a particular example is given Fig. 11(c). Clearly,

$$E^*_{I\bar{g}} \equiv \bar{E}_I$$

so that

$$P(I\bar{g}) = 0. \tag{4.5}$$

More generally, for every reducible graph G, obtained by dressing a covering R^N by a diagram contained in $\bar{g}$, or by juxtaposition of m coverings, there exists a set E^*_G: In the first case

$$E^*_G \equiv \bar{E}_{R^N},$$

while in the second

$$E^*_G \equiv (\bar{E}_{R^{N_1}}, \bar{E}_{R^{N_2}}, \ldots, \bar{E}_{R^{N_m}}).$$

The weight of such diagrams is therefore zero.

Thus the only reducible diagrams of nonzero weight are the coverings R^N, but we shall show that only a very restricted subclass thereof are topologically permissible and contribute to the n-body potential v_{2n}.

This is the content of Theorem 5.1 in the following section.

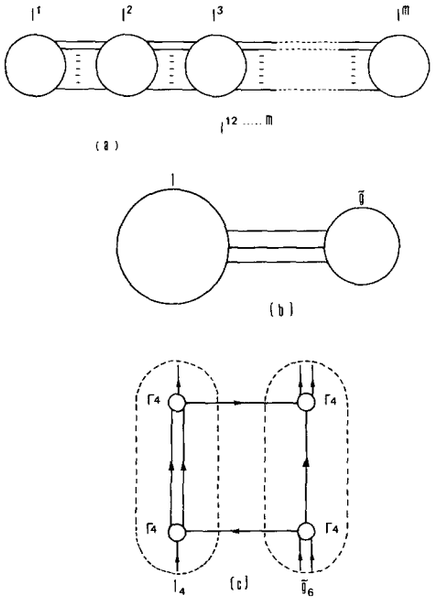


FIG. 11. (a) Diagram of type $I^{12\cdots m}; I^1, I^2, \ldots, I^m$ are the m irreducible graphs. (b) A diagram illustrating the dressing of an irreducible graph I by a member of $\bar{g}$. (c) A particular case of the general configuration 11(b).

5. TOPOLOGICAL ANALYSIS OF COVERINGS R^N.

Theorem 5.1: The only topologically permissible coverings R^N of dimension $2n$, are the linear convolutions L^N (modΓ_{2n}).

Denote by M and M_1 an arbitrary irreducible component of maximal dimension in R^N and $C_{R^N}M$, respectively.

Clearly $\dim M \geq \dim M_1$.

Lemma 5: For any covering R^N, we have

$$\dim R^N \leq \dim M,$$

the equality may hold only if

(i) $\dim I^i = \dim I^j, \quad 1 \leq i, j \leq N,$

(ii) there is no subdiagram common to more than two I^i,

(iii) there is no contribution to $\dim R^N$, originating from intersections of I^i,

(iv) all intersections of I^i and their relative complements are connected,

(v) no internal line of R^N is simultaneously external to more than one I^i,

(vi) there is no couple (M, M_1) such that $C_{R^N}(M, M_1)$ is connected and contributes to $\dim R^N$.

We show (Appendix A), that the lemma is verified for $N = 2$, and we proceed by induction to extend it for all N.

The induction hypothesis (H_i) is the following: Lemma 5 holds for each R^L, $L < N$.

The following two cases may arise.

Case 1: $C_{R^N}(M, M_1)$ is connected (Fig. 12).

We have

$$\dim R^N = m_c + m_1 + m$$
$$\dim \cap (M, C) = \mu + \nu,$$
$$\dim \cap (M_1, C) = \mu_1 + \nu_1,$$
$$\dim M_1 = m_1 + \nu_1,$$
$$\dim M = m + \nu.$$

Suppose that $\exists (i, j)$ such that

$$\dim I^i \neq \dim I^j. \tag{5.1a}$$

We shall show that the condition $\dim R^N \geq \dim M$ leads to a contradiction.

Indeed, if $\dim R^N \geqslant \dim M$, and *a fortiori* $\dim R^N \geqslant \dim M_1$, then

$$m_c + m_1 \geq \nu, \tag{5.1b}$$

$$m_c + m \geq \nu_1. \tag{5.2}$$

On the other hand, H_i applied to the coverings $M_1 \cup C_{R^N}(M, M_1)$ and $M \cup C_{R^N}(M, M_1)$ leads to

$$m_1 + m_c + \mu \leq m_1 + \nu_1,$$
$$m + m_c + \mu_1 \leq m + \nu,$$

so that

$$m_c + \mu \leq \nu_1, \tag{5.3}$$

$$m_c + \mu_1 \leq \nu. \tag{5.4}$$

The equalities in (5.3) and (5.4) are mutually exclusive; indeed, by (5.1a), the two coverings to which we apply H_i cannot simultaneously satisfy condition (i)

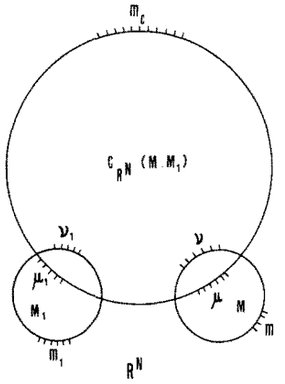


FIG. 12. A typical configuration illustrating Case 1, Sec. 5; m, m_1, m_c are contributions to $\dim R^N$ due to $M, M_1, C_{R^N}(M,M_1)$, respectively; μ and μ_1 are the number of lines internal to M, M_1, respectively, and external to $C_{R^N}(M,M_1)$; ν and ν_1 are the number of lines external to M and M_1, respectively, and internal to $C_{R^N}(M,M_1)$.

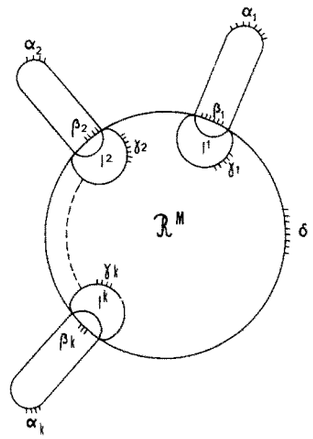


FIG. 13. A graph illustrating $\mathcal{R}^M; I^1, I^2, \ldots, I^k$ are the boundary components of $\mathcal{R}^M; \beta_1, \beta_2, \ldots, \beta_k$ are contributions to $\dim \mathcal{R}^M$ but not to $\dim R^N$; δ is the number of external lines of $\mathcal{R}^M$ contributing to $\dim R^N$.

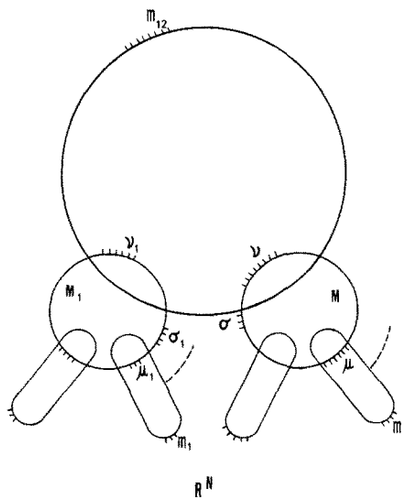


FIG. 14. A typical configuration illustrating Case 2, Sec. 5; m, m_1, m_{12} are contributions to $\dim R^N$ due to these connected components of $C_{R^N}(M,M_1)$ which intersect only M, M_1, and both, respectively; the remaining symbols have the same meaning as in Fig. 12.

of Lemma 5. Hence,

$$\mu + \mu_1 < \nu + \nu_1. \tag{5.5}$$

Combining, respectively, (5.1)–(5.4) it follows that

$$\mu \leq m, \tag{5.6}$$

$$\mu_1 \leq m_1, \tag{5.7}$$

which implies

$$\mu + \nu \leq M, \tag{5.8}$$

$$\mu_1 + \nu_1 \leq M_1, \tag{5.9}$$

The equalities in (5.8), (5.9) being mutually exclusive.

Since now $\cap (M, C)$ (Fig. 12) is a $(\mu + \nu)$–part of M, it must belong to $\bar{g}_{\mu+\nu}$ and can contain

(i) at most one Γ_M and no Γ_L, $L > M$, whenever (5.8) is an equality,

(ii) no Γ_L, $L \geq M$, whenever (5.8) is an inequality.

The same analysis applies to $\cap (M_1, C)$. From the mutually exclusive character of the equalities in (5.8) and (5.9), it follows that $\cap (M, C)$ and $\cap (M_1, C)$ cannot simultaneously contain a Γ_M and a Γ_{M_1}, respectively;

the above situation is therefore compatible with the irreducibility of M and M_1, only if

$$\mu + m \geq M, \tag{5.10}$$

$$\mu_1 + m_1 \geq M_1, \tag{5.11}$$

which implies

$$\mu \geq \nu, \tag{5.12}$$

$$\mu_1 \geq \nu_1. \tag{5.13}$$

From the previous discussion, it follows that the equalities in (5.12), (5.13) are, as usual, mutually exclusive, so that

$$\mu + \mu_1 > \nu + \nu_1. \tag{5.14}$$

The desired property follows from (5.14) and (5.5).

Case 2: $C_{R^N}(M, M_1)$ is not connected ($N > 3$) (Fig. 14).

We have

$$\dim R^N = m + m_1 + m_{12} + \sigma + \sigma_1,$$
$$\dim M = \mu + \nu + \sigma,$$
$$\dim M_1 = \mu_1 + \nu_1 + \sigma_1.$$

Suppose that condition (5.1a) holds. By repeated application of H_i, we obtain

$$\sigma + \sigma_1 + \mu + \mu_1 + m_{12} \leq M, \tag{5.15}$$

$$\sigma + \nu + m \leq \mu + \nu + \sigma, \tag{5.16}$$

$$\sigma_1 + \nu_1 + m_1 \leq \mu_1 + \nu_1 + \sigma_1. \tag{5.17}$$

The equalities in (5.15)–(5.17) are mutually exclusive. Indeed, by (5.1a) the three coverings to which we apply H_i cannot simultaneously satisfy (i) of Lemma 5. Summing (5.15)–(5.17), we have therefore

$$\dim R^N < M.$$

This completes the proof of Lemma 5, in the case where condition (5.1a) is satisfied.

If $\dim I^i = \dim I^j$ for all i, j between 1 and N, then since we can arbitrarily pick out M and M_1, we can follow the procedure of either Case 1 or Case 2, depending on our choice. It is easy to see that the procedure of Case 2 leads to a condition less restrictive than does Case 1; we shall therefore choose M and M_1 such that $C_{R^N}(M, M_1)$ is connected.

We intend to show that

$$\dim R^N = M$$

leads to a contradiction, only if

$$m_c \neq 0.$$

The proof follows the same pattern as in Case 1, except that (5.1) and (5.2) are replaced by equalities, and (5.3) and (5.4), and (5.6) and (5.7) are not mutually exclusive.

By suitably combining these relations, we obtain

$$\mu + \mu_1 \geq \nu + \nu_1, \tag{5.18}$$

$$\mu + \mu_1 + 2m_c \leq \nu + \nu_1. \tag{5.19}$$

Equations (5.18) and (5.19) are contradictory only if $m_c \neq 0$; therefore, $\dim R^N = M$ can only be realized if $m_c = 0$

The condition $\dim R^N > M$ leads of course to a contradiction in all cases.

This completes the proof of (i) and (vi) of Lemma 5.

Conditions (ii)–(iv) have their origin in the corresponding properties of a covering R^2 (Appendix A).

To prove (v), it suffices to remark that a covering R^N containing l such lines, is strictly equivalent, as regards our analysis, to a topologically identical covering, but of dimension $\dim R^N + 2l$.

Lemma 6: there exists no irreducible graphs I such that $\dim I > 2n, (\mathrm{mod}\Gamma_{2n})$. This lemma is an immediate consequence of the definition of an irreducible diagram.

Lemma 7: Let R^N be a covering of dimension $2n$, and I^i its irreducible components. Then $\dim I^i = \dim I^j = 2n$ for each i, j $(\mathrm{mod}\Gamma_{2n})$; in addition, conditions (i)–(vi) of Lemma 5 are fulfilled.

This lemma is a corollary of Lemmas 5 and 6; It expresses a very restrictive condition on the structure of a covering R^N of $\dim 2n (\mathrm{mod}\Gamma_{2n})$. We shall show that, under this last condition, R^N in fact satisfies a stronger condition.

To show the latter, we define the partial convering $\mathfrak{R}^M$ of an R^N to be a covering consisting of M irreducible components of R^N, $M < N$. We call the boundary components of $\mathfrak{R}^M$, the $I^i \in R^N$ containing lines external to $\mathfrak{R}^M$(internal or external to R^N).

By lemmas 7 and 5, if $\dim R^N = 2n$, then

$$\dim \mathfrak{R}^M \leq 2n. \tag{5.20}$$

We shall prove that (5.20) is always saturated.

Lemma 8: Each partial covering $\mathfrak{R}^M$ of a $2n$-dimensional R^N is of dimension $2n (\mathrm{mod}\Gamma_{2n})$. We shall first prove (Appendix B) that the only permissible configuration of an $\mathfrak{R}^M$, is the one represented on Fig. 13.

Let $I^1, I^2, \ldots, I^k$ be the boundary components of $\mathfrak{R}^M$. By hypothesis

$$\dim R^N = 2n = \sum_{i=1}^{k} \alpha_i + \delta. \tag{5.21}$$

On the other hand, by Lemma 7,

$$\gamma_i + \beta_i = 2n, \quad 1 \leq i \leq k. \tag{5.22}$$

Summing (5.21) and (5.22),

$$\sum_{i=1}^{k} (\alpha_i + \gamma_i) + \sum_{i=1}^{k} \beta_i + \delta = (k+1)2n. \tag{5.23}$$

By Lemma 5,

$$\alpha_i + \gamma_i \leq 2n, \quad 1 \leq i \leq k, \quad \sum_{i=1}^{k} \beta_i + \delta \leq 2n. \tag{5.24}$$

Equations (5.23) and (5.24) can be compatible, only if (5.24) are saturated.

Hence

$$\dim \mathfrak{R}^M = 2n.$$

Lemma 9: Let R^N be a covering, and I^k, I^l two arbitrary irreducible components. There exists (in R^N) at least one convolution whose endpoints are I^k and I^l. This property follows from the connectedness of R^N.

Define $\subset (I^k)$ to be the set of all endpoints of convolutions, whose second endpoint is I^k. Our aim is to prove that $\subset (I^k) \equiv R^N$. In the first instance $\subset (I^k) \neq \emptyset$; indeed, $\exists\, j$ such that $I^k \cap I^j \neq \emptyset$ by definition of R^N.

If $\subset (I^k)$ contained only I^k and I^j, R^N would be nonconnected. Hence, $\subset (I^k)$ contains at least a third element I^m and so forth. Step by step, we shall thus encompass in $\subset (I^k)$ all irreducible components of R^N. QED

It follows from Lemmas 5 and 8, that any convolution in R^N contains no external lines (external or internal to R^N) other than the ones external to the endpoints.

A corollary of this last property and of Lemma 9 is that the whole covering R^N takes the form of a linear

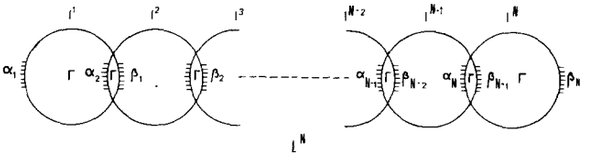


FIG. 15. A typical L^N; α_i, β_i refer, respectively, to the number of lines external to I^i and internal to I^{i-1} and I^{i+1}; only α_1, β_N contribute to $\dim L^N$.

convolution. This completes the proof of Theorem 5.1: The only topologically permissible $2n$-dimensional coverings R^N are the linear convolutions L^N (modΓ_{2n}). It follows from the preceding discussion that only the two endpoints of L^N contribute to dimL^N. In addition the L^N have the interesting property of conserving the number of lines. Indeed (Fig. 15), by Lemmas 8 and 5,

$$\alpha_i + \beta_i = 2n, \quad 1 \le i \le N-1,$$
$$\dim L^N = \alpha_1 + \beta_N = 2n.$$

Thus

$$\beta_i = \beta_j, \quad 1 \le i,j \le N.$$

We also have

$$\alpha_i + \beta_N = 2n, \quad 1 \le i \le N,$$

which leads to

$$\alpha_i = \alpha_j, \quad 1 \le i,j \le N.$$

On the other hand, by the property

$$\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = n$$

of L^2 (Appendix A), it follows that

$$\alpha_i = \beta_i = n, \quad 1 \le i \le N.$$

From this property and from the structure of an irreducible diagram, it follows that each irreducible component contains two Γ_{2n} located inside its respective intersections with its neighbors, except for the endpoints, which contain one Γ_{2n} in the intersection and one Γ_{2n} in its complement (Fig. 15).

In conclusion, we have demonstrated that only a very restricted set of Feynman diagrams survives in the complete expression for the n-body potential in terms of the Γ variables (modΓ_{2n}).

That is, collecting all previously obtained results:

$$V_{2n}(G_1;\Gamma_{2m}) = \Gamma_{2n} - I^{\Sigma}_{2n} + \sum_{N=2}^{\infty}(-1)^N L^{N\Sigma}_{2n},$$
$$\Gamma_{2m} = 0, \quad m > n.$$

Except for the single Γ_{2n}, the only survivors are the irreducible graphs, which appear either isolated or as building blocks of generalized ladder diagrams, i.e., our linear convolutions. The latter exhibit the property, that the number of lines entering at one endpoint is conserved from one element to the next, all the way to the second endpoint.

ACKNOWLEDGMENTS

I should like to express my gratitude to Professor F. Englert for directing my attention to this problem and for valuable discussions. I wish to thank Professor R. Brout for his continual interest and encouragement in this work, and Professor C. De Dominicis for a very fruitful discussion.

APPENDIX A

Let us consider the 2-convolution C^2 shown in Fig. 16:

$$C_{I^1}(\cap(I^1,I^2)) \equiv \bar{I}^1,$$
$$C_{I^2}(\cap(I^1,I^2)) \equiv \bar{I}^2,$$
$$\dim \cap(I^1,I^2) = \gamma + \delta + \rho = r,$$
$$\dim \bar{I}^1 = \alpha + \gamma + \mu = \bar{p},$$
$$\dim \bar{I}^2 = \delta + \beta + \mu = \bar{q},$$
$$\dim I^1 = \alpha + \delta + \rho + \mu = p,$$
$$\dim I^2 = \gamma + \beta + \rho + \mu = q,$$
$$\alpha + \beta = m.$$

Theorem A1: For any convolution C^2, we have

$$\dim C^2 \le \max[p,q] - 2\mu.$$

The equality holds only if the following conditions are satisfied:

(i) $\rho = 0$,

(ii) $p = q$,

(iii) $\gamma = \delta = \alpha + \mu = \beta + \mu$,

(iv) $\cap(I^1,I^2)$, $\bar{I}^1$, and $\bar{I}^2$ are connected.

We shall show that the contrary leads to incompatibility between the structures of the two irreducible components of C^2.

$$\textit{Case 1:} \quad \begin{cases} p - \rho - 2\mu < m \\ q - \rho - 2\mu < m \end{cases},$$

under these conditions

$$\begin{cases} \delta - \mu < \beta \\ \gamma - \mu < \alpha \end{cases}. \tag{A1}$$

The following two possibilities may arise:

(i) $\gamma \ne \delta$ say $\gamma < \delta$, which implies that

$$\bar{p} = \gamma + \alpha + \mu < \delta + \alpha + \mu \le p.$$

Since $\bar{I}^1$ is a $\bar{p}$-part of I^1, it is contained in $\tilde{g}_{\bar{p}}$ and can therefore contain no Γ_s, $s \ge p$; on the other hand, by (A1),

$$r = \gamma + \delta + \rho < \alpha + \mu + \delta + \rho = p.$$

$\cap(I^1,I^2)$, being an r-part of I^1, is contained in $\tilde{g}_r$ and can therefore contain no Γ_s, $s \ge p$; therefore neither $\bar{I}^1$ nor its complement, relative to I^1, viz. $\cap(I^1,I^2)$, can contain a Γ_s, $s \ge p$; this contradicts the irreducibility of I^1.

The situation $\gamma > \delta$ leads to the same result.

(ii) $\gamma = \delta$,

which implies

$$\bar{p} = \gamma + \alpha + \mu = \delta + \alpha + \mu \le p,$$

where the equality only holds if $\rho = 0$.

Therefore, if $\rho = 0$, $\bar{I}^1$ can contain at most a single Γ_p and no Γ_s, $s > p$, whereas if $\rho \ne 0$, $\bar{I}^1$ can contain no Γ_s, $s \ge p$; on the other hand, by (A1), $r < p$, and $\cap(I^1,I^2)$ can contain no Γ_s, $s \ge p$; once more we have

a contradiction to the definition of I^1. We must thus reject Case 1.

$$\textit{Case 2:} \quad \begin{cases} p-\rho-2\mu = m \\ q-\rho-2\mu < m \end{cases}.$$

This can be shown to be incompatible with the irreducibility of I^1 and I^2, by a similar analysis.

$$\textit{Case 3:} \quad \begin{cases} p-\rho-2\mu > m \\ q-\rho-2\mu < m \end{cases}.$$

$$\textit{Case 4:} \quad \begin{cases} p-\rho-2\mu > m \\ q-\rho-2\mu = m \end{cases}.$$

$$\textit{Case 5:} \quad \begin{cases} p-\rho-2\mu > m \\ q-\rho-2\mu > m \end{cases}.$$

Take, e.g., Case 4; we have

$$\begin{cases} \delta-\mu > \beta \\ \gamma-\mu = \alpha \end{cases}. \tag{A2}$$

The following situations may arise:

(i) $\gamma > \delta$, which implies

$$\bar{q} = \delta+\beta+\mu < \gamma+\beta+\mu \leq q$$

and so $\bar{I}^2$ can contain no Γ_s, $s \geq q$.

On the other hand, by (A2),

$$r = \gamma+\delta+\rho > \gamma+\beta+\mu+\rho = q$$

and $\cap(I^1, I^2)$ can contain more than one Γ_q or Γ_s, $s > q$, while simultaneously belonging to $\tilde{g}_r$; this condition is then compatible with the irreducible character of I^2. On the other hand, if $\gamma \geq \delta+\rho$, we have $\bar{p} \geq p$ and by (A2) $r = p$.

Thus, whenever $\gamma \geq \delta+\rho$, the situation is compatible with the irreducibility of I^1, as well, and so Case 4 is realizable.

Cases 3 and 5 can be analyzed in the same manner, and lead to analogous conclusions.

$$\textit{Case 6:} \quad \begin{cases} p-2\mu-\rho = m \\ q-2\mu-\rho = m \end{cases}.$$

Under these conditions,

$$\begin{cases} \delta-\mu = \beta \\ \gamma-\mu = \alpha \end{cases}, \tag{A3}$$
$$\gamma = \delta.$$

This implies $\bar{p} \leq p$, the equality being realized if $\rho = 0$. $\bar{I}^1$ can therefore contain at most one Γ_p and no Γ_s, $s > p$, if $\rho = 0$; in the event $\rho \neq 0$, $\bar{I}^1$ can contain no Γ_s, $s \geq p$. On the other hand, by (A3), $r = p$, so that $\cap(I^1, I^2)$ can contain at most one Γ_p. Consequently, this situation is compatible with the irreducibility of I^1, only if $\rho = 0$. Such is also the case for I^2, under the same condition.

Whatever ρ and μ may be, the situation $\gamma \neq \delta$ is not realizable.

It remains to show that the inequality is never saturated if $\cap(I^1, I^2)$ or $\bar{I}^1, \bar{I}^2$ are nonconnected (Fig. 17).

Consider the case where $\cap(I^1, I^2)$ is nonconnected and suppose, e.g., that it contains three connected pieces C_1, C_2, C_3:

$$\begin{aligned} \dim I^1 &= \alpha+\alpha_1+\alpha_2+\alpha_3 = p, \\ \dim I^2 &= \beta+\beta_1+\beta_2+\beta_3 = q, \\ \dim \bar{I}^1 &= \alpha+\beta_1+\beta_2+\beta_3 = \bar{p}, \\ \dim \bar{I}^2 &= \beta+\alpha_1+\alpha_2+\alpha_3 = \bar{q}, \\ \dim \cap(I^1, I^2) &= \alpha_1+\alpha_2+\alpha_3+\beta_1+\beta_2+\beta_3 = r, \\ \alpha+\beta &= m \\ \dim C_i &= \alpha_i+\beta_i, \quad 1 \leq i \leq 3. \end{aligned}$$

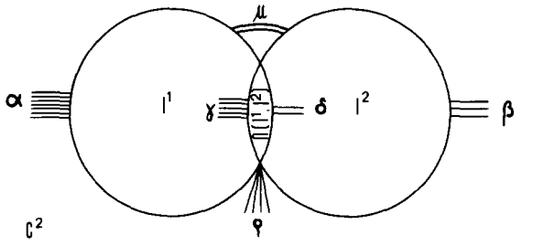


FIG. 16. A typical 2-convolution C^2; μ is the number of lines internal to C^2 while external to I^1 and I^2; ρ is the contribution to $\dim C^2$ originating from $\cap(I^1, I^2)$.

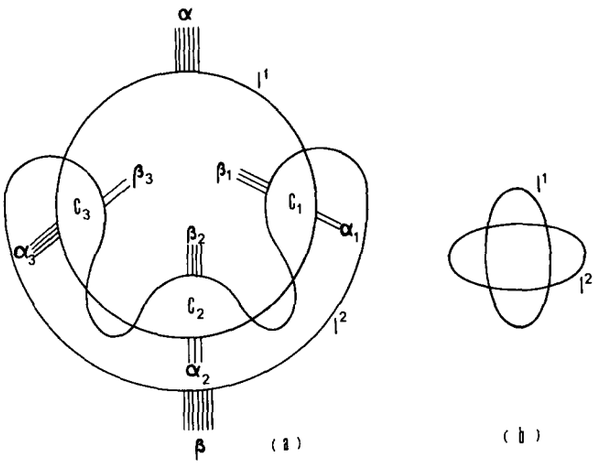


FIG. 17. (a) An R^2 with nonconnected $\cap(I^1, I^2)$ consisting of three pieces. (b) An R^2 with nonconnected $C_{I^1} \cap (I^1, I^2)$ and $C_{I^2} \cap (I^1, I^2)$.

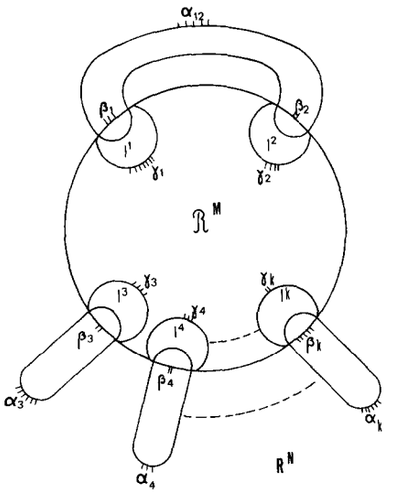


FIG. 18. (A configuration where two distinct boundary components (namely I^1, I^2) of $\mathfrak{R}^M$ intersect the same connected component of $C_{R^N}\mathfrak{R}^M$.

$$\textit{Case 1:} \quad \begin{cases} p < m \\ q < m \end{cases}.$$

Under these conditions,

$$\begin{cases} \alpha_1 + \alpha_2 + \alpha_3 < \beta \\ \beta_1 + \beta_2 + \beta_3 < \alpha \end{cases}, \qquad \text{(A4)}$$
$$\sum_i \beta_i < \sum_i \alpha_i,$$

which entails $\bar{p} < p$ and by (A4), $r < p$ and *a fortiori* $\dim C_i < p$.

This situation is therefore incompatible with the hypothesis of irreducibility of I^1; the situations $\sum \beta_i > \sum \alpha_i$ and $\sum \beta_i = \sum \alpha_i$ can be handled in the same way so that Case 1 is never realizable. The analyses of Cases 3–5 can be carried out in the same fashion and lead to the same conclusion; only the result of Case 6 is different

$$\textit{Case 6:} \quad \begin{cases} p = m \\ q = m \end{cases}.$$

In these conditions,

$$\begin{cases} \sum_i \alpha_i = \beta \\ \sum_i \beta_i = \alpha \end{cases}, \qquad \text{(A5)}$$
$$\sum_i \beta_i = \sum_i \alpha_i,$$

which entails $\bar{p} = p$ and, by (A5), $r = p$, so that

$$\dim C_i < p, \qquad i = 1, 2, 3,$$

and this case is not realizable.

Case (b) can be treated in precisely the same manner. This completes the proof of the theorem.

APPENDIX B

Let R^N be a $2n$-dimensional covering and $\mathcal{R}^M$ a partial covering of R^N. We shall prove the following lemma:

Lemma B1: Each connected component of $C_{R^N}\mathcal{R}^M$ can have intersections with no more than one boundary component of $\mathcal{R}^M$.

In fact, since the inequality of Lemma 5 is saturated, conditions (i)–(vi) hold and it is easy to convince oneself that the latter restrict drastically the number of possible configurations.

So as not to burden the reader with inessential details, we shall only give a brief sketch of the argument.

A possible configuration violating Lemma B1 is shown in Fig. 18.

By hypothesis,

$$\dim R^N = \alpha_{12} + \sum_{i=3}^{k} \alpha_i = 2n$$

and, by Lemma 7,

$$\gamma_i + \beta_i = 2n, \qquad 1 \le i \le k.$$

Summing these relations,

$$[\alpha_{12} + \gamma_1 + \gamma_2] + \sum_{i=3}^{k} [\alpha_i + \gamma_i] + \sum_{i=1}^{k} \beta_i = (k+1)2n.$$

Now, by Lemma 5, each of these k brackets are $\le 2n$. So that this configuration is not realizable.

By similar reasonings, we are led to Lemma B1.

1 F. Englert and C. De Dominicis, Nuovo Cimento 53A, 1007 (1968).
2 See Ref. 1, p. 1021.
3 F. Englert, R. Brout, and H. Stern, Nuovo Cimento 58A, 601 (1968).
4 R. Brout and F. Englert, Bull. Am. Phys. Soc. 11, 21 (1966).
5 R. Brout, *Topics in Bootstrap Theory, Cargèse Lectures* (Gordon and Breach, New York, 1966).
6 C. de Dominicis and P. C. Martin, J. Math. Phys. 5, 31 (1964).

Calculation of a Certain Type of 6*j*-Symbol

E. de Vries and A. J. van Zanten
Institute for Theoretical Physics, University of Groningen, The Netherlands

(Received 16 December 1971; Revised Manuscript Received 28 January 1972)

It is shown that, for an arbitrary finite simply reducible group, it is possible to express a certain type of 6*j*-symbol in which one of the representations is one-dimensional in terms of 3*j*-symbols $\{j\ k\ l\}$.

For an arbitrary finite simply reducible group, the 6*j*-symbol in which one of the irreducible representations is the trivial one (1_1) can be expressed in terms of 3*j*-symbols [see Ref. 1, Eq. (18b)]. We prove here the following similar result:

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_1 & j_2 & 1_i \end{Bmatrix} = (-1)^{2j_1 + j_3 \otimes (1_i)} \frac{1}{[j_1]} \{j_1\ j_2\ 1_i\}\{j_1\ j_2\ j_3\}. \quad (1)$$

In this equation $(-1)^{j_3 \otimes (1_i)}$ is the phase factor belonging to the irreducible representation, which is the Kronecker product of the irreducible representation (j_3) and the arbitrary one-dimensional representation (1_i). The 3*j*-symbol $\{j\ k\ l\}$ denotes the number of times that the irreducible representation (l) is contained in $(j) \otimes (k)$.

Proof: We start from Eq. (27a) of Ref. 1, taking $(j') = (j_3)$ and $(j) = (1_i)$, and substitute in the rhs $T = SR$. This gives

$$\textit{Case 1:} \quad \begin{cases} p < m \\ q < m \end{cases}.$$

Under these conditions,

$$\begin{cases} \alpha_1 + \alpha_2 + \alpha_3 < \beta \\ \beta_1 + \beta_2 + \beta_3 < \alpha \end{cases}, \qquad \text{(A4)}$$
$$\sum_i \beta_i < \sum_i \alpha_i,$$

which entails $\bar{p} < p$ and by (A4), $r < p$ and *a fortiori* $\dim C_i < p$.

This situation is therefore incompatible with the hypothesis of irreducibility of I^1; the situations $\sum \beta_i > \sum \alpha_i$ and $\sum \beta_i = \sum \alpha_i$ can be handled in the same way so that Case 1 is never realizable. The analyses of Cases 3–5 can be carried out in the same fashion and lead to the same conclusion; only the result of Case 6 is different

$$\textit{Case 6:} \quad \begin{cases} p = m \\ q = m \end{cases}.$$

In these conditions,

$$\begin{cases} \sum_i \alpha_i = \beta \\ \sum_i \beta_i = \alpha \end{cases}, \qquad \text{(A5)}$$
$$\sum_i \beta_i = \sum_i \alpha_i,$$

which entails $\bar{p} = p$ and, by (A5), $r = p$, so that

$$\dim C_i < p, \qquad i = 1, 2, 3,$$

and this case is not realizable.

Case (b) can be treated in precisely the same manner. This completes the proof of the theorem.

APPENDIX B

Let R^N be a $2n$-dimensional covering and $\mathcal{R}^M$ a partial covering of R^N. We shall prove the following lemma:

Lemma B1: Each connected component of $C_{R^N}\mathcal{R}^M$ can have intersections with no more than one boundary component of $\mathcal{R}^M$.

In fact, since the inequality of Lemma 5 is saturated, conditions (i)–(vi) hold and it is easy to convince oneself that the latter restrict drastically the number of possible configurations.

So as not to burden the reader with inessential details, we shall only give a brief sketch of the argument.

A possible configuration violating Lemma B1 is shown in Fig. 18.

By hypothesis,

$$\dim R^N = \alpha_{12} + \sum_{i=3}^{k} \alpha_i = 2n$$

and, by Lemma 7,

$$\gamma_i + \beta_i = 2n, \qquad 1 \le i \le k.$$

Summing these relations,

$$[\alpha_{12} + \gamma_1 + \gamma_2] + \sum_{i=3}^{k} [\alpha_i + \gamma_i] + \sum_{i=1}^{k} \beta_i = (k+1)2n.$$

Now, by Lemma 5, each of these k brackets are $\le 2n$. So that this configuration is not realizable.

By similar reasonings, we are led to Lemma B1.

1 F. Englert and C. De Dominicis, Nuovo Cimento 53A, 1007 (1968).
2 See Ref. 1, p. 1021.
3 F. Englert, R. Brout, and H. Stern, Nuovo Cimento 58A, 601 (1968).
4 R. Brout and F. Englert, Bull. Am. Phys. Soc. 11, 21 (1966).
5 R. Brout, *Topics in Bootstrap Theory, Cargèse Lectures* (Gordon and Breach, New York, 1966).
6 C. de Dominicis and P. C. Martin, J. Math. Phys. 5, 31 (1964).

Calculation of a Certain Type of 6*j*-Symbol

E. de Vries and A. J. van Zanten
Institute for Theoretical Physics, University of Groningen, The Netherlands

(Received 16 December 1971; Revised Manuscript Received 28 January 1972)

It is shown that, for an arbitrary finite simply reducible group, it is possible to express a certain type of 6*j*-symbol in which one of the representations is one-dimensional in terms of 3*j*-symbols $\{j\ k\ l\}$.

For an arbitrary finite simply reducible group, the 6*j*-symbol in which one of the irreducible representations is the trivial one (1_1) can be expressed in terms of 3*j*-symbols [see Ref. 1, Eq. (18b)]. We prove here the following similar result:

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_1 & j_2 & 1_i \end{Bmatrix} = (-1)^{2j_1 + j_3 \otimes (1_i)} \frac{1}{[j_1]} \{j_1\ j_2\ 1_i\}\{j_1\ j_2\ j_3\}. \quad (1)$$

In this equation $(-1)^{j_3 \otimes (1_i)}$ is the phase factor belonging to the irreducible representation, which is the Kronecker product of the irreducible representation (j_3) and the arbitrary one-dimensional representation (1_i). The 3*j*-symbol $\{j\ k\ l\}$ denotes the number of times that the irreducible representation (l) is contained in $(j) \otimes (k)$.

Proof: We start from Eq. (27a) of Ref. 1, taking $(j') = (j_3)$ and $(j) = (1_i)$, and substitute in the rhs $T = SR$. This gives

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_1 & j_2 & 1_i \end{Bmatrix} = \frac{(-1)^{2(1_i)}}{g^2} \sum_{RT} \chi^{(j_1)}(T)\chi^{(j_2)}(RT^{-1}R) \times \chi^{(1_i)}(TR^{-1})\chi^{(j_3)}(R)$$
$$= \frac{(-1)^{2(1_i)}}{g^2} \sum_{RT} \chi^{(j_1)}(T)\chi^{(j_2)}(RT^{-1}R) \times \chi^{(1_i)}(T)\chi^{(1_i)}(R)\chi^{(j_3)}(R)$$
$$= \frac{(-1)^{2(1_i)}}{g^2} \sum_{RT} \chi^{(j_1)\otimes(1_i)}(T) \times \chi^{(j_2)}(RT^{-1}R)\chi^{(j_3)\otimes(1_i)}(R), \quad (2)$$

where we used the facts that (1_i) is a one-dimensional representation and that the characters are real.

Now there are two possibilities for $(j_1) \otimes (1_i)$.
(a) (j_2) is not contained in $(j_1) \otimes (1_i)$. In this case the 6j-symbol in the lhs of Eq. (2) vanishes.
(b) (j_2) is contained in $(j_1) \otimes (1_i)$, and, because (1_i) is a one-dimensional representation, one even has $(j_2) = (j_1) \otimes (1_i)$.

From (a) and (b) we see that $\chi^{(j_1)\otimes(1_i)}(T)$ can be replaced by $\{j_1\ j_2\ 1_i\}\chi^{(j_2)}(T)$. We furthermore express the characters in terms of matrix elements and apply the definition formula of 3jm-symbols and an orthogonality relation (cf. also Ref. 2) and obtain

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_1 & j_2 & 1_i \end{Bmatrix} = \frac{(-1)^{2(1_i)}}{g^2} \{j_1\ j_2\ 1_i\} \sum_{RT} D^{(j_2)}_{\alpha\alpha}(T) D^{(j_2)}_{\beta\gamma}(R) \times D^{(j_2)*}_{\epsilon\gamma}(T) D_{\epsilon\beta}{}^{(j_2)}(R) D^{(j_3)\otimes(1_i)}_{\eta\eta}(R)$$
$$= (-1)^{2(1_i)}\{j_1\ j_2\ 1_i\}\delta_{\alpha\epsilon}\delta_{\alpha\gamma}\frac{1}{[j_2]} \times \begin{pmatrix} j_2 & j_2 & j_3\otimes 1_i \\ \beta & \epsilon & \eta \end{pmatrix}^* \begin{pmatrix} j_2 & j_2 & j_3\otimes 1_i \\ \gamma & \beta & \eta \end{pmatrix}. \quad (3)$$

By applying a symmetry relation for a 3jm-symbol and an orthogonality relation, we get

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_1 & j_2 & 1_i \end{Bmatrix} = (-1)^{2j_1+j_3\otimes(1_i)} \frac{1}{[j_2]}\{j_1\ j_2\ 1_i\}\{j_2\ j_2\ j_3\otimes 1_i\}. \quad (4)$$

The last 3j-symbol in Eq. (4) can be replaced by $\{j_1\ j_2\ j_3\}$, because $(j_2) \otimes (1_i) = (j_1)$ if the 6j-symbol does not vanish. For the same reason $[j_1] = [j_2]$, which gives Eq. (1).

It is not difficult to check that the rhs of Eq. (1) is symmetric in j_1 and j_2 as it must be.

Equation (1) can be used, e.g., for the calculation of the 6j-symbols of the Dirac matrix group, which play a role in Fierz transformations (see Ref. 3).

Along the same lines as in the proof of Eq. (1) one can derive an expression for the square of a more general 6j-symbol in terms of 3j-symbols

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_1' & j_2' & 1_i \end{Bmatrix}^2 = \frac{1}{[j_1][j_2]}\{j_1\ j_2\ j_3\}\{j_1\ j_2'\ 1_i\}\{j_1'\ j_2\ 1_i\}. \quad (5)$$

For the derivation of Eq. (5) one has to start from Eq. (27) of Ref. 1.

ACKNOWLEDGMENTS

The authors are indebted to Professor Dr. H. J. Groenewold for many discussions. One of them (E. de V.) has carried out this work as a scientific staff member of the Stichting F.O.M. (Foundation for Fundamental Research of Matter) which is financially supported by the Netherlands Organization for Pure Scientific Research (Z.W.O.).

[1] E. P. Wigner, "On the Matrices Which Reduce the Kronecker Product of Representations of S.R. Groups" (Princeton, 1951), in *Quantum Theory of Angular Momentum* edited by L. C. Biedenharn and H. van Dam (Academic, New York, 1965).

[2] A. J. van Zanten and E. de Vries, Physica **49**, 536 (1970).

[3] E. de Vries and A. J. van Zanten, Commun. Math. Phys. **17**, 322 (1970).

Local Existence of Solutions to the Equations $\Gamma^i_{jk;l} = 0, g_{ij;k} = 0$ When $R^i_{jkl} \neq 0$

M. Muraskin

Department of Physics, University of North Dakota, Grand Forks, North Dakota 58201

(Received 27 January 1971; Revised Manuscript Received 2 February 1972)

We prove that solutions to the equations $\Gamma^i_{jk;l} = 0$ and $g_{ij;k} = 0$ exist locally when $R^i{}_{jkl} \neq 0$. Previously this was only shown for the special case of $R^i{}_{jkl} = 0$.

1. INTRODUCTION AND SOLUTION OF INTEGRABILITY EQUATIONS

In a previous paper,[1] we introduced a four-dimensional nonsymmetric[2] Lorentz invariant field theory. We showed that solutions to the equations

$$\frac{\partial \Gamma^i_{jk}}{\partial x^l} + \Gamma^m_{jk}\Gamma^i_{ml} - \Gamma^i_{mk}\Gamma^m_{jl} - \Gamma^i_{jm}\Gamma^m_{kl} \equiv \Gamma^i_{jk;l} = 0, \quad (1)$$

$$\frac{\partial g_{ik}}{\partial x^l} - \Gamma^m_{il}g_{im} - \Gamma^m_{kl}g_{im} \equiv g_{ik;l} = 0 \quad (2)$$

exist locally. The problem of local existence depends on showing that there exists a consistent set of field components at the prescribed origin point.[3] This is not a trivial matter since the consistency conditions (integrability conditions) involve more equations than there are field components.

We introduce $e^{\alpha}{}_i$ by means of

$$\Gamma^i_{jk} = e_{\alpha}{}^i \frac{\partial e^{\alpha}{}_j}{\partial x^k}, \quad (3)$$

$$g_{ij} = e^{\alpha}{}_i e^{\beta}{}_j g_{\alpha\beta}, \quad (4)$$

where $g_{\alpha\beta}$ is the Minkowski metric. We get from

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_1 & j_2 & 1_i \end{Bmatrix} = \frac{(-1)^{2(1_i)}}{g^2} \sum_{RT} \chi^{(j_1)}(T)\chi^{(j_2)}(RT^{-1}R) \times \chi^{(1_i)}(TR^{-1})\chi^{(j_3)}(R)$$
$$= \frac{(-1)^{2(1_i)}}{g^2} \sum_{RT} \chi^{(j_1)}(T)\chi^{(j_2)}(RT^{-1}R) \times \chi^{(1_i)}(T)\chi^{(1_i)}(R)\chi^{(j_3)}(R)$$
$$= \frac{(-1)^{2(1_i)}}{g^2} \sum_{RT} \chi^{(j_1)\otimes(1_i)}(T) \times \chi^{(j_2)}(RT^{-1}R)\chi^{(j_3)\otimes(1_i)}(R), \quad (2)$$

where we used the facts that (1_i) is a one-dimensional representation and that the characters are real.

Now there are two possibilities for $(j_1) \otimes (1_i)$.
(a) (j_2) is not contained in $(j_1) \otimes (1_i)$. In this case the 6j-symbol in the lhs of Eq. (2) vanishes.
(b) (j_2) is contained in $(j_1) \otimes (1_i)$, and, because (1_i) is a one-dimensional representation, one even has $(j_2) = (j_1) \otimes (1_i)$.

From (a) and (b) we see that $\chi^{(j_1)\otimes(1_i)}(T)$ can be replaced by $\{j_1\ j_2\ 1_i\}\chi^{(j_2)}(T)$. We furthermore express the characters in terms of matrix elements and apply the definition formula of 3jm-symbols and an orthogonality relation (cf. also Ref. 2) and obtain

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_1 & j_2 & 1_i \end{Bmatrix} = \frac{(-1)^{2(1_i)}}{g^2} \{j_1\ j_2\ 1_i\} \sum_{RT} D^{(j_2)}_{\alpha\alpha}(T) D^{(j_2)}_{\beta\gamma}(R) \times D^{(j_2)*}_{\epsilon\gamma}(T) D_{\epsilon\beta}{}^{(j_2)}(R) D^{(j_3)\otimes(1_i)}_{\eta\eta}(R)$$
$$= (-1)^{2(1_i)}\{j_1\ j_2\ 1_i\}\delta_{\alpha\epsilon}\delta_{\alpha\gamma}\frac{1}{[j_2]} \times \begin{pmatrix} j_2 & j_2 & j_3\otimes 1_i \\ \beta & \epsilon & \eta \end{pmatrix}^* \begin{pmatrix} j_2 & j_2 & j_3\otimes 1_i \\ \gamma & \beta & \eta \end{pmatrix}. \quad (3)$$

By applying a symmetry relation for a 3jm-symbol and an orthogonality relation, we get

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_1 & j_2 & 1_i \end{Bmatrix} = (-1)^{2j_1+j_3\otimes(1_i)} \frac{1}{[j_2]}\{j_1\ j_2\ 1_i\}\{j_2\ j_2\ j_3\otimes 1_i\}. \quad (4)$$

The last 3j-symbol in Eq. (4) can be replaced by $\{j_1\ j_2\ j_3\}$, because $(j_2) \otimes (1_i) = (j_1)$ if the 6j-symbol does not vanish. For the same reason $[j_1] = [j_2]$, which gives Eq. (1).

It is not difficult to check that the rhs of Eq. (1) is symmetric in j_1 and j_2 as it must be.

Equation (1) can be used, e.g., for the calculation of the 6j-symbols of the Dirac matrix group, which play a role in Fierz transformations (see Ref. 3).

Along the same lines as in the proof of Eq. (1) one can derive an expression for the square of a more general 6j-symbol in terms of 3j-symbols

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_1' & j_2' & 1_i \end{Bmatrix}^2 = \frac{1}{[j_1][j_2]}\{j_1\ j_2\ j_3\}\{j_1\ j_2'\ 1_i\}\{j_1'\ j_2\ 1_i\}. \quad (5)$$

For the derivation of Eq. (5) one has to start from Eq. (27) of Ref. 1.

ACKNOWLEDGMENTS

The authors are indebted to Professor Dr. H. J. Groenewold for many discussions. One of them (E. de V.) has carried out this work as a scientific staff member of the Stichting F.O.M. (Foundation for Fundamental Research of Matter) which is financially supported by the Netherlands Organization for Pure Scientific Research (Z.W.O.).

[1] E. P. Wigner, "On the Matrices Which Reduce the Kronecker Product of Representations of S.R. Groups" (Princeton, 1951), in *Quantum Theory of Angular Momentum* edited by L. C. Biedenharn and H. van Dam (Academic, New York, 1965).

[2] A. J. van Zanten and E. de Vries, Physica **49**, 536 (1970).

[3] E. de Vries and A. J. van Zanten, Commun. Math. Phys. **17**, 322 (1970).

Local Existence of Solutions to the Equations $\Gamma^i_{jk;l} = 0, g_{ij;k} = 0$ When $R^i_{jkl} \neq 0$

M. Muraskin

Department of Physics, University of North Dakota, Grand Forks, North Dakota 58201

(Received 27 January 1971; Revised Manuscript Received 2 February 1972)

We prove that solutions to the equations $\Gamma^i_{jk;l} = 0$ and $g_{ij;k} = 0$ exist locally when $R^i{}_{jkl} \neq 0$. Previously this was only shown for the special case of $R^i{}_{jkl} = 0$.

1. INTRODUCTION AND SOLUTION OF INTEGRABILITY EQUATIONS

In a previous paper,[1] we introduced a four-dimensional nonsymmetric[2] Lorentz invariant field theory. We showed that solutions to the equations

$$\frac{\partial \Gamma^i_{jk}}{\partial x^l} + \Gamma^m_{jk}\Gamma^i_{ml} - \Gamma^i_{mk}\Gamma^m_{jl} - \Gamma^i_{jm}\Gamma^m_{kl} \equiv \Gamma^i_{jk;l} = 0, \quad (1)$$

$$\frac{\partial g_{ik}}{\partial x^l} - \Gamma^m_{il}g_{im} - \Gamma^m_{kl}g_{im} \equiv g_{ik;l} = 0 \quad (2)$$

exist locally. The problem of local existence depends on showing that there exists a consistent set of field components at the prescribed origin point.[3] This is not a trivial matter since the consistency conditions (integrability conditions) involve more equations than there are field components.

We introduce $e^{\alpha}{}_i$ by means of

$$\Gamma^i_{jk} = e_{\alpha}{}^i \frac{\partial e^{\alpha}{}_j}{\partial x^k}, \quad (3)$$

$$g_{ij} = e^{\alpha}{}_i e^{\beta}{}_j g_{\alpha\beta}, \quad (4)$$

where $g_{\alpha\beta}$ is the Minkowski metric. We get from

$$\frac{\partial^2 e^{\alpha}{}_i}{\partial x^j \partial x^k} = \frac{\partial^2 e^{\alpha}{}_i}{\partial x^k \partial x^j} \tag{5}$$

the following consistency conditions (integrability conditions):

$$R^i{}_{jkl} = 0. \tag{6}$$

Equation (6) implies, as well, that all the mixed derivatives of $g_{ij}, \Gamma^i_{jk}, e^{\alpha}{}_i$ are symmetric. Combining (6) with (1), we get

$$\Gamma^t_{ij}\Gamma^j_{km} - \Gamma^t_{ij}\Gamma^j_{mk} + \Gamma^j_{im}\Gamma^t_{jk} - \Gamma^j_{ik}\Gamma^t_{jm} = 0. \tag{7}$$

These are 96 relations for 64 Γ^i_{jk}. In our previous paper, we showed by explicit construction that nontrivial solutions to (7) exist. This means that solutions to (1), (2), and (6) exist locally.

In our previous work, we noted the possibility of a theory based on (1) and (2), but not (6) and (3). In this case, the consistency takes the form

$$g_{ih}R^t{}_{ipl} + g_{it}R^t{}_{hpl} = 0, \tag{8}$$

$$\Gamma^i_{mk}R^m{}_{jpl} + \Gamma^i_{jm}R^m{}_{kpl} - \Gamma^m_{jk}R^i{}_{mpl} = 0. \tag{9}$$

These equations follow from the requirement that all mixed derivatives of tensor functions of g_{ij}, Γ^i_{jk} be symmetric. In this paper, we show that solutions to (8) and (9) exist when $R^i{}_{jkl} \neq 0$. Thus, in fact, the field theory based on (1) and (2), but not (6), also has solutions existing locally. Equation (8) constitutes 60 equations and Eq. (9) is 384 equations (there is antisymmetry in the indices p and l. Thus, we have considerably more equations to solve than for the $R^i{}_{jkl} = 0$ case.

By trial and error, we have come up with solutions of (8) and (9). We first write

$$\Gamma^i_{jk} = a^{\alpha}{}_i a^{\beta}{}_j \alpha^{\gamma}{}_k \Gamma^{\alpha}_{\beta\gamma}(x), \tag{10}$$

$$g_{ij} = a^{\alpha}{}_i a^{\beta}{}_j g_{\alpha\beta}(x). \tag{11}$$

At the prescribed point P, we choose $g_{\alpha\beta}(P)$ to be the Minkowski metric $(1,-1,-1,-1)$ and $\Gamma^{\alpha}_{\beta\gamma}(P)$ is chosen to have the following nonvanishing components:

$$\begin{aligned} &\Gamma^1_{23} = -\Gamma^2_{13}, \quad \Gamma^1_{20} = -\Gamma^2_{10}, \\ &\Gamma^3_{33} = \Gamma^0_{03}, \quad \Gamma^0_{00} = \Gamma^3_{30}. \end{aligned} \tag{12}$$

The $a^{\alpha}{}_i(P)$ are chosen arbitrarily. Under these circumstances, we find $R^i{}_{jkl} \neq 0$ and (8) and (9) are satisfied.[4] Furthermore, all 64 Γ^i_{jk} and all 16 g_{ij} in (10) and (11) can be made nonzero. This occurs, for example, when we take $a^{\alpha}{}_i(P)$ to be

$$\begin{aligned} &a^1{}_1 = 0.8, \quad a^1{}_2 = 0.6, \quad a^1{}_3 = 0.5, \quad a^1{}_0 = 0.15, \\ &a^2{}_1 = 0.65, \quad a^2{}_2 = 0.9, \quad a^2{}_3 = 0.25, \quad a^2{}_0 = 0.55, \\ &a^3{}_1 = 0.72, \quad a^3{}_2 = 0.58, \quad a^3{}_3 = 0.13, \quad a^3{}_0 = 0.62, \\ &a^0{}_1 = 0.4, \quad a^0{}_2 = 0.2, \quad a^0{}_3 = 0.1, \quad a^0{}_0 = 2.0, \end{aligned} \tag{13}$$

We also take $\Gamma^1_{23} = 0.3$, $\Gamma^1_{20} = 0.4$, $\Gamma^3_{33} = 0.5$, $\Gamma^0_{00} = 0.27$. By means of a finite difference approximation, and with the help of a computer, we have also found that g_{ij}, Γ^i_{jk} all vary from point to point [using (13)] and, thus, constitute a nontrivial theory.

Furthermore, we have found, in all cases considered, that the field values are independent of path within a tolerance of nine decimal places. (The distance from the origin was taken to be of the order of 0.1).

2. INTEGRABILITY AND LOCAL EXISTENCE

We have made the statement that once we demonstrate integrability, then local existence will follow. We can ask, what does our computer work have to say about this statement. So far as the computer mapping is concerned, the only problem with respect to the existence of a computer finite-difference solution is whether the field value at a point is independent of path from the origin. Thus, the existence of computer solutions to the field equations in the vicinity of the origin (where the field is taken to be finite) depends only on integrability. Furthermore, so far as the computer solutions are concerned, we can continue, in principle, to lower the grid size indefinitely. Then, the finite difference system would go into the exact field equations [Eq. (2.3) of Ref. 1]. For any grid size, the only problem with respect to existence of solutions to the computer program is integrability. Therefore, when dx^l is infinitesimal, we conclude that the exact mathematical solutions exists provided that integrability can be established. This is a local result since we cannot infer that the field remains finite everywhere.

We also point out that once the integrability equations are established at the origin, then they hold in the vicinity of the origin as well. We define

$$A^i{}_{kjpl} \equiv \Gamma^i_{mk}R^m{}_{jpl} + \Gamma^i_{jk}R^m{}_{jpl} - \Gamma^m_{jk}R^i{}_{mpl}. \tag{14}$$

This has been established to be zero at the origin. Note that since $\Gamma^i_{jk;l} = 0$, it follows that $A^i{}_{kjpl;m} = 0$. Thus, it follows that $\partial A^i{}_{kjpl}/\partial x^m = 0$ when we use $A^i{}_{kjpl} = 0$ at the origin. We can do the same for all higher derivatives that are finite. Thus, the integrability equations are satisfied in the vicinity of the origin, once they are satisfied at the origin.

We may also prove that solutions to the field equations exist locally once integrability has been established, as follows. The field is assumed finite at the origin point. By means of the field equations, we can compute the higher derivatives of Γ^i_{jk} entirely in terms of Γ^i_{jk} at the origin point. In a similar fashion, the higher derivatives of g_{ij} are given in terms of Γ^i_{jk} and g^{ij} at the origin point. Due to the finiteness of the field at the origin, at least a finite number of derivatives will be finite as well (since they are given by finite products of the field at the origin). When we set up the field equations in Ref. 1, we assumed continuity for all finite tensor functions. Thus, the conditions for Taylor's theorem are satisfied (see Apostol,[5] Theorems 5–14 and 6–21) Taylor's theorem says[5] that the field can be approximated by a polynomial of degree $n-1$ if the nth-order derivative exists in the region containing P and R (where R is close to P). Thus, we may write

$$\Gamma^i_{jk}(R) = \Gamma^i_{jk}(P) + \left(\frac{\partial \Gamma^i_{jk}}{\partial x^l}\right)_P dx^l + \frac{1}{2}\left(\frac{\partial^2 \Gamma^i_{jk}}{\partial x^m \partial x^n}\right)_P \times dx^m dx^n + \cdots. \tag{15}$$

Now, by the field equations we see $\Gamma^i_{jk}(R)$ is entirely

given in terms of the field at the origin point P. Thus, we can calculate $\Gamma^i_{j\,k}(R)$[and by similar arguments $g_{ij}(R)$] provided a consistent set of Γ^i_{jk} exists at P. A consistent set of $\Gamma^i_{j\,k}$ at P must satisfy the requirement that the mixed derivatives[6] of all functions of $\Gamma^i_{j\,k}$ and g_{ij} be symmetric. These relations are just given by Eqs. (8) and (9). Thus, we conclude that local existence depends on being able to obtain solutions to (8) and (9) which, in fact, we have already found. A solution to (8) and (9) is given by (10), (11), and (12).

3. CONCLUSION

Thus, nontrivial solutions to (1) and (2) with $R^i{}_{jkl} \neq 0$ exist locally. Further investigations of the $\Gamma^i_{jk;\,l} = 0$, $g_{ij;\,k} = 0$ field theory appear elsewhere.[7]

---

[1] M. Muraskin, Ann. Phys. (N.Y.) **59**, 27 (1970). This reference gives background material for the present paper.

[2] We shall make a few comments about the $\Gamma^i_{jk} = \Gamma^i_{kj}$ situation. For symmetric Γ^i_{jk}, we can make a general coordinate transformation so that $\Gamma^i_{jk} = 0$ at the origin. Then it might appear that $\Gamma^i_{jk} = 0$ at all points as a consequence of (1) and, thus, no nontrivial solutions would be possible for symmetric Γ^i_{jk}. However, this argument is not correct since Eq. (1) is not covariant under general coordinate transformations. Thus, the transformation that leads to $\Gamma^i_{jk} = 0$ at the origin also implies $\Gamma^i_{jk;\,l} \neq 0$.

[3] See Sec. II for detailed discussion.

[4] $R^1{}_{230} = -R^1{}_{203}$ is nonzero.

[5] T. Apostol, *Mathematical Analysis* (Addison-Wesley, Reading, Mass., 1957), pp. 96, 123.

[6] The problem of consistency when the field depends on a number of parameters as well as x, y, z, x^0 is discussed in L. Eisenhart, *Continuous Groups of Transformation* (Dover, New York, 1961), Chap. 1.

[7] M. Muraskin and T. Clark, Ann. Phys. (N.Y.) **59**, 27 (1970). M. Muraskin, J. Math. Phys. **12**, 28 (1971); Intern. J. Theoret. Phys. **4**, 49 (1971).

A Class of Stationary Electromagnetic Vacuum Fields*

Werner Israel
Mathematics Department, University of Alberta, Edmonton, Canada

and

Gordon A. Wilson
Physics Department, University of Alberta, Edmonton, Canada

(Received 2 June 1971; Revised Manuscript Received 10 January 1972)

It is shown how a new class of stationary electromagnetic vacuum fields can be generated from solutions of Laplace's equation. These fields are a stationary generalization of the static electromagnetic vacuum fields of Weyl, Majumdar, and Papapetrou, and are plausibly interpreted as exterior fields of static or steadily moving distributions of charged dust having numerically equal charge and mass densities.

1. INTRODUCTION

Coulomb's law and Newton's law of gravity are formally identical apart from a sign. Hence, classically, any unstressed distribution of matter can, if suitably charged, be maintained in neutral equilibrium under a balance between the gravitational attraction and electrical repulsion of its parts.

Indications that this obvious Newtonian fact has a relativistic analog first emerged when Weyl[1] obtained a particular class of static electromagnetic vacuum fields, later generalized by Majumdar[2] and Papapetrou[3] to remove Weyl's original restriction to axial symmetry, and further studied by Bonnor[4] and Synge.[5] The Papapetrou–Majumdar fields are to all appearances the external fields of static sources whose charge and mass are numerically equal (in relativistic units: $G = c = 1$). That they are indeed interpretable as external fields of static distributions of charged dust having equal charge and mass densities has been shown by Das,[6] who has examined the corresponding interior fields.

Astrophysical bodies are electrically neutral to a good approximation, and the Papapetrou–Majumdar solutions have up to now received little attention. It seems to us, however, that they can play a useful, if limited, astrophysical role in providing simple quasistatic analogues for complex dynamical processes like the disappearance of asymmetries in gravitational collapse or the collision of black holes. In reality, such a process always involves large kinetic energies and at present can only be handled by elaborate numerical integrations under the assumption of small departures from spherical symmetry.[7,8] However, for *charged* bodies in neutral equilibrium the process can be made arbitrarily slow, and the details easily followed as a sequence of stationary configurations. While this procedure prevents us from considering features of undeniable observational importance, such as the emission of gravitational waves, it is for that very reason ideally suited for isolating and elucidating certain basic issues of principle relating to the final phases of the process.

Some of these questions are pursued in detail elsewhere.[9] Our purpose here is to demonstrate that the Papapetrou–Majumdar class can be extended straightforwardly from the static to the stationary realm.

2. STATIONARY FIELDS

The metric of an arbitrary stationary field is conveniently expressed in the form[10]

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -f^{-1}\gamma_{mn}dx^m dx^n + f(\omega_m dx^m + dx^4)^2, \quad (1)$$

in which f, γ_{mn}, and ω_m are independent of the time coordinate x^4. The inverse of $g_{\mu\nu}$ is given by

$$g^{\mu\nu}\frac{\partial}{\partial x^\mu}\frac{\partial}{\partial x^\nu} = -f\gamma^{mn}\frac{\partial}{\partial x^m}\frac{\partial}{\partial x^n} + 2f\omega^m\frac{\partial}{\partial x^m}\frac{\partial}{\partial x^4} + (f^{-1} - f\omega^2)\frac{\partial^2}{(\partial x^4)^2}, \quad (2)$$

given in terms of the field at the origin point P. Thus, we can calculate $\Gamma^i_{j\,k}(R)$[and by similar arguments $g_{ij}(R)$] provided a consistent set of Γ^i_{jk} exists at P. A consistent set of $\Gamma^i_{j\,k}$ at P must satisfy the requirement that the mixed derivatives[6] of all functions of $\Gamma^i_{j\,k}$ and g_{ij} be symmetric. These relations are just given by Eqs. (8) and (9). Thus, we conclude that local existence depends on being able to obtain solutions to (8) and (9) which, in fact, we have already found. A solution to (8) and (9) is given by (10), (11), and (12).

3. CONCLUSION

Thus, nontrivial solutions to (1) and (2) with $R^i{}_{jkl} \neq 0$ exist locally. Further investigations of the $\Gamma^i_{jk;\,l} = 0$, $g_{ij;\,k} = 0$ field theory appear elsewhere.[7]

---

[1] M. Muraskin, Ann. Phys. (N.Y.) **59**, 27 (1970). This reference gives background material for the present paper.

[2] We shall make a few comments about the $\Gamma^i_{jk} = \Gamma^i_{kj}$ situation. For symmetric Γ^i_{jk}, we can make a general coordinate transformation so that $\Gamma^i_{jk} = 0$ at the origin. Then it might appear that $\Gamma^i_{jk} = 0$ at all points as a consequence of (1) and, thus, no nontrivial solutions would be possible for symmetric Γ^i_{jk}. However, this argument is not correct since Eq. (1) is not covariant under general coordinate transformations. Thus, the transformation that leads to $\Gamma^i_{jk} = 0$ at the origin also implies $\Gamma^i_{jk;\,l} \neq 0$.

[3] See Sec. II for detailed discussion.

[4] $R^1{}_{230} = -R^1{}_{203}$ is nonzero.

[5] T. Apostol, *Mathematical Analysis* (Addison-Wesley, Reading, Mass., 1957), pp. 96, 123.

[6] The problem of consistency when the field depends on a number of parameters as well as x, y, z, x^0 is discussed in L. Eisenhart, *Continuous Groups of Transformation* (Dover, New York, 1961), Chap. 1.

[7] M. Muraskin and T. Clark, Ann. Phys. (N.Y.) **59**, 27 (1970). M. Muraskin, J. Math. Phys. **12**, 28 (1971); Intern. J. Theoret. Phys. **4**, 49 (1971).

A Class of Stationary Electromagnetic Vacuum Fields*

Werner Israel
Mathematics Department, University of Alberta, Edmonton, Canada

and

Gordon A. Wilson
Physics Department, University of Alberta, Edmonton, Canada

(Received 2 June 1971; Revised Manuscript Received 10 January 1972)

It is shown how a new class of stationary electromagnetic vacuum fields can be generated from solutions of Laplace's equation. These fields are a stationary generalization of the static electromagnetic vacuum fields of Weyl, Majumdar, and Papapetrou, and are plausibly interpreted as exterior fields of static or steadily moving distributions of charged dust having numerically equal charge and mass densities.

1. INTRODUCTION

Coulomb's law and Newton's law of gravity are formally identical apart from a sign. Hence, classically, any unstressed distribution of matter can, if suitably charged, be maintained in neutral equilibrium under a balance between the gravitational attraction and electrical repulsion of its parts.

Indications that this obvious Newtonian fact has a relativistic analog first emerged when Weyl[1] obtained a particular class of static electromagnetic vacuum fields, later generalized by Majumdar[2] and Papapetrou[3] to remove Weyl's original restriction to axial symmetry, and further studied by Bonnor[4] and Synge.[5] The Papapetrou–Majumdar fields are to all appearances the external fields of static sources whose charge and mass are numerically equal (in relativistic units: $G = c = 1$). That they are indeed interpretable as external fields of static distributions of charged dust having equal charge and mass densities has been shown by Das,[6] who has examined the corresponding interior fields.

Astrophysical bodies are electrically neutral to a good approximation, and the Papapetrou–Majumdar solutions have up to now received little attention. It seems to us, however, that they can play a useful, if limited, astrophysical role in providing simple quasistatic analogues for complex dynamical processes like the disappearance of asymmetries in gravitational collapse or the collision of black holes. In reality, such a process always involves large kinetic energies and at present can only be handled by elaborate numerical integrations under the assumption of small departures from spherical symmetry.[7,8] However, for *charged* bodies in neutral equilibrium the process can be made arbitrarily slow, and the details easily followed as a sequence of stationary configurations. While this procedure prevents us from considering features of undeniable observational importance, such as the emission of gravitational waves, it is for that very reason ideally suited for isolating and elucidating certain basic issues of principle relating to the final phases of the process.

Some of these questions are pursued in detail elsewhere.[9] Our purpose here is to demonstrate that the Papapetrou–Majumdar class can be extended straightforwardly from the static to the stationary realm.

2. STATIONARY FIELDS

The metric of an arbitrary stationary field is conveniently expressed in the form[10]

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -f^{-1}\gamma_{mn}dx^m dx^n + f(\omega_m dx^m + dx^4)^2, \quad (1)$$

in which f, γ_{mn}, and ω_m are independent of the time coordinate x^4. The inverse of $g_{\mu\nu}$ is given by

$$g^{\mu\nu}\frac{\partial}{\partial x^\mu}\frac{\partial}{\partial x^\nu} = -f\gamma^{mn}\frac{\partial}{\partial x^m}\frac{\partial}{\partial x^n} + 2f\omega^m\frac{\partial}{\partial x^m}\frac{\partial}{\partial x^4} + (f^{-1} - f\omega^2)\frac{\partial^2}{(\partial x^4)^2}, \quad (2)$$

where γ^{mn} is the 3×3 symmetric matrix inverse to γ_{mn}, $\omega^m = \gamma^{mn}\omega_n$ and $\omega^2 = \gamma^{mn}\omega_m\omega_n$. The determinants of $g_{\mu\nu}$ and γ_{mn} are related by

$$(-g)^{1/2} = f^{-1}\gamma^{1/2}. \tag{3}$$

The 3-vector ω_m in (1) is arbitrary up to an additive gradient $\partial_m\lambda(x^1,x^2,x^3)$, corresponding to the possibility of making arbitrary time translations $x^4 \to x^{4'} = x^4 - \lambda(x^1,x^2,x^3)$. However, we can derive from it an invariant "torsion vector"

$$f^{-2}\tau^m = -\gamma^{-1/2}\epsilon^{mpq}\partial_p\omega_q \quad \text{or} \quad f^{-2}\boldsymbol{\tau} = -\operatorname{curl}\boldsymbol{\omega} \tag{4}$$

in terms of a three-dimensional vector calculus employing $\gamma_{mn}dx^m dx^n$ as base metric.

We next consider a stationary electromagnetic field $F_{\mu\nu} = \partial_\nu A_\mu - \partial_\mu A_\nu$ in the space–time (1). The condition of time independence $\partial_4 A_\mu = 0$ yields for the "electric" components

$$F_{4n} = \partial_n A_4, \tag{5}$$

while the source-free Maxwell equations

$$\partial_n[(-g)^{1/2}\,{}^{(4)}F^{\mu n}] = 0 \tag{6}$$

for $\mu = m$ give the "magnetic" components

$${}^{(4)}F^{mn} = f\gamma^{-1/2}\epsilon^{mnp}\partial_p\Phi \tag{7}$$

in terms of a magnetic scalar potential Φ. All remaining components are then conveniently expressed in terms of these six; for example,

$${}^{(4)}F^{n4} = \omega_m\,{}^{(4)}F^{mn} + F_{4m}\gamma^{mn}, \tag{8}$$

an identity which follows readily from (1) or (2). Equation (6) with $\mu = 4$ now yields, on substituting (8), (7), (5) and (4),

$$\operatorname{div}(f^{-1}\nabla A_4) = -f^{-2}\,\boldsymbol{\tau}\cdot\nabla\Phi. \tag{9}$$

Next, writing $F_{mn}(= \partial_n A_m - \partial_m A_n)$ in terms of (5) and (7) and expressing the cyclic identity $\epsilon^{mnp}\partial_p F_{mn} = 0$, we obtain

$$\operatorname{div}(f^{-1}\nabla\Phi) = f^{-2}\,\boldsymbol{\tau}\cdot\nabla A_4. \tag{10}$$

If we now introduce[11] the complex scalar potential

$$\Psi = A_4 + i\Phi, \tag{11}$$

then (9) and (10) combine to give

$$\operatorname{div}(f^{-1}\nabla\Psi) = if^{-2}\,\boldsymbol{\tau}\cdot\nabla\Psi. \tag{12}$$

We have thus reduced the entire set of Maxwell's equations to the single complex equation (12).

3. GRAVITATIONAL FIELD EQUATIONS

The Ricci tensor

$$R_{\mu\nu} = \partial_\mu\Gamma^\alpha_{\nu\alpha} - \partial_\alpha\Gamma^\alpha_{\mu\nu} + \Gamma^\alpha_{\beta\mu}\Gamma^\beta_{\alpha\nu} - \Gamma^\alpha_{\beta\alpha}\Gamma^\beta_{\mu\nu}$$

for the general stationary metric (1) is conveniently expressed in terms of a complex 3-vector $\mathbf{G}$, defined by

$$2f\mathbf{G} = \nabla f + i\boldsymbol{\tau}. \tag{13}$$

Then[12]

$$-f^{-2}R_{44} = \operatorname{div}\mathbf{G} + (\mathbf{G}^* - \mathbf{G})\cdot\mathbf{G}, \tag{14a}$$

$$-2if^{-2}\,{}^{(4)}R^m_4 = \gamma^{-1/2}\epsilon^{mpq}(\partial_q G_p + G_p G^*_q), \tag{15a}$$

$$f^{-2}(\gamma_{pm}\gamma_{qn}\,{}^{(4)}R^{mn} - \gamma_{pq}R_{44}) = R_{pq}(\gamma) + G_pG^*_q + G^*_pG_q. \tag{16a}$$

Here, $R_{pq}(\gamma)$ denotes the Ricci tensor formed from the 3-metric $\gamma_{mn}dx^m dx^n$.

For the electromagnetic energy tensor

$$-4\pi\,T_{\mu\nu} = g^{\alpha\beta}F_{\mu\alpha}F_{\nu\beta} - \tfrac{1}{4}g_{\mu\nu}F_{\alpha\beta}F^{\alpha\beta},$$

one derives from the formulas of the previous section

$$\tfrac{1}{2}F_{\mu\nu}F^{\mu\nu} = (\nabla\Phi)^2 - (\nabla A_4)^2,$$
$$8\pi f^{-1}T_{44} = (\nabla\Phi)^2 + (\nabla A_4)^2, \tag{14b}$$

$$4\pi f^{-1}\,{}^{(4)}T^m_4 = \gamma^{-1/2}\epsilon^{mpq}(\partial_p\Phi)(\partial_q A_4), \tag{15b}$$

$$-4\pi f^{-1}\,{}^{(4)}T^{mn} = (\partial^m\Phi)(\partial^n\Phi) + (\partial^m A_4)(\partial^n A_4) - \tfrac{1}{2}\gamma^{mn}[(\nabla\Phi)^2 + (\nabla A_4)^2] \tag{16b}$$

with $\partial^m = \gamma^{mn}\partial_n$.

We can now impose the Einstein field equations $R_{\mu\nu} = -8\pi\,T_{\mu\nu}$. From (15a), (15b), we find

$$\operatorname{curl}\boldsymbol{\tau} = -4\nabla\Phi\times\nabla A_4 = i\operatorname{curl}(\Psi\nabla\Psi^* - \Psi^*\nabla\Psi),$$

so that the equation

$$\boldsymbol{\tau} + i(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*) = \nabla\psi \tag{17}$$

defines a real scalar ψ up to an additive constant.

We next define a complex function[11]

$$\mathcal{E} = f - \Psi\Psi^* + i\psi. \tag{18}$$

By virtue of (13) and (17),

$$f\mathbf{G} = \tfrac{1}{2}\nabla\mathcal{E} + \Psi^*\nabla\Psi. \tag{19}$$

Substituting (19) into the field equations (14a), (14b) and employing (12) leads to[11]

$$f\nabla^2\mathcal{E} = \nabla\mathcal{E}\cdot(\nabla\mathcal{E} + 2\Psi^*\nabla\Psi), \tag{20}$$

while (12) itself can be written

$$f\nabla^2\Psi = \nabla\Psi\cdot(\nabla\mathcal{E} + 2\Psi^*\nabla\Psi), \tag{21}$$

and we note from (18) that

$$f = \tfrac{1}{2}(\mathcal{E} + \mathcal{E}^*) + \Psi\Psi^*. \tag{22}$$

Finally the field equations (16a), (16b) reduce to

$$-f^2R_{mn}(\gamma) = \tfrac{1}{2}\mathcal{E}_{,(m}\mathcal{E}^*_{,n)} + \Psi\mathcal{E}_{,(m}\Psi^*_{,n)} + \Psi^*\mathcal{E}^*_{,(m}\Psi_{,n)} - (\mathcal{E} + \mathcal{E}^*)\Psi_{,(m}\Psi^*_{,n)}, \tag{23}$$

in which, for example,

$$2\mathcal{E}_{,(m}\mathcal{E}^*_{\ ,n)} \equiv (\partial_m \mathcal{E})(\partial_n \mathcal{E}^*) + (\partial_n \mathcal{E})(\partial_m \mathcal{E}^*).$$

The complete system of electromagnetic and gravitational field equations for an arbitrary electromagnetic vacuum field are summed up in (20), (21), and (23).

4. GENERALIZED PAPAPETROU–MAJUMDAR SOLUTIONS

So far, our considerations have been quite general. We now examine whether solutions of the system (20), (21), and (23) exist for which the background metric $\gamma_{mn}dx^m dx^n$ is *flat*. In this case equations (23) [with $R_{mn}(\gamma) = 0$] are satisfied if and only if there is a linear relation

$$\Psi = a + b\mathcal{E}, \qquad \text{with } a^*b + ab^* = -\tfrac{1}{2}$$

(as one easily verifies, for example, by choosing $\mathcal{E} = x^1$ and $\mathcal{E}^* = x^2$ as coordinates). Both $\mathcal{E}$ and Ψ contain arbitrary additive constants, and it is convenient to adjust these so that $\mathcal{E} \to 1$ when $\Psi \to 0$. We thus obtain

$$\Psi = \tfrac{1}{2}e^{i\alpha}(1 - \mathcal{E}), \tag{24}$$

in which the arbitrary real constant α represents the "complexion" of the electromagnetic field. We can submit this field to any constant duality rotation without affecting the geometry.

If we now substitute (24) into (20) and (21), *both* reduce to

$$\nabla^2[(1 + \mathcal{E})^{-1}] = 0 \tag{25}$$

which is Laplace's equation in Euclidean 3-space.

We conclude by summarizing the procedure for obtaining the complete field. (a) Write down a solution of (25) in terms of any convenient coordinates x^m. Suppose the Euclidean line element takes the form $\gamma_{mn}dx^m dx^n$ in these coordinates. (b) Obtain f, τ, and ω from the equations

$$f = \tfrac{1}{4}(1 + \mathcal{E})(1 + \mathcal{E}^*),$$
$$if^{-1}\,\tau = \nabla\{\ln[(1 + \mathcal{E})/(1 + \mathcal{E}^*)]\}, \quad \operatorname{curl}\omega = -f^{-2}\tau. \tag{26}$$

The space–time metric is given by (1). (c) Obtain $\Psi = A_4 + i\Phi$ from (24). The electromagnetic field can be found from (5) and (7).

5. EXAMPLE: CHARGED KERR-LIKE SOLUTIONS

The Kerr–Newman solution with $m^2 = e^2$ corresponds to the simplest complex solution of (25). We choose

$$2/(1 + \mathcal{E}) = 1 + m/R, \qquad \text{with } R^2 = x^2 + y^2 + (z - ia)^2, \tag{27}$$

where a and m are real constants and x, y, z Cartesian coordinates. In terms of oblate spheroidal coordinates r, θ, ϕ defined by

$$x + iy = [(r - m)^2 + a^2]^{1/2}\sin\theta\, e^{i\phi}, z = (r - m)\cos\theta,$$

the Euclidean 3-metric becomes

$$\gamma_{mn}dx^m dx^n = [(r - m)^2 + a^2\cos^2\theta]\{dr^2/[(r - m)^2 + a^2] + d\theta^2\} + [(r - m)^2 + a^2]\sin^2\theta\, d\phi^2.$$

Further, we find

$$R = r - m - ia\cos\theta,$$
$$f = [(r - m)^2 + a^2\cos^2\theta]/(r^2 + a^2\cos^2\theta),$$
$$\Psi = e^{i\alpha}\, m/(r - ia\cos\theta),$$

and, after a somewhat lengthy calculation,

$$\omega_m dx^m = \{[(2mr - m^2)a\sin^2\theta]/[(r - m)^2 + a^2\cos^2\theta]\}d\phi.$$

Putting everything together, we recover the charged Kerr metric with $m^2 = e^2$ in its usual form.[13]

As a natural generalization of (27), one may consider

$$\frac{2}{1 + \mathcal{E}} = 1 + \sum_{k=1}^{n}\frac{m_k}{R_k},$$

where $R_k^2 = (\mathbf{r} - \mathbf{c}_k)^2$, $\mathbf{r}$ is the Euclidean position vector, and $\mathbf{c}_k$ an arbitrary set of constant, complex vectors. The resulting metric will represent the field of a set of arbitrarily spinning, charged Kerr-like particles in neutral equilibrium. For the static analog of this solution, representing a set of Reissner–Nordström particles with $e_k = m_k$; see Ref. 5.

Note added in proof: The stationary extension of the Papapetrou–Majumdar solutions has since been obtained independently by Z. Perjés, Phys. Rev. Letters **27**, 1668 (1971).

---

* Work partially supported by the National Research Council of Canada.

1 H. Weyl, Ann. Physik **54**, 117 (1917).
2 S. D. Majumdar, Phys. Rev. **72**, 390 (1947).
3 A. Papapetrou, Proc. Roy. Irish Acad. **A51**, 191 (1947).
4 W. B. Bonnor, Proc. Phys. Soc. (London) **A66**, 1451 (1953); **A67**, 225 (1954).
5 J. L. Synge, *Relativity, The General Theory* (North-Holland, Amsterdam, 1960), p. 367.
6 A. Das, Proc. Roy. Soc. (London) **A267**, 1 (1962).
7 V. de la Cruz, J. E. Chase, and W. Israel, Phys. Rev. Letters **24**, 423 (1970).
8 R. Price, Phys. Rev. (to be published)
9 W. Israel and G. A. Wilson (to be published). Also J. B. Hartle and S. W. Hawking (to be published), R. Ruffini (to be published).

10 Greek indices run from 1 to 4, Latin indices from 1 to 3. Lowering and raising of Latin indices is always carried out with γ_{mn} and its inverse γ^{mn} unless specifically noted by a left superscript 4. Thus, if $F_{\mu\nu}$ is a given covariant tensor, we write $F^{ab} = \gamma^{am}\gamma^{bn}F_{mn}$ and ${}^{(4)}F^{ab} = g^{a\mu}g^{b\nu}F_{\mu\nu}$.

11 Cf., for the special case of axial symmetry, F. J. Ernst, Phys. Rev. **168**, 1415 (1968), where the idea of a complex potential is first introduced. We have been informed that B. K. Harrison (1968, unpublished) has cast the stationary electromagnetic vacuum equations into a form similar to that given in Secs. 2 and 3. See also B. K. Harrison, J. Math. Phys. **9**, 1744 (1968). A recent publication by Ernst, J. Math. Phys. **12**, 2395 (1971) treats the general stationary vacuum case.
12 Z. Perjes, J. Math. Phys. **11**, 3383 (1970).
13 See, e.g., B. Carter, Phys. Rev. **174**, 1559 (1968).

Constructing Examples in Relativistic Classical Particle Mechanics*

C. S. Shukre
Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania 15213

and

T. F. Jordan
Department of Physics, University of Minnesota, Duluth, Minnesota 55812

(Received 9 November 1971)

A new method is proposed for constructing examples in relativistic classical particle mechanics, starting with the Poincaré group and space reflections as transformations of position and velocity variables for two free particles and deriving equations for Lorentz covariant position variables for two interacting particles. The advantages are that the equations do not couple the position variables of the two particles, and solutions yield the particle world lines directly. Solutions are obtained for a special case.

I. INTRODUCTION

Requirements of relativistic invariance for interactions of classical particles have been developed, for example, as nonlinear partial differential equations for the force functions of Newtonian equations of motion.[1-5] Examples of physically acceptable solutions of these nonlinear equations have been difficult to find.[5,6]

Here we propose a new method of constructing examples in relativistic classical particle mechanics. A relativistic system of particles is a Poincaré group of transformations of position and velocity variables. We begin with Poincaré and space-reflection transformations of free-particle variables and look for functions of these variables to use as position variables for interacting particles. Lorentz covariance requirements in the form of nonlinear partial differential equations are the only nontrivial conditions these functions have to satisfy.

One advantage is that these equations do not couple the position variables of different particles, unlike the Lorentz-covariance conditions for Newtonian equations which do couple the forces of different particles.[1-5] Another advantage is that the solutions of these equations yield directly the particle world lines. This makes it possible to study their asymptotic properties.

In the next section we derive these equations for two particles. In Sec. III we obtain their solutions for a special case. Section IV contains conclusions and discussion.

II. EQUATIONS FOR RELATIVISTIC POSITIONS

We shall have occasion to consider only one particle at a time. Let its position be $\mathbf{y}$. After a Lorentz transformation with velocity $\tanh\epsilon$ in the kth direction, the jth component of the transformed position at time zero in the transformed frame is

$$y_j + \epsilon y_k \dot{y}_j$$

to first order in ϵ, where $\mathbf{y}$ and $\dot{\mathbf{y}}$ are the position and velocity at time zero in the original frame.[5,7,8] (Dots denote time derivatives.) We can use a bracket-generator symbol $[\ ,\mathbf{K}]$ for Lorentz transformations and $[\ ,H]$ for time derivatives,[9] and write

$$[y_j,K_k] = y_k\dot{y}_j = y_k[y_j,H], \qquad k,j = 1,2,3. \tag{II. 1}$$

The first-order part of the similarly transformed velocity is[5,7-9]

$$[\dot{y}_j,K_k] = y_k\ddot{y}_j + \dot{y}_k\dot{y}_j - \delta_{kj}. \tag{II. 2}$$

We shall use a Poincaré group of transformations defined for two free particles with positions $\mathbf{x}^{(n)}(t)$ and velocities $\mathbf{v}^{(n)}(t)$. Thus

$$\begin{aligned} &\mathbf{x}^{(n)}(t) = \mathbf{x}^{(n)} + \mathbf{v}^{(n)}t, \qquad \mathbf{v}^{(n)}(t) = \mathbf{v}^{(n)}, \\ &[x_j^{(n)},K_i] = x_i^{(n)}v_j^{(n)}, \qquad [v_j^{(n)},K_i] = v_i^{(n)}v_j^{(n)} - \delta_{ij}. \end{aligned} \tag{II. 3}$$

We also use a space-reflection transformation under which as usual $\mathbf{x}^{(n)}$ and $\mathbf{v}^{(n)}$ are odd. Then we have Poincaré and space reflection transformations of the correct number of independent variables for two interacting particles. It remains to find position variables for the interacting particles.

For the position of each of the interacting particles, we write[10]

$$\mathbf{y} = \mathbf{x}^{(1)} + f\mathbf{v}^{(1)} + g\mathbf{v}^{(2)} + h\mathbf{x}, \tag{II. 4}$$

where $\mathbf{x} = \mathbf{x}^{(1)} - \mathbf{x}^{(2)}$. We assume the usual transformations of positions for space translations, rotations, and reflections. This means that f, g, and h are scalars invariant under space translations. We can take them to be functions of the independent variables

$$\begin{aligned} &\alpha_1 = \mathbf{x}\cdot\mathbf{v}^{(1)}, \quad \alpha_2 = \mathbf{x}\cdot\mathbf{v}^{(2)}, \quad \alpha_3 = \mathbf{x}^2, \\ &\alpha_4 = 1 - \mathbf{v}^{(1)2}, \quad \alpha_5 = 1 - \mathbf{v}^{(2)2}, \\ &\alpha_6 = 1 - \mathbf{v}^{(1)}\cdot\mathbf{v}^{(2)}. \end{aligned} \tag{II. 5}$$

The whole problem now is to satisfy the Lorentz-transformation condition (II. 1) for the position (II. 4). This is

$$[y_j,K_i] = y_i\dot{y}_j$$

or

$$\begin{aligned} &[x_j^{(1)},K_i] + v_j^{(1)}[f,K_i] + v_j^{(2)}[g,K_i] + x_j[h,K_i] \\ &\quad + f[v_j^{(1)},K_i] + g[v_j^{(2)},K_i] + h[x_j,K_i] \\ &= (x_i^{(1)} + fv_i^{(1)} + gv_i^{(2)} + hx_i)[v_j^{(1)} + (\dot{f} + h)v_j^{(1)} \\ &\quad + (\dot{g} - h)v_j^{(2)} + \dot{h}x_j], \end{aligned} \tag{II. 6}$$

where $[f,K_i]$ can be calculated as $[f,K_i] = \sum_j (\partial f/\partial\alpha_j)[\alpha_j,K_i]$ with $[\alpha_j,K_i]$ calculated using the bracket relations of $\mathbf{x}$ and $\mathbf{v}^{(n)}$. We get

$$[f,K_i] = x_i^{(1)}\dot{f} + v_i^{(1)}D_1f + v_i^{(2)}D_2f + x_iD_3f, \tag{II. 7}$$

where

$$D_1 = \alpha_1\frac{\partial}{\partial\alpha_1} + 2\alpha_4\frac{\partial}{\partial\alpha_4} + \alpha_6\frac{\partial}{\partial\alpha_6},$$

$$D_2 = \alpha_2\frac{\partial}{\partial\alpha_2} + 2\alpha_5\frac{\partial}{\partial\alpha_5} + \alpha_6\frac{\partial}{\partial\alpha_6},$$

$$D_3 = 2\alpha_2 \frac{\partial}{\partial \alpha_3} - \alpha_6 \frac{\partial}{\partial \alpha_1} - \alpha_5 \frac{\partial}{\partial \alpha_2},$$

$$D_4 = 2\alpha_1 \frac{\partial}{\partial \alpha_3} - \alpha_4 \frac{\partial}{\partial \alpha_1} - \alpha_6 \frac{\partial}{\partial \alpha_2} = D_3 + \frac{d}{dt}.$$

Then substituting (II. 7) and the bracket relations for $\mathbf{x}$ and $\mathbf{v}^{(n)}$ in (II. 6) we obtain a tensor equation

$$T_{ij} = 0, \tag{II. 8}$$

where T_{ij} consists of ten terms with scalars multiplying $\delta_{ij}, x_i x_j, x_i v_j^{(1)}, x_i v_j^{(2)}$, etc.

We can extract equivalent scalar equations from the tensor equation (II. 8) by choosing a convenient set of three linearly independent vectors $\mathbf{a}^{(m)}$, $m = 1, 2, 3$. Then we obtain nine scalar equations

$$\sum_{i,j} a_i^{(m)} T_{ij} a_j^{(m')} = S^{mm'} = 0, \quad m, m' = 1, 2, 3. \tag{II. 9}$$

One choice for $\mathbf{a}^{(m)}$ is $\mathbf{a}^{(1)} = \tilde{\mathbf{v}}^{(1)}$, $\mathbf{a}^{(2)} = \tilde{\mathbf{v}}^{(2)}$, and $\mathbf{a}^{(3)} = \tilde{\mathbf{x}}$.[10] This has the advantage of giving equations with the least number of terms.

They are

$$\begin{aligned}
&D_1 f - f(D_4 - D_3)f - (f + g)(C_4/D) - hf = 0,\\
&D_1 g - f(D_4 - D_3)g - (f + g)(C_6/D) + hf = 0,\\
&D_1 h - f(D_4 - D_3)h - (f + g)C_1/D = 0,\\
&D_2 f - g(D_4 - D_3)f - (f + g)(C_6/D) - g(h + 1) = 0,\\
&D_2 g - g(D_4 - D_3)g - (f + g)(C_5/D) + g(h + 1) = 0,\\
&D_2 h - g(D_4 - D_3)h - (f + g)C_2/D = 0,\\
&D_3 f - h(D_4 - D_3)f - (f + g)(C_1/D) - h(h + 1) = 0,\\
&D_3 g - h(D_4 - D_3)g - (f + g)(C_2/D) + h(h + 1) = 0,\\
&D_3 h - h(D_4 - D_3)h - (f + g)C_3/D = 0,
\end{aligned} \tag{II. 10}$$

where

$$D = \det \begin{bmatrix} 1 - \alpha_4 & 1 - \alpha_6 & \alpha_1 \\ 1 - \alpha_6 & 1 - \alpha_5 & \alpha_2 \\ \alpha_1 & \alpha_2 & \alpha_3 \end{bmatrix}$$

and C_i are the cofactors of α_i or $1 - \alpha_i$ in D.

To solve these equations, we first bring the differential operators D_i to a simple form by making a transformation of independent variables.

If

$$\begin{aligned}
&\beta_1 = \sqrt{\alpha_4}, \quad \beta_2 = \sqrt{\alpha_5}, \quad \beta_5 = \alpha_6/\sqrt{\alpha_4 \alpha_5},\\
&\beta_3 = [(\alpha_1\alpha_6 - \alpha_2\alpha_4)/(\alpha_4\alpha_5 - \alpha_6^2)]\sqrt{\alpha_5},\\
&\beta_4 = [(\alpha_2\alpha_6 - \alpha_1\alpha_5)/(\alpha_4\alpha_5 - \alpha_6^2)]\sqrt{\alpha_4},
\end{aligned} \tag{II. 11}$$

and

$$\beta_6 = \alpha_3 + [(\alpha_1^2\alpha_5 + \alpha_2^2\alpha_4 - 2\alpha_1\alpha_2\alpha_6)/(\alpha_4\alpha_5 - \alpha_6^2)],$$

then

$$D_1 = \beta_1 \frac{\partial}{\partial \beta_1}, \quad D_2 = \beta_2 \frac{\partial}{\partial \beta_2}, \quad D_3 = \beta_1 \frac{\partial}{\partial \beta_3}, \quad \text{and} \quad D_4 = \beta_2 \frac{\partial}{\partial \beta_4}. \tag{II. 12}$$

We now face the second difficulty in the solution of Eq. (II. 10). This is the presence of the determinant D, which is clearly difficult to integrate. The simplest way out is to ignore this term by imposing the condition $f + g = 0$. In the next section we shall obtain solutions for this case.

The determinant is just $(\mathbf{v}^{(1)} \times \mathbf{v}^{(2)} \cdot \mathbf{x})^2$. Its presence is a direct consequence of the choice of vectors $\mathbf{a}^{(m)}$, i.e., $\tilde{\mathbf{v}}^{(1)}, \tilde{\mathbf{v}}^{(2)}$, and $\tilde{\mathbf{x}}$. This can be remedied by making another choice, e.g., $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}$, and $\mathbf{x}$, which is equivalent to expanding $\mathbf{v}$ in terms of $\tilde{\mathbf{v}}^{(n)}$ and $\tilde{\mathbf{x}}$ instead of $\mathbf{v}^{(n)}$ and $\mathbf{x}$ as in (II. 4).[10] This trick does give equations which have no determinant, but the corresponding differential operators like D_i are more complicated. It is quite difficult to bring these operators to a simple form by a transformation analogous to (II. 11). We have so far not succeeded in finding such a transformation, and the equations remain highly intractable.

III. SOLUTIONS FOR THE SPECIAL CASE, $f + g = 0$

After imposing the condition $f + g = 0$, the equations for f and g coalesce into the following six equations:

$$(D_1 + D_2)g = -g, \tag{III. 1}$$

$$(D_1 - D_2)g + 2g(D_4 - D_3)g - g(2h + 1) = 0, \tag{III. 2}$$

$$D_3 g - h(D_4 - D_3)g + h(h + 1) = 0, \tag{III. 3}$$

$$(D_1 + D_2)h = 0, \tag{III. 4}$$

$$(D_1 - D_2)h + 2g(D_4 - D_3)h = 0, \tag{III. 5}$$

$$D_3 h - h(D_4 - D_3)h = 0. \tag{III. 6}$$

We now make another transformation of independent variables (with hindsight):

$$\begin{aligned}
&\gamma_1 = 2\beta_2, \quad \gamma_2 = \beta_2/\beta_1, \quad \gamma_3 = (\beta_2/\beta_1)\beta_4,\\
&\gamma_4 = (\beta_2/\beta_1)\beta_4 + \beta_3, \quad \gamma_5 = \beta_5, \quad \gamma_6 = \beta_6.
\end{aligned} \tag{III. 7}$$

Then

$$D_1 + D_2 = \gamma_1 \frac{\partial}{\partial \gamma_1},$$

$$D_1 - D_2 = \gamma_1 \frac{\partial}{\partial \gamma_1} + 2\gamma_2 \frac{\partial}{\partial \gamma_2} + 2\gamma_3 \left(\frac{\partial}{\partial \gamma_3} + \frac{\partial}{\partial \gamma_4} \right),$$

$$D_3 = \frac{\gamma_1}{2} \frac{\partial}{\partial \gamma_4},$$

and

$$D_4 - D_3 = \frac{\gamma_1}{2} \frac{\partial}{\partial \gamma_3}.$$

The solution of (III. 1) and (III. 4) is trivial and gives

$$\begin{aligned}
&g(\gamma_1, \gamma_2 \cdots \gamma_6) = (1/\gamma_1)\bar{g}(\gamma_2 \cdots \gamma_6)\\
&h(\gamma_1, \gamma_2 \cdots \gamma_6) = \bar{h}(\gamma_2 \cdots \gamma_6)
\end{aligned}$$

The remaining four equations take a simple form after a change of dependent variables from $\bar{g}$ to ϕ:

$$\bar{g} = [\phi + 2(\gamma_3 - \gamma_4)]\bar{h}. \tag{III. 8}$$

Then the equations are

$$2\left(\gamma_2 \frac{\partial}{\partial \gamma_2} + \gamma_3 \frac{\partial}{\partial \gamma_3} + \gamma_4 \frac{\partial}{\partial \gamma_4}\right)\left[\frac{\phi}{\bar{h}}\right] - \phi \frac{\partial}{\partial \gamma_3}\left[\frac{\phi}{\bar{h}}\right] = 0 \tag{III. 9}$$

(III. 10)

and

$$\frac{\partial}{\partial \gamma_4}\left[\frac{\phi}{\bar{h}}\right] - \bar{h} \frac{\partial}{\partial \gamma_3}\left[\frac{\phi}{\bar{h}}\right] = 0. \tag{III. 11}$$

(III. 12)

Two cases arise according to whether $\partial\phi/\partial\gamma_3$ is zero or not.

Case I: $\partial\phi/\partial\gamma_3 \neq 0$: In this case $\bar{h}$ can be eliminated via (III. 11). Then (III. 12) becomes

$$\frac{\partial^2\phi}{\partial\gamma_4^2}\left(\frac{\partial\phi}{\partial\gamma_3}\right)^2 - 2\frac{\partial\phi}{\partial\gamma_4}\frac{\partial\phi}{\partial\gamma_3}\frac{\partial^2\phi}{\partial\gamma_4\partial\gamma_3} + \left(\frac{\partial\phi}{\partial\gamma_4}\right)^2\frac{\partial^2\phi}{\partial\gamma_3^2} = 0. \tag{III. 13}$$

After we use $\bar{h}$ in terms of ϕ from (III. 11), in (III. 10) we use (III. 9) also and get

$$\left(\frac{\partial\phi}{\partial\gamma_4} + \frac{\partial\phi}{\partial\gamma_3}\right)\frac{\partial\phi}{\partial\gamma_4}\left(\frac{\partial\phi}{\partial\gamma_3}\right)^2 = 0. \tag{III. 14}$$

Then (III. 9), (III. 13), and (III. 14) can be solved completely to give

$$\phi(\gamma_2 \cdots \gamma_6) = \{2\phi_0(\gamma_5, \gamma_6)/[\gamma_2 - \phi_0(\gamma_5, \gamma_6)]\}(\gamma_3 - \gamma_4), \tag{III. 15a}$$

which means that

$$\bar{h}(\gamma_2 \cdots \gamma_6) = -1 \tag{III. 15b}$$

$$\bar{g}(\gamma_2 \cdots \gamma_6) = 2\gamma_2(\gamma_4 - \gamma_3)/[\gamma_2 - \phi_0(\gamma_5, \gamma_6)]. \tag{III. 15c}$$

Case II: $\partial\phi/\partial\gamma_3 = 0$: In this case

$$\phi(\gamma_2 \cdots \gamma_6) = \phi_0(\gamma_5, \gamma_6)$$

and if we define

$$\delta_2 = \gamma_2, \quad \delta_3 = (2\gamma_3 - \phi_0)/2\gamma_2, \quad \text{and} \quad \delta_4 = \gamma_4/\gamma_2,$$

Eq. (III. 9)–(III. 12) become

$$\frac{\partial\bar{h}}{\partial\delta_2} = 0 \tag{III. 16a}$$

and

$$\frac{\partial\bar{h}}{\partial\delta_4} = \bar{h}\frac{\partial\bar{h}}{\partial\delta_3}. \tag{III. 16b}$$

Then

$$\bar{h}(\gamma_2 \cdots \gamma_6) = \hat{h}(\delta_3, \delta_4, \gamma_5\gamma_6) \tag{III. 17a}$$

and

$$\bar{g}(\gamma_2 \cdots \gamma_6) = (\phi_0 + 2\gamma_3 - 2\gamma_4)\bar{h} = 2\delta_2(\phi_0 + \delta_3 - \delta_4)\hat{h}. \tag{III. 17b}$$

Some solutions can be obtained in this case by trial and error, e.g.,

$$\hat{h} = [A(\gamma_5, \gamma_6) + \delta_3]/[B(\gamma_5, \gamma_6) - \delta_4]. \tag{III. 18}$$

One can try to get general solutions by expanding $\hat{h}$ in a power series,

$$\hat{h} = \sum_{j=0}^{\infty} a_j(\delta_3, \gamma_5, \gamma_6)(\delta_4)^j. \tag{III. 19}$$

Then, substituting in (III. 16b), one gets the following recursion relation for a_j (prime denotes $\partial/\partial\delta_3$):

$$\sum_{j=0}^{n} a_j a'_{n-j} = (n+1)a_{n+1}. \tag{III. 20}$$

Once a_0 is specified, all other a_j can be obtained in terms of it via (III. 20). However, summing the series (or knowing if it converges) is yet another problem.

If we assume $a_0 = e^{\delta_3}$ and use (III. 20), it turns out that $e^{-\delta_3}\hat{h}$ depends only on $\delta_4 e^{\delta_3}$. Using this one can integrate (III. 16b). The solution is given by the implicit equation

$$\hat{h} = ce^{\delta_3 + \delta_4\hat{h}}, \tag{III. 21}$$

where the integration constant c depends only on γ_5 and γ_6.

IV. CONCLUSIONS AND DISCUSSION

We have two classes of solutions. The first is (III. 15b), (III. 15c). Now γ_3 is the variable which carries time dependence. From (III. 15) it is clear that the position function (II. 4) is linear in γ_3. Therefore it is linear in time. Hence the second time derivative of the position, the acceleration, is zero. Thus this set of solutions describes only free particle motion.

In the second case we know two sets of solutions, (III. 18) and (III. 21). For (III. 18) the position (II. 3) looks quadratic in δ_3 (which carries time dependence), but due to cancellations it is really linear in time so that again we have only free particle motion.

The solution (III. 21) gives nonzero acceleration. But as $t \to \pm\infty$ (i.e., $\delta_3 \to \pm\infty$)

$$h \xrightarrow[t\to\infty]{} t$$

$$\xrightarrow[t\to-\infty]{} 0 \text{ or } |t|^n, \quad n \geq 1.$$

Hence the position function goes quadratically with time as $t \to +\infty$ and thus does not reduce to free motion asymptotically. As $t \to -\infty$ the position could behave either as a free-particle position or as $|t|^n$, $n \geq 2$. The latter is undesirable.

Thus of the two possible subcases of the special case when $f + g = 0$, one is completely solved. These solutions describe only freely moving particles. In the second subcase we have obtained only two solutions. These either describe free-particle motion or do not satisfy the asymptotic condition.

It seems possible that even for the special case $f + g = 0$ there would be some solutions which are completely physical.

---

* Supported in part by the U.S. Atomic Energy Commission under Contract No. AT-(30-1)-3829.

1 D. G. Currie, Phys. Rev. **142**, 817 (1966).
2 D. G. Currie and T. F. Jordan, Phys. Rev. Letters **16**, 1210 (1966).
3 R. N. Hill, J. Math. Phys. **8**, 201 (1967).
4 D. G. Currie and T. F. Jordan, Phys. Rev. **167**, 1178 (1968).
5 D. G. Currie and T. F. Jordan, in *Lectures in Theoretical Physics*, Vol. XA (Proceedings of the Tenth Boulder Summer Institute for Theoretical Physics), edited by A. O. Barut and W. E. Brittin (Gordon and Breach, New York, 1968), p. 91.

6 J. G. Wray, Phys. Rev. D 1, 2212 (1970).
7 D. G. Currie, T. F. Jordan, and E. C. G. Sudarshan, Rev. Mod. Phys. 35, 350 (1963).
8 D. G. Currie, J. Math. Phys. 4, 1470 (1963).
9 T. F. Jordan, Phys. Rev. 166, 1308 (1968).
10 Instead of expanding in terms of $\mathbf{x}$ and $\mathbf{v}^{(n)}$ in (II. 4) one could expand in terms of $\tilde{\mathbf{x}} = (\mathbf{v}^{(1)} \times \mathbf{v}^{(2)})(\mathbf{v}^{(1)} \times \mathbf{v}^{(2)}\cdot\mathbf{x})^{-1}$, $\tilde{\mathbf{v}}^{(1)} = (\mathbf{v}^{(2)} \times \mathbf{x})(\mathbf{v}^{(1)} \times \mathbf{v}^{(2)}\cdot\mathbf{x})^{-1}$ and $\tilde{\mathbf{v}}^{(2)} = (\mathbf{x} \times \mathbf{v}^{(1)})(\mathbf{v}^{(1)} \times \mathbf{v}^{(2)}\cdot\mathbf{x})^{-1}$. Then we would have

$$\mathbf{y} = \mathbf{x}^{(1)} + F\tilde{\mathbf{v}}^{(1)} + G\tilde{\mathbf{v}}^{(2)} + H\tilde{\mathbf{x}}.$$

These F, G, and H are linear combinations of f, g, and h with coefficients which depend on the α_i.

Identities Involving the Coefficients in the Asymptotic Expansion of the Outgoing Scattering Solution of the Schrödinger Equation*

Donald Gelman
Department of Physics, C. W. Post College, Greenvale, New York 11548
(Received 27 September 1971; Revised Manuscript Received 24 January 1972)

Scattering by a short-range potential is described by the wavefunction ψ having the asymptotic form $\psi \to e^{i\mathbf{k}\cdot\mathbf{r}} + \mathfrak{F}e^{ikr}, r \to \infty$, where $\mathfrak{F}$ has the series representation $\mathfrak{F} = f(\mathbf{k}, k\hat{r})/r + g(\mathbf{k}, k\hat{r})/r^2 + \cdots$. When only the first term in this series is taken into account, the requirement that the particle current be conserved leads to the optical theorem. By taking into account all of the remaining terms, we obtain a sequence of identities involving the coefficients $f, g, \cdots$. These identities are formally analogous to the optical theorem, which is seen to correspond to the first identity in the sequence.

1. INTRODUCTION

Scattering by a short-range potential is described by the wavefunction ψ having the asymptotic form

$$\psi \to e^{i\mathbf{k}\cdot\mathbf{r}} + \mathfrak{F}e^{ikr}, \quad r \to \infty, \tag{1.1}$$

where

$$\mathfrak{F} = f(\mathbf{k}, k\hat{r})r^{-1} + g(\mathbf{k}, k\hat{r})r^{-2} + O(r^{-3}). \tag{1.2}$$

Although the scattering amplitude $f(\mathbf{k}, k\hat{r})$ is the only coefficient in the expansion of $\mathfrak{F}$ with physical significance,[1] neglect of the higher-order terms may not always be justified. Such a situation arises, for example, in the analysis of the regularization $\Delta Z \equiv \mathrm{Tr}\,[\exp(-\beta H) - \exp(-\beta H^0)]$ of the divergent partition function $\mathrm{Tr}\exp(-\beta H)$.

In particular, the continuum contribution to ΔZ is given by[2]

$$\lim_{r\to\infty} (2\pi)^{-3} \int d\mathbf{k}\, e^{-\beta E_k}(2k)^{-1} \oint_{\sigma(r)} da\; C,$$

$$C = (\dot{\psi}\psi^{*\prime} - \psi^*\dot{\psi}') - (\psi \to \psi^0),$$

where $E_k = (\hbar^2/2m)\mathbf{k}^2$ and $\psi^0 = e^{i\mathbf{k}\cdot\mathbf{r}}$; $\sigma(r)$ denotes the surface of a sphere of radius r centered at the origin, and the dot and prime denote differentiation with respect to k and r, respectively. Since the area of $\sigma(r)$ increases as r^2, the asymptotic expansion of C must be correct through terms of order r^{-2}. Such an expansion is clearly not obtained if only the first term on the right-hand side of (1.2) is used.

When only the first term in the expansion of $\mathfrak{F}$ is taken into account, the requirement that the particle current be conserved leads to the optical theorem[3]

$$\int d\hat{r}\, |f(\mathbf{k}, k\hat{r})|^2 = (4\pi/k)\,\mathrm{Im} f(\mathbf{k}, \mathbf{k}), \tag{1.3}$$

the symbol $\int d\hat{r}$ denoting integration over all directions $\hat{r}$. The purpose of this paper is to determine the consequences of particle current conservation when account is taken of the remaining terms in the expansion of $\mathfrak{F}$.

The analysis is confined to the system consisting of a nonrelativistic, spinless particle of mass m moving in a short-range potential $V(\mathbf{r})$. There is no pretense of mathematical rigor. In Sec. 2, the fundamental identity (2.8) involving the surface integrals X and X^s is derived. Asymptotic expansions for X^s and X are obtained in Secs. 3 and 4, respectively. These are used in Sec. 5 to arrive at the identities[4] (5.8) and (5.9) which constitute the consequences referred to at the end of the preceding paragraph.

2. THE FUNDAMENTAL IDENTITY

The wavefunction $\psi = \psi(\mathbf{k}, \mathbf{r})$ having the asymptotic form given by (1.1) and (1.2) is a solution of the time-independent Schrödinger equation

$$[(-\hbar^2/2m)(\nabla^2 + k^2) + V(\mathbf{r})]\psi = 0. \tag{2.1}$$

Since the potential $V(\mathbf{r})$ is real, it follows from Eq. (2.1) that the quantity

$$\mathbf{J} \equiv (\psi^*\nabla\psi) - \mathrm{c.c.} \tag{2.2}$$

satisfies the condition

$$\nabla\cdot\mathbf{J} = 0. \tag{2.3}$$

Integrating both sides of Eq. (2.3) over the volume $\tau(R)$, with $\tau(R)$ a sphere of radius R centered on the origin $\mathbf{r} = 0$, and then using the divergence theorem, one obtains

$$\oint_{\sigma(R)} da\hat{r}\cdot\mathbf{J} = 0, \tag{2.4}$$

where $\sigma(R)$ denotes the surface of $\tau(R)$.

Let ψ be written in the form $\psi = \psi^0 + \psi^s$, with $\psi^0 = \exp(i\mathbf{k}\cdot\mathbf{r})$ denoting the free-particle wavefunction and $\psi^s = \psi^s(\mathbf{k}, \mathbf{r})$ denoting the scattered wave part of ψ. Substitution of this decomposition of ψ into Eq. (2.2) then yields for the integrand of Eq. (2.4) the expression

$$\hat{r}\cdot\mathbf{J} = \hat{r}\cdot\mathbf{J}^0 + W + W^s, \tag{2.5}$$

where $\mathbf{J}^0$ is the free-particle counterpart of $\mathbf{J}$ and where

$$W \equiv \left[\psi^{0*}\left(\frac{\partial\psi^s}{\partial r}\right) - \psi^s\left(\frac{\partial\psi^{0*}}{\partial r}\right)\right] - \mathrm{c.c.}, \tag{2.6}$$

6 J. G. Wray, Phys. Rev. D 1, 2212 (1970).
7 D. G. Currie, T. F. Jordan, and E. C. G. Sudarshan, Rev. Mod. Phys. 35, 350 (1963).
8 D. G. Currie, J. Math. Phys. 4, 1470 (1963).
9 T. F. Jordan, Phys. Rev. 166, 1308 (1968).
10 Instead of expanding in terms of $\mathbf{x}$ and $\mathbf{v}^{(n)}$ in (II. 4) one could expand in terms of $\tilde{\mathbf{x}} = (\mathbf{v}^{(1)} \times \mathbf{v}^{(2)})(\mathbf{v}^{(1)} \times \mathbf{v}^{(2)}\cdot\mathbf{x})^{-1}$, $\tilde{\mathbf{v}}^{(1)} = (\mathbf{v}^{(2)} \times \mathbf{x})(\mathbf{v}^{(1)} \times \mathbf{v}^{(2)}\cdot\mathbf{x})^{-1}$ and $\tilde{\mathbf{v}}^{(2)} = (\mathbf{x} \times \mathbf{v}^{(1)})(\mathbf{v}^{(1)} \times \mathbf{v}^{(2)}\cdot\mathbf{x})^{-1}$. Then we would have

$$\mathbf{y} = \mathbf{x}^{(1)} + F\tilde{\mathbf{v}}^{(1)} + G\tilde{\mathbf{v}}^{(2)} + H\tilde{\mathbf{x}}.$$

These F, G, and H are linear combinations of f, g, and h with coefficients which depend on the α_i.

Identities Involving the Coefficients in the Asymptotic Expansion of the Outgoing Scattering Solution of the Schrödinger Equation*

Donald Gelman
Department of Physics, C. W. Post College, Greenvale, New York 11548
(Received 27 September 1971; Revised Manuscript Received 24 January 1972)

Scattering by a short-range potential is described by the wavefunction ψ having the asymptotic form $\psi \to e^{i\mathbf{k}\cdot\mathbf{r}} + \mathfrak{F}e^{ikr}, r \to \infty$, where $\mathfrak{F}$ has the series representation $\mathfrak{F} = f(\mathbf{k}, k\hat{r})/r + g(\mathbf{k}, k\hat{r})/r^2 + \cdots$. When only the first term in this series is taken into account, the requirement that the particle current be conserved leads to the optical theorem. By taking into account all of the remaining terms, we obtain a sequence of identities involving the coefficients $f, g, \cdots$. These identities are formally analogous to the optical theorem, which is seen to correspond to the first identity in the sequence.

1. INTRODUCTION

Scattering by a short-range potential is described by the wavefunction ψ having the asymptotic form

$$\psi \to e^{i\mathbf{k}\cdot\mathbf{r}} + \mathfrak{F}e^{ikr}, \quad r \to \infty, \tag{1.1}$$

where

$$\mathfrak{F} = f(\mathbf{k}, k\hat{r})r^{-1} + g(\mathbf{k}, k\hat{r})r^{-2} + O(r^{-3}). \tag{1.2}$$

Although the scattering amplitude $f(\mathbf{k}, k\hat{r})$ is the only coefficient in the expansion of $\mathfrak{F}$ with physical significance,[1] neglect of the higher-order terms may not always be justified. Such a situation arises, for example, in the analysis of the regularization $\Delta Z \equiv \mathrm{Tr}\,[\exp(-\beta H) - \exp(-\beta H^0)]$ of the divergent partition function $\mathrm{Tr}\exp(-\beta H)$.

In particular, the continuum contribution to ΔZ is given by[2]

$$\lim_{r\to\infty} (2\pi)^{-3} \int d\mathbf{k}\, e^{-\beta E_k}(2k)^{-1} \oint_{\sigma(r)} da\; C,$$

$$C = (\dot{\psi}\psi^{*\prime} - \psi^*\dot{\psi}') - (\psi \to \psi^0),$$

where $E_k = (\hbar^2/2m)\mathbf{k}^2$ and $\psi^0 = e^{i\mathbf{k}\cdot\mathbf{r}}$; $\sigma(r)$ denotes the surface of a sphere of radius r centered at the origin, and the dot and prime denote differentiation with respect to k and r, respectively. Since the area of $\sigma(r)$ increases as r^2, the asymptotic expansion of C must be correct through terms of order r^{-2}. Such an expansion is clearly not obtained if only the first term on the right-hand side of (1.2) is used.

When only the first term in the expansion of $\mathfrak{F}$ is taken into account, the requirement that the particle current be conserved leads to the optical theorem[3]

$$\int d\hat{r}\, |f(\mathbf{k}, k\hat{r})|^2 = (4\pi/k)\,\mathrm{Im} f(\mathbf{k}, \mathbf{k}), \tag{1.3}$$

the symbol $\int d\hat{r}$ denoting integration over all directions $\hat{r}$. The purpose of this paper is to determine the consequences of particle current conservation when account is taken of the remaining terms in the expansion of $\mathfrak{F}$.

The analysis is confined to the system consisting of a nonrelativistic, spinless particle of mass m moving in a short-range potential $V(\mathbf{r})$. There is no pretense of mathematical rigor. In Sec. 2, the fundamental identity (2.8) involving the surface integrals X and X^s is derived. Asymptotic expansions for X^s and X are obtained in Secs. 3 and 4, respectively. These are used in Sec. 5 to arrive at the identities[4] (5.8) and (5.9) which constitute the consequences referred to at the end of the preceding paragraph.

2. THE FUNDAMENTAL IDENTITY

The wavefunction $\psi = \psi(\mathbf{k}, \mathbf{r})$ having the asymptotic form given by (1.1) and (1.2) is a solution of the time-independent Schrödinger equation

$$[(-\hbar^2/2m)(\nabla^2 + k^2) + V(\mathbf{r})]\psi = 0. \tag{2.1}$$

Since the potential $V(\mathbf{r})$ is real, it follows from Eq. (2.1) that the quantity

$$\mathbf{J} \equiv (\psi^*\nabla\psi) - \mathrm{c.c.} \tag{2.2}$$

satisfies the condition

$$\nabla\cdot\mathbf{J} = 0. \tag{2.3}$$

Integrating both sides of Eq. (2.3) over the volume $\tau(R)$, with $\tau(R)$ a sphere of radius R centered on the origin $\mathbf{r} = 0$, and then using the divergence theorem, one obtains

$$\oint_{\sigma(R)} da\hat{r}\cdot\mathbf{J} = 0, \tag{2.4}$$

where $\sigma(R)$ denotes the surface of $\tau(R)$.

Let ψ be written in the form $\psi = \psi^0 + \psi^s$, with $\psi^0 = \exp(i\mathbf{k}\cdot\mathbf{r})$ denoting the free-particle wavefunction and $\psi^s = \psi^s(\mathbf{k}, \mathbf{r})$ denoting the scattered wave part of ψ. Substitution of this decomposition of ψ into Eq. (2.2) then yields for the integrand of Eq. (2.4) the expression

$$\hat{r}\cdot\mathbf{J} = \hat{r}\cdot\mathbf{J}^0 + W + W^s, \tag{2.5}$$

where $\mathbf{J}^0$ is the free-particle counterpart of $\mathbf{J}$ and where

$$W \equiv \left[\psi^{0*}\left(\frac{\partial\psi^s}{\partial r}\right) - \psi^s\left(\frac{\partial\psi^{0*}}{\partial r}\right)\right] - \mathrm{c.c.}, \tag{2.6}$$

$$W^s \equiv \left[\psi^{s*}\left(\frac{\partial\psi^s}{\partial r}\right)\right] - \text{c.c.} \tag{2.7}$$

Since Eq. (2.4) is valid for $\mathbf{J}^0$ as well as for $\mathbf{J}$, it follows from Eq. (2.5) that

$$\oint_{\sigma(r)} da(W + W^s) \equiv X + X^s = 0. \tag{2.8}$$

(It should be noted that this equation is valid for all values of r, and is therefore an identity in r.)

3. ASYMPTOTIC FORM OF X^s

The quantity $\mathfrak{F}$ defined by (1.1) is assumed to have the series representation

$$\mathfrak{F} = \sum_{n=0}^{\infty} f_n r^{-(n+1)}, \tag{3.1}$$

with $f_n = f_n(\mathbf{k}, k\hat{r})$, and with $f_0 \equiv f$ and $f_1 \equiv g$. Using Eq. (3.1) and the relation[5]

$$\psi^s \to \mathfrak{F} e^{ikr}, \quad r \to \infty, \tag{3.2}$$

one can easily show that

$$\frac{\partial\psi^s}{\partial r} \to \mathfrak{F}_1 e^{ikr}, \tag{3.3}$$

where

$$\mathfrak{F}_1 \equiv ik\mathfrak{F} + \frac{\partial\mathfrak{F}}{\partial r} \tag{3.4a}$$

$$= \sum_{n=0}^{\infty} \gamma_n r^{-(n+1)}, \tag{3.4b}$$

$$\gamma_n \equiv ikf_n - nf_{n-1}. \tag{3.4c}$$

Substituting (3.2) and (3.3) into Eq. (2.7), and then using Eqs. (3.1) and (3.4b), we find that W^s has the asymptotic form

$$W^s \to \mathfrak{F}_2 - \text{c.c.}, \tag{3.5}$$

where

$$\mathfrak{F}_2 \equiv \mathfrak{F}^*\mathfrak{F}_1 = \sum_{n=0}^{\infty} c_n r^{-(n+2)}, \tag{3.6a}$$

$$c_n \equiv \sum_{j=0}^{n} f_j^* \gamma_{n-j}. \tag{3.6b}$$

The asymptotic form of the surface integral

$$X^s \equiv \oint_{\sigma(r)} da\, W^s$$

is therefore given by

$$X^s \to \sum_{n=0}^{\infty} \beta_n r^{-n}, \tag{3.7a}$$

where

$$\beta_n \equiv \int d\hat{r}\,(c_n - c_n^*). \tag{3.7b}$$

4. ASYMPTOTIC FORM OF X

The asymptotic form of W, obtained by substituting (3.2) and (3.3) into Eq. (2.6), is

$$W \to \mathfrak{F}_3 e^{ikr(1-\mu)} - \text{c.c.}, \tag{4.1}$$

$$\mathfrak{F}_3 \equiv \mathfrak{F}_1 + ik\mu\mathfrak{F} = \sum_{n=0}^{\infty} a_n r^{-(n+1)}, \tag{4.2a}$$

$$a_n \equiv ik(1+\mu)f_n - nf_{n-1}, \tag{4.2b}$$

where $\mu \equiv \hat{k}\cdot\hat{r}$ and where Eqs. (3.1) and (3.4) have been used. Introducing the usual spherical polar coordinates, with $\hat{k}$ as polar axis, we have from (4.1) and Eq. (2.8) that

$$X \equiv \oint_{\sigma(r)} daW = r^2 \int_0^{2\pi} d\phi \int_{-1}^{1} d\mu\, W \tag{4.3a}$$

$$\to \int_0^{2\pi} d\phi\, e^{ikr} r^2 I - \text{c.c.}, \tag{4.3b}$$

where

$$I \equiv \int_{-1}^{1} d\mu\, \mathfrak{F}_3 e^{-ikr\mu}. \tag{4.3c}$$

Substitution of Eq. (4.2a) into (4.3c) and interchange of the order of summation and integration then yields for I the series representation

$$I = \sum_{n=0}^{\infty} J[a_n] r^{-(n+1)} \tag{4.4}$$

where the functional J is defined by

$$J[w] \equiv \int_{-1}^{1} d\mu e^{-ikr\mu} w. \tag{4.5}$$

Integrating by parts N times, with N an arbitrary positive integer, we obtain for J the asymptotic series[6]

$$J[w] = \sum_{n=1}^{N} p_n[w](ikr)^{-n} + (ikr)^{-N} J[w^{(N)}], \tag{4.6a}$$

where the functional p_n is defined by

$$p_n[w] \equiv e^{-ikr\mu} w^{(n-1)} \Big|_{1}^{-1}, \tag{4.6b}$$

and where $w^{(m)}$ denotes the mth derivative of w with respect to μ.

Consider now the decomposition

$$I = S_N + R_N, \quad N \geq 2, \tag{4.7a}$$

of the series (4.4), with

$$S_N \equiv \sum_{n=0}^{N-2} J[a_n] r^{-(n+1)} \tag{4.7b}$$

and with R_N denoting the sum of the remaining terms in the series for I. From Eq. (4.6a), it follows that R_N is of order $r^{-(N+1)}$. It also follows from Eq. (4.6a) that in order to obtain S_N correct to order r^{-N}, the coefficient $J[a_n]$ in Eq. (4.7b) must be correct to order $r^{-(N-n-1)}$. We therefore make the substitution

$$J[a_n] = \sum_{j=1}^{N-n-1} p_j[a_n](ikr)^{-j} + O(r^{-N+n})$$

in Eq. (4.7b). This leads, after some algebra, to the relation

$$S_N = \sum_{n=0}^{N-2} q_n r^{-(n+2)} + O(r^{-N-1}), \tag{4.8a}$$

where

$$q_n \equiv \sum_{j=0}^{n} (ik)^{-(n-j+1)} p_{n-j+1}[a_j]. \tag{4.8b}$$

The asymptotic series for $r^2 I$, correct to order r^{-N} for any $N \geq 0$, is therefore given by[7]

$$r^2 I = \sum_{n=0}^{N} q_n r^{-n},$$

and, in view of (4.3b), the corresponding series for X is

$$X \to \sum_{n=0}^{N} (u_n - u_n^*) r^{-n}, \tag{4.9a}$$

$$u_n \equiv \int_0^{2\pi} d\phi\, e^{ikr} q_n. \tag{4.9b}$$

5. IDENTITIES INVOLVING THE COEFFICIENTS f_n

Since the relation $X + X^s = 0$ is satisfied identically for all values of r [see the remark at the end of Sec. 2], it follows from (3.7a) and (4.9a) that the equation

$$\beta_n + u_n - u_n^* = 0 \tag{5.1}$$

holds for all integers n from 0 to N. Since N is arbitrary, however, Eq. (5.1) must hold for every nonnegative integer n.

Substituting Eqs. (4.8b) and (4.6b) into Eq. (4.9b), we see that u_n can be written in the form

$$u_n = \sum_{j=0}^{n} (ik)^{-(n-j+1)} \int_0^{2\pi} d\phi\, e^{ikr}[e^{-ikr\mu}\, a_j^{(n-j)}]\,|_1^{-1}.$$

In view of the fact that the integrand is independent of ϕ, it follows that

$$u_n = 2\pi[e^{2ikr}\,\alpha_n(-1) - \alpha_n(1)], \tag{5.2a}$$

where

$$\alpha_n \equiv \sum_{j=0}^{n} (ik)^{-(n-j+1)}\, a_j^{(n-j)} \tag{5.2b}$$

and where $\alpha_n(\pm 1) \equiv \alpha_n|_{\mu=\pm 1}$. Substitution of Eq. (5.2a) into Eq. (5.1) then yields

$$\beta_n + 2\pi[\alpha_n^*(1) - \alpha_n(1)] + 2\pi[e^{2ikr}\,\alpha_n(-1) - \text{c.c.}] = 0.$$

Since this relation must hold for every value of r, we conclude that

$$\beta_n = 2\pi[\alpha_n(1) - \text{c.c.}] = 4\pi i\ \text{Im}\alpha_n(1) \tag{5.3}$$

and that

$$\alpha_n(-1) = 0. \tag{5.4}$$

Equations (5.3) and (5.4) are the essential results of this paper. Our final task is to express them directly in terms of the coefficients f_n.

From Eqs. (3.6b) and (3.4c), we find after performing some elementary manipulations on the summations that the integrand of Eq. (3.7b) is given by

$$c_n - c_n^* = 2ik \sum_{j=0}^{n} f_j^* f_{n-j} + \sum_{j=1}^{n} (n - 2j + 1) f_{n-j}^* f_{j-1}. \tag{5.5}$$

To express the quantity α_n in terms of the f_n, it should first be noted that the coefficient a_n defined by Eq. (4.2b) is, in addition to being explicitly dependent on μ, implicitly dependent on it through the factors f_n and f_{n-1}. With this in mind, the relation

$$a_n^{(m)} = ikm\, f_n^{(m-1)} + ik(1+\mu) f_n^{(m)} - n f_{n-1}^{(m)}, \qquad m \geq 1, \tag{5.6}$$

may easily be established by mathematical induction on the integer m. Substituting Eq. (5.6) into Eq. (5.2b) and juggling the summations, we obtain

$$\alpha_n = (1+\mu) \sum_{j=0}^{n} (ik)^{-j} f_{n-j}^{(j)} - \sum_{j=1}^{n} (ik)^{-j}(n - 2j + 1) f_{n-j}^{(j-1)}. \tag{5.7}$$

It then follows from Eqs. (3.7b), (5.5), and (5.7) that Eqs. (5.3) and (5.4) can be written as

$$2ik \sum_{j=0}^{n} \{ \int d\hat{r}\, f_j^* f_{n-j} - (4\pi/k)\ \text{Im}[(ik)^{-j} f_{n-j}^{(j)}(1)] \}$$
$$+ \sum_{j=1}^{n} (n - 2j + 1) \{ \int d\hat{r}\, f_{n-j}^* f_{j-1}$$
$$+ 4\pi i\ \text{Im}[(ik)^{-j} f_{n-j}^{(j-1)}(1)] \} = 0 \tag{5.8}$$

and

$$\sum_{j=1}^{n} (n - 2j + 1)(ik)^{-j} f_{n-j}^{(j-1)}(-1) = 0, \tag{5.9}$$

respectively.

When $n = 0$, Eq. (5.3) reduces to

$$2ik \int d\hat{r}\, f_0^* f_0 = 8\pi i\ \text{Im} f_0(1),$$

which is equivalent[8] to the optical theorem [see Eq. (1.3)], and Eq. (5.4) is satisfied identically since $\alpha_0 = (1+\mu) f_0$. For $n = 1$, Eq. (5.4) is again satisfied identically since α_1 is also proportional to $(1+\mu)$, and Eq. (5.8) becomes, after some rearrangement of terms and the substitution of f for f_0 and g for f_1,

$$\int d\hat{r}\,(f^* g + f g^*) = (4\pi/k)\ \text{Im} g(1) - (4\pi/k^2)\ \text{Re} f'(1). \tag{5.10}$$

---

* Work supported in part by the Research Committee of C.W. Post College.

1 M. Lieber (private communication).

2 S. Servadio, Phys. Rev. A **4**, 1256 (1971).

3 See, for example, L. Schiff, *Quantum Mechanics*, 3rd Edition (McGraw-Hill, New York, 1968), Chap. 5.

4 Equations (5.8) and (5.9) are identities in the sense that the optical theorem is an identity in the scattering amplitude.

5 Hereafter, the notation "$x \to y$" will be understood to mean "$x \to y$ as $r \to \infty$".

6 We are using the convention that a finite sum of the form $\sum_{n=m}^{n=N}$ is zero when $m > N$.

7 Note that N is being replaced by $N + 2$, so that the condition $N \geq 2$ [see Eq. (4.7a)] becomes the one just specified.

8 When $\mu = 1\ (-1)$, $\hat{r}$ is parallel (antiparallel) to $\hat{k}$. Thus, we have $f(1) \equiv f|_{\mu=1} = f(\mathbf{k}, \mathbf{k})$ and $f(-1) = f(\mathbf{k}, -\mathbf{k})$.

The Four-Dimensionality of Space and the Einstein Tensor

David Lovelock
Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada
(Received 10 January 1972)

All tensors of contravariant valency two, which are divergence free on one index and which are concomitants of the metric tensor, together with its first two derivatives, are constructed in the four-dimensional case. The Einstein and metric tensors are the only possibilities.

1. INTRODUCTION

In the general theory of relativity the Einstein field equations *in vacuo* are frequently introduced by solving the following problem: To construct all tensors of valency two, the components A^{hk} of which are:

(a) concomitants of the metric tensor g_{ij} and its first two derivatives, i.e., [1]

$$A^{hk} = A^{hk}(g_{ij}; g_{ij,r}, g_{ij,rs}); \tag{1.1}$$

(b) divergence free, i.e., [2]

$$A^{hk}{}_{|k} = 0; \tag{1.2}$$

(c) symmetric, i.e.,

$$A^{hk} = A^{kh}; \tag{1.3}$$

(d) linear in the second derivatives of g_{ij}.

The field equations *in vacuo* are then assumed to take the form

$$A^{hk} = 0. \tag{1.4}$$

Cartan[3] and Weyl and Vermeil[4] have shown that the only A^{hk} with the properties (a)–(d) is

$$A^{hk} = aG^{hk} + bg^{hk}, \tag{1.5}$$

where a, b are constants and G^{hk} is the Einstein tensor,[5] in which case (1.4) gives rise to the usual Einstein field equations with the cosmological term.

Recently the problem of constructing explicitly all A^{hk} which satisfy (a)–(c), without insisting on the severe restriction (d), was solved,[6] the general solution being

$$A^{hk} = \sum_{\mu=1}^{\infty} a_{(\mu)} g^{jk} \delta^{hh_1\cdots h_{2\mu}}_{jj_1\cdots j_{2\mu}} R_{h_1h_2}{}^{j_1j_2} \cdots R_{h_{2\mu-1}h_{2\mu}}{}^{j_{2\mu-1}j_{2\mu}} + ag^{hk}, \tag{1.6}$$

where $a_{(\mu)}$, a are constants and $\delta^{i_1\cdots i_N}_{j_1\cdots j_N}$ is the generalized Kronecker delta. By virtue of the fact that if $n < N$, then

$$\delta^{i_1\cdots i_N}_{j_1\cdots j_N} = 0 \quad \text{identically},[7] \tag{1.7}$$

it is easily seen that, for $n = 4$, (1.6) reduces exactly to (1.5). Consequently (a)–(c) and the assumption of the four-dimensionality of space implies the linearity condition (d), in which case, in general relativity, the apparently crucial assumption (d) is superfluous.

It is the purpose of this note to show that *in the four-dimensional case the symmetry condition (c) is also superfluous*, in the sense that it is a consequence of (1.1) and (1.2). More exactly we shall prove the following.

Theorem: If A^{hk} satisfies (1.1) and (1.2), then

$$A^{hk} = aG^{hk} + bg^{hk}$$

is an inevitable consequence for $n = 4$.

We should stress that no assumptions are made with respect to the first index of A^{hk} (for instance, we do *not* assume that $A^{hk}{}_{|h} = 0$).

2. PROOF OF THE THEOREM

The purpose of this section is to construct, for $n = 4$, all tensors A^{hk} which satisfy (1.1) and (1.2), without insisting that (1.3) be satisfied.

By virtue of the fact that A^{hk} is a tensor satisfying (1.1) the quantities

$$A^{hk;rs,tu} = \frac{\partial A^{hk}}{\partial g_{rs,tu}} \tag{2.1}$$

are the components of a tensor of contravariant valency six with the following symmetry properties,[8]

$$\begin{aligned} &A^{hk;rs,tu} = A^{hk;sr,tu} = A^{hk;rs,ut}, \\ &A^{hk;rs,tu} + A^{hk;ru,st} + A^{hk;rt,us} = 0, \end{aligned} \tag{2.2}$$

from which it can be shown that

$$A^{hk;rs,tu} = A^{hk;tu,rs}. \tag{2.3}$$

Furthermore, it is easily seen that (1.2) is equivalent to the condition

$$A^{hk;rs,tu} + A^{hu;rs,kt} + A^{ht;rs,uk} = 0. \tag{2.4}$$

In view of (2.3), $A^{hk;rs,tu}$ satisfies a cyclic identity similar to (2.4) on the indices krs.

We introduce the tensor $A^{hk;rs,tu;ab,cd}$ defined by

$$A^{hk;rs,tu;ab,cd} = \frac{\partial A^{hk;rs,tu}}{\partial g_{ab,cd}} \tag{2.5}$$

and, since

$$A^{hk;rs,tu;ab,cd} = A^{hk;ab,cd;rs,tu},$$

then $A^{hk;rs,tu;ab,cd}$ will have the same symmetry and cyclic properties in the indices $hkabcd$ as it has in $hkrstu$, viz., (2.2), (2.3), and (2.4). In Appendix A we prove that because of these properties

$$A^{hk;rs,tu;ab,cd} = 0 \tag{2.6}$$

when $n = 4$, the essence of the argument being that some of the twelve indices of $A^{hk;rs,tu;ab,cd}$ must coincide in a four-dimensional space. By virtue of (2.5), (2.6) implies that

$$A^{hk;rs,tu} = A^{hk;rs,tu}(g_{ab}; g_{ab,c}).$$

However, since $A^{hk;rs,tu}$ is a tensor, it must be independent[8] of $g_{ab,c}$, i.e.,

$$A^{hk;rs,tu} = \alpha^{hk;rs,tu}(g_{ab}), \tag{2.7}$$

where $\alpha^{hk,rs,tu}$ is a tensor satisfying (2.2)–(2.4) and is a concomitant of g_{ab} only.

From (2.7) and (2.1) we thus see that

$$A^{hk} = \alpha^{hk;rs,tu}g_{rs,tu} + \theta^{hk}(g_{ab}; g_{ab,c}),$$

which, as usual,[9] by (2.2) and (2.3), can be expressed in the form

$$A^{hk} = \tfrac{2}{3}\alpha^{hk;rs,tu}R_{trsu} + \beta^{hk}(g_{ab}), \tag{2.8}$$

where β^{hk} is a tensor and a concomitant of g_{ab} alone. Consequently, the problem of finding the most general A^{hk} satisfying (1.1) and (1.2), in a four-dimensional space, reduces to the problem of finding β^{hk} and $\alpha^{hk;rs,tu}$, which has been solved elsewhere. In fact it can be shown that[9]

$$\beta^{hk} = bg^{hk}, \tag{2.9}$$

where b is a constant, while any tensor of contravariant valency six which is a concomitant of g_{ab} alone must be of the form[10]

$$\alpha^{hk;rs,tu} = a^{hk;rs,tu} + b^{hk;rs,tu} \tag{2.10}$$

with

$$\begin{aligned} a^{hk;rs,tu} = {} & a_1 g^{hk}g^{rs}g^{tu} + a_2 g^{hk}g^{rt}g^{su} + a_3 g^{hk}g^{ru}g^{ts} \\ & + a_4 g^{hr}g^{ks}g^{tu} + a_5 g^{hr}g^{kt}g^{su} + a_6 g^{hr}g^{ku}g^{ts} \\ & + a_7 g^{hs}g^{kr}g^{tu} + a_8 g^{hs}g^{kt}g^{ru} + a_9 g^{hs}g^{ku}g^{tr} \\ & + a_{10} g^{ht}g^{kr}g^{su} + a_{11} g^{ht}g^{ks}g^{ru} \\ & + a_{12} g^{ht}g^{ku}g^{rs} + a_{13} g^{hu}g^{kr}g^{ts} \\ & + a_{14} g^{hu}g^{ks}g^{rt} + a_{15} g^{hu}g^{kt}g^{rs}, \end{aligned} \tag{2.11}$$

and

$$\begin{aligned} b^{hk;rs,tu} = {} & (b_1 g^{hk}\epsilon^{rstu} + b_2 g^{hr}\epsilon^{kstu} + b_3 g^{hs}\epsilon^{krtu} \\ & + b_4 g^{ht}\epsilon^{krsu} + b_5 g^{hu}\epsilon^{krts} + b_6 g^{kr}\epsilon^{hstu} \\ & + b_7 g^{ks}\epsilon^{hrtu} + b_8 g^{kt}\epsilon^{hrsu} + b_9 g^{ku}\epsilon^{hrts} \\ & + b_{10} g^{rs}\epsilon^{hktu} + b_{11} g^{rt}\epsilon^{hksu} + b_{12} g^{ru}\epsilon^{hkst} \\ & + b_{13} g^{st}\epsilon^{hkru} + b_{14} g^{su}\epsilon^{hkrt} \\ & + b_{15} g^{tu}\epsilon^{hkrs})/\sqrt{g}, \end{aligned} \tag{2.12}$$

where the a_μ, b_μ, $\mu = 1, \ldots, 15$, are constants, ϵ^{ijkl} is the Levi-Civita symbol, and $g = |\det(g_{ij})|$.

If we now impose (2.2)–(2.4) on $\alpha^{hk;rs,tu}$ and make use of Appendix B, we find

$$\begin{aligned} & a_5 = a_6 = a_8 = a_9 = a_{10} = a_{11} \\ & \qquad = a_{13} = a_{14} = \alpha \quad \text{(say)}, \\ & a_2 = a_3 = a_4 = a_7 = a_{12} = a_{15} = -2\alpha, \quad a_1 = 4\alpha. \end{aligned} \tag{2.13}$$

We substitute (2.4)–(2.13) in (2.8) and observe that

$$b^{hk;rs,tu}R_{trsu} = 0,$$

to find

$$A^{hk} = -8\alpha G^{hk} + bg^{hk},$$

which is the desired result, since α is a constant.

Finally we remark that in general (1.3) will not be a consequence of (1.1) and (1.2), $\epsilon^{hk}/\sqrt{g}$ and $\epsilon^{hkrstu}R_{rs}{}^{ab}R_{abtu}/\sqrt{g}$ being counter examples in spaces of dimension two and six, respectively.

ACKNOWLEDGMENTS

I would like to thank Professor J. Ehlers for drawing this problem to my attention. This work was supported by a grant from the National Research Council of Canada.

APPENDIX A

The purpose of this appendix is to show that, for $n = 4$,

$$A^{hk;rs,tu;ab,cd} = 0.$$

We shall adopt the following notation,

$$A^{hk;rs,tu;ab,cd} \equiv (hk; rs, tu; ab, cd),$$

and we shall suspend the use of the summation convention in this appendix.

If five or more of the indices $rstuabcd$ are equal then, by virtue of (2.2), $(hk; rs, tu; ab, cd) = 0$.

If four of the indices $rstuabcd$ are equal, then we also find $(hk; rs, tu; ab, cd) = 0$ by (2.2) and (2.4).

If three of the indices $rstuabcd$ are equal, the only quantities which are not trivially zero from (2.2) are equivalent to $(ba; ii, jk; jj, ik)$, $(ba; ii, jj; ij, kk)$, $(ba; ii, jj; ij, kl)$, $(ba; ii, jj; ik, jl)$, $(ba; ii, jk; jj, il)$, $(ba; ii, jj; il, kk)$ and $(ba; ii, kl; ik, jj)$, where i, j, k, l are distinct indices selected from 1, 2, 3, 4 and b, a are any of i, j, k, l. Clearly, by (2.4), if $a = i$ or $a = j$, then all these quantities vanish. By setting $a = k$ and $a = l$ in turn, and using (2.2) and (2.4), we find that the remaining quantities also vanish in this case.

Finally, if two of the indices $rstuabcd$ are equal, then it is not difficult to show, from (2.2), that $(bi; ij, kl; ij, kl)$ and $(bi; ij, kl; ik, jl)$ are the only independent quantities which are not trivially zero. The first of these is easily dealt with, but the second requires a little more analysis, which we now give in order to illustrate the technique. By repeated use of (2.4) we see that

$$\begin{aligned} (bi; ij, kl; ik, jl) &= -\tfrac{1}{2}(bj; ii, kl; ik, jl) \\ &= +\tfrac{1}{4}(bl; ii, kl:ik, jj) \\ &= -\tfrac{1}{4}(bk; ii, kl; il, jj) \\ &= \tfrac{1}{8}(bl; ii, kk; il, jj) \\ &= -\tfrac{1}{16}(bi; ii, kk; ll, jj) \\ &= 0. \end{aligned}$$

APPENDIX B

The result which is required in deriving (2.13) is the following: If

$$\begin{aligned} & \alpha_1 g^{hk}g^{rs}g^{tu} + \alpha_2 g^{hk}g^{rt}g^{su} + \alpha_3 g^{hk}g^{ru}g^{ts} \\ & \quad + \alpha_4 g^{hr}g^{ks}g^{tu} + \alpha_5 g^{hr}g^{kt}g^{su} + \alpha_6 g^{hr}g^{ku}g^{ts} \\ & \quad + \alpha_7 g^{hs}g^{kr}g^{tu} + \alpha_8 g^{hs}g^{kt}g^{ru} + \alpha_9 g^{hs}g^{ku}g^{tr} \\ & \quad + \alpha_{10} g^{ht}g^{kr}g^{su} + \alpha_{11} g^{ht}g^{ks}g^{ru} + \alpha_{12} g^{ht}g^{ku}g^{rs} \\ & \quad + \alpha_{13} g^{hu}g^{kr}g^{ts} + \alpha_{14} g^{hu}g^{ks}g^{rt} + \alpha_{15} g^{hu}g^{kt}g^{rs} \\ & \quad + (\beta_1 g^{hk}\epsilon^{rstu} + \beta_2 g^{hr}\epsilon^{kstu} + \beta_3 g^{hs}\epsilon^{krtu} \end{aligned}$$

$$+ \beta_4 g^{ht}\epsilon^{krsu} + \beta_5 g^{hu}\epsilon^{krts} + \beta_6 g^{kr}\epsilon^{hstu}$$
$$+ \beta_7 g^{ks}\epsilon^{hrtu} + \beta_8 g^{kt}\epsilon^{hrsu} + \beta_9 g^{ku}\epsilon^{hrts}$$
$$+ \beta_{10} g^{rs}\epsilon^{hktu} + \beta_{11} g^{rt}\epsilon^{hksu} + \beta_{12} g^{ru}\epsilon^{hkst}$$
$$+ \beta_{13} g^{st}\epsilon^{hkru} + \beta_{14} g^{su}\epsilon^{hkrt}$$
$$+ \beta_{15} g^{tu}\epsilon^{hkrs})/\sqrt{g} = 0, \qquad \text{(B1)}$$

where α_μ, β_μ, $\mu = 1, \ldots, 15$, are constants, then $\alpha_\mu = 0$. This is easily seen as follows. Consider three orthogonal vectors X_i, Y_i, Z_i each of nonzero length. If we multiply (B1) by $X_h X_k Y_r Y_s Z_t Z_u$, then clearly $\alpha_1 = 0$. The remaining α_μ vanish in a similar way.

We note that in general we cannot conclude from (B1) that $\beta_\mu = 0$, in view of the identity

$$g^{ai}\epsilon^{jkhl} + g^{al}\epsilon^{ijkh} + g^{ah}\epsilon^{lijk} + g^{ak}\epsilon^{hlij} + g^{aj}\epsilon^{khli} = 0$$

which follows from

$$\delta^{ijkhl}_{rstuv} g^{ar}\epsilon^{stuv} = 0.$$

---

[1] Unless otherwise specified, Latin indices run from 1 to n. A comma denotes partial differentiation.
[2] The summation convention is used throughout. The vertical bar denotes covariant differentiation.
[3] E. Cartan, J. Math. Pure Appl. **1**, 141 (1922).
[4] H. Weyl, *Space-Time-Matter* (Dover, New York, 1922), 4th ed., pp. 315 ff.; H. Vermeil, Nachr. Ges. Wiss. Göttingen, 334 (1917).
[5] If X^i is any contravariant vector field, then we define the Riemann curvature tensor $R_h{}^i{}_{jk}$, the Ricci tensor R_{hj}, the curvature scalar R, and the Einstein tensor G_{ij} by

$$X^i{}_{|jk} - X^i{}_{|kj} = R_h{}^i{}_{jk} X^h, \qquad R_{hj} = R_h{}^i{}_{ji}$$

$$R = g^{hj}R_{hj}, \qquad G_{ij} = R_{ij} - \tfrac{1}{2}g_{ij}R,$$

respectively.
[6] D. Lovelock, J. Math. Phys. **12**, 498 (1971).
[7] For applications of this result see D. Lovelock, Atti Accad. Naz. Lincei Rend. **42**, 187 (1967); Proc. Cambridge Phil. Soc. **68**, 345 (1970); Matrix Tensor Quart. **21**, 84 (1971).
[8] H. Rund, Abhandl. Math. Sem. Univ. Hamburg **29**, 243 (1966); J. C. du Plessis, Tensor **20**, 347 (1969).
[9] D. Lovelock, Arch. Ratl. Mech. Anal. **33**, 54 (1969).
[10] M. A. McKiernan, "Tensor Concomitants of the Metric Tensor," Demonstratio Mathematica (to be published).

The Weyl Basis of the Unitary Group $U(k)$*

T. H. Seligman†
Instituto de Física Universidad de México, México
(Received 19 November 1971)

Weyl's method for the construction of irreducible tensors of the unitary group is used to construct a basis for any irreducible representation of $U(k)$ or $GL(k)$ in terms of Bose creation operators. A simple way is indicated to select a complete but not over complete basis from the functions obtained. The basis obtained can be useful in nuclear or molecular calculations, as well as in some mathematical problems.

1. INTRODUCTION

The orthogonal Gel'fand basis[1] for the IR (irreducible representation) of $U(k)$ is well known and has found a wide range of applications.[1-3] Yet earlier a nonorthogonal basis of irreducible tensors was used. These tensors are obtained by applying Young symmetrizers[4] to the indices of reducible tensors. This basis was introduced by Weyl[5] using essentially arguments of the permutation group. We shall describe this basis in terms of Bose creation operators in the next chapter and we will call it for short the Weyl basis. Although Weyl used this basis to obtain many results,[5] the basis itself has never been investigated in similar detail as the Gel'fand basis. The first problem arises from the fact that, the basis is over complete; that is, the number of states obtained is larger than the dimension of the IR.

In Sec. 3 we shall therefore eliminate the redundant states and give a method to obtain a complete but not over complete basis, which we will call the reduced Weyl basis. In the following chapter we show that we could proceed similarly if we wish to use Fermi operators, which are important for many applications. Finally we point out some possible applications in molecular and nuclear physics, as well as in mathematical problems.

2. THE WEYL BASIS IN TERMS OF BOSE CREATION OPERATORS

We assume to have a set of Bose operators η_i^s, $i = 1, \ldots, k$, $s = 1, \ldots, k$ with the standard commutation relations

$$[\eta_i^s, \eta_j^t] = [(\eta_i^s)^\dagger (\eta_j^t)^\dagger] = 0 \qquad \text{(2.1a)}$$

and

$$[\eta_i^s, (\eta_j^t)^\dagger] = \delta_{st}\delta_{ij}. \qquad \text{(2.1b)}$$

We then know[2,3] that any product of N such operators forms a basis vector for the IR $[N]$ of $\mathbf{U}(k \times k)$. We may pass to the subgroup

$$\mathfrak{U}(k) \times U(k) \subset \mathbf{U}(k \times k), \qquad \text{(2.2)}$$

where $\mathfrak{U}(k)$ is acting on the upper and $U(k)$ on the lower indices alone. The IR in both groups have to be identical. We may have an arbitrary basis vector of a certain IR of $U(k)$ even if we chose the vector to be of highest weight with respect to $\mathfrak{U}(k)$.[3] We therefore must be able to construct any basis vector of the IR $f = [f_1, \ldots, f_k]$ of $U(k)$ as a homogeneous polynomial in which each term contains f_1 operators η_i^1, f_2 operators η_i^2, etc. and finally f_k operators η_i^k. Here i is an arbitrary index that for each operator η_i^s may take any value from 1 to k. Certain restrictions will appear later.

If we now take the construction prescription[4] for the Weyl basis, we immediately obtain such polynomials. We draw the Young diagram and fill in the creation operators in such a way that in the first row all upper indices are one, in the second two, etc. Thus the highest weight structure with respect to $U(k)$ is immediately apparent. The lower index has to be numbered and it proves to be convenient to give in turn an upper index denoting the row it stands in and a lower index for the column. Thus a creation operator reads as $\eta^s_{i\,s_m}$.

$$+ \beta_4 g^{ht}\epsilon^{krsu} + \beta_5 g^{hu}\epsilon^{krts} + \beta_6 g^{kr}\epsilon^{hstu}$$
$$+ \beta_7 g^{ks}\epsilon^{hrtu} + \beta_8 g^{kt}\epsilon^{hrsu} + \beta_9 g^{ku}\epsilon^{hrts}$$
$$+ \beta_{10} g^{rs}\epsilon^{hktu} + \beta_{11} g^{rt}\epsilon^{hksu} + \beta_{12} g^{ru}\epsilon^{hkst}$$
$$+ \beta_{13} g^{st}\epsilon^{hkru} + \beta_{14} g^{su}\epsilon^{hkrt}$$
$$+ \beta_{15} g^{tu}\epsilon^{hkrs})/\sqrt{g} = 0, \qquad \text{(B1)}$$

where α_μ, β_μ, $\mu = 1, \ldots, 15$, are constants, then $\alpha_\mu = 0$. This is easily seen as follows. Consider three orthogonal vectors X_i, Y_i, Z_i each of nonzero length. If we multiply (B1) by $X_h X_k Y_r Y_s Z_t Z_u$, then clearly $\alpha_1 = 0$. The remaining α_μ vanish in a similar way.

We note that in general we cannot conclude from (B1) that $\beta_\mu = 0$, in view of the identity

$$g^{ai}\epsilon^{jkhl} + g^{al}\epsilon^{ijkh} + g^{ah}\epsilon^{lijk} + g^{ak}\epsilon^{hlij} + g^{aj}\epsilon^{khli} = 0$$

which follows from

$$\delta^{ijkhl}_{rstuv} g^{ar}\epsilon^{stuv} = 0.$$

---

[1] Unless otherwise specified, Latin indices run from 1 to n. A comma denotes partial differentiation.
[2] The summation convention is used throughout. The vertical bar denotes covariant differentiation.
[3] E. Cartan, J. Math. Pure Appl. **1**, 141 (1922).
[4] H. Weyl, *Space-Time-Matter* (Dover, New York, 1922), 4th ed., pp. 315 ff.; H. Vermeil, Nachr. Ges. Wiss. Göttingen, 334 (1917).
[5] If X^i is any contravariant vector field, then we define the Riemann curvature tensor $R_h{}^i{}_{jk}$, the Ricci tensor R_{hj}, the curvature scalar R, and the Einstein tensor G_{ij} by

$$X^i{}_{|jk} - X^i{}_{|kj} = R_h{}^i{}_{jk} X^h, \qquad R_{hj} = R_h{}^i{}_{ji}$$

$$R = g^{hj}R_{hj}, \qquad G_{ij} = R_{ij} - \tfrac{1}{2}g_{ij}R,$$

respectively.
[6] D. Lovelock, J. Math. Phys. **12**, 498 (1971).
[7] For applications of this result see D. Lovelock, Atti Accad. Naz. Lincei Rend. **42**, 187 (1967); Proc. Cambridge Phil. Soc. **68**, 345 (1970); Matrix Tensor Quart. **21**, 84 (1971).
[8] H. Rund, Abhandl. Math. Sem. Univ. Hamburg **29**, 243 (1966); J. C. du Plessis, Tensor **20**, 347 (1969).
[9] D. Lovelock, Arch. Ratl. Mech. Anal. **33**, 54 (1969).
[10] M. A. McKiernan, "Tensor Concomitants of the Metric Tensor," Demonstratio Mathematica (to be published).

The Weyl Basis of the Unitary Group $U(k)$*

T. H. Seligman†
Instituto de Física Universidad de México, México
(Received 19 November 1971)

Weyl's method for the construction of irreducible tensors of the unitary group is used to construct a basis for any irreducible representation of $U(k)$ or $GL(k)$ in terms of Bose creation operators. A simple way is indicated to select a complete but not over complete basis from the functions obtained. The basis obtained can be useful in nuclear or molecular calculations, as well as in some mathematical problems.

1. INTRODUCTION

The orthogonal Gel'fand basis[1] for the IR (irreducible representation) of $U(k)$ is well known and has found a wide range of applications.[1-3] Yet earlier a nonorthogonal basis of irreducible tensors was used. These tensors are obtained by applying Young symmetrizers[4] to the indices of reducible tensors. This basis was introduced by Weyl[5] using essentially arguments of the permutation group. We shall describe this basis in terms of Bose creation operators in the next chapter and we will call it for short the Weyl basis. Although Weyl used this basis to obtain many results,[5] the basis itself has never been investigated in similar detail as the Gel'fand basis. The first problem arises from the fact that, the basis is over complete; that is, the number of states obtained is larger than the dimension of the IR.

In Sec. 3 we shall therefore eliminate the redundant states and give a method to obtain a complete but not over complete basis, which we will call the reduced Weyl basis. In the following chapter we show that we could proceed similarly if we wish to use Fermi operators, which are important for many applications. Finally we point out some possible applications in molecular and nuclear physics, as well as in mathematical problems.

2. THE WEYL BASIS IN TERMS OF BOSE CREATION OPERATORS

We assume to have a set of Bose operators η_i^s, $i = 1, \ldots, k$, $s = 1, \ldots, k$ with the standard commutation relations

$$[\eta_i^s, \eta_j^t] = [(\eta_i^s)^\dagger (\eta_j^t)^\dagger] = 0 \qquad \text{(2.1a)}$$

and

$$[\eta_i^s, (\eta_j^t)^\dagger] = \delta_{st}\delta_{ij}. \qquad \text{(2.1b)}$$

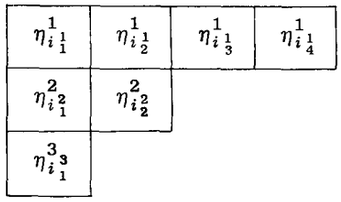
We then know[2,3] that any product of N such operators forms a basis vector for the IR $[N]$ of $\mathbf{U}(k \times k)$. We may pass to the subgroup

$$\mathfrak{U}(k) \times U(k) \subset \mathbf{U}(k \times k), \qquad \text{(2.2)}$$

where $\mathfrak{U}(k)$ is acting on the upper and $U(k)$ on the lower indices alone. The IR in both groups have to be identical. We may have an arbitrary basis vector of a certain IR of $U(k)$ even if we chose the vector to be of highest weight with respect to $\mathfrak{U}(k)$.[3] We therefore must be able to construct any basis vector of the IR $f = [f_1, \ldots, f_k]$ of $U(k)$ as a homogeneous polynomial in which each term contains f_1 operators η_i^1, f_2 operators η_i^2, etc. and finally f_k operators η_i^k. Here i is an arbitrary index that for each operator η_i^s may take any value from 1 to k. Certain restrictions will appear later.

If we now take the construction prescription[4] for the Weyl basis, we immediately obtain such polynomials. We draw the Young diagram and fill in the creation operators in such a way that in the first row all upper indices are one, in the second two, etc. Thus the highest weight structure with respect to $U(k)$ is immediately apparent. The lower index has to be numbered and it proves to be convenient to give in turn an upper index denoting the row it stands in and a lower index for the column. Thus a creation operator reads as $\eta^s_{i\,s_m}$.

As an example let us consider the IR $f = [421]$ of $U(4)$.



(2.3)

Here the lower indices i^s_m are still arbitrary. We now apply the Young operator as defined in[4] to these indices. Because of the Bose character of the η the row symmetrization is superfluous and we only have to antisymmetrize the columns; we may achieve that by forming the corresponding determinants, if we omit all factors of the symmetrizers and antisymmetrizers. We may do this as any normalization of nonorthogonal states is meaningless. For Example (2.3) the homogeneous polynomial then reads as

$$\Delta^{123}_{i^1_1 i^2_1 i^3_1}\Delta^{12}_{i^1_2 i^2_2}\Delta^{1}_{i^1_3}\Delta^{1}_{i^1_4}, \tag{2.4}$$

where the Δ are defined according to the notation of Ref. 2 as determinants:

$$\Delta^s_i = \eta^s_i, \qquad \Delta^{st}_{ij} = \eta^s_i\eta^t_j - \eta^s_j\eta^t_i, \text{etc.} \tag{2.5}$$

The polynomial (2.4) must be either a basis vector of the IR [421] of $U(4)$ or zero. The latter is obviously the case if any two lower indices standing in the same column are equal.

We obtain a basis vector of f if we impose the just mentioned condition, because the monomials obviously form tensors and the homogeneous polynomials are formed according to the standard Weyl procedure.[4] There cannot arise any loss of generality from putting all upper indices s into the sth row as any other ordering would either yield zero (if two equal upper indices appear in the same column) or would at most differ by a sign caused by the reordering in the determinants. The fact that the basis is complete if we choose all possible lower indices follows from the construction, but will be independently proved in the next chapter. We find that all states constructed in the described manner will have a well-defined weight,[3] which is very important for all applications.

3. THE REDUCED WEYL BASIS

Up to this point we essentially gave a reformulation of well-known facts in the second quantized picture. We now reach the main purpose of this paper; namely to formulate and prove a method for selecting a complete, but not over complete set of these states.

We will achieve this by imposing an ordering condition on the indices i^s_m similar to the conditions existing for a standard Young tableau.[4] These conditions are (as we shall prove in this chapter and the appendix)

$$i^s_m \leq i^s_{m'}, \quad \text{if } m < m' \tag{3.1a}$$

and

$$i^s_m < i^{s'}_m \quad \text{if } s < s'. \tag{3.1b}$$

We can see that the conditions vary from those for standard Young tableau by allowing also the equal sign in Eq. (3.1a).

We now define a set of numbers a^s_i that will uniquely characterize our generalized tableau. a^s_i indicates the number of times a lower index i appears in the row s. Conditions (3.1a) and (3.1b) immediately show that

$$a^s_i = 0 \quad \text{if } i < s. \tag{3.2}$$

Further restrictions on the a^s_i are imposed by Conditions (3.1a) and (3.1b) but we will not obtain them explicitly; rather we define a new set of numbers

$$H^s_g = \sum_{i=s}^{g} a^s_i. \tag{3.3}$$

Here $s \leq g \leq k$ must hold and we find

$$H^s_k = f_s. \tag{3.4}$$

The numbers H^s_g again characterize the generalized Young tableau uniquely. Further we can explicitly formulate Conditions (3.1a) and (3.1b) as

$$H^s_g \geq H^s_{g-1} \geq H^{s-1}_g. \tag{3.5}$$

These relations may be checked by comparison with (3.1a) and (3.1b) if we keep in mind that H^s_g is just the number of η^s_i with $i \leq g$ in the row s of the tableau defined by (3.3). Conditions (3.5) are identical to those for a Gel'fand scheme.[2] If we consider that the generalized Young tableau and therefore also the numbers H^s_g characterize uniquely our selected Weyl states, we obtain the important result that the number of states we selected is equal to the number of Gel'fand states and thus to the dimension of the IR of $U(k)$ considered. Note though that this by no means implies that the selected Weyl states are Gel'fand states; such an assumption is false, except for $U(2)$ where the states are determined by their weight.

We mentioned earlier that Weyl states are of definite weight. The components of the weight are given by

$$w_i = \sum_{s=1}^{i} a^s_i = \sum_{s=1}^{i} H^s_i - \sum_{s=1}^{i-1} H^s_{i-1}. \tag{3.6}$$

States of different weight are obviously orthogonal.

We have already shown that the number of basis states in the reduced Weyl basis defined by Eqs. (3.1a) and (3.1b) are equal to the dimension of the corresponding IR. In order to prove that we have a complete but not over complete basis we thus have to show that the states selected by Eqs. (3.1a) and (3.1b) are independent. This is done in the Appendix; the method consists in defining an order for states of equal weight and then showing that the leading term

$$\prod_{s=1}^{k}\prod_{s=1}^{s}(\eta^s_i)^{a^s_i} \tag{3.7}$$

of the product of determinants of a certain state does not appear in any state that is "lower" according to the ordering defined (the author is grateful to P. Kramer for suggesting the idea of this proof). This fact is sufficient for the independence of the states. We have thus shown that the states selected by Conditions (3.1a) and (3.1b) form a complete but not over complete basis. Note that the proof for completeness

does not make use of the knowledge that the general set of Weyl states is complete; but only that they belong to the IR f of $U(k)$.

4. OTHER PROPERTIES OF THE WEYL BASIS

We shall now proceed to show a few other properties of the Weyl basis. First we note that we made at no point in Sec. 3 or in the Appendix use of the Hermitian properties Eq. (2.1b) of the creation operators. We know[6] that if we replace Eq. (2.1b) by

$$[\eta_i^s,(\eta_j^t)^+] = A_{ij}^{st}, \tag{4.1}$$

where A is a number depending on the indices s,t,i,j we may construct from these operators a Lie algebra of $GL(k)$ rather than $U(k)$. The Weyl construction is valid for $GL(k)$ and the dimension of finite (nonunitary) IR of $GL(k)$ is equal to the dimension of the corresponding IR of $U(k)$. Therefore our prescription (3.1a) and (3.1b) will also select a reduced Weyl basis for finite IR of $GL(k)$ that is complete but not over complete. In a sense the Weyl basis seems particularly interesting in this case as nonorthogonality of the basis is unavoidable, and therefore no longer constitutes a drawback of the Weyl basis as compared to the Gel'fand basis.

Going back to the group $U(k)$ we may also ask ourselves what happens if we consider only states of weight

$$\mathbf{W} = (1,1,1,\ldots,1). \tag{4.2}$$

This was done by Moshinsky[7] for Gel'fand states and he finds that these "special" Gel'fand states form a basis for a Yamanouchi or orthogonal representation of the permutation group $S(k)$. As the Weyl states have well-defined weights, we may immediately single out the "special" Weyl states. These are a linear combination of the "special" Gel'fand states and thus also form a basis to the same IR of $S(k)$. Considering the way the basis is constructed we can identify it with the basis for the rational (natural) representation of $S(k)$.[8] Note that also in the derivation of the rational representation of $S(k)$, the problem of reducing the basis occurs and its solution is by no means trivial.

We finally wish to indicate that the whole procedure developed in Sec. 3 and in the Appendix may be equally well carried through if we use Fermi operators[2] $(b_i^s)^+$ instead of Bose operators η_i^s. The main differences that appear in the analysis are the following.

The IR of $U(k)$ and $U(k)$ are not equal but associate[2]; therefore we obtain the highest weight state in $U(k)$ if we fill the s column of the IR f of $U(k)$ with operators $(b_i^s)^+$. It then is adequate to apply first the column antisymmetrizers and then the row symmetrizers. As the columns are automatically antisymmetric, our states are products of permanents. We may then closely follow the reasoning for bosons to obtain the reduced Weyl basis. Note that in relations (3.1a) and (3.1b) the upper indices of i continue to characterize the row and the lower ones the column; thus the lower indices will coincide with the upper index of the Fermi operator. The proof that the number of states in the reduced basis is correct remains the same; while the independence proof of the Appendix may be applied after reflecting the pattern at the diagonal if we keep in mind that it does not matter whether the permutations considered stem from determinants or permanents; only the ordering will have to be done according to the reflected pattern. Note that the basis in terms of Bose and Fermi operators are not identical as row symmetrizer and column antisymmetrizer are interchanged in the second case, and the two operations do not commute.

5. APPLICATIONS

The Weyl basis may be applied in various ways in calculations of nuclear and molecular physics. First, it may be a valuable tool to choose physically significant trial wavefunctions. This is immediately obvious in molecular physics because, if we want to obtain the valence bond states, e.g., given in terms of the permutation group in Refs. 9 and 10, in a second quantized picture (Fermi operators) these turn out to be Weyl states. From this fact we recognize the importance of the Weyl basis; but the result has been known for a long time.

We may hope to find similar applications in nuclear physics in the framework of cluster and supermultiplet considerations. One possibility appears if we wish to extend Brink's model.[11] We make use here of the fact that Deenen[12] has found that the second quantized picture is compatible with the elimination of the center of mass motion, and gives explicit prescriptions how to proceed. For a more detailed description, especially of the non-α-clusters Deenen considers, it may be important to include excitations at the centers (e.g., to obtain a d-cluster). Yet we might still want to consider only orbitally symmetrical states at each center. This certainly will be possible with the Weyl basis. Similar applications may be considered if we are looking for α-like four-particle-four-hole excitations. Note that both the application in molecular physics and the one in Brink's model involve nonorthogonal orbitals and thus Weyl states of $GL(k)$.

The physically significant states, we may select in such a way, are not lying in general within the reduced basis defined in Sec. 3. We will thus in each case have to ensure independence of the chosen states. Nevertheless, the problem is not very critical if we chose only a few states from the basis and avoid obvious dependences arising from the interchange of indices in the determinants or permanents.

Quite a different type of applications may arise from the simplicity of the Weyl states, which are just products of determinants. For instance, the overlap of two states must be an integer if we consider orthogonal orbitals. We may hope that a simple algorithm to obtain these integers can be found, and this seems even more probable as we know for the case of molecules the Pauling numbers[10] that may be easily obtained. If we know these overlaps, the matrix elements of the generators as well as other quantities such as representation matrices may be readily obtained. It may then turn out that, in a similar way as is well known in molecular physics,[10] Weyl states also prove to be particularly efficient for computations in nuclear physics.

Another more mathematical application of the reduced Weyl basis was found by Sharp.[13] He points

out that this basis is a particularly simple solution of the internal labeling problem (labels to distinguish different states of the same weight within a basis of an IR). Thus he finds this basis to be particularly useful to solve the external labeling problem (multiplicity labels in the reduction of Kronecker products), using Speiser's theorem that connects the two problems.

APPENDIX

We shall proceed to prove the independence of the Weyl states selected by Conditions (3.1a) and (3.1b) in Sec. 3. To do this we first define an appropriate ordering of these states. This is achieved by arranging the numbers a_i^s that characterize a state into a triangular matrix, where s is the row and i the column index:

$$\begin{matrix} a_1^1 a_2^1 a_3^1 \cdots a_k^1 \\ a_2^2 a_3^2 \cdots \\ \cdots \\ a_k^k \end{matrix} \qquad \text{(A1)}$$

We then inspect the a column by column from top to bottom, and we say that a state **c** characterized by numbers c_i^s comes after a state **a** characterized by a_i^s if we find for the first $a_i^s \neq c_i^s$ that $a_i^s > c_i^s$ holds.

The monomial (3.7)

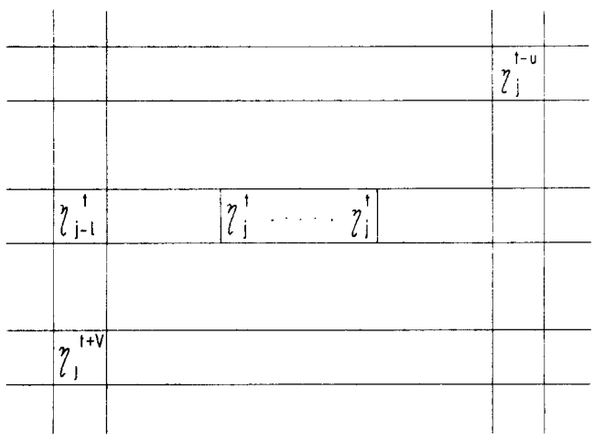


FIG. 1. The figure shows a section from a generalized Young tableau that characterizes a state from the reduced Weyl basis. Only parts of the rows $t-u$, t, and $t+v$ are shown.

$$\prod_{s=1}^{k} \prod_{i=1}^{s} (\eta_i^s)^{a_i^s} \qquad \text{(A2)}$$

is the leading term of the polynomial **a**. We now have to show that this term does not appear in any polynomial **c** that comes after **a**. This will be a sufficient condition for the set of polynomials to be independent.

As polynomials of different weight are orthogonal we may concentrate on the ones having equal weights. Therefore we have that $a_1^1 = c_1^1 = w_1$. We assume that all $c_i^s = a_i^s$ up to a certain $c_j^t < a_j^t$. This is exactly what is implied by the statement that **c** comes after **a**. We thus have less operators η_j^t in the row t of the pattern describing the polynomial **c** than in the same row of the pattern describing the polynomial **a**. We now have to find out whether any term in the column antisymmetrizer will yield the monomial Eq. (A2) when applied to the pattern characterizing **c**.

This will not happen because any permutation, that is contained in the column antisymmetrizer, will, when increasing, the power of η_j^t, simultaneously lower the power of some η_i^s that is equal in the leading term of both polynomials. This may be seen by considering a part of the pattern characterizing **c** which is shown in Fig. 1. There u, v, l are positive integers. The lower index j can only appear above to the right or below to the left, but never below to the right of the η_j^t according to the relations (3.1a) and (3.1b). We may now try to obtain a term in the polynomial **c** with a higher power of the η_j^t by considering a permutation involving η_j^{t-u}; but such a term would automatically have a lower power of η_j^{t-u} than the direct term. On the other hand according to the assumption $a_j^{t-u} = c_j^{t-u}$ and thus power of η_j^{t-u} may not be reduced, and we have by consequence to discard these permutations. We may in turn consider permutations involving η_j^{t+v}. The power of this operator is not fixed by our assumptions; but in order to obtain a η_j^t we have to permute the lower index with one of an operator η_{j-l}^t and by consequence lower the power of this operator. Yet our assumption was that also $a_{j-l}^t = c_{j-l}^t$ and thus also these permutations cannot yield the monomial Eq. (A2). Another possibility does not exist as we may easily see in Fig. 1. Therefore the leading term expression (A2) of **a** does not appear in the polynomial **c**.

ACKNOWLEDGEMENTS

I wish to thank A. Cantú, P. Kramer, M. Moshinsky, and R. T. Sharp for many useful suggestions and discussions.

* Work supported by the Comisión Nacional de Energía Nuclear, México.

† Fellow of the "Schweizerischer Nationalfonds der Wissenschaften."

1 I. M. Gel'fand and M. L. Zetlin, Dokl. Akad. Nauk. USSR 71, 825 (1950).

2 M. Moshinsky, *Group Theory and the Many Body Problem* (Gordon and Breach, New York, 1967).

3 P. Kramer and M. Moshinsky in *Group Theory and Its Applications*, edited by E. M. Loebl (Academic, New York, 1968).

4 M. Hammermesh, *Group Theory* (Addison-Wesley, Reading, 1962), p. 381.

5 H. Weyl, Math. Z. **23**, 271 (1925) (p. 277).

6 M. Moshinsky and T. H. Seligman, Ann. Phys. (N.Y.) **66**, 311 (1971).

7 M. Moshinsky, J. Math. Phys. **7**, 691 (1966).

8 G. Flach und R. Reif, *Gruppentheoretische Methoden im Schalenmodell der Atomkerne I Akademie* (Verlag, Berlin, 1964).

9 G. Rumer, E. Teller, and H. Weyl, Nachr. Ges. Wiss. Göttingen, Math.-Phys. Kl., 499 (1932).

10 F. A. Matson and A. Cantú, J. Phys. Chem. **72**, 21 (1968); A. Cantú and F. A. Matson (private communications).

11 D. M. Brink, "Generator Coordinate Method and Geometry of Alpha-clusters," in *Clustering Phenomena in Nuclei* (IAEA, Vienna, 1969).

12 J. Neenen, Université Libre de Bruxelles, Preprint.

13 R. T. Sharp (private communication).

Relation between the Three-Dimensional Fredholm Determinant and the Jost Functions*

Roger G. Newton†‡
Laboratoire de Physique Mathematique,§ Université des Sciences, Montpellier, France
and
Physics Department, Indiana University, Bloomington, Indiana
(Received 7 January 1972)

It is proved that the modified Fredholm determinant F of the three-dimensional Lippmann–Schwinger equation in the theory of scattering by spherically symmetric potentials is related to the Jost functions f_l of angular momentum l by

$$F = \prod_{l=0}^{\infty} \{ f_l^{2l+1} \exp[(2l+1)\,\mathrm{Tr}K_l] \},$$

where K_l is the kernel of the lth radial Lippmann–Schwinger equation. The relation between the multiplicity of the zeros of F and the degeneracy is discussed, and a relevant theorem for Hilbert–Schmidt operators is proved.

The theory of potential scattering is most commonly formulated in terms of the integral equation on R^3,

$$\psi(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} + \int (d\mathbf{r}') G^+(k;\mathbf{r},\mathbf{r}') V(r') \psi(\mathbf{r}'), \tag{1}$$

where the Green's function $G^+(k;\mathbf{r},\mathbf{r}')$ is the integral kernel of the boundary value of the resolvent operator

$$G(E) = (E - H)^{-1},$$

of the self-adjoint extension H to $L^2(R^3)$ of the negative Laplacean, $-\Delta$, at $E = k^2 + io$,

$$G^+(k;\mathbf{r},\mathbf{r}') = -\frac{1}{4\pi} \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}.$$

In the application of Fredholm methods to the solution of this integral equation a quantity of prime interest is the modified Fredholm determinant

$$F(E,\gamma) = \det{}_2[1 - \gamma G(E) V], \tag{2}$$

which is known to exist as an absolutely convergent power series in γ for all values of E and γ if the potential function $V(\mathbf{r})$ satisfies suitable conditions.[1]

If the potential is spherically symmetric, $V(\mathbf{r}) = V(r)$, $r = |\mathbf{r}|$, then the integral equation (1) is usually conveniently subjected to a spherical-wave analysis, that is, $\psi(\mathbf{r})$ is expanded on the basis of the Legendre polynomials

$$\psi(\mathbf{r}) = (kr)^{-1} \sum_{l=0}^{\infty} (2l+1) i^l \psi_l(\mathbf{r}) P_l(\cos\theta),$$

where θ is the angle between $\mathbf{r}$ and $\mathbf{k}$. The coefficients $\psi_l(r)$ then obey integral equations on R_+:

$$\psi_l(r) = u_l(kr) + \int_0^{\infty} dr' G_l^+(k;r,r') V(r') \psi_l(r'), \tag{3}$$

where u_l is a Riccati–Bessel function $u_l(z) = (\frac{1}{2}\pi z)^{1/2} J_{l+1/2}(z)$. The Green's function here is given by

$$G_l^+(k;r,r') = \begin{cases} (-1)^{l+1} k^{-1} u_l(kr) W_l^{(+)}(kr'), & r' > r, \\ (-1)^{l+1} k^{-1} u_l(kr') W_l^{(+)}(kr), & r' < r, \end{cases}$$

where $W_l^{(+)}$ is a Riccati–Hankel function $W_l^{(+)}(z) = i(-1)^l (\frac{1}{2}\pi z)^{1/2} H^{(1)}_{l+1/2}(z)$. This Green's function is the integral kernel of the boundary value of the resolvent $G_l(E) = (E - H_l)^{-1}$ of the self-adjoint extension H_l to $L^2(0,\infty)$ of the differential operator H_l',

$$H_l' = -\frac{d^2}{dr^2} + \frac{l(l+1)}{r^2},$$

as $E \to k^2 + io$.

In the context of solving the integral equation (3) one is much interested in the Jost function,[2] which is the (unmodified) Fredholm determinant

$$f_l(k,\gamma) = \det[1 - \gamma G_l(E) V]. \tag{4}$$

Under suitable conditions on the potential function $V(r)$, the Jost functions f_l are known[2] to exist for all values of E and γ, and all nonnegative integers l, as an absolutely convergent power series in γ.

The purpose of this paper is to give the relation between the function $F(E,\gamma)$ of (2) and the functions $f_l(E,\gamma)$ of (4), in case the potential is spherically symmetric. It may be regarded as a somewhat belated footnote to the paper by Jost and Pais.[1]

To start with, we shall take $E < 0$. Then the operator $-G(E)$ is positive and we form

$$g(E) = [-G(E)]^{1/2}$$

as a positive operator. The operator

$$K = gVg$$

is self-adjoint, and

$$F(E,\gamma) = \det{}_2[1 + \gamma K(E)]. \tag{5}$$

Let the eigenvalues of $K(E)$ be $-1/\gamma_n(E)$, $n = 1, 2, \cdots$. Then $F(E,\gamma_n(E)) = 0$, and the order of the zero (as a function of γ) equals the degeneracy of the eigenvalue. We shall include an eigenvalue as many times in the set $\{\gamma_n\}$ as its degeneracy, that is, as there are linearly independent solutions in $L^2(R^3)$ of the equation[4]

$$[-\Delta + \gamma_n(E) V]\phi = E\phi.$$

The modified Fredholm determinant may then be expressed as the convergent infinite product[5]

$$F(E,\gamma) = \prod_{n=1}^{\infty} \left(1 - \frac{\gamma}{\gamma_n(E)}\right) e^{\gamma/\gamma_n(E)}. \tag{6}$$

The same argument applies to the modified Fredholm determinant

$$\begin{aligned} F_l(E,\gamma) &= \det{}_2[1 + \gamma K_l(E)] \\ &= \det{}_2[1 - \gamma G_l(E) V] \end{aligned} \tag{7}$$

of Eq. (3), where $K_l = (-G_l)^{1/2} V (-G_l)^{1/2}$ for $E < 0$.

We have the convergent product representation

$$F_l(E,\gamma) = \prod_{n=1}^{\infty} \left(1 - \frac{\gamma}{\gamma_{ln}(E)}\right) e^{\gamma/\gamma_{ln}(E)} \tag{8}$$

in terms of the eigenvalues $\gamma_{ln}(E)$ of $K_l(E)$ in $L^2(0,\infty)$. In this case it is well known that there is no degeneracy, i.e., each zero γ_{ln} of $F_l(E,\gamma)$ is necessarily simple[6] as a function of γ.

If $\gamma_{ln}(E)$ is an eigenvalue of $K_l(E)$, then E is an eigenvalue[7] of $H_l + \gamma_{ln}(E)V$ on $L^2(0,\infty)$, with the unique eigenvector ϕ_{ln},

$$[H_l + \gamma_{ln}(E)V]\phi_{ln} = E\phi_{ln}.$$

Hence each of the $2l+1$ linearly independent functions[8]

$$\psi_{ln}^m(\mathbf{r}) = r^{-1}\phi_{ln}(r)\ Y_l^m(\theta,\varphi)$$

is an eigenvector in $L^2(R^3)$ of $H + \gamma_{ln}(E)V$ with the eigenvalue E, and hence of $K(E)$ with the eigenvalue $\gamma_{ln}(E)$. Unless $\gamma_{ln}(E)$ is also an eigenvalue of $K_{l'}(E)$, with $l' \neq l$, the multiplicity of $\gamma_{ln}(E)$ as an eigenvalue of $K(E)$ is therefore exactly $2l+1$. On the other hand, if m eigenvalues belonging to m different l values $l_1, \ldots, l_m$, coincide at E, then the multiplicity of the corresponding eigenvalue of $K(E)$ is $\sum_{i=1}^{m}(2l_i+1)$. In the first instance, $F(E,\gamma)$ must have γ_{ln} as an $(2l+1)$-fold zero, and in the second instance the zero is $[\sum(2l_i+1)]$-fold. In either case it follows that

$$F(E,\gamma) = \prod_{l=0}^{\infty}[F_l(E,\gamma)]^{2l+1} \tag{9}$$

and the product converges.

Since the unmodified Fredholm determinant $f_l(E,\gamma)$ also exists, it is related to the function $F_l(E,\gamma)$ by

$$f_l = F_l\ \exp\left(-\gamma\sum_n(1/\gamma_{ln})\right) = F_l e^{\gamma \mathrm{Tr} K_l}. \tag{10}$$

We therefore have

$$\begin{aligned} F(E,\gamma) &= \prod_{l=0}^{\infty}[f_l(E,\gamma)e^{\gamma \mathrm{Tr} K_l(E)}]^{2l+1} \\ &= \prod_{l=0}^{\infty}\{|f_l(E,\gamma)|^{2l+1}e^{(2l+1)\gamma \mathrm{Tr} K_l(E)}\} \end{aligned} \tag{11}$$

as the relation between the modified Fredholm determinant F and the Jost functions f_l. The terms $\exp[(2l+1)\gamma \mathrm{Tr} K_l]$ cannot be taken out of the product. They are needed for convergence.[9]

The relation (9) or (11) has been demonstrated so far only as $E < 0$. Let us now allow E to become complex. The operators $K_l(E)$ and $K(E)$ are analytic functions of E, and so are their eigenvalues[10] $\gamma_{ln}(E)$, with the negative real axis in the domain of regularity. The same is known for the modified Fredholm determinants[11] F and F_l. Moreover, the products in (6) and (8) converge for all complex values of E (and all γ), since $K(E)$ and $K_l(E)$ are Hilbert–Schmidt operators[1] for all complex E. Thus the convergence of the products is uniform in any compact domain of the E plane that does not contain any points of the nonnegative real axis. Therefore, the products are regular analytic in any such region, and (9) holds for all complex E as well. In addition, $f_l(E,\gamma)$ is an analytic function[12] of E regular in the cut E plane, and hence (10) holds there, and thus (11).

Finally, we let E approach the nonnegative real axis. Then $K(E)$ and $K_l(E)$ are still Hilbert–Schmidt operators,[1] the products (6) and (8) still converge, and hence (9) still holds. Similarly, the unmodified determinant f_l exists there and is the boundary value of $f_l(E)$ in the complex plane.[12] Hence (10) and (11) hold on the nonnegative real axis.

Since the result (11) is based on the zeros of the functions f_l and F, let us discuss these briefly. For negative E, the zeros $\gamma_{ln}(E)$, $n = 1, 2, \ldots$, are simple and have no finite accumulation point.[13] They are analytic functions of E. Hence as E is taken into the complex plane and to the nonnegative real axis, $\gamma_{ln}(E)$ remains a simple zero of $f_l(E,\gamma)$, unless two or more such zeros coincide. As E describes a path in the E plane, each $\gamma_{ln}(E)$ describes a trajectory in the complex plane, and each $\gamma_{ln}(E)$ remains a simple zero of F_l unless two (or more) such trajectories cross at some value of E. Consequently the resolvent $[1-\gamma K_l(E)]^{-1}$ has a simple pole at $\gamma = \gamma_{ln}(E)$, and the eigenvalue $1/\gamma_{ln}(E)$ of $K_l(E)$ is nondegenerate.[14]

Suppose now that for $E = E_0$, p eigenvalues $\gamma_{ln_i}(E)$, $i = 1, \ldots, p$, coincide, so that $f_l(E_0,\gamma)$, and hence $F_l(E_0,\gamma)$, has a p-fold zero at $\gamma = \gamma_{ln}(E_0) \equiv \Gamma$. Then the resolvent $[1-\gamma K_l(E_0)]^{-1}$ has a pole at $\gamma = \Gamma$, whose order M may have any value between 1 and p, $1 \leq M \leq p$, and the *ascent*[15] of $[1-\Gamma K_l(E_0)]$ equals M. We shall prove in the Appendix that the facts that $F_l(E_0,\gamma)$ has a p-fold zero at $\gamma = \Gamma$ and that the ascent of $[1-\Gamma K_l(E_0)]$ equals M, imply that the dimensionality of the nullspace of $[1-\Gamma K_l(E_0)]^M$ equals p. Hence[16] the dimensionality of the nullspace of $[1-\Gamma K_l(E_0)]$, i.e., the degeneracy equals *at most* $p-M+1$ (and at least one, of course). In the special case of the coincidence of two eigenvalues, this means that either the resolvent has a simple pole, the ascent is 1, and the degeneracy is 2, or the resolvent has a double pole, the ascent is 2, and there is no degeneracy.

Assuming that at $E = E_0$, no eigenvalues of K_l coincide with any eigenvalues of K_L, for all $l \neq L$, then the coincidence of p zeros of F_l at Γ implies, according to (9), that F has a (p^{2l+1})-fold zero at Γ. The fact that the resolvent $[1-\gamma K_l(E_0)]^{-1}$ has a pole of order M at $\gamma = \Gamma$ implies that $[1-\gamma K(E_0)]^{-1}$ also has a pole of order M there.[17] Thus the ascent of $[1-\Gamma K(E_0)]$ is M, and the theorem of the Appendix tells us that the dimensionality of the nullspace of $[1-\Gamma K(E_0)]^M$ equals p^{2l+1}.

Then there is the possibility of two or more eigenvalues for different values of l to coincide at a given energy $E = E_1$. Because of the angle functions, any two eigenfunctions in $L^2(0,\infty)$ of K_l and $K_{l'}$, respectively, with $l \neq l'$, lead to eigenfunctions of K, in $L^2(R^3)$, that are mutually orthogonal. Hence a coincidence of p eigenvalues of K_{l_i}, with p different l values, l_i, $i = 1, \ldots, p$, leads to a null space of K of dimensionality $D = \sum_{i=1}^{p}(2l_i+1)$, that is, to a D-fold degeneracy. The zero of F at $\gamma = \Gamma$, according to (9), is D-fold. Hence it follows from the corollary of the Appendix that the ascent of $[1-\Gamma K(E_1)]$ equals 1, and therefore the resolvent $[1-\gamma K(E_1)]^{-1}$ has a simple pole at $\gamma = \Gamma$.

The discussion of the effects of a coincidence, both of more then one zero of $K_l(E_0)$, and of zeros of $K_l(E_0)$ and of $K_L(E_0)$, $l \neq L$, at one and the same E_0, will be left as an exercise for the reader.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge a stimulating conversation on this subject with Dr. K. Chadan, and to thank

the Laboratoire de Physique Mathématique of the University of Montpellier for its hospitality.

APPENDIX

Theorem[18]: Let K be a Hilbert-Schmidt operator; let α be an eigenvalue of K, with the ascent[15] of $(K-\alpha)$ equal to M; and let $f(\gamma)=\det_2(1-\gamma K)$ be the modified Fredholm determinant. Then the order p of the zero of $f(\gamma)$ at $\gamma=1/\alpha$ equals the dimensionality m of the null space of $(K-\alpha)^M$.

Proof: We expand the resolvent

$$(z-K)^{-1}=\sum_{n=-M}^{\infty}A_n(z-\alpha)^n. \tag{A1}$$

Application of $(z-K)$ to both sides implies that

$$\sum (z-\alpha)^n[A_{n-1}-(K-\alpha)A_n]=\sum (z-\alpha)^n[A_{n-1}-A_n(K-\alpha)]=1,$$

and therefore

$$A_{n-1}=A_n(K-\alpha)=(K-\alpha)A_n, \qquad n\neq 0 \tag{A2}$$

$$A_{-M}(K-\alpha)=(K-\alpha)A_{-M}=0, \tag{A3}$$

$$A_{-1}-1=A_0(K-\alpha)=(K-\alpha)A_0. \tag{A4}$$

Repeated use of (A2) leads to

$$A_{-n-1}=(K-\alpha)^nA_{-1}=A_{-1}(K-\alpha)^n, \qquad n>0, \tag{A5}$$

$$A_0=(K-\alpha)^nA_n=A_n(K-\alpha)^n, \qquad n>0. \tag{A6}$$

Equation (A5) together with (A3) implies that

$$(K-\alpha)^MA_{-1}=A_{-1}(K-\alpha)^M=0 \tag{A7}$$

and (A4) together with (A6),

$$A_{-1}-1=A_{M-1}(K-\alpha)^M=(K-\alpha)^MA_{M-1}. \tag{A8}$$

Equations (A7) and (A8) show that

$$(A_{-1}-1)A_{-1}=0. \tag{A9}$$

Thus A_{-1} is a (not necessarily orthogonal) projection. Equation (A7) shows that its range is included in the nullspace of $(K-\alpha)^M$. Equation (A8) implies that the nullspace of $(K-\alpha)^M$ is included in its range. Hence the range of A_{-1} is equal to the nullspace of $(K-\alpha)^M$. Similarly, the nullspace of A_{-1} equals the range of $(K-\alpha)^M$. The operator A_{-1} being a projection whose range equals the dimensionality m of the nullspace of $(K-\alpha)^M$, it follows that[19]

$$\mathrm{Tr}A_{-1}=m. \tag{A10}$$

We now apply the well-known formula

$$\frac{d}{d\gamma}\ln\det A(\gamma)=\mathrm{Tr}A^{-1}(\gamma)\frac{d}{d\gamma}A(\gamma) \tag{A11}$$

to the modified Fredholm determinant and obtain[20]

$$\frac{d}{d\gamma}\ln f(\gamma)=-\gamma\,\mathrm{Tr}K^2(1-\gamma K)^{-1}. \tag{A12}$$

If $f(\gamma)$ has a zero at $\gamma=1/\alpha$ of order p, then the left-hand side of (A12) has a simple pole there, with

$$\lim_{\gamma\to 1/\alpha}\left(\gamma-\frac{1}{\alpha}\right)\frac{d}{d\gamma}\ln f(\gamma)=p. \tag{A13}$$

Hence according to (A12), $\mathrm{tr}K^2(z-K)^{-1}$ must have a simple pole at $z=\alpha$, with

$$\lim_{z\to\alpha}(z-\alpha)\,\mathrm{Tr}K^2(z-K)^{-1}=\alpha^2p. \tag{A14}$$

If the series

$$\mathrm{Tr}K^2(z-K)^{-1}=\sum_{-M}^{\infty}\mathrm{Tr}(K^2A_n)(z-\alpha)^n \tag{A15}$$

converges absolutely, then we may conclude from (A14) that for all $n<-1$,

$$\mathrm{Tr}(K^2A_n)=0, \tag{A16}$$

and that for $n=-1$,

$$\mathrm{Tr}(K^2A_{-1})=\alpha^2p. \tag{A17}$$

The absolute convergence of (A15) is easy to demonstrate. Since (A1) converges absolutely (i.e., in operator norm) in some neighborhood of α, there exist two positive numbers R and C such that $\|A_n\|\leq CR^{-n}$ for all $n\geq -M$. Hence,

$$|\mathrm{Tr}K^2A_n|\leq\|A_n\|(\mathrm{Tr}K^{2\dagger 2})^{1/2}\leq C'R^{-n}$$

and (A15) converges absolutely in some neighborhood of α.

According to (A3), Eq. (A16) implies

$$\mathrm{Tr}A_{-M}=\mathrm{Tr}KA_{-M}=0. \tag{A18}$$

Now it follows from (A2) that if $\mathrm{Tr}A_{n-1}=\mathrm{Tr}KA_{n-1}=0$, then (A16) implies that also $\mathrm{Tr}A_n=\mathrm{Tr}KA_n=0$. Hence (A18) implies by induction that for all $n<-1$,

$$\mathrm{Tr}A_n=\mathrm{Tr}KA_n=0. \tag{A19}$$

Therefore by (A2) for $n=-1$,

$$\begin{aligned}&\mathrm{Tr}KA_{-1}=\alpha\,\mathrm{Tr}A_{-1},\\ &\mathrm{Tr}K^2A_{-1}=\alpha^2\,\mathrm{Tr}A_{-1}=\alpha^2m,\end{aligned} \tag{A20}$$

because of (A10). Comparison of (A17) and (A20) proves that $m=p$, and hence the theorem.

Remark: Since the nullspace of $(K-\alpha)^n$ is a strictly increasing function[21] of n for $1\leq n\leq M$, we may conclude the following as a *corollary* of the theorem:

The dimensionality of the nullspace of $(K-\alpha)$, i.e., the degeneracy of the eigenvalue α, is less than or equal to $p-M+1$.

It would be nice to have a more exact statement about the degeneracy, but there does not appear to be one. If $p=2$, i.e., the zero of f is double, then there are only two possibilities, though: Either the resolvent has a double pole and the degeneracy is one, or the resolvent has a simple pole and the degeneracy is two.

---

* Supported in part by the National Science Foundation, the U.S. Army Research Office, Durham, N.C., and the French CNRS.

† On Sabbatical leave of absence from Indiana University, Bloomington, Indiana 47401.

‡ Postal Address: Département de Physique Mathématiques, Faculté des Sciences, 34—Montpellier, France.
§ Physique Mathématique et Théorique, Equipe de recherche associée au CNRS.
[1] R. Jost and A. Pais, Phys. Rev. **82**, 840 (1951).
[2] See Ref. 1, and for further discussion, for example, Ref. 3.
[3] R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966).
[4] It is easy to see that for $E < 0$, this equation is equivalent to the statement that $\gamma_n(E)$ is an eigenvalue of $K(E)$, because $g(E)$ is positive definite.
[5] G. Hoheisel, *Integral Equations* (Ungar, New York, 1968), p. 79; also W. Pogorzelski, *Integral Equations* (Pergamon, Oxford, 1966), Vol. I, p. 178.
[6] See Ref. 3, p. 345 and p. 375 for the proof that the zero is simple as a function of E. Equation (9. 10) on p. 236 of Ref. 3 shows that it is simple as a function of γ.
[7] The same remark as Ref. 4 applies here.
[8] $Y_l^m(\theta, \varphi)$ is a spherical harmonic as a function of the polar angles θ and φ.
[9] It should be noted that whereas TrK never exists, no matter how well behaved the potential, TrK_l on R_+ exists if the first absolute moment of V is finite.
[10] See, for example, Sec. 9. 1 of Ref. 3.
[11] See Refs. 1 and 3.
[12] See Chap. 12 of Ref. 3.
[13] Because f_l is an analytic function of γ.
[14] That the pole of the resolvent is simple at a simple zero of the Fredholm determinant follows immediately from the Fredholm construction. That there is no degeneracy if the determinant has a simple zero follows from the theorem of the Appendix.
[15] The ascent of an operator A is the smallest integer s such that the nullspaces of A^s and A^{s+1} are equal. Since K_l is compact, the ascent of $1 - \Gamma K_l$ is finite; see, for example, A. E. Taylor, *Introduction to Functional Analysis* (Wiley, New York, 1958), p. 279. For a demonstration that the ascent of $1 - \Gamma K_l$ equals the order of the pole of the resolvent at $\gamma = \Gamma$, see, for example, Ref. 3, p. 208.
[16] See the corollary stated in the Appendix.
[17] The resolvent $[1 - \gamma K]^{-1}$ may be decomposed according to the decomposition of $L^2(R^3)$ into a direct sum of the eigenspaces of the angular part of the Laplacean. Thus

$$[1 - \gamma K]^{-1} = \sum_{l=0}^{\infty} P_l [1 - \gamma K_l]^{-1},$$

where P_l are projections whose kernels are

$$P_l(\theta, \varphi; \theta', \varphi') = \sum_{m=-l}^{l} Y_l^m(\theta, \varphi) Y_l^{-m}(\theta', \varphi').$$

[18] This theorem is well known for finite-dimensional matrices; for Hilbert-Schmidt operators it does not seem to appear in the literature.
[19] "Tr" denotes the trace of the operator. A projection (not necessarily orthogonal) of finite-dimensional range is in the trace class and its trace equals the dimensionality of its range.
[20] The proof of (A12) from (A11) follows from the fact that if K is in the trace class, then $\det_2(1 - \gamma K) = \det(1 - \gamma K) \exp(\mathrm{Tr}\gamma K)$. Equation (A12) therefore holds for K in the trace class, with both sides defined as absolutely convergent power series in γ. Since neither side contains TrK and (A12) is an identity of the two power series whose convergence is assured if K is a Hilbert-Schmidt operator, it must hold also if TrK does not exist.
[21] The intersection of the range of L^m with the nullspace of L equals the image under L^m of the complement of the nullspace of L^m relative to the nullspace of L^{m+1}. Since the range of L^m is a nonincreasing function of m, the nullspace of L^m must be strictly increasing until the nullspaces of L^m and of L^{m+1} are equal. From then on they must remain equal as m increases.

Rank 1 Expansions

Robert Gilmore
Department of Physics and Center for Material Sciences and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
and
Institute for Theoretical Physics, University of Frankfurt, Frankfurt, Germany
(Received 19 August 1971; Revised Manuscript Received 13 December 1971)

A group expansion technique is presented which is valid for all homogeneous Riemannian symmetric coset spaces of rank 1. Examples are given.

I. INTRODUCTION

Group contraction[1-3] and expansion[4-12] has been the subject of great recent interest. This interest is of both a mathematical[4-12] and a physical[13-18] nature:

(i) Unitary irreducible representations of compact semisimple groups can be contracted to UIR for associated inhomogeneous groups.[8-10] Similarly, UIR for noncompact semisimple groups can be constructed from the UIR of an associated inhomogeneous form by the expansion process.[9,10]

(ii) Physically interesting groups can be constructed by contracting semisimple groups.[1-3] Thus, the Poincaré group $ISO(3, 1)$ is the "noncosmological" limit ($R_{\text{universe}} \to \infty$) of the de Sitter groups $SO(3, 2)$ and $SO(4, 1)$, while the Galilean group $G(3)$ is the nonrelativistic limit ($c \to \infty$) of $ISO(3, 1)$. A recently[13] proposed relativistic invariance group $G(3, 1)$ can be regarded as the analogously contracted limit of the conformal group $SO(4, 2)$ under ($m_{\text{photon}} \to 0$, $R_{\text{universe}} \to \infty$). Conversely, the expansion procedure has been used to construct[14-16] semisimple dynamical groups from inhomogeneous relativistic symmetry groups, and to construct[17,18] relativistic position operators.

Although the contraction procedure has been thoroughly studied,[3] only relatively few[5-12] expansions have been successfully carried out. Moreover, each successful expansion has involved a rank 1 coset. Below we show that the simple expansion procedure in common use is valid for all Riemannian symmetric spaces[19] of rank 1.

II. DEFINITION OF EXPANSION

Let M be a (pseudo)-Euclidean space on which a metric-preserving stability group K acts. The group of isometries of M is $G' = M \wedge K$ (semidirect product), and $M \cong G'/K$. We ask: under what conditions can a semisimple group G be found which, contracted with respect to K, gives the group G'? Under this contraction the limit of the (pseudo)-Riemannian symmetric space $P = G/K$ is M:

$$P = G/K \xrightarrow{\text{contraction}} G'/K = M. \qquad \text{(II. 1)}$$

Nonsingular linear transformations do not change the structure of a Lie algebra, since the structure constants are components of a tensor. Singular linear transformations which leave the structure constants well defined also lead to contractions. Therefore, to perform an expansion we must go outside the framework of linear transformations on the generators of a Lie algebra.

‡ Postal Address: Département de Physique Mathématiques, Faculté des Sciences, 34—Montpellier, France.
§ Physique Mathématique et Théorique, Equipe de recherche associée au CNRS.
[1] R. Jost and A. Pais, Phys. Rev. **82**, 840 (1951).
[2] See Ref. 1, and for further discussion, for example, Ref. 3.
[3] R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966).
[4] It is easy to see that for $E < 0$, this equation is equivalent to the statement that $\gamma_n(E)$ is an eigenvalue of $K(E)$, because $g(E)$ is positive definite.
[5] G. Hoheisel, *Integral Equations* (Ungar, New York, 1968), p. 79; also W. Pogorzelski, *Integral Equations* (Pergamon, Oxford, 1966), Vol. I, p. 178.
[6] See Ref. 3, p. 345 and p. 375 for the proof that the zero is simple as a function of E. Equation (9. 10) on p. 236 of Ref. 3 shows that it is simple as a function of γ.
[7] The same remark as Ref. 4 applies here.
[8] $Y_l^m(\theta, \varphi)$ is a spherical harmonic as a function of the polar angles θ and φ.
[9] It should be noted that whereas TrK never exists, no matter how well behaved the potential, TrK_l on R_+ exists if the first absolute moment of V is finite.
[10] See, for example, Sec. 9. 1 of Ref. 3.
[11] See Refs. 1 and 3.
[12] See Chap. 12 of Ref. 3.
[13] Because f_l is an analytic function of γ.
[14] That the pole of the resolvent is simple at a simple zero of the Fredholm determinant follows immediately from the Fredholm construction. That there is no degeneracy if the determinant has a simple zero follows from the theorem of the Appendix.
[15] The ascent of an operator A is the smallest integer s such that the nullspaces of A^s and A^{s+1} are equal. Since K_l is compact, the ascent of $1 - \Gamma K_l$ is finite; see, for example, A. E. Taylor, *Introduction to Functional Analysis* (Wiley, New York, 1958), p. 279. For a demonstration that the ascent of $1 - \Gamma K_l$ equals the order of the pole of the resolvent at $\gamma = \Gamma$, see, for example, Ref. 3, p. 208.
[16] See the corollary stated in the Appendix.
[17] The resolvent $[1 - \gamma K]^{-1}$ may be decomposed according to the decomposition of $L^2(R^3)$ into a direct sum of the eigenspaces of the angular part of the Laplacean. Thus

$$[1 - \gamma K]^{-1} = \sum_{l=0}^{\infty} P_l [1 - \gamma K_l]^{-1},$$

where P_l are projections whose kernels are

$$P_l(\theta, \varphi; \theta', \varphi') = \sum_{m=-l}^{l} Y_l^m(\theta, \varphi) Y_l^{-m}(\theta', \varphi').$$

[18] This theorem is well known for finite-dimensional matrices; for Hilbert-Schmidt operators it does not seem to appear in the literature.
[19] "Tr" denotes the trace of the operator. A projection (not necessarily orthogonal) of finite-dimensional range is in the trace class and its trace equals the dimensionality of its range.
[20] The proof of (A12) from (A11) follows from the fact that if K is in the trace class, then $\det_2(1 - \gamma K) = \det(1 - \gamma K) \exp(\mathrm{Tr}\gamma K)$. Equation (A12) therefore holds for K in the trace class, with both sides defined as absolutely convergent power series in γ. Since neither side contains TrK and (A12) is an identity of the two power series whose convergence is assured if K is a Hilbert-Schmidt operator, it must hold also if TrK does not exist.
[21] The intersection of the range of L^m with the nullspace of L equals the image under L^m of the complement of the nullspace of L^m relative to the nullspace of L^{m+1}. Since the range of L^m is a nonincreasing function of m, the nullspace of L^m must be strictly increasing until the nullspaces of L^m and of L^{m+1} are equal. From then on they must remain equal as m increases.

Rank 1 Expansions

Robert Gilmore
Department of Physics and Center for Material Sciences and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
and
Institute for Theoretical Physics, University of Frankfurt, Frankfurt, Germany
(Received 19 August 1971; Revised Manuscript Received 13 December 1971)

A group expansion technique is presented which is valid for all homogeneous Riemannian symmetric coset spaces of rank 1. Examples are given.

I. INTRODUCTION

Group contraction[1-3] and expansion[4-12] has been the subject of great recent interest. This interest is of both a mathematical[4-12] and a physical[13-18] nature:

(i) Unitary irreducible representations of compact semisimple groups can be contracted to UIR for associated inhomogeneous groups.[8-10] Similarly, UIR for noncompact semisimple groups can be constructed from the UIR of an associated inhomogeneous form by the expansion process.[9,10]

(ii) Physically interesting groups can be constructed by contracting semisimple groups.[1-3] Thus, the Poincaré group $ISO(3, 1)$ is the "noncosmological" limit ($R_{\text{universe}} \to \infty$) of the de Sitter groups $SO(3, 2)$ and $SO(4, 1)$, while the Galilean group $G(3)$ is the nonrelativistic limit ($c \to \infty$) of $ISO(3, 1)$. A recently[13] proposed relativistic invariance group $G(3, 1)$ can be regarded as the analogously contracted limit of the conformal group $SO(4, 2)$ under ($m_{\text{photon}} \to 0$, $R_{\text{universe}} \to \infty$). Conversely, the expansion procedure has been used to construct[14-16] semisimple dynamical groups from inhomogeneous relativistic symmetry groups, and to construct[17,18] relativistic position operators.

Although the contraction procedure has been thoroughly studied,[3] only relatively few[5-12] expansions have been successfully carried out. Moreover, each successful expansion has involved a rank 1 coset. Below we show that the simple expansion procedure in common use is valid for all Riemannian symmetric spaces[19] of rank 1.

II. DEFINITION OF EXPANSION

Let M be a (pseudo)-Euclidean space on which a metric-preserving stability group K acts. The group of isometries of M is $G' = M \wedge K$ (semidirect product), and $M \cong G'/K$. We ask: under what conditions can a semisimple group G be found which, contracted with respect to K, gives the group G'? Under this contraction the limit of the (pseudo)-Riemannian symmetric space $P = G/K$ is M:

$$P = G/K \xrightarrow{\text{contraction}} G'/K = M. \qquad \text{(II. 1)}$$

Nonsingular linear transformations do not change the structure of a Lie algebra, since the structure constants are components of a tensor. Singular linear transformations which leave the structure constants well defined also lead to contractions. Therefore, to perform an expansion we must go outside the framework of linear transformations on the generators of a Lie algebra.

Definition: An expansion of a (nonsemisimple) Lie algebra $\mathfrak{g}'$ is a nonlinear mapping of its generators X_μ:

$$X_\mu \to f_\mu(X), \tag{II. 2}$$

which closes under commutation

$$[f_\mu(X), f_\nu(X)] = c_{\mu\nu}^{\lambda} f_\lambda(X), \tag{II. 3}$$

When $c_{\mu\nu}^{\lambda}$ are structure constants for a (semisimple) algebra $\mathfrak{g}$ we call this a (semisimple) expansion of $\mathfrak{g}'$ to $\mathfrak{g}$.

We shall be specifically interested in the case where $\mathfrak{g}'$ has a semidirect-product structure $\mathfrak{m} \wedge \mathfrak{k}$, and $\mathfrak{g}$ is semisimple.

III. CARTAN DECOMPOSITION

Let $\mathfrak{k}$ be a maximal subalgebra of semisimple $\mathfrak{g}$ in the sense (III. 1)

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}:$$

$$[\mathfrak{k}, \mathfrak{k}] \subseteq \mathfrak{k}, \tag{III. 1a}$$

$$[\mathfrak{k}, \mathfrak{p}] = \mathfrak{p}, \tag{III. 1b}$$

$$[\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{k}, \tag{III. 1c}$$

$$g(\mathfrak{k}, \mathfrak{p}) = 0. \tag{III. 1d}$$

Here the Riemannian spaces G, K, P are related to the vector spaces $\mathfrak{g}, \mathfrak{k}, \mathfrak{p}$ by the exponential map.[19] In particular, a Cartan decomposition possesses this property and in addition the properties

$$g(\mathfrak{k}, \mathfrak{k}) < 0, \quad g(\mathfrak{p}, \mathfrak{p}) > 0. \tag{III. 2}$$

Therefore, the algebra $\mathfrak{g}_* = \mathfrak{k} + \mathfrak{p}_*$ ($\mathfrak{p}_* = i\mathfrak{p}$) is compact. The coset spaces P and P_* are dual.

Let bases for $\mathfrak{g}, \mathfrak{k}, \mathfrak{p}$ be given by $\mathfrak{g}$: $X_\mu, X_\nu, \ldots$, late Greek letters; $\mathfrak{k}$: $X_\alpha, X_\beta, \ldots$, early Greek letters; $\mathfrak{p}$: $X_i, X_j, \ldots$, Latin letters.

The structure constants for $\mathfrak{g}$ are

$$[X_\alpha, X_\beta] = c_{\alpha\beta}^{\gamma} X_\gamma + 0, \tag{III. 3a}$$

$$[X_\alpha, X_j] = 0 + c_{\alpha j}^{k} X_k, \tag{III. 3b}$$

$$[X_i, X_j] = c_{ij}^{\gamma} X_\gamma + 0. \tag{III. 3c}$$

The components of the metric tensor $g_{\mu\nu}$ are

$$g_{\alpha\beta} = c_{\alpha\mu}^{\nu} c_{\beta\nu}^{\mu} = c_{\alpha\gamma}^{\delta} c_{\beta\delta}^{\gamma} + c_{\alpha k}^{l} c_{\beta l}^{k}, \tag{III. 4a}$$

$$g_{\alpha j} = c_{\alpha\mu}^{\nu} c_{j\nu}^{\mu} = 0, \tag{III. 4b}$$

$$g_{ij} = c_{i\mu}^{\nu} c_{j\nu}^{\mu} = c_{ik}^{\gamma} c_{j\gamma}^{k} + c_{i\gamma}^{k} c_{jk}^{\gamma}. \tag{III. 4c}$$

Since $\mathfrak{g}$ is semisimple,

$$0 \neq \|g_{\mu\nu}\| = \det\begin{pmatrix} g_{\alpha\beta} & 0 \\ 0 & g_{ij} \end{pmatrix} = \|g_{\alpha\beta}\| \, \|g_{ij}\| \neq 0. \tag{III. 5}$$

Thus, the metric tensors g_{ij} defined on $\mathfrak{p}, P$ are nonsingular. The block diagonal structure of $g_{\mu\nu}$ implies also the block diagonal structure of its inverse

$$g^{\mu\nu} = \begin{pmatrix} g^{\alpha\beta} & 0 \\ 0 & g^{ij} \end{pmatrix}. \tag{III. 6}$$

IV. SOME USEFUL RELATIONS

The Casimir invariant $\mathcal{C}_2(\mathfrak{g})$ is given by

$$g^{\mu\nu} X_\mu X_\nu = g^{\alpha\beta} X_\alpha X_\beta + g^{ij} X_i X_j. \tag{IV. 1}$$

The commutation properties of this invariant with the subspace $\mathfrak{k}$ can be written

$$[X_\alpha, g^{\gamma\beta} X_\gamma X_\beta] = 0 = -[X_\alpha, g^{ij} X_i X_j]. \tag{IV. 2}$$

The left-hand side (lhs) is a homogeneous second-order polynomial in the generators of K, the right-hand side (rhs), in the bases of $\mathfrak{p}$. By separation of variables arguments, both sides vanish independently. This argument is not valid for the subspace $\mathfrak{p}$, since both sides are then homogeneous second-order polynomials, linear in both $\mathfrak{k}$ and $\mathfrak{p}$. Instead, we find the well-known identity

$$\begin{aligned} [X_i, g^{\gamma\beta} X_\gamma X_\beta] &= -[X_i, g^{kl} X_k X_l], \\ g^{\gamma\beta} \overset{\downarrow}{c}_{i\gamma}^{t} &= -g^{tl} \overset{\downarrow}{c}_{il}^{\beta}. \end{aligned} \tag{IV. 3}$$

The Jacobi identity is

$$\begin{aligned} c_{rs}^{\gamma} c_{\gamma\beta}^{\delta} &= c_{rs}^{\mu} c_{\mu\beta}^{\delta} \\ &= -(c_{\beta r}^{\mu} c_{\mu s}^{\delta} + c_{s\beta}^{\mu} c_{\mu r}^{\delta}) \\ &= -(c_{\beta r}^{t} c_{ts}^{\delta} + c_{s\beta}^{t} c_{tr}^{\delta}). \end{aligned} \tag{IV. 4}$$

V. RANK 1 COSETS

The curvature of a Riemannian space can be defined in a coordinate-free and a coordinate-dependent way:

$$\begin{aligned} & \overset{\qquad\qquad R(X_\mu, X_\nu)X_\lambda}{} \\ & \overset{19 \,\|\qquad\qquad\|\, 20}{} \\ & -[[X_\mu, X_\nu], X_\lambda] = R_{\mu\nu;\lambda}^{\kappa} X_\kappa. \end{aligned} \tag{V. 1}$$

For the Riemannian symmetric spaces $P_{(*)} = \exp \mathfrak{p}_{(*)}$ with metric g_{ij}, (V. 1) becomes

$$-c_{ij}^{\alpha} c_{\alpha k}^{l} = R_{ij,k}^{l}. \tag{V. 2}$$

Coset spaces of rank 1 have constant sectional curvature ρ when the decomposition (III. 1) is also a Cartan decomposition [obeys (III. 2)]. Such spaces have the additional property[19-21]

$$R(X, Y)Z = \rho\{g(Z, Y)X - g(Z, X)Y\}. \tag{V. 3}$$

The sectional curvature is given by[21]

$$\begin{aligned} \rho &= g(R(X_i, X_j)X_i, X_j) \\ &= -c_{ij}^{\gamma} c_{\gamma i}^{k} g_{kj} \\ &= -\sum_{\gamma, \xi} c_{ij}^{\gamma} c_{ij}^{\xi} g_{\gamma\xi} \quad (i \neq j \text{ fixed}). \end{aligned} \tag{V. 4}$$

So for spaces of rank 1 we have the following identity:

$$\begin{aligned} & R_{ij;k}^{l} \\ & \| \quad \| \\ & -c_{ij}^{\alpha} c_{\alpha k}^{l} = \rho\{g_{kj}\delta_i^{l} - g_{ki}\delta_j^{l}\}. \end{aligned} \tag{V. 5}$$

VI. A SIMPLE CONTRACTION

We now contract $\mathfrak{g} \to \mathfrak{g}'$ with respect to the subalgebra $\mathfrak{k}$:

$$\begin{aligned} &\mathfrak{k} \to \mathfrak{k}, \quad X_\alpha \to Y_\alpha = X_\alpha, \\ &\mathfrak{p} \to \mathfrak{m} = \lim_{\epsilon\to 0} \epsilon\mathfrak{p}, \quad X_i \to Y_i = \lim_{\epsilon\to 0} \epsilon X_i. \end{aligned} \tag{VI. 1}$$

The commutation relations of $\mathfrak{g}'$ are

$$[Y_\alpha, Y_\beta] = d^\gamma_{\alpha\beta} Y_\gamma, \quad d^\gamma_{\alpha\beta} = c^\gamma_{\alpha\beta}, \tag{VI. 2a}$$

$$[Y_\alpha, Y_j] = d^k_{\alpha j} Y_k, \quad d^k_{\alpha j} = c^k_{\alpha j}, \tag{VI. 2b}$$

$$[Y_i, Y_j] = 0, \quad d^\gamma_{ij} = \epsilon^2 c^\gamma_{ij} \to 0. \tag{VI. 2c}$$

The only information lost in the contraction $\mathfrak{g} \to \mathfrak{g}'$ is contained in the structure constants c^γ_{ij}. The metric tensor $g_{\alpha\beta}$ of $\mathfrak{g}$ can be reconstructed from the structure constants d using (III. 4a), but the metric g_{ij} cannot be reconstructed using (III. 4c) because $d^\gamma_{ij} = 0$.

The second-order Casimir invariant (IV. 1) on $\mathfrak{g}$ can be contracted to the second-order Casimir invariant on $\mathfrak{g}'$, which is *uniquely* defined up to a constant numerical factor [h in (VI. 4)]:

$$\begin{aligned} \mathcal{C}'_2(\mathfrak{g}') &= \lim \epsilon^2 \mathcal{C}_2(\mathfrak{g}) \\ &= \lim_{\epsilon\to 0} g^{\alpha\beta}(\epsilon X_\alpha)(\epsilon X_\beta) + \lim_{\epsilon\to 0} g^{ij}(\epsilon X_i)(\epsilon X_j) \\ &= g^{ij} Y_i Y_j. \end{aligned} \tag{VI. 3}$$

The uniqueness of $\mathcal{C}'_2(\mathfrak{g}')$ guarantees that *any* second-order homogeneous polynomial invariant operator $\mathcal{C}^2(Y_\alpha, Y_j)$ which commutes with all generators Y_β, Y_k of $\mathfrak{g}'$, must be proportional to $\mathcal{C}'_2(\mathfrak{g}')$

$$\mathcal{C}^2(Y_\alpha, Y_j) = h g^{ij} Y_i Y_j. \tag{VI. 4}$$

Without loss of generality, we can choose $h = +1$. The choice $h = -1$ is equivalent to a discussion of the dual algebra $\mathfrak{g}_*$.

Construction of an invariant (VI. 4) uniquely defines the metric tensor g^{ij} and its inverse. Thus, by (IV. 3) the information about the structure constants c^γ_{ij} is not really lost at all by contraction, since

$$c^\gamma_{ij} = g^{\gamma\beta} d^k_{\beta i} g_{kj}. \tag{VI. 5}$$

Therefore, it should be possible to construct the original algebras $\mathfrak{g}, \mathfrak{g}_*$ from the contracted algebras $\mathfrak{g}' = \mathfrak{g}'_*$.

VII. SOME COMMUTATION RELATIONS

The kernal of the contraction (VI. 3) is the invariant (IV. 2) of $\mathfrak{k}$. In $\mathfrak{g}'$ we have

$$\begin{aligned} &[Y_\alpha, g^{\gamma\beta} Y_\gamma Y_\beta] = 0, \\ &[Y_i, g^{\gamma\beta} Y_\gamma Y_\beta] = g^{\gamma\beta} d^\tau_{i\gamma} \{Y_\tau, Y_\beta\}. \end{aligned} \tag{VII. 1}$$

We define now a third set of generators by

$$\begin{aligned} &Z_\alpha = Y_\alpha, \quad (= X_\alpha) \\ &Z_i = [Y_i, g^{\gamma\beta} Y_\gamma Y_\beta]. \end{aligned} \tag{VII. 2}$$

The commutation properties (VII. 3) of these generators are easily obtained:

$$[Z_\alpha, Z_\beta] = c^\gamma_{\alpha\beta} Z_\gamma, \tag{VII. 3a}$$

$$\begin{aligned} [Z_\alpha, Z_j] &= [Y_\alpha, [Y_j, I]] \\ &= -[I, [Y_\alpha, Y_j]] - [Y_j[I, Y_\alpha]] \\ &= c^k_{\alpha j} Z_k \qquad 0. \end{aligned} \tag{VII. 3b}$$

We have indicated explicitly in the computation above that the result is valid for *any* invariant I of $\mathfrak{k}$.

Finally, we compute the commutator of Z_i with Z_j. The terms arising from

$$[\{Y_\tau, Y_\beta\}, \{Y_s, Y_\gamma\}] \tag{VII. 4}$$

are of four types:

$$\begin{aligned} &(1)\ [Y_\tau Y_\beta, Y_s Y_\eta] \to A_1^{ll'g} Y_l Y_{l'} Y_g, \\ &(2)\ [Y_\tau Y_\beta, Y_\eta Y_s] \to A_2^{ll'g} Y_l Y_g Y_{l'}, \\ &(3)\ [Y_\beta Y_\tau, Y_s Y_\eta] \to A_3^{ll'g} Y_{l'} Y_g Y_l, \\ &(4)\ [Y_\beta Y_\tau, Y_\eta Y_s] \to A_4^{ll'g} Y_g Y_l Y_{l'}. \end{aligned} \tag{VII. 5}$$

A moment's consideration will reveal that for fixed Latin (l, l') and Greek (g) indices

$$A_1^{ll'g} = A_2^{ll'g} = A_3^{ll'g} = A_4^{ll'g}.$$

The first and fourth terms together form a symmetrized combination, as do the second and third. Moreover,

$$Y_l Y_g Y_{l'} = Y_l([Y_g, Y_{l'}] + Y_{l'} Y_g) = Y_l Y_{l'} Y_g + Y_l c^t_{g l'} Y_t, \tag{VII. 6}$$

$$\begin{aligned} &Y_{l'} Y_g Y_l = ([Y_{l'}, Y_g] + Y_g Y_{l'}) Y_l = Y_g Y_l Y_{l'} + Y_t c^t_{l'g} Y_t, \\ &Y_l Y_g Y_{l'} + Y_{l'} Y_g Y_l = Y_l Y_{l'} Y_g + Y_g Y_l Y_{l'}. \end{aligned} \tag{VII. 7}$$

Therefore, it is sufficient to compute only the first type of term in the commutator.

We must study

$$\begin{aligned} &g^{\gamma\beta} d^\tau_{i\gamma} g^{\xi\eta} d^s_{j\xi} [Y_\tau Y_\beta, Y_s Y_\eta] \\ &\qquad = P^{\tau t,\eta}_{ij} Y_\tau Y_t Y_\eta + Q^{ts,\beta}_{ij} Y_t Y_s Y_\beta + R^{\tau s,\lambda}_{ij} Y_\tau Y_s Y_\lambda. \end{aligned} \tag{VII. 8}$$

We compute these tensors, term by term:

$$\begin{aligned} P^{\tau t,\eta}_{ij} &= \underline{g^{\gamma\beta} d^\tau_{i\gamma}} g^{\xi\eta} \underline{d^s_{j\xi} d^t_{\beta s}} \\ &= g^{\xi\eta} d^s_{j\xi} \underline{d^\tau_{i\gamma} (c^\gamma_{st'} g^{t't})} \qquad \text{[by (IV. 3)]} \\ &= g^{\xi\eta} d^s_{j\xi} \{\rho(g_{t'i} \delta^\tau_s - g_{is} \delta^\tau_{t'})\} g^{t't} \qquad \text{[by (V. 5)]} \\ &= -\rho(g^{\xi\eta} d^s_{j\xi} g_{si}) g^{\tau t} + \rho g^{\xi\eta} d^\tau_{j\xi} \delta^t_i; \end{aligned} \tag{VII. 9a}$$

similarly,

$$Q^{ts,\beta}_{ij} = \rho(g^{\gamma\beta} d^\tau_{i\gamma} g_{\tau j}) g^{ts} - \rho g^{\gamma\beta} d^s_{i\gamma} \delta^t_j. \tag{VII. 9b}$$

Finally, we come to the most involved calculation:

$$\begin{aligned} R^{\tau s,\lambda}_{ij} &= \underline{g^{\gamma\beta} d^\tau_{i\gamma} g^{\xi\eta} d^s_{j\xi} d^\lambda_{\beta\eta}} \\ &= \underline{(g^{\tau' \tau} c^\beta_{\tau' i})\ (g^{s's} c^\eta_{s' j})(-c^\lambda_{\eta\beta})} \qquad \text{[by (IV. 3)}^2\text{]} \\ &= g^{\tau'\tau} g^{s's} \underline{\underline{c^\beta_{\tau' i}}} (\underline{c^t_{\beta s}} c^\lambda_{tj} + \underline{c^t_{j\beta}} c^\lambda_{ts'}) \qquad \text{[by (IV. 4)]} \\ &= \rho g^{\tau'\tau} g^{s's} \{ c^\lambda_{tj} (g_{s'\tau} \delta^t_i - g_{s'i} \delta^t_\tau) \\ &\quad - c^\lambda_{ts'} (g_{j\tau} \delta^t_i - g_{ji} \delta^t_\tau) \} \qquad \text{[by (V. 5)}^2\text{]} \end{aligned}$$

TABLE I. Examples of rank 1 expansions

System	$G = \exp \mathfrak{g}$	$K = \exp \mathfrak{k}$	$\mathfrak{g}' = \mathfrak{m} \oplus \mathfrak{k}$ $G' = M \wedge K$	$\mathfrak{g}' = \mathfrak{m} \oplus \mathfrak{k}$ $G' = M \wedge K$	$K = \exp \mathfrak{k}$	$G = \exp \mathfrak{g}$	Comments and references
B_n, D_n	$SO(p, q)$	$SO(p-1, q)$ $SO(p, q-1)$	$ISO(p-1, q)$ $ISO(p, q-1)$	$ISO(p, q)$	$SO(p, q)$	$SO(p+1, q)$ $SO(p, q+1)$	Metric preserving groups groups in real linear vector spaces (Refs. 5–10)
A_n	$SU(p, q)$	$U(p-1, q)$ $U(p, q-1)$	$IU(p-1, q)$ $IU(p, q-1)$	$IU(p, q)$	$U(p, q)$	$SU(p+1, q)$ $SU(p, q+1)$	Metric preserving groups in complex linear vector spaces (Refs. 5–10)
C_n	$Sp(p, q)$	$Sp(p-1, q)$ $Sp(p, q-1)$	$ISp(p-1, q)$ $ISp(p, q-1)$	$ISp(p, q)$	$Sp(p, q)$	$Sp(p+1, q)$ $Sp(p, q+1)$	Metric preserving groups in quaternion linear vector spaces (Refs. 11 and 12)
F_4	$F_{4(-20)}$ $F_{4(-52)}$	$SO(9)$	$M \wedge SO(9)$	$M \wedge SO(9)$	$SO(9)$	$F_{4(-20)}$ $F_{4(-52)}$	

$$= \rho c_{ij}^{\lambda} g^{rs} - \rho g^{rt} c_{ij}^{\lambda} \delta_i^s - \rho g^{rt} c_{tj}^{\lambda} \delta_i^s + \rho g^{r'r} g^{s's} c_{r's'}^{\lambda} g_{ij}. \quad \text{(VII. 9c)}$$

Summing Eqs. (VII. 9) gives

$$(P_{ij}^{rs\lambda} + Q_{ij}^{rs\lambda} + R_{ij}^{rs\lambda}) Y_r Y_s Y_\lambda = -\rho c_{ij}^{\lambda} (g^{rs} Y_r Y_s) Y_\lambda. \quad \text{(VII. 10)}$$

Thus we have

$$[Z_i, Z_j] = c_{ij}^{\lambda} \{-2\rho g^{rs} Y_r Y_s, Y_\lambda\}. \quad \text{(VII. 3c)}$$

VIII. EXPANSION

The infinitesimal generators

$$Z'_\alpha = Z_\alpha, \qquad Z'_i = Z_i / \{-4\rho g^{rs} Y_r Y_s\}^{1/2} \quad \text{(VIII. 1)}$$

close under commutation to give the algebra $\mathfrak{g}$. Similarly, the generators

$$Z'_\alpha = Z_\alpha, \qquad Z'_i = Z_i / \{+4\rho g^{rs} Y_r Y_s\}^{1/2} \quad \text{(VIII. 2)}$$

give the algebra $\mathfrak{g}_*$.

IX. EXAMPLES

Examples of rank 1 expansions are given in Table I.

X. COMMENTS

(i) Strictly speaking, the Jacobi identity used in (VII. 3b) is valid only in a Lie algebra or in an associative algebra. We have used the Jacobi identity, not in the Lie algebra, but in its universal enveloping algebra. Since the enveloping algebra is associative,[19] the Jacobi identity can be proved by expanding the Lie brackets and rearranging terms.

(ii) The expansion procedure presented here is valid for Cartan decompositions [(III. 1) and (III. 2)] only, since it depends on constant sectional curvature. But it actually depends on the properties of the structure constants rather than the explicit real form chosen. Thus, the construction is valid for the complex extension $\tilde{\mathfrak{g}}^c$ of $\mathfrak{g}$ and any real form of $\mathfrak{g}^c$ with some real form of $\mathfrak{k}^c$ as a maximal [(III. 1)] subalgebra. This is why we are able to give as examples in Table I algebras with $\mathfrak{k}$ noncompact. All reductive coset decompositions (III. 1) have been given by Berger.[22]

(iii) Contractions with respect to a maximal subalgebra $\mathfrak{k}$ (III. 1) are information preserving in the sense of (VI. 2)–(VI. 5). Since no information is lost in the contraction $\mathfrak{g}_{(*)} \to \mathfrak{g}'$, it should be possible to construct rank r expansions $\mathfrak{g}' \to \mathfrak{g}_{(*)}$. It has not yet been possible to construct rank r expansions for $r > 1$.

XI. CONCLUSION

A unified approach has been given to the expansion problem for algebras with rank 1 cosets. It is seen why the expansion algorithm characteristic of all the recent literature is valid only for those cases studied: specifically, expansions involving Riemannian symmetric spaces with constant sectional curvature (rank 1).

1 I. E. Segal, Duke Math. J. **18**, 221 (1951).
2 E. Inönü and E. P. Wigner, Proc. Natl. Acad. Sci. (U.S.) **39**, 510 (1953).
3 E. J. Saletan, J. Math. Phys. **2**, 1 (1961).
4 M. A. Melvin, Bull. Amer. Phys. Soc. **7**, 493 (1962); **8**, 356 (1963).
5 J. Rosen, Nuovo Cimento **46B**, 1 (1966).
6 J. Rosen and P. Roman, J. Math. Phys. **7**, 2072 (1966).
7 J. Rosen, J. Math. Phys. **9**, 1305 (1968).
8 A. Chakrabarti, J. Math. Phys. **9**, 2087 (1968).
9 A. Sankaranarayanan, J. Math. Phys. **9**, 611 (1968).
10 K. B. Wolf, J. Math. Phys. **12**, 197 (1971).
11 J. G. Nagel and K. T. Shah, J. Math. Phys. **11**, 1483 (1970).
12 J. G. Nagel, J. Math. Phys. **11**, 1779 (1970).
13 J. J. Aghassi, P. Roman, and R. M. Santilli, Phys. Rev. D **1**, 2753 (1970).
14 A. Sankaranarayanan, Nuovo Cimento **38**, 1441 (1965).
15 A. Böhm, Phys. Rev. **145**, 1212 (1966).
16 M. Y. Han, Nuovo Cimento **42B**, 367 (1966).
17 A. Sankaranarayanan and R. H. Good, Jr., Phys. Rev. **140B**, 509 (1965).
18 L. Castel, Nuovo Cimento **49**, 285 (1967).
19 S. Helgason *Differential Geometry and Symmetric Spaces* (Academic, New York, 1962), p. 180.
20 S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* (Interscience, New York, 1963), Vol. I, p. 144.
21 Reference 20, Vol. I, p. 203.
22 M. Berger, Ann. Sci. Ecole Norm. Sup. **74**, 85 (1957).

Analytic Properties of the Free Energy for the "Ice" Models

M. L. Glasser
Battelle Memorial Institute, Columbus, Ohio 43201
and
D. B. Abraham*† and E. H. Lieb†
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
(Received 4 October 1971)

The calculation of the free energy for the two-dimensional lattice models which obey the "ice" condition is reviewed and summarized. The analytic properties of this function relevant to thermodynamics are obtained for the set of those models corresponding to the absence of external fields. Detailed complex temperature Reimann structures are presented for the KDP and F models.

1. INTRODUCTION

The purpose of this paper is to analyze the thermodynamic properties of the planar "ice rule" ferroelectric models, which are isomorphic to a lattice statistical problem defined by the following rules:

1. On every edge of an $N \times M$ square lattice an arrow is placed, pointing in one of two possible directions;

2. The canonical partition sum is restricted to those configurations in which there are two, and only two, arrows pointing towards each and every vertex. Thus each vertex must be one of the six types shown in Fig. 1.

3. The partition sum of the associated ferroelectric is given by

$$z = \sum_{\substack{\text{allowed} \\ \text{config.}}} \Pi_j w_j^{n_j},$$

where n_j is the number of times vertex j occurs in a given configuration; this configuration has energy e_j and the associated canonical weight

$$w_j = e^{-\beta e_j},$$

where $\beta = 1/kT$, in the customary notation. The e_j are shown in Fig. 1.

This paper reports the investigation of the dependence of the ordered state and associated phase transition, where it exists, on the two parameters ϵ_1 and ϵ_2 defined below. The exponents are the same throughout each region and there is no nonphysical dependence of an accidental nature on (ϵ_1/ϵ_2).

We also discuss the Riemann surfaces for the energy in the complex temperature plane. The free energy is evaluated in terms of elementary functions in two special cases.

2. REVIEW OF THE BASIC THEORY

To complete the description of the model, adumbrated in the introduction, we shall impose periodic boundary conditions both vertically and horizontally, although the former is not essential. We shall refer to a row of vertical arrows simply as a row, which then has the possibility of being in 2^N states ϕ. Rather than specifying the states of the horizontal arrows between the rows, it is sufficient to define a transfer matrix $T(\phi, \phi') = \sum e^{-\beta E}(\beta = 1/kT)$, where the sum is over all allowed configurations of horizontal arrows consistent with the lower row being in state ϕ' and the next higher row being in state ϕ; E is the sum of the vertex energies formed by the two states ϕ, ϕ' together with the horizontal arrows. From the definition of the partition function we have simply

$$Z(\beta) = \sum_{\phi_M} \cdots \sum_{\phi_1} T(\phi_1, \phi_M)\, T(\phi_M, \phi_{M-1}) \cdots T(\phi_2, \phi_1)$$
$$= \mathrm{Tr}[T^M] = \sum \lambda_j^M, \quad (2.1)$$

where λ_j are the eigenvalues of T. Since the elements of T are nonnegative, the Perron–Fröbenius theorem[1] guarantees that there is a positive maximum eigenvalue Λ which corresponds to an eigenvector having nonnegative components. We have simply

$$\lim_{M\to\infty} (1/M) \ln Z = \ln\Lambda. \quad (2.2)$$

We must thus formulate and solve the equation $T\Psi = \Lambda\Psi$.

It has been demonstrated[2] that horizontal periodicity and the ice rule are sufficient to ensure that $T(\phi,\phi') = 0$ unless ϕ and ϕ' have the same number n of down arrows, so that T decomposes into blocks of dimension $\binom{N}{n}, n = 0, 1, \ldots, N$. It is convenient to define an average polarization

$$y = 1 - 2n/N, \quad -1 \le y \le 1. \quad (2.3)$$

Since we need only solve the eigenvalue equation in each block separately, it is convenient to denote the corresponding eigenvector by $f(x_1, \ldots, x_n)$, where $1 \le x_1 \le x_2 \le \cdots \le x_n \le N$ denote the positions of the down arrows. We shall also denote the n-tuple simply by $\mathbf{X}$. In a given matrix element, the first down arrow can occur in either the lower or the upper row, and so accordingly we can write

$$T = T_R + T_L, \quad (2.4)$$

where

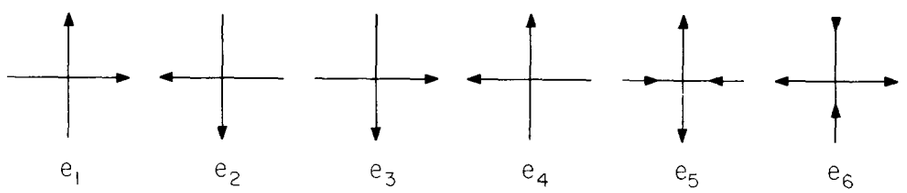


FIG. 1. The six allowed vertex configurations.

$$T_R(\mathbf{X},\mathbf{Y}) = \begin{cases} \neq 0, & \text{if} x_{i-1} \le y_i \le x_i, \quad i = 1,\ldots,n, \\ 0, & \text{otherwise}, \end{cases}$$

$$T_L(\mathbf{X},\mathbf{Y}) = \begin{cases} \neq 0, & \text{if} x_i \le y_i \le x_{i+1}, \\ 0, & \text{otherwise}. \end{cases} \tag{2.5}$$

(The subscripts R, L refer to the direction of the horizontal arrows at the extremities.)

Because the zero of energy may be chosen freely and n-conservation implies that the number of type 5 vertices must be the same as the number of type 6 vertices on every row, there can be only four independent vertex energies. We take as these a linear combination of the vertex energies for the KDP[3] and F[4] models together with energies corresponding to the presence of direct (horizontal and vertical) fields:

$$\begin{aligned} &e_1 = \epsilon_2 - h - v, \quad e_2 = \epsilon_2 + h + v, \\ &e_3 = \epsilon_1 + \epsilon_2 - h + v, \\ &e_4 = \epsilon_1 + \epsilon_2 + h - v, \quad e_5 = e_6 = \epsilon_1. \end{aligned} \tag{2.6}$$

Now, because the configuration energies are linear in the vertex energies, for each element of the transfer matrix, we have $T = T^{\mathrm{KDP}} T^{\mathrm{F}} T^{H} T^{V}$.

For convenience we define the quantities

$$K_i = \beta\epsilon_i, \quad H = \beta h, \quad V = \beta v. \tag{2.7}$$

Then, by inspection, we find

$$\begin{aligned} &T_R^V = e^{NVy} = T_L^V, \\ &T_R^H = e^{NH} e^{2H\sum_1^n (y_i - x_i)} = e^{2NH} T_L^H(\mathbf{X},\mathbf{Y}). \end{aligned} \tag{2.8}$$

The transfer matrices for the F and KDP models are[3,4]

$$\begin{aligned} &T_R^{\mathrm{F}}(\mathbf{X},\mathbf{Y}) = T_L^{\mathrm{F}}(\mathbf{X},\mathbf{Y}) \\ &\qquad = \exp\left[-K_2\left(N - 2n + 2\sum_1^n\sum_1^n \delta(x_i - y_i)\right)\right] \\ &T_R^{\mathrm{KDP}}(\mathbf{X},\mathbf{Y}) \\ &\qquad = \exp K_1\left(-n + \sum_1^n (y_i - x_i) + 2\sum_1^{n-1} \delta(y_{i+1} - x_i)\right) \\ &T_L^{\mathrm{KDP}}(\mathbf{X},\mathbf{Y}) \\ &\qquad = \exp K_1\left(-N - n + \sum_1^n [y_i - x_i + 2\delta(y_i - x_i)]\right). \end{aligned} \tag{2.9}$$

We must now solve the eigenvalue equation

$$\Lambda f(\mathbf{X}) = \sum_R T_R(\mathbf{X},\mathbf{Y}) f(\mathbf{Y}) + \sum_L T_L(\mathbf{X},\mathbf{Y}) f(\mathbf{Y}), \tag{2.10}$$

where the sums denoted R and L are over all $1 \le y_1 \le \cdots \le y_n \le N$ consistent with the first or second condition in (2.5), respectively. The solution of (2.10) follows precisely the treatment of the ice model[2] to which we refer for details. We find that

$$\begin{aligned} &\Lambda = \Lambda_R + \Lambda_L, \\ &\Lambda_R = e^{N(Vy - K_2 + H)} \prod_{j=1}^n \frac{2\Delta - e^{K_1} - e^{ik_j + 2H}}{1 - \exp(ik_j - K_1 - 2H)}, \\ &\Lambda_L = e^{N(Vy - K_1 - K_2 - H)} \prod_{j=1}^n \frac{2\Delta - e^{-K_1} - e^{-ik_j - 2H}}{1 - \exp(-ik_j - K_1 - 2H)}, \end{aligned} \tag{2.11}$$

where

$$2\Delta = e^{K_1} + e^{-K_1} - e^{2K_2 - K_1}. \tag{2.12}$$

The numbers k_j are obtained by solving the n-equations

$$e^{ik_j N} = \prod_{\substack{i=1 \\ i \ne j}}^n B(k_j, k_i), \quad j = 1,\ldots,n, \tag{2.13}$$

where

$$\begin{aligned} B(q,p)^{-1} = B(p,q) &= e^{-i\theta(p,q)} \\ &= \frac{1 + e^{4H + i(p+q)} - 2\Delta e^{2H + ip}}{1 + e^{4H + i(p+q)} - 2\Delta e^{2H + iq}}. \end{aligned} \tag{2.14}$$

When H is nonzero, the function θ is not real so that the numbers k_j are complex. What happens in this case has been explored briefly by Sutherland, Yang, and Yang,[5] but the details of their calculation are not available. In the remainder of this paper we shall assume that $H = 0$.

The important properties of the solution to (2.13), (2.14) which maximizes Λ for any n are:

(i) For $\Delta < 1$ the k_j are real.

(ii) For n even, if k_j is a root, then $-k_j$ is also a root.

(iii) It is assumed that for large even N the roots become densely distributed on some (finite or infinite) interval $(Q, -Q)$ with density $N\rho(k)dk$, where $\rho(k)$ is a smooth function.

(iv) For $\Delta = 1$, all the k_j vanish; i.e., $f(\mathbf{X}) = 1$.

Henceforth we shall adopt the notation

$$z(y) = \lim N^{-1} \ln[\max(\Lambda_R, \Lambda_L)] - Vy \tag{2.15}$$

so the free energy per vertex is given by

$$-\beta\mathfrak{F} = \max_y [z(y) + Vy]. \tag{2.16}$$

The discussions of the cases $\Delta \ge 1$ and $\Delta < 1$ are quite different, since in the former case the solution for $z(y)$ may be obtained by guile, avoiding the solution of (2.14). In general the k_j are complex but for $\Delta = 1$, from (iv) above, we have

$$z(y) = -K_2 + \max(0, -K_1) \tag{2.17}$$

independent of y. Now Eq. (2.17) implies: (i) $z(y)$ is convex and nonincreasing in K_1 and K_2, and thus is also convex and nonincreasing in the limit $N \to \infty$; (ii) by subtraction of ϵ_1 from the e_j, $z(y) \to z'(y) = z(y) + K_1$, where $z'(y)$ is convex in K_1 and K_2, nonincreasing in K_2 and nondecreasing in K_1. In the thermodynamic limit, the right-hand side of (2.17) is a lower bound on $z(y)$ for all Δ, because it can be achieved for $K_1 > 0$ by a single configuration which is a spiral band of type 1 vertices and a spiral band of type 2 vertices in the relative proportion $(1 + y) : (1 - y)$. These bands are separated by a ladder of alternating 5 and 6 vertices, but these make no contribution in the thermodynamic limit. The above argument is the same for $K_1 < 0$, with vertices 1 and 2 replaced by vertices 3 and 4. Using these results and the previous paragraph, we have the result:

Equation (2.17) is true for all y and all $\Delta \ge 1$.

It is clear that the system is completely frozen into homogeneous regions of vertex 1 and 4 when $\Delta \ge 1$,

the interfaces being composed of vertices 5 and 6. This complete ordering should be contrasted with the results for the Ising model.[6] Nagle[7] has given a rigorous proof of the freezing when $K_2 = 0$, using an argument due to Takahashi[8]; he also obtained a latent heat at $\Delta = 1$. The argument may be extended to cases with $K_2 \neq 0$, for which $\Delta = 1$ can be realized.

For $\Delta < 1$, (2.13) reduces, in the limit $N \to \infty$, to an integral equation for the density function $\rho(k)$, which we write in the form

$$1 = 2\pi\rho(p) - \int_{-Q}^{Q} \frac{\partial\theta(p,q)}{\partial p}\,\rho(q)dq, \tag{2.18}$$

where $Q(y)$ must be determined self-consistently by the condition

$$\tfrac{1}{2}(1-y) = \int_{-Q}^{Q} \rho(p)dp. \tag{2.19}$$

Once these equations have been solved for $\rho(p)$ and Q, the free energy may be determined. The calculation is simplified somewhat by noting that for $\Delta < 1$ either Λ_R or Λ_L vanishes in the thermodynamic limit, for

$$|1 - e^{ik+K_1}|^2 \left(\left| \frac{2\Delta - e^{K_1} - e^{ik}}{1 - \exp(ik + K_1)} \right|^2 - \left| e^{-K_1} \frac{2\Delta - e^{-K_1} - e^{-ik}}{1 - \exp(-ik - K_1)} \right|^2 \right)$$
$$= 4(\cosh K_1 - 2\Delta + \cos k)\sinh K_1, \tag{2.20a}$$

whence it follows readily that

$$\begin{aligned} &|\Lambda_R| > |\Lambda_L| \quad \text{for } K_1 > 0, \\ &|\Lambda_R| < |\Lambda_L| \quad \text{for } K_1 < 0, \\ &|\Lambda_R| = |\Lambda_L| \quad \text{for } K_1 = 0. \end{aligned} \tag{2.20b}$$

Consequently, in the thermodynamic limit $z(y)$ is given by

$$z(y) = -K_2 + \max(0, -K_1) + \tfrac{1}{2}\int_{-Q}^{Q} \rho(q) \times \log\left(\frac{(2\Delta - \eta)^2 + 1 - (2\Delta - \eta)\cos q}{1 + \eta^2 - 2\eta\cos q} \right) dq, \tag{2.21}$$

where $\eta = e^{|K_1|}$. This is equivalent to the result obtained by Sutherland.[9]

Equation (2.18) occurs in the theory of the Heisenberg–Ising chain[10]; this is not surprising since T has been shown to commute with the Hamiltonian[11]

$$\mathcal{H}(\Delta) = -\sum_{j=1}^{n} (\sigma_j^+\sigma_{j+1}^- + \sigma_j^-\sigma_{j+1}^+ + \tfrac{1}{2}\Delta\sigma_j^z\sigma_{j+1}^z) \tag{2.22}$$

provided Δ is given, as before, by Eq. (2.12). It was first shown by Hulthén[12] and Walker[13] that there exist transformations which reduce the kernel of the integral equation (2.18) to difference form. This work was unified and completed by Yang and Yang,[10] whose notation we follow. The transformations $k \to \alpha(k)$ needed differ for the three regions $\Delta < -1$, $-1 < \Delta < 1$, and $\Delta \geq 1$.

For $|\Delta| < 1$ the wavenumbers k_j are restricted to lie in the interval $|k_j| < (\pi - \mu)$, where

$$\cos\mu = -\Delta, \quad 0 < \mu < \pi. \tag{2.23}$$

The transformation which reduces the equation to difference type is

$$e^{ik} = (e^{i\mu} - e^{\alpha})/(e^{\alpha}e^{i\mu} - 1). \tag{2.24}$$

Even under this transformation we still have two coupled complicated integral equations to solve. In the case $y = 0$, which is applicable in the absence of a vertical external field, the quantity Q becomes infinite and the equation can be solved by Fourier analysis. Therefore, from now on we shall restrict ourselves to the cases $H = 0, V = 0$, where the free energy is given by $z(0)$. The important features of the behaviour in a vertical field can be determined by the means used for the F and KDP models. For $y = 0$ the density is

$$\rho(k)\,(dk/d\alpha) = (\tfrac{1}{4}\mu)\,\operatorname{sech}(\pi\alpha/2\mu), \quad |\Delta| < 1. \tag{2.25}$$

For $\Delta < -1$ the appropriate transformation is

$$e^{ik} = (e^{\lambda} - e^{-i\alpha})/(e^{\lambda - i\alpha} - 1), \tag{2.26}$$

where

$$\cosh\lambda = -\Delta, \quad 0 < \lambda < \infty. \tag{2.27}$$

In the case $y = 0$ we have $|\alpha| < \pi$ and the equation can be solved by Fourier series, leading to

$$\rho(k)\frac{dk}{d\alpha} = \frac{\pi}{4\lambda}\sum_{-\infty}^{\infty} \operatorname{sech}\left(\frac{\pi(\alpha + 2\pi n)}{2\lambda}\right). \tag{2.28}$$

In this case the density is closely related to the Jacobian elliptic function[14] cn, and, as we shall point out again later, there seems to be an ill-defined but pervasive connection between the ferroelectric problems and the theory of elliptic functions which may well have been finally established by Baxter.[15]

The free energy is given by (2.16), where $z(y)$ is given by (2.22) for $\Delta < 1$. Yang and Yang[10] have proven that, for $1 \geq y > 0$, $\rho(q)$ considered as a function of Δ and y is real analytic in Δ and y, which gives the following result:

> $z(y)$ is real analytic in the temperature and in y provided
>
> $0 < y \leq 1$ and $\Delta < 1$.

The case $y = 0$, which is pertinent if $V = 0$, is exceptional in this respect. For $\Delta < 1$ the kernel of the integral is not analytic in Δ at $\Delta = -1$. In the remainder of this paper we shall discuss this case. Our first conclusion is that singular behavior is only to be expected at $\Delta = \pm 1$. It is important, therefore, to decide what values of Δ can be realized on the real positive temperature axis for a given assignment of vertex energies. This is summarized in the following easily proven theorem:

Theorem: If the temperature T is real and positive, then $\eta \geq 1$ and:

(1) For no values of ϵ_1 and ϵ_2 can both $\Delta = +1$ and $\Delta = -1$ be realized as T is varied,

(2) If $\epsilon_2 > \max(0, \epsilon_1)$, then $-\infty < \Delta < \frac{1}{2}$ for real T, whereas if $\epsilon_2 < \max(0, \epsilon_1)$, $\frac{1}{2} < \Delta < \infty$. If $\epsilon_2 =$

$\max(0, \epsilon_1)$, then $0 \leq \Delta < \frac{1}{2}$; there is no singular behavior in this case.

3. PROPERTIES OF THE FREE ENERGY

Since both $\rho(k)$ and the argument of the logarithm in Eq. (2.21) are analytic with respect to Δ and η in the regions $|\Delta| < 1$ and $\Delta < -1$, critical behavior in the free energy can be expected only for $\Delta = \pm 1$. From an examination of the $\Delta = \pm 1$ contours in the (K_1, K_2) plane shown in Fig. 2, we can divide the plane into two broad regions, which we label I and II. Region I, in which, as will be seen, the model is antiferroelectric in character, contains the F model, and region II, in which the model is ferroelectric, we have the KDP model, both of which have been studied in detail by Lieb.[3,4] In Addition, there are two singular cases: the IKDP model discussed by Glasser[16] and the IF model, which we shall consider below. In addition to these models, which correspond to rays in the $K_1 - K_2$ plane, there are the ice model ($K_1 = K_2 = 0$) and a limiting case studied by Wu,[17] where $\Delta = 0$, which we shall not consider further.

The behavior of Δ as a function of inverse temperature is shown in Fig. 3 (for the IF model Δ attains 1 at $\beta = \infty$). We define a transition temperature T_0 by the condition $\Delta = \pm 1$; the behavior of T_0 as a function of the angle θ shown in Fig. 2 is given schematically in Fig. 4.

It is convenient to introduce the parameters

$$e^{i\varphi_0} = (1 + \eta e^{i\mu})/(e^{i\mu} + \eta), \quad 0 \leq \varphi_0 \leq \mu, \ |\Delta| < 1, \tag{3.1}$$

$$e^{\theta_0} = (1 + \eta e^{\lambda})/(e^{\lambda} + \eta), \quad 0 \leq \theta_0 \leq \lambda, \ \Delta < -1. \tag{3.2}$$

Then in the "high temperature" regime the free energy of our model may be expressed, by inserting (2.25) into (2.21), as

$$\mathfrak{F} = \epsilon_2 + \max(0, \epsilon_1) - \tfrac{1}{4} kT \int_{-\infty}^{\infty} \frac{d\alpha}{\cosh \pi\alpha} \times \ln\left(\frac{\cosh(2\mu\alpha) - \cos(2\mu - \varphi_0)}{\cosh(2\mu\alpha) - \cos\varphi_0}\right), \quad T > T_0, \tag{3.3}$$

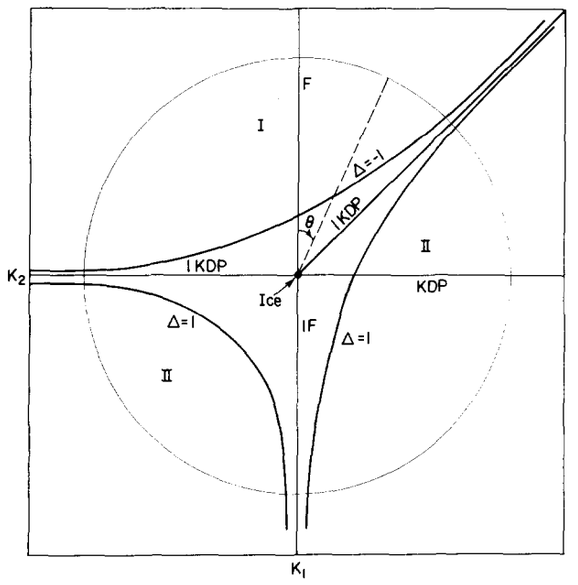


FIG. 2. Classification of models in the scaled (ϵ_1, ϵ_2) plane $(K_i = \beta\epsilon_i)$. I and II are the regions $\pi/2 > \theta > -\pi/4$ and $-3\pi/2 < \theta < -\pi/4$, respectively.

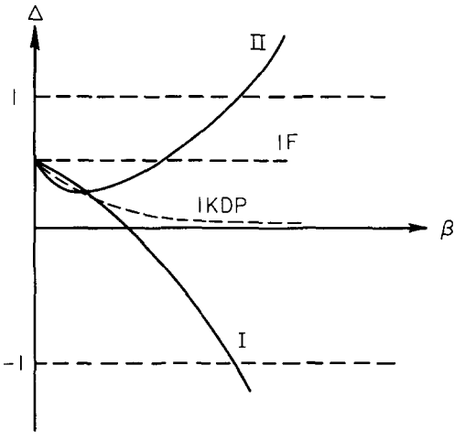


FIG. 3. Behavior of the parameter Δ as a function of inverse temperature for the models of Fig. 2.

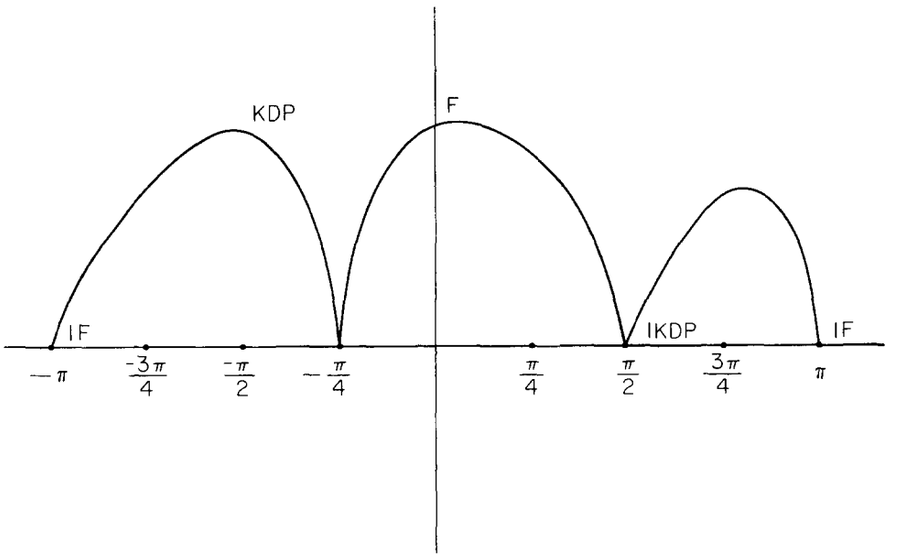


FIG. 4. Behavior of the transition temperature as a function of the angle θ of Fig. 2.

which is applicable to both regions I and II, which we discuss separately.

For region II we have that $\Delta = 1$ corresponds to $\mu = \pi$. At $\mu = \pi$ the argument of the logarithm in (3.3) becomes unity, so that the integral vanishes and the high temperature free energy joins continuously to the "low temperature" value given by (2.17). Therefore, we evidently have a first-order phase transition to a ferroelectrically ordered low-temperature phase. Since $\mathfrak{F}$ depends analytically on η about η_0, any singular behavior must be due to the dependence of the high temperature free energy on μ, since

$$\frac{d\mu}{d\Delta} = \csc\mu, \tag{3.4}$$

which is infinite at $\mu = \pi$. Indeed, a simple calculation gives

$$\mathfrak{F} = \frac{(\pi-\mu)^2}{4(\eta-1)} + \frac{(\pi-\mu)^3}{3(\eta-1)\pi} + O[(\pi-\mu)^4], \qquad T > T_0. \tag{3.5}$$

The singular part of the specific heat at vanishing vertical polarization is

$$C_0 \cong \frac{k}{4(\eta-1)\pi}\left(\frac{2\epsilon_1 \sinh K_1 - (2K_2 - K_1)e^{(2K_2-K_1)}}{kT_c}\right)^{3/2} \times\left(\frac{T}{T_c}-1\right)^{-1/2}. \tag{3.6}$$

The first-order phase transition in this case is unusual in that the low and high temperature-free energy functions cross where the latter has a vertical tangent as a function of temperature. For the special case of the KDP model, (3.6) reduces to give the behavior obtained previously.[3] This case will be treated in more detail in Sec. 4. This concludes the study of the free energy in the ferroelectric region II.

In the case of the IF model, for which $\eta = 1$, $e^{2K_2} < 1$, which represents a singular case in region II, we have $T_0 = 0$ and find for the expansion of the free energy about $\mu = \pi$

$$\mathfrak{F} \cong -\tfrac{1}{2} kT(\pi-\mu)[1 + \tfrac{1}{2}(\pi-\mu) + \cdots]. \tag{3.7}$$

From this we find for the internal energy

$$U \cong \tfrac{1}{2}|\epsilon_2| e^{2K_2}(\csc\mu)[1 + (\pi-\mu) + \cdots]. \tag{3.8}$$

Since, in this case

$$e^{2K_2} = 1 - \Delta, \tag{3.9}$$

we have

$$U \cong \tfrac{1}{2}|\epsilon_2|(\pi-\mu) + O[(\pi-\mu)^2], \tag{3.10}$$

and

$$C_0 \cong 2^{-1/2}|\epsilon_2|^2/(kT)^2. \tag{3.11}$$

In this case the evident vanishing of the latent heat means that there is a second-order phase transition at $T = 0$.

In region I, which includes the antiferroelectric F model, (3.3) again describes the high-temperature region $T > T_0$, but $T = T_0$ now corresponds to $\mu = 0$. This gives rise to an interesting phenomenon. Consider the function $H(\mu)$ defined by the integral

$$H(\mu) = \int_{-\infty}^{\infty} d\alpha \ \mathrm{sech}(\pi\alpha) f(\alpha, \mu). \tag{3.12}$$

The singularities of the integrand are simple poles due to $\mathrm{sech}(\pi\alpha)$ evenly spaced along the imaginary axis in the α plane together with the singularities of $f(\alpha, \mu)$ for positive real μ, which we represent as the blobs in Fig. 5. The location of these singularities depends on μ. (We assume f is real analytic, and so by the Schwartz reflection principle the singularity set of f is invariant under complex conjugation.) Let us now suppose that the singularities of f are rational in μ. As we allow μ to circle the origin in the μ plane, the blobs, which may also change shape, will circle the origin in the α plane. To avoid integrating through a singularity, we must deform the contour of integration more and more, until it gets pushed up against the imaginary axis. If the singularity set of f is

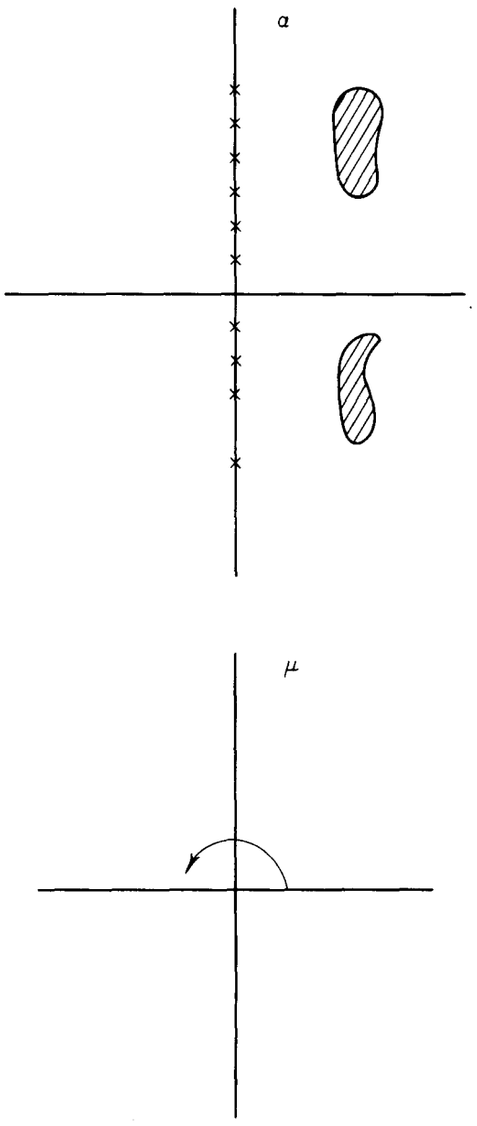


FIG. 5. Singularity structure for the hypothetical function $f(\alpha, \mu)$.

wider than the spacing of the poles along the imaginary axis, the contour will certainly become pinched, thereby forcing us to integrate through a singularity and causing there to be a singularity in $H(\mu)$ in the μ plane. In the case of the integral in (3.3) for small positive μ, the singularities of the logarithm form several lines of branch points. As we circle the origin in the μ plane these lines move in a complicated way determined in part by the dependence of φ_0 on μ. By the implicit function theorem, however, this dependence can be shown to be analytic in a neighborhood of $\mu = 0$, where $\varphi_0 = 0$. Thus for small μ we can ignore the presence of φ_0. Therefore, these lines of singularities simply rotate synchronously with μ, and, when μ approaches any small negative rational multiple of π, a pinch will occur. This shows that, at least in a small neighborhood of $\mu = 0$, the negative real axis forms a natural boundary (line segment having a dense set of singularities) with end point at $\mu = 0$. Since for larger μ the dependence of φ_0 on μ becomes more complex, it may happen that the natural boundary may disappear or take on some complicated form away from the origin in the μ plane. In Appendix A, however, we show that this does not happen and that the entire negative real axis is a natural boundary for the high temperature-free energy in region I. It may also happen, of course, that the function f in the integrand of $H(\mu)$ has zeros which coincide with the poles of $\operatorname{sech}(\pi\alpha)$ in which case the contour will not be pinched. In Appendix B we show that this actually happens in the case of the KDP and IKDP models, but for no others.

We may formally expand (3.3) in even powers of μ, where we take into account the fact that φ_0 is odd and analytic in μ; it is shown in Appendix D that the coefficients are all finite. However, as we have just demonstrated, $\mu = 0$ is a nonisolated singularity so that this series must have vanishing radius of convergence. We next note that Δ is analytic with respect to T about $T = T_0$, and, as can be proven from the implicit function theorem, $\mu^2 = [\arccos(-\Delta)]^2$ is an analytic function of Δ about $\Delta = -1$. { This surprising fact also follows from the identity $\arccos z = [2(1-z^2)^{1/2}/(1+z)]\,{}_2F_1[\frac{1}{2}, 1; \frac{3}{2}; (z-1)/(z+1)]$.} The free energy in region I is therefore infinitely differentiable, but nonanalytic as $T \to T_0^+$.

To characterize the thermodynamic behavior of the system completely, we must examine the behavior of the free energy as $T \to T_0^-$. From (2.28) and (2.21) we have in the low temperature region, where $\Delta < -1$

$$\mathfrak{F} = \epsilon_2 + \max(0, \epsilon_1) - \tfrac{1}{2} kT(\lambda - \theta_0) - kT \sum_{n=1}^{\infty} \frac{e^{-\lambda n} \sinh n(\lambda - \theta_0)}{n \cosh n\lambda}. \quad (3.13)$$

We note that this expression is nonsingular below $T_0(\lambda = 0)$ down to $T = 0(\lambda = \infty)$. Next we see that (3.8) and (3.3) agree at $T = T_0$ (where $\lambda = \theta_0 = 0$, $\mu = \varphi_0 = 0$). Finally, as is shown in Appendix A, (3.8) is the analytic continuation of (3.3) onto the imaginary μ axis. The entire free energy is thus described in terms of a single function, infinitely differentiable, but nonanalytic at $T = T_0$ and nonsingular at every other real temperature. We therefore conclude that in region I the system undergoes an infinite order phase transition. Since at low temperatures vertices 5 and 6, of which there must be an equal number, are favored, this transition is to an antiferroelectrically ordered state and occurs without any thermodynamic anomalies. To remove any doubt on these matters, we present in Appendix A a detailed calculation of the smooth matching of the high and low temperature-free energies at T_0.

We come finally to the IKDP model which separates region I from region II. In this case we have either $K_1 = K, K_2 = 0$ or $K_1 = K_2 = K \geq 0$ and $0 \leq \Delta \leq \frac{1}{2}$, which corresponds to $\frac{1}{2}\pi < \mu \leq 2\pi/3$, and $\varphi_0 = 2\pi - 3\mu$. Thus the system is essentially always in the "high temperature" region and the free energy is

$$\mathfrak{F} = \epsilon - \tfrac{1}{4} kT \int_{-\infty}^{\infty} d\alpha \operatorname{sech}(\pi\alpha) \ln\left(\frac{\cosh(2\mu\alpha) - \cos(5\mu)}{\cosh(2\mu\alpha) - \cos(3\mu)}\right) \quad (3.14)$$

As shown in Appendix B, the integral in this case can be evaluated in terms of elementary functions and we have

$$\mathfrak{F} = 2\max(0, \epsilon) + kT \ln[(\pi/2\mu) \sin\mu \tan(\pi^2/2\mu)]. \quad (3.15)$$

It will be seen below that (3.15) is the analytic continuation in ϵ of the free energy for KDP. These two models merge continuously at $T = \infty$. It is seen that (3.15) has no singularity for any real T. Because of the behavior of the $\Delta = \pm 1$ contours in Fig. 2, however, we define T_0 to be zero in this case.

This concludes our discussion of the thermodynamic properties of the general model in the absence of external fields, except to point out that as $T \to \infty$ in any special case we obtain the ice model for which

$$z(0) = \tfrac{3}{2} \ln\tfrac{4}{3}. \quad (3.16)$$

4. DETAILED ANALYTIC STRUCTURES FOR THE KDP AND F MODELS

The analytic properties of the Onsager solution for the free energy of the two dimensional Ising model have been studied in detail by Majumdar.[18] For comparison with the results for the ferroelectric models we have extended his results to obtain the behavior of the free energy in the complex temperature plane. This is shown in Fig. 6. The characteristic structure of this and succeeding diagrams arises in the following way. The partition function or free energy is most conveniently expressed in terms of an activity of the form $z = \exp(\beta J)$ and in the complex z plane will have a relatively simple analytic structure. When the logarithm is taken to get the analytic structure in the β plane, each singularity z_0 in the z plane becomes a sequence of singularities $\beta_0 + 2k\pi i$, $k = 0, \pm 1, \cdots$, which lie along a line parallel to the imaginary β axis. Finally, in taking the reciprocal to obtain the behavior in the temperature plane, lines which extend to infinity go into circles passing through $T = 0$, and the analytic structure consequently has a circular geometry. We note in Fig. 6 in the complex temperature Riemann structure for the Ising model a branch cut begins at $T = T_0$ and extends down to $T = 0$. As Majumdar showed this is due to the fact that the analytic continuation of the free energy from the high- to the low-temperature region does not remain real valued, but one must average the results taken above and below the branch cut in order to ob-

tain the physical result. This is similar to the situation for the F-like models in region I and is due to the fact that the free energy for real temperature depends on $|T - T_0|$, which is not an analytic relationship.

KDP Model

The free energy for the KDP model is given (up to an additive constant) by Eq. (B. 12). If we let $z = \mu/\pi$, we have $\Delta = -\cos\pi z$ and $\Delta = \frac{1}{2}e^{\beta\epsilon}$. Note that, for $T = \infty$, $z = \frac{2}{3}$ and (B. 12) gives $\frac{3}{2}\ln(\frac{4}{3})$, which is the residual entropy of square ice.[2]

The analytic structure below T_0 is trivial, and so we consider only the Riemann structure of the high-temperature free energy. The singularities of $\mathfrak{F}$ in

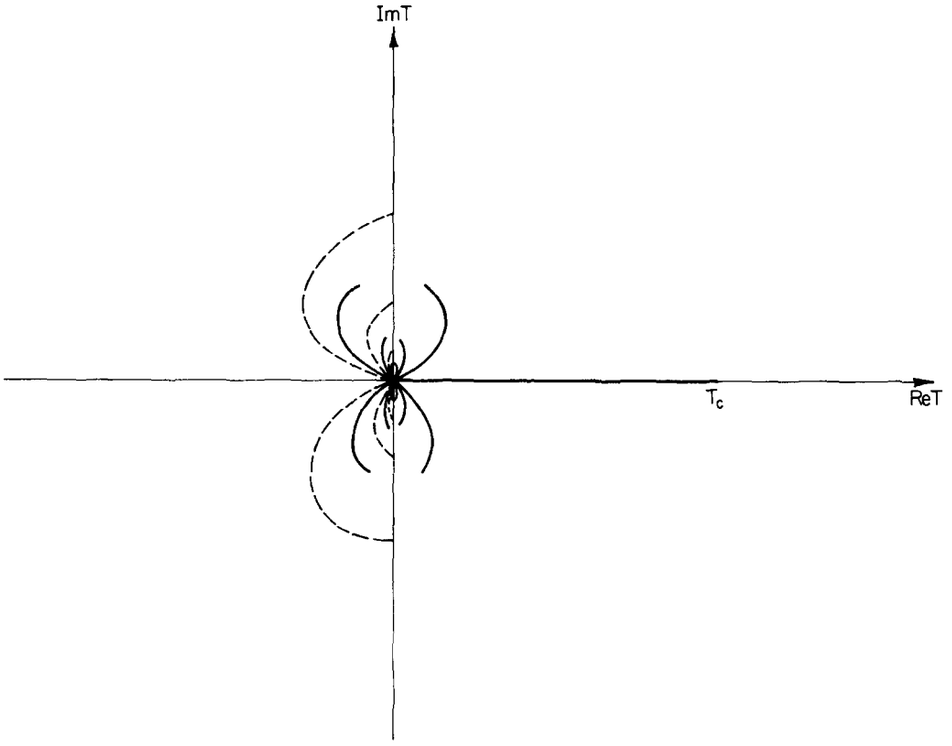


FIG. 6. Riemann structure for the two-dimensional Ising model in the complex temperature plane. The lines denote logarithmic branch cuts.

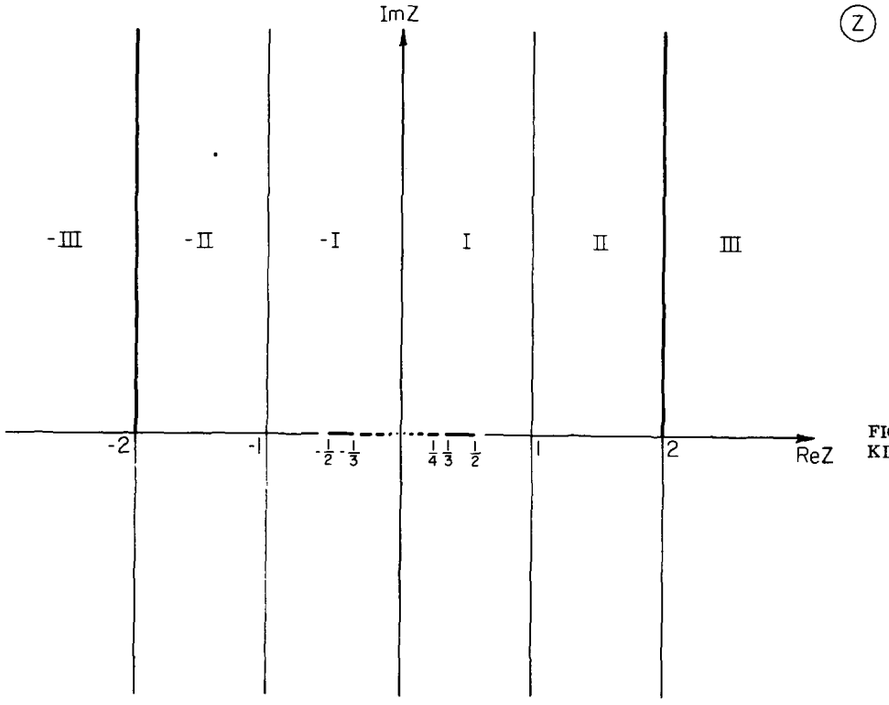


FIG. 7. Riemann structure for the KDP model in the complex z-plane.

the z plane are logarithmic branch points at $z = 1/n$, $n, n = \pm 2, \pm 3, \cdots$, and we have branch cuts $\pm(1/2|n|, 1/2(|n| + 1|))$ and cuts starting at $z = n$ which we take parallel to the imaginary z axis as shown in Fig. 7. $z = 0$ is an accumulation point of branch points and is an essential singularity. Next we note that each strip of unit width as indicated in the z plane maps onto an entire Riemann sheet in the Δ plane. We label the corresponding sheets Δ_N, $N = \pm \text{I}, \pm \text{II}, \cdots$. These sheets are connected along branch lines $(\pm 1, \pm \infty)$ as shown schematically in Fig. 8. The line $z = n$ corresponds to the two edges of the positive (negative) cut if n is odd (even). The small branch cuts on strips $\pm$ I of the z plane map into similar cuts $(-\cos(\pi/2|n|), -\cos[\pi/(2|n| + 1)])$ which occurs on Δ_I and Δ_{-I} only. The remaining Δ-structure is free of additional singularities. A small circle centered at $z = n\pi$ maps into a small circle about $\Delta = \pm 1$, and tracing the circle once in the positive sense in the z plane is equivalent to circling $\Delta = \pm 1$ twice in the positive sense.

The sheets $\Delta_{\pm 1}$ are each connected (along the small branch cuts) to a countable number of additional Δ planes which we denote $\Delta_{\pm\text{I}}^{(n)}$. On these planes the values of $\mathfrak{F}$ differ by $n\pi i$ from its values on $\Delta_{\pm\text{I}} \equiv \Delta_{\pm\text{I}}^{(0)}$. On $\Delta_{\text{I}}^{(0)}$, by circling $\Delta = 1$ clockwise, we go to $\Delta_{\text{II}} \equiv \Delta_{\text{II}}^{(0)}$ and again circling $\Delta = 1$ return us to $\Delta_{\text{I}}^{(0)}$. The same thing happens for $\Delta = -1$. However, if on Δ_{II} we go to -1 and circle this point we do not return to Δ_{I}, but spiral onto another countable set of Riemann sheets $\Delta_{\text{II}}^{(n)}$ on which the values of $\mathfrak{F}$ differ from those on $\Delta_{\text{II}}^{(0)}$ by $n\pi i$. This same situation occurs for all the remaining sheets $\Delta_N \equiv \Delta_N^{(0)}$. Thus, we have a doubly countable infinite collection of Riemann sheets $\Delta_N^{(n)}$, $N = \pm \text{I}, \pm \text{II}, \cdots$, $n = 0, \pm 1, \pm 2, \cdots$. For $N = \pm$ I, the $\Delta_N^{(n)}$ are joined along the small branch cuts, and for $N \neq \pm$ I they are joined along the long branch cuts.

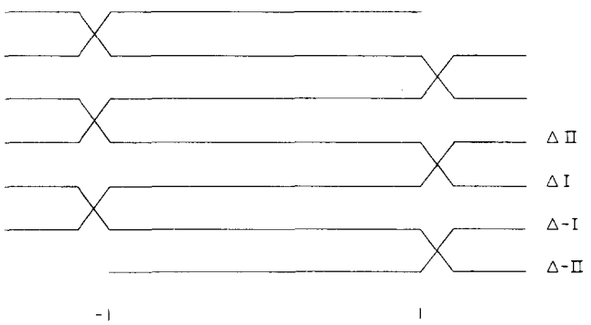


FIG. 8. Connectivity of the Δ sheets in the Riemann structures for the ice models.

Next we consider the Riemann structure for $\mathfrak{F}$ as a function of β. As indicated above, the analytic structure in the β-plane is periodic with period $2\pi i/\epsilon$, so we need only look at the principal strip $0 \leq \text{Im}\beta < 2\pi/\epsilon$. Each Riemann sheet $\Delta_N^{(n)}$ corresponds to a Riemann sheet $\beta_N^{(n)}$. Of these, $\beta_{\text{I}}^{(0)}$ and $\beta_{\text{I}}^{(0)}$ are typical and are shown in Fig. 9. $\beta_{\text{I}}^{(0)}$ is shown and $B_{\text{II}}^{(0)}$ differs only in the absence of the cuts labeled (3) and (4). The connectivity as regards the various sheets exactly mirrors that for the $\Delta_N^{(m)}$ except going from $\Delta = 1$ to $\Delta = -1$ means leaving $\beta = \ln(2/\epsilon)$ along the negative real axis and returning to $\ln(2/\epsilon) + \pi i/\epsilon$ parallel to the positive real axis. Finally, the behavior in the complex T plane is shown in Fig. 10.

F Model

For clarity we repeat here for the F model the general considerations applied to region I.

The solution for the free energy of the F model can be written (for $T > T_c$)

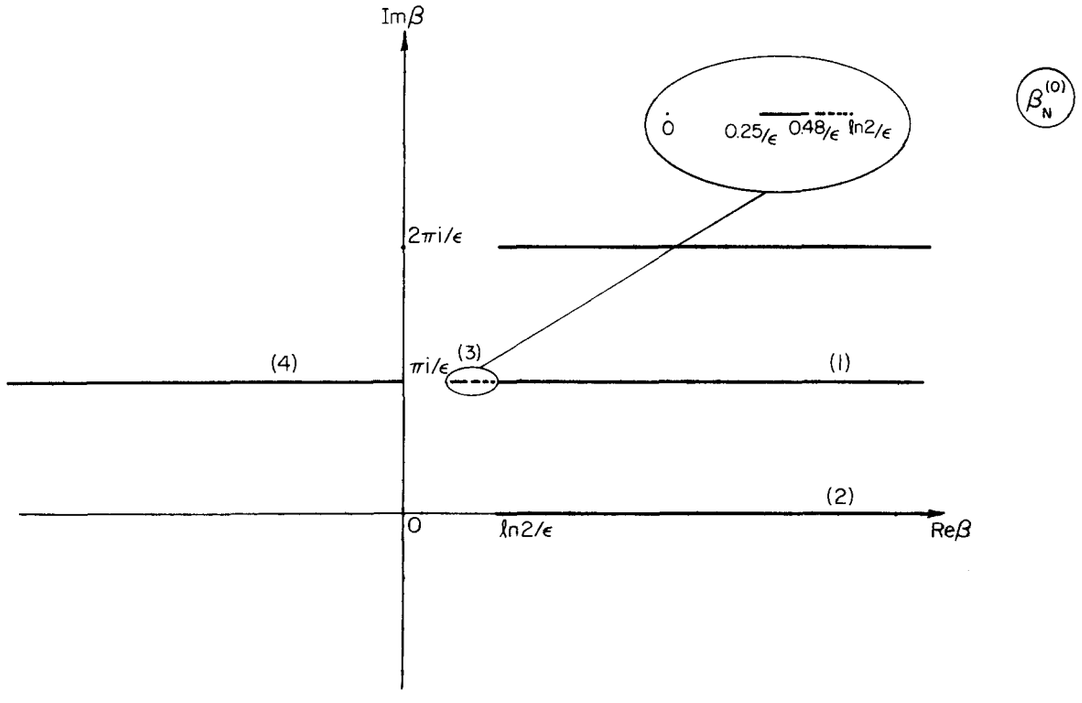


FIG. 9. Riemann structure for the KDP model in the complex β plane.

$$\Phi \equiv -\frac{F}{kT} = -\beta\epsilon + \frac{1}{8\mu}\int_{\infty}^{\infty} dx\ \mathrm{sech}\left(\frac{\pi x}{2\mu}\right) \times \ln\left(\frac{\cosh x - \cos 2\mu}{\cosh x - 1}\right),$$
$$\Delta = -\cos\mu = 1 - \tfrac{1}{2}e^{2\beta\epsilon}. \qquad (4.1)$$

Following the procedure in Ref. 16, Eq. (12), we find

$$I = \int_0^{\infty} dx\ \mathrm{sech}\left(\frac{\pi x}{2\mu}\right) \ln\left(\frac{\cosh x - \cos 2\mu}{\cosh x - 1}\right) = 2\mu \int_0^{\infty} \frac{dx}{x}\ \mathrm{sech}\mu x[\coth\pi x(1 - \cosh xy) + \sinh xy], \qquad (4.2)$$

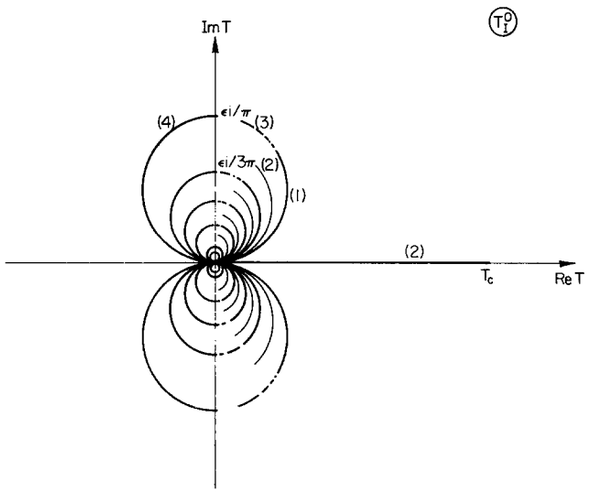


FIG. 10. Riemann structure for the KDP model in the complex temperature plane.

where $y = 2(\mu - n\pi)$ and $n\pi \le \mathrm{Re}\mu < (n+1)\pi$. Equation (4.2) can be rewritten

$$I = 4\mu \int_0^{\infty} \frac{dx}{x}\ \frac{\sinh(\mu - n\pi)x}{\cosh\mu x}\ \frac{\sinh[(n+1)\pi - \mu]x}{\sinh\pi x}, \qquad n\pi \le \mathrm{Re}\mu < (n+1)\pi, \qquad (4.3)$$

or

$$I = \begin{cases} 4\mu \int_0^{\infty} \frac{dx}{x}\ \tanh\mu x\ \frac{\sinh(\pi - \mu)x}{\sinh\pi x}, & 0 \le \mathrm{Re}\mu < \pi, \qquad (4.4a) \\ 4\mu \int_0^{\infty} \frac{dx}{x}\left(\frac{\tanh\mu x}{\tanh\pi x} - 1\right) \sinh[(2\pi - \mu)x], & \pi \le \mathrm{Re}\mu < 2\pi. \qquad (4.4b) \end{cases}$$

The first integral in (4.4) converges for all μ in the strip $0 \le \mathrm{Re}\mu < \pi$ and, therefore, defines an analytic function $\Phi_1(\mu)$. The singularities of the integrand are poles at the points $x_0(k) = ik$, $x_1(k) = (2k+1)\pi i/2\mu$, $k = 0, \pm 1, \pm 2, \cdots$ (excluding $x = 0$). The corresponding residues are

$$x_0(k): \frac{1}{k\pi i}\ \frac{\sin^2 k\mu}{\cos k\mu},$$
$$x_1(k): \frac{2(-1)^{k+1}}{\pi i(2k+1)} \cot\left((2k+1)\frac{\pi^2}{2\mu}\right). \qquad (4.5)$$

As μ moves counterclockwise along a semicircle from the right half-plane to $-\mu$ in the left half-plane, the pair of poles $x_1(k) = -x_1[-(k+1)]$ also exchange positions along semicircular paths. In doing this, the poles $x_1(k)$ break through the positive real axis (in the x plane), which we can deform into the positive real axis together with small counterclockwise circles

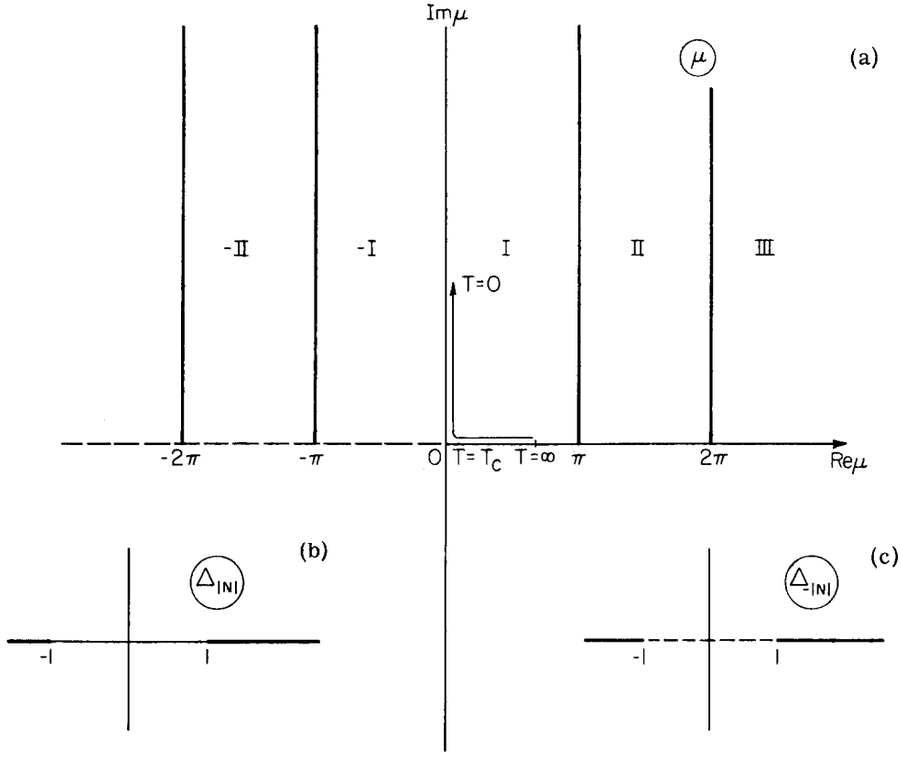


FIG. 11. Riemann Structure for the F model in the complex μ plane.

about the point $x_1(k)$. By noting that the pole structure remains completely unchanged, we have

$$\Phi_1(-\mu) = \Phi_1(\mu) + 2\pi i \sum \text{Residue at } \frac{k\pi i}{2\mu}$$
$$= \Phi_1(\mu) + 4 \sum_{k=1,\text{odd}}^{\infty} \frac{(-1)^{(k+1)/2}}{k} \cot\left(\frac{k\pi^2}{2\mu}\right)$$
$$= \Phi_1(\mu) + 4\Phi_2(\mu). \tag{4.6}$$

It is interesting to note that the series in (4.6) is closely related to the elliptic function[14] dn. Thus, we can continue $\Phi_1(\mu)$ from the strip $0 \le \text{Re}\mu < \pi$ to the strip $-\pi < \text{Re}\mu \le 0$. Next we note that $\Phi_2(\mu)$ is a uniformly convergent series everywhere off the real μ axis. However, for μ real $\Phi_2(\mu)$ has a pole at every rational multiple of π, $m\pi/n$, where n is odd. These points are dense on the real line so that any expansion of $\Phi_2(\mu)$ in a power series about any real point will have zero radius of convergence, i.e., $\Phi_2(\mu)$ has a natural boundary along the real axis.

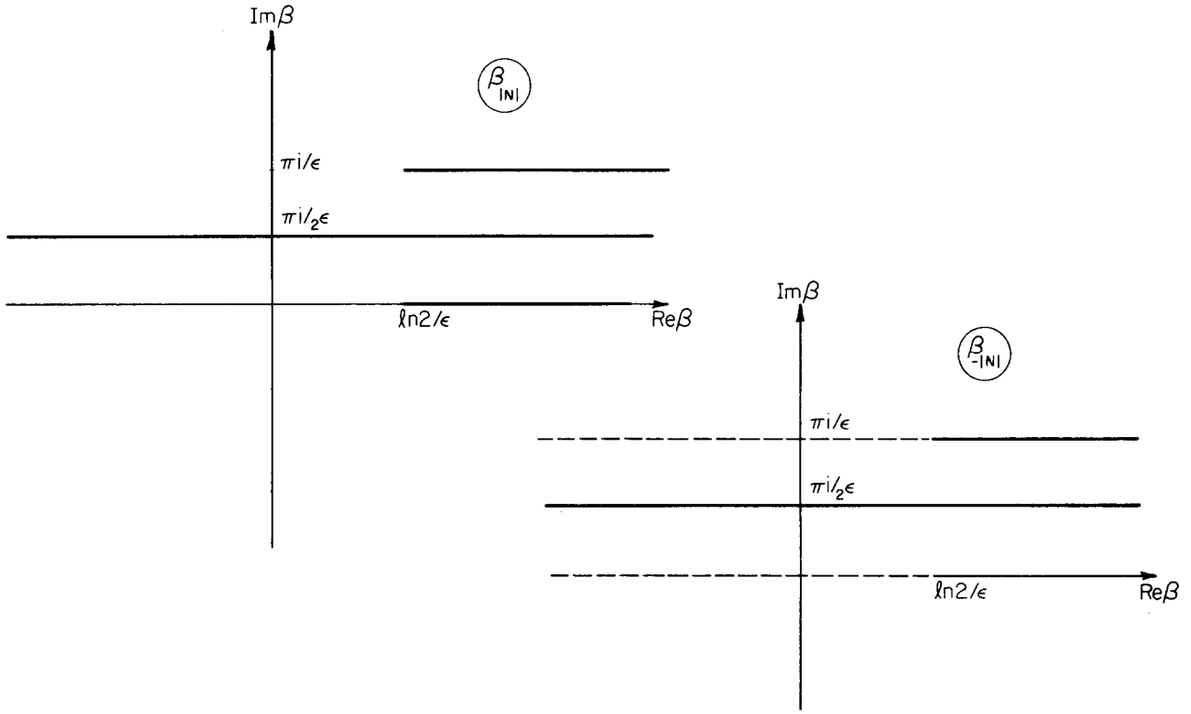


FIG. 12. Riemann Structure for the F model in the complex β-plane.

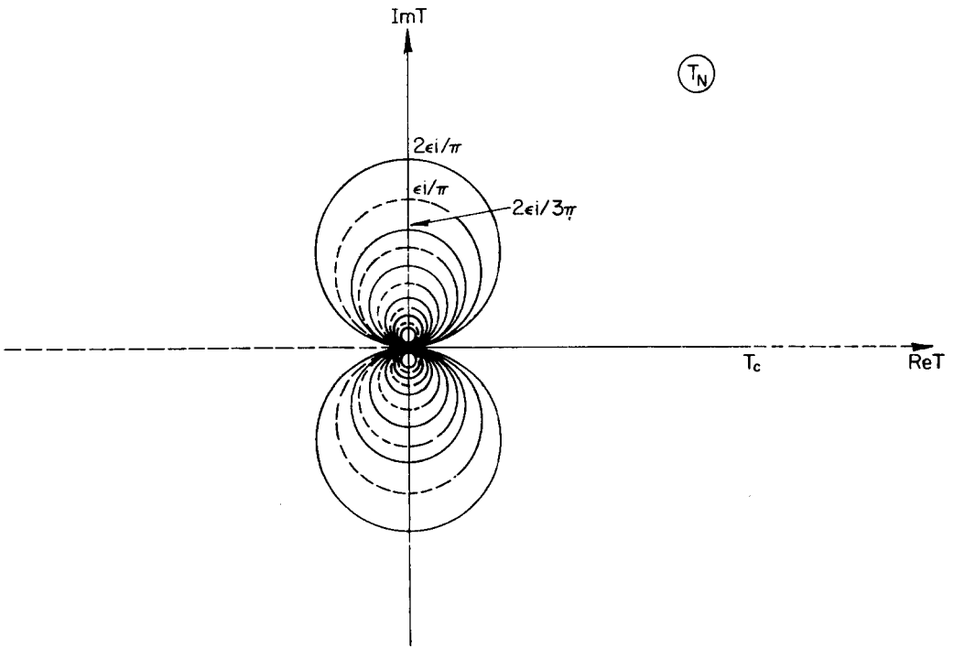


FIG. 13. Riemann structure for the F model in the complex temperature plane.

Since the same argument applied to each of the representations (4.3) leads to the same conclusion, we see that $\Phi(\mu)$ is analytic everywhere in the μ plane except for a natural boundary along the negative μ axis as shown in Fig. 11(a). Again, since $\Delta = -\cos\mu$, each strip of the μ plane of width π parallel to the imaginary axis maps into an entire Reimann sheet of the Δ-Riemann structure. These are connected as indicated in Fig. 8. Each Riemann sheet Δ_N has branch cuts $(-1,-\infty), (1,\infty)$. Thus, if one begins at $\Delta_{\rm I} = 0$ and circles $\Delta_{\rm I} = 1$, one goes onto $\Delta_{\rm II}$. By circling $\Delta = 1$ again, one returns to $\Delta_{\rm I}$ or by then circling $\Delta_{\rm II} = -1$ goes to $\Delta_{\rm III}$, etc. Other than these branch cuts, the sheets Δ_N, $N = {\rm I}, {\rm II}, \cdots$, are free of singularities. However, if we begin at $\Delta_{\rm I}$ (physical sheet above T_c) and circle -1, we go to $\Delta_{-{\rm I}}$. $\Delta_{-{\rm I}}$ has a natural boundary along the interval $[-1, 1]$, which cannot be crossed. By circling the entire interval, we go to $\Delta_{-{\rm II}}$, $\Delta_{-{\rm III}}$, etc., all of which have the same structure as $\Delta_{-{\rm I}}$. These two types of Riemann sheets are shown in Fig. 11(b), (c). Once again, the β-Riemann surface has the same singularity structure as the Δ-surface except that the behavior in the fundamental strip is repeated periodically throughout each sheet. The cuts $(\pm 1, \pm\infty)$ become $(-\infty + \pi i/2\epsilon, \infty + \pi i/2\epsilon)$ and $\ln(2/\epsilon) < \beta < \infty$. The natural boundary $(-1, 1)$ maps into $(-\infty, \ln(\epsilon/2))$. The resulting structures are shown in Fig. 12. Finally, the Riemann structure in the temperature plane is shown in Fig. 13.

Finally, we must consider the low temperature-free energy for the F model. In the original work[4] this was described by a different analytic expression than (4.1). However, for $T < T_c$, we have $\Delta < -1$, and μ becomes pure imaginary, $\mu = i\lambda$. Now if in (4.4) we set $\mu = i\lambda + \epsilon$ and take the limit as $\epsilon \to 0$, we find

$$\Phi = -\beta\epsilon + \tfrac{1}{2}\,\mathrm{Im}\,\not\int_{-\infty}^{\infty} \frac{dx}{x}\,\coth\pi x\,\tan\lambda x\,e^{i\lambda x}, \tag{4.7}$$

where the slash on the integral sign denotes that we take the Cauchy principal part in integrating past the singularities of the integrand. By closing the contour of integration in the upper half-plane, we obtain, by summing the residues at the singularities of $\coth\pi x$, ni, being careful to include only half the residue of the singularity at $x = 0$,

$$\mathrm{Re}\Phi = -\beta\epsilon + \tfrac{1}{2}\lambda + \sum_{n=1}^{\infty} \frac{1}{n}\,\tanh n\lambda\,e^{-n\lambda} \tag{4.8}$$

which is precisely the old expression. Therefore, the imaginary axis in the μ plane describes the low temperature behavior as indicated in Fig. 11(a). Once again the series in (4.8) is related to an elliptic function.[14]

ACKNOWLEDGMENT

The second named author wishes to thank Battelle Memorial Institute for its hospitality while this work was completed.

APPENDIX A

The function $\Phi(\mu)$ is defined by

$$\Phi(\mu) = \frac{1}{4}\int_{-\infty}^{\infty} \frac{dx}{\cosh(\pi x)}\,\log\left(\frac{\cosh(2\mu x) - \cos(2\mu - \phi_0)}{\cosh(2\mu x) - \cos\phi_0}\right), \tag{A1}$$

where $\phi_0(\mu)$ is defined by

$$e^{i\phi_0(\mu)} = (1 + \eta e^{i\mu})/(\eta + e^{i\mu}) \tag{A2}$$

as a function of the complex variable μ which we write as

$$\mu = \mu_1 + i\mu_2. \tag{A3}$$

In the range $-1 < \Delta < 1$, which corresponds to $0 < \mu_1 < \pi$, Eq. (2.21) gives

$$z(0) = -K_2 + \max(0, -K_1) + \Phi(\mu). \tag{A4}$$

Clearly there is a *single independent* variable in the physical problem; obvious choices for this are μ or T. Thus η depends on T, or indirectly on μ. Several remarks are relevant in this connection.

(1) For the F and IF cases, the problem is much more simple since $\eta = 1$ for all T; thus $\phi_0(\mu) \equiv 0$. Theorem 1 below discusses the analytic properties in the μ plane for this case.

(2) For the KDP and IKDP cases, the function $\Phi(\mu)$ can be evaluated in terms of elementary functions.[16]

(3) The physical properties of the model are evidently all obtained by considering a small neighborhood of the real positive temperature axis.

Theorem 1: Consider the case $\phi_0 \equiv 0$:

(1) $\Phi(\mu)$ can be analytically continued to the entire μ plane except the line $\mu_2 = 0, \mu_1 \le 0$, which is a natural boundary.

(2) $\Phi(\mu^*) = (\Phi(\mu))^*$. (A5)

(3) For real $\lambda > 0$,

$$\tfrac{1}{2}[\Phi(i\lambda) + \Phi(-i\lambda)] = \tfrac{1}{2}\lambda + \sum_{1}^{\infty} n^{-1}\,e^{-n\lambda}\,\tanh(n\lambda). \tag{A6}$$

This should be compared with Eq. (3.13).

(4) For all μ except points on the real axis,

$$\Phi(\mu) - \Phi(-\mu) = \sum_{0}^{\infty} \frac{(-1)^n}{2n+1}\,\cot\left(\frac{\pi^2(2n+1)}{2\mu}\right) - i\pi. \tag{A7}$$

Proof: The integrand in (A1) with $\phi_0 \equiv 0$ has simple poles at

$$x_0(n) = (n + \tfrac{1}{2})i, \quad n = 0, \pm 1, \cdots, \tag{A8}$$

and logarithmic branch points at

$$x_1^{\pm}(n) = \pm i + n\pi i/\mu, \tag{A9a}$$

$$x_2(n) = n\pi i/\mu, \quad n = 0, \pm 1, \cdots. \tag{A9b}$$

As μ moves away from the real axis, the branch points move in the x plane. Evidently $\Phi(\mu)$ may be analytically continued to the entire μ plane, except the negative real μ axis; in that case, the contour of integration is pinched by the branch points against the singularities $x_0(n)$, for *all* rational values of μ.

Since $\Phi(\mu)$ is real analytic and real on the line segment $(\mu_2 = 0, 0 < \mu_1 < \pi)$, the Schwarz principle gives (A5), the second part of Theorem 1.

In order to derive (A6) and (A7) and prove that $(\mu_2 = 0, \mu_1 \le 0)$ is a natural boundary, $\Phi(\mu)$ may be rewritten, using the Parseval formula, as

$$\Phi(\mu) = \tfrac{1}{2}\int_{-\infty}^{\infty} \frac{dx}{x}\,\frac{\sinh\mu x,\sinh(\pi-\mu)x}{\cosh\mu x,\sinh\pi x} \tag{A10}$$

provided $0 < \mu_1 < \pi$. (This result may be extended to other strips, but this will not be needed here.)

In this case, the integrand has simple poles at

$$p_1(n) = (n+\tfrac{1}{2})\,\pi i/\mu, \quad n = 0, \pm 1, \cdots, \tag{A11}$$

$$p_0(n) = ni, \quad n = \pm 1, \pm 2, \cdots. \tag{A12}$$

As μ rotates, the poles $p_1(n)$ rotate synchronously and impinge upon the real axis when $\mu = i\lambda$. An elementary, but tedious, calculation shows that

$$\Phi(i\lambda) = \frac{\lambda}{2} + \sum_1^{\infty} n^{-1}\, e^{-n\lambda}\tanh(n\lambda) + i\sum_0^{\infty}\frac{(-1)^n\, e^{-\pi^2(n+1/2)/\lambda}}{(n+\tfrac{1}{2})\sinh[(\pi^2/\lambda)(n+\tfrac{1}{2})]}. \tag{A13}$$

Using (2) of Theorem 1, one readily obtains (A6) from (A13) and the result

$$\Phi(i\lambda) - \Phi(-i\lambda) = -i\pi + 4i\sum_0^{\infty}\frac{(-1)^n}{(2n+1)} \times \coth\left(\frac{\pi^2(n+\tfrac{1}{2})}{\lambda}\right). \tag{A14}$$

According to the identity theorem for analytic functions, the analytic continuation of $\Phi(\mu)$, whenever it exists, must satisfy the equation

$$\Phi(\mu) - \Phi(-\mu) = -i\pi + 2\sum_0^{\infty}\frac{(-1)^n}{2n+1}\cot\left(\frac{\pi^2(2n+1)}{2\mu}\right). \tag{A15}$$

The series above represents an analytic function in the finite μ plane except for the line $\mu_2 = 0$, upon which there are simple poles at the points μ_0, given by

$$\mu_0 = \pi(2r+1)/2s, \tag{A16}$$

where r and s are any integers. In order to prove that the real axis is a natural boundary, we have to consider the approach to $\mu_2 = 0$ along a line

$$\mu = \pi(2r+1)/2s + i\mu_2. \tag{A17}$$

This gives

$$\Phi(\mu) - \Phi(-\mu) = \frac{4\mu(-1)^r}{\mu_2\pi(2r+1)^2}\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)^2} + G(\mu_2,\mu_0), \tag{A18}$$

where $G(\mu_2,\mu_0)$ can be bounded. The sum of the series is the Catalan constant, which is nonzero. The intercepts of the lines given by (A17) with the real axis form a dense set of points in $\mathcal{R}$. Thus the line $\{\mu_2 = 0, -\infty < \mu_1 \le 0\}$ is a natural boundary. This completes the proof of Theorem 1. The physical interpretation will be described later.

When $n > 1$, the motion of the singularities in the integration plane for $\Phi(\mu)$ is complicated by the explicit dependence of φ_0 on μ. Nevertheless, we shall be able to prove an analog of Theorem 1, although we shall be content with a more restricted domain of analyticity for $\Phi(\mu)$.

There are μ-dependent branch points in the integration plane at

$$x_2^{\pm}(n) = \pm i(1 - \phi_0/2\mu) + n\pi i/\mu \tag{A19a}$$

and

$$x_2^{\pm}(n) = \pm i\phi_0/2\mu + n\pi i/\mu, \quad n = 0, \pm 1, \cdots, \tag{A19b}$$

as μ varies the contour of integration pinches the stationary poles at

$$x_0(n) = (n+\tfrac{1}{2})i, \quad n = 0, \pm 1, \cdots. \tag{A20}$$

It is clear that the motion of the branch points has an epicyclic character which arises from the dependence of ϕ_0 on μ. It should also be noted that singular behavior of $\Phi(\mu)$ is given by the branch points of $\phi_0(\mu)$ at

$$\mu = i\,|K_1| + (2n+1)\pi, \quad n = 0, \pm 1, \cdots. \tag{A21}$$

For $K_1 > 0$, $\varphi_0(\mu)$ is analytic in a neighborhood of the real positive μ axis. Also $\varphi_0/2\mu$ is real for real μ, so that in that case the branch points $x_j^{\pm}(n)$ lie on the imaginary axis. Clearly there exists an $\epsilon > 0$ (for $K_1 > 0$) such that $\Phi(\mu)$ is analytic in the domain

$$D_r = \{\mu:\ \mu_1 > 0, -\epsilon < \mu_2 < \epsilon\}. \tag{A22}$$

On the principal branch of $\varphi_0(\mu)$, we have by Taylor's theorem at $\mu = 0$ that

$$\frac{\varphi_0}{2\mu} = \frac{1}{2}\frac{\eta-1}{\eta+1}\left(1 + \frac{\eta}{3(\eta+1)^2}\,\mu^2\right) + O(\mu^4). \tag{A23}$$

Thus the eccentric motion of the epicycle may be made arbitrarily small by confining μ to a small enough disc about $\mu = 0$. Evidently there exists a neighborhood D_0 of $\mu = 0$ for which the motion of $x_j^{\pm}(n)$ is essentially synchronous with μ. In this case, the analysis of Theorem 1 can be taken over directly to prove that $\Phi(\mu)$ is analytic for all μ in D_0, except the set of points $\{\mu:\ \mu_2 = 0, \mu_1 \le 0\} \cap D_0$, which is a natural boundary.

The function $\Phi(\mu)$ will now be evaluated for $\mu = i\lambda$, $\lambda > 0$. Using the Parseval transform, one obtains

$$\Phi(\mu) = \tfrac{1}{2}\int_{-\infty}^{\infty}\frac{dx}{x}\,\frac{\sinh(\pi-\mu)x\,\sinh(\mu-\varphi_0)x}{\cosh\mu x\,\sinh\pi x} \tag{A24}$$

for $0 \le \mu$, $0 < \pi$, $\mu_2 \ne 0$. By employing the same argument as for Theorem 1 (only the residues of the poles are changed), one may find

$$\Phi(i\lambda) = \frac{\lambda-\theta_0}{2} + \sum_1^{\infty}\frac{e^{-n\lambda}}{n}\,\frac{\sinh(\lambda-\theta_0)n}{\cosh(n\lambda)} + 2i\sum_0^{\infty}\frac{(-1)^n}{2n+1}\,\frac{\cos[(\pi\theta_0/\lambda)(n+\tfrac{1}{2})]}{\sin[(\pi^2/\lambda)(n+\tfrac{1}{2})]}\,e^{-(\pi^2/\lambda)(n+1/2)}, \tag{A25}$$

where θ_0 is given by (A2) with $\mu = i\lambda$. By using the analyticity of $\varphi_0(\mu)$ in a neighborhood of $\mu = i\lambda$, it is clear that there exists a $\delta > 0$ and a $\gamma(\delta) > 0$ such that $\Phi(\mu)$ is analytic in the domain D_i defined by

$$D_i = \{\mu: -\gamma(\delta) < \mu_1 < \gamma(\delta),\ |\mu_2| > \delta\}. \quad \text{(A26)}$$

The number δ in the above definition may be chosen so that $D_i \cap D_0 \neq \phi$. Thus $\Phi(\mu)$ is analytic everywhere in the domain $T = D_i \cup D_0 \cup D_r$ except for the set $\{\mu: \mu_2 = 0, \mu_1 \leq 0\} \cap D_0$, which is a natural boundary. Since $\Phi(\mu)$ is real analytic for $\mu_2 = 0$, $\mu_1 > 0$, it follows that

$$\Phi(\mu) = (\Phi(\mu^*))^* \quad \text{(A27)}$$

for $\mu \in T$, by the Schwartz reflection principle. Thus we have established the following theorem:

Theorem 2: If the variable η satisfies $\eta > 1$, then for all such values we have the following.

(1) $\Phi(\mu)$ may be analytically continued to a sector S of the μ plane given by

$$-(\tfrac{1}{2}\pi + \delta) \leq \arg\mu \leq (\tfrac{1}{2}\pi + \delta), \qquad \delta > 0.$$

There are singularities at the branch points of $\phi_0(\mu)$ given by

$$\mu_n = \pm i\,|K_1| + (2n+1)\pi, \qquad n = 0, \pm 1, \cdots.$$

(2) $\Phi(\mu^*) = (\phi(\mu))^*$.

(3) For $\lambda > 0$,

$$\tfrac{1}{2}[\Phi(i\lambda) + \Phi(-i\lambda)] = \frac{\lambda - \theta_0}{2} + \sum_1^\infty \frac{e^{-n\lambda}}{n}\,\frac{\sinh(\lambda - \theta_0)n}{\cosh n\lambda}$$

For fixed $\eta \geq 1$, in order to investigate the nature of the singularity of $\Phi(\mu)$ at $\mu = 0$ and its thermodynamic consequences, we have to find an asymptotic expansion for $\Phi(\mu)$ valid near $\mu = 0$. This may be done along the lines laid down by Yang and Yang.[10] Equation (A24) may be rewritten as

$$\Phi(\mu) = \tfrac{1}{2}\int_{-\infty}^{\infty} \frac{dx}{x\sinh\pi x}\,G(\mu, x), \quad \text{(A28)}$$

where

$$G(\mu, x) = \frac{\sinh(\pi - \mu)x\ \sinh(\mu - \varphi_0)x}{\cosh\mu x}. \quad \text{(A29)}$$

We define the set S of points in the μ plane by the prescription

$$|\mu| \leq r, \qquad -\delta_0 \leq \arg\mu \leq 0, \quad \text{(A30a)}$$

where

$$\delta_0 = \pi - \epsilon, \qquad \epsilon > 0. \quad \text{(A30b)}$$

Let the straight line P be given by

$$x = te^{i(\pi-\epsilon)/2}, \qquad -\alpha < t < \infty. \quad \text{(A31)}$$

For $\mu \in S$, P is always free of poles of $G(\mu, x)$; consequently the integration path may be deformed to P. Along P, however, all the derivatives with respect to μ of $G(\mu, x)$ are bounded for $\mu \in S$ and so we have the following.

(1) $G(\mu, x) = \sum_0^n c_j(x)\,\mu^j + g_n(x)\,\mu^{n+1}$, (A32)

where $g_n(x) \leq M_n$ for all $x \in P$.

(2) The functions $g_n(x)$ and $c_j(x)$ and even, with

$$g_n(0) = c_j(0) = 0. \quad \text{(A33)}$$

It follows that

$$\Phi(\mu) = \sum_0^n d_j\mu^j + O(\mu^{n+1}), \quad \text{(A34a)}$$

where

$$\alpha_j = \tfrac{1}{2}\int_{-\infty}^{\infty} \frac{c_j(x)\,dx}{x\sinh\pi x}\ . \quad \text{(A34b)}$$

The d_j must grow sufficiently fast to make the radius of convergence of (A34a) zero. The differential coefficients of $\Phi(\mu)$ with respect to μ and η at the point $(\mu = 0, \eta = \eta_0)$ are evidently obtained by termwise differentiation of (A34a). The coefficients d_j may be obtained by rewriting

$$\Phi(\mu) = J_1(\mu) + J_2(\mu), \quad \text{(A35a)}$$

where

$$J_1(\mu) = \tfrac{1}{2}\int_0^\infty \frac{dx}{x}\,\operatorname{sech}(\mu x)\,(e^{-\varphi_0 x} - e^{-(2\mu-\varphi_0)x})$$

$$= \log\left(\frac{\Gamma(1 + s/4)\,\Gamma(\frac{1}{2} - s/4)}{\Gamma(1 - s/4)\,\Gamma(\frac{1}{2} + s/4)}\right) \quad \text{(A35b)}$$

with

$$s = 1 - \varphi_0/\mu. \quad \text{(A35c)}$$

The other integral $J_2(\mu)$ is given by

$$J_2(\mu) = \int_0^\infty \frac{dx}{x}\,\frac{e^{-\pi x}}{\cosh(\mu x)\sinh\pi x} \times \sinh\varphi_0 x\ \sinh(\mu - \varphi_0)x \quad \text{(A35d)}$$

which may be rewritten as

$$J_2(\mu) = \int_0^\infty \frac{dy}{y}\left[\exp\left(\frac{2\pi y}{\mu}\right) - 1\right]^{-1} \times \frac{\cosh(1-s)y - \cosh(1+s)y}{\cosh y}. \quad \text{(A35e)}$$

The generating function

$$\frac{e^{(2\lambda-1)x}}{\cosh x} = \sum_0^\infty E_n(\lambda)\,\frac{(2x)^n}{n!}\ , \qquad |x| < \pi/2, \quad \text{(A36)}$$

for the Euler polynomials, and the formula

$$E_n(1 - x) = (-1)^n E_n(x) \quad \text{(A37)}$$

give the formula

$$J_2(\mu) = \int_0^\infty \frac{dy}{y}\left[\exp\left(\frac{2\pi y}{\mu}\right) - 1\right]^{-1} \sum_1^\infty \frac{y^{2n}}{(2n)!} \times \left[E_{2n}\left(\frac{s}{2}\right) - E_{2n}\left(\frac{-s}{2}\right)\right]. \quad \text{(A38)}$$

By using the standard integral

$$\int_0^\infty \frac{dy}{y}\left[\exp\left(\frac{2\pi y}{\mu}\right) - 1\right]^{-1} y^{2n} = \frac{2^{2n}(-1)^{n+1}}{4n}\,\mu^{2n}B_{2n}, \quad \text{(A39)}$$

one readily finds that

$$J_2(\mu) = \sum_1^\infty \frac{2^{2n}}{(2n)!}\left[E_{2n}\left(\frac{-s}{2}\right) - E_{2n}\left(\frac{s}{2}\right)\right] B_{2n}\,\frac{(i\mu)^{2n}}{4n}\ . \quad \text{(A40)}$$

Thus $\Phi(\mu)$ has the asymptotic expansion

$$\Phi(\mu) = \log \frac{\Gamma(1 + s/4)\,\Gamma(\frac{1}{2} - s/4)}{\Gamma(1 - s/4)\,\Gamma(\frac{1}{2} + s/4)} + \sum_1^\infty \frac{2^{2n}}{(2n)!}\left[E_{2n}\left(\frac{-s}{2}\right) - E_{2n}\left(\frac{s}{2}\right)\right] B_{2n}\,\frac{(i\mu)^{2n}}{4n}, \tag{A41}$$

valid for $\mu \in S$, defined by (A30).

The crucial point is that only even powers of μ appears in (A41), and $\Phi(\mu)$ belongs to the class C^∞ under differentiation with respect to μ^2 at $\mu = 0$ along any radial line in S. By the implicit function theorem, μ^2 is an analytic function of Δ in a neighborhood of $\Delta = -1$. Thus one concludes that, even though $\Phi(\mu)$ is not analytic at $\mu = 0$, nevertheless all temperature derivatives of $z(0)$ exist at $T = T_0$ (corresponding to $\Delta = -1$) and are continuous there. We see that the F model results are prototypical for any $\Delta = -1$ transition point; such transition points are of infinite order in the Ehrenfest scheme.

APPENDIX B

Consider the function

$$I(\alpha,\beta,\gamma) = \int_0^\infty \operatorname{sech}(\alpha x)\,\ln\left(\frac{\cosh x - \cos\beta}{\cosh x - \cos\gamma}\right) dx. \tag{B1}$$

Let the function $F(\beta, y)$ be defined by the integral

$$F(\beta,y) = \tfrac{1}{2}\int_{-\infty}^{\infty} \frac{e^{ixy}}{\cosh x - \cos\beta}\,dx. \tag{B2}$$

Then, for $0 < \mathrm{Re}\beta \le \pi$, we have by contour integration

$$F(\beta,y) = \frac{\pi}{\sin\beta}\,\frac{\sinh[(\pi-\beta)y]}{\sinh(\pi y)}. \tag{B3}$$

These results may be extended to other strips in the β plane by use of the relation

$$F(\beta,y) = F(2k\pi \pm \beta, y) \tag{B4}$$

although $F(\beta,y)$ is not defined for $\mathrm{Re}\beta = 2n\pi$.

The Fourier cosine transform of $\ln[(\cosh x - \cos\beta)/(\cosh x - \cos\gamma)]$ is readily obtained from this function. Since $\sin\beta F(\beta,y)$ is continuous for $0 \le \beta \le 2\pi$, we may integrate it with respect to β. Furthermore, the infinite integrand in (B2) is uniformly convergent with respect to β in $0 < \beta < 2\pi$ so that the order of integration may be reversed in that interval giving

$$\begin{aligned}\int_0^\infty dx\,\cos xy\,\ln\left(\frac{\cosh x - \cos\beta}{\cosh x - \cos\gamma}\right) &= \frac{2\pi}{y}\operatorname{csch}(\pi y)\sinh[(\pi\tfrac{1}{2}(\gamma+\beta))y]\sinh[\tfrac{1}{2}(\beta-\gamma)y] \\ &= \frac{\pi}{y}\{e^{-\gamma y} - e^{-\beta y} + 2e^{-\pi y}\operatorname{csch}(\pi y)\sinh(\tfrac{1}{2}(\gamma+\beta)y) \\ &\quad \times \sinh[\tfrac{1}{2}(\gamma-\beta)y]\}.\end{aligned} \tag{B5}$$

The Fourier cosine transform of $\operatorname{sech}(\alpha x)$ is

$$\int_0^\infty \cos(xy)\,\operatorname{sech}(\alpha x)\,dx = (\pi/2\alpha)\,\operatorname{sech}(\pi y/2\alpha). \tag{B6}$$

Application of the Parseval theorem to (B1) gives

$$\begin{aligned}I(\alpha,\beta,\gamma) = \frac{\pi}{\alpha}\int_0^\infty \frac{dy}{y}\,\operatorname{sech}\left(\frac{\pi y}{2\alpha}\right)\{&2e^{-\pi\gamma}\operatorname{csch}(\pi y) \\ &\times \sinh[\tfrac{1}{2}(\gamma+\beta)y]\sinh[\tfrac{1}{2}(\gamma-\beta)y] \\ &+ e^{-\gamma y} - e^{-\beta y}\}\end{aligned} \tag{B7}$$

for $0 < \mathrm{Re}(\beta,\gamma) < 2\pi$. The extension of this to other ranges of β and γ is achieved by using the relation

$$I(\alpha,\beta,\gamma) = I(\pm\alpha + 2k\pi i, \pm\beta + 2l\pi, \pm\gamma + 2m\pi). \tag{B8}$$

The evaluation of the free energy for KDP and IKDP will now be considered. By using the Laplace transform

$$\int_0^\infty e^{-pt}\left(\frac{1-\operatorname{sech} t}{t}\right)dt = 2\ln\frac{2\Gamma(\frac{1}{4}p + \frac{3}{4})}{p^{1/2}\Gamma(\frac{1}{4}p + \frac{1}{4})}. \tag{B9}$$

it follows that

$$\begin{aligned}\int_0^\infty \frac{dy}{y}\,\operatorname{sech}\left(\frac{\pi y}{2\alpha}\right)(e^{-\gamma y} - e^{-\beta y}) \\ = 2\ln\frac{\Gamma(\beta\alpha/2\pi + \frac{3}{4})\Gamma(\gamma\alpha/2\pi + \frac{1}{4})}{\Gamma(\beta\alpha/2\pi + \frac{1}{4})\Gamma(\gamma\alpha/2\pi + \frac{3}{4})}.\end{aligned} \tag{B10}$$

I. KDP

From (3.3)

$$-\beta\mathfrak{F}_{\mathrm{KDP}} = (1/4\mu)\,I(\pi/2\mu, \mu, 3\mu - 2\pi), \tag{B11}$$

where $2\pi/3 \le \mu \le \pi$. The remaining integral in (B7) may be evaluated by appealing to the table of Laplace transforms in Ref. 16. Thus

$$-\beta\mathfrak{F}_{\mathrm{KDP}} = \ln|(2\mu/\pi)\cot(\pi^2/2\mu)\csc\mu|. \tag{B12}$$

II. IKDP

In this case we have

$$\begin{aligned}-\beta\mathfrak{F}_{\mathrm{IKDP}} = -K + (1/4\mu)[&I(\pi/2\mu, 5\mu - 2\pi, \mu) \\ &+ I(\pi/2\mu, \mu, 3\mu)]\end{aligned} \tag{B13}$$

with $\pi/2 \le \mu \le 2\pi/3$. This gives

$$-\beta\mathfrak{F}_{\mathrm{IKDP}} = \ln|(2\mu/\pi)\cot(\pi^2/2\mu)\csc\mu|. \tag{B14}$$

This same formula therefore holds for KDP and IKDP, the values of μ lying in different ranges.

---

* Present address: D.A.M.T.P., Silver Street, Cambridge, England.
† Partially supported under NSF Grant GP 26526.

1 S. B. Frobenius, Preuss. Acad. Wiss., 514 (1909).
2 E. H Lieb, Phys. Rev. **162**, 162 (1967).
3 E. H. Lieb, Phys. Rev. Lett. **19**, 108 (1967).
4 E. H. Lieb, Phys. Rev. Lett. **18**, 1046 (1967).
5 B. Sutherland, C. N. Yang, and C. P. Yang, Phys. Rev. Lett. **19**, 588 (1967).
6 C. Domb, Advan. Phys. **9**, 149 (1960).
7 J. F. Nagle, Commun. Math. Phys. **13**, 62 (1969).
8 H. Takahashi, Proc. Math. Phys. Soc. Japan **23**, 1069 (1941).
9 B. Sutherland, Phys. Rev. Lett. **19**, 103 (1967).
10 C. N. Yang and C. P Yang, Phys. Rev. **150**, 321 (1966).
11 B. M. McCoy and T. T. Wu, Nuovo Cimento **56B**, 311 (1968).
12 L. Hulthén, Arkiv Mat. Astron. Fysik **26A**, No. 11 (1938).
13 L. R. Walker, Phys. Rev. **116**, 1089 (1959).
14 H. Hancock, *Theory of Elliptic Functions* (Dover, New York, 1958), p. 260.
15 R. J. Baxter, Phys. Rev. Lett. **26**, 832 (1971).
16 M. L. Glasser, Phys. Rev. **184**, 539 (1969). The partition function for IKDP given in this reference is incorrect due to the use of Λ_R rather than Λ_L. The correct value is presented in the Appendix of this paper.
17 F. Y. Wu, Phys. Rev. Lett. **18**, 605 (1967).
18 C. K. Majumdar, Phys. Rev. **145**, 158 (1966).

Canonical Transformations and the Radial Oscillator and Coulomb Problems*

M. Moshinsky, T. H. Seligman,† and K. B. Wolf
Instituto de Física and C.I.M.A.S.S., Universidad de México, México, D.F.
(Received 31 November 1971)

In a previous paper a discussion was given of linear canonical transformations and their unitary representation. We wish to extend this analysis to nonlinear canonical transformations, particularly those that are relevant to physically interesting many-body problems. As a first step in this direction we discuss the nonlinear canonical transformations associated with the radial oscillator and Coulomb problems in which the corresponding Hamiltonian has a centrifugal force of arbitrary strength. By embedding the radial oscillator problem in a higher dimensional configuration space, we obtain its dynamical group of canonical transformations as well as its unitary representation, from the $Sp(2)$ group of linear transformations and its representation in the higher-dimensional space. The results of the Coulomb problem can be derived from those of the oscillator with the help of the well-known canonical transformation that maps the first problem on the second in two-dimensional configuration space. Finally, we make use of these nonlinear canonical transformations, to derive the matrix elements of powers of r in the oscillator and Coulomb problems from a group theoretical standpoint.

1. INTRODUCTION

In previous publications,[1-3] we discussed the role of canonical transformations in quantum mechanics, and when the transformations were linear we obtained their unitary representations in appropriate spaces. These representations have also been derived by other authors from a more abstract standpoint.[4]

As our eventual aim is to obtain relevant canonical transformations and their unitary representations for physically significant many-body problems,[3] we must first deal with problems of one particle in one dimension that go beyond the harmonic oscillator case,[1] which was the starting point of our discussion. Thus in this paper we derive explicitly the dynamical Lie group (and not only the Lie algebra as is customary in the literature) of canonical transformations of the radial oscillator and Coulomb problems in which we have a centrifugal force of arbitrary strength. We then proceed to obtain the unitary representation of this group in configuration space and in the basis in which the Hamiltonian H is diagonal, and finally determine, as a group theoretical problem, the matrix elements of powers of the radial coordinate with respect to eigenstates of H.

By embedding our one-dimensional radial oscillator in a two-dimensional configuration space, we easily derive its dynamical group and the corresponding unitary representation from particular linear canonical transformations in the four-dimensional phase space of the latter problem. The well-known mappings[3,5] between the two-dimensional oscillator and Coulomb problems, allows us then to translate our results to the radial Coulomb case in a straightforward fashion.

2. THE RADIAL OSCILLATOR PROBLEM

We wish to consider a single particle one-dimensional problem whose Hamiltonian (in units in which the mass, frequency of the oscillator and $\hbar$ are 1) is

$$H = \tfrac{1}{2}(p_r^2 + \lambda^2 r^{-2} + r^2) \tag{2.1}$$

The coordinate r varies in the interval $0 \le r < \infty$, p_r is its canonically conjugate momentum, and λ is an arbitrary real constant. For reasons that will appear later we shall denote by μ a real positive constant related to λ through

$$\mu = (\lambda^2 + \tfrac{1}{4})^{1/2} \quad \text{or} \quad \lambda^2 = (\mu - \tfrac{1}{2})[(\mu - \tfrac{1}{2}) + 1]. \tag{2.2}$$

As the Poisson bracket $\{r, p_r\}$ is 1, we conclude that in the quantum mechanical picture $p_r = -i\partial/\partial r$ and thus the eigenstates $f(r)$ of (2.1) satisfy the equation

$$\frac{1}{2}\left(-\frac{d^2}{dr^2} + \frac{\lambda^2}{r^2} + r^2\right) f(r) = Ef(r). \tag{2.3}$$

It is well known that the eigenstates of (2.3) characterized by μ and an integer n have the form

$$f_n^\mu(r) \equiv [2(n!)]^{1/2}[\Gamma(n + \mu + 1)]^{-1/2} e^{-r^2/2} r^{\mu+1/2} L_n^\mu(r^2), \tag{2.4a}$$

where[6] L_n^μ is an associated Laguerre polynomial, and λ and μ are related as in (2.2). The states (2.4a) are orthonormal in the sense

$$\int_0^\infty f_{n'}^\mu(r) f_{n''}^\mu(r) dr = \delta_{n'n''}, \tag{2.4b}$$

and the eigenvalues of (2.3) are given by

$$E_n = (2n + \mu + 1), \quad n \text{ nonnegative integer.} \tag{2.5}$$

A. The Dynamical Group of Canonical Transformations

We wish now to obtain explicitly the dynamical Lie group associated with the Hamiltonian (2.1), and its unitary representation both in configuration space and in the basis where H is diagonal. For this purpose let us first replace λ in (2.1) by a momentum p_θ associated with an angle θ; we have then the two-dimensional Hamiltonian for an oscillator problem which in polar and cartesian coordinates takes the form

$$\mathbf{H} \equiv \tfrac{1}{2}(p_r^2 + r^{-2}p_\theta^2 + r^2) = \tfrac{1}{2}(\mathbf{p}^2 + \mathbf{r}^2) = \tfrac{1}{2}(p_1^2 + x_1^2 + p_2^2 + x_2^2). \tag{2.6}$$

We first recall[1,2] that the dynamical group of canonical transformations of $\mathbf{H}$ is the symplectic group in four dimensions $Sp(4)$. This group has a subgroup

$$Sp(4) \supset Sp(2) \times O(2), \tag{2.7}$$

where $O(2)$ is the rotation group in the two-dimensional space, while $Sp(2)$ is the symplectic group of linear canonical transformations

$$\bar{\mathbf{r}} = a\mathbf{r} + b\mathbf{p}, \quad \bar{\mathbf{p}} = c\mathbf{r} + d\mathbf{p}, \quad ad - bc = 1, \tag{2.8}$$

in which the constants a, b, c, d are real.

We now note that under the transformation (2.8) the angular momentum

$$p_\theta \equiv x_1 p_2 - x_2 p_1 \tag{2.9}$$

remains invariant. As, furthermore, we have that

$$\mathbf{r}\cdot\mathbf{p} = rp_r, \quad \mathbf{p}^2 = p_r^2 + r^{-2}p_\theta^2. \tag{2.10}$$

We see that the transformation (2.8) implies that the new radial coordinate and momentum $\bar{r}, \bar{p}_r$ are given in terms of the old ones r, p_r by

$$\bar{r} = [a^2r^2 + b^2(p_r^2 + \lambda^2 r^{-2}) + 2abrp_r]^{1/2}, \tag{2.11a}$$

$$\bar{p}_r = \frac{acr^2 + bd(p_r^2 + \lambda^2 r^{-2}) + (ad + bc)rp_r}{[a^2r^2 + b^2(p_r^2 + \lambda^2 r^{-2}) + 2abrp_r]^{1/2}}, \tag{2.11b}$$

where we replaced $p_\theta = \bar{p}_\theta$ by a constant value λ.

We have thus obtained the dynamical Lie group associated with the Hamiltonian (2.1) which is a representation of the group of unimodular real matrices

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad ad - bc = 1. \tag{2.12}$$

The subgroup $O(2)$ of (2.12) (not to be confused with the rotation group in two dimensions) for which

$$a = d = \cos\tfrac{1}{2}\alpha, \qquad b = -c = \sin\tfrac{1}{2}\alpha, \tag{2.13}$$

is the symmetry group of the Hamiltonian (2.1) as can be checked directly. We wish to determine the unitary representation of the canonical transformations (2.11) in a basis in which r is diagonal.

B. The Unitary Representation of the Dynamical Group in Configuration Space

We shall limit our discussion to the transformations (2.11) in which $b > 0$. The case $b < 0$ follows immediately[1,2] from it as well as the limit $b \to 0$. The analysis in Ref. 2 then indicates that for the group of linear canonical transformations (2.8) of the two-dimensional oscillator (2.6), the unitary representation is

$$\begin{aligned}\langle \mathbf{r}' | U | \mathbf{r}'' \rangle &= (2\pi b)^{-1} \exp[(-i/2b) \\ &\quad \times (ar'^2 - 2\mathbf{r}'\cdot\mathbf{r}'' + dr''^2)] \\ &= (2\pi b)^{-1} \exp[(-i/2b)(ar'^2 + dr''^2)] \\ &\quad \times \sum_{m=-\infty}^{\infty} i^m J_m(b^{-1}r'r'') e^{im(\theta' - \theta'')},\end{aligned} \tag{2.14}$$

where in the last term we have expanded the two-dimensional plane wave in polar coordinates.[7]

The eigenstates of the two-dimensional oscillator (2.6) in polar coordinates are characterized by the integer quantum numbers n, m and have the explicit form

$$\langle \mathbf{r} | nm \rangle = r^{-1/2} f_n^{|m|}(r)(2\pi)^{-1/2} e^{im\theta}, \tag{2.15}$$

where the radial function is given by (2.4a) with $\mu = (\lambda^2 + \frac{1}{4})^{1/2}$ being replaced by $|m|$. The unitary representation (2.14) with respect to these states is clearly diagonal in the m index as (2.14) is invariant under rotations and, thus, we can write

$$\begin{aligned}\langle n'm | U | n''m \rangle &= \iint \langle n'm | \mathbf{r}' \rangle d\mathbf{r}' \langle \mathbf{r}' | U | \mathbf{r}'' \rangle d\mathbf{r}'' \langle \mathbf{r}'' | n''m \rangle \\ &= \int_0^\infty \int_0^\infty f_{n'}^{|m|}(r') dr' \{ i^m b^{-1} \\ &\quad \times (r'r'')^{1/2} J_m(b^{-1}r'r'') \exp[(-i/2b) \\ &\quad \times (ar'^2 + dr''^2)] \} dr'' f_{n''}^{|m|}(r'').\end{aligned} \tag{2.16}$$

This equation immediately suggests that for $\mu = |m|$ or $\lambda = (m^2 - \frac{1}{4})^{1/2}$ the unitary representation between the radial variables $\langle r' | U_\mu | r'' \rangle$ of the canonical transformation (2.11) [which explicitly depends on μ through (2.2)] is given by the expression inside the curly bracket of (2.16). Thus we may expect that for an arbitrary μ we have for the unitary representation of (2.11)

$$\begin{aligned}\langle r' | U_\mu | r'' \rangle &= b^{-1}(r'r'')^{1/2} J_\mu(b^{-1}r'r'') \\ &\quad \times \exp[(-i/2b)(ar'^2 + dr''^2)].\end{aligned} \tag{2.17}$$

We suppressed the i^μ in (2.17) as, in any case, our representations will be ray representations.[1,2,8]

While (2.17) is rigorously true for $\mu = |m|$, it is only a surmise for other values of μ. We proceed to justify it by obtaining explicitly the unitary representation for arbitrary μ in the basis in which the Hamiltonian (2.1) is diagonal.

C. Unitary Representations of Canonical Transformations in the Basis in Which H Is Diagonal

If we want to have the unitary representation (2.17) in a basis in which H is diagonal, we must calculate the integral

$$\langle n' | U_\mu | n'' \rangle = \int_0^\infty \int_0^\infty f_{n'}^{\mu}(r') dr' \langle r' | U_\mu | r'' \rangle dr'' f_{n''}^{\mu}(r'') \tag{2.18}$$

This integral can be evaluated by exactly the same procedure that was followed in the determination of the matrix element (4.34) in Ref. 2. In fact, we just need to replace l by $\mu - \frac{1}{2}$ and suppress the factor i^l to get the value of the double integral (2.18).

To justify now the value (2.17) for $\langle r' | U_\mu | r'' \rangle$ for arbitrary μ, we notice first that the most general matrix (2.12) of the symplectic group can be written as[1]

$$\begin{aligned}\begin{pmatrix} a & b \\ c & d \end{pmatrix} &= \begin{pmatrix} \cos\frac{1}{2}\alpha & \sin\frac{1}{2}\alpha \\ -\sin\frac{1}{2}\alpha & \cos\frac{1}{2}\alpha \end{pmatrix} \begin{pmatrix} e^{1/2\beta} & 0 \\ 0 & e^{-1/2\beta} \end{pmatrix} \\ &\quad \times \begin{pmatrix} \cos\frac{1}{2}\gamma & \sin\frac{1}{2}\gamma \\ -\sin\frac{1}{2}\gamma & \cos\frac{1}{2}\gamma \end{pmatrix}.\end{aligned} \tag{2.19}$$

The transformations associated with the angles α and γ leave the Hamiltonian (2.1) invariant and thus from a classical standpoint we could identify these angles with time. Therefore when the elements of the matrix (2.19) are given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}, \tag{2.20}$$

the transformation (2.11) gives us the coordinate and momentum at time t from the coordinate and momentum at time 0. The corresponding unitary representation in the basis in which the Hamiltonian is diagonal must then be

$$\langle n' | U_\mu | n'' \rangle_{\alpha = 2t, \beta = \gamma = 0} = \delta_{n'n''} \exp[i(2n' + \mu + 1)t], \tag{2.21}$$

as the energy is given by (2.5). From (4.37) in Ref. 2 we note that we get exactly this value when we replace l by $\mu - (1/2)$ except for a constant phase which is irrelevant because we deal with ray representations. Thus the integral (2.18) gives the correct

unitary representation for any canonical transformation (2.11) in which the symplectic matrix has the form (2.20).

It remains then to check only if the integral (2.18) gives the unitary representation for a dilatation in which

$$a = d^{-1} = e^{\beta/2}, \quad b = c = 0. \tag{2.22}$$

Using the formula[6]

$$L_n^{\mu}(e^{\beta}x^2) = \sum_{k=0}^{n} [k!\,\Gamma(n-k+\mu+1)]^{-1}\Gamma(n+\mu+1) \times e^{(n-k)\beta}(1-e^{\beta})^k L_{n-k}^{\mu}(x^2), \tag{2.23}$$

we can immediately find out the expansion of the states

$$e^{-\beta/4} f_n^{\mu}(e^{\beta/2}r), \tag{2.24}$$

in terms of the states $f_n^{\mu}(r)$ of (2.4). The result turns out to be

$$\begin{aligned}\langle n'|U_{\mu}|n''\rangle_{\alpha=0,\beta,\gamma=0} &= i^{-\mu-1}[n'!\,n''!\,\Gamma(n'+\mu+1)\\ &\times \Gamma(n''+\mu+1)]^{1/2}(\cosh\tfrac{1}{2}\beta)^{-\mu-1}\\ &\times (1-e^{\beta})^{n'+n''}(1+e^{\beta})^{-n'-n''}\\ &\times \sum_{p}\{[p!\,(n'-p)!\,(n''-p)!\,\Gamma(p+\mu+1)]^{-1}\\ &\times (-1)^{n'-p}(\sinh\tfrac{1}{2}\beta)^{-2p}\},\end{aligned} \tag{2.25}$$

and up to a phase it is identical[2] to the one that comes out from the integral (2.18).

We have thus proved that (2.17) is the unitary representation in configuration space of the group of canonical transformations (2.11) for arbitrary $\lambda = (\mu^2 - \frac{1}{4})^{1/2}$. The unitary representation in a basis in which the Hamiltonian is diagonal is given by products of the matrices (2.21) and (2.25) using the decomposition (2.19).

3. THE RADIAL COULOMB PROBLEM

We now wish to consider a Hamiltonian which in atomic units has the form

$$\mathcal{H} = \tfrac{1}{2}(p_r^2 + \Lambda^2 r^{-2}) - r^{-1}, \tag{3.1}$$

with r, p_r having the same meaning as in Sec. 2 and Λ being an arbitrary real constant. As in (2.2) we introduce a positive constant M related to Λ through

$$\mathrm{M} \equiv (\Lambda^2 + \tfrac{1}{4})^{1/2} \quad \text{or} \quad \Lambda^2 = (M - \tfrac{1}{2})[(M - \tfrac{1}{2}) + 1]. \tag{3.2}$$

We shall denote the eigenvalue of the Hamiltonian (3.1) by

$$E \equiv -(2\nu^2)^{-1}. \tag{3.3a}$$

Introducing then the variable

$$\rho = (r/\nu), \tag{3.3b}$$

we see that the eigenstates of (3.1) satisfy the equation

$$\rho\left[\left(-\frac{d^2}{d\rho^2} + \frac{\Lambda^2}{\rho^2}\right) + 1\right]F(\rho) = 2\nu F(\rho). \tag{3.4}$$

The analysis of this equation indicates that the solutions will be regular at ∞ only when

$$\nu = n + \mathrm{M} + \tfrac{1}{2}, \tag{3.5}$$

with the radial quantum number n being a nonnegative integer. The eigenstates have then the explicit form

$$F_n^{\mathrm{M}}(\rho) = A_{n\mathrm{M}}\rho^{\mathrm{M}+1/2}e^{-\rho}L_n^{2\mathrm{M}}(2\rho), \tag{3.6}$$

where[6] $L_n^{2\mathrm{M}}$ are associated Laguerre polynomials in which M and n are related to Λ and ν through (3.2) and (3.5).

The coefficient $A_{n\mathrm{M}}$ can be determined by normalizations which can be achieved in two ways. If we consider F_n^{M} as a function of r, the requirement

$$\int_0^{\infty} F_{n'}^{\mathrm{M}}(r/\nu')F_{n''}^{\mathrm{M}}(r/\nu'')dr = \delta_{n'n''} \tag{3.7a}$$

gives

$$A_{n\mathrm{M}} \equiv A_{n\mathrm{M}}^{c} = \left[\frac{2^{2\mathrm{M}+1}(n-1)!}{(n+\mathrm{M}+\frac{1}{2})\Gamma(2\mathrm{M}+n+1)}\right]^{1/2}. \tag{3.7b}$$

On the other hand, if we consider F_n^{M} as a function of ρ, the operator on the left-hand side of (3.4) will be Hermitian (and thus give rise to orthonormalization), only for integrals of the form

$$\int_0^{\infty} F_{n'}^{\mathrm{M}}(\rho)F_{n''}^{\mathrm{M}}(\rho)\rho^{-1}d\rho = \delta_{n'n''}, \tag{3.8a}$$

which implies that the normalization coefficient becomes

$$A_{n\mathrm{M}} \equiv A_{n\mathrm{M}}^{p} = 2^{\mathrm{M}}[2(n!)/\Gamma(n+2M+1)]^{1/2}. \tag{3.8b}$$

The upper indices c and p distinguish between the two cases when necessary. When referring to the function $F_n^{\mathrm{M}}(\rho)$ without qualifications, we shall understand that it is given by (3.6) with the normalization (3.8b).

We wish now to obtain explicitly the dynamical Lie group associated with the problem (3.4) as well as its unitary representation. We can achieve both objectives through the mappings between the two-dimensional oscillator and Coulomb problems.

A. Mappings between the Two-Dimensional Oscillator and Coulomb Problems

From the two-dimensional coordinate and momentum vectors $\mathbf{r}, \mathbf{p}$ we can build the following independent quadratic expressions[1,2]

$$\begin{aligned}&I_1 \equiv \tfrac{1}{4}(\mathbf{p}^2 - \mathbf{r}^2), \quad I_2 \equiv \tfrac{1}{4}(\mathbf{r}\cdot\mathbf{p} + \mathbf{p}\cdot\mathbf{r}),\\ &I_3 \equiv \tfrac{1}{4}(\mathbf{p}^2 + \mathbf{r}^2) = \tfrac{1}{2}H.\end{aligned} \tag{3.9}$$

The Poisson brackets between the I_i, either when they are considered as classical observables or quantum mechanical operators, are given by

$$\{I_1, I_2\} = -I_3, \quad \{I_3, I_1\} = I_2, \quad \{I_2, I_3\} = I_1. \tag{3.10}$$

Thus they correspond to the generators[9] of a Lie algebra of the group $SU(1,1)$, which is isomorphic[1,3] to the $Sp(2)$ group of linear canonical transformations (2.8).

Turning now our attention to the two-dimensional Coulomb problem, we describe it in terms of the radial coordinate ρ and an angle φ. The corresponding Cartesian coordinates we designate by

$$\xi_1 = \rho\cos\varphi, \quad \xi_2 = \rho\sin\varphi \tag{3.11}$$

and their canonically conjugate momenta by π_1, π_2.

In terms of the two-dimensional vectors $\boldsymbol{\xi}, \boldsymbol{\pi}$ we can now construct the expressions[10]

$$K_1 \equiv \tfrac{1}{2}\rho(\boldsymbol{\pi}^2 - 1), \quad K_2 \equiv \tfrac{3}{4}\boldsymbol{\xi}\cdot\boldsymbol{\pi} + \tfrac{1}{4}\boldsymbol{\pi}\cdot\boldsymbol{\xi}, \quad K_3 \equiv \tfrac{1}{2}\rho(\boldsymbol{\pi}^2 + 1). \tag{3.12}$$

The Poisson brackets between the K_i, both classically and quantum mechanically, are given by

$$\{K_1, K_2\} = -K_3, \quad \{K_3, K_1\} = K_2, \quad \{K_2, K_3\} = K_1. \tag{3.13}$$

Thus they correspond to the generators[9] of the Lie algebra of $SU(1,1)$. The operators K_i are Hermitian under the measure used in (3.8a).

Before proceeding with the explicit construction of the group whose generators are the K_i of (3.13), we first indicate why we are interested in it. When we write K_3 as a quantum mechanical operator in polar coordinates, we immediately notice that its eigenstates are given by the wavefunction

$$\psi(\rho, \varphi) = \rho^{-1/2}F(\rho)e^{i\mathrm{M}\varphi}, \tag{3.14}$$

in which $F(\rho)$ satisfies Eq. (3.4) with $\mathrm{M} = (\Lambda^2 + \frac{1}{4})^{1/2}$ being an integer. Thus the relation of the problem (3.4) with the Lie algebra (3.13) is exactly of the same type as the one that exists between the radial equation (2.3) for the harmonic oscillator and the Lie algebra whose generators are the I_i given by (3.9). As the latter relation allowed us to determine,[1-3] the group of canonical transformations and its unitary representation for the harmonic oscillator problem, we expect that the former relation will achieve the same objectives for the Coulomb problem.

We now consider a canonical transformation that converts the I_i of (3.9) into K_i of (3.12) assuming them to be classical observables. The mapping appears in its simplest form in polar coordinates if we consider the relations[3,5]

$$r^2 = 2\rho, \tag{3.15a}$$

$$\theta = \tfrac{1}{2}\varphi \tag{3.15b}$$

This implies that in Cartesian coordinates we have

$$\xi_1 = \tfrac{1}{2}(x_1^2 - x_2^2), \tag{3.16a}$$

$$\xi_2 = x_1x_2. \tag{3.16b}$$

To determine the corresponding relation for momenta we recall that in classical mechanics the generalized velocities and momenta are connected by Hamilton's equation

$$\dot{q}_i = \partial H/\partial p_i. \tag{3.17}$$

For the two-dimensional oscillator and Coulomb problems the H are, respectively, $2I_3$ and $2K_3$ and thus

$$p_i = \dot{x}_i, \quad \pi_i = \dot{\xi}_i/2\rho, \quad i = 1, 2. \tag{3.18}$$

From (3.16a) and (3.16b) we obtain then that

$$\pi_1 = r^{-2}(x_1p_1 - x_2p_2), \tag{3.16c}$$

$$\pi_2 = r^{-2}(x_1p_2 + x_2p_1). \tag{3.16d}$$

We easily check that the transformation (3.16) is canonical and besides it maps the generators I_i of (3.9) of the Lie Algebra of the harmonic oscillator into the K_i of (3.12) of the Coulomb problem. From (3.16) we note also that

$$\pi_\varphi = \xi_1\pi_2 - \xi_2\pi_1 = \tfrac{1}{2}(x_1p_2 - x_2p_1) = \tfrac{1}{2}p_\theta, \tag{3.19a}$$

$$\rho\pi_\rho = \boldsymbol{\xi}\cdot\boldsymbol{\pi} = \tfrac{1}{2}\mathbf{r}\cdot\mathbf{p} = \tfrac{1}{2}rp_r, \tag{3.19b}$$

$$\boldsymbol{\pi}^2 = \mathbf{p}^2/r^2. \tag{3.19c}$$

In particular the transformation (3.15) maps I_i on K_i even when we interpret them as quantum mechanical operators.

With the help of the canonical transformations (3.15) and (3.19), we are now in position to obtain the dynamical group of canonical transformations associated with problem (3.4).

B. The Dynamical Group of the Coulomb Problem and Its Unitary Representation

The dynamical group of canonical transformations associated with the harmonic oscillator problem is given by (2.11). In the Coulomb problem our radial variable is ρ and its corresponding momentum π_ρ. From the relations (3.15) and (3.19) connecting ρ, π_ρ and r, p_r we conclude that (2.11) gives rise to the following nonlinear canonical transformations for the Coulomb problem

$$\bar{\rho} = \tfrac{1}{2}\bar{r}^2 = \tfrac{1}{2}(a^2r^2 + b^2p^2 + 2ab\mathbf{r}\cdot\mathbf{p}) = \rho[(a + b\pi_\rho)^2 + b^2\Lambda^2\rho^{-2}], \tag{3.20a}$$

$$\bar{\pi}_\rho = \bar{\rho}^{-1}\bar{\boldsymbol{\xi}}\cdot\bar{\boldsymbol{\pi}} = \frac{(a + b\pi_\rho)(c + d\pi_\rho) + bd\Lambda^2\rho^{-2}}{(a + b\pi_\rho)^2 + b^2\Lambda^2\rho^{-2}}, \tag{3.20b}$$

where we have replaced $\pi_\varphi = \bar{\pi}_\varphi$ by the value Λ; it takes in the classical picture. Thus we have obtained the dynamical Lie group of canonical transformations associated with the Coulomb system in the formulation (3.4).

To discuss the unitary representation of the canonical transformation (3.20), we first notice that from (3.15a) and (3.19b) we can map the classical oscillator problem (2.1) with $\lambda = 2\Lambda$, onto the Hamiltonian

$$\rho(\pi_\rho^2 + \Lambda^2\rho^{-2} + 1), \tag{3.21}$$

with the help of the canonical transformations

$$\rho = \tfrac{1}{2}r^2, \tag{3.22a}$$

$$\pi_\rho = r^{-1}p_r. \tag{3.22b}$$

Thus if we obtain the unitary representation V associated with (3.22), we can determine the corresponding one for (3.20) through the similarity transformation

$$W_{\mathrm{M}} \equiv V^{-1}U_\mu V, \tag{3.23}$$

where the matrix elements of U_μ are given by (2.17). In this expression we have to use the relation $\mu = 2\mathrm{M}$,

which is the quantum mechanical equivalent of $\lambda = 2\Lambda$, as in quantum mechanics μ and M are, respectively, the eigenvalues of p_θ and π_φ related by $p_\theta = 2\pi_\varphi$ as indicated in the previous subsection.

To get V, we proceed as in Refs. 1 and 2. Using the Dirac notation in which we indicate by $\langle r'|, |r''\rangle$ and $(\rho'|, |\rho'')$ bras and kets in which r and ρ are, respectively, diagonal, we look for the transformation bracket $\langle r'|\rho'\rangle$ that satisfies the equation[1,2]

$$\rho\langle r'|\rho'\rangle = \rho'\langle r'|\rho'\rangle. \tag{3.24}$$

From (3.22a) this equation implies that the transformation bracket is proportional to $\delta(\rho' - \frac{1}{2}r'^2)$. If we further require that the bracket should be orthonormal in the sense

$$\int_0^\infty (\rho'|r'\rangle dr'\langle r'|\rho'') = \delta(\rho' - \rho''), \tag{3.25}$$

we obtain that

$$\langle r'|\rho'\rangle = (2\rho')^{1/4}\delta(\rho' - \tfrac{1}{2}r'^2). \tag{3.26}$$

We can multiply the expression (3.26) by any phase factor which is a function of ρ'; but this proves unnecessary as the transformation bracket (3.26) already guarantees that

$$(\rho'|\rho|\rho'') = \int (\rho'|r'\rangle \tfrac{1}{2}r'^2\langle r'|\rho''\rangle dr' = \rho'\delta(\rho' - \rho''), \tag{3.27a}$$

$$\begin{aligned}(\rho'|\pi_\rho|\rho'') &= \int (\rho'|r'\rangle\left[\left(\frac{1}{ir'}\frac{\partial}{\partial r'} + \frac{i}{2r'^2}\right)\langle r'|\rho''\rangle\right]dr' \\ &= \int (\rho'|r'\rangle\left(-\frac{1}{i}\frac{\partial}{\partial\rho''}\langle r'|\rho''\rangle\right)dr' \\ &= -\frac{1}{i}\frac{\partial}{\partial\rho''}\delta(\rho' - \rho''),\end{aligned} \tag{3.27b}$$

where we made use of the fact that the Hermitian form of the quantum mechanical operator π_ρ of (3.22b) is

$$\tfrac{1}{2}(r^{-1}p_r + p_r r^{-1}) = \frac{1}{ir}\frac{\partial}{\partial r} + \frac{i}{2r^2}. \tag{3.27c}$$

The unitary representation of the canonical transformation (3.22) is thus given by the bracket (3.26) and therefore the matrix element of $V^{-1}U_{2M}V$ in the representation in which ρ is diagonal takes the form

$$\begin{aligned}(\rho'|W_M|\rho'') &= (\rho'|V^{-1}U_{2M}V|\rho'') \\ &= \iint (\rho'|r'\rangle dr'\langle r'|U_{2M}|r''\rangle dr''\langle r''|\rho''). \end{aligned} \tag{3.28}$$

Substituting the values (2.17) for $\langle r'|U_{2M}|r''\rangle$ and (3.26) for $\langle r'|\rho'\rangle$, we obtain

$$\begin{aligned}(\rho'|W_M|\rho'') = b^{-1}J_{2M}[2b^{-1}(\rho'\rho'')^{1/2}] \\ \times \exp[(-i/b)(a\rho' + d\rho'')].\end{aligned} \tag{3.29}$$

The unitary representation when the Hamiltonian (3.21) is diagonal, rather than the observable ρ, is given by

$$\begin{aligned}(n'|W_M|n'') = \int_0^\infty\int_0^\infty \rho'^{-1/2}F_{n'}^M(\rho') \\ \times d\rho'(\rho'|W_M|\rho'')d\rho''\,\rho''^{-1/2}F_{n''}^M(\rho''),\end{aligned} \tag{3.30}$$

where $F_n^M(\rho)$ is given by (3.6) and the extra factors $\rho'^{-1/2}, \rho''^{-1/2}$ come from the normalization (3.8). As from (2.4a) and (3.6) we have that

$$F_n^M(\rho) = (\rho/2)^{1/4}f_n^{2M}[(2\rho)^{1/2}], \tag{3.31}$$

we immediately obtain that

$$(n'|W_M|n'') = \langle n'|U_{2M}|n''\rangle, \tag{3.32}$$

with the latter expression being given by (2.21) when the transformation is of the type (2.20) and has the form (2.25) for a dilatation.

It is important to keep in mind that our canonical transformations and their representations are not so much connected with the Coulomb problem (3.1) as with the one whose Hamiltonian is (3.21). The latter is directly related, when $\Lambda^2 = l(l+1)$, with the stereographic projection of a four-dimensional point rotor on a three-dimensional momentum space as was first pointed out by Fock.[11] Thus we shall refer to the problem whose Hamiltonian is (3.21) as the pseudo-Coulomb problem and our analysis, so far, has been restricted to it.

4. RADIAL MATRIX ELEMENTS

The matrix elements of powers of the radial coordinate with respect to oscillator or Coulomb wavefunctions are easily evaluated using properties of the Laguerre polynomials or their generating functions.[6] We wish though to obtain their values through the use of the dynamical group of canonical transformations so as to develop a procedure susceptible to generalization to more complex problems.

For the oscillator case the radial integrals were already determined through the use of the dynamical group of canonical transformations[12] as well as by other group theoretical approaches.[13] We have thus to concentrate on the Coulomb problem, on which group theoretical methods have been developed,[14] but they do not use canonical transformations. Rather than enter into this problem directly it will prove more effective to discuss first the matrix elements of r^{2k}, k integer, in the oscillator problem, from an angle different from the one used in Ref. 12. Once we determine these matrix elements the extension to the pseudo-Coulomb problem will be achieved through the mappings (3.22a) and (3.31), while the actual Coulomb integrals of r^k can be obtained from those of the pseudo-Coulomb problem and the expression (2.25) for dilatations.

A. Matrix Elements of r^{2k} for Oscillator States

We start our discussion by noticing that if in the generators (3.9) of the dynamical group of the two-dimensional oscillator, we replace

$$p^2 = p_r^2 + r^{-2}p_\theta^2, \qquad p_\theta = \lambda, \tag{4.1}$$

we get the observables

$$\begin{aligned}&I_1 = \tfrac{1}{4}(p_r^2 + \lambda^2 r^{-2} - r^2), \qquad I_2 = \tfrac{1}{4}(rp_r + p_r r), \\ &I_3 = \tfrac{1}{4}(p_r^2 + \lambda^2 r^{-2} + r^2).\end{aligned} \tag{4.2}$$

The Poisson brackets of the I_i, both classically and quantum mechanically, have the value (3.10) and thus these observables are the generators of Lie Algebra

of a group $Sp(2)$ as well as groups isomorphic to it[1] such as $SU(1,1)$.

From I_1 and I_2 we can construct the operators

$$I_\pm \equiv I_i \pm iI_2 \tag{4.3}$$

and we easily check that for the $f_n^\mu(r)$ of (2.4a) we have

$$I_\pm f_n^\mu(r) = [(n + \mu + \tfrac{1}{2} \pm \tfrac{1}{2})(n + \tfrac{1}{2} \pm \tfrac{1}{2})]^{1/2} f_{n\pm1}^\mu(r). \tag{4.4}$$

Thus $I_\pm$ are raising and lowering operators for the index n and the set of $f_n^\mu(r), n = 0, 1, 2, \ldots$, belong to a single unitary irreducible representation of the $SU(1,1)$ group. The lowest weight element of the set corresponds to $n = 0$ and, thus, is an eigenstate of I_3 with eigenvalues $\frac{1}{2}(\mu + 1)$. We can use this number or more compactly μ itself to label the irreducible representation. If μ is 0 or a positive integer, the representations are part of the discrete series discussed by Bargmann[9]: They are single valued on the $SU(1,1)$ group manifold and are characterized by the integer or semi-integer numbers

$$\tfrac{1}{2}(\mu + 1) = \tfrac{1}{2}, 1, \tfrac{3}{2}, 2, \tfrac{5}{2}, \cdots. \tag{4.5}$$

When μ is an arbitrary nonnegative real number the representations are multiple valued on the $SU(1,1)$ group manifold and are not discussed in Bargmann's paper; but they are as straightforward to obtain as those when μ is integer. In fact, a basis for multi-valued-irreducible representations for arbitrary μ is precisely given by the functions $f_n^\mu(r), n = 0, 1, 2, \ldots$, with μ specifying the irreducible representation and n indicating its row.

We wish now to characterize r^{2k} as a linear combination of irreducible tensors of $SU(1,1)$. When we achieve this purpose we can make use of the Wigner–Eckart theorem to express the matrix elements of r^{2k} in terms of the Wigner coefficients of $SU(1,1)$ determined by Ui.[15] To reach our objective we notice from (4.2) that

$$\tfrac{1}{2}r^2 = I_3 - I_1. \tag{4.6}$$

Thus r^{2k} can be expressed as a polynomial in the generators I_i of the dynamical group. To develop this polynomial in terms of irreducible tensors of $SU(1,1)$, we first introduce the auxiliary generators

$$I_1' \equiv iI_1, \quad I_2' \equiv iI_2, \quad I_3' \equiv I_3. \tag{4.7}$$

The $SU(1,1)$ Casimir operator can then be written as

$$I_3^2 - I_1^2 - I_2^2 = I_1'^2 + I_2'^2 + I_3'^2, \tag{4.8}$$

so we can deal formally with the I_i' as generators of a rotation group. In order to express $\frac{1}{2}r^2$ as a lowering operator in $SU(2)$ we rotate the generators by $\pi/2$ around the axis 1, i.e.

$$I_1'' = I_1', \quad I_2'' = I_3', \quad I_3'' = -I_2', \tag{4.9}$$

and thus

$$\tfrac{1}{2}r^2 = I_3' + iI_1' = i(I_1'' - iI_2''). \tag{4.10}$$

As the rank 1 irreducible tensors in $SU(2)$ are

$$I''_{\pm1} = \mp (2)^{-1/2}(I_1'' \pm iI_2''), \quad I_0 = I_3'', \tag{4.11}$$

we can build the rank k and projection $-k$ irreducible tensor in the I_i'' as

$$(\tfrac{1}{2}r^2)^k = i^k(I_1'' - iI_2'')^k = i^k 2^{k/2}(I''_{-1})^k \equiv i^k 2^{k/2}\mathcal{T}^k_{-k}(I''). \tag{4.12}$$

We can now express r^{2k} in terms of I' undoing (4.9) through the rotation matrices $\mathfrak{D}^j_{mm'}(\alpha\beta\gamma)$, i.e.,[16]

$$\begin{aligned}(\tfrac{1}{2}r^2)^k &= i^k 2^{k/2} \sum_\tau \mathcal{T}^k_\tau(I')\mathfrak{D}^k_{\tau,-k}(\pi/2, \pi/2, -\pi/2) \\ &= \sum_\tau (-i)^\tau 2^{-k/2}[(2k)!]^{1/2} \\ &\quad \times [(k+\tau)!\,(k-\tau)!]^{-1/2}\mathcal{T}^k_\tau(I').\end{aligned} \tag{4.13}$$

Now in order to pass from the $SU(2)$ irreducible tensors $\mathcal{T}^k_\tau(I')$ of (4.13), to the $SU(1,1)$ irreducible tensors $\mathcal{T}^k_\tau(I)$, we notice that the former are defined by

$$SU(2)\begin{cases} I_\pm' = I_1' \pm iI_2', & [I_\pm', \mathcal{T}^k_\tau(I')] = [(k \mp \tau)(k \pm \tau + 1)]^{1/2}\mathcal{T}^k_{\tau\pm1}(I') \\ I_0' = I_3', & [I_0', \mathcal{T}^k_\tau(I')] = \tau\mathcal{T}^k_\tau(I'), \end{cases} \tag{4.14}$$

while the latter are characterized by[12,15]

$$SU(1,1)\begin{cases} I_\pm = I_1 \pm iI_2, & [I_\pm, \mathcal{T}^k_\tau(I)] = \pm\,[(k \mp \tau)(k \pm \tau + 1)]^{1/2}\mathcal{T}^k_{\tau\pm1}(I) \\ I_0 = I_3, & [I_0, \mathcal{T}^k_\tau(I)] = \tau\mathcal{T}^k_\tau(I). \end{cases} \tag{4.15}$$

From these relations and (4.7) we see that

$$\mathcal{T}^k_\tau(I') = i^\tau\mathcal{T}^k_\tau(I). \tag{4.16}$$

Carrying the corresponding substitution in (4.13), we finally obtain, using the Wigner–Eckart theorem,

$$\begin{aligned}I^{k\mu}_{n'n} &\equiv \int_0^\infty f^\mu_{n'}(r)(\tfrac{1}{2}r^2)^k f^\mu_n(r)dr \\ &= \sum_\tau \left[\left(\frac{(k!)^2}{(k+\tau)!\,(k-\tau)!}\right)^{1/2} \frac{\langle \mu, n, k, \tau \mid \mu, n'\rangle_{\text{n.c.}}}{\langle \mu, 0, k, 0 \mid \mu, 0\rangle_{\text{n.c.}}}\right] \\ &\quad \times \int_0^\infty f_0^\mu(r)(\tfrac{1}{2}r^2)^k f^\mu_0(r)dr.\end{aligned} \tag{4.17}$$

The brackets $\langle|\rangle_{\text{n.c.}}$ stand for the Wigner coefficients of the noncompact group $SU(1,1)$. These coefficients were given by Ui[15] for integer μ; but this formula is still valid for arbitrary μ. As the last integral is trivial to determine, we get from the explicit expression of $\langle|\rangle_{\text{n.c.}}$, and the selection rule[15] $n + \tau = n'$,

$$I^{k\mu}_{n'n} = \frac{(-1)^{n+n'}}{2^k}\frac{k!\,\Gamma(k+\mu+1)}{(k+n'-n)!}\left(\frac{n!\,n'!\,\Gamma(n'+\mu+1)}{\Gamma(n+\mu+1)}\right)^{1/2} \times \sum_{p=\max(0,n-n')}^{\min(n,\,k+n-n')}\left(\frac{(k+n'-n+p)!}{p!(k+n-n'-p)!\,(n-p)!\,(n'-n+p)!\,\Gamma(n'-n+\mu+p+1)}\right). \quad (4.18)$$

The relation $-k \le \tau \le k$ implies the selection rule $|n-n'| \le k$.

B. Matrix Elements of ρ^k for the Pseudo-Coulomb Problem

For the pseudo-Coulomb problem we need to calculate the matrix elements

$$\int_0^\infty F^{\mathrm{M}}_n(\rho)\rho^k F^{\mathrm{M}}_{n'}(\rho)\rho^{-1}d\rho, \quad (4.19)$$

as we are using the normalization condition (3.8). Taking then (3.22a) and (3.31) into account, we immediately see that the integral is identical to (4.17) and thus is given by the $I^{k2\mathrm{M}}_{n'n}$ of (4.18).

C. Matrix Elements of r^k for the Coulomb Problem

For the Coulomb problem the states are normalized according to (3.7) and thus we are interested in the integral

$$J^{k\mathrm{M}}_{nn'} \equiv (A^c_{n\,\mathrm{M}}/A^p_{n\mathrm{M}})(A^c_{n'\,\mathrm{M}}/A^p_{n'\mathrm{M}})\int_0^\infty F^M_n(r/\nu)r^k F^M_{n'}(r/\nu')dr, \quad (4.20)$$

where $A^c_{n\,l}, A^p_{n\,l}$ are given by (3.7b) and (3.8b), respectively, and ν, ν' and n, n' are related by (3.5). As before the functions $F^{\mathrm{M}}_n(\rho)$ are normalized in the sense (3.8). Introducing then the variable $\rho = (r/\nu)$, we can write

$$J^{k\mathrm{M}}_{nn'} = [nn'(n+\mathrm{M}+\tfrac{1}{2})(n'+M+\tfrac{1}{2})]^{-1/2} \times \nu^{k+1}\int_0^\infty F^{\mathrm{M}}_n(\rho)\rho^{k+1}F^{\mathrm{M}}_{n'}(\nu\rho/\nu')\rho^{-1}d\rho. \quad (4.21)$$

From (2.17) we easily see that

$$\lim_{\substack{b\to 0\\ d=a^{-1}}} \langle r'|U_\mu|r''\rangle = i^{-\mu-1}a^{-1/2}\delta(r'-a^{-1}r''), \quad (4.22)$$

and thus using (3.32) we obtain

$$F^{\mathrm{M}}_{n'}(\nu\rho/\nu') = \sum_{n''}(\nu/\nu')^{1/2}F^{\mathrm{M}}_{n''}(\rho)i^{\mu+1} \times \langle n''|U_{2M}|n'\rangle_{\alpha=0,\beta=\ln(\nu/\nu'),\gamma=0}, \quad (4.23)$$

where the matrix element is given by (2.25). Combining the previous results, we obtain

$$J^{k\mathrm{M}}_{nn'} = [nn'(n+\mathrm{M}+\tfrac{1}{2})(n'+\mathrm{M}+\tfrac{1}{2})]^{-1/2}(\nu/\nu')^{1/2}i^{\mu+1} \times \sum_{n''}\{I^{k2M}_{nn''}\langle n''|U_{2M}|n'\rangle_{\alpha=0,\beta=\ln(\nu/\nu'),\gamma=0}\}, \quad (4.24)$$

where, because $|n-n''| \le k$, the summation is a finite one.

We have thus achieved by group theoretical means the determination of the radial integrals in the pseudo-Coulomb and Coulomb problems. We note though that we have the same M in both radial wavefunctions. This is unavoidable if we consider the radial problem as one-dimensional, forgetting its relations with other coordinates in a higher-dimensional space in which it can be embedded. If we think in terms of the groups of canonical transformations in these higher-dimensional spaces, we can obtain matrix elements for different irreducible representations of $SU(1,1)$ in bra and ket as was shown for the oscillator case in Ref. 12.

* Work supported by the Comisión Nacional de Energía Nuclear, México.

† Fellow of the Schweizerischer Nationalfonds der Wissenschaften.

1 M. Moshinsky and C. Quesne, Proceedings of the XV Solvay Conference, Brussels (1970).

2 M. Moshinsky and C. Quesne, J. Math. Phys. **12**, 1772 (1971).

3 M. Moshinsky, "Canonical Transformations and Quantum Mechanics," *Notes of the Latin American School of Physics* (University of México, México, 1971).

4 C. Itzykson, Commun. Math. Phys. **4**, 92 (1967); V. Bargmann "Group Representations on Hilbert Spaces of Analytic Functions" in *Analytic Methods in Mathematical Physics*, edited by Gilbert and Newton (Gordon & Breach, New York, 1968).

5 D. Bergman and Y. Frischman, J. Math. Phys. **6**, 1855 (1965).

6 M. Abramowitz and I. A. Stegun, Editors, *Handbook of Mathematical Functions* (Dover, New York 1965), pp. 775, 781, 784, 785.

7 P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), p. 1371.

8 M. Hammermesh, *Group Theory* (Addison-Wesley, Reading, Mass., 1964), p. 469.

9 V. Bargmann, Ann. Math. **48**, 568 (1947).

10 K. B. Wolf, Nuovo Cimento Suppl. **5**, 1041 (1967)

11 V. Fock, Z. Physik **98**, 145 (1935).

12 C. Quesne and M. Moshinsky, J. Math. Phys. **12**, 1780 (1971).

13 L. Armstrong Jr., J. Math. Phys. **12**, 953 (1971): A. Crubeillier and S. Feneuille (to be published).

14 L. Armstrong Jr., Phys. Rev. **3**, 1546 (1971); S. Feneuille, Compt. Rend. **271**, 992 (1970); B. Wybourne (private communication).

15 H. Ui, Ann. Phys. (N.Y.) **49**, 69 (1968).

16 E. P. Wigner, *Group Theory* (Academic, New York, 1959), pp. 157–70.

Decision Procedures in Quantum Mechanics*

Paul A. Benioff
Chemistry Division, Argonne National Laboratory, Argonne, Illinois 60439
(Received 11 November 1971)

The results of an earlier paper on finite and infinite sequences of measurements are here extended to include decision procedures. It is shown that with each decision procedure Q there is uniquely associated a probability operator measure O^Q, which gives the statistical properties of Q. None, some, or all of the paths of Q can be infinitely long. A result of this association is that there are two methods of measuring the probability that carrying out Q on a system in state ρ gives an outcome sequence in some set F. A remarkable aspect of this equivalence is that the purely physical operation of one method is equivalent to, or can replace, the physical operation and mathematical decision procedure of the other method.

I. INTRODUCTION

In a previous paper,[1] hereafter called I, processes which consisted of finite or infinite sequences of measurements of observables were considered. The processes were restricted to be such that each observable was discrete, von Neumann's projection axiom[2] was applicable, and the process was "passive." That is, the sequence of observables and transformations was fixed at the outset and was independent of outcomes of prior measurements.

The main result of I was that, to each finite or infinite process which satisfied these requirements, there corresponds a unique probability operator measure which contains or encodes all the statistical properties of the process. It was also shown that, as a result of this correspondence, there are two ways to measure the probability that carrying out a process on a system in state ρ gives an outcome sequence in some set F. One method consists of an infinite repetition of carrying out the process on ρ, which generates an infinite sequence of outcome sequences. This sequence is then used to generate, by mathematical decision procedures, an infinite sequence of 0's and 1's whose limit mean is the desired probability. The other method consists in repeated measurements on ρ of the observable whose corresponding operator is the one the probability operator measure assigns to F. Again the limit mean of the resultant infinite sequence is the desired probability. (In this paper, as in I, infinite repetition of a procedure or observable measurement on a system in state ρ will always mean an infinite repetition of the following: Prepare a system in state ρ, carry out the procedure or observable measurement which gives an outcome (sequence), then discard the system.)

In this paper we extend these results to cover processes which are not necessarily passive; that is, processes consisting of discrete bounded observables and transformations and which use von Neumann's projection axiom are considered. However, the operation used for the $(j + 1)$th step of any path can depend on the outcomes of the measurements (if any) made in the first j steps. Since such processes require a choice of operations to be made at one or more steps they are referred to as "active" processes or decision procedures.

Section II begins with a description of finite and infinite decision procedures in terms of trees. After giving an example and some elementary properties of trees, the main result is obtained: that with each decision procedure consisting of infinite paths only, there is associated in a unique manner a probability operator measure which contains the statistical properties of the process. This result is then extended to decision procedures containing one or more finite paths, including those with finite paths only.

In Sec. III some consequences of this association are discussed. The two possible methods of measuring the probability that carrying out a decision procedure Q on a system in state ρ gives an outcome sequence in a set F are discussed. A brief sketch of the proof that these two methods are equivalent and do give the probability in question, is given. It is noted that the proof (like that given in I) requires the concept of a probability measure being "correct" for an outcome sequence. This concept, which forms the central part of a definition of agreement between theory and experiment, given elsewhere,[3] is based on the idea of a set being definable by a formula in the language of a formal theory.

Again as in I, it is noted that a remarkable aspect of this equivalence is that the purely physical operation of one of the methods is equivalent to the physical operation and mathematical decision procedures of the other method. It is pointed out that, for decision procedures, this equivalence appears to be more remarkable than for passive procedures.

The question may arise regarding why one should extend the results of I to include decision procedures. Besides the general importance of such procedures,[4] a basic reason for such an extension is that it becomes possible to discuss realizations of mathematical computation procedures in quantum mechanics. Clearly, any computer program is a decision procedure and, when broken down into elementary steps, would appear to be describable in terms of isometries (operations on the memory elements) and observables (reading the memory elements).

In particular, the possible relevance of this work to studies in the foundations of mathematics arises from the point that Turing machines appear to be describable by a decision procedure as defined here. The basic observation step of reading a square can be described by an observable and the basic transformations of right shift, left shift, change a 1 to a 0 and change a 0 to a 1, can be described by isometries. One aspect of the potential importance of this lies in the fact that Turing computability is equivalent to effective computability, and thus the theorems of any formal axiom system are Turing enumerable (relative to a decision procedure for whether or not any formula is an axiom).

II. DECISION PROCEDURES AND PROBABILITY OPERATOR MEASURES

A. Tree Structure

Here a decision procedure is defined to be any process where for each j the choice of the operation for the $(j + 1)$th step in any path can depend on the outcomes (if any) of the measurements made in the prior j steps.

The operations can be either transformations or discrete bounded observables. The detailed form of a decision procedure is irrelevant here. It may be written out in English or given in some language understandable by a computer. It must be meaningful, though, to who or whatever reads it.

With each decision procedure Q, one can associate a tree τ_Q consisting of all possible paths one could follow in carrying out Q. Each vertex of τ_Q, in our case, corresponds to carrying out either a transformation or measuring an observable. In the former case, only one line leads out of the vertex and in the latter case there is exactly one line for each possible outcome of measuring the observable. The fact that only one line leads out of a transformation vertex does not mean that in actually carrying out the procedure, one has looked to see if the transformation really is the one given in the procedure. Such a looking would be a measurement by some observable and would have to be included in the procedure Q, as an observation step.

A simple example of a decision procedure Q is the process "wait for a time t after the initial state preparation (which is not part of Q) and then measure observable $A = \sum_i a_i P_{a_i}$. If the outcome is a_i, translate and rotate state by $(\mathbf{x}_i, t_i, \Omega_i)$ and measure observable $B^i = \sum_j b_j^i P_{b_j^i}$. Stop unless outcome $b_{\frac{1}{2}}$ is observed. In this case, measure observable $C = \sum_j c_j P_{c_j}$ and stop."

The tree τ_Q describing this process is given by Fig. 1, where the small circles denote vertices and the first step (wait t) is step 0. The dotted lines denote other paths leading out of the vertices. The vertices denote operations and the line segments coming out of observation vertices denote outcomes.

Some definitions are useful in what follows. A path is a maximal, well-ordered set of vertices (operations), where the well ordering is defined in the usual way (step 0 is the initial step of all paths). Thus, in Fig. 1, $(t), A, a_2(\mathbf{x}_2, t_2, \Omega_2), B^2, b_1^2(S)$ with S denoting "stop" is a path. $a_2(\mathbf{x}_2, t_2, \Omega_2)$ and $b_1^2(S)$ denote the transformation vertex $(\mathbf{x}_2, t_2, \Omega_2)$ and the stop vertex immediately following outcomes a_2 and b_1^2, respectively. The stop vertex is a dummy vertex present at the end of finite paths only and does not correspond to an operation.

An initial segment of a path p is the set of all vertices occurring before some vertex v in p. Thus (t) and $(t), A, a_2(\mathbf{x}_2, t_2, \Omega_2)$ are two of the four initial segments of the above path. The length $L(p)$ of a path or initial segment equals the number of vertices in the path or initial segment. Thus the above example of a path has length 5.

A tree is a finite path tree if each path has a finite length. There may or may not be an upper bound to the path lengths of a finite path tree. An initial subtree of a tree τ is a tree resulting from the removal of one or more terminal segments from one or more paths of τ. τ_n denotes the initial subtree of τ in which all paths have length n [provided $L(p) > n$ for each path p in τ].

Another useful property is that to each path there corresponds exactly one outcome sequence. Thus each path p of a tree τ is uniquely described by an outcome sequence ϕ_p, and sums over paths are equivalent to sums over outcome sequences. Other properties of trees are given elsewhere.[4,5]

Finally, one notes that the decision procedures considered here are restricted to be such that every initial segment of each path in the procedure has finite length. Thus paths containing finite or countably infinite sequences of operations are allowed, but paths containing countably infinite sequences of operations followed by one or more operations are excluded. Decision procedures satisfying this restriction will be referred to as being of order type $\leq \omega$.

For each path p of a decision procedure Q, let $l^p(1), l^p(2), \ldots, l^p(k) \cdots$ denote the step number in p corresponding to the first, second, ..., kth $\cdots$ measurements in p. (If p is finite or there are only a finite number of measurements in p, these sequences terminate with some value k^p of k.) Clearly, $l^p(k+1) > l^p(k)$ for $k = 1, 2 \cdots$. All other steps in p correspond to transformations (or the terminal stop if p is finite). With each initial segment of length n of a path p, denoted by p_n (= step 0, step 1, ..., step $n-1$ of p), there is uniquely associated an outcome sequence φ_{p_n} whose length is given by the largest value of k for which $l^p(k) + 1 < n$. If p is finite or there are only a finite number of measurements in p, then the outcome sequence φ_p associated uniquely with p has a length k^p.

Let S^Q denote the set of all finite outcome sequences associated with all finite paths and all initial segments of all paths of Q. That is, for each initial segment of each path in Q and for each finite path in Q, the corresponding outcome sequence is in S^Q, and each sequence in S^Q corresponds to either a common initial segment of some paths or a finite path in Q.

S^Q is generated as follows: Let S_Φ denote the set of all outcomes associated with the initial measurement [step $l(1) = l^p(1)$ for all paths p] of Q. For each finite outcome sequence ϕ_n (the subscript n gives the length of ϕ_n), which corresponds to a common initial segment of some paths in Q, let S_{ϕ_n} be the set of outcomes of the measurement which Q assigns to step $l^p(n+1)$ of any path p in Q with associated initial sequence ϕ_n. If p is finite or contains only n measurements and $\phi_n = \phi_p$, then S_{ϕ_n} is empty.

Define S_1 and S_{n+1} by

$$S_1 = S_\Phi, \tag{1}$$

$$S_{n+1} = \bigcup_{\phi_n \subset S_n} \phi_n * S_{\phi_n}, \tag{2}$$

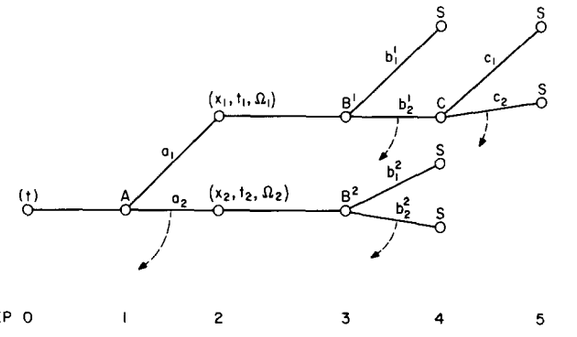


FIG. 1. The tree corresponding to the simple example of a decision procedure. The vertices correspond either to transformations $[(t), (\mathbf{x}_1, t_1, \Omega_1), (\mathbf{x}_2, t_2, \Omega_2)]$, measurements $[A, B^1, B^2, C]$ or the (dummy) stop operation $[S]$. The step numbers of the operations are given below the tree and the dashed arrows indicate other paths emanating from the preceding vertex.

where $n = 1, 2 \cdots$ and ${\phi_n}^* S_{\phi_n}$ denotes the set of sequences ϕ_{n+1} such that $\phi_{n+1}(j) = \phi_n(j)$ for $j = 0, 1, \ldots, n-1$ and $\phi_{n+1}(n) \in S_{\phi_n}$. Since each measurement operation in Q is descrete, each S_{ϕ_n} is at most countably infinite. Finally define S^Q by (ω is the first limit ordinal)

$$S^Q = \bigcup_{n=1}^{<\omega} S_n. \tag{3}$$

The basic correspondence assumption is that, for each decision procedure Q considered here, there is a Hilbert space $\mathcal{H}$ such that each transformation and each measurement procedure in Q are uniquely representable by a respective isometry (it is painless to generalize from unitary operators, so why not!) and discrete bounded self-adjoint operator on $\mathcal{H}$.

Under this assumption, it is clear that to each Q considered here there corresponds a unique tree of self-adjoint observables and isometries where for each step m of each path p, if $m = l^p(k)$ for some k, then the operator associated with vertex m of path p in the tree is self-adjoint, discrete, and given by

$$A^Q_{p_m} = \sum_{x \in S_{\phi_{k-1}}} x\, P^{Q,p}_{x\,m}, \tag{4}$$

where ϕ_k is the outcome sequence associated with the first m steps of p, and S_{ϕ_k} is the spectrum of $A^Q_{p_m}$. If $m \neq l^p(k)$ for all $k (\leq k^p$ if p is finite or contains only a finite number of measurements), then vertex m of p has the isometry $V^Q_{p_m}$ associated with it. If vertex m is the terminal stop vertex of a finite path, it has no operator associated with it. As before p_m denotes the first m steps of p and is used as an index to show that the choice of any operation in Q can depend on the steps and outcomes (if any) occurring up to but not including the step under consideration.

Also, under the correspondence assumption, S_{ϕ_k} becomes a subset of the set of real numbers, and S^Q is a subset of $\cup_{n=1}^{<\omega} R^n$, the set of all finite sequences of real numbers.

B. Probability Operator Measures

It is appropriate at this point to define and give some properties of probability operator measures. Other properties are given elsewhere.[1,6,7]

Let Σ be a σ-ring of subsets of some set Ω and $B(\mathcal{H})$ the set of all bounded linear operators on a Hilbert space $\mathcal{H}$. A mapping O: $\Sigma \to B(\mathcal{H})$ is an operator valued measure if

$$O(\Phi) = \mathbf{0}, \tag{5}$$

where Φ is the empty set and $\mathbf{0}$ the zero operator in $B(\mathcal{H})$ and where O is strongly countably additive. That is, if $E_1, E_2, \cdots$ is an infinite sequence of pairwise disjoint sets in Σ with $E = \cup_j E_j$, then

$$O(E) = \sum_j O(E_j), \tag{6}$$

where the implied limit is the strong limit.

An operator valued measure O is called self-adjoint (positive) if, for each E in Σ, $O(E)$ is a self-adjoint (positive) operator in $B(\mathcal{H})$. If O is positive, Σ is a σ-field, and

$$O(\Omega) = \mathbf{1}, \tag{7}$$

where $\mathbf{1}$ is the identity operator, then O is a probability operator measure. If O is a probability operator measure, then for each state ρ, the function $P_\rho: \Sigma \to R$ defined for each E in Σ by

$$P_\rho(E) = \mathrm{Tr}(\rho O(E)) \tag{8}$$

is a scalar probability measure.

An operator valued measure O has the following easy properties: O is finitely additive. If $E \subset F$, then

$$O(F) = O(E) + O(F - E). \tag{9}$$

If further O is positive, then

$$O(F) \geq O(E). \tag{10}$$

If O is finitely additive, positive, and strongly continuous from above at Φ or from below, then O is strongly countably additive. Weak and strong countable additivity are equivalent to one another but not to uniform countable additivity.[7] Note that the operators in the range set of a probability operator measure O do not have either to be projection operators or to commute with one another.

C. A Restricted Class of Q's

The association of probability operator measures to decision procedures will be carried out by first giving a direct construction and proof for a restricted class. By means of a natural correspondence between the restricted class and the class of all decision procedures (subject only to the restrictions of the correspondence assumption and that all paths must be of order type $\leq \omega$), a probability operator measure can be associated to any decision procedure in the larger class.

The restricted class of decision procedures considered here consists of only those for which (1) each path of the corresponding tree is infinite (and of order type $\leq \omega$), and (2) each path contains an infinite number of measurements or observations. The number of transformations in each path clearly is at most, countably infinite.

The reason for proceeding in this manner is primarily mathematical. It is easy to show that, to each decision procedure Q, all of whose paths are finite and of bounded length, a probability operator measure O^Q can be uniquely assigned. Also, as will be seen, with each procedure containing infinite paths only, a probability operator measure can be uniquely associated. However, there appear to be mathematical difficulties in carrying out the association directly for procedures whose paths either are partly finite and partly infinite or are all finite; but there is no upper bound on the number of measurements which can occur in a finite path. It is suspected that these difficulties, which consist of proving that Eq. (7) holds for the appropriate construction, are technical and will be overcome in the future.

The association of probability operator measures with decision procedures Q of the restricted type proceeds essentially as in I. For each path p in the tree τ_Q and each $n > 0$, the operator $\beta^Q_{p_n}$ is defined from Eq. (4) by

$$\beta^{Q}_{p_n} = V^{Q}_{p_{n-2}} \cdots V^{Q}_{p_{l(k)+1}} P^{Q,p_{l(k)}}_{\phi_p(k-1)} V^{Q}_{p_{l(k)}} \cdots \times P^{Q,p_{l(1)}}_{\phi_p(0)} V^{Q}_{p_{l(1)-1}} \cdots V^{Q}_{p_1} V^{Q}_{\Phi}, \quad (11)$$

where k is the largest integer such that $l(k) + 1 < n$. Note that in writing Eq. (11) for any initial path segment p_n, we do not regard step $n-1$ as having been carried out. If $n = l(k) + 2$, then the far left-hand sequence of isometries is absent from Eq. (11), and if $l(1) = 0$, the far right-hand sequence is absent. Von Neumann's projection axiom[2] has been used to give the form of the right-hand side of Eq. (11) as a product of the operators taken in the same order as the operations to which they correspond and, appear in path p of τ_Q.

Consider a sequence $\phi_n \in S_n$. As noted before this defines uniquely a set G of paths having a common initial segment of length $l(n+1)+1$ ($k = 1, 2 \cdots$ and length = 1+ step number). However, ϕ_n gives the outcomes of the measurements done at steps $l(1) \cdots l(n)$ only. [$l(j)$ is independent of any path p in the set so the index p is suppressed.] It does not give the outcome of the $(n+1)$th measurement but only defines, in Q, what observable is to be measured. Thus one can use Eq. (11) to define $\beta^{Q}_{\phi_n}$ by

$$\beta^{Q}_{\phi_n} = \beta^{Q}_{p_{l(n)+2}} = P^{Q,p_{l(n)}}_{\phi_n(n-1)} V^{Q}_{p_{l(n)-1}} \cdots V^{Q}_{p_{l(n-1)+1}} P^{Q,p_{l(n-1)}}_{\phi_n(n-2)} \cdots \times P^{Q,p_{l(1)}}_{\phi_n(0)} V^{Q}_{p_{l(1)-1}} \cdots V^{Q}_{p_1} V^{Q}_{\Phi}, \quad (12)$$

where p is any path in G. If desired the path index can be entirely suppressed on the right-hand side as ϕ_n determines uniquely, through Q, the operators and their ordering.

Define $O^{Q_n}(\{\phi_n\})$ by

$$O^{Q_n}(\{\phi_n\}) = \beta^{Q\dagger}_{\phi_n} \beta^{Q}_{\phi_n} \quad (13)$$

and for each subset E of S_n define $O^{Q_n}(E)$ by

$$O^{Q_n}(E) = \sum_{\phi_n \in E} O^{Q_n}(\{\phi_n\}), \quad (14)$$

where the convergence is in the strong operator topology. Finally, set $O^{Q_n}(\Phi) = \mathbf{0}$, the zero operator.

The proof that O^{Q_n} is a probability operator measure on S_n follows that given in I. For each ϕ_n, $O^{Q_n}(\{\phi_n\})$ is a positive operator and thus so is $O^{Q_n}(E)$ for all E for which $O^{Q_n}(E)$ exists. To show that $O^{Q_n}(E)$ exists for each E one first proves that Eq. (7) holds. One has

$$O^{Q_n}(S_n) = \sum_{\phi_n \in S_n} \beta^{Q\dagger}_{\phi_n} \beta^{Q}_{\phi_n} = \sum_{\phi_1 \in S_1} \sum_{\substack{\phi_2 \in S_2 \\ \phi_{2,1} = \phi_1}} \cdots \times \sum_{\substack{\phi_n \in S_n \\ \phi_{n,n-1} = \phi_{n-1}}} \beta^{Q\dagger}_{\phi_n} \beta^{Q}_{\phi_n}, \quad (15)$$

where $\phi_{n,j}$ denotes the first j elements of ϕ_n, i.e., $\phi_{n,j} = \phi_n(0) \cdots \phi_n(j-1)$. Here the second equality follows from the fact that the sequence of sums is just a rearrangement of the sum over all outcome sequences in S_n. A sum restricted to all $\phi_j \in S_j$ with $\phi_{j,j-1} = \phi_{j-1}$ is a sum over all sequences in S_j for which the first $j-1$ elements of ϕ_j are fixed and given by ϕ_{j-1}, and only the last element, $\phi_j(j-1)$, is free to vary.

Carrying out the far right-hand sum and using Eq. (12) along with the facts that the $V^{Q}_{p_{l(n)-1}} \cdots V^{Q}_{p_{l(n-1)}}$ are all isometries, $P^2 = P$, and the eigenprojectors of $A^{Q}_{p_{l(n)}}$ span all of $\mathcal{K}$ gives one the result that

$$\sum_{\substack{\phi_n \in S_n \\ \phi_{n,n-1} = \phi_{n-1}}} \beta^{Q\dagger}_{\phi_n} \beta^{Q}_{\phi_n} = \beta^{Q\dagger}_{\phi_{n-1}} \beta^{Q}_{\phi_{n-1}}.$$

But clearly this yields $O^{Q_n}(S_n) = O^{Q_{n-1}}(S_{n-1})$. Again carrying out the right-hand sum and repeating this over and over gives

$$O^{Q_n}(S_n) = O^{Q_{n-1}}(S_{n-1}) = \cdots = O^{Q_2}(S_2) = O^{Q_1}(S_1) = \sum_{\phi_1 \in S_1} V^{Q\dagger}_{\Phi} \cdots V^{Q\dagger}_{p_{l(1)-1}} P^{Q,p_{l(1)}}_{\phi_1(0)} V^{Q}_{p_{l(1)-1}} \cdots V^{Q}_{\Phi} = \mathbf{1}. \quad (16)$$

The existence of $O^{Q_n}(E)$ for each subset E of S_n follows from this as $O^{Q_n}(E)$ is the limit of a nondecreasing sequence of finite partial sums bounded from above by $\mathbf{1}$.[6] In this case, strong and weak convergence are equivalent.[6]

To see that O^{Q_n} is strongly countably additive [Eq. (6)], let $\{E_m | m = 1, \cdots\}$ be a sequence of pairwise disjoint subsets of S_n with $E = \cup_m E_m$. Then[1]

$$O^{Q_n}(E) = \sum_{\phi_n \in E} O^{Q_n}(\{\phi_n\}) = \sum_m \sum_{\phi_n \in E_m} O^{Q_n}(\{\phi_n\}) = \sum_m O^{Q_n}(E_m),$$

where the middle equality follows ultimately from the fact that if an infinite sum of nonnegative numbers exist, the value is independent of how the numbers are arranged in the sum. [This is also used to write the second equality of Eq. (15).]

Thus, for each n, Eqs. (5)-(7) are satisfied, and one has that Eqs. (12)–(14) define a unique probability operator measure $O^{Q_n}: \Sigma^n \to B(\mathcal{K})$. Σ^n is the set of all subsets of S_n. Furthermore, by construction, O^{Q_n} is the unique probability operator measure assigned to the initial subtree τ_{Q_n} of τ_Q, where τ_{Q_n} is obtained from τ_Q by cutting each path p between steps $l^p(n)$ and $l^p(n)+1$, adding terminal stop vertices, and discarding the infinite terminal segments. τ_{Q_n} corresponds to an "initial subdecision procedure" of Q which contains the first $l^p(n)+1$ steps of each path of Q. The uniqueness follows from the fact that, by construction, the assignment of O^{Q_n} to τ_{Q_n} is unique and the assignment of isometries and self-adjoint operators to the transformation and measurement procedures of Q is unique under the correspondence assumption.

The assignment of a probability operator measure O^{Q} to the whole procedure Q proceeds as follows: First consider the sequence $\{O^{Q_n} | n = 1, 2 \ldots\}$ of probability operator measures defined on the sequence $\{\Sigma^n | n = 1, 2 \cdots\}$ of σ-fields where Σ^n is the set of all subsets of S_n. We show that this sequence of operator measures is consistent. To this end, let $m > n$ and E be a subset of S_m [Eq. (2)], such that for some $F \subset S_n$

$$E = \bigcup_{\phi_n \in F} \bigcup_{\substack{\phi_m \in S_m \\ \phi_{m,n} = \phi_n}} \{\phi_m\}. \quad (17)$$

This is the appropriate generalization for decision procedures of the definition of cylinder sets which are appropriate to passive procedures and for which S_m reduces to the Cartesian product of m sets.

To see that the O^{Q_n} are consistent, one expands the definition of E in Eq. (17) and repeats the inductive procedures used to obtain Eq. (16). This gives

$$\begin{aligned} O^{Q_n}(E) &= \sum_{\phi_n \in F} \sum_{\substack{\phi_m \in S_m \\ \phi_{m,n} = \phi_n}} O^{Q_m}(\{\phi_m\}) \\ &= \sum_{\phi_n \in F} \sum_{\substack{\phi_{n+1} \in S_{n+1} \\ \phi_{n+1,n} = \phi_n}} \cdots \sum_{\substack{\phi_m \in S_m \\ \phi_{m,m-1} = \phi_{m-1}}} \beta_{\phi_m}^{Q\dagger} \beta_{\phi_m}^{Q} \\ &= \sum_{\phi_n \in F} \beta_{\phi_n}^{Q\dagger} \beta_{\phi_n}^{Q} = O^{Q_n}(F), \end{aligned} \tag{18}$$

which is the desired result.

Now let R^ω and $\mathfrak{B}(R^\omega)$ be the respective sets of all infinite sequences of real numbers and all Borel subsets of R^ω. Let $\mathfrak{F}$ be the field of all Borel cylinder subsets of R^ω. $\mathfrak{B}(R^\omega)$ is the minimal σ-field over $\mathfrak{F}$. For each n, define Ω_n by

$$\Omega_n = S_n \times R \times R \times \cdots \tag{19}$$

and Ω^Q by

$$\Omega^Q = \bigcap_n \Omega_n. \tag{20}$$

Ω^Q is the set of all possible real number outcome sequences of Q. Clearly, $\Omega^Q \in \mathfrak{B}(R^\omega)$ as $S_n \in \mathfrak{B}(R^n)$ and thus $\Omega_n \in \mathfrak{F}$ and $\mathfrak{B}(R^\omega)$ is closed under countable intersections.

For each n, define $O'^{Q}_n : \mathfrak{B}(R^n) \to B(\mathfrak{H})$ by

$$O'^{Q}_n(B) = O^{Q_n}(B \cap S_n) \tag{21}$$

for each $B \in \mathfrak{B}(R^n)$. O'^{Q}_n is well defined on $\mathfrak{B}(R^n)$ since if $B = B' \bmod S_n$, then $B \cap S_n = B' \cap S_n$ and $O'^{Q}_n(B) = O'^{Q}_n(B')$. Also $O'^{Q}_n(R^n) = 1$, and the countable additivity of O'^{Q}_n follows from that of O^{Q_n}. Finally let $m > n$ and $E = F \times R^{m-n}$ with $F \in \mathfrak{B}(R^n)$. Then

$$E \cap S_m = \bigcup_{\phi_n \in F \cap S_n} \bigcup_{\substack{\phi_m \in S_m \\ \phi_{m,n} = \phi_n}} \{\phi_m\}$$

and by Eqs. (7), (8)

$$O'^{Q}_m(E) = O^{Q_m}(E \cap S_m) = O^{Q_n}(F \cap S_n) = O'^{Q}_n(F).$$

Thus the sequence $\{O'^{Q}_n \mid n = 1, 2, \cdots\}$ is a consistent sequence of probability operator measures defined on $\{\mathfrak{B}(R^n) \mid n = 1, 2, \cdots\}$, where each O'^{Q}_n is the extension of O^{Q_n} onto $\mathfrak{B}(R^n)$, and O^{Q_n} is the restriction of O'^{Q}_n to $\Sigma^n \subset \mathfrak{B}(R^n)$.

By the operator valued equivalent[1] of the Kolmogorov extension theorem.[8] there exists a unique probability operator measure $O^{Q'}$ on $\mathfrak{B}(R^\omega)$ such that for each $E \in \mathfrak{F}$ with base $F \in \mathfrak{B}(R^n)$,

$$O^{Q'}(E) = O'^{Q}_n(F). \tag{22}$$

Define Σ^Q and $\mathfrak{F}^Q$ to be the respective σ-ring and ring of all subsets of R^ω of the form $B \cap \Omega^Q$ with $B \in \mathfrak{B}(R^\omega)$ and $B \in \mathfrak{F}$, respectively. That is, $\mathfrak{F}^Q$ is the set of all subsets E of Ω^Q such that for some n and base $F \subset S_n$

$$E = (F \times R \times R \times \cdots) \cap \Omega^Q \tag{23}$$

and Σ^Q is the minimal σ-ring over $\mathfrak{F}^Q$. Since $\Omega^Q \in \mathfrak{B}(R^\omega)$, Σ^Q is a sub σ-ring of $\mathfrak{B}(R^\omega)$.

Define O^Q to be the restriction of $O^{Q'}$ to Σ^Q. Since Σ^Q is a sub σ-ring of $\mathfrak{B}(R^\omega)$,

$$O^Q(E) = O^{Q'}(E) \tag{24}$$

for each $E \in \Sigma^Q$. Since $O^{Q'}$ is continuous from above[1] on $\mathfrak{B}(R^\omega)$ and $O^{Q'}(\Omega_n) = \mathbf{1}$ for each n, $O^Q(\Omega^Q) = O^{\phi'}(\Omega^\phi) = \lim_n O^{Q'}(\Omega_n) = \mathbf{1}$ and thus O^Q is a probability operator measure on Σ^Q. Also, for each $E \in \mathfrak{F}^Q$ with base $F \subset S_n$ for some n, Eqs. (21)–(24) give

$$O^Q(E) = O^{Q_n}(F). \tag{25}$$

Thus we have shown that with each infinite decision procedure Q of the type considered here there is associated a unique probability operator measure O^Q which satisfies Eq. (25) on the field $\mathfrak{F}^Q$ of subsets of Ω^ϕ.

Furthermore O^Q contains all the statistical properties of the decision procedure Q. That is, for each $E \in \mathfrak{F}^Q$ with base $F \subset S_n$ and for each state ρ, Eq. (25) gives that

$$\mathrm{Tr}[\rho O^Q(E)] = \mathrm{Tr}[\rho O^{Q_n}(F)] = \sum_{\phi_n \in F} \mathrm{Tr}(\rho \beta_{\phi_n}^{Q\dagger} \beta_{\phi_n}^{Q})$$

is the probability that carrying out steps of Q, until n measurements have been done (this is equivalent to carrying out Q_n), on a system in state ρ yields an outcome sequence in F. More generally for each $E \in \Sigma^Q$, $\mathrm{Tr}[\rho O^Q(E)]$ is the probability that carrying out Q on a system in state ρ yields an (infinite) outcome sequence in E.

D. Extension to All Q's

As stated, this result applied only to decision procedures in which each path is infinite and contains an infinite number of measurements. However, by making inessential changes, any decision procedure which contains one or more finite paths can be put in the form considered here. Thus this result holds for all decision procedures (which are made up of discrete observables and which satisfy Von Neumann's projection axiom and all of whose paths have order type $\leq \omega$).

To see this in more detail, let Q' be a decision procedure with one or more finite paths and let Q be the procedure obtained from Q' by adjoining to the end of each finite path of Q' an infinite sequence of measurements of the identity observable. For each n, let S_n^Q and $S_n^{Q'}$ be defined by Eq. (2) for Q and Q', respectively.

Note that if a particular ϕ_n in $S^{Q'}$ equals ϕ_p (the outcome sequence associated with path p) for some finite path which contains n measurements, then $S_{\phi_n}^{Q'}$ is empty. So for any $k > n$, there is no sequence in $S_k^{Q'}$ which has ϕ_n as an initial segment. From this, the construction of Q from Q', and Eq. (2), it is clear that, for each n, $S_n^{Q'} \subseteq S_n^Q$.

Under the correspondence assumption, S_n^Q and $S_n^{Q'}$ are subsets of R^n and are elements of $\mathfrak{B}(R^n)$. Thus one defines Ω_n^Q and $\Omega_n^{Q'}$ as before by Eq. (19) and Ω^Q and $\Omega^{Q'}$ by Eq. (20). Both Ω^Q and $\Omega^{Q'}$ are in $\mathfrak{B}(R^\omega)$. Let $\mathfrak{F}^Q$ and Σ^Q be defined as before from Eq. (23).

From the above it is clear that $\Omega^{Q'}$ contains the outcome sequences corresponding to all and only the infinite paths of Q'. Thus if ϕ_n is the outcome sequence corresponding to some finite path of Q' which contains n measurements, there is no sequence in $\Omega^{Q'}$ which contains ϕ_n as an initial segment. By the construction of Q, this is clearly not true for Ω^Q and one has that $\Omega^{Q'} \subset \Omega^Q$, where $\Omega^Q - \Omega^{Q'}$ contains all and only those outcome sequences which correspond to the extensions of the finite paths of Q'.

The σ-field $\Sigma^{Q'}$ of events for the procedure Q' is defined as follows. Let $B^{Q'}$ denote the σ-field of subsets of R^ω of the form $B \cap \Omega^{Q'}$ for some B in $\mathfrak{B}(R^\omega)$. Clearly $B^{Q'} \subset \Sigma^Q$. Let $Z^{Q'} = \Omega^{Q'} \cup F^{Q'}$ be the set of all possible outcome sequences of the process Q', where $F^{Q'}$ is the set of all outcome sequences associated with the finite paths of Q'. Then $\Sigma^{Q'}$ is defined to be the set of all sets of the form $E_1 \cup E_2$, where $E_1 \in B^{Q'}$ and $E_2 \subseteq F^{Q'}$. $\Sigma^{Q'}$ is clearly a σ-field because: (1) $Z^{Q'} \in \Sigma^{Q'}$; (2) if $E_1 \in B^{Q'}$ and $E_2 \subseteq F^{Q'}$, then the complement of $E_1 \cup E_2$ is (note that $E_1 \cap E_2 = \Phi$) $(\Omega^{Q'} - E_1) \cup (F^{Q'} - E_2)$ which is clearly in $\Sigma^{Q'}$; (3) let $\{E_{1j} \cup E_{2j} | j = 1, 2, \cdots\}$ be an infinite sequence of sets in $\Sigma^{Q'}$ with $E_{1j} \in B^{Q'}$ and $E_{2j} \subset F^{Q'}$ for each j. Then $\cup_j(E_{1j} \cup E_{2j}) = (\cup_j E_{1j}) \cup (\cup_k E_{2k})$, which is clearly in $\Sigma^{Q'}$.

Let m be a mapping from $Z^{Q'}$ to Ω^Q such that, for each $\phi \in \Omega^{Q'}$, $m(\phi) = \phi$ and, for each $\phi \in F^{Q'}$, $m(\phi)$ is the sequence obtained by adding an infinite sequence of 1's to the end of ϕ. By construction, m is a one-to-one mapping onto Ω^Q.

The association of a unique probability operator measure to the decision procedure Q' proceeds as follows: By the result of this section, under the correspondence assumption there is a unique probability operator measure O^Q associated with Q which satisfies Eqs. (13), (14), and (25) for Q. Define a mapping $O^{Q'}$: $\Sigma^{Q'} \to B(\mathcal{K})$ as follows: For each E in $\Sigma^{Q'}$ set

$$O^{Q'}(E) = O^Q(E_1) + O^Q(mE_2), \tag{26}$$

where $E = E_1 \cup E_2$ with $E \in B^{Q'}$ and $E_2 \subset F^{Q'}$ and $mE_2 = [m(\phi) | \phi \in E_2] \subset \Sigma^Q$.

Since O^Q is a probability operator measure, it is clear that $O^{Q'}$ is also, and we have thus shown that $O^{Q'}$, given by Eq. (26), is a probability operator measure associated with Q'. It remains to be shown that it satisfies Eqs. (13), (14), and (25) for Q'.

First, let E be any set of the form $F \times R \times R \times \cdots \cap \Omega^{Q'}$ with $F \subset S_n^{Q'}$ for some n. Clearly E is a subset of $\Omega^{Q'}$ and also $F \subset S_n^Q$ as $S_n^{Q'} \subset S_n^Q$. From Eqs. (13), (14), and (25) we have

$$O^Q(E) = O^Q{}_n(F) = \sum_{\phi_n \in F} \beta_{\phi_n}^{Q\dagger} \beta_{\phi_n}^Q.$$

By the construction of Q from Q', $\beta_{\phi_n}^{Q'} = \beta_{\phi_n}^Q$ for any such ϕ_n as the relevant parts of Q and Q' are identical. Thus one has from Eq. (26) ($E_2 = \Phi$ and $E_1 = E$)

$$O^{Q'}(E) = \sum_{\phi_n \in F} \beta_{\phi_n}^{Q'\dagger} \beta_{\phi_n}^{Q'}. \tag{27}$$

Next let E be any subset of $F^{Q'}$ ($E_1 = \Phi$ and $E = E_2$). Since $F^{Q'}$ is at most countably infinite and for each $E_2 \subset F^{Q'}$, $O^Q(mE_2)$ exists, one has

$$O^Q(mE_2) = \sum_{\phi \in mE_2} O^Q(\{\phi\}), \tag{28}$$

where the countable additivity of O^Q and the facts that m is one-to-one onto and Σ^Q contains the singleton sets have been used.

For each sequence ϕ in Ω^Q, let $E_{\phi,n}$ be the set of all sequences α in Ω^Q such that $\alpha(j) = \phi(j)$ for $j = 0, 1, \ldots, n-1$. One has that

$$O^Q(\{\phi\}) = \text{s-lim}_n O^Q(E_{\phi,n}) = \text{s-lim}_n \beta_{\phi_n}^{Q\dagger} \beta_{\phi_n}^Q. \tag{29}$$

The first equality follows from the facts that $\{E_{\phi,n} | n = 1, 2, \cdots\}$ is a nonincreasing sequence of sets with $\lim_n E_{\phi,n} = \{\phi\}$ and that O^Q is strongly continuous from above.[1] The second follows from Eqs. (13), (14), and (25) and the fact that $E_{\phi,n}$ has the form of Eq. (23) with $\Omega^{Q'}$ replacing Ω^Q and $F = \{\phi_n\}$.

By construction, each ϕ in mE_2 is the extension, by an infinite sequence of 1's, of a finite sequence θ in E_2 [with $m(\theta) = \phi$] which labels a finite path of Q' containing, say, l_θ measurements. Since the projection operator for eigenvalue 1 of the identity observable is the identity operator, Eq. (12) gives, for each $n > l_\theta$, $\beta_{\phi_n}^Q = \beta_\theta^{Q'}$. Use of this in Eq. (29) gives, along with Eqs. (26) and (28),

$$O^{Q'}(E_2) = \sum_{\theta \in E_2} \beta_\theta^{Q'\dagger} \beta_\theta^{Q'}, \tag{30}$$

which is the desired result.

So one sees from Eqs. (27) and (3) that $O^{Q'}$ does satisfy Eqs. (13), (14), and (25) for Q', and, by the theorem on the extension of positive operator measures from fields to σ-fields,[6] is unique. [The field $\mathfrak{F}^{Q'}$ is the set of all sets of the form $E_1 \cup E_2$ with $E_1 = (F \times R^\omega) \cap \Omega^{Q'}$ and $F \subset S_n^{Q'}$ and $E_2 \subset F^{Q'} \cap S_n^{Q'}$ for some n.]

Thus the result is obtained that under the correspondence assumption, the unique association of probability operator measures to decision procedures extends to those containing one or more finite paths. Obviously this includes procedures with finite paths only.

III. Discussion

One consequence of this association of probability operator measures to decision procedures is that the other results of I can be extended to include these procedures. Thus let Q be any decision procedure of the type discussed here and F any set in a certain subset (to be defined later) of Σ^Q. From the results of I there are two ways to measure the probability, that carrying out Q on a system in state ρ will yield an outcome sequence in F. One way (method 1) is to carry out Q on a system in state ρ an infinite number of times. This yields an infinite sequence α such that, for each $j = 0, 1, \ldots$, $\alpha(j) \in \Omega^Q$ is an outcome sequence for Q. One then generates an infinite 0-1 sequence $\theta_E\alpha$ from α by the prescription for each j, $(\theta_E\alpha)(j) = 1[0]$ if $\alpha(j)$ is in F [is not in F] and computes the limit mean $\overline{M}\theta_E\alpha$.

The other method (method 2) is to carry out an infinite sequence of repetitions of measurements of the observable whose self-adjoint operator is $O^Q(F)$ on a system in state ρ. (Here O^Q is the probability operator measure associated with Q by the methods of the last section). The result of this infinite sequence of measurements is a sequence $\beta \in R^\omega$.

It is clear that these two methods are equivalent if one can prove that

$$\overline{M}\theta_F\alpha = \mathrm{Tr}[\rho O^Q(F)] = \overline{M}\beta. \tag{31}$$

In brief, the proof using probability theoretic methods goes as follows: For the first method, one constructs the sample probability space $((\Omega^Q)^\omega, (\Sigma^Q)^\omega, \mathbf{P}_{\rho,Q})$, where $(\Omega^Q)^\omega$ is the set of all infinite sequences of elements of Ω^Q, $(\Sigma^Q)^\omega$ is the standard σ-field of subsets of $(\Omega^Q)^\omega$, and $\mathbf{P}_{\rho,Q}$ is a product probability measure on $(\Sigma^Q)^\omega$ defined by

$$\mathbf{P}_{\rho,Q}(E_{Fl}) = \mathrm{Tr}[\rho O^Q(F)] \tag{32}$$

for each l and F in Σ^Q, where $E_{Fl} = [\alpha \mid \alpha(l) \in F]$. By means of an ergodic theorem[9] and the indecomposability theorem[9] applied to the one-sided shift operator on $(\Omega^Q)^\omega$, one proves that

$$\overline{M}\theta_F(-) = \mathrm{Tr}[\rho O^Q(F)], \tag{33}$$

$\mathbf{P}_{\rho,Q}$ almost everywhere.

Similarly for method 2, one constructs the sample probability space $(R^\omega, \mathfrak{B}(R^\omega), \mathbf{P}_{\rho,Q,F})$, where $\mathbf{P}_{\rho,Q,F}$ is the product probability measure on $\mathfrak{B}(R^\omega)$ given by

$$\mathbf{P}_{\rho,Q,F}(E_{Bl}) = \int_B d[\mathrm{Tr}(\rho \mathfrak{E}^{O^Q(F)}_r)] \tag{34}$$

for each l and Borel subset B of R. Here $E_{Bl} = [\beta \mid \beta(l) \in B]$ and $\mathfrak{E}^{O^Q(F)}$ is the spectral measure for the operator $O^Q(F)$ and $\mathfrak{E}^{O^Q(F)}_r = \mathfrak{E}^{O^Q(F)}((-\infty, r])$ with $(-\infty, r]$ the set of all real numbers $\leq r$. Again by means of ergodic and indecomposability theorems[9] applied to the one-sided shift operator on R^ω, one proves that

$$\overline{M}(-) = \mathrm{Tr}[\rho O^Q(F)], \tag{35}$$

$\mathbf{P}_{\rho,Q,F}$ almost everywhere.

Now, as was stressed in I, in order to complete proofs of this type, one must assume that the scalar probability measure P assigned to a process is "correct" for the process. That is, all properties of outcome sequences which are τ-definable from P and which are true, P almost everywhere, must be true for the sequence obtained by actually carrying out the process. (This definition of correctness is discussed in detail elsewhere,[3] where it is the central part of a definition of agreement between a physical theory and experiment.)

A property of sequences is τ-definable from P if and only if it is expressible by some formula in $L(\tau)$ which contains a free sequence variable and may contain a name for P where $L(\tau)$ is the language of a formal theory τ. Also one requires τ to be strong enough to include formulas which relate limit means to expectation values. This brief description of "correct" measures is sufficient for our purposes.

In order to derive Eqs. (31) from Eqs. (33) and (35), one must assume that $\mathbf{P}_{\rho,Q}$ is "correct" for the infinite repetition of Q on ρ and that $\mathbf{P}_{\rho,Q,F}$ is "correct" for the infinite repetition of measuring $O^Q(F)$ on ρ. If F is τ-definable from the measure $\mathrm{Tr}[\rho O^Q(-)]$, which implies in this case that F is τ-definable from $\mathbf{P}_{\rho,Q}$ (this characterizes the "certain subset" of Σ^Q mentioned earlier), then by the definition of correctness Eq. (33) gives the left-hand equality of Eq. (31), and Eq. (35) gives the right-hand equality. Thus the two methods of measuring $\mathrm{Tr}[\rho O^Q(F)]$ are equivalent.

As we noted in I, a remarkable aspect of this equivalence is that the mathematical decision procedures used in method 1 to decide whether or not $\alpha(j) \in F$ for each j "disappear" into the physical operation of repeated measurements of $O^Q(F)$ on ρ in method 2. To see this in more detail, consider method 1. It consists of the physical operation of first repeatedly carrying out Q on ρ to give α. Then one generates $\theta_F\alpha$ by carrying out for each j, a mathematical decision procedure which gives the truth or falseness of $q[\alpha(j), \mathrm{Tr}[\rho O^Q(-)]]$, where $q[\alpha(j), \mathrm{Tr}[\rho O^Q(-)]]$ is a formula of τ which defines F. Finally one generates $\overline{M}\theta_E\alpha$ by means of a mathematical limit mean procedure. Method 2 consists of the physical operation of repeated measurements of $O^Q(F)$ on ρ to give a sequence β followed by the mathematical generation of $\overline{M}\beta$ by the equivalence $\overline{M}\beta = \overline{M}\theta_F\alpha$.

The point is that a mathematical procedure corresponding to the generation of $\theta_F\alpha$ in method 1 is missing in method 2. The physical operation of repeatedly measuring $O^Q(F)$ on ρ is equivalent to the physical operation of repeatedly carrying out Q on ρ followed by the mathematical operation of generating $\theta_F\alpha$, as the limit mean operation is common to both methods. It should be noted that other than having the same limit means, $\theta_F\alpha$ and β are quite different sequences: $\theta_F\alpha$ is an infinite sequence of 0's and 1's and β is an infinite sequence of real numbers all lying in the interval $[0, 1]$.

There is an interesting aspect of the fact that these results hold also for decision procedures. It is that if Q is a decision procedure, the physical operation in method 1 requires the intervention of an intelligent being or machine to make the decisions each time Q is carried out. No such decision making apparatus is required in method 2. There the physical operation of repeatedly carrying out the measurement of $O^Q(F)$ on ρ is a "passive" operation which requires no decisions to be made.

Thus one has the striking result that the infinite repetition of a decision procedure Q followed by an infinite sequence of mathematical decision procedures for F is equivalent to the passive process consisting of an infinite repetition of measurements of $O^Q(F)$ on ρ. Furthermore, this equivalence holds if Q is a decision procedure or if Q is a passive procedure which requires no decisions to be made.

One might argue that this is really not so striking because the procedure whereby one carries out a single measurement of $O^Q(F)$ is probably a decision procedure also; and, thus, the decisions in Method 2 are hidden, but still present. Although this is probably true, it is also just as true for each measurement procedure within Q. Furthermore, the argument that the decision procedure involved in carrying out a single measurement of $O^Q(F)$ is somehow of "higher order" than those needed for the procedures in Q would appear to be taken care of by the possibility that paths in Q can contain, as measurement steps, the measurement of $O^Q(F)$ [provided that $O^Q(F)$ is discrete].

Finally, it is to be briefly noted that many mathematical computation and decision procedures can

themselves be realized as decision procedures of the type considered here. Under the association of probability operator measures with these procedures, it becomes possible to talk about these realizations within quantum mechanics. In particular, this suggests the possibility that realizations of theorem proving procedures can be discussed within quantum mechanics. However, the consequences of this possibility and in particular its implications for the foundations of mathematics are left to future work.

* Work performed under the auspices of the U.S. Atomic Energy Commission.

1 P. A. Benioff, J. Math. Phys. **13**, 231 (1972).
2 J. Von Neuman, *Mathematical Foundations of Quantum Mechanics*, translated by R. T. Beyer (Princeton U.P., Princeton, N.J., 1955), pp. 335, 351.
3 P. A. Benioff, J. Math. Phys. **11**, 2553 (1970); **12**, 360 (1971).
4 C. H. Randall and D. J. Foulis, Amer. Math. Monthly **77**, 363 (1970).
5 H. Rogers, Jr., *Theory of Recursive Functions and Effective Computability* (McGraw-Hill, New York, 1967), Sec. 16. 3; T. J. Jech., J. Symbolic Logic **36**, 1 (1971).
6 S. K. Berberian, *Notes on Spectral Theory* (Van Nostrand Reinhold, New York, 1966).
7 I. Dobrakov, Czech. Math. J. **20**, 511 (1970).
8 A. N. Kolmogorov, *Foundations of Probability Theory*, 2nd English ed. (Chelsea, New York, 1956), pp. 29-33
9 M. Loève, *Probability Theory* (Van Nostrand Reinhold, New York, 1963), 3rd ed., Chap. IX.

Helicity and Canonical Spin Operators of the Poincaré Algebra

J. Beckers*
Service de Physique Mathématique, Université de Liège, Liège, Belgium
(Received 1 March 1971)

In the study of the irreducible unitary representations $[m, s]$ of the Poincaré group, we define an "helicity" spin operator through a fundamental connection between canonical and helicity bases. This operator, in covariant notation, is simply related to the well-known Bargmann–Wigner operator and to the canonical spin operator. Its spatial components generate an $SU(2)$-algebra and coincide with the elements of the Z-spin algebra recently proposed on different grounds by Braathen–Foldy. The very simple arguments developed here establish in a natural way the uniqueness of this algebra when helicity representations are studied.

1. INTRODUCTION

Since the fundamental paper of Wigner[1,2] on the irreducible unitary representations (IUR) of the Poincaré group (P), there have been many intimately related contributions which have appeared during the last fifteen years. Among all the different results referring to important fields,[3] let us simply mention that several *equivalent* bases (for the representation spaces associated with the IUR) have been proposed for the case of physical particles of nonzero rest mass and discrete spin (i.e., with timelike momenta). More precisely, we think of the canonical[1,2,4,5] and helicity[6] bases and of two more recent ones, the Ω- and Λ-bases.[7,8] So, the associated IUR have been called, respectively, the canonical-, helicity-, Ω-, and Λ-representations of P and their generators have been explicitly given.[1–8] Let us note that such bases and associated representations present different advantages and disadvantages, and therefore the choice of a particular one must be made according to the nature of the problem. On the one hand, the canonical basis which corresponds to the so-called "Foldy–Shirokov form" of the generators has the great advantage of giving the angular momentum operators as sums of two kinematically independent parts—the orbital and the spin angular momenta.[9] On the other hand, the helicity basis is particularly convenient for developing a general analysis[6] of binary reactions involving particles with arbitrary spin; it is also interesting in discussing[10] the so-called $[m, s]$–IUR with a view to connecting these representations with the $[0, s]$–ones through the limiting process $m \to 0$. For example, we can see how both aspects[6,10] have recently been applied to discussions[11] on vector meson dominance in photoproduction and the corresponding T matrix elements.

In this paper, we want to show through very simple arguments how, from a fundamental connection[10] between canonical and helicity bases, we can extract, *in a unique way*, privileged directions in spin space which lead to an interesting set of spin operators. In covariant notation, these operators can be written as the spatial components of an "helicity" spin 4-vector $S(p)$ which is directly related to the well-known Bargmann–Wigner[2] operator w and to the canonical spin operators. Furthermore, the new ones generate a structure isomorphic to the little group of p and coincide with the generators recently proposed on different grounds by Braathen and Foldy.[7] In this sense, the uniqueness of their developments is well established here. Let us also remark as a characteristic property of the so-used fundamental connection and the so-deduced spin algebra, that the Ω- and Λ-bases and representations are *both* "helicity" ones, particularly well adapted to the $[m, s] \to [0, s]$ correspondence. In fact, those representations can be easily found[12] through Coester's developments.[13]

In Sec. II, the main formulas[10] leading to the connection between canonical and helicity bases are collected, and privileged directions appear in a very simple manner when significant rotations in spin space are taken into account. In Sec. III, we explain briefly the meaning of the three effective little group generators and define the "canonical" spin 4-vector $S(p)$ in terms of the Bargmann–Wigner operator w. This is done in a slightly different but evidently equivalent way to well-known developments (for example, those of Hagedorn,[14] Moussa–Stora[15] or Gasiorowicz[16]). In Sec. IV, with the help of the "canonical" 4-vector and the results of Sec. II, we construct the "helicity" spin 4-vector $S(p)$. Finally, in Sec. V, we compare the present approach with that of Braathen–Foldy.[7]

In what follows, we essentially adopt Hagedorn's notations.[14] Therefore, since there is no risk of confusion, we do not distinguish here between the translation operators and their eigenvalues as is usually done. Furthermore, we shall confine ourselves to the nonzero rest mass case (the p^μ are always timelike, $p^0 > 0$) and to orthochronous proper homogeneous

themselves be realized as decision procedures of the type considered here. Under the association of probability operator measures with these procedures, it becomes possible to talk about these realizations within quantum mechanics. In particular, this suggests the possibility that realizations of theorem proving procedures can be discussed within quantum mechanics. However, the consequences of this possibility and in particular its implications for the foundations of mathematics are left to future work.

* Work performed under the auspices of the U.S. Atomic Energy Commission.

1 P. A. Benioff, J. Math. Phys. **13**, 231 (1972).
2 J. Von Neuman, *Mathematical Foundations of Quantum Mechanics*, translated by R. T. Beyer (Princeton U.P., Princeton, N.J., 1955), pp. 335, 351.
3 P. A. Benioff, J. Math. Phys. **11**, 2553 (1970); **12**, 360 (1971).
4 C. H. Randall and D. J. Foulis, Amer. Math. Monthly **77**, 363 (1970).
5 H. Rogers, Jr., *Theory of Recursive Functions and Effective Computability* (McGraw-Hill, New York, 1967), Sec. 16. 3; T. J. Jech., J. Symbolic Logic **36**, 1 (1971).
6 S. K. Berberian, *Notes on Spectral Theory* (Van Nostrand Reinhold, New York, 1966).
7 I. Dobrakov, Czech. Math. J. **20**, 511 (1970).
8 A. N. Kolmogorov, *Foundations of Probability Theory*, 2nd English ed. (Chelsea, New York, 1956), pp. 29-33
9 M. Loève, *Probability Theory* (Van Nostrand Reinhold, New York, 1963), 3rd ed., Chap. IX.

Helicity and Canonical Spin Operators of the Poincaré Algebra

J. Beckers*
Service de Physique Mathématique, Université de Liège, Liège, Belgium
(Received 1 March 1971)

In the study of the irreducible unitary representations $[m, s]$ of the Poincaré group, we define an "helicity" spin operator through a fundamental connection between canonical and helicity bases. This operator, in covariant notation, is simply related to the well-known Bargmann–Wigner operator and to the canonical spin operator. Its spatial components generate an $SU(2)$-algebra and coincide with the elements of the Z-spin algebra recently proposed on different grounds by Braathen–Foldy. The very simple arguments developed here establish in a natural way the uniqueness of this algebra when helicity representations are studied.

1. INTRODUCTION

Since the fundamental paper of Wigner[1,2] on the irreducible unitary representations (IUR) of the Poincaré group (P), there have been many intimately related contributions which have appeared during the last fifteen years. Among all the different results referring to important fields,[3] let us simply mention that several *equivalent* bases (for the representation spaces associated with the IUR) have been proposed for the case of physical particles of nonzero rest mass and discrete spin (i.e., with timelike momenta). More precisely, we think of the canonical[1,2,4,5] and helicity[6] bases and of two more recent ones, the Ω- and Λ-bases.[7,8] So, the associated IUR have been called, respectively, the canonical-, helicity-, Ω-, and Λ-representations of P and their generators have been explicitly given.[1–8] Let us note that such bases and associated representations present different advantages and disadvantages, and therefore the choice of a particular one must be made according to the nature of the problem. On the one hand, the canonical basis which corresponds to the so-called "Foldy–Shirokov form" of the generators has the great advantage of giving the angular momentum operators as sums of two kinematically independent parts—the orbital and the spin angular momenta.[9] On the other hand, the helicity basis is particularly convenient for developing a general analysis[6] of binary reactions involving particles with arbitrary spin; it is also interesting in discussing[10] the so-called $[m, s]$–IUR with a view to connecting these representations with the $[0, s]$–ones through the limiting process $m \to 0$. For example, we can see how both aspects[6,10] have recently been applied to discussions[11] on vector meson dominance in photoproduction and the corresponding T matrix elements.

In this paper, we want to show through very simple arguments how, from a fundamental connection[10] between canonical and helicity bases, we can extract, *in a unique way*, privileged directions in spin space which lead to an interesting set of spin operators. In covariant notation, these operators can be written as the spatial components of an "helicity" spin 4-vector $S(p)$ which is directly related to the well-known Bargmann–Wigner[2] operator w and to the canonical spin operators. Furthermore, the new ones generate a structure isomorphic to the little group of p and coincide with the generators recently proposed on different grounds by Braathen and Foldy.[7] In this sense, the uniqueness of their developments is well established here. Let us also remark as a characteristic property of the so-used fundamental connection and the so-deduced spin algebra, that the Ω- and Λ-bases and representations are *both* "helicity" ones, particularly well adapted to the $[m, s] \to [0, s]$ correspondence. In fact, those representations can be easily found[12] through Coester's developments.[13]

In Sec. II, the main formulas[10] leading to the connection between canonical and helicity bases are collected, and privileged directions appear in a very simple manner when significant rotations in spin space are taken into account. In Sec. III, we explain briefly the meaning of the three effective little group generators and define the "canonical" spin 4-vector $S(p)$ in terms of the Bargmann–Wigner operator w. This is done in a slightly different but evidently equivalent way to well-known developments (for example, those of Hagedorn,[14] Moussa–Stora[15] or Gasiorowicz[16]). In Sec. IV, with the help of the "canonical" 4-vector and the results of Sec. II, we construct the "helicity" spin 4-vector $S(p)$. Finally, in Sec. V, we compare the present approach with that of Braathen–Foldy.[7]

In what follows, we essentially adopt Hagedorn's notations.[14] Therefore, since there is no risk of confusion, we do not distinguish here between the translation operators and their eigenvalues as is usually done. Furthermore, we shall confine ourselves to the nonzero rest mass case (the p^μ are always timelike, $p^0 > 0$) and to orthochronous proper homogeneous

Lorentz transformations. The summation convention on repeated indices is always understood hereafter.

2. CONNECTION BETWEEN CANONICAL AND HELICITY BASES

In a recent discussion[10] of Coester's[13] and Chakrabarti's[17] transformations, we have singled out the particular interest of the Chakrabarti one in studying the transformation laws of state vectors in an helicity basis when the cases of nonzero as well as zero rest mass are considered. These developments gave us simple equivalences between different important contributions such as those of Foldy,[4] Shirokov,[5] Lomont–Moses,[18] and Fronsdal.[19,20] Here, let us simply recall we had rewritten the Chakrabarti transformation as a real rotation *in spin space* with Euler angles $(\phi, -\theta, -\phi)$:

$$U(\mathbf{p}) = R^{(s)}(\phi, -\theta, -\phi) = \exp(-i\phi S^3)\exp(i\theta S^2) \times \exp(i\phi S^3), \quad (1)$$

where (θ, ϕ) are the polar angles of $\mathbf{p}$ in an initial reference frame $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ and $\mathbf{S} \equiv (S^1, S^2, S^3)$ are the $(2s+1) \times (2s+1)$ matrices of arbitrary spin s such that

$$[S^i, S^j] = i\epsilon^{ijk}S^k, \quad i, j, k = 1, 2, 3, \quad (2)$$

the matrix S^3 being chosen diagonal. Furthermore, we had explicitly used[20] the rotation (1) when it was expressed in the form

$$U(\mathbf{p}) = \exp(i\phi \mathbf{S}\cdot\mathbf{n}), \quad (3)$$

$$\mathbf{n} \equiv \mathbf{e}_3 \wedge \mathbf{p}/|\mathbf{e}_3 \wedge \mathbf{p}|, \quad \phi = \arccos p^3/\mathbf{p}, \quad (4)$$

i.e., in a form which singles out the privileged direction $\mathbf{n}$.

So, if we remember, on the one hand, the essential role of $U(\mathbf{p}) \equiv (3)$ in the connection between the canonical basis $\{|[m,s];\mathbf{p},\sigma>\}$ and the helicity one $\{|[m,s];\mathbf{p},\lambda> = U(\mathbf{p})\,|[m,s];\mathbf{p},\sigma>\}$ and, on the other hand, the fundamental property

$$U(\mathbf{p})\Lambda\, U^{+}(\mathbf{p}) = U(\mathbf{p})[(\mathbf{S}\cdot\mathbf{p})/\mathbf{p}]U^{\dagger}(\mathbf{p}) = S^3, \quad (5)$$

we remark that the diagonalization of the helicity operator selects *and* encloses *two* privileged orthogonal directions: First, the particle displacement direction characterized by the unit vector $\mathbf{p}/|\mathbf{p}|$ and, secondly, the direction characterized by the unit vector $\mathbf{n} \equiv (4)$, orthogonal to $\mathbf{p}$ and to the third axis.[21] Thus, from the initial reference frame $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ used in typically "canonical" considerations, we get another one: $(\mathbf{n}_1 \equiv \mathbf{n}_2 \wedge \mathbf{n}_3, \mathbf{n}_2 \equiv \mathbf{p}/\mathbf{p}, \mathbf{n}_3 \equiv \mathbf{n})$ when "helicity" considerations became essential.[22] These two systems of unitary spatial vectors are then related through a simple three-dimensional rotation R:

$$\mathbf{n}_i = R\mathbf{e}_i, \quad i = 1, 2, 3,$$
$$R\tilde{R} = \tilde{R}R = 1, \quad \tilde{R} = R^{-1}, \quad \det R = 1, \quad (6)$$

with

$$R = \begin{pmatrix} -\cos\phi\cos\theta & \cos\phi\sin\theta & -\sin\phi \\ -\sin\phi\cos\theta & \sin\phi\sin\theta & \cos\phi \\ \sin\theta & \cos\theta & 0 \end{pmatrix}, \quad (7)$$

where (θ, ϕ) are the polar angles of $\mathbf{p}$ also.

3. THE "CANONICAL" SPIN OPERATOR

If we write the unitary operators $U(a, \Lambda)$ of P associated in the inhomogeneous Lorentz transformations (a, Λ) as usually[14]

$$U(a, \Lambda) = U(a)U(\Lambda), \quad U(a) = \exp(ip^\mu a_\mu),$$
$$U(\Lambda) = \exp(\tfrac{1}{2}i\alpha_{\mu\nu}J^{\mu\nu}), \quad \mu, \nu = 0, 1, 2, 3,$$

where the p^μ and $J^{\mu\nu}$ are the ten *Hermitian* generators associated, respectively, to translations (a) and homogeneous Lorentz transformations (Λ), we can resume the characteristic commutation relations of the Poincaré algebra (for infinitesimal transformations) in the form

$$[p^\mu, p^\nu] = 0, \quad [J^{\mu\nu}, p^\rho] = i(g^{\nu\rho}p^\mu - g^{\mu\rho}p^\nu),$$
$$[J^{\mu\nu}, J^{\rho\sigma}] = i(g^{\mu\sigma}J^{\nu\rho} + g^{\nu\rho}J^{\mu\sigma} - g^{\mu\rho}J^{\nu\sigma} - g^{\nu\sigma}J^{\mu\rho}), \quad (8)$$

with the metric tensor $g^{\mu\nu} = 0 (\mu \neq \nu)$, $g^{00} = -g^{ii} = 1$, $(i = 1, 2, 3)$.

Then, the well-known 4-vector of spin or Bargmann–Wigner[2] operator is defined by

$$w^\mu = \tfrac{1}{2}\epsilon^\mu{}_{\nu\rho\sigma}J^{\nu\rho}p^\sigma, \quad \mu = 0, 1, 2, 3, \quad \epsilon_{0123} = 1 \quad (9)$$

and satisfies

$$w^\mu p_\mu = 0, \quad (10a)$$
$$[w^\mu, p^\nu] = 0, \quad (10b)$$
$$[w^\mu, J^{\nu\rho}] = i(g^{\mu\nu}w^\rho - g^{\mu\rho}w^\nu), \quad (10c)$$
$$[w^\mu, w^\nu] = -i\epsilon^{\mu\nu}{}_{\rho\sigma}w^\rho p^\sigma. \quad (10d)$$

In particular, these relations show that, by (10a), w is only characterized by three linearly independent components and, by (10c), behaves like a 4-vector under Lorentz transformations. Furthermore, as we know that these three components generate the little group of p (the rotation group here), we must show they satisfy a structure isomorphic to Eq. (2) from the commutation relations (10c). Let us consider this important point now in a slightly different but equivalent way to other developments.[15,16]

If, as usual, the translation operators are taken diagonal with eigenvalue $p[p^2 = (p^0)^2 - (\mathbf{p})^2 = m^2]$, we can, after Moussa and Stora,[15] define $w(p)$ as the restriction of the operator w to the eigenspace of the $p^\mu (\mu = 0, 1, 2, 3)$ corresponding to the eigenvalue p. So, in the rest frame $[p_R \equiv (m, \mathbf{o})]$, we have the corresponding restriction $w(p_R)$, such that

$$w(p_R)\cdot p_R = 0,$$
$$w_i(p_R)\,|p_R, \cdots> = mS_i(p_R)\,|p_R, \cdots>.$$

It is straightforward to show that the three operators defined by

$$S_i(p_R) = w_i(p_R)/m, \quad i = 1, 2, 3, \quad (11)$$

verify the commutation relations (2), i.e., generate an $SU(2)$-structure. Now, let us extend Eq. (11) in covariant notation. For example, let us project $(1/m)w(p_R)$ on the following basis vectors:

$$m_0(p_R) \equiv \frac{p_R}{m} = (1, 0, 0, 0), \quad m_1(p_R) = (0, 1, 0, 0),$$
$$m_2(p_R) = (0, 0, 1, 0), \quad m_3(p_R) = (0, 0, 0, 1).$$

We have

$$S_\alpha(p_R) = (1/m)m^\mu_\alpha(p_R)w_\mu(p_R), \quad \alpha = 0, 1, 2, 3, \quad (12)$$

according to (11) and $S_0(p_R) \equiv 0$.

If we are interested in the restriction $w(p)$ with arbitrary p, we only consider the pure Lorentz transformation L_p, which takes the particle of mass m from rest to a state of momentum $\mathbf{p}$, i.e.,

$$\begin{aligned}
&(L_p)^0{}_0 = p^0/m, \quad p^0 = (|\mathbf{p}|^2 + m^2)^{1/2},\\
&(L_p)^i{}_0 = p^i/m, \quad (L_p)^0{}_i = -p_i/m,\\
&(L_p)^i{}_j = \delta^i{}_j - p^i p_j/m(p^0 + m),
\end{aligned} \quad (13)$$

such that

$$L_p p_R = p \equiv (p^0, \mathbf{p}). \quad (14)$$

Then, the operators $S_\alpha(p)$ are given by

$$S_\alpha(p) = (1/m)m^\mu_\alpha(p)w_\mu(p), \quad (15)$$

where

$$m^\mu_\alpha(p) = (L_p)^\mu{}_\nu m^\nu_\alpha(p) \quad (16)$$

or explicitly with $A = m(p^0 + m)$:

$$\begin{aligned}
m_0(p) &\equiv \frac{p}{m} = \left(\frac{p^0}{m}, \frac{p^1}{m}, \frac{p^2}{m}, \frac{p^3}{m}\right),\\
m_1(p) &= \left(\frac{p^1}{m}, 1 + \frac{(p^1)^2}{A}, \frac{p^1p^2}{A}, \frac{p^1p^3}{A}\right),\\
m_2(p) &= \left(\frac{p^2}{m}, \frac{p^1p^2}{A}, 1 + \frac{(p^2)^2}{A}, \frac{p^2p^3}{A}\right),\\
m_3(p) &= \left(\frac{p^3}{m}, \frac{p^1p^3}{A}, \frac{p^2p^3}{A}, 1 + \frac{(p^3)^2}{A}\right).
\end{aligned} \quad (17)$$

Let us note that these 4-vectors $m_\alpha(p)$ [as the $m_\alpha(p_R)$ ones], $\alpha = 0, 1, 2, 3$, verify the relations

$$\begin{aligned}
&m_\alpha(p)\cdot m_\beta(p) = g_{\alpha\beta},\\
&\det[m_0(p), m_1(p), m_2(p), m_3(p)] = 1.
\end{aligned} \quad (18)$$

In this way, we obtain, through (15) and (17), the only three fundamental spin operators [generating an $SU(2)$-algebra]:

$$S_i(p) = \frac{1}{m}\left(w_i - \frac{w_0 p_i}{p_0 + m}\right), \quad i = 1, 2, 3, \quad (19)$$

which, on the one hand, can be called "the spin operators relative to the m-basis" and, on the other hand, correspond to the definition of the well-known "canonical spin vector operators." Let us recall here that these operators have classical analogs, i.e., the Thomas vector.[23]

Following these results, we shall call up to now $S(p)$ the "canonical" spin 4-vector relative to p.

4. THE "HELICITY" SPIN OPERATOR

In connection with the remarks of Sec. II, let us now define a new spin operator by the projection of the "canonical" one given by (15) on the set of basis vectors characteristic of helicity developments. In fact, this set $\{n_\beta(p), \beta = 0, 1, 2, 3\}$ will be chosen as follows:

$$n_0(p) \equiv m_0(p_R) = (1, 0, 0, 0), \quad n_i(p) = (0, \mathbf{n}_i), \quad i = 1, 2, 3, \quad (20)$$

where the $\mathbf{n}_i$ $(i = 1, 2, 3)$ are given by (6) in such a way that the relations (18) still remain valid. Then, if we denote by $M(p)$ the Lorentz transformation which connects the n- and m-bases, we have

$$n_\beta(p) = M_\beta{}^\alpha(p)m_\alpha(p), \quad \beta = 0, 1, 2, 3, \quad (21)$$

where

$$M_\beta{}^\alpha(p) = \begin{pmatrix}
\frac{p^0}{m} & -\frac{p^1}{m} & -\frac{p^2}{m} & -\frac{p^3}{m}\\
0 & -\frac{p^1p^3}{B|\mathbf{p}|} & -\frac{p^2p^3}{B|\mathbf{p}|} & \frac{B}{|\mathbf{p}|}\\
-\frac{|\mathbf{p}|}{m} & \frac{p^0p^1}{m|\mathbf{p}|} & \frac{p^0p^2}{m|\mathbf{p}|} & \frac{p^0p^3}{m|\mathbf{p}|}\\
0 & -\frac{p^2}{B} & \frac{p^1}{B} & 0
\end{pmatrix} \quad (22)$$

with $B = [(p^1)^2 + (p^2)^2]^{1/2}$.

The new spin operators relative to the basis vectors $m_\beta(p)$ are directly obtained:

$$\begin{aligned}
\mathcal{S}_\beta(p) &= n^\alpha_\beta(p)S_\alpha(p) \quad (23)\\
&= (1/m)n^\alpha_\beta(p)m^\mu_\alpha(p)w_\mu(p)\\
&= (1/m)M_\beta{}^\mu(p)w_\mu(p), \quad \beta = 0, 1, 2, 3. \quad (24)
\end{aligned}$$

As expected, the last equation shows that $\mathcal{S}(p)$ could have been directly obtained by applying $M(p)$ to the restriction of the Bargmann-operator itself. Equations (22) and (24) give

$$\mathcal{S}_0(p) = 0 \quad (25)$$

and

$$\begin{aligned}
&\mathcal{S}_1(p) = (1/mB)\{|\mathbf{p}|\mathbf{w} - p^0[(\mathbf{J}\cdot\mathbf{p})/|\mathbf{p}|]\mathbf{p}\}\cdot\mathbf{e}_3,\\
&\mathcal{S}_2(p) \equiv \Lambda = \mathbf{J}\cdot\mathbf{p}/|\mathbf{p}|,\\
&\mathcal{S}_3(p) = (1/mB)\mathbf{w}\cdot(\mathbf{e}_3 \wedge \mathbf{p}).
\end{aligned} \quad (26)$$

So, on these simple grounds, we obtain what we want to call the "helicity" spin operator $\mathcal{S}(p)$ or "the spin operators relative to the n-basis". By using covariance arguments or performing the explicit calculations, we evidently obtain

$$[\mathcal{S}_i(p), \mathcal{S}_j(p)] = i\epsilon_{ijk}\mathcal{S}_k(p), \quad i, j, k = 1, 2, 3 \quad (27)$$

and maintain all the notions related to the little group structure of p.

5. CONNECTION WITH THE BRAATHEN–FOLDY ALGEBRA AND DISCUSSION

The structure (27) and the three $\mathcal{S}_i(p) \equiv (26)$ are identical to the Z spin algebra recently proposed by Braathen–Foldy[7] from a point of view which appears, in a sense, as the inverse of ours. Starting from the little group structure $\equiv (27)$, they *require* three explicit supplementary specifications on the $\mathcal{S}_i$, i.e.,

(1) one of these operators must be the helicity one: $\mathcal{S}_2 = \Lambda$,
(2) one of the remaining two commutes with $K_3(J^{03})$: $[\mathcal{S}_3, K_3] = 0$, (28)
(3) they all commute with $J_3(J^{12})$: $[\mathcal{S}_i, J_3] = 0$, $i = 1, 2, 3$.

These specifications are essentially introduced from physical and practical reasons. Furthermore, with the results (26) as a solution, they "suspect it is unique without being able to give a satisfactory proof."[7]

Now, from our point of view, starting with the characteristic triad (6) associated to the helicity character

of the considered representations, we find these operators *in a unique way* and the properties (28) directly follow. Then we can state that all the representations issued from the spin algebra [(26), (27)] will be typically *helicity* representations. So, as we know that a particular property of those representations is that they are well adapted for the transition to the case of zero rest mass, it is straightforward to show this property on the Ω- *and* Λ-representations.[7,8] Furthermore, the connection between the Ω-basis vectors and the Kotanski ones,[24] as shown by Braathen–Foldy, finds here a very natural explanation from the helicity character of Kotanski's developments.

Finally, let us note the already given[7] geometrical interpretation of the three $\mathcal{S}_i(p)$ $(i = 1, 2, 3)$ or as it follows from our Eq. (23),

$$\mathcal{S}_i(p) = \mathbf{n}_i(p)\cdot\mathbf{S}(p), \quad i = 1, 2, 3, \tag{29}$$

where $\mathbf{S}(p)$ is the canonical spin operator (19). More significantly here, we can also write in connection with the triad $(\mathbf{e}_i,\ i = 1, 2, 3)$:

$$\mathcal{S}_i(p) = \mathbf{e}_i\cdot\mathcal{S}(p) = (R\mathbf{e}_i)\cdot(R\mathcal{S}(p)),$$

such that with Eqs. (6) and (29), we have

$$\mathcal{S}(p) = R^{-1}\mathbf{S}(p), \tag{30}$$

when R is given by (7). This relation clearly shows the *specific* example constructed by Braathen and Foldy from the general little group structure relative to time like p.

ACKNOWLEDGMENT

It is a pleasure to thank Professor J. Serpe for his kind interest.

* Postal address: Institut de Physique, Sart-Tilman, B-4000 Liege 1, Belgium.
1 E. P. Wigner, Ann. Math. **40**, 149 (1939).
2 V. Bargmann and E. P. Wigner, Proc. Natl. Acad. Sci. (U.S.) **34**, 211 (1948).
3 Some aspects have been recently discussed by J. Beckers, in Phys. Rev. D **3**, 952 (1971); see also the references therein.
4 L. L. Foldy, Phys. Rev. **102**, 549 (1956).
5 Yu. M. Shirokov, Zh. Eksp. Teor. Fiz. **6**, 664, 919, 929 (1958).
6 M. Jacob and G. C. Wick, Ann. Phys. (N.Y.) **7**, 404 (1959).
7 H. J. Braathen and L. L. Foldy, Nucl. Phys. **B13**, 511 (1969).
8 H. J. Braathen, Nucl. Phys. **B15**, 477 (1970).
9 Recently, Mukunda has extended this discussion to the other IUR of P corresponding to lightlike and spacelike momenta: see N. Mukunda, Ann. Phys. (N.Y.) **61**, 329 (1970).
10 J. Beckers and C. Pirotte, Nuovo Cimento **A64**, 439 (1969).
11 G. Berlam and G. Eilam, Nucl. Phys. **B25**, 321 (1970).
12 J. Beckers and C. Pirotte, Lett. Nuovo Cimento, **2**, 147, (1971).
13 F. Coester, Phys. Rev. **129**, 2816 (1963).
14 R. Hagedorn, "Selected Topics on Scattering Theory, Part IV" (unpublished lectures given at the Max-Planck Institut für Physik, München, 1963).
15 P. Moussa and R. Stora, *Methods in Subnuclear Physics*, edited by N. Nikolic (Gordon and Breach, New York, 1968), Vol. II, p. 265.
16 S. Gasiorowicz, *Elementary Particle Physics* (Wiley, New York, 1966), Chap. IV.
17 A. Chakrabarti, J. Math. Phys. **7**, 949 (1966).
18 J. S. Lomont and Harry E. Moses, J. Math. Phys. **3**, 405 (1962).
19 C. Fronsdal, Phys. Rev. **113**, 1367 (1959).
20 See especially Ref. 10 for the interest of the helicity basis $(m \neq 0)$ in connection with Fronsdal's results $(m \neq 0, m = 0)$.
21 The third axis $(\mathbf{e}_3)$ plays an important role in the canonical basis since we usually refer to the third component of the spin as a member of the complete set of commuting observables.
22 These remarks could have been already obtained from the important paper of Jacob and Wick; see Ref. 6, Sec. II, when one particle states are discussed.
23 H. Bacry, *Leçons sur la théorie des groupes et les symétries des particules élémentaires* (Gordon and Breach, New York 1967); see especially Chap. 8.
24 A. Kotanski, Acta Phys. Polon. **29**, 699 (1966).

Renormalization of the Hilbert Space in the Mass Shift Model

Lon Rosen
Departments of Mathematics and Physics, Princeton University, Princeton, New Jersey 08540
(Received 4 January 1972)

The $(\phi^2)_{s+1}$ quantum field theory model can be solved exactly, where the number of space dimensions is $s \leq 3$. Thus this model explicitly illustrates many properties of nontrivial models such as $(\phi^{2n})_2$ and Yukawa$_2$. In particular we study hyperbolicity, the energy spectrum, local Fockness, and the change of Hilbert space as the spatial cut off is removed.

1. INTRODUCTION

In 1953, Friedrichs[1] derived an exact solution for the ϕ^2 field theory with a spatial cutoff describing scalar bosons. More recently, a number of authors,[2-6] some of them apparently unaware of Friedrichs' solution, have studied this model from a variety of viewpoints. The ϕ^2 Hamiltonian is

$$\begin{aligned} H(g) &= H_0 + \lambda H_I(g) \\ &= \tfrac{1}{2}\int :\pi(x)^2 + [\nabla\phi(x)]^2 + m_0^2\,\phi(x)^2 : dx \\ &\quad + \tfrac{1}{2}\lambda\int g(x) :\phi(x)^2 : dx . \end{aligned} \tag{1.1}$$

Here $g(x)$ is a smooth function of compact support, equal to 1 on a large set, and satisfying $0 \leq g(x) \leq 1$; $m_0 > 0$ is the bare mass; and $\lambda > -m_0^2$ is the coupling constant. The number of space dimension $s = 1, 2,$ or 3.

Because of its simplicity, this model has proved amenable to practically every known method and is consequently well understood. In the limit $g \to 1$, the theory approaches that of a free field with mass $m_1^2 = m_0^2 + \lambda$, as can be seen formally from the Hamiltonian (1.1). This limit necessitates a change or renormalization of the Hilbert space from the original Fock representation with mass m_0 to that of mass m_1. The scattering theory that results is of course trivial. When $g(x) \not\equiv 1$, nontrivial scattering does occur but this phenomenon may be regarded as spurious since essentially particles are bouncing off the "edges" of the cutoff function g. As a matter of fact, as we shall see, the structure of the cut-off theory is much like that of the quantum mechanical theory associated with the single particle operator

$$\mu_g^2 = -\Delta + m_0^2 + \lambda g(x). \tag{1.2}$$

The field equations for this model are linear.

of the considered representations, we find these operators *in a unique way* and the properties (28) directly follow. Then we can state that all the representations issued from the spin algebra [(26), (27)] will be typically *helicity* representations. So, as we know that a particular property of those representations is that they are well adapted for the transition to the case of zero rest mass, it is straightforward to show this property on the Ω- *and* Λ-representations.[7,8] Furthermore, the connection between the Ω-basis vectors and the Kotanski ones,[24] as shown by Braathen–Foldy, finds here a very natural explanation from the helicity character of Kotanski's developments.

Finally, let us note the already given[7] geometrical interpretation of the three $\mathcal{S}_i(p)$ $(i = 1, 2, 3)$ or as it follows from our Eq. (23),

$$\mathcal{S}_i(p) = \mathbf{n}_i(p)\cdot\mathbf{S}(p), \quad i = 1, 2, 3, \tag{29}$$

where $\mathbf{S}(p)$ is the canonical spin operator (19). More significantly here, we can also write in connection with the triad $(\mathbf{e}_i,\ i = 1, 2, 3)$:

$$\mathcal{S}_i(p) = \mathbf{e}_i\cdot\mathcal{S}(p) = (R\mathbf{e}_i)\cdot(R\mathcal{S}(p)),$$

such that with Eqs. (6) and (29), we have

$$\mathcal{S}(p) = R^{-1}\mathbf{S}(p), \tag{30}$$

when R is given by (7). This relation clearly shows the *specific* example constructed by Braathen and Foldy from the general little group structure relative to time like p.

ACKNOWLEDGMENT

It is a pleasure to thank Professor J. Serpe for his kind interest.

* Postal address: Institut de Physique, Sart-Tilman, B-4000 Liege 1, Belgium.
1 E. P. Wigner, Ann. Math. **40**, 149 (1939).
2 V. Bargmann and E. P. Wigner, Proc. Natl. Acad. Sci. (U.S.) **34**, 211 (1948).
3 Some aspects have been recently discussed by J. Beckers, in Phys. Rev. D **3**, 952 (1971); see also the references therein.
4 L. L. Foldy, Phys. Rev. **102**, 549 (1956).
5 Yu. M. Shirokov, Zh. Eksp. Teor. Fiz. **6**, 664, 919, 929 (1958).
6 M. Jacob and G. C. Wick, Ann. Phys. (N.Y.) **7**, 404 (1959).
7 H. J. Braathen and L. L. Foldy, Nucl. Phys. **B13**, 511 (1969).
8 H. J. Braathen, Nucl. Phys. **B15**, 477 (1970).
9 Recently, Mukunda has extended this discussion to the other IUR of P corresponding to lightlike and spacelike momenta: see N. Mukunda, Ann. Phys. (N.Y.) **61**, 329 (1970).
10 J. Beckers and C. Pirotte, Nuovo Cimento **A64**, 439 (1969).
11 G. Berlam and G. Eilam, Nucl. Phys. **B25**, 321 (1970).
12 J. Beckers and C. Pirotte, Lett. Nuovo Cimento, **2**, 147, (1971).
13 F. Coester, Phys. Rev. **129**, 2816 (1963).
14 R. Hagedorn, "Selected Topics on Scattering Theory, Part IV" (unpublished lectures given at the Max-Planck Institut für Physik, München, 1963).
15 P. Moussa and R. Stora, *Methods in Subnuclear Physics*, edited by N. Nikolic (Gordon and Breach, New York, 1968), Vol. II, p. 265.
16 S. Gasiorowicz, *Elementary Particle Physics* (Wiley, New York, 1966), Chap. IV.
17 A. Chakrabarti, J. Math. Phys. **7**, 949 (1966).
18 J. S. Lomont and Harry E. Moses, J. Math. Phys. **3**, 405 (1962).
19 C. Fronsdal, Phys. Rev. **113**, 1367 (1959).
20 See especially Ref. 10 for the interest of the helicity basis $(m \neq 0)$ in connection with Fronsdal's results $(m \neq 0, m = 0)$.
21 The third axis $(\mathbf{e}_3)$ plays an important role in the canonical basis since we usually refer to the third component of the spin as a member of the complete set of commuting observables.
22 These remarks could have been already obtained from the important paper of Jacob and Wick; see Ref. 6, Sec. II, when one particle states are discussed.
23 H. Bacry, *Leçons sur la théorie des groupes et les symétries des particules élémentaires* (Gordon and Breach, New York 1967); see especially Chap. 8.
24 A. Kotanski, Acta Phys. Polon. **29**, 699 (1966).

Renormalization of the Hilbert Space in the Mass Shift Model

Lon Rosen
Departments of Mathematics and Physics, Princeton University, Princeton, New Jersey 08540
(Received 4 January 1972)

The $(\phi^2)_{s+1}$ quantum field theory model can be solved exactly, where the number of space dimensions is $s \leq 3$. Thus this model explicitly illustrates many properties of nontrivial models such as $(\phi^{2n})_2$ and Yukawa$_2$. In particular we study hyperbolicity, the energy spectrum, local Fockness, and the change of Hilbert space as the spatial cut off is removed.

1. INTRODUCTION

In 1953, Friedrichs[1] derived an exact solution for the ϕ^2 field theory with a spatial cutoff describing scalar bosons. More recently, a number of authors,[2-6] some of them apparently unaware of Friedrichs' solution, have studied this model from a variety of viewpoints. The ϕ^2 Hamiltonian is

$$\begin{aligned} H(g) &= H_0 + \lambda H_I(g) \\ &= \tfrac{1}{2}\int :\pi(x)^2 + [\nabla\phi(x)]^2 + m_0^2\,\phi(x)^2 : dx \\ &\quad + \tfrac{1}{2}\lambda\int g(x) :\phi(x)^2 : dx . \end{aligned} \tag{1.1}$$

Here $g(x)$ is a smooth function of compact support, equal to 1 on a large set, and satisfying $0 \leq g(x) \leq 1$; $m_0 > 0$ is the bare mass; and $\lambda > -m_0^2$ is the coupling constant. The number of space dimension $s = 1, 2,$ or 3.

Because of its simplicity, this model has proved amenable to practically every known method and is consequently well understood. In the limit $g \to 1$, the theory approaches that of a free field with mass $m_1^2 = m_0^2 + \lambda$, as can be seen formally from the Hamiltonian (1.1). This limit necessitates a change or renormalization of the Hilbert space from the original Fock representation with mass m_0 to that of mass m_1. The scattering theory that results is of course trivial. When $g(x) \not\equiv 1$, nontrivial scattering does occur but this phenomenon may be regarded as spurious since essentially particles are bouncing off the "edges" of the cutoff function g. As a matter of fact, as we shall see, the structure of the cut-off theory is much like that of the quantum mechanical theory associated with the single particle operator

$$\mu_g^2 = -\Delta + m_0^2 + \lambda g(x). \tag{1.2}$$

The field equations for this model are linear.

The reason that the mass shift model can be solved exactly is that the Hamiltonian $H(g)$ is quadratic in creation and annihilation operators; hence it is possible to exhibit a unitary equivalence between $H(g)$ and the free Hamiltonian corresponding to the single particle energy μ_g. The intertwining operator which accomplishes this equivalence is also quadratic in creation and annihilation operators. Some of the results in Secs. 2, 4, and 5 have already been obtained by Eachus[5] in his comprehensive study of ϕ^2; but the proofs that we present are more direct than his.

In spite of its "triviality," the ϕ^2 model illuminates many aspects of the constructive field theory program initiated by Glimm and Jaffe (cf. Ref. 7 and the articles cited there): For $s \leq 3$, $(\phi^2)_{s+1}$ shares a number of properties with the much more difficult nonlinear theories provided by the $(\phi^{2n})_2$ model[8,9] and the Yukawa$_2$ model.[10,11] In particular these theories can all be realized in the standard Fock space representation for the bare masses, and when the spatial cutoff is removed the renormalized Hilbert space is obtained through a weak limit of vacuum states. The final renormalized representation is inequivalent to the original Fock one, but it is "locally Fock" in a sense to be described in Sec. 6.

In this paper we utilize the exact solution of the ϕ^2 model to provide a concise and explicit example of these phenomena. It is our hope that the ϕ^2 model can thus serve as a guide to the genuinely nonlinear theories. It remains to be seen whether the guidelines are too linear.

2. EXACT DRESSING TRANSFORMATION

Friedrichs[1] and Eachus[5] derive several forms of the dressing transformation T_g. Since our formulation differs somewhat from theirs, we outline the derivation of T_g in the Appendix.

Operators on the single-particle (momentum) space $\mathcal{K} = L^2(\mathbf{R}^s)$ will be denoted by small letters, and operators on Fock space $\mathfrak{F}_0$ by capital letters. We take the standard representation of the free field with mass m_0 on $\mathfrak{F}_0 = \sum_{n=0}^{\infty} \oplus \mathcal{K}_n$, where $\mathcal{K}_n$ is the n-fold symmetric tensor product of $\mathcal{K}$. In terms of the annihilation and creation operators $a(k)$ and $a^*(k)$ [cf. (A6)], the time zero field is given by

$$\phi(x) = (2\pi)^{-s/2} \int d^s k (2\mu)^{-1/2} e^{ikx} [a(k) + a^*(-k)], \quad (2.1)$$

where we shall write $\mu = \mu_0 = (k^2 + m_0^2)^{1/2}$ and $\mu_1 = (k^2 + m_1^2)^{1/2}$.

From the conditions on $g(x)$ and λ, it follows readily[12] that μ and μ_g are self-adjoint, and that $\mu \geq m_0$ and $\mu_g \geq m$ where $m = \min(m_0, m_1) > 0$; moreover, there are constants a and b such that $\mu^2 \leq a^2 \mu_g^2 \leq b^2 \mu^2$ on the domain $D(\mu) \times D(\mu)$. Thus $D(\mu) = D(\mu_g)$, and by an interpolation theorem of Löwner and Heinz,[13] for any $0 \leq \delta \leq 2$,

$$\mu^\delta \leq a^\delta \mu_g^\delta \leq b^\delta \mu^\delta. \quad (2.2)$$

T_g is a unitary operator on $\mathfrak{F}_0$ that intertwines the free and total Hamiltonians

$$H_{\text{ren}}(g) = T_g H_0(g) T_g^{-1} \quad (2.3)$$

Here $H_0(g)$ is the biquantization of the single particle operator μ_g:

$$H_0(g) = \int a^*(k_1) \mu_g(k_1, k_2) a(k_2) dk_1 dk_2, \quad (2.4)$$

where $\mu_g(k_1, k_2)$ is the kernel of μ_g. For a discussion of such operators and the proof that the self-adjointness of the single particle operator implies that of the biquantized operator, see either Ref. 5 or Ref. 8 (c). Formally $H_{\text{ren}}(g)$ is a renormalization of the Hamiltonian (1.1),

$$H_{\text{ren}}(g) = H_0 + \lambda H_I(g) + E(g), \quad (2.5)$$

where $E(g) = \frac{1}{4} \| (\mu - \mu_g) \mu^{-1/2} \|_2^2$ is a vacuum renormalization constant. Here $\| \cdot \|_2$ is the Hilbert–Schmidt norm.

It is not obvious that formula (2.5) gives a well-defined operator except when $s = 1$. Indeed a calculation shows that $E(g)$ is infinite unless $s \leq 2$, and even when $s = 2$ it is not clear that the sum $H_0 + \lambda H_I(g)$ can be defined as an operator. Thus the operator equality (2.3) can be established directly only when $s = 1$, and by a bilinear form argument when $s = 2$. One resolution of the difficulty when $s = 3$ would be to introduce and then remove a momentum cutoff σ, in the manner of Glimm[14] (see also Ref. 6). In the cut-off version of (2.3),

$$H_{\text{ren},\sigma}(g) T_{\sigma,g} = T_{\sigma,g} H_{0,\sigma}(g),$$

all the terms would be well defined. One could then show that as $\sigma \to \infty$, $H_{0,\sigma}(g) \to H_0(g)$ and $T_{\sigma,g} \to T_g$. This procedure would exhibit $H_{\text{ren}}(g)$ as the (graph) limit of the sequence $H_{\text{ren},\sigma}(g)$ and explicitly demonstrate the "infinite cancellations" in $H_{\text{ren}}(g)$. To expedite matters we shall simply take (2.3) as the definition of $H_{\text{ren}}(g)$. $H_{\text{ren}}(g)$ is thus a positive self-adjoint operator with domain $T_g D(H_0(g))$.

We now describe the form of T_g most useful for our purposes. $\mu^{-1/2} \mu_g^{1/2}$ can be decomposed in polar form

$$\mu^{-1/2} \mu_g^{1/2} = \rho u, \quad (2.6)$$

where $\rho = (\mu^{-1/2} \mu_g \mu^{-1/2})^{1/2}$. From (2.2) we see that $0 < a^{-1/2} \leq \rho \leq a^{-1/2} b^{1/2}$, so that ρ is a bounded invertible self-adjoint operator and u is unitary.

At this point it is necessary to impose a further condition on $g(x)$, namely that $g(x) = g(-x)$. Then if we regard g as a multiplication operator [i.e., convolution with $\hat{g}(k)$ on $\mathcal{K}$], we see that g commutes with the reflection operator r,

$$(rf)(k) = f(-k), \quad (2.7)$$

where $f \in \mathcal{K}$. It follows that r commutes with μ_g^2, hence with any function of μ_g; accordingly r commutes with ρ^2 and any function of ρ.

Now it is of critical importance for what follows that the operator

$$l = r \ln \rho \in \text{HS}. \quad (2.8)$$

where HS is the class of Hilbert–Schmidt operators on $\mathcal{K}$. (2.8) was proved by Eachus[5] by a perturbation series technique and we give a simpler proof in Lemma 2.1 below. This condition has been isolated by a number of other authors, for example Chadam,[3] who has proved in a more general formulation that (2.8) is a necessary and sufficient condition for the ϕ^2 dynamics to be unitarily implementable in $\mathfrak{F}_0$.

In terms of l we define

$$L = \tfrac{1}{2}i\int[l(k_1,k_2)a^*(k_1)a^*(k_2) - l(k_1,k_2)a(k_1)a(k_2)]dk_1dk_2. \quad (2.9)$$

It can be shown by an analytic vector argument,[5] that (2.8) implies that L is self-adjoint. T_g is written as a product of two unitary operators U_1U_2, the first of which is $U_1 = e^{iL}$. The second U_2 is the biquantization of u in the following sense. By the Spectral Theorem, we have $u = e^{iw}$ where w is a bounded self-adjoint operator on $\mathcal{K}$. The self-adjoint operator

$$W = \int a^*(k_1)w(k_1,k_2)a(k_2)dk_1dk_2$$

then generates $U_2 = e^{iW}$. In the notation of Ref. 15, we can write $U_2 = \Gamma(u)$ and $W = d\Gamma(w)$. Finally,

$$T_g = U_1U_2 = e^{iL}e^{iW}. \quad (2.10)$$

The main tool in the next lemma and in this paper is a representation for the fractional power of a positive self-adjoint operator obtained from the theory of the Dunford–Taylor integral[12]: Let $h \geq mI$ where $m > 0$. Then for $0 < \alpha < 1$,

$$h^\alpha = \frac{\sin\pi\alpha}{\pi}\int_0^\infty y^{\alpha-1}(h+y)^{-1}hdy \quad (2.11)$$

on $D(h)$. Note that the integral is convergent since

$$\|(h+y)^{-1}\| \leq (m+y)^{-1}. \quad (2.12)$$

We collect here the conditions that we impose on $g(x)$ even though all of them are not immediately used:

$$g(x) \in C_0^3; \quad g(x) = g(-x); \quad 0 \leq g(x) \leq 1, \quad (2.13)$$

where we also require that λ is not too negative in the sense that $m_1^2 = m_0^2 + \lambda > 0$. In the limit $g(x) \to 1$, we shall need some uniformity in the limit, such as $|g''(x)| \leq M < \infty$ and a bound on the ratio of the volumes

$$|\operatorname{supp} g|/|g^{-1}(1)| \leq R < \infty. \quad (2.14)$$

This is most easily arranged by considering the sequence

$$g_n(x) = h(x/n), \quad (2.15)$$

where $h(x)$ is a function satisfying (2.13) and (2.14).

Lemma 2.1: (Eachus[16]) The operator

$$\delta_g = \rho^2 - 1 = \mu^{-1/2}(\mu_g - \mu)\mu^{-1/2} \in \text{HS}. \quad (2.16)$$

Proof: Using (2.11) with $h = \mu_g^2$ we have on $D(\mu_g^2)$

$$\mu_g = \frac{1}{\pi}\int_0^\infty y^{-1/2}(\mu_g^2+y)^{-1}\mu_g^2dy.$$

Therefore on $D(\mu_g^2) \cap D(\mu^2)$

$$\mu_g - \mu = \frac{1}{\pi}\int_0^\infty y^{-1/2}[(\mu_g^2+y)^{-1}\mu_g^2 - (\mu^2+y)^{-1}\mu^2]dy$$

$$= \frac{1}{\pi}\int_0^\infty y^{1/2}(\mu_g^2+y)^{-1}[\mu_g^2-\mu^2](\mu^2+y)^{-1}dy$$

$$= \frac{\lambda}{\pi}\int_0^\infty y^{1/2}(\mu_g^2+y)^{-1}g(\mu^2+y)^{-1}dy, \quad (2.17)$$

where g in the preceding integral represents the operator of multiplication by $g(x)$. By (2.12) this last integral is norm convergent, so that $(\mu_g - \mu)$ extends to a bounded operator on $\mathcal{K}$.

To prove that $\delta_g \in \text{HS}$ we use the fact that if r, s, and t are bounded operators with $s \in \text{HS}$, then $rst \in \text{HS}$ and

$$\|rst\|_2 \leq \|r\|\|s\|_2\|t\|. \quad (2.18)$$

For $0 < \epsilon < \frac{1}{4}$ we write by (2.17)

$$\delta_g = \frac{\lambda}{\pi}\int_0^\infty y^{1/2}s_1s_2(y)s_3s_4s_5(y)dy, \quad (2.19)$$

where $s_1 = \mu^{-1/2}\mu_g^{1/2}$, $s_2 = \mu_g^{1/4+\epsilon}(\mu_g^2+y)^{-1}$, $s_3 = \mu_g^{-3/4-\epsilon}\mu^{3/4+\epsilon}$, $s_4 = \mu^{-3/4-\epsilon}g\mu^{-3/4-\epsilon}$, $s_5 = (\mu^2+y)^{-1}\mu^{1/4+\epsilon}$.

By (2.2) s_1 and s_3 are bounded. Clearly s_2 and s_5 are bounded; in fact a calculation of $\max_{x\geq m}[(y+x^2)^{-1}x^{1/4+\epsilon}]$ shows that for large y,

$$\|s_2(y)\| + \|s_5(y)\| \leq \text{const}\, y^{(-7/8)+(\epsilon/2)}.$$

As for s_4 we compute in momentum space that

$$\|s_4\|_2^2 \leq \int d^sk_1d^sk_2\mu(k_1)^{-(3/2)-2\epsilon}|\hat{g}(k_1-k_2)|^2\mu(k_2)^{-(3/2)-2\epsilon}$$

$$\leq \tfrac{1}{2}\int d^sk_1d^sk_2|\hat{g}(k_1-k_2)|^2[\mu(k_1)^{-3-4\epsilon} + \mu(k_2)^{-3-4\epsilon}]$$

by the arithmetic-geometric-mean inequality. Since by (2.13) $\hat{g} \in L_2$, and since $s \leq 3$, we see that $s_4 \in \text{HS}$. It follows from (2.18) that the integrand in (2.19) has finite HS norm which for large y is dominated by $O(y^{-5/4+\epsilon})$. Hence the integral (2.19) converges in HS norm.

Corollary 2.1: The operator $l \in \text{HS}$.

Proof: Note that $\ln\rho = \sinh^{-1}(h_-)$ where $h_- = \frac{1}{2}\rho^{-1}\delta_g$. By Lemma 2.1 and (2.18), h_- is in HS; since $|\sinh^{-1}(x)| \leq |x|$ so are $\sinh^{-1}(h_-)$ and l.

3. HYPERBOLICITY

The time-dependent field is

$$\phi(x,t) = e^{itH_{\text{ren}}(g)}\phi(x,0)e^{-itH_{\text{ren}}(g)},$$

where $\phi(x,0)$ is given by (2.1). Although $\phi(x,t)$ appears to be dependent on the space cutoff g, in fact it is not provided that $g \gg (x,t)$. By this notation we mean that $g(y) = 1$ in the ball $B_t(x) = \{y \mid |y-x| \leq |t|\}$. Similarly we write $g \gg (f,t)$ provided $g(y) = 1$ in the region $\cup_x B_t(x)$ where the union takes place over $x \in \operatorname{supp} f$.

When $s = 1$ this hyperbolicity, or independence of $\phi(x,t)$ on the values of g outside $B_t(x)$, is known for the ϕ^{2n} model. The proof relies on a theorem of Segal[7,17] and the fact that $H_{\text{ren}}(g)$ is essentially self-adjoint on the domain $D(H_0) \cap D(H_I(g))$.[9] In the present case when $s = 2$ or 3, this method is not available but it is easy enough to verify hyperbolicity directly.

Let $f(x)$ be a real C^3 function of compact support. We define the averaged field $\int \phi(x,0)f(x)dx$ as

$$\phi(f,0) = 2^{-1/2}[a(\mu^{-1/2}\hat{f}) + a^*(r\mu^{-1/2}\hat{f})]. \tag{3.1}$$

An analytic vector argument shows that $\phi(f,0)$ is self-adjoint on a domain including $D(N^{1/2})$ where N is the number operator. By (2.3) we have $\phi(f,t) = T_g e^{itH_0(g)} T_g^* \phi(f,0) T_g e^{-itH_0(g)} T_g^*$ on $D(N)$, say. Using (A14), its adjoint, and (A5), we find that

$$T_g^* \phi(f,0) T_g = \phi(\mu^{1/2}\mu_g^{-1/2} f, 0) \tag{3.2}$$

on $D(N)$. But by (A11) with $w = \mu_g$,

$$e^{itH_0(g)} a^{\#}(\hat{f}) e^{-itH_0(g)} = a^{\#}(e^{\mp it\mu_g}\hat{f}) \tag{3.3}$$

on $D(N^{1/2})$, where $a^{\#}$ stands for a or a^*. A straightforward calculation based on (3.1)–(3.3), (A.9), (A.7), and (A.5) yields

$$\phi(f,t) = \phi(f_1(t), 0) + \pi(f_2(t), 0), \tag{3.4}$$

where the conjugate field

$$\pi(f,0) = -i2^{-1/2}[a(\mu^{1/2}\hat{f}) - a^*(j\mu^{1/2}\hat{f})],$$

$f_1(t) = (\cos\mu_g t)f$, and $f_2(t) = (\mu_g^{-1}\sin\mu_g t)f$. Similarly we find that $\pi(f,t) = \phi(f,t)$; that is,

$$\pi(f,t) = \phi(f_3(t), 0) + \pi(f_1(t), 0), \tag{3.5}$$

where $f_3(t) = \dot{f}_1(t) = -(\mu_g \sin\mu_g t)f$.

But we recognize $f_1(x,t)$ as the solution to the classical initial value problem

$$\begin{aligned} &u_{tt} - \Delta u + m_0^2 u + \lambda g u = 0, \\ &u(x,0) = f(x), u_t(x,0) = 0 \end{aligned} \tag{3.6}$$

and similarly for $f_2(x,t)$ except that the initial conditions are $u(x,0) = 0$ and $u_t(x,0) = f(x)$. Now Eq. (3.6) is hyperbolic with speed of propagation equal to 1. This can be seen explicitly, for example, from the corresponding spherical mean integral equation. Hence f_1 and similarly f_2 are independent of g if $g \gg (f,t)$. We have proved:

Theorem 3.1: Let $f(x)$ be a real C^3 function of compact support. Then the sharp time fields $\phi(f,t)$ and $\pi(f,t)$ are self-adjoint operators whose domains include $D(N^{1/2})$, and which are independent of g provided that $g \gg (f,t)$.

The algebra of local observables (essentially functions of the fields) can thus be formulated independently of g. If B is a bounded open region of space $B \subset \mathbf{R}^s$, we define $\mathfrak{A}_0(B)$ to be the norm closure of the linear span generated by the operators $e^{i[\phi(f,0)+\pi(h,0)]}$, where f and h are real C^∞ functions with support in B. We define the local (von Neumann) algebra $\mathfrak{A}(B)$ as the strong closure of $\mathfrak{A}_0(B)$. The (C^*-algebra) of quasilocal observables is the norm closure of the union

$$\mathfrak{A} = [\cup_B \mathfrak{A}(B)]^-, \tag{3.7}$$

where B ranges over all bounded open subsets of $\mathbf{R}^s$. Alternatively we could define $\mathfrak{A}(B)$ for space–time B in terms of time-averaged fields, but by (3.4) and (3.5) we see that the resulting algebra $\mathfrak{A}$ would be the same. It is apparent at this point that we do not have genuinely interacting fields. The decision to take the algebra of observables to be $\mathfrak{A}$ rather than $\mathfrak{A}_0 = [\cup\mathfrak{A}_0(B)]^-$ is for mathematical convenience; it is justified by the results of Sec. 6 which imply the equivalence of the strong closures of the local algebras in the two representations of interest.

4. SPECTRUM OF $H_{\text{ren}}(g)$

By the unitary equivalence (2.3), $H_{\text{ren}}(g)$ and $H_0(g)$ have the same spectrum. But the spectrum of $H_0(g)$ is determined by that of μ_g: Suppose that μ_g has spectral decomposition $\mu_g = \int \lambda\, d e_\lambda$, then we have[18]

$$H_0(g) = \sum_{n=0}^{\infty} \oplus \int \left(\sum_{i=1}^{n} \lambda_i\right) dE_{\lambda_1\lambda_2\cdots\lambda_n}, \tag{4.1}$$

where $E_{\lambda_1\cdots\lambda_n}$ is the symmetric tensor product $e_{\lambda_1} \otimes e_{\lambda_2} \otimes \cdots \otimes e_{\lambda_n}$, and $\{E_{\lambda_1\cdots\lambda_n}\}$ gives an n-parameter resolution of the identity in $\mathcal{K}_n$.

Now the spectrum of μ_g is well understood. Consider the Schrödinger Hamiltonian $h(g) = \mu_g^2 - m_0^2 = -\Delta + \lambda g(x)$. Its spectrum $\sigma(h(g))$ is just what one expects. The positive real line $[0,\infty)$ forms the absolutely continuous part of the spectrum. If $\lambda \geq 0$, there is nothing else. If $\lambda < 0$ there are in addition a finite number of eigenvalues in $(\lambda, 0]$; as $g(x) \to 1$ the number of eigenvalues is of the order of $|\text{supp}g|$. This behavior is consistent with the fact that as $g \to 1$, $\sigma(h(g))$ approaches the spectrum of $(-\Delta + \lambda)$ in a lower semicontinuous fashion[12]; that is, the point spectrum of $h(g)$ must fill in the interval $[\lambda, 0]$.

The above conclusions follow at once from theorems of Kato,[12] Ikebe,[19] and Weidmann[20] and the rather strong conditions (2.13)–(2.14) on the potential $g(x)$. For example, Ikebe has shown that $h(g)$ has no singular spectrum by establishing the completeness of the system of normalized eigenfunctions and "distorted plane wave eigenfunctions" of $h(g)$. The estimate on the number of eigenvalues follows by comparing $h(g)$ with spherical well Hamiltonians for which the eigenvalues can be explicitly computed [here we are using (2.14)].

From the spectral theorem, we see that if $\lambda \geq 0$, then $\sigma(\mu_g) = [m_0, \infty)$ is absolutely continuous. Thus by (4.1), $H_{\text{ren}}(g)$ has an eigenvalue at 0 and absolutely continuous spectrum on $[m_0, \infty)$. If $\lambda < 0$, $\sigma(\mu_g)$ consists of an absolutely continuous part on $[m_0, \infty)$ together with a finite number of eigenvalues $\lambda_1, \ldots, \lambda_j$ in $(m_1, m_0]$. Here $j \sim |\text{supp}g|$. It follows from (4.1) that $H_{\text{ren}}(g)$ has eigenvalues $n_1\lambda_1 + n_2\lambda_2 + \cdots + n_j\lambda_j$, where $n_1, \ldots, n_j$ are nonnegative integers, and absolutely continuous spectrum on $[m_0, \infty)$. We note that in this case, $H_{\text{ren}}(g)$ has positive eigenvalues imbedded in the continuous spectrum. In either case the Fock vacuum Ω_0 is a unique ground state for $H_0(g)$, and hence

$$\Omega_g = T_g \Omega_0 \tag{4.2}$$

is a unique vacuum for $H_{\text{ren}}(g)$ corresponding to the eigenvalue 0.

A number of these features have already turned up in the $(\phi^{2n})_2$ theory: Glimm and Jaffe[8] have shown that

$H(g)$ has compact spectrum in $(-\infty, m_0)$, and Hoegh-Krohn[21] has established that $[m_0, \infty)$ forms the continuous spectrum.

In addition Glimm and Jaffe have verified the uniqueness of the vacuum for $H(g)$. On the basis of the results of this section one can conjecture that in the spatially cut-off $(\phi^{2n})_2$ model, there is no singular spectrum; that if 0 is the only eigenvalue in $[0, m_0)$ (corresponding to Ω_g), then $[m_0, \infty)$ is pure continuous spectrum; if there are more eigenvalues, then some point spectrum is imbedded in $[m_0, \infty)$; however, if this is the case it should be possible to remove these eigenvalues by a positive mass renormalization (corresponding to $\lambda \geq 0$ in the present model).

The above spectral information can be related to the question of whether $H_{\text{ren}}(g)$ is unitarily equivalent to H_0. As discussed by Eachus,[5] this question reduces to that of the unitary equivalence of μ_g and μ_0. For suppose there is a unitary operator u intertwining μ_g and μ_0:

$$\mu_g u = u\mu_0. \tag{4.3}$$

Then if we define U to be the unitary operator in $\mathfrak{F}_0$ which is the biquantization of u as in Sec. 2, we obtain by (A11)

$$H_0(g)U = UH_0 \tag{4.4}$$

and from (2.3)

$$H_{\text{ren}}(g)W = WH_0,$$

where W is the unitary operator $T_g U$.

The operator u of (4.3) is of course the wave operator of quantum mechanics; it exists,[12] and hence so does W, for $\lambda \geq 0$ and even for negative λ as long as $h(g)$ has no point spectrum.

When $\lambda < 0$ and $h(g)$ does have eigenvalues then u is only an isometry, mapping $\mathcal{H}$ onto $\mathcal{H}_{\text{a.c.}} = \mathcal{H} \ominus \mathcal{H}_p$; here $\mathcal{H}_p$ is the finite-dimensional subspace of $\mathcal{H}$ spanned by the corresponding eigenfunctions, and $\mathcal{H}_{\text{a.c.}}$ is the "absolutely continuous" part of $\mathcal{H}$ with respect to $h(g)$. However, in this case, we still have the relation (4.4) where now $U = \Gamma(u)$ (see Sec. 2) is an *isometry* mapping $\mathfrak{F}_0$ onto the subspace $\mathfrak{F}(\mathcal{H}_{\text{a.c.}})$, the Fock space constructed from the one-particle space $\mathcal{H}_{\text{a.c.}}$, i.e., $\mathfrak{F}(\mathcal{H}_{\text{a.c.}}) = \Sigma \oplus \mathcal{H}_{\text{a.c.}} \otimes \cdots \otimes \mathcal{H}_{\text{a.c.}}$. If we decompose $\mathfrak{F}_0 = \mathfrak{F}(\mathcal{H}_{\text{a.c.}}) \otimes \mathfrak{F}(\mathcal{H}_p)$ then $H_0(g)$ can be written as

$$H_0(g) = (UH_0U^* \otimes I) + (I \otimes H_p),$$

where H_p is the biquantization of $\mu_g \restriction \mathcal{H}_p$. This gives a corresponding decomposition for $H_{\text{ren}}(g)$.

Note that this latter case [$\sigma_p(h(g))$ nonempty and W an isometry] does not argue against the unitary equivalence of H_{ren} and H_0 in general in quantum field theory. This example indicates that it may simply be necessary to adjust H_{ren} by the addition of counterterms (in this case, a finite positive mass counterterm).

5. REMOVAL OF THE SPACE CUTOFF

As explained in the introduction, when $g \to 1$ the Hilbert space is renormalized by means of a weak limit of vacuum states. More precisely for any A in the algebra $\mathfrak{A}$ of observables, we define

$$\omega_g(A) = (\Omega_g, A\Omega_g) \tag{5.1}$$

in terms of the unique vacuum vector Ω_g of $H_{\text{ren}}(g)$. Clearly ω_g is positive linear functional on $\mathfrak{A}$ of norm 1, i.e., a state. As $g \to 1$ we show that ω_g converges in the w^*-topology to a state ω_1; that is, for each A in $\mathfrak{A}$,

$$\omega_g(A) \to \omega_1(A). \tag{5.2}$$

Moreover, for $A \in \mathfrak{A}(B)$, where B is a fixed bounded set in $\mathbf{R}^s$, (5.2) takes place in the sense of *norm convergence*. The limiting state ω_1 uniquely defines a new representation of $\mathfrak{A}$ in a "renormalized or physical Hilbert space" $\mathfrak{F}_1$ by means of the GNS construction.[22] In particular there is a cyclic vector $\Omega_1 \in \mathfrak{F}_1$, such that

$$\omega_1(A) = (\Omega_1, A\Omega_1).$$

Finally, we can identify this representation as the Fock representation for a free field of mass m_1 with no-particle state Ω_1.

Thus the arduous renormalization of $P(\phi)_2$ by Glimm and Jaffe[7,8] is simply illustrated by the above statements [except for the last which would be a disaster for $P(\phi)_2$ and the statement of local norm convergence where we rely on the $P(\phi)_2$ proof]. Note, however, that no space averaging is required in the definition of ω_g, that the original sequence $\{\omega_g\}$ and not just a subsequence converges, and that the limit ω_1 is unique. It has been conjectured that the same is true for $P(\phi)_2$.

Consider now a generator A of the local algebra $\mathfrak{A}(B)$, $A = e^{i[\phi(f)+\pi(h)]}$, where $\phi(f)$ and $\pi(h)$ are time-zero fields, and $f, h \in C_0^\infty(B)$ are real. From (3.2) and its analog for π, we see that

$$\begin{aligned}\omega_g(A) &= (\Omega_0, T_g^{-1} A T_g \Omega_0) \\ &= (\Omega_0, e^{i[\phi(\mu^{1/2}\mu_g^{-1/2}f)+\pi(\mu^{-1/2}\mu_g^{1/2}h)]}\Omega_0).\end{aligned} \tag{5.3}$$

It is a standard free field calculation that

$$\begin{aligned}\omega_0(A) &= (\Omega_0, e^{i[\phi(f)+\pi(h)]}\Omega_0) \\ &= \exp[-\tfrac{1}{4}(\langle f, \mu^{-1}f\rangle + \langle h, \mu h\rangle)],\end{aligned} \tag{5.4}$$

where $\langle \cdot, \cdot \rangle$ is the Hermitian inner product on $\mathcal{H}$. Hence from (5.3),

$$\omega_g(A) = \exp[-\tfrac{1}{4}(\langle f, \mu_g^{-1}f\rangle + \langle h, \mu_g h\rangle)]. \tag{5.5}$$

Formulas (5.4) and (5.5) provide the clearest demonstration of the statement that turning on the interaction $\lambda H_I(g)$ amounts to replacing the single particle energy μ by μ_g.

Formally it is obvious that in the limit $g = 1$,

$$\omega_1(A) = \exp[-\tfrac{1}{4}(\langle f, \mu_1^{-1}f\rangle + \langle h, \mu_1 h\rangle)]. \tag{5.6}$$

This knowledge of ω_1 on the generators A is enough for us to conclude that the representation of the canonical commutation relations determined by the state ω_1 is that of Fock space with mass m_1.[23,24]

We now prove the convergence of ω_g to ω_1. Note that

we are making some assumptions such as (2.15) about the uniformity of the limit $g(x) \to 1$. Let $\zeta(x)$ be a C_0^∞ function with support in the bonded open region $B \subset \mathbf{R}^s$, and let $\chi_j(x)$ be the characteristic function for the interval $[j, j+1) \subset \mathbf{R}^s$. Here $j = (j_1, \ldots, j_s)$, $[j, j+1) = \prod_{i=1}^s [j_i, j_i + 1)$, and $|j| = (j_1^2 + \cdots + j_s^2)^{1/2}$. We shall generally regard ζ and χ_j as multiplication operators on $\mathcal{H}$. To begin with we prove the following estimate on the loss of locality produced by the operator $\mu = (-\Delta + m_0^2)^{1/2}$ [cf. Ref. 8(c)].

Lemma 5.1: For $\alpha \leq 0$, there is a constant c independent of j such that

$$\|\chi_j \mu^\alpha \zeta\| \leq c(1 + |j|)^{-(\alpha+s)/2} e^{-m_0|j|}. \tag{5.7}$$

Proof: In coordinate space the operator μ^α is given by convolution with the (distribution) kernel

$$k_\alpha(x) = (2\pi)^{-s} \int e^{-ipx}(p^2 + m_0^2)^{\alpha/2} dp.$$

For $\alpha < -s$ this integral converges and in terms of the modified Bessel function K_ν we calculate that[25]

$$k_\alpha(x) = (2\pi)^{-s/2} 2^{\alpha/2+1} \Gamma(-\tfrac{1}{2}\alpha)^{-1} \times (m_0/|x|)^{(\alpha+s)/2} K_{(\alpha+s)/2}(m_0|x|). \tag{5.8}$$

In fact (5.8) is valid for all real $\alpha \neq 0, 2, 4, \ldots$, and from it we can read off the relevant properties of k_α; namely, $k_\alpha(x) \in C^\infty(\mathbf{R}^s)$ except at $x = 0$ (where it has a singularity of order $|x|^{-(\alpha+s)}$ when $\alpha + s > 0$), and as $|x| \to \infty$,

$$k_\alpha(x) = O[(m_0/|x|)^{(\alpha+s)/2} e^{-m_0|x|}]. \tag{5.9}$$

The lemma follows at once from (5.9).

Remark: Note that even when $\alpha + s < 0$ we have $\|\chi_j \mu^\alpha \zeta\| = O(e^{-(m_0-\epsilon)|j|})$ for any $\epsilon > 0$. Moreover, when the closures of the supports of χ_j and ζ are disjoint, $\|\chi_j \mu^\alpha \zeta\|_2 \leq c_1 e^{-(m_0-\epsilon)(|j|-c_2)}$ where c_1, c_2 are independent of both j and m_0. In particular,

$$\|\chi_j(\mu^2 + y)^{-1}\zeta\|_2 \leq c_1 e^{-(m_0+y-\epsilon)(|j|-c_2)}. \tag{5.10}$$

Lemma 5.2: As $g \to 1$, $\|(\mu_g - \mu_1)\zeta\| \to 0$.

Proof: The calculation (2.17) shows that $(\mu_g - \mu_1)$ is a bounded operator and that

$$(\mu_g - \mu_1)\zeta = \frac{\lambda}{\pi} \int_0^\infty y^{1/2}(\mu_g^2 + y)^{-1}(g-1)(\mu_1^2 + y)^{-1}\zeta \, dy.$$

When $g = 1$ on a sufficiently large set (say $|x| \leq c$), we obtain by (5.10)

$$\|(\mu_g - \mu_1)\zeta\| \leq \text{const} \sum_{|j| \geq c} \int_0^\infty \|\chi_j(\mu_1^2 + y)^{-1}\zeta\| dy \leq \text{const} \sum_{|j| \geq c} \int_0^\infty e^{-(m_1+y-\epsilon)(|j|-c_2)} dy.$$

The lemma follows upon choosing c large.

Lemma 5.3: As $g \to 1$, $\mu_g \overset{s}{\to} \mu_1$ and $\mu_g^{-1} \overset{s}{\to} \mu_1^{-1}$.

Proof: Note that by (2.2), $D(\mu_g) = D(\mu_1)$ and, by (2.17), $\|\mu_g - \mu_1\| \leq c < \infty$, where c is independent of g. We approximate any $f \in D(\mu_1)$ by ζf: $\|f - \zeta f\| < \epsilon$, where $\epsilon > 0$ is arbitrary and $\zeta \in C_0^\infty$ is chosen as an approximation to the identity. Next choose g sufficiently close to 1 that $\|(\mu_g - \mu_1)\zeta\| < \epsilon$. Then

$$\|(\mu_g - \mu_1)f\| \leq \|(\mu_g - \mu_1)(f - \zeta f)\| + \|(\mu_g - \mu_1)\zeta f\| < c\epsilon + \epsilon\|f\|.$$

The convergence of μ_g^{-1} follows from that of μ_g.

By the form (5.5) of $\omega_g(A)$, Lemma 5.3 establishes the convergence (5.2) for a generator A. This convergence obviously extends by closure to all of $\mathfrak{A}_0$:

Theorem 5.1: (Eachus) As states on $\mathfrak{A}_0$, $\omega_g \to \omega_1$ in the w^*-topology.

Finally we verify that ω_g converges in norm to ω_1 on each local algebra $\mathfrak{A}(B)$. We have been unable to prove this directly because of lack of a good underestimate for the norm $\|A\|$ of a local observable A. Instead the proof relies on results of the next section and a theorem of Glimm and Jaffe.[8] In Sec. 6 we introduce the local number-energy operator $N_{\tau,B}$ for the interval B and prove that $\omega_g(N_{\tau,B}) \leq c$ where the constant c depends on the volume of B but is otherwise independent of B and g. By Theorem 4.1 of Ref. 8(c), we conclude that on each $\mathfrak{A}(B)$, the ω_g have a norm convergent subsequence. Glimm and Jaffe prove this theorem for $s = 1$ but with appropriate modifications it remains valid in any number of space dimensions. We also need the fact that ω_1 is a normal state on $\mathfrak{A}_0(B)$ and hence extends to a normal state on $\mathfrak{A}(B)$. (A normal state ω is one for which $\omega(A_\nu) \to \omega(A)$ as $A_\nu \uparrow A$.) By Theorem 6.1, $\omega_1 \restriction \mathfrak{A}_0(B)$ is unitarily equivalent to a vector state and is hence normal.

Suppose then that ω_g does not converge in norm to ω_1 on $\mathfrak{A}(B)$. Then there exists a subsequence ω_{g_n} such that $\|(\omega_{g_n} - \omega_1)\restriction \mathfrak{A}(B)\| \geq \epsilon$ for some $\epsilon > 0$. But by the above result there is a further subsequence of the ω_{g_n} that converges in norm to a (normal) state ω_2 on $\mathfrak{A}(B)$. Clearly $\|(\omega_2 - \omega_1)\restriction \mathfrak{A}(B)\| \geq \epsilon$. However, by Theorem 5.1, $\omega_g(A) \to \omega_1(A)$ for every $A \in \mathfrak{A}_0(B)$; hence $\|(\omega_2 - \omega_1)\restriction \mathfrak{A}_0(B)\| = 0$. But $\|(\omega_2 - \omega_1)\restriction \mathfrak{A}_0(B)\| = \|(\omega_2 - \omega_1)\restriction \mathfrak{A}(B)\|$ since $(\omega_2 - \omega_1)\restriction \mathfrak{A}(B)$ is normal and the unit sphere of $\mathfrak{A}_0(B)$ is strongly dense in the unit sphere of $\mathfrak{A}(B)$.[26] As a result:

Theorem 5.2: As states on $\mathfrak{A}(B)$, $\omega_g \to \omega_1$ in norm.

By an approximation argument:

Corollary 5.1: As states on $\mathfrak{A}$, $\omega_g \to \omega_1$ in the w^*-topology.

6. LOCALLY FOCK PROPERTIES

Let ρ_1 be the representation of $\mathfrak{A}$ on $\mathfrak{F}_1$ determined by ω_1; and let ϕ_1, π_1, and H_1 be the standard fields and free Hamiltonian for the free theory with mass m_1 on $\mathfrak{F}_1$. By (5.6), exponentials of the time-zero fields $A = e^{i[\phi(f)+\pi(h)]} \in \mathfrak{A}$ are represented by $\rho_1(A) = e^{i[\phi_1(f)+\pi_1(h)]}$, and as can be checked from (3.8) and (3.9) this correspondence extends consistently to the time-propagated fields, e.g., $\rho_1(\phi(f,t)) = e^{itH_1}\phi_1(f,0)e^{-itH_1} = \phi_1(f,t)$.

Since representations of the CCR's with different masses are unitarily inequivalent,[24] ρ_1 is not unitarily implementable; however, ρ_1 is *locally* unitarily

implementable. It is this property that Glimm and Jaffe call "locally Fock" in the context of the $P(\phi)_2$ and Y_2 models. Here it might seem redundant to call a representation on Fock space $\mathfrak{F}_1$ "locally Fock," and in fact this redundancy is reflected in the proof. In the case of the $P(\phi)_2$ and Y_2 models, local Fockness depends critically on the local norm convergence of ω_g and the subsequent locally normal property of ρ_1. However, in the present case we need only note that by a result of Araki,[27] $\mathfrak{A}(B)$ and $\rho_1(\mathfrak{A}(B))$ are separable type III factors; hence the isomorphism ρ_1 is unitarily implemented.[28] We state this as:

Theorem 6.1: Let B be a bounded open subset of $\mathbf{R}^s$. There is a unitary operator U_B mapping from $\mathfrak{F}_0$ to $\mathfrak{F}_1$ such that

$$U_B A U_B^{-1} = \rho_1(A) \quad \text{for every } A \in \mathfrak{A}(B).$$

Unfortunately we have been unable to find an explicit representation for the mapping U_B. If we regard T_g as mapping from $\mathfrak{F}_1$ to $\mathfrak{F}_0$, then in a formal sense (since it does not exist) T_1 gives the canonical unitary equivalence between the action of $\mathfrak{A}$ on $\mathfrak{F}_0$ and $\rho_1(\mathfrak{A})$ on $\mathfrak{F}_1$; for support of this statement see formula (3.4). One might think then that U_B is given by T_g^{-1} with $g = 1$ on B but $g \neq 1$. However, this choice of U_B is only approximately correct; for no matter how large we take the region $\{x \mid g(x) = 1\}$ it is not true that $\mu_g^{-1/2} f = \mu_1^{-1/2} f$ for f with support in B'.

Perhaps the locally Fock property can be best understood in terms of local number-energy operator estimates which we now prove. Let $\chi(x)$ be the characteristic function for the bounded open region $B \subset \mathbf{R}^s$. If $\tau < \frac{1}{2}$, it is easy to see that $\chi \mu^\tau \chi$ is a well-defined positive operator on Schwartz space $\mathcal{S}(\mathbf{R}^s)$. We denote its Friedrichs extension by c_τ. The local number-energy operator $N_{\tau,B}$ for the region B is the biquantization of c_τ, $N_{\tau,B} = d\Gamma(c_\tau)$. Let $\{f_n\}$ be an orthonormal basis for $L^2(\mathbf{R}^s)$ such as the Hermite functions; then a convenient expression for $N_{\tau,B}$ is $\sum_n a^*(c_\tau^{1/2} f_n) a(j c_\tau^{1/2} f_n)$, where the sum certainly makes sense on the dense domain

$$\mathfrak{D} = \{\Psi \in \mathfrak{F}_0 : \Psi_n \in \mathcal{S}(\mathbf{R}^{sn}); \Psi_n = 0 \text{ for large } n\}. \quad (6.1)$$

If we refer to the particles with mass m_0 associated with the Fock space $\mathfrak{F}_0$ as "bare particles" and the particles of $\mathfrak{F}_1$ as "physical particles," then the observable $N_{\tau,B}$ corresponds to the number-energy of bare particles in the region B.

The $N_{\tau,B}$ estimates that we prove are of the form

$$N_{\tau,B}^j \leq c[H_{\text{ren}}(g) + 1]^j, \quad (6.2)$$

where τ is a small positive number, j an arbitrary positive integer, and c a constant depending on the volume of B but *otherwise independent of B and g*. The physical significance of (6.2) can be seen by passing to the limit $g = 1$: Any physical state of finite total energy has a finite number of *bare* particles in any bounded region. The inequality (6.2) is believed to be true for any locally Fock model such as $(\phi^{2n})_2$ or Y_2 but has not yet been proved for these models. Such inequalities would be useful for establishing the regularity of the physical vacuum (e.g., that it is in the domain of powers of the fields) and for determining the relation between normal ordering of powers of the field with respect to the bare vacuum and the physical vacuum.

We need some preliminary estimates.

Lemma 6.1: Let α be real and suppose the integer $r \geq |\alpha|/2$. Then if $\zeta \in C^{2r}(\mathbf{R}^s)$,

$$\|\mu^\alpha \zeta \mu^{-\alpha}\| \leq \text{const} \sup_{|j| \leq 2r} |D^j \zeta|. \quad (6.3)$$

Remark: On the left side of (6.3), ζ is interpreted as a multiplication operator, whereas on the right side $D^j \zeta = \partial_1^{j_1} \cdots \partial_s^{j_s} \zeta$ is the derivative of the function ζ of order $|j| = j_1 + \cdots + j_s$.

Proof: The idea is to commute the μ^α through the ζ. We shall consider the case $0 < \alpha < 2$; the other cases are similar. By (2.11)

$$\mu^\alpha = \frac{\sin \pi\alpha/2}{\pi} \int_0^\infty y^{(\alpha/2)-1} (\mu^2 + y)^{-1} \mu^2 \, dy$$

so that

$$[\mu^\alpha, \zeta] = \frac{\sin \pi\alpha/2}{\pi} \int_0^\infty y^{\alpha/2} [\zeta, (\mu^2 + y)^{-1}] dy.$$

Now

$$\begin{aligned}[\zeta, (\mu^2 + y)^{-1}] &= (\mu^2 + y)^{-1} [-\Delta, \zeta] (\mu^2 + y)^{-1} \\ &= (\mu^2 + y)^{-1} (-\Delta\zeta - 2\nabla\zeta \cdot \nabla) \\ &\quad \times (\mu^2 + y)^{-1}.\end{aligned}$$

We thus obtain

$$\|\mu^\alpha \zeta \mu^{-\alpha}\| \leq \sup|\zeta| + \frac{\sin \pi\alpha/2}{\pi} \int_0^\infty y^{\alpha/2} [a(y) + b(y)] d$$

where

$$a(y) = \|(\mu^2 + y)^{-1} (\Delta\zeta)(\mu^2 + y)^{-1} \mu^{-\alpha}\| = \sup|\Delta\zeta| \, O(y^{-2})$$

and

$$\begin{aligned}b(y) &= \|(\mu^2 + y)^{-1} 2\nabla\zeta \cdot \nabla (\mu^2 + y)^{-1} \mu^{-\alpha}\| \\ &= \sup|\nabla\zeta| \, O(y^{-(\alpha+3)/2})\end{aligned}$$

Therefore, the integral converges and $\mu^\alpha \zeta \mu^{-\alpha}$ is bounded as in (6.3).

Remark: When $(\text{supp}\zeta)^-$ is disjoint from $[j, j+1]$, then the method of the above lemma and the estimate (5.10) yield

$$\|\chi_j \mu^\alpha \zeta \mu^{-\alpha}\|_2 = O(e^{-(m_0 - \epsilon)|j|}) \quad (6.4)$$

for any $\epsilon > 0$. Moreover from (5.7) we see that when $\alpha \geq -1$, we may take $\epsilon = 0$ in (6.4).

If we replace ζ by the characteristic function χ of an interval in $\mathbf{R}^s$, then we obtain a somewhat weaker result.

Lemma 6.2: If $0 \leq \alpha < \frac{1}{2}$ and $\beta > \alpha s$, then $\mu^\alpha \chi \mu^{-\beta}$ is a bounded operator on $L^2(\mathbf{R}^s)$.

Proof: We sketch the proof which is similar to that of Theorem 3.3.1 of Ref. 8(c). As in the previous lemma, it is sufficient to prove that $[\mu^\alpha, \chi]\mu^{-\beta}$ is bounded. Let $f(k) \in \mathcal{S}$. Then

$$|[\mu^\alpha, \chi]\mu^{-\beta} f(k)| \leq \text{const}(\mu^\alpha |\hat{\chi}|) * \mu^{-\beta} |f|(k) \quad (6.5)$$

and estimating the norm of the convolution by L_p norms,[29]

$$\|[\mu^\alpha, \chi]\mu^{-\beta} f\| \leq \text{const}\| \mu^\alpha \hat{\chi} \|_p \|\mu^{-\beta} f\|_r,$$

where $1/p + 1/r = \frac{3}{2}$, $p \geq 1$, $r \geq 1$.

By an elementary computation $|\mu^\alpha \hat{\chi}(k)| = O((k_1 \cdots k_s)^{\alpha-1})$ as $k \to \infty$ so that $\mu^\alpha \hat{\chi} \in L_p$ if $1/p < 1 - \alpha$. By Hölder's inequality, $\| \mu^{-\beta} f\|_r \leq \|f\|_2 \| \mu^{-\beta}\|_{2r/(2-r)}$. This last norm is finite if $2\beta r/(2-r) > s$. The above restrictions on p and r are consistent if $\alpha < \frac{1}{2}$ and $\beta > \alpha s$.

Lemma 6.3: If $0 \leq \alpha < \frac{1}{2}$ and $\beta > \alpha + s/2$, then $\mu^\alpha \chi \mu^{-\beta} \in$ HS.

Proof: The operator $\chi \mu^{\alpha-\beta}$ has kernel $\hat{\chi}(k-k')\mu(k')^{\alpha-\beta}$ which is in $L^2(\mathbf{R}^{2s})$ if $\beta > \alpha + s/2$. Thus by (6.5) the lemma reduces to proving that $\mu^\alpha(k-k') \times \hat{\chi}(k-k')\mu(k')^{-\beta} \in L^2(\mathbf{R}^{2s})$ which follows from the fact that $\mu^\alpha \hat{\chi} \in L^2(\mathbf{R}^s)$.

By Lemma 2.1, $\delta_g \chi \in$ HS. In fact:

Lemma 6.4: If $\alpha < \frac{1}{2}$, then $\|\delta_g \chi \mu^\alpha\|_2$ is bounded independently of g.

Proof: Let $\nu(y) = \mu^{-1/2}(\mu_g^2 + y)^{-1} g (\mu^2 + y)^{-1}\mu^{-1/2} \chi \mu^\alpha$. By (2.17) it is sufficient to show that $\nu(y) \in$ HS and that for large y, $\|\nu(y)\|_2 \leq b|y|^{-(3/2)-\epsilon}$ for some $\epsilon > 0$ and b independent of g. Accordingly, for $0 < \epsilon < 1 - s/4 - \alpha/2$, we rewrite $\nu(y) = s_1(y)s_2 s_3(y)s_4$, where $s_1(y) = \mu^{-1/2}(\mu_g^2 + y)^{-1}\mu^{1/2}$, $s_2 = \mu^{-1/2} g \mu^{1/2}$, $s_3(y) = \mu^{-1/2}(\mu^2 + y)^{-1}\mu^{(3/2)-2\epsilon}$, and $s_4 = \mu^{-2+2\epsilon}\chi\mu^\alpha$. By Lemma 6.3, $s_4 \in$ HS, and by Lemma 6.1, s_2 is bounded independently of g. Clearly s_3 is bounded and by (2.2) so is s_1; but in fact for $y > 0$ we have $\|s_1(y)\| = O(y^{-1})$ and, by the arithmetic-geometric-mean inequality, $\|s_3(y)\| = O(y^{-1/2-\epsilon})$. Hence by (2.18), $\|\nu(y)\|_2 = O(y^{-(3/2)-\epsilon})$ for large y.

As applications of these estimates we note that if $0 \leq \tau < \frac{1}{2}$ and $\beta > s\tau/2$, then $c_\tau^{1/2}\mu^{-\beta}$ is a bounded operator. For if $f \in \mathcal{S}$, then

$$\|c_\tau^{1/2}\mu^{-\beta} f\| = \|\mu^{\tau/2}\chi\, \mu^{-\beta} f\| \leq \text{const}\|f\|, \tag{6.6}$$

by Lemma 6.2. Secondly, if $\alpha < \frac{1}{2}$, then

$$\|y_-\chi\mu^\alpha\|_2 \leq \text{const(independent of } g), \tag{6.7}$$

since by (A5), (2.6), and (2.16), $y_- = \frac{1}{2}u^*\rho^{-1}\delta_g$.

Lemma 6.5: If $\tau < \min(\frac{1}{2}, (1/s))$, then

$$\sum_n \|a^\#(jy_\pm c_\tau^{1/2} f_n)(H_0 + 1)^{-1/2}\Psi\|^2 \leq \text{const}\|\Psi\|^2, \tag{6.8}$$

where the constant is independent of g.

Proof: It is sufficient to prove the lemma for Ψ in the r-particle space $\mathcal{H}_r$ with estimates independent of r. From the definition (A6) of $a(f)$ we calculate that

$$\begin{aligned}\sum_n \|a(jy_+ c_\tau^{1/2} f_n)\Psi\|^2 &= r\int dp_1 \cdots dp_{r-1} \langle \Psi(p_1, \ldots, p_{r-1}, \cdot), \\ &\quad y_+ c_\tau y_+^* \Psi(p_1, \ldots, p_{r-1}, \cdot)\rangle \\ &\leq \|c_\tau^{1/2} y_+^* \mu^{-1/2}\|^2 \|(H_0 + 1)^{1/2}\Psi\|^2,\end{aligned}$$

since by symmetry

$$\|(H_0 + 1)^{1/2}\Psi\|^2 = r\int dp_1 \cdots dp_{r-1} \langle \Psi(p_1, \ldots, p_{r-1}, \cdot), \mu\Psi(p_1, \ldots, p_{r-1}, \cdot)\rangle.$$

Now by (6.6), $c_\tau^{1/2}\mu^{-1/2}$ is bounded if $\tau < \min(\frac{1}{2}, (1/s))$ and by (A5) and (2.2) so is $\mu^{1/2} y_\pm^* \mu^{-1/2}$. Thus $\|c_\tau^{1/2} y_\pm^* \mu^{-1/2}\| < \infty$ and (6.8)$_+$ is verified.

As for (6.8)$_-$ we have

$$\begin{aligned}\sum_n \|a^*(jy_- c_\tau^{1/2} f_n)\Psi\|^2 &= \sum_n (\Psi, a(y_- c_\tau^{1/2} f_n) a^*(jy_- c_\tau^{1/2} f_n)\Psi) \\ &= \sum \|a(y_- c_\tau^{1/2} f_n)\Psi\|^2 + \|\Psi\|^2 \sum \|y_- c_\tau^{1/2} f_n\|^2\end{aligned}$$

by the commutation relations for a and a^*. Now the first sum is bounded as before, and the second sum

$$\begin{aligned}\sum \|y_- c_\tau^{1/2} f_n\|^2 &= \|y_- c_\tau^{1/2}\|_2^2 \\ &= \|y_-\chi\mu^\alpha \mu^{-\alpha} c_\tau^{1/2}\|_2^2 \\ &\leq \|y_-\chi\mu^\alpha\|_2^2 \|\mu^{-\alpha} c_\tau^{1/2}\|^2,\end{aligned}$$

which is bounded independently of g by (6.6) and (6.7) provided that $s\tau/2 < \alpha < 1/2$.

Remark: If $\tau = 0$, then the above proof works if $(H_0 + 1)^{-1/2}$ in (6.8)$_\pm$ is replaced by $(N + 1)^{-1/2}$.

Theorem 6.2: Let $N_B = N_{O,B}$ be the local number operator for the bounded interval $B \subset \mathbf{R}^s$. If $\tau < \min(\frac{1}{2}, (1/s))$, then

$$N_{\tau,B} \leq c[H_{\text{ren}}(g) + 1] \tag{6.9}$$

and if j is a positive integer then

$$N_B^j \leq d[H_{\text{ren}}(g) + 1]^j, \tag{6.10}$$

where the constants c and d depend on the volume of B but are otherwise independent of B and g.

Proof: By (2.3), we see that (6.9) is equivalent to the inequality $T_g^* N_{\tau,B} T_g \leq c[H_0(g) + 1]$ which can be reformulated as

$$\sum_n \|a(jc_\tau^{1/2} f_n) T_g (H_0(g) + 1)^{-1/2}\Psi\|^2 \leq c\|\Psi\|^2. \tag{6.11}$$

But by (A14),

$$a(jc_\tau^{1/2} f_n)T_g = T_g\, a(jy_+ c_\tau^{1/2} f_n) - T_g\, a^*(y_- c_\tau^{1/2} f_n). \tag{6.12}$$

To verify (6.11), we insert (6.12) into the left side of (6.11) and appeal to the triangle inequality, the unitarity of T_g, and the previous lemma. Note that by (2.2), $H_0 \leq aH_0(g) \leq bH_0$.

To establish (6.10), we prove that

$$T_g^* N_B^j T_g \leq d(N + 1)^j, \tag{6.13}$$

which is actually a stronger inequality since $\mu_g \geq m$. The proof of (6.13) is by induction on j, the case $j = 1$ being covered by the remark after Lemma 6.5.

Assume then that (6.13) holds for j, and prove it for $j + 1$. Now

$$(\Psi, T_g^* N_B^{j+1} T_g \Psi) = \sum_n \|a(j\chi f_n) N_B^{j/2} T_g \Psi\|^2.$$

But $[a(f), N_{\tau,B}] = a(jc_\tau jf)$, and, if the region B is symmetric about the origin so that $j\chi = \chi j$, we have

$a(j\chi f_n)N_B = (N_B + 1)a(j\chi f_n)$ It follows by the functional calculus that $a(j\chi f_n)N_B^{j/2} = (N_B + 1)^{j/2}a(j\chi f_n)$.

It is no loss of generality to assume that B is symmetric since N_B^j is monotonically increasing with respect to B. Thus

$$(\Psi, T_g^* N_B^{j+1} T_g \Psi) = \sum_n \|(N_B + 1)^{j/2} a(j\chi f_n) T_g \Psi\|^2$$
$$= \sum_n \|(N_B + 1)^{j/2} T_g [a(y_+ j\chi f_n) - a^*(y_-\chi f_n)]\Psi\|^2 \quad (6.14)$$

by (6.12). We apply the inductive assumption and the triangle inequality to the right side of (6.14) to dominate it by

$$2d[\sum_n \|(N+1)^{j/2} a(y_+ j\chi f_n)\Psi\|^2 + \sum_n \|(N+1)^{j/2} a^*(y_-\chi f_n)\Psi\|^2].$$

Each of the above sums can be estimated as in Lemma 6.5; for example,

$$\sum_n \|(N+1)^{j/2} a(y_+ j\chi f_n)\Psi\|^2 = \sum_n \|a(y_+ j\chi f_n) N^{j/2}\Psi\|^2$$
$$\leq d_1 \|(N+1)^{1/2} N^{j/2}\Psi\|^2$$
$$\leq d_1 \|(N+1)^{(j+1)/2}\Psi\|^2,$$

where d_1 is independent of g.

Corollary 6.1: $\omega_g(N_{\tau,B}) \leq c$.

Finally we note that it is possible to interpolate between the results (6.9) and (6.10). For instance, by a variant of the above proof it follows that if $\tau < \min(\frac{1}{2}, (1/s))$, then

$$N_{\tau,B} N_B^j \leq d[H_{\rm ren}(g) + 1]^{j+1}$$

and from this that if $\nu < (j+1)^{-1} s^{-[j/2]}\tau$, then

$$N_{\nu,B}^{j+1} \leq d[H_{\rm ren}(g) + 1]^{j+1}.$$

APPENDIX: DERIVATION OF THE DRESSING TRANSFORMATION

Following Friedrichs,[1] we first give a heuristic derivation of the dressing transformation T_g that leads to its defining property (A9). We then sketch a rigorous proof that the T_g constructed in Sec. 2 possesses this property.

Friedrichs observes that turning on the interaction $H_I(g)$ has the effect of replacing the energy function μ of the theory by μ_g. It is reasonable to expect that $H_{\rm ren}(g)$ can be expressed in the form

$$H_{\rm ren}(g) = \int b^*(k_1)\mu_g(k_1, k_2) b(k_2) dk_1 dk_2, \quad (A1)$$

where the $b^\#(= b$ or $b^*)$ are linear combinations of the $a^\#$. To determine $b^\#$ we note that

$$\hat{a}^\#(x) = (2\pi)^{-s/2} \int e^{ikx} a^\#(\pm k) dk$$
$$= (\mu/2)^{1/2}\phi(x,0) \pm i(2\mu)^{-1/2}\pi(x,0), \quad (A2)$$

while we ought to have

$$\hat{b}^\#(x) = (\mu_g/2)^{1/2}\phi(x,0) \pm i(2\mu_g)^{-1/2}\pi(x,0). \quad (A3)$$

Eliminating ϕ and π from (A2) and (A3), we obtain

$$b(k) = y_+ a(k) + y_- a^*(-k),$$
$$b^*(k) = y_- a(-k) + y_+ a^*(k), \quad (A4)$$

where

$$y_\pm = \tfrac{1}{2}(\mu_g^{1/2}\mu^{-1/2} \pm \mu_g^{-1/2}\mu^{1/2}). \quad (A5)$$

One can verify formally that this choice of $b^\#$ gives equality between the expressions (A1) and (2.5) for $H_{\rm ren}(g)$.

We now reformulate the transformation (A4) more carefully. Let $f \in L^2(R^s)$. For $\Psi = (\Psi_0, \Psi_1, \cdots) \in \mathfrak{F}_0$, the destruction operator $a(f)$ is defined by

$$(a(f)\Psi)_n(p_1, \ldots, p_n) = (n+1)^{1/2} \int dk f(k) \Psi_{n+1}(p_1, \ldots, p_n, k). \quad (A6)$$

Both $a(f)$ and its adjoint $[a(f)]^* = a^*(jf)$ are densely defined operators in $\mathfrak{F}_0$ whose domains include $D(N^{1/2})$ where N is the number operator. [Here j is complex conjugation in momentum space; for $f(x)$ real, $jf = rf$.] We define modified annihilation and creation operators on $D(N^{1/2})$ by

$$b(f) = a(w_+ f) + a^*(w_- jf),$$
$$b^*(f) = a(w_- jf) + a^*(w_+ f), \quad (A7)$$

where $w_\pm = y_\pm^* = \frac{1}{2}(\mu^{-1/2}\mu_g^{1/2} \pm \mu^{1/2}\mu_g^{-1/2})$ are bounded operators on $\mathfrak{K}$. It is easy to verify that $[b(f)]^* = b^*(jf)$ and that b and b^* satisfy the canonical commutation relations. Hence (A7) is a canonical transformation (i.e., a Bogoliubov or symplectic transformation).

In terms of the operators u and ρ of (2.6),

$$w_\pm = \tfrac{1}{2}(\rho \pm \rho^{-1})u \equiv h_\pm u. \quad (A8)$$

As proved in Corollary 2.1, $h_- \in$ HS. Therefore $w_- j \in$ HS and by a well-known theorem (see for instance Theorem 4.1 of Ref. 2), the canonical transformation (A7) is proper. That is, there exists a unitary operator T_g (unique up to phase) such that

$$T_g a^\#(f) T_g^{-1} = b^\#(f). \quad (A9)$$

Of course when $g \equiv 1$, $w_- j \notin$ HS and the transformation (A7) is improper. A comparison of (A1), (A9), and (2.3) shows that T_g is the desired dressing transformation of Sec. 2 which effects the unitary equivalence between $H_{\rm ren}(g)$ and $H_0(g)$.

It remains to verify that T_g as constructed in (2.10) satisfies (A9). As in Sec. 2, let $W = d\Gamma(w)$ where $u = e^{iw}$. Using the fact that j anticommutes with w we calculate that the commutator $(ad\, W)a(f) = [W, a(f)] = a(wf)$, valid as an operator equality on $D(N^{3/2})$. Taking f real for convenience we similarly find that on $D(N^{3/2})$, $(ad\, iL)a(f) = a^*(lf)$ and $(ad\, iL)a^*(f) = a(lf)$.

A formal proof of (A9) can be given as follows by means of the Baker–Hausdorff formula. With U_1 and U_2 defined as in (2.10) we have

$$U_1 a(f) U_1^* = \sum_{n=0}^{\infty} (\mathrm{ad} iL)^n a(f)/n!$$
$$= a[(\cosh l) f] + a^*[(\sinh l) f]$$
$$= a(h_+ f) + a^*(j h_- f) \quad (A10)$$

by (2.8) and (A8). Similarly,

$$U_2 a(f) U_2^* = \sum_{n=0}^{\infty} i^n a(w^n f)/n! = a(uf). \quad \text{(A11)}$$

Since $T_g = U_1 U_2$, we obtain

$$T_g a(f) T_g^* = a(h_+ uf) + a^*(jh_- uf)$$
$$= a(w_+ f) + a^*(w_- jf) \equiv b(f).$$

We now convert the above argument into a rigorous one that takes into account convergence and domain questions. Since W commutes with N, it is clear that $U_2(\alpha) = e^{i\alpha W}$ and $U_2(\alpha)^*$ leave $D(N^n)$ invariant for arbitrary real α and positive integer n. Let Φ and Ψ be vectors in $C^\infty(N)$. $F(\alpha) = (\Phi, U_2(\alpha) a(f) U_2(\alpha)^* \Psi)$ is then a well-defined infinitely differentiable function of α. In fact $F^{(n)}(0) = (\Phi, \mathrm{ad}iW)^n a(f)\Psi) = (\Phi, i^n a(w^n f)\Psi)$. Hence $F(\alpha)$ is an analytic function of α and in particular, $F(1) = \sum F^{(n)}(0)/n! = (\Phi, a(uf)\Psi)$. This establishes (A11) as a form equation on $C^\infty(N) \times C^\infty(N)$ and by extension as an operator equation on $D(N^{1/2})$.

The relation (A10) is somewhat more troublesome. Again we write $U_1(\alpha) = e^{i\alpha L}$. Since by Corollary 2.1 $l \in \mathrm{HS}$, it is easy to see that the expansion $U_1(\alpha)\Psi = \sum (i\alpha L)^n/n!\Psi$ is convergent for Ψ a vector with a finite number of particles, provided that $|\alpha| < \alpha_0 = \|l\|_2^{-1}$. This convergence enables us to prove that for $|\alpha| < \alpha_0$,

$$a(f) U_1(\alpha)^* = U_1(\alpha)^* [a(h_+(\alpha) f) + a^*(jh_-(\alpha) f)], \quad \text{(A12)}$$

where $h_+(\alpha) = \cosh(\alpha l)$ and $h_-(\alpha) = \sinh(\alpha l)$. Equation (A12) is first proved as a form relation on $\mathfrak{D} \times \mathfrak{D}$ [see (6.1)], but it then extends to an operator equality on $D(N^{1/2})$.

From (A12) we can then deduce that if $|\alpha| < \alpha_0$,

$$U_1(\alpha)^* D(N) \subset D(N). \quad \text{(A13)}$$

The method is to write $NU_1(\alpha)^* = \sum a^*(f_n) a(f_n) U_1(\alpha)^*$ as in Sec. 6 and to use (A12) and its adjoint relation to "commute" the $U_1(\alpha)^*$ to the left; this displays $NU_1(\alpha)^*$ in the form $U_1(\alpha)^* A$, where A is an unbounded operator whose domain includes $D(N)$; and (A13) is proved. But $U_1(\alpha/n)^n = U_1(\alpha)$ so that, in fact, (A13) holds for all α.

It is now a simple matter to prove (A10). For $\Phi, \Psi \in D(N)$, the function $G(\alpha) = (U_1(\alpha)^* \Phi, a(f) U_1(\alpha)^* \Psi)$ is C^1 in α by (A13). When $G'(\alpha)$ is calculated, it too is seen to be C^1 and by continuing this argument we see that $G(\alpha)$ is actually analytic in α. Thus (A10) follows as did (A11).

Finally we mention that the same method can be applied to show that on $D(N^{1/2})$:

$$T_g^* a(f) T_g = a(y_+ f) - a^*(j y_- f). \quad \text{(A14)}$$

1 K.O. Friedrichs, Commun. Pure Appl. Math. **6**, 1 (1953).
2 F. A. Berezin, *The Method of Second Quantization* (Academic, New York, 1966).
3 J. M. Chadam, J. Math. Phys. **9**, 386 (1968).
4 J. Ginibre and G. Velo, Commun. Math. Phys. **18**, 65 (1970).
5 W. J. Eachus, thesis (Syracuse University, 1970).
6 A. Kleinstein, thesis (Columbia University, 1970).
7 J. Glimm and A. Jaffe, "Field Theory Models," in the *1970 Les Houches Lectures*, edited by C. DeWitt and R. Stora (Gordon and Breach, New York, 1971).
8 J. Glimm and A. Jaffe, (a) Phys. Rev. **176**, 1945 (1968); (b) Ann. Math. **91**, 362 (1970); (c) Acta. Math. **126**, 203 (1970).
9 L. Rosen, Commun. Math. Phys. **16**, 157 (1970); Commun. Pure Appl. Math. **24**, 417 (1971). For additional references see Ref. 7.
10 J. Glimm and A. Jaffe, Ann. Phys. (N.Y.) **60**, 321 (1970); "The Yukawa$_2$ Quantum Field Theory without Cutoffs," J. Functional Analysis (to be published).
11 R. Schrader, "Yukawa Quantum Field Theory in Two Space–Time Dimensions without Cutoffs," Harvard University preprint.
12 T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York. 1966).
13 K. Löwner, Math. Z. **38**, 177 (1934); E. Heinz, Math. Ann. **123**, 415 (1951).
14 J. Glimm, Advan. Math. **3**, 101 (1969).
15 I. Segal, Trans. Amer. Math. Soc. **81**, 106 (1956).
16 See Ref. 5. For the case of negative λ, Eachus must impose the artificial condition $\|\mu^{-1} g \mu^{-1}\|_2 < 1/2$, which is not satisfied if supp(g) is large.
17 I. Segal, Proc. Natl. Acad. Sci. (U.S.) **57**, 1178 (1967).
18 J. Cook, Trans. Amer. Math. Soc. **74**, 222 (1953).
19 T. Ikebe, Arch. Ratl. Mech. Anal. **5**, 1 (1960).
20 J. Weidmann, Math. Z. **98**, 268 (1967).
21 R. Hoegh-Krohn, "On the Spectrum of the Space Cut-Off: $P(\phi)$: Hamiltonian in Two Dimensions," (unpublished).
22 M. Naimark, *Normed Rings* (Groningen, Noordhoff, 1964).
23 H. Araki, J. Math. Phys. **1**, 492 (1960).
24 J. Klauder and L. Streit, J. Math. Phys. **10**, 1661 (1969).
25 G. Watson, *Theory of Bessel Functions* (Cambridge U.P., Cambridge, 1944); W. Grobner and N. Hofreiter, *Integraltafel* (Springer-Verlag, Vienna, 1961).
26 J. Dixmier, *Les algebres d'operateurs dans l'espace Hilbertien (Algebres de von Neumann)* (Gauthier-Villars, Paris, 1957).
27 H. Araki, Progr. Theoret. Phys. **32**, 956 (1964).
28 E. Griffin, Trans. Amer. Math. Soc. **79**, 389 (1955).
29 N. Dunford and S. Schwartz, *Linear Operators. I* (Interscience, New York, 1957), p. 528.

Erratum: Green's Function for the Nonlocal Wave Equation

[J. Maths. Phys. **11**, 1938 (1970)]

G. A. Baraff
Bell Telephone Laboratories, Inc., Murray Hill, New Jersey 07974
(Received 22 February 1972)

The line just before Eq. (4.4) should read: "x_0 by $(x - x_0) \geq 0$ in (4.3) gives." Consequently, Eq. (4.5) is valid only for $x > x_0$. For $x < x_0$, note that Eq. (4.3) shows that $G(x, x_0) = G(x_0, x)$ so that we have

$$G(x, x_0) = -\int_0^{x_0} ds\, e(s) e(s + x - x_0), \quad x > x_0, \quad \text{(4.5a)}$$

$$G(x, x_0) = -\int_0^{x} ds\, e(s) e(s + x_0 - x), \quad x < x_0. \quad \text{(4.5b)}$$

These two equations together should replace both Eqs. (4.5) and (1.3a). Equation (1.3b) is, of course, correct as it stands.

$$U_2 a(f) U_2^* = \sum_{n=0}^{\infty} i^n a(w^n f)/n! = a(uf). \quad \text{(A11)}$$

Since $T_g = U_1 U_2$, we obtain

$$T_g a(f) T_g^* = a(h_+ uf) + a^*(jh_- uf)$$
$$= a(w_+ f) + a^*(w_- jf) \equiv b(f).$$

We now convert the above argument into a rigorous one that takes into account convergence and domain questions. Since W commutes with N, it is clear that $U_2(\alpha) = e^{i\alpha W}$ and $U_2(\alpha)^*$ leave $D(N^n)$ invariant for arbitrary real α and positive integer n. Let Φ and Ψ be vectors in $C^\infty(N)$. $F(\alpha) = (\Phi, U_2(\alpha) a(f) U_2(\alpha)^* \Psi)$ is then a well-defined infinitely differentiable function of α. In fact $F^{(n)}(0) = (\Phi, \mathrm{ad}iW)^n a(f)\Psi) = (\Phi, i^n a(w^n f)\Psi)$. Hence $F(\alpha)$ is an analytic function of α and in particular, $F(1) = \sum F^{(n)}(0)/n! = (\Phi, a(uf)\Psi)$. This establishes (A11) as a form equation on $C^\infty(N) \times C^\infty(N)$ and by extension as an operator equation on $D(N^{1/2})$.

The relation (A10) is somewhat more troublesome. Again we write $U_1(\alpha) = e^{i\alpha L}$. Since by Corollary 2.1 $l \in \mathrm{HS}$, it is easy to see that the expansion $U_1(\alpha)\Psi = \sum (i\alpha L)^n/n!\Psi$ is convergent for Ψ a vector with a finite number of particles, provided that $|\alpha| < \alpha_0 = \|l\|_2^{-1}$. This convergence enables us to prove that for $|\alpha| < \alpha_0$,

$$a(f) U_1(\alpha)^* = U_1(\alpha)^* [a(h_+(\alpha) f) + a^*(jh_-(\alpha) f)], \quad \text{(A12)}$$

where $h_+(\alpha) = \cosh(\alpha l)$ and $h_-(\alpha) = \sinh(\alpha l)$. Equation (A12) is first proved as a form relation on $\mathfrak{D} \times \mathfrak{D}$ [see (6.1)], but it then extends to an operator equality on $D(N^{1/2})$.

From (A12) we can then deduce that if $|\alpha| < \alpha_0$,

$$U_1(\alpha)^* D(N) \subset D(N). \quad \text{(A13)}$$

The method is to write $NU_1(\alpha)^* = \sum a^*(f_n) a(f_n) U_1(\alpha)^*$ as in Sec. 6 and to use (A12) and its adjoint relation to "commute" the $U_1(\alpha)^*$ to the left; this displays $NU_1(\alpha)^*$ in the form $U_1(\alpha)^* A$, where A is an unbounded operator whose domain includes $D(N)$; and (A13) is proved. But $U_1(\alpha/n)^n = U_1(\alpha)$ so that, in fact, (A13) holds for all α.

It is now a simple matter to prove (A10). For $\Phi, \Psi \in D(N)$, the function $G(\alpha) = (U_1(\alpha)^* \Phi, a(f) U_1(\alpha)^* \Psi)$ is C^1 in α by (A13). When $G'(\alpha)$ is calculated, it too is seen to be C^1 and by continuing this argument we see that $G(\alpha)$ is actually analytic in α. Thus (A10) follows as did (A11).

Finally we mention that the same method can be applied to show that on $D(N^{1/2})$:

$$T_g^* a(f) T_g = a(y_+ f) - a^*(j y_- f). \quad \text{(A14)}$$

1 K.O. Friedrichs, Commun. Pure Appl. Math. **6**, 1 (1953).
2 F. A. Berezin, *The Method of Second Quantization* (Academic, New York, 1966).
3 J. M. Chadam, J. Math. Phys. **9**, 386 (1968).
4 J. Ginibre and G. Velo, Commun. Math. Phys. **18**, 65 (1970).
5 W. J. Eachus, thesis (Syracuse University, 1970).
6 A. Kleinstein, thesis (Columbia University, 1970).
7 J. Glimm and A. Jaffe, "Field Theory Models," in the *1970 Les Houches Lectures*, edited by C. DeWitt and R. Stora (Gordon and Breach, New York, 1971).
8 J. Glimm and A. Jaffe, (a) Phys. Rev. **176**, 1945 (1968); (b) Ann. Math. **91**, 362 (1970); (c) Acta. Math. **126**, 203 (1970).
9 L. Rosen, Commun. Math. Phys. **16**, 157 (1970); Commun. Pure Appl. Math. **24**, 417 (1971). For additional references see Ref. 7.
10 J. Glimm and A. Jaffe, Ann. Phys. (N.Y.) **60**, 321 (1970); "The Yukawa$_2$ Quantum Field Theory without Cutoffs," J. Functional Analysis (to be published).
11 R. Schrader, "Yukawa Quantum Field Theory in Two Space–Time Dimensions without Cutoffs," Harvard University preprint.
12 T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York. 1966).
13 K. Löwner, Math. Z. **38**, 177 (1934); E. Heinz, Math. Ann. **123**, 415 (1951).
14 J. Glimm, Advan. Math. **3**, 101 (1969).
15 I. Segal, Trans. Amer. Math. Soc. **81**, 106 (1956).
16 See Ref. 5. For the case of negative λ, Eachus must impose the artificial condition $\|\mu^{-1} g \mu^{-1}\|_2 < 1/2$, which is not satisfied if supp(g) is large.
17 I. Segal, Proc. Natl. Acad. Sci. (U.S.) **57**, 1178 (1967).
18 J. Cook, Trans. Amer. Math. Soc. **74**, 222 (1953).
19 T. Ikebe, Arch. Ratl. Mech. Anal. **5**, 1 (1960).
20 J. Weidmann, Math. Z. **98**, 268 (1967).
21 R. Hoegh-Krohn, "On the Spectrum of the Space Cut-Off: $P(\phi)$: Hamiltonian in Two Dimensions," (unpublished).
22 M. Naimark, *Normed Rings* (Groningen, Noordhoff, 1964).
23 H. Araki, J. Math. Phys. **1**, 492 (1960).
24 J. Klauder and L. Streit, J. Math. Phys. **10**, 1661 (1969).
25 G. Watson, *Theory of Bessel Functions* (Cambridge U.P., Cambridge, 1944); W. Grobner and N. Hofreiter, *Integraltafel* (Springer-Verlag, Vienna, 1961).
26 J. Dixmier, *Les algebres d'operateurs dans l'espace Hilbertien (Algebres de von Neumann)* (Gauthier-Villars, Paris, 1957).
27 H. Araki, Progr. Theoret. Phys. **32**, 956 (1964).
28 E. Griffin, Trans. Amer. Math. Soc. **79**, 389 (1955).
29 N. Dunford and S. Schwartz, *Linear Operators. I* (Interscience, New York, 1957), p. 528.

Erratum: Green's Function for the Nonlocal Wave Equation

[J. Maths. Phys. **11**, 1938 (1970)]

G. A. Baraff
Bell Telephone Laboratories, Inc., Murray Hill, New Jersey 07974
(Received 22 February 1972)

The line just before Eq. (4.4) should read: "x_0 by $(x - x_0) \geq 0$ in (4.3) gives." Consequently, Eq. (4.5) is valid only for $x > x_0$. For $x < x_0$, note that Eq. (4.3) shows that $G(x, x_0) = G(x_0, x)$ so that we have

$$G(x, x_0) = -\int_0^{x_0} ds\, e(s) e(s + x - x_0), \quad x > x_0, \quad \text{(4.5a)}$$

$$G(x, x_0) = -\int_0^{x} ds\, e(s) e(s + x_0 - x), \quad x < x_0. \quad \text{(4.5b)}$$

These two equations together should replace both Eqs. (4.5) and (1.3a). Equation (1.3b) is, of course, correct as it stands.